A presentation for the Thompson sporadic simple group

George Havas, Leonard H. Soicher and Robert A. Wilson

Abstract. We determine a presentation for the Thompson sporadic simple group Th. The proof of correctness of this presentation uses a coset enumeration of 143,127,000 cosets. In the process of our work, we determine presentations for ${}^3D_4(2)$, ${}^3D_4(2){:}3$, $G_2(3){:}2$, and $C_{Th}(2A)$ (of shape $2_+^{1+8}{\cdot}A_9$). We also provide, via the internet, matrices generating Th and satisfying our presentation.

1991 Mathematics Subject Classification: primary 20D08, 20F05, 20-04.

1. Introduction

Presentations for many sporadic simple groups and their automorphism groups have been published. For example, see [**PS97**] for M_{11}, M_{12}, M_{22}, M_{23}, M_{24}, J_1, J_2, HS, McL, He, Suz, $O'N$, Co_1, Co_2, Co_3, Fi_{22}, Fi_{23}, Fi_{24}; see [**HM69**] for J_3, [**Wei91**] for Ru, [**BN95**] for $HN{:}2$, [**JW96,HS99**] for Ly, [**SW88,Iva92**] for J_4, and [**Iva99**] for B and M. Indeed, until now, the Thompson group Th was the only sporadic simple group S for which there was no published presentation for S or $\mathrm{Aut}\,(S)$. In this paper, we apply computational group theory in various ways to determine a presentation for Th ($\cong \mathrm{Aut}\,(Th)$). In the process, we determine presentations for ${}^3D_4(2)$, ${}^3D_4(2){:}3$, $G_2(3){:}2$, and $C_{Th}(2A)$. We also provide, via the internet, matrices generating Th and satisfying our presentation.

Throughout, we use ATLAS notation [**ATLAS**] for conjugacy classes and group structures. The proof of correctness of the presentation for Th uses a coset enumeration of 143,127,000 cosets, first completed during the Conference.

2. The Thompson group

In the monster group M, the centralizer of an element in conjugacy class $3C$ is of shape $3{\times}Th$, where Th is the Thompson (sporadic simple) group. The group

Th has order $90,745,943,887,872,000 = 2^{15}.3^{10}.5^3.7^2.13.19.31$. The maximal subgroups of Th have been completely determined (see [**Wil88,Lin89,Lin91**]), the largest being ${}^3D_4(2){:}3$ of index 143,127,000. The next largest maximal subgroup is the Dempwolff group $2^5{}^{\cdot}L_5(2)$, of index 283,599,225. The Thompson group contains just one conjugacy class of involutions, and the centralizer of an element in this class is a maximal subgroup of shape $2_+^{1+8}{}^{\cdot}A_9$. There are exactly three conjugacy classes of elements of order 3 in Th, and the normalizer $N(3A)$ of the subgroup generated by an element in the smallest such class is a maximal subgroup of shape $(3 \times G_2(3)){:}2$.

3. The presentations

Theorem 3.1. *Consider the group T presented by generators a, b, c, d, e, s, t, u and relators $(1) - (6)$ below:*

$$a^2, b^2, c^2, d^2, e^2, (ab)^3, (ae)^2, (bc)^3, (bd)^2, (be)^2, a = (cd)^4, (ce)^2, (de)^3, (bcde)^8, \tag{1}$$

$$s^7, [s, a], [s, b], [s, c], (sd)^2, [e, s] = e^{s^3}, \tag{2}$$

$$t^3, [t, a], [t, b], [t, c], [t, d], [t, e], s^t = s^2, \tag{3}$$

$$u^2 = ac, [u, a], [u, c], [u, e], (ded^u)^2, [u, (ac)^b] = e, [u^d, (ac)^b] = ue(ac)^b u^d ec, \tag{4}$$

$$t^u = t^{-1}, \tag{5}$$

$$[e, u^{s^2}], ac = (us)^3 = [u, s]^4, (du^{s^2})^4 = acc^d c^{des^{-1}} c^{des^2}. \tag{6}$$

(Here, the equation $w_1 = w_2$ denotes the relator $w_1 w_2^{-1}$. The notation $w_1^{w_2}$ means $w_2^{-1} w_1 w_2$, and $[w_1, w_2]$ means $w_1^{-1} w_2^{-1} w_1 w_2$.)

Then the following hold:

1. *$T \cong Th$.*
2. *Generators a, b, c, d, e, s together with relators $(1), (2)$ give a presentation for ${}^3D_4(2)$.*
3. *Generators a, b, c, d, e, s, t together with relators $(1), (2), (3)$ give a presentation for ${}^3D_4(2){:}3$.*
4. *Generators a, b, c, d, e, u together with relators $(1), (4)$ give a presentation for $G_2(3){:}2$.*

5. *Generators* a, c, d, e, s, t, u *together with the relators from* $(1)-(6)$ *not involving* b *give a presentation for* $C_{Th}(2A)$.

Proof. Using the 248-dimensional $GF(2)$-matrix representation of Th available from the ATLAS of Finite Group Representations [**Wil98,W+99**], we have constructed non-identity 248-dimensional $GF(2)$-matrices $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{s}, \mathbf{t}, \mathbf{u} \in Th$ satisfying all the relators $(1)-(6)$ above (in the obvious way). The interested reader can find these matrices in GAP-format (see [**GAP4**]) at

`http://www.mat.bham.ac.uk/atlas/gap/Th/`

and verify this. Later, we will see that these matrices generate Th. The determination of these matrices is described in Section 4.

Now the abstract generators a, b, c, d, e together with the relators (1) are easily seen to give a presentation for the group $U_3(3){:}2$ (this is essentially the same presentation as that for $U_3(3){:}2$ which appears in [**Soi92**]). Indeed, the reader is encouraged to (at least) check that the group so presented has order 12,096. A coset enumeration shows that $\langle a, b, c, d, e\rangle$ has index 17,472 in the group D presented by generators a, b, c, d, e, s and relators $(1), (2)$. Thus D, as well as its matrix group image

$$\mathbf{D} := \langle \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{s}\rangle,$$

has order at most that of ${}^3D_4(2)$. Moreover, the subgroup $\mathbf{D}$ of Th contains $\langle \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}\rangle \cong U_3(3){:}2$ (as opposed to a proper quotient of this group, in which case the relators would imply $\mathbf{a} = 1$, a contradiction), and also $\mathbf{D}$ contains $7^2 \cong \langle \mathbf{bc}\mathbf{c}^{\mathbf{d}}, \mathbf{s}\rangle$. From the maximal subgroup structure of Th we conclude that the only possibility for $\mathbf{D}$ is ${}^3D_4(2)$. Since the order of D is bounded by that of ${}^3D_4(2)$ as well as having $\mathbf{D}$ as a homomorphic image, we conclude that $D \cong {}^3D_4(2)$, establishing Statement 2 of the theorem.

Given that $\mathbf{t}$ has order 3 and the relators (3), we conclude that t is an element of order 3 in T, normalizing the subgroup $\langle a, b, c, d, e, s\rangle \cong {}^3D_4(2)$ and centralizing its maximal subgroup $\langle a, b, c, d, e\rangle \cong U_3(3){:}2$. Since no inner element of ${}^3D_4(2)$ of order 3 has these properties, it follows that

$$\mathbf{H} := \langle \mathbf{D}, \mathbf{t}\rangle \cong {}^3D_4(2){:}3,$$

and Statement 3 follows.

Now $\mathbf{H}$ is a maximal subgroup of Th, and $\mathbf{u} \notin \mathbf{H}$ since, by relator (5), $\mathbf{u}$ must invert $\mathbf{t}$. Hence

$$Th = \langle \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{s}, \mathbf{t}, \mathbf{u}\rangle,$$

and we have that Th is a homomorphic image of the group T presented. Now the main Statement 1 of the theorem follows from the enumeration of 143,127,000 cosets of $H := \langle a, b, c, d, e, s, t\rangle$ in T. This enumeration is described in Section 5.

We now tackle Statement 4. The matrix $\mathbf{u}$ is in the normalizer of $\mathbf{t}$, with $\mathbf{u}$ inverting $\mathbf{t}$. Hence

$$\mathbf{G} := \langle \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{u} \rangle \leq N_{Th}(\mathbf{t}) \cong (3 \times G_2(3)){:}2,$$

and it is easy to see that $\mathbf{G} \cong G_2(3){:}2$ (with the full normalizer of $\langle \mathbf{t} \rangle$ being $\langle \mathbf{G}, \mathbf{t} \rangle$). A coset enumeration shows that $\langle a, b, c, d, e \rangle$ has index 702 in the group presented by generators a, b, c, d, e, u and relators $(1), (4)$, establishing Statement 4.

Finally, we prove Statement 5. It is easily deduced from relators $(1) - (3)$ that a commutes with c, d, e, s, and t. Then a coset enumeration shows that $\langle a, c, d, e, s, t \rangle$ has index 819 in $H = \langle a, b, c, d, e, s, t \rangle$, establishing that $\langle a, c, d, e, s, t \rangle = C_H(a) \cong 2_+^{1+8}{:}L_2(8){:}3 \cong \langle \mathbf{a}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{s}, \mathbf{t} \rangle$. Since $[\mathbf{u}, \mathbf{a}] = 1$ and $\mathbf{u} \notin \mathbf{H}$, we have that

$$\mathbf{C} := \langle \mathbf{a}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{s}, \mathbf{t}, \mathbf{u} \rangle = C_{Th}(\mathbf{a}).$$

Now let C be the group presented by generators a, c, d, e, s, t, u together with the relators from $(1) - (6)$ not involving b. It is easy to show, using coset enumeration, that in C, $L := \langle d, e, s, t \rangle \cong L_2(8){:}3$, and then a coset enumeration of the 61,440 cosets of L in C completes the proof of 5. $\square$

We now outline (roughly) how our presentation for Th was derived. We started by determining the presentation above for $H = \langle a, b, c, d, e, s, t \rangle \cong {}^3D_4(2){:}3$, starting with the known presentation for $U := \langle a, b, c, d, e \rangle \cong U_3(3){:}2$, and using the subgroup structure of ${}^3D_4(2){:}3$ given in [**ATLAS**] to determine the further generators and relators. In H, t is an outer element of order 3 centralized by U, and in Th, we can extend U to $G_2(3){:}2$ by an element u inverting t. Up to automorphisms of $G_2(3){:}2$ there is just one way to identify the elements a, b, c, d, e. Calculating in a permutation representation of $G_2(3){:}2$, we identified these elements, chose an outer element u centralizing a and satisfying other short relators, and then determined a presentation for $G_2(3){:}2$ on generators a, b, c, d, e, u. For this, the GAP4 function `PresentationViaCosetTable` was very helpful in determining the crucial relation $[u^d, (ac)^b] = ue(ac)^b u^d ec$, to make correct the subgroup $\langle a, c, (ac)^b, c^d, e, u, u^d \rangle \cong 2^3{\cdot}L_3(2).2$ in $G_2(3){:}2$. Defining relators for $C_{Th}(a) \cong 2_+^{1+8}{\cdot}A_9$ on generators a, c, d, e, s, t, u (for an appropriate choice of u) were more difficult to obtain, although starting with $\langle d, e, s, t \rangle \cong L_2(8){:}3$, we determined defining relators modulo the normal 2_+^{1+8} (which is the normal closure of $\langle c \rangle$ in $\langle a, c, d, e, s, t \rangle$). To exactly determine defining relators for $C_{Th}(a)$, we constructed the matrices generating Th and satisfying the relators $(1) - (5)$ above, and then deduced the exact form of the relators (6).

Some of our relators for $T \cong Th$ are known to be redundant, but are included for clarity or to aid coset enumeration. We also remark that, in T, $\langle a, c, (ac)^b, c^d, e, s, t, u, u^d \rangle$ is isomorphic to the Dempwolff group $2^5{\cdot}L_5(2)$.

4. Determining matrices in Th satisfying the presentation

Inside the Thompson group, a copy of ${}^3D_4(2){:}3$ can easily be found in a random search of subgroups generated by an involution and a $3C$-element. To find $3 \times U_3(3){:}2$, the centralizer of a $3C$-element in ${}^3D_4(2){:}3$, we first found two elements, of orders 21 and 12, powering up to two different $3C$-elements, and then in the group generated by these two 3-elements, found an element conjugating one to the other. (This could be done either by a random search, or by pure thought, since the group generated by these elements turned out to be $2^3{:}7{:}3$.) We thus obtain elements of orders 21 and 12 powering up to the same $3C$-element, and generating the whole of its centralizer, $3 \times U_3(3){:}2$. This $3C$-element may be taken as $\mathbf{t}^{\pm 1}$.

We then look for $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\mathbf{e}$ satisfying the given presentation for $U_3(3){:}2$. We first found an outer involution, which we can take to be $\mathbf{e}$, and used standard methods (see [**Lin95,Bra**]) to find its centralizer $2 \times S_4$. There are two subgroups isomorphic to S_4 in here, one of which is inside $U_3(3)$. We take $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ to be the transpositions (12), (23), (34) of the latter. It remains to find $\mathbf{d}$, which is unique up to conjugation by $\mathbf{e}$. We do this by searching in the centralizer of another involution, such as $\mathbf{a}$ (or $\mathbf{b}$). Since this centralizer is small, this does not take long.

The next step is to find $\mathbf{s}$ inside ${}^3D_4(2){:}3$. Now $\mathbf{s}$ is an element of order 7 commuting with $\langle \mathbf{a}, \mathbf{b}, \mathbf{c}\rangle \cong S_4$, so we work inside the normalizer of the fours-group $\langle \mathbf{ac}, (\mathbf{ac})^{\mathbf{b}}\rangle$, which is a group of shape $2^2.[2^9].(7 \times S_3)$. We take any element $\mathbf{s}_0$ of order 7 in this group, and then $\mathbf{s}_1 = \mathbf{s}_0[\mathbf{bc}, \mathbf{s}_0]$ commutes with $\mathbf{bc}$, so $\mathbf{s}$ is a suitable power of $\mathbf{s}_1$. Indeed, we may take $\mathbf{s} = \mathbf{s}_1$ without loss of generality. At this point we can use the relation $\mathbf{s}^{\mathbf{t}} = \mathbf{s}^2$ to determine which central element of $3 \times U_3(3){:}2$ is $\mathbf{t}$, and which is $\mathbf{t}^{-1}$.

Finally we must find $\mathbf{u}$, which we do inside the full centralizer $2_+^{1+8}{\cdot}A_9$ of $\mathbf{a}$. This centralizer is generated by $\mathbf{c}$, $\mathbf{d}$, $\mathbf{e}$, $\mathbf{s}$, $\mathbf{t}$ and one other element, found by standard methods. We chop the representation to find a constituent which represents A_9, and convert everything to the natural permutation representation where we can calculate easily. We are looking for an element which commutes with $\mathbf{e}$, and inverts $\mathbf{t}$. In the quotient A_9, there are just three involutions with these properties, but only one satisfies the relation $(\mathbf{us})^3 = 1$. If $\mathbf{u}_0$ is any lift of this element in $2_+^{1+8}{\cdot}A_9$, then $\mathbf{u}_1 = \mathbf{u}_0[\mathbf{t}, \mathbf{u}_0]$ inverts $\mathbf{t}$.

We now have $N(\mathbf{t}) \cong (3 \times G_2(3)){:}2$ generated by $\mathbf{t}$, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\mathbf{e}$ and $\mathbf{u}_1$, but $\mathbf{u}_1$ still does not satisfy all the required relations, because we have not yet specified precisely which preimage of the element of A_9 to take. Inside $G_2(3){:}2$ generated by $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\mathbf{e}$ and $\mathbf{u}_1$ we find that the normalizer of the fours-group $\langle \mathbf{a}, \mathbf{c}\rangle$ is generated by $\mathbf{bacb}$, $\mathbf{dcd}$, $\mathbf{e}$ and $\mathbf{u}_1$, and we search in this small group for the required element $\mathbf{u}$.

5. The large coset enumeration

The enumeration to determine $|T : H|$ requires the definition of more than 140 million cosets. Thus this step requires substantial resources. Recent descriptions of relevant aspects of coset enumeration are given in [**Sim94,HR**]. Of particular importance here is that the space required (memory locations) for a standard coset enumeration to complete is at least the maximum number of cosets defined times the number of columns in the coset table. The number of columns in the coset table is the number of involutory group generators plus twice the number of noninvolutory generators.

In some circumstances (eg, see [**HS99**]) this means we need to consider finding presentations with small generating sets. In our case here the given presentation for T leads to eleven columns, which is not too many in the context of memory resources available on some supercomputers. Our enumerations were performed on an SGI Origin 2000 computer which has 16 gigabytes of memory, and we were able to obtain access to some 8 gigabytes. This is enough to store over 180 million cosets with the given presentation.

For difficult coset enumeration problems there is sometimes the need for substantial experimentation to determine how to complete an enumeration with given resources (see [**HR**]). However our experience with the various subgroups of T discussed in this paper led us to believe that the enumeration in this case would be relatively easy, and so it turned out.

Using the ACE3 coset enumerator [**ACE**], with group presentation and subgroup generators as in this paper, we can readily compute the 143,127,000 cosets of H in T. It is worth noting that the final coset table (in standard internal representation) is too big to fit in a computer with 32 bit addressing architecture.

The ACE3 enumerator has a large number of parameters, with a wide choice of settings and enumeration styles (see [**ACE**]). We have completed the enumeration using various strategies without undue difficulty. Felsch-based enumerations seem best. They complete in as little as about 62 cpu hours. For example, with the styles `Hard`, `Felsch:0`, `Felsch:1` and `Sims:3` in each case the maximum number of cosets defined is the index. The total numbers are:

`Hard`	`Felsch:1`	`Felsch:0`	`Sims:3`
145,251,396	151,273,730	155,354,391	240,046,124

The `HLT` style (also available in [**ACE**]) seems less attractive, using a maximum of 179,631,904 cosets and a total of 473,745,756 in an enumeration given 8 gigabytes of memory.

Acknowledgements

The work of the first author was supported by the Australian Research Council. We are grateful to The University of Queensland's High Performance Computing Unit for access to the Origin 2000 supercomputer.

References

[Bra] J.N. Bray, An improved algorithm for finding the centralizer of an involution, Arch. Math., to appear.

[BN95] T. Breuer and S. Norton, Improvements to the ATLAS, Appendix 2 of: An ATLAS of Brauer Characters, (by C. Jansen, K. Lux, R. Parker and R. Wilson), Clarendon Press, Oxford, 1995, 297–327.

[ATLAS] J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker and R.A. Wilson, An ATLAS of Finite Groups, Clarendon Press, Oxford, 1985.

[GAP4] The GAP Group, Lehrstuhl D für Mathematik, RWTH Aachen, Germany and School of Mathematical and Computational Sciences, U. St. Andrews, Scotland, *GAP – Groups, Algorithms, and Programming, Version 4*, 1999. GAP is available at `http://www-gap.dcs.st-and.ac.uk/~gap/`.

[ACE] G. Havas and C. Ramsay, Coset enumeration: ACE version 3, 1999. Available as `http://www.csee.uq.edu.au/~havas/ace3.tar.gz`.

[HR] G. Havas and C. Ramsay, Experiments in coset enumeration, These proceedings.

[HS99] G. Havas and C.C. Sims, A presentation for the Lyons simple group, in: Computational Methods for Representations of Groups and Algebras, Progress in Mathematics **173**, Birkhäuser, Basel, 1999, 241–249.

[HM69] G. Higman and J. McKay, On Janko's simple group of order 50,232,960, Bull. London Math. Soc. **1** (1969), 89–94.

[Iva92] A.A. Ivanov, A presentation for J_4, Proc. London Math. Soc. **64** (1992), 369–396.

[Iva99] A.A. Ivanov, Geometry of Sporadic Groups I: Petersen and Tilde Geometries, Cambridge University Press, 1999.

[JW96] C. Jansen and R.A. Wilson, The minimal faithful 3-modular representation for the Lyons group, Comm. Algebra **24** (1996), 873–879.

[Lin89] S.A. Linton, The maximal subgroups of the Thompson group, J. London Math. Soc. (2) **39** (1989), 79–88.

[Lin91] S.A. Linton, Corrections to "The maximal subgroups of the Thompson group", J. London Math. Soc. (2) **43** (1991), 253–254.

[Lin95] S.A. Linton, The art and science of computing in large groups, in: Computational Algebra and Number Theory (W. Bosma et al., eds), Kluwer, Dordrecht, 1995, 91–109.

[PS97] C.E. Praeger and L.H. Soicher, Low Rank Representations and Graphs for Sporadic Groups, Cambridge University Press, 1997.

[Sim94] C.C. Sims, Computation with Finitely Presented Groups, Cambridge University Press, 1994.

[Soi92] L.H. Soicher, On simplicial complexes related to the Suzuki sequence graphs, in: Groups, Combinatorics and Geometry (M.W. Liebeck and J. Saxl, eds), LMS Lecture Note Series **165**, Cambridge University Press, 1992, 240–248.

[SW88] G. Stroth and R. Weiss, Modified Steinberg relations for the group J_4, Geom. Dedicata **25** (1988), 513–525.

[Wei91] R. Weiss, A geometric characterization of the groups M_{12}, He and Ru, J. Math. Soc. Japan **43** (1991), 795–814.

[Wil88] R.A. Wilson, Some subgroups of the Thompson group, J. Austral. Math. Soc. (A) **44** (1988), 17–32.

[Wil98] R.A. Wilson, An ATLAS of sporadic group representations, in: The ATLAS 10 Years On (R.T. Curtis and R.A. Wilson, eds), LMS Lecture Note Series **249**, Cambridge University Press, 1998, 261–273.

[W+99] R.A. Wilson et al., An ATLAS of Finite Group Representations, `http://www.mat.bham.ac.uk/atlas/`, 1999.

George Havas
Centre for Discrete Mathematics and Computing
Department of Computer Science and Electrical Engineering
The University of Queensland
Queensland 4072, Australia
havas@csee.uq.edu.au

Leonard H. Soicher
School of Mathematical Sciences
Queen Mary and Westfield College
Mile End Road, London E1 4NS, UK
L.H.Soicher@qmw.ac.uk

Robert A. Wilson
School of Mathematics and Statistics
The University of Birmingham
Edgbaston, Birmingham B15 2TT, UK
R.A.Wilson@bham.ac.uk