Optimizing double-base elliptic curve single-scalar multiplication

(Joint work with Daniel J. Bernstein, Peter Birkner, Tanja Lange)

Christiane Peters

Technische Universiteit Eindhoven

DIAMANT/EIDMA Symposium
November 29, 2007

Motivation: Elliptic Curve Cryptography

Given a finite field $\mathbb{F}_p$, an elliptic curve E over $\mathbb{F}_p$ and a point P of high order on E,
there is the Elliptic Curve Diffie-Hellman key exchange

- Alice and Bob choose $a, b \in \mathbb{F}_p$ privately,
- exchange $[a]P$ and $[b]P$
- and agree on a public key $[ab]P = [a]([b]P) = [b]([a]P)$.

Question: How do Alice and Bob perform all those scalar multiplications **efficiently**?

Speed-up techniques for elliptic-curve single-scalar multiplication

- choose different curve shapes (e.g. Edwards curves, Weierstrass form)
- choose different coordinate systems (e.g. inverted Edwards coordinates, Jacobian coordinates)
- use sliding-window methods
- use double-base chains

1. Edwards curves

2. Other curve shapes and coordinate systems

3. Double-base number systems

4. Experiments and results

Definition

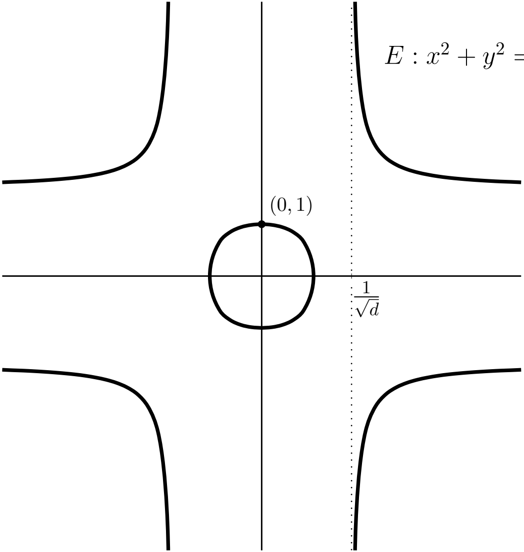
An elliptic curve E in Edwards form over $\mathbb{F}_p$ where $p \geq 3$ is given by the equation

$$x^2 + y^2 = 1 + d\,x^2y^2,$$

where $d \in \mathbb{F}_p \setminus \{0, 1\}$.

From now on we will call a curve in this shape an Edwards curve.

That's the way it looks over $\mathbb{R}$



$$E : x^2 + y^2 = 1 + d\, x^2 y^2$$

$$0 < d < 1$$

Addition on Edwards curves

We add two points (x_1, y_1), (x_2, y_2) on E according to the Edwards addition law

$$(x_1, y_1), (x_2, y_2) \mapsto \left(\frac{x_1 y_2 + x_2 y_1}{1 + d x_1 x_2 y_1 y_2}, \frac{y_1 y_2 - x_1 x_2}{1 - d x_1 x_2 y_1 y_2} \right).$$

- The addition law is strongly unified, i.e., it can be also used for doublings.
- If d is a non-square in $\mathbb{F}_p$ the addition law is complete.
- The point $(0, 1)$ is the neutral element of the addition law and
- the negative of $P = (x_1, y_1)$ is $-P = (-x_1, y_1)$.

Explicit fast doubling and tripling formulas

Doubling of a point (x_1, y_1) on $x^2 + y^2 = 1 + dx^2y^2$:

$$\begin{aligned}[2](x_1, y_1) &= \left(\frac{2x_1y_1}{1 + dx_1^2y_1^2}, \frac{y_1^2 - x_1^2}{1 - dx_1^2y_1^2}\right) \\ &= \left(\frac{2x_1y_1}{x_1^2 + y_1^2}, \frac{y_1^2 - x_1^2}{2 - (x_1^2 + y_1^2)}\right).\end{aligned}$$

Tripling:

$$[3](x_1, y_1) = \\ \left(\frac{((x_1^2 + y_1^2)^2 - (2y_1)^2)}{4(x_1^2 - 1)x_1^2 - (x_1^2 - y_1^2)^2}x_1, \frac{((x_1^2 + y_1^2)^2 - (2x_1)^2)}{-4(y_1^2 - 1)y_1^2 + (x_1^2 - y_1^2)^2}y_1\right).$$

Avoiding inversions

To avoid inversions we consider the homogenized Edwards equation

$$E:\ (X^2+Y^2)Z^2 = (Z^4 + dX^2Y^2)$$

A point $(X_1 : Y_1 : Z_1)$ with $Z_1 \neq 0$ on E corresponds to the affine point $(X_1/Z_1, Y_1/Z_1)$.

Bernstein/Lange (2007): Inverted Edwards coordinates

A point $(X_1 : Y_1 : Z_1)$ on

$$(X_1^2+Y_1^2)Z_1^2 = X_1^2Y_1^2 + dZ_1^4$$

where $X_1Y_1Z_1 \neq 0$ corresponds to $(Z_1/X_1, Z_1/Y_1)$ on the Edwards curve $x_1^2 + y_1^2 = 1 + dx_1^2y_1^2$.

Avoiding inversions

To avoid inversions we consider the homogenized Edwards equation

$$E:\ (X^2+Y^2)Z^2 = (Z^4 + dX^2Y^2)$$

A point $(X_1 : Y_1 : Z_1)$ with $Z_1 \neq 0$ on E corresponds to the affine point $(X_1/Z_1, Y_1/Z_1)$.

Bernstein/Lange (2007): Inverted Edwards coordinates

A point $(X_1 : Y_1 : Z_1)$ on

$$(X_1^2+Y_1^2)Z_1^2 = X_1^2Y_1^2 + dZ_1^4$$

where $X_1Y_1Z_1 \neq 0$ corresponds to $(Z_1/X_1, Z_1/Y_1)$ on the Edwards curve $x_1^2 + y_1^2 = 1 + dx_1^2y_1^2$.

1. Edwards curves

2. Other curve shapes and coordinate systems

3. Double-base number systems

4. Experiments and results

Weierstrass form over $\mathbb{F}_p$ $(p \geq 5)$

Short Weierstrass form $\quad E: y^2 = x^3 + a_4 x + a_6$
with $a_4, a_6 \in \mathbb{F}_p$, and $4a_4^3 + 27a_6^2 \neq 0$.

Jacobian coordinates: $(X_1 : Y_1 : Z_1)$ satisfying

$$Y_1^2 = X_1^3 + a_4 X_1 Z_1^2 + a_6 Z_1^6$$

corresponds to $(x_1, y_1) = (X_1/Z_1^2, Y_1/Z_1^3)$ on E.

The choice $a_4 = -3$ leads to the fastest arithmetic for curves in Jacobian coordinates.

More coordinate systems

- Jacobi quartics $Y^2 = X^4 + 2aX^2Z^2 + Z^4$,
- Hessian curves $X^3 + Y^3 + Z^3 = 3dXYZ$,
- Jacobi intersections $S^2 + C^2 = Z^2, aS^2 + D^2 = Z^2$,
- "tripling-oriented Doche/Icart/Kohel curves"
  $Y^2 = X^3 + a(X + Z^2)^2Z^2$.

Comparison

$\mathbf{M}$: general multiplications, $\mathbf{S}$: squarings, $\mathbf{D}$ multiplications by d

Curve shape	ADD	mADD	DBL	TRI
3DIK	$11\mathbf{M}+6\mathbf{S}$	$7\mathbf{M}+4\mathbf{S}$	$2\mathbf{M}+7\mathbf{S}$	$6\mathbf{M}+6\mathbf{S}$
Edwards	$10\mathbf{M}+1\mathbf{S}$	$9\mathbf{M}+1\mathbf{S}$	$3\mathbf{M}+4\mathbf{S}$	$9\mathbf{M}+4\mathbf{S}$
ExtJQuartic	$8\mathbf{M}+3\mathbf{S}$	$7\mathbf{M}+3\mathbf{S}$	$3\mathbf{M}+4\mathbf{S}$	$4\mathbf{M}+11\mathbf{S}$
Hessian	$12\mathbf{M}+0\mathbf{S}$	$10\mathbf{M}+0\mathbf{S}$	$7\mathbf{M}+1\mathbf{S}$	$8\mathbf{M}+6\mathbf{S}$
InvEdwards	$9\mathbf{M}+1\mathbf{S}$	$8\mathbf{M}+1\mathbf{S}$	$3\mathbf{M}+4\mathbf{S}$	$9\mathbf{M}+4\mathbf{S}$
JacIntersect	$13\mathbf{M}+2\mathbf{S}$	$11\mathbf{M}+2\mathbf{S}$	$3\mathbf{M}+4\mathbf{S}$	$4\mathbf{M}+10\mathbf{S}$
Jacobian	$11\mathbf{M}+5\mathbf{S}$	$7\mathbf{M}+4\mathbf{S}$	$1\mathbf{M}+8\mathbf{S}$	$5\mathbf{M}+10\mathbf{S}$
Jacobian-3	$11\mathbf{M}+5\mathbf{S}$	$7\mathbf{M}+4\mathbf{S}$	$3\mathbf{M}+5\mathbf{S}$	$7\mathbf{M}+7\mathbf{S}$
Std-Jac	$12\mathbf{M}+4\mathbf{S}$	$8\mathbf{M}+3\mathbf{S}$	$3\mathbf{M}+6\mathbf{S}$	$9\mathbf{M}+6\mathbf{S}$
Std-Jac-3	$12\mathbf{M}+4\mathbf{S}$	$8\mathbf{M}+3\mathbf{S}$	$4\mathbf{M}+4\mathbf{S}$	$9\mathbf{M}+6\mathbf{S}$

For details consider the Explicit-formulas database.
`http://www.hyperelliptic.org/EFD`.

1. Edwards curves

2. Other curve shapes and coordinate systems

3. Double-base number systems

4. Experiments and results

Given $n \in \mathbb{Z}$, compute $[n]P$

single-base: e.g. “signed double-and-add”: $n = \sum_{i \geq 1} c_i 2^i$ with $c_i \in \{0, \pm 1\}$

Example:

$$\begin{aligned} & 314159P \\ &= 2^{18}P + 2^{16}P - 2^{14}P + 2^{11}P + 2^{10}P - 2^{8}P + 2^{5}P + 2^{4}P - 2^{0}P \\ &= 2(2(2(2(2(2(2(2(2(2(2(2(2(2(2(2(2(2(P))+P))-P)))+P)+P)) \\ &\qquad -P)))+P)+P))))-P \end{aligned}$$

Double-bases: base $\{2, 3\}$

We express $n \in \mathbb{Z}$ as $\sum_{i \geq 1} c_i 2^{a_i} 3^{b_i}$ with e.g. $c_i = \pm 1$,

i.e. we express the point $[n]P$ as a sum of few points $[c_i 2^{a_i} 3^{b_i}]P$.

Dimitrov, Imbert and Mishra at Asiacrypt 2005:

$$a_1 \geq a_2 \geq a_3 \geq \ldots, \qquad b_1 \geq b_2 \geq b_3 \geq \ldots$$

$\Rightarrow$ Horner-like evaluation: only a_1 doublings and b_1 triplings needed.

$$\begin{aligned} 314159P &= 2^{15}3^2P + 2^{11}3^2P + 2^8 3^1 P + 2^4 3^1 P - 2^0 3^0 P \\ &= 3(2(2(2(2(2(2(2(2(3(2(2(2(2(2(2(2(P)))) + P)))) + P)))) + P))))) - P \end{aligned}$$

Double-bases: base $\{2, 3\}$

We express $n \in \mathbb{Z}$ as $\sum_{i \geq 1} c_i 2^{a_i} 3^{b_i}$ with e.g. $c_i = \pm 1$,

i.e. we express the point $[n]P$ as a sum of few points $[c_i 2^{a_i} 3^{b_i}]P$.

Dimitrov, Imbert and Mishra at Asiacrypt 2005:

$$a_1 \geq a_2 \geq a_3 \geq \ldots, \qquad b_1 \geq b_2 \geq b_3 \geq \ldots$$

$\Rightarrow$ Horner-like evaluation: only a_1 doublings and b_1 triplings needed.

$$\begin{aligned} 314159P &= 2^{15}3^2P + 2^{11}3^2P + 2^8 3^1 P + 2^4 3^1 P - 2^0 3^0 P \\ &= 3(2(2(2(2(2(2(2(2(3(2(2(2(2(2(2(2(P)))) + P)))) + P)))) + P))))) - P \end{aligned}$$

Expansion of the coefficient set

Doche and Imbert at Indocrypt 2006: additionally to $a_1 \geq a_2 \geq \dots$, $b_1 \geq b_2 \geq \dots$ in $n = \sum_{i \geq 1} c_i 2^{a_i} 3^{b_i}$ choose $c_i, -c_i$ from one of the sets

$$\{1\}, \{1, 2, 2^2, 3, 3^2\}, \dots, \{1, 2, \dots, 2^4, 3, \dots, 3^4\},$$
$$\{1, 5, 7\}, \dots \{1, 5, 7, 11, 13, 17, 19, 23, 25\}.$$

"Sliding-windows double-base-2-and-3":

$$\begin{aligned} 314159P &= 2^{12}3^3 3P - 2^7 3^3 5P - 2^4 3^1 7P - 2^0 3^0 P \\ &= 3(2(2(2(2(3(3(2(2(2(2(2(2(2(2(3P))))) - 5P))))) - 7P))))) - P \end{aligned}$$

2007

Bernstein/Birkner/Lange/P. 2007:

- more coordinate systems,
- inversion-free precomputations,
- new faster formulas for arithmetic for different coordinate systems,
- larger variety of coefficient sets S:

$$\{1\},\ \{1,2,3\},\{1,2,3,4,9\},\ldots\{1,2,3,4,8,9,16,27,81\},$$
$$\{1,5\},\{1,5,7\},\ldots,\{1,5,7,11,13,17,19,23,25\},$$
$$\{1,2,3,5\},\ \{1,2,3,5,7\},\ldots\{1,2,3,5,7,9,11,13,15,17,19,23,25\}$$

1. Edwards curves

2. Other curve shapes and coordinate systems

3. Double-base number systems

4. Experiments and results

Finding the chain $n = \sum_{i \geq 1} c_i 2^{a_i} 3^{b_i}$

Generalize how to find Thurber's base-2 sliding-window chain $\sum_i c_i 2^{a_i}$ with $\pm c_i \in \{1, 3, 5, 7\}$ and $a_1 > a_2 > a_3 > \ldots$:

Check which of the first bits of

$$\begin{array}{llllll}
\pm 1, & \pm 2, & \pm 2^2, & \pm 2^3, & \pm 2^4, & \ldots \\
\pm 3, & \pm 2 \cdot 3, & \pm 2^2 3, & \pm 2^3 3, & \pm 2^4 3, & \ldots \\
\pm 5, & \pm 2 \cdot 5, & \pm 2^2 5, & \pm 2^3 5, & \pm 2^4 5, & \ldots \\
\pm 7, & \pm 2 \cdot 7, & \pm 2^2 7, & \pm 2^3 7, & \pm 2^4 7, & \ldots
\end{array}$$

is closest to n.

Vary maximal power of 2 and 3 in the representation.

Upper bounds $a_0 \geq a_1$, $b_0 \geq b_1$: For an ℓ-bit number n, we choose $0 \leq a_0 \leq \ell$, $b_0 = \lceil (\ell - a_0)/\lg 3 \rceil$.

Optimal parameters for each curve shape for $\ell = 200$

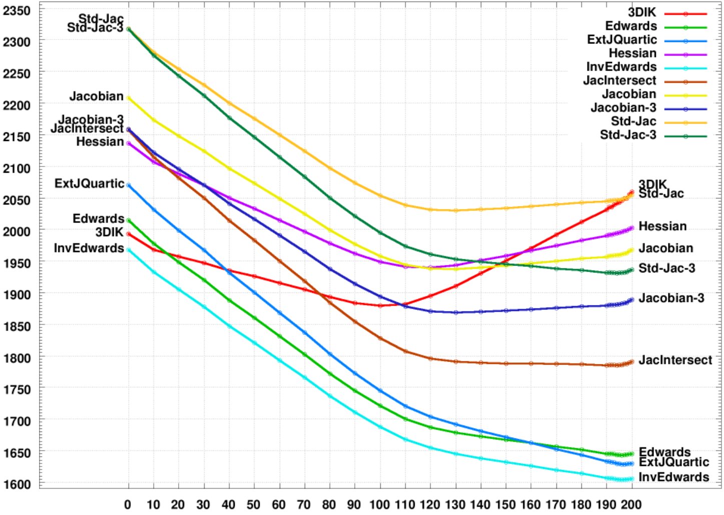
We assume $1\mathbf{S} = 0.8\mathbf{M}$ and $\mathbf{D} = 0M$.

Curve shape	Mults	a_0	a_0/ℓ	S
3DIK	1879.200960	100	0.5	$\{1, 2, 3, 5, 7\}$
Edwards	1642.867360	196	0.98	$\{1, 2, 3, 5, \ldots, 15\}$
ExtJQuartic	1628.386660	196	0.98	$\{1, 2, 3, 5, \ldots, 15\}$
Hessian	1939.682780	120	0.6	$\{1, 2, 3, 5, \ldots, 13\}$
InvEdwards	1603.737760	196	0.98	$\{1, 2, 3, 5, \ldots, 15\}$
JacIntersect	1784.742	190	0.95	$\{1, 2, 3, 5, \ldots, 15\}$
Jacobian	1937.129960	130	0.65	$\{1, 2, 3, 5, \ldots, 13\}$
Jacobian-3	1868.530560	130	0.65	$\{1, 2, 3, 5, \ldots, 13\}$

We got similar results for $\ell = 160, 256, 300, 400, 500$.

$\mathbf{M}$: general multiplications, $\mathbf{S}$: squarings, $\mathbf{D}$ multiplications by d

Choice of a_0 for $\ell = 200$



Conclusions

- For curves in Jacobian coordinates, tripling-oriented Doche/Icart/Kohel curves, Hessian curves we recommend using double-bases
- for Edwards curves, Jacobi intersections, extended Jacobi-quartic coordinates, and inverted Edwards coordinates we recommend single-bases

The latter use

- larger sets of precomputations,
- and fewer triplings,
- fast addition laws (precomputations less costly)
- and in particular very fast doublings.