| **Section A** |
| --- |

| **Institution:** University of St Andrews |
| --- |

| **Unit of Assessment:** UoA 11: Computer Science and Informatics |
| --- |

| **Title of case study:** More effective teaching of advanced mathematics through new algorithms incorporated in open source mathematical software |
| --- |

| **Period when the underpinning research was undertaken:** 2000 - 2018 |
| --- |

| **Details of staff conducting the underpinning research from the submitting unit:** |
| --- |

| Name(s): | Role(s) (e.g. job title): | Period(s) employed by submitting HEI: |
| --- | --- | --- |
| Steve Linton | Professor | 01 January 1994 - present |
| Christopher Jefferson | Reader | 01 April 2009 - present |
| Tom Kelsey | Professor | 01 December 1999 - present |
| Alexander Konovalov | Lecturer | 01 January 2007 - present |
| Max Neunhöffer | Lecturer | 01 August 2007 - 28 October 2013 |
| Colva Roney-Dougal | Professor | 15 September 2003 - present |

| **Period when the claimed impact occurred:** 01 August 2013 - 31 July 2020 |
| --- |

| **Is this case study continued from a case study submitted in 2014?** No |
| --- |

| |
| --- |

| **Section B** |
| --- |

**1. Summary of the impact** (indicative maximum 100 words)

The effectiveness and efficiency of university-level mathematics teaching and the training of teachers and researchers was increased by new algorithms, data and software developed in St Andrews which enhanced the performance, mathematical scope and usability of software systems widely used in teaching.  This had an impact on courses given to at least 5,700 people (students) at least 27 institutions in 23 countries, as well as weekly on many of the 20,000 people (users) of the CoCalc web-based cloud computing and course management platform for computational mathematics.

Advanced mathematics, and mathematical topics in related disciplines, are an essential, but pedagogically challenging, part of Higher Education teaching in programmes of study ranging from pure mathematics through computer science to physics, chemistry and engineering. The use of software that automates routine calculations is an established technique in mathematics education at all levels. The software allows learners to "come to grips" with the concepts directly without becoming "bogged down" in technical details of calculation. The GAP system, which encapsulates our research, is widely used for this purpose, both directly and through other systems and online services that incorporate it.

The specific research results listed in section 2 extend GAP's capabilities into new mathematical areas, which were the main topics of courses delivered to over 700 people (students), speed up critical aspects of the system allowing students to explore larger and more informative examples, add databases which are a valued source of examples for teaching, and link GAP to other systems providing a better experience for teachers and students.

**2. Underpinning research** (indicative maximum 500 words)

The underpinning research insights from the School of Computer Science can be loosely divided into four groups:

1. **Algorithms for semigroups:** St Andrews research on algorithms for semigroups enabled practically usable software for wide-ranging computation in this important mathematical area. **Semigroups** are a key mathematical structure, used to describe sets of transformations of physical and abstract objects and the ways in which they combine. [R1] is a foundational paper in the area, which has led to at least 30 published research papers (2000-2018) from

St Andrews, 14 St Andrews PhD theses and 5 GAP extension packages with authors from St Andrews.

2. **New and faster algorithms for algebraic problems:** St Andrews research in algorithms for permutation groups and minimal polynomials sped up these calculations dramatically. **Permutation groups** are ubiquitous in mathematics, describing symmetries of objects and structures. Jefferson et al. [R2] developed, implemented and released a family of new algorithms for some central problems in the area, outperforming the best existing implementations massively in some cases. On one test suite in [R2], it completed three times as many examples as the previous state of the art, in one-fifth of the time. These algorithms are widely used "behind the scenes" in mathematical software systems, so improving them speeds up user calculations in areas such as matrix groups, combinatorics and semigroups, not just those directly involving permutation groups. Computing **minimal polynomials** is a similarly fundamental problem for matrices. In 2008, Neunhöffer and Praeger (Western Australia) developed [R3] and implemented a new efficient algorithm, up to 200 times faster on difficult cases from real-world problems (examples in [R3]). This algorithm speeds up many common computations used in areas such as error-correcting codes and cryptography as well purely mathematical applications.

3. **New Classification Results:** St Andrews researchers led by Kelsey in 2012 classified all **the 12,418,001,077,381,302,684 non-equivalent semigroups** of order 10 and constructed a searchable library of the 1,843,973,431 of orders up to 8 [R4]. Roney-Dougal extended the classification of **all primitive permutation groups** from degree 1000 to degree 4095, adding 16,692 new groups [R5]. Permutation groups are ubiquitous mathematical objects, describing ways in which a collection can be reordered while preserving some property, and primitive permutation groups are foundational ones from which all others are built.

4. **Linking Mathematical Software Systems Together:** St Andrews researchers (with collaborators) developed new approaches to linking mathematical software systems to each other and to associated tools such as modern user interfaces. Doing this, without compromising speed of computation or correctness of results, is a challenging and long-standing problem. St Andrews researchers made major progress on this problem in two European projects: SCIEnce (2006-2011) and OpenDreamKit (2015-2019) [R6]. Our research has made major contributions to efficient, mathematically correct and reliable links between diverse mathematical systems and popular user interface tools, such as the Jupyter Notebook system.

These results form part of a much larger body of ongoing research in algorithms, systems and applications around computational algebra, hosted since 2000 in an interdisciplinary research centre (CIRCA – the Centre for Interdisciplinary Research in Computational Algebra) which bridges the University's School of Mathematics and Statistics and the School of Computer Science. One pillar of that research is our ongoing role as a major centre for the development of the GAP (Groups, Algorithms, Programming) system described below. Twenty-five St Andrews researchers have contributed directly to the project, publishing 34 journal and conference papers featuring improved algorithms or system features incorporated in GAP.

**3. References to the research** (indicative maximum of six references)

All outputs have been peer-reviewed in reputable journals or conferences. R6 was an accepted deliverable in a Horizon 2020 EUR7,600,000 project.

[R1] S. Linton, G. Pfeiffer, E.F. Robertson, and N. Ruškuc. Computing transformation semigroups, J. Symbolic Comput., 33 (2) (2002), 145–162. DOI: 10.1006/jsco.2000.0406.

[R2] C. Jefferson, E. Jonauskyte, M. Pfeiffer, and R. Waldecker. Minimal and canonical images. J. Algebra, 521 (2019), pp. 481–506. DOI: 10.1016/j.jalgebra.2018.11.009.

[R3] M. Neunhöffer and C.E. Praeger. Computing minimal polynomials of matrices. LMS J. Comput. Math., 11 (2008), pp. 252–279, DOI: 10.1112/S1461157000000590.

[R4] A. Distler, C.A. Jefferson, T. Kelsey, and L. Kotthoff. The semigroups of order 10. In 'Principles and Practice of Constraint Programming: 18th International Conference', Springer.

2012. pp. 883–899, DOI: [10.1007/978-3-642-33558-7_63](10.1007/978-3-642-33558-7_63).

[R5] H. J. Coutts, M. Quick, and C. M. Roney-Dougal. The primitive permutation groups of degree less than 4096. Communications in Algebra, 39(10) (2011), 3526-3546, DOI: [10.1080/00927872.2010.515521](10.1080/00927872.2010.515521).

[R6] A. Konovalov, S. Linton and M. Torpey. OpenDreamKit deliverable D5.15 – Final report and evaluation of all the GAP developments, 2019. PDF included, downloaded from: [https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c70b5a5c&appId=PPGMS.](https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c70b5a5c&appId=PPGMS.)

## 4. Details of the impact (indicative maximum 750 words).

By adding new capabilities for computing with semigroups, new and faster algorithms for key problems, new databases and new links to other systems, our research insights have enhanced software widely used in teaching and learning of advanced mathematics. Through this, our research has changed the practices of advanced mathematics teaching at more than 27 institutions in 23 countries, affecting courses taken by over 5,700 people (students), improving the students' learning and understanding. Specifically, more teachers are using computational algebra software to teach advanced mathematics, enabling them to make their courses more investigative and student-centred, leading to improved learning and understanding in their students. Our research has also improved the experience, learning and understanding of the approximately 20,000 weekly users of the CoCalc web-based cloud computing and course management platform for computational mathematics.

### Pathway from Research to Impact

The pathway from our research to this impact runs through an open-source software system called [GAP](GAP) (Groups, Algorithms, Programming) [S1]. GAP is a software package which allows the user to calculate interactively with complex mathematical objects, such as groups, semigroups, permutations, matrices and so on. It is highly extensible and, with appropriate extension packages, has functionality covering much of abstract algebra, combinatorics, representation theory and other key disciplines in pure mathematics. Our research insights are realised as algorithms, databases and software technologies, which are distributed as part of GAP, which is itself used for teaching.

Our development of new algorithms for computing with semigroups has expanded GAPs capabilities in semigroup theory from almost nothing, making it possible to use GAP for teaching (as well as researching) this key topic. As one measure of the scale of this expansion, since 2012 (the oldest version available, but already including some of our work) the number of documentation entries containing "semigroup" in their names increased from 700 to over 2,000 [S2]. This expanded capability provided the basis for subsequent changes in teaching practice and student learning.

Our new algorithms for permutation groups and linear algebra increased the speed of key computations across the whole system. This allows students to work flexibly and interactively with more and larger examples of the mathematical objects they are studying. This is especially important as the smallest examples in these fields (the only ones accessible to hand computation or with previous generations of software) are often atypical and can mislead students. A professor of mathematics at Colorado State University (Fort Collins) emphasised this in his letter, writing "*The use of computation makes it feasible to [study examples] in cases that are significantly larger than what could be done by hand, and allows [students] to study not only individual phenomena, but also to avoid misguided intuition, based on the availability of only small examples*" [S3, p. 5]. A professor of algebra from the University of Granada wrote, "*fast and responsive performance benefits their educational use, for example by allowing students to experiment with larger examples gaining deeper understanding*" [S3, p. 4].

Adding databases to the system builds still further on the ability of students to work with examples, allowing them to systematically study all examples and to look for patterns. It also increases the efficiency of teaching by making it trivial for the teacher to identify and supply examples for the students. A lecturer at Vrije Universiteit, Brussels, wrote "*Through the study of examples using the databases available in GAP... a natural approach to the formulation of*

*general theorems can be developed."* [S3, p. 6].

Finally, thanks to our research on linking mathematical software, GAP is now also part of the broader SageMath mathematics software system [S4] and the CoCalc on-line scientific collaboration and teaching system [S5], enabling thousands of additional students to benefit from its capabilities. As a direct output of our research [R6], GAP can now be used though the Jupyter notebook user interface, which many students and teachers are familiar with from the popular Python programming language and which is widely used in data science and scientific data analysis.

## Reach of Impact

Both GAP and SageMath are open-source software systems, which anyone is free to download and use, without informing the authors. Evidence for the reach of the impact through their direct use is thus unavoidably incomplete, based on surveys, requests for support and publications. Our most recent survey of the use of GAP, SageMath and CoCalc in teaching [S6] received 46 responses between 22 September 2020 and 01 December 2020, describing teaching in 23 countries to a total of 5,757 people (students) between August 2013 and December 2020. Of particular note are courses delivered to over 2,500 people (students) in at least seven African countries. Teaching included undergraduate (26 reports), postgraduate (26 reports) and professional development courses (10 reports). Among these are 740 people (students) taking computer supported courses in semigroup theory, something that would have been impossible without our research. Earlier data collections through the GAP website [S7] and GitHub [S8] definitively identify 7 additional sites, in 7 additional countries where GAP has been used in teaching between August 2013 and December 2020.

GAP and SageMath are key components of the CoCalc online service [S5] which provides additional tools for teaching, course management, assessment, etc. As of February 2020, CoCalc had over 20,000 unique weekly logins, growing at about 20% per year [S9].

## Significance of impact

Evidence for the significance of the specific changes flowing from our insights is found in our survey [S6]. As well as the 740 students whose semigroup theory courses would simply have been impossible without our algorithms and software, 88% of responses agreed that the ability to explore larger examples was an important or very important benefit of using the systems, and 72% said that the performance and responsiveness of the software was a significant factor in delivering the benefits of the software for their teaching. These benefits flow from our new and faster algorithms. Mathematical databases were rated "very" or "extremely" important by 50%. The significance of our work on linking software is seen in the extent of use of GAP via SageMath and CoCalc (52% of those who responded to our survey use it this way, and we know [S9] that the majority of CoCalc users use SageMath).

Evidence of widespread use and high value placed on computational tools for teaching in this area is found in the publication by independent authors, of textbooks for courses relying on the use of GAP and SageMath. We know of 8 editions of textbooks published between August 2013 and December 2020 which rely on GAP [S7] and 24 which rely on SageMath [S10]. In France, SageMath (under its previous name "Sage") is on the national list of software permitted for the "Aggregation", the qualifying examination for secondary school teachers [S11].

A pedagogical research project in 2018 at the Universitat d'Alacant and Universitat Politècnica de València studied the use of GAP in courses on Group Theory and Coding Theory. It reported a positive experience and included a recommendation that the University "*Integrate GAP classes into the subject from the beginning, thus giving time to ... see more examples and applications of the program*" [S12, p. 8, our translation].

Further evidence is found in the letters [S3] submitted as part of the follow-up to the survey:

"*I have made the experience that by giving the students the possibility to work concretely with the new algebraic objects they have seen in the courses the motivation to confront abstract algebraic problems hugely increases. .....*

*The interplay of theory with the calculations of more complex examples, for which manual computations are overly technical or simply not feasible, allows a much deeper insight into the underlying algebra.*" [S3, p. 6]

*" … the possibility of using structures as groups, rings and fields either finite or infinite allows the student to implement cryptographic protocols in platforms which the students could not construct themselves in a computer without an extraordinary effort."* [S3, p. 3] *(describing teaching a cryptography course).*

"*The combination of Sagemath with Jupyterlab is a useful tool that makes possible to combine GAP with other mathematical tools, for both the use of students (for teamwork and labs) and for the use of teachers in the class room. It offers a learning by experimentation which would be hardly achieved without these tools. ..... the inclusion of GAP in Sagemath has allowed to interact groups with other mathematical objects* " [S3, p. 2].

## 5. Sources to corroborate the impact (indicative maximum of ten references)

S1. Main GAP web page https://www.gap-system.org

S2. Search results showing numbers of documentation entries containing "semigroup" in their names in GAP manuals from 2012 (earliest available) and 2020. Evidence for the expansion of capabilities as a result of our body of work, which led to subsequent impact.

S3. Testimonials of 5 universities describing GAP's influence on teaching.

S4. PDF of web page http://doc.sagemath.org/html/en/reference/libs/index.html demonstrating GAP's incorporation into SageMath.

S5. https://cocalc.com/doc/software-executables.html demonstrating GAP's incorporation into CoCalc.

S6. Feedback from our survey: Qualtrics survey results, covering 46 submissions from teachers in 23 countries, showing breadth of use of GAP (directly and via SageMath) in teaching, and importance of various aspects of the system.

S7. https://www.gap-system.org/Doc/Teaching/teaching.html listing known resources, such as books and lecture notes demonstrating the widespread use of GAP in teaching.

S8. https://github.com/gap-system/GapWWW/issues/148 with responses to earlier request for information on use of GAP in teaching. [S6]

S9. Email from CEO of CoCalc Inc, confirming numbers of users of their system.

S10. https://github.com/carpentries-incubator/sage-lesson/wiki/SageMath-books - web page of book references showing widespread and varied use of the SageMath in teaching

S11. https://agreg.org/index.php?id=modelisation showing inclusion of SageMath (under its earlier name "Sage") for French national exam

S12. C. Sotomayor et al., paper "Recursos computacionales algebraicos en el Grado en Matemáticas", in "Redes de Investigación e Innovación en Docencia Universitaria". Volumen 2019. Alicante: Universidad de Alicante, Instituto de Ciencias de la Educación (ICE), 2019. ISBN 978-84-09-07186-9, 773 p. http://hdl.handle.net/10045/98732