Optimal Method of Integer Factorization

RUSLAN SKURATOVSKII
Igor Sikorsky Kiev Polytechnic Institute, av. Pobedy 37, Kiev, and National Aviational University Kiev,
UKRAINE
[0000-0002-5692-6123]

Abstract— The object of the research is performance of integer factorization algorithms and possibility of mathematical methods using in these algorithms. The subject of the research is cryptographic properties GNFS method. Methods of research are methods of the theory of elliptic curves, finite fields, abstract algebra and advanced the theory of factorization algorithms. As a result of this work, the dependence of the properties of a minimal time of factorization and choice of algebraic factor bases over the ring $\mathbb{Z}$n , where $n = pq$ was established. Moreover, we have implemented the general number field sieve (GNFS), which is the most efficient classical algorithm known for factoring integers.

Keywords — Method of integer factorization, GNFS algorithm.

Received: March 10, 2021. Revised: January 15, 2022. Accepted: February 19, 2022. Published: March 24, 2022.

1. Introduction

The first public key encryption algorithm (Public Key Encryption, hereinafter PKE) was proposed by Whitfield Diffie and Martin Hellman at Stanford University. They, as well as independently Ralph Merkel, developed its basic concepts in 1976. The advantage of PKE is that there is no need for secret key transfer. PKE is based on the unsolvability of the problem of decomposing a natural number into prime factors.

One of the first attacks on the RSA system was an attempt to factorize n. If the thief can do it, he will easily calculate $\varphi(n)$ and easily find the secret key d by the formula $d = e^{-1} mod(n)$. But the problem of factorizing large natural numbers is still unresolved. On the other hand, if it is possible to factorize $n-1$, and these factors are not greater than some number m, then the factorization of the number n can be carried out for a time not greater than m^3.

One of the most advanced and promising methods of number factorization is the GNFS algorithm, which will be described at the end of the monograph as the most complex and effective method.

2. Theoretical Foundations of the Gnfs Algorithm

Let us consider the essence of the algorithm. Firstly, we have to find numbers that are complete squares in the ring $\mathbb{Z}$ and its extension $\mathbb{Z}[\alpha]$ in this method of factoring numbers. Then we will find their homomorphic l^2, m^2 images in $\mathbb{Z}_n$. Accordingly to the property of homomorphism, they are also complete squares in $\mathbb{Z}_n$. Then the difference of these squares modulo n is formed . The expansion of these numbers by the difference of the squares and gives the expansion of the number by factors. To find square numbers in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$ will use expansions of numbers into elements from factor bases. Elements that decompose into the product of elements from factor bases are called smooth with respect to these factor bases.

To find items that are both relatively smooth factor database screening algorithm in which selected only those couples (a,b) that $\alpha + mb$ and $\alpha + \theta b$ are relatively smooth relevant factor bases. Then will apply homomorphic mapping $\phi(a+b\theta)$ and reduction modulo n. After that, congruent pairs of complete squares are obtained in $\mathbb{Z}_n$.

$$x^2 = \phi(\beta)\phi(\beta) = \phi^2(\beta) = \phi\left(\prod_{(a,b)\in U}(a+b\theta)\right) =$$

$$= \prod_{(a,b)\in U}(\phi(a+b\theta)) = \prod_{(a,b)\in U}(\alpha+bm) = y^2,$$

To find such smooth elements that are complete squares in the corresponding extensions, a system of linear equations is solved, where the coefficients are the degree of occurrence of a prime number from the factor base in the schedule of the selected number by the factor base.

We present ordered sets of degrees of decomposition of numbers into elements from the factor base in the form of vectors $(\nu_p)_{p\in P_0}$, where ν_i it is a vector of degrees of occurrence *and the* - th number from the factor base in the number selected for verification on a full square. $|P_0|$. is the number of coordinates in the vector $\sum_j z_j v_j = \bar{0}(mod 2)$, where $V = \{\nu_j\}$. It is clear that for the solvability of the system it is enough if $|V| > |P|$.

This can be achieved by increasing the size of factor bases. This is a large sparse system of equations over the field $\mathbb{Z}_2$. The solution of which is a subset $W \subseteq V$, for which:

$$\sum_{v\in W} v = \vec{0},$$

These numbers will be complete squares in the above rings. Whence we obtain

$$\boldsymbol{x}^2 \equiv \boldsymbol{y}^2(\text{mod}n).$$

This lead us to factorization

$$(\boldsymbol{x} - \boldsymbol{y})(\boldsymbol{x} + \boldsymbol{y})(\text{mod}n).$$

Only in case $x \neq y$ these multipliers equal to p and q.

3. Algebraic Structures and Statements for the Algorithm.

Definitions 3.1. A ring A is called a Noetherian if it satisfies the following three equivalent conditions:

- An arbitrary non-empty set in A is stabilized.
- An arbitrary growing chain of ideals A is stabilized.
- An arbitrary ideal A is completely generated.

Proposition 3.1. Let A be a Noetherian ring, is its homomorphic image, with some homomorphism φ. Then A is a non-shaded netting ring.

The well-known concept of a prime ideal is considered a generalization of the concept of a prime number. But the concept of phantom ideal is a generalization of the power of a prime number.

Definition 3.2 . An ideal q in the ring A is said to be primary ideal if $xy \in q$ it follows that either $x \in q$ or $y^n \in q$ for some $n \in \mathbb{N}$.

Definition 3.3 . Discrete field normalization K is called as the image of a group of v of group K^*, where K^* is a multiplicative group of a field K and v has the following properties:

1) $v(xy) = v(x) + v(y)$ that is that v is a homomorphism of groups.

2) $v(x+y) \geq \min(v(x), v(y))$.

Definition 3.4 . A discrete normalized ring is a set for which a is a field. $x \in K^*$ for which $v(x) \geq 0$ and K is a field.

Definition 3.5 . A Dedekind ring is a Noetherian one-dimensional region for which the following conditions are equivalent:

- A is closed.
- An arbitrary phantom ideal is a degree of a prime ideal.
- An arbitrary nonzero ring $A_{\wp}$, $\wp \neq 0$ is a discretely normalized ring.

Consider nonzero ideals in a Dedekind ring, for example in $\mathbb{Z}_m = \mathbb{Z}\left[\sqrt[5]{2}\right]$.

Proposition 3.1 . For every nonzero ideal in $\mathbb{Z}_K$ there exists a prime p and integer k such that $\mathbb{Z}_K / \mathfrak{I} = GF\left(p^k\right)$.

Definition 3.6. The norm of the ideal $\mathfrak{I}$ is determined by equality $Norm\,\mathfrak{I} = \left|\mathbb{Z}_K / \mathfrak{I}\right|$.

Definition 3.7. A prime ideal $\mathfrak{I}$ is called a prime ideal of the first degree if

$G_{k/\mathfrak{I} = GF\left(p^k\right)}$ where p is a prime number.

Definition 3.8. A rational factor base is a finite set of prime numbers. That it is

$$\mathfrak{R} = \{p : p \in \mathrm{P},\ p \leq M\},$$

where P is the set of primes.

Definition 3.9 . An algebraic factor base is a finite subset of an algebraic extension such that for $A = \{a + b\theta\} \in Z[\theta]$ satisfies the condition for $a, b \in G, \forall a + b\theta$ and $\forall (a,b) \in A, \exists c, d \in Z|\theta|$ and $c, d \notin A$ is such that $c, d : cd = a + b\theta$.

For this reason, it is customary to call the element $a + b\theta$ *generating a* prime *ideal.*

As is well known, in a quadratic field (and not only in a quadratic field) a prime ideal of a Dedekind ring of at most 2 is generated.

Example 3.1. An example of a prime ideal that cannot be generated by a single element. In a Dedekind ring, which is not a ring of principal ideals, namely, $\mathbb{Z}\left[\frac{1+\sqrt{7}}{2}\right]$ to consider an ideal, $p\left\langle 2, \frac{1+\sqrt{7}}{2}\right\rangle$ it, of course, cannot be generated by any one of these elements by means of which the extension was formed.

Definition 3.10 . The norm of the number of $z = a + b\sqrt{-r}$, $r \in G$ is called an integer number $a^2 + rb^2$ and is denoted as $N(z) = |z|^2 = a^2 + rb^2$ (for example, $a + b\sqrt{-7}$ (or equivalent to it $a + bi\sqrt{-7}$)) has a norm $a^2 + 7b^2$).

Remarks 3.1 . *Note that the norm of an element z is in fact equal to the product of all elements conjugate to it, taking into account itself.*

Example. 3.2 $N\left[\frac{1+\sqrt{5}}{2}\right] = \frac{1-5}{4} = -1$, $N[2] = 2 \cdot 2 = 4$, for a quadratic extension number of the main field has two conjugated elements. In general, or $a \in \mathbb{Q}$, $N_{K/\mathbb{Q}}(a) = a^n$, $n = (K : \mathbb{Q})$.

Examples of Euclidean rings with norms:

- Ring of integers $\mathbb{Z}$ with Euclidean norm $N(\alpha) = |\alpha|$, $\alpha \in \mathbb{Z}, a \neq 0$ [23].
- Ring of polynomials $N\left(f(x)\right) = deg f(x), f(x) \neq 0$.
- Gaussian integer ring $\mathbb{Z}[i]$ with Euclidean norm $N(a+bi) = a^2 + b^2, a + bi \in G[i], a + bi \neq 0$ ..

Each Euclidean ring is factorial, and therefore for arbitrary nonzero elements there is their greatest common divisor.

A partial case of the norm of numbers is the norm in the expansion of Galois fields. That is, in what is a normal and separable extension. Consider this rule on the example of the expansion of Galois $\lfloor \sqrt[5]{2} \rfloor$.The elements of this field look like $\alpha = \sum_{i=0}^{4} q_i \left(\sqrt[5]{2}\right)^i, q_i \in Q_i$. Let $\sigma_1,\ \sigma_2,\ \sigma_3, \sigma_4, \sigma_5$. are all isomorphisms from $\mathbb{Q}\left\lfloor \sqrt[5]{2} \right\rfloor$ in $\mathbb{C}$.

Lemma 3.1. Let $a, b \in \mathbb{Q}$. Then the norm of the element formed using the generative $2^{\frac{l}{5}}$ the basis of the extension has the following form

$$Norm\left(a - \left(\sqrt[5]{2}\right)^l b\right) = b^5\left(\left(\frac{a}{b}\right)^5 - 2^l\right)$$

To prove, consider that the elements $a, b \in \mathbb{Q}$ from the main field under the action of automorphisms remain motionless. And the fact that the roots of a polynomial pass into those roots that are associated with them, that is, differ by a factor that is an element of the group of roots from the unit of the fifth order. It is known that the product of all these roots is 1.

$$Norm\left(a - \left(\sqrt[5]{2}\right)^l b\right) = b^5 Norm\left(\frac{a}{b} - \left(\sqrt[5]{2}\right)^l\right) =$$

$$= b^5 \prod_{j=1}^{5}\left(\frac{a}{b} - \left(\sqrt[5]{2}\right)^l\right) = b^5\left(\left(\frac{a}{b}\right)^5 - 2^l\right).$$

Therefore, opening the brackets in the last equation and reducing such received the desired rate.

Assertion.3.3. The norm of an ideal generated by a number will be equal to the norm of that number.

Definition 3.11. In the general case, the norm of a number (element) is a determinant of a linear operator that acts on the elements of the extension base in the same way as this number when multiplying by these base elements.

In the expansion of the 3-rd degree, for example in $\breve{y}\left\lfloor d^{\frac{1}{3}} \right\rfloor$ where the elements have the form: $a = a + bd^{\frac{1}{3}} + cd^{2\!/\!3}$ the norm of the element is more convenient to calculate by the second definition, because there will be 3 conjugates to each element of the base, so the element α itself will have 27 conjugates. And by the second definition $N(\alpha)$ is the following:

$$|N| = \begin{vmatrix} a & dc & bd \\ b & a & cd \\ c & b & a \end{vmatrix} = a^3 + db^3 + d^2c^3 - 3dabc.$$

The determinant is constructed according to the action of the operator on the elements of the expansion base:

$$1\alpha = a + bd^{\frac{1}{3}} + cd^{\frac{2}{3}}$$

$$d^{\frac{1}{3}}\alpha = dc + ad^{\frac{1}{3}} + bd^{\frac{2}{3}}$$

$$d^{\frac{2}{3}}\alpha = bd + cdd^{\frac{1}{3}} + ad^{\frac{2}{3}}$$

Theorem 3.1 $N(ab) = N(a)N(b)$, in particular, if a | b then N (a) | N (b).

The norm of the element $1 + \sqrt{-7}$ in the Galois extension is the product of the conjugates to this element and itself, i.e. $N\left(1 + \sqrt{-7}\right) = \left(1 + \sqrt{-7}\right)\left(1 - \sqrt{-7}\right) = 1 + 7 = 8$ because $1 \in \mathbb{ZZ}$, then it is conjugate to itself and has no other conjugates, the number of conjugates is equal to the degree of expansion of the main field.

Definition 3.12. The number of elements in the factor ring $A\!/\!I$ called the norm of ideal I and denote $N(I)$.

The main property of the norm of the ideal is its multiplicity.

Theorem 3.2. For $\forall I,\ J \in A$ is a ring, the multiplicative property $N\ (IJ)\ =\ N\ (I)N\ (J)$ is fulfilled.

Here $\breve{y}[\theta] = \left\{x : x = a_{d-1}\theta^{d-1} + a_{d-2}\theta^{d-2} + \ldots + a_{0,} a_i \in \breve{y}\right\}$ is algebraic extension of the ring $\breve{y}$.

Definitions 3.13. Element $l \in Z\ [\theta]$ is called smooth over the algebraic factor base A if W⊂And such that $\Pi_{(c,d)\in W}(c + d\theta) = l$

4. Methods of Optimal Formation of a Rational Factor Base.

Definition 4.1. A rational factor base is a finite set of prime numbers that is no larger than a given prime number.

That is it $R = \{p : p \in \mathrm{P},\ p \le M\}$.

Definition 4.2. An algebraic factor base is a finite subset of an algebraic extension $A = \{a + b\theta\} \epsilon\, \breve{\boldsymbol{y}}[\theta]$ such that for $a, b \in \breve{y}, \forall a + b\theta$ satisfies the condition for $\forall (a,b) \epsilon A, \exists c, d \in \breve{y}[\theta] i\, c, d \notin A$ such that $c, d :\ c \cdot d = a + b\theta$ and $c, d \notin A$.

For this reason, it is customary to call the element $a\ +\ b\theta$ generating a prime ideal.

Definition 4.3. *Element* $l \in Z\ \lfloor \theta \rfloor$ is called smooth over the algebraic factor base A if $W \subset A$ and such that $\Pi_{(c,d)\in W}(c + d\theta) = l$.

Theorem 4.4.(On the bijection between the elements of an algebraic factor base and a finite set of pairs). Let the polynomial *f*(x) with integer coefficients such that

$\theta \in J,\ f\ (\theta) = 0$. Then the set of pairs $\{(p,\ r)\}$, where p is prime and r is such that $r \in \mathbb{Z}_n$, $f(r) \equiv (0 mod p)$ is in bijective correspondence with the algebraic factor base $a + b\theta \in \breve{y}[\theta]$.

This theorem makes it possible to represent the algebraic factor base as a set of pairs $\{(p,\ r)\}$ that satisfy the requirements of the theorem.

Remark 4.1. In expansion $\breve{y}[\theta]$ the factor base may become smaller than in $\breve{y}$ because the number that was prime in $\mathbb{Z}$ may not be the case in the extended field $\breve{y}[\theta]$. Technique of finding elements with a given value of smoothness.

To find complete squares you need to find pairs of numbers $a + b\theta \in \breve{y}[\theta]$, which is smooth in some algebraic factor base A and $a + bm$ is smooth in some rational factor base $\mathfrak{R}$.

Let the algebraic factor base A be represented by a set of pairs $\{(p_i, r_i)\}$, the rational factor base be represented by a set of prime numbers $\{p_i\}$.

Theorem 4.2. Let the element $c + d\theta \in A$, such that has a representation $(r,\ p)$, then the element $c + d\theta$ divides $a + b\theta \epsilon\ \breve{y}[\theta]$ then and only if $a \equiv -br(mod p)$.

Theorem 4.3. Finite set U of pairs $(r, p) \in \breve{y}[\theta]$ is a complete set of number divisors $a + b\theta \epsilon\ \breve{y}[\theta]$ if and only if: $\Pi_{(r_i, p_i) \in U} p_i = (-b)^d f\left(-a/b\right), d = \deg(f)$.

Remarks 4.2. The prime number q will divide $a + bm$ if and only if: $a \equiv -bm(\mathrm{mod} q)$.

5. Optimization of Other Parametrs of Alghoritm.

The main parameters of the algorithm are: the first parameter d is the degree of the polynomial that defines the mapping $f(x): X \to X$.

Remarks 5.1. In order for the chosen polynomial $f(x)$ to be optimal, it is expedient to first determine its degree d.

6. Experiments and Results

The main parameters of the algorithm are: the first parameter d is the degree of the polynomial that defines the mapping $f(x): X \to X$. Experimentally established its dependence on the value of n (table 6.1)

Table 6.1

The number of characters in n	<52	<52-82	<82-112	≈ 112
Power $f(x)$	2	3	4	5

The second parameter is a natural number m which satisfies the condition:

$$f(m) \equiv 0(mod n)$$

The number m is chosen after d is determined taking into account $f(m) \equiv 0(mod n)$ and so that it is performed $m \approx \sqrt{n^d}$, to optimize time estimates it is advisable to [5] put $d \approx \left(\frac{(3 + o(1)\log n)}{\log\log n}\right)^{1/3}$. In fact, we form a schedule of the number n: $n = f(m) = a_d m^d + a_{d-1} m^{d-1} + \ldots + a_0$

By definition, the function $f\ (x)$ is:

$$f(x) = a_d x^d + a_{d-1} x^{d-1} + \ldots + a_0$$

And $n = f(m)$ Ago $f(m) \equiv 0(mod n)$ and they are endowed with these properties everywhere in this text. You can better choose polynomials, as developed in [5], namely, by choosing d, we find the smallest integer k such that $kd \geq e$, we put $e = sr^{kd-e}$ and determine f and m by $f = X^d - t, m = r^k$.

Analytical approach to improving the time parameters of the algorithm and the possibilities of its parallelization.

Based on the table of timings (Table 6.1.) For all stages of the algorithm, we see that the process of sieving pairs (a, b) for smoothness takes the largest share of time, and namely, about 70-80%. Therefore, the algorithm will be significantly improved if the time parameters of the screening process can be improved.

The algorithm was tested on a cluster consisting of a main node on two AMD processors, with hyper trading, having 2 cores each. In addition, the cluster has 59 auxiliary nodes, which are built on AMD processors and also have 2 cores. The clock frequency of each node is 2 x 3.0 GHz AMD, DDR-512 ECC SDRAM (3 GB) RAM used.

Table 6.2

The number of digits in the number	Screening, c	Relation, with, c	Lancosh block, c	Square root, c	Total, c	Filtered total, %

30	26,4	3.6	0.1	2.0	32.3	82
39	15	3.1	0.1	1.4	19.7	76.1
45	184.2	45.8	4.1	15.7	250	74
51	222.4	63.9	7.3	18	311. 5	71.4
61	3620	591.7	32,6	57,4	4320 ,4	84
76	26477 ,8	8563,6	1226,3	904,2	3717 1,9	71,2
98	17300 ,7	2716,8	504,6	268,9	2079 ,09	83,2
30	26,4	3.6	0.1	2.0	32.3	82

Because there is no relationship between generating different pairs (a, b) for subsequent sieving, the parallel platform is ideal for the process of improving the GNFS algorithm. The only question left is the optimal number of connections between the slave workstations (nodes).

The parallel algorithm uses one server and the number of workstations is limited to 32. Workstations are not directly connected to each other only through a server. Each workstation has its own subspace of values of the number *b*, obtained by dividing the entire space of values of *b* into *p* stations by the formula:

$$b_j = \frac{b_1 - b_0}{p}.$$

Each workstation looks for a relationship within the range of its value space.

```
$b_0 = Min b, \quad b_1 = Max\_b$
$a_1 = -N,\ a_2 = N;$
$num\_of\_b_j = (b_1 - b_0)/k$
$ParalMode\_B\ (num\_of\_b_0)$
$for\ \ i\ \ in\, taskid:\ \ if\ \ (b == n):$
    $n = i \cdot num\_of\_b_j + b_0$
    $if\ \ (b < n):$
        $b_0 + (i+1) + num\_of\_b_j$
        $b = b + 1$
        $while(1):$
            $a = a_1$
            $a = a + 1$
            $if\ a < a_2:$
                $break$
    $if\ \ Smooth\_R(a,b) and\ Smooth\_A(a,b):$
        $if\ \ master:$
        $if\ master:$
            $MPI\_RE\_cu(a,b)$
            $save(a,b)$
        $else:$
            $MPI\_send(a,b)$
        $b_0 = Min b, \quad b_1 = Max\_b\,,$
```

where $\text{paralMod\_}B(num\_of\_bs)$ is a function that passes the bs of the intervals bs to all subordinate nodes for further processing of the corresponding values from these intervals, and task id is the identifier of the task which each node already knows.

The analysis of sequential sieving timings shows that the total time for large numbers increases noticeably faster than the total time corresponding to them increases with the growth of small numbers (Fig. 6.1).

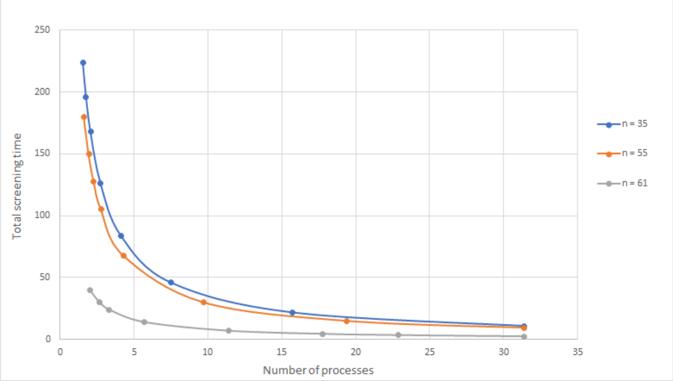

Figure 6.1

Graph on Fig. 6.1 of the dependence of the sifting time on the bit size of the factorized number and the number of processors. The bit size of *n* = 35, 55, 61. The conclusion for the asymptotic approximation to the OX graph is not enough just to increase the number of processors for screening, we need to synchronize these processes.

The graph of parallel execution of the sieving step has a shape similar to the branch of the hyperbola $y = \frac{1}{x}$, which in asymptotic, shows us that with increasing number of processors time of factorization decrease by hyperbolic law, approaches the *OX* axis, although the speed of this approximation depends on the bit size of the factorized number. The schedule of parallel execution of the whole algorithm is very similar to the schedule of parallel execution of the screening stage, because it is the most time-consuming in the algorithm. With one difference, the asymptote of this graph is not the OX axis, but a parallel line that passes slightly above the *OX* axis.

So we have n length intervals, each of which has bs values of the sifted value and n intervals as. Therefore, if we do the sieving process sequentially, the time complexity is $O\left(n^2\right)$, because we search 2 arrays of n elements.

If we have *k* processors for parallel processing, then each processor deals with the range of intervals $\left(n/k,\ b_j\right)$ and $\left(n/k,\ a_j\right)$ then the expected time complexity is as follows: $O(n^2/p)$.

Effective choice of polynomials to improve time efficiency

Parallel execution time can be improved by considering that there are many stages of information transfer between the server and slave nodes. In fact, each such node must send back to the server the sieving result for each *b*. The sieving results include 3 packets. Therefore, the total time of sending these packets to the server $3(b_1-b_0)(k-1)/k$ is significant and it is important that the transmission time will increase with increasing value of n.

The second reason for the significant delay is the asynchrony of steam processing. In fact, the sieving time is different for each pair, so the master node (server) cannot start the next sieving until all subordinate nodes have completed their smoothness check. Meanwhile, most of all processors are not busy. This asynchrony can be eliminated by better balancing the screening process. It is also important to choose the optimal parameters of the polynomial. For example, it is proved that if we reduce the coefficients of the polynomial $f(x)=a_dx^d+a_{d-1}x^{d-1}+...+a_0$, then we get not large numbers:

$$\prod_{(a_j,b_j)\in V}\left(a_j+b_j\theta\right)\cdot$$

This will speed up the work. The choice of the degree of the polynomial is as described above.

Maximum values $\left|F_i(a,b)\right|$ should not be large, where (a,b) - pairs mutually prime numbers such that $\prod_{(a_j,b_j)\in V}\left(a_j+b_j\alpha\right)$ and $\prod_{(a_j,b_j)\in V}\left(a_j+b_j\theta\right)$ are squares in ℤ[α] and ℤ [θ], respectively, it is to these squares and we will apply the homomorphism with ℤ [α] and ℤ[θ] in $\mathbb{Z}_n$. In addition $\left|F_i(a,b)\right|=y^{d_i}f_i\left(x/y\right)\in\mathbb{Z}[x,y]$ – these are homogeneous parts of the polynomial $f_i(x)$, and in the General case for NFS to form polynomials as follows:

$$f_1(x)=a_{d,1}x^d+a_{d-1,1}x^{d-1}+...+a_{d-1,0},$$

$$f_2(x)=a_{d,2}x^d+a_{d-1,2}x^{d-1}+...+a_{d-1,0}.$$

Where $a_{k-1,1}\in\mathbb{Z}$, $f_1(x)\neq\pm f_2(x)$ both are irreducible over ℤ.

In addition, their content $c\left(f_i(x)\right)$ in tableau (6.1) and the number m is the common root of mod n for $f_1(x)$, $f_2(x)$.

$$c\left(f_i(x)\right)=LCM\left(a_{d,i},a_{d-1,i},\dots,a_{d,0}\right)=1,\quad i\in\{1,2\} \ (6.1)$$

In particular, for SNFS, you can find the coefficients of polynomials without the help of a computer.

For GNFS, you can make a successful choice of polynomials to find polynomials as follows:

We put $m=[n^{1/d_1}]$.

Choose the coefficients of the polynomial such $0\le a_{d,i}\le m$ that n decomposes as follows:

$$n=a_{d,1}m^d+a_{d-1,1}m^{d-1}+...+a_{d-1,0}$$

Then

$$f_1(x)=a_{d,1}x^d+a_{d-1,1}x^{d-1}+...+a_{d-1,0},$$

$$f_2(x)=x-m.$$

From this method of the task it follows that $a_{d,1}$=1 [21, 22].

In addition, in [21, 22] there is another method (more general) for choosing polynomials with more than 1 senior coefficient and which allows negative coefficients here only $O\left(n^{1/(d_1+1)}\right)$ acceptable options for choosing the coefficient

$$f_1(x)=a_{2,1}x^2+a_{1,1}x+a_{0,1} \quad ,$$

$$f_2(x)=a_{2,2}x^2+a_{1,2}x+a_{0,2}\in\mathbb{Z}[x].$$

This method is based on Montgomery's idea to choose polynomials of the form (6.6.2) and (6.6.3), which must have a common root *m* modulo *n*, if and only when the vectors $a=\left(c_{0,1},c_{1,1},c_{12}\right)^T$ and $b=\left(c_{0,2},c_{1,2},c_{2,2}\right)^T$ are orthogonal to $\left(1,m,m^2\right)^T$ over $\mathbb{Z}_n$.

$$f_1(x)=a_{2,1}x^2+a_{1,1}x+a_{0,1}, \qquad (6.1.2)$$

$$f_2(x)=a_{2,2}x^2+a_{1,2}x+a_{0,2}\in\mathbb{Z}[x]. \quad (6.1.3)$$

We suppose that $f_1(x)\neq\pm f_2(x)$ both polynomials are irreducible over ℤ and their contents are equal to 1.

It is not difficult to show that it is practically possible to find a=$(c_{0,1},c_{1,1},c_{2,1})^T$ and $b=\left(c_{0,2},c_{1,2},c_{2,2}\right)^T$, whose coefficients are approximately equal $O(n^{1/4})$. Thus, space is orthogonal to vectors $a=\left(c_{0,1},c_{1,1},c_{2,1}\right)^T$ and $b=\left(c_{0,2},c_{1,2},c_{2,2}\right)^T$ has rank 1. If $c=a\times b$ (vector product), then *c* must be a multiple $(1,m,m^2)^T$ over $\mathbb{Z}_n$.

The fact that the polynomials

$$f_1(x)=a_{2,1}x^2+a_{1,1}x+a_{0,1}$$

and

$$f_2(x)=a_{2,2}x^2+a_{1,2}x+a_{0,2}\in\mathbb{Z}[x]\cdot$$

not multiples of each other, guarantees that $\vec{c}\neq\vec{0}$.

Good are polynomials that have real roots approximately equal $\frac{\max|a|}{\max|b|}$ This is shown by the diagram (Fig.6.2), constructed for two polynomials, one of them $f_1(x)=x^5+x^4-4x^3-3x^2+3x+1$, which has 5 valid roots, was 60,000 values are sifted a and 8625 values b , to factorize 119 significant numbers. It is those pairs (a,b) for which the values a/b are approximately equal to the roots of the polynomial, gave a larger number of ratios, as shown in the graph (Fig.6.2), which has 5 corresponding to the roots of the convex waves.

Polynomials with roots of small prime numbers (mostly different) are better than those that do not.

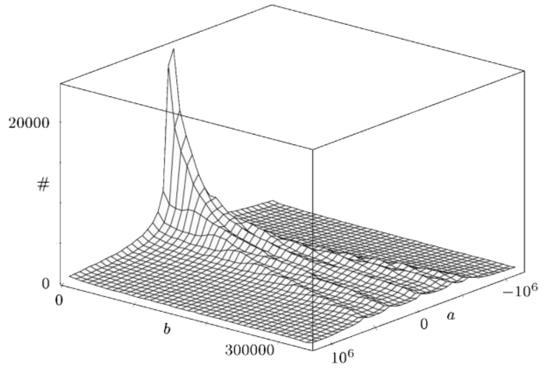

Fig 6.2

It is investigated that more dependences will be found if we choose such polynomials of fixed degree, the order of the Galois group of which is as small as possible.

Recall that the Galois group is a group of automorphisms (in the extended field) of the roots of the polynomial by which this extension is constructed.

For different types of polynomials, these groups are different. Thus, for a cyclotomic polynomial of prime order, the Galois group is cyclic of the same prime order. For instance the cyclotomic polynomial

$$\frac{x^7-1}{x-1}=x^6+x^5+...+x+1$$

has a Galois group a cyclic group of order 7. A non-decomposable cyclotomic polynomial of degree *n* has a Galois group of order *n*. A folding cyclotomic polynomial of degree *n* has a Galois group of order *n*!

And the cyclotomic polynomial of not prime order is set recurrently:

$$\frac{x^n-1}{\prod\limits_{\substack{d|n\\d<n}} f(x)}\frac{n!}{r!(n-r)!}$$

The order of his group $\varphi(n)$. A polynomial $f(x)=\prod\left(x-\varepsilon^j\right)$, $\varepsilon-$has a primitive root of 1.

In addition, it should be borne in mind that polynomials of degree which are irreducible and non-cyclotomic have the order of the Galois group *n*! Tha1t is, their Galois group is large enough.

Conclusion: It is advisable to choose a indecomposable cyclotomic polynomial.

References

[1] Laurence T. Yang, Li Xu, Sang-Soo Yeo, Sajid Hussain, An integrated parallel GNFS algorithm for integer factorization based on Linbox Montgomery block Lanczos method over GF(2), Computers & Mathematics with Applications, Volume 60, Issue 2, 2010, 60, pp. 338-346.

[2] L. Xu, L.T. Yang, M. Lin, Parallel general number field sieve method for integer factorization, in: Proceedings of the 2005 International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA-05, Las Vegas, USA, June 2005, pp. 1017–1023.

[3] L.T. Yang, L. Xu, M. Lin, J. Quinn, A parallel GNFS algorithm based on a reliable look-ahead block Lanczos method for integer factorization, in: Proceedings of the 2006 IFIP International Conference on Embedded and Ubiquitous Computing, EUC-06, Seoul, Korea, August 1–4 2006, pp. 110–120.

[4] C. Monico, General number field sieve documentation. GGNFS Documentation, Nov. 2004.

[5] Gad, Ibrahim & Daoud, Sameh. (2014). A parallel line sieve for the GNFS Algorithm. International Journal of Advanced Computer Science and Applications. 5. 178. 10.14569/IJACSA.2014.050727.

[6] Briggs, Matthew. (2021). An Introduction to the General Number Field Sieve. (ABSTRACT) With the proliferation of computers into homes and businesses and the explosive growth rate

[7] Buhler, J. & Lenstra, H. & Pomerance, Carl. (2006). Factoring integers with the number field sieve. 10.1007/BFb0091539.

[8] I. Niven, H. S. Zuckerman, and H. L. Montgomery, An Introduction To The Theory Of Numbers, 5th ed. Wiley, 1991.

[9] R. M. Huizing, “An implementation of the number field sieve, Tech. Rep. NM-R9511, 1995. [Online]. Available:citeseer.nj.nec.com/huizing95implementation.html

[10] Case, Michael A.. “A Beginner ’ s Guide To The General Number Field Sieve.” (2003).

[11] Skuratovskii R.V. Factorization of a number in the form *n* = *pq*. Journal of Mathematical and Computer Modeling. Series: Physical and Mathematical Sciences. 2017.pp. 201-208.

[12] Skuratovskii R.V. The Investigation of Euler’s Totient Function Preimages. Journal of Applied Mathematics and Computation (JAMC), 2019, 3(3), 591-598.

[13] Laurence T. Yang, Li XuAn integrated parallel GNFS algorithm for integer Computers and Mathematics with Applications Computers and Mathematics with Applications.

[14] Skuratovskii, R.V.: The timer compression of data and information. In: Proceedings of the 2020 IEEE 3rd International Conference on Data Stream Mining and Processing, DSMP 2020, pp. 455–459 (2020)

[15] Skuratovskii, R.V.: Employment of minimal generating sets and structure of sylow 2-subgroups alternating groups in block ciphers. In: Bhatia, S., Tiwari, S., Mishra, K., Trivedi, M. (eds.) Advances in Computer Communication and Computational Sciences. Advances in Intelligent Systems and Computing, vol. 759. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-0341-8_32

[16] Skuratovskii, R.V., Williams, A.: Irreducible bases and subgroups of a wreath product in applying to diffeomorphism groups acting on the Möbius band. Rend. Circ. Mat. Palermo Ser. 2, 1–19 (2020). https://doi.org/10.1007/s12215-020-00514-5

[17] Skuratovskii, R.V.: A method for fast timer coding of texts. Cybern. Syst. Anal. 49(1), 133–138 (2013).

[18] Skuratovskii, R., Osadchyy, V., Osadchyy, Y.: The timer inremental compression of data and information. WSEAS Trans. Math. 19, 398–406 (2020). DOI: 10.37394/23206.2020.19.41.

[19] Skuratovskii, R., Osadchyy, V. WSEAS Criterions of supersinguliarity and groups of Montgomery and Edwards curves in cryptography. Transactions on Mathematicsthis link is disabled, 2020, 19, pp. 709–722

[20]. Richard P. Brent. Some Integer Factorization Algorithms using Elliptic Curves Computer Sciences Laboratory Australian National University. Australian Computer Science Communications 8 (1986), 149-163 https://doi.org/10.1007/s10559-013-9493-4

[21] Gireesh Pandey, S.K. Pal,Polynomial selection in number field sieve for integer factorization, Perspectives in Science,Volume 8, 2016, Pages 101-103,.

[22] Laurence T. Yang, Gaoyuan Huang, Jun Feng, Li Xu, Parallel GNFS algorithm integrated with parallel block Wiedemann algorithm for RSA security in cloud computing, Information Sciences, Volume 387, 2017, Pages 254-265

[22] Bhat, Mohd & Giri, Kaiser. (2021). Impact of Computational Power on Cryptography. 10.1007/978-981-15-8711-5_4.

[23] Ruslan Skuratovskii, "The Investigation of Euler’s Totient Function Preimages for φ(n)=2mp1αp2β and the Cardinality of Pre-totients in General Case", WSEAS Transactions on Mathematics, vol. 21, pp. 44-52, 2022. DOI: 10.37394/23206.2022.21.7