Using Montgomery curve arithmetic over F_{p^2} for point scalar multiplication on short Weierstrass curve over F_p with exactly one 2-torsion point and order not divisible by 4

M. WROŃSKI
michal.wronski@wat.edu.pl

Military University of Technology, Faculty of Cybernetics
Institute of Mathematics and Cryptology
Kaliskiego Str. 2, 00-908 Warsaw, Poland

Montgomery curves are well known because of their efficiency and side channel attacks vulnerability. In this article it is showed how Montgomery curve arithmetic may be used for point scalar multiplication on short Weierstrass curve E_{SW} over F_p with exactly one 2-torsion point and $\#E_{SW}(F_p)$ not divisible by 4. If $P \in E_{SW}(F_p)$ then also $P \in E_{SW}(F_{p^2})$. Because $E_{SW}(F_{p^2})$ has three 2-torsion points (because $E_{SW}(F_p)$ has one 2-torsion point) it is possible to use 2-isogenous Montgomery curve $E_M(F_{p^2})$ to the curve $E_{SW}(F_{p^2})$ for counting point scalar multiplication on $E_{SW}(F_p)$. However arithmetic in F_{p^2} is much more complicated than arithmetic in F_p, in hardware implementations this method may be much more useful than standard methods, because it may be nearly 45% faster.

Keywords: elliptic curve cryptography, hardware implementations, Montgomery curves.

1. Introduction

There are many standards of elliptic curves, but not all of them allow to use fast point scalar mutliplication. Most of such elliptic curves over F_p, where p is large prime, are short Weierstrass curves $E_{SW}: y^2 = x^3 + ax + b$ with $a = -3$ (like all NIST curves over F_p) and with $\#E_{SW}$ prime. It is well known that there are special kinds of elliptic curves which allow to use much faster arithmetic, for example (twisted) Edwards curves, (twisted) Hessian curves and Montgomery curves. All of these curves always have order of points divisible by small integer bigger than 1 ((twisted) Hessian curves by 3, Montgomery and (twisted) Edwards curves by 4). That is why it is impossible to use for example Montgomery curve arithmetic to count point scalar multiplication on E_{SW} when $\#E_{SW}$ is prime.

Fortunately, this problem can be avoided by using field extension. It will be showed that for all short Weierstrass curves over F_p with exactly one 2-torsion point and $4 \nmid \#E_{SW}(F_p)$ it is possible to use 2-isogeny from $E_{SW}\left(F_{p^2}\right)$ to Montgomery curve $E_M\left(F_{p^2}\right)$. It will be also showed that using Montgomery ladder for point scalar multiplication on $E_M\left(F_{p^n}\right)$ may be almost 45% faster than using Brier–Joye ladder for elliptic curve $E_{SW}(F_p)$ for arbitrary a. Presented solution may be useful especially if it is possible to use parallel F_{p^2} arithmetic.

F_{p^2} arithmetic is useful for example in pairing or in GLS method. Such arithmetic is much faster than not-parallel F_{p^2} arithmetic (see [7]).

In this article will be presented how to speed-up arithmetic using XZ coordinates on short Weierstrass curve over F_p using 2-isogenous Montgomery curve over F_{p^2} using parallel F_{p^2} implementation.

2. Elliptic curves

Short Weierstrass curve

Elliptic curve in short Weierstrass form E_{SW} over field K, where $char(K) \neq 2$ is given by equation

$$E_{SW}: y^2 = x^3 + ax + b \qquad (1)$$

where $4a^3 + 27b^2 \neq 0$.

Motgomery curve

Montgomery curve E_M over field K, where $char(K) \neq 2$ is given by equation

$$E_M : by^2 = x^3 + ax^2 + x \tag{2}$$

where $b(a^2 - 4) \neq 0$. Every Montgomery curve $E_M(K)$ may be transformed into $E_{SW}(K)$. We should remember that it is not always possible to transform $E_{SW}(K)$ into $E_M(K)$.

3. Using ladder for point scalar multiplication

Arithmetic on short Weierstrass curve over F_p is the fastest if $a = -3$. Many of elliptic curves which are standards have $a = -3$. In hardware implementations it is often required to use solutions vulnerable for side channel attacks. The simplest way to achieve this condition is using Brier–Joye ladder for short Weierstrass curves and Montgomery ladder for Montgomery curves.

Very important fact is that both of these ladders use XZ coordinates, so to get Y (if necessary) it is needed to solve quadratic modular equation at the end of computations. There is one important difference between Brier–Joye ladder and Montgomery ladder. Brier–Joye ladder for arbitrary a requires 20 multiplications per step and Montgomery ladder requires only 11 multiplications per step. Because multiplication is crucial operation in F_q arithmetic, it is expected that using Montgomery ladder may be $100\% - \frac{11}{20} \cdot 100\% = 45\%$ faster than using Brier–Joye ladder.

Brier–Joye ladder

Brier–Joye ladder was firstly described in [2]. This method uses differential addition, so in every step from pair of points $([m]P, [m+1]P)$ is computed $\left([2m]P, [2m+1]P\right)$ if bit which is analyzed is equal to 0 or $\left([2m+1], [2m+2]P\right)$, if bit which is analyzed is equal to 1.

For arbitrary a and with assumptions that $b_4 = 4b$ and $X_1 = x_P$, $Z_1 = 1$, where $P = \left(x_P, y_P\right)$ is the generator, point scalar multiplication using Montgomery ladder may be computed as follow (presented formulas may be found in [4]):

$$XX = X_2^2 \tag{3}$$

$$ZZ = Z_2^2 \tag{4}$$

$$aZZ = a \cdot ZZ \tag{5}$$

$$E = (X_2 + Z_1)^2 - XX - ZZ \tag{6}$$

$$X_4 = (XX - aZZ)^2 - b_4 \cdot E \cdot ZZ \tag{7}$$

$$Z_4 = 2 \cdot E \cdot (XX + aZZ) + b_4 \cdot ZZ^2 \tag{8}$$

$$A = X_2 \cdot X_3 \tag{9}$$

$$B = Z_2 \cdot Z_3 \tag{10}$$

$$C = X_2 \cdot Z_3 \tag{11}$$

$$D = X_3 \cdot Z_2 \tag{12}$$

$$X_5 = (A - a \cdot B)^2 - b_4 \cdot B \cdot (C + D) \tag{13}$$

$$Z_5 = X_1 \cdot (C - D)^2 \tag{14}$$

These operations require 20 multiplications (8 multiplications, 7 squares, 5 multiplications by constant) and 12 additions/subtractions (11 additions/subtractions and one multiplication by 2).

Montgomery ladder

Montgomery ladder was firstly described in [5]. This method also uses differential addition, so in every step from pair of points $([m]P, [m+1]P)$ is computed $\left([2m]P, [2m+1]P\right)$ if bit which is analyzed is equal to 0 or $\left([2m+1], [2m+2]P\right)$, if bit which is analyzed is equal to 1. There are also assumptions that $4a_{24} = a + 2$. Then:

$$A = X_2 + Z_2 \tag{15}$$

$$AA = A^2 \tag{16}$$

$$B = X_2 - Z_2 \tag{17}$$

$$BB = B^2 \tag{18}$$

$$E = AA - BB \tag{19}$$

$$C = X_3 + Z_3 \tag{20}$$

$$D = X_3 - Z_3 \tag{21}$$

$$DA = D \cdot A \tag{22}$$

$$CB = C \cdot B \tag{23}$$

$$X_5 = Z_1 \cdot (DA + CB)^2 \tag{24}$$

$$Z_5 = X_1 \cdot (DA - CB)^2 \tag{25}$$

$$X_4 = AA \cdot BB \tag{26}$$

$$Z_4 = E \cdot (BB + a_{24} \cdot E) \tag{27}$$

These operations require 11 multiplications (6 multiplications, 4 squares and 1 multiplication by constant) and 8 additions/subtractions.

4. 2-isogeny from $E_{SW}(F_p{}^2)$ to $E_M(F_p{}^2)$

It is well known that if K is finite field, then every Montgomery curve $E_M(K)$ has it order $\#E_M(K)$ divisible by 4. Let see that if short Weierstrass curve $E_{SW}(K)$ has 2-torsion point $P=(x,y)$, then for such point always $y=0$. So in the case when the $E_{SW}(K)$ has exactly one 2-torsion point, then equation $x^3+ax+b=0$ has exactly one root in field K. If $K=F_p$ the root is also in F_p. It is showed in [1], that if curve E_{SW} has three 2-torsion points, it is possible to find for such curve 2-isogenous Montgomery curve. Now let see that if curve $E_{SW}(F_p)$ has exactly one 2-torsion point then $2\,|\,\#E_{SW}(F_p)$. But there was also made assumption that $4\nmid\#E_{SW}(F_p)$, so it is impossible to find for $E_{SW}(F_p)$ isomorphic or 2-isogenous $E_M(F_p)$.

Now let see that if $E_{SW}(F_p)$ has exactly one 2-torsion point then $x^3+ax+b=(x-r_0)(x^2+c_1x+c_0)$, where $c_0,c_1\in F_p$. It is easy to see that equation

$$x^2+c_1x+c_0 \tag{28}$$

cannot have roots in F_p because $E_{SW}(F_p)$ has only one 2-torsion point.

Now let see that using field F_{p^2} which is 2-nd degree field extension of field F_p, the equation (28) will have two roots $r_1,r_2\in F_{p^2}$ (see [6]). Because $r_0\in F_p$, then also $r_0\in F_{p^2}$, so equation (28) has exactly three roots in F_{p^2}.

It means that for curve $E_{SW}(F_{p^2})$ it is possible to find 2-isogenous curve $E_M(F_{p^2})$. Below is showed how to find such 2-isogeny $\phi:E_{SW}(F_{p^2})\rightarrow E_M(F_{p^2})$.

Firstly, there must be found all roots of polynomial given by equation (28). Then:

$$x^3+ax+b=(x-r_0)(x-r_1)(x-r_2) \tag{29}$$

So $r_0,r_1,r_2\in F_{p^2}$, $x_1,y_1,a,b\in F_p$ (but formally of course also $x_1,y_1,a,b\in F_{p^2}$) and:

$$R_0=0 \tag{30}$$

$$R_1=r_1-r_0 \tag{31}$$

$$R_2=r_2-r_0 \tag{32}$$

Now it is possible to construct elliptic curve E_2 which is isomorphic to E_1:

$$E_2: y^2=x^3-(R_1+R_2)x^2+R_1R_2x \tag{33}$$

with point

$$P_2=(x_2,y_2)=(x_1-r_0,y_1) \tag{34}$$

Then may be found curve E_3 which is 2-isogenous to E_2:

$$E_3: y^2=x^3+2(R_1+R_2)x^2+(R_1-R_2)^2x \tag{35}$$

with point

$$P_3=(x_3,y_3)=\left(\frac{y_2^2}{x_2^2},\frac{y_2(R_1R_2-x_2{}^2)}{x_2{}^2}\right) \tag{36}$$

Now it is easy to find Montgomery curve which is isomorphic to E_3:

$$E_4:\frac{1}{R_1-R_2}y^2=x^3+\frac{2(R_1+R_2)}{R_1-R_2}x^2+x \tag{37}$$

with point:

$$P_4=\left(\frac{x_3}{R_1-R_2},\frac{y_3}{R_1-R_2}\right). \tag{38}$$

Finally, it is easy to see that curve E_4 is 2-isogenous to curve E_1.

It is possible to use similar transformations to find dual isogeny $\phi^{-1}:E_M(F_{p^2})\rightarrow E_{SW}(F_{p^2})$

$$E_4:\frac{1}{R_1-R_2}y^2=x^3+\frac{2(R_1+R_2)}{R_1-R_2}x^2+x \tag{39}$$

with point

$$\overline{P_4}=(\overline{x_4},\overline{y_4}) \tag{40}$$

$$E_3: y^2=x^3+2(R_1+R_2)x^2+(R_1-R_2)^2x \tag{41}$$

with point

$$\begin{aligned}\overline{P_3}&=(\overline{x_3},\overline{y_3})=\\&=(\overline{x_4}(R_1-R_2),\overline{y_4}(R_1-R_2))\end{aligned} \tag{42}$$

$$E_2: y^2=x^3-(R_1+R_2)x^2+R_1R_2x \tag{43}$$

with point

$$\begin{aligned}\overline{P_2}&=(\overline{x_2},\overline{y_2})=\\&=\left(\frac{\overline{y_3}^2}{4\overline{x_3}^2},\frac{\overline{y_3}((R_1-R_2)^2-\overline{x_3}^2)}{8\overline{x_3}^3}\right)\end{aligned} \tag{44}$$

And finally:

$$E_1: y^2=x^3+ax+b \tag{45}$$

with point

$$\overline{P_1}=(\overline{x_1},\overline{y_1})=(\overline{x_2}+r_0,\overline{y_2}) \tag{46}$$

So for isogeny $\phi: E_1 \to E_4$ there is:

$$\phi(P_1) = P_4 = (x_4, y_4) = \tag{47}$$

$$= \left(\frac{y_1^2}{(x_1 - r_0)^2 \cdot (R_1 - R_2)}, \frac{y_1 \cdot (R_1 \cdot R_2 - (x_1 - r_0)^2)}{(x_1 - r_0)^2 \cdot (R_1 - R_2)} \right)$$

And of course for dual isogeny $\phi^{-1}: E_4 \to E_1$, there is:

$$\phi(\overline{P_4}) = \overline{P_1} = (\overline{x_1}, \overline{y_1}) = \tag{48}$$

$$= \left(\frac{\overline{y_4}^2}{4 \cdot \overline{y_4}^2} + r_0, \frac{(\overline{y_4} \cdot ((R_1 - R_2)^2 - \overline{x_4}^2 \cdot (R_1 - R_2)^2))}{8 \cdot \overline{x_4}^3 \cdot (R_1 - R_2)^2} \right)$$

It should be stated that because 2-isogeny has degree 2, then for every $P \in E_{SW}(F_{p^2})$ and $\overline{P} \in E_M(F_{p^2})$ there is always:

$$\phi^{-1}(\phi(P)) = [2]P \tag{49}$$

and:

$$\phi(\phi^{-1}(\overline{P})) = [2]\overline{P} \tag{50}$$

Let's see that using XZ coordinates it is lost information about y so this value is not known.

If it is not needed to know y it is the end of computations.

If it is needed to know the value y, it may be found because the value of y using equation (2) for Montgomery curve:

$$y^2 = \frac{x^3 + ax^2 + x}{b} \tag{51}$$

Because there are two possible solutions for y during computations using ladder information about the least significant bit of y should be remembered. Then it is easy to find which value of y is the proper one.

Now let see that procedure for using Montgomery curve arithmetic to count point scalar multiplication on $E_{SW}(F_p)$ with exactly one 2-torsion point and $4 \nmid \#E_{SW}(F_p)$ should be as follow:

- Extend field F_p to F_{p^2}. Then for generator $P_1 \in E_{SW}(F_p)$ there is also $P_1 \in E_{SW}(F_{p^2})$.
- For given generator $P_1 \in E_{SW}(F_{p^2})$ for which point scalar multiplication by k is required (so $Q_1 = [k]P_1$) use isogeny ϕ to obtain $P_4 = \phi(P_1), P \in E_M(F_{p^2})$.
- Use Montgomery ladder to obtain $Q_4 = [\frac{k}{2}]P_4$, where $\frac{k}{2} = k \cdot 2^{-1} (\mathrm{mod}\, Ord(P_1))$.
- Use the dual isogeny ϕ^{-1} to obtain: $Q_1 = \phi^{-1}(Q_4) = \phi^{-1}([\frac{k}{2}]P_4) =$ $= \phi^{-1}([\frac{k}{2}](\phi P_1)) = [\frac{k}{2}](\phi^{-1}(\phi P_1)) =$ $= [\frac{k}{2}]([2]P_1) = [k]P_1.$
- It is easy to see that the result $Q_1 \in E_{SW}(F_p)$ instead of that the computations are made in F_{p^2}.

5. Arithmetic in F_{p^2}

It is showed in [7] that in FPGA F_{p^2} arithmetic may be efficiently implemented using Karatsuba algorithm. Such implementation results in that all operations in F_{p^2} are almost as fast as in F_p. The biggest disadvantage of this solution is that F_{p^2} requires much more resources than F_p arithmetic. For more information about F_{p^2} arithmetic see [7].

6. Complexity of proposed solution

Speed of presented solution depends mostly on the time of multiplication in F_p which is denoted by T_M and on the number of additional processor cycles N_A required to compute multiplication in F_{p^2} Of course, the bigger is the degree of used field extension, the more resources is needed in hardware implementations.

Time required for point scalar multiplication in subgroup of order r using Brier–Joye ladder is equal to:

$$T_{BJL} = \left(\lfloor \log_2 r \rfloor + 1\right) \cdot \left(20T_M + 12\right) \tag{52}$$

Time required for point scalar multiplication of order r using Montgomery ladder over F_{p^n} is equal to:

$$T_{ML} = \left(\lfloor \log_2 r \rfloor + 1\right) \cdot \left(11\left(T_M + N_A\right) + 8\right) \tag{53}$$

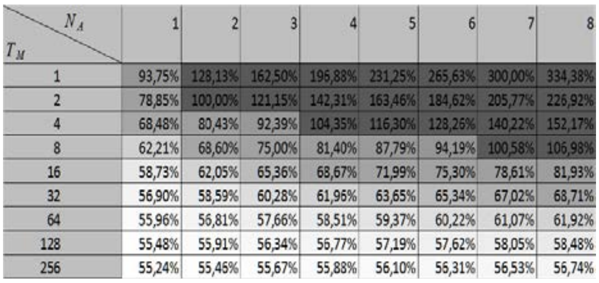
So let:

$$W(T_M, N_A) = \frac{T_{ML}}{T_{BJL}} \cdot 100\% = \qquad (54)$$

$$= \frac{11\left(T_M + N_A\right) + 8}{20T_M + 12} \cdot 100\%$$

In the table below we present values of W for different T_M and N_A.

Tab. 1. Values of $W(T_M, N_A)$ for different number of processor cycles of T_M and different number of additions N_A required for multiplication in F_{p^2}.



T_M \ N_A	1	2	3	4	5	6	7	8
1	93,75%	128,13%	162,50%	196,88%	231,25%	265,63%	300,00%	334,38%
2	78,85%	100,00%	121,15%	142,31%	163,46%	184,62%	205,77%	226,92%
4	68,48%	80,43%	92,39%	104,35%	116,30%	128,26%	140,22%	152,17%
8	62,21%	68,60%	75,00%	81,40%	87,79%	94,19%	100,58%	106,98%
16	58,73%	62,05%	65,36%	68,67%	71,99%	75,30%	78,61%	81,93%
32	56,90%	58,59%	60,28%	61,96%	63,65%	65,34%	67,02%	68,71%
64	55,96%	56,81%	57,66%	58,51%	59,37%	60,22%	61,07%	61,92%
128	55,48%	55,91%	56,34%	56,77%	57,19%	57,62%	58,05%	58,48%
256	55,24%	55,46%	55,67%	55,88%	56,10%	56,31%	56,53%	56,74%

7. Results

Using Montgomery curve arithmetic over F_{p^2} it is possible to compute point scalar multiplication on short Weierstrass curve over $E_{SW}(F_p)$ with exactly one 2-torsion point and $4 \nmid \#E_{SW}(F_p)$, up to 45% faster than using standard methods.

It should be also stated that implementation of elliptic curve arithmetic in F_{p^2} requires much more resources than arithmetic on elliptic curve over F_p.

Implementation of Montgomery arithmetic (with parallel addition and multiplication) in F_{p^2} should require up to four times more resources than classic solutions, so presented solution should be used only in some situations.

This method is suitable for example if it is needed to have arithmetic on two elliptic curves with different levels of security. Then one curve may use GLS (for details see [3]) method for which it is good to use parallel F_{p^2} arithmetic. Then the arithmetic on the second curve with smaller level of security (for example about $\sqrt{p}$), may be implemented with presented solution using the same F_{p^2} arithmetic which is used for the first curve.

8. Bibliography

[1] Birkner P., Joye M., Lange T., Peters Ch., Bernstein D., “Twisted Edwards Curves”, *eprint.iacr.org/2008/013*, 2008.

[2] Brier E., Joye M., “Weierstraß Elliptic Curves and Side-Channel Attacks”, *Lecture Notes in Computer Science*, Vol. 2274, 335–345 (2002).

[3] Galbraith S., Lin X., Scott M., “Endomorphisms for Faster Elliptic Curve Cryptography on a Large Class of Curves”, *eprint.iacr.org/2008/194*, 2008.

[4] Izu T., Takagi T., “A Fast Parallel Elliptic Curve Multiplication Resistant against Side Channel Attacks”, *Lecture Notes in Computer Science*, Vol. 2274, 280–296 (2002).

[5] Montgomery P., “Speeding the Pollard and elliptic curve methods of factorization”, *Mathematics of Computation*, Vol. 48, 243–264 (1987).

[6] Neunhöffer M., *Module MT 5826 Finite Fields*, RWTH Aachen University, Aachen, 2007.

[7] Wroński M., “Faster Point Scalar Multiplication on Short Weierstrass Elliptic Curves over F_p using Twisted Hessian Curves over F_{p^2}”, *Journal of Telecommunications and Information Technology*, Issue 3, 98–102 (2016).

Zastosowanie arytmetyki krzywych Montgomery'ego nad ciałem F_{p^2} w celu obliczenia krotności punktu na krzywej eliptycznej w skróconej postaci Weierstrassa nad ciałem F_p z dokładnie jednym punktem 2-torsyjnym i rzędem grupy punktów niepodzielnym przez 4

M. WROŃSKI

Krzywe Montgomery'ego są znane ze względu na efektywność wykonywanych na nich operacji i ich odporność na ataki typu „side channel". W artykule przedstawiono, w jaki sposób można wykorzystać arytmetykę krzywych Montgomery'ego w celu obliczenia krotności punktu na krzywej eliptycznej w skróconej postaci Weierstrassa E_{SW} nad ciałem F_p z dokładnie jednym punktem 2-torsyjnym oraz $\#E_{SW}(F_p)$ niepodzielnym przez 4. Jeżeli $P \in E_{SW}(F_p)$, wtedy również $P \in E_{SW}(F_p{}^2)$. Ponieważ $E_{SW}(F_p{}^2)$ posiada trzy punkty 2-torsyjne (wynika to z tego, że $E_{SW}(F_p)$ posiada jeden punkt 2-torsyjny), możliwe jest wykorzystanie krzywej Montgomery'ego $E_M(F_p{}^2)$ 2-izogenicznej do krzywej $E_{SW}(F_p{}^2)$, w celu obliczenia krotności punktu na krzywej eliptycznej na krzywej $E_{SW}(F_p)$. Jakkolwiek arytmetyka w ciałach $F_p{}^2$ jest bardziej skomplikowana niż arytmetyka w ciele F_p, w implementacjach sprzętowych metoda ta może być bardzo użyteczna i szybsza od metod klasycznych do 45%.

Słowa kluczowe: kryptografia oparta o krzywe eliptyczne, implementacje sprzętowe, krzywe Montgomery'ego.