E. Edwards curves

An *elliptic curve* over a field K is a projective curve E such that the function field $K(E)$ is of genus 1, and E contains at least one K-rational point. The curve E is often considered in its affine version. This is particularly true if the curve is smooth and there is only one point at infinity.

For each elliptic curve there exists a smooth Weierstraß curve C such that $K(E) \cong K(C)$. For each elliptic curve it is possible to define a group operation $\oplus$. The group is then denoted by $E(K)$. If $K(E) \cong K(C)$, then $E(K) \cong C(K)$. Are there any reasons why there should be considered other elliptic curves but the smooth Weierstraß curves? One reason may be computational, and this is why Montgomery curves have been considered. Another reason may be structural. In cryptographic applications the fact that doubling and adding proceeds differently makes an implementation vulnerable to side channel attacks. We would like to have an elliptic curve with only one formula for both doubling of a point, and addition of two distinct points. Such a formula is sometimes known as a *closed formula* or a *uniform formula*.

Edwards curves fulfil such a requirement. Some of the Edwards curves have no K-rational point at infinity, and these are those for which a closed formula can be used indiscriminately. This is also true for the so called twisted Edwards curves, which is a somewhat more general notion. Twisted Edwards curves correspond to Montgomery curves. Assume $\mathrm{char}(K) \neq 2$. Then for each twisted Edwards curve E there exists a smooth Montgomery curve M such that $M(K) \cong E(K)$, and vice versa.

To define the group $E(K)$, E an elliptic curve, in full generality, the notion of a place is needed. Elements of $E(K)$ are places of degree one. If $\alpha \in E$ is a smooth point, then there is only one place at α. However, if α is a singular point, then there are either more places at α, or there is a place of degree > 1. That makes the connection between K-rational projective points of E and elements of $E(K)$ a bit more complicated. If E is a twisted Edwards curve, then all affine points are smooth and all points at infinity are singular. If all places induced by K-rational points at infinity are of degree > 1, then each element of $E(K)$ can be represented by exactly one affine point. Thus in this case $E(K)$ may be constructed directly upon the set of all affine K-rational points. In the general case the affine points may be also used, but their set has to be extended by two or four extra elements that correspond to "places at infinity". There are several ways how to do that formally, and some of them have computational consequences. These are discussed below.

E.1. **Branches and the definition.** Consider a curve described by equation

$$y^2 + x^2 = 1 + dx^2y^2 \tag{E.1}$$

that is defined over real numbers. There are good reasons to expect that the corresponding projective curve will have one or two components of connectivity, each of them closed, similarly as in the case of Weierstraß curves. However, the number of connectivity components of an affine curve may be bigger than the number of components of its projective completion. (Think about a hyperbole which has only one component in projective coordinates, but two in affine coordinates.) For a while, for the sake of simplicity of expression, call an affine component of connectivity a *branch*. Denote the curve by E.

Suppose first that $d = s^2$ and $s > 1$. When (E.1) is written in the form $y^2 = (1 - x^2)/(1 - s^2x^2)$, then it is easy to deduce that in this case there exists exactly one branch which satisfies $y > 0$ and $x \in (-s^{-1}, s^{-1})$. This branch has a shape of the letter $\cup$, with $x = -s^{-1}$ and $x = s^{-1}$ being tangents at infinity, and $(0, 1)$

being the bottom point. By turning the branch bottom up, i.e. by reflecting it along the axis x (the line $y = 0$), we obtain another branch. This branch satisfies $y < 0$ and $x \in (-s^{-1}, s^{-1})$. Since the definition of the curve is $x \leftrightarrow y$ symmetric, the other two branches are obtained by right angle rotation of the branches that have been already described. So there are four branches, none of which is closed. Extreme points of these branches are $(0, 1)$, $(0, -1)$, $(1, 0)$ and $(-1, 0)$. Note that these points belong to E for any field K and any element $d \in K$.

At this point the reader might wish to guess the number of points at infinity without actually computing them.

If $0 < s < 1$ and $d = s^2$, then there are five branches, and the central branch is closed.

If $d < 0$, then any point $(\alpha, \beta) \in E$ fulfils $|\alpha| \leq 1$ since $\beta^2 = (1 - \alpha^2)/(1 - d\alpha^2)$ and $1 - d\alpha^2 \geq 1$ for every α. Similarly, $|\beta| \leq 1$. In this case there is only one branch. The branch is closed and resembles a somewhat smoothed star from the logo of an Orion chocolate bar. Note that if $d = 0$, then the curve coincides with a circle. With decreasing d, the circle gets more and more pressed crosswise towards the centre (the pressure comes along the quadrangle axes).

Consider now the projective curve induced by (E.1), for any field K, $\mathrm{char}(K) \neq 2$. The equation is $Y^2Z^2 + X^2Z^2 = Z^4 + dX^2Y^2$. Assume $d \neq 0$. If $Z = 0$, then $dX^2Y^2 = 0$. There are thus two points at infinity, $(0 : 1 : 0)$ and $(1 : 0 : 0)$. Both of them are singular. If $K = \mathbb{R}$ and $d = s^2 > 1$, then the two affine branches with $x \in (-s^{-1}, s)$ make one projective branch in the shape of the digit 8. The point of crossing is equal to the projective point $(0 : 1 : 0)$. There are two distinct places of degree 1 at this point. Think about the point as if consisting of two "ideal" points. Separating them "resolves the singularity" and changes the shape of 8 into a (topological) circle. If $1 > d = s^2 > 0$, then the situation is similar but somewhat different since there is a central closed affine branch. The other four affine branches form a single projective branch the shape of which can be represented by two circles that intersect in two points. The points of crossing are $(1 : 0 : 0)$ and $(0 : 1 : 0)$. Singularities can be resolved in a similar manner, and that makes this projective branch a topological circle too.

If $d < 0$, then each place of degree one of $K(E)$ corresponds to a (unique) affine point. In fact, this is true for any field K, $\mathrm{char}(K) \neq 2$, when d is not a square. How does this relate to the fact that the projective curve contains K-rational points $(1 : 0 : 0)$ and $(0 : 1 : 0)$ in this case too? The answer is that at each of these points there sits a single place, and this place is not of degree one, but of degree two. These points thus do not influence the structure of the group $E(K)$. Of course, the situation changes if the same curve is considered over the field $K[\sqrt{d}]$.

An *Edwards curve* over K, $\mathrm{char}(K) \neq 2$, is any curve given by (E.1), with $d \notin \{0, 1\}$. A *twisted Edwards curve* over K, $\mathrm{char}(K) \neq 2$ is a curve given by

$$ax^2 + y^2 = 1 + dx^2y^2, \text{ where } a, d \in K^* \text{ and } a \neq d. \tag{E.2}$$

Usage of the adjective "twisted" indicates that the class of twisted Edwards curves extends the class of Edwards curves only modestly. To see this note that if $a = b^2$, then $(bx)^2 + y^2 = 1 + db^{-2}(bx)^2y^2$. A twisted Edwards curve with parameters (b^2, d) is K-equivalent to the Edwards curve given by $x^2 + y^2 = 1 + db^{-2}x^2y^2$. More generally, there is a K-equivalence between parameters (b^2c, d) and $(c, b^{-2}d)$. To cover the class of twisted Edwards curves over a finite field $\mathbb{F}_q$, q odd, it is thus enough to consider the Edwards curves, and the curves given by $\vartheta x^2 + y^2 = 1 + dx^2y^2$, where ϑ is a preselected nonsquare.

Let us now verify that the polynomial $ax_1^2 + x_2^2 - 1 - dx_1^2x_2^2 \in K[x_1, x_2]$ is absolutely irreducible if $a, d \in K^*$ and $a \neq d$. Up to now we have tacitly assumed

that this is true. If it have not been true, we could not have had considered the function field $K(E)$ since this assumes that the polynomial defining E is irreducible.

Proposition E.1. *Let K be a field,* $\mathrm{char}(K) \neq 2$. *Assume that $a_1, a_2, d \in K^*$. The polynomial $f(x_1, x_2) = a_1x_1^2 + a_2x_2^2 - 1 - dx_1^2x_2^2$ is absolutely irreducible if and only if $d \neq a_1a_2$.*

Proof. If $d = a_1a_2$, then $f = (a_1x_1^2 - 1)(1 - a_2x_2^2)$. Let $f = g_1g_2$, where $g_1, g_2 \in \bar{K}[x_1, x_2]$. If $\deg_{x_1}(g_1) = \deg_{x_2}(g_1) = 2$, then $g_2 \in \bar{K}^*$. Assume that $g_1, g_2 \notin \bar{K}^*$. Suppose first that $g_i = \alpha_i x_i^2 + \beta_i x_i + \gamma_i \in \bar{K}[x_i]$, $i \in \{1, 2\}$. In the polynomial $f(x_1, x_2)$ the coefficients at both $x_1^2x_2$ and $x_1x_2^2$ vanish. We have $\alpha_1\alpha_2 = -d \neq 0$. Hence both β_1 and β_2 must vanish too. Now, $f = g_1g_2 = \alpha_1\alpha_2x_1^2x_2^2 + \alpha_1\gamma_2x_1^2 + \alpha_2\gamma_1x_2^2 + \gamma_1\gamma_2$. Therefore $\gamma_1\gamma_2 = -1$, $\alpha_1\gamma_2 = a_1$, $\alpha_2\gamma_1 = a_2$ and $-d = -\alpha_1\gamma_1\gamma_2\alpha_2 = -a_1a_2$.

Assume $d \neq a_1a_2$. We have shown that there cannot be $\deg_{x_i}(g_j) \in \{0, 2\}$ for all $i, j \in \{1, 2\}$. Hence there exists $i \in \{1, 2\}$ such that $\deg_{x_i}(g_1) = \deg_{x_i}(g_2) = 1$. Because of the $x_1 \leftrightarrow x_2$ symmetry it may be assumed that $i = 1$. This means that the polynomial f splits over the field of rational functions $\bar{K}(x_2)$ when regarded as a quadratic polynomial in one variable x_1. This can happen if and only if the discriminant $-4a_1(a_2x_2^2 - 1)(1 - da_1^{-1}x_2^2)$ is a square in $\bar{K}[x_2]$. If it is a square, then all of the roots have to have an even multiplicity. This is not possible since the polynomials $a_2x_2^2 - 1$ and $1 - da_1^{-1}x_2^2$ have no common root because $d \neq a_1a_2$ is assumed, and none of them has a double root because $\mathrm{char}(K) \neq 2$. □

Lemma E.2. *Let K be a field,* $\mathrm{char}(K) \neq 2$. *Assume that $a_1, a_2, d \in K^*$, $d \neq a_1a_2$ and $f(x_1, x_2) = a_1x_1^2 + a_2x_2^2 - 1 - dx_1^2x_2^2$. Let $\alpha_1, \alpha_2 \in \bar{K}$ be such that $f(\alpha_1, \alpha_2) = 0$. Then $(\partial f/\partial x_i)(\alpha_1, \alpha_2) \neq 0$ for at least one $i \in \{1, 2\}$.*

Proof. First note that $\partial f/\partial x_i = 2x_i(a_i - dx_j^2)$, where $i, j \in \{1, 2\}$ and $j \neq i$. If $\alpha_1 = 0$, then $\alpha_2^2 = a_2^{-1} \neq 0$ and $(\partial f/\partial x_2)(0, \alpha_2) = 2\alpha_2a_1 \neq 0$. Suppose that $\alpha_i \neq 0$ and $(\partial f/\partial x_i)(\alpha_1, \alpha_2) = 0$ for both $i \in \{1, 2\}$. Then $\alpha_1^2 = a_2d^{-1}$, $\alpha_2^2 = a_1d^{-1}$ and $f(\alpha_1, \alpha_2) = a_1a_2d^{-1} + a_1a_2d^{-1} - 1 - a_1a_2d^{-1} = d^{-1}(a_1a_2 - d) \neq 0$, a contradiction. □

Corollary E.3. *Let K be a field with* $\mathrm{char}(K) \neq 2$. *Any twisted Edwards curve over K is smooth at every of its affine points.*

If E is an elliptic curve over K, then any of its K-rational points may be chosen as the neutral element of $E(K)$. The choice is a matter of convention and is made so that the addition formula is as simple as possible. For twisted Edwards curves the neutral element has been chosen to be equal to $(0, 1)$. The closed formula for addition is

$$(\alpha_1, \alpha_2) \oplus (\beta_1, \beta_2) = \left(\frac{\alpha_1\beta_2 + \alpha_2\beta_1}{1 + d\alpha_1\alpha_2\beta_1\beta_2}, \frac{\alpha_2\beta_2 - a\alpha_1\beta_1}{1 - d\alpha_1\alpha_2\beta_1\beta_2} \right). \tag{E.3}$$

This formula works for any two affine points provided $d\alpha_1\alpha_2\beta_1\beta_2 \neq \pm 1$. If the latter condition is not satisfied, then the result is one of the places of infinity. Their number can be expressed as $2(\varepsilon_1 + \varepsilon_2)$, where $\varepsilon_1, \varepsilon_2 \in \{0, 1\}$ is defined so that $\varepsilon_1 = 1$ if d is a square, and $\varepsilon_2 = 1$ if ad^{-1} is square. Each of the places at infinity is an element of $E(K)$ that is either of order 2, or of order 4. Applications in cryptography are mainly concerned with points $P \in E(K)$ that are of large prime order. For these applications computation rules involving places at infinity are thus not needed. However, for other applications, like factorization algorithms, these formulas have to be established. This will be discussed later.

E.2. **Birational equivalence.** Let $C = V_f$, $f \in K[x_1, x_2]$ irreducible. Recall that $K(C)$ can be interpreted as a set of partial mappings $\rho\colon C \to K$ that can be represented by rational mappings $a(x_1, x_2)/b(x_1, x_2)$, $b \notin (f)$. If $\alpha \in C$ and $b(\alpha) \neq 0$, then $\rho(\alpha) = a(\alpha)/b(\alpha)$. Recall also that $a_1/b_1, a_2/b_2 \in K(x_1, x_2)$, $b_1, b_2 \notin (f)$, represent the same $\rho \in K(C)$ if and only if $a_1b_2 - a_2b_1 \in (f)$, i.e., if $(a_1 + (f))/(b_1 + (f))$ and $(a_2 + (f))/(b_2 + (f))$ denote the same element of $K(C)$. The partial mapping ρ is defined at $\alpha \in C$ whenever there exists a representative a/b such that $b(\alpha) \neq 0$. There are only finitely many $\alpha \in C$ at which $\rho(\alpha)$ is not defined. This is because if a/b represents ρ, $b \notin (f)$, then there are only finitely many $\alpha \in C$ such that $b(\alpha) = 0$. We say that ρ is defined *nearly everywhere*. This is meant as a synonym to *up to finitely many elements (or points)*. Use $\mathrm{Dom}(\rho)$ to denote the *domain* of ρ, i.e. the set of elements where ρ is defined.

Let $C_1 = V_{f_1}$ and $C_2 = V_{f_2}$, where $f_1, f_2 \in K[x_1, x_2]$ are irreducible. A pair $\rho = (\rho_1, \rho_2) \in K(C)^2$ is said to be a *rational map* $C_1 \to C_2$ if $(\rho_1(\alpha), \rho_2(\alpha)) \in C_2$ whenever $\alpha \in \mathrm{Dom}(\rho) = \mathrm{Dom}(\rho_1) \cap \mathrm{Dom}(\rho_2)$. The curves C_1 and C_2 are *birationally equivalent* (over K) if there exist rational maps $\rho\colon C_1 \to C_2$ and $\sigma\colon C_2 \to C_1$ such that $\sigma\rho(\alpha) = \alpha$ for nearly all $\alpha \in C_1$ and $\rho\sigma(\beta) = \beta$ for nearly all $\beta \in C_2$ (an equivalent condition: $\sigma\rho(\alpha) = \alpha$ whenever $\alpha \in \mathrm{Dom}(\rho) \cap \rho^{-1}(\mathrm{Dom}(\sigma))$, similarly for β).

If $\rho\colon C_1 \to C_2$ and $\sigma\colon C_2 \to C_1$ yield a birational equivalence, then there exist mutually inverse K-isomorphisms $\sigma^*\colon K(C_1) \cong K(C_2)$ and $\rho^*\colon K(C_2) \cong K(C_1)$ such that $x_i + (f_1) \mapsto \sigma_i$ and $x_i + (f_2) \mapsto \rho_i$. In fact, *$K(C_1)$ and $K(C_2)$ are K-isomorphic if and only if C_1 and C_2 are birationally equivalent over K.*

To see that a birational equivalence induces mutually inverse isomorphisms of function fields is not too difficult. Nevertheless it is technically somewhat demanding. For a reader who would like to verify the statement the following comments may be useful. If $\tau \in K(C_1)$, then $\sigma^*(\tau) = \sigma^*(\tau(x_1 + (f_1), x_2 + (f_1))) = \tau(\sigma^*(x_1 + (f_1)), \sigma^*(x_2 + (f_1))) = \tau(\sigma_1, \sigma_2)$. Hence $\sigma^*(\tau)(\beta) = \tau(\sigma_1(\beta), \sigma_2(\beta)) = \tau\sigma(\beta)$ for every $\beta \in C_2$. Since $\sigma^*\rho^*(x_i + (f_2)) = \sigma^*(\rho_i)$ we get $\sigma^*\rho^*(x_1 + (f_2))(\beta) = \sigma^*(\rho_1)(\beta) = \rho_1\sigma(\beta)$. Now $\rho\sigma(\beta) = (\rho_1\sigma(\beta), \rho_2\sigma(\beta))$ is assumed to be equal to $\beta = (\beta_1, \beta_2)$ nearly everywhere. Hence $\rho_1\sigma(\beta) = \beta_1$ nearly everywhere, and therefore $\sigma^*\rho^*(x_1 + (f_2)) = \sigma^*(\rho_1) = x_1 + (f_2)$. Similarly, $\sigma^*\rho^*(x_2 + (f_2)) = x_2 + (f_2)$, and hence $\sigma^*\rho^* = \mathrm{id}_{K(C_2)}$. The equality $\rho^*\sigma^* = \mathrm{id}_{K(C_1)}$ follows in the same way.

If C_1 and C_2 are birationally equivalent elliptic curves, then $C_1(K) \cong C_2(K)$. This is because the structure of the abelian group $C_i(K)$ fully depends upon the structure of the function field $K(C_i)$. If the fields are isomorphic, then the groups are isomorphic too.

Recall that equations $f_1(x_1, x_2) = 0$ and $f_2(x_1, x_2) = 0$ are said to be K-equivalent if the polynomials can be obtained one from another by a linear substitution. Such substitutions induce a birational equivalence between C_1 and C_2 that is realized by affine mappings, i.e. by a linear change of coordinates, like in the case of Weierstraß and Montgomery curves. However, not every birational equivalence is affine. Below we shall observe that twisted Edwards curves are birationally equivalent to Montgomery curves. The advantage of invertible affine (or linear) mappings is that they are defined globally for all $\alpha \in \mathbb{A}^2 = \bar{K} \times \bar{K}$, and their inversions are affine (linear) too. The birational equivalence may be thus obtained by restricting a global mapping to curves.

A *linear fractional mapping* $x \mapsto (ax + c)/(bx + d)$, $ad - bc \neq 0$, nearly permutes an affine line (it may be extended to a permutation of the projective line by $\infty \mapsto a/b$ and $-d/b \mapsto \infty$). Linear fractional mappings thus may serve as a tool to define transformations of $\mathbb{A}^2$ that are very close to permutations. One of such transformations is used to associate Montgomery and Edwards curves:

Lemma E.4. *Assume* $\mathrm{char}(K) \neq 2$. *Then* $\vartheta : \beta \mapsto (\beta+1)/(\beta-1)$ *permutes the set* $K' = K \setminus \{0, 1, -1\}$ *and* $\Psi\colon (\alpha, \beta) \mapsto (\vartheta(\beta), \vartheta(\beta)/\alpha)$ *is a bijection* $K^* \times K' \to K' \times K^*$.

Proof. If $\beta \neq 1$, then $\vartheta^2(\beta) = \beta$, $\vartheta(0) = -1$ and $\vartheta(-1) = 0$. Hence ϑ permutes K'. The mapping Ψ clearly sends $K^* \times K'$ to $K' \times K^*$ injectively. If $(\gamma, \delta) \in K' \times K^*$, then $(\gamma, \delta) = \Psi(\gamma/\delta, \vartheta^{-1}(\gamma))$. □

Lemma E.5. *Assume* $\mathrm{char}(K) \neq 2$. *The mappings*

$$(a, d) \mapsto \left(2\frac{a+d}{a-d}, 4\frac{1}{a-d}\right) \qquad (A, B) \mapsto \left(\frac{A+2}{B}, \frac{A-2}{B}\right) \tag{E.4}$$

are mutually inverse if $(a, d) \in K^* \times K^*$, $a \neq d$, *and* $(A, B) \in K \times K^*$, $A \neq \pm 2$.

Proof. Let $A = 2(a+d)/(a-d)$ and $B = 4/(a-d)$, where $a, d \in K$ and $a \neq d$. Then $B \neq 0$, $Aa - Ad = 2a + 2d$, $(A-2)a = (A+2)d$, $a = (4+Bd)/B$, $(A-2)(4+Bd) = ABd + 2Bd$, $-4 - Bd + 2A = Bd$, $d = (A-2)/B$, $4 + Bd = A + 2$ and $a = (A+2)/B$. This establishes a bijection between the set of all $(a, d) \in K \times K$, $a \neq d$, and the set $K \times K^*$. The rest is clear. □

Lemma E.6. *Let* $\mathrm{char}(K) \neq 2$ *and suppose that* $a, d \in K^*$ *are such that* $a \neq d$. *Set* $A = 2(a+d)/(a-d)$ *and* $B = 4/(a-d)$, *and assume that* $\alpha, \beta \in K$ *are such that* $\alpha \neq 0$ *and* $\beta \notin \{0, 1, -1\}$. *Put* $u = (1+\beta)/(1-\beta)$ *and* $v = u/\alpha$. *Then*

$$a\alpha^2 + \beta^2 = 1 + d\alpha^2\beta^2 \quad \Longleftrightarrow \quad Bv^2 = u^3 + Au^2 + u.$$

Proof. Multiplying the equality $Bv^2 = u^3 + Au^2 + u$ by $(1-\beta)^3$, dividing it by $1+\beta$, and using $(1+\beta)(1-\beta) = 1 - \beta^2$ yields an equivalent equation

$$B(1-\beta^2)\alpha^{-2} = (1+\beta)^2 + A(1-\beta^2) + (1-\beta)^2 = A(1-\beta^2) + 2(1+\beta^2).$$

Hence $(1-\beta^2)(B\alpha^{-2} - A) = 2(1+\beta^2)$. Therefore

$$2(1-\beta^2)(2\alpha^{-2} - (a+d)) = 2(a-d)(1+\beta^2),$$

which is the same as $2\alpha^{-2} - (a+d) - 2\alpha^{-2}\beta^2 + \beta^2 d = a - d - d\beta^2$ and as $\alpha^{-2} - \alpha^{-2}\beta^2 + \beta^2 d = a$. The latter can be written as $1 + d\alpha^2\beta^2 = a\alpha^2 + \beta^2$. Nothing else is needed since none of the transformations changes the set of solutions because $a \neq 0$ and $\beta \notin \{-1, 0, 1\}$ has been assumed. □

Theorem E.7. *Let* K *be a field of characteristic* $\neq 2$, *and let* $a, d \in K^*$ *be such that* $a \neq d$. *Set* $A = 2(a+d)/(a-d)$ *and* $B = 4/(a-d)$. *The twisted Edwards curve* E *given by* $1 + dx_1^2x_2^2 = ax_1^2 + x_2^2$ *is birationally equivalent over* K *to the Montgomery curve* M *given by* $Bx_2^2 = x_1^3 + Ax_1^2 + x_1$. *The rational map* $E \to M$ *may be represented by* $((1+x_2)/(1-x_2), (1+x_2)/x_1(1-x_2))$, *and the inverse rational map* $M \to E$ *by* $(x_1/x_2, (x_1-1)/(x_1+1))$.

Proof. The described rational map $E \to M$ sends nearly all elements of E upon M by Lemma E.6. The mapping is injective and its image covers nearly all elements of M, by Lemma E.4. It is immediately clear that the described rational map $M \to E$ behaves as an inverse mapping at each point where it is possible to define composition of the both mappings. □

Corollary E.8. *Let* K *be a field of characteristic* $\neq 2$. *For each twisted Edwards curve* E *over* K *there exists a smooth Montgomery curve* M *that is birationally equivalent over* K *to* E, *and for each smooth Montgomery curve* M *over* K *there exists a twisted Edwards curve* E *that is birationally equivalent over* K *to* M.

Proof. This immediately follows from Theorem E.7 and Lemma E.5. □

E.3. **Completed curves and various formulas.** Formula (E.3) is not the only way how the addition upon a twisted Edwards curve may be expressed. The so called *dual addition law*

$$(\alpha_1,\alpha_2)\oplus(\beta_1,\beta_2)=\left(\frac{\alpha_1\alpha_2+\beta_1\beta_2}{\alpha_2\beta_2+a\alpha_1\beta_1},\,\frac{\alpha_1\alpha_2-\beta_1\beta_2}{\alpha_1\beta_2-\alpha_2\beta_1}\right) \tag{E.5}$$

is an alternative. It gives the same result as (E.3) whenever the denominators in both (E.3) and (E.5) are nonzero. Obviously, (E.5) may never be used for doublings. However, it is important both theoretically and practically, since it is a source of various speed-ups. The speed-ups usually work differently for the doubling and for the addition of distinct points (which is often called a *generic addition*). They are used if the context does not require a closed formula that makes the computation resistant to side channel attacks.

Let us observe that the dual addition law really works. If the denominators are nonzero, then the equality

$$\frac{\alpha_1\alpha_2+\beta_1\beta_2}{\alpha_2\beta_2+a\alpha_1\beta_1}=\frac{\alpha_1\beta_2+\alpha_2\beta_1}{1+d\alpha_1\alpha_2\beta_1\beta_2}$$

holds if and only if

$$\begin{aligned}\alpha_1\alpha_2+d\alpha_1^2\alpha_2^2\beta_1\beta_2+\beta_1\beta_2+d\alpha_1\alpha_2\beta_1^2\beta_2^2&\\ =\alpha_1\alpha_2(1+d\beta_1^2\beta_2^2)+\beta_1\beta_2(1+d\alpha_1^2\alpha_2^2)&\\ =\alpha_1\alpha_2(a\beta_1^2+\beta_2^2)+\beta_1\beta_2(a\alpha_1^2+\alpha_2^2)&\end{aligned}$$

is equal to $(\alpha_2\beta_2+a\alpha_1\beta_1)(\alpha_1\beta_2+\alpha_2\beta_1)$. That is clearly true.

The proof for the second coordinate may be done similarly.

When the addition is computed upon **projective coordinates**, i.e. upon the zeros of $aX_1^2X_3^2+X_2^2X_3^2=X_3^4+dX_1^2X_2^2$, then it is possible to order the operations in such a way that the addition of distinct point (the *generic addition*) costs $10\mathrm{M}+1\mathrm{S}+1a+1d$, where $1a+1d$ refer to multiplications by a and d (which may be chosen small), while the cost of doubling is $3\mathrm{M}+4\mathrm{S}+1a$.

There have been also used **inverted coordinates** which correspond to the equation $aX_1^{-2}X_3^{-2}+X_2^{-2}X_3^{-2}=X_3^{-4}+dX_1^{-2}X_2^{-2}$, and thus also to $aX_2^2X_3^2+X_1^2X_3^2=X_1^2X_2^2+dX_3^4$. In these coordinates the cost of generic addition is $9\mathrm{M}+1\mathrm{S}+1a+1d$, and the doubling costs $3\mathrm{M}+4\mathrm{S}+1a+1d$.

We shall skip **extended coordinates** and turn directly to **completed coordinates**. They use projective coordinates, but not in $\mathbb{P}^2$ or $\mathbb{P}^3$, but in $\mathbb{P}^1\times\mathbb{P}^1$. The curve, say U, is formed by all $((\alpha_1:\alpha_2),(\beta_1:\beta_2))$ for which the substitutions $(X_1,X_2)\mapsto(\alpha_1,\alpha_2)$ and $(Y_1,Y_2)\mapsto(\beta_1,\beta_2)$ fulfil

$$aX_1^2Y_2^2+Y_1^2X_2^2=X_2^2Y_2^2+dX_1^2Y_1^2. \tag{E.6}$$

Note that $((\alpha_1:\alpha_2),(\beta_1:\beta_2))=((\mu\alpha_1:\mu\alpha_2),(\nu\beta_1:\nu\beta_2))$ for any $\mu,\nu\in\bar{K}^*$. The advantage of completed coordinates is that in this setting each K-rational point of U corresponds to exactly one place of degree one in the function field $K(E)$, where E is the curve given by $ax_1^2+x_2^2=1+dx_1^2x_2^2$. The points of $E(K)$ may hence be identified bijectively with the K-rational points of U. The affine points of E obviously embed into U by $(\alpha,\beta)\mapsto((\alpha:1),(\beta:1))$. If d is a square in K, $d=s^2$, then $((1:s),(1:0))\in U$ and $((1:-s),(1:0))\in U$ express the two places at infinity that sit in the singular projective point $(0:1:0)$. If a/d is a square in K, $a/d=t^2$, then $((1:0),(t:1))$ and $((1:0),(-t:1))$ correspond to the places at infinity at $(1:0:0)$.

The computation of

$$((\alpha_1 : \alpha_2), (\beta_1 : \beta_2)) \oplus ((\gamma_1 : \gamma_2), (\delta_1 : \delta_2))$$

requires two formulas. One yields $((\mu_1 : \mu_2), (\nu_1 : \nu_2))$, and the other $((\mu'_1 : \mu'_2), (\nu'_1 : \nu'_2))$. Since $\mu_1\mu'_2 = \mu'_1\mu_2$ and $\nu_1\nu'_2 = \nu'_1\nu_2$, both formulas yield the same result if both of them belong to $\mathbb{P}^1 \times \mathbb{P}^1$. However, it may happen that $\mu_1 = \mu_2 = 0$ or $\nu_1 = \nu_2 = 0$. In such a case both (μ'_1, μ'_2) and (ν'_1, ν'_2) are distinct from $(0, 0)$, and $((\mu'_1 : \mu'_2), (\nu'_1 : \nu'_2))$ is the result of the addition. Similarly, if $\mu'_1 = \mu'_2 = 0$ or $\nu'_1 = \nu'_2 = 0$, then the result is $((\mu_1 : \mu_2), (\nu_1 : \nu_2))$. The formulas are as follows:

$$\begin{aligned}
\mu_1 &= \alpha_1\beta_2\gamma_2\delta_1 + \alpha_2\beta_1\gamma_1\delta_2, & \mu'_1 &= \alpha_1\beta_1\gamma_2\delta_2 + \alpha_2\beta_2\gamma_1\delta_1,\\
\mu_2 &= \alpha_2\beta_2\gamma_2\delta_2 + d\alpha_1\beta_1\gamma_1\delta_1, & \mu'_2 &= a\alpha_1\beta_2\gamma_1\delta_2 + \alpha_2\beta_1\gamma_2\delta_1,\\
\nu_1 &= \alpha_2\beta_1\gamma_2\delta_1 - a\alpha_1\beta_2\gamma_1\delta_2, & \nu'_1 &= \alpha_1\beta_1\gamma_2\delta_2 - \alpha_2\beta_2\gamma_1\delta_1,\\
\nu_2 &= \alpha_2\beta_2\gamma_2\delta_2 - d\alpha_1\beta_1\gamma_1\delta_1, & \nu'_2 &= \alpha_1\beta_2\gamma_2\delta_1 - \alpha_2\beta_1\gamma_1\delta_2.
\end{aligned} \tag{E.7}$$

Let us now observe how these formulas correspond to formulas (E.3) and (E.5). Let $\sigma = (\sigma_1, \sigma_2)$ and $\tau = (\tau_1, \tau_2)$. By (E.3) and (E.5) $\sigma \oplus \tau$ is equal to

$$\left(\frac{\sigma_1\tau_2 + \sigma_2\tau_1}{1 + d\sigma_1\sigma_2\tau_1\tau_2}, \frac{\sigma_2\tau_2 - a\sigma_1\tau_1}{1 - d\sigma_1\sigma_2\tau_1\tau_2}\right) \text{ and } \left(\frac{\sigma_1\sigma_2 + \tau_1\tau_2}{\sigma_2\tau_2 + a\sigma_1\tau_1}, \frac{\sigma_1\sigma_2 - \tau_1\tau_2}{\sigma_1\tau_2 - \sigma_2\tau_1}\right),$$

respectively.

Insert $\sigma, \tau \in \mathbb{A}^2$ into $\mathbb{P}^1 \times \mathbb{P}^1$ by

$$(\sigma_1, \sigma_2) \mapsto ((\sigma_1 : 1), (\sigma_2 : 1)) \text{ and } (\tau_1, \tau_2) \mapsto ((\tau_1 : 1), (\tau_2 : 1)).$$

Apply now (E.7) with $\alpha_1 = \sigma_1, \beta_1 = \sigma_2, \gamma_1 = \tau_1$, $\delta_1 = \tau_2$, and the other values being equal to 1. We obtain

$$\begin{aligned}
\mu_1 &= \sigma_1\tau_2 + \sigma_2\tau_1, & \mu'_1 &= \sigma_1\sigma_2 + \tau_1\tau_2,\\
\mu_2 &= 1 + d\sigma_1\sigma_2\tau_1\tau_2, & \mu'_2 &= a\sigma_1\tau_1 + \sigma_2\tau_2,\\
\nu_1 &= \sigma_2\tau_2 - a\sigma_1\tau_1, & \nu'_1 &= \sigma_1\sigma_2 - \tau_1\tau_2,\\
\nu_2 &= 1 - d\sigma_1\sigma_2\tau_1\tau_2, & \nu'_2 &= \sigma_1\tau_2 - \sigma_2\tau_1.
\end{aligned}$$

We see that rules (E.7) can be interpreted as a transformation of the main addition law (E.3) and the dual addition law (E.5) to projective points. However, in addition to that, rules (E.7) may be applied to points and places at infinity. For example consider $((1 : s), (1 : 0)) \oplus ((1 : -s), (1 : 0))$, where $s^2 = d$. Then $(\mu_1, \mu_2, \nu_1, \nu_2) = (0, d, -d, -d)$ and $(\mu'_1, \mu'_2, \nu'_1, \nu'_2) = (0, -d, 0, 0)$. Hence only the former quadruple may be used to compute the result of the addition. The result is

$$((0 : d), (-d : -d)) = ((0 : 1), (1 : 1)), \text{ i.e., the affine point } (0, 1).$$

Recall that $(0, 1)$ is the neutral element of the group. Points $((1 : s), (1 : 0))$ and $((1 : -s), (1 : 0))$ are thus opposite each to other.