A Local-Global Principle for Diophantine Equations

(Extended Abstract)

Richard J. Lipton and Nisheeth Vishnoi
{rjl,nkv}@cc.gatech.edu

Georgia Institute of Technology, Atlanta, GA 30332, USA.

Abstract. We observe the following *global-local* principle for Diophantine equations: If an equation $f(x_1, \ldots, x_t) \in \mathbb{Z}[x_1, \ldots, x_t] = p$ can be solved efficiently for a *dense* set of primes p, then one can efficiently obtain solutions to equations of the form

$$f(x_1, \ldots, x_t) \equiv r \bmod n.$$

This is done without any knowledge of the factorization of n.
We apply this principle to get the following results:

- There is an efficient algorithm to solve a *modular* version of Fermat's equation
$$x^2 + y^2 \equiv r \pmod{n}.$$
- Assuming factoring is hard, it is hard to solve the following equation for a *dense* set of primes p :
$$x^2 - y^2 = p - 1.$$

Randomness and Extended Riemann Hypothesis are essential in the proofs.

1 Introduction

Around the third century AD, a Greek mathematician Diophantus from Alexandria started a systematic study of various number theoretic problems. One important question he raised was:

When do certain equations have rational or integral solutions ?

In his honor, these broad class of algebraic equations over integers or number fields are called *Diophantine equations*. The subject of Diophantine equations is the study of their solubility, or when do these equations have solutions, say over integers or rationals. Also included are questions of the kind:

When is an integer a sum of two squares ?

For more on the work of Diophantus see [8, 4].

Various Diophantine equations have been studied by various mathematicians, and their 1800 year study has a significant impact on the development of mathematics. Among the famous Diophantine equation questions are:

- **Fermat-Legendre**: When is an integer a sum of *two/four* squares ?
- **Pell**: For what d does $x^2 - dy^2 = 1$ have solutions over integers ?
- **Fermat**: Does $x^n + y^n = z^n$ have integral solutions for $n > 2$?
- **Mordell**: When does $y^2 = x^3 + k$ have rational solutions ?

For more and details on these equations see [14].

With the advent of computers, this subject has gained more significance. Prominent applications of Diophantine equations are factoring [5] and certain digital public key and signature schemes in Cryptography [12, 10]. These and other applications rely on *efficient* (non)-solubility of certain Diophantine equations. Thus investigation of efficient solubility of Diophantine equations is extremely important. (By *efficient* we mean that there is an algorithm running in polynomial time in the representation of the input. The common representation is in binary. So an integer n has size $\lceil \log n \rceil$.) Not many Diophantine problems are known to be solvable efficiently in the classical model, [5, 13]. The Quantum model of computation has seen more success in recent years, [7]. The main difficulty in finding efficient solutions to Diophantine problems is the fact that most of the proofs known to show existence of solutions usually are *brute force* search. There is however a useful class of problems whose solutions are shown to exist by an *infinite descent* argument, first used by Fermat [9]. Roughly this method reduces the solution to the original problem to one with *smaller* inputs, which in most cases can be converted to an efficient algorithm.

In this abstract we relate the complexity of finding efficient representations of a *dense* set of primes to solubility of the related *modular* Diophantine equations. For example, it is well known [9, 14] that for every prime p which is either 2 or of the form $4k + 1$, there are integers x, y such that $x^2 + y^2 = p$. In fact among many proofs of this fact (attributed to Fermat), one is via the method of *infinite descent* and leads to an algorithm that runs in time $\text{poly}(\log p)$. We show that under widely believed number theoretic conjectures, this implies that we can efficiently find solutions to modular equations of the form $x^2 + y^2 \equiv r \pmod{n}$, for integer r, n, whenever a solution exists. (It is important to note here that we *do not* have the factorization of n.) We also relate the complexity of factoring to efficient representability of a certain Diophantine equation.

1.1 Preliminaries

Before we proceed, we need some notation.

- (a, n) denotes the gcd of a and n.
- $a|b$ means $a = cb$, for some integer c.
- $\mathbb{Z}_n := \{a : \ 1 \leq a < n \text{ and } (a, n) = 1\}$. This is a group.

- For m, n such that $(m, n) = 1$, m_n^{-1} denotes the inverse of $m \pmod n$ in $\mathbb{Z}_n$.
- $\phi(n) := |\mathbb{Z}_n|$.
- For positive integers a, d with $(a, d) = 1$, $\mathcal{AP}_{a,d} := \{a + kd | k \geq 0\}$. Also $\mathcal{AP}_{a,d}(n) := \{a + kd | k \geq 0 \text{ and } a + kd \leq n\}$.
- $\pi(x; q, r) := \sum_{1 \leq p \leq x,\ p \equiv r \pmod q} 1$, where p are primes.
- An algorithm is said to be *efficient*, if its running time is polynomial in the size it takes to represent its input in binary. It is allowed access to a random source.
- A function $\mu(k)$ is said to be *negligible* if for every polynomial $\mathrm{p}(k)$, there is a k_0 such that for all $k > k_0$, $\mu(k) < \frac{1}{\mathrm{p}(k)}$.
- All logarithms are base 2.

1.2 Our results

The *global-local* principle we observe is that if we can efficiently find representations of a sufficiently *dense* set of primes by a Diophantine equation $f(x_1, \ldots, x_t)$, then we can efficiently solve equations of the form $f(x_1, \ldots, x_t) \equiv r \pmod n$. (We make no attempt to define this formally here.)

In this abstract we just present two concrete instances to illustrate this principle. Many other applications can be deduced by similar reasoning. We mention some in Section 3.

Theorem 1. *Given positive integers $1 \leq r < n$ satisfying one of the following conditions:*

1. $(4, n) = 1$,
2. $(4, n) = 2$ *and* $r \equiv 1 \pmod 2$ *or*
3. $(4, n) = 4$ *and* $r \equiv 1 \pmod 4$,

Assuming ERH, there is an algorithm which runs in time polynomial in $\log n$, *and with probability at least* $2/3$, *outputs* s, t *such that*

$$s^2 + t^2 \equiv r \pmod n.$$

It is important to note that we are not given factorizations of n, else the problem is trivial.

Theorem 2. *Assuming there is no randomized polynomial time algorithm for* factoring *and the ERH, for any fixed arithmetic progression $\mathcal{AP}_{a,d}$, (with $(a, d) = 1$) at least one of the following holds:*

1. *For infinitely many m, the fraction of primes in $\mathcal{AP}_{a,d}(m)$ which are representable as $s^2 - t^2 + 1$ for integers s, t is a negligible function in* $\mathrm{poly}(\log m, \log d)$.
2. *For infinitely many m, the number of primes in $\mathcal{AP}_{a,d}(m)$ for which there is polynomial time algorithm to find representations as $s^2 - t^2 + 1$ for integers s, t is a negligible function in* $\mathrm{poly}(\log m, \log d)$.

Theorem 1 is a positive application which helps solve a generalization of Fermat's equation, while Theorem 2 is (most likely) a negative application of the principle.

Its worth noticing that the two Diophantine equations in the two equations are strikingly similar except for the $+/-$ sign. The discrepancy in the results for the respective equations reflects the general state of affairs in this area of efficiently finding solutions to Diophantine equations!

The basic idea in both these proofs is very simple:

1. Assume that one can efficiently solve the equation $f(x, y) = p$ for a set of primes in some $\mathcal{AP}_{a,d}$.
2. Given r, n pick q uniformly at random from the set $\mathcal{AP}_{a',d'}(\text{poly}(n))$. Here a', d' are chosen such that $\mathcal{AP}_{a',d'} \in \mathcal{AP}_{a,d} \cap \mathcal{AP}_{r,n}$ and $\text{poly}(\cdot)$ is a suitable polynomial.
3. It follows from the ERH that q is a prime with high probability. Hence use the algorithm from Step (1) to find s, t such that $f(s, t) = q$. But since $q = r + ln$ for some l, $f(s, t) \equiv r \pmod{n}$.

The consequences however are surprising and new to the best of our knowledge. In Section 2 we present some number theoretic facts and in Section 3 we sketch proofs of Theorems 1 and 2.

2 Some number theoretic facts

In this section we state some well known facts in number theory that we will need for the proofs.

Lemma 1. *[2] There is a constant $c > 0$ such that for all $n \geq 2$, $\phi(n) \geq \frac{cn}{\log n}$.*

Lemma 2. *[14] For integer a, b, n, the congruence $ax \equiv b \pmod{n}$ is soluble if and only if $(a, n)|b$.*

Lemma 3. *[14] Given integers a, b, m, n such that $(m, n) = 1$, there is a solution to the simultaneous equations $x \equiv a \pmod{m}$ and $x \equiv b \pmod{n}$. One such solution is*

$$u(a, b, m, n) := a \cdot n \cdot n_m^{-1} + b \cdot m \cdot m_n^{-1}.$$

Also $u(a, b, m, n)$ is computable in time polynomial in $\log m, \log n, \log a, \log b$.

Lemma 4. *[14, 9] For a prime p, the equation $x^2+y^2 = p$ is soluble over integers if and only if $p = 2$ or $p \equiv 1 \pmod{4}$. Moreover there is randomized polynomial time algorithm (in $\log p$) to find such a solution.*

Lemma 5. *[6] Under the ERH, there is a constant $\lambda > 0, n_0$ such that for all $n_0 \leq x$ and all q, r such that $r < q$, $(r, q) = 1$ and $q \leq \frac{\sqrt{x}}{\log^2 x}$,*

$$\pi(x; q, r) > \frac{\lambda x}{\phi(q) \log x}.$$

3 Main results

In this section we outline proofs of Theorems 1 and 2. As noted in Section 1.2, the main idea of the proofs is straightforward. The consequences seem to be new and surprising.

Proof (Theorem 1). That the solution exists for these three cases follows from Lemma 2. We just consider the case when $(4, n) = 1$. The remaining are similar and we omit them. Compute a poisitive integer $u := u(1, r, 4, n)$ be as guaranteed by Lemma 3. Notice we can assume $u \leq 10n$. Consider $\mathcal{AP}_{u,4n}$. Pick k uniformly at random from $\{1, \ldots, n^2\}$. The probability that $q := u + 4kn$ is prime is at least $\frac{1}{\text{poly}(\log n)}$. This follows from Lemmata 1 and 5. Assuming q is a prime and Lemma 4, find s, t such that $s^2 + t^2 = q$. But $q = u + 4kn \equiv r \pmod{n}$. Hence $s^2 + t^2 \equiv r \pmod{n}$. This gives an efficient algorithm which succeeds with probability $\frac{1}{\text{poly}(\log n)}$. The error probability can be reduced to $1/3$ by at most a polynomially many independent repetitions of this algorithm.

Remark 1. There are many other equations for which an analogue of Lemma 4 (and hence Theorem 1) holds. We list a couple here:

- $x^2 + 3y^2 = p$ and
- $x^2 - xy + y^2 = p$ are soluble if and only if $p = 3$ or $p \equiv 1 \pmod{3}$.

Remark 2. Shamir [12] proposed a cryptosystem based on the equation

$$x^2 + ky^2.$$

This was broken by Pollard *et al.* [11]. It should be clear that our methodology can be used it to attack it too!

For the proof of Theorem 2, we just give the connection to factoring integers. The slightly tedious and technical details are omitted.

The connection to factoring lies in the algorithm for factoring attributed to Fermat [5]. First notice that given an odd integer n which is not a prime power (checking whether n is a prime power can be done efficiently, see [1]), it is sufficient to give an algorithm to find one non-trivial factor of n. This is based on the following observation:

Suppose for integers s, t we we know that $s^2 - t^2 \equiv 0 \pmod{n}$ and $s \not\equiv \pm t \pmod{n}$, then $(s - t, n)$ is a non-trivial factor of n. Moreover if n is an odd integer which is not a prime power, then there always exist such s, t.

Remark 3. We mention here as an open problem to study efficient solubility over primes p of the equation

$$x^2 - y^2 = p - 1.$$

As we noted in Theorem 2, that an efficient algorithm to solve this for a reasonable number of primes will lead to an efficient factorization solution.

References

1. M. Agrawal, N. Kayal and N. Saxena, *PRIMES in P*, (August 2002).
2. Tom M. Apostol, **Introduction to Analytic Number Theory,** Springer-Verlag, New York, 1997.
3. R. C. Baker, G. Harman, The Brun-Titchmarsh Theorem on average. *Analytic Number Theory, Proceedings of a Conference In Honor of Heini Halberstam.* Editors: B. C. Berndt, H. G Diamond, A. J. Hildebrand, Birkhäuser, 1996.
4. I. G. Bashmakova, *Diophantus and Diophantine equations*, (translated from Russian). Dolciani Math. Expositions **20**, MAA.
5. R. Crandall, C. Pomerance, **Prime Numbers: A computational perspective**, Springer-Verlag, 2002.
6. H. Davenport, **Multiplicative Number Theory**. Springer-Verlag, 2000.
7. S. Hallgren, *Polynomial time quantum algorithms for Pell's equation and the principal ideal problem,* ACM STOC 2002.
8. T. L. Heath, **Diophantus of Alexandria, A study in the History of Greek Algebra**. Dover Publications, Inc., 1964.
9. Y. Hellegouarch, **Invitation to the mathematics of Fermat-Wiles**, Academic Press, 2002.
10. C. H. Lin, C. C. Chang, R. C. T. Lee, *A New Public-Key Cipher System Based Upon the Diophantine Equations*, IEEE Transactions on Computers, Volume 44 , Issue 1 (January 1995).
11. J. M. Pollard, C. P. Schnorr, *An efficient solution of the congruence* $x^2 + ky^2 = m \pmod{n}$, IEEE Trans. Inf. Theory, IT-33, No. 5, Sept. 1987, pp – 702-709.
12. H. Ong, C. Schnorr, A. Shamir, *An efficient signature scheme based on polynomial equations*, Proc. of CRYPTO 1985, pp 37–46.
13. N. Smart, **The Algorithmic Resolution of Diophantine Equations**. London Mathematical Society Student Text, 41. Cambridge University Press, 19.
14. H. E. Rose, **A course in number theory**, Oxford Science Publ., 1999.