2.4 Normal form: Definitions and results

2.4.1 Invertible generator matrices

Let $t \in \mathbb{Z}_{>0}$ and let E_{ij} be the matrix in $M_{t\times t}(\mathbb{F})$ with 1 in the (i,j) entry and all other entries 0. Let

$$\mathbb{F}^\times = \{d \in \mathbb{F} \mid d \neq 0\}.$$

The *elementary diagonal matrices* are $h_i(d) = 1 + (d-1)E_{ii}$ for $i \in \{1, \ldots, t\}$ and $d \in \mathbb{F}^\times$.

$$h_i(d) = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& d &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \qquad \text{with} \qquad h_i(d)^{-1} = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& d^{-1} &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}$$

The *root matrices* are $x_{ij}(c) = 1 + cE_{ij}$ for $c \in \mathbb{F}$ and $i, j \in \{1, \ldots, t\}$ with $i < j$.

$$x_{ij}(c) = \begin{pmatrix} 1 &&&& \\ & \ddots && c & \\ && \ddots && \\ &&& \ddots & \\ &&&& 1 \end{pmatrix} \qquad \text{with} \qquad x_{ij}(c)^{-1} = \begin{pmatrix} 1 &&&& \\ & \ddots && -c & \\ && \ddots && \\ &&& \ddots & \\ &&&& 1 \end{pmatrix}$$

The *row reducers* are $s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{i,i+1}$ for $i \in \{1, \ldots, t-1\}$ and $c \in \mathbb{F}$.

$$s_i(c) = \begin{pmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 &&&&& \\ &&& c & 1 &&& \\ &&& 1 & 0 &&& \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{pmatrix} \qquad \text{with} \qquad s_i(c)^{-1} = \begin{pmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 &&&&& \\ &&& 0 & 1 &&& \\ &&& 1 & -c &&& \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{pmatrix}$$

2.4.2 Proof of the normal form theorem: The greedy reduction algorithm

Let $s, t \in \mathbb{Z}_{>0}$ and let $A \in M_{t\times s}(\mathbb{F})$ so that

$$A \text{ is a } t \times s \text{ matrix with entries in } \mathbb{F}.$$

The proof of the following theorem gives an explicit algorithm for writing A as a product of row reducers $s_i(c)$, diagonal generators $h_i(d)$ and upper triangular elementary matrices $x_{ij}(a_{ij})$. This procedure is no different than the usual row reduction procedure: namely, a way of writing an invertible matrix g in a 'normal form' as a product of elementary matrices by the 'row reduction' algorithm. To follow the proof it is helpful to be looking at examples (see Examples 2.4.3 and 2.4.4 and the examples in section 3.2).

Theorem 2.13. *Let $s, t \in \mathbb{Z}_{>0}$. Let E_{ij} be the $t \times s$ matrix with 1 in the (i,j) entry and 0 elsewhere. Let $A \in M_{t\times s}(\mathbb{F})$. Then there exist*

$$r \in \{1, \ldots, \min(s,t)\}, \qquad P \in GL_t(\mathbb{F}), \qquad Q \in GL_s(\mathbb{F})$$

$$\text{such that} \qquad A = P1_rQ, \qquad \text{where} \quad 1_r = E_{11} + \cdots + E_{rr}.$$

Proof. **Step 1.** (For an example of this step see Example 2.4.3)
Let j_1 be minimal such that column j_1 of A has a nonzero entry.
Let i_1 be maximal such that $A(i_1, j_1) \neq 0$. Let

$$A^{(1)} = s_1\left(\tfrac{A(1,j_1)}{A(i_1,j_1)}\right)^{-1} s_2\left(\tfrac{A(2,j_1)}{A(i_1,j_1)}\right)^{-1} \cdots s_{i_1-1}\left(\tfrac{A(i_1-1,j_1)}{A(i_1,j_1)}\right)^{-1} A.$$

Let j_2 be minimal such that column j_2 of $A^{(1)}$ has a nonzero entry below row 1.
Let $i_2 > 1$ be maximal such that $A^{(1)}(i_2, j_2) \neq 0$. Let

$$A^{(2)} = s_2\left(\tfrac{A^{(1)}(2,j_2)}{A^{(1)}(i_2,j_2)}\right)^{-1} s_3\left(\tfrac{A^{(1)}(3,j_2)}{A^{(1)}(i_2,j_2)}\right)^{-1} \cdots s_{i_2-1}\left(\tfrac{A^{(1)}(i_2-1,j_2)}{A^{(1)}(i_2,j_2)}\right)^{-1} A^{(1)}.$$

Let j_3 be minimal such that column j_3 of $A^{(2)}$ has a nonzero entry below row 2.
Let $i_3 > 2$ be maximal such that $A^{(2)}(i_3, j_3) \neq 0$. Let

$$A^{(3)} = s_3\left(\tfrac{A^{(2)}(3,j_3)}{A^{(2)}(i_3,j_3)}\right)^{-1} s_4\left(\tfrac{A^{(2)}(4,j_3)}{A^{(2)}(i_3,j_3)}\right)^{-1} \cdots s_{i_3-1}\left(\tfrac{A^{(2)}(i_2-1,j_2)}{A^{(2)}(i_3,j_3)}\right)^{-1} A^{(2)}.$$

Continue this process until it happens that there does not exist j_{r+1} such that column j_{r+1} of $A^{(r)}$ has a nonzero entry below row r. Then $A^{(r)}$ has the property that

the first nonzero entry in row $j+1$ is to the right of the first nonzero entry in row j.

and

$$\begin{aligned} A = &\left(s_{i_1-1}\left(\tfrac{A(i_1-1,j_11)}{A(i_1,j_1)}\right) \cdots s_2\left(\tfrac{A(2,j_1)}{A(i_1,j_1)}\right) s_1\left(\tfrac{A(1,j_1)}{A(i_1,j_1)}\right)\right) \\ &\cdot \left(s_{i_2-1}\left(\tfrac{A^{(1)}(i_2-1,j_2)}{A^{(1)}(i_2,j_2)}\right) \cdots s_3\left(\tfrac{A^{(1)}(3,j_2)}{A^{(1)}(i_2,j_2)}\right) s_2\left(\tfrac{A^{(1)}(2,j_2)}{A^{(1)}(i_2,j_2)}\right)\right) \\ &\cdots \left(s_{i_r-1}\left(\tfrac{A^{(1)}(j_r-1,j_r)}{A^{(1)}(i_r,j_r)}\right) \cdots s_{r+1}\left(\tfrac{A^{(1)}(r+1,j_r)}{A^{(1)}(i_r,j_r)}\right) s_r\left(\tfrac{A^{(1)}(r,j_r)}{A^{(1)}(i_r,j_r)}\right)\right) \cdot A^{(r)} \end{aligned} \tag{Afact}$$

Then

$$\begin{aligned} A^{(r)} = &\left(h_1(A^{(r)}(1,j_1)) \cdots h_r(A^{(r)}(r,j_r))\right) \\ &\cdot \left(x_{r-1,j_r}\left(\frac{A^{(r)}(r-1,j_r)}{A^{(r)}(r-1,j_{r-1})}\right) \cdots x_{1,j_r}\left(\frac{A^{(r)}(1,j_r)}{A^{(r)}(1,j_1)}\right)\right) \\ &\cdots \left(x_{2,j_3}\left(\frac{A^{(r)}(2,j_3)}{A^{(r)}(2,j_2)}\right) x_{1,j_3}\left(\frac{A^{(r)}(1,j_3)}{A^{(r)}(1,j_1)}\right)\right) \cdot x_{1,j_2}\left(\frac{A^{(r)}(1,j_2)}{A^{(r)}(1,j_1)}\right) \cdot R, \end{aligned}$$

where R is given by

$$R(k,j) = \begin{cases} \frac{A^{(r)}(k,j)}{A^{(r)}(k,j_k)}, & \text{if } k \in \{1,\ldots,r\} \text{ and } j \in \{j_k, j_k+1, \ldots, s\} \\ & \text{and } j \notin \{j_{k+1}, \ldots, j_r\}, \\ 0, & \text{otherwise.} \end{cases}$$

Step 2. (For example of this step see Example 2.4.4.)
Let $c_{r+1} < \cdots < c_{s-1} < c_s$ be such that $\{j_1, \ldots, j_r, c_{r+1}, \ldots, c_s\} = \{1, \ldots, s\}$. Then

$$R = 1_r \cdot Q, \quad \text{where} \quad Q \in GL_s(\mathbb{F})$$

is given by

$$\begin{aligned} Q = &\Big(x_{r,s}\Big(\frac{A^{(r)}(r,c_s)}{A^{(r)}(r,j_r)}\Big) \cdots x_{1,s}\Big(\frac{A^{(r)}(1,c_s)}{A^{(r)}1,j_1}\Big)\Big) \\ &\cdot \Big(x_{r,r+1}\Big(\frac{A^{(r)}(r,c_{r+1})}{A^{(r)}(r,j_r)}\Big) \cdots x_{1,r+1}\Big(\frac{A^{(r)}(1,c_{r+1})}{A^{(r)}(1,j_1)}\Big)\Big) \\ &\cdot (s_r \cdots s_{j_r-1}) \cdots (s_2 \cdots s_{j_2-1}) \cdot (s_1 \cdots s_{j_1-1}). \end{aligned} \tag{Qform}$$

Summary. In summary, $A = P 1_r Q$ where $P \in GL_t(\mathbb{F})$ and $Q \in GL_s(\mathbb{F})$ are given by

$$\begin{aligned} P = &\Big(s_{i_1-1}\Big(\tfrac{A(i_1-1,j_11)}{A(i_1,j_1)}\Big) \cdots s_2\Big(\tfrac{A(2,j_1)}{A(i_1,j_1)}\Big) s_1\Big(\tfrac{A(1,j_1)}{A(i_1,j_1)}\Big)\Big) \\ &\cdot \Big(s_{i_2-1}\Big(\tfrac{A^{(1)}(i_2-1,j_2)}{A^{(1)}(i_2,j_2)}\Big) \cdots s_3\Big(\tfrac{A^{(1)}(3,j_2)}{A^{(1)}(i_2,j_2)}\Big) s_2\Big(\tfrac{A^{(1)}(2,j_2)}{A^{(1)}(i_2,j_2)}\Big)\Big) \\ &\cdots \Big(s_{i_r-1}\Big(\tfrac{A^{(1)}(j_r-1,j_r)}{A^{(1)}(i_r,j_r)}\Big) \cdots s_{r+1}\Big(\tfrac{A^{(1)}(r+1,j_r)}{A^{(1)}(i_r,j_r)}\Big) s_r\Big(\tfrac{A^{(1)}(r,j_r)}{A^{(1)}(i_r,j_r)}\Big)\Big) \\ &\cdot \Big(h_1(A^{(r)}(1,j_1)) \cdots h_r(A^{(r)}(r,j_r))\Big) \\ &\cdot \Big(x_{r-1,j_r}\Big(\frac{A^{(r)}(r-1,j_r)}{A^{(r)}(r-1,j_{r-1})}\Big) \cdots x_{1,j_r}\Big(\frac{A^{(r)}(1,j_r)}{A^{(r)}(1,j_1)}\Big)\Big) \\ &\cdots \Big(x_{2,j_3}\Big(\frac{A^{(r)}(2,j_3)}{A^{(r)}(2,j_2)}\Big) x_{1,j_3}\Big(\frac{A^{(r)}(1,j_3)}{A^{(r)}(1,j_1)}\Big)\Big) \cdot x_{1,j_2}\Big(\frac{A^{(r)}(1,j_2)}{A^{(r)}(1,j_1)}\Big) \end{aligned}$$

and Q is as in (Qform). □

2.4.3 A greedy reduction example

This is an example of Step 1 of the proof of Theorem 2.13. Let

$$s_1(c) = \begin{pmatrix} c & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad s_2(c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad s_3(c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

and let

$$A = \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ -9 & -108 & -\frac{233}{2} & -114 & \frac{293}{2} & -167 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix}$$

Then

$$\begin{aligned}
A &= \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ -9 & -108 & -\frac{233}{2} & -114 & \frac{293}{2} & -167 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \\
&= s_2(-3) \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{2} & 12 & \frac{23}{2} & -23 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \\
&= s_2(-3)s_1(-1) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & \frac{1}{2} & 12 & \frac{23}{2} & -23 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \\
&= s_2(-3)s_1(-1)s_3(2) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \end{pmatrix} \\
&= s_2(-3)s_1(-1)s_3(2)s_2(0) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & -1 & -36 \end{pmatrix} \\
&= s_2(-3)s_1(-1)s_3(2)s_2(0)s_3(1) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

In this example,

$$A^{(1)} = \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & \frac{1}{2} & 12 & \frac{23}{2} & -23 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \qquad \text{and} \qquad A^{(1)} = s_1(-1)^{-1}s_2(-3)^{-1}A.$$

Then

$$A^{(2)} = \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & -1 & -36 \end{pmatrix} \qquad \text{and} \qquad A^{(2)} = s_2(0)^{-1}s_3(2)^{-1}A^{(1)}.$$

Then

$$A^{(3)} = \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad A^{(3)} = s_3(1)^{-1}A^{(2)}.$$

2.4.4 An example of the factorization a row echelon matrix

This is an example of Step 2 of the proof of Theorem 2.13. Let

$$h_1(d) = \begin{pmatrix} d & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad h_2(d) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad h_3(d) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and let $x_{ij}(c)$ denote the square matrix with 1s on the diagonal, c in the (i,j) entry and 0 elsewhere.

$$\text{If} \qquad A^{(3)} = \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

then

$$\begin{aligned} A^{(3)} &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \begin{pmatrix} 1 & 12 & 13 & 14 & 15 & 16 \\ 0 & 0 & 1 & 24 & 25 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \cdot x_{23}(25)x_{13}(15)x_{12}(13) \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \cdot x_{23}(25)x_{13}(15)x_{12}(13) \cdot R \end{aligned}$$

where

$$R = \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= 1_3 \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = 1_3 Q.$$

So

$$\begin{aligned} A^{(3)} &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \cdot x_{23}(25)x_{13}(15)x_{12}(13) \cdot R \\ &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \cdot x_{23}(25)x_{13}(15)x_{12}(13) \cdot 1_3 \cdot Q, \end{aligned}$$

where

$$Q = \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 12 & 14 & 16 \\ 0 & 1 & 0 & 0 & 24 & 26 \\ 0 & 0 & 1 & 0 & 0 & 36 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} s_3s_4s_2 = \begin{matrix} x_{36}(36)x_{26}(26)x_{16}(16) \\ \cdot x_{25}(24)x_{15}(14)x_{24}(12) \\ \cdot s_3s_4s_2. \end{matrix}$$

2.4.5 A normal form example

This is an example of the output of the proof of Theorem 2.13. Combining examples 2.4.3 and 2.4.4,

$$\text{if} \qquad A = \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ -9 & -108 & -\frac{233}{2} & -114 & \frac{293}{2} & -167 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \qquad \text{then} \qquad A = P1_3Q,$$

where

$$P = \begin{matrix} s_2(-3)s_1(-1)s_3(2)s_2(0)s_3(1) \\ \cdot h_1(3)h_2(\frac{1}{4})h_3(-1)\cdot x_{23}(25)x_{13}(15)x_{12}(13) \end{matrix} \quad \text{in } GL_4(\mathbb{F})$$

and

$$Q = \begin{matrix} x_{36}(36)x_{26}(26)x_{16}(16) \\ \cdot x_{25}(24)x_{15}(14)x_{24}(12) \\ \cdot s_3s_4s_2. \end{matrix} \quad \text{in } GL_6(\mathbb{F}).$$