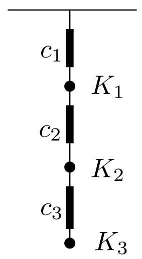
E Eigenwerte

Drei gleichschwere Kugeln K_1, K_2, K_3 mit der Masse 1 seien wie auf dem Bild durch Federn verbunden aufgehängt. Die Federkonstanten - sie geben an, wie stark die Feder ist - seien c_1, c_2, c_3. Wie schwingt dieses System, wenn man es einmal anstößt und wir annehmen, daß keine Dämpfung eintritt? Um es einfacher zu halten, betrachten wir nur genau senkrecht verlaufende Schwingungen.



Ist $y_i(t)$ die Auslenkung aus der Ruhelage der i-ten Kugel zur Zeit t, dann erhält man folgendes System von Differentialgleichungen für die y_i :

$$\begin{aligned} -y_1'' &= c_1 y_1 - c_2(y_2 - y_1) \\ -y_2'' &= c_2(y_2 - y_1) - c_3(y_3 - y_2) \\ -y_3'' &= c_3(y_3 - y_2) \end{aligned}$$

Da ungedämpfte Schwingungen schön Sinus-artig verlaufen, machen wir folgenden Ansatz:

$$y_1(t) := \xi_1 \sin \omega t,\ y_2(t) := \xi_2 \sin \omega t,\ y_3(t) := \xi_3 \sin \omega t.$$

Wegen

$$(\sin \omega t)'' = -\omega^2 \sin \omega t$$

wird unser System von Differentialgleichungen damit überführt in das System

$$\begin{aligned} \xi_1 \omega^2 \sin \omega t &= \big((c_1 + c_2)\xi_1 - c_2 \xi_2\big) \sin \omega t \\ \xi_2 \omega^2 \sin \omega t &= \big(-c_2 \xi_1 + (c_2 + c_3)\xi_2 - c_3 \xi_3\big) \sin \omega t \\ \xi_3 \omega^2 \sin \omega t &= \big(-c_3 \xi_2 + c_3 \xi_3\big) \sin \omega t, \end{aligned}$$

und hierin sind die Amplituden ξ_1, ξ_2, ξ_3 (nicht alle $= 0$) und die Frequenz ω zu bestimmen. Damit ist also zu lösen die Aufgabe

$$\begin{pmatrix} c_1 + c_2 & -c_2 & 0 \\ -c_2 & c_2 + c_3 & -c_3 \\ 0 & -c_3 & c_3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \lambda \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \qquad (\lambda = \omega^2.)$$

Gesucht ist also ein Vektor $x \neq 0$, der durch die gegebene Matrix A in ein Vielfaches von sich abgebildet wird, wobei aber der Faktor λ nicht bekannt ist. Man nennt dann x einen "Eigen-Vektor", λ den zugehörigen "Eigen-Wert".
Jeder Eigenwert $\lambda = \omega^2$ liefert in diesem Beispiel eine "Eigenfrequenz" des Systems und der zugehörige Eigenvektor x liefert die Amplituden der auftretenden Schwingungen. Jede entsprechend der Art, wie wir das System anstoßen, auftretende Schwingung des Gesamtsystems ergibt sich durch Überlagerung dieser Eigenschwingungen, d.h. ist als Linearkombination dieser aus dem Eigenwertproblem gewonnenen Sinus- und analoger Cosinus- Lösungen darstellbar.
Durch die Federn wird hier auf die Kugeln eine Kraft ausgeübt, die umso größer ist, je größer die Auslenkung aus der Ruhelage ist, und dadurch die Auslenkung

selbst korrigiert. Diese Situation, daß die Änderung einer Größe abhängig von der Größe selbst ist, finden Sie in allen Bereichen quantitativen Beschreibens und häufig resultiert daraus ein Eigenwertproblem.
Wir machen daraus die folgende Definition. K bezeichet wieder die reellen oder die komplexen Zahlen.

Definition E.1 (Eigenwert, Eigenvektor) *Zu einem Endomorphismus f in einem K-Vektorraum U heißt ein Element $x \in U$ "Eigenvektor" zum "Eigenwert" $\lambda \in K$ von f, wenn*

$$x \neq 0 \quad \text{und} \quad f(x) = \lambda x.$$

Die Eigenvektoren und Eigenwerte einer Matrix $A \in K^{n \times n}$ sind über den durch $x \mapsto Ax$ gegebenen Endomorphismus des K^n erklärt, d.h.
$x \in K^n$ ist "Eigenvektor" zum "Eigenwert" $\lambda \in K$ der Matrix A, wenn

$$x \neq 0 \quad \text{und} \quad Ax = \lambda x.$$

Im Englischen heißen diese Begriffe "eigenvector" bzw. "eigenvalue".
Für $\alpha \neq 0$ ist natürlich mit x auch αx Eigenvektor zum selben Eigenwert.

Wir beschränken uns im Weiteren auf die Behandlung von Eigenwerten zu Matrizen.

Ist λ Eigenwert zu A, so besitzt also die Gleichung $Ax = \lambda x$ eine nichttriviale Lösung. Dafür kann man auch (mit der Einheitsmatrix I) schreiben, $(A - \lambda I)x = 0$ besitzt eine nichttriviale Lösung oder $\ker(A - \lambda I) \neq 0$. Letzteres ist aber gleichbedeutend damit, daß $\det(A - \lambda I) = 0$.
Wir haben damit

Satz und Definition E.2 (Eigenraum) *Eine Zahl λ ist genau dann Eigenwert zur Matrix A, wenn $\det(A - \lambda I) = 0$, und jedes $x \neq 0$, $x \in \ker(A - \lambda I)$ ist Eigenvektor von A zum Eigenwert λ.*
$E(A; \lambda) := \ker(A - \lambda I)$ heißt "Eigenraum" zum Eigenwert λ.

Die Determinante der Matrix $(A - \lambda I)$ können wir aber nach den Sätzen des vorigen Kapitels berechnen.

Lemma E.3 *Sind in einer $n \times n$-Matrix F die Elemente φ_{ij} Polynome in einer Variablen λ und haben die Polynome in der j-ten Spalte einen Grad $\leq \gamma_j$, $(j = 1, ..., n)$, so ist die Determinante von F ein Polynom vom Grad $\leq \gamma_1 + \gamma_2 + ... + \gamma_n$ in λ.*

Beweis: Nach der Formel von LEIBNIZ (Satz D.22) ist $\det F$ eine Summe von Produkten, in denen aus jeder Zeile und jeder Spalte genau ein Element vorkommt, sodaß der Gesamtgrad solcher Produkte und damit auch der Determinante $\leq \gamma_1 + \gamma_2 + ... + \gamma_n$ ist.

□

Die Matrix $A - \lambda I$ hat mit $A = (\alpha_{ij})$ die Form

$$A - \lambda I = \begin{pmatrix} \alpha_{11} - \lambda & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} - \lambda & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} - \lambda \end{pmatrix},$$

d.h. sämtliche vorkommenden Polynome sind vom Grad ≤ 1 in λ.
Damit ergibt sich

Satz und Definition E.4 (Charakteristisches Polynom) *Zu einer $n \times n$ Matrix A ist*

$$\chi_A(\lambda) := \det(A - \lambda I)$$

ein Polynom vom Grad n, das "charakteristische Polynom" der Matrix A.
Sein Hauptkoeffizient ist $(-1)^n$, der nächste ist bis auf das Vorzeichen die Summe der Diagonalelemente, genannt "Spur" von A, also $(-1)^{n-1}\,\mathrm{spur}(A) := (-1)^{n-1}\sum_{i=1}^n \alpha_{ii}$, der Absolutkoeffizient ist $\det A$.
Ähnliche Matrizen haben dasselbe charakteristische Polynom.

Beweis: Die Gradaussage folgt aus Lemma E.3, die beiden oberen Koeffizienten erhält man per Koeffizientenvergleich aus der LAPLACE-Entwicklung (damit kann man auch die anderen berechnen), den Absolutkoeffizienten bekommt man durch Einsetzen von $\lambda = 0$.
Eine zu A ähnliche Matrix hat mit einer invertierbaren Matrix S die Gestalt $A' = S^{-1}AS$. Dafür ist dann

$$\begin{aligned}\chi_{A'}(\lambda) &:= \det(A' - \lambda I) = \det(S^{-1}AS - \lambda S^{-1}IS) = \det(S^{-1}(A - \lambda I)S) \\ &= \det(S^{-1})\det(A - \lambda I)\det(S) = \det(A - \lambda I) = \chi_A(\lambda).\end{aligned}$$

□

Den Satz und Definition E.2 können wir nun auch so formulieren:

Korollar E.5 *λ ist genau dann Eigenwert zu einer Matrix A, wenn λ Nullstelle des charakteristischen Polynoms $\chi_A(t)$ ist.*

Das Polynom $t^2 + 1$ besitzt in $\mathbb{R}$ keine Nullstelle, wohl aber jedes Polynom dessen Grad eine ungerade Zahl ist. Dagegen besitzt in $\mathbb{C}$ jedes nichtkonstante Polynom eine Nullstelle.
Damit haben wir

Satz E.6 *Jede $n \times n$ - Matrix A besitzt über $\mathbb{C}$ mindestens einen Eigenwert λ, über $\mathbb{R}$ sicher dann einen Eigenwert, wenn n eine ungerade Zahl ist.*

Untersuchen wir nun Eigenvektoren.

Lemma E.7 *Sind $\lambda_1, ..., \lambda_m$ verschiedene Eigenwerte zu A, und $u_1...., u_m$ zugehörige Eigenvektoren, so sind diese linear unabhängig.*

Beweis: mit Induktion.
Sei $m = 1$: Als Eigenvektor ist $u_1 \neq 0$ und damit die von ihm alleine gebildete Familie (u_1) unabhängig.
$m > 1$: Seien nun $(u_1...., u_{m-1})$ unabhängig. Wir untersuchen eine Linearkombination

$$\alpha_1 u_1 + ... + \alpha_{m-1}u_{m-1} + \alpha_m u_m = 0$$

mit $\alpha_i \in K$. Nach Multiplikation mit λ_m ist auch

$$\lambda_m\alpha_1 u_1 + ... + \lambda_m\alpha_{m-1}u_{m-1} + \lambda_m\alpha_m u_m = 0,$$

nach Anwenden von A auf die erste Gleichung ist auch

$$\alpha_1 Au_1 + ... + \alpha_{m-1}Au_{m-1} + \alpha_m Au_m = 0.$$

Da u_i Eigenvektor zum Eigenwert λ_i ist, d.h. $Au_i = \lambda_i u_i$, können wir die letzte Gleichung auch schreiben als

$$\lambda_1\alpha_1 u_1 + ... + \lambda_{m-1}\alpha_{m-1}u_{m-1} + \lambda_m\alpha_m u_m = 0.$$

Subtrahieren wir beide Linearkombinationen voneinander, so folgt

$$(\lambda_m - \lambda_1)\alpha_1 u_1 + ... + (\lambda_m - \lambda_{m-1})\alpha_{m-1} u_{m-1} = 0$$

und da $(u_1 ...., u_{m-1})$ unabhängig sind und $\lambda_m - \lambda_i \neq 0$ für $i < m$, ist notwendig $\alpha_1 = ... = \alpha_{m-1} = 0$ und dann aber auch $\alpha_m = 0$, was die Behauptung zeigt. □

Als Folgerung erhalten wir

Satz E.8 *Besitzt eine Matrix A genau n verschiedene Eigenwerte $\lambda_1, ..., \lambda_n$, so bilden die zugehörigen Eigenvektoren $(u_1 ...., u_n)$ eine Basis des K^n.*

Ist nun $(u_1 ...., u_n)$ eine Basis aus Eigenvektoren zu A mit zugehörigen Eigenwerten $(\lambda_1, ... \lambda_n)$, so gilt ja nach Definition

$$Au_i = \lambda_i u_i \quad (i = 1, ..., n)$$

und nach Satz und Bezeichnung H.22 bekommen wir damit

Satz E.9 *Bilden die Spalten $(u_1 ...., u_n)$ der Matrix S eine Basis aus Eigenvektoren von A zu Eigenwerten $\lambda_1, ..., \lambda_n$, so ist A ähnlich einer Diagonalmatrix, genauer*

$$A' := S^{-1}AS = \mathrm{diag}(\lambda_1, ..., \lambda_n) := \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Die Diagonalelemente sind dabei gerade die Eigenwerte. Man sagt: "A ist diagonalisierbar."

Man rechnet sofort nach, daß hier auch die Umkehrung gilt, d.h. ist $A' = S^{-1}AS$ eine Diagonalmatrix, so sind die Spalten von S eine Basis aus Eigenvektoren von A. Damit haben wir

Satz E.10 *Eine $n \times n$-Matrix A ist genau dann diagonalisierbar, wenn eine Basis aus Eigenvektoren von A existiert.*
A ist sicher dann diagonalisierbar, wenn n paarweise verschiedene Eigenwerte von A existieren.

Um Missverständnissen vorzubeugen: Leider sind keineswegs alle Matrizen diagonalisierbar.
Ehe wir untersuchen, welche häßlichen Matrizen denn sonst noch auftreten, ein paar Überlegungen zur Frage, wie man Eigenwerte und Eigenvektoren berechnet.
Ist λ Eigenwert, so ist nach Satz und Definition E.2 jedes $x \neq 0$ aus dem Kern von $(A - \lambda I)$ zugehöriger Eigenvektor. Den Kern zu bestimmen heißt aber ein homogenes lineares Gleichungssystem zu lösen und das sollte eigentlich kein Problem mehr sein. (GAUSS-JORDAN-Elimination!)
Demnach brauchen wir also "nur" noch die Eigenwerte zu bestimmen. Dafür legt Korollar E.5 folgenden Weg nahe:
Man bestimme das charakteristische Polynom und berechne dessen Nullstellen.
Für ganz kleine n oder für spezielle Fälle führt dies auch zum Ziel. Im allgemeinen ist aber schon die Bestimmung des charakteristischen Polynoms eine mühsame Sache und zur Berechnung seiner Nullstellen ist man auf numerische Methoden angewiesen. Dies ändert sich nicht prinzipiell, wenn man andere Methoden verfolgt.
Das Bestimmen der Eigenwerte einer Matrix ist i.a. nur mit numerischen Näherungsmethoden möglich.

Zwei Beispiele:

BE1: Untersuchen wir die Matrix

$$A = \begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}.$$

Das charakteristische Polynom lautet

$$\chi_A(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} -1-\lambda & 2 & 2 \\ 2 & 2-\lambda & 2 \\ -3 & -6 & -6-\lambda \end{pmatrix}.$$

Subtrahieren wir die letzte Spalte von der zweiten, so folgt

$$\chi_A(\lambda) = \det \begin{pmatrix} -(1+\lambda) & 0 & 2 \\ 2 & -\lambda & 2 \\ -3 & \lambda & -6-\lambda \end{pmatrix}.$$

Entwickeln nach der zweiten Spalte liefert

$$\begin{aligned} \chi_A(\lambda) &= -\lambda(-(1+\lambda)2 - 4) - \lambda((1+\lambda)(6+\lambda) + 6) \\ &= -\lambda(\lambda^2 + 5\lambda + 6) = -\lambda(\lambda+2)(\lambda+3). \end{aligned}$$

Demnach hat A die drei verschiedenen Eigenwerte $\lambda_1 = 0$, $\lambda_2 = -2$, $\lambda_3 = -3$. Nach Satz E.8 bilden die zugehörigen Eigenvektoren (u_1, u_2, u_3) eine Basis des $\mathbb{R}^3$. Also ist A diagonalisierbar.
Zur Bestimmung der Eigenvektoren betrachten wir das System $(A - \lambda I)x = 0$ für $\lambda = \lambda_1$ oder $\lambda = \lambda_2$ oder $\lambda = \lambda_3$, d.h. die Systeme:

$$\lambda_1 = 0: \quad \begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix} x = 0 \quad \text{mit einer Lösung } x = u_1 := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

$$\lambda_2 = -2: \quad \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 2 \\ -3 & -6 & -4 \end{pmatrix} x = 0 \quad \text{mit einer Lösung } x = u_2 := \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix},$$

$$\lambda_3 = -3: \quad \begin{pmatrix} 2 & 2 & 2 \\ 2 & 5 & 2 \\ -3 & -6 & -3 \end{pmatrix} x = 0 \quad \text{mit einer Lösung } x = u_3 := \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Es bezeichne G die mit diesen Eigenvektoren als Spalten gebildete Matrix:

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

Da diese eine Basis bilden, ist G invertierbar. Nach Satz E.9 ist dann

$$G^{-1}AG = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} =: A'.$$

Sie können dies "zu Fuß" nachrechnen oder nochmal so schließen:
Nach Konstruktion von G ist $u_i = Ge_i$ $(i = 1, 2, 3)$. Damit ist

$$G^{-1}AGe_i = G^{-1}Au_i = G^{-1}\lambda_i u_i = \lambda_i G^{-1}u_i = \lambda_i G^{-1}Ge_i = \lambda_i e_i = A'e_i.$$

BE2: In Satz E.6 können wir nicht auf die Voraussetzung verzichten, daß wir als Körper gerade $\mathbb{C}$ haben, (genauer, daß er algebraisch abgeschlossen ist):
Über $\mathbb{R}$ besitzt die Matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ keine Eigenwerte; denn es ist das charakteristische Polynom $\chi_A(\lambda) = \det(A - \lambda I) = \lambda^2 + 1 \neq 0$ für alle $\lambda \in \mathbb{R}$. Über den komplexen Zahlen $\mathbb{C}$ hat dagegen $\chi_A(\lambda)$ die beiden Nullstellen $\pm i$ und somit eine Basis aus Eigenvektoren. Wie sieht sie aus?

JORDANsche Normalform

Definition E.11 *Eine Matrix der Gestalt*

$$J = \begin{pmatrix} \lambda_1 & 1 & & & & \\ & \lambda_1 & 1 & & 0 & \\ & & \ddots & \ddots & & \\ & & & \ddots & 1 & \\ & 0 & & & \lambda_1 & 1 \\ & & & & & \lambda_1 \end{pmatrix}$$

*heißt "*JORDAN*-Kasten" zum Eigenwert* λ_1.

Sein charakteristisches Polynom hat offenbar die spezielle Form

$$\chi_J(\lambda) = (\lambda_1 - \lambda)^n.$$

sodaß λ_1 der einzige Eigenwert ist. Der Eigenraum zu λ_1 bestimmt sich zu

$$E(A, \lambda_1) = \ker(J - \lambda_1 I) = \ker \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & 0 & \\ & & \ddots & \ddots & & \\ & & & \ddots & 1 & \\ & 0 & & & 0 & 1 \\ & & & & & 0 \end{pmatrix}$$

und, da die Matrix $J - \lambda_1 I$ noch den Rang $n - 1$ hat, ist

$$E(A, \lambda_1) = \ker(J - \lambda_1 I) = \mathrm{span}(e_1),$$

d.h. eindimensional. Zu dieser Matrix gibt es also (wenn $n > 1$) keine Basis aus Eigenvektoren. Dafür gilt aber mit den kanonischen Einheitsvektoren für alle i:

$$(J - \lambda_1 I)e_i = \begin{cases} 0 & i = 1 \\ e_{i-1} & i > 1. \end{cases}$$

Damit haben wir ein Beispiel für folgende

Definition E.12 (Hauptvektoren) *Eine Familie* $(x_1, ..., x_m)$ *von Vektoren* $\neq 0$ *heißt "Kette von Hauptvektoren" zu dem Eigenwert* λ *der Matrix* A*, wenn für alle* $i = 1, ..., m$ *gilt*

$$(A - \lambda I)x_i = \begin{cases} 0 & i = 1 \\ x_{i-1} & i > 1, \end{cases}$$

oder äquivalent

$$Ax_i = \begin{cases} \lambda x_i & i = 1 \\ \lambda x_i + x_{i-1} & i > 1. \end{cases}$$

Man nennt dann x_i *auch einen "Hauptvektor* i*-ter Stufe". Hauptvektoren 1-ter Stufe sind gerade die Eigenvektoren.*

Satz E.13

1. *Ist x_i Hauptvektor i-ter Stufe zum Eigenwert λ von A, so ist*
$$x_i \in \ker(A-\lambda I)^i \text{ aber } x_i \notin \ker(A-\lambda I)^{i-1}.$$
2. *Eine Kette von Hauptvektoren ist linear unabhängig.*

Beweis:

1. Man prüft leicht nach, daß für alle i, k gilt
$$(A-\lambda I)^k x_i = \begin{cases} 0 & i \leq k \\ x_{i-k} & i > k. \end{cases}$$
Speziell für $k = i$ bzw. $k = i-1$ ergibt sich die Behauptung.
2. Es ist $(0) \subset \ker(A-\lambda I)^1 \subset \ker(A-\lambda I)^2 \subset \ldots \subset \ker(A-\lambda I)^m$ eine aufsteigende Kette von Unterräumen, für die nach 1. stets $x_i \in \ker(A-\lambda I)^i$, aber $\notin \ker(A-\lambda I)^{i-1}$ liegt. Dann müssen aber die x_i nach Lemma V.22 unabhängig sein. □

Satz E.14 *Es sei $(x_1, ..., x_m)$ eine Kette von Hauptvektoren zum Eigenwert λ der $n \times n$-Matrix A und $H := \operatorname{span}(x_1, ..., x_m)$. Dann gelten*

1. *$x \in H \Rightarrow Ax \in H$, d.h. H ist "invariant" unter A.*
2. *Die Hauptvektorenkette $(x_1, ..., x_m)$ ist Basis von H.*
3. *Ist $(u_1, ..., u_r, x_1, ..., x_m, v_1, ..., v_s)$ eine Basis des K^n, wobei $(x_1, ..., x_m)$ Hauptvektorenkette ist, so hat A bezüglich dieser Basis eine Darstellung*
$$A' = \left(\begin{array}{c|c|c} \ldots & 0 & \ldots \\ \hline \ldots & J & \ldots \\ \hline \ldots & 0 & \ldots \end{array}\right)$$
mit dem JORDAN-Kasten in den der Position der Kette in der Basis entsprechenden Spalten.

Beweis: 1. liest man aus Definition E.12 ab.
2. Die Hauptvektorkette ist unabhängig und damit Basis des von ihr aufgespannten Unterraumes.
3. Dies folgt wieder mit Satz und Bezeichnung H.22 aus der in Definition E.12 gegebenen Darstellung. □

Mit erheblich mehr Aufwand, als wir hier treiben wollen, beweist man den folgenden Satz, der nur für den komplexen Fall, also das Arbeiten im $\mathbb{C}^n$ uneingeschränkt richtig ist:

Satz und Definition E.15 *Zu jeder komplexen $n \times n$-Matrix $A \in \mathbb{C}^{n \times n}$ gibt es eine Zerlegung des Raumes*
$$\mathbb{C}^n = H_1 \oplus H_2 \oplus ... \oplus H_k$$
als direkte Summe von Unterräumen H_i, die jeweils eine Kette von Hauptvektoren als Basis besitzen. Hängt man alle diese Teilbasen aneinander, so entsteht eine Basis des $\mathbb{C}^n$, bezüglich der die Matrix A die Darstellung
$$A' = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}$$
hat, wobei die J_i gerade die zu H_i gehörenden JORDAN-Kästen sind.
Man nennt dies die "JORDANsche Normalform" der Matrix A.

Zum **Beweis:** Hat man eine Basis des Gesamtraumes, die aus lauter Ketten von Hauptvektoren zusammengesetzt ist, so bekommt man die behauptete Matrixdarstellung direkt aus dem vorigen Satz E.14. Die Hauptarbeit steckt also im Nachweis der Existenz solcher Basen und die schenken wir uns. □

Die JORDANsche Normalform liefert unter anderem den Schlüssel zur Lösung von linearen Differentialgleichungssystemen wie etwa dem eingangs betrachteten. Die Ausführung muß auf später verschoben werden.

Diese JORDANsche Normalform sei für ein – noch recht simples – Beispiel bestimmt.

BE3: Es sei

$$A = \begin{pmatrix} 17 & 0 & -25 \\ 0 & 3 & 0 \\ 9 & 0 & -13 \end{pmatrix}$$

Das charakteristische Polynom lautet

$$\begin{aligned} \chi_A(\lambda) = \det(A - \lambda I) &= \det \begin{pmatrix} 17-\lambda & 0 & -25 \\ 0 & 3-\lambda & 0 \\ 9 & 0 & -13-\lambda \end{pmatrix} \\ &= (3-\lambda)[-(17-\lambda)(13+\lambda) + 9 \cdot 25] = (3-\lambda)(\lambda-2)^2. \end{aligned}$$

Wir haben also zwei verschiedene Eigenwerte $\lambda_1 = 3, \lambda_2 = 2$.
$(A - 3I)x = 0$ besitzt genau eine unabhängige Lösung, nämlich $u_1 = (0,1,0)^T$.
Für

$$(A - 2I)x = \begin{pmatrix} 15 & 0 & -25 \\ 0 & 1 & 0 \\ 9 & 0 & -15 \end{pmatrix} x = 0$$

existiert wegen $\text{rg}(A - 2I) = 2$ auch nur genau eine unabhängige Lösung, nämlich $u_2 = (5,0,3)^T$. Da aber $u_2 \in \text{im}(A - 2I)$ ist, hat auch das inhomogene System

$$(A - 2I)x = u_2, \text{ d.h. } \begin{pmatrix} 15 & 0 & -25 \\ 0 & 1 & 0 \\ 9 & 0 & -15 \end{pmatrix} x = \begin{pmatrix} 5 \\ 0 \\ 3 \end{pmatrix}$$

eine Lösung, nämlich $u_3 = (\frac{1}{3}, 0, 0)^T$. Dann ist

$$(A - 2I)u_3 = u_2 \neq 0, \text{ aber } (A - 2I)^2 u_3 = (A - 2I)u_2 = 0,$$

d.h. u_3 ist ein Hauptvektor zweiter Stufe zum Eigenwert 2 von A. Die Vektoren (u_1, u_2, u_3) bilden eine Basis von $\mathbb{R}^3$, in der u_2, u_3 eine Kette bilden. Nach Satz und Definition E.15 bekommen wir dann für A bezüglich dieser Basis somit die Matrixdarstellung

$$A' = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

in JORDANscher Normalform.
Bilden Sie wie in BE1 die Matrix G mit den Spalten (u_1, u_2, u_3), dann können Sie nachrechnen oder analog oben schließen, daß $A' = G^{-1}FG$ ist.
Überlegen Sie mal, warum es keinen Sinn hat, für den Eigenwert 3 eine Hauptvektor zu suchen? Woran kann man das sehen?
Der Vektor $u_3' := (2,0,1)^T$ wäre auch ein Hauptvektor gewesen, der mit u_2 eine Kette bildet. Es gibt also zu dem Eigenvektor u_2 verschiedene (unabhängige) Hauptvektoren, die man in die gesuchte Basis aufnehmen kann.
Führen Sie die gleichen Überlegungen nochmal aus, nachdem Sie in A das mittlere Element von 3 in 2 geändert haben.

Normale Matrizen

Im $\mathbb{C}^n$ haben wir das Standard-Skalarprodukt

$$\langle x, y\rangle := \sum_{i=1}^{n} \overline{\xi_i}\eta_i$$

eingeführt, das wir jetzt zu den Eigenwertbetrachtungen hinzunehmen. In Definition M.7 hatten wir schon die adjungierte Matrix eingeführt:

Definition E.16 *Zu einer Matrix $A = (\alpha_{ij}) \in \mathbb{C}^{n\times m}$ heißt die Matrix*

$$A^* := A^H = (\alpha_{ij}^H) \in \mathbb{C}^{m\times n} \text{ mit } \alpha_{ij}^H := \overline{\alpha_{ji}}$$

die zu A hermitesch adjungierte Matrix. Sie entsteht also, indem man A zuerst transponiert und dann noch alle Elemente konjugiert komplex nimmt.

Aus der letzten Beschreibung erhält man sofort (siehe auch Satz M.8) die folgenden Rechenregeln:

Satz E.17 *Es ist*

1. $(A^H)^H = A, \quad I^H = I,$
2. $(AB)^H = B^H A^H$ *(Reihenfolge!)*
3. $(\alpha A + \beta B)^H = \overline{\alpha} A^H + \overline{\beta} B^H \quad (\alpha, \beta \in \mathbb{C}).$

Diese Matrixoperation bekommt ihre Bedeutung über den Zusammenhang mit dem Skalarprodukt.

Satz E.18 *Es ist $\langle x, y\rangle = x^H y$ als Matrix-Produkt. Ferner ist*

$$\langle x, Ay\rangle = \langle A^H x, y\rangle$$

und dadurch ist auch A^H festgelegt.

Beweis: Die erste Aussage folgt unmittelbar aus den Definitionen. Damit erhalten wir dann

$$\langle x, Ay\rangle = x^H(Ay) = (x^H A)y = (A^H x)^H y = \langle A^H x, y\rangle.$$

Setzt man für x, y die kanonischen Einheitsvektoren e_i, e_j ein, so bekommt man

$$\begin{aligned} \langle e_i, Ae_j\rangle &= \alpha_{ij} \\ \langle A^H e_i, e_j\rangle &= \overline{\alpha_{ji}^H}, \end{aligned}$$

sodaß die im Satz genannte Beziehung auch die adjungierte Matrix festlegt. □

Definition E.19 *Es sei $A \in \mathbb{C}^{n\times n}$. Dann heißt*

- *A normal, wenn $A^H A = AA^H$ und speziell*
- *A selbstadjungiert oder hermitesch, wenn $A^H = A$, (falls A reell spricht man von symmetrisch),*
- *A unitär,wenn $A^H = A^{-1}$, (falls A reell spricht man von orthogonal).*

Matrizen von diesem Typ tauchen bei vielen physikalischen Problemen auf, ferner ist ihre Eigenwerttheorie recht übersichtlich.
Beispiele selbstadjungierter Matrizen kann man sich natürlich trivial bauen, ferner erhält man direkt aus der Definition, daß eine Matrix genau dann unitär ist, wenn ihre Spalten eine Orthonormalbasis bilden.
Ein für das weitere wichtiges Beispiel selbstadjungierter Matrizen sind die orthogonalen Projektoren:
Es sei $U \subset \mathbb{C}^n$ ein Unterraum, darin $(x_1, ..., x_m)$ eine ONB. Wir bilden die $n \times n$ Matrix

$$P := \sum_{i=1}^{m} x_i x_i^H \qquad \text{(Spalte mal Zeile !)}$$

Es ist dann für jedes $y \in \mathbb{C}^n$

$$Py = \sum_{i=1}^{m} x_i x_i^H y = \sum_{i=1}^{m} x_i \langle x_i, y\rangle = \sum_{i=1}^{m} \langle x_i, y\rangle x_i$$

und dies ist nach Satz V.43 gerade das Proximum zu y in dem Unterraum U.

Wir nennen daher P die Matrix der orthogonalen Projektion auf U oder kurz den orthogonalen Projektor auf U.

Satz E.20 *Projektoren sind selbstdjungiert und idempotent, d.h. $P^H = P$ und $P^2 = P$ und durch diese Eigenschaften charakterisiert.*

Beweis: Es ist

$$\begin{aligned} P^H &= \sum (x_i x_i^H)^H = \sum (x_i^H)^H x_i^H = \sum x_i x_i^H = P, \\ P^2 &= \sum_{ij} x_i x_i^H x_j x_j^H = \sum_{ij} \langle x_i, x_j\rangle x_i x_j^H = \sum_{ij} \delta_{ij} x_i x_j^H = \sum_i x_i x_i^H = P. \end{aligned}$$

Der Nachweis der restlichen Behauptungen sei als Übung gelassen. □

Für das Arbeiten mit normalen Matrizen ist folgender Satz von Nutzen:

Satz E.21 *Für $A \in \mathbb{C}^{n\times n}$ gelten*

1. $(\operatorname{im} A)^\perp = \ker A^H$,
2. *A normal* $\Rightarrow \ker A = \ker A^H$,
3. *A normal* $\Rightarrow \mathbb{C}^n = \operatorname{im} A \oplus \ker A$.

Beweis: Zu 1.: $y \in (\operatorname{im} A)^\perp$ bedeutet, daß für alle $x \in \mathbb{C}^n$ stets $0 = \langle y, Ax\rangle = \langle A^H y, x\rangle$. Für $x = A^H y$ folgt dann notwendig $A^H y = 0$, also $y \in \ker A^H$. Dies kann man alles umdrehen.
Zu 2.: Es ist

$$\langle Ax, Ax\rangle = \langle A^H Ax, x\rangle \overset{\text{normal}}{=} \langle AA^H x, x\rangle = \langle (A^H)^H A^H x, x\rangle = \langle A^H x, A^H x\rangle,$$

also $Ax = 0$ genau wenn $A^H x = 0$.
Zu 3.: Es ist stets $C^n = \operatorname{im} A \oplus (\operatorname{im} A)^\perp \overset{1,2}{=} \operatorname{im} A \oplus \ker A$. □

Betrachten wir nun Eigenwerte von Normalen Matrizen.

Satz E.22 *Es sei $F \in \mathbb{C}^{n\times n}$ normal, $\lambda \in \mathbb{C}$. Dann gelten*

1. $(F - \lambda I)^H = F^H - \overline{\lambda} I$ *und* $(F - \lambda I)$ *ist normal.*

2. *$Fx = \lambda x \;\Leftrightarrow F^H x = \overline{\lambda} x$, insbesondere haben F und F^H dieselben Eigenvektoren.*
3. *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis: 1.: Dies rechnet man mit Satz E.17 aus.
2.: Nach Satz E.21 haben $(F - \lambda I)$ und $(F - \lambda I)^H$ denselben Kern.
3.: Sind $\lambda \neq \lambda'$ Eigenwerte zu F mit Eigenvektoren x, x', so ist

$$\lambda\langle x', x\rangle = \langle x', \lambda x\rangle = \langle x', Fx\rangle = \langle F^H x', x\rangle = \langle \overline{\lambda'} x', x\rangle = \lambda'\langle x', x\rangle,$$

woraus wegen $\lambda \neq \lambda'$ notwendig $\langle x', x\rangle = 0$ folgt □

Damit erhalten wir den zentralen Satz über die Eigenwerte normaler Matrizen, den sog. Spektralsatz:

Satz E.23 (Spektralsatz) *Für eine komplexe $n \times n$ - Matrix F sind äquivalent:*

1. *F ist normal.*
2. *Es existiert eine Orthonormalbasis aus Eigenvektoren von F.*
3. *Es gibt orthogonale Projektoren P_j und Zahlen λ_j, $(j = 1, ..., m)$, sodaß*
   (a) $P_i P_j = 0 \quad (i \neq j)$,
   (b) $\sum_{i=1}^m P_i = I$,
   (c) $\sum_{i=1}^m \lambda_i P_i = F$.

   Letzteres nennt man die Spektraldarstellung von F.

Beweis: 1. $\Rightarrow$ 2.: Es seien $\lambda_1, ..., \lambda_m$ sämtliche verschiedenen Eigenwerte von F (nach Satz E.6 gibt es welche), dazu jeweils $E_i := \ker(F - \lambda_i I)$ die entsprechenden Eigenräume. In jedem solchen Eigenraum wählen wir eine orthonormierte Basis. Da Eigenräume zu verschiedenen Eigenwerten orthogonal sind, bilden alle diese Basen zusammen ein ONS $(x_1, ..., x_r)$ in $\mathbb{C}^n$, das einen Unterraum $U := \mathrm{sp}(x_1, ..., x_r)$ aufspannt.
Ist $U = \mathbb{C}^n$ so haben wir eine ONB des $\mathbb{C}^n$ aus Eigenvektoren gefunden.
Nehmen wir also an, es sei $U \neq \mathbb{C}^n$. Die $(x_1, ..., x_r)$ sind auch Eigenvektoren von F^H und damit ist insbesondere $F^H U \subset U$. Dann ist für $u \in U, v \in U^\perp$:

$$\langle u, Fv\rangle = \langle F^H u, v\rangle = 0,$$

da ja $F^H u \in U$ und $v \perp U$. Dies bedeutet aber, daß für $v \in U^\perp$ auch $Fv \in U^\perp$ ist, somit $FU^\perp \subset U^\perp$. Also bildet F auch noch den Raum $U^\perp$ in sich ab und muß daher einen Eigenvektor $y \in U^\perp$ besitzen. Alle Eigenvektoren waren aber schon in U, sodaß $y \in U \cap U^\perp = (0)$ sein müßte, was nicht geht. Folglich ist unsere Annahme falsch, d.h. schon eine ONB gefunden.
2. $\Rightarrow$ 3.: Sind E_i $(i = 1, ..., m)$ die Eigenräume zu den verschiedenene Eigenwerten λ_i von F, so wähle man als P_i den orthogonalen Projektor auf den Eigenraum P_i. Ist $(x_{i1}, ..., x_{im_i})$ eine ONB im Eigenraum E_i, so ist also

$$P_i = \sum_{k=1}^{m_i} x_{ik} x_{ik}^H.$$

Für $i \neq j$ sind $\langle x_{ik}, x_{jk'}\rangle = 0$, also

$$P_i P_j = \sum_{k,k'} x_{ik} x_{ik}^H x_{jk'} x_{jk'}^H = \sum_{k,k'} \langle x_{ik}, x_{jk'}\rangle x_{ik} x_{jk'}^H = 0.$$

Ferner ist für beliebiges $y \in \mathbb{C}^n$

$$\sum_i P_i y = \sum_i \sum_k x_{ik} \langle x_{ik}, y \rangle = y,$$

da alle x_{ik} zusammen eine ONB von $\mathbb{C}^n$ bilden.
Schließlich ist

$$\begin{aligned} \sum_i \lambda_i P_i y &= \sum_i \sum_k \lambda_i \langle x_{ik}, y \rangle x_{ik} = \sum_i \sum_k \langle \overline{\lambda_i} x_{ik}, y \rangle x_{ik} \\ &= \sum_i \sum_k \langle F^H x_{ik}, y \rangle x_{ik} = \sum_i \sum_k \langle x_{ik}, Fy \rangle x_{ik} = Fy, \end{aligned}$$

womit 3. gezeigt ist.
3. $\Rightarrow$ 1. rechnet man mit Satz E.17 nach. □

Zum besseren Verständnis dieses Satzes sollten Sie sich unbedingt die folgende Aufgabe vornehmen!

Aufgabe E.24 *Wie sehen die Projektoren P_i aus, wenn die Basis-Eigenvektoren x_{ik} sämtlich kanonische Einheitsvektoren sind?*

Betrachten wir noch die Spezialfälle selbstadjungierter bzw. unitärer Matrizen.

Satz E.25

1. *Alle Eigenwerte einer selbstadjungierten (speziell symmetrischen) Matrix sind reell.*
2. *Symmetrische (die sind also reell) Matrizen besitzen eine ONB aus Eigenvektoren im $\mathbb{R}^n$, sind also mittels einer (reellen) orthogonalen Matrix auf Diagonalform transformierbar.*

Beweis: 1.: Ist $Fx = \lambda x$, und $F^H = F$, so ist

$$\lambda \langle x, x \rangle = \langle x, Fx \rangle = \langle F^H x, x \rangle = \langle Fx, x \rangle = \langle \lambda x, x \rangle = \overline{\lambda} \langle x, x \rangle,$$

somit ist $\lambda = \overline{\lambda}$ also reell.
2.: Der Eigenraum zu einem Eigenwert λ ist der Kern der reellen Matrix $F - \lambda I$, hat somit auch eine reelle ONB. Diese reellen Teilbasen kann man zusammensetzen und erhält eine reelle ONB aus Eigenvektoren. □

Satz E.26 *Es sei $F \in \mathbb{C}^{n \times n}$ oder $F \in \mathbb{R}^{n \times n}$. Im ersten Fall arbeiten wir im Raum $\mathbb{C}^n$, im zweiten im $\mathbb{R}^n$. Dann sind äquivalent:*

1. *F ist unitär bzw. orthogonal, d.h. $F^H = F^{-1}$ (bei $\mathbb{C}$) bzw. $F^T = F^{-1}$ (bei $\mathbb{R}$).*
2. *Die Spalten von F bilden eine ONB.*
3. *Ist $(x_1, ..., x_n)$ eine ONB, so auch $(Fx_1, ..., Fx_n)$.*
4. *Es ist stets $\langle Fx, Fy \rangle = \langle x, y \rangle$.*
5. *F ist isometrisch, d.h. es ist $|Fx| = |x|$ für alle x.*
6. *F ist normal und alle Eigenwerte haben den Betrag 1.*

Beweis: Wir zeigen nur, daß unitäre Matrizen alle die genannten Eigenschaften haben:
Für $x, y \in \mathbb{C}^n$ bzw. $\mathbb{R}^n$ ist

$$\langle Fx, Fy \rangle = \langle F^H Fx, y \rangle = \langle x, y \rangle.$$

Dies zeigt direkt 4. und für $x = y$ auch 5. . Sind x und y orthogonal, so damit auch Fx und Fy, was 3. liefert. Da die kanonischen Einheitsvektoren eine ONB bilden, haben wir auch 2. Schließlich folgt 6. aus 5. für Eigenvektoren über

$$|x| = |Fx| = |\lambda x| = |\lambda||x|,$$

sodaß notwendig $|\lambda| = 1$ ist. □

Betrachten wir die orthogonalen Matrizen noch etwas genauer:
<u>n = 2:</u> Die erste Spalte von F können wir schreiben als $\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$. Dann muß die zweite notwendig die Gestalt $\pm \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$ haben, sodaß entweder

$$F = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \text{ mit } \det F = +1,$$

oder

$$F = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \text{ mit } \det F = -1$$

ist. Im ersten Fall ist F die ebene Drehung um den Winkel φ, das charakteristische Polynom ist

$$\chi_F(\lambda) = (\cos\varphi - \lambda)^2 + \sin^2\varphi$$

mit den beiden Lösungen

$$\lambda_{1,2} = \cos\varphi \pm i\sin\varphi = e^{\pm i\varphi}.$$

Im zweiten Fall hat F das charakteristische Polynom

$$\chi_F(\lambda) = \lambda^2 - \cos^2\varphi - \sin^2\varphi = \lambda^2 - 1$$

mit den beiden Eigenwerten $\lambda_{1,2} = \pm 1$. Jetzt ist also F die orthogonale Spiegelung an der Nullpunktsgeraden in Richtung des Eigenvektors $\begin{pmatrix} \cos\left(\frac{\varphi}{2}\right) \\ \sin\left(\frac{\varphi}{2}\right) \end{pmatrix}$ zu $\lambda_1 = +1$.

<u>n = 3:</u> Das charakteristische Polynom hat Grad 3, hat, da F reell, auch reelle Koeffizienten und alle Nullstellen haben den Betrag 1. Somit folgt:
Entweder sind alle Eigenwerte reell, damit $\in \{\pm 1\}$ und dazu gibt es eine reelle ONB aus Eigenvektoren. Dann kann F sein die Identität oder eine Spiegelung an einer 0-Punkts-Ebene, einer 0-Punkts-Geraden oder die Punktspiegelung am Nullpunkt.
Andernfalls ist genau ein Eigenwert reell, also $\in \{\pm 1\}$ und die anderen beiden sind konjugiert komplex von der Form $\lambda_{1,2} = e^{\pm i\varphi}$. Ist a der reelle Eigenvektor zu dem reellen Eigenwert, $E \subset R^3$ die dazu orthogonale Ebene, so ist $E = (\operatorname{span}(a))^\perp$ und F muß diese Ebene wieder in sich selbst abbilden, wobei die konjugiert komplexen Eigenwerte $e^{\pm i\varphi}$ auftreten. Dies bedeutet: Mit einer geigneten reellen ONB kann man in diesem Fall F ähnlich transformieren auf

$$F \cong F' = \begin{pmatrix} \pm 1 & & \\ & \cos\varphi & -\sin\varphi \\ & \sin\varphi & \cos\varphi. \end{pmatrix}$$

Dies gilt nun sogar allgemein

Satz E.27 *Zu jeder orthogonalen Matrix $F \in \mathbb{R}^{n\times n}$ gibt es eine ONB $(x_1, ..., x_n)$ in $\mathbb{R}^n$, bezüglich der F die folgende Gestalt F' hat*

$$F' = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& R_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& R_s \end{pmatrix}.$$

Darin sind die R_j jeweils 2×2 Drehmatrizen.

Gruppen von Matrizen

An dieser Stelle ist es sinnvoll auf einige Gruppen von Abbildungen einzugehen:

Definition E.28 *Eine Teilmenge $\mathcal{G} \subset \mathbb{C}^{n\times n}$ oder $\subset \mathbb{R}^{n\times n}$ heißt Matrix-Gruppe, wenn gelten*

1. *mit $A, B \in \mathcal{G}$ ist auch $A \cdot B \in \mathcal{G}$,*
2. *$I \in \mathcal{G}$,*
3. *alle Matrizen in $\mathcal{G}$ sind invertierbar und die Inverse ist ebenfalls in $\mathcal{G}$.*

Beispiele

1. *Die Lineare Gruppe $GL(n)$ besteht aus allen invertierbaren $n \times n$ - Matrizen.*
2. *Die unitäre Gruppe U_n besteht aus allen unitären $n \times n$ - Matrizen.*
3. *Die orthogonale Gruppe O_n besteht aus allen orthogonalen $n \times n$ - Matrizen.*
4. *Die spezielle lineare Gruppe $SL(n)$, spezielle unitäre Gruppe SU_n, spezielle orthogonale Gruppe SO_n besteht jeweils aus allen entsprechenden Matrizen mit Determinante $= +1$.*
5. *Die Kreisgruppe besteht aus allen Matrizen der Form $e^{it} I$ mit $t \in \mathbb{R}$.*

Definitheit und quadratische Formen

Mit einer selbstadjungierten Matrix A ist für alle x

$$\langle x, Ax\rangle = \langle A^H x, x\rangle = \langle Ax, x\rangle = \overline{\langle x, Ax\rangle}$$

und somit reell. Also ist die folgende Definition sinnvoll:

Definition E.29 (Definit) *Die Matrix $A \in \mathbb{C}^{n\times n}$ sei selbstadjungiert, $U \subset \mathbb{C}^n$ sei ein Unterraum. Dann heißt*

- *A auf U "positiv definit", genau wenn für alle $x \in U, x \neq 0$ das Skalarprodukt $\langle x, Ax\rangle > 0$ ist.*
- *A auf U "negativ definit", genau wenn für alle $x \in U, x \neq 0$ das Skalarprodukt $\langle x, Ax\rangle < 0$ ist.*
- *Läßt man ≥ 0 bzw. ≤ 0 zu, so spricht man von "positiv bzw. negativ semi - definit".*

Ist kein Unterraum spezifiziert, so ist $U = \mathbb{C}^n$ gemeint.

Definition E.30 (Signatur) *Sind U_+, U_-, U_0 die Unterräume, die von allen Eigenvektoren zu sämtlichen positiven, sämtlichen negativen Eigenwerten bzw. zum Eigenwert Null einer selbstadjungierten Matrix A aufgespannt werden, so nennt man das Tupel (k_+, k_-, k_0) aus deren Dimensionen die Signatur von A. (Falls es keine entsprechenden Eigenwerte gibt, verwende man entsprechend den Nullraum (0).)*

Auf den Räumen U_+ bzw. U_- ist A offenbar positiv bzw. negativ definit. Genauer gilt

Satz E.31 *Ist A positiv definit auf U, so ist $\dim U \leq \dim U_+ = k_+$, ist A positiv semi-definit auf U, so ist $\dim U \leq \dim(U_+ + U_0) = k_+ + k_0$.*
Analoges für negativ (semi-)definit.
Man beachte, daß man die Aussage ***nicht*** *zu $U \subset U_+$ bzw. $U \subset U_+ + U_0$ etc. verschärfen kann.*

Beweis: Ist A positiv semidefinit auf U, so ist $\langle u, Au\rangle \geq 0$ für alle $u \in U$. Ist $x \in U_-, \neq 0$, so ist $\langle u, Au\rangle < 0$. Folglich ist $U \cap U_- = 0$ und damit $\dim U + \dim U_- \leq n = k_0 + k_+ + k_-$, also $\dim U \leq k_0 + k_+$. Der Rest geht analog. $\square$

Die Signatur hängt an den Eigenwerten, ändert sich also nicht bei Ähnlichkeitstransformationen. Es gilt sogar etwas mehr, nämlich der folgende

Satz E.32 (Trägheitssatz von Sylvester) *Ist $A \in \mathbb{C}^{n\times n}$ selbstadjungiert und $G \in \mathbb{C}^{n\times n}$ invertierbar, so ist auch $G^H AG$ selbstadjungiert und hat dieselbe Signatur wie A.*

Der **Beweis** ist eine einfache Anwendung von Satz E.31 und sei übergangen.
Als Test für positiv definit taugt das folgende Kriterium:

Satz E.33 *Eine selbstadjungierte Matrix A ist genau dann positiv definit, wenn alle Haupt-Unterdeterminanten der Form*

$$\det A_k := \det\left(\alpha_{ij}\right)_{k,k} \qquad (k = 1, ..., n)$$

positiv sind, sie ist genau dann negativ definit, wenn $(-1)^k \det A_k > 0$ für alle k.
Insbesondere sind damit positiv oder negativ definite Matrizen auch invertierbar.

Beweis: Ist A positiv definit, so sind alle Eigenwerte > 0 und damit auch die Determinante als Produkt aller Eigenwerte (siehe Diagonalform von A). Ferner ist mit $x = \begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix}$:

$$\langle x, Ax\rangle = \sum_{ij}^{n} \alpha_{ij}\overline{\xi_i}\xi_j.$$

Ist jetzt $U_k := \mathrm{sp}(e_1, ..., e_k)$, so ist für $x \in U_k$ also $\xi_{k+1} = ... = \xi_n = 0$ und damit für solche x

$$\langle x, Ax\rangle = \sum_{ij}^{k} \alpha_{ij}\overline{\xi_i}\xi_j.$$

Folglich ist A positiv definit auf U_k, genau wenn der entsprechende Minor A_k positiv definit auf $\mathbb{C}^k$ ist.

Ist nun A positiv definit, so auf jedem U_k und damit ist jeder Minor A_k positiv definit auf C^k und hat damit positive Determinante. Dies zeigt die eine Richtung des Satzes. Für die andere braucht man eine subtile Anwendung von Satz E.31. Wir übergehen die Details.
Für die letzte Behauptung beachte man, daß A genau dann positiv definit ist, wenn die Matrix $-A$ negativ definit ist. □

Ein hinreichendes Kriterium für Definitheit enthält

Satz E.34 *Es sei A selbstadjungiert mit "überwiegender Hauptdiagonale", d.h für alle k sei*

$$a_{kk} > \sum_{j\neq k} |a_{kj}|.$$

Dann ist A positiv definit.

Beweis: Sei λ Eigenwert mit Eigenvektor x und darin ξ_k eine Komponente mit größtem Betrag. Dann ist wegen $Ax = \lambda x$ insbesondere

$$(\lambda - a_{kk})\xi_k = \sum_{j\neq k} a_{kj}\xi_j, \text{ d.h.}$$

$$|\lambda - a_{kk}|\,|\xi_k| \leq \sum_{j\neq k} |a_{kj}|\,|\xi_j| \leq \Big(\sum_{j\neq k} |a_{kj}|\Big)|\xi_k| < |a_{kk}|\,|\xi_k|.$$

Folglich liegt λ im Innern des Kreises um a_{kk} mit dem Radius $|a_{kk}|$ und ist damit positiv. Da wir mit einem beliebigen Eigenwert begonnen haben, sind also alle Eigenwerte positiv und damit die Matrix positiv definit. □

Aufgabe E.35 *Hat die Matrix $F \in \mathbb{C}^{m\times n}$ den Rang n $(\leq m)$, so ist die $n \times n$ Matrix $F^H F$ positiv definit und damit invertierbar.*

Wir verwenden diese Überlegungen noch zur Klassifikation der Kurven bzw. Flächen zweiten Grades im R^n. Deren allgemeine Gleichung lautet

$$\sum_{i=1}^{n} \alpha_{ii}\xi_i^2 + 2\sum_{i<j} \alpha_{ij}\xi_i\xi_j + 2\sum_{i=1}^{n} \beta_i\xi_i + \gamma = 0,$$

wobei $\alpha_{ij}, \beta_i, \gamma \in \mathbb{R}$.
Für $n = 2$ liefert dies die sogenannten "Kegelschnitte", für $n \geq 3$ die "Quadriken" genannten Flächen.
Setzen wir A: $= (\alpha_{ij})_{n,n}$, wobei $a_{ij} := a_{ji}$ für $i > j$ gesetzt sei,

$$b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \qquad x := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

so kann man mit dem Standard-Skalarprodukt im $\mathbb{R}^n$ obige Gleichung kompakt schreiben als

$$\langle x, Ax\rangle + 2\,\langle x, b\rangle + \gamma = 0. \qquad (*)$$

Da A symmetrisch ist (nach Konstruktion), gibt es eine orthogonale Matrix G, d.h. $G^H = G^{-1}$, sodaß $G^H AG =: L$ Diagonalform hat. Wählen wir nun als neues Koordinatensystem gerade die durch die Spalten von G gegebene ONB in $\mathbb{R}^n$, so erhält jeder Vektor $x \in \mathbb{R}^n$ bezüglich der neuen Basis die Darstellung $x' = G^H x$, bzw. $x = Gx'$.

Damit wird $(*)$ zu

$$\begin{aligned}0 &= \langle Gx', AGx'\rangle + 2\,\langle Gx', b\rangle + \gamma = \langle x', G^H AGx'\rangle + 2\,\langle x', G^H b\rangle + \gamma \\ &= \langle x', Lx'\rangle + 2\,\langle x', G^H b\rangle + \gamma,\end{aligned}$$

d.h. mit $c := G^H b$:

$$0 = \langle x', Lx'\rangle + 2\,\langle x', c\rangle + \gamma. \qquad (**)$$

Nun haben wir wieder eine Gleichung der Gestalt $(*)$, wobei jetzt aber die Matrix L Diagonalform hat, d.h. in der ausgeschriebenen Form der zweite Term mit den "gemischten" Produkten $\xi_i\xi_j$ $(i \neq j)$ verschwunden ist. Als nächstes versuchen wir weitere Terme durch eine Translation der Form $x' = x'' - x_0$ zu beseitigen. Dies liefert

$$\begin{aligned}0 &= \langle x'' - x_0, L(x'' - x_0)\rangle + 2\,\langle x'' - x_0, c\rangle + \gamma \\ &= \langle x'', Lx''\rangle - (\langle x_0, Lx''\rangle + \langle x'', Lx_0\rangle - 2\,\langle x'', c\rangle) + (\langle x_0, Lx_0\rangle - 2\,\langle x_0, c\rangle + \gamma).\end{aligned}$$

Da L symmetrisch ist, ist $\langle x_0, Lx''\rangle = \langle x'', Lx_0\rangle$ und wir erhalten mit

$$\begin{aligned}c' &:= Lx_0 - c, \\ \gamma' &:= \langle x_0, Lx_0\rangle - 2\,\langle x_0, c\rangle + \gamma\end{aligned}$$

als Gleichung für unsere Fläche

$$0 = \langle x'', Lx''\rangle - 2\,\langle x'', c'\rangle + \gamma'. \qquad (***)$$

Nun gilt es noch, x_0 geschickt zu wählen.
Hat L (und ebenso A) die Signatur $(k_+, k_-, k_0) = (s, t, n - (s+t))$, so haben wir für L die Darstellung

$$L = \left(\begin{array}{c|c} \Lambda & 0 \\ \hline 0 & 0 \end{array}\right) := \left(\begin{array}{ccc|ccc} \lambda_1 & & & & & \\ & \ddots & & & 0 & \\ & & \lambda_{s+t} & & & \\ \hline & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{array}\right),$$

wobei die $\lambda_1, ..., \lambda_{s+t}$ gerade sämtliche Eigenwerte $\neq 0$ sind.
Wir zerlegen entsprechend

$$x_0 = \begin{pmatrix} x_0^{(1)} \\ x_0^{(2)} \end{pmatrix}, \quad c = \begin{pmatrix} c^{(1)} \\ c^{(2)} \end{pmatrix}, \quad c' = \begin{pmatrix} c'^{(1)} \\ c'^{(2)} \end{pmatrix}.$$

Dann ist

$$\begin{aligned}c'^{(1)} &= \Lambda x_0^{(1)} - c^{(1)} \\ c'^{(2)} &= c^{(2)} \\ \gamma' &= \langle x_0^{(1)}, \Lambda x_0^{(1)}\rangle - 2\langle x_0^{(1)}, c^{(1)}\rangle - 2\langle x_0^{(2)}, c^{(2)}\rangle + \gamma.\end{aligned}$$

Λ ist invertierbar und wir bestimmen $x_0^{(1)}$ so, daß $\Lambda x_0^{(1)} = c^{(1)}$ d.h. $c'^{(1)} = 0$ wird.
Es ist dann

$$\gamma' = \gamma - \langle x_0^{(1)}, c^{(1)}\rangle - 2\langle x_0^{(2)}, c^{(2)}\rangle.$$

Nun haben wir zwei Fälle:
Fall 1 : Es ist $c^{(2)} = 0$: (Hierzu gehört auch der Fall, daß $s + t = n$).
Dann ist jetzt $c' = 0$, d.h. der lineare Term ist beseitigt und wir haben ferner

$$\gamma' = \gamma - \langle x_0^{(1)}, c^{(1)} \rangle.$$

Fall 2 : Es ist $c^{(2)} \neq 0$:
Dann ist $c'^{(2)} \neq 0$, d.h. die zu Eigenwerten $= 0$ gehörigen Komponenten des linearen Termes verschwinden nicht alle. Dafür kann man $x_0^{(2)}$ so wählen, daß $\gamma' = 0$ wird.

Die erhaltene Gleichung $(***)$ können wir noch folgendermaßen kompakt schreiben:
Setzen wir zu $x'' = (\xi_1, ..., \xi_n)^T$ fest, daß $\hat{x} := (\xi_1, ..., \xi_n, 1)^T$ sei und

$$F := \left(\begin{array}{c|c} L & c' \\ \hline (c')^T & \gamma' \end{array}\right) = \left(\begin{array}{ccc|ccc|c} \lambda_1 & & & & & & \\ & \ddots & & & 0 & & 0 \\ & & \lambda_{s+t} & & & & \\ \hline & & & 0 & & & \\ & 0 & & & \ddots & & c'^{(2)} \\ & & & & & 0 & \\ \hline & 0 & & & (c'^{(2)})^T & & \gamma' \end{array}\right)_{(n+1,n+1)},$$

wobei also stets entweder $\gamma' = 0$ oder $c'^{(2)} = 0$ erreicht werden kann, so bekommt $(***)$ die Darstellung

$$\langle \hat{x}, F\hat{x} \rangle = 0. \qquad (+)$$

Bemerkung E.36 *Im Falle $c = 0$ spricht man von einer Kurve bzw. Fläche mit Zentrum, im Falle $c \neq 0$ von einer Kurve bzw. Fläche ohne Zentrum.*

Wir wollen nun die Fälle $n = 2, 3$ genauer studieren:
Im Falle $n = 2$ reden wir statt von Flächen von Kurven und erhalten die bekannten Kegelschnitte.
Die wichtigsten Spezialfälle für $n = 2$ sind:

$s = 2$:
Die Kurve hat ein Zentrum, also $c = 0, \lambda_1, \lambda_2 > 0$.
$(***)$ lautet: $\lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \gamma = 0$ oder nach Multiplikation mit $\alpha := -\frac{1}{\gamma}$

$$(\frac{\lambda_1}{-\gamma})\xi_1^2 + (\frac{\lambda_2}{-\gamma})\xi_2^2 = 1.$$

Für $\gamma > 0$ hat dies keine Lösung, für $\gamma < 0$ ist es die Gleichung einer *Ellipse* mit den Halbachsen $a_1 := \sqrt{|\frac{\gamma}{\lambda_1}|}$, $a_2 := \sqrt{|\frac{\gamma}{\lambda_2}|}$. Für $\lambda_1 = \lambda_2$ gehört hierzu auch der *Kreis.*

$s = 1$, $t = 1$:
Die Kurve hat ein Zentrum, also $c = 0, \lambda_1 > 0, \lambda_2 < 0$. Gleichung $(***)$ lautet: $|\lambda_1|\xi_1^2 - |\lambda_2|\xi_2^2 + \gamma = 0$ und dies liefert für $\gamma \neq 0$ analog oben eine *Hyperbel.*

$s = 1$, $t = 0$:
$c = \binom{0}{\gamma_2}$ mit $\gamma_2 \neq 0$. Die Kurve hat kein Zentrum.
Gleichung $(***)$ lautet: $\lambda_1\xi_1^2 + 2\gamma_2\xi_2 + \gamma = 0$ und dies ist die Gleichung einer *Parabel.*

Die weiteren Fälle liefern entweder die schon behandelten Kurven (etwa $s = 0$, $t = 2$) oder Ausartungen zu Geraden oder Punkten. Die Details seien dem Leser überlassen.

Spezialfälle für $n = 3$:
$s = 3$:
$\lambda_1, \lambda_2, \lambda_3 > 0$, die Fläche hat ein Zentrum, d.h. $c = 0$. Die Gleichung $(***)$ lautet: $\lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \lambda_3\xi_3^2 + \gamma = 0$.
Für $\gamma > 0$ hat dies keine Lösung, für $\gamma \leq 0$ ist es die Gleichung eines ***Ellipsoids***. Für den Fall, daß zwei oder gar alle drei Eigenwerte gleich sind, ergibt dies das ***Rotationsellipsoid*** bzw. die ***Kugel***. Das Ellipsoid ist ein beschränktes Gebilde und der Schnitt mit jeder Ebene ist eine Ellipse (oder leer).
Im folgenden seien die wichtigsten Fälle tabellarisch aufgeführt, wobei wir die Matrix F aus $(+)$ benutzen und dabei noch einige einfache Normierungen vornehmen. Die μ_i bezeichnen positive Zahlen. Die weiteren Fälle liefern außer Entartungen wie Ebenen, Geraden, Punkte, keine wesentlichen neuen Flächen und seien deshalb übergangen.

$$F = \begin{pmatrix} + &&& \\ & + && \\ && + & \\ &&& + \end{pmatrix}$$
(Leere Menge)
$\mu_1\xi_1^2 + \mu_2\xi_2^2 + \mu_3\xi_3^2 + 1 = 0$

$$F = \begin{pmatrix} + &&& \\ & + && \\ && + & \\ &&& - \end{pmatrix}$$
(Ellipsoid)
$\mu_1\xi_1^2 + \mu_2\xi_2^2 + \mu_3\xi_3^2 - 1 = 0$

$$F = \begin{pmatrix} + &&& \\ & - && \\ && - & \\ &&& + \end{pmatrix}, \begin{pmatrix} + &&& \\ & + && \\ && - & \\ &&& - \end{pmatrix}$$
(Einschaliges Hyperboloid)
$\mu_1\xi_1^2 + \mu_2\xi_2^2 - \mu_3\xi_3^2 - 1 = 0$

$$F = \begin{pmatrix} + &&& \\ & + && \\ && - & \\ &&& + \end{pmatrix}, \begin{pmatrix} + &&& \\ & - && \\ && - & \\ &&& - \end{pmatrix}$$
(Zweischaliges Hyperboloid)
$\mu_1\xi_1^2 - \mu_2\xi_2^2 - \mu_3\xi_3^2 - 1 = 0$

$$F = \begin{pmatrix} + &&& \\ & + && \\ && 0 & - \\ && - & 0 \end{pmatrix}$$
(Elliptisches Paraboloid)
$\mu_1\xi_1^2 + \mu_2\xi_2^2 = 2\xi_3$

$$F = \begin{pmatrix} + &&& \\ & - && \\ && 0 & - \\ && - & 0 \end{pmatrix}$$
(Hyperbolisches Paraboloid)
$\mu_1\xi_1^2 - \mu_2\xi_2^2 = 2\xi_3$

$$F = \begin{pmatrix} + &&& \\ & + && \\ && - & \\ &&& 0 \end{pmatrix}, \begin{pmatrix} + &&& \\ & - && \\ && - & \\ &&& 0 \end{pmatrix}$$
(Elliptischer Doppelkegel)
$\mu_1\xi_1^2 + \mu_2\xi_2^2 = \mu_3\xi_3^2$

$$F = \begin{pmatrix} + &&& \\ & + && \\ && 0 & \\ &&& - \end{pmatrix}$$
(Elliptischer Zylinder)
$\mu_1\xi_1^2 + \mu_2\xi_2^2 = 1$

$$F = \begin{pmatrix} + &&& \\ & - && \\ && 0 & \\ &&& - \end{pmatrix}, \begin{pmatrix} + &&& \\ & - && \\ && 0 & \\ &&& + \end{pmatrix}$$
(Hyperbolischer Zylinder)
$\mu_1\xi_1^2 - \mu_2\xi_2^2 = 1$

$$F = \begin{pmatrix} + &&& \\ & 0 && - \\ && 0 & \\ & - && 0 \end{pmatrix}$$
(Parabolischer Zylinder)
$\mu_1\xi_1^2 = 2\xi_2$