Breaking SIDH in polynomial time

DAMIEN ROBERT

Abstract. We show that we can break SIDH in polynomial time, even with a random starting curve E_0.

1. Introduction

We extend the recent attacks by [CD22; MM22] and prove that there exists a proven polynomial time attack on SIDH, even with a random starting curve E_0. Both papers had the independent beautiful idea to use isogenies between abelian surfaces to break a large class of parameter on SIDH. Namely, on a random starting curve E_0, if the degree of the secret isogenies are $N_A > N_B$, their attack essentially apply whenever $a := N_A - N_B$ is smooth. This is highly unlikely, however they use the fact that it is possible to tweak the parameters N_A and N_B to augment the probability of success (or reduce the smoothness bound on a), see Remark 1.2. In the case where $\mathrm{End}(E_0)$ is known, [CD22] also have a (heuristic) polynomial time attack, essentially because one can use the endomorphism ring to compute an a-isogeny on E_0 even if it is not smooth.

A natural idea is to go in even higher dimension to extend the range of parameters on which an attack is possible, even on a random curve E_0. We show that by going to dimension 8, it is possible to break in polynomial time all parameters for SIDH.

In an upcoming version, we will also show how to break a large class of parameters N_A, N_B by going to dimension 4 rather than 8. Namely, this is possible whenever we can write $N_A = bN_B + a$ with $a, b > 0$ sum of two squares (along with some slight technical conditions). This is a much more likely condition than smoothness of $N_A - N_B$, hence (if possible tweaking N_A and N_B), we expect this attack to be highly likely and more efficient than the one in this paper in practice.

The idea of the present attack is that we can always write a, b as a sum of four squares, hence we always get an attack in dimension 8. A rough version of this article was originally published late at night with many errors and mistakes…I originally intended to carefully write a more thorough version of this article fully spelling out the dimension 4 attack too. However, given the interest on this subject, I thought it would be better to first correct the existing mistakes, and leave the generalisation for later.

Many thanks are due to the persons who commented on the prior version. Special thanks to Benjamin Wesolowski and Marco Streng, for suggesting to simply use $b = 1$ in the dimension 8 attack. This significantly simplify the description of the attack in this case. (Although as noted above the general $b > 0$ case will still be useful for the dimension 4 attack, see Remark 1.2).

Theorem 1.1. *We suppose that we are given the following input: we are given a secret N_B-isogeny over a finite field $\phi_B : E_0 \to E_B$ along with its images on (a basis of) the N_A-torsion points of E_0, where N_A and N_B are smooth coprime integers and $N_A > N_B$. We also assume*

Date: August 11, 2022.

that we are given the factorisations of N_A and N_B and (for simplicity) that we are given a basis of $E_B[N_B]$.

Let $\mathbb{F}_q$ be the smallest field such that ϕ_B, and the points of $E_0[N_A N_B]$ are defined[1]*. Then we can recover ϕ_B in time $O(\ell_A^8 \log \ell_A \log N_A + \log^2 N_A + \log^2 N_B)$ arithmetic operations in $\mathbb{F}_q$ where ℓ_A is the largest prime divisor of N_A.*

Note that in in the context of SIDH, if $N_B > N_A$ we will simply try to recover Alice's secret isogeny Φ_A instead.

Remark 1.2. We can tweak the parameters N_A and N_B as follow, as in the strategies of [CD22; MM22]: we can replace N_A by $N'_A = eN_A/d_A$ where e is a small integer (this will require to guess the image of Φ_B on the $N_A e$ torsion), and d_A any divisor of N_A, and N_B by $N'_B = fN_B/d_B$ where f is any smooth integer prime to N_A (this requires prolonging Φ_B by an f-isogeny) and d_B a small divisor of N_B (this requires guessing the first d_B-isogeny step of Φ_B). We can hope to find integers N'_A and N'_B, $a' > 0$, $b' > 0$ such that $N'_A > N'_B$ and $N'_A = b'N'_B + a'$ where both b' and a' are a sum of two squares. In this case, suppose for simplicity that we can also find a decomposition $b' = {b'_1}^2 + {b'_2}^2$ where $\gcd(b'_1, b'_2)$ is prime to N'_B (the general case will be tackled in an upcoming revision of this paper). Once these parameter tweaks are found, the complexity of Theorem 1.1 is reduced to $O({\ell'_A}^4 \log \ell'_A \log N'_A + \log^2 N'_A + \log^2 N'_B)$ arithmetic operations, because we can replace the endomorphism computation F from an N_A-endomorphism in dimension 8 to an N'_A-endomorphism in dimension 4.

2. Proof

Since $N_A > N_B$, write $N_A = N_B + a$ for a positive integer $a > 0$. Since N_A is prime to N_B, $\gcd(N_A, a) = 1$.

Let $M \in M_4(\mathbb{Z})$ be a 4×4 matrix such that $t_M M = a\,\mathrm{Id}$, Explicitly we write $a = a_1^2 + a_2^2 + a_3^2 + a_4^2$ and take M the matrix of the multiplication of $a_1 + a_2 i + a_3 j + a_4 k$ in the standard quaternion algebra $\mathbb{Z}[i, j, k]$. Let α_0 be the endomorphism on E_0^4 given matricially by M, The dual $\tilde{\alpha}_0$ of α_0 is given matricially by t_M (since integer multiplications are their own dual), so $\tilde{\alpha}_0 \alpha_0 = a\,\mathrm{Id}$, hence α_0 is an a-isogeny. We let α_B be the endomorphism of E_B^4 given by the same matrix M.

Remark 2.1. The decomposition of a as a sum of four squares is a precomputation step that only depends on N_A and N_B and can be done by solving a norm equation.

Let $F = \begin{pmatrix} \alpha_0 & \hat{\phi_B} \\ -\phi_B & \widetilde{\alpha_B} \end{pmatrix}$, where $\hat{\phi_B}$ is the dual isogeny $E_B \to E_0$ of ϕ_B. F is an endomorphism on the 8-dimensional abelian variety $A = E_0^4 \times E_B^4$. Since N_A is prime to N_B, we know how $\hat{\phi_B}$ acts on $E_B[N_A]$, hence we know how F acts on $A[N_A]$ (we actually won't need to compute $\hat{\phi_B}$ on $E_B[N_A]$). Furthermore, since α_0 is given by an integral matrix, it commutes with ϕ_B in the sense that we have the equation: $\phi_B \alpha_0 = \alpha_B \phi_B$.

Since the dual $\tilde{F}$ of F is given by $\tilde{F} = \begin{pmatrix} \widetilde{\alpha_0} & -\hat{\phi_B} \\ \phi_B & \alpha_B \end{pmatrix}$, we compute

$$\tilde{F} F = F \tilde{F} = \begin{pmatrix} N_B + a & 0 \\ 0 & N_B + a \end{pmatrix} = N_A\,\mathrm{Id}\,.$$

[1]We make no further assumptions on E_0 and E_B: we do not require them to be supersingular. In the context of SIDH, $\mathbb{F}_q$ will be the base field $\mathbb{F}_{p^2}$.

Hence F is an N_A-isogeny on A (with respect to the product polarisations), and we can compute its action on the N_A-torsion.

It is easy to compute its kernel: it is given by the image of $\tilde{F}$ on $A[N_1]$. In fact, since a is prime to N_A, the kernel of F is exactly the image of $\tilde{F}$ on $E_0^4[N_1] \times 0$, so we immediately get the four generators (g_1, g_2, g_3, g_4) of the kernel $\operatorname{Ker} F$. This step costs $O(\log a)$ arithmetic operations in $A(\mathbb{F}_q)$.

We can then compute F (on any point $P \in A(\mathbb{F}_q)$) using an isogeny algorithm in dimension 8, decomposing the N_A-endomorphism F as a chain of ℓ-isogeny for ℓ the prime factors of N_A. If ℓ_A is the largest prime divisor of N_A, the complexity of the first ℓ_A-isogeny computation will first be $\widetilde{O}(\log N_A)$ arithmetic operations in $A(\mathbb{F}_q)$ to compute the multiples $\frac{N_A}{\ell_A} g_i$, followed by the individual ℓ_A-isogeny computation on P and the g_i. This isogeny computation costs $O(\ell^8 \log \ell)$ operations over $\mathbb{F}_q$ using [LR22].

Remark 2.2. The isogeny computation in [LR22; BCR10] uses a (level $n = 4$) theta model of A, which we can compute as the (fourfold) product theta structure of the theta models of E_0 and E_B. It is also well known how to switch between the theta model and the Weierstrass model on an elliptic curve, and it is not hard to extend the conversion to the product of elliptic curves. The arithmetic on the theta models can be done in $O(1)$ arithmetic operations in a $O(1)$-extension of $\mathbb{F}_q$ (if $8 \mid N_A N_B$ the theta model will already be rational). However the big $O()$ notation hides an exponential complexity in the dimension g. In dimension 8 and level $n = 4$, the theta model uses 2^{16} coordinates, so we would need in practice to switch to the *Kummer* model by working in level $n = 2$ which "only" requires 2^8 coordinates. This is another reason why we would prefer to compute an endomorphism in dimension $g = 4$ rather than $g = 8$: in dimension 4 we would only need 2^8 coordinates in level $n = 4$, or 2^4 coordinates in level $n = 2$.

Since we compute a composition of at most $O(\log N_A)$ isogenies, the total cost of evaluating F on P is $O(\log^2 N_A + \log N_A \ell^8 \log \ell)$.

Thus we can evaluate F on any point of A, so we can evaluate ϕ_B or $\hat{\phi}_B$ on any point of E_0 (resp. E_B). We can now recover the kernel of ϕ_B on E_0 as the image of $\hat{\phi}_B$ on $E_B[N_B]$. If (Q_1, Q_2) is a basis of $E_B[N_B]$, we compute $Q_i' = \hat{\phi}_B(Q_i)$ by evaluating F on the point $(0, 0, 0, 0, Q_i, 0, 0, 0)$, and the kernel of ϕ_B is generated by whichever Q_i' has order N_B. This step costs $O(\log^2 N_B)$ operations in $E_0(\mathbb{F}_q)$, along with two calls to the evaluation of F. This concludes the complexity analysis.

References

[BCR10] G. Bisson, R. Cosset, and D. Robert. "AVIsogenies". Magma package devoted to the computation of isogenies between abelian varieties. 2010. URL: https://www.math.u-bordeaux.fr/~damienrobert/avisogenies/. Free software (LGPLv2+), registered to APP (reference IDDN.FR.001.440011.000.R.P.2010.-000.10000). Latest version 0.6, released on 2012-11-28.

[CD22] W. Castryck and T. Decru. *An efficient key recovery attack on SIDH (preliminary version)*. Cryptology ePrint Archive, Paper 2022/975. https://eprint.iacr.org/2022/975. 2022. URL: https://eprint.iacr.org/2022/975.

[LR22] D. Lubicz and D. Robert. "Fast change of level and applications to isogenies". Accepted for publication at ANTS XV Conference — Proceedings. Aug. 2022. URL: http://www.normalesup.org/~robert/pro/publications/articles/change_level.pdf.

[MM22] L. Maino and C. Martindale. *An attack on SIDH with arbitrary starting curve*. Cryptology ePrint Archive, Paper 2022/1026. https://eprint.iacr.org/2022/1026. 2022. URL: https://eprint.iacr.org/2022/1026.

INRIA Bordeaux–Sud-Ouest, 200 avenue de la Vieille Tour, 33405 Talence Cedex FRANCE
Email address: damien.robert@inria.fr
URL: http://www.normalesup.org/~robert/

Institut de Mathématiques de Bordeaux, 351 cours de la liberation, 33405 Talence cedex FRANCE