Postmodern Primality Proving

Preda Mihăilescu

Mathematical Institute, University of Göttingen, Germany

June 28, 2013

Present talk focuses on the problem of **distinguishing rational primes from composites**.

Thus $n \in \mathbb{N}$ is always a *test - number*.

The algorithms for doing this may fulfill one ore more of the following purposes:

A. Ad hoc (trial division is sufficient).

B. Practical applications – high reliability required, proofs not necessary (e.g. cryptography).

C. (Reproducible) proofs for very large numbers.

D. Achieve complexity theoretical goals (polynomial, deterministic, etc.)

Pocklington - Morrison:

Theorem

Suppose that I know some large factored part:

$$F = \prod_i q_i = \prod_i \ell_i^{m_i} | (n-1).$$

Furthermore, $a_i \in \mathbb{Z}$ *with* $(a_i, n) = 1$ *and*

$$\boxed{a_i^{n-1} \equiv 1 \mod n, \quad \left(a_i^{(n-1)/\ell_i} - 1, n\right) = 1 \quad \forall i.}$$

Then $p \equiv 1 \mod F$ *for all primes* $p | n$*. Similar in a quadratic extension, for* $q | (n+1)$*.*

In particular, if $F > \sqrt{n}$**, then** n **is prime**.

Consequences:

Together with some not too surprizing tricks for extensions of degree 2 and 4: origin of the **Lucas – Lehmer** *family* of tests. These are **deterministic tests**, requiring some massive additional information (factor F).

Certificates Idea: Let the first run of a primality test find some *information on n* which allows it, in later runs, to quick(er) prove its primality (if it does hold).

Pratt Certificates: Recursive tree rooted at n and based on the previous Theorem:

- Et each level, a prime m to be certified comes with a list of triples

  $$(a_i, \ell_i, e_i) \quad \text{such that} \quad q_i = \ell_i^{e_i} \quad \text{and} \quad F = (\prod_i q_i^{e_i}) \mid (m-1),$$

  and the Pocklington - Morrison test is verified.
- The values q_i are pseudo - primes and nodes for a primality certificate at the next level.
- Sufficiently small (e.g. < 1000) primes are certified by trial and error division. This are the terminal primes of the certificate tree.

Compositeness tests revisited

- Solovay – Strassen

$$C : a^{(n-1)/2} = \left(\frac{a}{n}\right), \quad \delta_C = 1/2.$$

- Strong pseudoprime test (Selfridge, Miller, Rabin et. al.).
  Let $n - 1 = 2^h \cdot m$ with odd m.

$$C : \begin{cases} a^m \equiv 1 \mod n \quad \text{or} \\ a^{2^{k-1}\cdot m} \equiv -1 \mod n \quad \text{and} \quad a^{2^k \cdot m} \equiv 1 \mod n \end{cases}$$

  for some $0 < k \le h$. For this $\delta_C = 1/4$.

- Quadratic (Frobenius !) test of Grantham. C : .... more complicated, essentially Lucas in quadratic extensions. $\delta_C < 1/7710$.

Alternative estimate of Damgård, Landrock, Pomerance

Rather than worst case, average case error probability - tables for the strong pseudoprime test.

k / t	1	2	3	4	5	6	7	8	9	10
100	5	14	20	25	29	33	36	39	41	44
150	8	20	28	34	39	43	47	51	54	57
200	11	25	34	41	47	52	57	61	65	69
250	14	29	39	47	54	60	65	70	75	79
300	19	33	44	53	60	67	73	78	83	88
350	28	38	48	58	66	73	80	86	91	97
400	37	46	55	63	72	80	87	93	99	105
450	46	54	62	70	78	85	93	100	106	112
500	56	63	70	78	85	92	99	106	113	119
550	65	72	79	86	93	100	107	113	119	126
600	75	82	88	95	102	108	115	121	127	133

Table: Lower bounds for $p_{k,t}$: from [DLP]

The problem of *general* primality *proving*.

Problem statement. Input a number n, decide and prove in (wishfully) polynomial time, whether n is prime or not. No false outputs, no (or "few") undecisions allowed.

Known approaches:

- Cyclotomy (Adleman, Pomerance, Lenstra, Bosma, M., et. al.)
- Elliptic curve Pocklington (Goldwasser, Kilian, Atkin, Morain)
- Hyperelliptic curve Pocklington (Adleman, Huang).
- "Introspection group cyclotomy" (Agrawal, Kayal, Saxena).
- CIDE - Cyclotomy Improved by Dual Ellptic Primes.

In the Lucas – Lehmer test, the values $b_i = a_i^{(n-1)/\ell_i}$ are *primitive* $q_i - \mathrm{th}$ *roots of unity* modulo n (in some sense ...). Their product $b = \prod_i b_i$ is an $F-\mathrm{th}$ p.r.u. Generalize this idea to extension algebras over $\mathbb{Z}/(n \cdot \mathbb{Z})$!

Theorem (Lenstra,1981)

Let $s \in \mathbb{Z}_{>0}$. Let $\mathbf{A}$ be a ring containing $\mathbb{Z}/(n \cdot \mathbb{Z})$ as a subring. Suppose that there exists $\alpha \in \mathbf{A}$ satisfying the following conditions:

$$\begin{aligned} \alpha^s &= 1, \\ \alpha^{s/q} - 1 &\in \mathbf{A}^*, \text{ for every prime } q|s, \\ \Psi_\alpha(X) = \prod_{i=0}^{t-1} (X - \alpha^{n^i}) &\in \mathbb{Z}/(n \cdot \mathbb{Z})[X], \text{ for some } t \in \mathbb{Z}_{>0} \end{aligned} \tag{1}$$

Then, for every divisor r of n there exists $i(r)$ such that

$$1 \leq i(r) < t : r = n^{i(r)} \mod s, \tag{2}$$

and in particular if r is a prime $< \sqrt{n}$, it is equal to the minimal positive representant of $n^{i(r)} \mod s$.

Consequence: Cyclotomy test CPP

- Analytic number theory shows that there is a
$$t = O\left((\log n)^{c \log\log\log n}\right), \quad \text{with } c < 1+\epsilon,$$
such that
$$s = \prod_{q \ : \ q-1|t} q > \sqrt{\log n},$$
for prime powers q.
- For such t, s, the cyclotomy test *implicitely* proves the existence of the algebra $\mathbf{A}$ and α verifying Lenstra's theorem. It uses Jacobi sums and exponentiation in small extensions of $\mathbb{Z}/(n \cdot \mathbb{Z})$.
- Asymptotic runtime **overpolynomial**, $O(t)$.
- **De facto runtime** for $\log_{10}(n) < 10^6$ is $O\left(\log(n)^4\right)$.
- For input the size of the Universe ($\log(n)$ 10^{100}, the run time still is
$$\boxed{T = O\left(\log(n)^7\right).}$$

Elliptic curves - ECPP

- Uses Pocklington for "elliptic curves"

$$E_n(a, b) : y^2 \equiv x^3 + ax + b \mod n.$$

(defined as varieties only if n is prime ...)

- Recursive: search a, b such that $|E_n(a, b)| = q.r$, with q some large pseudoprime. Use Pocklington, then recurse to prove primality of q.
- Initial Goldwasser - Kilian variant: $O(\log n)^{11}$, "random polynomial" for all but an exponentially thin subset of the inputs. Counts points using Schoof's algorithm. Impractical.
- Improvement due to Atkin and implemented by Morain: $O\left((\log n)^6\right)$, but not provable random polynomial any more - it works in practice with very few exceptions.

Comparing General Primality Proving Methods

Complexity theoretic, de facto performance marked $1 - 5$ and use of random decisions (yes/no).

Alg. / Quality	Complexity	Perf. de facto	Random (0/1)
Cyclotomy	1	5	0
ECPP	2	4	1
Hyper Elliptic	4	1	1
AKS	5	3	0

Table: Quality Marks for General Primality Proving Algorithms

The Agrawal, Kayal, Saxena (AKS) test.

Theorem (AKS)

Let n be an odd integer and $r \in \mathbb{N}$ such that:

- $$\mathrm{ord}_r(n) > 4 \cdot \log^2(n), \quad \textit{and} \quad (r, n) = 1.$$
- *The number n has no prime factor $< r$.*
- *The number n is not a prime power.*

Let $\ell = \lfloor 2\sqrt{\varphi(r)} \cdot \log(n) \rfloor$ *and* $\zeta = \zeta_r \in \mathbb{C}$ *a primitive* $r-\mathrm{th}$ *root of unity. If*

$$\boxed{(\zeta - a)^n \equiv \zeta^n - a \mod (n, \mathbb{Z}[\zeta]), \quad \forall\, 1 \leq a \leq \ell,}$$

then n is prime.

Run time count.

Lemma

There is an $r \in \mathbb{N}$ satisfying the conditions and such that $r < (2\ \log n)^5$.

Let $M(\ell)$ be the time for a multiplication in an extension of degree ℓ of $\mathbb{Z}/(n \cdot \mathbb{Z})$; then run-time is

$$T = O\left(\ell \cdot \log n \cdot M(\ell)\right) \sim O(\ell \cdot \log n)^{\rho},$$

for some $2 < \rho < 3$ Thus, for some $3 \leq k \leq 6$

$$T = O(\log n)^{k \cdot \rho}, \quad \text{for some} \quad 2 < \rho < 3.$$

Run time count.

Lemma

There is an $r \in \mathbb{N}$ satisfying the conditions and such that $r < (2\ \log n)^5$.

Let $M(\ell)$ be the time for a multiplication in an extension of degree ℓ of $\mathbb{Z}/(n \cdot \mathbb{Z})$; then run-time is

$$T = O\left(\ell \cdot \log n \cdot M(\ell)\right) \sim O(\ell \cdot \log n)^{\rho},$$

for some $2 < \rho < 3$. Thus, for some $3 \leq k \leq 6$

$$T = O(\log n)^{k \cdot \rho}, \quad \text{for some} \quad 2 < \rho < 3.$$

We gather the **lower bound**

The proof of theorem (AKS)

Definition

For fixed n, r and $\alpha = f(\zeta_r) \in \mathbb{Z}[\zeta_r]$ we say that m is **introspective** with respect to α, if

$$\sigma_m(\alpha) \equiv \alpha^m \mod n\mathbb{Z}[\zeta_r].$$

Introspection is multiplicative with respect both to m and α.

Clue of the proof: Find two groups $\mathcal{G} \subset \mathbb{Z}[\zeta_r]/(n\mathbb{Z}[\zeta_r])$ and $I \subset \mathbb{Z}/(r \cdot \mathbb{Z})$ such that $\mathcal{G}$ is introspective for I and then derive contradictory bounds for the size of $\mathcal{G} \mod p$, for any possible prime $p \mid n$, provided that n is not a prime power.

Some details

Assume that $p|n$ is a prime divisor and let $p \in \wp \subset \mathbb{Z}[\zeta]$ be a maximal ideal.

- Let $G \subset \mathbb{Z}[\zeta]$ be the group generated by $\{a + \zeta_r : 1 \le a \le \ell\}$ and $\mathcal{G} = G \mod \wp$, so $\mathcal{G}$ is a multiplicative group in a field of characteristic p: let $o(\mathcal{G})$ be its order.
- Consider the set $I_0 = \{m : \alpha^m \equiv \sigma_m(\alpha)) \mod n, \ \forall\ \alpha \in G\} \subset \mathbb{N}$ and $I = I_0 \mod r \subset \mathbb{Z}/(r \cdot \mathbb{Z})$.

With these definitions, one proves:

- If $m, m' \in I_0$ are such that $m \equiv m' \mod r$ then $m \equiv m' \mod o(G)$. Define $t = |I|$.
- $1, n^i, p^j \in I$.
- Let $E = \{n^i \cdot p^j : 0 \le i, j \le \lfloor \sqrt{r} \rfloor\} \subset I$. We have $|E| > r$: pigeon hole implies

$$\boxed{n^{i_1} p^{j_1} \equiv n^{i_2} p^{j_2} \mod r.}$$

- Above congruence holds aslo mod $o(G) > n^{2\cdot\sqrt{r}}$. Since both terms are $< n^{2\cdot\sqrt{r}}$, it must be an equality:

$$\boxed{n^{i_1-i_2} = p^{j_2-j_1},}$$

If n is not a power of p, we gather the **upper bound**

$$\boxed{\mid \mathcal{G} \mid \leq n^{\sqrt{t}}.}$$

For the upper bound, let $\mathbb{F}_q = \mathbb{Z}[\zeta]/\wp$ and prove that $\zeta + a \mod \wp \in \mathcal{G}$ are pairwise distinct in $\mathbb{F}_q$, for $1 \leq a \leq \ell$. Together with the group structure and the definition of ζ, this leads to the **lower bound**:

$$\boxed{\mid \mathcal{G} \mid \geq \binom{t+\ell}{\ell-1}.}$$

The two bounds are contradictory, so n must be a prime power.

The group with generators $\mathcal{G}$ replaces the cycle of a root of unity which was used in all previous, essentially Pocklingotn based tests.

Berrizbeitia: Uses Kummer extensions and their Galois theory and drops the condition of a deterministic test. Obtains a variant which is faster then AKS by a factor of $(\log n)^2$.

Theorem (Berrizbeitia, M.)

Let $m > \log^2(n)$ and $\mathbf{A} \supset \mathbb{Z}/(n \cdot \mathbb{Z})$ an algebra with some $\zeta \in \mathbf{A}$, $\Phi_m(\zeta) = 0$, where $\Phi_m(x) \in \mathbb{Z}[x]$ is the $m-\mathrm{th}$ cyclotomic polynomial. Let $\mathbf{R} = \mathbf{A}[X]/\left(X^m - \zeta\right)$ and $\xi \in \mathbf{R}$ be the image of X in $\mathbf{R}$. If

$$1 + \xi^n = (1 + \xi)^n,$$

then n is a prime power.

Certificates for CPP

- In (1) we have identities $\alpha^{(n^d-1)/p^r} = \zeta_{p^r}^m$ in some algebra $\mathbf{A}$. Let $E = (n^d - 1)/p^r$ and $m \equiv E \cdot u \bmod p^r$ (assumption on r required!). Then

$$\left(\alpha\zeta^{-u}\right)^E = 1,$$

- If n is prime, then there exists a $\beta \in \mathbf{A}$ with $\beta^{p^r} = \alpha\zeta^{-u}$.
- This leads to the certificate idea: attempt to compute β; if computation fails, then n is composite. Otherwise $\beta \in \mathbf{A}$ certifies the test of (1) for α.
- One proves explicitly that if $\beta \in \mathbf{A}$ verifies its defining identity, then the tests (1) are correct, so the central part of the cyclotomy test is verified.

- The resulting certificate is verified in time $O(\log(n))$ faster than it was obtained. It is conceptually an extension of the Pratt certificates to the setting of CPP.
- The certification method has been implemented, works – requires rather large certificates. Not a problem with modern computer in the realm of up to one million decimal digits, say.

CIDE - A combination of CPP and ECPP

- **CIDE**: Cyclotomy improved with *dual elliptic* primes. A variant using elliptic curves.
- Two primes p, q are *dual elliptic*, if there is an ordinary elliptic curve over $\mathbb{F}_p$ which has q points. Then there also exists an elliptic curve over $\mathbb{F}_q$ with p points!
- CIDE uses integers which have some related property, without being certified primes.
- The test is random polynomial with run time (heuristically) $O\left(\log(n)^{3+\varepsilon}\right)$.

CIDE - Main Lemmata

Lemma

Two integers m, n are dual elliptic, if there is an imaginary quadratic field $\mathbb{K} = \mathbb{Q}[\sqrt{-d}]$ in which both split in principal ideals, and $m = \mu \cdot \overline{\mu}, n = \nu \cdot \overline{\nu}$ with $\nu = \mu \pm 1$. Then m, n are simultaneoulsy prime or composite. In the second case, there are prime factors $p|m, q|n$, which are dual elliptic primes. Moreover $|p - q| \leq 2 \cdot \sqrt[4]{\max(m, n)}$.

CIDE - A definition

Definition

Suppose that ℓ is a prime, $\mathcal{E}: Y^2 = X^3 + aX + b$ an elliptic curve and $f(X)$ is a divisor of the $\ell-\mathrm{th}$ division polynomial of $\mathcal{E}$ which has a zero modulo n. Let

$$P = (X + (f(X), n), Y + (Y^2 - (X^3 + aX + b)))$$

and $\tau(\chi) = \sum_{k=1}^{\ell-1} \chi(k)[kP]_x$. We say n allows an ℓ-th elliptic extension for $\mathcal{E}$, iff $\tau(\chi)^n = \chi^{-n}(\lambda) \cdot \tau(\chi^n) \bmod (n, f(X))$.

CIDE - Main Theorem

Theorem

Let m, n be dual elliptic pseudoprimes and suppose that $s-\mathrm{th}$ cyclotomic extensions $\mathfrak{M}, \mathfrak{N}$ exist for both m, n and $s \geq 2\max(m^{1/4}, n^{1/4})$ (CPP - tests!). Let $\mu \cdot \overline{\mu} = m$; $\nu \cdot \overline{\nu} = n$ be the decomposition in $\mathbb{K} = \mathbb{Q}[\sqrt{-d}]$. Let L be a square free integer all the prime factors of which split in $\mathbb{K}$ and suppose that there is an elliptic curve $\mathcal{E}$ together with an L-th elliptic extension for $\mathcal{E}$ with respect to both m and n. Then there are two integers k, k' such that

$$(\mu + 1)^{k'} - \mu^k \equiv \pm 1 \bmod L\mathcal{O}(\mathbb{K}). \tag{3}$$

CIDE - Algorithm

- For given n find a dual m, with some preprocessing step of ECPP. Let $\mathbb{K}$ be the imaginary quadratic extension in which the two split, so that $\mu = \nu \pm 1$, as above. Let $\mathcal{E}, \mathcal{E}'$ be corresponding CM curves.
- Choose the paramters s, t for the cyclotomic extensions $\mathfrak{M}, \mathfrak{N}$ and prove their existence.
- Find an integer L for which the identity (3) has no solution (combinatorial problem, $L = O(\log\log(n))$).
- Perform the elliptic Gauss sum verifications for all primes $\ell | L$.
- If all these steps are performed successfully, declare n, m primes. Otherwise either no decision or composite (simultaneously).
- Extend the certificates for $\mathfrak{N}, \mathfrak{M}$ by some for the elliptic Gauss sums.

The computations of Jens Franke et. al.

We have confirmed the primality of the Leyland numbers $3110^{63} + 63^{3110}$ (5596 digits) and $8656^{2929} + 2929^{8656}$ (**30008** digits) by an implementation of a version of Mihăilescu's CIDE. The certificates may be found at
http://www.math.uni-bonn.de:people/franke/ptest/x3110y63.cert.tar.bz2
and
http://www.math.uni-bonn.de:people/franke/ptest/x8656y2929.cert.tar.bz2
.

Damgard I; Landrock, P; Pomerance, C.: "Average Case Bounds for the Strong Probable Prime Test", Math. Comp. **61**, no.203, pp.177-194.