Complete addition laws for elliptic curves

D. J. Bernstein University of Illinois at Chicago

Tanja Lange

Technische Universiteit Eindhoven

<u>Weierstrass coordinates</u>

Fix a field k with $2 \neq 0$. Fix $a, b \in k$ with $4a^3 + 27b^2 \neq 0$. Well-known fact: The points of the "elliptic curve" $E: y^2 = x^3 + ax + b$ over k form a commutative group E(k).

"So the group is $\{(x, y) \in k imes k : y^2 = x^3 + ax + b\}$?"

Not exactly! It's $\{(x, y) \in k imes k : y^2 = x^3 + ax + b\} \cup \{\infty\}.$

To add (x_1, y_1) , $(x_2, y_2) \in E(k)$: Define $x_3 = \lambda^2 - x_1 - x_2$ and $y_3 = \lambda(x_1 - x_3) - y_1$ where $\lambda = (y_2 - y_1)/(x_2 - x_1)$. Then $(x_3, y_3) \in E(k)$.

Geometric interpretation: $(x_1, y_1), (x_2, y_2), (x_3, -y_3)$ are on the curve $y^2 = x^3 + ax + b$ and on a line; $(x_3, y_3), (x_3, -y_3)$ are on a vertical line.

"So that's the group law? $(x_1, y_1) + (x_2, y_2) = (x_3, y_3)$?"

Not exactly! Definition of λ assumes that $x_2
eq x_1$.

To add (x_1, y_1) , $(x_1, y_1) \in E(k)$: Define $x_3 = \lambda^2 - x_1 - x_2$ and $y_3 = \lambda(x_1 - x_3) - y_1$ where $\lambda = (3x_1^2 + a)/2y_1$. Then $(x_3, y_3) \in E(k)$.

Geometric interpretation: The curve's tangent line at (x_1, y_1) passes through $(x_3, -y_3)$. "So that's the group law? One special case for doubling?" Not exactly! More exceptions: e.g., y_1 could be 0.

Six cases overall: $\infty + \infty = \infty$: $\infty + (x_2, y_2) = (x_2, y_2);$ $(x_1, y_1) + \infty = (x_1, y_1);$ $(x_1, y_1) + (x_1, -y_1) = \infty;$ for $y_1
eq 0$, $(x_1, y_1) + (x_1, y_1) =$ (x_3,y_3) with $x_3=\lambda^2-x_1-x_2$, $y_3=\lambda(x_1-x_3)-y_1$, $\lambda = (3x_1^2 + a)/2y_1;$ for $x_1
eq x_2$, $(x_1, y_1) + (x_2, y_2) =$ (x_3, y_3) with $x_3 = \lambda^2 - x_1 - x_2$, $y_3=\lambda(x_1-x_3)-y_1$, $\lambda = (y_2 - y_1)/(x_2 - x_1).$

E(k) is a commutative group: Has neutral element ∞ , and -: $-\infty = \infty; -(x, y) = (x, -y).$ Commutativity: P + Q = Q + P. Associativity: (P+Q) + R = P + (Q+R).Straightforward but tedious: use a computer-algebra system to check each possible case. Or relate each P + Q case to "ideal-class product." Many other proofs, but can't escape case analysis.

Projective coordinates

Can eliminate some exceptions.

Define (X : Y : Z), for $(X, Y, Z) \in k \times k \times k - \{(0, 0, 0)\},\$ as $\{(rX, rY, rZ) : r \in k - \{0\}\}.$

Could split into cases: (X : Y : Z) = (X/Z : Y/Z : 1) if $Z \neq 0$; (X : Y : 0) = (X/Y : 1 : 0) if $Y \neq 0$; (X : 0 : 0) = (1 : 0 : 0). But scaling unifies all cases. Write $\mathbf{P}^2(k) = \{(X : Y : Z)\}.$ Revised definition: E(k) = $\{(X : Y : Z) \in \mathbf{P}^2(k) :$ $Y^2Z = X^3 + aXZ^2 + bZ^3\}.$

Could split into cases:

If $(X : Y : Z) \in E(k)$ and $Z \neq 0$: (X : Y : Z) = (x : y : 1)where x = X/Z, y = Y/Z. Note that $y^2 = x^3 + ax + b$. Corresponds to previous (x, y).

If $(X : Y : Z) \in E(k)$ and Z = 0: $X^3 = 0$ so X = 0 so $Y \neq 0$ so (X : Y : Z) = (0 : 1 : 0). Corresponds to previous ∞ .

 $(X_1:Y_1:Z_1)+(X_2:Y_2:Z_2)$ $= (X_3 : Y_3 : Z_3)$ where $U = Y_2 Z_1 - Y_1 Z_2$ $V = X_2 Z_1 - X_1 Z_2$ $W = U^2 Z_1 Z_2 - V^3 - 2V^2 X_1 Z_2$ $X_3 = VW$, $Y_3 = U(V^2 X_1 Z_2 - W) - V^3 Y_1 Z_2,$ $Z_3 = V^3 Z_1 Z_2$.

"Aha! No more divisions by 0."

Compare to previous formulas: $x_3 = \lambda^2 - x_1 - x_2$ and $y_3 = \lambda(x_1 - x_3) - y_1$ where $\lambda = (y_2 - y_1)/(x_2 - x_1)$. Oops, still have exceptions!

Formulas give bogus $(X_3, Y_3, Z_3) = (0, 0, 0)$ if $(X_1 : Y_1 : Z_1) = (0 : 1 : 0).$

Same problem for doubling.

Formulas produce (0 : 1 : 0) for $(X_1 : Y_1 : Z_1) + (X_1 : -Y_1 : Z_1)$ if $Y_1 \neq 0$ and $Z_1 \neq 0$ but not if $Y_1 = 0$.

To define complete group law, use six cases as before.

Jacobian coordinates

"Weighted projective coordinates using weights 2, 3, 1":

Redefine (X : Y : Z) as $\{(r^2X, r^3Y, rZ) : r \in k - \{0\}\}.$

Redefine E(k)using $Y^2 = X^3 + aXZ^4 + bZ^6$.

Could again split into cases for $(X : Y : Z) \in E(k)$: if $Z \neq 0$ then (X : Y : Z) = $(X/Z^2 : Y/Z^3 : 1)$; if Z = 0then (X : Y : Z) = (1 : 1 : 0).

 $(X_1:Y_1:Z_1)+(X_2:Y_2:Z_2)$ $= (X_3 : Y_3 : Z_3)$ where $U_1 = X_1 Z_2^2, U_2 = X_2 Z_1^2,$ $S_1 = Y_1 Z_2^3$, $S_2 = Y_2 Z_1^3$, $H = U_2 - U_1$, $J = S_2 - S_1$, $X_3 = -H^3 - 2U_1H^2 + J^2$, $Y_3 = -S_1 H^3 + J(U_1 H^2 - X_3),$ $Z_3 = Z_1 Z_2 H$.

Streamlined algorithm uses $12\mathbf{M} + 4\mathbf{S}$, where \mathbf{S} is squaring in k and \mathbf{M} is general multiplication in k. (1986 Chudnovsky–Chudnovsky) $11\mathbf{M} + 5\mathbf{S}$. (2001 Bernstein) Still need all six cases.

Why use Jacobian coordinates? Answer: Only $3\mathbf{M} + 5\mathbf{S}$ for Jacobian-coordinate doubling if a = -3 (e.g. NIST curves).

Formulas: If $Y_1 \neq 0$ then $(X_1 : Y_1 : Z_1) + (X_1 : Y_1 : Z_1)$ $= (X_3, Y_3, Z_3)$ where $T = Z_1^2, U = Y_1^2, V = X_1U,$ $W = 3(X_1 - T)(X_1 + T),$ $X_3 = W^2 - 8V,$ $Z_3 = (Y_1 + Z_1)^2 - U - T,$ $Y_3 = W(4V - X_3) - 8U^2.$

Unified addition laws

Do addition laws have to fail for doublings? Not necessarily!

Example: "Jacobi intersection" $s^2 + c^2 = 1$, $as^2 + d^2 = 1$ has 17M addition formula that works for doublings. (1986 Chudnovsky–Chudnovsky) 16**M**. (2001 Liardet–Smart) Many more "unified formulas." But always find exceptions:

points not added by formulas.

"Is this Jacobi intersection related to $y^2 = x^3 + \cdots$?" Yes: $s^2 + c^2 = 1$, $as^2 + d^2 = 1$ is birationally equivalent to $y^2 = x^3 + (2-a)x^2 + (1-a)x$. $(s, c, d) \mapsto (x, y)$: x = (d-1)(1-a)/(ca-d+1-a);y = s(1-a)a/(ca-d+1-a). $(x, y) \mapsto (s, c, d)$: $s = -2y/((y^2/x^2 + a)x);$ $c = 1 - 2/(y^2/x^2 + a) - c$ $2(1-a)/((y^2/x^2+a)x);$ $d = 1 - 2a/(y^2/x^2 + a).$

Do we need 6 cases? No!

Can cover $E(k) \times E(k)$ using 3 addition laws. (1985 H. Lange–Ruppert)

How about just one law that covers $E(k) \times E(k)$? One complete addition law?

Bad news: "Theorem 1. The smallest cardinality of a complete system of addition laws on *E* equals two." (1995 Bosma–H. Lenstra)

Edwards curves

2007 Edwards:

Every elliptic curve over \mathbf{Q} is birationally equivalent to $x^2 + y^2 = c^2(1 + x^2y^2)$ for some $c \in \overline{\mathbf{Q}} - \{0, \pm 1, \pm i\}$. $x^2 + y^2 = c^2(1 + x^2y^2)$ has neutral element (0, c), addition $(x_1, y_1) + (x_2, y_2) = (x_3, y_3)$ with $x_3 = rac{x_1y_2 + y_1x_2}{c(1 + x_1x_2y_1y_2)},$

$$y_3 = rac{y_1y_2 - x_1x_2}{c(1 - x_1x_2y_1y_2)}$$

2007 Bernstein–Lange:

Over a non-binary finite field k, $x^2 + y^2 = c^2(1 + dx^2y^2)$ covers more elliptic curves. Here $c, d \in k^*$ with $dc^4 \neq 1$.

$$x_3=rac{x_1y_2+y_1x_2}{c(1+dx_1x_2y_1y_2)},$$

$$y_3 = rac{y_1y_2 - x_1x_2}{c(1 - dx_1x_2y_1y_2)}$$

Can always take c = 1. Then $10\mathbf{M} + 1\mathbf{S} + 1\mathbf{D}$ for addition, $3\mathbf{M} + 4\mathbf{S}$ for doubling.

Latest news, comparisons: hyperelliptic.org/EFD

<u>Completeness</u>

2007 Bernstein–Lange:

If d is not a square in k then $\{(x,y)\in k imes k:\ x^2+y^2=c^2(1+dx^2y^2)\}$

is a commutative group under this addition law.

The denominators $c(1 + dx_1x_2y_1y_2),$ $c(1 - dx_1x_2y_1y_2)$ are never zero. No exceptional cases! Recall Bosma–Lenstra theorem: "The smallest cardinality of a complete system of addition laws on *E* equals two."

Recall Bosma–Lenstra theorem: "The smallest cardinality of a complete system of addition laws on *E* equals two." . . . meaning: Any addition formula for a Weierstrass curve Ein projective coordinates must have exceptional cases in $E(k) \times E(k)$, where k = algebraic closure of k.

Recall Bosma–Lenstra theorem: "The smallest cardinality of a complete system of addition laws on E equals two." ... meaning: Any addition formula for a Weierstrass curve Ein projective coordinates must have exceptional cases in $E(k) \times E(k)$, where k = algebraic closure of k.

Edwards addition formula has exceptional cases for $E(\overline{k})$

... but not for E(k). We do computations in E(k).

Cryptographic impact

Advantages for cryptography of choosing Edwards curves:

Very high speed.

Completeness eases implementations, avoids simple side-channel attacks.

Cryptographic impact

Advantages for cryptography of choosing Edwards curves:

Very high speed.

Completeness eases implementations, avoids simple side-channel attacks.

Oops, hardware people want binary fields!

2008 B.–L.–Rezaeian Farashahi: binary analogue to Edwards curves; complete, very fast. Still one reason for complaint.

Edwards curves always have point of order 4.

Standard NIST curves were chosen to have prime order. Still one reason for complaint.

Edwards curves always have point of order 4.

Standard NIST curves were chosen to have prime order.

NIST curves can't take advantage of Edwards speed *and* don't have complete addition formulas. Still one reason for complaint.

Edwards curves always have point of order 4.

Standard NIST curves were chosen to have prime order.

NIST curves can't take advantage of Edwards speed *and* don't have complete addition formulas.

2009 Bernstein–Lange, this talk: Have a complete addition law for all of these curves.

Today's curve shape

Fix a field k with $2 \neq 0$.

Fix $t, d \in k$ with $d \neq 0$, $d \neq (t+2)^2$, $27d \neq (2-t)^3$.

Consider the curve $x^2 + y^2 = x + y + txy + dx^2y^2$ with neutral element (0, 0).

Warning: We're still studying choices of curve shapes; we don't promise that this is the best.

For comparison, Edwards: $x^2 + y^2 = 1 + dx^2y^2$ with neutral element (0, 1).

Birational equivalence from $x^2+y^2=x+y+txy+dx^2y^2$ to $v^2 - (t+2)uv + dv =$ $u^3 - (t+2)u^2 - du + (t+2)d$ i.e. $v^2 - (t+2)uv + dv =$ $(u^2 - d)(u - (t + 2))$: u = (dxy + t + 2)/(x + y); $v=\frac{((t+2)^2-d)x}{(t+2)xy+x+y}.$

Assuming t + 2 square, d not: only exceptional point is (0,0), mapping to ∞ .

Inverse: $x = v/(u^2 - d)$; $y = ((t+2)u - v - d)/(u^2 - d)$.

Example: the NIST curves

Consider curve with d = -1 and $t = rac{77856058252666544098227759201}{8607150561834371823249249461}$ over \mathbf{F}_p where $p = 2^{192} - 2^{64} - 1$. Note: d is non-square in \mathbf{F}_{p} . Birationally equivalent to standard "NIST P-192" curve $v^2 = u^3 - 3u + a_6$ where $a_6 = rac{24551555460089438177402939151}{97451784769108058161191238065}.$

Consider curve with d = 11 and

 $t = 59653337798320066750209233023 \\ 6009670$

over \mathbf{F}_p where $p = 2^{224} - 2^{96} + 1$. Note: d is non-square in \mathbf{F}_p .

Birationally equivalent to standard "NIST P-224" curve $v^2 = u^3 - 3u + a_6$ where 18958286285566608000408668544

 $a_6 = 49392641550468096867932107578.$ 7234672564 Consider curve with d = -1 and

 $t = \frac{78751018041117252545420999954}{1395651175859201799}$

over \mathbf{F}_p where $p = 2^{256} - 2^{224} + 2^{192} + 2^{96} - 1$.

Note: d is non-square in \mathbf{F}_p .

Birationally equivalent to standard "NIST P-256" curve $v^2 = u^3 - 3u + a_6$ where $a_6 = {}^{41058363725152142129326129780}_{04726840911444101599372555483.}_{5256314039467401291}$

Consider curve with d = -1 and

 $t = \frac{85909296364310935634030366769}{19596768294026516624086336448} \\ 0501907705272975221538249252$

over \mathbf{F}_p where $p = 2^{384} - 2^{128} - 2^{96} + 2^{32} - 1$.

Note: d is non-square in \mathbf{F}_p .

Birationally equivalent to standard "NIST P-384" curve $v^2 = u^3 - 3u + a_6$ where 75801935599597058778490118403 $a_6 = \frac{89048093056905856361568521428}{70730198868924130986086513626}$.

Consider curve with d = 3 and

 $t = {28255491549159851139291566929 \ 14423222253417506441326327182 \ 78098467340130883832560776891 \ 27881593298389934213527989123 \ 13871892632272472360900308353 \ 04279675250$

over \mathbf{F}_p where $p = 2^{521} - 1$. Note: d is non-square in \mathbf{F}_p .

Birationally equivalent to standard "NIST P-521" curve $v^2 = u^3 - 3u + a_6$ where

 $a_6 = { \begin{smallmatrix} 10938490380737342745111123907\ 66805569936207598951683748994\ 58639449595311615073501601370\ 87375737596232485921322967063\ 13309438452531591012912142327\ 488478985984 \end{split}$

Today's addition law

$$x_1+x_2+(t-2)x_1x_2+\ (x_1-y_1)(x_2-y_2)+\ x_3=rac{dx_1^2(x_2y_1+x_2y_2-y_1y_2)}{1-2dx_1x_2y_2-};\ dx_1^2(x_2+y_2+(t-2)x_2y_2)$$
;

$$y_1+y_2+(t-2)y_1y_2+\ (y_1-x_1)(y_2-x_2)+\ y_3=rac{dy_1^2(y_2x_1+y_2x_2-x_1x_2)}{1-2dy_1y_2x_2-}.$$

Exercise: On curve, if denominators are nonzero. Exercise: (x, y) + (0, 0) = (x, y). Exercise: (x, y) + (y, x) = (0, 0). Exercise: Compute projectively using 26M + 8S + 8D. ... Clearly can be improved; we're not done optimizing yet. Exercise: Corresponds to addition on Weierstrass curve.

<u>Completeness</u>

$$x_1+x_2+(t-2)x_1x_2+\ (x_1-y_1)(x_2-y_2)+\ x_3=rac{dx_1^2(x_2y_1+x_2y_2-y_1y_2)}{1-2dx_1x_2y_2-};\ dx_1^2(x_2+y_2+(t-2)x_2y_2)$$
;

$$egin{aligned} &y_1+y_2+(t-2)y_1y_2+\ &(y_1-x_1)(y_2-x_2)+\ &y_3=&rac{dy_1^2(y_2x_1+y_2x_2-x_1x_2)}{1-2dy_1y_2x_2-}\ &dy_1^2(y_2+x_2+(t-2)y_2x_2) \end{aligned}$$

Can denominators be 0?

Only if *d* is a square!

Theorem: Assume that k is a field with $2 \neq 0$; d, t, x_1 , y_1 , x_2 , $y_2 \in k$; d is not a square in k; $27d \neq (2-t)^3$; $x_1^2 + y_1^2 = x_1 + y_1 + tx_1y_1 + dx_1^2y_1^2;$ $x_2^2 + y_2^2 = x_2 + y_2 + tx_2y_2 + dx_2^2y_2^2$. Then $1 - 2dx_1x_2y_2$ $dx_1^2(x_2+y_2+(t-2)x_2y_2) \neq 0.$

Only if *d* is a square!

Theorem: Assume that k is a field with $2 \neq 0$; $d,t,x_1,y_1,x_2,y_2\in k;$ d is not a square in k; $27d \neq (2-t)^3$; $x_1^2 + y_1^2 = x_1 + y_1 + tx_1y_1 + dx_1^2y_1^2;$ $x_2^2 + y_2^2 = x_2 + y_2 + tx_2y_2 + dx_2^2y_2^2$. Then $1 - 2dx_1x_2y_2$ $dx_1^2(x_2+y_2+(t-2)x_2y_2) \neq 0.$

By $x \leftrightarrow y$ symmetry also $1 - 2dy_1y_2x_2 - dy_1^2(y_2 + x_2 + (t-2)y_2x_2) \neq 0.$ Proof: Suppose that $1 - 2dx_1x_2y_2 - dx_1^2(x_2 + y_2 + (t - 2)x_2y_2) = 0.$ Proof: Suppose that $1 - 2dx_1x_2y_2 - dx_1^2(x_2 + y_2 + (t - 2)x_2y_2) = 0.$

Note that $x_1 \neq 0$.

Proof: Suppose that $1-2dx_1x_2y_2-dx_1^2(x_2+y_2+(t-2)x_2y_2)=0.$ Note that $x_1
eq 0.$

Use curve equation₂ to see that $(1 - dx_1x_2y_2)^2 = dx_1^2(x_2 - y_2)^2.$

Proof: Suppose that $1-2dx_1x_2y_2-dx_1^2(x_2+y_2+(t-2)x_2y_2)=0.$ Note that $x_1
eq 0.$

Use curve equation₂ to see that $(1 - dx_1x_2y_2)^2 = dx_1^2(x_2 - y_2)^2.$

By hypothesis d is non-square so $x_1^2(x_2 - y_2)^2 = 0$ and $(1 - dx_1x_2y_2)^2 = 0$. Proof: Suppose that $1-2dx_1x_2y_2-dx_1^2(x_2+y_2+(t-2)x_2y_2)=0.$ Note that $x_1
eq 0.$ Use curve equation₂ to see that $(1-dx_1x_2y_2)^2=dx_1^2(x_2-y_2)^2.$

By hypothesis d is non-square so $x_1^2(x_2 - y_2)^2 = 0$ and $(1 - dx_1x_2y_2)^2 = 0$.

Hence $x_2 = y_2$ and $1 = dx_1x_2y_2$.

Curve equation₁ times $1/x_1^2$: $1 + y_1^2/x_1^2 =$ $1/x_1 + y_1(1/x_1^2 + t/x_1) + dy_1^2$. Curve equation₁ times $1/x_1^2$: $1 + y_1^2/x_1^2 =$ $1/x_1 + y_1(1/x_1^2 + t/x_1) + dy_1^2$. Substitute $1/x_1 = dx_2^2$: $1 + d^2y_1^2x_2^4 =$ $dx_2^2 + dy_1(dx_2^4 + x_2^2t) + dy_1^2$.

Curve equation₁ times $1/x_1^2$: $1 + y_1^2 / x_1^2 =$ $1/x_1 + y_1(1/x_1^2 + t/x_1) + dy_1^2$. Substitute $1/x_1 = dx_2^2$: $1 + d^2 y_1^2 x_2^4 =$ $dx_2^2 + dy_1(dx_2^4 + x_2^2t) + dy_1^2.$ Substitute $2x_2^2 = 2x_2 + tx_2^2 + dx_2^4$: $(1 - dy_1 x_2^2)^2 = d(x_2 - y_1)^2.$

Curve equation₁ times $1/x_1^2$: $1 + y_1^2 / x_1^2 =$ $1/x_1 + y_1(1/x_1^2 + t/x_1) + dy_1^2$. Substitute $1/x_1 = dx_2^2$: $1 + d^2 y_1^2 x_2^4 =$ $dx_2^2 + dy_1(dx_2^4 + x_2^2t) + dy_1^2.$ Substitute $2x_2^2 = 2x_2 + tx_2^2 + dx_2^4$: $(1 - dy_1 x_2^2)^2 = d(x_2 - y_1)^2.$ Thus $x_2 = y_1$ and $1 = dy_1 x_2^2$. Hence $1 = dx_2^3$.

Curve equation₁ times $1/x_1^2$: $1 + y_1^2 / x_1^2 =$ $1/x_1 + y_1(1/x_1^2 + t/x_1) + dy_1^2$. Substitute $1/x_1 = dx_2^2$: $1 + d^2 y_1^2 x_2^4 =$ $dx_2^2 + dy_1(dx_2^4 + x_2^2t) + dy_1^2.$ Substitute $2x_2^2 = 2x_2 + tx_2^2 + dx_2^4$: $(1 - dy_1 x_2^2)^2 = d(x_2 - y_1)^2.$ Thus $x_2 = y_1$ and $1 = dy_1 x_2^2$. Hence $1 = dx_2^3$. Now $2x_2^2 = 2x_2 + tx_2^2 + x_2$

so $3 = (2-t)x_2$ so $27d = (2-t)^3$. Contradiction.