An algorithm for solving

\[ \sum_{i=1}^{n} a_i x_i^n = b \]

over finite fields

Christiaan van de Woestijne, Universiteit Leiden

Oberwolfach workshop on Finite Fields
7 December 2004
Surroundings

Currently known algorithms for solving equations over finite fields include:

- brute force search
- algorithms for factoring polynomials
- Shanks’ algorithm for taking square (and higher) roots
- methods for multivariate equations based on the above
- Schoof’s algorithm for taking square roots in prime fields

However, all of these are either probabilistic (barring a proof of GRH for some) or take more than polynomial time.
Overview: a tower of algorithms

(This is part of my PhD project with H. W. Lenstra, Jr.)

I. Computing field generators in multiplicative subgroups:

for $G \subseteq \mathbb{F}^*$, find $\alpha \in G$ such that $\mathbb{F} = \mathbb{F}_p(\alpha)$.

II. Writing field elements as sums of like powers:

given $b \in \mathbb{F}^*$, find $x_1, \ldots, x_n \in \mathbb{F}$ such that $b = \sum_{i=1}^{n} x_i^n$.

III. Finding representations by diagonal forms in many variables:

given $a_1, \ldots, a_n \in \mathbb{F}^*$, and $b \in \mathbb{F}^*$, find $x_1, \ldots, x_n \in \mathbb{F}$ such that

$b = \sum_{i=1}^{n} a_i x_i^n$. 
Overview: building blocks

I. A multiplicative version of the primitive element theorem (using elementary linear algebra)

II. Selective root extraction (a generalisation of the Tonelli-Shanks algorithm)

III. Reducing the number of terms in a sum of like powers (a bisection-like idea)

IV. Dealing with coefficients other than 1 by means of the “trapezium algorithm” (an algorithmic version of an idea of Dem’yanov and Kneser)
It can be shown that...

- the set of sums of $n$th powers of elements, $S_n$, in $\mathbb{F}$ is a subfield of $\mathbb{F}$.
- $S_n = \mathbb{F}$ iff $\mathbb{F}$ can be generated over $\mathbb{F}_p$ by an $n$th power in $\mathbb{F}$.
- if $S_n \neq \mathbb{F}$, we have $n^2 > q$.
- if $S_n = \mathbb{F}$, then every equation of the form
  \[ \sum_{i=1}^{n} a_i x_i^n = b \]
  for $a_1, \ldots, a_n$ and $b$ in $\mathbb{F}^*$ is solvable.

The homogeneous variant $\sum_{i=0}^{n} a_i x_i^n = 0$ is always solvable by the Chevalley-Warning theorem.
By comparison...

- the results from the last slide can be much improved if \( q \) is much larger than \( n^2 \). For example, if \( q > n^4 \), then every equation of the form

\[
ax^n + by^n = c
\]

is solvable (Weil 1948).

- the algorithms I will present are not unpractical but probabilistic algorithms will probably do better if \( q \) is much larger than \( n \).
Conventions

In this talk, the phrase “we can compute $X$” means:

“we know explicitly a deterministic polynomial time algorithm to compute $X$”.

The same goes for “we can decide $Y$”.

We will denote by $\mathbb{F}$ a finite field of $q$ elements and characteristic $p$, given by a polynomial $f$ that is irreducible over the prime field $\mathbb{F}_p$.

Our algorithms take $\mathbb{F}$ as input; thus the input size is about $\log q$, and our algorithms must finish in time polynomial in $\log q$. 
**Algorithm I: a generator in a given subgroup (1)**

**Theorem.** Let $G \subseteq \mathbb{F}^*$ be a multiplicative subgroup; we can compute $\beta \in G$ such that $\beta$ generates $\mathbb{F}$ over its prime field, or decide that no such $\alpha$ exists.

Main (in fact only) example: $G = \mathbb{F}^n$ for some positive integer $n$.

**Proof.** Let $n = [\mathbb{F}^* : G]$ and let $\alpha$ be the given generator of $\mathbb{F}$.

If $K_1 = \mathbb{F}_p(\gamma_1^n)$ and $K_2 = \mathbb{F}_p(\gamma_2^n)$ are subfields of $\mathbb{F}$, we can compute $\gamma \in \langle \gamma_1, \gamma_2 \rangle$ such that

$$\gamma^n \text{ generates } \mathbb{F}_p(\gamma_1^n, \gamma_2^n) \text{ over } \mathbb{F}_p$$

by means of a elementary linear algebra.
Building block I: A “multiplicative” primitive element theorem

Lemma. Let $L/K$ be a cyclic extension of fields of degree $d$, and let $b_1,\ldots,b_d$ be a $K$-basis for $L$. Then at least $\varphi(d)$ of the $b_i$ generate $L$ as a field over $K$.

Now suppose $\alpha \in L$ has degree $e$ over $K$ and $\beta$ has degree $f$. The degree of $\beta$ over $K(\alpha)$ is given by $g = \text{lcm}(e,f)/e = f/g\text{d}(e,f)$, so a basis of $K(\alpha,\beta)$ is given by

$$(\alpha^i \beta^j \mid i = 0,\ldots,e-1, j = 0,\ldots,g-1).$$

One of these elements generates $K(\alpha,\beta)$ over $K$!

Obviously, by induction we may extend this result to systems of more than two generators.
Algorithm I: a generator in a given subgroup (2)

Proof (ctd.) We start induction with \( K = \mathbb{F}_p = \mathbb{F}_p(1^n) \). Assume now we have \( K = \mathbb{F}_p(\gamma_1^n) \). If \( |K| \leq n \), we find \( \gamma_2 \in \mathbb{F}^* \) with \( \gamma_2^n \notin K \).

If no such \( \gamma_2 \) exists, the algorithm fails (and rightly so)!

If \( |K| > n \), then at least one of \( (\alpha + c_i)^n \), where \( c_0, \ldots, c_n \) are distinct elements of \( K \), is not in \( K \); now put \( \gamma_2 = \alpha + c_i \). (Recall that \( \mathbb{F} = \mathbb{F}_p(\alpha) \).

Now in either case, adjoin \( \gamma_2^n \) to \( K \) and compute \( \gamma \) with \( K = \mathbb{F}_p(\gamma^n) \), using Building block I. \( \square \)
Building block II: selective root extraction

**Theorem.** If $a_0, a_1, \ldots, a_n$ are in $\mathbb{F}^*$, then we can compute some $\beta \in \mathbb{F}^*$ such that, for some $i, j$ with $0 \leq i < j \leq n$, we have

$$a_i/a_j = \beta^n.$$  

**Proof.** Let $H = \langle a_0, \ldots, a_n \rangle$. The $a_i$ cover the cosets of $H$ modulo $H^n$, so there exist $i$ and $j$ such that $a_i/a_j \in H^n$.

We can factor $n$ into primes $\ell$ and use this to compute generators $\gamma_\ell$ for the $\ell$-parts of $H$. Now, we compute an $n$th root $\beta$ of $a_i/a_j$ using these generators $\gamma_\ell$, by means of the Tonelli-Shanks algorithm. \qed
Algorithm II: sums of like powers

Theorem. Let $b$ be in $\mathbb{F}^*$ and $n$ a positive integer. We can decide if $b$ is in $S_n$ and if so, we can compute $x_1, \ldots, x_n$ such that $b = \sum_{i=1}^{n} x_i^n$.

Proof. If $n^2 \geq q$, we have enough time to enumerate all possibilities.

If $n^2 < q$, then $S_n = \mathbb{F}$, so the answer is yes. We use Algorithm I to compute $\gamma \in \mathbb{F}$ such that $\gamma^n$ generates $\mathbb{F}$ over $\mathbb{F}_p$; this gives us

$$b = \sum_{i=0}^{[\mathbb{F} : \mathbb{F}_p] - 1} b_i \gamma^{ni}.$$ 

This is a sum of $n$th powers with at most $(p - 1) \cdot [\mathbb{F} : \mathbb{F}_p]$ terms!

Now use Building blocks II and III to come down to just $n$ terms. \qed
Building block III: reducing sums of like powers

**Theorem.** Given $y_1, \ldots, y_N$ and $b \in \mathbb{F}^*$ with $\sum y_i^n = b$, we can compute $x_1, \ldots, x_n \in \mathbb{F}^*$ such that $\sum_{i=1}^n x_i^n = b$.

**Proof.** Divide $y_1, \ldots, y_N$ into $n+1$ roughly equal groups $G_0, \ldots, G_n$. Let $S_i$ denote the sum of all terms in the first $i+1$ groups.

If one of the $S_i$ is zero, we discard all terms in the first $i+1$ groups. Otherwise, we use **selective root extraction** to compute $\beta \in \mathbb{F}^*$ with

$$S_i / S_j = \beta^n.$$  

(assume $i > j$). This means we can **discard the groups** $G_{j+1}$ up to $G_i$, provided we multiply all terms in the first $i+1$ groups by $\beta$. This trick is applicable as long as we have at least $n+1$ terms. $\square$
Algorithm III: representations by diagonal forms

**Theorem.** Let $b$ be in $\mathbb{F}^*$ and $n$ a positive integer. For any $a_1, \ldots, a_n \in \mathbb{F}^*$ we can decide if the equation

$$b = \sum_{i=1}^{n} a_i x_i^n$$

is solvable, and if so, we can compute a solution.

**Proof.** Again, if $n^2 \geq q$, we can just enumerate all possibilities.

If $n^2 < q$, there is a solution. Write $a_0 = -b$. We use now Algorithm II to write the elements $b/a_i$ (for $i = 1, \ldots, n$) as sums of $n$th powers, so we get

$$-a_i \sum_j y_{ij}^n = -b = a_0 \cdot 1^n.$$
Building block IV: the trapezium algorithm (1)

We now have a system of the form

\[
\begin{align*}
-a_0(y_{0,1}^n + \cdots + y_{0,n}^0) &= 0 \\
-a_1(y_{1,1}^n + \cdots + y_{1,n}^1) &= a_0x_{1,0}^n \\
& \vdots \\
-a_n(y_{n,1}^n + \cdots + y_{n,n}^n) &= a_0x_{n,0}^n + \cdots + a_{n-1}x_{n,n-1}^n
\end{align*}
\]

Recall that we wrote \(a_0 = -b\). If \(h_i = 0\) for some \(i \geq 1\), we are done!

We try to lower the \(h_i\) by bringing the last term \(a_i y_{i,h_i}^n\) to the other side. We get the sequence

\[
(a_0 y_{0,h_0}^n, a_0 x_{1,0}^n + a_1 y_{1,h_1}^n, \ldots, a_0 x_{n,0}^n + \cdots + a_{n-1} x_{n,n-1}^n + a_n y_{n,h_n}^n).
\]
Building block IV: the trapezium algorithm (2)

The sequence

\[
(a_0 y_0, h_0, a_0 x_1, 0 + a_1 y_1, h_1, \ldots, a_0 x_n, 0 + \ldots + a_{n-1} x_n, n-1 + a_n y_n, h_n).
\]

has \(n + 1\) elements, say \(c_0, \ldots, c_n\). If one is zero, we are done!

Otherwise, use selective root extraction to compute \(\beta \in \mathbb{F}^*\) with

\[
\beta^n = c_i / c_j, \quad \text{i.e.} \quad c_i = \beta^n c_j
\]

(assume \(i > j\)).

Replace now the \(i\)th term in the sequence by \(\beta^n\) times the \(j\)th term, and we can reduce \(h_i\) by one!

Thus, in at most \(n^2\) steps, we will get one of the \(h_i\) down to zero. □
Applications (for $n = 2$)

If $n = 2$ and the characteristic of $\mathbb{F}$ is odd, then every form is diagonal. Furthermore, in characteristic 2, zeros of quadratic forms can be found by means of linear algebra.

**Corollary.** Given a quadric hypersurface over a finite field $\mathbb{F}$, we can compute a rational point on it.

**Corollary.** Given two regular quadratic spaces $V$ and $W$ over a finite field $\mathbb{F}$ (char. $\neq 2$), such that $\dim V \geq \dim W + 1$, we can compute an isometric embedding of $W$ into $V$.

On the other hand, if $\dim V = \dim W$, we can reduce the problem of finding an isometry from $V$ to $W$ to the computation of just one square root in $\mathbb{F}$. 
More applications (for $n = 2$)

**Corollary.** (Bumby) Given a prime $p$, we can compute integers $x_1, \ldots, x_4$ such that $p = x^2 + y^2 + z^2 + w^2$. This works also for any other Euclidean quaternion orders.

**Corollary.** Given a central simple algebra $A$ of degree 2 over a finite field $\mathbb{F}$, we can compute an explicit isomorphism from $A$ to a $2 \times 2$-matrix algebra over $\mathbb{F}$. 