

10.4 The Real Jordan Form and the Weyr Form

The Jordan form of Section 3 exists only when p_T splits, and over $\mathbb{R}$ that often fails: a rotation of the plane has no real eigenvalue at all. But a real matrix is also a complex matrix, and its complex eigenvalues come in conjugate pairs (Theorem 9.2.5). This section turns a complex Jordan basis into a **real** one by pairing each chain with its conjugate; the price for a conjugate pair is a 2×2 rotation-scaling block instead of a scalar, and the result is the best real matrix there is. The second half rearranges the same data in another way, the Weyr form, which is the form to reach for when a second operator commutes with the first.

10.4.1 A block for a conjugate pair

The obstruction is visible in the smallest case. The rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ has $p_{\mathbf{R}} = x^2 + 1$, so $\mathbf{R}$ has no real eigenvalue and no real eigenvector (Example (A Rotation Has No Invariant Lines)). Every real matrix similar to $\mathbf{R}$ therefore has no real eigenvalue either, so it cannot be triangular: whatever basis of $\mathbb{R}^2$ we choose, the matrix of $\mathbf{R}$ keeps two off-diagonal entries. So $\mathbf{R}$ is already as simple as it will ever get over $\mathbb{R}$, and the honest thing to do is to promote it to a building block.

More generally, for $\lambda = a + bi \in \mathbb{C}$ with $b \neq 0$, consider

$$\mathbf{\Lambda} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M_2(\mathbb{R}), \quad p_{\mathbf{\Lambda}} = (x - a)^2 + b^2 = (x - (a + bi))(x - (a - bi)).$$

This matrix carries the conjugate pair $\lambda, \bar{\lambda}$ as its complex eigenvalues, has trace $2a = 2 \operatorname{Re} \lambda$ and determinant $a^2 + b^2 = |\lambda|^2$, and is a scaled rotation: it stretches by $|\lambda|$ and turns by the argument of λ . It is the real substitute for the 1×1 block (λ) , and it costs one extra dimension.

A real Jordan block is a Jordan block with the scalar λ replaced by the 2×2 rotation-scaling matrix it needs over $\mathbb{R}$, and the superdiagonal 1 replaced by $\mathbf{I}_2$.

Definition 10.4.1. Rotation-Scaling Block, Real Jordan Block

Let $\lambda = a + bi \in \mathbb{C}$ with $a, b \in \mathbb{R}$ and $b \neq 0$. The **rotation-scaling block** of the conjugate pair $\{\lambda, \bar{\lambda}\}$ is

$$\mathbf{\Lambda}(\lambda) := \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M_2(\mathbb{R}).$$

For an integer $k \geq 1$, the **real Jordan block** of size k for the pair $\{\lambda, \bar{\lambda}\}$ is the matrix $\mathbf{C}_k(\lambda) \in M_{2k}(\mathbb{R})$, partitioned into $k \times k$ blocks of size 2×2 , with $\mathbf{\Lambda}(\lambda)$ in every diagonal block, $\mathbf{I}_2$ in every **superdiagonal** block, and the zero matrix elsewhere:

$$\mathbf{C}_k(\lambda) := \begin{pmatrix} \mathbf{\Lambda}(\lambda) & \mathbf{I}_2 & & \\ & \mathbf{\Lambda}(\lambda) & \ddots & \\ & & \ddots & \mathbf{I}_2 \\ & & & \mathbf{\Lambda}(\lambda) \end{pmatrix}.$$

A **real Jordan matrix** is a direct sum of ordinary Jordan blocks $\mathbf{J}_k(\mu)$ with $\mu \in \mathbb{R}$ and real

Jordan blocks $\mathbf{C}_k(\lambda)$ with $\lambda \notin \mathbb{R}$ and $\operatorname{Im} \lambda > 0$.

In words: $\mathbf{C}_k(\lambda)$ is built exactly like $\mathbf{J}_k(\lambda)$ (Definition 10.3.1), but every entry of $\mathbf{J}_k(\lambda)$ is replaced by a 2×2 block — the diagonal λ by $\mathbf{A}(\lambda)$, the superdiagonal 1 by $\mathbf{I}_2$, and each 0 by the 2×2 zero matrix. It has size $2k$, not k . The last clause matters: $\mathbf{A}(\bar{\lambda}) = \mathbf{A}(\lambda)^\top$, so the pair $\{\lambda, \bar{\lambda}\}$ offers two candidate blocks of each size, and a real Jordan matrix is required to use the member with $b > 0$, which makes $\mathbf{A}(\lambda)$ a rotation through a positive angle. Without that normalization the uniqueness below would fail, since $\mathbf{C}_k(\lambda)$ and $\mathbf{C}_k(\bar{\lambda})$ are similar over $\mathbb{R}$.

Examples.

- **Size one.** $\mathbf{C}_1(\lambda) = \mathbf{A}(\lambda)$. For $\lambda = i$ this is the rotation $\mathbf{R}$ above; for $\lambda = 2 + 3i$ it is $\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$.
- **Size two.** For $\lambda = 1 + i$,

$$\mathbf{C}_2(1 + i) = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in M_4(\mathbb{R}).$$

Its characteristic polynomial is $((x - 1)^2 + 1)^2 = (x^2 - 2x + 2)^2$, by Theorem 7.8.4 applied to the block upper triangular shape.

- **A real Jordan matrix mixing the two kinds.** $\mathbf{J}_2(3) \oplus \mathbf{C}_1(i) \in M_4(\mathbb{R})$ has characteristic polynomial $(x - 3)^2(x^2 + 1)$.

Nonexample by minimal change. Change the sign of the lower-left entry: $\begin{pmatrix} a & -b \\ -b & a \end{pmatrix}$ is still real, still has trace $2a$, and looks just as harmless. But it is symmetric, its characteristic polynomial is $(x - a)^2 - b^2$, and its eigenvalues $a \pm b$ are **real**; for $a = 0, b = 1$ it is the coordinate swap, which is diagonalizable over $\mathbb{R}$. The clause that fails is that the two off-diagonal entries must have **opposite** signs; that is what makes the eigenvalues nonreal.

Why this definition. Two conventions are being fixed, and both are only conventions. First, the sign of b : replacing $\mathbf{A}(\lambda)$ by $\mathbf{A}(\bar{\lambda}) = \mathbf{A}(\lambda)^\top$ amounts to changing the sign of every second basis vector, so the two choices give similar matrices. Second, the ones: a real Jordan block could be written with $\mathbf{I}_2$ below the diagonal instead, giving the transpose. What is **not** a convention is the size: a conjugate pair cannot be carried by a 1×1 real block, because a 1×1 real matrix has a real eigenvalue.

The following computation is the whole idea of the section, and we will use it twice: once to see what $\mathbf{C}_k(\lambda)$ is over $\mathbb{C}$, and once to build the real basis.

Lemma 10.4.1. Realifying a Conjugate Pair of Chains

Let $\mathbf{A} \in M_n(\mathbb{R})$, let $\lambda = a + bi$ with $a, b \in \mathbb{R}$ and $b \neq 0$, and let $(\mathbf{w}_1, \dots, \mathbf{w}_k)$ be a Jordan chain of $\mathbf{A}$ for λ in $\mathbb{C}^n$, so that

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{w}_1 = \mathbf{0} \quad \text{and} \quad (\mathbf{A} - \lambda \mathbf{I})\mathbf{w}_j = \mathbf{w}_{j-1} \quad (2 \leq j \leq k).$$

Write $\mathbf{w}_j = \mathbf{u}_j - i\mathbf{v}_j$ with $\mathbf{u}_j, \mathbf{v}_j \in \mathbb{R}^n$, and put $\mathbf{u}_0 = \mathbf{v}_0 = \mathbf{0}$. Then:

- (a) $\mathbf{A}\mathbf{u}_j = a\mathbf{u}_j + b\mathbf{v}_j + \mathbf{u}_{j-1}$ and $\mathbf{A}\mathbf{v}_j = -b\mathbf{u}_j + a\mathbf{v}_j + \mathbf{v}_{j-1}$ for $1 \leq j \leq k$;

- (b) $(\overline{\mathbf{w}}_1, \dots, \overline{\mathbf{w}}_k)$ is a Jordan chain of $\mathbf{A}$ for $\overline{\lambda}$;
- (c) the $2k$ vectors $\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k$ span, over $\mathbb{C}$, the same subspace of $\mathbb{C}^n$ as the $2k$ vectors $\mathbf{w}_1, \dots, \mathbf{w}_k, \overline{\mathbf{w}}_1, \dots, \overline{\mathbf{w}}_k$, and one of the two lists is linearly independent over $\mathbb{C}$ if and only if the other is;
- (d) if the $2k$ vectors $\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k$ are linearly independent over $\mathbb{R}$, then $U = \text{span}(\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k) \subseteq \mathbb{R}^n$ is invariant under $\mathbf{A}$, and the matrix of $T_{\mathbf{A}}|_U$ in the ordered basis $(\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k)$ is $\mathbf{C}_k(\lambda)$.

Idea. For (a), write the chain relation as $\mathbf{A}\mathbf{w}_j = \lambda\mathbf{w}_j + \mathbf{w}_{j-1}$ and compare real and imaginary parts; since $\mathbf{A}$ is real, $\mathbf{A}$ acts on the two parts separately, and the only mixing comes from multiplying by $a + bi$. The minus sign in $\mathbf{w}_j = \mathbf{u}_j - i\mathbf{v}_j$ is chosen so that the answer comes out with the signs of $\mathbf{A}(\lambda)$ rather than its transpose. Part (c) is bookkeeping: $(\mathbf{u}_j, \mathbf{v}_j)$ and $(\mathbf{w}_j, \overline{\mathbf{w}}_j)$ are obtained from each other by an invertible 2×2 matrix of scalars. Part (d) just reads off the columns from (a).

Proof.

- (a) Every $\mathbf{z} \in \mathbb{C}^n$ is uniquely $\mathbf{x} + i\mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, by Definition 0.6.1 applied entry by entry; so $\mathbf{u}_j, \mathbf{v}_j$ are well defined. The chain relation says $\mathbf{A}\mathbf{w}_j = \lambda\mathbf{w}_j + \mathbf{w}_{j-1}$ for $1 \leq j \leq k$. Expanding both sides with $\mathbf{w}_j = \mathbf{u}_j - i\mathbf{v}_j$ and $\lambda = a + bi$,

$$\begin{aligned} \mathbf{A}\mathbf{u}_j - i\mathbf{A}\mathbf{v}_j &= (a + bi)(\mathbf{u}_j - i\mathbf{v}_j) + \mathbf{u}_{j-1} - i\mathbf{v}_{j-1} \\ &= (a\mathbf{u}_j + b\mathbf{v}_j + \mathbf{u}_{j-1}) + i(b\mathbf{u}_j - a\mathbf{v}_j - \mathbf{v}_{j-1}), \end{aligned}$$

where we used $i^2 = -1$ and, on the left, that $\mathbf{A}$ has real entries, so $\mathbf{A}(\mathbf{u}_j - i\mathbf{v}_j) = \mathbf{A}\mathbf{u}_j - i\mathbf{A}\mathbf{v}_j$ with $\mathbf{A}\mathbf{u}_j, \mathbf{A}\mathbf{v}_j \in \mathbb{R}^n$. Comparing real parts gives the first identity, and comparing imaginary parts gives $-\mathbf{A}\mathbf{v}_j = b\mathbf{u}_j - a\mathbf{v}_j - \mathbf{v}_{j-1}$, that is, the second.

- (b) Conjugating the identity $(\mathbf{A} - \lambda\mathbf{I})\mathbf{w}_j = \mathbf{w}_{j-1}$ entry by entry and using that $\mathbf{A}$ is real, $\overline{\mathbf{A}\mathbf{z}} = \mathbf{A}\overline{\mathbf{z}}$ and $\overline{\lambda\mathbf{z}} = \overline{\lambda}\overline{\mathbf{z}}$ for every $\mathbf{z} \in \mathbb{C}^n$ (Theorem 0.6.2 (a), (b)). Hence $(\mathbf{A} - \overline{\lambda}\mathbf{I})\overline{\mathbf{w}}_j = \overline{\mathbf{w}_{j-1}}$ for $j \geq 2$ and $(\mathbf{A} - \overline{\lambda}\mathbf{I})\overline{\mathbf{w}}_1 = \mathbf{0}$, which is the chain condition for $\overline{\lambda}$. (That $\overline{\mathbf{w}}_1 \neq \mathbf{0}$ follows from $\mathbf{w}_1 \neq \mathbf{0}$.)
- (c) Since $\mathbf{w}_j = \mathbf{u}_j - i\mathbf{v}_j$ and $\overline{\mathbf{w}}_j = \mathbf{u}_j + i\mathbf{v}_j$, we have

$$\mathbf{u}_j = \frac{1}{2}(\mathbf{w}_j + \overline{\mathbf{w}}_j), \quad \mathbf{v}_j = \frac{i}{2}(\mathbf{w}_j - \overline{\mathbf{w}}_j).$$

So each vector of either list is a $\mathbb{C}$ -combination of the other list, and the two spans agree (Theorem 1.4.2). The passage from $(\mathbf{w}_j, \overline{\mathbf{w}}_j)$ to $(\mathbf{u}_j, \mathbf{v}_j)$ is given, for each j separately, by the matrix $\begin{pmatrix} 1/2 & 1/2 \\ i/2 & -i/2 \end{pmatrix}$ of determinant $-i/2 \neq 0$, which is invertible; so the change of lists is an invertible linear substitution of the coefficients, and a nontrivial vanishing combination of one list yields one of the other. Hence one list is independent if and only if the other is.

- (d) By (a), every $\mathbf{A}\mathbf{u}_j$ and every $\mathbf{A}\mathbf{v}_j$ lies in U , so U is $\mathbf{A}$ -invariant (Lemma 9.1.1). Order the basis as $(\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k)$ and read off the columns (Definition 4.1.1). By (a), the column of $\mathbf{u}_j$ has entry a at $\mathbf{u}_j$, b at $\mathbf{v}_j$, 1 at $\mathbf{u}_{j-1}$, and 0 elsewhere; the column of $\mathbf{v}_j$ has $-b$ at $\mathbf{u}_j$, a at $\mathbf{v}_j$, 1 at $\mathbf{v}_{j-1}$, and 0 elsewhere. Grouping the basis into the consecutive pairs $(\mathbf{u}_j, \mathbf{v}_j)$, the (j, j) block is $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \mathbf{A}(\lambda)$, the $(j-1, j)$ block is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}_2, \text{ and all other blocks are zero. That is } \mathbf{C}_k(\lambda) \text{ (Definition 10.4.1).}$$

■

Applying the lemma to the block itself tells us what a real Jordan block looks like over $\mathbb{C}$.

Corollary 10.4.2. A Real Jordan Block Splits into Two Complex Ones

Let $\lambda \notin \mathbb{R}$ and $k \geq 1$. Then, as matrices over $\mathbb{C}$,

$$\mathbf{C}_k(\lambda) \sim \mathbf{J}_k(\lambda) \oplus \mathbf{J}_k(\bar{\lambda}).$$

In particular $p_{\mathbf{C}_k(\lambda)} = ((x - \lambda)(x - \bar{\lambda}))^k$, the polynomial $((x - a)^2 + b^2)^k \in \mathbb{R}[x]$, and $m_{\mathbf{C}_k(\lambda)}$ is the same polynomial.

Proof. Put $\mathbf{A} = \mathbf{C}_k(\lambda) \in M_{2k}(\mathbb{R})$ and let $(\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k)$ be the standard basis of $\mathbb{R}^{2k}$ in this order. Reading the columns of $\mathbf{C}_k(\lambda)$ gives exactly the identities of Lemma 10.4.1 (a). Put $\mathbf{w}_j = \mathbf{u}_j - i\mathbf{v}_j$, so that $\mathbf{w}_1 = \mathbf{e}_1 - i\mathbf{e}_2 \neq \mathbf{0}$. Combining the two real identities,

$$\begin{aligned} \mathbf{A}\mathbf{w}_j &= \mathbf{A}\mathbf{u}_j - i\mathbf{A}\mathbf{v}_j = (a\mathbf{u}_j + b\mathbf{v}_j + \mathbf{u}_{j-1}) - i(-b\mathbf{u}_j + a\mathbf{v}_j + \mathbf{v}_{j-1}) \\ &= (a + bi)\mathbf{u}_j + (b - ai)\mathbf{v}_j + \mathbf{w}_{j-1} \\ &= \lambda\mathbf{w}_j + \mathbf{w}_{j-1}, \end{aligned}$$

since $\lambda(-i\mathbf{v}_j) = (b - ai)\mathbf{v}_j$. So $(\mathbf{w}_1, \dots, \mathbf{w}_k)$ is a Jordan chain of $\mathbf{A}$ for λ , and by (b) so is $(\bar{\mathbf{w}}_1, \dots, \bar{\mathbf{w}}_k)$ for $\bar{\lambda}$. By (c), the $2k$ vectors $\mathbf{w}_1, \dots, \mathbf{w}_k, \bar{\mathbf{w}}_1, \dots, \bar{\mathbf{w}}_k$ are independent over $\mathbb{C}$, because the standard basis is; being $2k$ independent vectors in $\mathbb{C}^{2k}$, they form a basis (Theorem 1.6.4 (a)). On the span of one chain, $\mathbf{A} - \lambda\mathbf{I}$ has matrix $\mathbf{J}_k(0)$ by Lemma 10.2.2 (b), so $\mathbf{A}$ has matrix $\lambda\mathbf{I}_k + \mathbf{J}_k(0) = \mathbf{J}_k(\lambda)$ there, the matrix of a sum of operators being the sum of their matrices (Theorem 4.1.4, Definition 10.3.1); likewise $\mathbf{J}_k(\bar{\lambda})$ on the span of the conjugate chain. By Theorem 9.1.5 the matrix of $T_{\mathbf{A}}$ in the combined basis is $\mathbf{J}_k(\lambda) \oplus \mathbf{J}_k(\bar{\lambda})$. Hence $\mathbf{C}_k(\lambda) \sim \mathbf{J}_k(\lambda) \oplus \mathbf{J}_k(\bar{\lambda})$ over $\mathbb{C}$ (Theorem 4.2.4).

The characteristic polynomial of $\mathbf{J}_k(\mu)$ is $(x - \mu)^k$ (Theorem 7.4.6), and similar matrices have the same characteristic polynomial (Theorem 7.10.3), so $p_{\mathbf{C}_k(\lambda)} = (x - \lambda)^k(x - \bar{\lambda})^k$. For the minimal polynomial, $m_{\mathbf{J}_k(\mu)} = (x - \mu)^k$ by Corollary 10.3.2 (c), a single Jordan block being a Jordan form of itself with one block, of size k ; so $m_{\mathbf{C}_k(\lambda)} = \text{lcm}((x - \lambda)^k, (x - \bar{\lambda})^k) = (x - \lambda)^k(x - \bar{\lambda})^k$ by Proposition 9.5.6 and Theorem 9.5.5, the factors being powers of distinct irreducibles.

■

Quick check. What are the complex eigenvalues of $\mathbf{C}_2(1 + i)$, with their algebraic and geometric multiplicities?

Solution. By Corollary 10.4.2, $\mathbf{C}_2(1 + i) \sim \mathbf{J}_2(1 + i) \oplus \mathbf{J}_2(1 - i)$ over $\mathbb{C}$. So the eigenvalues are $1 \pm i$, each with $a(\lambda) = 2$ (the total size of its blocks) and $g(\lambda) = 1$ (the number of its

blocks), by Corollary 10.3.2 (a), (b).

10.4.2 The real Jordan form

Two facts about real matrices are needed before the theorem. The first is that a real matrix has the same kernel over $\mathbb{R}$ as over $\mathbb{C}$, in the sense that a real basis of the one is a complex basis of the other. Once one knows that $\tilde{U}$ is the complexification of U inside $\mathbb{C}^n = (\mathbb{R}^n)_{\mathbb{C}}$ (Definition 1.9.2, Example (The Complexification of $\mathbb{R}^n$ Is $\mathbb{C}^n$)), this is Theorem 1.9.3. It is worth proving directly, because the identification is itself part of what has to be checked, and because the statement is used both to find the real chains and to compare ranks.

Lemma 10.4.3. Real and Complex Null Spaces of a Real Matrix

Let $\mathbf{M} \in M_n(\mathbb{R})$, let $U = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{M}\mathbf{x} = \mathbf{0}\}$ be its null space over $\mathbb{R}$, and let $\tilde{U} = \{\mathbf{z} \in \mathbb{C}^n : \mathbf{M}\mathbf{z} = \mathbf{0}\}$ be its null space over $\mathbb{C}$. Then:

- (a) for $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$: $\mathbf{z} \in \tilde{U}$ if and only if $\mathbf{x} \in U$ and $\mathbf{y} \in U$;
- (b) every basis of U over $\mathbb{R}$ is a basis of $\tilde{U}$ over $\mathbb{C}$; in particular $\dim_{\mathbb{R}} U = \dim_{\mathbb{C}} \tilde{U}$ and $\text{rank } \mathbf{M}$ is the same computed over $\mathbb{R}$ and over $\mathbb{C}$.

Proof.

- (a) Since $\mathbf{M}$ is real, $\mathbf{M}\mathbf{z} = \mathbf{M}\mathbf{x} + i\mathbf{M}\mathbf{y}$ with $\mathbf{M}\mathbf{x}, \mathbf{M}\mathbf{y} \in \mathbb{R}^n$. By the uniqueness of real and imaginary parts (Definition 0.6.1), $\mathbf{M}\mathbf{z} = \mathbf{0}$ if and only if $\mathbf{M}\mathbf{x} = \mathbf{0}$ and $\mathbf{M}\mathbf{y} = \mathbf{0}$.
- (b) Let $(\mathbf{x}_1, \dots, \mathbf{x}_r)$ be a basis of U over $\mathbb{R}$. *Spanning.* Let $\mathbf{z} = \mathbf{x} + i\mathbf{y} \in \tilde{U}$. By (a), $\mathbf{x}, \mathbf{y} \in U$, so $\mathbf{x} = \sum_j c_j \mathbf{x}_j$ and $\mathbf{y} = \sum_j d_j \mathbf{x}_j$ with $c_j, d_j \in \mathbb{R}$; hence $\mathbf{z} = \sum_j (c_j + id_j) \mathbf{x}_j$. *Independence.* Let $\sum_j \alpha_j \mathbf{x}_j = \mathbf{0}$ with $\alpha_j = c_j + id_j \in \mathbb{C}$. Separating real and imaginary parts, which is legitimate because the $\mathbf{x}_j$ are real, gives $\sum_j c_j \mathbf{x}_j = \mathbf{0}$ and $\sum_j d_j \mathbf{x}_j = \mathbf{0}$; independence over $\mathbb{R}$ forces every $c_j = d_j = 0$. So $(\mathbf{x}_1, \dots, \mathbf{x}_r)$ is a basis of $\tilde{U}$ and $\dim_{\mathbb{C}} \tilde{U} = r = \dim_{\mathbb{R}} U$. The rank statement follows from Theorem 3.5.7 over each field. ■

Now the theorem. Read it as the Jordan form with one amendment: a conjugate pair of complex blocks of size k is replaced by a single real block of size $2k$.

Every real matrix is similar, over the reals, to a direct sum of Jordan blocks for its real eigenvalues and real Jordan blocks for its conjugate pairs, and that matrix is unique up to the order of its blocks.

Theorem 10.4.4. Real Jordan Form

Let $\mathbf{A} \in M_n(\mathbb{R})$ with $n \geq 1$. List the eigenvalues of $\mathbf{A}$ in $\mathbb{C}$ as

$$\begin{aligned} \mu_1, \dots, \mu_p &\in \mathbb{R} \quad (\text{the real ones, distinct}), \\ \lambda_1, \overline{\lambda_1}, \dots, \lambda_q, \overline{\lambda_q} &\quad (\text{the conjugate pairs, with } \text{Im } \lambda_t > 0). \end{aligned}$$

- (a) **(Existence)** There is an invertible $\mathbf{P} \in M_n(\mathbb{R})$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is a real Jordan matrix (Definition 10.4.1), in which the blocks for a real eigenvalue μ_l are ordinary

Jordan blocks $\mathbf{J}_k(\mu_l)$ and the blocks for a pair $\{\lambda_t, \bar{\lambda}_t\}$ are real Jordan blocks $\mathbf{C}_k(\lambda_t)$.

- (b) **(Uniqueness)** In any such matrix, the number of blocks $\mathbf{J}_k(\mu)$ with $k \geq j$, for $\mu \in \mathbb{R}$, and the number of blocks $\mathbf{C}_k(\lambda)$ with $k \geq j$, for $\lambda \notin \mathbb{R}$, both equal

$$\text{rank}(\mathbf{A} - \nu \mathbf{I})^{j-1} - \text{rank}(\mathbf{A} - \nu \mathbf{I})^j \quad (\nu = \mu \text{ resp. } \nu = \lambda),$$

the ranks being computed over $\mathbb{C}$ in the second case. Consequently the real Jordan matrix is determined by $\mathbf{A}$ up to the order of its blocks.

Idea. Work over $\mathbb{C}$, then come home. ① Split $\mathbb{C}^n = \bigoplus G_\nu(\mathbf{A})$ over all complex eigenvalues ν (Theorem 10.1.3). ② For a **real** eigenvalue, the generalized eigenspace is the complexification of a real subspace, by Lemma 10.4.3, and Section 2's structure theorem runs over $\mathbb{R}$ on that real subspace: nothing complex is needed. ③ For a **pair**, choose the chains in G_{λ_t} freely, and take their conjugates in $G_{\bar{\lambda}_t}$ — no choice there. ④ Replace each conjugate pair of chains by its real and imaginary parts; Lemma 10.4.1 says the new matrix is $\mathbf{C}_k(\lambda_t)$, and a count shows the real vectors collected are exactly n in number and independent. Uniqueness is the uniqueness of the complex Jordan form, since by Corollary 10.4.2 a real Jordan matrix has a known complex Jordan form.

Proof. Since $p_{\mathbf{A}} \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$ has degree $n \geq 1$, it splits over $\mathbb{C}$ (Corollary 6.5.5), and its nonreal roots come in conjugate pairs with equal multiplicities (Theorem 9.2.5 (a)); so the list of eigenvalues has the displayed shape. Write $a(\nu)$ for the algebraic multiplicity of ν in $p_{\mathbf{A}}$, so $a(\lambda_t) = a(\bar{\lambda}_t)$ and

$$\sum_{l=1}^p a(\mu_l) + \sum_{t=1}^q 2a(\lambda_t) = n. \quad (*)$$

By Theorem 10.1.3 applied to $T_{\mathbf{A}}$ over $\mathbb{C}$,

$$\mathbb{C}^n = \bigoplus_{l=1}^p G_{\mu_l}(\mathbf{A}) \oplus \bigoplus_{t=1}^q (G_{\lambda_t}(\mathbf{A}) \oplus G_{\bar{\lambda}_t}(\mathbf{A})),$$

with $\dim_{\mathbb{C}} G_\nu(\mathbf{A}) = a(\nu)$ and $G_\nu(\mathbf{A}) = \ker(\mathbf{A} - \nu \mathbf{I})^n$ (Corollary 10.1.2 (a)).

Step 1. Real eigenvalues, handled entirely over $\mathbb{R}$. Fix a real eigenvalue $\mu = \mu_l$. The matrix $(\mathbf{A} - \mu \mathbf{I})^n$ is real, so by Lemma 10.4.3 its real null space

$$U_\mu := \{ \mathbf{x} \in \mathbb{R}^n : (\mathbf{A} - \mu \mathbf{I})^n \mathbf{x} = \mathbf{0} \}$$

has a real basis which is a $\mathbb{C}$ -basis of $G_\mu(\mathbf{A})$; in particular $\dim_{\mathbb{R}} U_\mu = a(\mu)$. The subspace U_μ is $\mathbf{A}$ -invariant, since $(\mathbf{A} - \mu \mathbf{I})^n (\mathbf{A} \mathbf{x}) = \mathbf{A} (\mathbf{A} - \mu \mathbf{I})^n \mathbf{x} = \mathbf{0}$ for $\mathbf{x} \in U_\mu$ (Theorem 6.7.4 (b)), and $N_\mu := (\mathbf{A} - \mu \mathbf{I})|_{U_\mu}$ satisfies $N_\mu^n = 0$, so N_μ is nilpotent (Definition 10.2.1). By Theorem 10.2.3 there is a basis $\mathcal{B}_\mu$ of the **real** space U_μ , a concatenation of Jordan chains of N_μ , with $[N_\mu]_{\mathcal{B}_\mu}^{\mathcal{B}_\mu} = \mathbf{J}_{k_1}(0) \oplus \cdots \oplus \mathbf{J}_{k_s}(0)$. Since $T_{\mathbf{A}}|_{U_\mu} = \mu \text{id}_{U_\mu} + N_\mu$ and $[\text{id}_{U_\mu}]_{\mathcal{B}_\mu}^{\mathcal{B}_\mu} = \mathbf{I}_{k_1} \oplus \cdots \oplus \mathbf{I}_{k_s}$, adding the matrices (Theorem 4.1.4) blockwise (Theorem 8.1.4 (a), Definition 10.3.1) gives

$$[T_{\mathbf{A}}|_{U_\mu}]_{\mathcal{B}_\mu}^{\mathcal{B}_\mu} = \mathbf{J}_{k_1}(\mu) \oplus \cdots \oplus \mathbf{J}_{k_s}(\mu),$$

a direct sum of ordinary Jordan blocks. Moreover $\mathcal{B}_\mu$, viewed in $\mathbb{C}^n$, is a Jordan basis of $G_\mu(\mathbf{A})$ for $T_{\mathbf{A}}$ over $\mathbb{C}$: the chain relations are the same equations, and $\mathcal{B}_\mu$ is a $\mathbb{C}$ -basis of $G_\mu(\mathbf{A})$ by Lemma 10.4.3 (b).

Step 2. Conjugate pairs, chosen on one side. Fix t and write $\lambda = \lambda_t$, $G = G_\lambda(\mathbf{A})$, $\overline{G} = G_{\overline{\lambda}}(\mathbf{A})$. By Theorem 10.1.3 (d), $(\mathbf{A} - \lambda\mathbf{I})|_G$ is nilpotent, so Theorem 10.2.3 provides chains whose concatenation is a $\mathbb{C}$ -basis of G ; as in Step 1 these are Jordan chains of $\mathbf{A}$ for λ , of lengths $k_1, \dots, k_m$ summing to $a(\lambda)$.

Claim: conjugating gives a Jordan basis of $\overline{G}$. Indeed, conjugation maps G onto $\overline{G}$: if $(\mathbf{A} - \lambda\mathbf{I})^n \mathbf{z} = \mathbf{0}$, then $(\mathbf{A} - \overline{\lambda}\mathbf{I})^n \overline{\mathbf{z}} = \mathbf{0}$ by the computation in Lemma 10.4.1 (b), and conjugating twice returns $\mathbf{z}$. If $(\mathbf{z}_1, \dots, \mathbf{z}_d)$ is a $\mathbb{C}$ -basis of G , then $(\overline{\mathbf{z}}_1, \dots, \overline{\mathbf{z}}_d)$ spans $\overline{G}$ — any $\mathbf{z} \in \overline{G}$ has $\overline{\mathbf{z}} \in G$, say $\overline{\mathbf{z}} = \sum_j c_j \mathbf{z}_j$, whence $\mathbf{z} = \sum_j \overline{c_j} \overline{\mathbf{z}}_j$ — and it has $d = \dim_{\mathbb{C}} \overline{G}$ entries, so it is a basis (Theorem 1.6.4 (b)). Applying this to the concatenated chains, and using Lemma 10.4.1 (b) chain by chain, the conjugated chains form a Jordan basis of $\overline{G}$ consisting of chains for $\overline{\lambda}$ of the same lengths $k_1, \dots, k_m$.

Step 3. Realify each pair of conjugate chains. Keep t fixed and let $(\mathbf{w}_1, \dots, \mathbf{w}_k)$ be one of the chains of Step 2, with $\mathbf{w}_j = \mathbf{u}_j - i\mathbf{v}_j$, $\mathbf{u}_j, \mathbf{v}_j \in \mathbb{R}^n$. By Lemma 10.4.1 (c), the $2k$ real vectors $\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{u}_k, \mathbf{v}_k$ have the same $\mathbb{C}$ -span as the $2k$ chain vectors $\mathbf{w}_1, \dots, \mathbf{w}_k, \overline{\mathbf{w}}_1, \dots, \overline{\mathbf{w}}_k$. Do this for each of the m chains attached to λ_t , and for each t .

Step 4. The collected real vectors form a basis of $\mathbb{R}^n$. Let $\mathcal{B}$ be the list obtained by writing, one after another: the real Jordan bases $\mathcal{B}_{\mu_1}, \dots, \mathcal{B}_{\mu_p}$ of Step 1, and then, for each t and each chain of Step 2, the block of $2k$ real vectors of Step 3. Its length is

$$\sum_{l=1}^p a(\mu_l) + \sum_{t=1}^q \sum_{\text{chains}} 2k = \sum_{l=1}^p a(\mu_l) + \sum_{t=1}^q 2a(\lambda_t) = n$$

by (*), since the chain lengths for λ_t sum to $a(\lambda_t)$. Its $\mathbb{C}$ -span contains a $\mathbb{C}$ -basis of each $G_{\mu_l}(\mathbf{A})$ (Step 1) and, for each t , all the vectors $\mathbf{w}_j$ and $\overline{\mathbf{w}}_j$ (Step 3), which together form a $\mathbb{C}$ -basis of $G_{\lambda_t}(\mathbf{A}) \oplus G_{\overline{\lambda}_t}(\mathbf{A})$ by Step 2. So the $\mathbb{C}$ -span of $\mathcal{B}$ is all of $\mathbb{C}^n$, and $\mathcal{B}$ has exactly n entries, so $\mathcal{B}$ is a $\mathbb{C}$ -basis of $\mathbb{C}^n$ (Theorem 1.6.4 (b)). In particular $\mathcal{B}$ is linearly independent over $\mathbb{C}$, hence over $\mathbb{R}$, and being a list of n independent vectors in $\mathbb{R}^n$ it is a basis of $\mathbb{R}^n$ (Theorem 1.6.4 (a)).

Step 5. The matrix. Each segment of $\mathcal{B}$ spans an $\mathbf{A}$ -invariant real subspace: the spaces U_{μ_l} by Step 1, and the spaces U of Lemma 10.4.1 (d) by Step 3 and the independence just proved. Their direct sum is $\mathbb{R}^n$, since $\mathcal{B}$ is a basis. By Theorem 9.1.5, $[T_{\mathbf{A}}]_{\mathcal{B}}^{\mathcal{B}}$ is block diagonal with the matrices of the restrictions as diagonal blocks, which are $\mathbf{J}_k(\mu_l)$'s by Step 1 and $\mathbf{C}_k(\lambda_t)$'s by Lemma 10.4.1 (d). Taking $\mathbf{P}$ to be the real matrix whose columns are the vectors of $\mathcal{B}$ gives $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = [T_{\mathbf{A}}]_{\mathcal{B}}^{\mathcal{B}}$ (Theorem 4.2.4), which proves (a).

- (b) Let $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ be any real Jordan matrix similar to $\mathbf{A}$ over $\mathbb{R}$, and regard the similarity over $\mathbb{C}$. By Corollary 10.4.2 and Theorem 8.1.4, replacing each block $\mathbf{C}_k(\lambda)$ of $\mathbf{B}$ by $\mathbf{J}_k(\lambda) \oplus \mathbf{J}_k(\overline{\lambda})$ produces a Jordan matrix similar to $\mathbf{B}$, hence to $\mathbf{A}$, over $\mathbb{C}$. By Theorem 10.3.1 (b), for every $\nu \in \mathbb{C}$ and $j \geq 1$ the number of blocks $\mathbf{J}_k(\nu)$ with $k \geq j$ in that Jordan matrix equals $\text{rank}(\mathbf{A} - \nu\mathbf{I})^{j-1} - \text{rank}(\mathbf{A} - \nu\mathbf{I})^j$, computed over $\mathbb{C}$. For a real μ , the blocks $\mathbf{J}_k(\mu)$ of the Jordan matrix are exactly the blocks $\mathbf{J}_k(\mu)$ of $\mathbf{B}$, since the replaced blocks contribute only nonreal eigenvalues; and the ranks over $\mathbb{C}$ agree with the ranks over $\mathbb{R}$ by Lemma 10.4.3 (b), applied to the real matrix $(\mathbf{A} - \mu\mathbf{I})^j$. For $\lambda \notin \mathbb{R}$ with $\text{Im } \lambda > 0$, the blocks $\mathbf{J}_k(\lambda)$ of the Jordan matrix come one each from the blocks $\mathbf{C}_k(\lambda)$ of $\mathbf{B}$: a block $\mathbf{C}_k(\lambda')$ of $\mathbf{B}$ with $\lambda' \neq \lambda$ contributes $\mathbf{J}_k(\lambda')$ and $\mathbf{J}_k(\overline{\lambda}')$, and $\overline{\lambda}' \neq \lambda$ because λ and λ' both have **positive** imaginary part (Definition 10.4.1). In both cases the count is determined by $\mathbf{A}$, so any two real Jordan matrices similar to $\mathbf{A}$ have the same blocks up to order.

So over $\mathbb{R}$ the dream of Chapter 2 comes true in the amended form: **every real square matrix is similar over $\mathbb{R}$ to a matrix that is block diagonal with blocks $\mathbf{J}_k(\mu)$ and $\mathbf{C}_k(\lambda)$** , and the only new feature is the 2×2 rotation-scaling block. Counting: a real eigenvalue μ contributes $a(\mu)$ to the size and $g(\mu)$ blocks; a pair $\{\lambda, \bar{\lambda}\}$ contributes $2a(\lambda)$ to the size and $g(\lambda)$ real blocks. For an operator T on a real vector space V the same statement holds, applied to $[T]_{\mathcal{C}}^{\mathcal{C}}$ for any basis $\mathcal{C}$ of V : the theorem produces a real Jordan matrix similar to it, hence equal to $[T]_{\mathcal{B}}^{\mathcal{B}}$ for another basis $\mathcal{B}$ of V (Theorem 4.2.4). The complexification $V_{\mathbb{C}}$ of Definition 1.9.2 is where the complex chains live; for $V = \mathbb{R}^n$ it is $\mathbb{C}^n$ (Example (The Complexification of $\mathbb{R}^n$ Is $\mathbb{C}^n$)), which is the case written out above.

Warning. A real Jordan matrix is not triangular, and the real form is not obtained by “ignoring” the imaginary parts. The real Jordan form of a matrix with a nonreal eigenvalue **cannot** be triangular, since a triangular real matrix has its diagonal entries as eigenvalues, all real (Theorem 9.7.4). Nor is $\mathbf{\Lambda}(\lambda)$ a diagonal matrix in disguise: it has no real eigenvector. And the real basis is not the real part of the complex one: each complex chain vector $\mathbf{w}_j$ contributes **two** real vectors, $\operatorname{Re} \mathbf{w}_j$ and $-\operatorname{Im} \mathbf{w}_j$, and dropping either of them destroys the invariance.

Example. The Real Jordan Form of a 4×4 Matrix

Let

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & 1 & 2 \end{pmatrix} \in M_4(\mathbb{R}), \quad p_{\mathbf{A}} = (x^2 - 2x + 2)^2.$$

Find the real Jordan form of $\mathbf{A}$ (Theorem 10.4.4) and an invertible $\mathbf{P} \in M_4(\mathbb{R})$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ in that form.

Solution. The eigenvalues. $x^2 - 2x + 2 = (x - 1)^2 + 1$ has roots $1 \pm i$, so over $\mathbb{C}$ the eigenvalues are $\lambda = 1 + i$ and $\bar{\lambda} = 1 - i$, each with $a(\lambda) = 2$. There is no real eigenvalue, so the real Jordan form consists of real blocks $\mathbf{C}_k(1 + i)$ only, with sizes k summing to $a(\lambda) = 2$.

A chain for $\lambda = 1 + i$. Put $\mathbf{M} = \mathbf{A} - (1 + i)\mathbf{I}$, so that

$$\mathbf{M} = \begin{pmatrix} -1 - i & -1 & 1 & 0 \\ 1 & -1 - i & 0 & 1 \\ -1 & 0 & 1 - i & -1 \\ 0 & -1 & 1 & 1 - i \end{pmatrix}.$$

Solve $\mathbf{M}\mathbf{z} = \mathbf{0}$ using rows 2, 1, and 4 in that order. Row 2 gives $z_1 = (1 + i)z_2 - z_4$; row 1 gives $z_3 = (1 + i)z_1 + z_2$. Substituting the first into the second, and writing $s = z_2$, $u = z_4$,

$$z_1 = (1 + i)s - u, \quad z_3 = (1 + i)((1 + i)s - u) + s = (1 + 2i)s - (1 + i)u,$$

using $(1+i)^2 = 2i$. Row 4 reads $-z_2 + z_3 + (1-i)z_4 = 0$, which becomes

$$-s + (1+2i)s - (1+i)u + (1-i)u = 2is - 2iu = 0,$$

so $s = u$. Row 3 is then automatic: with $s = u = 1$ we get $\mathbf{z} = (i, 1, i, 1)$, and row 3 gives $-i + (1-i)i - 1 = -i + i + 1 - 1 = 0$. So

$$\ker \mathbf{M} = \text{span}((i, 1, i, 1)), \quad g(\lambda) = 1.$$

Hence there is exactly one λ -block over $\mathbb{C}$ (Corollary 10.3.2 (b)), of size $a(\lambda) = 2$, and the real Jordan form is $\mathbf{C}_2(1+i)$. For the basis we need the chain: take $\mathbf{w}_1 = (i, 1, i, 1)$ and solve $\mathbf{M}\mathbf{w}_2 = \mathbf{w}_1$. Try $\mathbf{w}_2 = (0, 0, z_3, z_4)$, so that only columns 3 and 4 of $\mathbf{M}$ are used. Row 1 then reads $z_3 = i$ and row 2 reads $z_4 = 1$, so $\mathbf{w}_2 = (0, 0, i, 1)$ is the only candidate of this shape. Rows 3 and 4 confirm it:

$$(1-i)i - 1 = (i+1) - 1 = i \quad \text{and} \quad i + (1-i) = 1,$$

which are the third and fourth entries of $\mathbf{w}_1$. So $(\mathbf{w}_1, \mathbf{w}_2)$ is a Jordan chain of $\mathbf{A}$ for $1+i$.

Normalizing the chain. If $(\mathbf{w}_1, \mathbf{w}_2)$ is a chain, so is $(c\mathbf{w}_1, c\mathbf{w}_2)$ for any $c \neq 0$, since the chain relations are linear. Taking $c = -i$ clears the denominators of the realification:

$$\mathbf{w}'_1 = -i(i, 1, i, 1) = (1, -i, 1, -i), \quad \mathbf{w}'_2 = -i(0, 0, i, 1) = (0, 0, 1, -i).$$

Now read off $\mathbf{w}'_j = \mathbf{u}_j - i\mathbf{v}_j$ with $\mathbf{u}_j, \mathbf{v}_j$ real, as Lemma 10.4.1 requires:

$$\mathbf{u}_1 = (1, 0, 1, 0), \quad \mathbf{v}_1 = (0, 1, 0, 1), \quad \mathbf{u}_2 = (0, 0, 1, 0), \quad \mathbf{v}_2 = (0, 0, 0, 1).$$

The answer. With $\mathcal{B} = (\mathbf{u}_1, \mathbf{v}_1, \mathbf{u}_2, \mathbf{v}_2)$,

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{C}_2(1+i) = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

$\mathbf{P}$ is lower triangular with ones on the diagonal, so $\det \mathbf{P} = 1$ (Theorem 7.4.6) and $\mathbf{P}$ is invertible.

The check, column by column. Writing $\mathbf{c}_1, \dots, \mathbf{c}_4$ for the columns of $\mathbf{P}$, the identity $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{C}_2(1+i)$ says

$$\begin{aligned} \mathbf{A}\mathbf{c}_1 &= \mathbf{c}_1 + \mathbf{c}_2, \\ \mathbf{A}\mathbf{c}_2 &= -\mathbf{c}_1 + \mathbf{c}_2, \\ \mathbf{A}\mathbf{c}_3 &= \mathbf{c}_1 + \mathbf{c}_3 + \mathbf{c}_4, \\ \mathbf{A}\mathbf{c}_4 &= \mathbf{c}_2 - \mathbf{c}_3 + \mathbf{c}_4, \end{aligned}$$

read off the four columns of $\mathbf{C}_2(1+i)$; these are exactly the identities of Lemma 10.4.1 (a) with $a = b = 1$. Two of them in full:

$$\begin{aligned} \mathbf{A}\mathbf{c}_1 &= \mathbf{A}(1, 0, 1, 0) = (0+1, 1+0, -1+2, 0+1) = (1, 1, 1, 1) \\ &= (1, 0, 1, 0) + (0, 1, 0, 1) = \mathbf{c}_1 + \mathbf{c}_2, \end{aligned}$$

adding columns 1 and 3 of $\mathbf{A}$; and

$$\begin{aligned}\mathbf{A}\mathbf{c}_3 &= \mathbf{A}(0, 0, 1, 0) = (1, 0, 2, 1) = (1, 0, 1, 0) + (0, 0, 1, 0) + (0, 0, 0, 1) \\ &= \mathbf{c}_1 + \mathbf{c}_3 + \mathbf{c}_4,\end{aligned}$$

the third column of $\mathbf{A}$. The remaining two are the same kind of one-line computation.

Cross-checks. $\text{tr } \mathbf{A} = 0 + 0 + 2 + 2 = 4$ and $\text{tr } \mathbf{C}_2(1 + i) = 4 \cdot 1 = 4$. Over $\mathbb{C}$, [Corollary 10.4.2](#) gives the complex Jordan form $\mathbf{J}_2(1 + i) \oplus \mathbf{J}_2(1 - i)$, whose characteristic polynomial is $((x - 1)^2 + 1)^2 = p_{\mathbf{A}}$.

10.4.3 The Weyr form

The Jordan form of a nilpotent operator is the Young diagram of Section 2 read **by rows**: one row per chain, one block per row. The same diagram can be read **by columns**, and the columns are the jumps of the kernel chain (Definition 10.2.3 and the discussion after Theorem 10.2.4). Reordering a Jordan basis accordingly costs nothing and gives a second normal form, the Weyr form. It carries exactly the same information — a partition for each eigenvalue — but it wears it differently, and for one purpose it is strictly better: the matrices that commute with a Weyr matrix are block upper triangular, which is false for the Jordan form.

Here is the diagram for the partition $(2, 2, 1)$ of 5, whose dual partition is $(3, 2)$:

$$\begin{array}{l} w_1 = 3, w_2 = 2 \\ k_1 = 2 \\ k_2 = 2 \\ k_3 = 1 \end{array} \quad \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \\ \hline \end{array} \quad \begin{array}{l} \text{rows: the Jordan block sizes } (2, 2, 1) \\ \text{columns: the Weyr structure } (3, 2) \\ \text{the operator moves each box one step left} \end{array}$$

The basis is the same; only its order changes. Instead of listing chain by chain, list the first column (a basis of the kernel), then the second column, and so on. Since the operator moves each box one step to the left, it maps the j -th group of basis vectors into the $(j - 1)$ -th, injectively, and the matrix becomes block upper triangular with a very rigid pattern of ones.

Definition 10.4.2. Basic Weyr Matrix, Weyr Structure, Weyr Form

Let $\lambda \in F$ and let $w_1 \geq w_2 \geq \cdots \geq w_s \geq 1$ be a partition of an integer $m \geq 1$ (Definition 10.2.3). The **basic Weyr matrix** with eigenvalue λ and **Weyr structure** $(w_1, \dots, w_s)$ is the matrix $\mathbf{W} \in M_m(F)$, partitioned into blocks according to $m = w_1 + \cdots + w_s$, whose blocks are

$$\mathbf{W}_{jj} = \lambda \mathbf{I}_{w_j}, \quad \mathbf{W}_{j,j+1} = \begin{pmatrix} \mathbf{I}_{w_{j+1}} \\ \mathbf{O} \end{pmatrix} \in M_{w_j \times w_{j+1}}(F), \quad \mathbf{W}_{jl} = \mathbf{O} \text{ otherwise.}$$

A **Weyr matrix** is a direct sum of basic Weyr matrices. A **Weyr form** of $T \in \mathcal{L}(V)$ is a Weyr matrix $\mathbf{W}$ with $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{W}$ for some basis $\mathcal{B}$ of V ; for $\mathbf{A} \in M_n(F)$, a Weyr form of $\mathbf{A}$ is a Weyr matrix similar to $\mathbf{A}$.

In words: the diagonal blocks are scalar, of the sizes given by the partition, and the only other nonzero blocks sit **immediately to the right** of the diagonal ones, each being an identity matrix stacked on top of a zero block — so $\mathbf{W}_{j,j+1}$ has full column rank w_{j+1} , which is possible

exactly because $w_j \geq w_{j+1}$. The matrix is upper triangular, with λ repeated m times on the diagonal.

Examples.

- **One column.** If $s = 1$, then $\mathbf{W} = \lambda \mathbf{I}_{w_1}$: a scalar matrix. This is the diagonalizable case.
- **All parts equal to 1.** If $w_1 = \cdots = w_s = 1$, then $\mathbf{W} = \mathbf{J}_s(\lambda)$: a single Jordan block. So Jordan blocks are the basic Weyr matrices of the thinnest possible structure.
- **The structure $(3, 2)$ with $\lambda = 5$.** Here $m = 5$ and

$$\mathbf{W} = \begin{pmatrix} 5 & 0 & 0 & 1 & 0 \\ 0 & 5 & 0 & 0 & 1 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix},$$

with the 3×2 block $\begin{pmatrix} \mathbf{I}_2 \\ \mathbf{O} \end{pmatrix}$ in the upper right corner. As we will see, this is the Weyr form of $\mathbf{J}_2(5) \oplus \mathbf{J}_2(5) \oplus \mathbf{J}_1(5)$, the diagram drawn above.

Nonexample by minimal change. Put the identity block in the **lower** part of $\mathbf{W}_{12}$ instead: with structure $(2, 1)$ and $\lambda = 0$, compare

$$\mathbf{W} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{W}' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Both are nilpotent of rank 1, and indeed $\mathbf{W} \sim \mathbf{W}'$ (Corollary 10.2.5): the shape is not an extra invariant, and the two matrices represent the same operator in different bases. But only $\mathbf{W}$ is a Weyr matrix. The clause that fails for $\mathbf{W}'$ is that the block $\mathbf{W}_{j,j+1}$ must be an identity matrix stacked **on top of** a zero block, not below it. The ordering is a normalization, and it is the one the commutant property below is proved with: there the block $\mathbf{W}_{j,j+1}$ is canceled on the left, which uses that its columns are the **first** w_{j+1} standard basis vectors, in order. Concretely, every matrix commuting with $\mathbf{W}$ turns out to be upper triangular, while the commutant of $\mathbf{W}'$ contains the matrix unit $\mathbf{E}_{21}$.

The existence and uniqueness of the form are now a repackaging of Section 3.

Theorem 10.4.5. Weyr Form

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, and let $T \in \mathcal{L}(V)$.

- (**Existence**) T has a Weyr form if and only if p_T splits over F . One can be arranged with one basic Weyr block per eigenvalue.
- (**Uniqueness**) In any Weyr form of T with one basic block per eigenvalue, the block for λ has Weyr structure

$$w_j = \dim \ker(T - \lambda \operatorname{id}_V)^j - \dim \ker(T - \lambda \operatorname{id}_V)^{j-1} \quad (j \geq 1),$$

the nonzero terms of this sequence; equivalently $(w_1, w_2, \dots)$ is the **dual partition** of the partition formed by the sizes of the λ -blocks of the Jordan form of T . So the Weyr form is determined by T up to the order of its basic blocks.

Idea. Both parts come from one observation: a Weyr basis is a Jordan basis in a different order. Arrange the λ -chains of a Jordan basis as the rows of the Young diagram, sorted by decreasing length, and read the diagram **column by column**. The operator $T - \lambda \text{id}$ moves each box one step to the left, so it carries the j -th column into the $(j - 1)$ -th, matching the i -th box of one with the i -th box of the other — which is exactly the block $\begin{pmatrix} \mathbf{I} \\ \mathbf{O} \end{pmatrix}$. For uniqueness, count the boxes: the first j columns of the diagram are a basis of $\ker(T - \lambda \text{id})^j$, so the column lengths are the jumps of the kernel chain.

Proof. ($\Rightarrow$) A Weyr matrix is upper triangular, so if T has a Weyr form then T is triangularizable and p_T splits (Definition 9.7.1, Theorem 9.7.2).

($\Leftarrow$) Suppose p_T splits, and let λ be an eigenvalue. By Theorem 10.3.1 there is a Jordan basis; fix its part inside $G_\lambda(T)$, a concatenation of Jordan chains of $N = T - \lambda \text{id}_V$ of lengths $k_1 \geq k_2 \geq \dots \geq k_r \geq 1$ summing to $a(\lambda)$ (Theorem 10.2.3, after reindexing so that the lengths decrease). Write $b_{i,1}, \dots, b_{i,k_i}$ for the i -th chain in the order of Definition 10.2.2, so that

$$Nb_{i,1} = 0, \quad Nb_{i,j} = b_{i,j-1} \quad (2 \leq j \leq k_i).$$

Put $s = k_1$ and, for $1 \leq j \leq s$, let $w_j = \#\{i : k_i \geq j\}$, the length of the j -th column of the Young diagram; then $w_1 \geq \dots \geq w_s \geq 1$ is the dual partition of $(k_1, \dots, k_r)$ (Definition 10.2.3), a partition of $a(\lambda)$. Because the k_i decrease, $\{i : k_i \geq j\} = \{1, 2, \dots, w_j\}$. Let

$$\mathcal{C}_j = (b_{1,j}, b_{2,j}, \dots, b_{w_j,j})$$

be the j -th column, and let $\mathcal{B}_\lambda$ be the list $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_s$ written one after another: a reordering of the chosen basis of $G_\lambda(T)$, hence a basis of it.

Read the columns of $[T|_{G_\lambda}]_{\mathcal{B}_\lambda}^{\mathcal{B}_\lambda}$. For $j = 1$, $Tb_{i,1} = \lambda b_{i,1}$. For $j \geq 2$, $Tb_{i,j} = \lambda b_{i,j} + b_{i,j-1}$, and $b_{i,j-1}$ is the i -th entry of $\mathcal{C}_{j-1}$, which exists because $i \leq w_j \leq w_{j-1}$. In the block partition of $\mathcal{B}_\lambda$ by the groups $\mathcal{C}_1, \dots, \mathcal{C}_s$, this says: the (j, j) block is $\lambda \mathbf{I}_{w_j}$; the $(j - 1, j)$ block sends the i -th basis vector of group j to the i -th of group $j - 1$ for $i = 1, \dots, w_j$, so it is $\begin{pmatrix} \mathbf{I}_{w_j} \\ \mathbf{O} \end{pmatrix}$; and every other block is $\mathbf{O}$. That is the basic Weyr matrix with eigenvalue λ and structure $(w_1, \dots, w_s)$ (Definition 10.4.2). Writing the lists $\mathcal{B}_\lambda$ for the distinct eigenvalues one after another and using $V = \bigoplus_\lambda G_\lambda(T)$ (Theorem 10.1.3) with Theorem 9.1.5, the resulting basis $\mathcal{B}$ makes $[T]_{\mathcal{B}}^{\mathcal{B}}$ a direct sum of basic Weyr matrices, one per eigenvalue.

- (b) Let $\mathbf{W}$ be a Weyr form of T with one basic block per eigenvalue, and fix λ with basic block of structure $(w_1, \dots, w_s)$ and size m . Let $\mathcal{B}$ be a basis with $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{W}$, and name the basis vectors inside the λ -block $c_{i,j}$, where j indexes the groups and $i = 1, \dots, w_j$ the entries within a group. The blocks of Definition 10.4.2 say precisely that

$$Nc_{i,1} = 0, \quad Nc_{i,j} = c_{i,j-1} \quad (j \geq 2), \quad N = T - \lambda \text{id}_V,$$

while on the span V' of the basis vectors belonging to the other basic blocks the operator N is invertible: V' is T -invariant and the matrix of $N|_{V'}$ is $\mathbf{W}' - \lambda \mathbf{I}$, where $\mathbf{W}'$ is a direct sum of basic Weyr matrices with eigenvalues $\mu \neq \lambda$, hence upper triangular with all diagonal entries $\mu - \lambda \neq 0$, so of nonzero determinant (Theorem 7.4.6, Theorem 7.5.2). Hence N^j kills $c_{i,l}$ for $l \leq j$ and sends $c_{i,l} \mapsto c_{i,l-j}$ for $l > j$. Write $V = U \oplus V'$ with $U = \text{span}(c_{i,l})$; both summands are N -invariant, so $\ker N^j = (\ker N^j \cap U) \oplus (\ker N^j \cap V')$,

and the second summand is $\{0\}$ because $N^j|_{V'}$ is invertible. Inside U the displayed action gives

$$\ker N^j = \text{span}(c_{i,l} : l \leq j), \quad \dim \ker N^j = w_1 + \cdots + w_j$$

for $0 \leq j \leq s$, the spanning vectors being members of the basis $\mathcal{B}$ and hence independent; and $\ker N^j = \ker N^s$ for $j \geq s$. Subtracting consecutive values gives $w_j = \dim \ker N^j - \dim \ker N^{j-1}$ as claimed, a number depending only on T and λ . By Corollary 10.3.2 (d) this is also $\#\{i : k_i \geq j\}$ for the λ -block sizes k_i of the Jordan form, that is, the dual partition. Since the structures are determined, so is $\mathbf{W}$ up to the order in which its basic blocks are listed. ■

Existence and uniqueness are thus free; what makes the form worth the trouble is the next property. Here $\mathbf{K} \in M_n(F)$ is called **block upper triangular** for a partition $n = n_1 + \cdots + n_s$, taken as both the row and the column partition (Definition 8.1.1), when every block $\mathbf{K}_{pq}$ with $p > q$ is zero; for $s = 2$ this is the shape of Chapter 8.

Proposition 10.4.6. The Commutant of a Basic Weyr Matrix Is Block Triangular

Let $\mathbf{W} \in M_m(F)$ be a basic Weyr matrix with eigenvalue λ and Weyr structure $(w_1, \dots, w_s)$, and let $\mathbf{K} \in M_m(F)$ satisfy $\mathbf{WK} = \mathbf{KW}$. Then $\mathbf{K}$ is block upper triangular for the partition $m = w_1 + \cdots + w_s$: its blocks $\mathbf{K}_{pq}$ vanish for $p > q$.

Idea. Write both products blockwise. Because $\mathbf{W}$ has only one nonzero block in each block row, each block equation is short: $\mathbf{E}_p \mathbf{K}_{p+1,q} = \mathbf{K}_{p,q-1} \mathbf{E}_{q-1}$, where $\mathbf{E}_j = \begin{pmatrix} \mathbf{I}_{w_{j+1}} \\ \mathbf{O} \end{pmatrix}$. Every $\mathbf{E}_j$ has full column rank, so it can be canceled on the left. Now sweep the block columns from left to right: the first block column forces $\mathbf{K}_{p1} = \mathbf{O}$ for $p > 1$, and each new column is forced by the previous one.

Proof. Put $\mathbf{N} = \mathbf{W} - \lambda \mathbf{I}_m$ and $\mathbf{E}_j = \begin{pmatrix} \mathbf{I}_{w_{j+1}} \\ \mathbf{O} \end{pmatrix} \in M_{w_j \times w_{j+1}}(F)$ for $1 \leq j \leq s-1$, so that $\mathbf{N}$ has blocks $\mathbf{N}_{j,j+1} = \mathbf{E}_j$ and all other blocks zero (Definition 10.4.2). Since $\lambda \mathbf{I}_m$ commutes with everything, $\mathbf{WK} = \mathbf{KW}$ is equivalent to $\mathbf{NK} = \mathbf{KN}$. Comparing the (p, q) blocks with Theorem 8.1.1, and reading an undefined symbol as $\mathbf{O}$,

$$\begin{aligned} \mathbf{E}_p \mathbf{K}_{p+1,q} &= (\mathbf{NK})_{pq} = (\mathbf{KN})_{pq} = \mathbf{K}_{p,q-1} \mathbf{E}_{q-1} \\ &\text{for all } 1 \leq p, q \leq s. \end{aligned} \tag{*}$$

Each $\mathbf{E}_j$ is injective as a map $F^{w_{j+1}} \rightarrow F^{w_j}$, so $\mathbf{E}_j \mathbf{X} = \mathbf{O}$ forces $\mathbf{X} = \mathbf{O}$ for every matrix $\mathbf{X}$ with w_{j+1} rows.

We prove $\mathbf{K}_{pq} = \mathbf{O}$ for $p > q$ by induction on $q \geq 1$.

Case $q = 1$. Let $p > 1$. Taking the indices $(p-1, 1)$ in $(*)$, the right-hand side involves $\mathbf{K}_{p-1,0}$, which is undefined and therefore read as $\mathbf{O}$. So $\mathbf{E}_{p-1} \mathbf{K}_{p,1} = \mathbf{O}$, and $p-1 \leq s-1$ makes $\mathbf{E}_{p-1}$ one of the listed blocks; injectivity gives $\mathbf{K}_{p,1} = \mathbf{O}$.

Induction step. Let $q \geq 2$ and assume $\mathbf{K}_{p'q'} = \mathbf{O}$ whenever $p' > q'$ and $q' < q$. Let $p > q$. Taking the indices $(p-1, q)$ in (*),

$$\mathbf{E}_{p-1} \mathbf{K}_{pq} = \mathbf{K}_{p-1, q-1} \mathbf{E}_{q-1}.$$

Here $p-1 > q-1$ and $q-1 < q$, so $\mathbf{K}_{p-1, q-1} = \mathbf{O}$ by the induction hypothesis, and the right-hand side is $\mathbf{O}$. Again $1 \leq p-1 \leq s-1$, so injectivity of $\mathbf{E}_{p-1}$ gives $\mathbf{K}_{pq} = \mathbf{O}$. This completes the induction and the proof. ■

Remark. The Jordan form has no such property, which is the reason the Weyr form exists at all. Take $\mathbf{J} = \mathbf{J}_2(0) \oplus \mathbf{J}_1(0)$ and $\mathbf{W}$ the Weyr matrix of the same operator, of structure $(2, 1)$:

$$\mathbf{J} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Solving $\mathbf{J}\mathbf{K} = \mathbf{K}\mathbf{J}$ entry by entry gives the commutant $\left\{ \begin{pmatrix} a & b & c \\ 0 & a & 0 \\ 0 & h & i \end{pmatrix} \right\}$, which contains ma-

trices with a nonzero entry **below** the diagonal, in position $(3, 2)$: the commutant of $\mathbf{J}$ is not contained in the upper triangular matrices, and in particular it is not block upper triangular

for the Weyr partition $3 = 2 + 1$ of this operator. Solving $\mathbf{W}\mathbf{K} = \mathbf{K}\mathbf{W}$ gives $\left\{ \begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & a \end{pmatrix} \right\}$,

every member of which is upper triangular, in agreement with [Proposition 10.4.6](#). This is the property that makes the Weyr form the right one when two commuting operators must be brought to normal form at the same time.

Quick check. An operator T on a 7-dimensional complex space has one eigenvalue λ , with Jordan block sizes $(3, 2, 1, 1)$. What is the Weyr structure of its Weyr form, and what is $\dim \ker(T - \lambda \text{id})^2$?

Solution. The Weyr structure is the dual partition: $w_1 = \#\{i : k_i \geq 1\} = 4$, $w_2 = \#\{i : k_i \geq 2\} = 2$, $w_3 = \#\{i : k_i \geq 3\} = 1$, so $(4, 2, 1)$, a partition of 7 as it must be. By [Theorem 10.4.5](#) (b), $\dim \ker(T - \lambda \text{id})^2 = w_1 + w_2 = 6$. (Equivalently, $\sum_i \min(k_i, 2) = 2 + 2 + 1 + 1 = 6$, by Corollary 10.3.2 (d).)

10.4.4 Exercises

10.4.4.1 A. Check your understanding

Exercise. A1

- (a) Write down $\mathbf{A}(\lambda)$ for $\lambda = a + bi$ and $\mathbf{C}_2(\lambda)$, and say what their sizes are.
- (b) State the Real Jordan Form Theorem.
- (c) True or false: every matrix in $M_n(\mathbb{R})$ is similar over $\mathbb{R}$ to an upper triangular real matrix. Justify your answer.
- (d) Let $\mathbf{A} \in M_6(\mathbb{R})$ have complex eigenvalues $2, 2, 1 + i, 1 - i, 1 + i, 1 - i$. What are the possible real Jordan forms?
- (e) Define the Weyr structure of an eigenvalue and say how it is computed from T .
- (f) What is the Weyr form of $\mathbf{J}_3(7)$? And of $7\mathbf{I}_3$?

Solution.

- (a) $\mathbf{A}(\lambda) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M_2(\mathbb{R})$ and $\mathbf{C}_2(\lambda) = \begin{pmatrix} \mathbf{A}(\lambda) & \mathbf{I}_2 \\ \mathbf{0} & \mathbf{A}(\lambda) \end{pmatrix} \in M_4(\mathbb{R})$; the real Jordan block $\mathbf{C}_k(\lambda)$ has size $2k$ (Definition 10.4.1).
- (b) Every $\mathbf{A} \in M_n(\mathbb{R})$ is similar, by a **real** invertible matrix, to a direct sum of blocks $\mathbf{J}_k(\mu)$ for the real eigenvalues μ and $\mathbf{C}_k(\lambda)$ for the conjugate pairs $\{\lambda, \bar{\lambda}\}$, and that matrix is unique up to the order of the blocks (Theorem 10.4.4).
- (c) False. A triangular real matrix has real eigenvalues (Theorem 9.7.4), and similar matrices have the same characteristic polynomial (Theorem 7.10.3); so a real matrix with a nonreal eigenvalue, such as the rotation by a right angle, is similar to no real triangular matrix.
- (d) The real eigenvalue 2 has $a(2) = 2$, so its blocks are $\mathbf{J}_2(2)$ or $\mathbf{J}_1(2) \oplus \mathbf{J}_1(2)$. The pair $\{1 \pm i\}$ has $a(1 + i) = 2$, so its real blocks are $\mathbf{C}_2(1 + i)$ (size 4) or $\mathbf{C}_1(1 + i) \oplus \mathbf{C}_1(1 + i)$. That gives four possible real Jordan forms, of total size $2 + 4 = 6$ in each case.
- (e) The Weyr structure of λ is the partition $(w_1, w_2, \dots)$ with $w_j = \dim \ker(T - \lambda \text{id}_V)^j - \dim \ker(T - \lambda \text{id}_V)^{j-1}$, that is, the dual partition of the λ -block sizes of the Jordan form (Theorem 10.4.5).
- (f) $\mathbf{J}_3(7)$ has one 7-block of size 3, so the Weyr structure is $(1, 1, 1)$ and the Weyr form is $\mathbf{J}_3(7)$ itself. For $7\mathbf{I}_3$ the block sizes are $(1, 1, 1)$, the dual partition is (3) , and the Weyr form is $7\mathbf{I}_3$.

10.4.4.2 B. Practice

Exercise. B1: Real forms of two by two matrices

For each real matrix, find the complex eigenvalues, the real Jordan form (Theorem 10.4.4), and an invertible $\mathbf{P} \in M_2(\mathbb{R})$ putting it in that form. Verify $\mathbf{AP} = \mathbf{PA}$ column by column.

- (a) $\mathbf{A}_1 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$.
- (b) $\mathbf{A}_2 = \begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix}$.

Solution.

- (a) $p_{\mathbf{A}_1} = x^2 - (\text{tr } \mathbf{A}_1)x + \det \mathbf{A}_1 = x^2 - 2x + 5 = (x - 1)^2 + 4$, so the eigenvalues are $1 \pm 2i$ and

the real Jordan form is $\mathbf{A}(1+2i) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$. For $\lambda = 1+2i$, $\mathbf{A}_1 - \lambda\mathbf{I} = \begin{pmatrix} -2i & 2 \\ -2 & -2i \end{pmatrix}$, and the first row gives $z_2 = iz_1$; take $\mathbf{w} = (1, i)$. Then $\mathbf{w} = \mathbf{u} - i\mathbf{v}$ with $\mathbf{u} = (1, 0)$ and $\mathbf{v} = (0, -1)$, so

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{P}^{-1}\mathbf{A}_1\mathbf{P} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}.$$

Check: $\mathbf{A}_1\mathbf{u} = (1, -2) = \mathbf{u} + 2\mathbf{v}$ and $\mathbf{A}_1\mathbf{v} = (-2, -1) = -2\mathbf{u} + \mathbf{v}$, which are the two columns of $\mathbf{A}(1+2i)$ as Lemma 10.4.1 (a) predicts with $a = 1, b = 2$.

- (b) $p_{\mathbf{A}_2} = x^2 - 2x + (-3 + 5) = x^2 - 2x + 2$, with roots $1 \pm i$; the real Jordan form is $\mathbf{A}(1+i) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. For $\lambda = 1+i$, the second row of $\mathbf{A}_2 - \lambda\mathbf{I} = \begin{pmatrix} 2-i & -5 \\ 1 & -2-i \end{pmatrix}$ gives $z_1 = (2+i)z_2$; take $\mathbf{w} = (2+i, 1)$. Then $\mathbf{u} = (2, 1)$ and $\mathbf{v} = -\text{Im } \mathbf{w} = (-1, 0)$, so

$$\mathbf{P} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad \det \mathbf{P} = 1, \quad \mathbf{P}^{-1}\mathbf{A}_2\mathbf{P} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Check: $\mathbf{A}_2\mathbf{u} = (6-5, 2-1) = (1, 1) = \mathbf{u} + \mathbf{v}$ and $\mathbf{A}_2\mathbf{v} = (-3, -1) = -\mathbf{u} + \mathbf{v}$.

Exercise. B2: Weyr forms of Jordan matrices

For each Jordan matrix, write down the block sizes, the Weyr structure of each eigenvalue, and the Weyr form (Theorem 10.4.5).

- (a) $\mathbf{J}_3(0) \oplus \mathbf{J}_2(0) \oplus \mathbf{J}_1(0) \in M_6(F)$.
- (b) $\mathbf{J}_2(4) \oplus \mathbf{J}_1(4) \oplus \mathbf{J}_1(9) \in M_4(F)$.
- (c) $\mathbf{J}_2(0) \oplus \mathbf{J}_2(0) \in M_4(F)$.

Solution.

- (a) Block sizes $(3, 2, 1)$, a partition of 6. Dual partition: $w_1 = 3, w_2 = 2, w_3 = 1$. So the Weyr form is the basic Weyr matrix with $\lambda = 0$ and structure $(3, 2, 1)$,

$$\mathbf{W} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

with the 3×2 block $\begin{pmatrix} \mathbf{I}_2 \\ \mathbf{O} \end{pmatrix}$ and the 2×1 block $\begin{pmatrix} \mathbf{I}_1 \\ \mathbf{O} \end{pmatrix}$ above the diagonal. As a check, $\text{rank } \mathbf{W} = 3 = \text{rank } \mathbf{J}$ and $\text{rank } \mathbf{W}^2 = 1 = \text{rank } \mathbf{J}^2$.

- (b) For $\lambda = 4$: sizes $(2, 1)$, dual $(2, 1)$. For $\lambda = 9$: sizes (1) , dual (1) . The Weyr form is the direct sum of the basic blocks,

$$\begin{pmatrix} 4 & 0 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \oplus (9).$$

- (c) Sizes $(2, 2)$, dual $w_1 = 2$, $w_2 = 2$. The basic Weyr matrix with $\lambda = 0$ and structure $(2, 2)$ is

$$\mathbf{W} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which is $\mathbf{J}_2(0) \oplus \mathbf{J}_2(0)$ with its basis reordered as $(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_4)$.

Exercise. B3: Real forms of three by three matrices from the characteristic polynomial

For each characteristic polynomial, list all possible real Jordan forms (Theorem 10.4.4) of a matrix $\mathbf{A} \in M_3(\mathbb{R})$ with that polynomial.

- (a) $p_{\mathbf{A}} = (x - 2)(x^2 + 1)$.
- (b) $p_{\mathbf{A}} = (x + 1)(x^2 - 2x + 5)$.
- (c) $p_{\mathbf{A}} = (x - 1)^3$.
- (d) $p_{\mathbf{A}} = (x^2 + 1)(x - 1)$ but with $\mathbf{A}$ required to be diagonalizable over $\mathbb{R}$. Is this possible?

Solution.

- (a) The real eigenvalue 2 has $a(2) = 1$, giving the block $\mathbf{J}_1(2) = (2)$. The pair $\{i, -i\}$ has $a(i) = 1$, giving $\mathbf{C}_1(i) = \mathbf{A}(i) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. So the real Jordan form is $(2) \oplus \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the only possibility.
- (b) $x^2 - 2x + 5$ has roots $1 \pm 2i$, so the form is $(-1) \oplus \mathbf{A}(1 + 2i) = (-1) \oplus \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$, again the only possibility.
- (c) All eigenvalues are real, so the real Jordan form is an ordinary Jordan form: the 1-block sizes form a partition of 3, giving $\mathbf{J}_3(1)$, $\mathbf{J}_2(1) \oplus \mathbf{J}_1(1)$, or $\mathbf{I}_3$.
- (d) No. A diagonalizable $\mathbf{A} \in M_3(\mathbb{R})$ is similar over $\mathbb{R}$ to a diagonal real matrix, whose characteristic polynomial splits over $\mathbb{R}$ (Theorem 7.4.6); but $x^2 + 1$ has no real root. In fact whenever $\mathbf{A}$ has a nonreal eigenvalue its real Jordan form contains a block $\mathbf{A}(\lambda)$, which is not diagonal.

10.4.4.3 C. Going deeper

Exercise. C1: The real form of a companion matrix

Let $p = (x^2 + 1)^2 = x^4 + 2x^2 + 1 \in \mathbb{R}[x]$ and let $\mathbf{C} = \mathbf{C}(p) \in M_4(\mathbb{R})$ be its companion matrix, with 1 in each position just below the diagonal and last column $(-1, 0, -2, 0)$ (Exercise (C1: Companion matrices)).

- (a) Write down $\mathbf{C}$, and explain why $p_{\mathbf{C}} = m_{\mathbf{C}} = p$.
- (b) Determine the complex Jordan form of $\mathbf{C}$.
- (c) Deduce the real Jordan form of $\mathbf{C}$.
- (d) Verify your answer by comparing traces and characteristic polynomials.

Hint: for (b), the exponent of $x - i$ in $m_{\mathbf{C}}$ is the largest size of an i -block.

Solution.

- (a) With $a_0 = 1$, $a_1 = 0$, $a_2 = 2$, $a_3 = 0$,

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

By Exercise (C1: Companion matrices) (b), $p_{\mathbf{C}} = p$, and by Exercise (B3: Companion matrices) (d), $m_{\mathbf{C}} = p$ as well.

- (b) Over $\mathbb{C}$, $p = (x - i)^2(x + i)^2$, so $a(i) = a(-i) = 2$. By Corollary 10.3.2 (c), the exponent 2 of $x - i$ in $m_{\mathbf{C}}$ is the largest size of an i -block; the sizes form a partition of 2 with largest part 2, so there is one block $\mathbf{J}_2(i)$. Likewise one block $\mathbf{J}_2(-i)$. The complex Jordan form is $\mathbf{J}_2(i) \oplus \mathbf{J}_2(-i)$.
- (c) By Theorem 10.4.4, the real Jordan form has no ordinary Jordan blocks (there is no real eigenvalue) and one real block per complex i -block, of the same size parameter. So it is

$$\mathbf{C}_2(i) = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

(Consistently, Corollary 10.4.2 turns $\mathbf{C}_2(i)$ back into $\mathbf{J}_2(i) \oplus \mathbf{J}_2(-i)$ over $\mathbb{C}$.)

- (d) $\text{tr } \mathbf{C} = 0 = \text{tr } \mathbf{C}_2(i)$, and both matrices have characteristic polynomial $(x^2 + 1)^2$: for $\mathbf{C}$ by (a), and for $\mathbf{C}_2(i)$ by Corollary 10.4.2. Their minimal polynomials agree too, both being $(x^2 + 1)^2$.

Exercise. C2: Commuting with a Weyr matrix, and with a Jordan matrix

Let F be a field.

- (a) Let $\mathbf{W} \in M_3(F)$ be the basic Weyr matrix with $\lambda = 0$ and structure $(2, 1)$ (Definition 10.4.2). Find all $\mathbf{K} \in M_3(F)$ with $\mathbf{WK} = \mathbf{KW}$, and check that they are block upper triangular for the partition $3 = 2 + 1$.
- (b) Let $\mathbf{J} = \mathbf{J}_2(0) \oplus \mathbf{J}_1(0)$. Find all $\mathbf{K} \in M_3(F)$ with $\mathbf{JK} = \mathbf{KJ}$, and exhibit one that is **not** upper triangular. Which partition of 3 does it violate?
- (c) Compute the dimension of each of the two commutants, and explain why they are equal.

Solution.

- (a) $\mathbf{W} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Write $\mathbf{K} = (k_{ij})$. Then $\mathbf{WK}$ has first row equal to the third row of $\mathbf{K}$ and other rows zero, while $\mathbf{KW}$ has third column equal to the first column of $\mathbf{K}$

and other columns zero. Equating,

$$(k_{31}, k_{32}, k_{33}) = (0, 0, k_{11}), \quad k_{21} = k_{31} = 0.$$

So the commutant is

$$\left\{ \begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & a \end{pmatrix} : a, b, c, e, f \in F \right\},$$

of dimension 5. The lower-left 1×2 block is $(k_{31}, k_{32}) = (0, 0)$, so every such $\mathbf{K}$ is block upper triangular for $3 = 2 + 1$, as [Proposition 10.4.6](#) asserts.

- (b) $\mathbf{J} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Now $\mathbf{JK}$ has first row equal to the second row of $\mathbf{K}$ and other rows zero, while $\mathbf{KJ}$ has second column equal to the first column of $\mathbf{K}$ and other columns zero. Equating gives $k_{21} = 0$, $k_{22} = k_{11}$, $k_{23} = 0$, and $k_{31} = 0$, so the commutant is

$$\left\{ \begin{pmatrix} a & b & c \\ 0 & a & 0 \\ 0 & h & i \end{pmatrix} \right\},$$

also of dimension 5. Take $\mathbf{K}_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, which commutes with $\mathbf{J}$ (it is the case

$a = b = c = i = 0$, $h = 1$). Its only nonzero entry sits in position $(3, 2)$, below the diagonal, so $\mathbf{K}_0$ is not upper triangular. The partition it violates is $3 = 2 + 1$, the Weyr partition of this operator: the lower-left 1×2 block of $\mathbf{K}_0$ is $(0, 1) \neq \mathbf{0}$. So the conclusion of [Proposition 10.4.6](#) genuinely fails for the Jordan form, while (a) confirms it for the Weyr form ([Theorem 10.4.5](#)) of the same operator.

- (c) Both have dimension 5. They must agree, because $\mathbf{J}$ and $\mathbf{W}$ are similar (both are nilpotent with $\text{rank} = 1$ and $\text{rank}^2 = 0$, so [Corollary 10.2.5](#) applies): if $\mathbf{W} = \mathbf{P}^{-1}\mathbf{JP}$, then $\mathbf{K} \mapsto \mathbf{P}^{-1}\mathbf{KP}$ is a bijection from the commutant of $\mathbf{J}$ to that of $\mathbf{W}$, and it is linear and invertible, hence preserves dimension.