On the Optimality of Linear-Time Enumeration of $E(\mathbb{F}_p)$

— draft note —

Setting. Fix an odd prime $p > 3$ and a non-singular elliptic curve

$$\mathcal{E}/\mathbb{F}_p : \quad y^2 \equiv x^3 + Ax + B \pmod{p}, \qquad \Delta \neq 0.$$

We consider the task of enumerating all points $\mathcal{E}(\mathbb{F}_p)$. Write $\chi : \mathbb{F}_p \to \{0, \pm 1\}$ for the quadratic character with $\chi(0) = 0$ and $\chi(u) = \left(\frac{u}{p}\right)$ for $u \neq 0$.

Algorithms. The workhorse implementation we analyse is:

1. For each $x \in \mathbb{F}_p$, set $f(x) = x^3 + Ax + B \in \mathbb{F}_p$ and compute $\chi(f(x))$.
2. If $\chi(f(x)) = 0$ output $(x, 0)$. If $\chi(f(x)) = 1$ recover a square root $y \equiv \sqrt{f(x)} \pmod{p}$ and output (x, y) and $(x, -y)$. If $\chi(f(x)) = -1$ output nothing.

Square roots are found either (i) by Tonelli–Shanks in time $\tilde{O}(1)$ field operations[1] when needed, or (ii) by a precomputed table (one pass over $y \in \mathbb{F}_p$ storing the first y seen for each residue $r = y^2$).

Main statements

Theorem 1 (Output-size lower bound)**.** *For any elliptic curve $\mathcal{E}/\mathbb{F}_p$ one has*

$$\#\mathcal{E}(\mathbb{F}_p) \;=\; p + 1 - t, \qquad |t| \leq 2\sqrt{p},$$

so $\#\mathcal{E}(\mathbb{F}_p) = \Theta(p)$. In particular, any correct algorithm must produce *$\Theta(p)$ point records, hence takes $\Omega(p)$ time on any model where emitting a record costs $\Omega(1)$.*

Proof. This is Hasse's bound; see e.g. [1, Thm. V.1.1]. The $\Omega(p)$ lower bound is a trivial consequence of having to write $\Theta(p)$ outputs. □

Theorem 2 (Decision lower bound)**.** *Consider algorithms that, given $A, B \in \mathbb{F}_p$, may use field operations and evaluations of the quadratic character $\chi(\cdot)$ on values of their choice in $\mathbb{F}_p$. Any algorithm that always outputs* exactly *$\mathcal{E}(\mathbb{F}_p)$ must perform $\Omega(p)$ distinct evaluations of $\chi(f(x))$ (or equivalent work) in the worst case.*

Proof sketch (adversary/indistinguishability). Fix (A, B) with $\Delta \neq 0$. The set of abscissae for which $\mathcal{E}$ has rational points is

$$X^* \;=\; \{\, x \in \mathbb{F}_p : \; \chi(f(x)) \in \{0, 1\} \,\}.$$

[1] Formally $O(\log^2 p)$ bit operations on the RAM/word-RAM; here and below $\tilde{O}(\cdot)$ hides polylogarithms.

An algorithm that queries $\chi(f(x))$ on fewer than $p-c$ distinct x leaves at least c abscissae *unprobed*. For any unprobed x_0 with $f(x_0) \neq 0$ one can find a constant shift $\Delta \in \mathbb{F}_p^\times$ such that the modified curve $\mathcal{E}_\Delta : \; y^2 = x^3 + Ax + (B+\Delta)$ agrees with $\mathcal{E}$ on all probed abscissae (i.e. $\chi(f(x)+\Delta) = \chi(f(x))$ for those x) yet *flips* the quadratic character at x_0, i.e. $\chi(f(x_0)+\Delta) \neq \chi(f(x_0))$. (Heuristically, each constraint $\chi(f(x_i)+\Delta) = \chi(f(x_i))$ removes a factor ≈ 2 of the admissible Δ; with fewer than p constraints some Δ remain.) Thus an algorithm that probes fewer than p abscissae cannot distinguish inputs (A, B) from $(A, B+\Delta)$ that induce different membership of x_0 in X^*, yet must output different point sets to be correct—contradiction. Hence $\Omega(p)$ probes are necessary. □

The (standard) proof above can be made fully rigorous using multiplicative character orthogonality to count the number of Δ satisfying the probe constraints; see e.g. Weil bounds for character sums.

Corollary 1 (Optimality up to polylog factors)**.** *The x-scan algorithm (Legendre test per x, plus Tonelli–Shanks or a precomputed $\sqrt{\cdot}$ table) runs in time $T(p) = \Theta(p)$ field operations with the table, or $T(p) = \Theta(p) \cdot \tilde{O}(1)$ without it. By Theorems 1 and 2, no correct algorithm can asymptotically improve the* work *below $\Omega(p)$, hence the approach is optimal up to polylogarithmic factors and constant improvements and trivially parallelises across x.*

Why “line-exclusion” cannot asymptotically help

Several practical heuristics try to exclude lattice points $(x, y) \in \mathbb{F}_p^2$ en masse by reasoning about lines of integral slope (including vertical and horizontal). We record two simple facts.

Lemma 1 (Vertical and horizontal lines carry no mass advantage)**.** *For $\mathcal{E}/\mathbb{F}_p$ non-singular and $p > 3$, each abscissa $x \in \mathbb{F}_p$ supports either 0, 1 (when $f(x) \equiv 0$), or 2 points of $\mathcal{E}(\mathbb{F}_p)$ with that x-coordinate. In particular, no vertical line contains 3 distinct affine points of $\mathcal{E}(\mathbb{F}_p)$.*

Proof. Immediate from $y^2 \equiv f(x)$: for fixed x, y is determined up to sign, with the unique $y = 0$ case when $f(x) = 0$. □

Lemma 2 (Collinearity is a group law identity)**.** *Three affine points $P, Q, R \in \mathcal{E}(\mathbb{F}_p)$ are collinear if and only if $P + Q + R = \mathcal{O}$ in the group law. Consequently, on any fixed non-vertical line in $\mathbb{F}_p^2$, the number of intersections with $\mathcal{E}$ is at most 3 counted with multiplicity.*

Proof. Standard; see e.g. [1, Ch. III]. □

These lemmas show that any exclusion scheme driven by *lines* can only infer that, once you have identified the (up to) 2 abscissae where a line meets $\mathcal{E}$ besides a given point, no *other* lattice point on that line belongs to $\mathcal{E}$. But to obtain those seed points, one must already solve instances of $y^2 = f(x)$ for representative x on that line. There is no mechanism by which line sweeps can certify large *blocks* of abscissae as non-productive without, in effect, learning the quadratic character $\chi(f(x))$ for almost all x.

We can formalise this obstruction:

Proposition 1 (Line-based exclusion cannot beat $\Theta(p)$ work)**.** *Fix any algorithm that, in addition to field operations, may enumerate (and mark as excluded) all lattice points on a finite collection of $\mathbb{F}_p$-lines, except for the (at most three) points where the line meets $\mathcal{E}$. Then, in the worst case over $\mathcal{E}/\mathbb{F}_p$, the algorithm still must determine $\chi(f(x))$ for $\Omega(p)$ distinct abscissae x to output $\mathcal{E}(\mathbb{F}_p)$ correctly.*

Proof idea. By Lemma 1, vertical lines never cover more than two $\mathcal{E}$-points per abscissa; horizontal lines likewise do not certify absence of points at a given x without knowing $\chi(f(x))$. By Lemma 2, any other line contains at most 3 points of $\mathcal{E}$ (counted with multiplicity); using such lines is equivalent to repeatedly applying the group law. In particular, "exclusion by lines" can only rule out lattice points *conditional on* already having discovered the true intersections—which requires solving $y^2 = f(x)$ at those abscissae.[2] Thus the adversary argument of Theorem 2 applies verbatim: unless $\chi(f(x))$ is effectively learned for $\Omega(p)$ distinct x, one can produce two curves indistinguishable by the algorithm's probes but with different point sets on some unprobed abscissa; correctness then fails. $\square$

Parallelism and memory

The lower bounds above are on *total work*. They do not preclude strong wall-clock speedups by parallelising the independent per x tests. Your implementation assigns ranges of x to workers. With the precomputed square-root table the cost is:

$$\text{build time } \Theta(p), \quad \text{query time } \Theta(1) \text{ per } x, \quad \text{total } \Theta(p),$$

which is optimal by Cor. 1. Without the table, replacing table lookups by Tonelli–Shanks gives total time $\Theta(p) \cdot \tilde{O}(1)$, still optimal up to polylog factors and constant improvements. The memory–time tradeoff is clean: the table uses $\Theta(p)$ words and removes the (rare) expensive square-root steps; when RAM is constrained, the on-the-fly variant remains optimal up to polylog factors.

Takeaway. Enumerating $E(\mathbb{F}_p)$ in *linear work* by scanning $x \in \mathbb{F}_p$ and testing quadratic residuosity is optimal up to polylogarithmic factors. Line-based exclusion cannot asymptotically reduce the necessary information one must obtain (the quadratic character of $f(x)$ for almost all x); if implemented with an explicit grid, it is in fact $\Omega(p^2)$ in the worst case.

References

[1] J. H. Silverman and J. T. Tate, *Rational Points on Elliptic Curves*, 2nd ed., Springer, 2015. (Hasse bound; group law; basic facts.)

[2] D. Shanks, "Five Number-Theoretic Algorithms," *Proc. Second Manitoba Conf. Numer. Math.* (1971), pp. 51–70. (Tonelli–Shanks.)

[3] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed., Springer, 1990. (Quadratic characters; orthogonality.)

[2] If one maintains an explicit $p \times p$ grid to mark exclusions, each processed non-vertical line touches $\Theta(p)$ cells, so even a bounded number of lines already costs $\Theta(p)$ operations; if one avoids the grid, the exclusion carries no asymptotic informational gain beyond what the quadratic character per x provides.