Book reviews

Mathematics by experiment: plausible reasoning in the 21st century

J. Borwein and D. Bailey
AK Peters Wellesley MA 2003
ISBN 1-56881-211-6

Experimentation in mathematics: computational paths to discovery

J. Borwein, D. Bailey and R. Girgensohn
AK Peters Wellesley MA 2004
ISBN 1-56881-136-5

Perhaps the first piece of advice I was given by my PhD supervisor about doing mathematics was

> "*It is always easier to prove something when you know it is true.*"

For me, this statement highlights the difference between the way that most of us *do* mathematics and the way that we *present* it.

When we publish a proof we are often like a magician showing their latest trick — a slick, polished and beautiful performance, that (hopefully) entertains the audience and leaves them in awe of the spectacle and impressed by our ingenuity. We move in a crescendo of logic from definitions to lemmas to our main theorem; always onwards and upwards.

On the other hand, we do not actually *do* mathematics in this way. In fact it is probably more accurate to say that we start with the theorem and work backwards:

> "*I have the result, but I do not yet know how to get it.*"
>
> — C. F. Gauss

All mathematicians build their intuition about problems and objects by looking at examples. When this intuition becomes strong enough, we *know* the result before we have a proof of it. And once we have the theorem we set about proving it. Arguably mathematics has been this way for a very long time; while it is different from other areas of knowledge in that it has the certainty of proof, as practitioners we are not immune from getting our hands a little dirty with some experimentation — though we are generally not willing to admit it and tend to hide it if possible.

The message of these two books is that it is time to embrace experimentation as part of mathematics, not hide it. And this is now possible because the computer has made wide-spread, systematic experimentation a reality.

Moore's law, and all the engineers who have perpetuated it, have made computers powerful, cheap and ubiquitous. We now have an abundance of raw computing power and sophisticated mathematical software, both free (such as GAP) and commercial (such as Maple and Mathematica), at their disposal. This gives us a "laboratory" in which we can perform numerical and symbolic explorations on a grand scale without pages of painful by-hand calculations. In so doing we can test conjectures and guess entirely new ones. This approach is leading to new results discovered and even proved partially or entirely with the aid of computers.

These two books are devoted to exploring this approach and are aimed at a wide cross-section of mathematically-trained readers from (roughly) honours level and up.

The first, *Mathematics by experiment*, describes the ideas behind experimental mathematics (giving those words a far more concrete definition than I have done here) and devotes some time to the more philosophical implications suggested by the title — including a very interesting section on paradigm shifts. The bulk of the book is made up of a series of examples of experimental mathematics in action. Reflecting the expertise of the authors, the examples come mainly from number theory and combinatorics, covering topics such as the digits of π and normality, but also included are smaller examples from a wide range of other topics such as chaos theory and knot theory. The book is perhaps best read with a computer nearby so that you can tinker and try things out as you read. The examples are also complemented by commentaries and exercises for the reader.

The chapter on π gives an account of various computational methods that have been employed to find more and more terms of its expansion in different bases. This culminates in the famous result by Peter Borwein, David Bailey and Simon Plouffe:

$$\pi = \sum_{i=0}^{\infty} \frac{1}{16^i}\Big(\frac{4}{8i+1} - \frac{2}{8i+4} - \frac{1}{8i+5} - \frac{1}{8i+6}\Big).$$

The importance of this result lies in that it can be used to compute individual digits of π — one can compute the d^{th} binary digit without having to compute any of the previous digits! The formula was not discovered by formal reasoning or symbolic calculations, but rather was discovered by high precision *numerical* calculations.

The computation of digits of π then leads naturally into a more general discussion of the *normality* of numbers — a number is *normal* if its expansion (in a given base) behaves like a random sequence. While it has been shown that almost all numbers are normal, the only numbers *proved* to be normal are very contrived. On the other hand, numerical analyses show many common mathematical constants, such as π and $\sqrt{2}$, *appear* to be normal. The authors describe how results like that given for π above, have lead to significant progress in the theory of normal numbers. A careful experimental "verification" of the normality of algebraic numbers is described later in the book.

The second-last chapter of *Mathematics by experiment* describes many of the numerical tools and techniques that are required (such as high-precision arithmetic and integer relation detection) and also give many useful links to implementations of these methods. The book then ends with a reprint of the article "*Making sense of experimental mathematics*" by J.M. Borwein, P.B. Borwein, R. Girgensohn and S. Parnes.

The second text, *Experimentation in mathematics* continues with further chapters of examples and numerical and symbolic techniques. This text moves through a broader range of topics and is more demanding of the reader than the first. Number theory and combinatorics are prominent amongst the topics covered; the chapter on zeta functions details some of the work by the authors on special values of zeta and multizeta functions and shows how this work grew from and inspired their interest in experimental methods. Again, while topics closest to the authors' experience, form the bulk of the book, there are many sections that cover problems from other areas of mathematics including probability theory and Fourier series.

While the both books may be read and used independently, I would suggest reading the *Mathematics by experiment* (particularly because of the philosophical and historical discussions) before delving into the heavier *Experimentation in mathematics*.

In addition to a large quantity of mathematical theory, examples and problems they contain, the books are littered with numerous interesting (and often colourful) stories and histories of people and theorems.

To summarise, I do not think I that have had the good fortune to read two more entertaining and informative mathematics texts.

Andrew Rechnitzer
Department of Mathematics, University of Melbourne, VIC 3010
E-mail: A.Rechnitzer@ms.unimelb.edu.au

◇ ◇ ◇ ◇ ◇ ◇

Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond

B. Schölkopf and A.J. Smola
MIT Press 2002
ISBN 0-262-19475-9

According to Tomaso Poggio and Steve Smale [6]:
"We believe that a set of techniques based on a new area of science and engineering becoming known as "supervised learning" will become a key technology to extract information from the ocean of bits around us and to make sense of it."
When mathematical scientists of this stature make such strong statements, we should take note. In broad terms, learning theory (also known as machine learning) deals with developing rules for making decisions based on learning from examples or data. For example, we may develop rules for classifying a new email message as junk mail by using the experience of previous messages.

In recent years there have been several books written on learning theory that may appeal to readers of the *Gazette*: see the references below for a selection. Some, such as Hastie *et al.* [3], may appeal to statisticians; others, such as those by Cristianini and Shawe-Taylor [2], [8], or Kecman [4] may appeal to those with a computer science bent. The present book under review by Bernhard Schölkopf (Max Planck Institute for Biological Cybernetics) and Alexander Smola (ANU) may appeal to mathematicians. Learning theory brings together many aspects of pure mathematics, mathematical modelling, computational mathematics, probability, statistics, and computer science. Reading through books on the subject makes one wonder about the wisdom of the ways in which universities compartmentalise these parts of mathematics in the curriculum or in different departments.

Perhaps many of us associate the phrase "machine learning" with neural networks and probably don't think more about it. Schölkopf and Smola have written a book of 626 pages on machine learning and rarely mention neural networks. They have produced an introduction to kernel based methods of machine learning which use ideas from classical machine learning theory, optimization, and mathematical analysis—including the ubiquitous reproducing kernel Hilbert spaces. The canonical example of this approach is the Support Vector Machine (SVM) about which they write:
"... successful applications have demonstrated that SVMs not only have a more solid foundation than artificial neural networks, but are able to serve as a replacement for neural networks that perform as well or better, in a wide variety of fields."
(For an extensive discussion of the links between approximation theory and neural networks, see [5].)

The book opens with a tutorial. Chapter 1 is an informal introduction to SVMs and associated ideas intended to orient the reader. The authors use a simple example of classifying certain objects into one of two classes (e.g. classify an email message as junk or not junk). A subset of the objects is used for training, and each of these objects is mapped into an inner product space known as the feature space. Classification is then a process of determining a hyperplane which separates the vectors representing the objects.

If there is more than one such separating hyperplane, we look for the optimal one in some well-defined sense (roughly corresponding to keeping the hyperplane as far from both classes as possible). Reverting back to the input space where the objects live, we have a nonlinear decision boundary corresponding to the hyperplane in the feature space. The hyperplane can be specified, up to translation, by a vector orthogonal to it. It turns out that this vector is a linear combination of the training vectors: and those training vectors which have a non-zero contribution to this linear combination are called "support vectors". Essentially these support vectors are those that are closest to the separating hyperplane. On the other hand, if there is no separating hyperplane, then we compromise by choosing a hyperplane that misclassifies a small number of objects.

Thus far, the constrained optimization to compute the hyperplane has occurred in the feature space, but we can avoid this and work directly in the object space by introducing the "kernel trick". The kernel in question is defined on a pair of objects as being the inner product of their representations in the feature space, and the optimization is rewritten in terms of the object space. The optimization performed in the feature space can then be rewritten so as to occur in the object space. In fact, we can choose the kernel first, so long as the implicitly defined mapping of objects to feature vectors satisfies certain conditions. This freedom in choosing the kernel is an attractive feature of these algorithms: in the words of the authors (p. 34),

"Given an algorithm which is formulated in terms of a positive definite kernel k, one can construct an alternative algorithm by replacing k by another positive definite kernel $\tilde{k}$."

This disarmingly simple statement opens the way to algorithms on sets much more general than inner product spaces, by effectively embedding the object space in a linear space through the use of an appropriate kernel.

The rest of the book is divided into three parts. Part I brings together the necessary mathematical tools; ideas from probability, approximation theory, and optimization are important. Part II treats SVMs in detail, and Part III deals with kernel methods more generally. There is an extensive bibliography of more than 600 items, and there are many exercises ranging in difficulty from simple exercises to suggestions for further research.

Although the "study of patterns in data is as old as science" [8, p. xi], studying patterns through machine learning is an important, exciting part of contemporary science. This book is an excellent work by leaders in the field; it would make an appropriate text for a seminar series for those who wanted to learn more about support vector machines. Given the importance of the subject and the quality of the book, we recommend that *Learning with Kernels* by Schölkopf and Smola should be in every university library.

References

[1] C. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press Oxford 1995).

[2] N. Cristianini and J. Shawe-Taylor, *Support Vector Machines and Other Kernel-Based Learning Methods* (Cambridge University Press Cambridge 2000).

[3] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Verlag New York 2001).

[4] V. Kecman, *Learning and Soft Computing: Support Vector Machines, Neural Networks and Fuzzy Logic Models* (MIT Press Cambridge Mass. 2001).

[5] A. Pinkus, *Approximation theory of the MLP model in neural networks*, Acta Numerica, **8** (1999), 143–195.

[6] T. Poggio and S. Smale, *The mathematics of learning: Dealing with data*, Notices of Amer. Math. Soc. **50** (5) (2003), 537–544.

[7] B. Ripley, *Neural Networks for Pattern Recognition* (Cambridge University Press Cambridge 1996).
[8] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis* (Cambridge University Press Cambridge 2004).
[9] J. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines* (World Scientific Singapore 2002).

C.T. Lenard and T.M. Mills
La Trobe University, PO Box 199, Bendigo VIC 3552
E-mail: c.lenard@latrobe.edu.au
and t.mills@latrobe.edu.au

◇ ◇ ◇ ◇ ◇ ◇

Computational Algebraic Geometry

Hal Schenck
Cambridge University Press Cambridge 2004
ISBN 0-521-53650-2

Algebraic geometry can be a forbidding area of study. It requires a background in commutative algebra, homological algebra and topology. And although some examples and definitions can be easily motivated with pictures, the definition of schemes, cohomology groups and so on are abstract and not easy to assimilate. This can make reading the standard references such as Hartshorne [4] or EGA [3] daunting indeed.

Now Hal Schenck has written a short computationally inspired book providing lots of hands on examples. In the preface he writes that his aim is to include "snapshots" from commutative algebra, algebraic geometry, algebraic topology and algebraic combinatorics, with homological algebra as a common backdrop. The free computer algebra system Macaulay 2 is used throughout the book to carry out the computations; a brief introduction to its syntax is given in an appendix.

The author's enthusiasm is clear throughout and he succeeds very well in his primary goal of giving explicit computational examples.

The text is quite informal. Many definitions are found within examples or discussions: (quasi)compact, Euler characteristic, intersection multiplicity etc. The exposition is fairly terse. To keep the book short, many results are stated without proof, such as the Nullstellensatz, and the existence of long exact sequences in cohomology. Other proofs are relegated to exercises (even exercises appearing many chapters later).

Prerequisites for the book include some commutative algebra and point-set topology. In practice a fairly high degree of mathematical maturity is required as well.

Chapter 1 is on affine algebraic geometry: affine varieties, primary decomposition of ideals, the Zariski topology and the correspondence between ideals and affine algebraic sets.

Chapter 2 introduces projective space and graded objects, and then defines the Hilbert series of a graded module. As an example of the computational theme, the dimension of a projective variety is then defined via the Hilbert series.

Chapter 3 discusses free resolutions. Chapter 4 gives an introduction to Gröbner bases, and discusses how they may be used to compute Hilbert polynomials. Chapter 5 introduces simplicial complexes and simplicial homology.

Chapters 6 and 8 define localization, Hom and tensor products. The exactness properties of these functors are discussed and the derived functors Tor and Ext are introduced. Chapter 7 discusses regularity and the Hilbert function.

Chapter 9 introduces sheaves and Čech cohomology. Chapter 10 returns to commutative algebra, with discussion of Cohen-Macaulay rings and related topics. There are two appendices, one on abstract algebra and an introduction to Macaulay, and one on complex analysis.

The index and bibliography are reasonably comprehensive, and directions to the

literature are given regularly in the text. Many pictures are included throughout. A small number of typos are listed on the author's webpage at http://www.math.tamu.edu/~schenck/errata.pdf.

The competition for this book is probably the two books of Cox, Little and O'Shea [1], [2]. These cover basic algebraic geometry, Gröbner bases, elimination theory, Hilbert polynomials and free resolutions, also with calculations in Macaulay. Each of these texts is 500 pages long, and cover less than Schenck does in 200. However they are much more detailed and self-contained and require less effort from the reader.

Schenck's text cannot be read by itself as an introduction to algebraic geometry since so many fundamental results are not proved. Instead, it could be read in parallel with a text such as [4], [5] or [6]. Indeed, Schenck claims that his book is an "advertisement" for more comprehensive texts.

In this role, the book would provide a useful study aide for a motivated and computationally minded graduate student. It is brief, informal and provides many examples—all attributes that many other algebraic geometry books lack! It would also provide a useful source of concrete problems for an instructor. It is not a complete reference in itself; its role is to stimulate and outline rather than to given an encyclopedic treatment, and in this it succeeds well.

References

[1] D. Cox, J. Little, D. O'Shea, *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Second Edition, (Springer-Verlag 1997).

[2] D. Cox, J. Little, D. O'Shea, *Using Algebraic Geometry*, (Springer-Verlag 1998).

[3] A. Grothendieck and J. Dieudonné, *Éléments de Géométrie Algébrique*, Publications mathematiques d'IHES 1960–67.

[4] R. Hartshorne *Algebraic Geometry,* (Springer-Verlag 1992).

[5] M. Miyanishi *Algebraic Geometry*, Translations Mathematical Monographs **136** (AMS 1994).

[6] *Algebraic Geometry 1, 2, 3*, Translations Mathematical Mongographs **185**, **197**, **218** (AMS 1999–2003).

Victor Scharaschkin
Department of Mathematics, University of Queensland, St Lucia, QLD 4072
E-mail: victors@math.uq.edu.au

◇ ◇ ◇ ◇ ◇ ◇

The Cross-Entropy Method: A Unified Approach to Combinatorial Optimisation, Monte-Carlo Simulation, and Machine Learning

R.Y. Rubinstein and D.P. Kroese
Springer Heidelberg 2004
ISBN 0-387-21240-X

The Cross-Entropy (CE) method was invented by the first author of this book, and was initially conceived as a simulation-based procedure for estimating probabilities of rare events in stochastic networks. However, it was soon realised that the method could also be used to solve challenging optimisation problems, both discrete and continuous. This is the first published book on the topic, and it provides a comprehensive introduction to all aspects of the CE method.

The book is based on an advanced undergraduate course given in recent years at the Israel Institute of Technology, and in the author's words, is aimed at engineers, computer scientists, mathematicians and statisticians who are "interested in smart simulation, fast optimisation, learning algorithms, image processing, etc.". Indeed, the book certainly does have a strong interdisciplinary flavour, and aims to give the reader a broad view of the topic; the content ranges from basic theory and pseudo-code for algorithm implementation, through to new ideas and applications which are still under development by researchers in the field.

The first chapter provides a very brief summary of some topics in probability and simulation that feature in subsequent chapters, including probability distributions, estimation of parameters, and the generation of random variables. The reader is assumed to have had prior exposure to elementary ideas in probability. Chapter 2 gives a concise and accessible tutorial-style introduction to the main aspects of the CE method, starting with two concrete examples which allow the reader to quickly get a feel for the nature and scope of the method.

The CE method essentially has two distinct "branches", or areas of application:

- efficient simulation via importance sampling, and
- stochastic optimisation.

The former is presented in detail in Chapter 3, while the latter is covered in Chapter 4-8.

Indeed, the term "Cross-Entropy" originates from the first branch. Specifically, it is well-known that many quantities of interest, such as rare event probabilities, can be estimated by performing importance sampling within a simulation environment. The main idea behind the CE method is to iteratively update the parameters of some nominal importance sampling density, in such a way as to minimise the Cross-Entropy, also known as Kullback-Leibler "distance", between the current density and the optimal importance sampling density. In this regard, the CE method is an alternative to other importance sampling techniques, such as variance minimisation and exponential twisting. Comparisons and advantages of the CE method versus other techniques are described, and a number of applications are presented, including the estimation of probabilities of rare events in single and tandem $GI/G/1$ queues.

A minor modification of the CE method can also be used as a stochastic search method for solving difficult discrete and continuous optimisation problems. In this setting, one maintains and updates a sampling probability density on the set of all candidate solutions, which effectively reflects the estimated likelihood of the location of the optimal solution. In the absence of any prior information about an optimal solution, the initial density is naturally taken to be uniform. Thus, the event of generating the optimal solution is effectively a rare event, and the "machinery" from the rare event setting can be carried over, with some modifications, to the optimisation setting.

In Chapter 4, the authors show how to apply the CE method to a variety of well-known combinatorial optimisation problems, including the well-known travelling salesman, max-cut, and quadratic assignment problems. There is no explicit comparison of the performance of the CE method versus the performance of other optimisation techniques, as this is not the claimed purpose of the book. However, the authors do compare the results obtained via the CE method with the best available estimates of the optimal solutions for established "benchmark" problems, and they demonstrate that the CE method is typically able to obtain very good or optimal solutions with relatively little computational effort.

Chapter 5 is dedicated to continuous unconstrained optimisation, and is relatively short, reflecting that fact that this is a more recent development of the CE method. In particular, the method is used to find minima of the (multi-extremal) Rosenbrock and trigonometric functions.

Chapters 6 and 7 contain more examples of applications to discrete optimisation problems, including DNA sequence alignment, a shortest path Markov decision problem, data clustering and vector quantisation. The treatment of each example is brief, but serves to demonstrate the broad scope of the CE method for simulation-based optimisation.

At the end of each chapter, the authors have compiled a set of exercises and

questions which challenge the reader to engage in both theoretical and practical (computer programming) aspects of the material. Some of these exercises would be suitable for students at an honours-level course, while others lead nicely into research topics. Matlab code for a number of the examples presented in book is provided in the Appendix, and is also available online at http://www.cemethod.org. Relevant computational issues and useful tips for modifying and experimenting with the algorithms are discussed, as the need arises, throughout the text. In addition, the research-oriented reader will find that many parts of the book provide solid starting points and inspiration for further investigation. Indeed, this appears to be one of the authors' aims; to stimulate ideas and research in this relatively new area.

I wholeheartedly recommend this book to anybody who is interested in stochastic optimisation or simulation-based performance analysis of stochastic systems.

Andre Costa
University of Melbourne, VIC 3010
E-mail: acosta@ms.unimelb.edu.au

◇ ◇ ◇ ◇ ◇ ◇

Quantization, Classical and Quantum Field Theory and Theta Functions

Andrei Tyurin
CRM Monograph Series **21**
AMS Providence 2003
ISBN 0-8218-3240-9

This monograph is based on a series of lectures, given by Andrei Nikolaevich Tyurin, at the 'Centre de Recherches Mathématiques' of the University of Montreal, in October 2001, a year before his fatal heart attack. Final editing of the manuscript was carried out by colleagues.

The manuscript is strongly influenced, both in topic and style of presentation, by Arnaud Beauville's classic survey "Vector bundles on curves and generalized theta functions: recent results and open problems" [1], which appeared 10 years ago. In fact, the present monograph aims to answer the final question of Beauville's survey, namely how to generalize the concept of ordinary (abelian) theta functions to the non-abelian case. Whereas abelian theta functions describe properties of moduli spaces of one-dimensional vector bundles (i.e. line-bundles) on algebraic curves, non-abelian theta functions play a similar role in the study of higher-dimensional vector bundles. The monograph describes various aspects of the theory of non-abelian theta functions and the moduli spaces of vector bundles, including their applications to problems of quantization, classical and quantum conformal field theories, spin networks, three-dimensional topology, gauge theory on graphs, and abelian gauge theory.

The book is definitely written for specialists. In order to fully comprehend its contents the reader needs to have more than a rudimentary background in algebraic geometry, geometric quantization and the theory of abelian theta functions. At the same time, in order to understand the applications, one needs a thorough knowledge of mathematical aspects of two-dimensional conformal field theories, in particular those corresponding to the Wess-Zumino-Witten model, e.g. at the level of [2]. The book, by and large, is a compendium of results (mostly from the author's papers on the subject) and lacks explanation, background and proofs. Therefore the monograph is not very suitable to be used as a textbook, but more as introduction, albeit a very advanced one, into a specialised research field and should stimulate further reading to fill in the gaps and details.

From the reviewers point of view the monograph would have benefited from a lengthier introduction into the theory of

abelian theta functions, which might have been the original intention of the author. Nevertheless I think the book provides a useful summary of the theory, and it is commendable that the author's colleagues have brought the draft manuscript to completion. It definitely stimulated my own interest in the field.

References

[1] A. Beauville, *Vector bundles on curves and generalized theta functions: recent results and open problems*, Current Topics in Complex Algebraic Geometry (Berkeley CA 1992/93), Math. Sci. Res. Inst. Publ. **28** (Cambridge University Press Cambridge 1995), 17–33.

[2] A. Tsuchiya and Y. Kanie, *Vertex operators in conformal field theory on $\mathbb{P}^1$ and monodromy representation of braid group*, Adv. Stud. Pure Math. **16** (Academic Press Boston 1988), 297–372.

Peter Bouwknegt
Department of Theoretical Physics, RSPhysSE, and Department of Mathematics, MSI, The Australian National University
E-mail: peter.bouwknegt@anu.edu.au

◇ ◇ ◇ ◇ ◇ ◇

Regular neighbourhoods and canonical decompositions for groups

Peter Scott and Gadde Swarup
Astérisque
Societé Mathématique de France 2003
ISBN 2-85629-146-5

Consider a closed orientable manifold M and a codimension one π_1-injective submanifold N. Then N defines a splitting of $G = \pi_1(M)$ over $H = \pi_1(N)$ as an amalgamated product or HNN extension. When N is only immersed in M, then there is no splitting correponding to N. Instead, the algebraic object corresponding to this situation is what the authors call an *H-almost invariant subset of G* which should be thought as the analogue of a splitting in the case of an immersion.

The goal of the authors is to generalize to finitely presented groups the JSJ decomposition of an orientable Haken 3-manifold M by Jaco-Shalen and Johannson. If M is a closed 3-manifold, this decomposition consists in a maximal disjoint union of non-parallel π_1-injective embedded tori $\mathcal{T}$ having the property that any immersed torus of M can be homotoped to be disjoint with $\mathcal{T}$.

A central notion in this work is the notion of *enclosing*: consider $\mathcal{T}$ a disjoint union of submanifolds of M. If an immersion of N into M can be homotoped to $M \setminus \mathcal{T}$, then N is *enclosed* in the corresponding component of $M \setminus N$. This notion has a natural generalization saying when a H-almost invariant subset of G is enclosed in a vertex of a splitting of G. Enclosure is closely related to the *compatibility* of two splittings: two splittings Γ_1, Γ_2 of G are *compatible* if one can split G into a larger graph of groups Γ such that Γ_1 and Γ_2 can be obtained from Γ by collapsing some edges; a one-edge splitting Γ_1 of G is compatible with another splitting Γ_2 if and only if Γ_1 is enclosed in a vertex of Γ_2.

The main construction of the paper is a *regular neighbourhood* for a family of almost invariant subsets. This is the algebraic counterpart of the regular neighbourhood $N(C)$ of a family of immersed codimension one submanifolds C of M. The subset $N(C)$ satisfies that

- each component of C is contained in a component of $N(C)$ (up to homotopy),
- each component of $N(C)$ contains a component of C (up to homotopy),
- and any submanifold disjoint from C can be homotoped out of $N(C)$.

The boundary of $N(C)$ defines a splitting Γ of $\pi_1(M)$, with two kinds of vertices coming from the components $N(C)$ and from the components of $C \setminus N(C)$. Note that the properties make sense even if C is an infinite collection of immersions, although existence of $N(C)$ is not clear in this case.

In this sense, the JSJ decomposition of a 3-manifold can be seen as a regular neighbourhood of all its immersed tori.

The three properties above can be formulated in terms of enclosing of almost invariant subsets into vertices of Γ. The authors prove in a very general setting, that a splitting Γ satisfying these properties (with some additional technical assumptions) is unique. This is what the author call the *regular neighbourhood* of a family of almost invariant subsets. The uniqueness claimed here is a strong uniqueness, in particular, not up to certain moves.

The main theorem claims that for a finitely presented group G which does not split over a virtually polycyclic group of Hirsch length $< n$, there exists a (unique) regular neighbourhood Γ of all the H-almost invariant subsets of G where H varies among virtually polycyclic groups of Hirsch length n. Moreover, there is a description of this splitting. For simplicity, we only give a rough description for a one-ended torsion free group G where one only consider almost invariant subsets over $\mathbb{Z}$. For each vertex $v \in \Gamma$ of the first kind, one of the following holds:

- Γ_v is cyclic
- Γ_v is the fundamental group of a surface with boundary Σ such that the incident edge groups correspond to the fundamental groups of the components of $\partial\Sigma$.
- Γ_v is the full commensurizer of a cyclic group H such that G/H has more than one end.

Some vertex groups of commensurizer type might fail to be finitely generated, and edge groups may fail to be finitely generated (and thus may fail to be cyclic).

The authors also extend their results to include almost invariant subsets over virtually abelian groups of bounded rank (several ranks may occur) and almost invariant subsets over virtually polycyclic groups of two successive Hirsch lengths.

The splitting obtained is unique (in a strong sense), and therefore invariant under automorphisms. This is one of the main features of the construction. This makes it adapted for the study of automorphisms of G. In this sense, this splitting may be seen as a generalization of Bowditch's JSJ splitting for one-ended hyperbolic group [Bowditch, Cut points and canonical splittings of hyperbolic groups, Acta Math. 1998]. Compared to other JSJ theories by Sela, Rips-Sela, Dunwoody-Sageev and Fujiwara-Papasoglu, the main difference is that here, the central notion is enclosure (or compatibility of splittings). Instead, the central notion in the JSJ theories cited above is the fact that a splitting is *elliptic* with respect to another, which means that edge groups of one splitting are contained in vertex groups of the other splitting (up to conjugacy). Ellipticity of a splitting with respect to another one is a much less rigid requirement than compatibility. This strengthened rigidity is the main reason of the strong uniqueness of the obtained splitting and of its invariance under automorphisms; there is no such strong uniqueness for the JSJ splittings obtained by the authors above. A paper in preparation by Levitt and the reviewer will give a unified setting for those JSJ theories.

This long paper explains in detail the motivations and the topological analogies with immersions. It develops many technical tools devoted to the study of almost invariant subsets. The authors first construct the regular neighbourhood of a finite collection of almost invariant subsets, and prove the uniqueness in general. Under a small commensurizer hypothesis, some finiteness properties are proved allowing to deduce the existence of the regular neighbourhood of all the $\mathbb{Z}$-invariant subsets. The generalization of this finiteness property without this hypothesis is more tricky. Then the authors extend the results to include almost invariant subsets over virtually polycyclic groups. The paper includes, as an appendix, two

preceding papers by the authors concerning the (symmetric) intersection of almost invariant subsets.

Note that P. Scott maintains a web page with errata and corrections on http://www.math.lsa.umich.edu/~pscott/.

Vincent Guirardel
Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4 France
E-mail: guirardel@picard.ups-tlse.fr