Proceedings of the 3rd International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/13/20240801

Discrete logarithms and primitive roots: Algorithms, properties, and typical solution methods

Junchi Yang

University of Waterloo, Ontario, Waterloo, Canada

j647yang@uwaterloo.ca

Abstract. In mathematics, the logarithm, $\log_a b$, where a$\in (0,1) \cup (1,\infty)$ and b>0, is always defined as the real number x, such that a^x=b. Moreover, in the field of number theory, a similar concept called the discrete logarithm can be defined as follows: For a given positive integer m(m≥ 2), let a$\in N^+$ $with$ $(a,m) = 1$, and r is the primitive root of m, x=$ind_r a$ if $r^x \equiv a\ (mod\ m)$. Here, x is the discrete logarithm. The Discrete Logarithm Problem, which is a famous problem in number theory, is formulized as: For a positive integer b and a prime number p, and a is the primitive root of p, the goal is to find the exact value of i, such that $a^i \equiv b\ (mod\ p)$, in other words, it is targeted at finding the exact value of $ind_a b$. The goal of this research is to give several solutions to the Discrete Logarithm Problem, so firstly, some background concept like order and primitive root will be introduced with the proof of some foundational theories of these two concepts, then this essay will give two methods that can solve the Discrete Logarithm Problem called Shanks' Babystep-Giantstep Algorithm and Pohlig-Hellman Discrete Logarithm Algorithm.

Keywords: Discrete Logarithm, The Discrete Logarithm Problem, Order, Primitive Root.

1. Introduction

In cryptographic circles, the discrete logarithm remains a topic of intrigue. Although the discrete logarithm can be computed in specific scenarios, finding efficient solutions for general cases remains a formidable challenge. Notably, some algorithms tackle this problem and hold paramount significance in public-key cryptography, exemplified by systems like Elgamal [1]. This research endeavors to illuminate the intricacies of the Shanks' Babystep-Giantstep Algorithm and the Pohlig-Hellman Discrete Logarithm Algorithm. Both stand as robust solutions to the Discrete Logarithm Problem. To lay a foundation, it's imperative first to delve into fundamental concepts such as order and primitive root. By understanding these, one can better appreciate their applications to the focal problem. The crux of this study revolves around the operational mechanics of these two algorithms, exploring their methodologies in solving the Discrete Logarithm Problem, and discerning their connections to foundational tenets of elementary number theory.

DOI: 10.54254/2753-8818/13/20240801

2. Foundational Theories of Orders and Primitive Roots

2.1. Order

Definition 1: Let $m \in N^+, and\ a \in N^+\ with\ (a,\ m) = 1$, the order (or the multiplicative order) of a modulo m is the smallest positive integer r satisfying $a^r \equiv 1\ (mod\ m)$ [2].

The order of a modulo m is always written as $\delta_m(a)$ or $ord_m(a)$ [3]. Also, order always exists due to the Euler's Theorem: Let $m \in N^+, and\ a \in N^+\ with\ (a,m) = 1$, then $a^{\varphi(m)} \equiv 1\ (mod\ m)$ [4]. Euler's Theorem is too basic so the proof is skipped here. The Euler's Theorem says that for $m \in N^+, and\ a \in N^+\ with\ (a,m) = 1$, the set {r$\in N^+|\ a^r \equiv 1(mod\ m)$} is not empty so this set must have the smallest element, which is the (multiplicative) order, due to the Well-Ordering Principle.

Proposition 1: Let $m \in N^+, and\ a \in N^+\ with\ (a,m) = 1,\ k \in N^+$, then $a^k \equiv 1\ (mod\ m)$ if and only if $\delta_m(a) \mid k$ [5].

Proof: If $a^k \equiv 1(mod\ m)$, let, then $a^r \equiv 1\ (mod\ m)$.

By Division Algorithm, there exists q, t$\in N^+$ with 0$\leq t$<r-1 such that k=qr+t.

This means t=k-qr.

Also, notice that $a^{qr} \equiv 1\ (modm)$.

Thus, $a^t = a^{k-qr} \equiv a^{k-qr}a^{qr} = a^k \equiv 1(modm)$.

But r is the smallest positive integer satisfying $a^r \equiv 1\ (mod\ m)$ and t<r, so it means t=0

So k=qr, which means r|k.

Therefore, δ_m(a)|k.

On the other hand, if $\delta_m(a) \mid k$, so r |k, thus there exists $l \in N^+$, such that k= lr.

Since $a^r \equiv 1\ (mod\ m)$, thus, $a^k = a^{lr} = (a^r)^l \equiv 1\ (mod\ m)$.

Hence, $a^k \equiv 1\ (mod\ m)$ if and only if $\delta_m(a) \mid k$.

By Proposition1 and Euler's Theorem, a result can be got easily:

Corollary 1: Let $m \in N^+, and\ a \in N^+\ with\ (a,m) = 1$, then $\delta_m(a) \mid \varphi(m)$.

Proof: By Euler's Theorem, $a^{\varphi(m)} \equiv 1\ (mod\ m)$.

By Proposition1 and let k= $\varphi(m)$, $\delta_m(a) \mid \varphi(m)$.

So the Corollary1 holds.

Next, another important result about (multiplicative) order will be introduced.

Proposition 2: Let $m \in N^+, and\ a \in N^+\ with\ (a,m) = 1$, then a, $a^2, \dots, a^r$ are distinct modulo m, where r= δ_m(a) [6].

Proof: Suppose $\exists 1 \leq i < j \leq r, s.t.\ a^i \equiv a^j(mod\ m), then\ a^i(a^{j-i} - 1) \equiv 0(mod\ m)$.

This means m | $a^i(a^{j-i} - 1)$.

Since (a, m) =1, so (a^i, m) =1.

Thus, m | $a^{j-i} - 1$, which means $a^{j-i} \equiv 1(mod\ m)$.

Hence, r | j-i, so j-i $\geq r$.

But $1 \leq i < j \leq r$, which says j-i<r, it is a contradiction.

Hence, a, $a^2, \dots, a^r$ are distinct modulo m.

Proposition 2: Let $m \in N^+, and\ a \in N^+\ with\ (a,m) = 1$, let δ_m(a) = r, then $\delta_m(a^n) = \frac{r}{(r,n)}$ $(n \in N^+)$ [7].

Proof: Let $\delta_m(a^n) = l$, then $a^{nl} \equiv 1(mod\ m)$.

By Proposition1, r | ln, so $\exists\ q \in N^+, s.t.\ nl = rq$

Thus, $\frac{ln}{(r,n)} = \frac{r}{(r,n)}q$.

So $\frac{r}{(r,n)} \mid \frac{ln}{(r,n)}$.

Notice that $(\frac{r}{(r,n)}, \frac{n}{(r,n)}) = 1$.

Hence, $\frac{r}{(r,n)} \mid \frac{l}{(r,n)}$., so $\frac{r}{(r,n)} \mid l$.

On the other hand, notice that $(a^n)^{\frac{r}{(r,n)}} = (a^r)^{\frac{n}{(r,n)}} \equiv 1(mod\ m)$.

DOI: 10.54254/2753-8818/13/20240801

By Proposition1, $l|\ \frac{r}{(r,n)}$.

Therefore, $l = \frac{r}{(r,n)}$.

Hence, $\delta_m(a^n)$=$\frac{r}{(r,n)}$.

2.2. Primitive root

Definition 2: Let m$\in N^+(m \geq 2)$. The primitive root mod m is an integer g, such that δ_m(g) =$\varphi(m)\ and\ (g,m) = 1$ [8].

By Definition 2 and Proposition2, it is easy to get the following corollary.

Corollary 2: Let m$\in N^+(m \geq 2)$, a be an primitive root of m, then the list a,$a^2, \ldots, a^{\varphi(m)}$ picks up every element of Z_m^*.

What is worth saying is that the corollary2 is also another version of the definition of primitive root [7].

The most important result about primitive root is the Primitive Root Theorem.

Theorem 1(Primitive Root Theorem): Let p be a prime number, then Z_p^* has a primitive root.

Proof: [Lemma]: Let p be a prime number and a, b$\in Z_p^*$, denote $\delta_p(a) = k, and\ \delta_p(b) = l. If\ (k,l) = 1\ , then\ \delta_p(ab) = kl.$

[Proof of Lemma]: Let $\delta_p(ab) = r$.

Since $(ab)^{kl} = a^{kl}b^{kl} = (a^k)^l(b^l)^k \equiv 1\ (mod\ p)$.

By Proposition 1, $r|kl$.

The following is to prove $k|r\ and\ l|r$.

Because $(a^r)^k = a^{rk} = (a^k)^r \equiv 1(modp), and\ (b^r)^l = b^{rl} = (b^l)^r \equiv 1\ (modp)$.

Also, notice that $(a^r)^l \equiv (a^r)^l\ (b^r)^l = (ab^r)^l \equiv 1\ (mod\ p)$.

Thus, $k|rl$, combined with $(k,l) = 1, so\ k|r$.

Similarly, $l|r$.

Since $(k,l) = 1, this\ leads\ to\ kl \mid r$.

Hence, $kl = r$.

Therefore, $\delta_p(ab) = kl$.

[Back to Primitive Root Theorem]: For any a in Z_p^*, then (a, p) =1.

By Fermart's Little Theorem, $a^{p-1} \equiv 1\ (mod\ p)$.

By Proposition 1, $\delta_p(a)|p-1$.

If $\delta_p(a) = p-1 = \varphi(p)$, then a is the primitive root of p, the Primitive Root Theorem holds.

If $\delta_p(a) < p-1$, let $\delta_p(a) = k$, the main idea of the following part is to find some b in Z_p^*, such that the order of b modulo p is greater than the order of a.

By Proposition 2 (or Corollary 2), the list a, $a^2, \ldots, a^k$ can pick up all the roots of the polynomial f(x)=$x^k - 1\ in\ Z_p^*$, since k < p-1, so there exists c in Z_p^* and c is not in the list above.

Let $\delta_p(c) = l, if\ l \mid k\ , then\ c^k \equiv 1(mod\ p)$, thus $l\ does\ not\ divide\ k\ since\ c\ is\ not\ the\ root\ of\ f$

Consider the prime factorization of $k\ and\ l,$ there must be an unique prime number q who appears more often in l than it appears in k, in other words, $v_q(l) > v_q(k)$, here $v_q(x)$ represents the power of q in the prime factorization of the positive integer x.

Let k=$q^d k_1$ and $l = q^e l_1, where\ 0 \leq d < e$ and both $k_1\ and\ l_1$ does not contain the prime factor q.

Pick b = $a^{q^d} c^{l_1}$, By Proposition 2, it tells that $\delta_p\left(a^{q^d}\right) = \frac{k}{(k,q^d)} = \frac{k}{q^d}$=$k_1$ and $\delta_p\left(c^{l_1}\right) = \frac{l}{(l,l_1)}) = \frac{l}{l_1} = q^e$.

Now, notice that $(k_1, q^e) = 1$.

DOI: 10.54254/2753-8818/13/20240801

By Lemma, $\delta_p(b) = \delta_p\left(a^{q^d}c^{l_1}\right) = \delta_p\left(a^{q^d}\right)\delta_p\left(c^{l_1}\right) = k_1q^e > q^dk_1 = k = \delta_p(a)$.

Thus, an element b in Z_p^* with greater order than a is found.

Following this way, new elements in Z_p^* with strictly increasing order can be found until find an element with the order p-1 and that element is just the primitive root.

In general, Z_p^* has a primitive root.

The Primitive Root Theorem tells that every prime number has its own primitive root but there are still many problems about primitive root cannot be solved by this theorem although it has already been an amazing result. Also, the Primitive Root Theorem can describe why the assumption of Discrete Logarithm Problem always holds and this point will be discussed in the following session of this essay. The following is to introduce several results of primitive roots without proof since it does not have a close relation to the main topic of this research.

Theorem 2: Let m$\in N^+(m \geq 2)$. If Z_m^* has primitive roots, then the number of primitive roots in Z_m^* is $\varphi\big(\varphi(m)\big)$ [8].

In particular, if m = p is a prime number, then $\varphi\big(\varphi(m)\big) = \varphi(p-1)$, so it can tell that for any prime number p, the total number of primitive roots of p is $\varphi(p-1)$.

Theorem 3: Let m $\in N^+(m \geq 2)$. Then Z_m^* has primitive roots if and only if m $\in \{2,4,p^k,2p^k \mid p \text{ is an odd prime and } k \in N^+\}$ [9].

This result tells the structure of m that has primitive roots of m.

3. Definition and Properties of Discrete Logarithms

3.1. Discrete Logarithms and its properties

Definition 3: For a given positive integer m (m≥ 2), let a$\in N^+$ $with\ (a,m) = 1$, and r is the primitive root of m, x=$ind_r a$ if $r^x \equiv a\ (mod\ m)$.

The discrete logarithms have the following 5 properties:

Proposition 3: Let p be a prime number, and a is the primitive root of p, then: $x \equiv y\ (mod\ p)\ \ if\ and\ only\ if\ ind_a x \equiv ind_a y\ \ (mod\ p-1)$;

$$ind_a a^k \equiv k\ (mod\ p-1);$$

$$ind_a a = 1;$$

$ind_a xy \equiv ind_a x + ind_a y\ (\ mod\ p-1)$;

$$ind_a x^k \equiv k ind_a x\ (mod\ p-1).$$

To prove these properties, an easy lemma should be used:

[Lemma]: Let p be an prime number and a is the primitive root of p. Let b, c be positive integers, then $b \equiv c\ (mod\ p-1)\ if\ and\ only\ if\ a^b \equiv a^c\ (mod\ p)$.

[Proof of Lemma]: Since p is a prime number and a is primitive root of p, so $\delta_p(a) = p-1$

Also, By Fermat's Little Theorem, $a^{p-1} \equiv 1\ (mod\ p)$ (Also this holds since its order is p-1).

If $b \equiv c\ (mod\ p-1)$, then there exists positive integers k, such that b-c =k(p-1)

Thus, $a^{b-c} \equiv a^{k(p-1)} \equiv 1(mod\ p)$, so $a^b \equiv a^c(mod\ p)$.

On the other hand, if $a^b \equiv a^c\ (mod\ p)$, since (a, p) =1, so (p, a^c) $= 1$.

Therefore, $a^{b-c} \equiv 1(mod\ p)$.

Since $\delta_p(a) = p-1$, By Proposition 1, p-1 | b-c.

Hence, $b \equiv c\ (mod\ p-1)$.

In general, $b \equiv c\ (mod\ p-1)\ if\ and\ only\ if\ a^b \equiv a^c\ (mod\ p)$.

This Lemma is proved.

By using this lemma, it is not difficult to prove the above five properties and here only the property iv) will be proved. The rest of them can just be showed by the similar way of using lemma and the direct use of definition of the discrete logarithm.

Proof of iv): Let $l = ind_a xy,\ l_1 = ind_a x,\ l_2 = ind_a y$.

By definition, $a^l \equiv xy\ (mod\ p), a^{l_1} \equiv x\ (mod\ p), a^{l_2} \equiv\ y\ (mod\ p)$

DOI: 10.54254/2753-8818/13/20240801

Thus, $a^l \equiv xy \equiv a^{l_1}a^{l_2} = a^{l_1+l_2} \ (mod\ p)$.
By the Lemma above, $l_1 + l_2 \equiv l \ (mod\ p-1)$.
Hence, $ind_a xy \equiv ind_a x + ind_a y \ (\ mod\ p-1)$.

4. Solution Methods for the Discrete Logarithm Problem

4.1. The Discrete Logarithm Problem

The main idea of this session is to introduce the Discrete Logarithm Problem and its solutions including. The two main algorithms mentioned before. Firstly, the target is seeing what is Discrete Logarithm Problem.

Definition 4: The Discrete Logarithm Problem can be formulized as follows:

Given a positive integer b, and a large prime number p, let a be a primitive root of p, there exists an.

unique index i ($0 \ \leq i \leq p-1$), such that $b \equiv a^i \ (mod\ p)$. The problem is targeted at finding the exact value of this i.

The assumption of this problem holds because such a must exist by Theorem 1 (Primitive Root Theorem). Also the index i satisfying such property must be unique since by the congruence $b \equiv a^i \ (mod\ p)$, it is clear that (b, p) = (a, p) =1, so the remainder of b divides p is not 0 so the remainder, called r, is in Z_p^*.

By Corollary 2, the list a, $a^2, \dots, \ a^{p-1}$ picks up every element of $Z_p^* = \{1,2, \dots, p-1\}$. This means that there must exists an unique i, such that $r \ \equiv \ b \equiv \ a^i \ (mod\ p)$. Also, from the argument above, it is clear that there exists a bijection between two sets Z_p^* and the set {a, $a^2, \dots, \ a^{p-1}$} under modulo p.

Moreover, it is necessary that a should be the primitive root of p, otherwise, the set {a, $a^2, \dots, \ a^{p-1}$} under modulo p has the competitive element so that it is impossible to encrypt it.

Another problem is that why this problem should need a big prime number. This is because if just take a small prime number, it is very easy to encrypt it so the large prime ensures the difficulty of this problem. Then, this essay will discuss about two algorithms to solve this problem: Shanks' Babystep-Giantstep Algorithm and Pohlig-Hellman Discrete Logarithm Algorithm.

4.2. Shanks' Babystep-Giantstep Algorithm

Algorithm 1(Shanks' Babystep-Giantstep Algorithm): Consider the given congruence $b \equiv \ a^x \ (mod\ p)$, where p is a large prime number. Let N = p-1. The process of the algorithm is as follows:

i) Calculate $n = [\sqrt{N}]+1$;

ii) Construct the two sets A = {1, a, $a^2, \dots, a^n\}$ $and\ B = \{b, ba^{-n}, ba^{-2n}, \dots, ba^{-n^2}\}$;

iii) A and B actually have the same element, so there exists i, j $\in \{0,1,2, \dots, n\}, such\ that\ a^i \equiv ba^{-jn} \ (mod\ p)$;

iv) Let x = i + jn, then, x is the solution to the congruence $b \equiv \ a^x \ (mod\ p)$.

v) It is very clear that this algorithm works and its run time is O ($\sqrt{N}$), which greatly decreases the run time compared with calculating each value [10].

4.3. Pohlig-Hellman Discrete Logarithm Algorithm

To introduce the Pohlig-Hellman Discrete Logarithm Algorithm, firstly, the Chinese Remainder Theorem should be reviewed.

Theorem 4 (Chinese Remainder Theorem): Let $m_1, m_2, \dots, m_k \in N^+ \ (k \ \geq 2, k \ \in N^+)$ and they are pairwise coprime. (That is, $(\ m_i, m_j) = 1 \ for\ all\ 1 < i \leq j \leq k)$. If $a_1, a_2, \dots, a_k \in Z$, and consider the system of congruences:

$$\begin{aligned} x &\equiv \ a_1 \ (mod\ m_1) \\ x &\equiv \ a_2 \ (mod\ m_2) \\ &\dots \\ x &\equiv \ a_k \ (mod\ m_k) \end{aligned} \tag{1}$$

DOI: 10.54254/2753-8818/13/20240801

This system of congruences has an unique solution modulo $m_1\ m_2\ ...\ m_k$. In other words, if $x = x_0$ is a particular solution of this system, then all the solutions are given by all the integers $x\ satisfying\ x \equiv x_0\ (mod\ m_1\ m_2\ ...\ m_k)$ [7].

This theorem is a very basic result in number theory so the proof does not present in this essay but the proof can tell the algorithm to solve the system of congruences by using Chinese Remainder Theorem.

The solutions can be given by the formula:

$$x \equiv a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_n M_n y_n\ (mod\ m_1\ m_2\ ...\ m_k) \tag{2}$$

Where, $M_i = \frac{M}{m_i},\ M = m_1\ m_2\ ...\ m_k, y_i = (M_i)^{-1}(mod\ m_i), i = 1,2, ..., k.$

Now, it is the time to present the Pohlig-Hellman Discrete Logarithm Algorithm [11]:

Algorithm 2(Pohlig-Hellman Discrete Logarithm Algorithm):

Consider the prime factorization of p-1 = ${p_1}^{k_1}{p_2}^{k_2} ... {p_m}^{k_m}$;

For each prime factor $p_i\ (1 \le i \le m)$, let x = $a_0 + a_1 p_i + \cdots + a_{k_{i-1}}{p_i}^{k_{i-1}}\ \left(mod\ {p_i}^{k_i}\right)$;

Let r =1, compute $(a^x)^{\frac{p-1}{{p_i}^r}} \equiv b^{\frac{p-1}{{p_i}^r}}\ (mod\ p)$; Substitute x, and expand it, notice that from the second term, all the values are since due to the Fermat's Little Theorem, so it leads to $a^{a_0\frac{p-1}{p_i}} \equiv b^{\frac{p-1}{{p_i}^r}}\ (mod\ p)$; By the former steps, a_0 can be computed in the run-time of O (p_i), then let $r_1 = r + 1$, and go back to the third step; Continue the operation above until all the $a_i (1 \le i \le m)$ are computed; For each i, a congruence can be got in the form of second step, then use the Chinese Remainder Theorem to solve x. The above two algorithms are the two main effective algorithms to solve the Discrete Logarithm Problem.

5. Conclusion

This research targeted at solving the Discrete Logarithm Problem so to introduce the algorithm to solve this famous problem, first of all, several important concepts in the field of Elementary Number Theory are introduced, including the multiplicative order and the primitive root. In addition, several important theorems are given the rigorous proof like the Primitive Root Theorem, and then this essay turn to focus on the discrete logarithm, which is the base of the Discrete Logarithm Problem, and the most important properties of discrete logarithm are introduced. Finally, this research starts to give the solutions to the Discrete Logarithm Problem but before this, it discusses about why such this problem is designed in such way and how the previous concept and theories in number theory play an important role in this problem. Then, the two main algorithms are demonstrated including the Shanks' Babystep-Giantstep Algorithm and Pohlig-Hellman Discrete Logarithm Algorithm. This research gives the effective solutions to the Discrete Logarithm Problem and they can work much more efficiently than compute each value of power, which greatly reduce the run-time of solving this problem.

References

[1] Menezes, A.J., van Oorschot, P.C., Vanstone, S.A. Handbook of Applied Cryptography. CRC Press.
[2] Burton, D.M. (1989). The Order of an Integer Modulo n. Elementary Number Theory, 4th ed.
[3] Von zur Gathen, J., Jurgen, G. (2013). Modern Computer Algebra. Cambridge University Press.
[4] Gauss, C.F., Clarke, A.A. (translated into English) (1986). Disquisitiones Arithemeticae (Second, corrected edition), New York: Springer.
[5] Davidson, K.R. (2012). Integers, Polynomials and Finite Fields. University of Waterloo.
[6] Davidson, K.R. (1994). Integer and Polynomial Algebra. University of Waterloo.
[7] Zorzitto, F. (2016). A Taste of Number Theory.
[8] Stromquist, W. (2017). What are Primitive Roots? Mathematics. Bryn Mawr College.
[9] Gauss & Clarke. (1986). Arts 92.

DOI: 10.54254/2753-8818/13/20240801

[10] Shanks, D. (1971). Class number, a theory of factorization and genera. In Proc. Symp. Pure Math., Providence, R.I.: American Mathematical Society, vol. 20, pp. 415–440.
[11] Mollin, R. (2006). An Introduction to Cryptography. Chapman and Hall/CRC.p.344.