Jordan Canonical Form Let T be a linear operator on a finite dimensional vector space V, whose characteristic polynomial splits. A matrix J is called a *Jordan Canonical form* if J is block diagonal, and all of its diagonal blocks are Jordan blocks. For example: the matrix

$$\begin{pmatrix} 3 & 1 & 0 & 0 & & & \\ 0 & 3 & 1 & 0 & & & \\ 0 & 0 & 3 & 1 & & & \\ 0 & 0 & 0 & 3 & & & \\ & & & & 3 & & \\ & & & & & 0 & 1 \\ & & & & & 0 & 0 \end{pmatrix}$$

is a Jordan canonical form. If β is a basis for V such that $[T]_\beta$ is a Jordan canonical form, then we say β is a *Jordan canonical basis*.

A *cycle of generalized eigenvectors* is defined to be an ordered set of vectors

$$\{x, x', x'', \ldots, x^{(l)}\}$$

where $x \in \mathsf{E}_\lambda$ is an eigenvector, and $(T - \lambda I)x' = x$, $(T - \lambda I)x'' = x'$, etc. The vectors $x, x', x'', \ldots$ belong to K_λ, the generalized λ-eigenspace. A basis β is a Jordan canonical basis if and only if β is an ordered union of cycles of generalized eigenvectors. Each cycle of generalized eigenvectors spans a T-cyclic subspace of V. Hence, the blocks of a Jordan canonical form for T correspond to T-cyclic subspaces of V, and a Jordan canonical basis yields a direct sum decomposition of V into T-cyclic subspaces.

Algorithm (Algorithm for finding a Jordan canonical basis)**.**

1. Find the eigenvalues of T.
2. For each eigenvalue λ:
   - Compute bases for $\operatorname{im}(T-\lambda I)$, $\operatorname{im}(T-\lambda I)^2$, ..., as well as for E_λ, $\operatorname{im}(T-\lambda I) \cap \mathsf{E}_\lambda$, $\operatorname{im}(T-\lambda I)^2 \cap \mathsf{E}_\lambda$, ..., until we find the smallest number p such that $\operatorname{im}(T-\lambda I)^p \cap \mathsf{E}_\lambda = \{0\}$.

     Note: p is also equal to:
     - the largest number such that $\operatorname{im}(T - \lambda I)^{p-1} \cap \mathsf{E}_\lambda \neq \{0\}$;
     - the size of the largest Jordan block with eigenvalue λ;
     - the multiplicity of λ as a root of the minimal polynomial.
   - *Stage* 1:
     - Find a basis $S_1 = \{x_1, \ldots, x_{k_1}\}$ for $\operatorname{im}(T - \lambda I)^{p-1} \cap \mathsf{E}_\lambda$.
   - *Stage* 2:
     - Extend $\{x_1, \ldots, x_{k_1}\}$ to a basis $\{x_1, \ldots, x_{k_2}\}$, for $\operatorname{im}(T - \lambda I)^{p-2} \cap \mathsf{E}_\lambda$.
     - For each $x \in S_1$, find $x' \in \operatorname{im}(T - \lambda I)^{p-2}$ such that $(T - \lambda I)x' = x$.
     - Let $S_2 = \{x'_1, \ldots, x'_{k_1}, x_{k_1+1}, \ldots, x_{k_2}\}$ be the set of vectors found at stage 2.

- *Stage* 3:
  - Extend $\{x_1, \ldots, x_{k_2}\}$ to a basis $\{x_1, \ldots, x_{k_3}\}$, for $\mathrm{im}(T - \lambda I)^{p-3} \cap \mathsf{E}_\lambda$.
  - For each $x \in S_2$, find $x' \in \mathrm{im}(T - \lambda I)^{p-3}$ such that $(T - \lambda I)x' = x$.
  - Let $S_3 = \{x_1'', \ldots, x_{k_3}\}$ be the set of vectors found at stage 3.

  $\vdots$

- *Stage* p:
  - Extend $\{x_1, \ldots, x_{k_{p-1}}\}$ to a basis $\{x_1, \ldots, x_{k_p}\}$, for E_λ.
  - For each $x \in S_{p-1}$, find $x' \in V$ such that $(T - \lambda I)x' = x$.
  - Let $S_p = \{x_1^{(p-1)}, \ldots, x_{k_p}\}$ be the set of vectors found at stage p.
- If x_j was added at stage q_j, then the ordered set $\gamma_j = \{x_j, x_j', x_j'', \ldots, x_j^{(p-q_j)}\}$ is a cycle of generalized eigenvectors.

3. Take the (ordered) union of all cycles of generalized eigenvectors, for all eigenvalues. This will be a Jordan canonical basis.

Example. Let

$$A = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 1 & 3 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 \end{pmatrix}$$

The eigenvalues are $1, 1, 1, 3$.

- $\lambda = 1$
  - First compute a basis for E_1:

    $$\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

    Next compute a basis for $\mathrm{col}(A - I)$:

    $$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

    Next compute a basis for $\mathrm{col}(A - I) \cap \mathsf{E}_1$. For this, we need to find all vectors of the form

    $$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \in \mathsf{E}_1 \,.$$

Thus we solve

$$(A-I)\begin{pmatrix}1&0\\0&1\\0&0\\1&0\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

and find that $\left(\begin{smallmatrix}a\\b\end{smallmatrix}\right)\in\operatorname{col}\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)$. Therefore

$$\begin{pmatrix}1&0\\0&1\\0&0\\1&0\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}\in\operatorname{col}\begin{pmatrix}1&0\\0&1\\0&0\\1&0\end{pmatrix}\begin{pmatrix}1\\-2\end{pmatrix}=\operatorname{col}\begin{pmatrix}1\\-2\\0\\1\end{pmatrix}.$$

Thus a basis for $\operatorname{col}(A-I)\cap\mathsf{E}_1$ is given by

$$\left\{\begin{pmatrix}1\\-2\\0\\1\end{pmatrix}\right\}$$

Proceeding similarly, we compute $\operatorname{col}(A-I)^2$, and $\operatorname{col}(A-I)^2\cap\mathsf{E}_1$, and find that $\operatorname{col}(A-I)^2\cap\mathsf{E}_1=\{0\}$. Therefore $p=2$ and we can stop here.

– *Stage 1.* Let $S_1=\{x_1\}$ be the basis for $\operatorname{col}(A-I)\cap\mathsf{E}_1$ above.

$$x_1=\begin{pmatrix}1\\-2\\0\\1\end{pmatrix}$$

– *Stage 2.* Extend S_1 to a basis for E_1. We can do this by choosing either of the two vectors in the basis for E_1 above: We'll take

$$x_2=\begin{pmatrix}-2\\1\\0\\0\end{pmatrix}$$

Next find a vector $x_1'\in\operatorname{col}(A-I)^{p-2}$ such that $(A-I)x_1'=x_1$. Solving this gives:

$$x_1'=\begin{pmatrix}-2\\0\\\frac{1}{3}\\0\end{pmatrix}$$

(Note: This step is easy in this example, because $p=2$ and so $\operatorname{col}(A-I)^{p-2}$ is all of $\mathbb{C}^4$. In examples where $p>2$, we must be careful to ensure that $x_1'\in\operatorname{col}(A-I)^{p-2}$.)

– We have two cycles of generalized eigenvectors:

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ \frac{1}{3} \\ 0 \end{pmatrix} \right\} \quad \text{and} \quad \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

- $\lambda = 3$.
  - Since the multiplicity of λ is 1, we know that $p = 1$, and we only have to compute a basis for E_3, which is

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

  - Hence we have one cycle of generalized eigenvectors (of length 1), which is

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

A Jordan canonical basis is therefore:

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ \frac{1}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

which yields the Jordan canonical form

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

In other words:

$$A = \begin{pmatrix} 1 & -2 & -2 & 0 \\ -2 & 0 & 1 & 1 \\ 0 & \frac{1}{3} & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 & -2 & 0 \\ -2 & 0 & 1 & 1 \\ 0 & \frac{1}{3} & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{-1}.$$

We'll now prove that the algorithm produces a basis for V. This is basis is a Jordan canonical basis, since it is a union of cycles of generalized eigenvectors. The proof below is based on the following elementary (but not too easy) exercise.

Exercise. *Let V be a finite dimensional vector space, and let $A, B \in \mathsf{L}(V)$. Prove that* $\text{nullity}(AB) = \dim(\text{im}(B) \cap \ker(A)) + \text{nullity}(B)$.

Proof that the algorithm is correct. We need to show that $S_1 \cup \cdots \cup S_p$ is a basis for K_λ. To do this, we show that

(a) $|S_1 \cup \cdots \cup S_p| = \dim \mathsf{K}_\lambda$, and

(b) $S_1 \cup \cdots \cup S_p$ is linearly independent.

(a) From the exercise, with $A = T - \lambda I$ and $B = (T - \lambda I)^{k-1}$, we obtain

$$\text{nullity}(T - \lambda I)^k = \dim(\text{im}(T - \lambda I)^{k-1} \cap \mathsf{E}_\lambda) + \text{nullity}(T - \lambda I)^{k-1} . \tag{1}$$

Now for all $k \geq p$, we have

$$\text{im}(T - \lambda I)^k \cap \mathsf{E}_\lambda \subset \text{im}(T - \lambda I)^p \cap \mathsf{E}_\lambda = \{0\} ,$$

so $\text{nullity}(T - \lambda I)^{k+1} = \text{nullity}(T - \lambda I)^k$, and therefore $\ker(T - \lambda)^p = \mathsf{K}_\lambda$. Therefore applying (1) inductively, we have

$$\begin{aligned} \dim \mathsf{K}_\lambda &= \text{nullity}(T - \lambda I)^p \\ &= \sum_{q=1}^{p} \dim(\text{im}(T - \lambda I)^{p-q} \cap \mathsf{E}_\lambda) \\ &= \sum_{q=1}^{p} |S_q| \\ &= |S_1 \cup \cdots \cup S_p| . \end{aligned}$$

(b) We can write $S_1 \cup \cdots \cup S_p = \bigcup_{j=1}^{k_p} \{x_j, x'_j, \ldots, x_j^{(p-q_j)}\}$ where $x_j \in S_{q_j}$. Suppose that

$$0 = \sum_{i=1}^{k_p} (a_i x_i + a'_i x'_i + \cdots + a_i^{(p-q_i)} x_i^{(p-q_i)}) \tag{2}$$

and that not all coefficients $a_i^{(j)}$ are equal to 0. Let $k \geq 0$ be the largest integer such that $a_i^{(k)} \neq 0$ for some i. Then applying $(T - \lambda I)^k$ to both sides of (2) yields

$$0 = \sum a_i^{(k)} x_i ,$$

where the sum is taken over all i for which a vector $x_i^{(k)}$ has been constructed in the algorithm. Since $\{x_1, \ldots, x_{k_p}\}$ is a basis for E_λ, it is a linearly independent set, so $a_i^{(k)} = 0$ for all i. This is a contradiction; we conclude that coefficients of (2) must equal 0, which shows that $S_1 \cup \cdots \cup S_p$ is linearly independent. □

Rational Canonical Form (an overview). If the characteristic polynomial of T does not split, we might wish to find a *rational canonical form*. This is a block diagonal matrix, where each block is the companion matrix of a power of an irreducible polynomial. For example:

$$\begin{pmatrix} 0 & 0 & 0 & -1 & & & \\ 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & -2 & & & \\ 0 & 0 & 1 & 0 & & & \\ & & & & 3 & & \\ & & & & & 0 & -5 \\ & & & & & 1 & 1 \end{pmatrix}$$

is a rational canonical form over $\mathbb{Q}$: the blocks are the companion matrices of $(t^2+1)^2$, $t-3$, and t^2-t+5, which are all powers of irreducible polynomials. If β is a basis for V and $[T]_\beta$ is a rational canonical form, we say that β is a *rational canonical basis*.

Like the Jordan canonical form, the blocks of a rational canonical form for T correspond to T-cyclic subspaces of V. A rational canonical basis corresponds to a direct sum decomposition of V into the maximum possible number of T-cyclic subspaces.

The algorithm for producing a rational canonical basis is very similar to the algorithm for producing a Jordan canonical basis. The proof that it works is also similar. The modifications one needs to make are summarized in the table below. The main point is that the operator $T - \lambda I$ (where λ is an eigenvalue) is replaced by $q(T)$ (where $q(t)$ is an irreducible factor of the characteristic polynomial).

Jordan Canonical Basis	Rational Canonical Basis
$f_T(t) = (t-\lambda_1)^{d_1} \dots (t-\lambda_k)^{d_k}$	$f_T(t) = q_1(t)^{d_1} \dots q_k(t)^{d_k}$ $q_1(t), \dots, q_k(t)$ irreducible
Eigenspaces E_λ	$F_i = \ker q_i(T)$
Generalized eigenspaces K_λ	Primary subspaces $W_i = \ker q_i(T)^{d_i}$
$\operatorname{im}(T-\lambda I)^j, \quad j = 0, 1, 2, \dots$	$\operatorname{im} q_i(T)^j, \quad j = 0, 1, 2, \dots$
Extend basis for $\mathsf{E}_\lambda \cap \operatorname{im}(T-\lambda I)^j$ to a basis for $\mathsf{E}_\lambda \cap \operatorname{im}(T-\lambda I)^{j-1}$	$(*)$ Extend basis in a special way
Organize vectors found into cycles of generalized eigenvectors	$(**)$ Organize vectors into bases for T-cyclic subspaces, and replace with cyclic basis

The only place where things are notably trickier is $(*)$. When we extend a basis for $F_i \cap \operatorname{im} q_i(T)^j$ to a basis for $F_i \cap \operatorname{im} q_i(T)^{j-1}$, we do the following. Whenever we add a vector x to our basis, we also add $Tx, T^2x, T^{l_i-1}x$, where $l_i = \deg q_i(t)$. The hard part here is the theory: one needs to prove this procedure actually gives a basis.

Things are also a bit different at $(**)$. In the Jordan Canonical basis algorithm, the vectors we obtain can be organized into bases for T-cyclic subspaces of V, and these bases happen to be cycles of generalized eigenvectors (which is exactly what we want). If we do the analogous thing with Rational Canonical form, we can still organize the vectors into bases for T-cyclic subspaces, but they are not the bases we want. We therefore need to take the extra step of replacing each of these bases by the basis we actually want (the cyclic basis).

The rational canonical form enjoys similar properties to the Jordan canonical form: e.g. any two rational canonical forms for T are equivalent (i.e. they are the same, up to permuting the blocks.) Hence many theorems involving Jordan Canonical form (in either the statement or the proof) also work with rational canonical form.

There is alternate version of the rational canonical form, in which one decomposes V into the *minimum* possible number of T-cyclic subspaces (instead of the maximum possible). The downside of this is that the blocks are bigger, so it's "less close" to being a diagonal matrix. However, there are also some advantages of the alternate version:

1. It can be computed without factoring the characteristic polynomial.
2. It does not depend on the field of definition: we get exactly the same answer if we work in a field extension.
3. There is a natural way to order the blocks: it's 100% canonical.
4. The largest block is the companion matrix for the minimal polynomial.

Because of these facts, many computer algebra packages use this alternate version of the rational canonical form.