Rings and modules - Problem set 6 solutions

Solved on Tuesday 21.11

Problem 1. (After Chapter 14.2.) Let R be a domain (i.e. a commutative unital integral domain) and M an R-module. Show that the set

$$\operatorname{Tor} M = \{m \in M \mid \text{there exists nonzero } r \in R \text{ with } rm = 0\}$$

is a submodule of M.

Solution. Since $1 \cdot 0 = 0$ we have that $0 \in \operatorname{Tor} M$ and so $\operatorname{Tor} M \neq \varnothing$. Let $m, n \in \operatorname{Tor} M$ and $r \in R$. Since $m, n \in \operatorname{Tor} M$, there exist nonzero $r_1, r_2 \in R$ such that $r_1 m = 0$ and $r_2 n = 0$. Since R is an integral domain, we have $r_1 r_2 \neq 0$. Then, since R is commutative, we have

$$r_1 r_2 (m - n) = (r_1 r_2) m - (r_1 r_2) n = r_2 (r_1 m) - r_1 (r_2 n) = r_2 0 - r_1 0 = 0,$$

and so $m - n \in \operatorname{Tor} M$, and we also have that

$$r_1 (rm) = r(r_1 m) = r0 = 0,$$

and so $rm \in \operatorname{Tor} M$. This shows that $\operatorname{Tor} M$ is a submodule of M.

Problem 2. (After Chapter 21.4.) Let K be a field and $A \in M_{n \times n}(K)$ be an $n \times n$ matrix with coefficients in K. Show that there exists a unique monic polynomial $m_A(X) \in K[X]$ which generates the ideal $I_A = \{p(X) \in K[X] \mid p(A) = 0\}$.

Solution. The set $\{I, A, A^2, \ldots, A^{n^2}\}$ has $n^2 + 1$ elements. Since the dimension of $M_{n \times n}(K)$ is n^2, it follows that the above set is linearly dependent and so there exist $\lambda_0, \ldots, \lambda_{n^2} \in K$, not all equal to zero, such that

$$\lambda_0 I + \lambda_1 A + \cdots + \lambda_{n^2} A^{n^2} = 0.$$

Setting $p(X) = \lambda_0 + \lambda_1 X + \cdots + \lambda_{n^2} X^{n^2} \in K[X]$ we obtain that $p(A) = 0$ and so I_A is not the zero ideal. Since $K[X]$ is a PID, we have that I_A can be generated by a single nonzero element $q(X) \in K[X]$. Let $q(X) = a_0 + a_1 X + \cdots + a_k X^k$ and set $m_A(X) = a_n^{-1} q(X)$. Then $m_A(X)$ is monic and

$$(m_A(X)) = (a_n^{-1} q(X)) = (q(X)) = I_A.$$

Assume now that $b(X)$ is another monic polynomial such that $(b(X)) = I_A$. Then $b(X) \in I_A = (m_A(X))$ implies that $b(X) = m_A(X) r(X)$ for some $r(X) \in K[X]$. Similarly, $m_A(X) \in I_A = (b(X))$ implies that $m_A(X) = b(X) s(X)$. Then

$$m_A(X) = b(X) s(X) = m_A(X) r(X) s(X)$$

implies that $r(X) s(X)$ is a nonzero constant polynomial, since the degree of the two polynomials $m_A(X)$ and $m_A(X) r(X) s(X)$ must be the same. Moreover, since the coefficient of the highest degree term of $m_A(X) = m_A(X) r(X) s(X)$ is 1, it follows that $r(X) s(X) = 1$. Hence $r(X) = r \in K$. Then $b(X) = m_A(x) r$ implies that $r = 1$ since both $b(X)$ and $m_A(X)$ are monic. We conclude that $m_A(X) = b(X)$, showing uniqueness of $m_A(X)$.

Problem 3. (After Chapter 20.3.) (Exam August 2013, Problem 1.) Find the Smith normal form over the integers $\mathbb{Z}$ for the matrix $\left(\begin{smallmatrix} 4 & 2 & 4 \\ 12 & -6 & 6 \\ -16 & 10 & -4 \end{smallmatrix}\right)$.

Solution. We observe that the greatest common divisor of all the elements in the matrix is 2. Hence we proceed by obtaining a 2 at the top left corner of the matrix and then obtaining zeros in its row and column:

$$\begin{pmatrix} 4 & 2 & 4 \\ 12 & -6 & 6 \\ -16 & 10 & -4 \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} 2 & 4 & 4 \\ -6 & 12 & 6 \\ 10 & -16 & -4 \end{pmatrix} \xrightarrow[C_3 \to C_3 - 2C_1]{C_2 \to C_2 - 2C_1} \begin{pmatrix} 2 & 0 & 0 \\ -6 & 24 & 18 \\ 10 & -36 & -24 \end{pmatrix} \xrightarrow[R_3 \to R_3 - 5R_1]{R_2 \to R_2 + 3R_1} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 24 & 18 \\ 0 & -36 & -24 \end{pmatrix}.$$

Now we observe that the greatest common divisor of all the elements in the 2×2 matrix block $\left(\begin{smallmatrix} 24 & 18 \\ -36 & -24 \end{smallmatrix}\right)$ is 6. Hence we proceed by obtaining a 6 at the top left corner of this matrix block and then obtaining zeros in its row and column:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 24 & 18 \\ 0 & -36 & -24 \end{pmatrix} \xrightarrow{C_2 \to C_2 - C_3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 18 \\ 0 & -12 & -24 \end{pmatrix} \xrightarrow{C_3 \to C_3 - 3C_2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & -12 & 12 \end{pmatrix} \xrightarrow{R_3 \to R_3 + 2R_2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix},$$

which is the Smith normal form of A since $2 \mid 6 \mid 12$.

Problem 4. (After Chapter 21.4.) (Exercise 21.4.1 in the book.) Find rational canonical forms of the following matrices over $\mathbb{Q}$:

(a) $\left(\begin{smallmatrix} 1 & 5 & 7 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{smallmatrix}\right)$.

(b) $\left(\begin{smallmatrix} 2 & 4 & 0 \\ 1 & 4 & 1 \\ 3 & 8 & 3 \end{smallmatrix}\right)$.

(c) $\left(\begin{smallmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right)$.

Solution.

(a) Let the matrix be A. We start with the Smith normal form of $A - XI_3$:

$$\begin{pmatrix} 1-X & 5 & 7 \\ 0 & 4-X & 3 \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} 5 & 1-X & 7 \\ 4-X & 0 & 3 \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow{C_1 \to \frac{1}{5}C_1} \begin{pmatrix} 1 & 1-X & 7 \\ \frac{4-X}{5} & 0 & 3 \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow[C_3 \to C_3 - 7C_1]{C_2 \to C_2 - (1-X)C_1}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ \frac{4-X}{5} & \frac{(4-X)(X-1)}{5} & \frac{-13+7X}{5} \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow{R_2 \to R_2 - \frac{4-X}{5}R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{(4-X)(X-1)}{5} & \frac{-13+7X}{5} \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow{R_2 \to 5R_2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & (4-X)(X-1) & -13+7X \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow{R_2 \to R_2 + 7R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (4-X)(X-1) & -6 \\ 0 & 0 & 1-X \end{pmatrix} \xrightarrow{C_2 \leftrightarrow C_3}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -6 & (4-X)(X-1) \\ 0 & 1-X & 0 \end{pmatrix} \xrightarrow{C_2 \to -\frac{1}{6}C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & (4-X)(X-1) \\ 0 & \frac{X-1}{6} & 0 \end{pmatrix} \xrightarrow{C_3 \to C_3 - (4-X)(X-1)C_2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{X-1}{6} & \frac{(X-1)^2(X-4)}{6} \end{pmatrix} \xrightarrow{R_3 \to R_3 - \frac{X-1}{6}R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{(X-1)^2(X-4)}{6} \end{pmatrix} \xrightarrow{R_3 \to 6R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (X-1)^2(X-4) \end{pmatrix}.$$

Hence the invariant factors of $A - XI_3$ are 1, 1 and $(X-1)^2(X-4) = -4 + 9X - 6X^2 + X^3$. The only non-unit is $-4 + 9X - 6X^2 + X^3$ and so the rational canonical form of A is

$$C_{-4+9X-6X^2+X^3} = \begin{pmatrix} 0 & 0 & 4 \\ 1 & 0 & -9 \\ 0 & 1 & 6 \end{pmatrix}.$$

(b) Let the matrix be B. We start with the Smith normal form of $B - XI_3$:

$$\begin{pmatrix} 2-X & 4 & 0 \\ 1 & 4-X & 1 \\ 3 & 8 & 3-X \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & 4-X & 1 \\ 2-X & 4 & 0 \\ 3 & 8 & 3-X \end{pmatrix} \xrightarrow[C_3 \to C_3 - C_1]{C_2 \to C_2 - (4-X)C_1} \begin{pmatrix} 1 & 0 & 0 \\ 2-X & -4+6X-X^2 & X-2 \\ 3 & -4+3X & -X \end{pmatrix}$$

$$\xrightarrow[R_3 \to R_3 - 3R_1]{R_2 \to R_2 - (2-X)R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4+6X-X^2 & X-2 \\ 0 & -4+3X & -X \end{pmatrix} \xrightarrow{R_2 \to R_2 + R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -8+9X-X^2 & -2 \\ 0 & -4+3X & -X \end{pmatrix} \xrightarrow{C_2 \leftrightarrow C_3}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & -8+9X-X^2 \\ 0 & -X & -4+3X \end{pmatrix} \xrightarrow{C_2 \to -\frac{1}{2}C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -8+9X-X^2 \\ 0 & \frac{X}{2} & -4+3X \end{pmatrix} \xrightarrow{R_3 \to R_3 - \frac{X}{2}R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -8+9X-X^2 \\ 0 & 0 & -4+7X-\frac{9}{2}X^2+\frac{X^3}{2} \end{pmatrix}$$

$$\xrightarrow{C_3 \to C_3 - (-8+9X-X^2)C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4+7X-\frac{9}{2}X^2+\frac{X^3}{2} \end{pmatrix} \xrightarrow{C_3 \to \frac{1}{2}C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -8+14X-9X^2+X^3 \end{pmatrix}.$$

Hence the invariant factors of $B - XI_3$ are 1, 1 and $-8+14X-9X^2+X^3$. The only non-unit is $-8+14X-9X^2+X^3$ and so the rational canonical form of B is

$$C_{-8+14X-9X^2+X^3} = \begin{pmatrix} 0 & 0 & 8 \\ 1 & 0 & -14 \\ 0 & 1 & 9 \end{pmatrix}.$$

(c) Let the matrix be C. We start with the Smith normal form of $C - XI_4$:

$$\begin{pmatrix} 1-X & 1 & -2 & 4 \\ 0 & 1-X & 2 & 2 \\ 0 & 0 & 1-X & 3 \\ 0 & 0 & 0 & -X \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} 1 & 1-X & -2 & 4 \\ 1-X & 0 & 2 & 2 \\ 0 & 0 & 1-X & 3 \\ 0 & 0 & 0 & -X \end{pmatrix} \xrightarrow[C_4 \to C_4 - 4C_1]{\substack{C_2 \to C_2 - (1-X)C_1 \\ C_3 \to C_3 + 2C_1}}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1-X & -(1-X)^2 & 4-2X & -2+4X \\ 0 & 0 & 1-X & 3 \\ 0 & 0 & 0 & -X \end{pmatrix} \xrightarrow{R_2 \to R_2 - (1-X)R_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -(1-X)^2 & 4-2X & -2+4X \\ 0 & 0 & 1-X & 3 \\ 0 & 0 & 0 & -X \end{pmatrix} \xrightarrow{C_2 \leftrightarrow C_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -2+4X & 4-2X & -(1-X)^2 \\ 0 & 3 & 1-X & 0 \\ 0 & -X & 0 & 0 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 1-X & 0 \\ 0 & -2+4X & 4-2X & -(1-X)^2 \\ 0 & -X & 0 & 0 \end{pmatrix} \xrightarrow{R_2 \to \frac{1}{3}R_2}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{1-X}{3} & 0 \\ 0 & -2+4X & 4-2X & -(1-X)^2 \\ 0 & -X & 0 & 0 \end{pmatrix} \xrightarrow{C_2 \to C_2 - \frac{1-X}{3}C_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2+4X & \frac{14-12X+4X^2}{3} & -(1-X)^2 \\ 0 & -X & \frac{X-X^2}{3} & 0 \end{pmatrix} \xrightarrow[R_4 \to R_4 + XR_1]{R_3 \to R_3 - (-2+4X)R_2}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{14-12X+4X^2}{3} & -(1-X)^2 \\ 0 & 0 & \frac{X-X^2}{3} & 0 \end{pmatrix} \xrightarrow[C_4 \to -C_4]{C_3 \to 3C_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 14-12X+4X^2 & (1-X)^2 \\ 0 & 0 & X-X^2 & 0 \end{pmatrix} \xrightarrow{R_3 \to R_3 + 4R_4}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 14-8X & 1-2X+X^2 \\ 0 & 0 & X-X^2 & 0 \end{pmatrix} \xrightarrow{C_4 \to 8C_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 14-8X & 8-16X+8X^2 \\ 0 & 0 & X-X^2 & 0 \end{pmatrix} \xrightarrow{C_4 \to C_4 + XC_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 14-8X & 8-2X \\ 0 & 0 & X-X^2 & X^2-X^3 \end{pmatrix} \xrightarrow{C_3 \to C_3 - 4C_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -18 & 8-2X \\ 0 & 0 & X-5X^2+4X^3 & X^2-X^3 \end{pmatrix} \xrightarrow{R_3 \to \frac{-1}{18}R_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{X-4}{9} \\ 0 & 0 & X-5X^2+4X^3 & X^2-X^3 \end{pmatrix} \xrightarrow{R_4 \to R_4 - (X-5X^2+4X^3)R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{X-4}{9} \\ 0 & 0 & 0 & \frac{4X-12X^2+12X^3-4X^4}{9} \end{pmatrix} \xrightarrow{C_4 \to C_4 - \frac{X-4}{9}C_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{4X-12X^2+12X^3-4X^4}{9} \end{pmatrix} \xrightarrow{C_4 \to -\frac{9}{4}C_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -X+3X^2-3X^3+X^4 \end{pmatrix}$$

Hence the invariant factors of $C - XI_3$ are 1, 1, 1 and $-X + 3X^2 - 3X^3 + X^4$. The only non-unit is $-X + 3X^2 - 3X^3 + X^4$ and so the rational canonical form of C is

$$C_{-X+3X^2-3X^3+X^4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

Problem 5. (After Chapter 21.5) (Exercise 21.5.1 in the book.) Find the Jordan canonical form of the following matrices:

(a) $\begin{pmatrix} 0 & 4 & 2 \\ -3 & 8 & 3 \\ 4 & -8 & -2 \end{pmatrix}$.

(b) $\begin{pmatrix} 5 & \frac{1}{2} & -2 & 4 \\ 0 & 5 & 4 & 4 \\ 0 & 0 & 5 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$.

Solution.

(a) Let the matrix be A. We start with the Smith normal form of $A - XI_3$:

$$\begin{pmatrix} -X & 4 & 2 \\ -3 & 8-X & 3 \\ 4 & -8 & -2-X \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} 4 & -X & 2 \\ 8-X & -3 & 3 \\ -8 & 4 & -2-X \end{pmatrix} \xrightarrow{C_1 \to \frac{1}{4}C_1} \begin{pmatrix} 1 & -X & 2 \\ 2-\frac{1}{4}X & -3 & 3 \\ -2 & 4 & -2-X \end{pmatrix} \xrightarrow[C_3 \to C_3 - 2C_1]{C_2 \to C_2 + XC_1}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2-\frac{1}{4}X & -3+2X-\frac{1}{4}X^2 & -1+\frac{1}{2}X \\ -2 & 4-2X & 2-X \end{pmatrix} \xrightarrow[R_3 \to R_3 + 2R_1]{R_2 \to R_2 - (2-\frac{1}{4}X)R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3+2X-\frac{1}{4}X^2 & -1+\frac{1}{2}X \\ 0 & 4-2X & 2-X \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4-2X & 2-X \\ 0 & -3+2X-\frac{1}{4}X^2 & -1+\frac{1}{2}X \end{pmatrix} \xrightarrow{C_2 \leftrightarrow C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2-X & 4-2X \\ 0 & -1+\frac{1}{2}X & -3+2X-\frac{1}{4}X^2 \end{pmatrix} \xrightarrow{C_3 \to C_3 - 2C_2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2-X & 0 \\ 0 & -1+\frac{1}{2}X & -1+X-\frac{1}{4}X^2 \end{pmatrix} \xrightarrow{R_3 \to R_3 - \frac{1}{2}R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2-X & 0 \\ 0 & 0 & -1+X-\frac{1}{4}X^2 \end{pmatrix} \xrightarrow[R_3 \to -4R_3]{R_2 \to -R_2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & X-2 & 0 \\ 0 & 0 & 4-4X+X^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-2 & 0 \\ 0 & 0 & (X-2)^2 \end{pmatrix}.$$

Hence the elementary divisors are $X-2$ and $(X-2)^2$. It follows that the Jordan canonical form of A is

$$\left(\begin{array}{c|cc} J_{1,X-2} & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{$J_{2,X-2}$}} \\ 0 & & \end{array}\right) = \left(\begin{array}{c|cc} 2 & 0 & 0 \\ \hline 0 & 2 & 0 \\ 0 & 1 & 2 \end{array}\right).$$

(b) Let the matrix be B. We start with the Smith normal form of $B - XI_4$:

$$\begin{pmatrix} 5-X & \frac{1}{2} & -2 & 4 \\ 0 & 5-X & 4 & 4 \\ 0 & 0 & 5-X & 3 \\ 0 & 0 & 0 & 4-X \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} \frac{1}{2} & 5-X & -2 & 4 \\ 5-X & 0 & 4 & 4 \\ 0 & 0 & 5-X & 3 \\ 0 & 0 & 0 & 4-X \end{pmatrix} \xrightarrow{C_1 \to 2C_1}$$

$$\begin{pmatrix} 1 & 5-X & -2 & 4 \\ 10-2X & 0 & 4 & 4 \\ 0 & 0 & 5-X & 3 \\ 0 & 0 & 0 & 4-X \end{pmatrix} \xrightarrow{R_2 \to R_2 - (10-2X)R_1} \begin{pmatrix} 1 & 5-X & -2 & 4 \\ 0 & -2(X-5)^2 & 4(6-X) & 4(2X-9) \\ 0 & 0 & 5-X & 3 \\ 0 & 0 & 0 & 4-X \end{pmatrix}$$

$$\xrightarrow[\substack{C_3 \to C_3 + 2C_1 \\ C_4 \to C_4 - 4C_1}]{C_2 \to C_2 - (5-X)C_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -2(X-5)^2 & 4(6-X) & 4(2X-9) \\ 0 & 0 & 5-X & 3 \\ 0 & 0 & 0 & 4-X \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 5-X & 3 \\ 0 & -2(X-5)^2 & 4(6-X) & 4(2X-9) \\ 0 & 0 & 0 & 4-X \end{pmatrix} \xrightarrow{C_2 \leftrightarrow C_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 5-X & 0 \\ 0 & 4(2X-9) & 4(6-X) & -2(X-5)^2 \\ 0 & 4-X & 0 & 0 \end{pmatrix} \xrightarrow{R_2 \to \frac{1}{3}R_2}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{3}(5-X) & 0 \\ 0 & 4(2X-9) & 4(6-X) & -2(X-5)^2 \\ 0 & 4-X & 0 & 0 \end{pmatrix} \xrightarrow{C_3 \to C_3 - \frac{1}{3}(5-X)C_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4(2X-9) & \frac{4}{3}(2X^2-22X+63) & -2(X-5)^2 \\ 0 & 4-X & -\frac{1}{3}(5-X)(4-X) & 0 \end{pmatrix}$$

$$\xrightarrow[R_4 \to R_4 - (4-X)R_2]{R_3 \to R_3 - 4(2X-9)R_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{4}{3}(2X^2-22X+63) & -2(X-5)^2 \\ 0 & 0 & -\frac{1}{3}(5-X)(4-X) & 0 \end{pmatrix}$$

Let us now focus on the 2×2 matrix on the lower right hand-side. The polynomial $2(X-5)^2$ has only 5 as a root. By evaluating $2X^2 - 22X + 63$ on 5 we see that 5 is not a root and hence $(X-5)$ does not divide this polynomial. We conclude that the greatest common divisor of the two polynomials in row 3 is 1. Hence by theory we know that we can obtain a 1 in position $(3,3)$ in this matrix. Moreover, after using this 1 to eliminate the polynomials in positions $(3,4)$ and $(4,3)$, we obtain that the polynomial in the position $(4,4)$ has to be the characteristic polynomial of B by Theorem 19.7. We may compute the characteristic polynomial of B immediately to be

$$c_B(X) = \det(B - XI_4) = \begin{vmatrix} 5-X & \frac{1}{2} & -2 & 4 \\ 0 & 5-X & 4 & 4 \\ 0 & 0 & 5-X & 3 \\ 0 & 0 & 0 & 4-X \end{vmatrix} = (4-X)(5-X)^3 = (X-4)(X-5)^3.$$

We conclude that the Smith normal form of $B - XI_4$ is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & (X-4)(X-5)^3 \end{pmatrix}$$

Otherwise, we could also proceed from where we stopped our computation of the Smith normal form as follows

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{4}{3}(2X^2-22X+63) & -2(X-5)^2 \\ 0 & 0 & -\frac{1}{3}(5-X)(4-X) & 0 \end{pmatrix} \xrightarrow[C_4 \to -\frac{1}{2}C_4]{C_3 \to -3C_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -8X^2+88X-252 & (X-5)^2 \\ 0 & 0 & X^2-9X+20 & 0 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 + 8R_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 16X-92 & (X-5)^2 \\ 0 & 0 & X^2-9X+20 & 0 \end{pmatrix} \xrightarrow{R_3 \to \frac{1}{16}R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X-\frac{23}{4} & \frac{1}{16}(X-5)^2 \\ 0 & 0 & X^2-9X+20 & 0 \end{pmatrix}$$

$$\xrightarrow{R_4 \to R_4 - XR_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X-\frac{23}{4} & \frac{1}{16}(X-5)^2 \\ 0 & 0 & \frac{1}{4}(80-13X) & -\frac{1}{16}X(X-5)^2 \end{pmatrix} \xrightarrow{C_4 \to 16C_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X-\frac{23}{4} & (X-5)^2 \\ 0 & 0 & \frac{1}{4}(80-13X) & -X(X-5)^2 \end{pmatrix}$$

$$\xrightarrow{R_4 \to 4R_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X-\frac{23}{4} & (X-5)^2 \\ 0 & 0 & 80-13X & -4X(X-5)^2 \end{pmatrix} \xrightarrow{R_4 \to R_4 + 13R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X-\frac{23}{4} & (X-5)^2 \\ 0 & 0 & \frac{21}{4} & (X-5)^2(13-4X) \end{pmatrix}$$

$$\xrightarrow{R_3 \leftrightarrow R_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{21}{4} & (X-5)^2(13-4X) \\ 0 & 0 & X-\frac{23}{4} & (X-5)^2 \end{pmatrix} \xrightarrow{R_3 \to \frac{4}{21}R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{4}{21}(X-5)^2(13-4X) \\ 0 & 0 & X-\frac{23}{4} & (X-5)^2 \end{pmatrix}$$

$$\xrightarrow{R_4 \to R_4 - (X-\frac{23}{4})R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{4}{21}(X-5)^2(4X-13) \\ 0 & 0 & 0 & \frac{16}{21}(X-5)^3(X-4) \end{pmatrix} \xrightarrow{C_4 \to C_4 + \frac{4}{21}(X-5)^2(4X-13)C_3}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{16}{21}(X-4)(X-5)^3 \end{pmatrix} \xrightarrow{R_4 \to \frac{21}{16}R_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & (X-4)(X-5)^3 \end{pmatrix}$$

and we obtain the same matrix as before. Hence the elementary divisors are $(X-4)$ and $(X-5)^3$. It follows that the Jordan canonical form of B is

$$\left(\begin{array}{c|ccc} J_{1,X-4} & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & J_{3,X-5} & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} 4 & 0 & 0 & 0 \\ \hline 0 & 5 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 5 \end{array}\right).$$

Problem 6. (After Chapter 21.4.) (Exam August 2013, Problem 2.) Let A be a 4×4 matrix over $\mathbb{R}$ with minimal polynomial $m_A(X) = (X-3)^2$. What are the possible forms that the rational canonical form of A can have?

Solution. Let $p_1(X), p_2(X), p_3(X), p_4(X)$ be the monic invariant factors of $A - XI_4$. Then

$$p_1(X) \mid p_2(X) \mid p_3(X) \mid p_4(X) \tag{1}$$

Then we know that $p_4(X) = m_A(X) = (X-3)^2$ by Theorem 19.7(4). We also now that the characteristic polynomial $c_A(X)$ of A satisfies $c_A(X) = p_1(X)p_2(X)p_3(X)p_4(X)$ by Theorem 19.7(3) and that it is of degree 4 by Theorem 19.7(1). Since $p_4(X) = (X-3)^2$, we conclude that $p_3(X)$ is of degree 1 or 2 (it cannot be constant since otherwise $p_1(X)$ and $p_2(X)$ must also be constant by (1)). Since $p_3(X) \mid p_4(X) = (X-3)^2$, we conclude that $p_3(X) = (X-3)^2$ or $p_3(X) = (X-3)$. If $p_3(X) = (X-3)^2$, then $p_2(X)$ and $p_1(X)$ are units because the degree of $c_A(X)$ is 4. Therefore, in this case we obtain that the non-unit invariant factors of $A - XI_4$ are $(X-3)^2$, $(X-3)^2$ and so the rational canonical form of A is

$$\left(\begin{array}{c|c} C_{(X-3)^2} & 0 \\ \hline 0 & C_{(X-3)^2} \end{array}\right) = \begin{pmatrix} 0 & -9 & 0 & 0 \\ 1 & 6 & 0 & 0 \\ 0 & 0 & 0 & -9 \\ 0 & 0 & 1 & 6 \end{pmatrix}.$$

If $p_3(X) = (X-3)$, then $p_2(X)$ must be of degree 1 so that the degree of $c_A(X)$ is 4. Since $p_2(X) \mid p_3(X) = (X-3)$ and both are monic, we conclude that $p_2(X) = (X-3)$. Thus in this case the non-unit invariant factors of $A - XI_4$ are $(X-3)$, $(X-3)$ and $(X-3)^2$ and so the rational canonical form of A is

$$\left(\begin{array}{c|c|cc} C_{X-3} & 0 & 0 & 0 \\ \hline 0 & C_{X-3} & 0 & 0 \\ \hline 0 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{$C_{(X-3)^2}$}} \\ 0 & 0 & & \end{array}\right) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -9 \\ 0 & 0 & 1 & 6 \end{pmatrix}.$$

Problem 7. (After Chapter 21.4.) (Exam December 2013, Problem 1.)

(a) Find the Smith normal form of the matrix $\begin{pmatrix} 2-X & 1 & 2 \\ 0 & 1-X & 2 \\ 1 & 0 & 1-X \end{pmatrix}$ over $\mathbb{Z}_3[X]$.

(b) Find the rational canonical form of the matrix $A = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ over $\mathbb{Z}_3$.

(c) Let $M_{3\times 3}(\mathbb{Z}_3)$ be the 3×3 matrix ring over $\mathbb{Z}_3$ and define $\Phi_A : \mathbb{Z}_3[X] \to M_{3\times 3}(\mathbb{Z}_3)$ by letting $\Phi_A(P(X)) = P(A)$ for each polynomial $P(X)$ in $\mathbb{Z}_3[X]$. The image of Φ_A is then the subring of $M_{3\times 3}(\mathbb{Z}_3)$ generated by the matrix A. Prove that this subring is a field.

Solution.

(a) We have

$$\begin{pmatrix} 2-X & 1 & 2 \\ 0 & 1-X & 2 \\ 1 & 0 & 1-X \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 1 & 0 & 1-X \\ 0 & 1-X & 2 \\ 2-X & 1 & 2 \end{pmatrix} \xrightarrow{C_3 \to C_3-(1-X)C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-X & 2 \\ 2-X & 1 & -X^2 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3-(2-X)R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-X & 2 \\ 0 & 1 & -X^2 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -X^2 \\ 0 & 1-X & 2 \end{pmatrix} \xrightarrow{C_3 \to C_3+X^2C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1-X & 2+X^2-X^3 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3-(1-X)R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2+X^2-X^3 \end{pmatrix} \xrightarrow{R_3 \to -R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2-X^2+X^3 \end{pmatrix} = S,$$

and S is the Smith normal form of A (note that $3 = 0$ since we work over $\mathbb{Z}_3[X]$).

(b) By part (a) we have that the only non-unit monic invariant factor of A is $-2 - X^2 + X^3$. Hence the rational canonical form of A is the companion matrix $C_{-2-X^2+X^3}$, or

$$\begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

(c) We first show that Φ_A is indeed a ring homomorphism. For polynomials $p(X), q(X) \in \mathbb{Z}_3[X]$ we have

$$\Phi_A(p(X) + q(X)) = \Phi_A((p+q)(X)) = (p+q)(A) = p(A) + q(A) = \Phi_A(p(X)) + \Phi_A(q(X)),$$

and

$$\Phi_A(p(X)q(X)) = \Phi_A((pq)(X)) = (pq)(A) = p(A)q(A) = \Phi_A(p(X))\Phi_A(q(X)),$$

and so indeed Φ_A is a ring homomorphism. By the first isomorphism theorem for ring homomorphisms we have

$$\operatorname{Im} \Phi_A \cong \mathbb{Z}_3[X]/\ker \Phi_A.$$

We want to show that $\operatorname{Im} \Phi_A$ is a field, so it is enough to show that $\mathbb{Z}_3[X]/\ker \Phi_A$ is a field. Let $m_A(X)$ be the minimal polynomial of A. Then by its definition we have that

$$\ker \Phi_A = \{p(X) \in \mathbb{Z}_3[X] \mid p(A) = 0\} = (m_A(X)).$$

We know that $m_A(X)$ is the last monic invariant factor of $A - XI_3$ by Theorem 19.7(4), and so by part (a) we have that $m_A(X) = -2 - X^2 - X^3$. We claim that $m_A(X)$ is an irreducible polynomial. Assume to a contradiction that $m_A(X)$ is not irreducible. Then there is a factorization

$$m_A(X) = p(X)q(X),$$

with $p(X), q(X) \in \mathbb{Z}_3[X]$ and where neither $p(X)$ nor $q(X)$ are units. Since $\mathbb{Z}_3$ is a field, it follows that $p(X)$ and $q(X)$ have degree at least 1. Since

$$3 = \deg(m_A(X)) = \deg(p(X)q(X)) = \deg(p(X)) + \deg(q(X)),$$

we conclude that one of the two polynomials $p(X)$ and $q(X)$ is of degree 1 and one is of degree 2. Then the polynomial of degree 1 is of the form $aX + b$ with $a \neq 0$ and so it has $a^{-1}b$ as a root. It follows that $a^{-1}b$ is a root of $m_A(X)$ as well. But

$$m_A(0) = -2, \;\; m_A(1) = -2, \;\; m_A(2) = 1$$

and so $m_A(X)$ has no root in $\mathbb{Z}_3[X]$. We conclude that $m_A(X)$ is indeed irreducible.

We now claim that $(m_A(X))$ is a maximal ideal. Assume that there exists an ideal I with $(m_A(X)) \subseteq I \subsetneq \mathbb{Z}_3[X]$ and we show that $I = (m_A(X))$. Since $\mathbb{Z}_3[X]$ is a PID, we have that $I = (p(X))$ for some polynomial $p(X) \in \mathbb{Z}_3[X]$. Then $m_A(X) \in (m_A(X)) \subseteq (p(X))$ implies that $m_A(X) = p(X)q(X)$ for some $q(X) \in \mathbb{Z}_3[X]$. Since $m_A(X)$ is irreducible, we conclude that $p(X)$ is a unit or $q(X)$ is a unit. Since $(p(X)) \neq \mathbb{Z}_3[X]$, we have that $p(X)$ cannot be a unit and so $q(X) = u \in \mathbb{Z}_3[X]$ is a unit. Then $p(X) = u^{-1}m_A(X)$ and so $p(X) \in (m_A(X))$ as well. We thus obtain that

$$I = (p(X)) \subseteq (m_A(X)) \subseteq I,$$

and so $I = (m_A(X))$ as claimed. Therefore $(m_A(X))$ is a maximal ideal. Hence the quotient

$$\operatorname{Im}\Phi_A \cong \mathbb{Z}_3[X]/\ker\Phi_A = \mathbb{Z}_3[X]/(m_A(X))$$

is a field.

Problem 8. (After Chapter 21.2.) Let R be a PID and $A, B \in M_{n\times m}(R)$ be $n \times m$ matrices with coefficients in R. Let $C(A)$ respectively $C(B)$ be the column module of A respectively B, that is the submodule of R^n generated by the columns of A respectively B.

(a) Show that

$$C(A) = \left\{ A\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle|\, r_1, \ldots, r_m \in R \right\}.$$

(b) Let $Q \in M_{m\times m}(R)$ be an invertible $m \times m$ matrix. Show that if $A = BQ$, then $C(A) = C(B)$.

(c) Let $Q \in M_{n\times n}(R)$ be an invertible $n \times n$-matrix. Let $T_P : R^n \to R^n$ be the R-module homomorphism given by $T_P\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = P\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$. Show that if $A = PB$, then T_P gives an isomorphism from $C(B)$ to $C(A)$.

(d) Show that if A is equivalent to B then $\frac{R^m}{C(A)}$ is isomorphic to $\frac{R^m}{C(B)}$.

(e) Use the Smith normal form to show that the converse of part (d) also holds, that is if $\frac{R^m}{C(A)}$ is isomorphic to $\frac{R^m}{C(B)}$, then A is equivalent to B.

Solution.

(a) Let

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Then

$$C(A) = \left\{ \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} r_1 + \cdots + \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix} r_m \,\middle|\, r_1, \ldots, r_m \in R \right\} = \left\{ \begin{pmatrix} a_{11}r_1 & \cdots & a_{1m}r_1 \\ \vdots & & \vdots \\ a_{n1}r_m & \cdots & a_{nm}r_m \end{pmatrix} \middle|\, r_1, \ldots, r_m \in R \right\}$$
$$= \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \middle|\, r_1, \ldots, r_m \in R \right\} = \left\{ A\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle|\, r_1, \ldots, r_m \in R \right\},$$

as required.

(b) Using part (a) we have

$$C(A) = \left\{ A\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle|\, r_1, \ldots, r_m \in R \right\} = \left\{ BQ\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle|\, r_1, \ldots, r_m \in R \right\}.$$

Since $Q\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \begin{pmatrix} r_1' \\ \vdots \\ r_m' \end{pmatrix}$ for some $r_1', \dots, r_m' \in R$, we have that

$$\left\{ BQ \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle| \ r_1, \dots, r_m \in R \right\} \subseteq \left\{ B \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle| \ r_1, \dots, r_m \in R \right\} = C(B).$$

Hence we have showed that $C(A) \subseteq C(B)$. Now notice that $A = BQ$ implies that $B = AQ^{-1}$ where $Q \in M_{m\times m}(R)$ is again invertible. We thus also obtain that $C(B) \subseteq C(A)$ and so $C(A) = C(B)$ as required.

(c) Using part (a) we have

$$C(A) = \left\{ A \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle| \ r_1, \dots, r_m \in R \right\} \text{ and } C(B) = \left\{ B \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \middle| \ r_1, \dots, r_m \in R \right\}.$$

Let $B\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in C(B)$. Then

$$T_P \left(B \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \right) = PB \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in C(A).$$

Hence $T_P(C(B)) \subseteq C(A)$. In particular, T_P defines an R-module homomorphism from $C(B)$ to $C(A)$. Let $A\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in C(A)$. Since $A = PB$ and P is invertible, we have that $P^{-1}B = A$. Hence $P^{-1}A\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = B\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in C(B)$ and

$$T_P \left(B \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \right) = PB \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Therefore $T_P : C(B) \to C(A)$ is surjective. Moreover, for any $v \in R^n$ we have that if $T_P(v) = 0$ then $Pv = 0$ and so $v = P^{-1}0 = 0$. Hence T_P is injective. We conclude that $T_P : C(B) \to C(A)$ is an isomorphism of R-modules.

(d) Since A is equivalent to B, there exist invertible matrices $P \in M_{n\times n}(R)$ and $Q \in M_{m\times m}(R)$ such that $A = PBQ$. Then by part (b) we obtain that $C((PB)Q) = C(PB)$, while by part (d) we obtain that $C(PB) \cong C(B)$. Then

$$\frac{R^m}{C(A)} = \frac{R^m}{C(PBQ)} = \frac{R^m}{C(PB)} \cong \frac{R^m}{C(B)},$$

as required.

(e) Let D_A be the Smith normal form of A and D_B be the Smith normal form of B, that is

$$D_A = \begin{pmatrix} d_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & d_k & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \text{ and } D_B = \begin{pmatrix} f_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & f_l & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

where $d_1, \dots, d_k, f_1, \dots, f_l \in R$ are nonzero, $d_1 \mid \cdots \mid d_k$ and $f_1 \mid \cdots \mid f_l$. Then by part (d) we have that

$$\frac{R^m}{C(D_A)} \cong \frac{R^m}{C(A)} \cong \frac{R^m}{C(B)} \cong \frac{R^m}{C(D_B)}.$$

Therefore we obtain an isomorphism of R-modules

$$\frac{R}{(d_1)} \oplus \cdots \frac{R}{(d_k)} \oplus R^s \cong R^m C(D_A) \cong \frac{R^m}{C(D_B)} \cong \frac{R}{(f_1)} \oplus \cdots \oplus \frac{R}{(f_l)} \oplus R^t.$$

By the uniqueness of finitely generated modules over a PID, we conclude that $k = l$ and that $(d_i) = (f_i)$ for $1 \leq i \leq k$. In particular the matrices D_A and D_B are the same up to multiplication of the diagonal elements by units. By the uniqueness of the Smith normal form, we conclude that A is equivalent to D_B and so A is equivalent to B as required.

Problem 9. (After Chapter 21.4.) Let K be a field and let $A \in M_{n\times n}(K)$ be an $n \times n$ matrix with coefficients in K. Consider the $K[X]$-module M_A where $M_A = K^n$ and for $p(X) \in K[X]$ and $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in K^n$ the multiplication is given by

$$p(X) \cdot_A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := p(A) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Define a map $\phi : K[X]^n \to M_A$ by

$$\phi((p_1(X), \dots, p_n(X))) = p_1(A) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + p_n(A) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

(a) Show that ϕ is a homomorphism of $K[X]$-modules.

(b) Show that $\operatorname{Im} \phi = M_A$.

(c) Show that $\ker \phi = C(A - XI_n)$, where $C(A - XI_n)$ is the column module of the matrix $A - XI_n \in M_{n\times n}(K[X])$. *(Hint: Show first that $C(A - XI_n) \subseteq \ker \phi$. Then for $(p_1(X), \dots, p_n(X)) \in \ker \phi$, show that $(p_1(X), \dots, p_n(X)) \in C(A - XI_n)$ by setting $\tilde{p}_j(X) = \frac{p_j(X) - b_{j0}}{X}$ and using induction on $d = \max\{\deg(p_1(X), \dots, p_n(X))\}$. This may be a bit challenging.)*

(d) Conclude that $\frac{K[X]^n}{C(A - XI_n)} \cong M_A$.

Solution.

(a) Let $(p_1(X), \dots, p_n(X)), (q_1(X), \dots, q_n(X)) \in K[X]^n$ and $f(X) \in K[X]$. We have

$$\begin{aligned}
\phi((p_1(X), \dots, p_n(X)) + (q_1(X), \dots, q_n(X))) &= \phi((p_1(X) + q_1(X), \dots, p_n(X) + q_n(X))) \\
&= (p_1(A) + q_1(A)) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (p_n(A) + q_1(A)) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \\
&= p_1(A) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + p_n(A) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} + q_1(A) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + q_n(A) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \\
&= \phi((p_1(X), \dots, p_n(X))) + \phi((q_1(X), \dots, q_n(X))),
\end{aligned}$$

and

$$\begin{aligned}
\phi(f(X)(p_1(X),\dots,p_n(X))) &= \phi((f(X)p_1(X),\dots,f(X)p_n(X))) \\
&= f(A)p_1(A)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+\cdots+f(A)p_n(A)\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix} \\
&= f(A)\left(p_1(A)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+\cdots+p_n(A)\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}\right) \\
&= f(X)\cdot_A\left(p_1(A)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+\cdots+p_n(A)\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}\right) \\
&= f(X)\cdot_A\phi((p_1(X),\dots,p_n(X))).
\end{aligned}$$

Hence ϕ is indeed a $K[X]$-module homomorphism.

(b) For any $\begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix}\in M_A=K^n$ we have that $v_i\in K[X]$ (the constant polynomial v_i). Then

$$\phi((v_1,\dots,v_n))=v_1I_n\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+\cdots+v_nI_n\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}=\begin{pmatrix}v_1\\0\\\vdots\\0\end{pmatrix}+\cdots+\begin{pmatrix}0\\\vdots\\0\\v_n\end{pmatrix}=\begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix},$$

showing that ϕ is a surjective map and so $\operatorname{Im}\phi=M_A$.

(c) Next we claim that $\ker\phi=C(A-XI_n)$. Let $A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{pmatrix}$ and set

$$A_1(X)=(a_{11}-X,a_{21},\dots,a_{n1}),\dots,A_n(X)=(a_{1n},\dots,a_{(n-1)n},a_{nn}-X),$$

that is $A_j(X)$ is the j-th column of the matrix $A-XI_n$ (written in row form). Then

$$\phi(A_j(X))=a_{1j}I_n\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+\cdots+(a_{jj}I_n-A)\begin{pmatrix}0\\\vdots\\0\\1\\0\\\vdots\\0\end{pmatrix}+a_{nj}I_n\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}=\begin{pmatrix}a_{1j}\\0\\\vdots\\0\end{pmatrix}+\cdots+\begin{pmatrix}-a_{1j}\\\vdots\\-a(j-1)j\\0\\-a_{(j+1)j}\\\vdots\\-a_{nj}\end{pmatrix}+\cdots\begin{pmatrix}0\\\vdots\\0\\a_{nj}\end{pmatrix}=\begin{pmatrix}0\\\vdots\\0\end{pmatrix}.$$

Hence $A_1(X),\dots,A_n(X)\in\ker\phi$ and so

$$C(A-XI_n)=(A_1(X),\dots,A_n(X))\subseteq\ker\phi.$$

For $1\le j\le n$ let e_j be the unit vector in $K[X]^n$. Then we have

$$Xe_j=(0,\dots,0,X,0,\dots,0)=(a_{1j},\dots,a_{nj})-A_j(X).$$

Now let $(p_1(X),\dots,p_n(X))\in\ker\phi$ and we show that $(p_1(X),\dots,p_n(X))\in C(A-XI_n)$. For this we use induction on $d=\max\{\deg(p_1(X)),\dots,\deg(p_n(X))\}$ (where the degree of the zero polynomial is defined to be $-\infty$.) If $d=-\infty$ then we obtain that $p_1(X)=\cdots=p_n(X)=0$ and so clearly $(0,\ldots,0)\in C(A-XI_n)$. If $d=0$, then all polynomials $p_1(X),\ldots,p_n(X)$ are constant. For $1\le j\le n$ let b_{j0} be the constant term of $p_j(X)$. Then

$$0=\phi(p_1(X),\ldots,p_n(X))=b_{10}I_n\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+\cdots+b_{n0}I_n\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}=\begin{pmatrix}b_{10}\\\vdots\\b_{n0}\end{pmatrix}$$

and so $(p_1(X),\ldots,p_n(X))=(0,\ldots,0)\in C(A-XI_n)$. Now assume that $d\ge 1$. For any $1\le j\le n$ set

$$\tilde{p}_j(X)=\frac{p_j(X)-b_{j0}}{X}.$$

Then we have

$$\begin{aligned}(p_1(X),\ldots,p_n(X))&=(b_{10}+\tilde{p}_1(X)X,\ldots,b_{n0}+\tilde{p}_n(X)X)\\&=(b_{10},\ldots,b_{n0})+\tilde{p}_1(X)(X,0,\ldots,0)+\cdots+\tilde{p}_n(X)(0,\ldots,0,X)\\&=(b_{10},\ldots,b_{n0})+\tilde{p}_1(X)((a_{11},\ldots,a_{n1})-A_1(X))+\cdots+\tilde{p}_n(X)((a_{1n},\ldots,a_{nn})-A_n(X))\\&=(b_{10},\ldots,b_{n0})+\tilde{p}_1(X)(a_{11},\ldots,a_{n1})+\cdots+\tilde{p}_n(X)(a_{1n},\ldots,a_{nn})\\&\quad-(\tilde{p}_1(X)A_1(X)+\cdots+\tilde{p}_n(X)A_n(X)).\end{aligned}$$

Since $\tilde{p}_1(X)A_1(X)+\cdots+\tilde{p}_n(X)A_n(X)\in C(A-XI_n)$, we obtain that

$$(b_{10},\ldots,b_{n0})+\tilde{p}_1(X)(a_{11},\ldots,a_{n1})+\cdots+\tilde{p}_n(X)(a_{1n},\ldots,a_{nn})\in C(A-XI_n)\subseteq\ker\phi.$$

Setting $q_j(X)=b_{j0}+\tilde{p}_1(X)a_{j1}+\cdots+\tilde{p}_n(X)a_{jn}$, we have shown that

$$(p_1(X),\ldots,p_n(X))=\underbrace{(q_1(X),\ldots,q_n(X))}_{\in\ker\phi}-\underbrace{(\tilde{p}_1(X)A_1(X)+\cdots+\tilde{p}_n(X)A_n(X))}_{\in C(A-XI_n)}.$$

Moreover, by construction we have

$$\max\{\deg(q_1(X)),\ldots,\deg(q_n(X))\}\le\max\{\deg(\tilde{p}_1(X)),\ldots,\deg(\tilde{p}_n(X))\}=d-1,$$

since by construction we have that

$$\deg(\tilde{p}_j(X))\begin{cases}\deg(p_j(X))-1, & \text{if }\deg(p_j(X))\ge 1,\\ -\infty, & \text{if }p_j(X)=b_{j0}.\end{cases}$$

By induction hypothesis we also obtain that $(q_1(X),\ldots,q_n(X))\in C(A-XI_n)$. We conclude that $(p_1(X),\ldots,p_n(X))\in C(A-XI_n)$, which shows that $\ker\phi=C(A-XI_n)$.

(d) By the first isomorphism theorem for modules we have that

$$\frac{K[X]^n}{C(A-XI_n)}=\frac{K[X]^n}{\ker\phi}\cong\operatorname{Im}\phi=M_A,$$

as required.

Problem 10. (After Chapter 14.3.) Let R be a PID and let $I_1\subseteq I_2\subseteq\cdots\subseteq I_n\subsetneq R$ be ideals in R. Assume that N is an R-module such that

$$N\cong R/I_1\times\cdots\times R/I_n.$$

Show that the following hold.

(a) There exists a generating set of N with n elements and every generating set of N has at least n elements. *(Hint: to show that every generating set of N has at least n elements, let M be a maximal ideal in R containing I_n and show that every generating set of N gives a generating set of $(R/M)^n$. Then, use the fact that R/M is a field to conclude that any generating set of $(R/M)^n$ has at least n elements.)*

(b) Let $r \in R$ and let $\phi_r : R \to R$ be the map $\phi_r(x) = rx$. Show that

$$rN \cong R/\phi_r^{-1}(I_1) \times \cdots \times R/\phi_r^{-1}(I_n).$$

(c) For $1 \leq i \leq n$ we have that

$$I_i = \{r \in R \mid rN \text{ can be generated by fewer than } i \text{ elements}\}.$$

(Hint: apply part (a) to part (b).)

Solution.

(a) Clearly the set

$$\{(1 + I_1, 0, \dots, 0), \dots, (0, \dots, 0, 1 + I_n)\}$$

is a generating set of N with n elements. Now assume to a contradiction that there exists a generating set of N with k elements where $k < n$. By Theorem 7.4 we have that there exists a maximal ideal M in R such that $I_n \subseteq M$. Then $I_i \subseteq I_n \subseteq M$ for all $1 \leq i \leq n$ and so by the third isomorphism theorem for modules (Problem 10 in Problem Set 3) we obtain a surjective homomorphism of R-modules $\pi_i : R/I_i \to R/M$ with $\ker(\pi) = M/I_i$, given by $\pi(r + I_i) = r + M$. It follows that the map

$$\pi : R/I_1 \times \cdots \times R/I_n \to R/M \times \cdots \times R/M$$

given by $\pi(r_1 + I_1, \dots, r_n + I_n) = (r_1 + M, \dots, r_n + M)$ is a surjective homomorphism of R-modules. By Lemma 12.3(2) it follows that there is a generating set of $R/M \times \cdots \times R/M$ as an R-module with k elements. Let $\{b_1, \dots, b_k\}$ be that generating set, so that for $1 \leq i \leq k$ we have

$$b_i = (b_{1i} + M, \dots, b_{ni} + M)$$

with $b_{1i}, \dots, b_{ni} \in R$. Since $\{b_1, \dots, b_k\}$ is an R-generating set for $R/M \times \cdots \times R/M$, we have that for any $(r_1 + M, \dots, r_n + M) \in R/M \times \cdots \times R/M$ there exist $s_1, \dots, s_k \in R$ such that

$$\begin{aligned}
(r_1 + M, \dots, r_n + M) &= \sum_{i=1}^{k} s_i b_i \\
&= \sum_{i=1}^{k} s_i (b_{1i} + M, \dots, b_{ni} + M) \\
&= \sum_{i=1}^{k} (s_i(b_{1i} + M), \dots, s_i(b_{ni} + M)) \\
&= \sum_{i=1}^{k} (s_i b_{1i} + M, \dots, s_i b_{ni} + M) \\
&= \sum_{i=1}^{k} ((s_i + M)(b_{1i} + M), \dots, (s_i + M)(b_{ni} + M)) \\
&= \sum_{i=1}^{k} (s_i + M)(b_{1i} + M, \dots, b_{ni} + M) \\
&= \sum_{i=1}^{k} (s_i + M) b_i.
\end{aligned}$$

It follows that $\{b_1, \dots, b_k\}$ generate $R/M \times \cdots \times R/M$ as an R/M-module as well. But M is a maximal ideal, and so R/M is a field. Then $R/M \times \cdots \times R/M \cong (R/M)^n$ has a generating set of k elements with $k < n$, which is a contradiction. This shows part (a).

(b) Since

$$N \cong \frac{R}{I_1} \times \cdots \times \frac{R}{I_n} \cong \frac{R \times \cdots \times R}{I_1 \times \cdots \times I_n},$$

there exists an isomorphism of R-modules $\eta : (R \times \cdots \times R)/(I_1 \times \cdots \times I_n) \to N$. Let

$$\pi : R^n = R \times \cdots \times R \to \frac{R \times \cdots \times R}{I_1 \times \cdots \times I_n}$$

be the canonical projection. Composing with η we obtain a surjective homomorphism of R-modules $\psi = \eta \circ \pi : R^n \to N$ with kernel equal to $I_1 \times \cdots \times I_n$.

Now let $r \in R$ and define a map $\zeta : R^n \to rN$ by $\zeta(x) = r\psi(x)$. For $x_1, x_2 \in R^n$ and $s \in R$ we have that

$$\zeta(x_1 + x_2) = r\psi(x_1 + x_2) = r(\psi(x_1) + \psi(x_2)) = r\psi(x_1) + r\psi(x_2) = \zeta(x_1) + \zeta(x_2),$$

and

$$\zeta(sx) = r(\psi(sx)) = r(s\psi(x)) = s(r\psi(x)) = s\zeta(x),$$

and so ζ is a homomorphism of R-modules. We know compute the kernel and image of ζ.

Assume first that $\zeta(x) = 0$ for some $x \in R^n$. Then

$$0 = \zeta(x) = r\psi(x) = \psi(rx),$$

and so $rx \in \ker \psi = I_1 \times \cdots \times I_n$. Writing $x = (x_1, \ldots, x_n)$, we obtain that $rx_i \in I_i$ for $1 \leq i \leq n$. But then $x_i \in \phi_r^{-1}(I_i)$ for $1 \leq i \leq n$. Hence $\ker \zeta \subseteq \phi_r^{-1}(I_1) \times \cdots \times \phi_r^{-1}(I_n)$. The other inclusion follows immediately, since if $(a_1, \ldots, a_n) \in \phi_r^{-1}(I_1) \times \cdots \times \phi_r^{-1}(I_n)$, then $ra_i = \psi_r(a_i) \in I_i$ and so

$$\zeta(a_1, \ldots, a_n) = r\psi(a_1, \ldots, a_n) = \psi(ra_i, \ldots, ra_n) = 0,$$

since $(ra_1, \ldots, ra_n) \in I_1 \times \cdots \times I_n = \ker \psi$. This shows that $\ker \zeta I_1 \times \cdots \times I_n$.

Now let $ry \in rN$. Since ψ is surjective, there exists $x \in R^n$ with $\psi(x) = y$. Then $\zeta(x) = r\psi(x) = ry$ and so ζ is surjective too. We obtain that $\operatorname{Im} \zeta = rN$.

Using the first isomorphism theorem for modules, we obtain an isomorphism

$$\frac{R \times \cdots \times R}{I_1 \times \cdots \times I_n} = R^n / \ker \zeta \cong \operatorname{Im} \zeta = rN,$$

as required.

(c) Let $1 \leq i \leq n$ and $r \in R$. We claim that $r \in I_i$ if and only if $\phi_r^{-1}(I_i) = R$. Indeed, if $\phi_r^{-1}(I_i) = R$, then $1 \in R = \phi_r^{-1}(I_i)$ implies that $r = \phi_r(1) \in I_i$. For the other direction, assume that $r \in I_i$ and let $x \in R$. Since I_i is an ideal, we have that $\phi_r(x) = rx \in I_i$. But then $x \in \phi_r^{-1}(I_i)$ showing that $R \subseteq \phi_r^{-1}(I_i)$. We conclude that $\phi_r^{-1}(I_i) = R$.

Now let $r \in I_i$ and we want to show that rN is generated by fewer than i elements. Since $r \in I_i \subseteq I_{i+1} \subseteq \cdots \subseteq I_n$, we obtain that

$$\phi_r^{-1}(I_i) = \phi_r^{-1}(I_{i+1}) = \cdots = \phi_r^{-1}(I_n) = R.$$

Then by part (b) we have that

$$rN \cong R/\phi_r^{-1}(I_1) \times \cdots \times R/\phi_r^{-1}(I_{i-1}) \times \underbrace{R/\phi_r^{-1}(I_i)}_{0} \times \cdots \times \underbrace{R/\phi_r^{-1}(I_n)}_{0} \cong R/\phi_r^{-1}(I_1) \times \cdots \times R/\phi_r^{-1}(I_{i-1}),$$

and so by part (a) we obtain that rN can be generated by fewer than i elements.

Finally, let r be such that rN is generated by fewer than i elements and we want to show that $r \in I_i$. Assume to a contradiction that $r \notin I_i$. Then

$$rN \cong R/\phi_r^{-1}(I_1) \times \cdots \times R/\phi_r^{-1}(I_{i-1}) \times \underbrace{R/\phi_r^{-1}(I_i)}_{\neq 0} \times \cdots \times \underbrace{R/\phi_r^{-1}(I_n)}_{0},$$

and so by part (a) any set of generators of rN has at least i elements. But this is a contradiction and so $r \in I_i$.

Extra problems

The following problems may be a bit more challenging, in case you feel like you need something more.

Problem 11. (After Chapter 21.4.) Let F be a field and assume that F is a subfield of another field K. Let $A, B \in M_{n\times n}(F)$ be two $n \times n$ matrices with coefficients in F. Show that A and B are similar over $M_{n\times n}(F)$ if and only if they are similar over $M_{n\times n}(K)$.

Solution. Assume first that A and B are similar over $M_{n\times n}(F)$. Then there exists an invertible matrix $P \in M_{n\times n}(F)$ such that

$$A = PBP^{-1}.$$

Since $P \in M_{n\times n}(F)$ and $F \subseteq K$, we have that $P \in M_{n\times n}(K)$ as well. Since F is a subfield of K, we still have that $A = PBP^{-}$ and so A and B are similar over $M_{n\times n}(K)$ as well.

Assume now that A and B are similar over $M_{n\times n}(K)$. Let R be the rational canonical form of A over F. Then this rational canonical form has the companion matrices of the non-unit monic invariant factors of A as a matrix in $M_{n\times n}(F)$. These invariant factors of A are unique, and they still remain invariant factors of A as matrix in $M_{n\times n}(K)$ (as they only depend on the elements of A). It follows that R is the rational canonical form of A as a matrix in $M_{n\times n}(K)$. But then the same is true of B and its rational canonical form R'. Since A and B are similar over $M_{n\times n}(K)$, we have that $R = R'$. But then, and since this common matrix is the rational canonical form of both A and B as matrices in $M_{n\times n}(F)$, we conclude that A and B are similar over $M_{n\times n}(F)$.

Problem 12. (After Chapter 21.5.) For a matrix $M \in M_{n\times n}(\mathbb{R})$, we define

$$\exp(M) = e^M := I_n + M + \tfrac{1}{2!}M^2 + \tfrac{1}{3!}M^3 + \cdots .$$

(It can be shown that this is well-defined.) Let $y_1(t), \ldots, y_n(t)$ be functions from $\mathbb{R}$ to $\mathbb{R}$ and let $\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$. Consider the initial value problem

$$\begin{cases} \mathbf{y}'(t) &= A\mathbf{y}(t), \\ \mathbf{y}(0) &= y_0, \end{cases}$$

where $A \in M_{n\times n}(\mathbb{R})$, $y_0 \in \mathbb{R}^n$ and $\mathbf{y}'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_n'(t) \end{pmatrix}$. It can be shown that this problem has a unique solution given by $\mathbf{y}(t) = e^{At}y_0$. Assume that the Jordan canonical form of A has only Jordan blocks corresponding to polynomials of degree 1. Use this Jordan canonical form to compute e^{At}.

Solution. Let $P \in M_{n\times n}(\mathbb{R})$ be such that

$$PAP^{-1} = J = \left(\begin{array}{c|c|c|c} J_1 & 0 & 0 & 0 \\ \hline 0 & J_2 & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & 0 & J_k \end{array}\right)$$

is the canonical Jordan form of A, where

$$J_i = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 \\ 0 & 1 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}$$

for some $\lambda_i \in \mathbb{R}$ (and of varying size). We substitute $\mathbf{z}(t) = P\mathbf{y}(t)$ (and so $\mathbf{y}(t) = P^{-1}\mathbf{z}(t)$ to obtain

$$\mathbf{z}'(t) = P\mathbf{y}'(t) = PA\mathbf{y}(t) = PAP^{-1}\mathbf{z}(t) = J\mathbf{z}(t),$$

and

$$\mathbf{z}(0) = P\mathbf{y}(0) = Py_0$$

Then we know by the statement in the problem that the general solution to the initial value problem

$$\begin{cases} \mathbf{z}'(t) &= J\mathbf{z}(t), \\ \mathbf{z}(0) &= Py_0, \end{cases}$$

is given by

$$\mathbf{z}(t) = e^{Jt}Py_0.$$

Therefore, we obtain that

$$\mathbf{y}(t) = P^{-1}\mathbf{z}(t) = P^{-1}e^{Jt}Py_0,$$

and so it is enough to compute e^{Jt}. Since J is a block diagonal matrix, we have that

$$J^m = \left(\begin{array}{c|c|c|c} J_1 & 0 & 0 & 0 \\ \hline 0 & J_2 & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & 0 & J_k \end{array}\right)^m = \left(\begin{array}{c|c|c|c} J_1^m & 0 & 0 & 0 \\ \hline 0 & J_2^m & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & 0 & J_k^m \end{array}\right)$$

for all $m \geq 1$. Since e^J is just a linear combination of powers of J, it follows that

$$e^{Jt} = \exp\left(\left(\begin{array}{c|c|c|c} J_1t & 0 & 0 & 0 \\ \hline 0 & J_2t & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & 0 & J_kt \end{array}\right)\right) = \left(\begin{array}{c|c|c|c} e^{J_1t} & 0 & 0 & 0 \\ \hline 0 & e^{J_2t} & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & 0 & e^{J_kt} \end{array}\right).$$

Hence it is enough to compute e^{J_it}. We may write

$$J_i = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 \\ 0 & 1 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}}_{=\lambda_i I_{s_i}} + \underbrace{\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}}_{:=V_i}.$$

Then clearly we have that

$$e^{\lambda_i t I_{s_i}} = \begin{pmatrix} e^{\lambda_i t} & 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_i t} & 0 & \cdots & 0 \\ 0 & 0 & e^{\lambda_i t} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda_i t} \end{pmatrix}.$$

Moreover, we have that each successive power of the V_i moves the sub-diagonal with 1's one step to the lower left, until $V_i^{s_i-1} = 0$. Then a direct computation gives that

$$e^tV_i = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ t & 1 & 0 & \cdots & 0 \\ \frac{t^2}{2!} & t & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{t^{s_i-1}}{(s_i-1)!} & \frac{t^{s_i-2}}{(s_i-2)!} & \frac{t^{(s_i-3)}}{(s_i-3)!} & \cdots & 1 \end{pmatrix}.$$

Finally, note also that $\lambda_i I_{s_i}$ and V_i commute. Therefore we may use the binomial theorem and the Cauchy product together with the above calculations to obtain

$$
\begin{aligned}
e^{J_i t} &= e^{\lambda_i t I_{s_i} + tV_i} = \sum_{j=0}^{\infty} \frac{(\lambda_i t I_{s_i} + tV_i)^j}{j!} = \sum_{j=0}^{\infty} \left(\sum_{a+b=j} \frac{(\lambda_i t I_{s_i})^a}{a!} \frac{(tV_i)^b}{b!} \right) = \left(\sum_{a=0}^{\infty} \frac{(\lambda_i t I_{s_i})^a}{a!} \right) \left(\sum_{b=0}^{\infty} \frac{(tV_i)^b}{b!} \right) \\
&= e^{\lambda_i t I_{s_i}} e^{tV_i} = \begin{pmatrix} e^{\lambda_i t} & 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_i t} & 0 & \cdots & 0 \\ 0 & 0 & e^{\lambda_i t} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda_i t} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ t & 1 & 0 & \cdots & 0 \\ \frac{t^2}{2!} & t & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{t^{s_i-1}}{(s_i-1)!} & \frac{t^{s_i-2}}{(s_i-2)!} & \frac{t^{(s_i-3)}}{(s_i-3)!} & \cdots & 1 \end{pmatrix}.
\end{aligned}
$$

This is now the value of $e^{J_i t}$ which we may replace in the computation of e^{Jt}, letting us obtain the required solution.