A database of genus 2 curves over the rational numbers

Andy Booker (Bristol)
Jeroen Sijsling (Ulm)
Drew Sutherland (MIT)
John Voight (Dartmouth)
Dan Yasaki (UNCG)

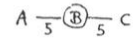
ANTS-XII
Technische Universität Kaiserslautern
30 August 2016

The goal

To generalize John Cremona's elliptic curve tables to genus 2...

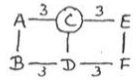
11 = 11

A	0	-1	1	0	0 - 1	I1	1	5 0
B	0	-1	1	-10	-20 - 5	I5	5	5 0
C	0	-1	1	-7820	-263580 - 1	I1	1	1 0



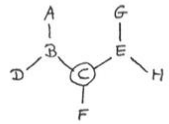
14 = 2.7

A	1	0	1	-1	0 - 2, 1	I2, I1	2, 1	6 0
B	1	0	1	-11	12 + 1, 2	I1, I2	1, 2	6 0
C	1	0	1	4	-6 - 6, 3	I6, I3	6, 3	6 0
D	1	0	1	-36	-70 + 3, 6	I3, I6	3, 6	6 0
E	1	0	1	-171	-874 - 18, 1	I18, I1	18, 1	2 0
F	1	0	1	-2731	-55146 + 9, 2	I9, I2	9, 2	2 0



15 = 3.5

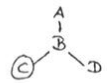
A	1	1	1	0	0 - 1, 1	I1, I1	1, 1	4 0
B	1	1	1	-5	2 + 2, 2	I2, I2	2, 2	8 0
C	1	1	1	-10	-10 + 4, 4	I4, I4	4, 4	8 0
D	1	1	1	-80	242 + 1, 1	I1, I1	1, 1	4 0
E	1	1	1	-135	-660 + 8, 2	I8, I2	8, 2	8 0
F	1	1	1	35	-28 - 2, 8	I2, I8	2, 8	4 0
G	1	1	1	-110	-880 - 16, 1	I16, I1	16, 1	2 0
H	1	1	1	-2160	-39540 + 4, 1	I4, I1	4, 1	2 0



88

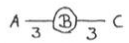
17 = 17

A	1	-1	1	-1	0 + 1	I1	1	4 0
B	1	-1	1	-6	-4 + 2	I2	2	4 0
C	1	-1	1	-1	-14 - 4	I4	4	4 0
D	1	-1	1	-91	-310 + 1	I1	1	2 0



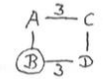
19 = 19

A	0	1	1	1	0 - 1	I1	1	3 0
B	0	1	1	-9	-15 - 3	I3	3	3 0
C	0	1	1	-769	-8470 - 1	I1	1	1 0



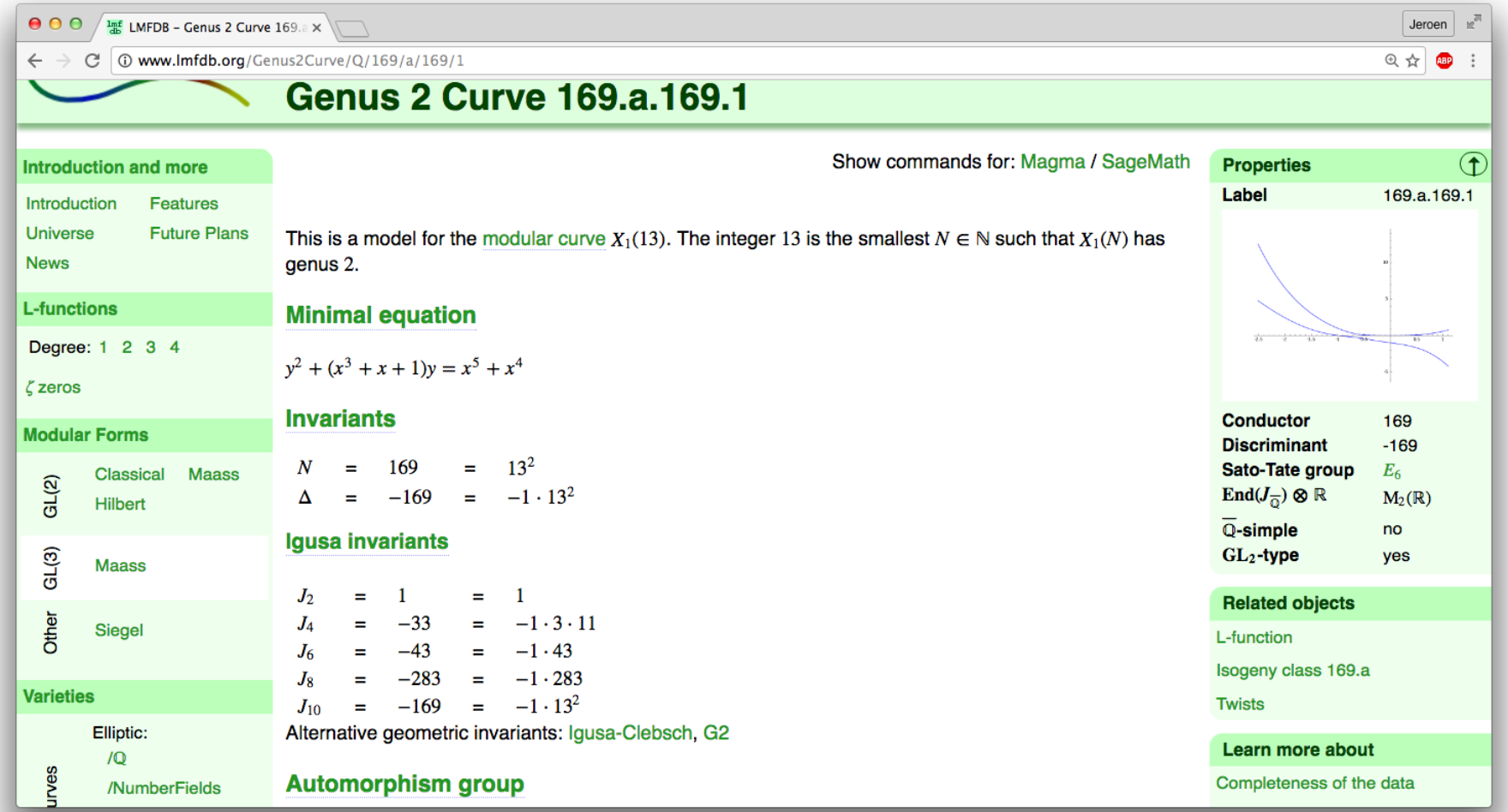
20 = 2.2.5

A	0	1	0	-1	0 + 4, 1	IV, I1	0, 1	6 0
B	0	1	0	4	4 - 8, 2	IV*, I2	0, 2	6 0
C	0	1	0	-41	-116 + 4, 3	IV, I3	0, 3	2 0
D	0	1	0	-36	-140 - 8, 6	IV*, I6	0, 6	2 0



The goal

in a modern format.



What the database contains

We tabulate smooth genus 2 curves X over $\mathbb{Q}$ by arithmetic complexity.

What the database contains

We tabulate smooth genus 2 curves X over $\mathbb{Q}$ by arithmetic complexity.

Such a curve is given by an equation

$$X : y^2 + hy = f$$

with

$$h = h_3x^3 + h_2x^2 + h_1x + h_0$$

of degree ≤ 3 and

$$f = f_6x^6 + f_5x^5 + f_4x^4 + f_3x^3 + f_2x^2 + f_1x + f_0$$

of degree ≤ 6.

We can and will take $h, f \in \mathbb{Z}[x]$.
(As we should, since we care about arithmetic aspects.)

Measuring arithmetic complexity

Given $X : y^2 + hy = f$, we can form the associated sextic

$$g = 4f + h^2 \in \mathbb{Z}[x].$$

We measure the complexity of this equation of the curve by the absolute value of its discriminant

$$\Delta(h, f) = 2^{-12} \operatorname{disc}(g) \in \mathbb{Z} \setminus \{0\}.$$

Measuring arithmetic complexity

Given $X : y^2 + hy = f$, we can form the associated sextic

$$g = 4f + h^2 \in \mathbb{Z}[x].$$

We measure the complexity of this equation of the curve by the absolute value of its discriminant

$$\Delta(h, f) = 2^{-12} \operatorname{disc}(g) \in \mathbb{Z} \setminus \{0\}.$$

Curves over $\mathbb{Q}$ admit a minimal Weierstrass model, that is, a model with $h, f \in \mathbb{Z}[x]$ for which $|\Delta(h, f)|$ is minimal. (The `Magma` algorithm that implements this is due to Stoll.)

We define the absolute discriminant $\Delta_{\text{abs}}(X) > 0$ of X accordingly; it is our measure for the geometric complexity of the curve X.

What is new

So far:

- Stoll (2013) found 823 isomorphism classes of curves of odd absolute discriminant up to 11^4;
- Merriman and Smart (1993) found all 427 isomorphism class of curves with good reduction away from 2 (some of huge absolute discriminant);
- Gonzalez *et al.* (2000s): databases of modular curves (those admitting a morphism $X_1(N) \to X$ or those whose Jacobians are abelian subvarieties of $J_1(N)$) and Shimura curves;
- Brumer *et al.* (2014), Farmer *et al.* (2015): identification of possible conductors via conjectural paramodular methods, resp. restrictions on L-functions.

What is new

So far:

- Stoll (2013) found 823 isomorphism classes of curves of odd absolute discriminant up to 11^4;
- Merriman and Smart (1993) found all 427 isomorphism class of curves with good reduction away from 2 (some of huge absolute discriminant);
- Gonzalez *et al.* (2000s): databases of modular curves (those admitting a morphism $X_1(N) \to X$ or those whose Jacobians are abelian subvarieties of $J_1(N)$) and Shimura curves;
- Brumer *et al.* (2014), Farmer *et al.* (2015): identification of possible conductors via conjectural paramodular methods, resp. restrictions on L-functions.

We searched for curves of absolute discriminant up to 10^6 and found 66158 isomorphism classes, including all those found by the methods above.

How to find curves

We look for curves in boxes of four shapes:

$$S_1(B) = \{(f,h) : |f_i| \le B,\ h_i = 0,1\} \quad (\text{"flat"})$$
$$S_2(a,b) = \{(f,h) : |f_i| \le ab^{6-i},\ h_i = 0,1\} \quad (\text{"weighted"})$$
$$S_3(b) = \{(f,h) : |f_i| \le b^{4-|i-3|},\ h_i = 0,1\} \quad (\text{"crested"})$$
$$S_4(b,d) = \{(f,h) : \textstyle\sum_i \lceil \log_b(|f_i|+1) \rceil \le d,\ h_i = 0,1\} \quad (\text{"weird"})$$

In the end we considered

$$S_1(90) \cup S_2(2,3.51) \cup S_3(7.14) \cup S_4(10,10).$$

This leads to over 3×10^{17} curves.

Computing the discriminant

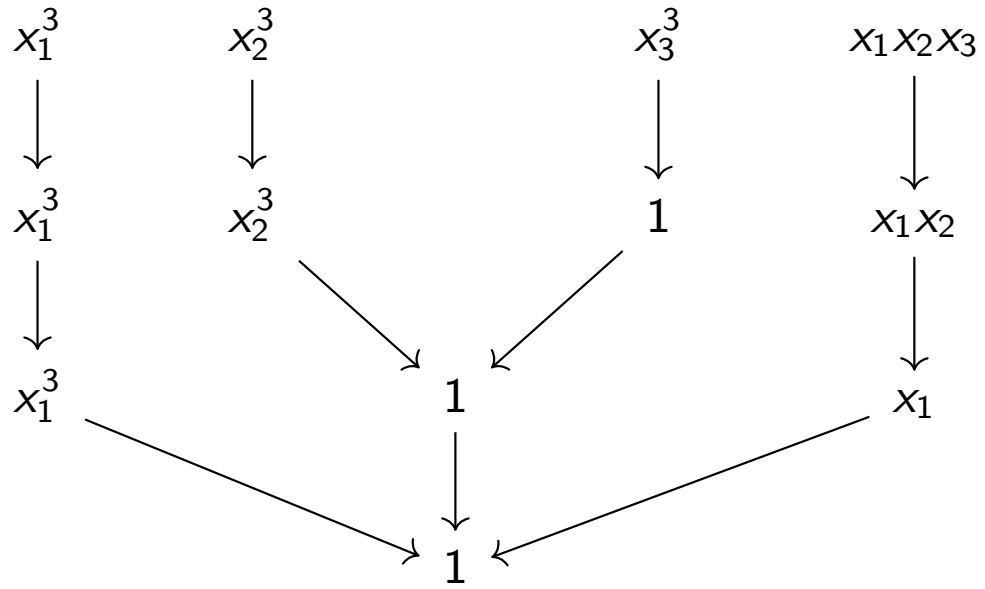
The reason for using boxes is that the discriminant can be calculated quickly by using a monomial tree. More generally, suppose that we want to evaluate a polynomial $p(x_1, \ldots, x_n)$ on a box $A_1 \times \cdots \times A_n \subset \mathbb{Z}^n$. Then this tree is as follows:

- Nodes at level n (leaves): monomials in x_i of $p(x_1, \ldots, x_n)$.
- Nodes at level $n-1$: monomials in x_i of $p(x_1, \ldots, x_{n-1}, a_n)$.
- ...
- Nodes at level 1: monomials in x_i of $p(x_1, a_2, \ldots, a_n)$.
- Node at level 0: root!

Nodes at level $m+1$ are connected to those at level m via an edge corresponding to the substitution $x_{m+1} = a_{m+1}$.
Total number of nodes of tree for the discriminant: 703.

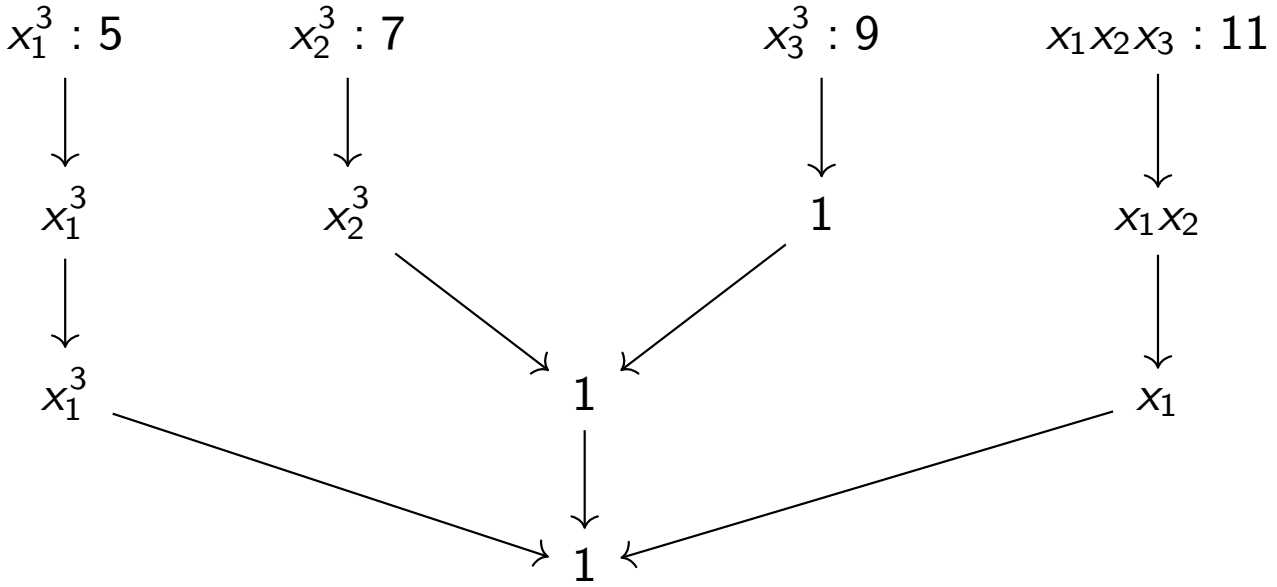
Computing the discriminant

Monomial tree for $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



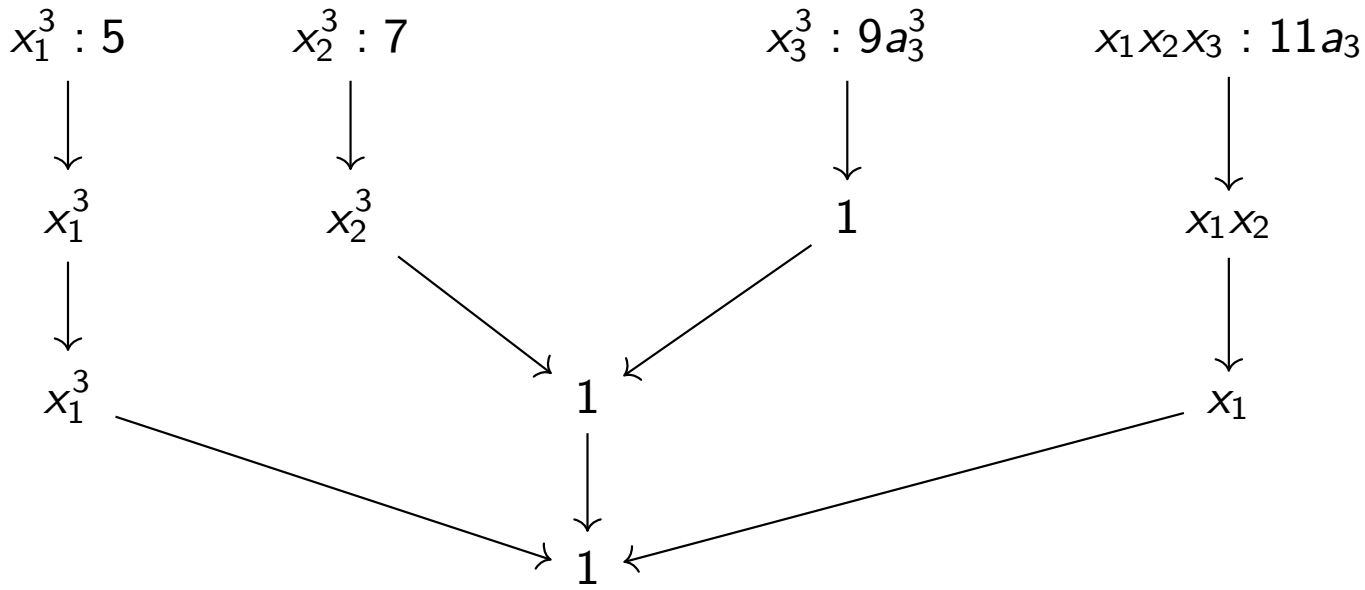
Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



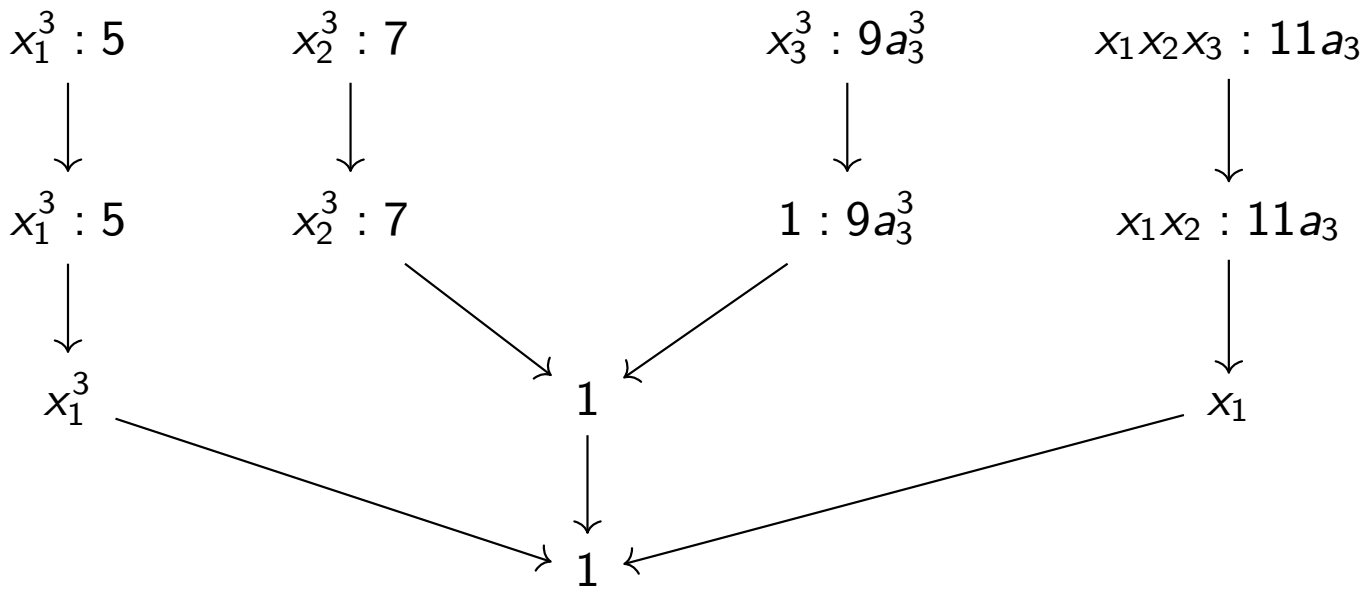
Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



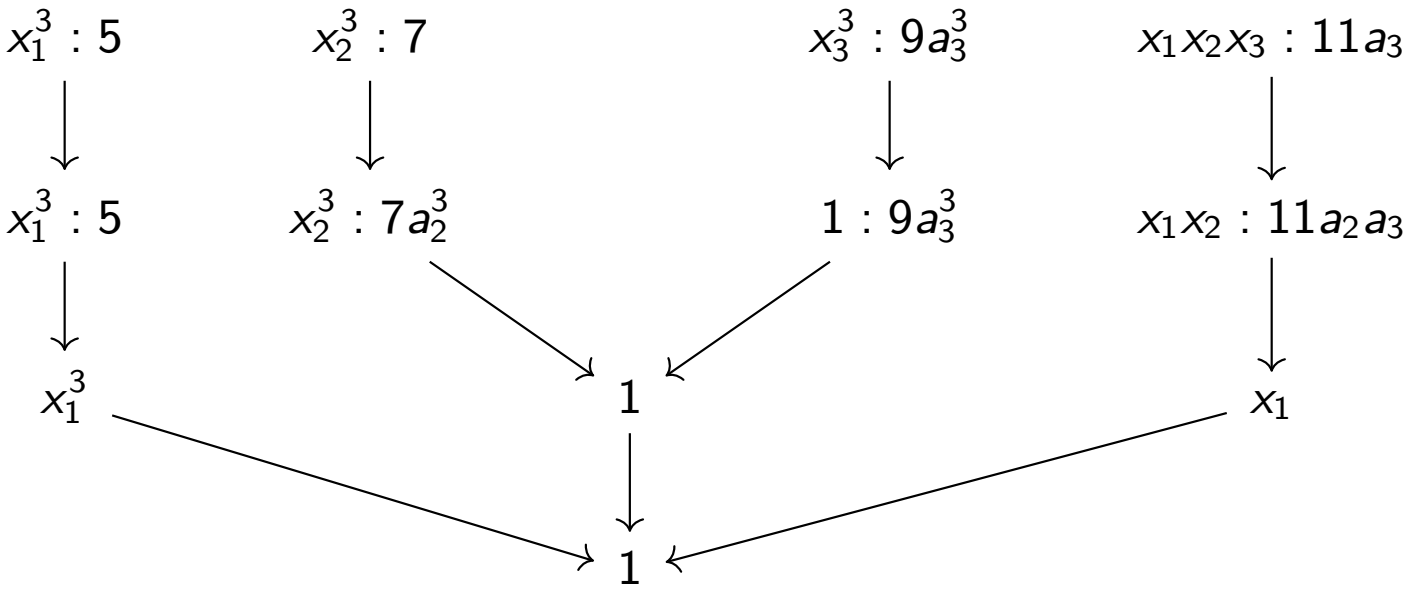
Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



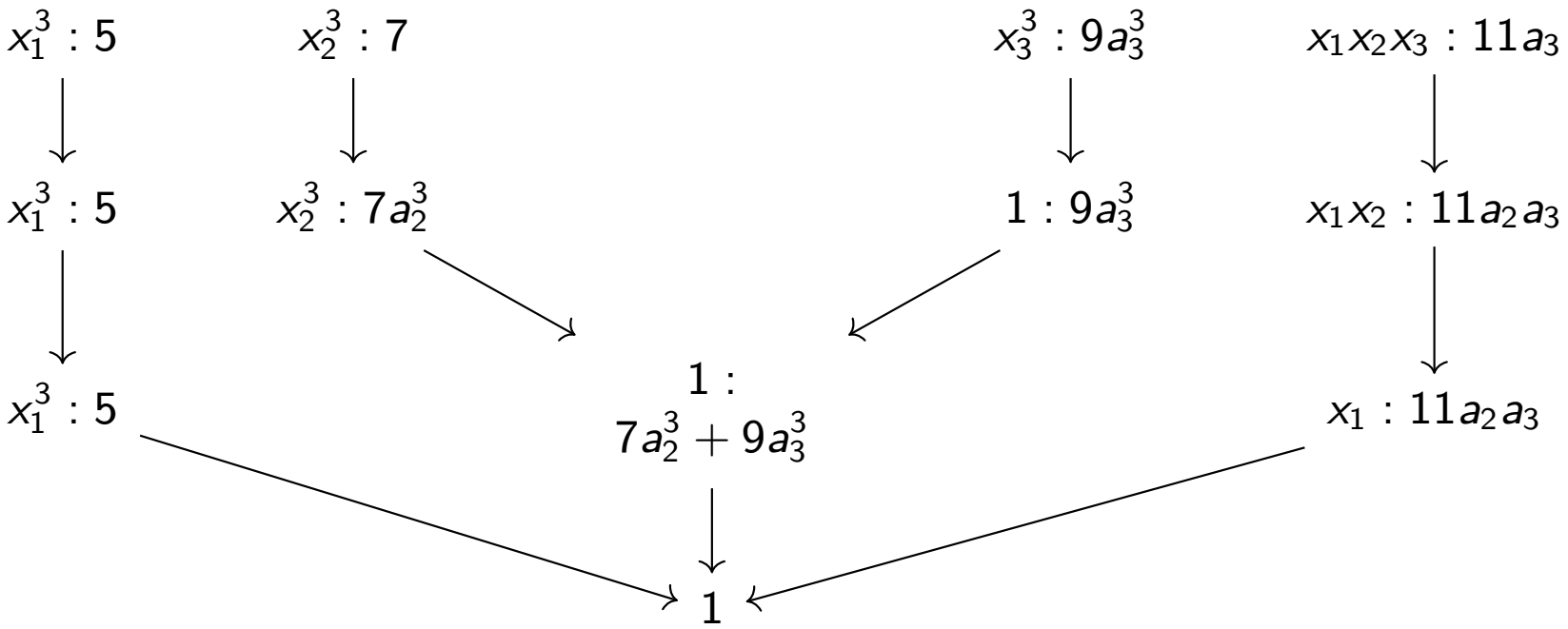
Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



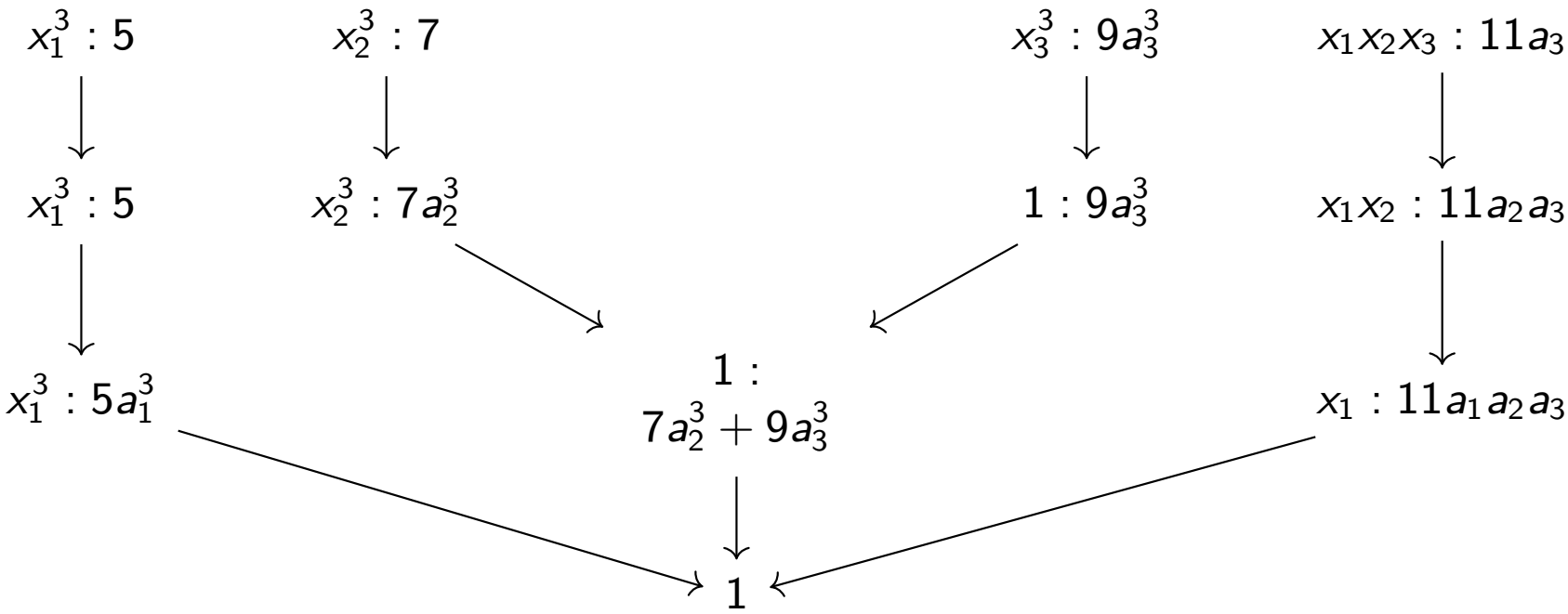
Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



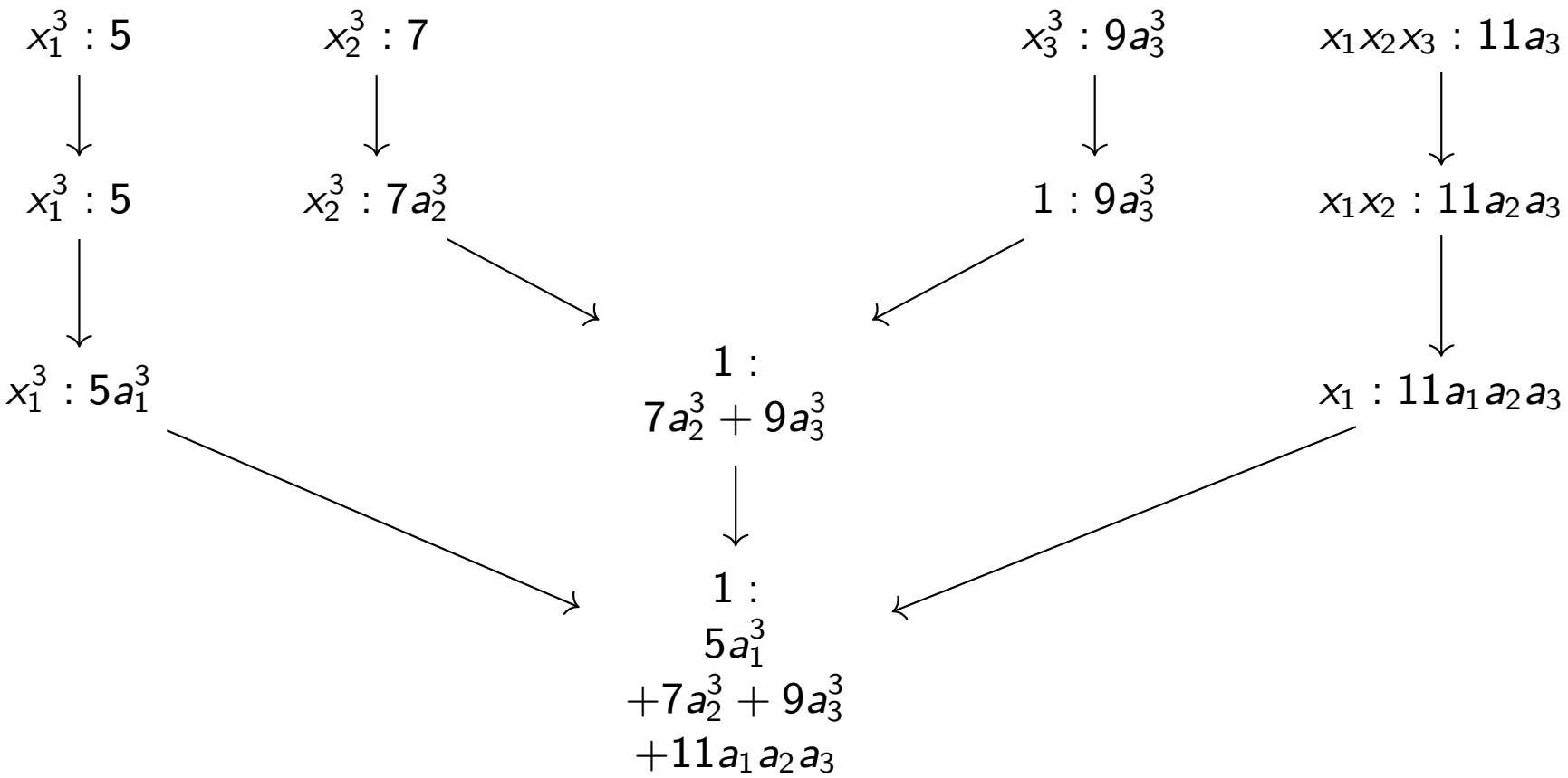
Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



Computing the discriminant

Evaluation of $5x_1^3 + 7x_2^3 + 9x_3^3 + 11x_1x_2x_3$:



Computing the discriminant

In words: we run over the values $(a_1, \ldots, a_n)$ in $A_1 \times \cdots \times A_n$ lexicographically, so incrementing the final entries last, and inductively pass on the values via the tree.

This approach is dominated by the evaluation of the univariate monomials in x_1, which can be done by a finite difference method, as developed by Kedlaya-Sutherland (2008). In total this approach has a main cost of 5 additions per discriminant, rather than > 1000 arithmetic operations by the naive method.

Parallel computation

The computation was parallellized by cutting up into subboxes and then run on Google's Cloud Platform Compute Engine:

- 2250 high-CPU compute nodes in 6 regions (5 in the US, 1 in Europe);
- 32 Intel Haswell cores and 28.8 GB of memory per node;
- Wall time: 8 hours.
- Total CPU time: 2250 CPUs $\times$ 32 cores/CPU $\times$ 8 hours $\approx$ 66 core-years.

Conductor

Let $J = \mathrm{Jac}(X)$ be the Jacobian of X. An effective algebraic approach to computing the algebraic conductor N_{alg} of J is under development by Bouw–Wewers. Currently, one has to use Qing Liu's `genus2reduction` package from 1998 which only determines the valuation of N_{alg} at odd primes.

Conductor

Let $J = \mathrm{Jac}(X)$ be the Jacobian of X. An effective algebraic approach to computing the algebraic conductor N_{alg} of J is under development by Bouw–Wewers. Currently, one has to use Qing Liu's `genus2reduction` package from 1998 which only determines the valuation of N_{alg} at odd primes.

Conjecturally, N_{alg} equals the analytic conductor N_{an} of J, which appears in the functional equation of the L-function X. An argument due to Booker (2005) and independently Dokchitser (2004) shows that there is at most one conductor compatible with this equality, and gives an explicit method to obtain it.

Conductor

Let $J = \mathrm{Jac}(X)$ be the Jacobian of X. An effective algebraic approach to computing the algebraic conductor N_{alg} of J is under development by Bouw–Wewers. Currently, one has to use Qing Liu's `genus2reduction` package from 1998 which only determines the valuation of N_{alg} at odd primes.

Conjecturally, N_{alg} equals the analytic conductor N_{an} of J, which appears in the functional equation of the L-function X. An argument due to Booker (2005) and independently Dokchitser (2004) shows that there is at most one conductor compatible with this equality, and gives an explicit method to obtain it.

So suppose that $N = N_{\mathrm{alg}} = N_{\mathrm{an}}$. We define the completed L-function $\Lambda(J, s) = \Gamma_{\mathbb{C}}(s)^2 \sum_{n=1}^{\infty} a_n n^{-s}$, where $a_p = p + 1 - \#X(\mathbb{F}_p)$ for a prime p.

Conductor

Conjecturally, $\Lambda(J,s)$ continues to an entire function and satisfies the functional equation

$$\Lambda(J,s) = wN^{1-s}\Lambda(J,2-s)$$

for some $w \in \{\pm 1\}$. Define

$$S(y) = \frac{1}{y}\sum_{n=1}^{\infty} a_n K_0(4\pi\sqrt{n/y}) \quad \text{for } y > 0.$$

Then one has the integral representation

$$\Lambda(J,s) = 8\int_0^{\infty} S(y)y^{-s}\,dy \quad \text{for } \Re(s) > \frac{3}{2},$$

which can be used to show that the functional equation of $\Lambda(J,s)$ amounts to the identity

$$S(y) = wS(N/y) \quad \text{for all } y > 0.$$

Conductor

In particular, if $a \in \{0, 1\}$ with $w = (-1)^{a+1}$, then

$$S^{(a)}(\sqrt{N}) = 0.$$

Given a purported set of data $\{a_n\}$, N and w, this gives a falsifiable identity that we use to test consistency with the Hasse–Weil conjecture, by using estimates on truncations of the series representation of S.

If some numbers are not known (e.g. the 2-part of N and the bad Euler factors), we simply try all possibilities for $L_2(J, T)$; remember that only one can work!

Sato-Tate group

Another important invariant of X is its Sato–Tate group $\mathrm{ST}(X)$; conjecturally, the distribution of the normalized L-polynomials of X is determined by the Haar measure of $\mathrm{ST}(X)$. In genus 2, this distribution in turn determines $\mathrm{ST}(X)$.

The Sato–Tate group can be approximated by computing L-polynomials at good primes p up to some bound. Using the average polynomial-time implementation of Harvey and Sutherland (2014), one can heuristically infer the Sato–Tate distribution, and from it, $\mathrm{ST}(X)$ itself. Knowing $\mathrm{ST}(X)$ is equivalent to knowing the Galois module structure of $\mathrm{End}(J_{\overline{\mathbb{Q}}}) \otimes \mathbb{R}$ by FKRS (2012).

Endomorphisms

It is also possible to calculate the Galois module $\mathrm{End}(J_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q}$, and even the corresponding endomorphism rings, as follows:

- Choose a symplectic basis $\gamma_1, \ldots, \gamma_4$ of $H_1(X, \mathbb{Z})$ and a basis ω_1, ω_2 of $H^0(X, \omega_X)$ over $\mathbb{Q}$;

Endomorphisms

It is also possible to calculate the Galois module $\mathrm{End}(J_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q}$, and even the corresponding endomorphism rings, as follows:

- Choose a symplectic basis $\gamma_1, \ldots, \gamma_4$ of $H_1(X, \mathbb{Z})$ and a basis ω_1, ω_2 of $H^0(X, \omega_X)$ over $\mathbb{Q}$;
- Realize $J(\mathbb{C})$ as a complex torus $\mathbb{C}^2/\Lambda$ by computing the period matrix $P = (\int_{\gamma_j} \omega_i)_{i,j}$;

Endomorphisms

It is also possible to calculate the Galois module $\mathrm{End}(J_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q}$, and even the corresponding endomorphism rings, as follows:

- Choose a symplectic basis $\gamma_1, \ldots, \gamma_4$ of $H_1(X, \mathbb{Z})$ and a basis ω_1, ω_2 of $H^0(X, \omega_X)$ over $\mathbb{Q}$;
- Realize $J(\mathbb{C})$ as a complex torus $\mathbb{C}^2/\Lambda$ by computing the period matrix $P = (\int_{\gamma_j} \omega_i)_{i,j}$;
- Use LLL to determine a basis of the $\mathbb{Z}$-module of matrices $R \in M_4(\mathbb{Z})$ such that $\Lambda R = R$;

Endomorphisms

It is also possible to calculate the Galois module $\mathrm{End}(J_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q}$, and even the corresponding endomorphism rings, as follows:

- Choose a symplectic basis $\gamma_1, \ldots, \gamma_4$ of $H_1(X, \mathbb{Z})$ and a basis ω_1, ω_2 of $H^0(X, \omega_X)$ over $\mathbb{Q}$;
- Realize $J(\mathbb{C})$ as a complex torus $\mathbb{C}^2/\Lambda$ by computing the period matrix $P = (\int_{\gamma_j} \omega_i)_{i,j}$;
- Use LLL to determine a basis of the $\mathbb{Z}$-module of matrices $R \in M_4(\mathbb{Z})$ such that $\Lambda R = R$;
- Determine the matrices $M \in M_2(\overline{\mathbb{Q}})$ in the equality $M\Pi = \Pi R$ to obtain the representation of $\mathrm{End}(J_{\overline{\mathbb{Q}}})$ on the tangent space at 0 of $J_{\overline{\mathbb{Q}}}$.

Endomorphisms

This gives $\mathrm{End}(J_{\overline{\mathbb{Q}}})$ as a ring. The matrices M are all defined over the field of definition of the corresponding endomorphism, because we used a $\mathbb{Q}$-basis of $H^0(X, \omega_X)$. We can therefore recognize them algebraically by once more using LLL.

For a subfield K of the field of definition L of $\mathrm{End}(J_{\overline{\mathbb{Q}}})$, we can compute $\mathrm{End}(J_K)$ by finding the integral solutions of

$$\sum_{i=1}^{d} n_i (M_i^\sigma - M_i)$$

for $\sigma \in \mathrm{Gal}(L|K)$. Thus we find the Galois module structure of $\mathrm{End}(J_{\overline{\mathbb{Q}}})$.

Endomorphisms

This gives $\mathrm{End}(J_{\overline{\mathbb{Q}}})$ as a ring. The matrices M are all defined over the field of definition of the corresponding endomorphism, because we used a $\mathbb{Q}$-basis of $H^0(X, \omega_X)$. We can therefore recognize them algebraically by once more using LLL.

For a subfield K of the field of definition L of $\mathrm{End}(J_{\overline{\mathbb{Q}}})$, we can compute $\mathrm{End}(J_K)$ by finding the integral solutions of

$$\sum_{i=1}^{d} n_i (M_i^\sigma - M_i)$$

for $\sigma \in \mathrm{Gal}(L|K)$. Thus we find the Galois module structure of $\mathrm{End}(J_{\overline{\mathbb{Q}}})$.

The result thus obtained always agreed with the heuristic determination of the Sato–Tate group. We can also determine elliptic curve factors of $J_{\overline{\mathbb{Q}}}$ in case it splits up to isogeny.

An example

http://www.lmfdb.org/Genus2Curve/Q/15360/f/983040/2