In this problem we establish the real Jordan form. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear. The *complexification* of T is defined as $T_\mathbb{C} : \mathbb{C}^n \to \mathbb{C}^n$ by

$$T_\mathbb{C}(v + iw) = T(v) + iT(w) \ .$$

1. **Show that $T_\mathbb{C}$ is a linear transformation on $\mathbb{C}^n$. If $\lambda \in \mathbb{C}$ is one of its eigenvalues and $\hat{E}_\lambda$ is the corresponding generalized eigenspace, show that $\mathfrak{c}(\hat{E}_\lambda) = \hat{E}_{\overline{\lambda}}$. (Here $\mathfrak{c}$ is the complex conjugation map from last problem.)**

   Solution. If $v \in \mathbb{C}^n$ we can write it as $u + iw$ for $u, w \in \mathbb{R}^n$. Then if $z = x + iy$ is a complex scalar,

$$z(u + iw) = (x + iy)(u + iw) = xu - yw + i(yu + xw) \ .$$

   Therefore if $v_1, v_2 \in \mathbb{C}^n$, written as $v_1 = u_1 + iw_1$ and $v_2 = u_2 + iw_2$, then

$$\begin{aligned} T_\mathbb{C}(zv_1 + v_2) &= T_\mathbb{C}(xu_1 - yw_1 + u_2 + i(yu_1 + xw_1 + w_2)) \\ &= T(xu_1 - yw_1 + u_2) + iT(yu_1 + xw_1 + w_2) \\ &= [xT(u_1) - yT(w_1) + i(yT(u_1) - xT(w_1))] + [T(u_2) + iT(w_2)] \\ &= (x + iy)(T(u_1) + iT(w_1)) + T(u_2) + iT(w_2) \\ &= (x + iy)T_\mathbb{C}(u_1 + iw_1) + T_\mathbb{C}(u_2 + iw_2) \ . \end{aligned}$$

   This means $T_\mathbb{C}$ is linear.

   Next let λ be an eigenvalue of $T_\mathbb{C}$ and $v = u + iw$ be in the generalized eigenspace for $T_\mathbb{C}$ corresponding to eigenvalue λ. Then note that

$$T_\mathbb{C}(\mathfrak{c}(v)) = T_\mathbb{C}(u + i(-w)) = T(u) + iT(-w) = T(u) - iT(w) = \mathfrak{c}(T_\mathbb{C}(v)) \ .$$

   Therefore for any $m \geq 1$, $T_\mathbb{C}^m(\mathfrak{c}(v)) = \mathfrak{c}(T_\mathbb{C}^m(v))$. Now since v is a generalized eigenvector for eigenvalue λ, there exists $k \geq 1$ such that $(T_\mathbb{C} - \lambda I)^k(v) = \vec{0}$. Write $(T_\mathbb{C} - \lambda I)^k = a_k T_\mathbb{C}^k + \cdots + a_0 I$ so that

$$\begin{aligned} \vec{0} = (T_\mathbb{C} - \lambda I)^k(v) &= (T_\mathbb{C} - \lambda I)^k(\mathfrak{c}(\mathfrak{c}(v))) \\ &= a_k T_\mathbb{C}^k(\mathfrak{c}(\mathfrak{c}(v))) + \cdots + a_0 I(\mathfrak{c}(\mathfrak{c}(v))) \\ &= \mathfrak{c}\left(\overline{a_k} T_\mathbb{C}^k(\mathfrak{c}(v)) + \cdots + \overline{a_0} I(\mathfrak{c}(v))\right) \\ &= \mathfrak{c}\left((T_\mathbb{C} - \overline{\lambda} I)^k(\mathfrak{c}(v))\right) \ . \end{aligned}$$

   However $\mathfrak{c}$ is injective, so we find that $\mathfrak{c}(v)$ is a generalized eigenvector for eigenvalue $\overline{\lambda}$. This means

$$\mathfrak{c}(\hat{E}_\lambda) \subset \hat{E}_{\overline{\lambda}} \ .$$

   Repeating the argument with $\overline{\lambda}$ in place of λ gives

$$\mathfrak{c}(\hat{E}_{\overline{\lambda}}) \subset \hat{E}_\lambda \ .$$

   Now take complex conjugate of both sides to get

$$\hat{E}_{\overline{\lambda}} \subset \mathfrak{c}(\hat{E}_\lambda) \ .$$

2. **Show that the non-real eigenvalues of $T_{\mathbb{C}}$ come in pairs. In other words, show that we can list the distinct eigenvalues of $T_{\mathbb{C}}$ as**

$$\lambda_1, \ldots, \lambda_r, \sigma_1, \ldots, \sigma_{2m} \ ,$$

where for each $j = 1, \ldots, r$, $\overline{\lambda_j} = \lambda_j$ and for each $i = 1, \ldots, m$, $\sigma_{2i-1} = \overline{\sigma_{2i}}$.

Solution. Let σ be an eigenvalue of $T_{\mathbb{C}}$. Then there is a nonzero vector v in $\hat{E}_\sigma$. By the previous part, $\mathfrak{c}(v)$, which is also a nonzero vector, is an element of $\hat{E}_{\overline{\sigma}}$. Thus $T_{\mathbb{C}}$ has an eigenvector associated to $\overline{\sigma}$, meaning $\overline{\sigma}$ is an eigenvalue as well.

3. **Because $\mathbb{C}$ is algebraically closed, the proof of Jordan form shows that**

$$\mathbb{C}^n = \hat{E}_{\lambda_1} \oplus \cdots \oplus \hat{E}_{\lambda_r} \oplus \hat{E}_{\sigma_1} \oplus \cdots \oplus \hat{E}_{\sigma_{2m}} \ .$$

Using the previous two parts, show that for $j = 1, \ldots, r$ and $i = 1, \ldots, m$, the subspaces of $\mathbb{C}^n$

$$\hat{E}_{\lambda_j} \text{ and } \hat{E}_{\sigma_{2i-1}} \oplus \hat{E}_{\sigma_{2i}}$$

are $\mathfrak{c}$-invariant.

Solution. If $v \in \hat{E}_{\lambda_j}$ then part 1 shows that $\mathfrak{c}(v) \in \hat{E}_{\overline{\lambda_j}}$. But λ_j is real, so it equals its complex conjugate, implying that $\mathfrak{c}(v) \in \hat{E}_{\lambda_j}$. Thus $\hat{E}_{\lambda_j}$ is $\mathfrak{c}$-invariant.

If however $v \in \hat{E}_{\sigma_{2i-1}} \oplus \hat{E}_{\sigma_{2i}}$ then we can write $v = v_1 + v_2$ where

$$v_1 \in \hat{E}_{\sigma_{2i-1}} \text{ and } v_2 \in \hat{E}_{\sigma_{2i}} \ .$$

Then $\mathfrak{c}(v) = \mathfrak{c}(v_1 + v_2) = \mathfrak{c}(v_1) + \mathfrak{c}(v_2)$. But by part (a),

$$\mathfrak{c}(v_1) \in \hat{E}_{\sigma_{2i}} \text{ and } \mathfrak{c}(v_2) \in \hat{E}_{\sigma_{2i-1}} \ ,$$

so $\mathfrak{c}(v) \in \hat{E}_{\sigma_{2i-1}} \oplus \hat{E}_{\sigma_{2i}}$, giving $\mathfrak{c}$-invariance.

4. **Deduce from the previous problem that there exist subspaces $X_1, \ldots, X_r$ and $Y_1, \ldots, Y_m$ of $\mathbb{R}^n$ such that for each $j = 1, \ldots, r$ and $i = 1, \ldots, m$,**

$$\hat{E}_{\lambda_j} = \mathbf{Span}(\iota(X_j)) \text{ and } \hat{E}_{\sigma_{2i-1}} \oplus \hat{E}_{\sigma_{2i}} = \mathbf{Span}(\iota(Y_i)) \ .$$

Show that $\mathbb{R}^n = X_1 \oplus \cdots \oplus X_r \oplus Y_1 \oplus \cdots \oplus Y_m$.

Solution. It was shown in the last problem (on HW 10) that a subspace W of $\mathbb{C}^n$ is $\mathfrak{c}$-invariant if and only if there exists a subspace U of $\mathbb{R}^n$ such that $W = \mathrm{Span}(\iota(U))$. Applying this result to the subspaces in the last part gives the existence of $X_1, \ldots, X_r$ and $Y_1, \ldots, Y_m$.

To prove independence of these subspaces, we will use independence of their embeddings in $\mathbb{C}^n$. Let $v_1, \ldots, v_r$ and $v_1', \ldots, v_m'$ be vectors in $\mathbb{R}^n$ such that

$$v_i \in X_i \text{ for } i = 1, \ldots r \text{ and } v_i' \in Y_i \text{ for } i = 1, \ldots, m$$

and $v_1+\cdots+v_r+v_1'+\cdots+v_m'=\vec{0}$. Then we may embed both sides an use $\mathbb{R}$-linearity:

$$\vec{0}=\iota(v_1+\cdots+v_r+v_1'+\cdots+v_m')=\iota(v_1)+\cdots+\iota(v_r)+\iota(v_1')+\cdots+\iota(v_m') \ .$$

Since these vectors lie in independent spaces in $\mathbb{C}^n$, they are all zero. But if for some vector $v\in\mathbb{R}^n$ we have $\iota(v)=\vec{0}$ we must have $v=\vec{0}$; this reason is that $\iota(v)=(v,\vec{0})$, so both components must be $\vec{0}$, giving $v=\vec{0}$. Therefore

$$v_1=\cdots=v_r=\vec{0} \text{ and } v_1'=\cdots=v_m'=\vec{0}$$

and the spaces $X_1,\ldots,X_r,Y_1,\ldots,Y_m$ are independent.

Last we must show that $\mathbb{R}^n=X_1+\cdots+X_r+Y_1+\cdots+Y_m$. So let $v\in\mathbb{R}^n$ and consider the embedding $\iota(v)$. Since it is in $\mathbb{C}^n$ we can write it as a sum of elements of the generalized eigenspaces. We then use the previous part to find vectors $v_1,\ldots,v_r$ and $v_1',\ldots,v_m'$ with $v_i\in X_i$ and $v_j'\in Y_j$ and complex scalars $z_1,\ldots,z_r,z_1',\ldots,z_m'$ such that

$$\iota(v)=z_1\iota(v_1)+\cdots+z_r\iota(v_r)+z_1'\iota(v_1')+\cdots+z_m'\iota(v_m')$$

or

$$(v,\vec{0})=z_1(v_1,\vec{0})+\cdots+z_r(v_r,\vec{0})+z_1'(v_1',\vec{0})+\cdots+z_m'(v_m',\vec{0}) \ .$$

Now writing $z_1=x_1+iy_1,\ldots,z_r=x_r+iy_r$ and $z_1'=x_1'+iy_1',\ldots,z_m'=x_m'+iy_m'$, we find

$$(v,\vec{0})=(x_1v_1+\cdots+x_rv_r+x_1'v_1'+\cdots+x_m'v_m',y_1v_1+\cdots+y_rv_r+y_1'v_1'+\cdots+y_m'v_m') \ .$$

Finally,

$$v=x_1v_1+\cdots+x_rv_r+x_1'v_1'+\cdots+x_m'v_m' \ ,$$

proving that $v\in X_1+\cdots+X_r+Y_1+\cdots+Y_m$.

5. **Prove that for each $j=1,\ldots,r$, the transformation $T-\lambda_jI$ restricted to X_j is nilpotent and thus we can find a basis B_j for X_j consisting entirely of chains for $T-\lambda_jI$.**

Solution. First we prove that each X_j is T-invariant, so that the restriction is a linear transformation. Let $v\in X_j$; by definition $\iota(v)\in\hat{E}_{\lambda_j}$, a $T_{\mathbb{C}}$-invariant space. So we can write

$$(T(v),\vec{0})=T_{\mathbb{C}}(\iota(v))\in\hat{E}_{\lambda_j}=\mathrm{Span}(\iota(X_j)) \ .$$

This means that we can write $(T(v),\vec{0})=\sum_{k=1}^{\ell}z_k\iota(v_k)$ for complex scalars z_k and vectors $v_k\in X_j$. Writing $z_k=x_k+iy_k$, this becomes

$$(T(v),\vec{0})=(x_1v_1+\cdots+x_\ell v_\ell,y_1v_1+\cdots+y_\ell v_\ell) \ ,$$

or $T(v)=x_1v_1+\cdots+x_\ell v_\ell$. Therefore $T(v)\in X_j$.

Next we show that if $v\in X_j$ then there exists $k\geq 1$ such that $(T-\lambda_jI)^k(v)-\vec{0}$. Assuming we show that, then if $\{v_1,\ldots,v_s\}$ is a basis for X_j then we can choose

$k_1, \ldots, k_s$ such that $(T - \lambda I)^{k_p}(v_p) = \vec{0}$ for all $p = 1, \ldots, s$ and then $(T - \lambda_j I)^{k^*} = 0$ where $k^* = \max\{k_1, \ldots, k_s\}$, giving nilpotency.

So let $v \in X_j$. Then $\iota(v) \in \hat{E}_{\lambda_j}$ so there exists $k \geq 1$ such that $(T_{\mathbb{C}} - \lambda_j I)^k(\iota(v)) = \vec{0}$. because λ_j is real, we can find $a_0, \ldots, a_k \in \mathbb{R}$ such that

$$(T_{\mathbb{C}} - \lambda_j I)^k = a_k T_{\mathbb{C}}^k + \cdots + a_0 I$$

so

$$\begin{aligned}
\vec{0} = (T_{\mathbb{C}} - \lambda_j I)^k(\iota(v)) &= a_k T_{\mathbb{C}}^k(\iota(v)) + \cdots + a_0 \iota(v) \\
&= a_k T_{\mathbb{C}}^k((v, \vec{0})) + \cdots + a_0 (v, \vec{0}) \\
&= a_k (T^k(v), \vec{0}) + \cdots + a_0 (v, \vec{0}) \\
&= (a_k T^k(v) + \cdots + a_0 v, \vec{0}) \\
&= ((T - \lambda_j I)^k(v), \vec{0}) \ .
\end{aligned}$$

This implies $(T - \lambda_j I)^k(v) = \vec{0}$ and completes the proof.

6. **For each $k = 1, \ldots, m$, let**

$$C_k = \{v_1^{(k)} + i w_1^{(k)}, \ldots, v_{n_k}^{(k)} + i w_{n_k}^{(k)}\}$$

be a basis of $\hat{E}_{\sigma_{2k-1}}$ consisting of chains for $T_{\mathbb{C}} - \sigma_{2k-1} I$. Prove that

$$\hat{C}_k = \{v_1^{(k)}, w_1^{(k)}, \ldots, v_{n_k}^{(k)}, w_{n_k}^{(k)}\}$$

is a basis for Y_k. Describe the form of the matrix representation of T restricted to Y_k, relative to the basis $\hat{C}_k$.

Solution. This problem is worded a bit strangely (my bad!), because we do not want to assume that Y_k is fixed by part 4. We will first show that $\mathrm{Span}(\iota(\hat{C}_k)) = \hat{E}_{\sigma_{2k-1}} \oplus \hat{E}_{\sigma_{2k}}$, so then define Y_k as the span in $\mathbb{R}^n$ of $\hat{C}_k$ (recall part 4 only said "there exists").

First let $u \in \hat{C}_k$. Then either $u = v_p^{(k)}$ or $w_p^{(k)}$ for some p. In the first case,

$$\iota(u) = \iota(v_p^{(k)}) = (v_p^{(k)}, \vec{0}) = \frac{1}{2}\left[(v_p^{(k)}, w_p^{(k)}) + \mathfrak{c}((v_p^{(k)}, w_p^{(k)}))\right] \ .$$

However $(v_p^{(k)}, w_p^{(k)}) \in \hat{E}_{\sigma_{2k-1}}$, so its conjugate is in $\hat{E}_{\sigma_{2k}}$. This means $\iota(u) \in \hat{E}_{\sigma_{2k-1}} \oplus \hat{E}_{\sigma_{2k}}$. In the case that $u = w_p^{(k)}$ we can represent

$$\iota(u) = \frac{1}{2}\left[(v_p^{(k)}, w_p^{(k)}) - \mathfrak{c}((v_p^{(k)}, w_p^{(k)}))\right]$$

and arrive at the same conclusion. Either way, $\iota(\hat{C}_k) \subset \hat{E}_{\sigma_{2k-1}} \oplus \hat{E}_{\sigma_{2k}}$. Taking the span gives one inclusion.

For the other inclusion, let $v_p^{(k)} + iw_p^{(k)}$ be a vector in C_k. Then we can write

$$v_p^{(k)} + iw_p^{(k)} = (v_p^{(k)}, \vec{0}) + i(w_p^{(k)}, \vec{0}) = \iota(v_p^{(k)}) + i\iota(w_p^{(k)}) \ .$$

This is a combination of vectors of the form $\iota(u)$ for $u \in \hat{C}_k$, so each element of C_k is in the span of $\iota(\hat{C}_k)$. Because C_k spans $\hat{E}_{\sigma_{2k-1}} \oplus \hat{E}_{\sigma_{2k}}$, we get the other inclusion, proving

$$\mathrm{Span}(\iota(\hat{C}_k)) = \hat{E}_{\sigma_{2k-1}} \oplus \hat{E}_{\sigma_{2k}} \ .$$

We have now defined $Y_k = \mathrm{Span}(\hat{C}_k)$, so clearly $\hat{C}_k$ is a spanning set. We must then show linear independence. Let $a_1, \ldots, a_{n_k}, b_1, \ldots, b_{n_k}$ be real numbers and suppose that

$$a_1 v_1^{(k)} + \cdots + a_{n_k} v_{n_k}^{(k)} + b_1 w_1^{(k)} + \cdots + b_{n_k} w_{n_k}^{(k)} = \vec{0} \ .$$

We now apply ι and use $\mathbb{R}$-linearity, getting

$$\vec{0} = a_1(v_1^{(k)}, \vec{0}) + \cdots + a_{n_k}(v_{n_k}^{(k)}, \vec{0}) + b_1(w_1^{(k)}, \vec{0}) + \cdots + b_{n_k}(w_{n_k}^{(k)}, \vec{0})$$

For $j = 1, \ldots, n_k$ write $u_j = v_j^{(k)} + iw_j^{(k)}$. Then we can rewrite the right side as

$$\frac{1}{2}\left[a_1(u_1 + \mathfrak{c}(u_1)) + \cdots + a_{n_k}(u_{n_k} + \mathfrak{c}(u_{n_k}))\right] - \frac{i}{2}\left[b_1(u_1 + \mathfrak{c}(u_1)) + \cdots + b_{n_k}(u_{n_k} + \mathfrak{c}(u_{n_k}))\right] \ .$$

This can be decomposed into a sum of two vectors:

$$\frac{1}{2}\left[(a_1 - ib_1)u_1 + \cdots + (a_{n_k} - ib_{n_k})u_{n_k}\right]$$

and

$$\mathfrak{c}\left(\frac{1}{2}\left[(a_1 + ib_1)u_1 + \cdots + (a_{n_k} + ib_{n_k})u_{n_k}\right]\right) \ .$$

The first vector is in $\hat{E}_{\sigma_{2k-1}}$ and the second is in $\hat{E}_{\sigma_{2k}}$. Since these spaces are independent, it follows that they are both zero. By linear independence of the u_i's, we conclude that

$$a_1 - ib_1 = \cdots = a_{n_k} - ib_{n_k} = 0 \ ,$$

so $a_1 = \cdots = a_{n_k} = b_1 = \cdots = b_{n_k} = 0$. This completes the proof.

For the last part of the question, write again the vectors in C_k as $u_1, \ldots, u_{n_k}$ and consider a chain of generalized eigenvectors $u_1, \ldots, u_p$, where

$$T_{\mathbb{C}}(u_j) = \sigma_{2k-1} u_j + u_{j-1} \text{ for } j = 2, \ldots, p \text{ and } T_{\mathbb{C}}(u_1) = \sigma_{2k-1} u_1 \ .$$

Writing $\sigma_{2k-1} = a + bi$ for some $a, b \in \mathbb{R}$, we can then write this in real and imaginary parts as

$$(T(v_j^{(k)}, T(w_j^{(k)})) = (av_j^{(k)} - bw_j^{(k)}, aw_j^{(k)} + bv_j^{(k)}) + (v_j^{(k-1)}, w_j^{(k-1)}) \text{ for } j = 2, \ldots, p$$

and $(T(v_1^{(k)}), T(w_1^{(k)})) = (av_1^{(k)} - bw_1^{(k)}, aw_1^{(k)} + bv_1^{(k)})$. Splitting into real and imaginary parts, we find

$$T(v_j^{(k)}) = \begin{cases} av_j^{(k)} - bw_j^{(k)} + v_j^{(k-1)} & \text{if } j = 2, \ldots, p \\ av_1^{(k)} - bw_1^{(k)} & \text{if } j = 1 \end{cases}$$

and

$$T(w_j^{(k)}) = \begin{cases} bv_j^{(k)} + aw_j^{(k)} + w_j^{(k-1)} & \text{if } j = 2, \ldots, p \\ bv_1^{(k)} + aw_1^{(k)} & \text{if } j = 1 \end{cases} .$$

Finally, this means that corresponding to the vectors $\{v_1^{(k)}, w_1^{(k)}, \ldots, v_p^{(k)}, w_p^{(k)}\}$, the block in our matrix is

$$\begin{pmatrix} \Sigma & I & & & \\ & \Sigma & I & & \\ & & \cdots & & \\ & & & \Sigma & I \\ & & & & \Sigma \end{pmatrix} ,$$

where Σ and I are 2×2 block matrices

$$\Sigma = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ and } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

7. **Gathering the previous parts, state and prove a version of Jordan form for linear transformations on $\mathbb{R}^n$. Your version should be of the form "If $T : \mathbb{R}^n \to \mathbb{R}^n$ is linear then there exists a basis B such that $[T]_B^B$ has the form ..."**

   Solution.

   Theorem 0.1 (Real Jordan form)**.** *Let $T : V \to V$ be linear with $V = \mathbb{R}^n$. There exists a basis B of $\mathbb{R}^n$ such that $[T]_B^B$ is a block diagonal matrix. The blocks are of one of two types: J_σ and J_λ, the second of which is a standard Jordan block for an eigenvalue $\lambda \in \mathbb{R}$:*

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdots & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} .$$

   The first type corresponds to non-real eigenvalues σ of T (that is, non-real roots of the characteristic polynomial c_T):

$$J_\sigma = \begin{pmatrix} \Sigma & I & & & \\ & \Sigma & I & & \\ & & \cdots & & \\ & & & \Sigma & I \\ & & & & \Sigma \end{pmatrix} ,$$

where Σ *and* I *are* 2×2 *block matrices*

$$\Sigma = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ and } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $\sigma = a + bi$.

Proof. It was shown above that there is a decomposition of $\mathbb{R}^n$ into spaces

$$\mathbb{R}^n = X_1 \oplus \cdots \oplus X_r \oplus Y_1 \oplus \cdots \oplus Y_m \ .$$

These spaces are T-invariant (we proved this for the X_i's and it follows for the Y_i's from the last part, since we determined the action of T on a basis of Y_i). So if we build a basis of $\mathbb{R}^n$ as a union of bases for these spaces, our matrix relative to this basis will be block diagonal. So it suffices to prove real Jordan form for T restricted to these spaces separately.

$X_1, \ldots, X_r$ and $Y_1, \ldots, Y_m$ have the properties that, writing the real roots of c_T as $\lambda_1, \ldots, \lambda_r$ and the non-real roots as $\sigma_1, \ldots, \sigma_m$ with $\sigma_{2i} = \overline{\sigma_{2i-1}}$, $T - \lambda_j I$ is nilpotent on X_j and one can choose bases $\hat{C}_1, \ldots, \hat{C}_m$ of $Y_1, \ldots, Y_m$ such that the transformation T restricted to Y_i has a matrix representation which is block diagonal using only blocks of the form J_{σ_i}. Choosing a basis C of chains of generalized eigenvectors for $X_1 \oplus \cdots \oplus X_r$ (which is allowed by nilpotency) and setting $B = C \cup \left(\cup_{j=1}^m \hat{C}_j \right)$ gives the form stated in the theorem. $\square$