

**Verschiedene Anwendungen  
kombinatorischer und algebraischer  
Strukturen in der Topologie**

Simon A. King

Habilitationsschrift  
eingereicht an der  
Technischen Universität Darmstadt  
Mai 2007

### **Danksagung**

Ich danke Professor Jürgen Bokowski für seine stetige Unterstützung, Professor Gert-Martin Greuel für die Chance, fast ein Jahr in Oberwolfach zu verbringen, dem Fachbereich Mathematik der TU Darmstadt für die Möglichkeit, mich in Forschung und Lehre weiterzuentwickeln, und meiner Familie für ihre große Geduld.

# Contents

Introduction	5
Chapter 1. Triangulations of Compact Surfaces	11
1. Regular flip equivalence	11
2. Proof sketch	12
Chapter 2. Transformations of Triangulations of $\mathbb{RP}^3$	15
1. The result	15
2. Proof sketch of Theorem 2	15
Chapter 3. Continued fractions and the Unwinding Lemma	17
1. Exposition of the problem	17
2. Approximation by continued fractions	19
3. Connections with the Unwinding Technique	20
Chapter 4. Phyllotaxis	25
1. Description of the problem	25
2. Various approaches to explain phyllotaxis	25
3. The light capture of plants	27
4. A simplified light capture model and the Golden Angle	29
Chapter 5. The Topological Representation of Oriented Matroids	31
1. Introduction	31
2. Hyperline sequences	32
3. A potential generalisation	36
Chapter 6. Ideal state sum invariants	39
1. Introduction	39
2. Ideal Turaev-Viro invariants of 3-manifolds	41
3. State sum invariants for the Andrews–Curtis problem	50
4. Ideal link invariants	55
Chapter 7. Fast Computation of Secondary Invariants	59
1. Motivation	59
2. Non-modular invariant rings	59
3. The New Algorithm	62
4. Benchmark Tests for the Computation of Secondary Invariants	65
Chapter 8. Minimal generating sets of non-modular invariant rings	71
1. The Algorithm	72
2. Benchmark Tests for the Computation of Minimal Generating Sets	74
3. Application to irreducible secondary invariants	80
Bibliography	83
Index	87



## Introduction

A very classical and fruitful branch of mathematics is located at the meeting points of two topics that initially look fairly different: *Algebraic Topology*. This Habilitationsschrift is concerned with various points of interaction of topology with algebraic or combinatorial approaches.

The interactions of topology and algebra are bidirectional. On the one hand, algebra can provide tools for solving topological problems. E.g., if one tries to find out whether two topological spaces are homeomorphic or not, one can read algebraic data (e.g., a group) off the spaces such that the algebraic data are invariant under homeomorphism. Hence, if one obtains different algebraic data then the spaces are not homeomorphic. Under very lucky circumstances, one even obtains an “if and only if”: then, the algebraic data contain the complete topological information. This is the case, e.g., for the classification of closed surfaces. On the other hand, the topological application can motivate algebraic studies.

Three problems have to be solved in that context, that I will illustrate with some well-known homeomorphism invariants.

- (1) How can one “read algebraic data off a space”? This very much depends on the way of presenting the topological data. E.g., if a topological space has been given the structure of a simplicial complex (a triangulation of the space), then it is very easy to write down a presentation of its fundamental group.
- (2) How can one prove invariance of the algebraic data under homeomorphism? Sometimes this is immediate from the definition, as in the case of singular homology, sometimes it is very hard, as in the case of simplicial homology.
- (3) How can one solve the algebraic classification problem? Assume we have two triangulated spaces  $X$  and  $Y$ , yielding presentations of the fundamental groups of  $X$  and  $Y$ . Then in general it is still a very difficult problem to test whether the two presentations define isomorphic groups.

In the first two of these problems, topological theorems play a prominent rôle. E.g., it depends on topological theorems whether an approach via triangulations works (not every topological space has a triangulation), and in some cases it will be difficult to provide the topological data in a form that is easy to deal with. And if the first problem is solved and one can try and define a homeomorphism invariant of compact 3-manifolds by means of triangulations, then one still needs to know how different triangulations of a manifold are related with each other.

The third problem looks purely algebraic. It may happen that this problem is unsolvable. In fact, any finitely presented group occurs as the fundamental group of a compact 4-dimensional manifold, and the isomorphism problem of finitely presented groups is algorithmically unsolvable. However, topological insight may provide additional information about the algebraic data, which in some cases may help. Here, topology can be a source of inspiration for the study of algebraic

structures, or even motivates the definition of new algebraic structures that, in some cases, become of independent interest.

Last but not least, algebraic structures may be directly blended with topological structures, as is the case, e.g., for Lie groups. This is another obvious way how topology and algebra may interact.

In the remainder of the introductory chapter, we describe how the contents of this Habilitationsschrift fit in that framework.

In the first two chapters, we study the question how different triangulations of a manifold can be related with each other. In the first chapter, the focus is on triangulations of surfaces. It was known before that any two triangulations can be related by diagonal flips and by subdivision of 2-simplices (or the inverse process) [78]. In contrast to the subdivision, a diagonal flip leaves the number of vertices unchanged. So a question naturally occurs: Provided two triangulations of a closed surface have the same number of vertices, is it possible to relate them by diagonal flips, without subdivisions and without leaving the class of triangulations?

In general, the answer is “no”, as there are examples of several non-isomorphic triangulations of surfaces over the same number of vertices that do not admit any diagonal flip at all [2]. Negami [77] has shown by a non-constructive proof that if the number of vertices is sufficiently large, then a transformation by diagonal flips alone is always possible. We provide a constructive version of Negami’s statement: If two triangulations of a closed surface have the same number of vertices and the number of vertices exceeds some bound that is linear in the Euler characteristic, then a transformation by diagonal flips is possible.

In the second chapter, the dimension increases by one, as we consider triangulations of the 3-dimensional projective space. We give an upper bound in terms of the number of tetrahedra for the length of a minimal transformation sequence relating any two triangulations of  $\mathbb{RP}^3$  (see [51]). Here, the elementary transformations are contraction of edges and the inverse process. The basic approach of tackling that problem was exposed in [48] and [50], that was part of our doctoral thesis. There, we studied triangulations of  $\mathbb{S}^3$ . The extension of our approach from  $\mathbb{S}^3$  to  $\mathbb{RP}^3$  involves techniques that have not been used in the doctoral thesis.

The most important tool for the proof, that is only outlined in Chapter 2, is the theory of normal surfaces and their generalisations. Originally introduced by H. Kneser [57], normal surfaces became a very powerful tool for the algorithmic classification of a large class of compact 3-dimensional spaces by work of W. Haken [31], [32]. Haken’s techniques assume the existence of *incompressible surfaces*. Therefore the somehow “simplemost” closed 3-manifold, namely the sphere  $\mathbb{S}^3$ , is inaccessible for Haken’s algorithms. This inconvenient situation was improved when H. Rubinstein suggested a recognition algorithm for  $\mathbb{S}^3$ , based on a generalisation of normal surfaces [84]. The correctness of Rubinstein’s algorithm was proven by A. Thompson [95]. A main textbook on algorithmic methods in low-dimensional topology is [68].

The set of normal surfaces is equipped with a partial addition. Haken [31], [32] showed that there is a finite set of so-called *fundamental surfaces* that additively generate the set of normal surfaces. The fundamental surfaces can be constructed by means of Integer Programming; see also [33], [48].

The set of normal surfaces is infinite. However, often it is possible to restrict the quest for an interesting surface  $F$  in a 3-manifold  $M$  to a *finite* set of normal surfaces: The idea is to prove the existence of a homeomorphism of  $M$  taking  $F$  to a normal surface that can be expressed as a sum of fundamental surfaces with bounded coefficients. One approach to find such homeomorphism is Hemion’s Unwinding Technique. In G. Hemion’s book [35], the proof of the Unwinding

Technique is left as an exercise to the reader, which is very unfortunate since the technical details of the proof soon turn out to be extremely nasty. Therefore it is common belief of the experts that the Unwinding Technique can not be taken for granted.

In Chapter 3, we provide some evidence that the Unwinding Technique is erroneous. Although we did not succeed in constructing an actual counterexample, our considerations show at least that a proof of the Unwinding Technique would go far beyond the scope of an exercise in a textbook. Our approach is based on the approximation of irrational numbers by continued fractions.

Chapter 4 is an intermezzo on phytology. The aim is to explain a long-standing observation about the leaf arrangement of higher plants: If the leaves are arranged along a spiral around the stem, then the angle between two adjacent leaves (*divergence angle*) strongly tends towards  $137.5^\circ$ , which is the Golden Section times  $180^\circ$ . There have been various explanations of that observation. E.g., experimental results indicate that a divergence angle of  $137.5^\circ$  maximizes the light capture and thus the photosynthetic activity of the plant, which is a strong benefit and thus explains the evolution of that divergence angle. Based on very mild and natural assumptions, we construct a simple model for the light capture of plants with helical leaf arrangement. The model only has three parameters, one of them a proportionality constant. It turns out that, when choosing two of the parameters according to the experimental setting, an adjustment of the proportionality constant by least square fit suffices to perfectly explain the available experimental data. Hence, simple though the model is, it is very realistic. An even more simplistic version of our model allows an analytic study, and this is how it connects with the previous chapter: Again, we can use approximation of irrational numbers by continued fractions. It turns out that the simplified light capture model attains the optimum exactly for the “Golden” divergence angle.

These results are exposed in a joint paper [53]. I contributed the construction of the model and the number theoretical analysis. F. Beck provided the numerical simulations, and U. Lüttge provided the phytological background.

Chapter 5 returns to topology. Here, the dimension is arbitrary, and the algebraic structure has a very combinatorial flavour: Oriented Matroids. Oriented Matroids did occur in various contexts, therefore yielding various axiomatics that in the end all turned out to be equivalent. We add one more axiomatics, allowing a new proof of the Topological Representation Theorem for Oriented Matroids. That theorem states a one-one correspondence of the set of Oriented Matroids with the set of equivalence classes of pseudosphere arrangements up to homeomorphism. A pseudosphere is a tame oriented sub-sphere of co-dimension one in a (high-dimensional) oriented sphere. A pseudosphere arrangement is a set of pseudospheres so that any “small” sub-arrangement is equivalent (by a homeomorphism) to an arrangement of hyperspheres. Not any Oriented Matroid can be realised by an arrangement of hyperspheres, but any Oriented Matroid can be realised by an arrangement of pseudospheres.

The new axiomatics, which is based on so-called *hyperline sequences*, not only yields a new proof of the Topological Representation Theorem. It also gives a very short proof of the fact that even arrangements of “wild” spheres allow to read off an Oriented Matroid.

A natural generalisation of arrangements of pseudospheres appear to be “arrangements of co-dimension one sub-manifolds”. Along that line, Oriented Matroids are (weakly) related to normal surfaces discussed in previous chapters. We discuss the question to what extent a generalisation of the topological notion of

arrangement corresponds to a natural generalisation of the combinatorial notion of Oriented Matroids defined in terms of hyperline sequences.

Most of the contents of Chapter 5 are published in a joint paper with J. Bokowski, S. Mock and I. Streinu. Hyperline sequences have been introduced by Bokowski; Bokowski, Mock and Streinu used hyperline sequences for a proof of the Topological Representation Theorem in rank 3. My contribution was the formulation of the axiomatics based on Bokowski’s ideas, a proof of the equivalence to previous definitions of Oriented Matroids, an extension of the rank-3-proof to arbitrary rank, which involved some advanced results from topology, and the study of “wild” arrangements.

Oriented Matroids can be understood as invariants of oriented pseudosphere arrangements under homeomorphisms. The Topological Representation Theorem even implies that they are complete invariants. That point of view connects Chapter 5 with the next chapter.

Chapter 6 is devoted to a certain type of homeomorphism invariants for compact 3-dimensional manifolds. The prototype of these invariants has been introduced by V. Turaev and O. Viro [96]. Originally, the Turaev–Viro invariants associate to any compact 3-manifold an element of an algebraic extension of the rational numbers, only depending on the homeomorphism type of the manifold. These invariants are defined in terms of state sums, that depend on the choice of a triangulation of the manifold. The state sum is a polynomial, and a Turaev–Viro invariant is obtained by evaluation of the state sum at certain values. These values are provided by the representation theory of Quantum Groups at roots of unity.

We generalise the Turaev–Viro invariants, avoiding the use of representation theory. Let  $R$  be the ring in which the state sums live. We define an ideal  $I \subset R$  such that the image of the state sum in the quotient ring  $R/I$  is a homeomorphism invariant of compact 3-manifolds. We call these *ideal Turaev–Viro invariants*. By definition, they are at least as strong as the “classical” Turaev–Viro invariants. A problem, however, is the practical computation, as we have to deal with quotient rings of polynomial rings with many variables (usually more than twenty). Fortunately, modern computer algebra systems (such as SINGULAR [21]) allow to compute ideal Turaev–Viro invariants, based on Gröbner basis methods. It turns out that the ideal invariants are considerably stronger than the classical invariants associated to Quantum Groups.

Ideal Turaev–Viro invariants are a potential tool to disprove the Andrews–Curtis conjecture. This 40 year old conjecture from combinatorial group theory has a topological counterpart. Essentially, one defines an equivalence relation (3-deformation) on compact 2-dimensional cellular complexes, and the conjecture states that there is only one equivalence class containing contractible complexes.

Nearly as old as the conjecture are several infinite classes of potential counterexamples. Obviously, a 3-deformation invariant could be used to verify whether one of the candidates is actually a counterexample for the Andrews–Curtis conjecture. Turaev–Viro invariants can be modified to yield 3-deformation invariants. However, by a result of I. Bobtcheva and F. Quinn [6], it is impossible to detect Andrews–Curtis counterexamples by Turaev–Viro invariants associated to Quantum Groups. The main reason is that these are multiplicative under connected sum of 3-manifolds.

Examples show that our ideal Turaev–Viro invariants are *not* multiplicative under connected sum. Hence, even after the result of Bobtcheva and Quinn, the possibility remains to detect Andrews–Curtis counterexamples by *ideal* Turaev–Viro invariants. Unfortunately, so far we did not succeed in finding a counterexample.

In the last part of Chapter 6, we briefly discuss an analogous approach towards the construction of new invariants for knots and links.

The polynomial ring  $R$  containing the state sums is naturally equipped with an action of some finite permutation group, acting on the variables in some non-standard way. The state sums are invariant under the group action, and so is the ideal  $I$  used in the definition of ideal Turaev–Viro invariants. This motivated the attempt to define and compute everything in the *invariant ring*, i.e., in the sub-algebra of  $R$  formed by polynomials that are invariant under the group action. Since we are in characteristic 0, we have the problem of finding generators of a non-modular invariant ring of an action of a finite group. There is software for that purpose, but neither SINGULAR [21] nor MAGMA [11] was strong enough to compute our examples, except in the most easy cases.

This motivated us to try and find new algorithms for the computation of generators of invariant rings of non-modular finite group actions. Our algorithms rely on a result about homogeneous Gröbner bases with degree bounds. All our algorithms are implemented in the SINGULAR library `finvar.lib`. First, we implemented an algorithm for the computation of secondary and irreducible secondary invariants. This is explained Chapter 7. Later, we developed an algorithm for the computation of minimal generating sets of invariant rings. That algorithm also provides an alternative way for the computation of irreducible secondary invariants. This is explained in the final Chapter 8. We made extensive comparative benchmark tests with our algorithms implemented in SINGULAR and the algorithms of Derksen, Kemper and Steel [22], [45], [47] implemented in MAGMA. For our benchmarks, we used classical examples such as all transitive permutation groups on 7 or 8 variables, and also some of the examples motivated by our study of ideal Turaev–Viro invariants.

It turns out that our algorithms mark a dramatic breakthrough. In quite a few of the examples, our algorithms are more than 1000 times faster than previously known algorithms. Although our study of invariant rings was originally motivated by applications in topology, it became of independent interest.



## Triangulations of Compact Surfaces

### 1. Regular flip equivalence

Let  $F$  be a closed surface and let  $\chi(F)$  be its Euler characteristic. A **singular triangulation** of  $F$  is a graph  $T$  embedded in  $F$  such that each face of  $F \setminus T$  is bounded by an edge path of length three. We denote by  $v(T)$ ,  $e(T)$  and  $f(T)$  the number of vertices, edges and faces of  $T$ . If  $T$  is without loops and multiple edges and has more than three faces, then  $T$  corresponds to a triangulation of  $F$  in the classical meaning of the word; in order to avoid confusions, we use for it the term **regular triangulation** in this section.

Let  $e$  be an edge of a singular triangulation  $T$  and suppose that there are two distinct faces  $\delta_1$  and  $\delta_2$  adjacent to  $e$ . The faces  $\delta_1$  and  $\delta_2$  form a (possibly degenerate) quadrilateral, containing  $e$  as a diagonal. A **flip** of  $T$  along  $e$  replaces  $e$  by the opposite diagonal of this quadrilateral, see Figure 1. The flip is called

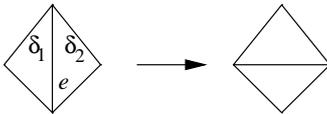


FIGURE 1. A flip

**regular**, if both  $T$  and the result of the flip are regular triangulations. Two singular (resp. regular) triangulations  $T_1, T_2$  of a closed surface are called flip equivalent (resp. regularly flip equivalent), if they are related by a finite sequence of flips (resp. regular flips) and isotopy.

The following result is well known, and there are many proofs for it. There are interesting applications to the automatic structure of mapping class groups, see [74] or [83].

LEMMA 1. *Any two singular triangulations  $T_1$  and  $T_2$  of a closed surface  $F$  with  $v(T_1) = v(T_2)$  are flip equivalent.*  $\square$

One might ask whether any two regular triangulations of  $F$  with the same number of vertices are *regularly* flip equivalent. The answer is “Yes” in special cases: any two regular triangulations of the sphere [98], the torus [23], the projective plane or the Klein bottle [76] with the same number of vertices are regularly flip equivalent. But in general, the answer is “No”: it is known that there are 59 different triangulations of the closed oriented surface of genus six based on the complete graph with 12 vertices, see [2]. Such a triangulation does not admit any regular flip, thus the different triangulations are not regularly flip equivalent.

In [49], we obtain the following result.

THEOREM 1. *Let  $F$  be a closed surface and  $N(F) = 9450 - 6020\chi(F)$ . Any two regular triangulations  $T_1$  and  $T_2$  of  $F$  with  $v(T_1) = v(T_2) \geq N(F)$  are regularly flip equivalent.*

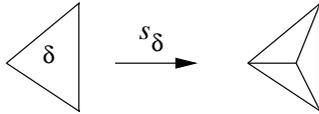


FIGURE 2. A face subdivision

Negami [77] stated the mere existence of  $N(F)$  without an estimate. The estimate in Theorem 1 is far from being best possible, at least for the surfaces up to genus one. The number  $N(F)$  is negative if and only if  $F$  is a sphere, in which case the statement is true since the transformation by regular flips is always possible, by Wagner’s Theorem [98]. We assume in the following that  $F$  is not the sphere.

## 2. Proof sketch

Let  $\mathcal{T}'$  denote the barycentric subdivision of a singular triangulation  $\mathcal{T}$  of a closed 2-manifold  $F$ .

LEMMA 2 (Lemma 5 in [49]). *Let  $\mathcal{T}_1$  and  $\mathcal{T}_2$  be two singular triangulations of  $F$  with the same number of vertices. Then  $\mathcal{T}_1''$  and  $\mathcal{T}_2''$  are regularly flip equivalent.*  $\square$

Let  $\delta$  be a face of a regular triangulation  $\mathcal{T}$ . A **face subdivision** of  $\mathcal{T}$  along  $\delta$  replaces  $\delta$  by the cone over its boundary, see Figure 2, and the result is denoted  $s_\delta\mathcal{T}$ . If  $\delta$  and  $\delta'$  are two faces of  $\mathcal{T}$ , then  $s_\delta\mathcal{T}$  and  $s_{\delta'}\mathcal{T}$  are regularly flip equivalent, which is easy to see. Hence, up to regular flip equivalence, the result of a sequence of face subdivisions only depends on the number of subdivisions. If  $\mathcal{T}_2$  is obtained from  $\mathcal{T}_1$  by  $m$  successive face subdivisions, we write  $\mathcal{T}_2 = s^m(\mathcal{T}_1)$ , which is well-defined up to regular flip equivalence.

We need a further notion, that also plays a role in some other of our results.

DEFINITION 1. Let  $M$  be a closed PL-manifold with PL-triangulations  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , and let  $e$  be an edge of  $\mathcal{T}_1$  with  $\partial e = \{a, b\}$ . Suppose that  $\mathcal{T}_2$  is obtained from  $\mathcal{T}_1$  by removing the open star of  $e$  and identifying  $a * \sigma$  with  $b * \sigma$  for any simplex  $\sigma$  in the link of  $e$ . Then  $\mathcal{T}_2$  is obtained from  $\mathcal{T}_1$  by a **contraction** along  $e$ , and  $\mathcal{T}_1$  is obtained from  $\mathcal{T}_2$  by an **expansion** along  $e$ .

In general, there are edges of  $\mathcal{T}_1$  along which contractions are impossible. This is the case, e.g., if an edge  $e$  of  $\mathcal{T}_1$  is part of an edge path of length 3 that does not bound a 2-simplex of  $\mathcal{T}_1$ . Indeed then  $\mathcal{T}_2$  has multiple edges and is not a simplicial complex. We only consider manifolds of dimension 2 and 3, so any triangulation is PL, so for simplicity, we write “triangulation” in the place of “PL-triangulation”. After these preliminaries, we can cite a lemma of Negami [77].

LEMMA 3. *Let  $\mathcal{T}_1$  and  $\mathcal{T}_2$  be regular triangulations of  $F$ . If  $\mathcal{T}_2$  is obtained by contraction along some edges of  $\mathcal{T}_1$ , then  $\mathcal{T}_1$  is regularly flip equivalent to  $s^m(\mathcal{T}_2)$ , with  $m = v(\mathcal{T}_1) - v(\mathcal{T}_2)$ .*  $\square$

The next brick in our proof of Theorem 1 is a result of Nakamoto and Ota on irreducible triangulations. A triangulation is called **irreducible**, if it has no edge along which a contraction is possible. Nakamoto and Ota [75] found the following bound for the number of vertices of irreducible triangulations of a closed surface is bounded in terms of the Euler characteristic.

PROPOSITION 1 (see [75]). *If  $T$  is an irreducible triangulation of a closed surface  $F$  which is not the sphere, then  $v(T) \leq 270 - 171\chi(F)$ .*  $\square$

The last brick in the proof of Theorem 1 is the following lemma that relates singular with regular flips. Let  $T'$  denote the barycentric subdivision of a singular triangulation  $T$  of  $F$ .

LEMMA 4 (see [49]). *Let  $T_1$  and  $T_2$  be two singular triangulations of  $F$  with  $v(T_1) = v(T_2)$ . Then  $T_1''$  and  $T_2''$  are regularly flip equivalent.*  $\square$

COROLLARY 1. *Let  $T_1$  and  $T_2$  be two regular triangulations of  $F$  with  $v(T_1) = v(T_2)$ . Then  $s^m(T_1)$  and  $s^m(T_2)$  are regularly flip equivalent, with*

$$m = 35(v(T_1) - \chi(F)).$$

To obtain Corollary 1, one first verifies that  $m = v(T_1) - v(T_1'') = v(T_2) - v(T_2'')$ ; then one shows that  $s^m(T_i)$  is regularly flip equivalent to  $T_i''$ , for  $i = 1, 2$ ; by Lemma 1 and Lemma 4,  $T_1''$  and  $T_2''$  are regularly flip equivalent, and thus the corollary follows.

Now, we can finish the proof of Theorem 1. Let  $N(F) = 9450 - 6020\chi(F)$  and let  $T_1, T_2$  be two triangulations of  $F$  with  $v(T_1) = v(T_2) = N_0 \geq N(F)$ . By contractions along some edges,  $T_i$  ( $i \in \{1, 2\}$ ) can be transformed into an irreducible triangulation  $S_i$ . By Lemma 3,  $T_i$  is regularly flip equivalent to  $s^{N_0 - v(S_i)}S_i$ . By Proposition 1, we have  $N(F) \geq 35(v(S_i) - \chi(F))$ . Hence by Corollary 1,  $s^{N(F) - v(S_1)}S_1$  and  $s^{N(F) - v(S_2)}S_2$  are regularly flip equivalent, and so are also  $s^{M - v(S_1)}S_1$  and  $s^{M - v(S_2)}S_2$  after further face subdivisions. Therefore also  $T_1$  and  $T_2$  are regularly flip equivalent.  $\square$



## Transformations of Triangulations of $\mathbb{RP}^3$

### 1. The result

Let  $M$  be a compact 3-manifold with two triangulations  $\mathcal{T}_1, \mathcal{T}_2$ . Recall the definition of edge expansions and contractions (Definition 1). Moise [72] has shown in 1952 that any 3-manifold has a unique PL-structure and all its triangulations are PL. So it a consequence of results of Pachner [78] that one can transform  $\mathcal{T}_1$  into  $\mathcal{T}_2$  by a finite sequence of edge expansions and contractions. For a given number  $n$ , there are only finitely many triangulations of  $M$  with at most  $n$  tetrahedra. Hence, there is some number  $\alpha_M(n)$  such that any two triangulations of  $M$  with at most  $n$  tetrahedra can be related by a sequence of less than  $\alpha_M(n)$  edge expansions and contractions. Since for any triangulation there are only finitely many ways to perform a single expansion or contraction, one obtains a recognition algorithm<sup>1</sup> for  $M$ , provided  $\alpha_M(n)$  is computable.

By [48] and [50], we obtain  $\alpha_{\mathbb{S}^3}(n) \leq 2^{cn^2}$  for some explicitly given constant  $c > 0$ ; this was part of our doctoral thesis. In [51], we extend our methods and obtain the following result on triangulations of the projective space  $\mathbb{RP}^3$ .

**THEOREM 2.** *Any two triangulations of  $\mathbb{RP}^3$  with at most  $n$  tetrahedra are related by a sequence of less than  $2^{27000n^2}$  edge contractions and expansions.*

The constant factor in the exponent is certainly not optimal. According to the examples in [50], concerning the minimal number of edge expansions needed to transform a triangulation of  $\mathbb{S}^3$  into a *polytopal* triangulation, we believe that the bound in Theorem 2 can not be replaced by a subexponential bound.

Based on our results in [48], Mijatović obtained results similar to Theorem 2. He deals with Pachner moves rather than edge expansions and contractions. His result is more general, since it concerns triangulations of a large class of manifolds, namely fibre free Haken manifolds [71]. However, his bound for the number of moves is much weaker, as it is expressed by the  $2^{an}$ -fold composition of exponential functions, for some  $a \approx 200$ .

### 2. Proof sketch of Theorem 2

All details can be found in [51]. Let  $\mathcal{T}$  be a triangulation of  $\mathbb{RP}^3$  with  $n$  tetrahedra. Our aim is to transform  $\mathcal{T}$  into one of two standard triangulations of  $\mathbb{RP}^3$  (to be defined below). Let  $\mathcal{C}$  be the dual cellular decomposition of  $\mathcal{T}$ . The first step is the construction of a normal projective plane  $P \subset \mathbb{RP}^3$  with respect to  $\mathcal{T}$ , with the additional property that  $P \setminus \mathcal{C}^2$  is a disjoint union of disks, and for any 2-dimensional cell  $c$  of  $\mathcal{C}$  the intersection  $P \cap c$  is a disjoint union of arcs, each of which connects different edges of  $c$ . We find an upper bound for  $\|P\|$  in terms of  $n$ , which is mainly by Theorem 4.

The complement of  $P$  in  $\mathbb{RP}^3$  is a ball. By the methods that we developed in [48], we construct a sweep-out  $H: \mathbb{S}^2 \times I \rightarrow \mathbb{RP}^3 \setminus U(P)$ , i.e., an embedding in

---

<sup>1</sup>This means an algorithm that, for any triangulated 3-manifold  $N$ , recognizes whether  $M \approx N$  or not.

general position with respect to  $\mathcal{C}$ , with level surfaces  $H_\xi = H(\mathbb{S}^2 \times \{\xi\})$  for  $\xi \in I$ . Moreover,  $H_0 = \partial U(x)$  for some vertex  $x$  of  $\mathcal{C}$ , and  $H_1 = 2P$ . There are only finitely many  $0 < \xi_1 < \xi_2 < \dots < \xi_N < 1$  (the ‘‘critical parameters’’) such that  $H_{\xi_i}$  is not transversal to  $\mathcal{C}^2$ , and  $N < 2^{cn^2}$  for some explicitly given constant  $c > 0$ . By the methods described in [50], we can additionally achieve that for all non-critical parameters  $\xi \in I$  the special polyhedron  $H_\xi \cup (\mathcal{C}^2 \setminus H(\mathbb{S}^2 \times [0, \xi])) \subset \mathbb{RP}^3$  is the 2-skeleton of the dual cellular decomposition of some triangulation  $\mathcal{T}_\xi$  of  $\mathbb{RP}^3$ .

It is not difficult to provide a sequence of edge expansions that changes  $\mathcal{T}$  into  $\mathcal{T}_0$ . We study how  $\mathcal{T}_\xi$  changes for increasing  $\xi$ . First, if  $\xi$  does not pass a critical parameter, then the isomorphism type of  $\mathcal{T}_\xi$  as a simplicial complex does not change. When  $\xi$  passes a critical parameter  $\xi_i$ , we obtain an explicit sequence of edge expansions and contractions transforming  $\mathcal{T}_{\xi_i - \epsilon}$  into  $\mathcal{T}_{\xi_i + \epsilon}$ , for small  $\epsilon > 0$ , by looking at different types of critical parameters. In conclusion, we obtain a sequence of edge expansions and contractions transforming  $\mathcal{T}$  into  $\mathcal{T}_1$ , which an upper bound in terms of  $n$  for the number of moves. The triangulation  $\mathcal{T}_1$  depends on  $P \cap \mathcal{C}^2$ , as is studied in the following paragraphs.

Let  $\mathcal{Z}$  be a cellular decomposition of  $P$  that is dual to some triangulation; for example, take  $\mathcal{Z}$  with  $\mathcal{Z}^1 = P \cap \mathcal{C}^2$ , which can be shown to be dual to some triangulation  $T$ . A regular neighbourhood  $U(P) \subset \mathbb{RP}^3$  fibers over  $P$ . By lifting all cells of  $\mathcal{Z}$  along the fibers, adding  $U(P)$  and  $2P = \partial U(P)$ , we obtain a cellular decomposition of  $\mathbb{RP}^3$ , and it turns out that its barycentric subdivision is a triangulation  $\mathcal{T}(T)$  of  $\mathbb{RP}^3$ . In the case  $\mathcal{Z}^1 = P \cap \mathcal{C}^2$ , we obtain in that way the barycentric subdivision of  $\mathcal{T}_1$ , which is easily obtained from  $\mathcal{T}_1$  by edge expansions.

The next step relies on results of Barnette [4] on triangulations of the projective plane. The idea is to simplify  $T$  by edge contractions (this time, not in the 3-dimensional, but in the 2-dimensional setting) until a further edge contraction is impossible since it would introduce multiple edges. This yields a sequence  $T = T_1, T_2, \dots, T_K$  of triangulations of  $P$ , where, of course,  $K$  is bounded by the number of edges of  $T$ .

A triangulation of a surface which does not allow an edge contraction is called **irreducible**. By [4], there are exactly two irreducible triangulations  $\tilde{T}_1, \tilde{T}_2$  of the projective plane (up to isomorphism of simplicial complexes). If  $T_{i+1}$  is obtained from  $T_i$  by an edge contraction (in dimension 2), then  $\mathcal{T}(T_{i+1})$  is obtained from  $\mathcal{T}(T_i)$  by a sequence of 18 edge contractions (in dimension 3).

In conclusion, we can transform any triangulation  $\mathcal{T}$  of  $\mathbb{RP}^3$  with at most  $n$  tetrahedra into either  $\mathcal{T}(\tilde{T}_1)$  or  $\mathcal{T}(\tilde{T}_2)$ , by a sequence of edge expansions and contractions, whose length is bounded in terms of  $n$ . Finally, we provide a sequence of edge contractions and expansions relating  $\mathcal{T}(\tilde{T}_1)$  with  $\mathcal{T}(\tilde{T}_2)$ . Hence, we can relate any two triangulations of  $\mathbb{RP}^3$  by a sequence of edge expansions and contractions, via  $\mathcal{T}(\tilde{T}_1)$  or  $\mathcal{T}(\tilde{T}_2)$ . This finishes the proof of Theorem 2.  $\square$

## Continued fractions and the Unwinding Lemma

### 1. Exposition of the problem

We recall here the main features of the theory of normal surfaces. For details, see [68], for instance. We denote the number of connected components of a compact topological space  $X$  by  $\#(X)$ . Let  $\mathcal{T}$  be a triangulation of a compact orientable 3-manifold  $M$  with  $n$  tetrahedra. A **normal isotopy** with respect to  $\mathcal{T}$  is an isotopy of  $M$  that preserves any simplex of  $\mathcal{T}$  set-wise. A **normal surface**  $F \subset M$  is an embedded not necessarily connected surface with  $\partial F \subset \partial M$  that is in general position with respect to  $\mathcal{T}$  such that the intersection of  $F$  with any tetrahedron of  $\mathcal{T}$  is a disjoint union of triangles and quadrilaterals as shown in Figure 1. We refer to them as **normal pieces**. Let  $\|F\| = \#(F \cap \mathcal{T}^1)$ . By slight abuse of notation, if we have two normal surfaces  $F_1, F_2$ , we denote by  $\#(F_1 \cap F_2)$  the minimal number of lines of intersection of  $F_1$  with  $F_2$ , up to normal isotopy of  $F_1$  and  $F_2$ .

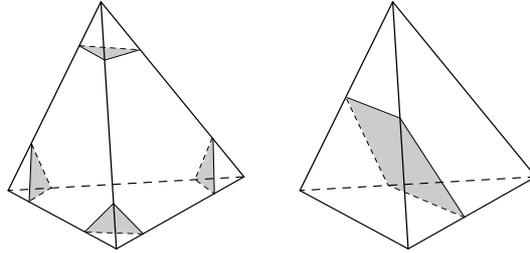


FIGURE 1. Normal Triangles (left) and Normal Quadrilateral (right)

By the **class** of a normal piece, we mean its normal isotopy class. There are 7 classes of normal pieces in each tetrahedron: Four classes of normal triangles (one for edge vertex of the tetrahedron) and three classes of normal quadrilaterals (one for each pair of opposite edges of the tetrahedron). So in total, there are  $7n$  classes of normal pieces in  $\mathcal{T}$ . Any normal surface  $F$  gives rise to a vector  $\mathbf{v}(F) \in \mathbb{Z}_{\geq 0}^{7n}$ , whose coefficients correspond to the  $7n$  classes of normal pieces and indicate how many copies of each normal piece occur in  $F$ . It is not difficult to show that  $F$  is determined by  $\mathbf{v}(F)$  up to normal isotopy. Actually one can characterise the vectors that correspond to normal surfaces.

The set  $\mathcal{F}$  of normal surfaces is equipped with a partial addition. Let  $F_1, F_2 \subset M$  be two normal surfaces, and assume that whenever  $t$  is a tetrahedron of  $\mathcal{T}$  such that both  $F_1 \cap t$  and  $F_2 \cap t$  comprise quadrilaterals, then the quadrilaterals of  $F_1 \cap t$  belong to the same class than those of  $F_2 \cap t$ ; this is called the *compatibility condition*. Then, there is a normal surface  $G \subset M$  with  $\mathbf{v}(G) = \mathbf{v}(F_1) + \mathbf{v}(F_2)$ , and one denotes  $G = F_1 + F_2$ .

There is a geometric interpretation of the addition of normal surfaces: First, up to normal isotopy we can assume that  $F_1$  intersects  $F_2$  in general position. Then one

obtains  $F_1 + F_2$  from  $F_1 \cup F_2$  by cutting along  $F_1 \cap F_2$  and connecting the components of  $(F_1 \cup F_2) \setminus U(F_1 \cap F_2)$  by annuli, obtaining an embedded surfaces. A priori, there are two ways to do so. But there is precisely one way that yields a normal surface, provided  $F_1$  and  $F_2$  satisfy the compatibility condition; compare Figure 2. This geometric interpretation makes it easy to see that  $\chi(F_1 + F_2) = \chi(F_1) + \chi(F_2)$  and  $\|F_1 + F_2\| = \|F_1\| + \|F_2\|$ .

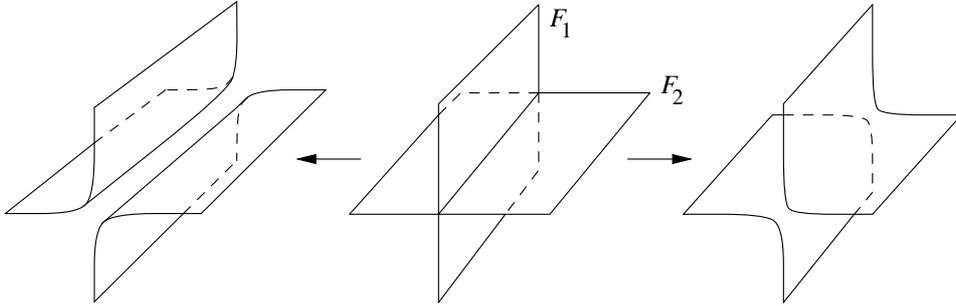


FIGURE 2. The two ways to switch

The following two finiteness results are essential for the theory of normal surfaces.

**THEOREM 3** (Kneser's Lemma, Lemma 4 in [32]). *Let  $F \subset M$  be a normal surface with more than  $10n$  two-sided components. Then two connected components of  $F$  are normally isotopic to each other.*  $\square$

**THEOREM 4.** *There is a system  $F_1, \dots, F_q$  of normal surfaces such that the components of  $\mathfrak{v}(F_i)$  are bounded from above by  $n \cdot 2^{7n+2}$  for  $i = 1, \dots, q$ , and any normal surface  $F \subset M$  can be written as a sum  $F = \sum_{i=1}^q k_i F_i$  with non-negative integers  $k_1, \dots, k_q$ .*  $\square$

Theorem 4 is proven in [33]. It is based on results on Integer Programming. Weaker versions of this theorem, without explicit bounds for the components of  $\mathfrak{v}(F_i)$ , had been known long before by work of W. Haken [31]. The surfaces  $F_1, \dots, F_q$  are called **fundamental surfaces**.

Let  $F = \sum_{i=1}^k a_i F_i$  be a closed connected normal surface expressed as a sum of fundamental surfaces. Assume that  $\|F\| \leq \|F'\|$  for all normal surfaces  $F'$  that are related with  $F$  by isotopy and Dehn twists along incompressible tori. An important problem is to find an upper bound for  $\|F\|$  in terms of  $n$  and  $\chi(F)$ . We indicate the importance of that problem by two applications.

1. A solution of the problem can be used as part of an algorithm for the classification of Haken manifolds — this is indicated in Hemion's book [35]. A thorough exposition of Haken theory, that also fixes a flaw in the original algorithm, can be found in [68].

2. Waldhausen's conjecture states that any closed orientable 3-manifold only has a finite number of Heegaard splittings of minimal genus, up to isotopy and Dehn twists along incompressible tori. This conjecture was proved in the case of Haken manifolds by Johannson [41]. Only recently, Tao Li [62] gave a proof of Waldhausen's conjecture also for the non-Haken case. The case of non-Haken manifolds could be dealt with based on the theory of *almost normal surfaces*; see [68] for a detailed account. This approach is roughly as follows.

- (1) Any minimal genus Heegaard surface of a non-Haken manifold is *strongly irreducible*, by [19].
- (2) Any strongly irreducible Heegaard surface is isotopic to an almost normal surface.
- (3) Similarly to normal surfaces, almost normal surfaces can be constructed as sums of *fundamental* almost normal surfaces (actually all but one summand is normal).
- (4) A finiteness result for Heegaard surfaces of minimal genus is then obtained by a solution of the above-stated problem.

One can assume that no fundamental surface occurring in the sum is of positive Euler characteristic; see [35]. Since the Euler characteristic is additive, it follows that  $|\chi(F)| \leq a_i |\chi(F_i)|$  for all  $i = 1, \dots, k$ .

Hence, if  $\chi(F_i) < 0$  then there is an upper bound for  $a_i$  in terms of  $\chi(F)$ . Let  $G' = \sum_{\chi(F_i) < 0} a_i F_i$ . Since there is an upper bound for  $\|F_i\|$  in terms of  $n$ , there is an upper bound for  $\|G'\|$  in terms of  $\chi(F)$  and  $n$ . If  $F_i$  is a Klein bottle then  $2F_i$  is a torus, since  $M$  is orientable and thus  $F_i$  is one-sided. Let  $G$  be obtained from  $G'$  by adding one copy of  $F_i$  whenever  $c_i$  is odd. We obtain  $F = G + \sum_{i=1}^N c_i T_i$ , where  $T_1, \dots, T_N$  are orientable normal surfaces of vanishing Euler characteristic which are either fundamental tori or the double of a fundamental Klein bottles. We both have upper bounds for  $\|G\|$  in terms of  $n$  and the genus of  $F$ , and for  $\|T_i\|$  in terms of  $n$ .

In conclusion, there remains to provide an upper bound for  $c_i$  in terms of  $n$ . Hemion [35] suggests the following approach.

STATEMENT 1 (“Unwinding Lemma”). *Let  $G$  be a normal surface and  $T$  a normal torus such that  $G + T$  is defined. If  $k > \#(G \cap T)$  then  $G + kT$  is related with  $G + (k - \#(G \cap T))T$  by isotopy and Dehn twists along  $T$ .*

Hemion leaves the preceding lemma as an exercise; since we are not going to prove it here, we classify it just as a “Statement”, but we refer to it as the “Unwinding Lemma”. We have an upper bound for  $\#(G \cap T_i)$  in terms of  $n$ , and Hemion claims that therefore the Unwinding Lemma allows one to find an upper bound for  $c_i$  ( $i = 1, \dots, N$ ). The idea is to prove that if  $\sum c_i$  is sufficiently large then an application of Lemma 1 is possible. This, again, is left as an exercise by Hemion. We refer to this as the “Unwinding Technique”.

We will not discuss here whether the Unwinding Lemma is true. However, experts are very much in doubt about the correctness of the Unwinding Technique [67], [85]. The aim of this chapter is to study Hemion’s Unwinding Technique by means of *continued fractions*. Although our considerations should certainly increase the doubts about the Unwinding Technique, we can not give an explicit counterexample. In the following section we expose the necessary background on continued fractions. In the last section of this chapter, we expose its connections to the Unwinding Technique.

## 2. Approximation by continued fractions

We recall here some basic facts on continued fractions and refer to [18] for any further details.

For  $a_1, \dots, a_n \in \mathbb{N} \setminus \{0\}$ , the **continued fraction** with coefficients  $a_1, \dots, a_n$  is defined by

$$[a_1, \dots, a_n] = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_{n-1} + \frac{1}{a_n}}}}}$$

Since  $[a_1, \dots, a_{n-1}, 1] = [a_1, \dots, a_{n-1} + 1]$ , we will always assume that the last coefficient of a continued fraction is different from 1.

It is well known that for any infinite sequence  $a_1, a_2, \dots$  of positive integers, the sequence  $([a_1, \dots, a_i])_{i \in \mathbb{N}}$  converges to some irrational number that is denoted by  $[a_1, a_2, \dots]$ . If  $a_1, a_2, \dots$  is a finite sequence then  $[a_1, a_2, \dots]$  is rational. Any real number in the interval  $]0, 1[$  can be expressed in this way. Since in the finite case we assume that the last coefficient is different from 1, the sequence  $a_1, a_2, \dots$  is uniquely determined by  $[a_1, a_2, \dots]$ .

**DEFINITION 2.** For  $x \in \mathbb{R}$  define  $\|x\| = \min_{p \in \mathbb{Z}} |x - p|$  and  $[x] = \max\{k \in \mathbb{Z} : k \leq x\}$ . Let  $\alpha \in ]-1, 1[$  and let  $p, q$  be co-prime integers,  $q > 0$ . The fraction  $\frac{p}{q}$  is called a **best approximation** of  $\alpha$  if  $q|\frac{p}{q} - \alpha| = \|q\alpha\|$  and  $\|q\alpha\| < \|q'\alpha\|$  for all integers  $q'$  with  $0 < q' < q$ .

For the rest of this section, let  $\alpha = [a_1, a_2, \dots]$  and  $i \in \mathbb{N}$ . We implicitly assume that the sequence  $a_1, a_2, \dots$  is long enough so that all terms we consider are defined.

**LEMMA 5** ([18], Section I.2). *Let  $p_i, q_i$  be the unique coprime integers with  $q_i > 0$  and  $\frac{p_i}{q_i} = [a_1, \dots, a_i]$ . Then,  $\frac{p_i}{q_i}$  is a best approximation of  $\alpha$ , and moreover  $\|q_i\alpha\| \leq \|q\alpha\|$  for all integers  $q \in ]0, q_{i+1}[$  (not only for  $0 < q < q_i$ ).*  $\square$

Note that in our use of  $p_i, q_i$  there is an index shift compared with [18]. With the definition  $p_0 = 0, q_0 = 1$  and the observation  $p_1 = 1, q_1 = a_1$ , one obtains  $p_{i+1} = a_{i+1}p_i + p_{i-1}$  and  $q_{i+1} = a_{i+1}q_i + q_{i-1}$  for any  $i \in \mathbb{N}$ .

**PROPOSITION 2** ([18], Section I.2, Equation 15). *If  $a_1, a_2, \dots$  is a sequence of length at least  $i + 2$  of positive integers then*

$$([a_i, a_{i-1}, \dots, a_1] + a_{i+1} + [a_{i+2}, a_{i+3}, \dots]) \cdot \|q_i \cdot [a_1, a_2, \dots]\| = \frac{1}{q_i}$$

*If  $a_1, a_2, \dots$  is of length  $i + 1$  then*

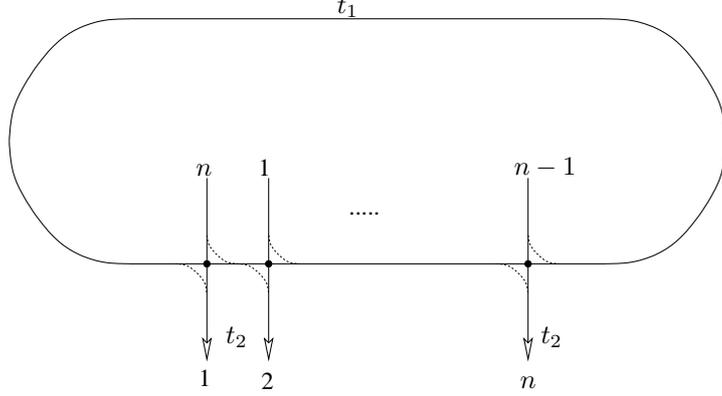
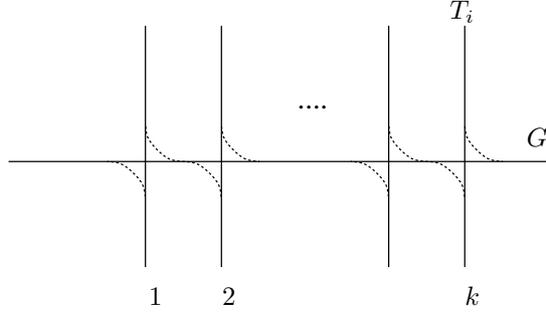
$$([a_i, a_{i-1}, \dots, a_1] + a_{i+1}) \cdot \|q_i \cdot [a_1, a_2, \dots]\| = \frac{1}{q_i} \quad \square$$

### 3. Connections with the Unwinding Technique

We explain here the construction of a triangulated 3-manifold containing some normal surface  $G$  and two normal tori  $T_1, T_2$ , such that for all co-prime  $p, q \in \mathbb{N}$ ,  $pT_1 + qT_2$  is a normal torus  $T_{p,q}$ , and  $G + pT_1 + qT_2$  is a connected normal surface for all  $P, Q \in \mathbb{N}$  (not necessarily co-prime). Figure 3 shows two intersecting circles  $t_1, t_2$  on a surface  $\Sigma$ , where the numbering indicates how various arcs combine to form  $t_2$ . Then,  $T_1, T_2$  are obtained by taking the product of  $t_1, t_2$  with the 1-dimensional sphere  $\mathbb{S}^1$ . The thick dots in the figure become circles that are the connected components of  $T_1 \cap T_2$ , and the dotted lines indicate the switches yielding the sum  $T_1 + T_2$ . We do not prescribe  $n = \#(T_1 \cap T_2)$ , but it is clear that one can make  $\#(T_1 \cap T_2)$  arbitrary large.

It is not difficult to construct some triangulated compact 3-manifold  $M$  containing  $T_1$  and  $T_2$  as two-sided normal surfaces such that  $T_1 + T_2$  is obtained from switching  $T_1 \cup T_2$  along  $T_1 \cap T_2$  in the way that is indicated by the dotted lines close to the thick dots. The switches are chosen so that  $pT_1 + qT_2$  is a normal torus (in particular, connected) for all co-prime  $p, q \in \mathbb{N}$ .

If  $T_1, T_2$  each intersect a normal surface  $G \subset M$  in  $k$  lines such that the regular switches are as indicated in Figure 4, one obtains that  $\#(T_i \cap (G + pT_1 + qT_2)) \geq k - 1$ . Let  $k > 2$ . We can construct  $M$  so that  $G$  as above exists, and so that  $G \setminus (T_1 \cup T_2)$  is connected. Since  $G \cap T_1 \cap T_2 = \emptyset$  and  $T_i \cap (G + pT_1 + qT_2) \neq \emptyset$ ,  $G + pT_1 + qT_2$  is a connected normal surface, for all  $P, Q \in \mathbb{N}$ .

FIGURE 3. Construction of  $T_1$  and  $T_2$ FIGURE 4. A surface  $G$  intersecting  $T_i$ 

For  $i \in \mathbb{N}$ , let  $P_i, Q_i$  be co-prime positive integers with  $\frac{P_i}{Q_i} = \underbrace{[1, \dots, 1, 2]}_{i-1}$ . In other words,  $\frac{P_i}{Q_i} = \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{5}{8}, \dots$ . Note that  $\lim_{i \rightarrow \infty} \frac{P_i}{Q_i} = \frac{1}{2}(\sqrt{5} - 1)$ , the **Golden Section**.

For non-negative co-prime integers  $p \leq P_i, q \leq Q_i$ , let  $k_{p,q}^{(i)}$  be the maximal number of copies of  $T_{p,q}$  that can be build as a sum of at most  $P_i$  copies  $T_1$  and at most  $Q_i$  copies  $T_2$ . In other words,  $k_{p,q}^{(i)} = \min\left(\lfloor \frac{P_i}{p} \rfloor, \lfloor \frac{Q_i}{q} \rfloor\right)$ .

By the following theorem, the number  $k_{p,q}^{(i)}$  of copies of  $T_{p,q}$  is always less than the number of lines of intersection of  $pT_1 \cup qT_2$  with the sum of the remaining copies of  $T_1, T_2$  with  $G$ . Hemion's Unwinding Technique would fail, if one could modify the above construction so that  $\#(T_{p,q} \cap ((P_i - pk_{p,q}^{(i)})T_1 + (Q_i - qk_{p,q}^{(i)})T_2 + G))$  is not too much smaller than  $\#((pT_1 \cup qT_2) \cap ((P_i - pk_{p,q}^{(i)})T_1 + (Q_i - qk_{p,q}^{(i)})T_2 + G))$ , for  $p < P_i, q < Q_i, p, q$  coprime. Unfortunately, the number of lines of intersection can be decreased by a normal isotopy after switching. See, e.g., Figure 5, illustrating the intersection of  $2T_1 + 3T_2$  (thin lines) with  $T_1 + T_2$  (thick lines); by normal isotopy, the number of lines of intersections decreases from 5 to 1. So, we can not prove that our example is a counterexample to the Unwinding Technique, although it indicates that a proper proof of the Unwinding Technique will certainly not be easy to obtain.

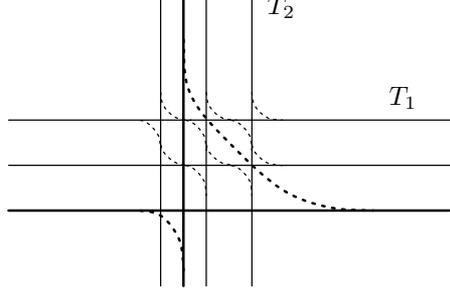


FIGURE 5. Removing intersections by normal isotopy

**THEOREM 5.** *Both  $P_i$  and  $Q_i$  go to infinity for  $i \rightarrow \infty$ . If  $\#(G \cap T_1) \geq 2$ ,  $\#(G \cap T_2) \geq 2$  and  $\#(T_1 \cap T_2) \geq 5$  then for all co-prime integers  $p, q$  with  $0 \leq p \leq P_i$ ,  $0 \leq q \leq Q_i$  holds  $k_{p,q}^{(i)} < \#((pT_1 \cup qT_2) \cap (G + (P_i - k_{p,q}^{(i)})T_1 + (Q_i - k_{p,q}^{(i)})T_2))$*

**PROOF.** It is clear from the definition that  $P_i$  and  $Q_i$  become arbitrarily large. Since the lines of intersection of  $T_1 \cap G$ ,  $T_2 \cap G$  and  $T_1 \cap T_2$  are pairwise disjoint, we obtain

$$\begin{aligned} & \#((pT_1 \cup qT_2) \cap (G + (P_i - k_{p,q}^{(i)})T_1 + (Q_i - k_{p,q}^{(i)})T_2)) \\ &= p \cdot \#(G \cap T_1) + q \cdot \#(G \cap T_2) \\ & \quad + (q(P_i - k_{p,q}^{(i)}) + p(Q_i - k_{p,q}^{(i)})) \cdot \#(T_1 \cap T_2). \end{aligned}$$

We first consider the case  $p = 0$  (or, analogously,  $q = 0$ ). Then,  $q = 1$ ,  $(pT_1 \cup qT_2) = T_2$  and  $k_{p,q}^{(i)} = Q_i$ . It is easy to prove by induction that  $Q_i \leq 3P_i$ . Hence

$$\begin{aligned} & \#((pT_1 \cup qT_2) \cap (G + (P_i - k_{p,q}^{(i)})T_1 + (Q_i - k_{p,q}^{(i)})T_2)) \Big|_{q=1} \\ &= \#(T_2 \cap G) + P_i \#(T_2 \cap T_1) > Q_i = k_{p,1}^{(i)}. \end{aligned}$$

Secondly, we consider the case  $p = P_i$  (or, analogously,  $q = Q_i$ ). Then,  $k_{p,q}^{(i)} = 1$ , and

$$\begin{aligned} & \#((pT_1 \cup qT_2) \cap (G + (P_i - k_{p,q}^{(i)})T_1 + (Q_i - k_{p,q}^{(i)})T_2)) \\ &> \min(\#(T_1 \cap G), \#(T_2 \cap G)) \geq k - 1 \geq 2 > k_{p,q}^{(i)}. \end{aligned}$$

We now consider the remaining case  $0 < p < P_i$ ,  $0 < q < Q_i$ . For  $1 \leq j < i$ , let  $p_j, q_j$  be coprime positive integers with  $\frac{p_i}{q_i} = \underbrace{[1, \dots, 1]}_j$ . Hence,  $\frac{p_i}{q_i} = \frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \dots$

Specifically, we have  $p_{j+1} = q_j$ ,  $p_{j+1} = p_j + p_{j-1} \leq 2p_j$  and  $q_{j+1} = q_j + q_{j-1} \leq 2q_j$ . Moreover,  $P_i = 2p_{i-1} + p_{i-2} \leq 3p_{i-1}$  and  $Q_i = 2q_{i-1} + q_{i-2} \leq 3q_{i-1}$ . We choose  $j$  maximal with  $q_j \leq q$ . It follows  $q < a_{j+1}q_j + q_{j-1} < (a_{j+1} + 1)q_j$ . Define  $\beta$  such that  $\frac{P_i}{Q_i} = \frac{p}{q} + \beta$ .

First, let us assume  $\beta \geq 0$ . Hence,  $k_{p,q}^{(i)} = \lfloor \frac{Q_i}{q} \rfloor$ . By Definition of  $\|\cdot\|$ , by Lemma 5, by Proposition 2, since the coefficients of the continued fraction corresponding to  $\frac{P_i}{Q_i}$  are bounded by 2, and since each continued fraction contributes at

most 1 to the first factor of the left hand side of Proposition 2, we obtain

$$\begin{aligned} q \cdot \left| \frac{P_i}{Q_i} - \frac{p}{q} \right| &\geq \left\| q \frac{P_i}{Q_i} \right\| \\ &\geq \left\| q_j \frac{P_i}{Q_i} \right\| > \frac{1}{4q_j}, \end{aligned}$$

hence  $\beta \geq \frac{1}{4qq_j} \geq \frac{1}{4q^2}$ .

We substitute  $P_i = Q_i \frac{p}{q} + Q_i \beta$  and obtain

$$\begin{aligned} q(P_i - k_{p,q}^{(i)} p) &= q \cdot \left( \frac{p}{q} Q_i + \beta Q_i - \lfloor \frac{Q_i}{q} \rfloor p \right) \\ &\geq \beta q Q_i \geq \frac{Q_i}{4q} \\ &\geq \frac{k_{p,q}^{(i)}}{4} > \frac{k_{p,q}^{(i)}}{\#(T_1 \cap T_2)}, \end{aligned}$$

which proves Theorem 5 in the case  $\beta \geq 0$ .

There remains the case  $\beta < 0$ . The idea is to exchange  $q$  with  $p$ ,  $q_j$  with  $p_j$  and  $Q_i$  with  $P_i$ , which is possible by the following arguments.

If  $\beta < 0$  then we have  $\frac{Q_i}{P_i} = \frac{q}{p} + \beta'$  for some  $\beta' > 0$ . We can estimate  $\beta'$  as follows. By our choice of the coefficients of continued fractions, we have

$$\frac{Q_i}{P_i} = \frac{P_{i+1}}{Q_{i-1}} = \frac{P_i + P_{i-1}}{Q_{i-1}} = 1 + \frac{P_{i-1}}{Q_{i-1}},$$

hence  $\left\| p \frac{Q_i}{P_i} \right\| = \left\| p \frac{P_{i-1}}{Q_{i-1}} \right\|$ . Let  $j$  be maximal such that  $p_j \leq p$ . Hence,  $p_{j+1} = q_j > p$ , and so  $\left\| p \frac{Q_i}{P_i} \right\| \geq \left\| q_j \frac{P_{i-1}}{Q_{i-1}} \right\|$ . As in the case  $\beta \geq 0$ , we obtain  $\left\| q_j \frac{P_{i-1}}{Q_{i-1}} \right\| \geq \frac{1}{4q_j}$  and hence  $\beta' \geq \frac{1}{4pq_j} \geq \frac{1}{4p^2}$ . The rest of the proof is analogous to the case  $\beta \geq 0$ .  $\square$



## Phyllotaxis

### 1. Description of the problem

Phyllotaxis (or *phyllotaxy*) denotes the helical arrangement of plant leaves around the shoot axis. The angle between two adjacent leaves is called **divergence angle**. Phyllotaxis is shown by the vast majority of higher plants, approximately 250,000 species. It is a long-standing observation that in most cases the divergence angle in most cases is  $\frac{p}{q} \cdot 360^\circ$ , where  $\frac{p}{q} = \frac{1}{2}, \frac{1}{3}, \frac{2}{5}, \frac{3}{8}, \dots$ , which is called **Schimper–Braun series**; in the limit, this yields a divergence angle of roughly  $137.508^\circ$ , which is  $[2, 1, 1, 1, \dots] \cdot 360^\circ$ . This angle is related to the Golden Section (we have  $1 - [1, 1, 1, 1, \dots] = [2, 1, 1, 1, 1, \dots]$ ) and is therefore called **Golden Angle**.

Jean [40] estimates that among those plants displaying helical phyllotaxis, the Schimper–Braun series prevails for about 92%. In exceptional cases, there occur divergence angles with  $\frac{p}{q} = \frac{1}{3}, \frac{1}{4}, \frac{2}{7}, \dots$ , sometimes called **Lucas series**, which, in the limit, roughly yields  $99.502^\circ$ . Both Schimper–Braun and Lucas series correspond to continued fractions, namely those of the form  $[2, 1, 1, 1, \dots]$  for the Schimper–Braun and  $[3, 1, 1, 1, \dots]$  for the Lucas series.

Figures 1 and 2 illustrate helical phyllotaxis. In both figures, we have the same 18 leaves, and one looks at the figures parallel to the shoot axis of the plant. The leaves are numbered (starting from the bottom end of the stem), and the angle between successive leaves is the divergence angle. In the figures, the leaflets get smaller with increasing number; this is often the case in nature, as the leaves at the top of the shoot axis are the youngest. In Figure 1, the divergence angle is the Golden Angle. In Figure 2, one has a divergence angle that would not occur in nature, but is close to a divergence angle occurring as one of the first members of the Schimper–Braun or Lucas series, namely  $120^\circ$ .

Even if one accepts the fact that leaf arrangements are helical, these observations need an explanation. There are two aspects that might be particularly worthwhile to explain:

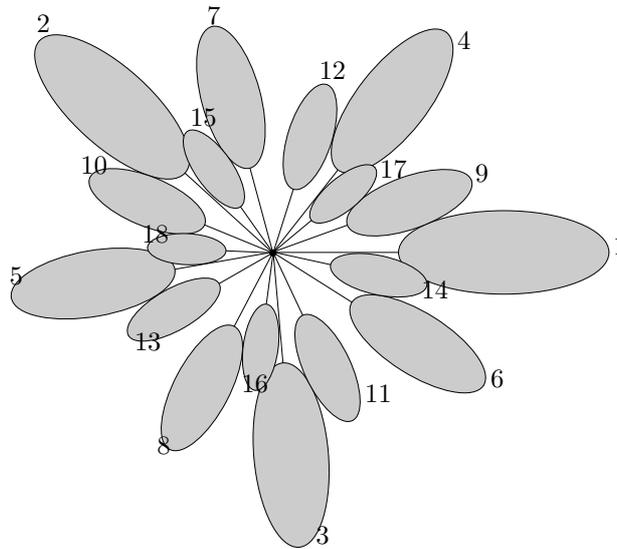
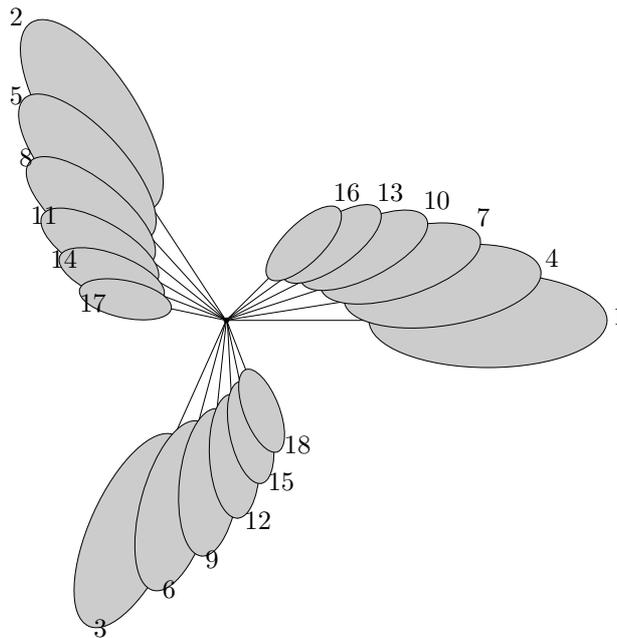
- (1) How does the plant manage to produce the observed divergence angles?
- (2) Why is there a preference in nature for specific divergence angles?

### 2. Various approaches to explain phyllotaxis

The rules of phyllotaxis have been known already since 19th century, by pioneering work of Braun [13], Schimper [86], and Bravais-Bravais [12]. So it is no surprise that there have been numerous proposed explanations; see [40] for an overview.

One approach is the Gierer–Meinhardt activator–inhibitor model [70]. This is a biochemical model that explains how a growing plant can produce a helical leaf arrangement. However, the Gierer–Meinhardt model does not explain the preference for a specific divergence angle, since in principal *any* divergence angle can be produced.

A very elegant approach tries to answer both the “How?” and the “Why?” of phyllotaxis. It is based on the hypothesis that the arrangement of leaflets is a

FIGURE 1. Phyllotaxis with divergence angle  $\approx 137.5^\circ$ FIGURE 2. Phyllotaxis with divergence angle  $123^\circ$ 

consequence of an optimal (i.e., densest) packing of the leaf buds on the growth cone of the shoot axis. It is obvious for economic reasons why there should be a preference for densest packings. In fact, leaf primordia can be found on the growth cone, and their position determines the position of the fully-grown leaves.

However, there are reasons to doubt of that explanation of regular helices. Firstly, we consider here optimal packings of finitely many objects (leaf buds) in a finite area (growth cone). Even if one models the packing of leaf buds on the growth cone by a packing of circles of equal size inside a circle-shaped area, this is a very difficult mathematical problem. Although the optimal packing of infinitely many equally-sized circles in the plane is a regular hexagonal lattice, it can be shown [87]

that finite densest packings have no lattice symmetry, for a large class of notions of “density”. But how should a regular helix emerge from a non-regular densest packing?

Secondly, it is already a considerable simplification to model the leaf buds by circles of equal size. They do not emerge all at the same time, hence we will find leaf buds of different size. But the optimal packings of circles of *different* size in a circle-shaped area are even more irregular than those of equal size. For example, by [26] one should expect that an optimal packing is obtained by placing young (small) leaf buds into the gaps between older (bigger) leaf buds. But this is not the case in nature.

Thirdly, it is questionable if the advantage of saving resources by placing the leaf buds in an optimal way is really significant. In fact, the saved resources seem marginal compared with the resources used for the subsequent growth of the plant.

Another approach, that we study in [53], does not explain how the plant produces a particular divergence angle, but it explains why it is favourable for the plant to have a divergence angle close to  $137.5^\circ$ . The basic idea is to study the light capture of the plant as a function of the divergence angle. If one shows that the optimal light capture is attained for a divergence angle of  $137.5^\circ$  and if other divergence angles lead to a *significantly* reduced light capture, then one has a good reason to expect that the evolution by “survival of the fittest” led to the observed divergence angles.

However, this just explains why *light capturing* plant structures show phyllotaxis. But phyllotaxis also occurs in other structures, such as cones of cycads and conifers, angiosperm flowers, capitulae (Asteraceae), or spines (Cactaceae). But all these structures originally emerged from light-capturing leaves. So, it makes sense to argue that phyllotaxis had evolved under the economic need of optimal light capture, and when the light-capturing leaves evolved into other structures, phyllotaxis has simply been preserved. Note that prickles (Rosaceae) did *not* emerge from leaves; they are outgrowths of epidermis, and in fact they do not show phyllotaxis.

### 3. The light capture of plants

The leaves of a plant cast shadow on each other. Since the main direction of light usually is parallel to the stem, it is obvious from the above figures that with a divergence angle of  $123^\circ$  there is much more shadow than with a divergence angle of  $137.5^\circ$ , since there is much more overlap in Figure 2 than in Figure 1.

This observation was quantified by Percy and Yang [79]. By experiments and very careful simulations based on specimens on the Redwood forest understory plant *Adenocaulon bicolor*, they determined the light capture and even the photosynthetic carbon gain (which is not simply proportional to the light capture) for 17 different divergence angles ranging roughly from  $30^\circ$  to  $175^\circ$ . They found that indeed optimal carbon gain is obtained for divergence angles close to the Golden Angle, and the photosynthetic activity depends on the divergence angle by as much as 30%. They also studied the influence of small perturbations in the leaf arrangements.

The subject of our paper [53] is an a-priori-model for the light capture of plants. Our model only has two parameters and one scaling factor and reproduces the experimental data almost perfectly. Moreover, a simplified version of the model is accessible by the number theoretic methods exposed in Section 3.2. One can in fact prove that the simplified model attains its global optimum if the divergence angle is the Golden Angle  $137.5\dots^\circ = [2, 1, 1, 1, 1, \dots] \cdot 360^\circ$  (Schimper-Braun) and attains a local optimum at  $99.50\dots^\circ = [3, 1, 1, 1, 1, \dots] \cdot 360^\circ$  (Lucas).

Moreover, our model predicts a dependence of the size and of the total number of leaves. According to our model, the global optimum of carbon gain for divergence angles close to  $137.5^\circ$  is most pronounced if the plant has many long and thin leaves. Hence, in the case of narrow leaves the evolutionary pressure towards the Golden Angle should be strongest. In this context, it is interesting to note that Lycopodiophyta, forming the oldest extant vascular plant division (about 420 million years old) are *microphylls*, hence, have narrow leaves.

In the rest of this section, we expose the basic ideas of our model. For details, see [53]. The starting point is the following hypothesis, that appears to be satisfied in nature.

HYPOTHESIS 1.

- (1) Plants grow towards the light, hence, the main direction of light is roughly parallel to the shoot axis.
- (2) Natural light is somehow diffuse.

There are cases in which Hypothesis 1.(1) does not hold. E.g., the branches of the fir grow horizontally. Hence, in this case the main direction of light is perpendicular to the shoot axis. In the first place, the needle leaves of the fir grow according to the rules of phyllotaxis (which might be a heritage of evolution), but then the leaves turn horizontally — which clearly is the best way to capture vertical light and hence is still in accordance to the basic approach.

By symmetry we can assume that the divergence angle  $\alpha$  is at most  $180^\circ$ . Let  $x \in [0, \frac{1}{2}]$ , and  $\alpha = x \cdot 360^\circ$ . Hypothesis 1.(1) implies that leaf number  $n$  casts shadow onto leaf number  $m$  only if  $n > m$ , and the shadow decreases with the angle between these leaves. By symmetry, it suffices to minimize the shadow onto leaf number 0. The angle between leaf number  $n$  and leaf number 0 is  $w_x(n) = |xn - [xn]| \cdot 360^\circ$ , where  $[xn]$  denotes the integer that is closest to  $xn$ .

We denote by  $s_x(n)$  the shadow cast onto leaf number 0 by leaf number  $n$ . By Hypothesis 1.(2), there will always be shadow, even if  $w_x(n) = 180^\circ$ . Hence, at least for sufficiently large angles, it is reasonable to assume  $s_x(n) \propto \frac{1}{w_x(n)}$  in our model.

However, the singularity occurring for  $w_x(n) = 0$  means that  $s_x(n) \propto \frac{1}{w_x(n)}$  is not realistic for small angles. This has two reasons: Firstly, by Hypothesis 1.(2), there is always light coming from different directions (not only parallel to the stem), and so leaf number 0 will catch some light even if  $w_x(n) = 0$ . Secondly, leaves are not completely opaque, and so leaf number  $n$  can not fully block all light. If  $w_x(n)$  is small, then  $s_x(n)$  should mainly depend on the transparency and the size of leaf number  $n$ , and shouldn't much depend on the angle.

Therefore, we introduce a model parameter  $B$ , which we call **effective leaf width**. With some scaling factor  $c$ , we then assume

$$s_x(n) = \frac{c}{\min(B, w_x(n))}$$

The parameter  $B$  is empirical and summarizes many aspects, such as size and transparency of the leaves and the proportion of indirect light.

It seems natural to measure the total shadow on leaf number 0 by  $S(x) = \sum_n s_x(n)$  and the total light capture by  $\frac{1}{S(x)}$ . If one chooses  $B$  and the leaf number in our model in accordance with the leaf size and leaf number of *Adenocaulon bicolor* then it suffices to choose the scaling constant  $c$  by least square fit, and our model can perfectly reproduce (within the error margins) the data of Percy and Yang [79] (see [53] for details).

Hence, our model, that is based on fairly general hypothesis, gives not only a qualitative but even a quantitative description of light capture of plants. Global

and local maxima of light capture in our model are obtained for those divergence angles that are found in nature. Moreover, we find that for small values of the effective leaf width the maxima of light capture are more pronounced. All this gives evidence for the theory that the main reason for phyllotaxis is light capture.

#### 4. A simplified light capture model and the Golden Angle

The light capture model introduced in the preceding subsection seems to be a good description of nature, and can be easily studied numerically using computers. But we also studied a simplified model that shows the same qualitative behaviour as the non-simplified model, and can be investigated using continued fractions.

In our model, we chose  $S(x) = \sum_n s_x(n)$ , and we would like to minimize  $S(x)$ . The basic idea for our simplified model is that the sum will not be very big if all summands are small. Therefore, for our simplified model we chose the shadow function  $S'(x) = \max_n s_x(n)$ , and try to find  $x \in [0, \frac{1}{2}]$  such that  $\max_n s_x(n)$  is minimal.

Of course,  $S(x)$  fits better to the data than  $S'(x)$ . But still, the arguments of the maxima and minima of  $S(x)$  and  $S'(x)$  are roughly the same (see [53]). Minimizing  $S'(x)$  means maximizing  $\frac{1}{S'(x)} = \min_n (\min(B, w_x(n)))$ . As a further simplification, we let  $B$  vanish, and let the number of leaves go to infinity. Then, there remains to determine  $x \in [0, \frac{1}{2}]$  such that

$$\min_{n \in \mathbb{N}} w_x(n) = \min_{n \in \mathbb{N}} |xn - [xn]| = \min_{n \in \mathbb{N}} \|xn\|$$

is maximal. Subsection 2 provides the means to estimate  $\|xn\|$ . It turns out that the maximum is attained for  $x = [2, 1, 1, 1, \dots]$ ; this is part of the Theorem of Hurwitz [81]. But this means that the optimal divergence angle is the Golden Angle.



# The Topological Representation of Oriented Matroids

## 1. Introduction

Oriented matroids are encountered in many fields: Vector configurations or central hyperplane configurations, point sets on a sphere or great hypersphere arrangements, vector spaces or their duals, points on grassmannians, polytopes and their corresponding cellular decompositions in projective space, etc. Accordingly, there is a multitude of definitions for an oriented matroid, reflecting the variety of objects that an oriented matroid can represent. In the research monograph on oriented matroids, [5], three chapters are devoted to axiomatics concerning oriented matroids and to the Topological Representation Theorem for oriented matroids (TRT, for brevity), that we will state below. The TRT is a very useful connection of topological and combinatorial approaches.

The various definitions, via circuit or cocircuit axioms ([5], p. 103), sphere systems ([5], p. 227), Graßmann Plücker relations (chirotope axioms) ([5], p. 126, p. 138, [30]), hull systems ([58]), to mention just a few of them, differ a lot with respect to their motivational aspects, their algorithmic efficiency or their relation to the actual application. Each definition in general provides an additional insight for the motivating problem.

In [7], we treat three aspects of oriented matroid theory. First, we give a new axiomatic related to the concept of hyperline sequences and show that it is equivalent to the axiomatic of chirotopes. Hyperline sequences provide a rather efficient representation of an oriented matroid. Secondly, we present a new proof of the TRT, based on hyperline sequences. Thirdly, we show that one can read off an oriented matroid from arrangements of embedded spheres of codimension one, even if wild spheres are involved. This was proven by Hochstättler [38] by a much longer argument.

The TRT due to Lawrence is central in the theory of oriented matroids. It shows the equivalence of oriented matroids defined via sphere system axioms with oriented matroids defined via covector axioms. This remarkable result asserts that each oriented matroid has a topological representation as an oriented pseudosphere arrangement, even a piecewise-linear one, cf. Edmonds and Mandel [24]. A topological representation of matroids (rather than *oriented* matroids) was found in [89].

Other authors ([5], [38]) have later simplified or complemented the original proof of the TRT, but all use fundamentally the same approach: the face lattice (tope) formalism for oriented matroids and a shelling order to carry through the construction. Finding a reasonably direct proof in the planar case (rank 3) has been posed as an open problem in the research monograph [5] (Exercise 6.3). In [8] such a proof was given, based on hyperline sequences, that are particularly natural in rank 3.

In [7] we generalise this proof to the arbitrary rank case. This generalisation is not straight-forward. First of all, it was needed to give a precise definition of Oriented Matroids in terms of hyperline sequences, and to show that this definition

is equivalent to more classical definitions of Oriented Matroids. Then, the proof of the TRT required a careful use of advanced results from topology, in particular of the generalized Schönflies theorem.

The rest of this chapter is organised as follows. In Section 2.1, we give a geometric motivation for the notion of Oriented Matroids yielding the notion of a hyperline sequence. In Section 2.2 and 2.3, we provide a definition of Oriented Matroids based on hyperline sequences, and relate that definition to our geometric motivation. In Section 2.4 we define pseudosphere arrangements. Our proof of the TRT is outlined in Section 2.5. In the final Section 3, we discuss a tempting approach towards a generalisation of Oriented Matroids.

## 2. Hyperline sequences

**2.1. Geometric motivation.** Before we give a formal definition, we formulate the basic idea. A hyperline sequence represents a rank 2 contraction of an oriented matroid. To motivate the notion of hyperline sequences geometrically, we consider a vector arrangement  $V = \{v_1, \dots, v_n\} \subset \mathbb{R}^r$  of unit vectors that span  $\mathbb{R}^r$ , with  $r \geq 2$ . Let  $B \subset \mathbb{R}^r$  be an oriented subspace of codimension 2 spanned by  $V \cap B$ . We obtain a vector arrangement  $V_B = V \cap B$  in  $\mathbb{R}^{r-2}$ . The orthogonal complement  $C$  of  $B$  is a plane, that is oriented according to the orientation of  $B$  and of  $\mathbb{R}^r$ . The orthogonal projection of  $V \setminus V_B$  to  $C$  is an ordered set of non-zero vectors which give rise to an ordered set  $L_B$  of oriented lines in  $\mathbb{R}^2$ . When we move along a circle in  $C$  around the origin according to the orientation of  $C$ , we meet the elements of  $V_C$  in a circular sequence  $Z_B$ , where any element of  $V_C$  is met twice (in positive and negative orientation). By an inductive definition, the **hyperline sequence** of rank  $r$  associated to  $B$  is the pair  $(Y_B|Z_B)$ , where  $Y_B$  is the oriented matroid of rank  $r - 2$  associated to  $V_B$ . The oriented matroid of rank  $r$  associated to  $V$  is the set of all hyperline sequences that can be read off from  $V$ .

We chose an inductive definition since it naturally fits into the framework of our proof of the TRT (sketched in Subsection 2.5). In fact it is not necessary to keep complete information on the rank  $r - 2$  oriented matroids  $Y_B$ . If one chooses a single positively oriented base  $\sigma_B$  for any hyperline sequence  $(Y_B|Z_B)$  of an oriented matroid  $X$  then  $X$  is determined by the set of pairs  $(\sigma_B|Z_B)$ , see [7]. Hence a direct definition of oriented matroids in terms of hyperline sequences is possible, which might be preferable for algorithmic problems, e.g. the extension of oriented matroids [61].

**2.2. Definition.** Let  $(E, <)$  be a finite totally ordered set. Let  $\overline{E} = \{\overline{e} | e \in E\}$  be a copy of  $E$ . The set  $\mathbf{E}$  of **signed indices** is defined as the disjoint union of  $E$  and  $\overline{E}$ . By extending the map  $e \mapsto \overline{e}$  to  $\overline{e} \mapsto \overline{\overline{e}} = e$  for  $e \in E$ , we get an involution on  $\mathbf{E}$ . We define  $e^* = \overline{e^*} = e$ . For  $X \subset \mathbf{E}$ , define  $\overline{X} = \{\overline{x} | x \in X\}$  and  $X^* = \{x^* | x \in X\}$ .

An **oriented  $d$ -simplex** in  $E$  is a  $(d + 1)$ -tuple  $\sigma = [x_1, \dots, x_{d+1}]$  of elements of  $\mathbf{E}$  such that  $x_1^*, \dots, x_{d+1}^*$  are pairwise distinct. Let an equivalence relation  $\sim$  on oriented  $d$ -simplices in  $E$  be generated by

$$[x_1, \dots, x_{d+1}] \sim [x_1, \dots, x_{i-1}, \overline{x_{i+1}}, x_i, x_{i+2}, \dots, x_{d+1}],$$

for  $i = 1, \dots, d$ . As usual, any oriented  $d$ -simplex is equivalent to one of the form  $[e_1, \dots, e_{d+1}]$  or  $[e_1, \dots, e_d, \overline{e_{d+1}}]$ , with elements  $e_1 < e_2 < \dots < e_{d+1}$  of  $E$ . Define  $-[x_1, \dots, x_{d+1}] = [x_1, \dots, x_d, \overline{x_{d+1}}]$ .

Let  $(E, <)$  be a finite totally ordered set. Let  $\overline{E} = \{\overline{e} | e \in E\}$  be a copy of  $E$ . The set  $\mathbf{E}$  of **signed indices** is defined as the disjoint union of  $E$  and  $\overline{E}$ . By extending the map  $e \mapsto \overline{e}$  to  $\overline{e} \mapsto \overline{\overline{e}} = e$  for  $e \in E$ , we get an involution on  $\mathbf{E}$ . We define  $e^* = \overline{e^*} = e$ . For  $X \subset \mathbf{E}$ , define  $\overline{X} = \{\overline{x} | x \in X\}$  and  $X^* = \{x^* | x \in X\}$ .

An **oriented  $d$ -simplex** in  $E$  is a  $(d+1)$ -tuple  $\sigma = [x_1, \dots, x_{d+1}]$  of elements of  $\mathbf{E}$  such that  $x_1^*, \dots, x_{d+1}^*$  are pairwise distinct. Let an equivalence relation  $\sim$  on oriented  $d$ -simplices in  $E$  be generated by

$$[x_1, \dots, x_{d+1}] \sim [x_1, \dots, x_{i-1}, \overline{x_{i+1}}, x_i, x_{i+2}, \dots, x_{d+1}],$$

for  $i = 1, \dots, d$ . As usual, any oriented  $d$ -simplex is equivalent to one of the form  $[e_1, \dots, e_{d+1}]$  or  $[e_1, \dots, e_d, \overline{e_{d+1}}]$ , with elements  $e_1 < e_2 < \dots < e_{d+1}$  of  $E$ . Define  $-[x_1, \dots, x_{d+1}] = [x_1, \dots, x_d, \overline{x_{d+1}}]$ .

In the following inductive definition of hyperline sequences and oriented matroids, we denote with  $C_m = (\{0, 1, \dots, m-1\}, +)$  the cyclic group of order  $m$ .

**DEFINITION 3 (Rank 1).** An **oriented matroid**  $X$  over  $E(X) \subset E$  of rank 1 is a non-empty subset  $X \subset E(X) \cup \overline{E(X)}$  such that  $|X| = |X^*|$  and  $X^* = E(X)$ .

The oriented simplex  $[x]$  is by definition a **positively oriented base** of  $X$  for any  $x \in X$ . We define  $-X = \overline{X}$ .

**DEFINITION 4 (Rank 2).** Let  $k \in \mathbb{N}$ ,  $k \geq 2$ . A **hyperline sequence**  $X$  of rank 2 over  $E(X) \subset E$  is a map from  $C_{2k}$  to oriented matroids of rank one,  $a \mapsto X^a$ , such that  $X^{a+k} = -X^a$  for all  $a \in C_{2k}$ , and  $E(X) \cup \overline{E(X)}$  is a disjoint union of  $X^0, \dots, X^{2k-1}$ .

An oriented matroid of rank 2 is by definition a hyperline sequence of rank 2. We refer to  $X^0, \dots, X^{2k-1}$  as the **atoms** of  $X$  and to  $2k$  as the **period length** of  $X$ . We say that  $e \in E(X)$  is **incident** to an atom  $X^a$  of  $X$  if  $e \in (X^a)^*$ . Let  $x_1, x_2 \in E(X) \cup \overline{E(X)}$  such that  $x_1^*$  and  $x_2^*$  are not incident to a single atom of  $X$ , and  $X$  induces the cyclic order  $(x_1, x_2, \overline{x_1}, \overline{x_2})$ . Then, the oriented simplex  $[x_1, x_2]$  is by definition a **positively oriented base** of  $X$ . We define the hyperline sequence  $-X$  over  $E(-X) = E(X)$  of rank 2 as the map  $a \mapsto (-X)^a = X^{-a}$  for  $a \in C_{2k}$ .

A hyperline sequence  $X$  of rank 2 is determined by a sequence  $(X^0, \dots, X^{2k-1})$  of atoms. We define that two hyperline sequences  $X_1$  and  $X_2$  of rank 2 are equal,  $X_1 = X_2$ , if  $E(X_1) = E(X_2)$ , the number  $2k$  of atoms coincides, and  $X_1$  is obtained from  $X_2$  by a shift, i.e., there is an  $s \in C_{2k}$  with  $X_1^{a+s} = X_2^a$  for all  $a \in C_{2k}$ .

We prepare the axioms for oriented matroids of rank  $r > 2$  with the following definitions. A **hyperline sequence**  $X$  of rank  $r$  is a pair  $(Y|Z)$ , where  $Y$  is an oriented matroid of rank  $r-2$  and  $Z$  is a hyperline sequence of rank 2. If  $X$  is a set of hyperline sequences of rank  $r$ , a **positively oriented base** of  $X$  in  $(Y|Z) \in X$  is an oriented simplex  $[x_1, \dots, x_r]$  in  $E(X)$ , where  $[x_1, \dots, x_{r-2}]$  is a positively oriented base of  $Y$  and  $[x_{r-1}, x_r]$  is a positively oriented base of  $Z$ . Then,  $-[x_1, \dots, x_r]$  is a negatively oriented base of  $X$ . We define  $-X = \{(Y|-Z) \mid (Y|Z) \in X\}$ . An **atom** of  $X$  in a hyperline sequence  $(Y|Z) \in X$  is the pair  $(Y|Z^a)$ , where  $Z^a$  is an atom of  $Z$ .

**DEFINITION 5 (Rank  $r > 2$ ).** A set  $X$  of hyperline sequences of rank  $r$  is an **oriented matroid** of rank  $r > 2$  over  $E(X) \subset E$  if it satisfies the following axioms.

- (H1)  $E(X)$  is a disjoint union of  $E(Y)$  and  $E(Z)$ , for all  $(Y|Z) \in X$ .
- (H2) Let  $(Y_1|Z_1), (Y_2|Z_2) \in X$  and let  $[x_1, \dots, x_{r-2}]$  be a positively oriented base of  $Y_1$ . If  $\{x_1^*, \dots, x_{r-2}^*\} \subset E(Y_2)$  then  $(Y_1|Z_1) = (Y_2|Z_2)$  or  $(Y_1|Z_1) = (-Y_2|-Z_2)$ .
- (H3) For all positively oriented bases  $[x_1, \dots, x_r]$  and  $[y_1, \dots, y_r]$  of  $X$ , there is some  $j \in \{1, \dots, r\}$  such that  $[x_1, \dots, x_{r-1}, y_j]$  is a positively or negatively oriented base of  $X$ .
- (H4) For any positively oriented base  $[x_1, \dots, x_r]$  of  $X$ ,

$$[x_1, \dots, x_{r-3}, \overline{x_{r-1}}, x_{r-2}, x_r]$$

is a positively oriented base of  $X$ .

At the end of Section 2 in [7], we provide a more direct (i.e., non-inductive) definition of oriented matroids via hyperline sequences. Storing an oriented matroid by hyperline sequences seems more economical than just storing all positively oriented bases, even in the non-uniform case. Moreover, as is clear from the proof of Theorem 6 in [7], the cyclic order of a single hyperline sequence already captures many instances of the Grassmann-Plücker relations. Thus when enumerating oriented matroids, it seems easier to produce a set of hyperline sequences and verify Axioms (H1)–(H4) than to produce a list of oriented simplices and verify, say, the chirotope axioms.

### 2.3. Connection of definition and geometric motivation.

Let  $V = \{v_1, \dots, v_n\} \subset \mathbb{R}^r$ ,  $B \subset \mathbb{R}^r$ ,  $Y_B$ ,  $L_B$  and  $Z_B$  be as in Subsection 2.1. To any  $v_k \in V \setminus Y_B$  we get an oriented line  $l_k \in L_B$ . We move along a circle in the oriented plane and store the letter  $k$  in the circular sequence  $Z_B$  when  $l_k$  is met in positive orientation, and  $\bar{k}$  if  $l_k$  is met in negative orientation. Obviously  $k$  and  $\bar{k}$  appear on opposite places of the circular sequence. Hence  $Z_B$  is a hyperline sequence of rank 2. By induction and abuse of notation, the vector arrangement  $Y_B$  “is” an oriented matroid  $Y_B$ , and  $(Y_B|Z_B)$  is a hyperline sequence of rank  $r$ . Axiom (H1) means that  $V$  is a disjoint union of  $V \cap B$  and  $V \setminus B$ . Axiom (H2) corresponds to the fact that  $B$  is determined by any oriented base of  $V_B$ . Axiom (H3) is the Steinitz–McLane exchange lemma, stating that one can replace any vector in a base by some vector of any other base. Axiom (H4) ensures that the definition of oriented bases is compatible with the equivalence relation on oriented simplices; this is part of Theorem 6 below. Axiom (H4) is related to the “consistent abstract sign of determinant” in [8]. It means that if  $r$  vectors span an  $(r-1)$ -simplex, then any subset of  $r-2$  vectors spans a hyperline, and the orientation of the  $(r-1)$ -simplex does not depend on the hyperline on which we consider the  $r$  points. A hyperline sequence stores information on a rank 2 contractions of an oriented matroid.

A classical way to define oriented matroids is via **chirotope** axioms (see [5], p. 126, p. 138, [30]). In fact chirotopes and hyperline sequences yield equivalent notions of oriented matroids, as in the following theorem. This connects our concept of hyperline sequences with other ways to look at oriented matroids.

**THEOREM 6** (Theorem 1 in [7]). *The set of positively oriented bases of an oriented matroid of rank  $r$  over  $E$  given by hyperline sequences is the set of positively oriented bases of a chirotope of rank  $r$  over  $E$ , and vice versa.*  $\square$

The cyclic structure of a hyperline sequence captures many instances of the 3-term Graßmann–Plücker relations at once. Certainly it is easier to deal with a few cyclic structures than with a multitude of Graßmann–Plücker relations, specifically in algorithmic applications. There is a price to pay for the simplification in the representation of oriented matroids: the proof of the preceding theorem becomes rather long and tedious if it is carried out in detail, and also it seems impossible to give a brief resume of it. So we simply refer to [7].

**2.4. Arrangements of oriented pseudospheres.** A submanifold  $N$  of codimension  $m$  in a  $d$ -dimensional manifold  $M$  is **tame** if any  $x \in N$  has an open neighborhood  $U(x) \subset M$  such that there is a homeomorphism  $\overline{U(x)} \rightarrow B^d$  sending  $U(x) \cap N$  to  $B^{d-m} \subset B^d$ .

An **oriented pseudosphere**  $S \subset \mathbb{S}^d$  is a tame embedded  $(d-1)$ -dimensional sphere with a choice of an orientation. Any oriented hypersphere is an oriented pseudosphere. The image of an oriented pseudosphere  $S$  under a homeomorphism  $\phi: \mathbb{S}^d \rightarrow \mathbb{S}^d$  obviously is an oriented pseudosphere as well.

The **generalized Schönflies theorem** assures that if  $S$  is a pseudosphere in  $\mathbb{S}^d$  then there is a homeomorphism  $\phi: \mathbb{S}^d \rightarrow \mathbb{S}^d$  such that  $\phi(S) = \mathbb{S}^{d-1} \subset \mathbb{S}^d$ . It

was proven by M. Brown [15]. Similarly, if  $B \subset B^d$  is a tame embedded ball of codimension 1 in the  $d$ -dimensional unit ball  $B^d \subset \mathbb{R}^3$  and  $\partial B \subset \partial B^d$  then there is a homeomorphism  $B^d \rightarrow B^d$  sending  $B$  to  $B^{d-1} \subset B^d$ . It is well known that in dimension  $d \geq 3$  there are embedded spheres of codimension 1 whose complement is not formed by two balls, so-called wild spheres.

Let  $\psi: \mathbb{S}^{d-1} \rightarrow \mathbb{S}^d$  be an embedding with image  $S$ , inducing the correct orientation on  $S$ . By another result of M. Brown [16], the image of  $\psi$  is tame if and only if  $\psi$  can be extended to an orientation preserving embedding

$$\tilde{\psi}: \mathbb{S}^{d-1} \times [-1, 1] \rightarrow \mathbb{S}^d \text{ with } \psi(\cdot) = \tilde{\psi}(\cdot, 0).$$

We call the connected component of  $\mathbb{S}^d \setminus S$  containing  $\tilde{\psi}(\mathbb{S}^{d-1} \times \{1\})$  (resp.  $\tilde{\psi}(\mathbb{S}^{d-1} \times \{-1\})$ ) the **positive side**  $S^+$  (resp. **negative side**  $S^-$ ) of  $S$ .

Arrangements of oriented pseudospheres are defined in [5], p. 227. Recall  $E_n = \{1, \dots, n\}$ . By Theorem 3 in [7], one can equivalently define arrangements of oriented pseudospheres as an ordered multiset  $\mathcal{A} = \{S_1, \dots, S_n\}$  of oriented pseudospheres in  $\mathbb{S}^d$  such that any ‘‘small enough’’ subarrangement is equivalent to an arrangement of oriented hyperspheres. Formally, for  $R \subset E_n$ , denote  $\mathcal{A}_R = \{S_j | j \in R\}$ . We obtain that  $\mathcal{A}$  is an arrangement of oriented pseudospheres in the sense of [5] if and only if the following holds: Let  $R \subset E_n$  such that  $S_{R'} \neq S_R$  for any proper subset  $R'$  of  $R$  (i.e.,  $R$  is an independent set); then,  $\mathcal{A}_R$  is equivalent to an arrangement of  $|R|$  oriented hyperspheres in  $\mathbb{S}^d$ . An arrangement of oriented pseudospheres is called **of full rank** if the intersection of its members is empty.

It turns out that any arrangement  $\mathcal{A}$  of oriented pseudospheres in  $\mathbb{S}^d$  of full rank yields a cellular decomposition  $\mathcal{C}(\mathcal{A})$  of  $\mathbb{S}^d$  that is non-degenerate in the sense that the closure of any open cell of  $\mathcal{C}(\mathcal{A})$  in  $\mathbb{S}^d$  is a closed cell.

Any cell  $c$  of  $\mathcal{C}(\mathcal{A})$  is provided with its index  $I \subset E_n$ , which is inclusion maximal with the property  $c \subset S_I$ . Let  $\mathcal{C}(I, \mathcal{A})$  be the union of all open cells of  $\mathcal{C}(\mathcal{A})$  with index  $I$ .

**DEFINITION 6.** Two ordered multisets  $\{S_1, \dots, S_n\}$  and  $\{S'_1, \dots, S'_n\}$  of oriented pseudospheres in  $\mathbb{S}^d$  are **equivalent** if there is an orientation preserving homeomorphism  $\mathbb{S}^d \rightarrow \mathbb{S}^d$  sending  $S_i^+$  to  $(S'_i)^+$  and  $S_i^-$  to  $(S'_i)^-$ , simultaneously for all  $i \in E_n$ . We do not allow renumbering of the pseudospheres. We do distinguish the two orientations of  $\mathbb{S}^d$ .

**2.5. The Topological Representation Theorem.** Let  $\mathcal{A} = \{S_1, \dots, S_n\}$  an arrangement of oriented pseudospheres in  $\mathbb{S}^d$  of full rank. Let  $R \subset E_n$  be inclusion minimal with the property that  $S_R = \bigcap_{i \in R} S_i$  is a circle. The orientations of the pseudospheres and the order of  $E_n$  yields a circular orientation for  $S_R$ . Let  $R' \subset E_n$  be the inclusion maximal set with  $S_{R'} = S_R$ . On  $S_R$  we transversally intersect all pseudospheres  $S_j$  with  $j \in E_n \setminus R'$ , in cyclic order. In fact, each of them is met twice: One time we pass from  $S_j^-$  to  $S_j^+$ , the other time the other way around. This yields a cyclic sequence of signed indices. Moreover, one obtains an oriented matroid on  $R'$ . When we do this for all appropriate  $R \subset E_n$ , we obtain a set of hyperline sequences satisfying the axioms of oriented matroids provided in Section 2.2. We denote this oriented matroid by  $X(\mathcal{A})$ ; see [7, Section 6] for details.

Both for oriented matroids  $X$  and for arrangements  $\mathcal{A}$  of oriented pseudospheres, we have two operations called **deletion**  $X \setminus R$  (resp.  $\mathcal{A} \setminus R$ ) of a subset  $R \subset E_n$  and **contraction**  $X/R$  (resp.  $\mathcal{A}/R$ ) on a subset  $R \subset E_n$ . These are standard notions, and we give no definition here, but refer to [7].

**THEOREM 7 (Topological Representation Theorem).** *To any oriented matroid  $X$  of rank  $r$  over  $E_n$ , there is an arrangement  $\mathcal{A}(X)$  of  $n$  oriented pseudo hyperspheres in  $\mathbb{S}^{r-1}$  of full rank with  $X = X(\mathcal{A}(X))$ . The equivalence class of  $\mathcal{A}(X)$  is unique.*

In [7] we prove Theorem 7 by induction on the number of elements and the rank of  $X$ . The base cases  $r \leq 2$  and  $n = r$  are rather easy. We outline here the proof of the induction step. Let  $n > r > 2$ . Suppose that Theorem 7 holds for all oriented matroids of rank  $r$  with less than  $n$  elements and for all oriented matroids of rank less than  $r$ .

Thus, for any non-empty  $R \subset E_n$  for which the contraction  $X/R$  (resp. the deletion  $X \setminus R$ ) is defined, there is an essentially unique arrangement  $\mathcal{A}(X/R)$  (resp.  $\mathcal{A}(X \setminus R)$ ) of oriented pseudospheres in  $\mathbb{S}^{r-1-|R|}$  (resp. in  $\mathbb{S}^{r-1}$ ) of full rank with  $X/R = X(\mathcal{A}(X/R))$  (resp. with  $X \setminus R = X(\mathcal{A}(X \setminus R))$ ).

There is some element of  $X$ , say,  $n$  for simplicity, such that the deletion  $X \setminus \{n\}$  is an oriented matroid of rank  $r$ . Denote  $\{S_1, \dots, S_{n-1}\} = \mathcal{A}(X \setminus \{n\})$ . Our aim is to construct an oriented pseudosphere  $S_n \subset \mathbb{S}^{r-1}$  as the image of a tame embedding  $\psi: \mathbb{S}^{r-2} \rightarrow \mathbb{S}^{r-1}$ , so that  $\{S_1, \dots, S_n\}$  is an arrangement of oriented pseudospheres with  $X(\{S_1, \dots, S_n\}) = X$ .

The construction of  $\psi$  is roughly as follows. We start with the arrangement  $\mathcal{A}(X \setminus \{n\})$  in  $\mathbb{S}^{r-2}$ . We require that  $\psi$  maps this arrangement “consistently” to the arrangement  $\mathcal{A}(X \setminus \{n\})$ , in the sense that any cell in  $\mathcal{C}(I, \mathcal{A}(X \setminus \{n\}))$  is mapped to a cell in  $\mathcal{C}(I, \mathcal{A}(X \setminus \{n\}))$  in the correct orientation. It turns out that this forces  $\{S_1, \dots, S_n\}$  to be an arrangement of oriented pseudospheres. Moreover, we show that if  $S_n$  intersects the cycles of  $\mathcal{A}(X \setminus \{n\})$  in a way consistent with the rank 2 contractions of  $X$  (i.e., the cyclic order on its hyperline sequences), then  $X(\{S_1, \dots, S_n\}) = X$ . Our construction of  $\psi$  is iterative. We start with defining  $\psi$  on 0-dimensional cells of  $\mathcal{A}(X \setminus \{n\})$  and show that if it is defined on  $d$ -dimensional cells then it can be consistently extended to  $(d+1)$ -dimensional cells. It turns out that this is possible in an essentially unique way.

A formalisation of this idea is given in [7].

### 3. A potential generalisation

By the Topological Representation Theorem, there is a bijection between Oriented Matroids and equivalence classes of arrangements of oriented pseudospheres, i.e., arrangements of oriented tame sub-spheres of co-dimension one embedded in spheres. The topology admits a natural generalisation: Arrangements of oriented tame sub-manifolds of co-dimension one in compact oriented manifolds. What might be a good notion of *arrangement* in this context, and how could one try to generalise the hyperline axioms in order to get a notion of Generalised Oriented Matroids corresponding to sub-manifolds arrangements? In this section, we will discuss that question, although we can not give a good definition of Generalised Oriented Matroids.

Of course, a sub-manifold *arrangement*  $\mathcal{A}$  should be a system  $\{F_i: i \in E_n\}$  of tame two-sided co-dimension one sub-manifolds in a compact manifold  $M$ . As for pseudosphere arrangements, there should be a notion of contraction onto one of the sub-manifolds. Hence, for appropriate  $R \subset E_n$ , one should obtain a sub-manifold arrangement  $\mathcal{A}/R$  on  $F_R = \bigcap_{i \in R} F_i$  formed by  $\{F_i \cap F_R: i \in E_n \text{ with } F_i \not\supset F_R\}$ . In particular, for any two different sub-manifolds  $\mathcal{A} \ni F_1, F_2 \subset M$ , the intersection  $F_1 \cap F_2$  should be a tame two-sided co-dimension one sub-manifold of  $F_1$  and of  $F_2$ , and all  $F_R$  are either connected or formed by exactly two points.

It might be worth-while to take non-orientable manifolds into account. In fact, any Oriented Matroid can also be realised by an arrangement of tame co-dimension one *projective spaces*. Here, if  $F_R$  is of dimension 0 then it is exactly one (not two) points, and if  $F_R$  is a circle then one obtains a hyperline sequence by running twice along it.

On the part of Oriented Matroids, modifications of Axiom (H4) appear to be a natural source for generalisations. Figure 1 illustrates the geometric meaning of Axiom (H4): A set of three hyperline sequences satisfying Axioms (H1), (H2)

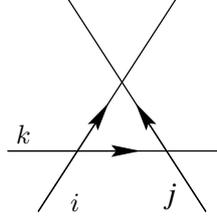


FIGURE 1. Geometric meaning of Axiom (H4)

and (H3) can be realised by an arrangement of circles in  $\mathbb{S}^2$  if and only if it satisfies Axiom (H4). Namely, the orientation of the intersection of  $k$  with  $i$  and  $j$  and the order of these two points of intersection on  $k$  determine the orientation of the intersection of  $i$  with  $j$ .

An appropriate weakening of Axiom (H4) would still imply that the three hyperline sequences can be realised by an arrangement of circles in an *oriented* surfaces. By work of Bokowski and Pisanski [9], Oriented Matroids of rank 3 can indeed be generalised so that one obtains a bijection to equivalence classes of curve arrangements in surfaces. However, there seem to be obstructions to lift this approach to arbitrary rank, as we point out in the remainder of this section.

Let  $\mathcal{A} = \{F_i : i \in E_n\}$  be sub-manifold arrangement in a closed  $d$ -dimensional manifold  $M$ , and assume  $F_{E_n} = \emptyset$  (in the setting of pseudosphere arrangements, this would mean that  $\mathcal{A}$  is of full rank). Without additional assumptions,  $M$  is not uniquely determined by the combinatorics of  $\mathcal{A}$ . Indeed, if  $N$  is any closed  $d$ -dimensional manifold, then  $\mathcal{A}$  could also be realised in the connected sum  $M \# N$ , simply by replacing a ball in  $M \setminus \bigcup_{i \in E_n} F_i$  by  $N$  minus a ball. So, it seems reasonable to additionally assume that  $M \setminus \bigcup_{i \in E_n} F_i$  is a disjoint union of balls —that hypothesis would exclude to plug  $N \neq \mathbb{S}^d$  into  $M \setminus \bigcup_{i \in E_n} F_i$ .

Additionally, this hypothesis would help to generalise our proof of the Topological Representation Theorem: Our proof depends on the generalised Schönflies theorem, since this implies that for  $i \in E_n$  there is a unique way to plug in a  $(d-1)$ -dimensional cell of the arrangement  $\mathcal{A}/\{i\}$  into a  $d$ -dimensional cell of the arrangement  $\mathcal{A} \setminus \{i\} = \{F_k : k \in E_n \setminus \{i\}\}$ . So, if we are not dealing with cells, an application of the generalised Schönflies theorem is impossible, and a generalisation of our proof of the Topological Representation Theorem would fail.

However, this hypothesis would make it difficult to maintain the notion of *deletion* of a sub-manifold in an arrangement: It may be that  $M \setminus \bigcup_{k \in E_n, k \neq i} F_k$  is not a disjoint union of cells, although  $M \setminus \bigcup_{k \in E_n} F_k$  is, namely if there is a  $d$ -cell  $c$  in  $M \setminus \bigcup_{k \in E_n} F_k$  that touches both sides of a  $(d-1)$ -cell  $c_i$  in  $F_i \setminus \bigcup_{k \in E_n, k \neq i} F_k$ . For  $d = 2$ , this would mean that  $c \cup c_i$  is an annulus or a Möbius strip.

A possible solution in the case  $d = 2$  is to replace  $c \cup c_i$  by one or two disks along  $\partial(c \cup c_i)$ , changing  $M$  into a different surface  $\tilde{M}$ . So,  $\mathcal{A} \setminus \{i\}$  would be defined in  $\tilde{M}$  rather than in  $M$ .

Unfortunately, that solution does not apply to the case  $d > 3$ . Then,  $c \cup c_i$  would be a solid torus or the twisted product of  $\mathbb{S}^1$  with a disk. We can not alter  $M$  by replacing  $c \cup c_i$  with cells in a different manifold, as  $\partial(c \cup c_i)$  is not a disjoint union of spheres.

So, for  $d > 2$ , either one has to drop the hypothesis that arrangements of full rank define a cellular decomposition (thereby losing the uniqueness of  $M$ ), or one

has to do without the notion of “deletion of a sub-manifold”. The second alternative would destroy our inductive proof of the Topological Representation Theorem.

The first alternative might even be worse, for a different reason: For an inductive construction of a topological realisation of a generalised Oriented Matroid  $X$  (whatever this should be) of rank  $d + 1$  by an arrangement  $\mathcal{A}$  of co-dimension one sub-manifolds in a  $d$ -manifold  $M$ , one would first construct the realisations  $\mathcal{A}/\{i\}$  of  $X/\{i\}$ . These would be  $(d-1)$ -manifolds  $F_i$  with an arrangement of co-dimension one sub-manifolds that decompose  $F_i$  into parts that are not supposed to be cells. Some parts of  $F_1, \dots, F_n$  must fit together in order to form a closed  $(d-1)$ -manifold  $B$  that is the boundary of a component of  $M \setminus \bigcup_{k \in E_n} F_k$ , and the combinatorics of  $X$  would tell us what parts of  $F_1, \dots, F_n$  we need to take for forming  $B$ .

By a theorem of V.A. Rokhlin, there are closed 4-manifolds that are not null-bordant, i.e., not homeomorphic to the boundary of any compact 5-dimensional manifold (see, e.g. [90]). Hence, if  $d = 5$ , it may be that the combinatorics of  $X$  forces us to combine parts of  $F_1, \dots, F_n$  yielding a 4-manifold  $B$  that is not the boundary of a 5-manifold. Hence, in that situation, there would not be a 5-manifold  $M$  containing  $F_1, \dots, F_n$  in a way that is compatible with  $X$ .

Hence, if one wants to have a Topological Representation Theorem for Generalised Oriented Matroids  $X$  of rank  $r > 6$ , one can not simply drop Axiom (H4). One needs to introduce other axioms excluding that the parts occurring in rank  $r-1$  contractions of  $X$  form a manifold that is not null-bordant. Such axioms seem to be out of reach.

## Ideal state sum invariants

### 1. Introduction

State sum invariants play a prominent rôle in modern 3-dimensional topology. In particular the quantum invariants are to mention. The interest in quantum invariants was first raised by the **Jones polynomial** of knots and links. It was the first invariant that was both easy to compute and powerful enough to detect chiral knots (i.e., knots that are inequivalent to their mirror images). Still, there remain important unsolved questions on the Jones polynomial. E.g., it is still unknown whether there is a non-trivial knot with trivial Jones polynomial; and there is the famous Volume Conjecture originally formulated by Kashaev [43], stating a relation between the Jones polynomial of a hyperbolic knot and the hyperbolic volume of the knot complement.

Originally, the Jones polynomial was inspired by certain physical state sum models for ferro-magnetism [42]. Later, it was discovered that the Jones polynomial is related to the quantum group  $U_q(\mathfrak{sl}_2)$ , and is but one example of a large class of knot invariants. The study of quantum groups also led to the construction of new homeomorphism invariants of low-dimensional manifolds, specifically the Reshetikhin–Turaev and the Turaev–Viro invariants of closed 3-dimensional manifolds.

The Turaev–Viro invariant of a closed oriented 3-manifold equals the product of its Reshetikhin–Turaev invariant and the Reshetikhin–Turaev invariant of the oppositely oriented manifold [97]; so both invariants are closely related. However, their definitions look fairly different, at the first glance: Reshetikhin–Turaev invariants are based on the Kirby calculus (i.e., the representation of closed 3-manifolds as the boundary of 4-manifolds that are represented by “decorated” links in  $\mathbb{S}^3$ ). But the Turaev–Viro invariants were defined in terms of triangulations of 3-manifolds: A certain polynomial (the *state sum*) is read off from the triangulation, the polynomial is evaluated in a specific way, and the result of evaluation is a homeomorphism invariant of the manifold, which has been proven using Pachner moves. Turaev–Viro invariants are very powerful and interesting, both in theoretical studies and in the practical algorithmic classification of closed orientable irreducible 3-manifolds, which now is done for manifolds having special spines with up to 12 vertices.

The 2-skeleton of the dual cellular decomposition of a triangulated 3-manifold  $(M, \mathcal{T})$  is a special spine of  $M \setminus U(\mathcal{T}^0)$ . So, it makes sense to define Turaev–Viro invariants in terms of special spines rather than of triangulations. We believe that this point of view has several advantages. Firstly, for proving the invariance one just needs to consider *one* local transformation of special spines, while one needs *two* types of Pachner moves for triangulations [78]. Secondly, since the invariance conditions are relaxed, it seems possible that the class of Turaev–Viro invariants defined via special spines provide a *proper* generalisation of the class of Turaev–Viro invariants defined via triangulations. Thirdly, one could change the focus by extending the definition of Turaev–Viro invariants to *all* special 2-polyhedra, rather

than only to those that occur as special spines of 3-manifolds. In that way, one can try to obtain invariants for the *Andrews–Curtis problem*.

One formulation of the Andrews–Curtis conjecture states that any two contractible special 2-polyhedra are related with each other by certain types of local transformations (called 3-deformations). 3-deformations do not change the Euler characteristic of a special 2-polyhedron. It is known that 3-deformations together with *stabilisation* (i.e., connected sum with  $S^2$ ) suffice to relate any two contractible special 2-polyhedron with each other. The Andrews–Curtis conjecture is about 40 years old, and nearly as old are potential counterexamples to the conjecture. However, so far there is no proof that they actually *are* counterexamples. A proof might be found using invariants of special 2-polyhedra (e.g., generalised Turaev–Viro invariants). This must be an invariant under 3-deformations that is not invariant under stabilisation.

Modifications of Turaev–Viro invariants associated to  $U_q(\mathfrak{sl}_2)$  provide such invariants. However, there is a problem that holds for a more general class of invariants, namely *reductions of modular invariants*. These are invariants that take values in commutative rings and satisfy certain conditions, among them the *multiplicativity* under connected sums: If  $P_1$  and  $P_2$  are special 2-polyhedra, then a modular invariant  $|\cdot|$  satisfies  $|P_1\#P_2| = |P_1||P_2|$ . So, under stabilisation one obtains  $|P\#S^2| = |P||S^2|$  for any special 2-polyhedron  $P$ . Note that the notation  $|S^2|$  is a bit sloppy:  $|\cdot|$  is only defined for special 2-polyhedra, and so  $S^2$  stands for some special 2-polyhedron that is a special spine of  $S^2 \times [0, 1]$ . Any two contractible special 2-polyhedra  $P_1$  and  $P_2$  are related by Andrews–Curtis moves and stabilisations. Hence, for some  $n \in \mathbb{N}$ , we have

$$|P_1||S^2|^n = |P_1\#^n S^2| = |P_2\#^n S^2| = |P_2||S^2|^n.$$

Therefore, if  $|S^2|$  is invertible then  $|P_1| = |P_2|$ , and one can not detect if  $P_1, P_2$  are counterexamples for the Andrews–Curtis conjecture. One may hope that the situation is better for those invariants for which  $|S^2|$  is not invertible. But a Theorem of I. Bobtcheva and F. Quinn [6] states that, even in this case, the value of a reduction of a modular invariant of a special 2-polyhedron  $P$  is completely determined by the homology of  $P$ , provided the Euler characteristic of  $P$  is at least 1.

We succeeded to generalise Turaev–Viro invariants so that we obtain *non-multiplicative* invariants for Andrews–Curtis moves. So, in principle, we provide a tool to detect counterexamples for the Andrews–Curtis conjecture; see Section 3. However, our method had no success for several notorious conjectured counterexamples. We also provide generalised Turaev–Viro invariants for compact 3-manifolds. They are computed using computer algebra. It turns out that these so-called *ideal Turaev–Viro invariants* are considerably stronger than those obtained from quantum groups; see Section 2. Similarly, we tried to generalise state sum invariants for knots and links. Unfortunately, we did not find anything stronger than the Jones polynomial; see Section 4 for a very short account.

Ideal Turaev–Viro invariants are defined via state sum polynomials in some polynomial ring. This ring is provided with the action of some symmetric group, and the state sum polynomial actually belongs to the invariant ring. When we tried to compute the invariant ring, we encountered problems that have been unsolvable with existing software. This motivated us to study algorithms for the computation of invariant rings, and we came up with a new algorithm that provides a dramatic improvement. Our work on the computation of invariant rings is exposed in Chapter 7.

## 2. Ideal Turaev-Viro invariants of 3-manifolds

This section is a summary of [55] and [54].

**2.1. The basic idea.** Let  $M$  be compact 3-manifolds, represented by a special spine  $P$  (see Definition 7). The Turaev–Viro invariants of  $M$ , originally formulated for triangulations rather than special spines [96], can be read off from  $P$ : One computes the state sum, i.e., a polynomial whose summands correspond to different “colourings” of the  $P$ . The state sum polynomial depends not only on the set of colours, but also on the choice of  $P$ . However, when the state sum is evaluated at a solution of the so-called Biedenharn–Elliott equations known from quantum physics [60], one obtains a homeomorphism invariant of  $M$ ; this result is due to the fact that any two special spines of  $M$  are related by a finite sequence of certain local transformations (see Theorem 8 below). We call this homeomorphism invariant a “numerical Turaev–Viro invariant”.

It is difficult to find solutions of the Biedenharn–Elliott equations, but an important class of solutions is provided by the representation theory of Quantum Groups [97]. The invariants obtained in that way are rather strong and are an important tool in the census of closed orientable irreducible 3-manifolds pursued by different research groups.

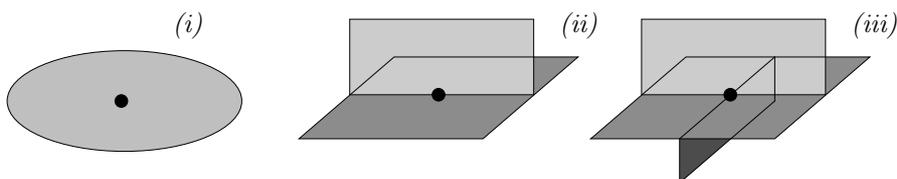
The Biedenharn–Elliott equations generate an ideal (the so-called *Turaev–Viro ideal*) in some polynomial ring. The ring and the ideal only depend on the set of colours. Our starting point is the observation that the coset of the state sum of  $P$  with respect to the Turaev–Viro ideal is a homeomorphism invariant of  $M$ . We call this an “ideal Turaev–Viro invariant”. Obviously any numerical Turaev–Viro invariant is obtained by evaluation of some ideal Turaev–Viro invariant, and one may hope for a proper generalisation of the numerical Turaev–Viro invariants. Indeed, we show that ideal Turaev–Viro invariants are considerably stronger than the numerical Turaev–Viro invariants associated to the quantum group  $U_q(\mathfrak{sl}_2)$ .

In the next subsection, we outline the definition of ideal Turaev–Viro invariants and recall some properties of the numerical Turaev–Viro invariants associated with  $U_q(\mathfrak{sl}_2)$ . Since ideal Turaev–Viro invariants are cosets with respect to ideals in a polynomial ring, we need the possibility to compare those cosets. This is possible by the theory of Gröbner bases, and we recall the necessary results in Subsection 2.3. In Subsection 2.4, introduce some simplifying assumptions that help to construct the basic data of interesting ideal Turaev–Viro invariants. We made some computation of ideal Turaev–Viro invariants for closed orientable irreducible 3-manifolds that possess a special spine with at most 9 vertices. Our computational results are presented in Subsection 2.5

### 2.2. Definition of ideal Turaev–Viro invariants.

#### 2.2.1. Special Spines.

**DEFINITION 7.** A **simple 2-polyhedron**  $P$  is a compact connected hausdorff space such that any point has an open neighbourhood of one of the following three homeomorphism types (where the point under consideration is marked by a thick dot):



The connected components of points of type (i) are the **2-strata** of  $P$ , the connected components of points of type (ii) are the **true edges** of  $P$ , and the points of type (iii) are the **true vertices** of  $P$ . The set of 2-strata of  $P$  is denoted by  $\mathcal{C}(P)$ , the set of true edges of  $P$  is denoted by  $\mathcal{E}(P)$ , and the set of true vertices of  $P$  is denoted by  $\mathcal{V}(P)$ . The true edges and true vertices of  $P$  form the **singular graph**  $S(P)$  of  $P$ .

A simple 2-polyhedron is special, if it has a true vertex, its singular graph is connected, and its 2-strata are homeomorphic to open discs. Let  $M$  be a compact 3-manifold. A special 2-polyhedron  $P$  embedded in  $M$  is a **special spine** of  $M$ , if  $\partial M = \emptyset$  and  $M \setminus P$  is homeomorphic to a 3-ball, or if  $\partial M \neq \emptyset$  and  $M \setminus P \approx (\partial M) \times [0, 1)$  (where  $[0, 1)$  denotes a half-open interval).

A general reference for the theory of special spines of compact 3-manifolds is [68]. Any compact 3-manifold has a special spine, which can be deduced from the fact that any compact 3-manifold admits a triangulation [73]. Moreover, the homeomorphism type of a special spine uniquely determines the homeomorphism type of the 3-manifold [17]. The following classical result explains how all special spines of a compact 3-manifold are related with each other.

**THEOREM 8** (Matveev [64], Piergallini [80]). *Let  $M$  be a compact 3-manifold with special spines  $P_1, P_2$ , and assume that  $P_1$  and  $P_2$  both have at least two true vertices. Then  $P_1$  and  $P_2$  are related by a finite sequence of a local transformation called  $T$  move and its inverse. The  $T$  move is shown in Figure 1, where true vertices are marked by a thick dot and true edges are drawn bold.*

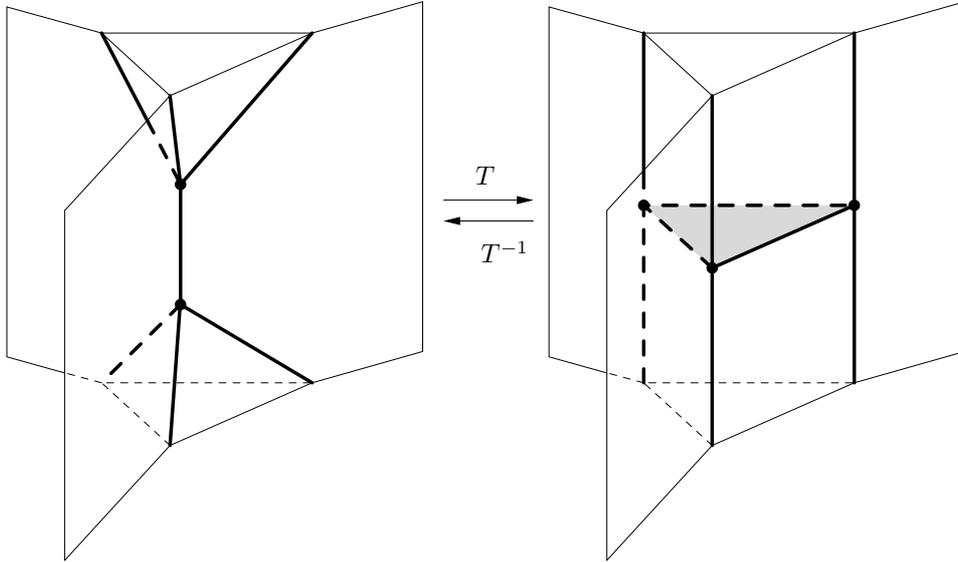


FIGURE 1. The Matveev-Piergallini move

“Local transformation” means that the special 2-polyhedron remains unchanged outside of the depicted part. The assumption on the number of vertices is no restriction, as *any* compact 3-manifold has a special spine with at least two vertices. Note that originally Theorem 8 was formulated without restriction on the number of vertices and involved an additional type of local transformation (compare Section 3). However, if the special spines all have at least two true vertices, the additional local transformation factorises by  $T$  and  $T^{-1}$ , see [68].

2.2.2. *Turaev–Viro State Sums and Biedenharn–Elliott equations.* Let  $P$  be a special 2-polyhedron with a choice of orientation for each 2-stratum. Let  $\mathcal{F}$  be a finite set, to whose elements we will refer by **2-strata colours**, and let “ $-$ ” be an involution on  $\mathcal{F}$ . Let  $\mathcal{G}$  be another finite set. A  $\mathcal{F}, \mathcal{G}$ -colouring of  $P$  is any pair  $(\varphi, \psi)$  of maps  $\varphi: \mathcal{C}(P) \rightarrow \mathcal{F}$ ,  $\psi: \mathcal{E}(P) \rightarrow \mathcal{G}$ .

For an oriented 2-stratum of colour  $f \in \mathcal{F}$ , the oppositely oriented 2-stratum shall have the colour  $-f \in \mathcal{F}$ . We denote by  $\Phi_{\mathcal{F}, \mathcal{G}}(P)$  the set of all  $\mathcal{F}, \mathcal{G}$ -colourings of  $P$ . If  $\mathcal{G}$  contains only one element then obviously an  $\mathcal{F}, \mathcal{G}$ -colouring is determined by  $\varphi$  alone, and we refer to it as a  $\mathcal{F}$ -colouring.

For any  $f \in \mathcal{F}$ , the symbol  $w(f)$  is referred to as the **weight** of  $f$ . At a true vertex of  $P$ , six 2-strata and four true edges meet (counted with multiplicities). A  $\mathcal{F}, \mathcal{G}$ -colouring  $(\varphi, \psi)$  of  $P$  thus assigns to each true vertex of  $P$  a 6-tuple of 2-strata colours together with a 4-tuple of edge colours. Let  $\varphi$  and  $\psi$  assign the colours  $a, \dots, f \in \mathcal{F}$  and  $A, \dots, D \in \mathcal{G}$  to the 2-strata and true edges in the neighbourhood of a true vertex  $v$ , as depicted in Figure 2 (where circular orientations of the 2-strata are indicated by arrows, the true vertex is marked by a thick dot, and the true edges are drawn bold); then we associate to  $v$  a symbol  $v^{\varphi, \psi} := \left| \begin{smallmatrix} a & b & c \\ f & e & d \end{smallmatrix} \right|_{C, D}^{A, B}$ ,

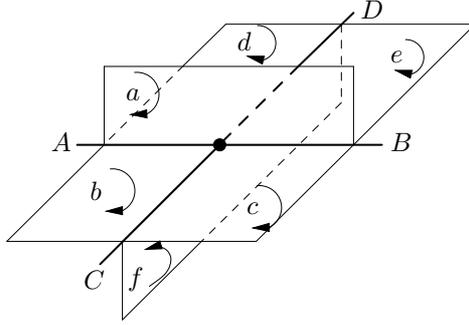


FIGURE 2. A true vertex  $v$  with coloured neighbourhood

called **6j4k-symbol**. If essentially only the 2-strata are coloured, i.e., if  $\mathcal{G} = \{*\}$  contains only one element, we just have a **6j-symbol**  $v^\varphi := \left| \begin{smallmatrix} a & b & c \\ f & e & d \end{smallmatrix} \right| = \left| \begin{smallmatrix} a & b & c \\ f & e & d \end{smallmatrix} \right|_{*, *}$ , to simplify the notation. This will be the case in most of our examples.

Due to a tetrahedral symmetry, there are 12 different ways to draw the neighbourhood of a true vertex as in Figure 2. We want the 6j4k-symbol to be independent of that choice, and thus impose that for all  $a, b, c, d, e, f \in \mathcal{F}$  and all  $A, B, C, D \in \mathcal{G}$  holds

$$\begin{aligned} \left| \begin{smallmatrix} a & b & c \\ f & e & d \end{smallmatrix} \right|_{C, D}^{A, B} &= \left| \begin{smallmatrix} b & c & a \\ -e & -d & f \end{smallmatrix} \right|_{B, D}^{C, A} \\ &= \left| \begin{smallmatrix} a & -d & -e \\ -f & -c & -b \end{smallmatrix} \right|_{D, C}^{A, B} \end{aligned}$$

Similarly, we impose that  $w(f) = w(-f)$  for all  $f \in \mathcal{F}$ .

Let  $R$  be the polynomial ring over some field  $\mathbb{F}$  whose variables are the equivalence classes of colour weights and 6j4k-symbols. In most cases, we will simply take  $\mathbb{F} = \mathbb{Q}$ . Let  $m = |\mathcal{F}|$  and  $n = |\mathcal{G}|$ . The following polynomial only depends on the homeomorphism type of  $P$ :

$$TV_{m, n}(P) := \sum_{(\varphi, \psi) \in \Phi_{\mathcal{F}, \mathcal{G}}(P)} \left( \prod_{C \in \mathcal{C}(P)} w(\varphi(C)) \right) \cdot \left( \prod_{v \in \mathcal{V}(P)} v^{\varphi, \psi} \right) \in R$$

This polynomial is the **Turaev–Viro state sum** of  $P$  of type  $(m, n)$ .

Of course, if  $P$  is a special spine of a 3-manifold  $M$ , the state sum does not only depend on the homeomorphism type of  $M$  but also on the choice of  $P$ . The following ideal captures the effect of choosing  $P$ . The **Turaev–Viro ideal**  $I_{m,n} \subset R$  of type  $(m,n)$  as the ideal in  $R$  that is generated by

$$(1) \quad \sum_{A \in \mathcal{G}} \begin{vmatrix} j_1 & j_2 & j_3 \\ j_9 & j_8 & j_7 \end{vmatrix}_{k_3, A}^{k_1, k_2} \cdot \begin{vmatrix} j_4 & j_5 & j_6 \\ -j_9 & -j_8 & -j_7 \end{vmatrix}_{k_6, A}^{k_4, k_5} -$$

$$(2) \quad \sum_{A_1, A_2, A_3 \in \mathcal{G}} \sum_{j \in \mathcal{F}} w(j) \cdot \begin{vmatrix} j & j_1 & j_2 \\ j_7 & -j_5 & -j_4 \end{vmatrix}_{k_1, k_4}^{A_1, A_2} \cdot \begin{vmatrix} j & j_2 & j_3 \\ j_9 & -j_6 & -j_5 \end{vmatrix}_{k_3, k_6}^{A_2, A_3} \cdot \begin{vmatrix} j & j_3 & j_1 \\ -j_8 & -j_4 & -j_6 \end{vmatrix}_{k_2, k_5}^{A_3, A_1},$$

for all  $j_1, \dots, j_9 \in \mathcal{F}$  and all  $k_1, \dots, k_6 \in \mathcal{G}$ . Note that these generators are known from quantum mechanics and are called “Biedenharn–Elliott equations” [60]. Let  $tv_{m,n}(P)$  be the coset of the Turaev–Viro state sum with respect to the Turaev–Viro ideal, i.e.

$$tv_{m,n}(P) = TV_{m,n}(P) + I_{m,n} \in R/I_{m,n}.$$

**THEOREM 9** (and Definition (see [55])). *If  $P$  is any special spine of a compact 3-manifold  $M$  with at least two true vertices, then the coset  $tv_{m,n}(P)$  only depends on the homeomorphism type of  $M$ . We call  $tv_{m,n}(M) = tv_{m,n}(P)$  an **ideal Turaev–Viro invariant** of  $M$  of type  $(m,n)$ .  $\square$*

Let  $N$  be the number of variables of  $R$ , let  $\hat{\mathbb{F}}$  be the algebraic closure of  $\mathbb{F}$ , and let  $\mathfrak{v}(I_{m,n}) \subset \hat{\mathbb{F}}^N$  be the (affine) variety associated to  $I_{m,n}$ . Any numerical Turaev–Viro invariant associated to  $tv_{m,n}$  is obtained by evaluation of  $TV_{m,n}(\cdot)$  at some parameter  $x \in \mathfrak{v}(I_{m,n})$ .

Recall that the **radical**  $\sqrt{I}$  of an ideal  $I \subset R$  is the ideal formed by all polynomials  $p \in R$  with  $p^n \in I$  for some  $n \in \mathbb{N}$ . An ideal is called radical if it coincides with its radical.

**DEFINITION 8.** Let  $M$  be a compact 3-manifold with a special spine  $P$ . Let  $tv_{m,n}(\cdot)$  be the ideal Turaev–Viro invariant obtained from the Turaev–Viro ideal  $I_{m,n}$ . The coset

$$\hat{t}v_{m,n}(M) = TV_{m,n}(P) + \sqrt{I_{m,n}} \in R/\sqrt{I_{m,n}}$$

is called the **universal numerical Turaev–Viro invariant** of  $M$  associated to  $tv_{m,n}$ .

The name “universal numerical Turaev–Viro invariant” is justified by the following theorem.

**THEOREM 10** (see [55]). *Let  $tv_{m,n}(\cdot)$  be an ideal Turaev–Viro invariant. Then for all compact 3-manifolds  $M_1, M_2$  holds  $\hat{t}v_{m,n}(M_1) = \hat{t}v_{m,n}(M_2)$  if and only if all numerical Turaev–Viro invariants associated to  $tv_{m,n}(\cdot)$  coincide on  $M_1$  and  $M_2$ .  $\square$*

Since  $\sqrt{I_{m,n}} \supset I_{m,n}$ , the computation of  $\hat{t}v_{m,n}(\cdot)$  does not yield an improvement of  $tv_{m,n}(\cdot)$ . However, since numerical Turaev–Viro invariants came first and are well studied, it seems interesting to compare the strength of an ideal Turaev–Viro invariant  $tv_{m,n}(\cdot)$  with the strength of its numerical descendants — and  $\hat{t}v_{m,n}(\cdot)$  is the right tool to do this in full generality.

**2.2.3. Numerical Turaev–Viro invariants associated to  $U_q(\mathfrak{sl}_2)$ .** It is well known that interesting numerical Turaev–Viro invariants exist for arbitrarily large colour sets. First examples have been presented by V. Turaev and O. Viro [96]. The representation theory of quantum groups yields a very successful machinery for constructing numerical Turaev–Viro invariants [97].

For the numerical Turaev–Viro invariants associated to  $U_q(\mathfrak{sl}_2)$  for  $q = e^{\frac{\pi i}{k+1}}$ , one has  $\mathcal{F} = \{\frac{i}{2} \mid i = 0, \dots, k-1\}$ , trivial involution and trivial edge colours. The  $6j$ -symbol  $\begin{vmatrix} a & b & c \\ f & e & d \end{vmatrix}$  vanishes unless each of the four triples  $(a, b, d)$ ,  $(a, c, e)$ ,  $(b, c, f)$  and  $(d, e, f)$  satisfies the triangle inequalities and has integer sum; moreover,  $w(0) = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} = 1$ . For details, see [96] or [97].

Note that the colouring rules for the  $6j$ -symbols of  $U_q(\mathfrak{sl}_2)$  are related to the theory of normal surfaces (compare 3). The notion of normal surfaces is not only defined for triangulations but also for special spines [68]; in fact, a special spine of a closed 3-manifold is dual to a one-vertex triangulation of that manifold.

If the 2-strata of a special spine  $P$  of a closed 3-manifold  $M$  is coloured such that the  $U_q(\mathfrak{sl}_2)$   $6j$ -symbols associated to the true vertices of  $P$  do not vanish, then there is a so-called *pre-normal* surface  $F \subset M$  (pre-normal is the terminology in [48]; it is  $k$ -normal in the terminology of [68]), such that the number of copies of  $c$  in  $F$  is twice the colour of  $c$ , for each 2-stratum  $c$  of  $P$ .

Any normal surface is pre-normal. There is an addition on the set of pre-normal surfaces, yielding the structure of a finitely generated semi-group [48]. The generators are called *fundamental surfaces*, in analogy to the corresponding property of normal surfaces (compare Chapter 3).

If the parameter  $q$  in  $U_q(\mathfrak{sl}_2)$  is not a root of unity, then we obtain  $6j$ -symbols over an infinite set of colours. So, the notion of a state sum comprises an infinite number of summands and is thus not defined. However, it was shown by Frohman and Kania-Bartoszyńska [28] that the state sum of  $P$  is absolutely convergent for  $q \in ]0, 1[$ , if there is only one normal sphere (namely the boundary of a regular neighbourhood of  $P$  in  $M$ ) and no normal torus.

**2.3. How to compute ideal Turaev–Viro invariants.** How can one distinguish manifolds using ideal Turaev–Viro invariants? The first task is to present a special 2-polyhedron in a form that is accessible for computers. Matveev [68, Sec. 7.1] introduced a way of encoding special 2-polyhedra by lists of cyclic sequences of integers. From the Matveev representation of  $P$ , it is not difficult to deduce the incidences of oriented 2-strata and true edges of  $P$  at the true vertices. Hence, one can easily implement the computation of the Turaev–Viro state sum.

Let  $P_1$  and  $P_2$  be special spines of compact 3-manifolds  $M_1$  and  $M_2$ . We are now able to compute  $TV_{m,n}(P_1)$  and  $TV_{m,n}(P_2)$ . But how can we determine whether  $tv_{m,n}(M_1) = tv_{m,n}(M_2)$  or not? In other words, we need to compare cosets with respect to ideals in a polynomial ring over a field. This is algorithmically possible by the theory of Gröbner bases. For an introduction to that subject, we refer the reader to [27] or [59], among many other possible sources.

Firstly, we need to choose an **admissible monomial ordering**  $<$  on  $R$ ; this is a total order on the set of monomials (i.e., products of variables) of  $R$  such that  $1 < \mathfrak{m}$  for any monomial  $\mathfrak{m} \in R$  and such that  $\mathfrak{m}_1 < \mathfrak{m}_2$  implies  $\mathfrak{m}\mathfrak{m}_1 < \mathfrak{m}\mathfrak{m}_2$  for all monomials  $\mathfrak{m}, \mathfrak{m}_1, \mathfrak{m}_2 \in R$ . For a polynomial  $f \in R$ , the **leading monomial** of  $f$  with respect to  $>$  is denoted by  $lm(f)$ .

A **Gröbner basis** with respect to  $>$  of an ideal  $I \subset R$  is a finite set  $B \subset I$  such that

$$\langle \{lm(f) : f \in B\} \rangle = \langle \{lm(f) : f \in I\} \rangle,$$

where  $\langle X \rangle$  denotes the ideal generated by a set  $X \subset R$ . It turns out that any Gröbner basis of  $I$  is a generating subset of  $I$ , and that any ideal in  $R$  has a Gröbner basis (specifically, any ideal is finitely generated). If  $B$  satisfies some additional hypothesis (see [27, Sec. 3.7] for details), it is called *reduced* Gröbner basis, and turns out to be unique, hence depends only on  $I$  and  $>$ . The reduced Gröbner basis can be algorithmically constructed, given an arbitrary finite generating subset of  $I$ .

If an admissible ordering on  $R$  is given then one can generalise the usual division algorithm of univariate polynomials to multivariate polynomials and can define the remainder  $\text{rem}(f; g) \in R$  of a polynomial  $f \in R$  with respect to a polynomial  $g \in R$ . In general the remainder will depend on the chosen ordering.

Let  $I = \langle g_1, \dots, g_k \rangle \subset R$ . If one wants to test whether some polynomial  $f \in R$  belongs to  $I$  (this is the “ideal membership problem”), it is a reasonable idea to iteratively compute the remainder  $\text{rem}(f; g_1, \dots, g_k)$  of  $f$  with respect to  $g_1, \dots, g_k$ , i.e.,  $\text{rem}(\dots \text{rem}(\text{rem}(f; g_1); g_2) \dots; g_k)$ . Certainly if  $\text{rem}(f; g_1, \dots, g_k) = 0$  then  $f \in I$ . In general the converse is not true. However, for Gröbner bases, it is, by the following theorem, that solves the ideal membership problem. For a proof, we refer to [27] or [59].

**THEOREM 11.** *Let  $\mathcal{G}$  be a Gröbner basis for  $\langle \mathcal{G} \rangle \subset R$ , and let  $p \in R$ . Then,  $\text{rem}(p; \mathcal{G})$  does not depend on the order of polynomials in  $\mathcal{G}$ , and we have  $\text{rem}(p; \mathcal{G}) = 0$  if and only if  $p \in \langle \mathcal{G} \rangle$ .  $\square$*

Our computations involve the following three steps.

- (1) Produce the list of variables of  $R$  and the list of generators of  $I_{m,n}$  defined in the previous section.
- (2) Compute a Gröbner basis of  $I_{m,n}$  for the chosen admissible monomial ordering  $>$ .
- (3) For any special 2-polyhedron  $P$ , compute  $TV_{m,n}(P)$ , and
- (4) compute the normal form of  $TV_{m,n}(P) + I_{m,n}$  using the Gröbner basis obtained in step 2.

For step 1 and 3, we wrote `maple`<sup>TM</sup> programs [63]. For step 2 and 4, we used SINGULAR [29]. For the computation of Gröbner bases of Turaev–Viro ideals, the algorithm `slimgb`, implemented in SINGULAR by M. Brickenstein [14], turned out to be particularly well-performing. If we want to compute the universal numerical Turaev–Viro invariant associated to  $tv_{m,n}$ , we simply replace  $I_{m,n}$  by  $\sqrt{I_{m,n}}$ , which is possible since one can compute a finite set of generators of  $\sqrt{I_{m,n}}$  for any finite set of generators of  $I_{m,n}$  (we used the `primdec.lib` library of SINGULAR [21] for that purpose).

**2.4. Simplifying assumptions for Turaev–Viro invariants.** The number of variables of  $R$  and the number of generators of  $I_{m,n}$  grow rapidly with increasing  $m$  and  $n$ . Moreover, the complexity of a Gröbner basis computation grows rapidly with the number of variables. So obviously the computation of ideal Turaev–Viro invariants is a difficult task. One way to overcome these complexity problems is to introduce simplifying assumptions, that reduce the number of variables. For instance, we can restrict the set of colourings by sending some equivalence classes of  $6j4k$ -symbols to some element of  $\mathbb{F}$ , e.g., to zero.

At the end of Subsection 2.2.2, we stated some properties of the numerical Turaev–Viro invariants associated with  $U_q(\mathfrak{sl}_2)$ . These properties show that one obtains a non-trivial ideal Turaev–Viro invariant by defining  $\mathcal{F} = \{\frac{i}{2} \mid i = 0, \dots, k-1\}$  (for some  $k \in \mathbb{N}$ ) with trivial involution on  $\mathcal{F}$ ,  $\mathcal{G} = \{*\}$ , and assuming that the  $6j$ -symbol  $\left| \begin{smallmatrix} a & b & c \\ f & e & d \end{smallmatrix} \right|$  vanishes unless each of the four triples  $(a, b, d)$ ,  $(a, c, e)$ ,  $(b, c, f)$  and  $(d, e, f)$  satisfies the triangle inequalities and has integer sum; moreover, one may also define  $w(0) = \left| \begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix} \right| = 1$ .

In our computations, we even weakened this condition: Instead of the triangle inequalities for  $(a, b, d)$  (and analogously for  $(a, c, e)$ ,  $(b, c, f)$  and  $(d, e, f)$ ) we require  $a \leq b + d + 1$ ,  $b \leq a + d + 1$  and  $d \leq a + b + 1$ , and  $a + b + d \in \mathbb{Z}$ . We will refer to this as the *admissibility condition*, and we denote the ideal Turaev–Viro invariant thus obtained by  $\mathfrak{tv}_k$  and the Turaev–Viro ideal by  $\mathcal{J}_k$ .

It is not difficult to see that one obtains  $\mathfrak{t}\mathfrak{v}_{k-1}$  from  $\mathfrak{t}\mathfrak{v}_k$  by a partial evaluation. Hence,  $\mathfrak{t}\mathfrak{v}_k$  is at least as strong as  $\mathfrak{t}\mathfrak{v}_{k-1}$ , and we get a hierarchy of increasingly strong invariants. After matching our notations and conventions on the order of indices of  $6j$ -symbols with those of [96] or [97, Sec. XII.8.5], the formulas provided there allow to express the  $U_q(\mathfrak{sl}_2)$  invariant with  $q = e^{\frac{\pi i}{k+1}}$  as a numerical Turaev–Viro invariant associated to  $\mathfrak{t}\mathfrak{v}_k$ .

We have another example  $\tilde{t}v_{2,1}$ , that is constructed according to Matveev’s  $\epsilon$ -invariant [66], [68], so that the  $\epsilon$ -invariant is a numerical Turaev–Viro invariant associated to  $tv_{2,1}$ . We provide the basic data for  $\tilde{t}v_{2,1}$  in [55], and will not repeat them here, for the sake of brevity. In [55], we also describe simplifying assumptions that are useful for the construction of ideal Turaev–Viro invariants with non-trivial edge colours.

**2.5. Results.** Let  $\tilde{c}(M)$  be the minimal number of true vertices of a special spine of a compact 3-manifold  $M$ . This is related to Matveev’s notion of **complexity** of manifolds,  $c(M)$ : If  $M$  is a closed irreducible 3-manifold different from the 3-sphere, the projective space and the lens space  $L(3, 1)$  then  $c(M) = \tilde{c}(M)$ , by Theorem 2.2.4 in [68].

Let  $p \in R$  be a polynomial. Let  $\deg_w(p)$  be the total degree of  $p$  in the colour weights, and let  $\deg_{6j}(p)$  be the total degree of  $p$  in the  $6j4k$ -symbols. For any subset  $A \subset R$ , let  $\deg_w(A) = \min\{\deg_w(p) : p \in A\}$  and  $\deg_{6j}(A) = \min\{\deg_{6j}(p) : p \in A\}$ . For any special spine  $P$  of a compact 3-manifold  $M$ ,  $\deg_{6j}(TV_{m,n}(P))$  is the number of true vertices of  $P$ . Hence,  $c(M)$  is an upper bound for  $\deg_{6j}(tv_{m,n}(M))$ , and similarly one can estimate  $\deg_w(tv_{m,n}(M))$ . One obtains the following result.

LEMMA 6 (see [55]). *Let  $tv_{m,n}(\cdot)$  be an ideal Turaev–Viro invariant. For any closed 3-manifold  $M$  with  $\tilde{c}(M) > 1$ , we have*

$$\tilde{c}(M) \geq \max \{ \deg_w(tv_{m,n}(M)) - 1, \deg_{6j}(tv_{m,n}(M)) \} \quad \square$$

We tested the following ideal Turaev–Viro invariants (for more details, see [55]):

- (1)  $\tilde{t}v_{2,1}$ , an invariant of type  $(2, 1)$  that generalises Matveev’s  $\epsilon$ -invariant [66], [68]. After a simplifying assumption, there remain 4  $6j$ -symbols and 1 colour weight. Here one has 12 generators for the Turaev–Viro ideal, and we obtain a Gröbner basis of 22 polynomials.
- (2)  $\mathfrak{t}\mathfrak{v}_k$  for  $k < 5$ . The case  $k = 1$  is trivial. For  $k = 2$ , only two equivalence classes of  $6j$ -symbols remain, and  $\mathfrak{J}_2$  is generated by four polynomials; a Gröbner basis with 6 polynomials is easily obtained. For  $k = 3$ , we have 17 classes of  $6j$ -symbols and two colour weights; the Turaev–Viro ideal  $\mathfrak{J}_3$  is generated by 130 polynomials, and there is a Gröbner basis with respect to some degree reverse lexicographic order, formed by 496 polynomials; this computation is non-trivial, and different algorithms differ widely in their performance. For  $k = 4$ , we have 49 classes of  $6j$ -symbols and 3 colour weights, 892 generators of  $\mathfrak{J}_4$  and a Gröbner basis formed by 13642 polynomials. For  $k = 5$ , we have 136 classes of  $6j$ -symbols, 4 colour weights, and 4830 generators of  $\mathfrak{J}_5$ ; a Gröbner basis is not known, up to now.
- (3)  $tv_{3,1}^\pm(\cdot)$ , an invariant of type  $(3, 1)$ , with  $\mathcal{F} = \{-1, 0, 1\}$ , the usual involution  $-(-1) = 1$ ,  $-0 = 0$ , and trivial edge colours. After a mild simplifying assumption, we have 41 equivalence classes of  $6j$ -symbols and one remaining colour weight. We obtain 1661 generators for the Turaev–Viro ideal, and SINGULAR [29] succeeds with finding a Gröbner basis with respect to some degree reverse lexicographic order formed by 1297 polynomials.

- (4) Another invariant of type 3, 1,  $\tilde{tv}_{3,1}^+(\cdot)$ , has the same colours sets and the same involution as  $tv_{3,1}^+(\cdot)$ , but is subject to a stronger simplifying assumption. There remain 21 equivalence classes of  $6j$ -symbols. The Turaev–Viro ideal is generated by 474 polynomials, and SINGULAR [29] easily finds a Gröbner basis of 337 polynomials.
- (5) Our biggest example is of type (5, 1) with non-trivial involution. Here, we have  $\mathcal{F} = \{-2, \dots, 2\}$ , and a simplifying assumption very similar to the admissibility condition. We have 72 classes of  $6j$ -symbols, 5667 generators of the Turaev–Viro ideal, and a Gröbner basis formed by 4403 polynomials. We denote the resulting invariant by  $\tilde{tv}_{5,1}^+(\cdot)$ .
- (6) We considered one invariant of type (2, 2). We made simplifying assumptions that reduced the number of  $6j$ -symbols to 22, and the number of generators of the Turaev–Viro ideal to 353. But this was still too difficult for the computer. Therefore we added 22 *additional* generators to the Turaev–Viro ideal. These are gotten from modifying a Gröbner basis for our invariant of type (2, 1). The additional generators simplify the computations, and Singular [29] finds a Gröbner basis formed by 449 polynomials. We denote the resulting invariant by  $\tilde{tv}_{2,2}(\cdot)$ .

We computed  $\tilde{tv}_{2,1}$ ,  $\mathfrak{tv}_3$ ,  $\tilde{tv}_{3,1}^+$ ,  $tv_{3,1}^+$  and  $\tilde{tv}_{5,1}^+$  for closed orientable irreducible manifolds up to complexity 9,  $\mathfrak{tv}_4$  up to complexity 8, and  $\tilde{tv}_{2,2}$  up to complexity 6. It is known that there are precisely 1900 closed orientable irreducible 3-manifolds of complexity at most 9, up to homeomorphism. A list containing exactly one special spine for each of these manifolds was provided to us by Sergei Matveev. A tabulation of all closed orientable irreducible 3-manifolds of complexity at most 11 can be found in [69].

The following statements result from our computations.

PROPOSITION 3 (see [55]).

- (1) *The Turaev–Viro ideals involved in the construction of  $\tilde{tv}_{2,1}(\cdot)$ ,  $\tilde{tv}_{3,1}^+(\cdot)$  and  $tv_{3,1}^+(\cdot)$  are not radical.*
- (2) *We measure the strength of an invariant by the number of different values that it assumes on the 1900 closed irreducible orientable 3-manifolds of complexity  $\leq 9$ . We obtained the following.*
  - *The  $\epsilon$ -invariant assumes 35 values; its generalisation  $\tilde{tv}_{2,1}$  assumes 134 different values, hence it is stronger roughly by a factor 3.8.*
  - *The  $U_q(\mathfrak{sl}_2)$  invariant with  $q = e^{\pi i/4}$  assumes 29 different values; its generalisation  $\mathfrak{tv}_3$  assumes 250 different values, hence it is stronger roughly by a factor 8.6.*
  - *$\tilde{tv}_{3,1}^+$  assumes 242 and  $\tilde{tv}_{5,1}^+$  assumes 387 different values.*
  - *Homology assumes 272 different values.*
- (3) *Using the combination of homology with  $\tilde{tv}_{2,1}$ ,  $\tilde{tv}_{3,1}^+$ ,  $\mathfrak{tv}_3$  or  $\tilde{tv}_{5,1}^+$ , one can distinguish respectively 764, 764, 879 or 879 homeomorphism types of closed irreducible orientable 3-manifolds of complexity  $\leq 9$ . A combination of different ideal Turaev–Viro invariants did not yield a further improvement.*
- (4) *On closed irreducible orientable 3-manifolds of complexity at most 9,  $\tilde{tv}_{3,1}^+$  and  $tv_{3,1}^+$  are equivalent invariants. On closed irreducible orientable 3-manifolds of complexity  $\leq 8$ ,  $\mathfrak{tv}_3$  and  $\mathfrak{tv}_4$  are equivalent. On closed irreducible orientable 3-manifolds of complexity  $\leq 6$ ,  $\tilde{tv}_{2,1}$  and  $\tilde{tv}_{2,2}$  are equivalent.*

- (5) On the closed irreducible orientable 3-manifolds that we considered, the ideal Turaev–Viro invariants  $\tilde{tv}_{2,1}(\cdot)$ ,  $\tilde{tv}_{3,1}^+(\cdot)$  and  $tv_{3,1}^+(\cdot)$  are equivalent to their associated universal numerical Turaev–Viro invariant, although their associated Turaev–Viro ideals are not radical, by Statement (1).
- (6) The lower bound for the complexity stated in Lemma 6 is trivial in all examples that we computed.
- (7) Ideal Turaev–Viro invariants are, in general, not multiplicative under connected sum of compact 3-manifolds.  $\square$

The first statement of Proposition 3 says, in combination with Theorem 10, that one should expect that ideal Turaev–Viro invariants are, in general, stronger than a combination of all associated numerical Turaev–Viro invariants.

The second and third statement of Proposition 3 shows that  $\tilde{tv}_{2,1}(\cdot)$  sees properties of manifolds that are invisible for homology, and vice versa.

Statement (4) is surprising, because one would expect that one obtains a stronger invariant if one avoids to impose simplifying assumptions. Statement (5) is even more surprising, because by statement (1) the Turaev–Viro ideals are not radical — hence there are elements of  $R$  so that the cosets with respect to  $\sqrt{I_{m,n}}$  coincide, but the cosets with respect to  $I_{m,n}$  are different. Are there compact 3-manifolds  $M_1, M_2$  that can be distinguished by some ideal Turaev–Viro invariant  $tv(\cdot)$  but can not be distinguished by all associated numerical Turaev–Viro invariants, i.e., can not be distinguished by  $\hat{tv}(\cdot)$ ? Note that  $\tilde{tv}_{2,1}$  is stronger than the  $\epsilon$ -invariant; but the  $\epsilon$ -invariant is not the only numerical Turaev–Viro invariant associated to  $\tilde{tv}_{2,1}$  (see [68, Sec. 8.1]).

**2.6. Multiplicative invariants.** The last statement of Proposition 3 is a bad news if one wants to construct a Topological Quantum Field Theory. But it is a good news if one aims to construct invariants that potentially detect counterexamples of the Andrews–Curtis conjecture, as we will explain in Section 3.

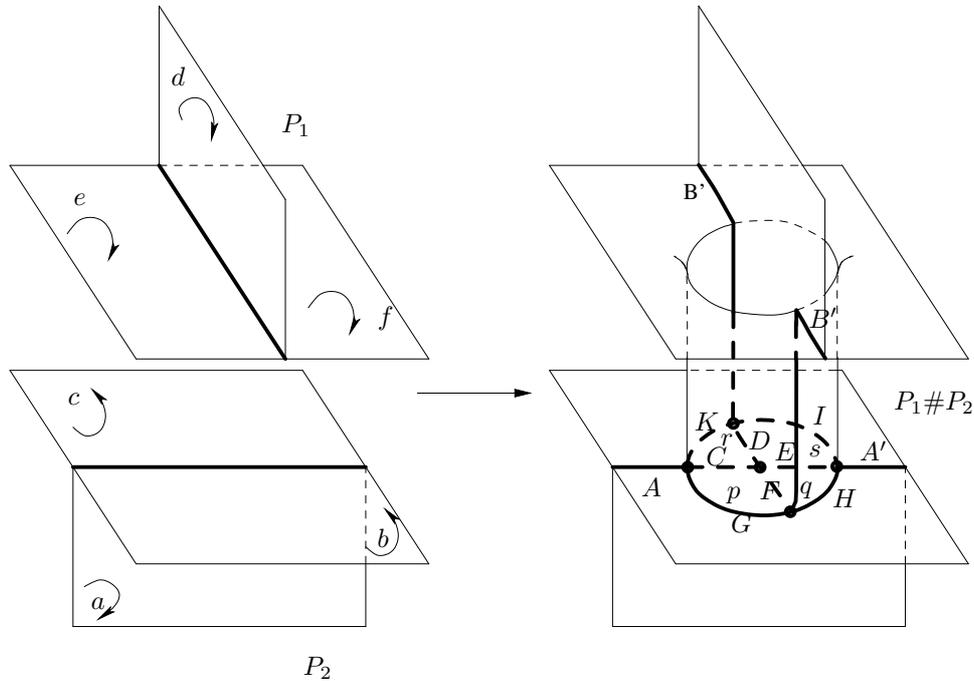


FIGURE 3. Connected sum of two special spines

However, it is not difficult to construct ideal Turaev–Viro invariants that are multiplicative under connected sums. If  $P_1, P_2$  are special spines of closed 3–manifolds  $M_1, M_2$ , then one obtains a special spine  $P_1\#P_2$  of the connected sum  $M_1\#M_2$  by the local process depicted in Figure 3. In the figure, 2–strata are labeled by lower case letters and true edges by upper case letters. The 2–strata  $a, b, c, d, e, f$  on the left hand side are circularly oriented as indicated by the arrows, and correspond to six 2–strata on the right hand side in the obvious way. The true edges  $A, A'$  (resp.  $B, B'$ ) on the right hand side correspond to only one true edge on the left hand side. Hence, there must be no contribution unless  $A = A'$  and  $B = B'$ , hence, using Kronecker symbols,  $\delta_{A,A'} = \delta_{B,B'} = 1$ .

The 2–strata  $p, q, r, s$  on the right hand side have counterclockwise orientation. These four 2–strata and the eight intrinsic edges  $C, D, E, F, G, H, I, K$  are completely contained inside the local picture, hence, give rise to a summation (we abbreviate this in the following formula). An inspection of the figure shows that  $tv_{m,n}(M_1\#M_2) = tv_{m,n}(M_1) \cdot tv_{m,n}(M_2)$  if we add the following generators to the Turaev–Viro ideal, for  $a, b, c, d, e, f \in \mathcal{F}$  and  $A, A', B, B' \in \mathcal{G}$ :

$$\delta_{A,A'}\delta_{B,B'} - \sum_{p,q,r,s \in \mathcal{F}} \sum_{C,\dots,K \in \mathcal{G}} \left( w(p)w(q)w(r)w(s) \begin{vmatrix} a & b & q \\ -f & s & c \end{vmatrix} \begin{vmatrix} A',E \\ H,I \end{vmatrix} \begin{vmatrix} a & p & b \\ e & c & r \end{vmatrix} \begin{vmatrix} C,A \\ G,K \end{vmatrix} \right. \\ \left. \begin{vmatrix} b & p & q \\ d & -f & -e \end{vmatrix} \begin{vmatrix} G,H \\ F,B \end{vmatrix} \begin{vmatrix} c & s & r \\ -d & -e & -f \end{vmatrix} \begin{vmatrix} I,K \\ D,B' \end{vmatrix} \begin{vmatrix} a & q & p \\ -d & r & s \end{vmatrix} \begin{vmatrix} E,C \\ F,D \end{vmatrix} \right)$$

The rest of the machinery (computation of Gröbner bases etc.) is as usual, and yields an invariant that is multiplicative under connected sum. We made no extensive computations in this context.

### 3. State sum invariants for the Andrews–Curtis problem

**3.1. Introduction.** Let  $\mathcal{P} = \langle g_1, \dots, g_m \mid r_1, \dots, r_n \rangle$  a finite presentation of a group  $G$ ; hence, the  $r_i$  (“relators”) are reduced words in the free group  $F(g_1, \dots, g_m)$  and generate a normal subgroup  $\langle r_1, \dots, r_n \rangle$  of  $F(g_1, \dots, g_m)$ , and  $G$  is isomorphic to  $F(g_1, \dots, g_m)/\langle r_1, \dots, r_n \rangle$ . The **deficiency** of  $\mathcal{P}$  is  $m - n$ , and if the deficiency vanishes, then  $\mathcal{P}$  is called a **balanced presentation**. Sometimes, for convenience, a relator is given in the form  $w_1 = w_2$  for words  $w_1, w_2 \in F(g_1, \dots, g_m)$ ; this stands for the relator  $w_1 w_2^{-1}$ .

It is well known [39] that any two finite presentations of  $G$  are related by finite sequences of the following transformations and their inverses.

- (1) Renumbering of generators or relators,
- (2) Free reduction of relators,
- (3) Replacement of  $g_i$  by  $g_i^{-1}$  in all relators ( $i = 1, \dots, m$ ),
- (4) Replacement of  $g_i$  by  $g_i g_j$  in all relators ( $i, j = 1, \dots, m, j \neq i$ ),
- (5) Replacement of  $g_i$  by  $g_j g_i$  in all relators ( $i, j = 1, \dots, m, j \neq i$ ),
- (6) Addition of a new generator  $g_{m+1}$  and a new relator  $r_{n+1} = g_{m+1}$ ,
- (7) Replacement of  $r_i$  by  $r_i^{-1}$  ( $i = 1, \dots, n$ ),
- (8) Replacement of  $r_i$  by  $r_i r_j$  ( $i, j = 1, \dots, n, j \neq i$ ),
- (9) Replacement of  $r_i$  by  $w r_i w^{-1}$  ( $i = 1, \dots, n, w \in F(g_1, \dots, g_m)$ ),
- (10) Addition of a trivial relator  $r_{n+1} = 1$ .

The last type of transformations obviously plays a special rôle: It is the only one that changes the deficiency of a presentation. So it is a natural question if this transformation can be avoided if one considers two presentations of the same deficiency. According to the Andrews–Curtis conjecture, this is indeed possible, at least for balanced presentation of the trivial group.

CONJECTURE 1 (Andrews–Curtis, [3]). *Any balanced presentation of the trivial group can be transformed into the trivial presentation  $\langle g_1 | g_1 \rangle$  by a finite sequence of transformations of type (1)–(9).*

The equivalence relation generated by transformations (1)–(9) is called **Andrews–Curtis equivalence** or  $Q^{**}$  equivalence in [39]. A balanced presentation of the trivial group that is Andrews–Curtis equivalent to the trivial presentation is called **Andrews–Curtis trivial**. The Andrews–Curtis conjecture is 40 years old, and nearly as old are infinite classes of potential counterexamples. The following presentations are balanced presentations of the trivial group, but for all of them it is unknown whether they are Andrews–Curtis trivial.

- (1)  $\langle a, b, c \mid c^{-1}bc = b^2, a^{-1}ca = c^2, b^{-1}ab = a^2 \rangle$ , see [82]
- (2)  $\langle a, b \mid aba = bab, a^k = b^{k+1} \rangle$ , for  $k \in \mathbb{N}, k \geq 3$ . See [39], generalising [1].
- (3)  $\langle a, b \mid ab^m = b^{m+1}a, ba^n = a^{n+1}b \rangle$ , for  $m, n \in \mathbb{N}, m, n \geq 3$ . See [39], generalising [20].

Example (2) with  $k = 3$  is, to the current knowledge, the smallest potential counterexample [10], [34], namely with relators of total length 13.

It is known [39] that the Andrews–Curtis conjecture is equivalent to a statement on transformations of special 2–polyhedra. Let  $P$  be a special 2–polyhedron. We do not assume that it is a special spine of some compact 3–manifold. Let  $c$  be a 2–stratum of  $P$ . Since  $P$  is special,  $c$  is a disc. Let  $D^2$  be the closed disc, and let  $\phi: D^2 \rightarrow P$  such that its restriction to the interior is a homeomorphism onto  $c$ . Walking along  $\phi(\partial D^2)$ , one obtains a not necessarily simple closed path in the singular graph  $S(P)$ . For  $x \in \partial D^2$  with  $\phi(x) \notin V(P)$ , a small neighbourhood of  $\phi(x)$  in  $P$  is formed by the  $\phi$ –image of a small neighbourhood of  $x$  in  $D^2$  together with a small stripe in  $P$ . When walking across a true vertex, these stripes fit together in a natural way. Hence, along  $\partial c$ , the stripes close up either to a locally embedded annulus or to a locally embedded Möbius strip in  $P$ . In the first case, we call  $c$  a **straight** 2–stratum, in the second case, we call  $c$  a **twisted** 2–stratum. It is known [68] that a special 2–polyhedron is special spine of some compact 3–manifold if and only if it has no twisted 2–stratum.

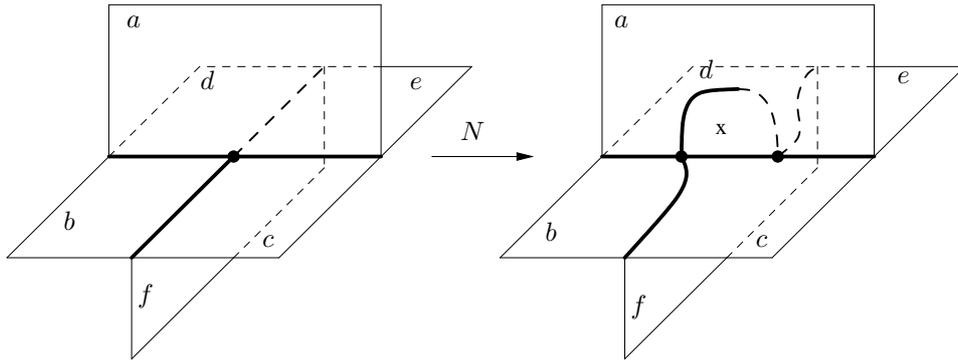
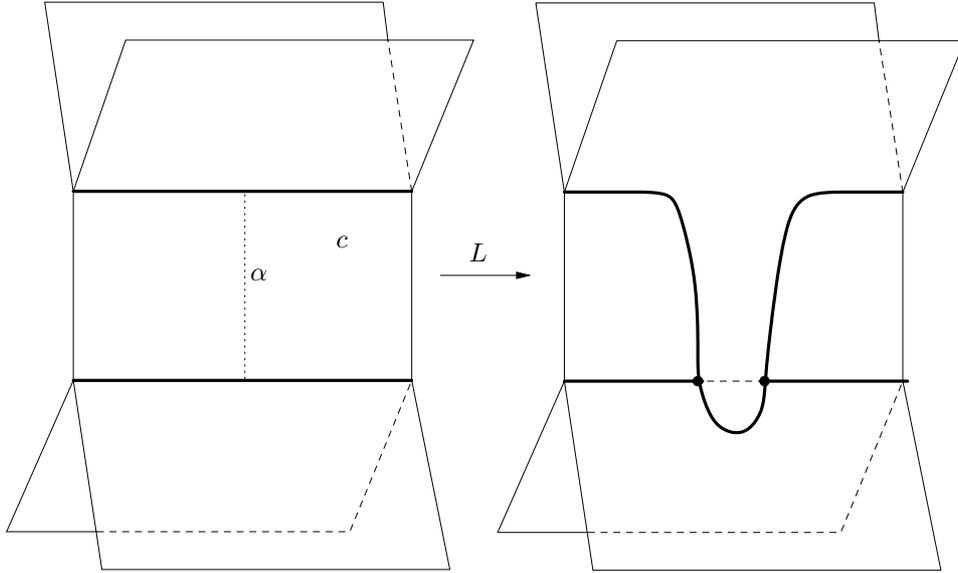
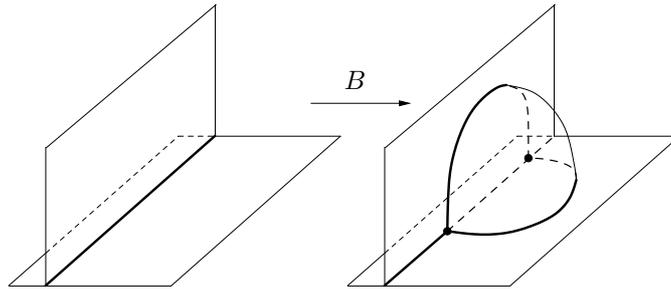


FIGURE 4. The move  $N$

Consider the following transformations.

- (1) The move  $T$ , depicted in Figure 1 on Page 42
- (2) The move  $N$ , depicted in Figure 4
- (3) The move  $L$ , depicted in Figure 5; here,  $\alpha$  is some arc embedded in a 2–stratum  $c$ ,  $\partial\alpha \subset \partial c$ , and the move takes place in a regular neighbourhood of  $\alpha$

FIGURE 5. The move  $L$ FIGURE 6. The move  $B$ 

(4) The move  $B$ , depicted in Figure 6.

The local picture of the right side of move  $N$  is *not* embeddable in  $\mathbb{R}^3$  (this is what the letter  $N$  stands for). As one can see in the figure, the new 2-stratum  $x$  introduced by  $N$  is twisted, while the 2-strata  $a, \dots, f$  change from twisted to straight and vice-versa. The move  $L$  is also known as *loon move*. Note that an application of the inverse of move  $L$  to a special 2-polyhedron might result in a simple 2-polyhedron that is not *special* — namely when it creates a 2-stratum that is not a disc. In that situation, we do not allow to perform the move. With that restriction, the moves  $T$ ,  $N$ ,  $L$  and their inverses generate an equivalence relation on special 2-polyhedra that we call **equivalence by 3-deformations**. Note that the same equivalence relation sometimes is called equivalence by 2-deformations in the literature; we follow here the terminology in [39]. It can be shown that the Andrews–Curtis conjecture for group presentations is equivalent to the following statement about special 2-polyhedra.

**THEOREM 12** (see [39]). *The Andrews–Curtis conjecture holds if and only if any two contractible<sup>1</sup> special 2-polyhedra are equivalent by 3-deformations.*  $\square$

<sup>1</sup>A *contractible* 2-complex has trivial fundamental group and Euler characteristic 1.

The move  $B$  adds a “bubble” (this is what the letter  $B$  stands for), and in contrast to  $T$ ,  $N$  and  $L$  it changes the Euler characteristic. It is known that any two contractible special 2–polyhedra  $P_1$  and  $P_2$  become equivalent by 3–deformation after finitely many applications of the move  $B$  to both  $P_1$  and  $P_2$ .

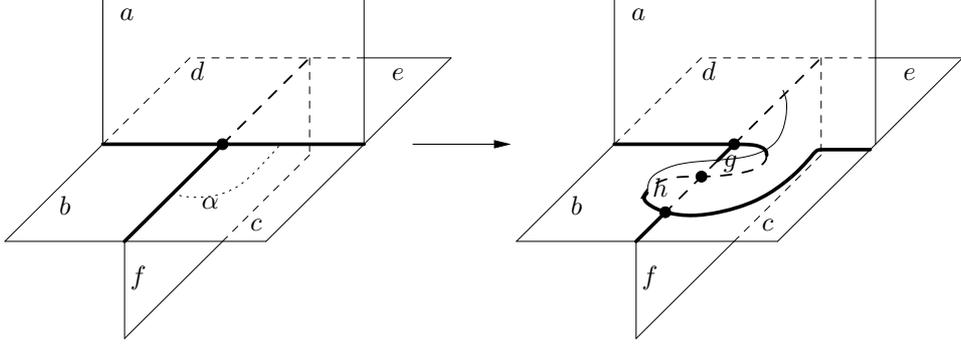
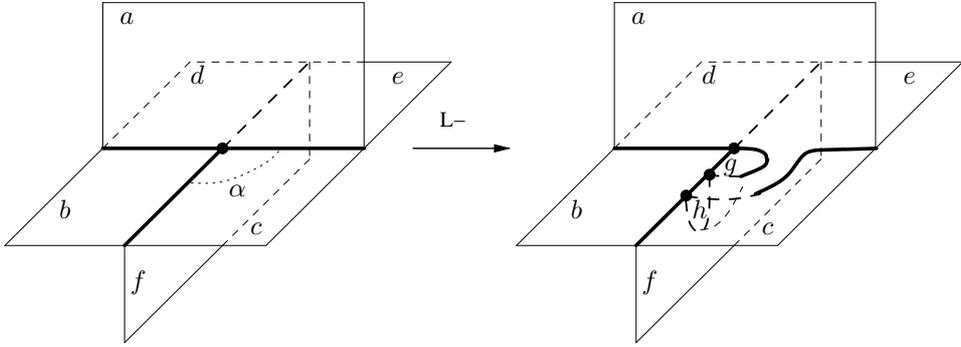
Any counterexample to the Andrews–Curtis conjecture gives rise to a pair of contractible special 2–polyhedra that are not equivalent by 3–deformation, and *vice versa*. A reasonable approach for disproving the Andrews–Curtis conjecture is to construct a **3–deformation invariant** for special 2–polyhedra, *i.e.*, to associate to any special 2–polyhedron some algebraic object that does not change under the moves  $T$ ,  $N$ ,  $L$  and their inverses. Such invariant could detect two contractible special 2–polyhedra that are not equivalent by 3–deformations. However, by the result of Bobtcheva and Quinn [6] mentioned in Section 1, this approach fails for a large class of invariants called *reductions of modular invariants*. We will not provide a proper definition of modular invariants here. Note, however, that modular invariants are multiplicative under connected sum of special 2–complexes (depicted in Figure 3).

Starting with numerical Turaev–Viro invariants associated to quantum groups, it is possible to construct 3–deformation invariants. However, these are reductions of modular invariants, hence, are unable to detect a counterexample to the Andrews–Curtis conjecture. By Statement (7) of Theorem 3, ideal Turaev–Viro invariants are, in general, not multiplicative under connected sums. Hence, there is some hope to obtain an invariant that goes beyond the restrictions provided by the result of Bobtcheva and Quinn [6]. In the next subsection, we will prove that one can replace the move  $L$  by another move that seems better suited for the construction of invariants. Then we modify the definition of ideal Turaev–Viro invariants and obtain 3–deformation invariants. By direct computations, we can show that our invariants are not multiplicative under connected sums; hence, they do not satisfy the formulas provided by Bobtcheva and Quinn [6]. Nevertheless, we have not been able to distinguish any two special 2–polyhedra that have the same homology groups.

**3.2. 3–deformation invariants.** Let  $P_1$  be a special 2–polyhedron. Let  $c$  be a 2–stratum of  $P_2$ , and let  $\alpha \subset c$  be an embedded arc with  $\partial\alpha = \{x, y\} \subset \partial c$ . Let  $P_2$  be the result of  $P_1$  under a move  $L$  along  $\alpha$  (compare Figure 5). We study the effect of a different choice of  $\alpha$ . Let  $\alpha' \subset c$  be another embedded arc, with  $\partial\alpha' = \{x, z\} \subset \partial c$ , and let  $P'_2$  be the result of  $P_1$  under a move  $L$  along  $\alpha'$ . Since  $P_1$  is special,  $c$  is a disc. Therefore, if  $y$  and  $z$  belong to the same true edge of  $P_1$  then  $\alpha$  and  $\alpha'$  are related by an isotopy that fixes any stratum of  $P_1$  set-wise. Hence, in that case,  $P_2$  and  $P'_2$  are homeomorphic. If there is a segment  $\beta \subset \partial c$  with  $\partial\beta = \{y, z\}$  that is disjoint from  $x$  and contains exactly one true vertex of  $P_1$  then  $P_2$  can be transformed into  $P'_2$  by a move of type  $T$  and the inverse of a move of type  $T$ .

Since  $P_1$  is special, there is some true vertex of  $P_1$  in  $\partial c$ . So, up to a sequence of moves of type  $T$  and their inverses, we can choose  $\alpha$  so that there is a segment  $\gamma \subset \partial c$  with  $\partial\gamma = \{x, y\}$  that contains precisely one true vertex,  $v$ , of  $P_1$ . Let  $D \subset c$  be the sub-disc of  $c$  bounded by  $\alpha \cup \gamma$ . After move  $L$ ,  $D$  becomes a 2–stratum  $g$  of  $P_2$ . Either  $g$  is straight or it is twisted. If it is straight, we have the situation of Figure 7. It is known that this move factorises by the move  $T$  and its inverse, provided that  $P_1$  has at least two true vertices — see [68]. If  $g$  is twisted, we obtain the situation in Figure 8. We denote this special case of move  $L$  by  $\tilde{L}$ .

The move  $\tilde{L}$  is local, in contrast to the move  $L$  that depends on the choice of an arc  $\alpha$  and whose inverse is not always defined. So, the move  $\tilde{L}$  is better suited for definition of 3–deformation invariants than the move  $L$ . The preceding paragraphs

FIGURE 7. A “straight” move  $L$  near a vertexFIGURE 8. A “twisted” move  $L$  near a vertex — the move  $\tilde{L}$ 

immediately imply that two special 2–polyhedra with at least two true vertices are 3–deformation equivalent if and only if they can be related by a finite sequence of moves  $T$ ,  $N$ ,  $\tilde{L}$  and their inverses.

How to construct 3–deformation invariants for special 2–complexes? Our basic construction, namely a state sum of Turaev–Viro type, remains as before. So, we have a finite set  $\mathcal{F}$  with an involution, providing colourings  $\Phi_{\mathcal{F}}(\mathcal{P})$  of 2–strata of special spines. For simplicity, we use no edge colours. The six coloured wings around each true vertex give rise to  $6j$ –symbols, and each coloured 2–stratum has a weight, and equivalence classes of the  $6j$ –symbols and colour weights are variables of some polynomial ring  $R$  over a field. However, in the context of special 2–polyhedra that are not special spines of 3–manifolds, we have an additional property of 2–strata that we may use in our construction: A 2–stratum can be either twisted or straight. Therefore, we add variables  $t(f)$  to  $R$ , for all  $f \in \mathcal{F}$ , and  $t(f) = t(-f)$  for all  $f \in \mathcal{F}$ . For any 2–stratum  $c$  of a special 2–polyhedron  $\mathcal{P}$ , let  $\epsilon(c) = 0$  if  $c$  is straight and  $\epsilon(c) = 1$  if  $c$  is twisted. Now, we form a modified **state sum**  $TV_{\mathcal{F}}(\mathcal{P})$  as follows:

$$TV_{\mathcal{F}}(\mathcal{P}) = \sum_{\varphi \in \Phi_{\mathcal{F}}(\mathcal{P})} \left( \prod_{C \in \mathcal{C}(\mathcal{P})} t(\varphi(C))^{\epsilon(C)} w(\varphi(C)) \right) \cdot \left( \prod_{v \in \mathcal{V}(\mathcal{P})} v^{\varphi, \psi} \right) \in R$$

Note that a very similar construction was used by Turaev for his theory of shadows of 4–manifolds [97].

According to Figures 4 and 8, the moves  $N$  and  $\tilde{L}$  change some 2–strata from straight to twisted and vice versa. Changing a 2–stratum of colour  $f$  from straight to twisted evidently changes the corresponding summand in the state sum by the

factor  $t(f)$ , and a change from twisted to straight changes the summand state sum by a factor  $t(f)^{-1}$ . Hence, if we use the relation  $t(f)^2 = 1$  for all  $f \in \mathcal{F}$ , both the change from straight to twisted and vice versa contribute a factor  $t(f)$  to the summands of the state sum.

Let  $I_{\mathcal{F}}^{AC} \subset R$  be the ideal generated by the Biedenharn–Elliott equations (see Eqn. 1), by  $t(f)^2 - 1$  for all  $f \in \mathcal{F}$ , by

$$(3) \quad \left| \begin{array}{ccc} a & b & c \\ f & e & d \end{array} \right| = \sum_{x \in \mathcal{F}} w(x)t(a)t(b)t(c)t(d)t(e)t(f)t(x) \left| \begin{array}{cc} f & c-x \\ -d & a \end{array} \right| \left| \begin{array}{cc} f & -a \\ c & d \end{array} \right| \left| \begin{array}{cc} x & e \\ -a & d \end{array} \right|$$

and by

$$(4) \quad \left| \begin{array}{ccc} a & b & c \\ f & e & d \end{array} \right| = \sum_{g, h \in \mathcal{F}} w(g)w(h)t(g)t(c) \left| \begin{array}{cc} a & b \\ f & e \end{array} \right| \left| \begin{array}{cc} g & c \\ a & f \end{array} \right| \left| \begin{array}{cc} h & e \\ -a & h \end{array} \right|$$

for all  $a, b, c, d, e, f \in \mathcal{F}$ .

**THEOREM 13.** *For special 2-polyhedra with at least two true vertices, the coset*

$$tv_{\mathcal{F}}(\mathcal{P}) = TV_{\mathcal{F}}(\mathcal{P}) + I_{\mathcal{F}}^{AC} \in R/I_{\mathcal{F}}^{AC}$$

*only depends on the 3-deformation type of  $\mathcal{P}$ .*

**PROOF.** For special 2-polyhedra with at least two true vertices, 3-deformation equivalence is generated by the moves  $T$ ,  $N$  and  $\tilde{L}$ . Invariance under the move  $T$  is provided by the Biedenharn–Elliott equations. Invariance under the move  $N$  (resp.  $\tilde{L}$ ) is provided by  $t(f)^2 - 1$  together with the generators 3 (resp. 4), which can be seen by an inspection of Figure 4 (resp. Figure 8).  $\square$

Modifications of Turaev–Viro invariants look like tools to detect counterexamples of the Andrews–Curtis conjecture. However, it follows from a result of Bobtcheva and Quinn [6] that classical Turaev–Viro invariants will fail. In fact, Bobtcheva and Quinn show that a large class of invariants for the Andrews–Curtis problem is determined by homology, by some explicit formula. This is mainly since these invariants are multiplicative under connected sums.

However, by Theorem 3, ideal Turaev–Viro invariants are not multiplicative under connected sums, in contrast to the classical numerical Turaev–Viro invariants that arise from representation theory of Quantum Groups. Hence, there is some hope ideal 3-deformation invariants are not just determined by homology.

We succeeded to compute Gröbner bases of  $I_{\mathcal{F}}^{AC}$  in various cases, partially involving simplifying assumptions similar to those studied in Section 2.4. We computed our invariants for various notorious potential counterexamples of the Andrews–Curtis conjecture, but we were not able to disprove the conjecture. However, ideal Turaev–Viro invariants remain a possible way towards a disprove of the Andrews–Curtis conjecture.

#### 4. Ideal link invariants

In previous sections, we constructed homeomorphism invariants of 3-manifolds or 3-deformation invariants of special 2-polyhedra taking values in quotients of polynomial rings. In this section, we briefly discuss an approach towards the construction of new invariants of knots and links in  $\mathbb{S}^3$ .

The basic strategy is as follows. Let a link  $L \subset \mathbb{S}^3$  be represented by a link diagram  $\mathcal{D}$  in  $\mathbb{S}^2$ . Read off a state sum  $|\mathcal{D}| \in R$ , where  $R$  is some polynomial ring. Any two diagrams of  $L$  are related by sequences of Reidemeister moves. The Reidemeister moves give rise to generators of some ideal  $\Omega \subset R$ , such that the state sums of two diagrams of  $L$  only differ by an element of  $\Omega$ . Hence,  $|\mathcal{D}| + \Omega \in R/\Omega$  only depends on the ambient isotopy type of  $L$ .

A diagram  $\mathcal{D} \subset \mathbb{S}^2$  of a link  $L$  is formed by **bridges** (i.e., arcs in the projection of  $L$  that start and end at undercrossings but may pass several overcrossings) and decomposes  $\mathbb{S}^2$  into 2-cells, to which we refer as **regions** of  $\mathcal{D}$ . Let  $\mathcal{F}$  and  $\mathcal{G}$  be finite sets. An  $(\mathcal{F}, \mathcal{G})$ -colouring of  $\mathcal{D}$  assigns an element of  $\mathcal{F}$  to any bridge and an element of  $\mathcal{G}$  to any region of  $\mathcal{D}$ .

Hence, at any crossing of a diagram with an  $(\mathcal{F}, \mathcal{G})$ -colouring, three elements of  $\mathcal{F}$  (corresponding to the overcrossing bridge and the two bridges starting at the undercrossing) and four elements of  $\mathcal{G}$  occur. To each triple  $x, y, z \in \mathcal{F}$  and quadruple  $A, B, C, D \in \mathcal{G}$ , a crossing as in Figure 9 gives rise to a variable  $X(x, y, z; A, B, C, D)$  of some polynomial ring  $R$  over  $\mathbb{Q}$ . The variables of  $R$  shall respect the symmetry of a crossing. Hence, we identify  $X(x, y, z; A, B, C, D)$  with  $X(x, z, y; C, D, A, B)$ . In addition, we have a variable  $|A|$  of  $R$  for any  $A \in \mathcal{G}$ .

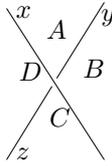


FIGURE 9. A coloured crossing of a link diagram

Now, for any colouring  $\phi$  of  $\mathcal{D}$ , we let  $\mathcal{D}^\phi \in R$  be the product of all variables  $X(\dots; \dots)$  and  $|\cdot|$  corresponding to the coloured crossings and regions of  $\mathcal{D}$ . By the symmetry of  $X(\dots; \dots)$ ,  $\mathcal{D}^\phi$  is well-defined. Let the **state sum**  $|\mathcal{D}| \in R$  of  $\mathcal{D}$  be the sum of  $\mathcal{D}^\phi$  over all  $(\mathcal{F}, \mathcal{G})$ -colourings of  $\mathcal{D}$ .

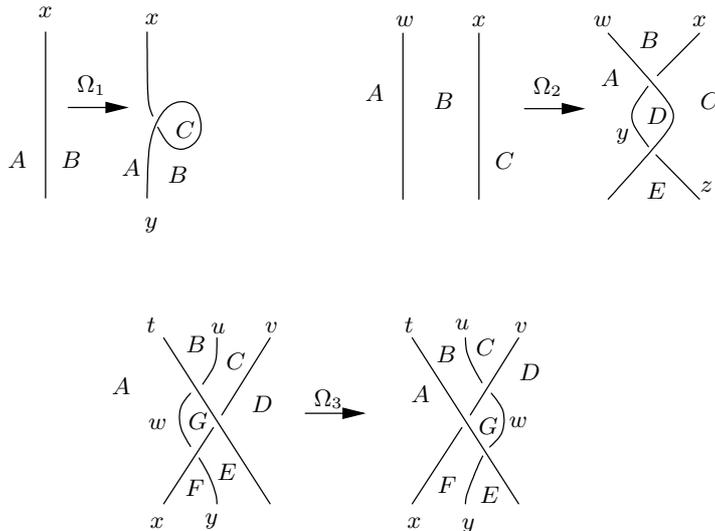


FIGURE 10. The Reidemeister moves

Let the ideal  $\Omega \in R$  be generated by the following polynomials, for all  $t, \dots, z \in \mathcal{F}$ ,  $A, \dots, G \in \mathcal{G}$ , where  $\delta_{x,y}$  denotes the Kronecker symbol.

$$\begin{aligned} & \delta_{x,y} - \sum_{C \in \mathcal{G}} |C|X(y, x, x; A, B, C, B) \\ & \delta_{x,z}\delta_{B,E} - \sum_{D \in \mathcal{G}} \sum_{y \in \mathcal{F}} |D|X(w, x, y; B, C, D, A)X(w, y, z; A, D, C, E) \\ & \sum_{G \in \mathcal{G}} \sum_{w \in \mathcal{F}} |G|X(t, u, w; B, C, G, A)X(t, v, x; C, D, E, G)X(x, w, y; A, G, E, F) \\ & - \sum_{G \in \mathcal{G}} \sum_{w \in \mathcal{F}} |G|X(v, u, w; B, C, D, G)X(t, v, x; B, G, F, A)X(t, w, y; G, D, E, F) \end{aligned}$$

The three types of generators correspond to the three Reidemeister moves, as indicated in Figure 10. One obtains

**THEOREM 14** (and Definition). *For any link  $L \subset \mathbb{S}^3$  with a diagram  $\mathcal{D}$ , the coset  $|L| = |\mathcal{D}| + \Omega \in R/\Omega$  is independent of the choice of  $\mathcal{D}$ .*

Now, the following steps are as in the case of ideal Turaev–Viro invariants:

- (1) Compute a Gröbner basis  $G$  of  $\Omega$ .
- (2) Compute the state sum  $|\mathcal{D}|$  of any link diagram  $\mathcal{D}$  of a link  $L$ .
- (3) Compute  $\text{rem}(\mathcal{D}; G)$ , which is a unique representative of  $|L|$ .

Unfortunately, the computation of Gröbner bases seems to be much more difficult for  $\Omega$  than for the Turaev–Viro ideals studied in the previous sections. So, there is urgent need for simplifying assumptions. We tested various simplifying assumptions motivated by the Jones polynomial [44]. However, we did not succeed to find a proper generalisation of the Jones polynomial. In spite of that negative outcome, we believe that with better simplifying assumptions one should obtain new strong link invariants.



## Fast Computation of Secondary Invariants

### 1. Motivation

Ideal Turaev–Viro invariants (see Section 2 in Chapter 6) have many symmetries. For simplicity, let us assume that the edge colours are trivial. Let  $\mathcal{F}$  be the set of 2–strata colours. Let a group  $G$  act on  $\mathcal{F}$  by permutation of colours. If there are no simplifying assumptions then this action induces an action on the set of  $6j$ -symbols and colour weights, hence, an action on the variables of the polynomial ring  $R$  that we introduced in the definition of ideal Turaev–Viro invariants. Let  $R^G = \{p \in R: g.r = r, \forall g \in G\}$  be the *invariant ring* of that action. Obviously,  $R^G$  is a sub-algebra of  $R$ .

Let  $P$  be any special 2-polyhedron. We obtain an action of  $G$  on the set of all  $\mathcal{F}$ -colourings of  $P$ . The state sum  $TV_{m,1}(P)$  is a sum over all  $\mathcal{F}$ -colourings of  $P$ . Hence,  $TV_{m,1}(P) \in R^G$ . Let  $I_{m,1} \subset R$  be the Turaev–Viro ideal, and let  $I_{m,1}^G = I_{m,1} \cap R^G$  be the invariant ideal. We readily obtain the following estimate.

LEMMA 7. *For any closed 3-manifold  $M$  with  $\tilde{c}(M) > 1$  and a special spine  $P$  of  $M$ , we have*

$$\tilde{c}(M) \geq \max \left\{ \deg_w (TV_{m,1}(P) + I_{m,1}^G) - 1, \deg_{6j} (TV_{m,1}(P) + I_{m,1}^G) \right\}.$$

□

Since  $I_{m,1}^G \subset I_{m,1}$ , we have  $TV_{m,1}(P) + I_{m,1}^G \subset tv_{m,1}(M)$ , hence,

$$\begin{aligned} \deg_w (TV_{m,1}(P) + I_{m,1}^G) &\geq \deg_w (tv_{m,1}(M)) \text{ and} \\ \deg_{6j} (TV_{m,1}(P) + I_{m,1}^G) &\geq \deg_{6j} (tv_{m,1}(M)). \end{aligned}$$

So, Lemma 6 is an improvement of Lemma 6.

Using SINGULAR [29], we succeeded to compute  $R^G$  and  $I_{m,1}^G \subset R^G$ , in some basic cases. This did not provide non-trivial bounds for the complexity of 3–manifolds. However, this attempt was the motivation to develop and implement a new algorithm for the computation of invariant rings. We recall classical results on invariant rings in the next section.

Our algorithm is based on three new ideas, as described in Section 3 or in our preprint [56]. The benchmark tests that we expose in Section 4 show that the implementation of our algorithm in the SINGULAR library `finvar.lib` marks a dramatic improvement in the manageable problem size, compared with previous algorithms in SINGULAR or in MAGMA [11]. After the first version of our manuscript [56] was posted, there was a new release of MAGMA with a new algorithm due to Kemper, that seems to be competitive with our algorithm, but was not described yet. A bit later, we came up with a second new algorithm that in some cases provides a further drastic improvement; this is described in Chapter 8.

### 2. Non-modular invariant rings

Let  $G$  be a finite group, linearly acting on a polynomial ring  $R$  with  $n$  variables over some field  $K$ . We denote the action of  $g \in G$  on  $r \in R$  by  $g.r \in R$ .

Let  $R^G = \{r \in R: g.r = r, \forall g \in G\}$  be the . Obviously, it is a sub-algebra of  $R$ , and one would like to compute generators for  $R^G$ . We study here the **non-modular case**, i.e., the characteristic of  $K$  does not divide the order of  $G$ . Note that according to [45], algorithms for the non-modular case are useful also in the modular case.

For any subset  $S \subset R$ , we denote by  $\langle\langle S \rangle\rangle \subset R$  the sub-algebra generated by  $S$ , and by  $\langle S \rangle \subset R$  the ideal generated by  $S$ . It is well known [25] that there are  $n$  (the number of variables) algebraically independent homogeneous invariant polynomials  $P = \{p_1, \dots, p_n\} \subset R^G$  such that  $R^G$  is a finitely generated free  $\langle\langle P \rangle\rangle$ -module (this is not always true in the modular case!). The elements of  $P$  are called **primary invariants**. Of course, they are not uniquely determined. There are various algorithms to compute primary invariants [45]. Since the primary invariants are algebraically independent, the sub-algebra  $\langle\langle P \rangle\rangle$  is isomorphic to a polynomial ring with  $n$  variables. It is called **(homogeneous) Noetherian normalization** of  $R^G$ .

Let  $S \subset R^G$  be a minimal set of homogeneous  $\langle\langle P \rangle\rangle$ -module generators of  $R^G$ . The elements of  $S$  are called **secondary invariants**. The number of secondary invariants increases with the product of the degrees of the primary invariants. Hence, it is advisable to minimize the degrees of the primary invariants. **Irreducible** secondary invariants are those non-constant secondary invariants that can not be written as a polynomial expression in the primary invariants and the other secondary invariants. The set of secondary invariants is not unique, even if one fixes the primary invariants. It is easy to see that one can choose secondary invariants so that all of them are power products of irreducible secondary invariants.

Our aim is to present a new algorithm for the computation of (irreducible) homogeneous secondary invariants, if homogeneous primary invariants  $P$  are given. The key theorem for our algorithm concerns Gröbner bases and holds in arbitrary characteristic; however, the algorithm assumes that we are in the non-modular case.

Since we are in the non-modular case, we can use the Reynolds operator  $\text{Rey}: R \rightarrow R^G$ , which is defined by

$$\text{Rey}(r) = \frac{1}{|G|} \sum_{g \in G} g.r$$

for  $r \in R$ . By construction, the restriction of the Reynolds operator to  $R^G$  is the identity. The Reynolds operator does not commute with the ring multiplication. However, it does commute, if one of the factors is invariant, as in the following lemma. In other words, the Reynolds operator is a module homomorphism, where  $R$  and  $R^G$  are considered as  $R^G$ -modules.

LEMMA 8. *Let  $p \in R$  and  $q \in R^G$ . Then,  $\text{Rey}(pq) = \text{Rey}(p)q$ .*

PROOF. For any  $g \in G$ , we have  $g.(pq) = (g.p)(g.q)$ . But  $q \in R^G$ , and thus  $g.(pq) = (g.p)q$ . It follows

$$\begin{aligned} \text{Rey}(pq) &= \frac{1}{|G|} \sum_{g \in G} g.(pq) \\ &= \frac{1}{|G|} \sum_{g \in G} (g.p)q = \text{Rey}(p)q \end{aligned}$$

□

Let  $B_d \subset R^G$  be the images under the Reynolds operator of all monomials of  $R$  of degree  $d$  that do not occur as leading monomials in  $\langle P \rangle$ . Using a Gröbner basis of  $\langle P \rangle$ ,  $B_d$  is easy to compute. It is well known that one can find a system of homogeneous secondary invariants of degree  $d$  in  $B_d$  (see Lemma 3.5.1 and Remark 3.5.3

in [22]). But how can one determine what elements of  $B_d$  are eligible as secondary invariants? The solution is provided by the following lemma. For completeness, we give a proof, although it is well known [88].

LEMMA 9. *Let  $d > 0$ , let  $S_0, S_1, S_2, \dots, S_{d-1} \subset R^G$  be the homogeneous secondary invariants of degree  $0, 1, 2, \dots, d-1$ , let  $s_1, \dots, s_m \in R^G$  be some homogeneous secondary invariants of degree  $d$ , and let  $b \in B_d$ . We can choose  $b$  as a new homogeneous secondary invariant of  $R^G$ , if and only if  $b$  is not contained in the ideal  $\langle P \cup \{s_1, \dots, s_m\} \rangle \subset R$ .*

PROOF. By definition, we can choose  $b$  as a new homogeneous secondary invariant of  $R^G$ , if and only if  $b$  is not contained in the  $\langle\langle P \rangle\rangle$ -module generated by  $S_0 \cup S_1 \cup \dots \cup S_{d-1} \cup \{s_1, \dots, s_m\}$ . Since  $b$  is homogeneous of degree  $d$ , if  $b$  belongs to the  $\langle\langle P \rangle\rangle$ -module generated by  $S_0 \cup S_1 \cup \dots \cup S_{d-1} \cup \{s_1, \dots, s_m\}$  then it is an element of  $\langle P \cup \{s_1, \dots, s_m\} \rangle \subset R$ .

Conversely, let  $b \in \langle P \cup \{s_1, \dots, s_m\} \rangle \subset R$  be a homogeneous invariant polynomial of degree  $d$ . As an element of the ideal, we may write it as a finite sum  $b = \sum_i p_i q_i$ , with homogeneous polynomials  $p_i \in R$  and  $q_i \in P \cup \{s_1, \dots, s_m\} \subset R^G$ . Since  $b$  is invariant, since the Reynolds operator is additive and by Lemma 8, we obtain  $b = \text{Rey}(b) = \sum_i \text{Rey}(p_i q_i) = \sum_i \text{Rey}(p_i) q_i$ . If  $q_i \in P$  then  $\text{Rey}(p_i)$  is a non-constant homogeneous invariant polynomial of degree at most  $d-1$ . Hence, by hypothesis,  $\text{Rey}(p_i)$  belongs to the  $\langle\langle P \rangle\rangle$ -module generated by  $S_0 \cup S_1 \cup \dots \cup S_{d-1}$ . If  $q_i \in \{s_1, \dots, s_m\}$  then  $p_i = \text{Rey}(p_i)$  is a constant. Hence, in both cases,  $\text{Rey}(p_i) q_i$  belongs to the  $\langle\langle P \rangle\rangle$ -module generated by  $S_0 \cup S_1 \cup \dots \cup S_{d-1} \cup \{s_1, \dots, s_m\}$ ; and so does  $b$ .  $\square$

By Theorem 11, to test ideal membership, it suffices to know a Gröbner basis for the ideal and to compute the remainder with respect to this Gröbner basis. We thus obtain the following very basic algorithm for finding homogeneous secondary invariants of degree  $d$ .

#### BASIC ALGORITHM

- (1) Let  $S_d = \emptyset$ . Let  $\mathcal{G}$  be a Gröbner basis of  $\langle P \rangle$ .
- (2) For all  $b \in B_d$ :  
If  $b \notin \langle P \cup S_d \rangle$  (which is tested by reduction modulo  $\mathcal{G}$ ) then replace  $S_d$  by  $S_d \cup \{b\}$ ; compute a Gröbner basis of  $\langle P \cup S_d \rangle$  and replace  $\mathcal{G}$  with it.
- (3) Return  $S_d$ .

There are several ways to improve this algorithm. One way is an application of Molien's Theorem [88], [45], [37]. We will not go into details here. Molien's Theorem allows to compute the number  $m_d$  of secondary invariants of degree  $d$ . In other words, if in the above algorithm we got  $m_d$  secondary invariants, we can immediately break the loop in Step (2).

We also would like to see which of the secondary invariants in  $S_d$  are irreducible, since these, together with  $P$ , generate  $R^G$  as a sub-algebra of  $R$ . Let  $IS_i \subset S_i$  be the irreducible secondary invariants, for  $i = 1, \dots, d-1$ . For computing the irreducibles in degree  $d$ , usually one first forms all power products of degree  $d$  of elements of  $IS_1 \cup IS_2 \cup \dots \cup IS_{d-1}$  and chooses from them as many secondary invariants as possible (see [45], [47] or [22]). If there are further secondary invariants (which we know from computation of  $m_d$ ), then one proceeds as above with  $B_d$ , and obtains all *irreducible* secondary invariants  $IS_d$  of degree  $d$ . So, the algorithm is as follows.

#### REFINED ALGORITHM

- (1) Compute  $m_d$ . Let  $S_d = IS_d = \emptyset$  and let  $\mathcal{G}$  be a Gröbner basis of  $\langle P \rangle$ .
- (2) For all power products  $b$  of degree  $d$  of elements of  $IS_1 \cup IS_2 \cup \dots \cup IS_{d-1}$ :  
(a) If  $b \notin \langle P \cup S_d \rangle$  (which is tested using  $\mathcal{G}$ ) then replace  $S_d$  by  $S_d \cup \{b\}$ ; compute a Gröbner basis of  $\langle P \cup S_d \rangle$  and replace  $\mathcal{G}$  with it.

- (b) If  $|S_d| = m_d$  then break and return  $(S_d, IS_d)$ .
- (3) For all  $b \in B_d$ :
  - (a) If  $b \notin \langle P \cup S_d \rangle$  (which is tested using  $\mathcal{G}$ ) then replace  $S_d$  by  $S_d \cup \{b\}$ , and  $IS_d$  by  $IS_d \cup \{b\}$ ; compute a Gröbner basis of  $\langle P \cup S_d \rangle$  and replace  $\mathcal{G}$  with it.
  - (b) If  $|S_d| = m_d$  then break and return  $(S_d, IS_d)$ .

Eventually,  $S_d$  contains homogeneous secondary invariants of degree  $d$ , and  $IS_d$  contains the irreducible ones. In this form, the algorithm has been implemented in 1998 by A. Heydtmann [37] as the procedure `secondary_char0` of the library `finvar` of SINGULAR. In Step (2), the ideal membership is tested by computing the remainder modulo some Gröbner basis of the ideal. This ideal changes once a new secondary invariant has been found. So, the algorithm involves many Gröbner basis computations. This is its main disadvantage and limits the applicability of the Basic and the Refined Algorithm.

An alternative algorithm, but with essentially the same structure, was proposed by Kemper and Steel (see [45], [47] or [22]) and implemented in MAGMA [11]. Here, new secondary invariants are detected not by a general solution of the ideal membership problem but by testing linear independency of normal forms with respect to a Gröbner basis of  $\langle P \rangle$ , hence, by Linear Algebra. This algorithm only involves one Gröbner basis computation, namely for the ideal  $\langle P \rangle$ . But for computing some of the invariant rings that arise in our study of ideal Turaev–Viro invariants, this does not suffice either.

### 3. The New Algorithm

The main feature of our new algorithm is that, after computing some (homogeneous) Gröbner basis of  $\langle P \rangle$ , we can directly write down a *homogeneous Gröbner basis up to degree  $d$*  of  $\langle P \cup S_d \rangle$ , once a new secondary invariant of degree  $d$  has been found. We can do so without any lengthy computations (in contrast to [37]), and we also avoid to deal with huge systems of linear equations (in contrast to [47], [45], [22]). This allows to solve the ideal membership problem in a very quick way. We recall the notion of “homogeneous Gröbner bases up to degree  $d$ ” in the following paragraphs. At the end of the section, we provide our key theorem and formulate our new algorithm.

For  $p \in R$ , let  $lm(p)$  the **leading monomial** of  $p$ , let  $lc(p)$  be the coefficient of  $lm(p)$  in  $p$ , and let  $lt(p) = lc(p)lm(p)$  be the **leading term** of  $p$ . The least common multiple is denoted by  $LCM(\cdot, \cdot)$ . Now we can recall the definition of the  $S$ -**polynomial** of  $p, q \in R$ :

$$S(p, q) = \frac{LCM(lm(p), lm(q))}{lt(p)}p - \frac{LCM(lm(p), lm(q))}{lt(q)}q$$

Obviously, the  $S$ -polynomial of  $p$  and  $q$  belongs to the ideal  $\langle p, q \rangle \subset R$ . The leading terms of  $p$  and  $q$  are canceling one another, so, the leading monomial of  $S(p, q)$  corresponds to monomials of  $p$  or  $q$  that are not leading. The following result can be found, e.g., in [27] or [59].

**THEOREM 15** (Buchberger’s Criterion). *A set  $g_1, \dots, g_k \in R$  of polynomials is a Gröbner basis of the ideal  $\langle g_1, \dots, g_k \rangle \subset R$  if and only if  $\text{rem}(S(g_i, g_j); g_1, \dots, g_k) = 0$  for all  $i, j = 1, \dots, k$ .  $\square$*

Buchberger’s Criterion directly leads to Buchberger’s algorithm for the construction of a Gröbner basis of an ideal: One starts with any generating set of the ideal. If the remainder modulo the generators of the  $S$ -polynomial of some pair of generators does not vanish, then the remainder is added as a new generator. This

will be repeated until all  $S$ -polynomials reduce to 0; it can be shown that this will eventually be the case, after finitely many steps.

Here, we are in a special situation: We work with homogeneous polynomials. It is easy to see that if  $p$  and  $q$  are homogeneous then so is  $S(p, q)$ , and its degree is higher than the maximum of the degrees of  $p$  and  $q$ , unless the leading monomials of  $p$  and  $q$  are divisible by each other. If  $p, g_1, g_2, \dots, g_k \in R$  are homogeneous then so is  $\text{rem}(p; g_1, \dots, g_k)$ . Moreover, either  $\text{rem}(p; g_1, \dots, g_k) = 0$  or  $\deg(\text{rem}(p; g_1, \dots, g_k)) = \deg(p)$ . For computing  $\text{rem}(p; g_1, \dots, g_k)$ , only those  $g_i$  play a role with  $\deg(g_i) \leq \deg(p)$ , for  $i = 1, \dots, k$ . It follows: If an ideal  $I \subset R$  is homogeneous (i.e., it can be generated by homogeneous polynomials) then it has a Gröbner basis of homogeneous polynomials. Such a Gröbner basis can be constructed degree-wise.

**DEFINITION 9.** A finite set  $\{g_1, \dots, g_k\} \subset R$  of homogeneous polynomials is a **homogeneous Gröbner basis up to degree  $d$**  of the ideal  $\langle g_1, \dots, g_k \rangle$ , if

$$\text{rem}(S(g_i, g_j); g_1, \dots, g_k) = 0$$

or  $\deg(S(g_i, g_j)) > d$ , for all  $i, j = 1, \dots, k$ .

**LEMMA 10.** Let  $\{g_1, \dots, g_k\} \subset R$  be a homogeneous Gröbner basis up to degree  $d$ , and let  $p \in R$  be a homogeneous polynomial of degree at most  $d$ . Then,  $p \in \langle g_1, \dots, g_k \rangle$  if and only if  $\text{rem}(p; g_1, \dots, g_k) = 0$ .

**PROOF.** The paragraph preceding the definition implies that  $\{g_1, \dots, g_k\}$  can be extended to a Gröbner basis  $\mathcal{G}$  of  $\langle g_1, \dots, g_k \rangle$  by adding homogeneous polynomials whose degrees exceed  $d$ . Since  $\deg(p) \leq d$ , we have  $\text{rem}(p; \mathcal{G}) = \text{rem}(p; g_1, \dots, g_k)$ . Since  $p \in \langle \mathcal{G} \rangle$  if and only if  $\text{rem}(p; \mathcal{G}) = 0$  by Theorem 11, the result follows.  $\square$

We see that in order to do Step (2) in the BASIC ALGORITHM (or the corresponding steps in the REFINED ALGORITHM) it suffices to know a homogeneous Gröbner basis up to degree  $d$  of  $\langle P \cup S_d \rangle$ . Our key theorem states that this Gröbner basis can be constructed iteratively, as follows.

**THEOREM 16.** Let  $\mathcal{G} \subset R$  be a homogeneous Gröbner basis up to degree  $d$  of  $\langle \mathcal{G} \rangle$ . Let  $p \in R$  be a homogeneous polynomial of degree  $d$ , and  $p \notin \langle \mathcal{G} \rangle$ . Then  $\mathcal{G} \cup \{\text{rem}(p; \mathcal{G})\}$  is a homogeneous Gröbner basis up to degree  $d$  of  $\langle \mathcal{G} \cup \{p\} \rangle$ .

**PROOF.** Let  $r = \text{rem}(p; \mathcal{G})$ . Since  $p \notin \langle \mathcal{G} \rangle$  and all polynomials are homogeneous, we have  $r \neq 0$ ,  $\deg(r) = d$ , and  $\langle \mathcal{G} \cup \{p\} \rangle = \langle \mathcal{G} \cup \{r\} \rangle$ .

By hypothesis, the  $S$ -polynomials of pairs of elements of  $\mathcal{G}$  are of degree  $> d$  or reduce to 0 modulo  $\mathcal{G}$ . We now consider the  $S$ -polynomials of  $r$  and elements of  $\mathcal{G}$ . Let  $g \in \mathcal{G}$ . By definition of the remainder,  $lm(r)$  does not divide  $lm(g)$ . Therefore the  $S$ -polynomial of  $r$  and  $g$  is of degree  $> d = \deg(r)$ . Thus the claim follows.  $\square$

We obtain the  
NEW ALGORITHM

- (1) Compute  $m_d$  and a homogeneous Gröbner basis  $\mathcal{G}$  of  $\langle P \rangle$ .  
Let  $S_d = IS_d = \emptyset$ .
- (2) For all power products  $b$  of degree  $d$  of elements of  $IS_1 \cup IS_2 \cup \dots \cup IS_{d-1}$ :
  - (a) If  $\text{rem}(b; \mathcal{G}) \neq 0$  then replace  $S_d$  by  $S_d \cup \{b\}$ , and  $\mathcal{G}$  by  $\mathcal{G} \cup \{\text{rem}(b; \mathcal{G})\}$ .
  - (b) If  $|S_d| = m_d$  then break and return  $(S_d, IS_d)$ .
- (3) For all  $b \in B_d$ :
  - (a) If  $\text{rem}(b; \mathcal{G}) \neq 0$  then replace  $S_d$  by  $S_d \cup \{b\}$ ,  $IS_d$  by  $IS_d \cup \{b\}$ , and  $\mathcal{G}$  by  $\mathcal{G} \cup \{\text{rem}(b; \mathcal{G})\}$ .
  - (b) If  $|S_d| = m_d$  then break and return  $(S_d, IS_d)$ .

Although the NEW ALGORITHM is a dramatic improvement of the REFINED ALGORITHM, one should take more care in Step (2) of the NEW ALGORITHM. It simply says “For all power products  $b$  of degree  $d$  of elements of  $IS_1 \cup IS_2 \cup \dots \cup IS_{d-1}$ ”. Two questions arise:

- (1) How shall one generate the power products?
- (2) Is it necessary to generate *all* possible power products, or can one restrict the search?

In very complex computations, the number of power products is gigantic. But usually only a small proportion of them will be eligible as secondary invariant. So, for saving computer’s memory, it is advisable to generate the power products one after the other (or in small packages), rather than generating all power products at once; this answers Question (1).

Apparently Question (2) was never addressed in the literature. However, it turns out that a careful choice of power products provides another dramatic improvement of the performance of the algorithm. Our choice is based on the following lemma. It seems to be well-known to the experts. However, we did not find it in the literature, and it was not used in the `finvar` library of SINGULAR up to version 3-0-1.

LEMMA 11. *Assume that secondary invariants of degree  $< d$  are computed such that all of them are power products of irreducible secondary invariants. In the quest for reducible homogeneous secondary invariants of degree  $d$ , it suffices to consider power products of the form  $i \cdot s$ , where  $i$  is a homogeneous irreducible secondary invariant of degree  $< d$ , and  $s$  is some secondary invariant of degree  $d - \deg(i)$ .*

PROOF. Let  $p \in R$  be a power product of degree  $d$  of irreducible secondary invariants. Hence, it can be written as  $p = iq$ , with an irreducible homogeneous secondary invariant  $i$  of degree  $< d$  and some homogeneous  $G$ -invariant polynomial  $q$  of degree  $d - \deg(i)$  (we do not use here that  $q$  is a power product itself).

Recall that the secondary invariants generate the invariant ring as a  $\langle\langle P \rangle\rangle$ -module. Hence one can rewrite  $q = q_0 + k_1 s_1 + \dots + k_t s_t$ , where  $q_0 \in \langle P \rangle \cap R^G$  (i.e.,  $q_0$  belongs to the  $\langle\langle P \rangle\rangle$ -module generated by secondary invariants of lower degree),  $k_1, \dots, k_t \in K$ , and  $s_1, \dots, s_t$  are homogeneous secondary invariants of degree  $\deg(q)$ . We obtain  $p = iq_0 + k_1(i s_1) + \dots + k_t(i s_t)$ . Hence, rather than choosing  $p$  as a  $\langle\langle P \rangle\rangle$ -module generator of  $R^G$ , we may choose  $i s_1, \dots, i s_t$ , which, by induction, are power products of irreducible secondary invariants.  $\square$

#### IMPROVED NEW ALGORITHM

- (1) Compute  $m_d$ . Let  $\mathcal{G}$  be a Gröbner basis of  $\langle P \rangle$ . Let  $S_d = IS_d = \emptyset$ .
- (2) For all products  $b = i \cdot s$  with  $i \in IS_1 \cup \dots \cup IS_{d-1}$  and  $s \in S_{d-\deg(i)}$ :
  - (a) If  $\text{rem}(b; \mathcal{G}) \neq 0$  then replace  $S_d$  by  $S_d \cup \{b\}$ , and  $\mathcal{G}$  by  $\mathcal{G} \cup \{\text{rem}(b; \mathcal{G})\}$ .
  - (b) If  $|S_d| = m_d$  then break and return  $(S_d, IS_d)$ .
- (3) For all  $b \in B_d$ :
  - (a) If  $\text{rem}(b; \mathcal{G}) \neq 0$  then replace  $S_d$  by  $S_d \cup \{b\}$ ,  $IS_d$  by  $IS_d \cup \{b\}$ , and  $\mathcal{G}$  by  $\mathcal{G} \cup \{\text{rem}(b; \mathcal{G})\}$ .
  - (b) If  $|S_d| = m_d$  then break and return  $(S_d, IS_d)$ .

This is the algorithm that is implemented as `secondary_char0` in the library `finvar` of SINGULAR 3-0-2 [29], released in Juli 2006. In Step (2), the secondary invariant  $s$  may be a non-trivial powerproduct itself, hence, can be expressed as  $s = i_s s'$ , where  $i_s$  is an irreducible secondary invariant and  $s'$  is (by induction) some other secondary invariant. Of course one should consider only one of the two products  $i_s(i_s s')$  and  $i(i_s s')$  in the enumeration.

Often one is only interested in the irreducible secondary invariants, which, together with the primary invariants, generate the invariant ring as a sub-algebra.

Therefore we implemented yet another version of the IMPROVED NEW ALGORITHM in SINGULAR 3-0-2, namely `irred_secondary_char0`. This algorithm computes irreducible secondary invariants, but does not explicitly compute the reducible secondary invariants. That works as follows.

Let  $\mathcal{G}_P$  be a Gröbner basis of  $\langle P \rangle$ . In Step (2)(a) of the IMPROVED NEW ALGORITHM, one replaces  $S_d$  by  $S_d \cup \{\text{rem}(b; \mathcal{G}_P)\}$ , rather than by  $S_d \cup \{b\}$ . In Step (3)(a) one replaces  $S_d$  by  $S_d \cup \{\text{rem}(b; \mathcal{G}_P)\}$  and  $IS_d$  by  $IS_d \cup \{b\}$ . In the end,  $S_d$  does not contain secondary invariants, but *normal forms* of secondary invariants with respect to  $\mathcal{G}_P$ . Since  $\text{rem}(\text{rem}(p_1; \mathcal{G}_P) \cdot \text{rem}(p_2; \mathcal{G}_P); \mathcal{G}_P) = \text{rem}(p_1 \cdot p_2; \mathcal{G}_P)$  and since a reduction modulo  $\mathcal{G}$  in Steps (2)(a) and (3)(a) also comprises a reduction modulo  $\mathcal{G}_P$ , this maintains all informations that one needs for determining how many secondary invariants are reducible in Step (2) and for finding the irreducible secondary invariants in Step (3). So in the end,  $IS_d$  contains the irreducible secondary invariants in degree  $d$ .

This third new idea for our algorithm very often saves much memory and computation time, as can be seen in Table 1 in Examples (1) and (6)–(9). In Example (8), we can compute the irreducible secondary invariants although the computation of all 31104 secondary invariants exceeds the resources.

An example of Kemper (example (9) in the next Section) motivated us to further refine the implementation of the IMPROVED NEW ALGORITHM. It concerns the generation of  $B_d$ : If there are irreducible secondary invariants in rather high degrees  $d$  (in Kemper’s example, there are two irreducible secondary invariants of degree 9), it is advisable to generate not all of  $B_d$  at once, but in small portions. This will be part of release 3-0-3 of SINGULAR.

#### 4. Benchmark Tests for the Computation of Secondary Invariants

**4.1. The Test Examples.** We already mentioned that the study of ideal Turaev–Viro invariants is a source of test examples for the computation of invariant rings. Let  $\mathcal{F} = \{1, \dots, m\}$ , and let  $G = S_m$  be the symmetric group of  $\mathcal{F}$ , and let  $R$  be the ring whose variables are colour weights and  $6j$ -symbols, as in Section 6.2, without simplifying assumptions. We obtain an action on the set of colour weights by  $g.w(a) = w(g.a)$ , and an action on the set of  $6j$ -symbols by

$$g \cdot \left| \begin{array}{ccc} a & b & c \\ f & e & d \end{array} \right| = \left| \begin{array}{ccc} g.a & g.b & g.c \\ g.f & g.e & g.d \end{array} \right|,$$

for  $a, \dots, f \in \mathcal{F}$  and  $g \in G$ . For computing primary invariants, it helps to use the fact that the set of variables of  $R$  is decomposed into several  $G$ -orbits. Of course, this also works in the case of a simplifying assumption for which  $g.w(a)$  (resp.  $g \cdot \left| \begin{array}{ccc} a & b & c \\ f & e & d \end{array} \right|$ ) is a variable of  $R$  if and only if  $w(a)$  (resp.  $\left| \begin{array}{ccc} a & b & c \\ f & e & d \end{array} \right|$ ) is a variable of  $R$ , for all  $g \in G$ . It is not difficult to invent appropriate simplifying assumptions. Examples (1), (7) and (8) come from that source. We will not go into details here, but just provide the matrices and primary invariants of our nine test examples. They are roughly ordered by increasing computation time. The ring variables are called  $x_1, x_2, \dots$ . Let  $e_i$  be the column vector with 1 in position  $i$  and 0 otherwise.

- (1) A 13-dimensional representation of the symmetric group  $S_2$  is given by the matrix

$$M = (e_2 e_1 e_{13} e_{12} e_{11} e_8 e_{10} e_6 e_9 e_7 e_5 e_4 e_3)$$

Our primary invariants are

$$\begin{aligned} &x_9, \quad x_7 + x_{10}, \quad x_6 + x_8, \quad x_5 + x_{11}, \quad x_4 + x_{12}, \quad x_3 + x_{13}, \\ &x_1 + x_2, \quad x_3 x_{13}, \quad x_4 x_{12}, \quad x_5 x_{11}, \quad x_7 x_{10}, \quad x_6 x_8, \quad x_1 x_2 \end{aligned}$$

There are 32 secondary invariants of maximal degree 6, among which are 15 irreducible secondary invariants up to degree 2.

- (2) A 6–dimensional representation of  $S_4$  is given by the matrices

$$\begin{aligned} M_1 &= (e_1 e_4 e_5 e_2 e_3 e_6) \\ M_2 &= (e_4 e_1 e_5 e_2 e_6 e_3) \end{aligned}$$

Our primary invariants are

$$\begin{aligned} &x_3 + x_5 + x_6, \quad x_1 + x_2 + x_4, \quad x_3 x_5 + x_3 x_6 + x_5 x_6, \\ &x_3 x_4 + x_2 x_5 + x_1 x_6, \quad x_1 x_2 x_4, \quad x_1^3 x_2^3 + x_1^3 x_4^3 + x_2^3 x_4^3 + x_3^2 x_5^2 x_6^2 \end{aligned}$$

There are 12 secondary invariants of maximal degree 9, among which are 4 irreducible secondary invariants of maximal degree 3.

- (3) A 6–dimensional representation of the alternating group  $A_4$  is given by the matrices

$$\begin{aligned} M_1 &= (e_4 e_1 e_5 e_2 e_6 e_3) \\ M_2 &= (e_2 e_3 e_1 e_6 e_4 e_5) \end{aligned}$$

Our primary invariants are

$$\begin{aligned} &x_1 + x_2 + x_3 + x_4 + x_5 + x_6, \quad x_3 x_4 + x_2 x_5 + x_1 x_6, \\ &x_1 x_2 + x_1 x_3 + x_2 x_3 + x_1 x_4 + x_2 x_4 + x_1 x_5 + x_3 x_5 \\ &\quad + x_4 x_5 + x_2 x_6 + x_3 x_6 + x_4 x_6 + x_5 x_6, \\ &x_3^2 x_4 + x_3 x_4^2 + x_2^2 x_5 + x_2 x_5^2 + x_1^2 x_6 + x_1 x_6^2, \\ &x_1 x_2 x_4 + x_1 x_3 x_5 + x_2 x_3 x_6 + x_4 x_5 x_6, \\ &x_1^2 x_2^4 + x_1^4 x_3^2 + x_2^2 x_3^4 + x_1^4 x_4^2 + x_2^2 x_4^4 \\ &\quad + x_3^4 x_5^2 + x_4^4 x_5^2 + x_1^2 x_5^4 + x_2^2 x_6^4 + x_5^4 x_6^2 + x_3^2 x_6^4 + x_4^2 x_6^4 \end{aligned}$$

There are 18 secondary invariants of maximal degree 11, among which are 8 irreducible secondary invariants of maximal degree 5.

- (4) A 6–dimensional representation of the dihedral group  $D_6$  is given by the matrices

$$\begin{aligned} M_1 &= (e_6 e_5 e_4 e_3 e_2 e_1) \\ M_2 &= (e_3 e_1 e_2 e_6 e_4 e_5) \end{aligned}$$

Our primary invariants are the elementary symmetric polynomials. There are 120 secondary invariants of maximal degree 10, among which are 14 irreducible secondary invariants of maximal degree 4.

- (5) A 8–dimensional representation of  $D_8$  is given by the matrices

$$\begin{aligned} M_1 &= (e_8 e_7 e_6 e_5 e_4 e_3 e_2 e_1) \\ M_2 &= (e_4 e_1 e_2 e_3 e_8 e_5 e_6 e_7) \end{aligned}$$

Our primary invariants are

$$\begin{aligned} &x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8, \\ &x_4 x_5 + x_1 x_6 + x_2 x_7 + x_3 x_8, \quad x_3 x_5 + x_4 x_6 + x_1 x_7 + x_2 x_8, \\ &x_2 x_5 + x_3 x_6 + x_4 x_7 + x_1 x_8, \quad x_1 x_5 + x_2 x_6 + x_3 x_7 + x_4 x_8, \\ &x_1 x_3 + x_2 x_4 + x_5 x_7 + x_6 x_8, \quad x_1 x_2 x_3 x_4 + x_5 x_6 x_7 x_8, \\ &x_1 x_2^3 + x_2 x_3^3 + x_1^3 x_4 + x_3 x_4^3 + x_5^3 x_6 + x_6^3 x_7 + x_7^3 x_8 + x_5 x_8^3 \end{aligned}$$

There are 64 secondary invariants of maximal degree 11, among which are 24 irreducible secondary invariants of maximal degree 5.

(6) A 7–dimensional representation of  $D_{14}$  is given by the matrices

$$\begin{aligned} M_1 &= (e_2e_3e_4e_5e_6e_7e_1) \\ M_2 &= (e_1e_7e_6e_5e_4e_3e_2) \end{aligned}$$

Our primary invariants are the elementary symmetric polynomials. There are 360 secondary invariants of maximal degree 18, among which are 19 irreducible secondary invariants of maximal degree 7.

(7) A 15–dimensional representation of  $S_3$  is given by the matrices

$$\begin{aligned} M_1 &= (e_2e_1e_3e_4e_7e_{14}e_5e_8e_{11}e_{13}e_9e_{15}e_{10}e_6e_{12}) \\ M_2 &= (e_1e_3e_2e_4e_5e_9e_8e_7e_6e_{13}e_{12}e_{11}e_{10}e_{15}e_{14}) \end{aligned}$$

Our primary invariants are

$$\begin{aligned} &x_1 + x_2 + x_3, \quad x_1x_2 + x_1x_3 + x_2x_3, \quad x_1x_2x_3, \\ &x_{10} + x_{13}, \quad x_{10}x_{13}, \quad x_6 + x_9 + x_{11} + x_{12} + x_{14} + x_{15}, \\ &x_{11}x_{12} + x_6x_{14} + x_9x_{15}, \quad x_9x_{11} + x_6x_{12} + x_{14}x_{15}, \\ &x_6x_{11} + x_9x_{12} + x_9x_{14} + x_{12}x_{14} + x_6x_{15} + x_{11}x_{15}, \\ &x_6x_9x_{14} + x_6x_{11}x_{14} + x_{11}x_{12}x_{14} + x_6x_9x_{15} + x_9x_{12}x_{15} + x_{11}x_{12}x_{15}, \\ &x_6^6 + x_9^6 + x_{11}^6 + x_{12}^6 + x_{14}^6 + x_{15}^6, \quad x_4, \quad x_5 + x_7 + x_8, \\ &x_5x_7 + x_5x_8 + x_7x_8, \quad x_5x_7x_8 \end{aligned}$$

There are 1728 secondary invariants of maximal degree 17, among which are 76 irreducible secondary invariants of maximal degree 4.

(8) A 18–dimensional representation of  $S_3$  is given by the matrices

$$\begin{aligned} M_1 &= (e_2e_1e_3e_4e_{12}e_{10}e_7e_{11}e_{14}e_6e_8e_5e_{15}e_9e_{13}e_{17}e_{16}e_{18}) \\ M_2 &= (e_1e_3e_2e_{14}e_8e_7e_6e_5e_9e_{10}e_{15}e_{13}e_{12}e_4e_{11}e_{16}e_{18}e_{17}) \end{aligned}$$

Our primary invariants are

$$\begin{aligned} &x_1 + x_2 + x_3, \quad x_1x_2 + x_1x_3 + x_2x_3, \quad x_1x_2x_3, \\ &x_4 + x_9 + x_{14}, \quad x_4x_9 + x_4x_{14} + x_9x_{14}, \quad x_4x_9x_{14}, \\ &x_{16} + x_{17} + x_{18}, \quad x_{16}x_{17} + x_{16}x_{18} + x_{17}x_{18}, \quad x_{16}x_{17}x_{18}, \\ &x_6 + x_7 + x_{10}, \quad x_6x_7 + x_6x_{10} + x_7x_{10}, \\ &x_6x_7x_{10}, \quad x_5 + x_8 + x_{11} + x_{12} + x_{13} + x_{15}, \\ &x_5x_{12} + x_8x_{13} + x_{11}x_{15}, \quad x_8x_{11} + x_{12}x_{13} + x_5x_{15}, \\ &x_5x_{11} + x_8x_{12} + x_5x_{13} + x_{11}x_{13} + x_8x_{15} + x_{12}x_{15}, \\ &x_5x_8x_{12} + x_5x_{11}x_{12} + x_5x_8x_{13} + x_{11}x_{12}x_{15} + x_8x_{13}x_{15} + x_{11}x_{13}x_{15}, \\ &x_5^6 + x_8^6 + x_{11}^6 + x_{12}^6 + x_{13}^6 + x_{15}^6 \end{aligned}$$

There are 31104 secondary invariants of maximal degree 22, among which are 137 irreducible secondary invariants of maximal degree 4.

(9) A 10–dimensional representation of  $S_5$  is given by the matrices

$$M_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{2}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{2}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \end{pmatrix}$$

We are not listing the primary invariants here, as they are too big polynomials. There are 720 secondary invariants of maximal degree 22, among which are 46 irreducible secondary invariants of maximal degree 9.

Examples (2), (3) and (9) belong to a very interesting class of examples that was shown to us by G. Kemper [46]. For  $n \in \mathbb{N}$ , Let  $M_n$  be the set of two-element subsets of  $\{1, \dots, n\}$ . Then, one studies the obvious  $S_n$  action on  $M_n$  (or similarly, the obvious  $A_n$  action), and one can try to compute the invariant ring  $\mathbb{Q}[M_n]^{S_n}$  (resp.  $\mathbb{Q}[M_n]^{A_n}$ ).

The 10–dimensional representation of  $S_5$  in Example (9) is a surprisingly challenging problem. To simplify the computations, Kemper provided a decomposition of the representation into a direct sum of a 1–, a 4– and a 5–dimensional representation. Without ad-hoc methods, the computation of secondary invariants for that problem has been beyond reach. The procedure `(Irreducible)SecondaryInvariants` of MAGMA V2.13-8 breaks immediately, since it requests 55.62 GB memory, while the memory limit of our computer is 16 GB. Our algorithm `irred_secondary_char0` in SINGULAR version 3-0-2 exceeds the limit of 16 GB while computing secondary invariants in degree 8.

The total number of secondary invariants in Example (9) is not particularly large. The difficulties in Example (9) come from the fact that there are *irreducible* secondary invariants of rather high degrees. A slightly refined implementation of `secondary_char0`, that will be part of SINGULAR version 3-0-3, can compute Example (9). Here, when finding irreducible secondary invariants of high degree  $d$ , we do not generate all of  $B_d$  at once, but we decompose it into handy blocks, in order to save memory.

**4.2. Comparison.** We describe here how different algorithms perform on Examples (1) up to (9). All computations had been done on a Linux x86\_64 platform with two AMD Opteron 248 processors (2,2 GHz) and a memory limit of 16 GB. The computation of primary invariants is not part of our tests. Hence, in each example we use the same primary invariants for all algorithms. We compare the following implementations:

- (1) `secondary_char0` as in SINGULAR release 2-0-6. In Table 1, we refer to it as “SINGULAR (1998)”.
- (2) `secondary_char0` as in SINGULAR release 3-0-2, with a slight refinement. In Table 1, we refer to it as “SINGULAR (all sec.)”.
- (3) `irred_secondary_char0`, as in SINGULAR release 3-0-2, with a slight refinement. In Table 1, we refer to it as “SINGULAR (irr. sec.)”.
- (4) `SecondaryInvariants` in MAGMA V2.13-8.

Implementation (1) is due to A. Heydtmann [37] (1998) and has been part of SINGULAR up to release 3-0-1.

Implementations (2) and (3) are our implementations of the IMPROVED NEW ALGORITHM explained in Section 3. They are part of SINGULAR 3-0-2, released in Juli, 2006. They do not belong to the SINGULAR kernel but are interpreted code. Here, we test a slightly improved version, that will be part of the next SINGULAR release and saves memory when generating irreducible secondary invariants in high degrees. However, this only affects example (9); the performance in the other eight examples remains essentially the same, as the degrees of their irreducible secondary invariants are not high enough. SINGULAR 3-0-3 will contain another algorithm, that we describe in Chapter 8. The timings include the computation of Reynolds operator and Molien series, which belong to the input of the SINGULAR procedures.

Implementation (4) is due to A. Steel, based on [47], [45] or [22]. We consider here the MAGMA-version V2.13-8, released in October, 2006. There is also a function `IrreducibleSecondaryInvariants` in MAGMA, but computation time and memory consumption are essentially the same, in our examples. So, for the sake of simplicity, we do not provide separate timings for that function.

Note that, after posting the first version of our paper [56], there was a new release of MAGMA containing an algorithm that G. Kemper developed in 2006. It provides a major improvement in the performance of MAGMA and seems to be competitive with our algorithm. However, Kemper's new algorithm apparently is not described in the literature yet. We provide comparative benchmarks for the new versions of SINGULAR and MAGMA in the last section of Chapter 8.

Note that, in contrast to the corresponding MAGMA functions, `irred_secondary_char0` often works much faster and needs much less memory than `secondary_char0`; see Examples (1) and (6)–(9). However, this is not always the case, as can be seen in Examples (4) and (5).

In Table 1, “—” means that the computation fails since the process exceeds the memory limit; in examples (8) and (9), MAGMA requests the amount of memory that we indicate in round brackets. In some cases, we stopped the computation when it was clear that it takes too much time; this is indicated in the table by “> ...”.

In conclusion, our benchmarks provide some evidence that the IMPROVED NEW ALGORITHM has great advantages in the computation of invariant rings with many secondary invariants. Here, it marks a dramatic improvement compared with previous algorithms in SINGULAR or algorithms in MAGMA. In 3 of our 9 examples, it is the only algorithm that terminates in reasonable time with a memory limit of 16 GB. A particular benefit of our algorithm is that the computation of irreducible secondary invariants does not involve the explicit computation of *reducible* secondary invariants, which may save resources. After writing [56], we found another algorithm for the computation of algebra generators of  $R^G$ . It also provides an improvement for the computation of irreducible secondary invariants, as we explain in Chapter 8.

TABLE 1. Comparative benchmark for the computation of secondary invariants

	(1) SINGULAR (1998)	(2) SINGULAR (all sec.)	(3) SINGULAR (irr. sec.)	(4) MAGMA
Expl. (1)	0.55 s 8.62 MB	0.05 s 1.49 MB	0.03 s 1.0 MB	0.05 s 10.3 MB
Expl. (2)	0.05 s 0.99 MB	0.04 s 0.96 MB	0.04 s 0.97 MB	0.01 s 7.05 MB
Expl. (3)	0.48 s 2.97 MB	0.33 s 1.95 MB	0.3 s 1.96 MB	0.19 s 8.96 MB
Expl. (4)	6.55 s 12.29 MB	0.63 s 2.47 MB	0.32 s 2.97 MB	0.48 s 9.09 MB
Expl. (5)	18.15 s 45.79 MB	10.53s 10.61 MB	9.69 s 17.0 MB	6.66 s 31.82 MB
Expl. (6)	> 984 m > 167 MB	100.4 s 110.0 MB	16.55 s 39.0 MB	118.51 s 54.0 MB
Expl. (7)	— —	268.9s 872.7 MB	20.94 s 35.1 MB	> 7 h > 15 GB
Expl. (8)	— —	> 10 h > 10 GB	50.7 m 3.36 GB	— (259.5 GB)
Expl. (9)	— —	6.42 h 10.74 GB	99.2 m 7.35 GB	— (55.62 GB)

## Minimal generating sets of non-modular invariant rings

In the previous chapter, we provided an algorithm to compute module generators of the invariant ring  $R^G$  of a non-modular finite group action. In this chapter, we provide an algorithm to compute a minimal set of homogeneous invariant polynomials that generate  $R^G$  as a sub-algebra of  $R$ . Such generators are also known as **fundamental invariants**.

In principal, this can be solved as explained in the previous chapter: First, one computes primary invariants of  $R^G$  and then irreducible secondary invariants. Primary and irreducible secondary invariants together generate  $R^G$  as an algebra, and (potentially after removing some primary invariants) they form an inclusion minimal generating set [45]. Thiéry [93] suggests another algorithm for the computation of a minimal generating set in the special case of permutation groups, i.e., of groups acting on  $R$  as subgroup of the permutation group of the variables of  $R$ . Thiéry's algorithm is not based on the computation of primary invariants, but uses the incremental construction of so-called *SAGBI-Gröbner bases*. His algorithm is implemented in the library PERMUVAR of MUPAD [91]. There is extensive benchmark on MAGMA and MUPAD, using the transitive permutation groups on up to nine variables [92].

Our algorithm comes in one version for permutation groups and one version for finite matrix groups. We present comparative benchmarks based on transitive permutation groups on 7 or 8 variables. We implemented our algorithm in a library (i.e., as interpreted code) in SINGULAR [29]. In most of the examples, our algorithm is at least 50 times faster than the algorithm implemented in a pre-compiled MAGMA [11] library, often more than 1000 times. We also computed minimal generating sets for some transitive permutations groups on 9 and 10 variables. Moreover, we compute minimal generating sets for the natural action of the cyclic groups of order up to 12 in characteristic zero and for the cyclic groups of order up to 15 in prime characteristic (but, of course, still in the non-modular case).

The key ingredient of our algorithm is Theorem 16 in Chapter 7. Our algorithm does not involve solving linear algebra problems that may become rather huge, in contrast to the algorithm exposed in [22]. Instead, we use Gröbner bases. Of course, the computation of a Gröbner basis can be, in general, a very difficult business. The main feature of our algorithm is that it involves at most one computation of a Gröbner basis in each degree. It turns out that this yields a very well-performant algorithm.

Another peculiarity of our algorithm is the fact that it does not rely on *a-priori* bounds for the maximal degree  $\beta(R^G)$  of elements of a minimal generating set of  $R^G$ . For other algorithms, like the one presented in [93], the performance crucially depends on good estimates for  $\beta(R^G)$ . Unfortunately, well known a-priori bounds like Noether's  $\beta(R^G) \leq |G|$  are, in general, far from being optimal. In contrast, we rely on more realistic *a-posteriori* bounds: While incrementally constructing the set of generators, we obtain informations allowing to estimate  $\beta(R^G)$ .

We outline our algorithm. In the case of finite matrix groups, candidates for generators are found by applying the Reynolds operator to some monomials. In the case of permutation groups, candidates are found among the *orbit sums*. For testing whether a candidate is already contained in the algebra generated by previously found generators, an ideal membership problem needs to be solved. The solution is provided by computing the normal form with respect to a homogeneous Gröbner basis up to degree  $d$  of the ideal spanned by the previously found generators. When starting in a new degree, the Gröbner basis is computed by standard procedures (e.g., Buchberger's algorithm), and when a new generator of  $R^G$  of degree  $d$  has been found, we can directly write down a new Gröbner basis up to degree  $d$ , as in Chapter 7. While incrementally constructing the set of generators, one obtains informations allowing to estimate the maximal degree  $\beta(R^G)$  of elements of a minimal generating set of  $R^G$ . Hence, after finishing in that degree, one can stop the quest for more generators.

A modification of our algorithm can be used to compute irreducible secondary invariants. According to our comparative benchmarks, this algorithm is even faster than our algorithm presented in Chapter 7 and the algorithm recently implemented in MAGMA V2.13-9 that appears to be not described in a paper yet.

The rest of this chapter is organized as follows. In the next section, we explain our algorithm in more detail. In Subsection 2.1, we do some benchmark tests, comparing the implementation of our algorithm in SINGULAR [29] with the function `FundamentalInvariants` of MAGMA [11]. In Subsection 2.2, we expose some additional examples that seem to be out of reach for other known algorithms. In the final section, we modify our algorithm in order to compute irreducible secondary invariants, and do some benchmarks with that algorithm.

## 1. The Algorithm

Let  $G$  be a finite group, linearly acting on a polynomial ring  $R$  with  $n$  variables over some field  $K$ . We denote the action of  $g \in G$  on  $r \in R$  by  $g.r \in R$ . For  $d > 0$ , let  $R_d^G$  be the set of homogeneous  $G$ -invariant polynomials of degree  $d$ . For an ideal  $I \subset R$ , let  $lm(I)$  be the set of leading monomials occurring in  $I$ . As before, let  $\text{Rey}: R \rightarrow R^G$  be the Reynolds operator. For  $S \subset R$ , let  $mon_d(S) \subset R$  be the set of monomials of degree  $d$  that are not contained in  $lm(\langle S \rangle)$ . Let  $B_d(S) = \text{Rey}(mon_d(S))$ , which is easy to compute if  $S$  is a Gröbner basis at least up to degree  $d$ . By Lemma 3.5.1 and Remark 3.5.3 in [22], if  $\langle\langle S \rangle\rangle \not\supset R_d^G$  then  $\langle\langle S \rangle\rangle \not\supset B_d(S)$ ; hence, we may restrict the quest for new generators to the finite set  $B_d(S)$ .

So, in increasing degree  $d$  starting with  $d = 1$  and  $S = \emptyset$ , we may loop through all  $b \in B_d(S)$ , and add  $b$  to the set  $S$  of previously found generators if  $b \notin \langle\langle S \rangle\rangle$ . In that way, one incrementally constructs a generating set of  $R^G$ , consisting of homogeneous invariant polynomials. In fact, it is a *minimal* generating set [93]. We can test whether  $b \in \langle\langle S \rangle\rangle$  according to the following lemma. Although the lemma is well known, we include a proof for completeness.

**LEMMA 12.** *Let  $S \subset R^G$  be a set of homogeneous invariant non-constant polynomials. Assume that  $R_{d'}^G \subset \langle\langle S \rangle\rangle$  for all  $d' < d$ , and assume that we are in the non-modular case. Let  $b \in R_d^G$ . We have  $b \in \langle\langle S \rangle\rangle$  if and only if  $b \in \langle S \rangle$ .*

**PROOF.** If  $b \in \langle\langle S \rangle\rangle$  then  $b \in \langle S \rangle$ . If  $b \in \langle S \rangle$  then we can write  $b$  as a finite sum,

$$b = \sum_i p_i q_i$$

with homogeneous polynomials  $p_i \in R$  and  $q_i \in S$ . As in the proof of Lemma 9, it follows by Lemma 8 that  $b = \text{Rey}(b) = \sum_i \text{Rey}(p_i)q_i$ . Since the elements of  $S$  are non-constant, the  $p_i$  are of degree at most  $d - 1$ . Hence,  $\text{Rey}(p_i) \in R_{d'}^G$  for some  $d' < d$ . Thus  $\text{Rey}(p_i) \in \langle\langle S \rangle\rangle$  by hypothesis. Therefore,  $b \in \langle\langle S \rangle\rangle$ .  $\square$

As in Chapter 7, we test whether  $b \in \langle S \rangle$  by reduction of  $b$  with respect to a homogeneous Gröbner basis  $\mathcal{G}$  up to degree  $d$ . Moreover, after adding  $b$  to the set of generators, we easily obtain a homogeneous Gröbner basis up to degree  $d$  of  $\langle S \cup \{b\} \rangle$ , by Theorem 16 in Chapter 7.

There is a problem, though. We can incrementally construct a minimal generating set of  $R^G$ , in increasing degrees — but in what degree shall we stop the construction? By definition, we can stop after having found the generators in degree  $\beta(R^G)$ . So, we could adopt a general estimate for  $\beta(R^G)$  like Noether's bound  $\beta(R^G) \leq |G|$ . However, such general a-priori estimates are very often far from being optimal.

Therefore, we prefer to derive an estimate for  $\beta(R^G)$  from the previously constructed generators. If  $S$  is a generating set of  $R^G$ , then it follows that  $\langle S \rangle$  is zero-dimensional, as in the proof of Proposition 3.3.1 in [22]. Hence, there are only finitely many monomials outside  $lm(\langle S \rangle)$ , of maximal degree  $d_{max}$ . Since we can restrict the quest for generators of  $R^G$  of degree  $d$  to the Reynolds images of monomials of degree  $d$  outside  $lm(\langle S \rangle)$ , it follows  $\beta(R^G) \leq d_{max}$ .

Our strategy was to work with a homogeneous Gröbner basis up to degree  $d$  of  $\langle S \rangle$ . However, for testing whether  $\langle S \rangle$  is of dimension 0, one needs a Gröbner basis of  $\langle S \rangle$  — without degree restriction. To avoid needless computations, we use the following trick.

By definition, in degree  $\beta(R^G)$  we will find a homogeneous generator of  $R^G$ , but in degree  $\beta(R^G) + 1$  we don't. Hence, only if our incremental construction of  $S$  arrives at some degree  $d$ , such that there is an element of  $S$  in degree  $d - 2$  but none in  $d - 1$ , it makes sense to compute a Gröbner basis of  $\langle S \rangle$  without degree restriction (although certainly we can not exclude that there are generators above degree  $d$ ). If  $\dim(\langle S \rangle) = 0$ , which is tested using the Gröbner basis, then we obtain an estimate for  $\beta(R^G)$  that tells us in what degree we can stop the incremental search. We thus obtain the following algorithm for the computation of a minimal generating set of  $R^G$ , where  $G$  is a finite matrix group.

Algorithm INVARIANT ALGEBRA

- (1) Construct the Reynolds operator  $\text{Rey}: R \rightarrow R^G$ .  
Let  $S = \mathcal{G} = \emptyset$ . Let  $d_{max} = 0$ .
- (2) For increasing degree  $d$ , starting with  $d = 1$ :
  - (a) If  $S$  contains elements of degree  $d - 2$  but no elements of degree  $d - 1$ :
    - (i) Replace  $\mathcal{G}$  by a (complete) Gröbner basis of  $\langle S \rangle$ .
    - (ii) If  $\dim(\langle S \rangle) = 0$  (which is tested using  $\mathcal{G}$ ), then replace  $d_{max}$  by the maximal degree of polynomials outside  $lm(\langle S \rangle)$ , and if, moreover,  $d$  exceeds the new  $d_{max}$  then break and return  $S$ .

If  $S$  contains elements of degree  $d - 1$ , replace  $\mathcal{G}$  by a homogeneous Gröbner basis  $\mathcal{G}$  of  $\langle S \rangle$  up to degree  $d$ .
  - (b) Compute  $B_d(S)$  using  $\mathcal{G}$  and  $\text{Rey}$ .
  - (c) For all  $b \in B_d(S)$ :  
If  $\text{rem}(b; \mathcal{G}) \neq 0$  then replace  $S$  by  $S \cup \{b\}$  and  $\mathcal{G}$  by  $\mathcal{G} \cup \{\text{rem}(b; \mathcal{G})\}$ .
  - (d) If  $d = d_{max}$  then break and return  $S$ .

By Theorem 16, in all steps,  $\mathcal{G}$  is a homogeneous Gröbner basis of  $\langle S \rangle$  up to degree  $d$ . Note that the algorithm used in MAGMA [11] (see [22]) involves the computation of primary and irreducible secondary invariants. Our approach is more straight

forward: Why should one construct primary invariants if one is not primarily interested in them? Moreover, our algorithm uses much more elementary methods than the alternative algorithm described in [22] based on linear algebra. No huge systems of linear equations occur, only few explicit Gröbner basis computations are needed (one per degree), and apart from that the most time consuming operation is the computation of normal forms. So it is not surprising that usually our implementation of INVARIANT ALGEBRA in SINGULAR [29] is much faster than the algorithm implemented in MAGMA [11].

In most of our examples, the computation of homogeneous Gröbner bases up to degree  $d$  is not a big deal (there are exceptions, though). However, for large group orders, the computation of the Reynolds operator exceeds the resources. So, the use of the Reynolds operator can be a problem. In the case of permutation groups, it helps to replace it by so-called *orbit sums*, which is also used in [93]. The *orbit* of a monomial  $m \in R$  is  $G.m = \{g.m : g \in G\}$ . The **orbit sum** of  $m$  is  $m^\circ = \sum_{m' \in G.m} m'$ . Of course,  $m^\circ \in R^G$ .

In contrast to the Reynolds operator, the orbit sums are defined even in the modular case, i.e., if the characteristic of  $R$  divides  $|G|$ . In the non-modular case,  $m^\circ$  is just a scalar multiple of  $\text{Rey}(m)$ . In conclusion, if  $G$  is a permutation group, we can also define  $B_d(S)$  to be the orbit sums of the monomials in  $\text{mon}_d(S)$ . Note, however, that even when using orbit sums, the algorithm INVARIANT ALGEBRA only works for the non-modular case, since it relies on Lemma 12.

## 2. Benchmark Tests for the Computation of Minimal Generating Sets

A classical test bed for the computation of minimal generating sets of invariant rings of finite groups is provided by **transitive permutation groups** [93], [92]. These are groups acting on a polynomial ring  $R$  over a field  $K$  by permuting variables, such that any two variables are related by the group action. The MAGMA function `TransitiveGroups(i)` provides a list of all classes of transitive permutation groups on  $i$  variables.

In our comparative benchmark (Subsection 2.1), we consider transitive permutation groups on 7 and 8 variables in characteristic 0. In Subsection 2.2, we present some more examples of transitive permutation groups, with up to 12 variables in characteristic 0 and up to 15 variables in prime characteristic. Our benchmarks are based on a Linux x86\_64 platform with two AMD Opteron 248 processors (2,2 GHz) and a memory limit of 16 GB.

**2.1. Comparative Benchmark: Transitive Permutation Groups.** We study here minimal generating sets of invariant rings of transitive permutation groups on 7 and 8 variables, in characteristic 0. We compare the following algorithms.

- (1) Our implementation of INVARIANT ALGEBRA using orbit sums. This is part of the `finvar.lib` library of SINGULAR-3-0-3 (to be released soon) and is called `invariant_algebra_perm`. We tested a  $\beta$ -version of SINGULAR-3-0-3.
- (2) The function `FundamentalInvariants` of MAGMA V2.13-9 (released January 2007), which, to the best of the author's knowledge, is either based on the algorithms described in [22] or unpublished.

Note that our implementation in SINGULAR is interpreted code, without any pre-compilation. As far as known to the author, `FundamentalInvariants` in MAGMA is pre-compiled.

Usually (but not thoroughly) we stopped the computations of an example after two hours CPU time. Moreover, we stopped the computation by one algorithm if it

took more than about 1000 times longer than by the other algorithm. The results are provided in Table 1 for the 7 transitive permutation groups on 7 variables, and in Table 2 for 45 transitive permutation groups on 8 variables. In the first column of the tables, the group is defined by its generators in disjoint cycle presentation. The rounded CPU times for SINGULAR or MAGMA in seconds are provided in the next two columns. The last column of the tables indicates the number of generators of a minimal generating set of  $R^G$ , sorted degree-wise.

TABLE 1. Transitive permutation groups on 7 variables (characteristic 0)

Group	SINGULAR time [s]	MAGMA time [s]	# generators (sorted by degree)
(1,2,3,4,5,6,7)	0.52	25.3	1,3,8,12,12,6,6
(1,2,3,4,5,6,7), (1,6)(2,5)(3,4)	0.67	11	1,3,4,6,6,3,3
(1,2,3,4,5,6,7), (1,2,4)(3,6,5)	6.6	239	1,1,4,5,8,8,6
(1,2,3,4,5,6,7), (1,2)(3,6)	16.9	107	1,1,2,2,2,2,2
(1,2,3,4,5,6,7), (1,3,2,6,4,5)	81.5	600	1,1,2,3,4,7,7,5,1
(3,4,5,6,7), (1,2,3)	117	474	1,1,1,1,1,1,1,0,0,0, 0,0,0,0,0,0,0,0,0,1
(1,2,3,4,5,6,7), (1,2)	198	0.04	1,1,1,1,1,1,1

Table 2: Transitive permutation groups on 8 variables (characteristic 0)

Group	SINGULAR time [s]	MAGMA time [s]	# generators (sorted by degree)
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8)	0.14	0.07	1,7,7,7
(1,2,3,8)(4,5,6,7), (1,6)(2,5)(3,4)(7,8)	0.24	11.6	1,6,8,12,5
(1,2,3,8)(4,5,6,7), (1,5)(2,6)(3,7)(4,8)	0.35	15	1,5,9,16,8
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (4,5)(6,7)	0.35	10.8	1,5,5,8,4
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (4,5)(6,7), (4,6)(5,7)	0.55	34.6	1,4,4,7,3
(1,2,3,8)(4,5,6,7), (1,7,3,5)(2,6,8,4)	0.65	137	1,4,10,19,15,7
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (2,3)(4,5), (2,3)(6,7)	0.65	52.2	1,4,4,7,3,1
(1,5)(3,7), (1,2,3,8)(4,5,6,7)	0.77	73.9	1,4,6,11,7,2
(1,5)(3,7), (1,3,5,7)(2,4,6,8), (1,4,5,8)(2,3,6,7)	0.8	167	1,4,6,11,7,3
(1,5)(3,7), (1,4,5,8)(2,3)(6,7), (1,3)(2,8)(4,6)(5,7)	1.2	60.3	1,4,4,6,4,3,2,1
(4,8), (1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7)	1.4	7.38	1,4,4,6,3,1

*Continued on the next page*

Group	SINGULAR	MAGMA	# generators (sorted by degree)
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,3)(4,5,6,7), (1,3)(5,7)	1.9	318	1,3,3,6,3,2,1
(1,2,3,4,5,6,7,8)	2.2	2608	1,4,10,18,16,8,4,4
(2,6)(3,7), (1,2,3,4,5,6,7,8)	2.3	> 2200	1,3,5,8,7,7,4,4
(2,6)(3,7), (1,2,3,8)(4,5,6,7)	2.3	385	1,3,5,9,6,4,2,1
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,3)(4,5,6,7)	2.4	649	1,3,3,7,6,7,5,1
(1,2,3,4,5,6,7,8), (1,5)(3,7)	2.8	> 2800	1,3,7,12,13,9,4,4
(1,2,3,4,5,6,7,8), (1,6)(2,5)(3,4)(7,8)	3	1040	1,4,5,9,8,4,2,2
(1,2,3,4,5,6,7,8), (1,3)(2,6)(5,7)	3.3	> 3300	1,3,6,11,12,7,2,2
(4,8), (1,2,3,8)(4,5,6,7)	3.7	580	1,3,5,8,6,4,2,2
(2,6)(3,7), (1,3)(5,7), (1,2,3,4,5,6,7,8)	3.7	> 3600	1,3,3,5,4,4,2,2
(1,2,3,8), (1,5)(2,6)(3,7)(4,8)	4	> 4000	1,3,4,7,6,4,2,2
(1,2,3,4,5,6,7,8), (1,5)(3,7), (1,6)(2,5)(3,4)(7,8)	4.3	5440	1,3,4,7,6,4,2,2
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,2,3)(4,6,5)	4.9	703	1,3,3,7,8,11,7
(1,2,3,4,5,6,7,8), (1,5)(4,8), (1,7)(3,4,8)	5	4780.6	1,3,3,5,3,3,2,3,1
(4,8), (1,3)(5,7), (1,2,3,8)(4,5,6,7)	5.4	444	1,3,3,5,3,2,1,1
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,2,3)(4,6,5), (2,3)(4,5)	6.5	1995	1,3,3,6,4,3,1
(2,6)(3,7), (1,3)(4,8)(5,7), (1,2,3,8)(4,5,6,7)	7.5	> 10800	1,3,3,5,3,2,3,4,3,2,1,1
(1,3)(2,8)(4,6)(5,7), (1,2,3)(5,6,7), (1,4)(2,6)(3,7)(5,8)	8.3	2410	1,3,3,8,7,9,6,1,1
(1,8)(2,3), (1,2,3)(5,6,7), (1,5)(2,7)(3,6)(4,8)	17.5	> 7200	1,2,2,5,2,5,4,3,3
(1,3)(2,8), (1,2,3), (1,5)(2,6)(3,7)(4,8)	31	> 7200	1,2,2,3,2,3,2,2,1,1
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,2,3)(4,6,5), (2,5)(3,4)	36.5	> 7200	1,2,2,4,3,5,4,2,2,1,1,1
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,2,3)(4,6,5), (1,6)(2,3,5,4)	37	3454	1,2,2,4,2,2,1
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,2,3)(4,6,5), (1,3)(4,5,6,7)	37	> 7200	1,2,2,4,2,3,2,2,1
(1,3,5,7)(2,4,6,8), (1,3,8)(4,5,7)	39	> 7200	1,2,4,8,11,12,7

Continued on the next page

Group	SINGULAR	MAGMA	# generators (sorted by degree)
(4,8), (1,8)(4,5), (1,2,3,8)(4,5,6,7)	39	> 7200	1,2,2,3,2,2,1,1
(1,8)(2,3)(4,5)(6,7), (1,3)(2,8)(4,6)(5,7), (1,5)(2,6)(3,7)(4,8), (1,2,3)(4,6,5), (4,6)(5,7)	44	> 7200	1,2,2,4,3,6,5,5,3
(1,3)(2,8), (1,2,3), (1,8)(4,5), (1,5)(2,6)(3,7)(4,8)	47	> 7200	1,2,2,3,2,2,1,1,0,0,0,1
(1,3)(2,8), (1,2,3), (1,8)(4,5), (1,5)(2,7,3,6)(4,8)	50	> 7200	1,2,2,3,2,2,1,1,0,0,0,1,1,1,1
(4,8), (1,8)(2,3)(4,5)(6,7), (1,2,3)(5,6,7)	51	> 7200	1,2,2,3,3,5,4,3,2,1,1,1
(1,2,3,8), (2,3), (1,5)(2,6)(3,7)(4,8)	51	73	1,2,2,3,2,2,1,1
(1,5)(4,8), (1,8)(2,3)(4,5)(6,7), (1,2,3)(5,6,7), (2,3)(4,8)(6,7)	56	> 7200	1,2,2,3,2,2,1,1,1,3,3,2,2,1,1,1
(1,2,3,4,5,6,7,8), (1,3,8)(4,5,7)	161.5	> 7200	1,2,3,5,6,6,5,2
(1,2)(3,4,5,6,7,8), (1,2,3)	17410	> 20000	1,1,1,1,1,1,1,1,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,1
(1,2,3,4,5,6,7,8), (1,2)	24629	0.18	1,1,1,1,1,1,1,1

In total, there are 50 classes of transitive permutation groups on 8 variables, but for five of them, neither SINGULAR nor MAGMA succeeded with the computation in the realm of our time and memory limits. Note that, according to [93], MUPAD can manage one of these five exceptions with the library PERMUVAR; with a memory limit of 500 MB and a time limit of 2 days, it can compute 17 of the 50 examples.

In the majority of the examples, SINGULAR-3-0-3 is at least 50 times faster than MAGMA V2.13-9, in some cases even more than 1000 times faster. There appears to be only one class of exceptions: The symmetric group on  $n$  variables (the last example on Tables 1 or 2, respectively). This is a special case with a well known theoretical solution. Since MAGMA knows that `TransitiveGroup(7,7)` and `TransitiveGroup(8,50)` are symmetric groups, it seems very likely to the author that `FundamentalInvariants` simply returns the well known solution in this case, without computation.

An extensive comparative benchmark of MUPAD and MAGMA on transitive permutation groups is provided by [92]. There, a different machine is used, the memory limit is more restrictive (500 MB), and the time limit is more generous (2 days).

Note that in the case of small group orders, it sometimes turned out to be faster to use images of the reynolds operator (the function `invariant_algebra_reynolds` in SINGULAR-3-0-3) rather than orbit sums. However, for groups of order greater than 1000, SINGULAR is hardly able to compute the reynolds operator in reasonable time. Of course, a pre-compilation would yield a considerable speed-up of our implementation.

**2.2. Further computational results.** In this subsection, we consider some more examples of transitive permutation groups, acting on up to 15 variables. Given the results exposed in the preceding subsection, it seems very unlikely to us that MAGMA V2.13-9 is able to compute these examples in reasonable time. Hence, we

TABLE 3. Some transitive permutation groups on 9 variables (characteristic 0)

Group	time [s]	# generators (sorted by degree)
$(1,2,9)(3,4,5)(6,7,8),$ $(1,4,7)(2,5,8)(3,6,9)$	6.24	1,4,16,24,24
$(1,2,3,4,5,6,7,8,9)$	38.19	1,4,14,26,32,18,12,6,6
$(1,2,9)(3,4,5)(6,7,8),$ $(1,4,7)(2,5,8)(3,6,9),$ $(1,2)(3,6)(4,8)(5,7)$	45.5	1,4,8,12,12,10
$(1,2,9)(3,4,5)(6,7,8),$ $(1,2)(4,5)(7,8),$ $(1,4,7)(2,5,8)(3,6,9)$	55.3	1,3,10,14,19,9,2
$(1,2,9)(3,4,5)(6,7,8),$ $(1,4,7)(2,5,8)(3,6,9),$ $(3,4,5)(6,8,7)$	84.3	1,2,8,9,16,18,14,4,2
$(1,2,3,4,5,6,7,8,9),$ $(1,8)(2,7)(3,6)(4,5)$	141.6	1,4,7,13,16,12,6,3,3
$(1,2,9)(3,4,5)(6,7,8),$ $(1,2)(4,5)(7,8),$ $(1,4,7)(2,5,8)(3,6,9),$ $(3,6)(4,7)(5,8)$	280.7	1,3,6,8,9,8,2
$(1,4,7)(2,5,8)(3,6,9),$ $(3,6)(4,7)(5,8)$	290.5	1,2,6,6,9,8,4
$(1,4,7)(2,8,5),$ $(1,2,3,4,5,6,7,8,9)$	455.1	1,2,6,11,20,25,26,10,8

TABLE 4. Some transitive permutation groups on 10 variables (characteristic 0)

Group	time [s]	# generators (sorted by degree)
$(1,3,5,7,9)(2,4,6,8,10),$ $(1,4)(2,3)(5,10)(6,9)(7,8)$	12.3	1,7,14,29,28,12
$(1,2,3,4,5,6,7,8,9,10)$	306	1,5,16,36,48,32,12,8,4,4
$(2,7)(5,10),$ $(1,3,5,7,9)(2,4,6,8,10)$	478	1,3,8,14,21,16,12,8,4,3
$(1,3,5,7,9)(2,4,6,8,10),$ $(1,2,9,8)(3,6,7,4)(5,10)$	1294	1,4,9,20,31,23,8
$(1,2,3,4,5,6,7,8,9,10),$ $(1,8)(2,7)(3,6)(4,5)(9,10)$	1425	1,5,8,18,24,17,6,4,2,2

only tried with our implementation of INVARIANT ALGEBRA in SINGULAR. Table 3 and Table 4 provide the results for some transitive permutation groups on 9 and 10 variables, in characteristic 0; here, we used orbit sums. According to [93], MUPAD can manage 5 of the transitive permutation groups on 9 variables (in total, there are 34 of them) using the library PERMUVAR, with a memory limit of 500 Mb and a time limit of 2 days.

TABLE 5. Natural action of  $C_n$  on  $n$  variables (characteristic 0)

$n$	time [s]	mem. [Mb]	# generators (sorted by degree)
6	0.05	0.746	1,3,6,6,2,2
7	0.17	1.25	1,3,8,12,12,6,6
8	1.54	2.25	1,4,10,18,16,8,4,4
9	35.6	11.92	1,4,14,26,32,18,12,6,6
10	298.3	54.16	1,5,16,36,48,32,12,8,4,4
11	1187	116	1,5,20,50,82,70,50,30,20,10,10
12(*)	2010 min	2160	1,6,24,64,104,84,36,20,12,8,4,4

A rather harmlessly looking class of transitive permutation groups is the natural action of the cyclic group  $C_n$  of order  $n$  on  $n$  variables. The maximal degree occurring in a minimal generating set is, by Noether's bound, of course at most

TABLE 6. Natural action of  $C_n$  on  $n$  variables (characteristic  $p > 0$ )

$n$	$p$	time [s]	mem. [Mb]	# generators (sorted by degree)
6	5	0.03	0.746	1,3,6,6,2,2
7	2	0.09	0.746	1,3,8,12,12,6,6
8	3	0.34	1.25	1,4,10,18,16,8,4,4
9	2	1.65	1.86	1,4,14,26,32,18,12,6,6
10	3	12.7	4.48	1,5,16,36,48,32,12,8,4,4
11	2	73.5	9.33	1,5,20,50,82,70,50,30,20,10,10
12	5	693	33.2	1,6,24,64,104,84,36,20,12,8,4,4
13	2	4079	81.1	1,6,28,84,168,180,132,84,60,36,24,12,12
14	3	25280	304.3	1,7,32,104,216,242,162,96,42,30,18,12,6,6
15	2	99873	780.4	1,7,38,130,306,388,264,120,88,56,40,24,16,8,8

$|C_n| = n$ , hence, quite small. However, the minimal number of generators of  $R^{C_n}$  is surprisingly large. According to [92], the invariant ring of  $C_{10}$  in characteristic 0 was out of reach. Still in release MAGMA V2.13-10, the function `FundamentalInvariant` needs at least 2.6 Gb and more than 10 hours for computing the invariant ring of  $C_9$  both in characteristic 0 and in characteristic 2. In contrast, our algorithm computes a minimal generating set in less than 36 respectively less than 2 seconds. It needs less than 30 minutes and less than 150 Mb even for the invariant ring of  $C_{11}$  — this is a major progress.

Since in this class of examples the group orders are very small, we use the Reynolds operator rather than orbit sums for the generation of invariants. For  $n \leq 5$  the computation is finished in almost no time, so we omit them in our tables. Table 5 provides the result for  $n = 6, \dots, 12$  in characteristic 0. It turns out that for  $n = 12$  the computation of the Gröbner basis needed in degree 5 is very hard to compute for SINGULAR 3-0-3  $\beta$ , so we used SINGULAR 3-0-2 instead. Recall that for the timings in Tables 1–4 we used orbit sums and not the Reynolds operator — this explains the different computation times in the case of cyclic groups.

Table 6 provides the results for  $n = 6, \dots, 15$  in small prime characteristic  $p > 0$ , of course such that  $p$  does not divide  $n$  (non-modular case). Apparently this is much easier than characteristic 0. The reason is that in characteristic 0 the coefficients occurring in the Gröbner bases become very huge. By consequence, it takes too long to compute normal forms.

Note that the in all examples, the number of generators in each degree is the same in characteristic 0 and in non-modular prime characteristic. It is in fact conjectured that this is always the case [94].

To work in prime characteristic is not the only way to simplify the computations. As a last example, we study here the action of  $S_5$  on pairs, which yields a 10-dimensional representation of  $S_5$ . One can decompose it into a 1-, a 4- and a 5-dimensional irreducible representation, and in this form, the representation is given by the matrices of Example (9) in Section 7.4 on page 68.

We could describe that representation of  $S_5$  by a transitive permutation group on 10 variables. However, in that formulation of the problem, our algorithm would take a very long time to find a minimal generating set. But after the decomposition, our algorithm INVARIANT ALGEBRA executed in SINGULAR 3-0-2 finds a minimal generating set after 47.8 minutes using 4.4 Gb in characteristic 0 respectively after only 84.2 seconds using 81.7 Mb in characteristic 7. In both cases, there is a minimal number of 1, 2, 4, 7, 10, 13, 13, 4, 2 generators sorted by degree.

Even using the decomposition, the MAGMA V2.13-9 function `FundamentalInvariants` is unable to find a minimal generating set in less than 4 hours, both in characteristic 0 and in characteristic 7.

### 3. Application to irreducible secondary invariants

In Chapter 7 (our paper [56]), we presented an algorithm for the computation of secondary invariants and a specialised version for the computation of *irreducible* secondary invariants. Shortly after the first version of [56] was posted, there was a new release of MAGMA containing a new algorithm of G. Kemper for the computation of secondary invariants. Unfortunately, to the best of the author's knowledge, Kemper did not describe his new algorithm in a manuscript, yet. So it is not clear how that algorithm differs from the one described in [45], [47] and [22] or the one described in Chapter 7 (see [56]).

Our algorithm for the computation of minimal generating sets can be easily modified to yield yet another algorithm for the computation of irreducible secondary invariants. For this, let  $P$  be a system of primary invariants. In Step (1) of algorithm INVARIANT ALGEBRA, let  $S = P$  and let  $G$  be a Gröbner basis of  $P$ . The rest of the algorithm remains unchanged. In the end, it returns the union of  $P$  with a system of irreducible secondary invariants. Note that this algorithm does not involve an application of Molien's Theorem. So, it applies also to cases when the Molien series is difficult to compute.

In the new version of `irred_secondary_char0` in SINGULAR-3-0-3, we combine both algorithms, i.e., we use the Molien series and power products as described in [56] in low degrees, and the algorithm INVARIANT ALGEBRA in higher degrees.

In Table 7, we compare a  $\beta$ -version of SINGULAR-3-0-3 with MAGMA V2.13-9 (released in January, 2007). As in the examples presented in Chapter 7, the ring variables are called  $x_1, x_2, \dots$ . Let  $e_i$  be the column vector with 1 in position  $i$  and 0 otherwise. For our benchmark, we use Expl. (4)–(9) from Chapter 7 and one additional example (again in characteristic 0), that was originally motivated by our study of ideal Turaev–Viro invariants (compare Section 2 in Chapter 6).

(10) A 20-dimensional representation of  $S_3$  is given by the matrices

$$\begin{aligned} M_1 &= (e_2e_1e_3e_{19}e_9e_{13}e_{17}e_{11}e_5e_{15}e_8e_{16}e_6e_{14}e_{10}e_{12}e_7e_{20}e_4e_{18}) \\ M_2 &= (e_1e_3e_2e_4e_6e_5e_{10}e_9e_8e_7e_{13}e_{16}e_{11}e_{19}e_{20}e_{12}e_{18}e_{17}e_{14}e_{15}) \end{aligned}$$

We use the following sub-optimal primary invariants:

$$\begin{aligned} &x_1 + x_2 + x_3, \quad x_1x_2 + x_1x_3 + x_2x_3, \quad x_1x_2x_3, \quad x_4 + x_{14} + x_{19}, \\ &x_4x_{14} + x_4x_{19} + x_{14}x_{19}, \quad x_4x_{14}x_{19}, \quad x_5 + x_6 + x_8 + x_9 + x_{11} + x_{13}, \\ &x_8x_9 + x_5x_{11} + x_6x_{13}, \quad x_6x_8 + x_5x_9 + x_{11}x_{13}, \\ &x_5x_8 + x_6x_9 + x_6x_{11} + x_9x_{11} + x_5x_{13} + x_8x_{13}, \\ &x_5x_6x_{11} + x_5x_8x_{11} + x_8x_9x_{11} + x_5x_6x_{13} + x_6x_9x_{13} + x_8x_9x_{13}, \\ &x_5^6 + x_6^6 + x_8^6 + x_9^6 + x_{11}^6 + x_{13}^6, \quad x_{12} + x_{16}, \quad x_{12}x_{16}, \\ &x_7 + x_{10} + x_{15} + x_{17} + x_{18} + x_{20}, \quad x_7x_{17} + x_{10}x_{18} + x_{15}x_{20}, \\ &x_{10}x_{15} + x_{17}x_{18} + x_7x_{20}, \quad x_7x_{15} + x_{10}x_{17} + x_7x_{18} + x_{15}x_{18} + x_{10}x_{20} + x_{17}x_{20}, \\ &x_7x_{10}x_{17} + x_7x_{15}x_{17} + x_7x_{10}x_{18} + x_{15}x_{17}x_{20} + x_{10}x_{18}x_{20} + x_{15}x_{18}x_{20}, \\ &x_7^6 + x_{10}^6 + x_{15}^6 + x_{17}^6 + x_{18}^6 + x_{20}^6 \end{aligned}$$

In this example, there are 248832 secondary invariants of maximal degree 26, among which are 283 irreducible secondary invariants of maximal degree 4. The sheer number of secondary invariants (which can be computed by Molien's Theorem) makes the computations hardly manageable for any algorithm that is based on

the generation of power products, as the one described in [45], [47] and [22], or the one described in Chapter 7. It is in fact too much for MAGMA V2.13-9 and for SINGULAR-3-0-2. However, our new algorithm implemented in SINGULAR-3-0-3 just needs few seconds to find the irreducible secondary invariants.

In Table 7, we compare a  $\beta$ -version of SINGULAR-3-0-3 (function `irred_secondary_char0`) with MAGMA V2.13-9 (function `IrreducibleSecondaryInvariants`, released in January, 2007). The only exception is Example (9), that we compute with our new algorithm, but based on SINGULAR-3-0-2. For convenience, we repeat in Table 7 the timings for SINGULAR-3-0-2 and MAGMA V2.13-8 from Table 1 in Section 7.4.

TABLE 7. Comparative benchmark for the computation of irreducible secondary invariants

	SINGULAR 3-0-3	MAGMA V2.13-9	MAGMA V2.13-8	SINGULAR 3-0-2
Expl. (4)	0.07 s 1.48 MB	0.09 7.35 MB	0.48 s 9.09 MB	0.32 s 2.97 MB
Expl. (5)	5.71 s 12.0 MB	0.49 s 9.06 MB	6.66 s 31.82 MB	9.69 s 17.0 MB
Expl. (6)	1.32 s 7.49 MB	2.49 s 19.8 MB	118.51 s 54.0 MB	16.55 s 39.0 MB
Expl. (7)	0.34 s 3.74 MB	36.57 s 30.1 MB	> 7 h > 15 GB	20.94 s 35.1 MB
Expl. (8)	1.06 s 9.37 MB	> 72 min > 2.5 GB	— (259.5 GB)	50.7 min 3.36 GB
Expl. (9)	17.2 min 4.67 GB	29.9 min 399.5 MB	— (55.62 GB)	99.2 min 7.35 GB
Expl. (10)	6.54 s 34.3 MB	> 280 min > 9.9 GB		— —

The outcome of these benchmarks is less clear than of our benchmarks on minimal generating sets. In 3 of the 7 examples, our algorithm and the one used in MAGMA V2.13-9 show more or less the same performance, in one example MAGMA is faster by a factor of about 10, and in 3 examples our algorithm is faster by factors between 100 and at least 4000.

Note that in Expl. (9), the critical part is the computation of a Gröbner basis of primary and irreducible secondary invariants. The rest of the computations just takes about 5 minutes. The beta version of SINGULAR-3-0-3 spends much more than 30 minutes with the computation of a Gröbner basis. Here, the old version SINGULAR-3-0-2 happens to be quicker.



## Bibliography

- [1] S. Akbulut and R. Kirby: A potential smooth counterexample in dimension 4 to the Poincaré conjecture, the Schoenflies conjecture, and the Andrews–Curtis conjecture. *Topology* **24**, pp.375–390 (1985).
- [2] A. Altshuler, J. Bokowski and J. Schuchert: Neighborly 2-manifolds with 12 vertices. *J. Combinatorial Theory (A)*, **75**, pp. 148–162 (1996).
- [3] J. J. Andrews and M. L. Curtis: Free groups and handlebodies. *Proc. Amer. Math. Soc.* **16**, pp. 192–195 (1965).
- [4] D. Barnette: Generating the triangulations of the projective plane. *J. Combin. Theory Ser. B* **33**, pp. 222–230 (1982).
- [5] A. Björner, M. Las Vergnas, B. Sturmfels, N. White and G. Ziegler: *Oriented Matroids*. Cambridge University Press, (1993) second edition 1999.
- [6] I. Bobtcheva and F. Quinn: The reduction of quantum invariants of 4-thickenings. *Fund. Math.* **188**, pp. 21–43 (2005).
- [7] J. Bokowski, S. King, S. Mock and I. Streinu: The topological representation of oriented matroids. *Discrete Comput. Geom.* **33,4**, pp. 645–668 (2005).
- [8] J. Bokowski, S. Mock and I. Streinu: On the Folkman–Lawrence topological representation theorem for oriented matroids of rank 3. *European J. of Combinatorics*, **22,5**, pp. 601–615 (2001).
- [9] J. Bokowski, T. Pisanski: Oriented matroids and complete-graph embeddings on surfaces. To appear in *J. Combin. Theory Ser. A*
- [10] A. V. Borovik, E. I. Khukhro and A. G. Myasnikov: The Andrews–Curtis conjecture and black box groups. *Intern. J. Algebr. Comput.* **13,4**, pp. 415–436 (2003).
- [11] W. Bosma, J. Cannon and C. Playoust: The Magma algebra system I: The user language. *J. Symb. Comp.*, **24, 3/4**, pp. 235–265 (1997).
- [12] L. Bravais and A. Bravais: Essai sur la disposition des feuilles curvisériées. *Annals of Science Nat Botany Biology Vég* **7**, pp. 42–110, 193–221, 291–348 and **8**, pp. 11–42 (1837).
- [13] A. Braun: Vergleichende Untersuchung über die Ordnung der Schuppen an den Tannenpflanzen als Einleitung zur Untersuchung der Blattstellung überhaupt. *Nova Acta Ph Medical Academy of Cesar Leop Carolina Nat Curiosorum* **15**, pp. 195–402 (1831).
- [14] M. Brickenstein: Slimgb: Gröbner Bases with Slim Polynomials. In: *Rhine Workshop on Computer Algebra*. Proceedings of RWCA’06, Basel, March 2006.
- [15] M. Brown: A proof of the generalized Schoenflies theorem. *Bull. Amer. Math. Soc.*, **66**, pp. 74–76 (1960).
- [16] M. Brown: Locally flat imbeddings of topological manifolds. *Ann. of Math. (2)*, **77**, pp. 331–341 (1962).
- [17] B. G. Casler: An imbedding theorem for connected 3-manifolds with boundary. *Proceedings of the AMS* **16**, pp. 559–566 (1965).
- [18] J. W. S. Cassels: *An Introduction to Diophantine Approximation*. Cambridge Tracts in Mathematics and Mathematical Physics, 45. Cambridge University Press 1957.
- [19] A. J. Casson and C. McA. Gordon: Reducing Heegaard splittings. *Topology Appl.* **27,3**, pp. 275–283 (1987).
- [20] R. H. Crowell and R. H. Fox: *Introduction to knot theory*. Ginn and Co. (1963).
- [21] W. Decker, G. Pfister and H. Schönemann: `primdec.lib`. A SINGULAR 3.0 library for computing Primary Decomposition and Radical of Ideals (2005).
- [22] H. Derksen and G. Kemper: *Computational invariant theory*. Invariant Theory and Algebraic Transformation Groups, I. Encyclopaedia of Mathematical Sciences, **130**. Springer-Verlag, Berlin, 2002.
- [23] A. K. Dewdney: Wagner’s theorem for torus graphs. *Discrete Math.*, **4**, pp. 139–149 (1973).
- [24] J. Edmonds and A. Mandel: Topology of Oriented Matroids. Abstract 758-05-9, *Notices Amer.Math.Soc.* **25**, pp. A-410 (1978); also Ph.D. Thesis of A. Mandel, Univ. of Waterloo, 1982.
- [25] D. Eisenbud: *Commutative algebra. With a view toward algebraic geometry*. Graduate Texts in Mathematics, **150**. Springer-Verlag, New York, 1995.

- [26] L. Fejes Tóth: *Lagerungen in der Ebene, auf der Kugel und im Raum*. Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Band LXV. Springer-Verlag, Berlin-Göttingen-Heidelberg (1953).
- [27] R. Fröberg: *An introduction to Gröbner bases*. Pure and Applied Mathematics. John Wiley & Sons, Ltd., Chichester, 1997.
- [28] C. Frohman, J. Kania-Bartoszyńska: The Quantum Content of the Normal Surfaces in a Three-Manifold. Preprint, [arXiv:math/0310273](https://arxiv.org/abs/math/0310273), 2003.
- [29] G.-M. Greuel, G. Pfister and H. Schönemann: SINGULAR 3-0-2. A Computer Algebra System for Polynomial Computations. Centre for Computer Algebra, University of Kaiserslautern (2005). <http://www.singular.uni-kl.de>.
- [30] L. Gutierrez Novoa: On  $n$ -ordered sets and order completeness. *Pacific J. Math.* **15**, pp. 1337-1345 (1965).
- [31] W. Haken: Theorie der Normalflächen. *Acta Math.* **105** pp. 245-375 (1961).
- [32] W. Haken: *Some results on surfaces in 3-manifolds*. In: Studies in Modern Topology (MAA Stud. Math. 5, Prentice Hall 1968), pp. 39-98. Editor Hilton, P.
- [33] J. Hass and J. C. Lagarias: The number of Reidemeister moves needed for unknotting. *J. Amer. Math. Soc.* **14,2**, pp. 399-428 (2001).
- [34] G. Havas and C. Ramsay: Breadth-first search and the Andrews-Curtis. *Intern. J. Algebr. Comput.* **13,1**, pp. 61-68 (2003).
- [35] G. Hemion: *The Classification of Knots and 3-Dimensional Spaces*. Oxford University Press 1992.
- [36] G. Hemion: On the classification of homeomorphisms of 2-manifolds and the classification of 3-manifolds. *Acta Math.* **142, 1-2**, pp. 123-155 (1979).
- [37] A. Heydtmann: *Generating Invariant Rings of Finite Groups*. Diploma Thesis, Saarbrücken, 1996.
- [38] W. Hochstättler: *Seitenflächenverbände orientierter Matroide*, PhD Thesis, Köln, 1992.
- [39] C. Hog-Angeloni, W. Metzler and A.J. Sieradski (eds.): *Two-dimensional homotopy and combinatorial group theory*. London Mathematical Society Lecture Note Series, 197. Cambridge University Press, Cambridge, 1993.
- [40] R. V. Jean: *Phyllotaxis — a Systematic Study of Plant Morphogenesis*. Cambridge University Press, Cambridge/UK, 1994.
- [41] K. Johannson: *Topology and combinatorics of 3-manifolds*. Lecture Notes in Mathematics 1599, Springer-Verlag 1995.
- [42] V. F. R. Jones: A polynomial invariant for knots via von Neumann algebras. *Bull. Amer. Math. Soc. (N.S.)* **12**, pp. 103-111 (1985).
- [43] R. M. Kashaev: The hyperbolic volume of knots from the quantum dilogarithm. *Lett. Math. Phys.* **39**, pp. 269-275 (1997).
- [44] L.H. Kauffman: *Knots and physics*. Series on Knots and Everything, 1. World Scientific Publishing Co., Inc., River Edge, NJ, 1991.
- [45] G. Kemper: Computational Invariant Theory. In: *The Curves Seminar at Queen's*. Vol. XII (Kingston, ON, 1998), pp. 5-26. Queen's Papers in Pure and Appl. Math., **114**. Queen's Univ., Kingston, ON, 1998.
- [46] G. Kemper: Personal communication (Oberwolfach, 2006).
- [47] G. Kemper and A. Steel: Some Algorithms in Invariant Theory of Finite Groups. In: *Proceedings of the Euroconference on Computational Methods for Representations of Groups and Algebras, Essen, 1997*, pp. 267-285. Eds.: P. Dräxler, G. O. Michler and C. M. Ringel. Progr. Math. **173**, Birkhäuser, Basel, 1999.
- [48] S. King: The size of triangulations supporting a given link. *Geom. Topol.* **5**, pp. 369-398 (2001).
- [49] S. King: Regular flip equivalence of surface triangulations. *Topology Appl.* **127,1-2**, pp. 169-173 (2003).
- [50] S. King: How to make a triangulation of  $\mathbb{S}^3$  polytopal. *Trans. Amer. Math. Soc.* **356,11**, pp. 4519-4542 (2004).
- [51] S. King: Complexity of triangulations of the projective space. *Topology Appl.* **132,3**, pp. 261-270 (2003).
- [52] S. King: Crossing number of links formed by edges of a triangulation. *J. Knot Theory Ramifications* **12,3**, pp. 281-286 (2003).
- [53] S. King, F. Beck and U. Lüttge: The Mystery of the Golden Angle in Phyllotaxis. *Plant, Cell & Environment* **27**, pp. 685-695 (2004).
- [54] S. King: Ideal Turaev-Viro invariants. *Sib. Elektron. Mat. Izv.* **3**, pp. 62-66 (2006).
- [55] S. King: Ideal Turaev-Viro invariants. *Topology and Its Applications* **154** (2007), pp. 1141-1156.

- [56] S. King: Fast Computation of Secondary Invariants. [arXiv:math.AC/0701270](https://arxiv.org/abs/math/0701270). Preprint, January 2007.
- [57] Kneser, H.: Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten. *Jahresber. Deut. Math. Ver.* **38**, 248–260 (1929).
- [58] D. E. Knuth: *Axioms and Hulls*. Springer Verlag, New York, 1992.
- [59] M. Kreuzer and L. Robbiano: *Computational commutative algebra 1*. Springer-Verlag, 2000.
- [60] L. D. Landau and E. M. Lifshitz: *Course of Theoretical Physics*, Vol. 3, Quantum mechanics: non-relativistic theory. Translated from the Russian by J. B. Sykes and J. S. Bell Addison-Wesley Publishing Co. (1958).
- [61] M. Las Vergnas: *Extensions ponctuelles d'une géométrie combinatoire orientée*. Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), pp. 265–270, Colloq. Internat. CNRS, 260, CNRS, Paris, 1978.
- [62] Tao Li: Heegaard Surfaces and Measured Laminations, II: Non-Haken 3-Manifolds. *Journal of the Amer. Math. Soc.* **19**,**3**, pp. 625–657 (2006).
- [63] Maple 10. Maple is a trademark of Waterloo Maple Inc.
- [64] S. V. Matveev: Transformations of special spines and the Zeeman conjecture. *Math. USSR-Izvestiya* **31**, pp. 423–434 (1988).
- [65] S. V. Matveev and T. Nowik. On 3-manifolds having the same Turaev-Viro invariants. *Russian journal of mathematical physics* **2**(**3**), pp. 317–324 (1994).
- [66] S. V. Matveev, M. A. Ovchinnikov and M. V. Sokolov: On a simple invariant of Turaev–Viro type. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* **234**, pp. 137–142 (1996). English translation in *J. Math. Sci.* **94**, pp. 1226–1229 (1999).
- [67] S. V. Matveev: Personal communication (Orsay, 1999).
- [68] S. V. Matveev: *Algorithmic topology and classification of 3-manifolds*. Algorithms and Computation in Mathematics, 9. Springer-Verlag, Berlin, 2003.
- [69] S. V. Matveev: Tabulation of three-dimensional manifolds. *Uspekhi Mat. Nauk* **60**,**4**, pp. 97–122 (2005); translation in *Russian Math. Surveys* **60**,**4**, pp. 673–698 (2005).
- [70] H. Meinhardt: Models of pattern formation and their application to plant development. In *Positional Controls in Plant Development*, pp. 1–32 (eds. P. W. Barlow and D. J. Carr) Cambridge University Press, Cambridge/UK 1984.
- [71] A. Mijatović: Triangulations of fibre-free Haken 3-manifolds. *Pacific J. Math.* **219**,**1**, pp. 139–186 (2005).
- [72] E. E. Moise: Affine structures in 3-manifolds. *Ann. Math.* **56**, pp. 96–114 (1952).
- [73] E. E. Moise: Affine structures in 3-manifolds. V. The triangulation theorem and Hauptvermutung. *Ann. of Math* **57**, 547–560 (1952).
- [74] L. Mosher: Mapping class groups are automatic. *Ann. of Math.*, **142**, pp. 303–384 (1995).
- [75] A. Nakamoto and K. Ota: Note on irreducible triangulations of surfaces. *J. Graph Theory*, **20**,**2**, pp. 227–233 (1995).
- [76] S. Negami and S. Watanabe: Diagonal transformations of triangulations of surfaces. *Tsukuba J. Math.*, **14**, pp. 155–166 (1990).
- [77] S. Negami: Diagonal flips in triangulations of surfaces. *Discrete Math.*, **135**, pp. 225–232 (1994).
- [78] U. Pachner: P.L. Homeomorphic Manifolds are Equivalent by Elementary Shellings. *Europ. J. Combinatorics* **12**, pp. 129–145 (1991).
- [79] R. W. Pearcy and W. Yang: The functional morphology of light capture and carbon gain in the Redwood forest understory plant *Adenocaulon bicolor* Hook. *Functional Ecology* **12**, pp. 543–552 (1996).
- [80] R. Piergallini: Standard moves for standard polyhedra and spines. *Rendiconti del circolo matematico di Palermo* **37**, pp. 391–414 (1988).
- [81] A. V. Prasad: Note on a theorem of Hurwitz. *Journal of London Mathematical Society, Series I*, **23**, pp. 169–171 (1948).
- [82] E. S. Rapaport: Groups of order 1, some properties of presentations. *Acta Math.* **121**, pp. 127–150 (1968).
- [83] C. Rourke and B. Wiest: Order automatic mapping class groups. *Pacific J. Math.* **194**, **1**, pp. 209–227 (2000).
- [84] Rubinstein, J. H.: *Polyhedral minimal surfaces, Heegaard splittings and decision problems for 3-dimensional manifolds*. Proc. Georgia Topology Conf. (Amer. Math. Soc./ Intl. Press, 1993).
- [85] M. Scharlemann: Mathematical Reviews, MR1211184 (94g:57015).
- [86] C. F. Schimper: Geometrische Anordnung der um eine Axe periferischen Blattgebilde. *Verhandl. Schweiz. Naturf. Ges.* **21**, pp. 113–117 (1836).
- [87] A. Schürmann: On extremal finite packings. *Discrete Comput. Geom.* **28**, **3**, pp. 389–403 (2002).
- [88] B. Sturmfels: *Algorithms in Invariant Theory*. Springer-Verlag, Wien/New York 1993.

- [89] E. Swartz: Representations of matroids, *J. Amer. Math. Soc.* **16**, pp. 427–442 (2003).
- [90] A. Szűcs: The fourth oriented cobordism group  $\Omega_4$  is isomorphic to  $\mathbb{Z}$ . *J. Math. Sci., New York* **113**, pp. 893–897 (2003).
- [91] N. M. Thiéry: Library PERMUVAR of MuPAD. <http://PerMuVAR.sourceforge.net/>
- [92] N. M. Thiéry: Summary of computations of minimal generating sets of invariant rings of transitive permutation groups. Comparative benchmark (2001). <http://permuvar.sourceforge.net/Groups/>
- [93] N. M. Thiéry: Computing minimal generating sets of invariant rings of permutation groups with SAGBI-Gröbner basis. In: *Discrete models: combinatorics, computation, and geometry (Paris, 2001)*. Discrete Math. Theor. Comput. Sci. Proc., AA, Maison Inform. Math. Discrèt. (MIMD), Paris, pp. 315–328 (2001).
- [94] N. M. Thiéry: Personal communication (2007).
- [95] Thompson, A.: Thin position and the recognition problem for  $S^3$ . *Mathematical Research Letters* **1**, 613–630 (1994).
- [96] V. G. Turaev, O. Y. Viro: State sum invariants of 3-manifolds and quantum  $6j$ -symbols. *Topology* **31**, pp. 865–902 (1992).
- [97] V. G. Turaev: *Quantum invariants of knots and 3-manifolds*. de Gruyter Studies in Mathematics, 18. Walter de Gruyter & Co. 1994.
- [98] K. Wagner: Bemerkungen zum Vierfarbenproblem. *Jahresber. Deutsch. Math.-Verein.*, **46** Abt. 1, pp. 26–32 (1936).

# Index

- 2-stratum, **42**
- 3-deformation, **52**, 51–55
- 3-deformation invariant, **53**
- $6j$ -symbol, **43**
- $6j4k$ -symbol, **43**
  
- admissibility condition, 46
- Andrews–Curtis equivalence, **51**
- Andrews–Curtis conjecture, 50–55
- Andrews–Curtis trivial, **51**
  
- balanced presentation, **50**
- best approximation, **20**
  
- chirotope, **34**
- complexity, **47**
- continued fraction, **19**, 19–20, 22, 25
- contraction
  - oriented matroid, **35**
  - triangulation, **12**
  
- divergence angle, **25**
  
- $\epsilon$ -invariant, 47
- effective leaf width, **28**
- expansion, **12**
  
- face subdivision, **12**
- flip, **11**
- fundamental invariants, **71**
- fundamental surfaces, **18**
  
- generating set, 71–80
- Golden Angle, **25**
- Golden Section, **21**, 25
- Gröbner basis, **45**
  - with degree bound, 63, **63**
  
- hyperline sequence, **32**, **33**
  
- ideal membership, 46
- ideal Turaev–Viro invariant, **44**
- invariant ring, **60**, 59–81
  - non-modular, 60
  
- Jones polynomial, 39, 57
  
- leading monomial, **45**, **62**
- leading term, **62**
- Lucas series, **25**
  
- modular
  - invariant, 53
  - invariant ring, 60
  - monomial ordering, 45
  
- Noetherian normalization, **60**
- non-modular case, **60**
- normal form, **46**
- normal isotopy, **17**
- normal surface, 15, **17**, 17–20
  - fundamental, **18**
- normalization, **60**
  
- orbit sum, **74**
- oriented matroid, **33**, 31–38
  - contraction, **35**
  - deletion, **35**
  - of full rank, **35**
- oriented pseudosphere, **34**
  
- phyllotaxis, 25–29
- primary invariants, **60**
- pseudosphere, **34**
  
- quantum group, 44
  
- radical, **44**, 48
  
- $S$ -polynomial, **62**
- Schönflies theorem, 34
- Schimper–Braun series, **25**
- secondary invariants, **60**, 59–69
  - irreducible, 60, 80–81
- simple 2-polyhedron, **41**
- singular graph, **42**
- special spine, 41, **42**
- state sum, **43**, **54**, **56**
  
- tame, **34**
- transitive permutation groups, **74**
- triangulation
  - irreducible, **12**, 16
  - regular/singular, 11
- true edges, **42**
- true vertices, **42**
- Turaev–Viro ideal, **44**
- Turaev–Viro invariant
  - ideal, **44**, 41–55
  - numerical, 41, 44, **44**
  - universal numerical, **44**, 46
- Turaev–Viro state sum, **43**