LECTURE 14

CHAPTER 8 : THE FOUR SQUARE THM
Stillwell's Elements of Number Theory.

§8.1 Real matrices and $\mathbb{C}$

One method to construct $\mathbb{C} = \{a+bi \mid a,b \in \mathbb{R},\ i^2=-1\}$ is to use matrices in $\mathbb{R}^{2\times 2}$ of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$

$$M(a+bi) = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \left[\text{defines } M : \mathbb{C} \longrightarrow \mathbb{R}^{2\times 2}\right]$$

For example,

$$M(i) = M(0+1\cdot i) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$M(1) = M(1+0\cdot i) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$M(a+bi) = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = b\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = aM(1) + bM(i)$$

$$M(a+bi)M(c+di) = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}\begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \left[\begin{array}{c|c} ac-bd & ad+bc \\ -bc-ad & -bd+ac \end{array}\right]$$

$$= M[ac-bd+i(ad+bc)]$$

$$= M((a+bi)(c+di))$$

We may denote

$$M(1) = \mathbb{1} \quad \text{and} \quad M(i) = \mathbb{i}$$

$$M(a+bi) = a\mathbb{1} + b\mathbb{i}$$

Notice,

$$\text{norm}(a+bi) = a^2+b^2 = \det\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

So, multiplicativity is seen as $\det(AB) = \det A \det B$ consequence,

$$\text{norm}[(a+bi)(c+di)] = \det[M((a+bi)(c+di))]$$

$$= \det[M(a+bi)M(c+di)] \quad \text{exercise 8.1.1.}$$

$$= \det[M(a+bi)]\det[M(c+di)]$$

$$= \text{norm}(a+bi)\,\text{norm}(c+di).$$

§8.2 Complex matrices and $\mathbb{H}$

Defn/ Let $\alpha, \beta \in \mathbb{C}$ then $\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$ is a quaternion

Consider,

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ -\bar{\beta}_1 & \bar{\alpha}_1 \end{pmatrix}\begin{pmatrix} \alpha_2 & \beta_2 \\ -\bar{\beta}_2 & \bar{\alpha}_2 \end{pmatrix} = \left[\begin{array}{c|c} \alpha_1\alpha_2 - \beta_1\bar{\beta}_2 & \alpha_1\beta_2 + \beta_1\bar{\alpha}_2 \\ \hline -\bar{\beta}_1\alpha_2 - \bar{\alpha}_1\bar{\beta}_2 & -\bar{\beta}_1\beta_2 + \bar{\alpha}_1\bar{\alpha}_2 \end{array}\right] = \begin{bmatrix} \alpha_3 & \beta_3 \\ -\bar{\beta}_3 & \bar{\alpha}_3 \end{bmatrix}$$

$$\alpha_3 = \alpha_1\alpha_2 - \beta_1\bar{\beta}_2$$

$$\beta_3 = \alpha_1\beta_2 + \beta_1\bar{\alpha}_2$$

$$\boxed{\text{Defn/ } q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \text{ then define } \operatorname{norm}(q) = \det\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} = \alpha\bar{\alpha} + \beta\bar{\beta} \quad \operatorname{norm}(q) = |\alpha|^2 + |\beta|^2}$$

Since we use matrices to define quaternion algebra we have very natural properties (although, I would offer, the structure of quaternions transcends this representation or model for $\mathbb{H}$). Let $\mathbb{H}$ denote set of quaternions, if $q_1, q_2, q_3 \in \mathbb{H}$ then

$$q_1(q_2 + q_3) = q_1q_2 + q_1q_3$$

$$(q_1 + q_2)q_3 = q_1q_3 + q_2q_3$$

] distributive properties

$$q_1(q_2 q_3) = (q_1 q_2)q_3$$

] associative multiplication

$$\operatorname{norm}(q_1q_2) = \operatorname{norm}(q_1)\operatorname{norm}(q_2)$$

] multiplicative norm

However, generally, $q_1q_2 \neq q_2q_1$] non abelian

Remark: looking at $\mathbb{H}$ as 2×2 complex matrices may be worthwhile to remove some of the weirdness of $\mathbb{H}$. But in practice the notation is what is preferred for applications of $\mathbb{H}$. Similarly, we use $a+ib$ rather than $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ in our application & study of $\mathbb{C}$ most times.

A good notation for $\mathbb{H}$

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} a+di & b+ci \\ -b+ci & a-di \end{pmatrix}$$

$$= a\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mathbf{1}} + b\underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{i} + c\underbrace{\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}}_{j} + d\underbrace{\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}}_{k}$$

$$ii = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -1$$

$$jj = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -1$$

$$ij = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = k$$

$$ji = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = -k$$

We can derive by similar calculations, let's collect all for sake of discussion,

$$* \left\{\begin{aligned} i^2 = j^2 = k^2 &= -1 \\ ij = k &= -ji \\ jk = i &= -kj \\ ki = j &= -ik \end{aligned}\right\}$$

algebra for imaginary units i, j, k in $\mathbb{H}$

Remark: formally we have $\mathbb{H} = \{t + xi + yj + zk \mid t,x,y,z \in \mathbb{R}\}$ and we add, subtract and multiply as usual except we do not assume commutative and the units i, j, k satisfy $*$. I probably took this formal approach in Lecture to give you some contrast to §8.1 → 8.3 of Stillwell.

Quaternions a formal introduction

Generally $\alpha = t + xi + yj + zk$ and we denote $Re(\alpha) = t$ and $Im(\alpha) = xi + yj + zk$ thus $\alpha = Re(\alpha) + Im(\alpha)$. Define also the <u>conjugate</u> $\bar{\alpha} = Re(\alpha) - Im(\alpha)$. A quaternion α is <u>real</u> iff $Im(\alpha) = 0$; a quaternion α is <u>pure</u> <u>imaginary</u> iff $Re(\alpha) = 0$.

Remark: we assume,
$i^2 = j^2 = k^2 = -1$
$ij = -ji = k$
$jk = -kj = i$
$ki = -ki = j$

$$\text{norm}(\alpha) = \alpha\bar{\alpha} = t^2 + x^2 + y^2 + z^2$$

<u>Special Case</u>: product of pure imaginary quaternions

$$\alpha_1\alpha_2 = (x_1 i + y_1 j + z_1 k)(x_2 i + y_2 j + z_2 k) =$$

$$= \underline{x_1x_2 i^2} + x_1y_2\, ij + x_1z_2\, ik + y_1x_2\, ji + \underline{y_1y_2 j^2} + y_1z_2\, jk + z_1x_2\, ki + z_1y_2\, kj + \underline{z_1z_2 k^2}$$

$$= \underline{-x_1x_2 - y_1y_2 - z_1z_2} + (x_1y_2 - y_1x_2)k + (z_1x_2 - x_1z_2)j + (y_1z_2 - z_1y_2)i$$

$$= \underbrace{-\langle x_1, y_1, z_1\rangle \cdot \langle x_2, y_2, z_2\rangle}_{-\vec{\alpha}_1\cdot\vec{\alpha}_2} + \underbrace{\langle x_1, y_1, z_1\rangle \times \langle x_2, y_2, z_2\rangle}_{\vec{\alpha}_1\times\vec{\alpha}_2}$$

$$\boxed{-\vec{\alpha}_1 \cdot \vec{\alpha}_2 + \vec{\alpha}_1 \times \vec{\alpha}_2 = \alpha_1\alpha_2}$$

Multiplication in $\mathbb{H}$ encodes both <u>dot</u> and <u>cross</u> products of vectors. We used these for about 50 years before the modern vector notation supplanted $\mathbb{H}$.

Continuing from the special case calculation

$$\alpha_1 = t_1 + \vec{\alpha}_1 \quad \text{and} \quad \alpha_2 = t_2 + \vec{\alpha}_2$$

it occurs to me the special case is concisely denoted $\alpha_1 = \vec{\alpha}_1$. and $\alpha_2 = \vec{\alpha}_2$

$$\begin{aligned}
\alpha_1\alpha_2 &= (t_1 + \vec{\alpha}_1)(t_2 + \vec{\alpha}_2) \\
&= t_1\vec{\alpha}_2 + t_2\vec{\alpha}_1 + t_1 t_2 + \underline{\vec{\alpha}_1\vec{\alpha}_2} \\
&= \underline{t_1 t_2 + \vec{\alpha}_1\vec{\alpha}_2 + t_1\vec{\alpha}_2 + t_2\vec{\alpha}_1}\ *
\end{aligned}$$

we calculated this on 4

Notice if $\alpha_1 = \alpha_2 = \alpha = t + \vec{\alpha}$ we have,
$\bar{\alpha} = t - \vec{\alpha}$ and so,

$$\begin{aligned}
\alpha\bar{\alpha} &= (t + \vec{\alpha})(t - \vec{\alpha}) \\
&= t^2 + \cancel{t\vec{\alpha}} - \cancel{t\vec{\alpha}} - \vec{\alpha}\vec{\alpha} \\
&= t^2 - [-\vec{\alpha}\cdot\vec{\alpha} + \underbrace{\vec{\alpha}\times\vec{\alpha}}_{0}] \\
&= t^2 + \vec{\alpha}\cdot\vec{\alpha}
\end{aligned}$$

$$\therefore \boxed{\text{norm}(t + xi + yj + zk) = t^2 + x^2 + y^2 + z^2}$$

Another interesting calculation then follows,

$$\begin{aligned}
\text{norm}(q_1 q_2) &= q_1 q_2 \overline{q_1 q_2} \\
&= q_1 q_2 \bar{q}_2 \bar{q}_1 \\
&= q_1 |q_2|^2 \bar{q}_1 \quad : \text{note, } |q_2|^2 \in \mathbb{R}, \text{ factor out.} \\
&= |q_2|^2 q_1 \bar{q}_1 \\
&= |q_2|^2 |q_1|^2 \\
&= \underline{\text{norm}(q_1)\,\text{norm}(q_2)}.
\end{aligned}$$

Lemma. $\overline{q_1 q_2} = \bar{q}_2 \bar{q}_1$

~~Lemma~~: $\cancel{\overline{q_1 q_2} = \overline{t_1 t_2 + \vec{\alpha}_1\vec{\alpha}_2 + t_1\vec{\alpha}_2 + t_2\vec{\alpha}_1}}$
$= t_1 t_2 - \vec{\alpha}_1(-\vec{\alpha}_2)$

I leave the Lemma to you guys ☺

THE FOUR SQUARE IDENTITY:

$$\text{norm}(q_1)\,\text{norm}(q_2) = \text{norm}(q_1 q_2)$$

$$* \quad \underline{(a_1^2+b_1^2+c_1^2+d_1^2)(a_2^2+b_2^2+c_2^2+d_2^2)} =$$

$$= \left|(a_1+b_1 i+c_1 j+d_1 k)(a_2+b_2 i+c_2 j+d_2 k)\right|^2$$

$$\begin{aligned}
= \big| & (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) \\
& + (b_1a_2 + c_1d_2 - d_1c_2 + a_1b_2)\, i \\
& + (a_1c_2 + c_1a_2 + d_1b_2 - b_1d_2)\, j \\
& + (a_1d_2 + d_1a_2 + b_1c_2 - c_1b_2)\, k \,\big|^2
\end{aligned}$$

$$\left.\begin{aligned}
= & (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2)^2 + \\
& + (b_1a_2 + a_1b_2 + c_1d_2 - d_1c_2)^2 \\
& + (a_1c_2 + c_1a_2 + d_1b_2 - b_1d_2)^2 \\
& + (a_1d_2 + d_1a_2 + b_1c_2 - c_1b_2)^2
\end{aligned}\right\} **$$

The equality of * and ** is the 4-square identity. We see it is merely a consequence of the multiplicativity of the $\mathbb{H}$-norm.

<u>Remark</u>: Euler found in 1748 w/o help of $\mathbb{H}$. Euler wanted to prove all $n \in \mathbb{N}$ are expressed as sums of four squares... turns out Lagrange did it in 1770. We give proof in §8.4 → 8.8 which mere's our proof of two-square thm built off the structure of $\mathbb{Z}[i]$.