Research Article

Andrea Basso and Fabien Pazuki*

On the supersingular GPST attack

<https://doi.org/10.1515/jmc-2021-0020>

received February 11, 2020; accepted August 01, 2021

Abstract: The main attack against static-key supersingular isogeny Diffie–Hellman (SIDH) is the Galbraith–Petit–Shani–Ti (GPST) attack, which also prevents the application of SIDH to other constructions such as non-interactive key-exchange. In this paper, we identify and study a specific assumption on which the GPST attack relies that does not necessarily hold in all circumstances. We show that in some circumstances the attack fails to recover part of the secret key. We also characterize the conditions necessary for the attack to fail and show that it rarely happens in real cases. We give a link with collisions in the Charles–Goren–Lauter (CGL) hash function.

Keywords: isogenies, supersingular elliptic curves, modular invariants

MSC 2020: 14H52, 14K02, 11T71, 94A60, 81P94, 65P25

1 Introduction

In 2011, Jao and De Feo [1] introduced supersingular isogeny Diffie–Hellman (SIDH), a post-quantum key exchange protocol that mimics the Diffie–Hellman protocol in the settings of isogenies between supersingular curves. Its security relies on the hardness of the supersingular isogeny path finding problem with additional torsion image information. A detailed study of its security is reported in ref. [2]. The authors also report three attacks under different scenarios. The first attack, the so-called GPST attack, is an active attack against static keys, where one of the two parties sends maliciously chosen torsion information that lead to the protocol failing or succeeding based on a single bit of the secret key of the other party. This can then be used to fully recover the secret key by repeating the process for each bit of the key.

The problem can be solved by applying the Fujisaki–Okamoto transform [3] to obtain a CCA-secure key-encapsulation mechanism called SIKE. However, this prevents static–static key exchanges that can be used to create a non-interactive key exchange (NIKE) protocol. Indeed, developing an efficient SIDH-based NIKE remains partially an open problem. There have been some proposed solutions, such as k -SIDH [4], but their performance and communication costs are almost impractically large.

The GPST attack thus plays an essential role in the security and the development of SIDH and other isogeny-based protocols. In this paper, we identify a specific property of the j -invariant of elliptic curves that is used by the GPST attacks which is not guaranteed to hold in all circumstances. In such circumstances, we show the attack fails to recover the secret key. We then characterize these specific circumstances and show they occur with negligible probability in real cases. We conclude by analyzing special starting curves, such as those used in SIDH.

* **Corresponding author: Fabien Pazuki**, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark, e-mail: fpazuki@math.ku.dk

Andrea Basso: Centre for Cyber Security and Privacy, School of Computer Science, University of Birmingham, Birmingham, UK, e-mail: a.basso@cs.bham.ac.uk

1.1 Preliminaries

For an elliptic curve E defined over a field k , we denote the neutral element of the elliptic curve by O and the modular invariant of E by $j(E)$. The invariant is an element of k that characterizes the $\bar{k}$ -isomorphism class of E . For $P \in E(k)$ rational over k , we denote by $\langle P \rangle$ the subgroup of $E(k)$ generated by P .

The parameters of the SIDH protocol are the following:

- A prime $p = 2^n 3^m f - 1$, where n, m, f are natural numbers, f is small and $2^n \approx 3^m$.
- A supersingular elliptic curve E_0 defined over $\mathbb{F}_{p^2}$.
- Points $P_A, Q_A \in E_0$ which form a basis of $E_0[2^n]$ and points $P_B, Q_B \in E_0$ which form a basis of $E_0[3^m]$.

We refer the reader to the original paper [2] for the presentation of the attack. We introduce here the two assumptions that are required for the GPST attack to take place:

- (a) Alice uses a static key (a_1, a_2) . The values a_1, a_2 are elements of $\mathbb{Z}/2^n\mathbb{Z}$, not both divisible by 2.
- (b) The attacker has access to an oracle O such that, if E, E' denote supersingular elliptic curves defined over $\mathbb{F}_{p^2}$ and R, S denote any two points on E ,

$$O(E, R, S, E') = \begin{cases} \text{true}, & \text{if } j(E/\langle [a_1]R + [a_2]S \rangle) = j(E'), \\ \text{false}, & \text{otherwise.} \end{cases} \quad (1)$$

Such an oracle can be realized in practice by having a mechanism in place that ensures the two parties obtain the same shared key.

The attacker does not need any additional information, which makes the GPST attack one of the most powerful attacks against SIDH known in the literature.

2 Attacking the attack

Alice's static key (a_1, a_2) is always equivalent (i.e., it leads to the same shared secret) to a key of the form $(1, \alpha)$ or $(\alpha, 1)$, where α is again an element of $\mathbb{Z}/2^n\mathbb{Z}$ [2, Lemma 2.1]. Without loss of generality, we may assume we are in the former case.

Each iteration of the GPST attack relies on the following implication, used in the paragraph *first step of the attack* of ref. [2]:

$$(j(E/\langle T_1 \rangle) = j(E/\langle T_2 \rangle)) \Rightarrow (\langle T_1 \rangle = \langle T_2 \rangle) \quad (2)$$

to recover a single bit of the key. In particular, the oracle query (1) is used to recover whether $\langle R' + \alpha S' \rangle$ is equal to $\langle R + \alpha S \rangle$, where R, S, R', S' are points on an elliptic curve computed by the attacker. The points R' and S' are chosen such that $\langle R' + \alpha S' \rangle = \langle R + \alpha S \rangle$ implies the i th bit α_i of the key is a zero-bit. Thus, if the oracle returns *true*, then $\alpha_i = 0$, otherwise $\alpha_i = 1$.

However, implication (2) does not hold in all cases, because there exist pairs of elliptic curves E_1, E_2 with multiple isogenies between them. Thus, if ϕ_1 and ϕ_2 are distinct cyclic isogenies between E_1 and E_2 and K_1 and K_2 are their respective kernels, we have that E_1/K_1 is isomorphic to E_1/K_2 because they are both isomorphic to E_2 , but K_1 and K_2 may not be equal and need not even be isomorphic. Over $\mathbb{Q}$, a classical example can be found in [5, page 110], namely the pair (E, E) , where E is the curve defined over $\mathbb{Q}$ with $j(E) = 8,000$, which is given by the affine Weierstrass equation $y^2 = x^3 + 4x^2 + 2x$. This curve E admits a rational 2-isogeny φ to itself (hence cyclic, and given explicitly in *loc. cit.*), and one computes $j(E/\ker \varphi) = 8,000 = j(E)$, hence E and $E/\ker \varphi$ are in fact $\bar{\mathbb{Q}}$ -isomorphic, but clearly $\ker \varphi \neq \{O\}$.

Over a finite field of characteristic $p > 0$, if one starts with a supersingular elliptic curve E , then for any prime $\ell \neq p$, for any positive integer m and for any cyclic subgroup G of order ℓ^m , the elliptic curve E/G will also be supersingular, because there are still no points of order p on E/G . As there are only finitely many

supersingular curves over $\overline{\mathbb{F}}_p$ (see for instance [6, Theorem 4.1, pages 148–149]), there exist pairs of cyclic groups (G_1, G_2) where G_1 and G_2 are not isomorphic, such that E/G_1 is isomorphic to E/G_2 . This argument is used to prove that the endomorphism ring of supersingular elliptic curves is an order in a quaternion algebra, see [6, page 146] or [7, page 267].

Let us give now a detailed example where the GPST attack fails to recover the private key of a SIDH key exchange. The example was found through direct search with small primes and it has been experimentally validated.

Example 1. Let $p = 2^5 3^3 - 1 = 863$ be a prime and k be the finite field $\mathbb{F}_{p^2}$, considered as $\mathbb{F}_p(\beta)$, where β satisfies the quadratic equation $\beta^2 - \beta + 5 = 0$. Let E_0 be the supersingular elliptic curve with affine Weierstrass model $y^2 = x^3 + (531\beta + 538)x + (720\beta + 375)$ over k . Consider the points

$$\begin{aligned} P_A &= (834\beta + 726, 642\beta + 130), & Q_A &= (583\beta + 276, 180\beta + 854), \\ P_B &= (254\beta + 697, 516\beta + 268), & Q_B &= (753\beta + 317, 234\beta + 532). \end{aligned}$$

The points P_A, Q_A form a basis of $E_0[2^5]$, and P_B, Q_B form a basis of $E_0[3^3]$. Let Alice's key be of the form $(1, \alpha)$, with $\alpha = 10$ (written in base ten). The value α can also be expressed in bits as 1010 (written in base two). From now on, by a slight abuse of notation we refer to α as the key.

Let us now carry on the GPST attack. Assume the randomly generated values b_1, b_2 are $b_1 = 1, b_2 = 6$ (for simplicity, we assume b_1 and b_2 stay constant across iterations, but we only need $b_1 = 1, b_2 = 6$ in the second round for the attack to fail). Let ϕ_A be the isogeny with kernel $K_A = \langle P_A + [\alpha]Q_A \rangle$. Then $E_A = E_0/K_A$ has affine Weierstrass model $y^2 = x^3 + (40\beta + 535)x + (720\beta + 768)$. Let us begin the attack and let ϕ_B be the isogeny with kernel $K_B = \langle P_B + [6]Q_B \rangle$. Thus, $E_B = E_0/K_B$ has affine Weierstrass model $y^2 = x^3 + 105x + 254$. Furthermore, we have

$$\begin{aligned} R &= \phi_B(P_A) = (151\beta + 257, 594\beta + 2), \\ S &= \phi_B(Q_A) = (98\beta + 386, 286\beta + 58). \end{aligned}$$

Note that since the degree of ϕ_B is coprime with the order of P_A and Q_A , the points R and S have also order 2^5 .

The first iteration of the attack starts by computing

$$\begin{aligned} \theta_0 &= \sqrt{(1 + 2^4)^{-1}} \pmod{2^5} = 7, \\ R'_0 &= [\theta_0]R = (527\beta + 129, 700\beta + 163), \\ S'_0 &= [\theta_0][1 + 2^4]S = (164\beta + 377, 566\beta + 641). \end{aligned}$$

It then proceeds by querying the oracle with $O(E_B, R'_0, S'_0, E_{AB}) = E_B/\langle \phi_A(P_B) + [6]\phi_A(Q_B) \rangle$. The curve $E_{AB} : y^2 = x^3 + 698x + (516\beta + 605)$ has j -invariant 117. The curve $E_B/\langle R'_0 + [\alpha]S'_0 \rangle$ also has j -invariant 117, thus the oracle response is *true* and the first (least significant) bit of the key is a zero-bit. Hence, the attack correctly obtains the first bit of the key.

For the second round, the attacker computes

$$\begin{aligned} \theta_1 &= \sqrt{(1 + 2^3)^{-1}} \pmod{2^5} = 5, \\ R'_1 &= [\theta_1]R = (97\beta + 261, 795\beta + 545), \\ S'_1 &= [\theta_1][1 + 2^3]S = (718\beta + 214, 450\beta + 844), \end{aligned}$$

and queries the oracle on $O(E_B, R'_1, S'_1, E_{AB})$. As before, the curves E_{AB} and $E_B/\langle R'_1 + [\alpha]S'_1 \rangle$ both have j -invariant 117, thus the oracle responds *true*. Hence, this time the attacker incorrectly deduces the second bit of the key to be zero.

Brute-forcing the remaining bits does not yield any solution, since there is no key α ending with two zero-bits (00) such that $E_0/\langle P_A + [\alpha]Q_A \rangle$ is isomorphic to E_A .

Note that each iteration of the attack depends on the previous one since the modified value R'_i of iteration $i + 1$ relies on K_i , whose last bit is obtained in the i th iteration. To see how the error propagates,

assume the attack fails at iteration i . Thus α_i , the i th bit of the key, is a one-bit but the attacker deduces it to be a zero-bit. It follows that $K_i = K_{i-1}$ for the attacker since prepending zero-bits does not affect the value of K . The attacker then computes the points

$$R'_i = [\theta]R - [\theta][2^{n-i-1}K_{i-1}]S, \quad S'_i = [\theta][1 + 2^{n-i-1}]S$$

and queries the oracle on $O(E_B, R'_i, S'_i, E_{AB})$.

Thus, omitting θ since it does not affect the subgroup, we have

$$\begin{aligned} \langle R'_i + [\alpha]S'_i \rangle &= \langle R + [-2^{n-i-1}K_{i-1} + \alpha(1 + 2^{n-i-1})]S \rangle \\ &= \langle R + [\alpha]S + [-2^{n-i-1}K_{i-1} + 2^{n-i-1}(K_{i-1} + 2^{i-1} + 2^i\alpha_i + 2^{i+1}\alpha')]S \rangle \\ &= \langle R + [\alpha]S + [2^{n-2} + 2^{n-1}\alpha_i]S \rangle \\ &\neq \langle R + [\alpha]S \rangle. \end{aligned}$$

To see why the inequality in the last line holds, consider that if the two cyclic subgroups are the same, their generators must be scalar multiples of each other. Assume then $R + [\alpha]S + [2^{n-2} + 2^{n-1}\alpha_i]S = [h](R + [\alpha]S)$ for a natural integer h . The linear independence of R and S implies that $R = [h]R$, which means $h \equiv 1 \pmod{2^n}$. Thus, the equality on S implies that $[2^{n-2} + 2^{n-1}\alpha_i]S = O$, which is impossible because $\alpha_i \in \{0, 1\}$ and we have neither $0 \neq 2^{n-2} \pmod{2^n}$ nor $0 \neq -2^{n-1} \pmod{2^n}$. This means that the oracle will return *false* (unless the two subgroups, $\langle R + [\alpha]S \rangle$ and $\langle R'_i + [\alpha]S'_i \rangle$, give rise again to two isogenies with isomorphic codomain).

The subsequent iterations further propagate the error and the oracle will respond *false* to every further query (unless, again, the isogenies with different kernels have the same codomain, which is believed to be a rare occurrence). Thus, if the attack fails at the i th iteration, all the following key bits are deduced to be a one-bit. This allows the attacker to approximately identify the part of the key that has been correctly deduced and target, with a different choice of ϕ_B , only the remaining part.

3 Failure case conditions

3.1 General analysis and a lemma for 1,728

In this paragraph, we investigate the specific conditions that lead to the attack failing. Example 1 in the previous section establishes that there exist cases where the attack fails. The following are sufficient conditions (when considered together) for the attack to fail at the i th iteration of the attack, when we choose $1 \leq i < n - 3$ (the attack brute-forces the last three bits of the key and thus cannot fail after $n - 3$):

- (i) There exist two distinct isogenies ϕ_1 and ϕ_2 of degree 2^n between E_B and E_{AB} .
- (ii) One has $R + [\alpha]S = P_1$ and $R'_i + [\alpha]S'_i = P_2$, where P_1 and P_2 are generators of the kernel of ϕ_1 and ϕ_2 , respectively.
- (iii) The i th key bit α_i is 1.

From the point of view of the attacker, determining whether condition (i) is satisfied is computationally hard since it is hard to compute isogenies between two given curves. Moreover, the points P_1 and P_2 are dependent on the key α and thus unknown to the attacker. Hence, while the attacker influences the points P_1 and P_2 by choosing ϕ_B , they cannot know whether his choice of ϕ_B would cause the attack to fail.

Condition (ii) ensures that the attack is taking place in the case where the points R and S give rise to the isogenies considered in the first condition. Given the points P_1, P_2 , we have

$$\begin{cases} P_1 = R + [\alpha]S, \\ P_2 = R'_i + [\alpha]S'_i = [\theta_i]R + [\theta_i][\alpha + \alpha_i 2^{n-i-1}]S, \end{cases}$$

thus $[\theta_i]P_1 - P_2 = [\theta_i][2^{n-i-1}][\alpha_i]S = [\theta_i][2^{n-i-1}]S$ (because of condition (iii)). If a θ_i exists such that the previous equation is satisfied, then suitable points R and S also exist.

Condition (iii) is necessary because the attack fails by incorrectly deducing a zero-bit instead of a one-bit.

If ϕ_1 and ϕ_2 are such isogenies, then $\hat{\phi}_1 \circ \phi_2$ is an endomorphism of degree 2^{2n} in $\text{End}(E_B)$, and if it is not a classical multiplication, this gives rise to non-trivial short cycles. Thus, studying the endomorphism ring of the starting curve can reveal useful information on the conditions that lead to failures of the attack.

One curve that is often used is the curve with j -invariant 1,728. In real cases, we have that $2^n \approx p^{1/2}$. Thus, thanks to a remark from the referee, we can show that when the curve E_B has j -invariant 1,728 multiple isogenies between E_B and E_{AB} are possible when the isogeny paths only overlap in E_B and E_{AB} . We do so by showing that multiedges shorter than 2^n are not possible.

Lemma 1. *Let p be a prime such that $p \equiv 3 \pmod{4}$ and $\frac{p+1}{4} > 2^{2e}$ for some integer e , and let E be a supersingular curve over $\mathbb{F}_p$. If $j(E) = 1,728$, there is no endomorphism of norm 2^{2e} in $\text{End}(E)$ besides the classical multiplications.*

Proof. By Lemma 2 of [8], $\text{End}(E)$ is isomorphic to an order in a quaternion algebra $B(-1, -p)$. The order has basis $\left\{1, i, \frac{1+j}{2}, i\frac{1+j}{2}\right\}$, where $i^2 = -1$ and $j^2 = -p$. Let ϕ be an endomorphism of E of norm 2^{2e} , and write $\phi = w + xi + y\frac{1+j}{2} + zi\frac{1+j}{2}$. The norm of ϕ is given by $w^2 + x^2 + (y^2 + z^2)\frac{p+1}{4}$. Thus, $y^2 + z^2 = 0$, which implies $y = z = 0$ and $w^2 + x^2 = 2^{2e}$. By considering the last equation modulo 4, we have that $w \equiv x \equiv 0$ since the only squares modulo 4 are $\{0, 1\}$. This means both sides can be divided by 4, and if the process is repeated e times, we see that either $w = 2^e$ and $x = 0$, or $w = 0$ and $x = 2^e$. $\square$

3.2 Collisions in the CGL hash function and SIDH parameters

There is an important link between these failures of the GPST attack and the following work, which helps in understanding when these failures occur. From a cryptographic perspective, the CGL function [9] is a hash function based on isogenies between supersingular curves. It constructs a hash value by traveling the supersingular graph, choosing the next curve at each step depending on a single bit of the input. The mixing property of the supersingular graphs ensures the desired properties of hash functions. We note that having two distinct isogenies of the same degree between two supersingular curves leads to a collision in the CGL hash function.

The authors of ref. [9] showed that such collisions can be avoided by carefully choosing the base prime p . As explained in their paragraph 5.3.4, if a cycle is represented by an element x of $\text{End}(E)$ of norm ℓ^{2e} , the polynomial $X^2 + \text{Tr}(x)X + \ell^{2e}$ is irreducible and there are only finitely many values of $\text{Tr}(x)$ such that the discriminant $\text{Tr}(x)^2 - 4\ell^{2e}$ is negative. For each of these values, we obtain an imaginary quadratic field generated by the quadratic polynomial. It is thus sufficient to pick a prime p which splits in all those fields to guarantee that no cycle of length $2e$ exists. This can be done for sufficiently many values of e to guarantee that no short cycle exists.

More specifically to SIDH, Ghantous [10] has analyzed the presence of multiedges in the supersingular ℓ -isogeny graph. Theorem 3.2 of ref. [10] shows that the 2-isogeny graph has no multiedges only if the base prime $p \equiv 1, 121, 169, 289, 361, 529 \pmod{840}$. Similar modular conditions exist for the 3-isogeny graph. However, they do not apply to SIDH parameters. Ghantous [10] has also proved that the ℓ -isogeny graph has at most $\ell^2 + \ell/4$ multiedges. Thus, there is an upper bound in the 2-isogeny graph of 4 multiedges and 9 multiedges in the 3-isogeny graph. In the case of SIDH parameters for the two lowest security levels, Onuki et al. [11] have shown that there are only two cycles in the 3-isogeny graph in SIKEp434, and there are no cycles in the other graphs in SIKEp434 and SIKEp503. Thus in these real cases, the GPST attack may never fail or fail in negligible cases.

Acknowledgements: Both authors thank Christophe Petit and Christophe Ritzenthaler for interesting feedback. The authors also thank the reviewers for the useful suggestions. The second author is a member of the project ANR-17-CE40-0012 Flair and ANR-20-CE40-0003 Jinvariant.

Conflict of interest: Authors state no conflict of interest.

References

- [1] Jao D, De Feo L. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In: Proceedings of the 4th International Conference on Post-Quantum Cryptography, PQCrypto'11. Berlin, Heidelberg: Springer-Verlag; 2011. p. 19–34.
- [2] Galbraith S, Petit C, Shani B, Ti YB. On the security of supersingular isogeny cryptosystems. In: Advances in Cryptology - ASIACRYPT. Vol. 10031. Berlin, Heidelberg: Springer; 2016. p. 63–91.
- [3] Hofheinz D, Hövelmanns K, Kiltz E. A modular analysis of the Fujisaki-Okamoto transformation. In: Theory of Cryptography – 15th International Conference, TCC 2017, Lecture Notes in Computer Science. Vol. 10677. Springer; 2017. p. 341–71.
- [4] Azarderakhsh R, Jao D, Leonardi C. Post-quantum static-static key agreement using multiple protocol instances. In: Selected Areas in Cryptography - SAC 2017. Vol. 10719. Cham: Springer International Publishing; 2017. p. 45–63.
- [5] Silverman J. Advanced topics in the arithmetic of elliptic curves. New York: GTM Springer-Verlag; 1994. p. 151.
- [6] Silverman J. The arithmetic of elliptic curves. Second edition, New York: GTM Springer-Verlag; 1992. p. 106.
- [7] Husemöller D. Elliptic curves. Second edition. New York: GTM Springer-Verlag; 2004. p. 111.
- [8] Kohel DR, Lauter K, Petit C, Tignol JP. On the quaternion-isogeny path problem. LMS J Comput Math. 2014;17(suppl A):418–32.
- [9] Charles DX, Goren EZ, and Lauter KE. Cryptographic hash functions from expander graphs, J Cryptol. 2009;22(1):93–113.
- [10] Ghantous W. Loops, multi-edges and collisions in supersingular isogeny graphs. 2021. arXiv:2101.08761.
- [11] Onuki H, Aikawa Y, Takagi T. The existence of cycles in the supersingular isogeny graphs used in SIKE, Cryptology ePrint Archive, Report 2020/439. 2020. <https://eprint.iacr.org/2020/439>