p-adics in Sage

Jean-Pierre Flori
ANSSI, France

Sage Days 53, Oxford
September 24, 2013

Section 1

p-adics in Sage

p-adics in Sage

- Representing p-adic numbers require an infinite amount of data. . .
- . . . which our computers do not currently provide.
- Sage provides:
  1. Different representation of p-adics integers.
  2. One representation of p-adic fields.
  3. Unramified and Eisenstein extensions thereof.
  4. General extensions currently being worked on by Julian Rüth.

Precision tracking

- Currently, three different ways to represent p-adics in Sage:
  1. fixed modulus (only for integral element),
  2. capped absolute (only for integral element),
  3. capped relative (both for p-adic ring and field).
- These implementations basically represent p-adics using integers with bounded precision for the unit part.

- Another representation using power series is possible and can for example be used to implement lazy p-adics.
- Currently, there is only one way to deal with the precision of each type of non-basic p-adic object.
- Both these points are being addressed by Xavier Caruso and David Roe.

Implementation details

- p-adics numbers:
  - Fixed-modulus elements share a common precision and are represented using one `mpz_t`.
  - Capped-absolute elements track their own precision and are represented using one `mpz_t`.
  - Capped-relative elements track their own precision and are represented using one `mpz_t` for the unit part and an additional integer for the valuation.
- q-adics numbers:
  - Through NTL's `ZZ_pX` class.
  - Further details for precision tracking similar to those for p-adic numbers.

Functionalities

- All basic functionalities for p-adic and q-adic numbers.
- And much more.
- Demo!

Current and future developments

- Refactoring of the p-adic code to use templates similar to what is done for polynomials by David Roe and Julian Rüth.
- This makes the code easier to maintain and make it possible to use different low-level implementations much more easily.
- This is trac ticket 12555 and is positively reviewed; hopefully to be included in Sage 6.0.

- Much more flexible precision models by Xavier Caruso and David Roe.
- Experimental code available on CETHop's project website.

- General extension of p-adics numbers by Julian Rüth.

Section 2

p-adics in FLINT

FLINT: Fast Library for Number Theory

- C library on top of GMP/MPIR, MPFR (with support for NTL).
- FLINT 1 (2007/xx – 2010/12) originally developed by Hart, Harvey and Novocin.
- FLINT 2 (2011/01 –) is a completele rewrite by Bill Hart, Frederik Johansson and Sebastian Pancratz.
- About 130k lines of C code.
- Used by Sage since 2007.
- Used by Singular since 2011/12, code by Martin Lee; not used in Sage, see trac ticket 13331.

p-adics in FLINT

- `padic` module in FLINT 2 since version 2.2 (released 2011/06/04), mostly by Sebastiab Pancratz.
- `padic_poly`, `padic_matrix` and `qadic` modules on Pancratz's github since a few years, to be included into version 2.4.
- About 14k lines of C code.
- Backward incompatible changes between versions 2.3 and 2.4 (more on that later).

p-adics in Sage using FLINT

- Unramified p-adics implementation using the new template interface.
- See trac ticket 14304 and `https://github.com/saraedum/sage-renamed/tree/Zq`.
- This relies on the `fmpz_mod_poly` module.
- No implementation using the `padic`, `padic_poly` and `qadic` modules yet?

Other applications

- Point counting using deformation theory available on Sebastian Pancratz's github.
- Point counting à la Satoh, . . . , Harley using a custom `qadic_dense` module available on Jean-Pierre Flori's github.
- Both of these are based on version 2.3, so have to be rebased on top of the new `padic` structure..

Design decisions

Decision.

- Each p-adic operation treats the input as exact data and requires the desired output precision as a separate argument.

Rationale.

- A number is just a number.
- The intrinsic difficulty in p-adic arithmetic stems from the precision loss, which depends on the particular operation.
- Note that it would be straightforward to implement various precision models on top of this.

Design decisions

An element $x \neq 0$ is typically stored as $x = pu$ with $v = \mathrm{ord}_p(x) \in \mathbb{Z}$ and $u \in \mathbb{Z}$ with $p \nmid u$.
In 2.3 and before.

```
typedef struct {
    fmpz u ;
    long v ;
} padic_struct ;
```

After 2.3.

```
typedef struct {
    fmpz u;
    slong v;
    slong N;
} padic_struct;
```

Design decisions

Additional information stored in a context object.
In 2.3 and before.

```
typedef struct {
    fmpz_t p;
    long N;

    double pinv;

    fmpz *pow;
    long min;
    long max;

    enum padic_print_mode mode;
} padic_ctx_struct;
```

From 2.3 onward the precision is not stored anymore.

Design decisions

Remarks.

- Improved maintainability by having one data type; no special case depending on the size of p or p^N;
- One could consider a different implementation performing basic arithmetic to base p^k with k s.t. such that p^k fits in a word. This would allow replacing mod p^N operations by mod p^k operations (with a precomputed word-sized inverse) in many algorithms.

Functions for $\mathbb{Q}_p$

- Addition, subtraction, negation
- Multiplication, powers
- Inversion
- Inversion (with precomputed lifting structure)
- Division
- Square root
- Exponential
- Logarithm
- Teichmüller lift

Addition

Signature

```
void padic_add(z, x, y, ctx)
```

Contract

Assumes that x and y are reduced modulo p^N and returns z in reduced form, too.

Algorithm

Avoids expensive modulo operation, replacing this by one comparison and at most one subtraction.

Multiplication

Signature

```
void padic_mul(z, x, y, ctx)
```

Contract

Makes no assumptions on x and y, returns z reduced modulo p^N.

Inversion

Signature

```
void padic_inv(z, x, ctx)
```

Contract

Makes no assumptions on x.

Algorithm

Hensel lifting on $g(X) = xX - 1$, starting from an inverse in $\mathbb{F}_p$ and using the update formula $z = z + z(1 - xz)$.

Square root

Signature

```
int padic_sqrt(z, x, ctx)
```

Contract

Makes no assumptions on x. Returns whether x is actually a square and if so computes its square root.

Algorithm

- Hensel lifting to compute an inverse square root to half precision.
- The final step performs the needed inversion as well.

Teichmüller lift

Signature

```
void padic_teichmuller(z, x, ctx)
```

Contract

Assumes only that $\mathrm{ord}_p(x) = 0$.

Algorithm

Hensel lifting, avoiding inversions.

Exponential

Signature

```
int padic_exp(z, x, ctx)
```

Contract

Return whether the series converges, and if so computes the exponential.

Algorithm

Evaluate the truncated series, multiplying by the common factorial in denominators, hence requiring only one inversion.

- Rectangular splitting.
- Balanced splitting.

Logarithm

Signature

```
int padic_log(z, x, ctx)
```

Contract

Return whether the series converges, and if so computes the logarithm.

Algorithm

Evaluate the truncated series, performing an inversion for each summand.

- Rectangular splitting.
- Balanced splitting (quasi-linear in N when p is fixed).
- à la SST.

Polynomials over $\mathbb{Q}_p$

We represent a non-zero polynomial $f(X) \in \mathbb{Q}_p[X]$ as

$$f(X) = p^v(a_0 + a_1 X + \cdots + a_n X^n)$$

where a_0, . . . , $a_n \in \mathbb{Z}$ and, for at least one i , p does not divide a_i.

Functions for $\mathbb{Q}_p[X]$

- Conversions to polynomials over $\mathbb{Z}$ and $\mathbb{Q}$
- Coefficient manipulation
- Addition, subtraction, negation
- Scalar multiplication
- Multiplication
- Powers
- Series inversion
- Derivative
- Evaluation
- Composition

Unramified extensions $\mathbb{Q}_q$

We represent an unramified extension of $\mathbb{Q}_p$ as

$$\mathbb{Q}_q = \mathbb{Q}_p[X]/(f(X))$$

where $f(X) \mod p$ is separable, storing $f(X)$ in a data structure for sparse polynomials.
This allows for the reduction of a degree n polynomial modulo $f(X)$ in linear time $O(n)$ (but slow Frobenius substitutions...).

Functions for $\mathbb{Q}_q$

- Addition, subtraction, negation
- Multiplication
- Powers
- Inversion
- Exponential
- Logarithm
- Frobenius
- Teichmüller lift
- Trace
- Norm

Exponential

Signature

```
int qadic_exp(z, x, ctx)
```

Contract

Return whether the series converges, and if so computes the exponential.

Algorithm

Evaluate the truncated series, performing an inversion at each step.

- Rectangular splitting.
- Balanced splitting.

Logarithm

Signature

```
int qadic_log(z, x, ctx)
```

Contract

Return whether the series converges, and if so computes the logarithm.

Algorithm

Evaluate the truncated series, performing an inversion for each summand.

- Rectangular splitting.
- Balanced splitting.

Frobenius

Signature

```
void qadic_frobenius(z, x, k, ctx)
```

Contract

Computes $z = \Sigma^k(x)$.

Algorithm

- Compute $\Sigma^k(X)$ using Hensel lifting.
- Perform polynomal composition modulo p^N and $f(X)$.
- Generalize to use rectangular splitting.

Trace

Signature

```
void qadic_trace(z, x, ctx)
```

Contract

No assumptions are made on x.

Algorithm

- Compute the traces of X^i iteratively.
- Compute the trace of x.

Norm

Signature

```
void qadic_norm(z, x, ctx)
```

Contract

No assumptions are made on x.

Algorithm

- Using an analytical formula.
- Using resultants.

Future features?

- Specialize code for finite fields.
- Modular reduction for non-sparse modulus.
- Other types of extensions.
- Specific implementations for $p = 2$.
- Wrap into Sage using the new template interface.

Section 3

p-adics in Mathemagix

Mathemagix

- Computer algebra and analysis system coded in C++.
- Main developers: Joris van der Hoeven, Gégoire Lecerf, Bernard Mourrain, and others.
- lines of code for all modules.
- 120k lines of code for the three modules (`basix`, `numerix`, `algebramix`) needed for p-adics.
- Planned to be integrated into Sage, see branch `u/jpflori/mmx` for xperimentations.

p-adics in Mathemagix

- p-adics are represented as power series.
- Both naive and relaxed implementation available.
- Possibility to group power series terms into blocks.
- Choice between the different implementations mostly done through templating.

Advantages over zealous implementations:

- Increase precision as needed, no need to double it at each iteration.
- Solve recursive equations, avoiding costly inversion of Jacobians.

- A minimalistic wrapper was developed during Sage Days 52 by Jérémy Berthomieux, Xavier Caruso, and Jean-Pierre Flori.
- Unfortunately, Cython does not support templated functions (yet?), which makes the wrapper code quite ugly and hard to extend.
- And the wrapper code seems buggy!
- Demo!

Section 4

p-adics in PARI

PARI

- C library geared toward number theory.
- Currently maintained by Karim Belabas and Bill Allombert.
- 175k lines of code.
- Offers stable (2.5.x) and development (2.6.x) branches.
- One of the main pieces of Sage.
- Sage currently ships the stable version of PARI.

p-adics in PARI

- PARI exposes a `t_PADIC` type.
- A `t_PADIC` object contains:
  1. precision and valuation,
  2. unit part,
  3. powers of p.

- Usual basic functionalities.
- `sqrt`, `sqrtn`, `exp`, `log`, `gamma`, AGM.
- Machinery for Hensel lifting on p-adic numbers and unramified extensions.
- Elliptic curves over p-adic numbers.
- Very fast implementation of point counting à la Satoh on elliptic curves in small characteristic.

Section 5

Benchmarks

Benchmarks for $\mathbb{Q}_p$

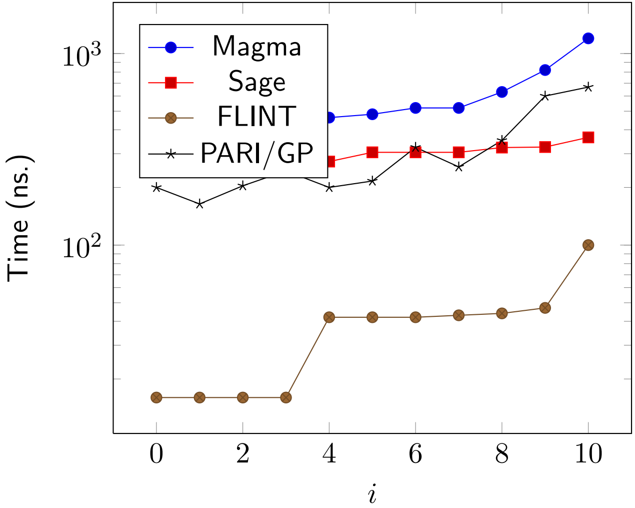
We present some timings for arithmetic in $\mathbb{Q}_p \mod p^N$ where $p = 17$, $N = 2^i$, $i = 0, \ldots, 10$, comparing the three systems Magma (V2.19-2), Sage (version 5.12.beta4, pulled from github), FLINT (pulled from github), and PARI/GP (version 2.5.4) on a machine with Intel Core i7-2620M CPU running at 2.70GHz.

To avoid worrying about taking the same random sequences of elements, we instead fix elements $a = 3^{3N}$, $b = 5^{2N}$ (and variations thereof) modulo p^N .

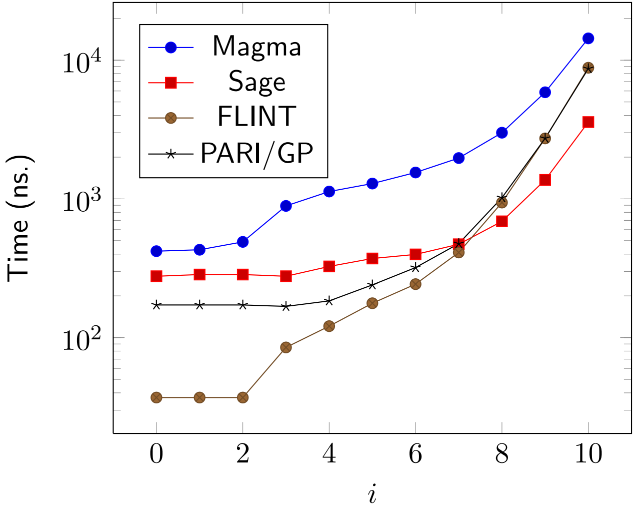
We consider the following operations:

- Addition
- Multiplication
- Inversion
- Square root
- Teichmüller lift
- Exponential
- Logarithm

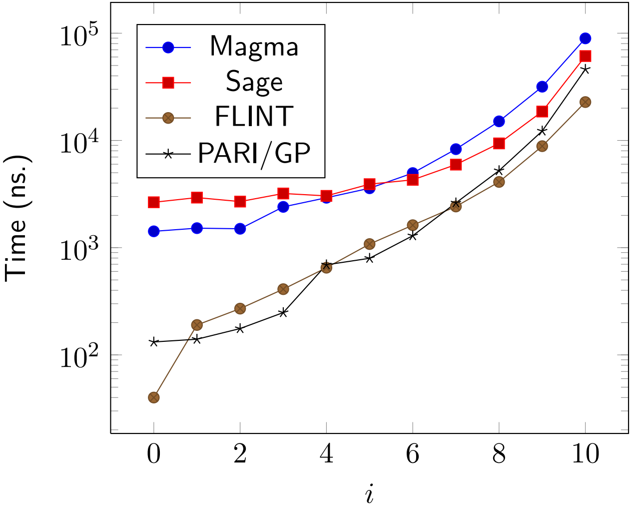
Addition



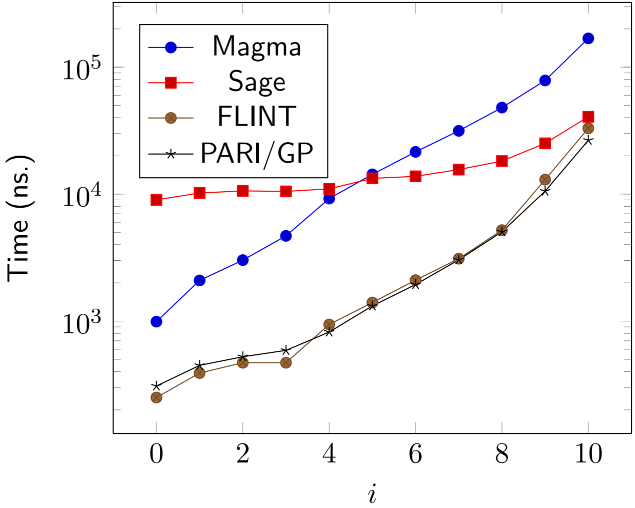
Multiplication



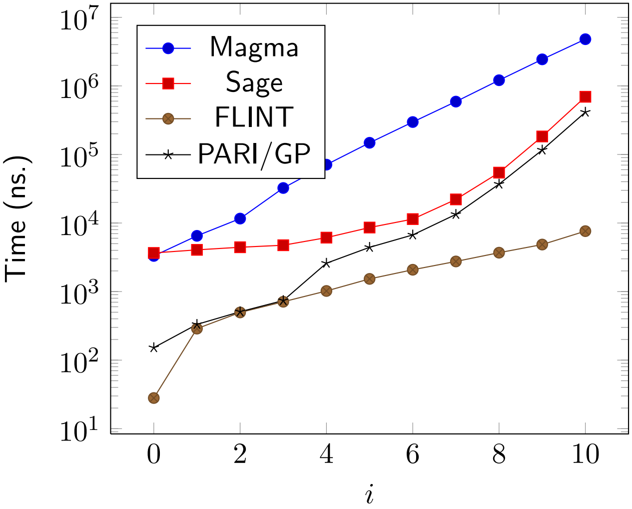
Inversion



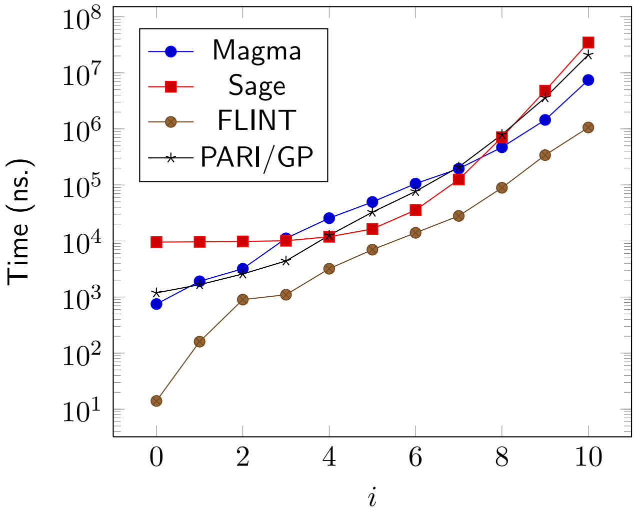
Square root



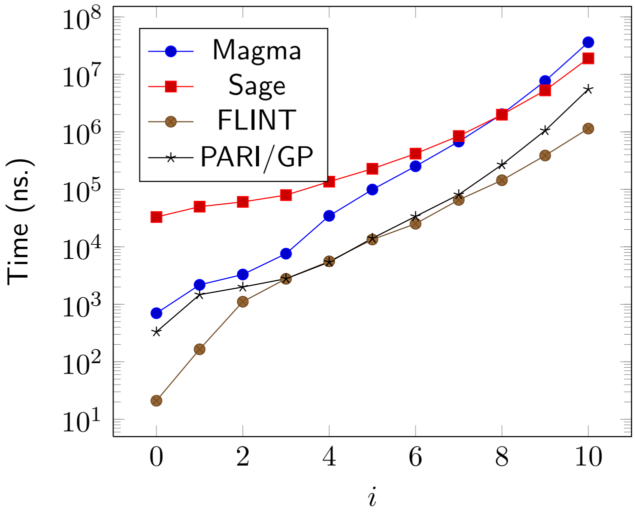
Teichmüller lift



Exponential



Logarithm



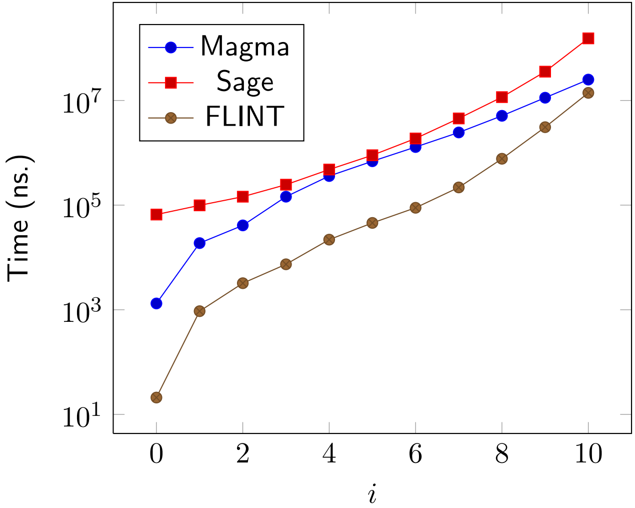
Magma
Sage
FLINT
PARI/GP
Time (ns.)
10^1
10^2
10^3
10^4
10^5
10^6
10^7
10^8
0
2
4
6
8
10
i

Benchmarks for $\mathbb{Q}_q$

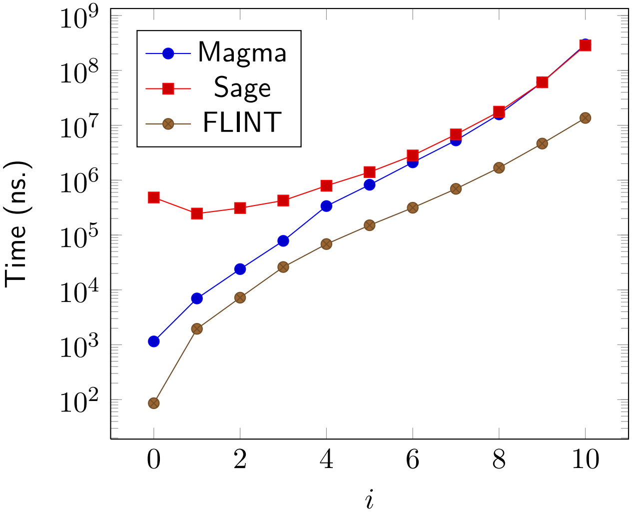
We present some timings for arithmetic in $\mathbb{Q}_q \mod p^N$ where $p = 17$, $N = 2^i$, $i = 0, ..., 10$, comparing the three systems Magma (V2.19-2), Sage (current github, 5.12.beta4) and FLINT (current github) on a machine with Intel Core i7-2620M CPU running at 2.70GHz.
To avoid worrying about taking the same random sequences of elements, we instead fix elements as before.
We consider the following operations:

- Exponential
- Logarithm
- Frobenius
- Trace
- Norm

Exponential

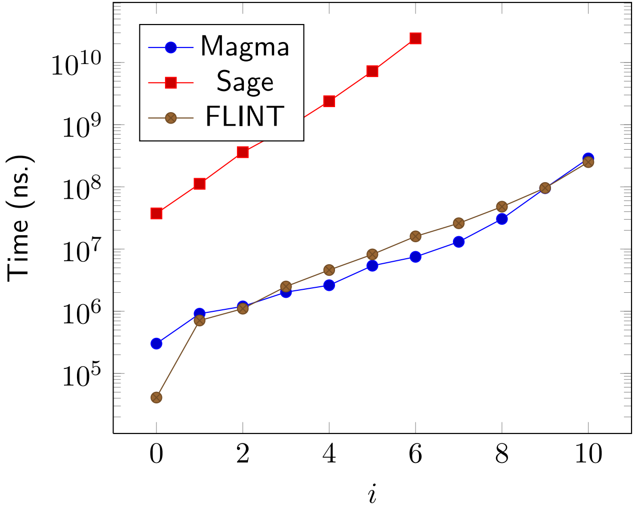


Logarithm

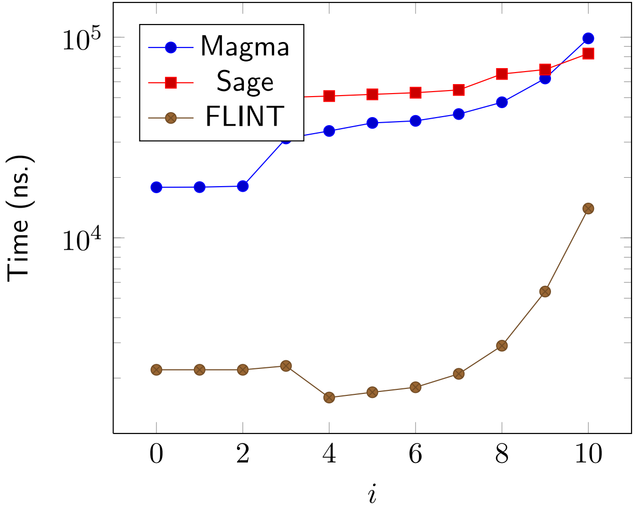


Magma
Sage
FLINT
Time (ns.)
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
0
2
4
6
8
10
i

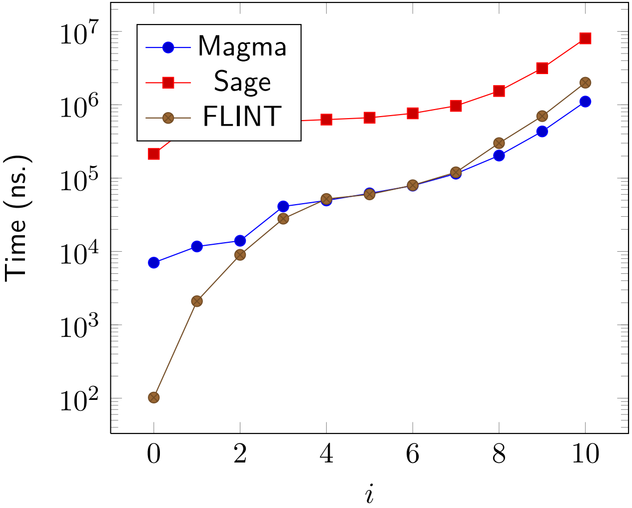
Frobenius



Trace



Norm



Section 6

Precision

Current limitations

- A p-adic object is represented by a unique object storing both:
  1. the approximation data,
  2. the precision information.
- You only have one way to deal with precision of non-basic p-adic objects.

A new model for precision

- Xavier Caruso and David Roe have been working on this.
- Approximation and precision are two completely separated objects, wrapped into an inexact p-adic object.
- As a side-effect, you also get lazy p-adics for free.
- Experimental precision package available from CETHop's website (plus trac ticket 6667).
- Demo!