

Introducing Quaternions to Integer Factorisation

HuiKang Tong

4500 Ang Mo Kio Avenue 6, 569843, Singapore

Abstract: The key purpose of this paper is to open up the concepts of the sum of four squares and the algebra of quaternions into the attempts of factoring semiprimes, the product of two prime numbers. However, the application of these concepts here has been clumsy, and would be better explored by those with a more rigorous mathematical background. There may be real immediate implications on some RSA numbers that are slightly larger than a perfect square.

Key words: Integer factorisation, RSA, quaternions, sum of four squares, euler factorisation method.

Nomenclature

p, q :	prime factors
n :	semiprime pq , the product of two primes
P :	quaternion with norm p
a, b, c, d :	components of a quaternion

1. Introduction

We assume that the reader know the RSA cryptosystem [1]. Notably, the ability to factorise a random and large semiprime n (the product of two prime numbers p and q) efficiently can completely break RSA, which is integral to many cryptographic systems worldwide. We also assume that the reader is familiar with the state-of-the-art factoring methods [2], although its knowledge is not required to understand our ideas in this paper, but is useful to judge the novelty of our work.

1.1 Outline of the Paper

The Euler factorisation method is introduced in Section 2, while our probably new pathway to the solution with the Gaussian integers is explained in Section 2.2, after the introduction to Gaussian integers in Section 2.1. Then we state the limitations of the Euler factorisation method and address one of the work previously done to make Euler factorisation method workable.

In Section 3, we extend the Euler factoring method to one using the sum of four squares and the algebra of quaternions. We comment on the development of the mathematics in Section 3.1, and introduce the integral quaternions in Section 3.2, and its relationship with the sum of four squares in Section 3.3. In Section 3.4, we mention an algorithm to generate the sum of four squares.

In Section 4, we propose the usage of concepts of the algebra of quaternions into the factorisation of semiprimes. The method in Section 2.2 using Gaussian integers is brought into Section 4.1. Then we subsequently loosen one of its constrains in Section 4.2. We will also show the Euclidean algorithm for quaternions. Then we propose an alternative method in Section 4.3, the one which we feel has better potential to contribute to the research on the integer factorisation problem.

Finally we discuss the results in Section 5 and conclude in Section 6.

1.2 Contributions of the Paper

The main contribution of this paper is to open up the concepts of sum of four squares in quaternions into the attempts of factoring semiprimes. Sections 2.1, 4.1 and 4.2 are adapted from standard mathematical texts. Expressing the Euler's factoring method with Gaussian integers has not been specifically published on, so content in Section 2.2 may be new. To the

author's knowledge, quaternions has never been employed in integer factorisation, so works from Section 4 onwards are original.

2. The Euler Factorisation Method

We will now look at the much neglected Euler factorisation method [3]. In essence, it instantly obtains the factors of a semiprime when it is written as a sum of two square in two distinct ways.

$$pq = n = a^2 + b^2 = c^2 + d^2$$

2.1 The Gaussian Integers

The Gaussian integers are described in this section so that the reader can draw parallels with the algebra of quaternions, where the main ideas are built on.

Gaussian integers are complex numbers with rational integers as coefficients for its real and imaginary components [4]:

$$\mathbb{Z}[i] = \{\alpha = a + bi : a, b \in \mathbb{Z}\}$$

Similar to complex numbers, Gaussian integers has its conjugate:

$$\bar{\alpha} = a - bi$$

The norm of any Gaussian integer is defined as:

$$Nm(\alpha) = \alpha\bar{\alpha} = (a - bi)(a + bi) = a^2 + b^2 \in \mathbb{R}$$

For Gaussian integers, the product of the norm is equal to the norm of its product:

$$\begin{aligned} Nm(\alpha)Nm(\beta) &= (\alpha\bar{\alpha})(\beta\bar{\beta}) = \alpha\beta\bar{\beta}\bar{\alpha} = \alpha\beta\bar{\alpha}\bar{\beta} \\ &= Nm(\alpha\beta) \end{aligned}$$

A Gaussian prime cannot be expressed in terms of two Gaussian integers that each has a smaller norm [5]. Hence, any prime number p congruent to $1 \pmod{4}$ is not a Gaussian prime, as it is factorable into:

$$p = \rho\bar{\rho} = (r + si)(r - si)$$

2.2 A Complex Approach

The Euler factorisation method can be done with Gaussian integers. We have independently derived this relationship, although we expect one to have made this minor result. However, to the author's knowledge, nothing is specifically published on this concept.

An alert reader would have noticed that norm of a

Gaussian integer is a sum of two squares, which is the basis of the Euler factorisation method. The explanation will be illustrated by a numerical example that is first factored by Euler:

$$\begin{aligned} pq = n &= a^2 + b^2 = c^2 + d^2 \\ 1000009 &= 1000^2 + 3^2 = 972^2 + 235^2 \end{aligned}$$

The sum of two squares can be factorised into Gaussian integers:

$$\begin{aligned} pq &= (a + bi)(a - bi) = (c + di)(c - di) \\ 1000009 &= (1000 + 3i)(1000 - 3i) \\ &= (972 + 235i)(972 - 235i) \end{aligned}$$

The Gaussian integers can be further factorised. The two forms of the semiprime are actually re-arrangements of each other:

$$\begin{aligned} pq &= [(r + si)(t + ui)][(t - ui)(r - si)] \\ &= [(r + si)(t - ui)][(t + ui)(r - si)] \end{aligned}$$

where $p = (r + si)(r - si)$ and $q = (t + ui)(t - ui)$

The norm of one of the terms is one of the factors. The term is obtained by the Euclidean algorithm, which works for complex numbers [6]. The only difference from the algorithm for rational integers is that the remainder can be negative, so that the size of the remainder can be continually reduced.

$$(1000 + 3i) = (1)(972 + 235i) + (28 - 232i)$$

$$\begin{aligned} (972 + 235i) &= (-1 + 4i)(28 - 232i) \\ &\quad + (72 - 109i) \end{aligned}$$

$$(28 - 232i) = (2 - i)(72 - 109i) + (-7 + 58i)$$

$$(72 - 109i) = (-2 - i)(-7 + 58i) + (0)$$

The norm of the greatest common divisor $Nm(-7 + 58i) = 3413$ is a prime factor.

2.3 Evaluation of the Method

The Euler factorisation method works only when both prime factors is congruent to $1 \pmod{4}$, because only such semiprimes have the two representations [7]. More importantly, there has not been a feasible way to find the two representations of sum of two squares for large semiprimes.

However, in our literature review, we came across a work that made the Euler factorisation method feasible [8]. The result was an algorithm that factorises

in $O(n^{1/3})$ time. Our proposal focuses on quaternions and the sum of four squares which should be totally different from his ideas.

3. Extending to Quaternions

Extending the concept of the Euler factorisation method beyond Gaussian integers, we subsequently explore the quaternions, and the sum of four squares. Here is our justification why this may be novel.

3.1 Comment on Mathematical Development

Much rigorous work has already been done on the separate topics. RSA was patented in 1977, and the RSA factoring challenge was put up on 1991. Attention on factorising large semiprimes with factors of similar length, only then, has been emphasized. Before that, factorisation is acknowledged as difficult, and it is only interested in special probable prime numbers like the Fermat numbers and the Mersenne prime candidates.

On the other hand, the theory of integral quaternions and octonions are developed before the 1930s by Lipchitz and Hurwitz [9]. Modern usage on quaternions only centres on rotation manipulation. This may well explain and support the lack of usage of quaternions on the integer factorisation problem.

3.2 Quaternions

We begin by defining integral quaternions [10], so as to draw parallels between them and Gaussian integers.

Quaternions has three instead of one imaginary component:

$$\mathbb{H} = \{A = a + bi + cj + dk : a, b, c, d \in \mathbb{R}\}$$

Multiplication of quaternions is non-commutative, with the following properties:

$$i^2 = j^2 = k^2 = ijk = -1$$

More specifically:

$$ij = -ji = k$$

$$jk = -kj = i$$

$$ki = -ik = j$$

The multiplication of two quaternions is thus:

$$\begin{aligned} (a_1 + b_1i + c_1j + d_1k)(a_2 + b_2i + c_2j + d_2k) = \\ (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) \\ + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)i \\ + (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)j \\ + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)k \end{aligned}$$

The conjugate of a quaternion has its imaginary components negated:

$$\bar{P} = a - bi - cj - dk$$

Our focus is on quaternions with integral coefficients, called a Lipchitz quaternion:

$$\mathbb{H}_{\mathbb{Z}} = \{P_x = a + bi + cj + dk : a, b, c, d \in \mathbb{Z}\}$$

The concept of the norm is applicable to quaternions:

$$Nm(P_x) = a^2 + b^2 + c^2 + d^2$$

3.3 Sum of Four Squares

The norm of the product of the quaternions is equal to the product of the norm of the quaternions:

$$\begin{aligned} Nm(PQ) = PQ\bar{P}\bar{Q} = PQ\bar{Q}\bar{P} = P \cdot Nm(Q) \cdot \bar{P} \\ = P\bar{P} \cdot Nm(Q) = Nm(P) \cdot Nm(Q) \end{aligned}$$

The result is the Euler's four-square identity [11]:

$$\begin{aligned} pq &= Nm(P) \cdot Nm(Q) \\ &= (a_1^2 + b_1^2 + c_1^2 + d_1^2)(a_2^2 + b_2^2 + c_2^2 + d_2^2) \\ &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2)^2 \\ &\quad + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)^2 \\ &\quad + (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)^2 \\ &\quad + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)^2 \\ &= Nm(PQ) = n \end{aligned}$$

After knowing this set of information, the author feels that the algebra of quaternions can be applied into the problem of integer factorisation.

In this paper, we denote a quaternion by an upper-case letter, and the norm of the quaternion the corresponding lower case. So quaternion P_x has a norm of p . Different subscripts distinguishes each quaternion with the same norm.

The Jacobi's Theorem of Four Squares [7] specifies the number of distinct representations of a number as a sum of four squares, which is equivalent to the number of quaternions with the same norm - the number of ways to represent n as the sum of four

squares is eight times the sum of the divisors of n if n is odd, and 24 times the sum of the odd divisors of n if n is even. Therefore, there are $8(pq + p + q + 1)$ quaternions of norm pq , and $8(p + 1)$ quaternions of norm p .

3.4 Finding a Quaternion $P_x Q_y$

Finding any quaternion $P_x Q_y$ of norm $n = pq$ is equivalent to expressing the semiprime n as a sum of four squares, a result derived from the integer norm of any quaternion and its conjugate.

$$n = P_x Q_y \overline{P_x Q_y} = a^2 + b^2 + c^2 + d^2$$

The methods of finding a set of representation as a sum of four squares has been published and covered [12]. We will include an outline here.

The first two squares a^2 and b^2 is first chosen, which may be tailored according to specifications. Then the remainder $n - a^2 - b^2$ is checked on whether it is expressible as a sum of two squares with its complete factorisation [7]. There is an algorithm [13] which efficiently express each prime factor congruent to 1 mod 4 into a sum of two squares. After which the Brahmagupta–Fibonacci identity is used to obtain the full set of representation of $n - a^2 - b^2$ as a sum of two squares, and pq is now expressed as a sum of four squares in various ways.

4. Proposed Directions

The key purpose of this paper is to open up the concepts of sum of four squares and quaternions into the attempts of factoring semiprimes. This may be an interesting insight, but our applications were rather clumsy.

4.1 Direct Extension of the Method

Firstly, we will logically extend the method mentioned in Section 2.2. When the semiprime is factorised into quaternions:

$$pq = (PQ)(\overline{QP}) = (P\overline{Q})(Q\overline{P})$$

When one found a pair of quaternions PQ and $P\overline{Q}$, factorisation can be done, similarly with Euclidean

algorithm (explained in the following section).

However, there does not seem to be an easy method to find the set of quaternions PQ and $P\overline{Q}$. Moreover, it seems to be unnecessary restrictive to specify quaternion Q to be the conjugate of each other.

4.2 With a Slight Modification

For the Euclidean algorithm to generate the result, a common left or right divisor is already sufficient. Hence any pair of Q will suffice.

If we happen to find two quaternions and that has the same left ($P_x Q_y$ and $P_x Q_z$) or right ($Q_y P_x$ and $Q_z P_x$) divisor, factorisation can be done by taking the greatest (left/right) common divisor. For quaternions, multiplication is non-commutative. We have to note if it is a left or right divisor.

We show the resulting factorisation when a pair with a common left divisor P_x is found. Firstly, we introduce the division of a quaternion:

$$\begin{aligned} AB^{-1} &= A(\overline{BB^{-1}})B^{-1} \\ &= A\overline{B}/Nm(B) = a_d + b_d i + c_d j + d_d k \end{aligned}$$

One first finds a quotient λ that is integral, and the remainder is evaluated accordingly:

$$\lambda = [A\overline{B}/Nm(B)] = [a_d] + [b_d]i + [c_d]j + [d_d]k$$

We now show that the semiprime can be factored when a pair $P_x Q_y$ and $P_x Q_z$ is found:

$$P_x Q_y = (996 + 30i - 62j - 57k)$$

$$P_x Q_z = (962 + 10i - 247j - 116k)$$

Here is a Euclidean algorithm for quaternions [6]:

$$\begin{aligned} &(996 + 30i - 62j - 57k) = \\ &(962 + 10i - 247j - 116k)(1 + 0i + 0j + 0k) \\ &\quad + (34 + 20i + 185j + 59k) \\ &(962 + 10i - 247j - 116k) = \\ &(34 + 20i + 185j + 59k)(-1 + 0i - 5j - 1k) \\ &\quad + (12 - 80i + 88j + 77k) \\ &(34 + 20i + 185j + 59k) = \\ &(12 - 80i + 88j + 77k)(1 + 1i + 0j + 1k) \\ &\quad + (19 + 0i - 60j + 58k) \\ &(12 - 80i + 88j + 77k) = \\ &(19 + 0i - 60j + 58k)(0 + 1i + 1j + 1k) \\ &\quad + (10 + 19i + 11j - 2k) \end{aligned}$$

$$\begin{aligned}
 &(19 + 0i - 60j + 58k) = \\
 &(10 + 19i + 11j - 2k)(-1 - 2i + 1j + 3k) \\
 &\quad + (-4 + 4i - 6j - 15k) \\
 &(10 + 19i + 11j - 2k) = \\
 &(-4 + 4i - 6j - 15k)(0 - 1i + 1j + 0k) \\
 &\quad + (0 + 0i + 0j + 0k)
 \end{aligned}$$

The author has no remedy to find out the sum of two squares in the first place. To hit upon one pair at random, it requires an infeasible amount of computations in the order of $(pq)^{1/4}$; the above example was in fact generated with the factorisation of the semiprime in mind. Nevertheless, unless proven otherwise, there may be a much more efficient method to finding such a pair.

4.3 A More Likely Method

The reader may be perplexed on the need to find a pair of quaternions. The author highly doubts that the direct factorisation of a quaternion $P_x Q_y$ with a norm $n = pq$ will ever have an efficient method. Works focused on the factorisation of quaternions merely stated [9] and reiterated [14] that factorisations of quaternions are equivalent - they can be permuted from one form to another through a series of meta-commutation, recombination and unit-migration. Once again, they gave minimal or zero comment on its application on the integer factorisation problem.

However, here is a method that uses the same quaternion factors P_x and Q_y . The author believes that this approach may well have the greatest potential among this set of clumsily proposed directions.

The pair is $P_x Q_y$ and $Q_y P_x$. As multiplication of quaternions is non-commutative, they are usually different. However, they have the same real part (a subscript to the components is added to distinguish between the components from the different quaternionic factors):

$$a_p a_q - b_p b_q - c_p c_q - d_p d_q$$

The search for quaternions with norm pq should start with the largest real part first, thus there are few combinations to represent the smaller remaining part

as a sum of three squares. Elaboration on finding such as quaternions has been done in Section 3.4. Among the quaternions with the same real part, there are pairs of the quaternions $P_x Q_y$ and $Q_y P_x$. Just any one of them is sufficient for the factorisation of the semiprime.

We sketch a method to factor the semiprime when a pair of $P_x Q_y$ and $Q_y P_x$ is found:

$$\begin{aligned}
 P_x Q_y &= (996 + 30i - 62j - 57k) \\
 Q_y P_x &= (996 + 6i + 6j - 89k)
 \end{aligned}$$

Finding the pair of quaternions is equivalent to generating this set of 7 bilinear equations of 8 unknowns and the solutions produces P and Q :

$$\begin{aligned}
 a_p a_q - b_p b_q - c_p c_q - d_p d_q &= w \\
 a_p b_q + b_p a_q + c_p d_q - d_p c_q &= x_1 \\
 a_p b_q + b_p a_q - c_p d_q + d_p c_q &= x_2 \\
 a_p c_q - b_p d_q + c_p a_q + d_p b_q &= y_1 \\
 a_p c_q + b_p d_q + c_p a_q - d_p b_q &= y_2 \\
 a_p d_q + b_p c_q - c_p b_q + d_p a_q &= z_1 \\
 a_p d_q - b_p c_q + c_p b_q + d_p a_q &= z_2
 \end{aligned}$$

To the author's knowledge, there is no general solution to bilinear equations. Assuming the components are non-zero and distinct, the author could only use Gaussian elimination to reduce the set of equations into:

$$\begin{aligned}
 w &= a_p a_q - b_p b_q - c_p c_q - d_p d_q \\
 a_p &= \left(\frac{z_1 + z_2}{x_1 - x_2} \right) c_p + \left(-\frac{y_1 + y_2}{x_1 - x_2} \right) d_p \\
 a_q &= \left(-\frac{z_1 + z_2}{x_1 - x_2} \right) c_q + \left(\frac{y_1 + y_2}{x_1 - x_2} \right) d_q \\
 b_p &= \left(-\frac{y_1 - y_2}{x_1 - x_2} \right) c_p + \left(-\frac{z_1 - z_2}{x_1 - x_2} \right) d_p \\
 b_q &= \left(-\frac{y_1 - y_2}{x_1 - x_2} \right) c_q + \left(-\frac{z_1 - z_2}{x_1 - x_2} \right) d_q
 \end{aligned}$$

Solving the reduced set of solution will still require a certain amount of brute force. The author would appreciate if someone could come out with a much more efficient solution to the set of bilinear equations, and also for a general one that take into account of zeroes in the components of the pair $P_x Q_y$ and $Q_y P_x$ and its quaternionic factors.

5. Discussion

5.1 Possible Immediate Implications

Semiprimes that are just slightly larger a perfect square may be vulnerable to our method with $P_x Q_y$ and $Q_y P_x$. After subtracting the largest perfect square a^2 from the semiprime, one is left with a small number $n - a^2$. The list of representation as a sum of three squares will be short. Permuting the order and sign of the imaginary components in this short list one may easily generate a quaternion pair $P_x Q_y$ and $Q_y P_x$.

Factorisation is then obtained from solving the set of bilinear equations, which however has yet to have an efficient method. Moreover, one needs to check whether the pair picked are indeed the quaternion pair $P_x Q_y$ and $Q_y P_x$, and there is yet to have a shortcut in checking them. If these concerns are addressed, then our work will truly be of a substantial result, adding to the list of criteria that semiprime pq need to satisfy.

The main idea of our work is that if one can express the semiprime into a suitable pair of representations as a sum of four squares, then factorisation can be achieved.

5.2 Open Questions

There are some questions that one would like to consider. Firstly, it is on whether these methods are worth undertaking - considering the lack of an efficient algorithm to determine pairs of suitable quaternions that will factorise the corresponding semiprime. Unless a proof that show that our approach will be fruitless, this leaves possibilities open.

Secondly, we would like to ask whether it is possible to efficiently derive a quaternion $Q_y P_x$ given the corresponding quaternion $P_x Q_y$. It may be an interesting question by itself.

5.3 Possible Future Work

The author is a recent high school graduate, and as of yet may not have a sufficiently rigorous mathematical background to contribute substantially to these already

extensively-covered topics mentioned in the paper, and further opportunities may yield significant insights to the application of quaternions to semiprime factorisation. The use of eight-dimensional octonions is also a possibility, if justifiable.

6. Conclusions

The main motivation of our research is explore new approaches to tackle the integer factorisation problem, and we are pleased to have managed to apply some theorems of quaternions on this problem. Our application of quaternions in this paper is brief and perhaps trivial in nature. Yet, to our knowledge, we believe that our approach of generalising the Euler factorisation method is novel and the result is accountable - if one is able to find a suitable pair of representations of the semiprime as a sum of four squares, the semiprime can then be factorised easily.

The author calls for anyone interested in our application of quaternions to factorisation to perform a more rigorous analysis of our methods, or devise more creative approaches, thereby pushing for greater cryptographic standards in view of faster processing speeds over time. Ultimately, the integer factorisation problem is one of the most notorious, yet most elementary unsolved problems in number theory.

Acknowledgments

I would like to thank Peh Yu Xiang for exploring these mathematical ideas with me. I would also like to appreciate the support from several personnel from Anderson Junior College where I developed these ideas, notably Mr Poh Wei Leong who have supported by endeavour all along and Ms Goh Lay Hoon for her comments on the initial idea. Useful comments were solicited from Dr Toh Pee Choon.

References

- [1] Rivest, R. L., Shamir, A., and Adleman, L. 1978. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems." *Communications of the ACM* 21 (2): 120-6.

- [2] Kleinjung, T., Aoki, K., Franke, J., Lenstra, A., Thomé, E., Bos, J., and et al. 2010. "Factorization of a 768-Bit RSA Modulus." *Cryptology ePrint Archive: Report 006*.
- [3] Wolfram Research. "Euler's Factorization Method." MathWorld - A Wolfram Web Resource, Accessed June 5, 2015.
<http://mathworld.wolfram.com/EulersFactorizationMethod.html>.
- [4] Wolfram Research. "Gaussian Integer." MathWorld - A Wolfram Web Resource, Accessed June 5, 2015.
<http://mathworld.wolfram.com/GaussianInteger.html>.
- [5] Wolfram Research. "Gaussian Prime." MathWorld - A Wolfram Web Resource, Accessed June 5, 2015.
<http://mathworld.wolfram.com/GaussianPrime.html>.
- [6] Hardy, G. H., and Wright, E. M. 1938. *An Introduction to the Theory of Numbers*. Oxford: Clarendon Press.
- [7] Wolfram Research. "Sum of Squares Function." MathWorld - A Wolfram Web Resource, Accessed June 5, 2015.
<http://mathworld.wolfram.com/SumofSquaresFunction.html>.
- [8] McKee, J. 1996. "Turning Euler's Factoring Method Into a Factoring Algorithm." *Bulletin of the London Mathematical Society* 28: 351-5.
- [9] Conway, J. H., and Smith, D. A. 2003. *On Quaternions and Octonions*. A K Peters.
- [10] Wolfram Research. "Quaternion." MathWorld - A Wolfram Web Resource, Accessed June 5, 2015.
<http://mathworld.wolfram.com/Quaternion.html>.
- [11] Wolfram Research. "Euler Four-Square Identity." MathWorld - A Wolfram Web Resource, Accessed June 5, 2015.
<http://mathworld.wolfram.com/EulerFour-SquareIdentity.html>.
- [12] Moreno, C. J., and Wagstaff Jr., S. S. 2006. *Sum of Squares of Integers*. CRC Press.
- [13] Williams, K. S. 1995. "Some Refinements of an Algorithm of Brillhart." In: *Canadian Mathematical Society Conference Proceedings*, 15: 409-416.
- [14] Coan, B., and Perng, C. T. 2012. "Factorization of Hurwitz Quaternions." In: *International Mathematical Forum*, 7 (43): 2143-2156.