

# The $L$ -functions and Modular Forms DataBase (LMFDB)

John Voight  
Dartmouth College

Simons Collaboration on Arithmetic Geometry,  
Number Theory and Computation  
Annual Meeting  
Simons Foundation, NYC  
10 January 2019

# Outline of this talk

# Outline of this talk

1. (Re)Introduce the LMFDB;
2. New classical modular forms functionality  
(private screening/world premiere!);
3. Future plans.

# The LMFDB, in a nutshell

## The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

# The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

modular forms,  $L$ -functions, number fields, zeta zeros, elliptic curves, Maass forms, Artin representations, genus 2 curves, ...

# The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

modular forms,  $L$ -functions, number fields, zeta zeros, elliptic curves, Maass forms, Artin representations, genus 2 curves, ...

The LMFDB was first conceived at an AIM workshop in 2007.

## The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

modular forms,  $L$ -functions, number fields, zeta zeros, elliptic curves, Maass forms, Artin representations, genus 2 curves, ...

The LMFDB was first conceived at an AIM workshop in 2007. The acknowledgements list 100 contributors.

# The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

modular forms,  $L$ -functions, number fields, zeta zeros, elliptic curves, Maass forms, Artin representations, genus 2 curves, ...

The LMFDB was first conceived at an AIM workshop in 2007. The acknowledgements list 100 contributors. It has received substantial funding by the NSF, EPSRC, and the Simons Foundation, in addition to other smaller grants.

# The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

modular forms,  $L$ -functions, number fields, zeta zeros, elliptic curves, Maass forms, Artin representations, genus 2 curves, ...

The LMFDB was first conceived at an AIM workshop in 2007. The acknowledgements list 100 contributors. It has received substantial funding by the NSF, EPSRC, and the Simons Foundation, in addition to other smaller grants. (Thank you!)

# The LMFDB, in a nutshell

The LMFDB is a catalogue of mathematical objects arising in connection with the Langlands program, including

modular forms,  $L$ -functions, number fields, zeta zeros, elliptic curves, Maass forms, Artin representations, genus 2 curves, ...

The LMFDB was first conceived at an AIM workshop in 2007. The acknowledgements list 100 contributors. It has received substantial funding by the NSF, EPSRC, and the Simons Foundation, in addition to other smaller grants. (Thank you!)

It is now managed like a journal, with managing editors (John Cremona, John Jones, Drew Sutherland, and me) and a team of associate editors.

## Motivation and desiderata

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes.

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online.

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)
- ▶ Data sets should be connected to one another.

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)
- ▶ Data sets should be connected to one another. (Sometimes the point is to make connections across data sets.)

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)
- ▶ Data sets should be connected to one another. (Sometimes the point is to make connections across data sets.)
- ▶ Data should be reliable and have a glossary/appropriate annotations.

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)
- ▶ Data sets should be connected to one another. (Sometimes the point is to make connections across data sets.)
- ▶ Data should be reliable and have a glossary/appropriate annotations.
- ▶ Data should be searchable, so you can find particular examples of interest.

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)
- ▶ Data sets should be connected to one another. (Sometimes the point is to make connections across data sets.)
- ▶ Data should be reliable and have a glossary/appropriate annotations.
- ▶ Data should be searchable, so you can find particular examples of interest. (Before you prove something, it's helpful to know if there are counterexamples!)

## Motivation and desiderata

- ▶ Number theory has historically been in part an experimental science: even Gauss used tables of primes. Big databases are now ubiquitous in science.
- ▶ By exhaustive computation, we can catch some issues (“sweep in all corners”) and bugs in code (“compute all examples”).
- ▶ Databases should be easily accessible online. (Not easy if data is scattered on homepages: learn format, install program, ... Worse still if a paper says “E-mail me if you want the data”.)
- ▶ Data sets should be connected to one another. (Sometimes the point is to make connections across data sets.)
- ▶ Data should be reliable and have a glossary/appropriate annotations.
- ▶ Data should be searchable, so you can find particular examples of interest. (Before you prove something, it's helpful to know if there are counterexamples!)
- ▶ Data should have statistics, to notice patterns in aggregate.

# Technical specs

## Technical specs

- ▶ Size: The LMFDB includes a database of zeros of  $\zeta(s)$  (1 TB), class groups of imaginary quadratic fields (2 TB), with the rest of the data totalling 600+ GB. [More detailed stats](#) are available.

## Technical specs

- ▶ Size: The LMFDB includes a database of zeros of  $\zeta(s)$  (1 TB), class groups of imaginary quadratic fields (2 TB), with the rest of the data totalling 600+ GB. [More detailed stats](#) are available.
- ▶ The LMFDB is currently hosted in the Google cloud rather than on a university-hosted server.

## Technical specs

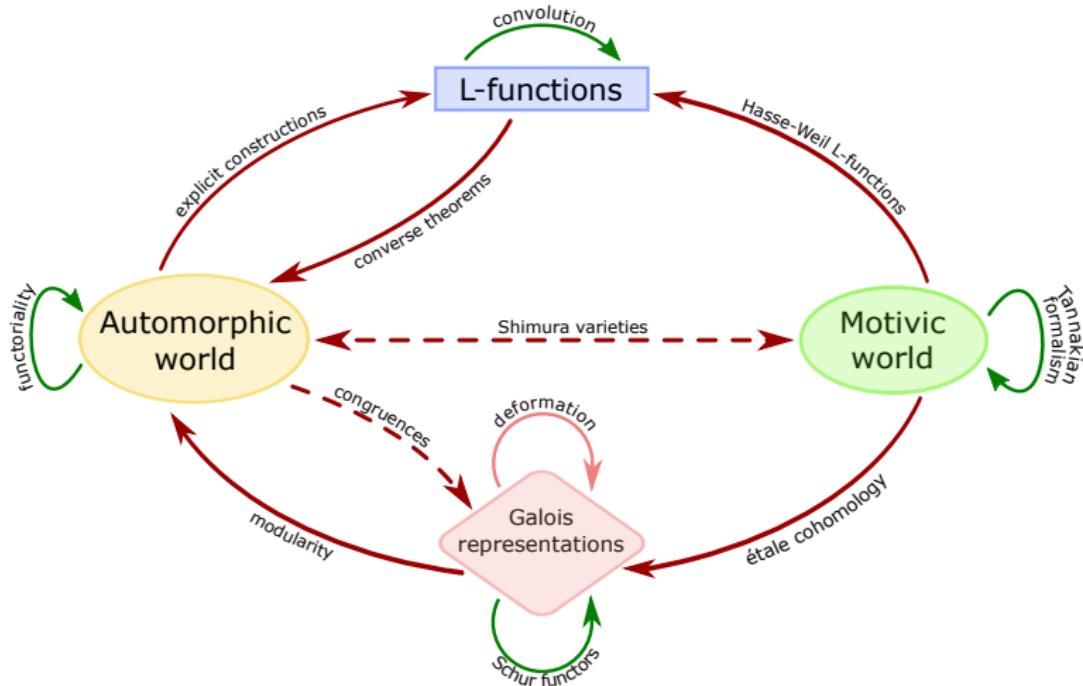
- ▶ Size: The LMFDB includes a database of zeros of  $\zeta(s)$  (1 TB), class groups of imaginary quadratic fields (2 TB), with the rest of the data totalling 600+ GB. [More detailed stats](#) are available.
- ▶ The LMFDB is currently hosted in the Google cloud rather than on a university-hosted server.
- ▶ The underlying database will soon be PostgreSQL, an object-relational database management system.

## Technical specs

- ▶ Size: The LMFDB includes a database of zeros of  $\zeta(s)$  (1 TB), class groups of imaginary quadratic fields (2 TB), with the rest of the data totalling 600+ GB. [More detailed stats](#) are available.
- ▶ The LMFDB is currently hosted in the Google cloud rather than on a university-hosted server.
- ▶ The underlying database will soon be PostgreSQL, an object-relational database management system. (We moved from MongoDB, which is a document database; talk to Edgar Costa and David Roe for more details!)

Big picture

# Big picture



LMFDB Universe

# Classical modular forms

## Classical modular forms

At a week-long workshop in August at MIT funded by the Simons Foundation, a significant upgrade was made to the MFs in the LMFDB.

## Classical modular forms

At a week-long workshop in August at MIT funded by the Simons Foundation, a significant upgrade was made to the MFs in the LMFDB. Edgar Costa and David Roe (Simons postdocs at MIT, led by Drew Sutherland) have continued the charge through this year.

## Classical modular forms

At a week-long workshop in August at MIT funded by the Simons Foundation, a significant upgrade was made to the MFs in the LMFDB. Edgar Costa and David Roe (Simons postdocs at MIT, led by Drew Sutherland) have continued the charge through this year. Thanks to Alex Best, Jonathan Bober, Andy Booker, John Cremona, and David Lowry-Duda.

# Classical modular forms

At a week-long workshop in August at MIT funded by the Simons Foundation, a significant upgrade was made to the MFs in the LMFDB. Edgar Costa and David Roe (Simons postdocs at MIT, led by Drew Sutherland) have continued the charge through this year. Thanks to Alex Best, Jonathan Bober, Andy Booker, John Cremona, and David Lowry-Duda.

Objects:

- ① space  $S_k(N, \chi)$  identifier  $N, k, \chi$  or  $N, k, n$  Dirichlet characters
- dimension  $\ell$  old siblings  $\ell$  from database
- first minimal  $\ell$  from  $S_k(N, \chi)$ ,  $n \leq 100$  (limited range)
- sturm bound

② Hecke orbits  $N, k, \chi, \ell$   $\ell$  is over  $\mathbb{Q}$   $\ell$  is prime  $\ell$  is CM  $\ell$  is discriminant (Hecke character) Dirichlet characters

- dimension  $\ell$  field poly, label, polredabs?  $\ell$  is prime  $\ell$  is CM
- label  $N, k, \chi, \ell$   $\ell$  is prime  $\ell$  is CM
- field poly, label, polredabs?
- CM discriminant (Hecke character)

③ complex  $a_n$ 's?  $a_n$   $\ell$  label, orbit label,  $\ell$  some well-known others? TBD  $\ell$  precision, digits enough data to reconstruct alg.  $a_n$  up to Sturm bound

- embedding, label, orbit label,  $\ell$  some well-known others? TBD

④ primitive deg 2 L-function  $N, k, \chi$   $\ell$   $\ell$  from

⑤ rational L-function (degree 2!)  $N, k, \chi$   $\ell$  from

Magma  $\hookrightarrow$  Drew  
Pari  $\hookrightarrow$  JC  
Analytic  $\hookrightarrow$  JV  
(LMFDB/Mango)

① lookup polredabs polys in LMFDB } Edgar  
② polredabs Field polys

Database schema  $\hookrightarrow$  David, Alex

Web page design/layout

Rigor/completeness/source knobs

Dirichlet characters  $\hookrightarrow$  LMFDB Nisoprop

Parity, Dolgachev, congr. labels/orbit labels

Browse/search  $\hookrightarrow$  David, JV weight 1

L-function data (Andy, JB) complex data?

Algebraic coeffs  $\hookrightarrow$  JC, JV

History: Wada 1971

# History: Wada 1971

269 ?

$$\begin{aligned} 277 \quad & (x^{11}-21x^9+152x^7-13x^6-445x^5+102x^4+450x^3-147x^2-56x+4)^2 \\ 281 \quad & x^{22}+50x^{20}+1081x^{18}+13285x^{16}+102611x^{14}+519821x^{12}+1748638x^{10} \\ & +3865180x^8+5420534x^6+4487135x^4+1884424x^2+265883 \end{aligned}$$

293 ?

$$\begin{aligned} 313 \quad & (x^{12}-21x^{10}-5x^9+158x^8+63x^7-509x^6-241x^5+648x^4+272x^3-250x^2 \\ & -81x-4)^2 \end{aligned}$$

317 ?

$$\begin{aligned} 337 \quad & (x^{13}-x^{12}-22x^{11}+16x^{10}+182x^9-91x^8-697x^7+221x^6+1217x^5-215x^4 \\ & -808x^3+61x^2+175x)^2 \end{aligned}$$

349 ?

$$\begin{aligned} 353 \quad & x^{28}+57x^{26}+1429x^{24}+20814x^{22}+196024x^{20}+1256811x^{18} \\ & +5621908x^{16}+17671372x^{14}+38764478x^{12}+58122566x^{10} \\ & +57447036x^8+35475661x^6+12722068x^4+2334034x^2+160344 \end{aligned}$$

373 ?

389 ?

397 ?

$$\begin{aligned} 401 \quad & x^{32}+71x^{30}+2276x^{28}+43639x^{26}+558857x^{24}+5055254x^{22} \\ & +33307004x^{20}+162429517x^{18}+589806266x^{16}+1590754165x^{14} \\ & +3152518983x^{12}+4498023798x^{10}+4468102725x^8+2926806000x^6 \\ & +1151600000x^4+225600000x^2+12800000 \end{aligned}$$

$$\begin{aligned} 409 \quad & (x^{16}+x^{15}-27x^{14}-27x^{13}+284x^{12}+272x^{11}-1487x^{10}-1296x^9+4094x^8 \\ & +2998x^7-5757x^6-3006x^5+3774x^4+907x^3-964x^2+32)^2 \end{aligned}$$

421 ?

History: Antwerp IV 1975

# History: Antwerp IV 1975

## SOURCES AND RELIABILITY OF THE TABLES.

### 1. Tables 2-6.

Table 2 has been prepared for this volume by N.M. Stephens and James Davenport using the ATLAS at the Chilton laboratory. No errors are likely, but it is possible that in one or two cases a multiple of a generator has been listed rather than a generator.

Table 3 was prepared by J. Vélu; except for the eigenvalues of the  $W_q$ , the calculations have been performed independently by Stephens and Vélu, and no discrepancy has been observed; there are also strong requirements of consistency. The  $W_q$ -eigenvalues were calculated separately by D.J.

Tingley, by operating on the one-dimensional homology; his results were incorporated in Vélu's tables by reading in cards. Consistency requirements make any errors in these particular calculations of Tingley hardly conceivable; copying errors are possible but unlikely. The identifying letters of the curves in this table have been added in manuscript; I hope they have never been jumbled.

Table 4 has been extracted from the unpublished 1966 Manchester thesis of F.B. Coghlan. Unless some mathematical error has remained undetected, the table should be complete.

The third column of Table 5 comes from a much larger table supplied by A.O.L. Atkin; the final column was supplied by D.J. Tingley, who has completely computed the action of the Hecke algebra on the 1-dimensional homology for  $N \leq 300$ . His complete tables of eigenvalues will be published elsewhere. Table 6, too, is copied from a table of Atkin's; this table is new, note that the much shorter table of Deuring contains errors. Tables 5 and 6 were produced by typewriter rather than line-printer; it was very difficult to avoid slips.

11 = 11					
A	0	-1	1	0	0 - 1
B	0	-1	1	-10	-20 - 5
C	0	-1	1	-7820	-263560 - 1

14 = 2.7					
A	1	0	1	-1	0 - 2, 1
B	1	0	1	-11	12 + 1, 2
C	1	0	1	4	-6 - 6, 3
D	1	0	1	-36	-70 + 3, 6
E	1	0	1	-171	-874 - 18, 1
F	1	0	1	-2731	-55146 + 9, 2

15 = 3.5					
A	1	1	1	0	0 - 1, 1
B	1	1	1	-5	2 + 2, 2
C	1	1	1	-10	-10 + 4, 4
D	1	1	1	-80	242 + 1, 1
E	1	1	1	-135	-660 + 8, 2
F	1	1	1	35	-28 - 2, 8
G	1	1	1	-110	-880 + 16, 1
H	1	1	1	-2160	-39540 + 4, 1

17 = 17					
A	1	-1	1	-1	0 + 1
B	1	-1	1	-6	-4 + 2
C	1	-1	1	-1	-14 - 4
D	1	-1	1	-91	-310 + 1

19 = 19					
A	0	1	1	1	0 - 1
B	0	1	1	-9	-15 - 3
C	0	1	1	-769	-8470 - 1

20 = 2.2.5					
A	C	1	0	-1	0 + 4, 1
B	0	1	0	4	-8, 2
C	0	1	0	-41	-116 + 4, 3
D	0	1	0	-36	-140 - 8, 6

Stein 2000s

# Stein 2000s

https://wstein.org/tables/tables.html

The Modular Forms Database: Tables

William Stein

My Tables

**WARNING:** The dynamic query-based tables no longer work, since I have not had time to maintain this database in years. Please see also the newer LMFDB.

Modular Abelian Varieties  $A_f$  that are Jacobians of Curves

- Basic Data

Hilbert Modular Forms

Elliptic Curves

- Neum Elkies: Tables of Elliptic curves of unit discriminant over real quadratic number fields

Tables Over The Field  $\text{Q(sqrt(5))}$

- Dimensions of spaces of Hilbert modular forms, by level
- Factored characteristic polynomials of Hecke operators
- Rational Hilbert modular forms (these correspond to isogeny classes of elliptic curves over  $\text{Q(sqrt(5))}$ )

Dimensions of Spaces of Cusforms

- Dimensions of  $S_k(\Gamma_0(N))$ ,  $S_k(\Gamma_1(N))$ , and  $S_k(\Gamma_2(N))$  (eps).
- PARI-readable dimension tables for  $\Gamma_0(N)$  and  $\Gamma_1(N)$
- MAGMA-readable dimension tables for  $\Gamma_0(N)$  with character

Characteristic Polynomials

- Characteristic polynomials of  $T_p$  on  $S_k(\text{SL}_2(\mathbb{Z}))$
- Characteristic polynomials of  $T_p$  on  $\Gamma_0(N)$  and  $\Gamma_1(N)$
- Discriminants of Hecke algebras
- Arithmetic data about every weight 2 newform on  $\Gamma_0(N)$  for all  $N \leq 135$  (contains charpoly's of many eigenvalues).

Eigenvalues and q-expansions of Cusforms

- q-expansions of newforms
- Basis of eigenforms
- q-expansions of eigenforms on  $\Gamma_0(N)$  of weight k <= 14
- q-expansions of eigenforms on  $\Gamma_1(N)$  of weight k <= 50
- Eigenvalues of modular forms on  $\Gamma_0(N)$  of weight k <= 100 and high level, and of weight k <= 100 and low level
- Eigenforms on  $\Gamma_0(N)$  in terms of the free abelian group on supersingular j-invariants

Elliptic Curves

- PARI Tables of Cremona Elliptic Curves

Stein tables

# LMFDB: Scope of data

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ ,

# LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

Data computed using Magma, Pari, and code of Bober–Booker.

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

Data computed using Magma, Pari, and code of Bober–Booker.

- ▶ Data for 316 236 newspaces  $S_k(\Gamma_0(N), \chi)$  (50 098 nonzero)

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

Data computed using Magma, Pari, and code of Bober–Booker.

- ▶ Data for 316 236 newspaces  $S_k(\Gamma_0(N), \chi)$  (50 098 nonzero)
- ▶ 236 555 (Galois orbits of) newforms (48 324 with rational coefficients, 19 306 weight 1)

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

Data computed using Magma, Pari, and code of Bober–Booker.

- ▶ Data for 316 236 newspaces  $S_k(\Gamma_0(N), \chi)$  (50 098 nonzero)
- ▶ 236 555 (Galois orbits of) newforms (48 324 with rational coefficients, 19 306 weight 1)
- ▶ 13 710 564 complex embedded newforms

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

Data computed using Magma, Pari, and code of Bober–Booker.

- ▶ Data for 316 236 newspaces  $S_k(\Gamma_0(N), \chi)$  (50 098 nonzero)
- ▶ 236 555 (Galois orbits of) newforms (48 324 with rational coefficients, 19 306 weight 1)
- ▶ 13 710 564 complex embedded newforms
- ▶ Total size (including  $L$ -functions) about 300 GB

## LMFDB: Scope of data

To be systematic, our data includes two big “boxes”:

- ▶ All spaces  $S_k(\Gamma_1(N))$  with  $N, k \in \mathbb{Z}_{\geq 1}$  and  $Nk^2 \leq 4000$ , and
- ▶ All spaces  $S_k(\Gamma_0(N))$  with  $N, k \in \mathbb{Z}_{\geq 2}$  and  $Nk^2 \leq 40000$  (trivial character).

With additional boxes to cover the spread, we easily majorize Stein’s database (and current LMFDB offerings).

Data computed using Magma, Pari, and code of Bober–Booker.

- ▶ Data for 316 236 newspaces  $S_k(\Gamma_0(N), \chi)$  (50 098 nonzero)
- ▶ 236 555 (Galois orbits of) newforms (48 324 with rational coefficients, 19 306 weight 1)
- ▶ 13 710 564 complex embedded newforms
- ▶ Total size (including  $L$ -functions) about 300 GB

[Search page](#) and [11.2.a.a](#) (to show homepages)

## A sampling of new features (items 1–5)

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts).

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts). In particular, our data is complete in these boxes.  
[Search page](#)

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts). In particular, our data is complete in these boxes.  
[Search page](#)
2. Coefficients  $a_n$  are represented compactly: either sparse cyclotomic ([1620.1.bp.a](#)) or

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts). In particular, our data is complete in these boxes.  
[Search page](#)
2. Coefficients  $a_n$  are represented compactly: either sparse cyclotomic ([1620.1.bp.a](#)) or in terms of an LLL-reduced basis for the Hecke ring ([6877.2.a.s](#)).

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts). In particular, our data is complete in these boxes.  
[Search page](#)
2. Coefficients  $a_n$  are represented compactly: either sparse cyclotomic ([1620.1.bp.a](#)) or in terms of an LLL-reduced basis for the Hecke ring ([6877.2.a.s](#)).
3. We have complex eigenvalues when exact eigenvalues would be computationally infeasible ([983.2.c.a](#)), allowing for the computation of all  $L$ -functions, and can customize the display the embeddings ([2412.1.b.b](#)).

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts). In particular, our data is complete in these boxes.  
[Search page](#)
2. Coefficients  $a_n$  are represented compactly: either sparse cyclotomic ([1620.1.bp.a](#)) or in terms of an LLL-reduced basis for the Hecke ring ([6877.2.a.s](#)).
3. We have complex eigenvalues when exact eigenvalues would be computationally infeasible ([983.2.c.a](#)), allowing for the computation of all  $L$ -functions, and can customize the display the embeddings ([2412.1.b.b](#)).
4. Self-twists (all rigorously proved), inner twists (indicated when proved or not), Atkin-Lehner eigenvalues, Hecke cutters available ([441.4.a.p](#)).

## A sampling of new features (items 1–5)

1. We bound our boxes by  $Nk^2$ , which is a good approximation to the analytic conductor  $N||k||^2$  which governs the total complexity of the  $L$ -function (for more, talk to David Roberts). In particular, our data is complete in these boxes.  
[Search page](#)
2. Coefficients  $a_n$  are represented compactly: either sparse cyclotomic ([1620.1.bp.a](#)) or in terms of an LLL-reduced basis for the Hecke ring ([6877.2.a.s](#)).
3. We have complex eigenvalues when exact eigenvalues would be computationally infeasible ([983.2.c.a](#)), allowing for the computation of all  $L$ -functions, and can customize the display the embeddings ([2412.1.b.b](#)).
4. Self-twists (all rigorously proved), inner twists (indicated when proved or not), Atkin-Lehner eigenvalues, Hecke cutters available ([441.4.a.p](#)).
5. Can download the form to get more coefficients, reduce, ...

## Item 6: Trace tables (matching classical modular forms)

## Item 6: Trace tables (matching classical modular forms)

Dear John, please forgive a cold call.

With colleagues, I have been calculating local zeta functions for Calabi-Yau varieties in a family depending on  $t$ . In general this gives rise to a factor

$$1 + aT + bpT^2 + ap^3T^3 + p^6T^4.$$

For  $t = 33 \pm 8\sqrt{17}$ , the quartic factors in the form

$$(1 - \alpha pT + p^3T^2)(1 - \beta T + p^3T^2)$$

whenever 17 is a square mod  $p$ .

We have lists of data of the form  $p, t, \{\alpha, \beta\}$  for when this happens, and we believe that the coefficients  $\{\alpha, \beta\}$  are coefficients of a (Hilbert) modular form.

So finally I come to my question, which is: is there a practical way of extracting lists of coefficients of such modular forms from LMFDB with which to compare our lists?

We have a number of cases similar to the example above so we hope for a simple procedure that we can repeat as necessary.

## Narrowing the search

## Narrowing the search

- ▶ There is bad reduction at 17, and possibly at 2.

## Narrowing the search

- ▶ There is bad reduction at 17, and possibly at 2.
- ▶ The quartic factor is independent of the choice of  $\sqrt{17}$  modulo  $p$ , so we guess that there is descent from  $F = \mathbb{Q}(\sqrt{17})$  to  $\mathbb{Q}$ .

## Narrowing the search

- ▶ There is bad reduction at 17, and possibly at 2.
- ▶ The quartic factor is independent of the choice of  $\sqrt{17}$  modulo  $p$ , so we guess that there is descent from  $F = \mathbb{Q}(\sqrt{17})$  to  $\mathbb{Q}$ .
- ▶ We guess that the Hodge structure splits as

$$(1, 1, 1, 1) = (0, 1, 1, 0) + (1, 0, 0, 1)$$

## Narrowing the search

- ▶ There is bad reduction at 17, and possibly at 2.
- ▶ The quartic factor is independent of the choice of  $\sqrt{17}$  modulo  $p$ , so we guess that there is descent from  $F = \mathbb{Q}(\sqrt{17})$  to  $\mathbb{Q}$ .
- ▶ We guess that the Hodge structure splits as

$$(1, 1, 1, 1) = (0, 1, 1, 0) + (1, 0, 0, 1)$$

and accordingly the (likely)  $\ell$ -adic Galois representation

$$\rho: \text{Gal}_F \rightarrow \text{GL}_4(\mathbb{Q}_\ell)$$

is the base change from  $\mathbb{Q}$  to  $F$  of

$$\rho_f(-1) \oplus \rho_g$$

where  $f, g$  are classical modular forms of weights 2, 4.

## Narrowing the search

- ▶ There is bad reduction at 17, and possibly at 2.
- ▶ The quartic factor is independent of the choice of  $\sqrt{17}$  modulo  $p$ , so we guess that there is descent from  $F = \mathbb{Q}(\sqrt{17})$  to  $\mathbb{Q}$ .
- ▶ We guess that the Hodge structure splits as

$$(1, 1, 1, 1) = (0, 1, 1, 0) + (1, 0, 0, 1)$$

and accordingly the (likely)  $\ell$ -adic Galois representation

$$\rho: \text{Gal}_F \rightarrow \text{GL}_4(\mathbb{Q}_\ell)$$

is the base change from  $\mathbb{Q}$  to  $F$  of

$$\rho_f(-1) \oplus \rho_g$$

where  $f, g$  are classical modular forms of weights 2, 4.

- ▶ We have a fair amount of eigenvalue data:

$p$	13	19	43	67	...
$a_p(f)$	2	-4	-4	-4	...
$a_p(g)$	-42	60	508	-676	...

Let's [search for the form!](#)

## Item 7: Weight 1 forms

## Item 7: Weight 1 forms

We treat weight 1 forms.

## Item 7: Weight 1 forms

We treat weight 1 forms.

- a. The data set majorizes the Buzzard–Lauder database.

## Item 7: Weight 1 forms

We treat weight 1 forms.

- a. The data set majorizes the Buzzard–Lauder database.
- b. The projective image types have been proved for all weight 1 forms, requiring a couple of new tricks. [124.1.i.a](#) (smallest  $A_4$ ), [148.1.f.a](#) (smallest  $S_4$ ), [1763.1.p.b](#) (new  $A_5$ ); [3600.1.e.a](#) ( $D_2$ , having RM and CM), [3997.1.cz.a](#) ( $D_{285}$ ).

## Item 7: Weight 1 forms

We treat weight 1 forms.

- a. The data set majorizes the Buzzard–Lauder database.
- b. The projective image types have been proved for all weight 1 forms, requiring a couple of new tricks. [124.1.i.a](#) (smallest  $A_4$ ), [148.1.f.a](#) (smallest  $S_4$ ), [1763.1.p.b](#) (new  $A_5$ ); [3600.1.e.a](#) ( $D_2$ , having RM and CM), [3997.1.cz.a](#) ( $D_{285}$ ).
- c. Artin representations are matched whenever these were already present in the LMFDB. [385.1.q.b](#)

## Item 7: Weight 1 forms

We treat weight 1 forms.

- a. The data set majorizes the Buzzard–Lauder database.
- b. The projective image types have been proved for all weight 1 forms, requiring a couple of new tricks. [124.1.i.a](#) (smallest  $A_4$ ), [148.1.f.a](#) (smallest  $S_4$ ), [1763.1.p.b](#) (new  $A_5$ ); [3600.1.e.a](#) ( $D_2$ , having RM and CM), [3997.1.cz.a](#) ( $D_{285}$ ).
- c. Artin representations are matched whenever these were already present in the LMFDB. [385.1.q.b](#)
- d. Dimension data for subspaces by projective image type.  
[1161.1.i](#).

## A sampling of new features (items 8–10)

## A sampling of new features (items 8–10)

8. Can view dimension tables, with constraints. For example, rational forms.

## A sampling of new features (items 8–10)

8. Can view dimension tables, with constraints. For example, rational forms.
9. Dynamic statistics: you can tabulate statistics for a specified subset of the data. For example, what are the **statistics** for weight and level for those forms having self-twist?

## A sampling of new features (items 8–10)

8. Can view dimension tables, with constraints. For example, rational forms.
9. Dynamic statistics: you can tabulate statistics for a specified subset of the data. For example, what are the statistics for weight and level for those forms having self-twist?
10. Provable analytic ranks (including non-self dual forms of analytic rank 1).

## A sampling of new features (items 8–10)

8. Can view dimension tables, with constraints. For example, rational forms.
9. Dynamic statistics: you can tabulate statistics for a specified subset of the data. For example, what are the statistics for weight and level for those forms having self-twist?
10. Provable analytic ranks (including non-self dual forms of analytic rank 1). We found a rational form (2.42.a.a) of weight 42 and a form (8.14.b.a) of weight 14 each having analytic rank 1.

## Future plans

## Future plans

1. Annotations for interesting examples (IAS, March 2019).

## Future plans

1. Annotations for interesting examples (IAS, March 2019).
2. Think of the LMFDB as like an “OEIS (Online Encyclopedia of Integer Sequences) for  $L$ -functions.”

## Future plans

1. Annotations for interesting examples (IAS, March 2019).
2. Think of the LMFDB as like an “OEIS (Online Encyclopedia of Integer Sequences) for  $L$ -functions.”  
 $L$ -functions can have many origins: [2.2.56.1-32.1-c](#) has at least 12 origins!

## Future plans

1. Annotations for interesting examples (IAS, March 2019).
2. Think of the LMFDB as like an “OEIS (Online Encyclopedia of Integer Sequences) for  $L$ -functions.”  
 $L$ -functions can have many origins: [2.2.56.1-32.1-c](#) has at least 12 origins!
3. New data sets: mod  $p$  Galois representations, Siegel paramodular forms, half-integer weight modular forms, hypergeometric motives, Belyi maps, small groups, ...

## Future plans

1. Annotations for interesting examples (IAS, March 2019).
2. Think of the LMFDB as like an “OEIS (Online Encyclopedia of Integer Sequences) for  $L$ -functions.”  
 $L$ -functions can have many origins: [2.2.56.1-32.1-c](#) has at least 12 origins!
3. New data sets: mod  $p$  Galois representations, Siegel paramodular forms, half-integer weight modular forms, hypergeometric motives, Belyi maps, small groups, ...

We hope the LMFDB will be a useful tool on your workbench!