

# La forma canónica de Jordan

Fernando Chamizo

11 de abril de 2019

## 1. Repaso de diagonalización

En breve, la diagonalización de aplicaciones lineales busca un sistema de coordenadas en que adquieran una forma especialmente sencilla. Esto es muy útil en física y matemáticas. Por ejemplo, supongamos que un problema mecánico nos lleva a unas trayectorias  $(x, y) = (x(t), y(t))$  que satisfacen ecuaciones de movimiento  $\dot{x} = 7x - 10y$ ,  $\dot{y} = 3x - 4y$  donde  $\dot{x}$  e  $\dot{y}$  indican derivadas con respecto al tiempo. Es difícil a simple vista determinar las trayectorias con  $x(0) = 7$ ,  $y(0) = 4$ . Sin embargo con el cambio de coordenadas (de base)  $x = 2X + 5Y$ ,  $y = X + 3Y$  las ecuaciones se diagonalizan a  $X' = 2X$ ,  $Y' = Y$  bajo  $X(0) = Y(0) = 1$  que tienen  $X = e^{2t}$ ,  $Y = e^t$  como solución. Deshaciendo el cambio se sigue la trayectoria  $(2e^{2t} + 5e^t, e^{2t} + 3e^t)$  para el problema original.

La “sencillez” que perseguimos es que la aplicación lineal se reduzca a dilataciones y contracciones. Con esta idea, se dice que  $\vec{v} \in V - \{\vec{0}\}$  es un *vector propio* o *autovector* de un endomorfismo  $L : V \rightarrow V$  si  $L(\vec{v}) = \lambda\vec{v}$  con  $\lambda \in \mathbb{K}$ . Este número  $\lambda$  se dice que es un *autovalor*. Aunque estas definiciones son generales y por cierto no necesitan de productos escalares, nos restringimos una vez más al caso  $\dim V < \infty$ . La única referencia vaga que haremos al caso infinita es:

**Ejemplo.** La ecuación de Schrödinger independiente del tiempo es  $\hat{H}\Psi = E\Psi$  donde  $\hat{H}$  es el *hamiltoniano*, una aplicación lineal típicamente en un espacio de dimensión infinita. El autovalor  $E$  correspondiente a la *función de ondas*  $\Psi$  es la *energía*. Por ejemplo, para una partícula de masa  $m$  bajo un potencial  $V$  en una dimensión  $\hat{H}$  es el operador que aplica cada función ondas en  $\Psi$  en  $-\frac{\hbar^2}{2m}\Psi'' + V\Psi$ . El espacio vectorial es un espacio de funciones (que a menudo no se especifica en física y se sobreentiende que es el necesario para que contenga todas las funciones de ondas empleadas en un problema).

Si  $V = 0$  (partícula libre), los autovalores para  $E > 0$  son de la forma  $\lambda \cos(px/\hbar) + \mu \sin(px/\hbar)$  con  $p = \sqrt{2mE}$  (el momento), ondas puras muy diferentes de la concentración que uno espera de una partícula. Esto es una manifestación del principio de incertidumbre: Especificar completamente la energía y por tanto  $p$ , implica una falta total de localización en el espacio. Para otros potenciales se pueden tener autovalores aislados correspondientes a funciones de onda *normalizadas* ( $\int |\Psi|^2 = 1$ ) lo cual significa que la energía está cuantizada.

Cuando tenemos suficientes vectores como para formar una base de  $V$  se dice que la aplicación lineal es *diagonalizable*. Por la fórmula del cambio de base,  $L$  es diagonalizable si y solo si su matriz, que también llamaremos  $L$ , verifica que existe  $C$  no singular y  $D$  diagonal tal que

$$(1) \quad C^{-1}LC = D \quad \text{además:} \quad \begin{cases} \text{columnas de } C = \text{coordenadas de los autovectores,} \\ \text{diagonal de } D = \text{autovalores respectivos.} \end{cases}$$

Un vector propio  $\vec{v}_0$  de un autovalor  $\lambda_0$  es un elemento no nulo de  $\text{Ker}(L - \lambda_0 \text{Id})$ . Por tanto necesariamente  $L - \lambda_0 \text{Id}$  pierde la inyectividad. Esto da un método para calcular autovalores y autovectores a partir de las matrices. Los autovalores son los ceros del *polinomio característico*  $p(\lambda) = \det(L - \lambda I)$ , las raíces de la *ecuación característica*  $\det(L - \lambda I) = 0$ .

Observación. El método indicado se vuelve muy poco eficiente para un ordenador cuando la dimensión es grande. Un tema central del *álgebra lineal numérica* es diseñar métodos alternativos con este propósito. El algoritmo matemático de ordenación de resultados en el buscador de internet en el que estás pensando, está basado en la aproximación de un autovector en dimensión mayor que  $10^{12}$ . El cálculo se repite cada cierto número de meses para reflejar los cambios en la red.

**Ejemplo.** Si tenemos una aplicación lineal  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$  cuya matriz es:

$$(2) \quad A = \begin{pmatrix} 3 & -2 & 2 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{se tiene} \quad p(\lambda) = \begin{vmatrix} 3 - \lambda & -2 & 2 \\ 1 & -\lambda & 1 \\ 0 & 0 & 1 - \lambda \end{vmatrix} = (2 - \lambda)(\lambda - 1)^2.$$

A pesar de que solo hay dos autovalores, queda la posibilidad de que sea diagonalizable si podemos encontrar en el núcleo suficientes autovectores linealmente independientes. Para  $\lambda = 2$

$$(3) \quad A - 2I = \begin{pmatrix} 1 & -2 & 2 \\ 1 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix} \Rightarrow \text{Base del núcleo} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Y para  $\lambda = 1$ ,

$$(4) \quad A - I = \begin{pmatrix} 2 & -2 & 2 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \text{Base del núcleo} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

De esta forma se consigue

$$(5) \quad C^{-1}AC = D \quad \text{con} \quad D = \begin{pmatrix} 2 & & \\ & 1 & \\ & & 1 \end{pmatrix} \quad \text{y} \quad C = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Es importante respetar la ordenación respectiva de los vectores y valores propios. En el ejemplo anterior, la igualdad  $C^{-1}AC = D$  no se vería afectada si intercambiamos las dos

últimas columnas de  $C$  porque corresponden al mismo autovalor pero sería incorrecta intercambiando la primera y la segunda.

Una consecuencia de lo que veremos más adelante es que  $\dim \text{Ker}(L - \lambda \text{Id})$  es menor o igual que la multiplicidad del cero  $\lambda$  en el polinomio característico pero la igualdad no siempre se da. Justamente es eso lo que impide que todos los endomorfismos en espacios vectoriales de dimensión finita sobre  $\mathbb{C}$  sean diagonalizables. En el caso de  $\mathbb{K} = \mathbb{R}$  se añade la dificultad de que una ecuación polinómica sobre  $\mathbb{R}$  no siempre tiene soluciones en  $\mathbb{R}$ , como ocurre con  $x^2 + 1 = 0$ .

**Ejemplo.** La aplicación  $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por  $L((x, y)^t) = (3x + 5y, 3y)$  tiene polinomio característico  $p(\lambda) = (\lambda - 3)^2$  y  $\text{Ker}(L - 3\text{Id})$  es el subespacio generado por  $\vec{e}_1 = (1, 0)^t$  por tanto no es posible construir una base de vectores propios y  $L$  no es diagonalizable.

**Ejemplo.** La aplicación  $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por  $L((x, y)^t) = (-3x - 2y, 5x + 3y)^t$  tiene como polinomio característico  $p(\lambda) = \lambda^2 + 1$  por tanto no es diagonalizable. Si consideramos la aplicación  $\tilde{L} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  con la misma ecuación sí es lícito considerar  $\lambda_1 = i$ ,  $\lambda_2 = -i$ . En este caso,  $\text{Ker}(\tilde{L} - \lambda \text{Id})$  está generado por  $(2, -3 - i)^t$  para  $\lambda = i$  y por  $(2, -3 + i)^t$  para  $\lambda = -i$  (esto último se sigue conjugando ya que la expresión para  $\tilde{L}$  es real). En términos matriciales se tiene

$$(6) \quad \begin{pmatrix} 2 & 2 \\ -3 - i & -3 + i \end{pmatrix}^{-1} \begin{pmatrix} -3 & -2 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ -3 - i & -3 + i \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

## 2. El teorema espectral

El resultado principal que veremos en esta sección es que hay una clase de aplicaciones lineales, ubicuas en física, que siempre son diagonalizables.

Recordemos que una propiedad que tenían las aplicaciones autoadjuntas es que en la descomposición  $V = W \oplus W^\perp$  si cumplían  $L(W) \subset W$  entonces también cumplían  $L(W^\perp) \subset W^\perp$ . En particular, si  $W$  es el subespacio generado por un vector propio, entonces también actúa sobre su ortogonal. Por un procedimiento inductivo se van obteniendo vectores propios ortogonales, al menos sobre  $\mathbb{C}$  donde siempre hay un autovector porque la ecuación característica siempre tiene raíces. Por ello vamos a centrarnos momentáneamente en el caso  $\mathbb{K} = \mathbb{C}$ .

Si  $V$  es un espacio unitario (de dimensión finita) la generalidad que buscamos pasa por considerar las aplicaciones lineales  $L : V \rightarrow V$  *normales*, lo que significa que conmutan con su adjunta

$$(7) \quad L^* \circ L = L \circ L^* \quad \text{o en términos matriciales, } \overline{L}^t L = L \overline{L}^t.$$

Es obvio que las aplicaciones autoadjuntas y las aplicaciones unitarias son normales (así como las ortogonales si considerásemos el caso real).

**Ejemplo.** La aplicación  $L : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  dada por  $L((x, y)^t) = (-2y, 2x)^t$  no es ni autoadjunta ni unitaria pero sí normal. Esto es fácil pensando en que su matriz es antisimétrica real. Hay otros contraejemplos que no involucran matrices antisimétricas.

El *teorema espectral* afirma que en un espacio unitario *las aplicaciones normales siempre son diagonalizables en una base ortonormal*.

La prueba se basa en concretar los pasos esbozados antes tras comprobar que las aplicaciones normales (como las autoadjuntas) cumplen que si  $W$  es el espacio generado por un autovector  $\vec{v}$  de  $L$ , entonces  $L(W^\perp) \subset W^\perp$ . Si  $\vec{v}$  también fuera autovector de  $L^*$ , esto sería fácil ya que para todo  $\vec{x} \in W^\perp$  se tendría  $\langle L(\vec{x}), \vec{v} \rangle = \langle \vec{x}, L^*(\vec{v}) \rangle = \mu \langle \vec{x}, \vec{v} \rangle = 0$ . El hecho aparentemente milagroso de que siempre  $\vec{v}$  sea autovector de  $L^*$  tiene lugar gracias a la siguiente igualdad donde se ha abreviado  $L - \lambda \text{Id}$  por  $M$ ,

$$(8) \quad \|M^*(\vec{v})\|^2 = \langle M^*(\vec{v}), M^*(\vec{v}) \rangle = \langle (M \circ M^*)(\vec{v}), \vec{v} \rangle = \langle M^*(M(\vec{v})), \vec{v} \rangle$$

y  $M(\vec{v}) = \vec{0}$  si  $\lambda$  es el autovalor de  $\vec{v}$  por lo cual  $M^*(\vec{v}) = \vec{0}$ . Como  $M^* = L^* - \bar{\lambda} \text{Id}$ , se concluye que  $\vec{v}$  es autovector de  $L^*$  con autovalor  $\bar{\lambda}$ .

Si  $\vec{v}$  es un autovector con autovalor  $\lambda$  de una aplicación autoadjunta  $L$ , se tiene la siguiente cadena de igualdades:

$$(9) \quad \lambda \|\vec{v}\|^2 = \langle L(\vec{v}), \vec{v} \rangle = \langle \vec{v}, L(\vec{v}) \rangle = \bar{\lambda} \|\vec{v}\|^2$$

que implica que *los autovalores de una aplicación autoadjunta son siempre reales*. En el caso real podríamos *complejificar* el espacio añadiendo escalares para tener un espacio vectorial sobre  $\mathbb{C}$  y se deduce de (9) la sorprendente propiedad de que siempre la ecuación característica de una matriz simétrica tiene todas sus raíces reales, que es lo que nos faltaba en el caso euclídeo para saber que son diagonalizables.

En resumen, *las aplicaciones autoadjuntas son siempre diagonalizables en una base ortonormal, tanto en el caso real como en el complejo, y los autovalores son siempre reales*. Las aplicaciones normales, en particular las unitarias, también se diagonalizan en una base ortonormal cuando  $\mathbb{K} = \mathbb{C}$  y los autovalores son en general complejos.

Observación. En física cuántica los valores propios de operadores autoadjuntos corresponden a los posibles resultados de las mediciones, por tanto es natural que sean reales.

**Ejemplo.** Analicemos en qué se manifiesta que una matriz simétrica arbitraria real  $2 \times 2$  siempre tenga valores propios reales. Debemos resolver

$$(10) \quad \begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = 0 \quad \text{que equivale a} \quad \lambda^2 - (a + c)\lambda - b^2 = 0.$$

El discriminante de esta ecuación es  $(a + c)^2 + 4b^2$  que claramente es no negativo y por tanto nunca lleva a soluciones complejas. Dar una prueba similar en el caso  $3 \times 3$  sería mucho más difícil.

El hecho de que podamos diagonalizar en una base ortonormal no implica que la base que elijamos lo sea. Necesariamente debemos normalizar los vectores, y si  $\dim \text{Ker}(L - \lambda \text{Id}) > 1$  debemos ortonormalizar porque habrá más de un vector. En relación con esto, un subproducto de la prueba anterior es que *para operadores autoadjuntos los autovectores correspondientes a autovalores distintos son siempre ortogonales*, sin mediar ninguna elección. Una prueba directa de ello es una variante de (9):

$$(11) \quad \lambda_j \langle \vec{v}_j, \vec{v}_k \rangle = \langle L(\vec{v}_j), \vec{v}_k \rangle = \langle \vec{v}_j, L(\vec{v}_k) \rangle = \lambda_k \langle \vec{v}_j, \vec{v}_k \rangle$$

así que  $\lambda_j \neq \lambda_k$  implica  $\langle \vec{v}_j, \vec{v}_k \rangle = 0$ .

**Ejemplo.** En la molécula de amoniaco  $\text{NH}_3$  el nitrógeno puede estar en dos posiciones: por encima o por debajo del plano que determinan los hidrógenos. Según los modelos cuánticos las energías y estados corresponden a autovalores y autovectores de cierta matriz  $2 \times 2$ , que por ciertas simetrías debe ser de la forma

$$(12) \quad H = \begin{pmatrix} E_0 & \Delta \\ \Delta & E_0 \end{pmatrix} \quad \text{con } E_0, \Delta \in \mathbb{R}^+.$$

Es la matriz de una aplicación autoadjunta (hermítica)  $\mathbb{C}^2 \rightarrow \mathbb{C}^2$ , por tanto debe diagonalizarse en una base ortonormal. Los autovalores son raíces de  $(\lambda - E_0)^2 - \Delta^2 = 0$ , resultando los niveles de energía  $\lambda_1 = E_0 + \Delta$  y  $\lambda_2 = E_0 - \Delta$ . Resolviendo las ecuaciones  $(H - \lambda_j I)\vec{v}_j = \vec{0}$  y normalizando los vectores se obtiene  $\vec{v}_1 = (1, 1)^t / \sqrt{2}$  y  $\vec{v}_2 = (1, -1)^t / \sqrt{2}$ . En la base ortonormal  $\{\vec{v}_1, \vec{v}_2\}$  la matriz se diagonaliza y podemos escribir

$$(13) \quad H = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} E_0 + \Delta & 0 \\ 0 & E_0 - \Delta \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}^{-1}.$$

Como la matriz de cambio es necesariamente ortogonal la inversa es en realidad la traspuesta, que casualmente coincide con ella misma en este ejemplo.

Físicamente podríamos determinar  $2\Delta$  midiendo la energía de un fotón emitido al pasar de uno a otro estado. El módulo de las coordenadas al cuadrado de los estados representa probabilidades y así la profusión de  $1/\sqrt{2}$  indica que es igualmente probable que el nitrógeno esté arriba o abajo.

**Ejemplo.** Para ver la necesidad de ortonormalizar, diagonalicemos la matriz simétrica

$$(14) \quad A = \begin{pmatrix} 3 & 2 & \sqrt{2} \\ 2 & 3 & \sqrt{2} \\ \sqrt{2} & \sqrt{2} & 2 \end{pmatrix}.$$

El polinomio característico es  $|A - \lambda I| = (\lambda - 1)^2(6 - \lambda)$ , entonces los valores propios son  $\lambda_1 = \lambda_2 = 1$  y  $\lambda_3 = 6$ . Al resolver  $(A - I)\vec{x} = \vec{0}$  y  $(A - 6I)\vec{x} = \vec{0}$  se obtienen autovectores que dan bases de los núcleos, por ejemplo:

$$(15) \quad \text{Ker}(A - I) \rightarrow \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{2}/2 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \text{y} \quad \text{Ker}(A - 6I) \rightarrow \left\{ \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ 1 \end{pmatrix} \right\}.$$

Los vectores de  $\text{Ker}(A-I)$  son necesariamente ortogonales a los de  $\text{Ker}(A-6I)$  por (11) pero los vectores elegidos como base para el primer subespacio no lo son entre ellos. Con el proceso de Gram-Schmidt, podemos sustituir el segundo por

$$(16) \quad \begin{pmatrix} -\sqrt{2}/2 \\ 0 \\ 1 \end{pmatrix} - \frac{\sqrt{2}/2}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\sqrt{2}/4 \\ -\sqrt{2}/4 \\ 1 \end{pmatrix}.$$

Con estos tres vectores tenemos una base ortogonal de  $\mathbb{R}^3$  formada por vectores propios de  $A$ , que ortonormalizada (dividiendo cada vector por su norma) es

$$(17) \quad B = \left\{ \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{2}/2\sqrt{5} \\ -\sqrt{2}/2\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}, \begin{pmatrix} \sqrt{2}/\sqrt{5} \\ \sqrt{2}/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \right\}.$$

La matriz  $C$  de cambio de base de  $B$  a la canónica es la matriz (ortogonal) que tiene como columnas estos vectores.

### 3. Exponenciales de matrices

La diagonalización, cuando es posible, produce una simplificación en nuestra manera de ver una aplicación lineal que es de gran importancia en muchas situaciones. Una de ellas es calcular potencias de matrices. Si sabemos que una matriz  $A$  se diagonaliza como  $A = CDC^{-1}$  entonces para  $k \in \mathbb{Z}^+$

$$(18) \quad A^k = CDC^{-1} \cdot CDC^{-1} \cdot \dots \cdot CDC^{-1} \cdot CDC^{-1} = CD^kC^{-1}$$

porque cada una de las matrices de  $C$  interiores se cancela con su inversa. El cálculo de  $D^k$  es trivial frente al cálculo directo de  $A^k$  que en general parece inasequible. Una aplicación típica es la solución de ecuaciones de recurrencia que se explora a continuación en un ejemplo famoso.

**Ejemplo.** La *sucesión de Fibonacci* es  $\{F_n\}_{n=0}^{\infty}$  donde  $F_0 = 0$ ,  $F_1 = 1$  y en general  $F_{n+2} = F_{n+1} + F_n$ . Los primeros elementos son 0, 1, 1, 2, 3, 5, 8, 13, 21... y no es difícil sospechar un crecimiento exponencial. Veamos cómo las potencias de matrices permiten dar una fórmula explícita para  $F_n$ . La relación que define  $F_n$  se escribe matricialmente como

$$(19) \quad \vec{x}_{n+1} = A\vec{x}_n \quad \text{con} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{y} \quad \vec{x}_n = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix}.$$

Por tanto se tiene la fórmula  $\vec{x}_n = A^n \vec{x}_0$ . Los autovalores de  $A$  son  $\lambda_{\pm} = (1 \pm \sqrt{5})/2$  y se tiene

$$(20) \quad A^n = \begin{pmatrix} \lambda_+ & \lambda_- \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}^n \begin{pmatrix} \lambda_+ & \lambda_- \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_+ & \lambda_- \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix} \begin{pmatrix} 1 & -\lambda_- \\ -1 & \lambda_+ \end{pmatrix}.$$

De  $A^n \vec{x}_0$ , recordando  $\vec{x}_0 = (1, 0)^t$ , se sigue que  $F_n$  es el elemento inferior izquierdo de  $A^n$  y así operando se deduce

$$(21) \quad F_n = \frac{\lambda_+^n - \lambda_-^n}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right).$$

Como  $-1 < \lambda_- < 0$ , para  $n$  grande  $\lambda_-^n$  es despreciable.

Este triunfo con potencias arbitrarias abre la puerta a tratar de definir funciones más generales de matrices. A pesar de que esto parece muy artificial y abstracto, en física es absolutamente crucial considerar exponenciales de matrices (e incluso de operadores en dimensión infinita).

La *exponencial de una matriz* cuadrada  $A$  se define como

$$(22) \quad \exp(A) = \text{Id} + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

Para cada elemento se tiene una serie que parece complicadísima y además no parece nada evidente que la convergencia esté garantizada. Sin embargo para una matriz diagonal digamos  $D = \text{diag}(d_1, \dots, d_n)$ , donde  $d_j$  son los elementos de la diagonal, es muy fácil ver con esta definición que  $\exp(D) = \text{diag}(e^{d_1}, \dots, e^{d_n})$ . Gracias a (18), si  $A$  es diagonalizable tendremos  $\exp(A) = C \exp(D) C^{-1}$ .

En física aparecen sobre todo exponenciales de matrices hermíticas multiplicadas por  $i$ . Digamos que  $H$  es hermítica y diagonaliza como  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  en una base ortonormal, entonces  $\exp(iH)$  en esa misma base tendrá una matriz  $\text{diag}(e^{i\lambda_1}, \dots, e^{i\lambda_n})$  que claramente cumple que su inversa es su traspuesta conjugada. Por tanto es unitaria. De hecho aplicando la operación  $H \mapsto \exp(iH)$  obtenemos todas las matrices unitarias a partir de las hermíticas. En física esta es la base para pasar de Hamiltonianos a operadores de evolución. El análogo en mecánica clásica sería pasar de la energía a la ecuación de movimiento.

**Ejemplo.** La evolución a lo largo del tiempo de la molécula de amoníaco descrita por (12) viene dada por  $U(t) = \exp(-itH)$ . Para calcular esta matriz, basta usar (13) para obtener

$$(23) \quad U(t) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e^{-it(E_0+\Delta)} & 0 \\ 0 & e^{-it(E_0-\Delta)} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = e^{-itE_0} \begin{pmatrix} \cos(\Delta t) & -i \text{sen}(\Delta t) \\ -i \text{sen}(\Delta t) & \cos(\Delta t) \end{pmatrix}.$$

Está claro que la matriz es unitaria. Físicamente esto implica que si el nitrógeno está inicialmente arriba, la probabilidad de que en un tiempo  $t$  lo midamos arriba o abajo es  $\cos^2(\Delta t/\hbar)$  y  $\text{sen}^2(\Delta t/\hbar)$ . Las oscilaciones están ligadas a la teoría del *máser*, el antecesor del *láser*.

En esencia, hay dos razones (relacionadas aunque no lo parezca) por las que  $\exp(A)$  aparece con frecuencia en física, ya dentro del grado. La primera es que la ecuación de Schrödinger

lleva a estudiar cierto tipo de ecuaciones diferenciales para las cuales la relación

$$(24) \quad \frac{d}{dt} \exp(tA) = A \exp(tA)$$

es un buen modelo. La segunda razón es más abstracta y es que la exponencial permite en algunos casos notables establecer un vínculo entre grupos de transformaciones (“difícil”, nivel de tercero) y espacios vectoriales (“fácil”, nivel de primero). Lo dicho antes sobre las matrices hermíticas y las unitarias es un ejemplo de ello, las primeras forman un espacio vectorial sobre  $\mathbb{R}$  sin embargo la suma de matrices unitarias no puede ser unitaria.

**Ejemplo.** La ecuación diferencial de la primera sección se reescribe como

$$(25) \quad \frac{d}{dt} \vec{x} = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 7 & -10 \\ 3 & -4 \end{pmatrix}, \quad \vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \quad \text{y} \quad \vec{x}(0) = \begin{pmatrix} 7 \\ 4 \end{pmatrix}.$$

Según la relación (24),  $\vec{x}(t) = \exp(tA)\vec{x}(0)$  resuelve esta ecuación. Como  $\{(2, 1)^t, (5, 3)^t\}$  es una base de vectores propios con valores propios respectivos 2 y 1, se sigue

$$(26) \quad \vec{x}(t) = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 7 \\ 4 \end{pmatrix} = \begin{pmatrix} 2e^{2t} + 5e^t \\ e^{2t} + 3e^t \end{pmatrix}.$$

**Observación.** La ecuación de Schrödinger  $i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi$  formalmente tiene como solución  $\Psi(t) = \exp(-it\hat{H}/\hbar)\Psi(0)$  cuando  $\hat{H}$  no depende de  $t$ . El problema es que la dimensión es infinita en muchos casos y  $\hat{H}$  está lejos de ser una matriz. El modelo de la molécula de amoníaco es un ejemplo en dimensión 2 de lo que se llama *sistema de dos estados*. Si  $\hat{H}$  depende de  $t$  incluso el caso matricial es complicado, por ejemplo es el que aparece en las resonancias magnéticas, porque  $\hat{H}$  refleja que hay un campo magnético variable externo que atraviesa a la muestra (al paciente en el caso médico). En ese caso se hace un cambio de base previo que depende del tiempo, algo así como los sistemas de referencia rotatorios de la *fuerza de Coriolis*.

## 4. El teorema de Jordan

Según hemos visto, con propiedades adicionales sobre la aplicación lineal relacionadas con el producto escalar, asegurábamos que era diagonalizable. Volvamos ahora al caso general de dimensión finita  $L : V \rightarrow V$  donde no se supone en  $V$  un producto escalar ni propiedades de  $L$ . Para que no haya problema con la existencia de autovalores pensemos que trabajamos con  $\mathbb{K} = \mathbb{C}$  y no hay restricción de generalidad en considerar  $V = \mathbb{C}^n$ . Fijada la base canónica, identificamos la aplicación y su matriz, como es habitual.

Ya habíamos visto que hay matrices que no son diagonalizables. El *teorema de Jordan* afirma que siempre se puede conseguir en cierta base una matriz “casi diagonal”  $D$  llamada



la *forma canónica de Jordan* en la que aparecen unos fuera de la diagonal en los lugares  $d_{i i+1}$  siempre intercalados entre autovalores iguales. Concretamente, que existe una base en la que la matriz de la aplicación es una matriz diagonal por bloques de la forma:

$$(27) \quad \begin{pmatrix} J_{n_1}(\lambda_1) & & & & \\ & J_{n_2}(\lambda_2) & & & \\ & & J_{n_3}(\lambda_3) & & \\ & & & \ddots & \\ & & & & J_{n_k}(\lambda_k) \end{pmatrix} \quad \text{donde} \quad J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

En la matriz, los lugares vacíos indican bloques nulos. A  $J_n(\lambda)$  se le llama *celda de Jordan* o *matriz elemental de Jordan*. Aquí los  $n_j$  y los  $\lambda_j$  podrían tomar valores iguales. La suma de todos los  $n_r$  que corresponden a un mismo autovalor  $\lambda_j$  dará la multiplicidad de  $\lambda_j$  como cero del polinomio característico mientras que  $\dim \text{Ker}(L - \lambda_j \text{Id})$  es el número de celdas en las que aparece un mismo  $\lambda_j$ .

Por la forma diagonal por bloques de (27), cada celda está asociada a un *subespacio invariante*, es decir, tal que  $L(W) \subset W$ , de hecho  $L(W) = W$  si  $\lambda \neq 0$ . De esta forma,  $V = \bigoplus W_j$  con  $W_j$  subespacios invariantes lo más pequeños posibles de forma que se tenga la igualdad.

Observación. Los subespacios que quedan invariantes por muchas (un grupo de) aplicaciones aparecen en el concepto de *representación* que es un instrumento matemático bastante avanzado que ha sido empleado por los físicos teóricos para clasificar las partículas.

Una manera muy sintética de enunciar el teorema evitando lo aparatoso de las matrices (27) es definir para  $L : V \rightarrow V$  un *autovector generalizado* para cierto autovalor  $\lambda$  como un vector  $\vec{v} \in \text{Ker}(L - \lambda \text{Id})^k - \{\vec{0}\}$  para algún  $k \in \mathbb{Z}^+$ . El teorema de Jordan es equivalente (de una manera no obvia) a decir que siempre existe una base de  $V$  formada por autovectores generalizados de  $L$ .

Calcular la base que lleva a la forma canónica de Jordan puede ser muy trabajoso en cuanto la dimensión crece un poco. Vamos a tratar de entender los casos de dimensión 2 y 3 de los cuales se infiere el procedimiento general. Según (27), las posibilidades en dimensión 2 son:

$$(28) \quad \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

El primer caso es diagonalizable (sea  $\lambda_1$  igual a  $\lambda_2$  o no). En dimensión 3, aparte del caso diagonalizable, aparecen otros dos más:

$$(29) \quad \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

La prueba habitual del teorema de Jordan es larga y no la daremos aquí<sup>1</sup> solo veremos cómo se probaría cuando  $\dim V = 2$ . En este caso, llamemos  $\lambda_1$  y  $\lambda_2$  a los autovalores repetidos con su multiplicidad. Si  $\lambda_1 \neq \lambda_2$ ,  $L$  es diagonalizable porque hay al menos un autovector por autovalor y no puede ser uno múltiplo del otro. Si  $\dim \text{Ker}(L - \lambda_1 \text{Id}) = 2$ , lo cual ocurre solo si  $L = \lambda_1 \text{Id}$ , tenemos suficientes autovectores y volvemos al primer caso de (28). Entonces  $L$  no es diagonalizable si y solo si  $\lambda_1 = \lambda_2$  y  $\dim \text{Ker}(L - \lambda_1 \text{Id}) = 1$ . Sea  $\vec{v}$  un autovector y  $\vec{w}$  un vector cualquiera linealmente independiente, entonces en la base  $\{\vec{v}, \vec{w}\}$  la matriz de la aplicación debe ser de la forma

$$(30) \quad \begin{pmatrix} \lambda_1 & a \\ 0 & b \end{pmatrix} \quad \text{con} \quad b = \lambda_1$$

porque  $b \neq \lambda_1$  implicaría  $\text{Ker}(L - b \text{Id}) \neq \{\vec{0}\}$  y tendríamos un autovalor distinto de  $\lambda_1$ . Un cálculo con esta matriz prueba que  $(L - \lambda_1 \text{Id})^2$  es la aplicación nula. Tomando  $\vec{v}_2 \notin \text{Ker}(L - \lambda_1 \text{Id})$  y definiendo  $\vec{v}_1 = (L - \lambda_1 \text{Id})\vec{v}_2$ , se tiene  $\vec{v}_1 \neq \vec{0}$  y  $(L - \lambda_1 \text{Id})\vec{v}_1 = \vec{0}$ , es decir,  $\vec{v}_1$  es autovector. Estas ecuaciones se pueden reescribir como

$$(31) \quad L(\vec{v}_1) = \lambda_1 \vec{v}_1 \quad \text{y} \quad L(\vec{v}_2) = \lambda_1 \vec{v}_2 + \vec{v}_1,$$

lo que significa que en la base  $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$  la matriz de  $L$  es la segunda en (28).

La idea para hallar la forma canónica de Jordan y el cambio de base correspondiente radica en que se puede comprobar que

$$(32) \quad \dim \text{Ker}(J_n(\lambda) - \lambda I)^k = k \quad \text{para} \quad 0 < k \leq n.$$

así pues, como en el caso  $n = 2$ , cualquier matriz  $A$  que tenga una sola celda de Jordan adquiere su forma canónica usando la base  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$  con

$$(33) \quad \vec{v}_j = (A - \lambda I)^{n-j} \vec{v}_n \quad \text{donde} \quad \vec{v}_n \in \text{Ker}(A - \lambda I)^n - \text{Ker}(A - \lambda I)^{n-1}.$$

En el caso de una sola celda de Jordan,  $\text{Ker}(A - \lambda I)^n$  es el espacio total,  $(A - \lambda I)^n$  es la aplicación nula. En el caso en que hay varias celdas de Jordan no es necesariamente así, no obstante, seguiremos utilizando (33) para encontrar bases correspondientes a cada una de ellas. La mayor precaución que hay que tener en el caso general es no mezclar vectores correspondientes a diferentes celdas de dimensión mayor que 1 con el mismo  $\lambda$ . En (28) y (29) esto no ocurre. Más adelante veremos un ejemplo en dimensión 4. La condición que generaliza la elección de  $\vec{v}_n$  en (33) cuando hay varias cajas de la misma dimensión correspondientes

---

<sup>1</sup>Hay una prueba breve debida a Lucía Contreras, antigua profesora de la UAM, ya jubilada, que procede por inducción utilizando que  $W = \text{Im}(L - \lambda_1 \text{Id})$  tiene dimensión menor que  $\dim V$  y la forma canónica de Jordan en  $W$  permite construir la de  $L$  en  $V$ .

a un mismo autovalor es que siempre hay que escoger vectores que forman una base de un subespacio  $W$  con  $\text{Ker}(A - \lambda I)^n = W \oplus \text{Ker}(A - \lambda I)^{n-1}$ . en el caso de (33), necesariamente  $\dim W = 1$  y por eso cualquier vector de  $\text{Ker}(A - \lambda I)^n$  que no esté en  $\text{Ker}(A - \lambda I)^{n-1}$  sirve para construir  $W$ .

Dicho sea de paso, del hecho de que cada celda de Jordan cumpla que  $(J_n(\lambda) - \lambda I)^n$  sea la aplicación nula se deduce el curioso resultado de que una matriz sustituida en su polinomio característico siempre da la matriz nula porque se puede escribir  $p(A) = \prod(\lambda_i I - A)^{n_i}$  con cualquier orden de los términos del producto y en la base que da la forma canónica de Jordan  $J$  esto es  $\prod(\lambda_i I - J)^{n_i}$ . Este resultado,  $p(A) = 0$ , se conoce con el nombre de *teorema de Cayley-Hamilton*. En realidad se usa habitualmente en la prueba del teorema de Jordan y por tanto la deducción anterior no es una demostración genuina.

Sabíamos calcular potencias arbitrarias y exponenciales de matrices diagonalizables. La forma canónica de Jordan permite extender estos cálculos al caso no diagonalizable aunque las fórmulas son más complicadas. Veremos solo un ejemplo en dimensión 2.

**Ejemplo.** Consideramos la ecuación diferencial  $x'' - 2x' + x = 0$ ,  $x(0) = 1$ ,  $x'(0) = 0$ . definiendo  $y' = x$  se puede escribir matricialmente como:

$$(34) \quad \vec{x}' = A\vec{x}, \quad \vec{x}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{donde} \quad \vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \quad \text{y} \quad A = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

Según lo visto antes, la fórmula para la solución es  $\vec{x} = \exp(tA)\vec{x}(0)$ . El problema es que  $A$  no es diagonalizable y por tanto no está claro cómo hallar  $\exp(tA)$ . Escribiendo la forma canónica de Jordan de  $A$  es fácil hallar sus potencias con un sencillo proceso inductivo:

$$(35) \quad A = C \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} C^{-1} \quad \text{con} \quad C = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{implica} \quad A^n = C \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} C^{-1}.$$

De esta forma, sustituyendo en la definición (22) se obtiene

$$(36) \quad \exp(tA) = C \begin{pmatrix} \sum \frac{1}{n!} t^n & \sum \frac{n}{n!} t^n \\ 0 & \sum \frac{1}{n!} t^n \end{pmatrix} C^{-1} = C \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} C^{-1} = \begin{pmatrix} (1-t)e^t & te^t \\ -te^t & (t+1)e^t \end{pmatrix}$$

y se deduce  $(x(t), y(t)) = ((1-t)e^t, -te^t)$ .

Ahora vamos con un surtido de ejemplos de formas canónicas.

**Ejemplo.** Calculemos la forma canónica de Jordan de

$$(37) \quad A = \begin{pmatrix} -1 & 0 & -2 \\ 0 & -1 & -3 \\ 2 & -1 & -3 \end{pmatrix} \quad \text{que cumple} \quad |A - \lambda I| = -(\lambda + 2)^2(\lambda + 1).$$

Se verifica  $\dim \text{Ker}(A + 2I) = \dim \text{Ker}(A + I) = 1$  (esto es equivalente a que el rango de  $A + 2I$  y el de  $A + I$  es dos), por tanto solo tenemos un autovector y sus múltiplos para  $\lambda = -2$  y otro para  $\lambda = -1$ ,

así que  $A$  no es diagonalizable. Necesariamente la forma canónica de Jordan es la segunda de (29) con  $\lambda_1 = -2$  y  $\lambda_2 = -1$ . La base será  $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$  donde  $\vec{v}_3$  es autovector de  $\lambda_2 = -1$  y  $\vec{v}_1 = (A + 2I)\vec{v}_2$  con  $\vec{v}_2 \in \text{Ker}(A - \lambda I)^2 - \text{Ker}(A - \lambda I)$  según (33). Los cálculos relevantes son

$$(38) \quad A + 2I = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 2 & -1 & -1 \end{pmatrix}, \quad (A + 2I)^2 = \begin{pmatrix} -3 & 2 & 0 \\ -6 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A + 2I = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & -3 \\ 2 & -1 & -2 \end{pmatrix}.$$

Del último se sigue que  $\vec{v}_3 = (1, 2, 0)^t$  es una elección válida. También lo es  $\vec{v}_2 = (2, 3, 0)^t$  que cumple  $(A + 2I)^2\vec{v}_2 = \vec{0}$  pero  $(A + 2I)\vec{v}_2 \neq \vec{0}$  (por supuesto hay muchas otras posibilidades), lo que lleva a  $\vec{v}_1 = (A + 2I)\vec{v}_2 = (2, 3, 1)^t$ . con todo ello

$$(39) \quad C^{-1}AC = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{donde } C = \begin{pmatrix} 2 & 2 & 1 \\ 3 & 3 & 2 \\ 1 & 0 & 0 \end{pmatrix}.$$

Veamos un caso similar pero con las dos celdas correspondientes al mismo valor propio.

**Ejemplo.** Partimos ahora de

$$(40) \quad A = \begin{pmatrix} 3 & -4 & 4 \\ 2 & -3 & 4 \\ 1 & -2 & 3 \end{pmatrix} \quad \text{que tiene } |A - \lambda I| = (1 - \lambda)^3.$$

Se comprueba que  $\dim \text{Ker}(A - I) = 2$ , por tanto hay dos vectores propios linealmente independientes y la forma canónica de Jordan es de nuevo la segunda de (29) pero ahora con  $\lambda_1 = \lambda_2 = 1$ . Sin hacer cálculos se sigue  $\dim \text{ker}(A - I)^2 = 3$ , es todo el espacio, y podemos elegir una base  $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$  con  $\vec{v}_1 = (A - I)\vec{v}_2$ ,  $\vec{v}_3$  autovector y  $\vec{v}_2$  arbitrario mientras no esté en el subespacio  $\text{Ker}(A - I)$ , que es el generado por los autovectores. Se tiene

$$(41) \quad A - I = \begin{pmatrix} 2 & -4 & 4 \\ 2 & -4 & 4 \\ 1 & -2 & 2 \end{pmatrix} \quad \text{y entonces } \text{Ker}(A - I) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x - 2y + 2z = 0 \right\}.$$

Tomemos por ejemplo  $\vec{v}_2 = (1, 0, 0)^t$ , de este modo  $\vec{v}_1 = (A - I)\vec{v}_2 = (2, 2, 1)^t$  y  $\vec{v}_3$  es cualquier otro vector propio que no sea múltiplo de  $\vec{v}_1$ , por ejemplo  $\vec{v}_3 = (2, 1, 0)^t$ . En definitiva se tiene

$$(42) \quad C^{-1}AC = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{donde } C = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

El caso con una sola celda de Jordan es algo más sencillo.

**Ejemplo.** Para la matriz

$$(43) \quad A = \begin{pmatrix} -2 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & -1 & -1 \end{pmatrix} \quad \text{se verifica } |A - \lambda I| = -(\lambda + 1)^3.$$

Como  $\dim \text{Ker}(A + I) = 1$ , porque  $\text{rg}(A + I) = 2$ , la forma canónica de Jordan es del tercer tipo de (29) con  $\lambda_1 = -1$ . La base será  $\{(A + I)^2 \vec{v}, (A + I) \vec{v}, \vec{v}\}$ , según (33), con  $\vec{v}$  fuera del núcleo de  $A + I$ . eligiendo por ejemplo  $\vec{v} = (1, 0, 0)^t$  se tiene  $(A + I) \vec{v} = (-1, -1, 2)^t$  y  $(A + I)^2 \vec{v} = (0, 0, -1)^t$ .

Veamos un ejemplo en dimensión mayor para ilustrar que el algoritmo es similar.

**Ejemplo.** Partimos de

$$(44) \quad A = \begin{pmatrix} 0 & 2 & 1 & 4 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 1 & 0 & 4 \end{pmatrix} \quad \text{que cumple} \quad |A - \lambda I| = (\lambda - 2)^4.$$

Un cálculo prueba que  $(A - 2I)^2$  es la matriz nula y

$$(45) \quad A - 2I = \begin{pmatrix} -2 & 2 & 1 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 2 \end{pmatrix} \Rightarrow \text{Ker}(A - 2I) = \{(x_1, x_2, x_3, x_4)^t \in \mathbb{R}^4 : -x_1 + x_2 + 2x_4 = 0, x_3 = 0\}.$$

Escojamos  $\vec{v}_2, \vec{v}_4 \in \mathbb{R}^4 - \text{Ker}(A - 2I) = \text{Ker}(A - 2I)^2 - \text{Ker}(A - 2I)$  que generen un subespacio de dimensión 2 que tenga intersección trivial con  $\text{Ker}(A - 2I)$ , entonces es seguro que  $\vec{v}_1 = (A - 2I) \vec{v}_2$  y  $\vec{v}_3 = (A - 2I) \vec{v}_4$  son autovectores linealmente independientes y se cumple  $A \vec{v}_2 = 2 \vec{v}_2 + \vec{v}_1$ ,  $A \vec{v}_4 = 2 \vec{v}_4 + \vec{v}_3$ . Tomando por ejemplo  $\vec{v}_2 = (0, 0, 1, 0)^t$  y  $\vec{v}_4 = (1, 2, 0, 0)^t$ , se obtiene  $\vec{v}_1 = (1, 1, 0, 0)^t$  y  $\vec{v}_3 = (2, 0, 0, 1)^t$ . De esta forma

$$(46) \quad C^{-1}AC = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{con} \quad C = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

La forma canónica de Jordan necesita que el polinomio característico se descomponga en factores lineales, por tanto está restringida al caso  $\mathbb{K} = \mathbb{C}$ . Sin embargo a veces es conveniente tener algo parecido cuando  $\mathbb{K} = \mathbb{R}$ , lo que se llama la *forma de Jordan real*. Si la matriz es real los autovalores complejos no reales aparecerán por parejas de complejos conjugados y en  $\mathbb{C}$  los autovectores también se pueden elegir conjugados. Consideremos el caso  $2 \times 2$ . Digamos que  $\lambda_1 = a + bi$  y  $\lambda_2 = a - bi$  con  $b \neq 0$  y autovectores respectivos  $\vec{v}$  y  $\bar{\vec{v}}$ , entonces se prueba que

$$(47) \quad A \quad \text{se convierte en} \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{al cambiar a la base} \quad \{\Re \vec{v}, \Im \vec{v}\},$$

donde  $\Re$  e  $\Im$  indican parte real e imaginaria, con lo cual es una base real.

En consecuencia, en el caso real  $n \times n$  diagonalizable en los complejos, siempre se puede hallar una base real en la que se convierte en una matriz diagonal por bloques donde los

bloques son de dimensión 1, correspondientes a un autovalor real o de dimensión 2 como (47), correspondientes a autovalores complejos conjugados.

**Ejemplo.** Ya vimos en un ejemplo anterior que la aplicación  $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por  $L((x, y)^t) = (-3x - 2y, 5x + 3y)$  no es diagonalizable porque su polinomio característico es  $\lambda^2 + 1$ . En  $\mathbb{C}^2$  sin embargo sí era diagonalizable con una base de vectores propios  $\{\vec{v}, \bar{\vec{v}}\}$  donde  $\vec{v} = (2, -3 - i)^t$ , correspondientes a  $\lambda_1 = i$  y  $\lambda_2 = \bar{\lambda}_1$ . Tomando las partes reales e imaginarias de  $\lambda_1$  y de las coordenadas de  $\vec{v}$  se llega a la forma de Jordan real que corresponde a la identidad matricial

$$(48) \quad \begin{pmatrix} 2 & 0 \\ -3 & -1 \end{pmatrix}^{-1} \begin{pmatrix} -3 & -2 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -3 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$