Polynomial time reduction from 3SAT to solving low first fall degree multivariable cubic equations system (Draft)

Koh-ichi Nagao (nagao@kanto-gakuin.ac.jp)

Faculty of Science and Engineering, Kanto Gakuin Univ.,

Abstract. Recently, there are many researches [5] [3] [7] [4] that, under the first fall degree assumption, the complexity of ECDLP over $\mathbb{F}_{p^n}$ where p is small prime and the extension degree n is input size, is subexponential. However, from the recent research, the first fall degree assumption seems to be doubtful. Koster [2] shows that the problem for deciding whether the value of Semaev's formula $S_m(x_1, ..., x_m)$ is 0 or not, is NP-complete. This result directly does not means ECDLP being NP-complete, but, it suggests ECDLP being NP-complete. Further, in [7], Semaev shows that the equations system using $m-2$ number of $S_3(x_1, x_2, x_3)$, which is equivalent to decide whether the value of Semaev's formula $S_m(x_1, ..., x_m)$ is 0 or not, has constant(not depend on m and n) first fall degree. So, under the first fall degree assumption, its complexity is poly in n ($O(n^{Const})$). And so, suppose $P \neq NP$, which almost all researcher assume this, it has a contradiction and we see that first fall degree assumption is not true.
Koster shows the NP-completeness from the group belonging problem, which is NP-complete, reduces to the problem for deciding whether the value of Semaev's formula $S_m(x_1, ..., x_m)$ is 0 or not, in polynomial time.
In this paper, from another point of view, we discuss this situation. Here, we construct some equations system defined over arbitrary field K and its first fall degree is small, from any 3SAT problem. The cost for solving this equations system is polynomial times under the first fall degree assumption. So, 3SAT problem, which is NP-complete, reduced to the problem in P under the first fall degree assumption.
Almost all researcher assume $P \neq NP$, and so, it concludes that the first fall degree assumption is not true. However, we can take $K = \mathbb{R}$(not finite field!!!). It means that 3SAT reduces to solving multivariable equations system defined over $\mathbb{R}$ and there are many method for solving this by numerical computation.
So, I must point out the very small possibility that NP complete problem is reduces to solving cubic equations equations system over $\mathbb{R}$ which can be solved in polynomial time.

1 Boolean Algebra

Let $X_1, ..., X_N$ be Boolean variables and $x_1, ..., x_N \in \{0, 1\}$ be the Boolean values. Boolean tables (Here $\neg, \vee, \wedge$ mean NOT, OR, AND respectively, $0, 1$ means **False, True** respectively) is written like as follows:

X_1	X_2	$\neg X_1$	$X_1 \vee X_2$	$X_1 \wedge X_2$
0	0	1	0	0
0	1	1	1	0
1	0	0	1	0
1	1	0	1	1

Definition 1. *$X_1, ..., X_N, \neg X_1, ..., \neg X_N$ are called literals.*
The formula connecting literals by OR ($\vee$)is called node.
The node including exact 3 literals is called 3L-node.
The formula connecting nodes by AND ($\wedge$) is called CNF.
The formula connecting 3L-nodes by AND ($\wedge$) is called 3CNF.

Note: we do not consider the node including both of X_i and $\neg X_i$, since this node equals to 1 and can be omitted(for example $X_1 \vee X_2 \vee \neg X_1 \equiv 1$).

Definition 2. *Let $F = F(X_1, ..., X_N)$ be a Boolean formula. If there exists some $x_1, ..., x_N \in \{0, 1\}$ such that $F(x_1, ..., x_N) = 1$, F is called satisfiable.*

Example 1 *$X_1 \vee \neg X_2 \vee X_3$, $X_2 \vee X_3 \vee X_4$ are 3L-nodes.*
$F = (X_1 \vee \neg X_2 \vee X_3) \wedge (X_2 \vee X_3 \vee X_4)$ is 3CNF.
$F(0, 0, 0, 1) = 1$. So F is satisfiable.

Problem 1 *The problem, deciding Boolean formula F is satisfiable or not, is called SAT. The problem, deciding 3CNF F is satisfiable or not, is called 3SAT.*

Proposition 1 (Cook [1], cf [9]) *1) 3SAT is NP-complete.*
2) SAT reduces to 3SAT in polynomial time and so SAT is also NP-complete.

Personal Note: I have a question why 3L-node has exact 3 literals and the nodes having 1 or 2 literals are omitted. Note that $X_1 \wedge (X_1 \vee X_2 \vee X_3) \wedge \cdots$ can be transformed into $X_1 \wedge \cdots$ and $X_1 \wedge (\neg X_1 \vee X_2 \vee X_3) \wedge \cdots$ can be transformed into $X_1 \wedge (X_2 \vee X_3) \wedge \cdots$. Let F be a CNF having nodes including 1,2 and 3 literals. From this observations, there are some variables $X_{i1}, .., X_{il}$ and the literal $L_1 = X_{i1}$ or $\neg X_{i1}$,...,$L_l = X_{il}$ or $\neg X_{il}$, and the nodes $R_1, ..., R_{l'}$ consists of 2 or 3 literals and variables $X_{i1}, .., X_{il}$ do not appear, such that $F \equiv L_1 \wedge .... \wedge L_l \wedge R_1 \wedge ..... \wedge R_{l'}$. Then we see that the condition F being satisfiable is equivalent to $R_1 \wedge ..... \wedge R_{l'}$ being satisfiable. And so, the nodes consists of only one literal can be ignored. Suppose F is a CNF whose nodes have 2 or 3 literals. For example, $F = (X_1 \vee \neg X_2) \wedge (X_2 \vee X_3) \wedge (X_1 \vee X_4 \vee \neg X_6)$..... Let X_{dummy} be a variable which is not included in the nodes having 2 literals. Put $\overline{F} := (X_1 \vee \neg X_2 \vee X_{dummy}) \wedge (X_1 \vee \neg X_2 \vee \neg X_{dummy}) \wedge (X_2 \vee X_3 \vee X_{dummy}) \wedge (X_2 \vee X_3 \vee \neg X_{dummy}) \wedge (X_1 \vee X_4 \vee \neg X_6)$.... So, we see $\overline{F}$ is 3CNF and the condition F being satisfiable is equivalent to $\overline{F}$ being satisfiable.

2 Embedding into polynomial ring

Here, we consider the set of Boolean value$\{0, 1\} = \mathbb{F}_2$ and Boolean variable $X_1, ..., X_N$ as normal variables over $\mathbb{F}_2$ and consider the polynomial ring $\mathbb{F}_2[X_1, ..., X_N]$. Boolean table of X_1, X_2 and addition and multiplication in $\mathbb{F}_2[X_1, X_2]$ is written as follows

X_1	X_2	$\neg X_1$	$X_1 \vee X_2$	$X_1 \wedge X_2 = X_1 \cdot X_2$	$X_1 + X_2$
0	0	1	0	0	0
0	1	1	1	0	1
1	0	0	1	0	1
1	1	0	1	1	0

Note that

$$\neg X_1 = 1 - X_1,$$
$$X_1 \vee X_2 = X_1 + X_2 + X_1 \cdot X_2,$$
$$X_1 + X_2 = (X_1 \vee X_2) \wedge (\neg(X_1 \wedge X_2)).$$

So, from the operations $\neg, \vee, \wedge$, the operations $1 - *, \cdot, +$ are obtained and conversely, from the operations, $1 - *, \cdot, +$ the operations $\neg, \vee, \wedge$ are obtained

Let K be an arbitrary field and

$$|_K : \mathbb{F}_2 \to K$$

be a map $0|_K = 0, 1|_K = 1$. Here , we once more show the Boolean table of X_1, X_2, X_3 (here 3 variables).

X_1	X_2	X_3	1	X_1X_2	X_2X_3	X_3X_1	$X_1X_2X_3$
0	0	0	1	0	0	0	0
0	0	1	1	0	0	0	0
0	1	0	1	0	0	0	0
0	1	1	1	0	1	0	0
1	0	0	1	0	0	0	0
1	0	1	1	0	0	1	0
1	1	0	1	1	0	0	0
1	1	1	1	1	1	1	1

Let

$$M := \begin{pmatrix} 0\,0\,0\,1\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,0\,0 \\ 0\,1\,0\,1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,1\,0\,0 \\ 1\,0\,0\,1\,0\,0\,0\,0 \\ 1\,0\,1\,1\,0\,0\,1\,0 \\ 1\,1\,0\,1\,1\,0\,0\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1 \end{pmatrix} \in GL_8(K).$$

Note that this matrix is coming from above table. From direct calculation, we have determinant of M is -1 and

$$M^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{pmatrix}.$$

Let $3LN$ is the set of 3L-nodes. We define the map $\phi : 3LN \to K[Y_1, ..., Y_N]$ as follows. Let $F = F(X_{i1}, X_{i2}, X_{i3})$ be a 3L-node, written by literals $X_{i1}, X_{i2}, X_{i3}, \neg X_{i1}, \neg X_{i2}, \neg X_{i3}$ (For a while, we consider F is a function of only 3 variables X_{i1}, X_{i2}, X_{i3}). Put
$a_0 := F(0,0,0)|_K, a_1 := F(0,0,1)|_K, a_2 := F(0,1,0)|_K, a_3 := F(0,1,1)|_K,$
$a_4 := F(1,0,0)|_K, a_5 := F(1,0,1)|_K, a_6 := F(1,1,0)|_K, a_7 := F(1,1,1)|_K$ and
$\overrightarrow{a_F} :=^t (a_0, ..., a_7)$. Also put $\overrightarrow{b_F} =^t (b_0, ..., b_7)$ by $\overrightarrow{b_F} := M^{-1}\overrightarrow{a_F}$. From this preparation, we define

$$\phi(F) := 1 - (Y_{i1}, Y_{i2}, Y_{i3}, 1, Y_{i1}Y_{i2}, Y_{i2}Y_{i3}, Y_{i3}Y_{i21}, Y_{i1}Y_{i2}Y_{i3}) \cdot^t (b_0, ..., b_7).$$

Now, we stop to consider F is a function of 3 variables X_{i1}, X_{i2}, X_{i3} and consider F is a function of whole variables $X_1, ..., X_N$.

Example 2 *Let $F_1 = X_1 \vee \neg X_2 \vee \neg X_3$ and Boolean tables is written as follows:*

X_1	X_2	X_3	$\neg X_2$	$\neg X_3$	F
0	0	0	1	1	1
0	0	1	1	0	1
0	1	0	0	1	1
0	1	1	0	0	0
1	0	0	1	1	1
1	0	1	1	0	1
1	1	0	0	1	1
1	1	1	0	0	1

So, $\overrightarrow{a_{F_1}} =^t (1,1,1,0,1,1,1,1)$ and $\overrightarrow{b_{F_1}} =^t (0,0,0,1,0,-1,0,1)$. Thus we have $\phi(F_1) = Y_2Y_3 - Y_1Y_2Y_3$.

Example 3 *Let $F_2 = X_2 \vee X_3 \vee X_4$ and Boolean tables is written as follows:*

X_2	X_3	X_4	F_2
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

So, $\overrightarrow{a_{F_2}} =^t (0,1,1,1,1,1,1,1)$ and $\overrightarrow{b_{F_1}} =^t (1,1,1,0,-1,-1,-1,1)$. Thus we have $\phi(F_2) = 1 - Y_2 - Y_3 - Y_4 + Y_2Y_3 + Y_3Y_4 + Y_4Y_2 - Y_2Y_3Y_4$.

By a direct calculation, we have the following table;

L_1	L_2	L_3	$\phi(L_1 \vee L_2 \vee L_3)$
X_1	X_2	X_3	$-Y_1 - Y_2 - Y_3 + Y_1Y_2 + Y_2Y_3 + Y_3Y_1 - Y_1Y_2Y_3 + 1$
X_1	X_2	$\neg X_3$	$Y_3 - Y_2Y_3 - Y_3Y_1 + Y_1Y_2Y_3$
X_1	$\neg X_2$	X_3	$Y_2 - Y_1Y_2 - Y_2Y_3 + Y_1Y_2Y_3$
X_1	$\neg X_2$	$\neg X_3$	$Y_2Y_3 - Y_1Y_2Y_3$
$\neg X_1$	X_2	X_3	$Y_1 - Y_1Y_2 - Y_3Y_1 + Y_1Y_2Y_3$
$\neg X_1$	X_2	$\neg X_3$	$Y_3Y_1 - Y_1Y_2Y_3$
$\neg X_1$	$\neg X_2$	X_3	$Y_1Y_2 - Y_1Y_2Y_3$
$\neg X_1$	$\neg X_2$	$\neg X_3$	$Y_1Y_2Y_3$

We can consider ϕ as a function from the set of nodes including less than 3 literals. The images of the nodes including less than 2 literals are written by the following table;

L_1	L_2	$\phi(L_1 \vee L_2)$
X_1	X_2	$-Y_1 - Y_2 + Y_1Y_2 + 1$
X_1	$\neg X_2$	$Y_2 - Y_1Y_2$
X_1	X_3	$-Y_1 - Y_3 + Y_3Y_1 + 1$
X_1	$\neg X_3$	$Y_3 - Y_3Y_1$
$\neg X_1$	X_2	$Y_1 - Y_1Y_2$
$\neg X_1$	$\neg X_2$	Y_1Y_2
$\neg X_1$	X_3	$Y_1 - Y_3Y_1$
$\neg X_1$	$\neg X_3$	Y_3Y_1
X_2	X_3	$-Y_2 - Y_3 + Y_2Y_3 + 1$
X_2	$\neg X_3$	$Y_3 - Y_2Y_3$
$\neg X_2$	X_3	$Y_2 - Y_2Y_3$
$\neg X_2$	$\neg X_3$	Y_2Y_3

$$\phi(X_1) = -Y_1 + 1, \phi(\neg X_1) = Y_1, \phi(X_2) = -Y_2 + 1, \phi(\neg X_2) = Y_2,$$

$$\phi(X_3) = -Y_3 + 1, \phi(\neg X_3) = Y_3, \phi(1) = 0, \phi(0) = 1.$$

From the construction of ϕ, we have this key Lemma.

Lemma 1 (Key Lemma). *For any $(x_1, ..., x_N) \in \mathbb{A}^N(\mathbb{F}_2)$ and $F \in 3LN$,*

$$(\neg F(x_1, ..., x_N))|_K = \phi(x_1|_K, ..., x_N|_K).$$

Note and check that above two examples satisfy this lemma.

Definition 3. *Put*

$$S_{BE} := \{Y_i^2 - Y_i \,|\, i = 1, ..., N\} \subset K[Y_1, ..., Y_N].$$

S_{BE} is called Boolean equations.

Let $3CNF$ is the set of 3CNFs. Now, we will define the map $\Phi : 3CNF \to \wp(K[Y_1, ..., Y_N])$ from 3CNF to finite subset of $K[Y_1, ..., Y_N]$.

Let $F = F_1 \wedge ... \wedge F_l$ be a 3CNF. Remember that F_i's are 3L-nodes. Put

$$\Phi(F) = \{\phi(F_1), ..., \phi(F_l)\} \cup S_{BE}.$$

From the construction of Φ and Lemma 1, we have the following:

Theorem 2. *Let $F = F(X_1, ..., X_N)$ be a 3CNF. Then the conditions 1) and 2) are equivalent.*
1) F is satisfiable.
2) There exists some $y_1, ..., y_N \in K$ such that for any $f = f(Y_1, ..., Y_N) \in \Phi(F)$, $f(y_1, ..., y_N) = 0$.

Example 4 *Let* $F = (X_1 \vee \neg X_2 \vee \neg X_3) \wedge (X_2 \vee X_3 \vee X_4)$.
We have

$$\Phi(F) := \{Y_2Y_3 - Y_1Y_2Y_3, 1 - Y_2 - Y_3 - Y_4 + Y_2Y_3 + Y_3Y_4 + Y_4Y_2 - Y_2Y_3Y_4\} \cup S_{BE}.$$

and Boolean tables is written as follows:

X_1	X_2	X_3	X_4	$X_1 \vee \neg X_2 \vee \neg X_3$	$X_2 \vee X_3 \vee X_4$	F
0	0	0	0	1	0	0
0	0	0	1	1	1	1
0	0	1	0	1	1	1
0	0	1	1	1	1	1
0	1	0	0	1	1	1
0	1	0	1	1	1	1
0	1	1	0	0	1	0
0	1	1	1	0	1	0
1	0	0	0	1	0	0
1	0	0	1	1	1	1
1	0	1	0	1	1	1
1	0	1	1	1	1	1
1	1	0	0	1	1	1
1	1	0	1	1	1	1
1	1	1	0	1	1	1
1	1	1	1	1	1	1

We can easily check equations system $\{f(Y_1, ..., Y_4) = 0 \,|\, f \in \Phi(F)\}$ *has solution* $\{(0,0,0,1), (0,0,1,0), (0,0,1,1), (0,1,0,0), (0,1,0,1), (1,0,0,1), (1,0,1,0), (1,0,1,1), (1,1,0,0), (1,1,0,1), (1,1,1,0), (1,1,1,1)\}$ *and the notation of Theorem 2 holds.*

3 First fall degree assumption

Definition 4 (First fall degree). *Let* K *be a field and* $f_1, ..., f_M \in K[Y_1, ..., Y_N]$. *First fall degree of* $\{f_1, ..., f_M\}$ *is the minimal integer* d_F *satisfying the following.*
There exists $g_1, ..., g_M \in K[Y_1, ..., Y_N]$ *such that*
1) $\max_i\{\deg g_i f_i\} \geq d_F$,
2) $\deg(\sum_{i=1}^{M} g_i f_i) < d_F$,
3) $\sum_{i=1}^{M} g_i f_i \neq 0$.

Under the following assumption, the algorithm for solving ECDLP in sub-exponential complexity are proposed [5], [3], [7].

Assumption 1 $\{f_1, ..., f_M\}$ *Degree of the polynomial appears in the Gröbner basis computation (by F4 algorithm) of* $\{f_1, ..., f_M\}$ *is* $\leq d_F$.

From this assumption, the number of the monomial appears in the Gröbner basis computation is $\leq O(N^{d_F})$ So, we have the following;

Lemma 2. *The complexity of Gröbner basis computation (by F4 algorithm) of* $\{f_1, ..., f_M\}$ *is* $\leq O(N^{d_F w})$, *where* $w \sim 2.7$ *is the linear algebra constant.*

Proposition 3 *Let* F *be a 3L-node. Then the first fall degree of* $\{\phi(F)\} \cup S_{BE}$ *is* ≤ 4.

Proof. From the table before Lemma 1, we have $\deg \phi(F) = 3$. So, $\phi(F)$ is written by

$$\phi(F) = Y_{i1}Y_{i2}Y_{i3} + \sum \text{ of the terms degree } \leq 2.$$

When $\phi(F) \neq Y_{i1}Y_{i2}Y_{i3} + Y_{i2}Y_{i3}$,

$$Y_{i1}\phi(F) - Y_{i2}Y_{i3}(Y_{i1}^2 - Y_{i1}) = \sum \text{ of the terms degree } \leq 3 \neq 0.$$

So, the first fall degree is ≤ 4. When $\phi(F) = Y_{i1}Y_{i2}Y_{i3} + Y_{i2}Y_{i3}$,

$$Y_{i2}\phi(F) - Y_{i1}Y_{i3}(Y_{i2}^2 - Y_{i2}) = \sum \text{ of the terms degree } \leq 3 \neq 0.$$

So, the first fall degree is ≤ 4. Thus we finish the proof.

From this Proposition, we have the following theorem.

Theorem 4. *Let $F = F(X_1, ..., X_N)$ be a 3CNF. Then the first fall degree of $\Phi(F)$ is ≤ 4.*

4 Conclusion

Here, we construct some equations system defined over arbitrary field K and its first fall degree is ≤ 4, from any 3SAT problem. The important trick of this paper is as follows; The Boolean equation can easily be transformed to the equations system over $\mathbb{F}_2$. However, by using Boolean equation of the form $\{Y_i^2 - Y_i = 0\}$, 3CNF can be transformed to the equations system over arbitrary field K.

The cost for solving this equations system is polynomial times under the first fall degree assumption. So, 3SAT problem, which is NP-complete, reduced to the problem in P under the first fall degree assumption.

Almost all researcher assume $P \neq NP$, and so, it concludes that the first fall degree assumption is not true. However, we can take $K = \mathbb{R}$(not finite field). It means that 3SAT reduces to solving multivariable equations system defined over $\mathbb{R}$ and there are many method for solving this by numerical computation. So, I must point out that there are some (but very very small) possibility that $P = NP$ is true (it means any NP problem reduces to solving some multivariable cubic equations system which can be solved in polynomial time).

Acknowledgement I would like to have great thanks to Professor Kazuto Matsuo in Kanagawa University for useful advices and coments.

References

1. S. A. Cook, The Complexity of Theorem Proving Procedures. Proceedings Third Annual ACM Symposium on Theory of Computing, May 1971, pp 151-158. `http://4mhz.de/cook.html`
2. M. Kosters, NOTES ON SUMMATION POLYNOMIALS, http://arxiv.org/pdf/1503.08001.pdf 2015.
3. K. Nagao, Equations System coming from Weil descent and subexponential attack for algebraic curve cryptosystem, https://eprint.iacr.org/2013/549
4. K. Nagao, Complexity of ECDLP under the First Fall Degree Assumption, draft, 2015.
5. C. Petit and J-J. Quisquater. On Polynomial Systems Arising from a Weil Descent, Asiacrypt 2012, Springer LNCS **7658**, Springer, pp.451-466.
6. I. Semaev. Summation polynomials and the discrete logarithm problem on elliptic curves. https://eprint.iacr.org/2004/031.pdf
7. I. Semaev, New algorithm for the discrete logarithm problem on elliptic curves, https://eprint.iacr.org/2015/310.pdf
8. G. Takeuchi, P and NP, Nipponhyouronsha, 1996.
9. O. Watanabe, P $\neq$ NP Conjecture, Koudansha, 2014.