Proof and certification for an accurate discriminant

S. Boldo M. Daumas W. Kahan G. Melquiond

Proval, LRI, INRIA–UPSud
LIRMM, UMII–CNRS
Mathematics and EECS, UC Berkeley
Arénaire, LIP, CNRS–ENSL–INRIA–UCBL

SCAN'2006: 12th GAMM - IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics

2006-09-29

Introduction

William Kahan has proposed a way to accurately compute the roots of a quadratic equation with floating-point arithmetic.

http://www.cs.berkeley.edu/~wkahan/Qdrtcs.pdf

Introduction

William Kahan has proposed a way to accurately compute the roots of a quadratic equation with floating-point arithmetic.

http://www.cs.berkeley.edu/~wkahan/Qdrtcs.pdf

Given a, b, and c, three floating-point numbers (precision t), the roots of $a \cdot x^2 + b \cdot x + c = 0$ are given by

$$\frac{-b + s \cdot \sqrt{D}}{a} \quad \text{and} \quad \frac{c}{-b + s \cdot \sqrt{D}}$$

with $D = b^2 - a \cdot c$ and $s \in \{-1, +1\}$.

Introduction

William Kahan has proposed a way to accurately compute the roots of a quadratic equation with floating-point arithmetic.

http://www.cs.berkeley.edu/~wkahan/Qdrtcs.pdf

Given a, b, and c, three floating-point numbers (precision t), the roots of $a \cdot x^2 + b \cdot x + c = 0$ are given by

$$\frac{-b + s \cdot \sqrt{D}}{a} \quad \text{and} \quad \frac{c}{-b + s \cdot \sqrt{D}}$$

with $D = b^2 - a \cdot c$ and $s \in \{-1, +1\}$.

> *The shortest proofs found so far for the correctness of the algorithms are, as usual for floating-point, far longer and trickier than the algorithms in question.*

Requirement: an accurate discriminant

The tricky part lies in the computation of the discriminant $D = b^2 - a \cdot c$.

Goal: Assuming there is neither overflow nor underflow, compute a floating-point number d such that the error $\delta = |d - D|$ is relatively small: $\delta \leq 2 \cdot \mathrm{ulp}(d)$.

Requirement: an accurate discriminant

The tricky part lies in the computation of the discriminant $D = b^2 - a \cdot c$.

Goal: Assuming there is neither overflow nor underflow, compute a floating-point number d such that the error $\delta = |d - D|$ is relatively small: $\delta \leq 2 \cdot \mathrm{ulp}(d)$.

Difficulty: When $b \otimes b$ and $a \otimes c$ are close, their subtraction is computed exactly. Hence the rounding errors from $b \otimes b$ and $a \otimes c$ are amplified.

Discriminant algorithm

Originally some Matlab vector code. Here is a C version:

```c
double discr(double a, double b, double c) {
  double p, q, d;
  p = b * b;
  q = a * c;
  d = p - q;
  if (3 * fabs(d) < p + q) {
    // slow path, d is not accurate enough
    double dp, dq;
    dp = fma(b, b, -p); // p + dp = b^2
    dq = fma(a, c, -q); // q + dq = a · c
    d = (p - q) + (dp - dq);
  }
  return d;
}
```

Execution paths

Two execution paths:

1. $3 \otimes |d| \geq p \oplus q$ (fast path).
   - Returned value: $d = p \ominus q$.

Execution paths

Two execution paths:

1. $3 \otimes |d| \geq p \oplus q$ (fast path).
   - Returned value: $d = p \ominus q$.
   - Error bound: $\delta \leq \frac{1}{2}\mathrm{ulp}(d) + \frac{1}{2}\mathrm{ulp}(p) + \frac{1}{2}\mathrm{ulp}(q)$.

Execution paths

Two execution paths:

1. $3 \otimes |d| \geq p \oplus q$ (fast path).
   - Returned value: $d = p \ominus q$.
   - Error bound: $\delta \leq \frac{1}{2}\text{ulp}(d) + \frac{1}{2}\text{ulp}(p) + \frac{1}{2}\text{ulp}(q)$.
   - Proof goal: $\text{ulp}(p) + \text{ulp}(q) \leq 3 \cdot \text{ulp}(d)$.

Execution paths

Two execution paths:

1. $3 \otimes |d| \geq p \oplus q$ (fast path).
   - Returned value: $d = p \ominus q$.
   - Error bound: $\delta \leq \frac{1}{2}\mathrm{ulp}(d) + \frac{1}{2}\mathrm{ulp}(p) + \frac{1}{2}\mathrm{ulp}(q)$.
   - Proof goal: $\mathrm{ulp}(p) + \mathrm{ulp}(q) \leq 3 \cdot \mathrm{ulp}(d)$.

2. $3 \otimes |d| < p \oplus q$ (slow path).
   Assuming $p \ominus q = p - q$.
   - Returned value: $d = (p - q) \oplus (dp \ominus dq)$.
     Mathematical value: $D = (p - q) + (dp - dq)$.

Execution paths

Two execution paths:

1. $3 \otimes |d| \geq p \oplus q$ (fast path).
   - Returned value: $d = p \ominus q$.
   - Error bound: $\delta \leq \frac{1}{2}\text{ulp}(d) + \frac{1}{2}\text{ulp}(p) + \frac{1}{2}\text{ulp}(q)$.
   - Proof goal: $\text{ulp}(p) + \text{ulp}(q) \leq 3 \cdot \text{ulp}(d)$.

2. $3 \otimes |d| < p \oplus q$ (slow path).
   Assuming $p \ominus q = p - q$.
   - Returned value: $d = (p - q) \oplus (dp \ominus dq)$.
     Mathematical value: $D = (p - q) + (dp - dq)$.
   - Error bound: $\delta \leq \frac{1}{2}\text{ulp}(d) + \frac{1}{2}\text{ulp}(dp \ominus dq)$.

Execution paths

Two execution paths:

1. $3 \otimes |d| \geq p \oplus q$ (fast path).
   - Returned value: $d = p \ominus q$.
   - Error bound: $\delta \leq \frac{1}{2}\text{ulp}(d) + \frac{1}{2}\text{ulp}(p) + \frac{1}{2}\text{ulp}(q)$.
   - Proof goal: $\text{ulp}(p) + \text{ulp}(q) \leq 3 \cdot \text{ulp}(d)$.

2. $3 \otimes |d| < p \oplus q$ (slow path).
   Assuming $p \ominus q = p - q$.
   - Returned value: $d = (p - q) \oplus (dp \ominus dq)$.
     Mathematical value: $D = (p - q) + (dp - dq)$.
   - Error bound: $\delta \leq \frac{1}{2}\text{ulp}(d) + \frac{1}{2}\text{ulp}(dp \ominus dq)$.
   - Proof goal: $\text{ulp}(dp \ominus dq) \leq 3 \cdot \text{ulp}(d)$.

Motivation for a formal proof

The proof relies on numerous case studies and constantly goes forth and back between real arithmetic and floating-point arithmetic.

A formal certification guarantees no case was forgotten and each step of the proof is sound.

The Coq proof assistant was used with the formalization by Daumas, Rideau, and Théry.

Formalization of floating-point arithmetic

A pair of integers (m, e) is associated to the real value $m \cdot 2^e$.
This is a representable floating-point number if $|m| < 2^t$ holds.

IEEE-754 specifies that the results of floating-point operations are the same as if they were first computed in infinite precision and then rounded to precision t.

Execution paths

Let us suppose that the hypotheses for the execution paths are:

1. $q \le \frac{p}{2}$ or $q \ge 2 \cdot p$ (fast path).
   - Equivalent to $3 \cdot |p - q| \ge p + q$.
   - Not equivalent to $3 \otimes |d| \ge p \oplus q$.

Execution paths

Let us suppose that the hypotheses for the execution paths are:

1. $q \leq \frac{p}{2}$ or $q \geq 2 \cdot p$ (fast path).
   - Equivalent to $3 \cdot |p - q| \geq p + q$.
   - Not equivalent to $3 \otimes |d| \geq p \oplus q$.

2. $\frac{p}{2} \leq q \leq 2 \cdot p$ (slow path).
   - The assumption $p \ominus q = p - q$ now holds.
   - It implies $3 \otimes |d| \leq p \oplus q$.

The fast path: $q \leq \frac{p}{2}$ or $q \geq 2 \cdot p$

Returned value: $d = p \ominus d$ with $p = b \otimes b$ and $q = a \otimes c$.
Proof goal: $\mathrm{ulp}(p) + \mathrm{ulp}(q) \leq 3 \cdot \mathrm{ulp}(d)$.
Trivial case: $q = 0$; assuming $q \neq 0$.

The fast path: $q \le \frac{p}{2}$ or $q \ge 2 \cdot p$

Returned value: $d = p \ominus d$ with $p = b \otimes b$ and $q = a \otimes c$.
Proof goal: $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 3 \cdot \mathrm{ulp}(d)$.
Trivial case: $q = 0$; assuming $q \neq 0$.

1. When q is negative,
   - $d \ge p$ and $d \ge |q|$,
   - $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 2 \cdot \mathrm{ulp}(d)$.

The fast path: $q \le \frac{p}{2}$ or $q \ge 2 \cdot p$

Returned value: $d = p \ominus d$ with $p = b \otimes b$ and $q = a \otimes c$.
Proof goal: $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 3 \cdot \mathrm{ulp}(d)$.
Trivial case: $q = 0$; assuming $q \ne 0$.

1. When q is negative,
   - $d \ge p$ and $d \ge |q|$,
   - $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 2 \cdot \mathrm{ulp}(d)$.
2. When q is positive and $q \le \frac{p}{2}$,
   - $p - q \ge \frac{p}{2} \ge q \ge 0$,
   - $\mathrm{ulp}(d) \ge \frac{1}{2}\mathrm{ulp}(p) \ge \mathrm{ulp}(q)$,
   - $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 3 \cdot \mathrm{ulp}(d)$.

The fast path: $q \le \frac{p}{2}$ or $q \ge 2 \cdot p$

Returned value: $d = p \ominus d$ with $p = b \otimes b$ and $q = a \otimes c$.
Proof goal: $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 3 \cdot \mathrm{ulp}(d)$.
Trivial case: $q = 0$; assuming $q \ne 0$.

1. When q is negative,
   - $d \ge p$ and $d \ge |q|$,
   - $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 2 \cdot \mathrm{ulp}(d)$.
2. When q is positive and $q \le \frac{p}{2}$,
   - $p - q \ge \frac{p}{2} \ge q \ge 0$,
   - $\mathrm{ulp}(d) \ge \frac{1}{2}\mathrm{ulp}(p) \ge \mathrm{ulp}(q)$,
   - $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 3 \cdot \mathrm{ulp}(d)$.
3. When q is positive and $q \ge 2 \cdot p$,
   - $q - p \ge \frac{q}{2} \ge p \ge 0$,
   - $\mathrm{ulp}(d) \ge \frac{1}{2}\mathrm{ulp}(q) \ge \mathrm{ulp}(p)$,
   - $\mathrm{ulp}(p) + \mathrm{ulp}(q) \le 3 \cdot \mathrm{ulp}(d)$.

The slow path: $\frac{p}{2} \le q \le 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.

Trivial case: $p = q$; assuming $p \ne q$.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Trivial case: $p = q$; assuming $p \neq q$.

1. When $dp \ominus dq = dp - dq$,
   - $d = (p - q) \oplus (dp - dq)$, so $\delta \leq \frac{1}{2}\text{ulp}(d)$.
   - Note: when $\text{ulp}(p) = \text{ulp}(q)$, $dp \ominus dq = dp - dq$.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Trivial case: $p = q$; assuming $p \neq q$.

1. When $dp \ominus dq = dp - dq$,
   - $d = (p - q) \oplus (dp - dq)$, so $\delta \leq \frac{1}{2}\text{ulp}(d)$.
   - Note: when $\text{ulp}(p) = \text{ulp}(q)$, $dp \ominus dq = dp - dq$.
2. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| \geq 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$.
3. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| < 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Trivial case: $p = q$; assuming $p \neq q$.

1. When $dp \ominus dq = dp - dq$,
   - $d = (p - q) \oplus (dp - dq)$, so $\delta \leq \frac{1}{2}\text{ulp}(d)$.
   - Note: when $\text{ulp}(p) = \text{ulp}(q)$, $dp \ominus dq = dp - dq$.
2. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| \geq 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$.
3. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| < 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$.

Why $3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$?

- $\max(\text{ulp}(p), \text{ulp}(q)) = 2 \cdot \min(\text{ulp}(p), \text{ulp}(q))$, since $\frac{p}{2} \leq q \leq 2 \cdot p$.
- $|dp - dq| \leq |dp| + |dq| \leq \frac{1}{2}\text{ulp}(p) + \frac{1}{2}\text{ulp}(q) \leq \frac{3}{2}\min(\text{ulp}(p), \text{ulp}(q))$.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.

Proof goal: $\mathrm{ulp}(dp \ominus dq) \leq 3 \cdot \mathrm{ulp}(d)$.

2. When $\mathrm{ulp}(p) \neq \mathrm{ulp}(q)$ and $|p - q| \geq 3 \cdot \min(\mathrm{ulp}(p), \mathrm{ulp}(q))$,
   - $|dp - dq| \leq \frac{3}{2}\min(\mathrm{ulp}(p), \mathrm{ulp}(q)) \leq \frac{1}{2}|p - q|$,
   - $|(p - q) + (dp \ominus dq)| \geq \frac{1}{2}|p - q| \geq |dp - dq|$,
   - $\mathrm{ulp}(d) \geq \mathrm{ulp}(dp \ominus dq)$.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Proof goal: $\text{ulp}(dp \ominus dq) \leq 3 \cdot \text{ulp}(d)$.

2. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| \geq 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$,
   - $|dp - dq| \leq \frac{3}{2} \min(\text{ulp}(p), \text{ulp}(q)) \leq \frac{1}{2}|p - q|$,
   - $|(p - q) + (dp \ominus dq)| \geq \frac{1}{2}|p - q| \geq |dp - dq|$,
   - $\text{ulp}(d) \geq \text{ulp}(dp \ominus dq)$.

3. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| < 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$,
   - $|p - q|$ is a multiple of $\min(\text{ulp}(p), \text{ulp}(q))$, so either $2\times$ or $1\times$.
   - p and q have to be close to a power of 2.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.

Proof goal: $\text{ulp}(dp \ominus dq) \leq 3 \cdot \text{ulp}(d)$.

2. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| \geq 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$,
   - $|dp - dq| \leq \frac{3}{2} \min(\text{ulp}(p), \text{ulp}(q)) \leq \frac{1}{2}|p - q|$,
   - $|(p - q) + (dp \ominus dq)| \geq \frac{1}{2}|p - q| \geq |dp - dq|$,
   - $\text{ulp}(d) \geq \text{ulp}(dp \ominus dq)$.

3. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| < 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$,
   - $|p - q|$ is a multiple of $\min(\text{ulp}(p), \text{ulp}(q))$, so either $2\times$ or $1\times$.
   - p and q have to be close to a power of 2.

   Without any loss of generality, let us assume $p = 1$ and $p \geq q$.
   Only two cases:
   - $p = 1$ and $q = 1^- = 1 - \frac{1}{2}\text{ulp}(1)$,
   - $p = 1$ and $q = 1^{--} = 1 - \text{ulp}(1)$.

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

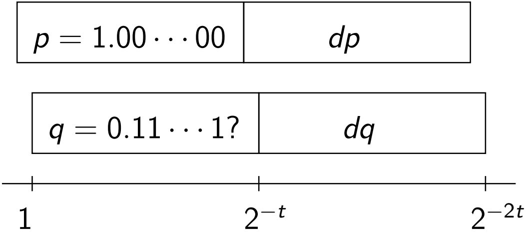
Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Proof goal: $\mathrm{ulp}(dp \ominus dq) \leq 3 \cdot \mathrm{ulp}(d)$ or $dp \ominus dq = dp - dq$.

3. When $\mathrm{ulp}(p) \neq \mathrm{ulp}(q)$ and $|p - q| < 3 \cdot \min(\mathrm{ulp}(p), \mathrm{ulp}(q))$, (assuming $p = 1$ and either $q = 1^-$ or $q = 1^{--}$)

The slow path: $\frac{p}{2} \le q \le 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Proof goal: $\mathrm{ulp}(dp \ominus dq) \le 3 \cdot \mathrm{ulp}(d)$ or $dp \ominus dq = dp - dq$.

3. When $\mathrm{ulp}(p) \neq \mathrm{ulp}(q)$ and $|p - q| < 3 \cdot \min(\mathrm{ulp}(p), \mathrm{ulp}(q))$, (assuming $p = 1$ and either $q = 1^-$ or $q = 1^{--}$)



The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Proof goal: $\text{ulp}(dp \ominus dq) \leq 3 \cdot \text{ulp}(d)$ or $dp \ominus dq = dp - dq$.

3. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| < 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$, (assuming $p = 1$ and either $q = 1^{-}$ or $q = 1^{--}$)

   3.1 When dp and dq have the same sign, $dp \ominus dq = dp - dq$. (Both dp and dq are multiple of 2^{-2t}, and $|dp - dq| < 2^{-t}$.)

The slow path: $\frac{p}{2} \leq q \leq 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Proof goal: $\text{ulp}(dp \ominus dq) \leq 3 \cdot \text{ulp}(d)$ or $dp \ominus dq = dp - dq$.

3. When $\text{ulp}(p) \neq \text{ulp}(q)$ and $|p - q| < 3 \cdot \min(\text{ulp}(p), \text{ulp}(q))$, (assuming $p = 1$ and either $q = 1^-$ or $q = 1^{--}$)

   3.1 When dp and dq have the same sign, $dp \ominus dq = dp - dq$. (Both dp and dq are multiple of 2^{-2t}, and $|dp - dq| < 2^{-t}$.)

   3.2 When dp and dq have opposite signs and $dp - dq \geq 0$,
   $d \geq p - q \leq 2^{-t}$ and $dp - dq \leq 2^{-t} + 2^{-1-t}$,
   so $\text{ulp}(dp \ominus dq) \leq 2^{1-2t} \leq \text{ulp}(d)$.

The slow path: $\frac{p}{2} \le q \le 2 \cdot p$

Returned value: $d = (p - q) \oplus (dp \ominus dq)$
with $p + dp = b^2$ and $q + dq = a \cdot c$.
Proof goal: $\mathrm{ulp}(dp \ominus dq) \le 3 \cdot \mathrm{ulp}(d)$ or $dp \ominus dq = dp - dq$.

3. When $\mathrm{ulp}(p) \ne \mathrm{ulp}(q)$ and $|p - q| < 3 \cdot \min(\mathrm{ulp}(p), \mathrm{ulp}(q))$,
(assuming $p = 1$ and either $q = 1^-$ or $q = 1^{--}$)

   3.1 When dp and dq have the same sign, $dp \ominus dq = dp - dq$.
   (Both dp and dq are multiple of 2^{-2t}, and $|dp - dq| < 2^{-t}$.)

   3.2 When dp and dq have opposite signs and $dp - dq \ge 0$,
   $d \ge p - q \le 2^{-t}$ and $dp - dq \le 2^{-t} + 2^{-1-t}$,
   so $\mathrm{ulp}(dp \ominus dq) \le 2^{1-2t} \le \mathrm{ulp}(d)$.

   3.3 When dp and dq have opposite signs and $dp - dq \le 0$,
   $dp \ominus dq = dp - dq$.
   (Either $dp - dq = -2^{-t}$ or as in 3.1).

What is missing?

The proof is not complete since the hypotheses do not match the conditional execution.

Two steps may be required:

1. Modify the algorithm: change the boolean expression to $3 \otimes |d| \leq p \oplus q$ for the slow path.
2. Prove that the slow path still works correctly when both $3 \cdot |p - q| > p + q$ and $3 \otimes |p \ominus q| \leq p \oplus q$ hold.

What is missing?

The proof is not complete since the hypotheses do not match the conditional execution.

Two steps may be required:

1. Modify the algorithm: change the boolean expression to $3 \otimes |d| \leq p \oplus q$ for the slow path.
2. Prove that the slow path still works correctly when both $3 \cdot |p - q| > p + q$ and $3 \otimes |p \ominus q| \leq p \oplus q$ hold.

Two potential reasons why the slow path should still work:

- $p \ominus q = p - q$ when $\frac{p}{2+2^{-t}} \leq q \leq (2 + 2^{-t}) \cdot p$,
- if the subtraction is not exact, the additional error-term is bounded by $\text{ulp}(d)$.

Conclusion

- The paper proof is not meant to convince you: it has an educational purpose.

Conclusion

- The paper proof is not meant to convince you: it has an educational purpose.
- The correctness is guaranteed by a formal proof that can be mechanically checked.

Conclusion

- The paper proof is not meant to convince you: it has an educational purpose.
- The correctness is guaranteed by a formal proof that can be mechanically checked.
- Formal proofs are the only way to ensure that tricky arithmetic codes are reliable.

Conclusion

- The paper proof is not meant to convince you: it has an educational purpose.
- The correctness is guaranteed by a formal proof that can be mechanically checked.
- Formal proofs are the only way to ensure that tricky arithmetic codes are reliable.

Future work:

- Complete the proof.

Conclusion

- The paper proof is not meant to convince you: it has an educational purpose.
- The correctness is guaranteed by a formal proof that can be mechanically checked.
- Formal proofs are the only way to ensure that tricky arithmetic codes are reliable.

Future work:

- Complete the proof.
- Prove the modified algorithm ($3 \longrightarrow 1024$) when extra-precision is available for intermediate results.

Questions?

Formal development: http://lipforge.ens-lyon.fr/www/pff/

Mail: guillaume.melquiond@ens-lyon.fr