Introduction to Modern Algebra

Section 7.4. Integral Quaternions and the Four-Square Theorem—Proofs of Theorems





Table of contents

- Lemma 7.4.1
- 2 Lemma 7.4.2
- 3 Lemma 7.4.3. Lagrange Identity
- 4 Lemma 7.4.5. Left-Division Algorithm
- 5 Lemma 7.4.6
- 6 Lemma 7.4.7
- Theorem 7.4.1. Lagrange's Four-Square Theorem

Lemma 7.4.1. The adjoint in Q satisfies:

1.
$$x^{**} = x$$
,
2. $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$, and
3. $(xy)^* = y^* x^*$

for all $x, y \in Q$ and for all real δ and γ .

Proof. (1) If
$$x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$$
 then

$$x^{**} = (x^*)^* = (\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k)^* = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k,$$

Lemma 7.4.1. The adjoint in Q satisfies:

1.
$$x^{**} = x$$
,
2. $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$, and
3. $(xy)^* = y^* x^*$

for all $x, y \in Q$ and for all real δ and γ .

Proof. (1) If $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ then

$$x^{**} = (x^*)^* = (\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k)^* = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k,$$

as claimed.

(2) Let $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ and $y = \beta_0 + \beta_1 i \beta_2 j + \beta_3 k$ be in Q and let δ and γ be real numbers. Then

 $\delta x + \gamma y = (\delta \alpha_0 + \gamma \beta_0) + (\delta \alpha_1 + \gamma \beta_1)i + (\delta \alpha_2 + \gamma \beta_2)j + (\delta \alpha_3 + \gamma \beta_3)k, \dots$

Lemma 7.4.1. The adjoint in Q satisfies:

1.
$$x^{**} = x$$
,
2. $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$, and
3. $(xy)^* = y^* x^*$

for all $x, y \in Q$ and for all real δ and γ .

Proof. (1) If $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ then

$$x^{**} = (x^*)^* = (\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k)^* = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k,$$

as claimed.

(2) Let $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ and $y = \beta_0 + \beta_1 i \beta_2 j + \beta_3 k$ be in Q and let δ and γ be real numbers. Then

$$\delta x + \gamma y = (\delta \alpha_0 + \gamma \beta_0) + (\delta \alpha_1 + \gamma \beta_1)i + (\delta \alpha_2 + \gamma \beta_2)j + (\delta \alpha_3 + \gamma \beta_3)k, \dots$$

Lemma 7.4.1 (continued 1)

Lemma 7.4.1. $(xy)^* = y^*x^*$

Proof (continued). ... and so

$$(\delta x + \gamma y)^* = (\delta \alpha_0 + \gamma \beta_0) - (\delta \alpha_1 + \gamma \beta_1)i - (\delta \alpha_2 + \gamma \beta_2)j - (\delta \alpha_3 + \gamma \beta_3)k$$
$$= \delta(\alpha_0 - \alpha_1 j - \alpha_2 j - \alpha_2)k) + \gamma(\beta_0 - \beta_1 i - \beta_2 j - \beta_3 k) = \delta x^* + \gamma y^*,$$
as claimed

(3) We prove the result for the basis elements 1, i, j, k of Q (as a real vector space). This requires several cases. We have ii = k and ii = -k, so by (2) we have $(ij) = k^* = -k = ii = (-i)(-i) = i^*i^*$. We have ki = iand ik = -i, so by (2) we have $(ik)* = (-i)^* = i = ki = (-k)(-i) = k^*i^*$. We have ik = i and ki = -i, so by (2) we have $(ik)* = i^* = -i = ki = (-k)(-i) = k^*i^*$. Also, $(i^2)^* = (-1)^* = -1 = (-i)^2 = (i^*)^2, (i^2)^* = (-1)^* = -1 = (-i)^2$ $(i^*)^2$, and $(k^2)^* = (-1)^* = -1 = (-k)^2 = (k^*)^2$.

Lemma 7.4.1 (continued 1)

Lemma 7.4.1. $(xy)^* = y^*x^*$

Proof (continued). ... and so

$$(\delta x + \gamma y)^* = (\delta \alpha_0 + \gamma \beta_0) - (\delta \alpha_1 + \gamma \beta_1)i - (\delta \alpha_2 + \gamma \beta_2)j - (\delta \alpha_3 + \gamma \beta_3)k$$
$$= \delta(\alpha_0 - \alpha_1 j - \alpha_2 j - \alpha_2)k) + \gamma(\beta_0 - \beta_1 i - \beta_2 j - \beta_3 k) = \delta x^* + \gamma y^*,$$

as claimed.

(3) We prove the result for the basis elements 1, *i*, *j*, *k* of *Q* (as a real vector space). This requires several cases. We have ij = k and ji = -k, so by (2) we have $(ij)* = k^* = -k = ji = (-j)(-i) = j^*i^*$. We have ki = j and ik = -j, so by (2) we have $(ik)* = (-j)^* = j = ki = (-k)(-i) = k^*i^*$. We have jk = i and kj = -i, so by (2) we have $(jk)* = i^* = -i = kj = (-k)(-j) = k^*j^*$. Also, $(i^2)^* = (-1)^* = -1 = (-i)^2 = (i^*)^2$, $(j^2)^* = (-1)^* = -1 = (-j)^2 = (j^*)^2$, and $(k^2)^* = (-1)^* = -1 = (-k)^2 = (k^*)^2$.

Lemma 7.4.1 (continued 2)

Lemma 7.4.1. The adjoint in Q satisfies:

1.
$$x^{**} = x$$
,
2. $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$, and
3. $(xy)^* = y^* x^*$

for all $x, y \in Q$ and for all real δ and γ .

Proof (continued). Let
$$x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$$
 and
 $y = \beta_0 + \beta_1 i \beta_2 j + \beta_3 k$ be in Q . Then by (2)
 $(xy)^* = ((\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)(\beta_0 + \beta_1 i \beta_2 j + \beta_3 k))^*$
 $= ((\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)\beta_0 + (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)\beta_1 i + (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)\beta_2 j + (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)\beta_3 k)^*$
 $= (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* \beta_0 + (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* \beta_1 i^* + (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* \beta_2 j^* + (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* \beta_3 k^*$
 $= \beta_0 (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* + \beta_1 i^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* + \beta_2 j^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* + \beta_3 k^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^*$

Lemma 7.4.1 (continued 3)

Lemma 7.4.1. The adjoint in Q satisfies:

1.
$$x^{**} = x$$
,
2. $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$, and
3. $(xy)^* = y^* x^*$

for all $x, y \in Q$ and for all real δ and γ .

Proof (continued). ...

$$(xy)^* = \beta_0(\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* + \beta_1 i^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* + \beta_2 j^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* + \beta_3 k^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* = (\beta_0 + \beta_1 i + \beta_2 j + \beta_3 k)^* (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)^* = y^* x^*,$$

Lemma 7.4.2. For all $x, y \in Q$ we have N(xy) = N(x)N(y).

Proof. By the definition of norm, $N(xy) = (xy)(xy)^*$. By Lemma 7.4.1(3), $(xy)^* = y^*x^*$ and so (since norms are real and real numbers commute with all quaternions; that is, the reals are in the center of the quaternions)

$$N(xy) = (xy)(xy)^* = xy(y^*x^*) = x(yy^*)x^*$$
$$= xN(y)x^* = xx^*N(y) = N(x)N(y),$$

Lemma 7.4.2. For all $x, y \in Q$ we have N(xy) = N(x)N(y).

Proof. By the definition of norm, $N(xy) = (xy)(xy)^*$. By Lemma 7.4.1(3), $(xy)^* = y^*x^*$ and so (since norms are real and real numbers commute with all quaternions; that is, the reals are in the center of the quaternions)

$$N(xy) = (xy)(xy)^* = xy(y^*x^*) = x(yy^*)x^*$$

= $xN(y)x^* = xx^*N(y) = N(x)N(y),$

Lemma 7.4.3. Lagrange Identity

Lemma 7.4.3. Lagrange Identity. If $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and $\beta_0, \beta_1, \beta_2, \beta_3$ are real numbers then $(\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2)(\beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2) = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3)^2$ $+(\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_0\beta_2 - \alpha_1\beta_3 + \alpha_2\beta_0 + \alpha_3\beta_1)^2$ $+(\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_0)^2.$

Proof. With $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ and $y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$ in Q, we have $N(x) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2$ and $N(y) = \beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2$. So the left-hand side of the equation in the claim equals N(x)N(y). Also (see Quaternions–An Algebraic View (Supplement); the product is part of the definition of the quaternions):

$$xy = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3) + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)i$$

 $+(\alpha_0\beta_2+\alpha_2\beta_0+\alpha_3\beta_1-\alpha_1\beta_3)j+(\alpha_0\beta_3+\alpha_3\beta_0+\alpha_1\beta_2-\alpha_2\beta_1)k.$

Lemma 7.4.3. Lagrange Identity

Lemma 7.4.3. Lagrange Identity. If $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and $\beta_0, \beta_1, \beta_2, \beta_3$ are real numbers then $(\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2)(\beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2) = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3)^2$ $+(\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_0\beta_2 - \alpha_1\beta_3 + \alpha_2\beta_0 + \alpha_3\beta_1)^2$ $+(\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_0)^2.$

Proof. With $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ and $y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$ in Q, we have $N(x) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2$ and $N(y) = \beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2$. So the left-hand side of the equation in the claim equals N(x)N(y). Also (see Quaternions–An Algebraic View (Supplement); the product is part of the definition of the quaternions):

$$xy = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3) + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)i + (\alpha_0\beta_2 + \alpha_2\beta_0 + \alpha_3\beta_1 - \alpha_1\beta_3)j + (\alpha_0\beta_3 + \alpha_3\beta_0 + \alpha_1\beta_2 - \alpha_2\beta_1)k.$$

Lemma 7.4.3. Lagrange Identity (continued)

Lemma 7.4.3. Lagrange Identity. If $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and $\beta_0, \beta_1, \beta_2, \beta_3$ are real numbers then $(\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2)(\beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2) = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3)^2$ $+(\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_0\beta_2 - \alpha_1\beta_3 + \alpha_2\beta_0 + \alpha_3\beta_1)^2$ $+(\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_0)^2.$

Proof (continued). ...

 $xy = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3) + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)i + (\alpha_0\beta_2 + \alpha_2\beta_0 + \alpha_3\beta_1 - \alpha_1\beta_3)j + (\alpha_0\beta_3 + \alpha_3\beta_0 + \alpha_1\beta_2 - \alpha_2\beta_1)k.$

So the right-hand side of the equation in the claim equals N(xy). Since N(x)N(y) = N(xy) by Lemma 7.4.2, then we have Lagrange's Identity.

Lemma 7.4.5. Left-Division Algorithm

Lemma 7.4.5. Left-Division Algorithm.

Let $a, b \in H$ with $b \neq 0$. Then there exists two elements $c, d \in H$ such that a = cb + d and N(d) < N(b).

Proof. We prove the result in two steps. First, suppose $a \in H$ and let b > 0 be real (i.e., $b \in \mathbb{Z}$, b > 0). Let $a = t_0\zeta + t_1i + t_2j + t_3k$ where $t_0, t_1, t_2, t_3 \in \mathbb{Z}$ and b = n where n is a positive integer. Let $c = x_0\zeta + x_1i + x_1j + x_3k$ where $x_0, x_1, x_2, x_3 \in \mathbb{Z}$ (but are yet to be determined; we want them to satisfy the condition $N(d) = N(a - cb) = N(a - cn) < N(b) = N(n) = n^2$).

Lemma 7.4.5. Left-Division Algorithm

Lemma 7.4.5. Left-Division Algorithm.

Let $a, b \in H$ with $b \neq 0$. Then there exists two elements $c, d \in H$ such that a = cb + d and N(d) < N(b).

Proof. We prove the result in two steps. First, suppose $a \in H$ and let b > 0 be real (i.e., $b \in \mathbb{Z}$, b > 0). Let $a = t_0\zeta + t_1i + t_2j + t_3k$ where $t_0, t_1, t_2, t_3 \in \mathbb{Z}$ and b = n where n is a positive integer. Let $c = x_0\zeta + x_1i + x_1j + x_3k$ where $x_0, x_1, x_2, x_3 \in \mathbb{Z}$ (but are yet to be determined; we want them to satisfy the condition $N(d) = N(a - cb) = N(a - cn) < N(b) = N(n) = n^2$). Now

$$a - cn = \left(t_0\left(\frac{1 + i + j + k}{2}\right) + t_1i + t_2j + t_3k\right) \\ -nx_0\left(\frac{1 + i + j + k}{2}\right) - nx_1i - nx_2j - nx_3k$$



Lemma 7.4.5. Left-Division Algorithm

Lemma 7.4.5. Left-Division Algorithm.

Let $a, b \in H$ with $b \neq 0$. Then there exists two elements $c, d \in H$ such that a = cb + d and N(d) < N(b).

Proof. We prove the result in two steps. First, suppose $a \in H$ and let b > 0 be real (i.e., $b \in \mathbb{Z}$, b > 0). Let $a = t_0\zeta + t_1i + t_2j + t_3k$ where $t_0, t_1, t_2, t_3 \in \mathbb{Z}$ and b = n where n is a positive integer. Let $c = x_0\zeta + x_1i + x_1j + x_3k$ where $x_0, x_1, x_2, x_3 \in \mathbb{Z}$ (but are yet to be determined; we want them to satisfy the condition $N(d) = N(a - cb) = N(a - cn) < N(b) = N(n) = n^2$). Now

$$a - cn = \left(t_0 \left(\frac{1 + i + j + k}{2} \right) + t_1 i + t_2 j + t_3 k \right) \\ -nx_0 \left(\frac{1 + i + j + k}{2} \right) - nx_1 i - nx_2 j - nx_3 k$$

Lemma 7.4.5. Left-Division Algorithm (continued 1)

Proof (continued). ...

$$\begin{aligned} a - cn &= \frac{1}{2}(t_0 - nx_0) + \frac{1}{2}(t_0 + 2t_1 - n(t_0 + 2x_1))i \\ &+ \frac{1}{2}(t_0 + 2t_1 - n(t_0 + 2x_2))j + \frac{1}{2}(t_0 + 2t_3 - n(t_0 + 2x_3))k. \end{aligned}$$

We now seek to choose x_0, x_1, x_2, x_3 such that $|t_0 - nx_0| \le n/2$, $|t_0 + 2t_1 - n(t_0 + 2x_1)| \le n$, $t_0 + 2t_2 - n(t_0 + 2x_2)| \le n$, and $|t_0 + 2t_3 - n(t_0 + 2x_3)| \le n$ then we would have

$$N(a - cn) = \frac{(t_0 - nx_0)^2}{4} + \frac{(t_0 + 2t_1 - n(t_0 + 2x_1))^2}{4} + \frac{(t_0 + 2t_2 - n(t_0 + 2x_2))^2}{4} + \frac{(t_0 + 2t_3 - n(t_0 + 2x_3))^2}{4} \le n^2/16 + n^2/4 + n^2/4 + n^2/4 < n^2 = N(n),$$

as desired.

Lemma 7.4.5. Left-Division Algorithm (continued 1)

Proof (continued). ...

$$\begin{aligned} \mathbf{a} - \mathbf{c}\mathbf{n} &= \frac{1}{2}(t_0 - \mathbf{n}x_0) + \frac{1}{2}(t_0 + 2t_1 - \mathbf{n}(t_0 + 2x_1))\mathbf{i} \\ &+ \frac{1}{2}(t_0 + 2t_1 - \mathbf{n}(t_0 + 2x_2))\mathbf{j} + \frac{1}{2}(t_0 + 2t_3 - \mathbf{n}(t_0 + 2x_3))\mathbf{k}. \end{aligned}$$

We now seek to choose x_0, x_1, x_2, x_3 such that $|t_0 - nx_0| \le n/2$, $|t_0 + 2t_1 - n(t_0 + 2x_1)| \le n$, $t_0 + 2t_2 - n(t_0 + 2x_2)| \le n$, and $|t_0 + 2t_3 - n(t_0 + 2x_3)| \le n$ then we would have

$$\begin{split} \mathsf{N}(\mathsf{a}-\mathsf{cn}) &= \frac{(t_0-\mathsf{n} x_0)^2}{4} + \frac{(t_0+2t_1-\mathsf{n}(t_0+2x_1))^2}{4} \\ &+ \frac{(t_0+2t_2-\mathsf{n}(t_0+2x_2))^2}{4} + \frac{(t_0+2t_3-\mathsf{n}(t_0+2x_3))^2}{4} \\ &\leq \mathsf{n}^2/16 + \mathsf{n}^2/4 + \mathsf{n}^2/4 + \mathsf{n}^2/4 < \mathsf{n}^2 = \mathsf{N}(\mathsf{n}), \end{split}$$

as desired.

Lemma 7.4.5. Left-Division Algorithm (continued 2)

Proof (continued). The existence of desired x_0, x_1, x_2, x_3 are given as follows:

- 1. By the Division Algorithm in \mathbb{Z} , there is an integer x_0 such that $t_0 = x_0 n + r$ where $-n/2 \le r \le n/2$. For this x_0 , we have $|t_0 x_0 n| = |r| \le n/2$.
- 2. By the Division Algorithm in \mathbb{Z} , there is an integer k such that $t_0 + 2t_1 = kn + r$ and $0 \le r \le n$. If $k t_0$ is even, set $2x_1 = k t_0$ so that $t_0 + 2t_1 = (2x_0 + t_0)n + r$ and $|t_0 + 2t_1 (2x_1 + t_0)n| = r < n$. If $k t_0$ is odd, set $2x_1 = k t_0 + 1$ so that $t_0 + 2t_1 = (2x_1 + t_0 1)n + r = (2x_1 + t_0)n + r n$ and $|t_0 + 2t_1 (2x_1 + t_0)n| = |r n| \le n$ since $0 \le r < n$. There (regardless of the parity of $k t_0$) there is integer x_1 for which $|t_0 + 2t_1 (2x_1 + t_0)n| \le n$.

Lemma 7.4.5. Left-Division Algorithm (continued 2)

Proof (continued). The existence of desired x_0, x_1, x_2, x_3 are given as follows:

- 1. By the Division Algorithm in \mathbb{Z} , there is an integer x_0 such that $t_0 = x_0 n + r$ where $-n/2 \le r \le n/2$. For this x_0 , we have $|t_0 x_0 n| = |r| \le n/2$.
- 2. By the Division Algorithm in \mathbb{Z} , there is an integer k such that $t_0 + 2t_1 = kn + r$ and $0 \le r \le n$. If $k t_0$ is even, set $2x_1 = k t_0$ so that $t_0 + 2t_1 = (2x_0 + t_0)n + r$ and $|t_0 + 2t_1 (2x_1 + t_0)n| = r < n$. If $k t_0$ is odd, set $2x_1 = k t_0 + 1$ so that $t_0 + 2t_1 = (2x_1 + t_0 1)n + r = (2x_1 + t_0)n + r n$ and $|t_0 + 2t_1 (2x_1 + t_0)n| = |r n| \le n$ since $0 \le r < n$. There (regardless of the parity of $k t_0$) there is integer x_1 for which $|t_0 + 2t_1 (2x_1 + t_0)n| \le n$.
- 3. As in part 2, we can find integers x_2 and x_3 which satisfy $|t_0 + 2t_2 (2x_2 + t_0)n| \le n$ and $|t_0 + 2t_3 (2x_3 + t_0)n| \le n$.

Lemma 7.4.5. Left-Division Algorithm (continued 2)

Proof (continued). The existence of desired x_0, x_1, x_2, x_3 are given as follows:

- 1. By the Division Algorithm in \mathbb{Z} , there is an integer x_0 such that $t_0 = x_0 n + r$ where $-n/2 \le r \le n/2$. For this x_0 , we have $|t_0 x_0 n| = |r| \le n/2$.
- 2. By the Division Algorithm in \mathbb{Z} , there is an integer k such that $t_0 + 2t_1 = kn + r$ and $0 \le r \le n$. If $k t_0$ is even, set $2x_1 = k t_0$ so that $t_0 + 2t_1 = (2x_0 + t_0)n + r$ and $|t_0 + 2t_1 (2x_1 + t_0)n| = r < n$. If $k t_0$ is odd, set $2x_1 = k t_0 + 1$ so that $t_0 + 2t_1 = (2x_1 + t_0 1)n + r = (2x_1 + t_0)n + r n$ and $|t_0 + 2t_1 (2x_1 + t_0)n| = |r n| \le n$ since $0 \le r < n$. There (regardless of the parity of $k t_0$) there is integer x_1 for which $|t_0 + 2t_1 (2x_1 + t_0)n| \le n$.
- 3. As in part 2, we can find integers x_2 and x_3 which satisfy $|t_0 + 2t_2 (2x_2 + t_0)n| \le n$ and $|t_0 + 2t_3 (2x_3 + t_0)n| \le n$.

Lemma 7.4.5. Left-Division Algorithm (continued 3)

Lemma 7.4.5. Left-Division Algorithm.

Let $a, b \in H$ with $b \neq 0$. Then there exists two elements $c, d \in H$ such that a = cb + d and N(d) < N(b).

Proof (continued). So the claim holds for $a \in H$ and b > 0 real. We now consider the general case where $a, b \in H$ and $b \neq 0$. By Lemma 7.4.4 $n = bb^*$ is a positive integer, so by the first part of the proof there is $c \in H$ such that $ab^* = cn + d_1$ where $N(d_1) < N(n)$; that is $N(d_1) = N(ab^* - cn) < N(n)$. But $n = bb^*$ we have $N(ab^* - cbb^*) < N(n)$, or $N((a - cb)b^*) < N(n) = N(bb^*)$. By Lemma 7.4.2, this implies $N(a - cb)N(b^*) < N(b)N(b^*)$ or (since $b \neq 0$ and $N(b^*) > 0$) N(a - cb) < N(b). Set d = a - cb and we have a = cb + dwhere N(d) < N(b), so that the general case holds.

Lemma 7.4.5. Left-Division Algorithm (continued 3)

Lemma 7.4.5. Left-Division Algorithm.

Let $a, b \in H$ with $b \neq 0$. Then there exists two elements $c, d \in H$ such that a = cb + d and N(d) < N(b).

Proof (continued). So the claim holds for $a \in H$ and b > 0 real. We now consider the general case where $a, b \in H$ and $b \neq 0$. By Lemma 7.4.4 $n = bb^*$ is a positive integer, so by the first part of the proof there is $c \in H$ such that $ab^* = cn + d_1$ where $N(d_1) < N(n)$; that is $N(d_1) = N(ab^* - cn) < N(n)$. But $n = bb^*$ we have $N(ab^* - cbb^*) < N(n)$, or $N((a - cb)b^*) < N(n) = N(bb^*)$. By Lemma 7.4.2, this implies $N(a - cb)N(b^*) < N(b)N(b^*)$ or (since $b \neq 0$ and $N(b^*) > 0$) N(a - cb) < N(b). Set d = a - cb and we have a = cb + d where N(d) < N(b), so that the general case holds.

Lemma 7.4.6. Let *L* be a left-ideal of *H*. Then there exists an element $u \in L$ such that every element in *L* is a left-multiple of *u*; in other words, there exists $u \in L$ such that every $x \in L$ is of the form x = ru where $r \in H$.

Proof. If *L* is the trivial ideal, $L = \{0\}$, then we take u = 0. We now suppose that *L* has nonzero elements. By Lemma 7.4.4, the norms of nonzero elements are positive integers, so there is an element $u \neq 0$ in *L* whose norm is minimum over the nonzero elements of *L*.

Lemma 7.4.6. Let *L* be a left-ideal of *H*. Then there exists an element $u \in L$ such that every element in *L* is a left-multiple of *u*; in other words, there exists $u \in L$ such that every $x \in L$ is of the form x = ru where $r \in H$.

Proof. If *L* is the trivial ideal, $L = \{0\}$, then we take u = 0. We now suppose that *L* has nonzero elements. By Lemma 7.4.4, the norms of nonzero elements are positive integers, so there is an element $u \neq 0$ in *L* whose norm is minimum over the nonzero elements of *L*. For $x \in L$, by the Left-Division Algorithm (Lemma 7.4.5), x = cu + d where N(d) < N(u). Now d = x - cu where x and u are in *L* (and hence $cu \in L$ since it is a left-ideal), so $d \in L$. Since N(u) is the minimum positive norm of nonzero elements of *L*, then we must have N(d) = 0 and so d = 0. Therefore x = cu and (replacing $c \in H$ here with $r \in H$ in the statement of the lemma) the claim holds.

Lemma 7.4.6. Let *L* be a left-ideal of *H*. Then there exists an element $u \in L$ such that every element in *L* is a left-multiple of *u*; in other words, there exists $u \in L$ such that every $x \in L$ is of the form x = ru where $r \in H$.

Proof. If *L* is the trivial ideal, $L = \{0\}$, then we take u = 0. We now suppose that *L* has nonzero elements. By Lemma 7.4.4, the norms of nonzero elements are positive integers, so there is an element $u \neq 0$ in *L* whose norm is minimum over the nonzero elements of *L*. For $x \in L$, by the Left-Division Algorithm (Lemma 7.4.5), x = cu + d where N(d) < N(u). Now d = x - cu where x and u are in *L* (and hence $cu \in L$ since it is a left-ideal), so $d \in L$. Since N(u) is the minimum positive norm of nonzero elements of *L*, then we must have N(d) = 0 and so d = 0. Therefore x = cu and (replacing $c \in H$ here with $r \in H$ in the statement of the lemma) the claim holds.

Lemma 7.4.7. If $a \in H$ then $a^{-1} \in H$ if and only if N(a) = 1.

Proof. If both *a* and a^{-1} are in *H*, then by Lemma 7.4.4 both N(a) and $N(a^{-1})$ are positive integers. However, $aa^{-1} = 1$, so by Lemma 7.4.2 we have $N(a)N(a^{-1}) = N(aa^{-1}) = N(1) = 1$. But then N(a) = 1, as claimed.



Lemma 7.4.7. If $a \in H$ then $a^{-1} \in H$ if and only if N(a) = 1.

Proof. If both *a* and a^{-1} are in *H*, then by Lemma 7.4.4 both N(a) and $N(a^{-1})$ are positive integers. However, $aa^{-1} = 1$, so by Lemma 7.4.2 we have $N(a)N(a^{-1}) = N(aa^{-1}) = N(1) = 1$. But then N(a) = 1, as claimed.

If $a \in H$ and N(a) = 1, then $aa^* = N(a) = 1$ and so $a^{-1} = a^*$. By Lemma 7.4.4, since $a \in H$ then $a^* \in H$, so that $a^{-1} \in H$ as claimed.



Lemma 7.4.7. If $a \in H$ then $a^{-1} \in H$ if and only if N(a) = 1.

Proof. If both *a* and a^{-1} are in *H*, then by Lemma 7.4.4 both N(a) and $N(a^{-1})$ are positive integers. However, $aa^{-1} = 1$, so by Lemma 7.4.2 we have $N(a)N(a^{-1}) = N(aa^{-1}) = N(1) = 1$. But then N(a) = 1, as claimed. If $a \in H$ and N(a) = 1, then $aa^* = N(a) = 1$ and so $a^{-1} = a^*$. By Lemma 7.4.4, since $a \in H$ then $a^* \in H$, so that $a^{-1} \in H$ as claimed.

Theorem 7.4.1. Lagrange's Four-Square Theorem

Theorem 7.4.1. Lagrange's Four-Square Theorem. Every positive integer can be expressed as the sum of squares of four integers.

Proof. Let *n* be a positive integer. By the Fundamental Theorem of Arithmetic, *n* is a product of powers of prime numbers and by Lagrange's Identity (Lemma 7.4.3) a product of integers expressible as a sum of four squares is itself a sum of four squares. So it is sufficient to prove that every prime number is a sum of four squares. Of course prime number 2 equals $0^2 + 0^2 + 1^2 + 1^2$, so we only need to consider odd primes.

Theorem 7.4.1. Lagrange's Four-Square Theorem

Theorem 7.4.1. Lagrange's Four-Square Theorem. Every positive integer can be expressed as the sum of squares of four integers.

Proof. Let *n* be a positive integer. By the Fundamental Theorem of Arithmetic, *n* is a product of powers of prime numbers and by Lagrange's Identity (Lemma 7.4.3) a product of integers expressible as a sum of four squares is itself a sum of four squares. So it is sufficient to prove that every prime number is a sum of four squares. Of course prime number 2 equals $0^2 + 0^2 + 1^2 + 1^2$, so we only need to consider odd primes.

Let *p* be an odd prime. With \mathbb{Z}_p as the integers modulo *p*, consider the set of quaternions $W_p = \{\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \mid \alpha_0, \alpha_1, \alpha_3, \alpha_3 \in \mathbb{Z}_p\}$. The W_p is finite (in fact, $|W_p| = p^4$) and forms a ring. Since $p \neq 2$, the W_p is not commutative because $ij = -ji \neq ji$ (if p = 2 then, so to speak, "-1 = 1").

Theorem 7.4.1. Lagrange's Four-Square Theorem

Theorem 7.4.1. Lagrange's Four-Square Theorem. Every positive integer can be expressed as the sum of squares of four integers.

Proof. Let *n* be a positive integer. By the Fundamental Theorem of Arithmetic, *n* is a product of powers of prime numbers and by Lagrange's Identity (Lemma 7.4.3) a product of integers expressible as a sum of four squares is itself a sum of four squares. So it is sufficient to prove that every prime number is a sum of four squares. Of course prime number 2 equals $0^2 + 0^2 + 1^2 + 1^2$, so we only need to consider odd primes.

Let *p* be an odd prime. With \mathbb{Z}_p as the integers modulo *p*, consider the set of quaternions $W_p = \{\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \mid \alpha_0, \alpha_1, \alpha_3, \alpha_3 \in \mathbb{Z}_p\}$. The W_p is finite (in fact, $|W_p| = p^4$) and forms a ring. Since $p \neq 2$, the W_p is not commutative because $ij = -ji \neq ji$ (if p = 2 then, so to speak, "-1 = 1").

Theorem 7.4.1 (continued 1)

Proof (continued). Thus, by Wedderburn's Theorem W_p is not a division ring. By Lemma 7.4.A, W_p has a proper, nontrivial left-ideal. The two-sided ideal V in H defined as

$$V = \{x_0\zeta + x_1i + x_2j + x_2k \mid p \text{ divides all of } x_0, x_1, x_2, x_3\}$$

has the property that H/V is isomorphic to W_p by Note 7.4.A. If V were a maximal left-ideal in H, then $H/V \cong W_p$ would have no left ideals other the the trivial one and $H/V \cong W_p$ (remember, "bigger" ideals yield "smaller" quotient rings). Therefore there is some left ideal L of H such that $L \neq H$, $L \neq V$, and $L \supset V$. By Lemma 7.4.6, there is an element $u \in L$ such that every element in L is a left multiple of u. Since $p \in V$ then $p \in L$ and hence p = cu for some $c \in H$. If $u \in V$ then, since V is a two-sided ideal, every multiple of u would be in V and this cannot be the case since V is a proper subset of L and every element in L is a left multiple of u. So $u \notin V$.

Theorem 7.4.1 (continued 1)

Proof (continued). Thus, by Wedderburn's Theorem W_p is not a division ring. By Lemma 7.4.A, W_p has a proper, nontrivial left-ideal. The two-sided ideal V in H defined as

$$V = \{x_0\zeta + x_1i + x_2j + x_2k \mid p \text{ divides all of } x_0, x_1, x_2, x_3\}$$

has the property that H/V is isomorphic to W_p by Note 7.4.A. If V were a maximal left-ideal in H, then $H/V \cong W_p$ would have no left ideals other the the trivial one and $H/V \cong W_p$ (remember, "bigger" ideals yield "smaller" quotient rings). Therefore there is some left ideal L of H such that $L \neq H$, $L \neq V$, and $L \supset V$. By Lemma 7.4.6, there is an element $u \in L$ such that every element in L is a left multiple of u. Since $p \in V$ then $p \in L$ and hence p = cu for some $c \in H$. If $u \in V$ then, since V is a two-sided ideal, every multiple of u would be in V and this cannot be the case since V is a proper subset of L and every element in L is a left multiple of u. So $u \notin V$.

()

Theorem 7.4.1 (continued 2)

Proof (continued). Now *c* cannot have an inverse in *H*, or else $u = c^{-1}p$ would be in *V*. By Lemma 7.4.7, we now have that N(c) > 1. Next *u* cannot have an inverse in *H* or else the left-multiple of *u* by this inverse would imply that $1 \in L$ and, since *L* is a left ideal of *H*, we would have L = H in contradiction to the fact that $L \neq H$. Again by Lemma 7.4.7, we have N(u) > 1. Since off prime *p* satisfies p = cu, then $p^2 = N(p) = N(cu) = N(c)N(u)$. But N(c) and N(u) are integers (since $c, u \in H$) greater than 1, hence N(c) = N(u) = p.

Since $u \in H$, the $u = m_0\zeta + m_1i + m_2j + m_3k$ where m_0, m_1, m_2, m_3 are integers. Thus (by the definition of ζ):

$$2u = 2m_0\zeta + 2m_1i + 2m_2j + 2m_3k = (m_0 + m_0i + m_0j + m_0k)$$

 $+2m_1i+2m_2j+2m_3k = m_0+(2m_1+m_0)i+(2m_2+m_0)j+(2m_3+m_0)k.$

Therefore $N(2u) = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2$.

Theorem 7.4.1 (continued 2)

Proof (continued). Now *c* cannot have an inverse in *H*, or else $u = c^{-1}p$ would be in *V*. By Lemma 7.4.7, we now have that N(c) > 1. Next *u* cannot have an inverse in *H* or else the left-multiple of *u* by this inverse would imply that $1 \in L$ and, since *L* is a left ideal of *H*, we would have L = H in contradiction to the fact that $L \neq H$. Again by Lemma 7.4.7, we have N(u) > 1. Since off prime *p* satisfies p = cu, then $p^2 = N(p) = N(cu) = N(c)N(u)$. But N(c) and N(u) are integers (since $c, u \in H$) greater than 1, hence N(c) = N(u) = p.

Since $u \in H$, the $u = m_0\zeta + m_1i + m_2j + m_3k$ where m_0, m_1, m_2, m_3 are integers. Thus (by the definition of ζ):

$$2u = 2m_0\zeta + 2m_1i + 2m_2j + 2m_3k = (m_0 + m_0i + m_0j + m_0k)$$

 $+2m_1i + 2m_2j + 2m_3k = m_0 + (2m_1 + m_0)i + (2m_2 + m_0)j + (2m_3 + m_0)k.$ Therefore $N(2u) = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2.$

Theorem 7.4.1 (continued 3)

Proof (continued). But N(2u) = N(2)N(u) = 4p since $N(2) = 2^2 = 4$ and N(u) = p. We now have

$$4p = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2. \quad (*)$$

Next, notice that if $2a = x_0^2 + x_1^2 + x_2^2 + x_3^2$ where $a, x_0, x_1, x_2, x_3 \in \mathbb{Z}$ then all the x_i 's are even, all are odd, or two are even and two odd. In all three cases, the x_i 's can be paired in such a way that

$$y_0 = \frac{x_0 + x_1}{2}, \ y_1 = \frac{x_0 - x_1}{2}, \ y_0 = \frac{x_2 + x_3}{2}, \ \text{and} \ y_0 = \frac{x_2 - x_3}{2},$$

are all integers. Then

$$y_0^2 + y_1^2 + y_2^2 + y_3^2 = \left(\frac{x_0 + x_1}{2}\right)^2 + \left(\frac{x_0 - x_1}{2}\right)^2 + \left(\frac{x_2 + x_3}{2}\right)^2 + \left(\frac{x_2 - x_3}{2}\right)^2$$
$$= (x_0^2 + x_1^2 + x_2^2 + x_3^2)/2 = (2a)/2 = a.$$

That is, if 2a is a sum of four squares, then so is a.

()

Theorem 7.4.1 (continued 3)

Proof (continued). But N(2u) = N(2)N(u) = 4p since $N(2) = 2^2 = 4$ and N(u) = p. We now have

$$4p = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2. \quad (*)$$

Next, notice that if $2a = x_0^2 + x_1^2 + x_2^2 + x_3^2$ where $a, x_0, x_1, x_2, x_3 \in \mathbb{Z}$ then all the x_i 's are even, all are odd, or two are even and two odd. In all three cases, the x_i 's can be paired in such a way that

$$y_0 = \frac{x_0 + x_1}{2}, \ y_1 = \frac{x_0 - x_1}{2}, \ y_0 = \frac{x_2 + x_3}{2}, \ \text{and} \ y_0 = \frac{x_2 - x_3}{2},$$

are all integers. Then

$$y_0^2 + y_1^2 + y_2^2 + y_3^2 = \left(\frac{x_0 + x_1}{2}\right)^2 + \left(\frac{x_0 - x_1}{2}\right)^2 + \left(\frac{x_2 + x_3}{2}\right)^2 + \left(\frac{x_2 - x_3}{2}\right)^2$$
$$= (x_0^2 + x_1^2 + x_2^2 + x_3^2)/2 = (2a)/2 = a.$$

That is, if 2a is a sum of four squares, then so is a.

Theorem 7.4.1 (continued 4)

Theorem 7.4.1. Lagrange's Four-Square Theorem.

Every positive integer can be expressed as the sum of squares of four integers.

Proof (continued). Now 4*p* is a sum of four squares by (*), so the the previous comment we have that 2*p* is a sum of four squares and, again by the previous comment, *p* itself is a sum of four square. That is, odd prime *p* satisfies $p = a_1^2 + a_1^2 + a_2^2 + a_3^2$ for some integers a_0, a_1, a_2, a_3 . So Lagrange's Four-Square Theorem holds for all primes and, as commented at the start of the proof, holds for all positive integers.