Kapitel 4
Differentialgleichungen

Differentialgleichungssysteme sind so ziemlich *das* wichtigste mathematischen Hilfsmittel der Naturwissenschaften und der Technik. Die dahinterstehende Grundidee ist einfach: Man kann zwar nur selten *a priori* sagen, wie sich ein System über einen längeren Zeitraum hinweg entwickeln wird, aber man hat oft aufgrund von Naturgesetzen eine klare Vorstellung über die Zustandsänderung *im nächsten Augenblick,* d.h. also über den Wert der zeitlichen Ableitung der Zustandsgrößen in Abhängigkeit vom gegenwärtigen Zustand des Systems.

§1: Definitionen und erste Beispiele

a) Wurfparabel

Ein einfaches Beispiel hierfür liefert das zweite NEWTONsche Gesetz, wonach die zeitliche Ableitung des Impuls eines Teilchens gleich der auf das Teilchen wirkenden Kraft ist.

Ein in die Luft geworfener Gegenstand bewegt sich unter gewissen Bedingungen näherungsweise auf einer parabelförmigen Bahn. Wir wollen diese etwas vage Aussage präzisieren und mathematisch herleiten.

Es gibt viele Wurftechniken, und nur wenige davon können auf einfache Weise durch ein mathematisches Modell beschrieben werden; wir ignorieren daher den genauen Vorgang des Abwurfs und gehen davon aus, daß der Gegenstand *irgendwie* eine Anfangsgeschwindigkeit $\vec{v}$ erreicht hat im Abwurfpunkt mit Koordinaten (x_0, y_0, z_0); den Zeitpunkt des Abwurfs bezeichnen wir mit t_0.

Als nächstes nehmen wir an, daß wir den Luftwiderstand vernachlässigen können, eine Annahme, die beim Kugelstoßen kaum zu Fehlern führt, die aber beispielsweise für einen Fallschirmspringer (auch mit geschlossenem Fallschirm) oder einen Papierflieger völlig unrealistisch ist. Als nächstes wollen wir auch noch annehmen, daß wir nur relativ geringe Wurfhöhen erreichen, so daß die Erdanziehung als konstant angenommen werden kann.

Die Bewegung des Gegenstandes wird dann durch zwei Naturgesetze bestimmt: Das Gravitationsgesetz beschreibt den Effekt der Erdanziehung, und das zweite NEWTONsche Gesetz sagt uns, wie sich diese Kraft auf die Bewegung des Gegenstands auswirkt. Die Gravitation können wir aufgrund der gemachten Annahmen als konstant annehmen, d.h. auf einen Körper der Masse m wirkt die Kraft gm, wobei $g \approx 9{,}8\, m/s^2$ die Gravitationsbeschleunigung an der Erdoberfläche ist; bei „üblicher" Ausrichtung des Koordinatensystems wirkt sie in Richtung der negativen z-Achse. Diese Gravitationskraft ist nach dem zweiten NEWTONschen Gesetz gleich der Ableitung des Impulses nach der Zeit; wenn wir die Masse m als konstant voraussetzen, ist das also gleich m mal der Ableitung der Geschwindigkeit oder m mal der zweiten Ableitung des Orts. Wir haben somit das Differentialgleichungssystem

$$\ddot{x}(t) = 0, \quad \ddot{y}(t) = 0 \quad \text{und} \quad \ddot{z}(t) = -g\,.$$

Diese Gleichungen sind erfüllt, wann immer $x(t)$ und $y(t)$ lineare Funktionen von t sind und $z(t)$ eine quadratische Funktion mit führendem Koeffizienten $-g$. Die sechs noch fehlenden Koeffizienten dieser drei Polynomfunktionen geben uns die Anfangsbedingungen: Zum Zeitpunkt $t = t_0$ ist

$$x(t_0) = x_0, \quad y(t_0) = y_0 \quad \text{und} \quad z(t_0) = z_0\,,$$

und die Geschwindigkeit ist $\vec{v}$, d.h.

$$\dot{x}(t_0) = v_1, \quad \dot{y}(t_0) = v_2 \quad \text{und} \quad \dot{z}(t_0) = v_3\,.$$

Also ist

$$x(t) = v_1(t - t_0) + x_0, \quad y(t) = v_2(t - t_0) + y_0$$

und

$$z(t) = -g(t - t_0)^2 + v_3(t - t_0) + z_0\,.$$

Diese Gleichungen beschreiben in der Tat fast immer eine Parabel: Falls wir die x-Achse des Koordinatensystems so wählen, daß die Anfangsgeschwindigkeit $\vec{v}$ in der (x, z)-Ebene liegt, ist $v_2 = 0$. Falls auch v_1 verschwindet, falls wir den Gegenstand also senkrecht nach oben (oder gar unten) werfen, sind $x(t) = x_0$ und $y(t) = y_0$ konstant und nur

$$z(t) = -g(t - t_0)^2 + v_3(t - t_0) + z_0$$

hängt von der Zeit ab. Andernfalls können wir durch v_1 dividieren; wir erhalten

$$t - t_0 = \frac{x(t) - x_0}{v_1} \quad \text{und} \quad z(t) = \frac{-g}{v_1}\big(x(t) - x_0\big)^2 + \frac{v_3}{v_1}\big(x(t) - x_0\big) + z_0\,,$$

die Punkte $\big(x(t), z(t)\big)$ liegen also in der Tat auf einer Parabel.

b) Radioaktiver Zerfall

Das gerade durchgerechnete Beispiel war insofern untypisch für Differentialgleichungen, als auf den rechten Seite der Gleichungen nur Konstanten standen; üblicherweise wird man dort Funktionen erwarten, die nicht nur von t abhängen (so daß man einfach integrieren kann), sondern auch noch von den gesuchten Funktionen. Beim radioaktiven Zerfall etwa ist die pro (kleiner) Zeiteinheit zerfallende Masse proportional zur noch vorhandenen Masse, es gibt also eine Konstante $\lambda > 0$, die sogenannte Zerfallskonstante, so daß die zum Zeitpunkt t vorhandene Masse $m(t)$ der Gleichung

$$\dot{m}(t) = -\lambda m(t)$$

genügt – zumindest, wenn diese Masse hinreichend groß ist. (Im atomaren Bereich muß man auch statistische Effekte berücksichtigen, aber ab etwa 10^{10} Atomen können die für alle praktischen Fälle vernachlässigt werden.)

Wir kennen bereits eine Funktion, die sich so verhält, wie es die obige Differentialgleichung angibt, nämlich die Exponentialfunktion $e^{-\lambda t}$, und natürlich entspricht auch für jedes konstante Vielfache dieser Funktion die Differentiation einfach der Multiplikation mit $-\lambda$. Das sind dann aber bereits alle Funktionen mit dieser Eigenschaft, denn der Quotient

$$q(t) = \frac{m(t)}{e^{-\lambda t}} = m(t) \cdot e^{\lambda t}$$

einer Lösungsfunktion und der Funktion $e^{-\lambda t}$ hat die Ableitung

$$\dot{q}(t) = \dot{m}(t) \cdot e^{\lambda t} + m(t) \cdot \lambda e^{\lambda t} = -\lambda m(t) \cdot e^{\lambda t} + \lambda m(t) \cdot e^{\lambda t} = 0\,,$$

ist also gleich einer Konstanten c, so daß

$$m(t) = c \cdot e^{-\lambda t}$$

ist. Indem wir $t = 0$ setzen, sehen wir, daß die Konstante $c = m(0)$ gleich der zum Zeitpunkt 0 vorhandenen Masse ist; falls wir stattdessen die Masse $m_0 = m(t_0)$ zu einem anderen Zeitpunkt t_0 kennen, können wir analog zum obigen Beispiel auch schreiben

$$m(t) = m_0 e^{-\lambda(t - t_0)} = (m_0 e^{\lambda t_0}) \cdot e^{-\lambda t}\,.$$

c) Differentialgleichungen und Differentialgleichungssysteme

Wir betrachten ein System, das durch n zeitlich veränderliche Größen $y_1(t), \ldots, y_n(t)$ beschrieben wird; unter einem System von Differentialgleichungen oder kurz einer Differentialgleichung verstehen wir eine Vorschrift, die die zeitlichen Ableitungen $\dot{y}_1(t), \ldots, \dot{y}_n(t)$ aus den Funktionswerten berechnet:

$$\begin{aligned}
\dot{y}_1(t) &= f_1\big(t, y_1(t), \ldots, y_n(t)\big)\\
\dot{y}_2(t) &= f_2\big(t, y_1(t), \ldots, y_n(t)\big)\\
&\;\;\vdots \quad \vdots\\
\dot{y}_n(t) &= f_n\big(t, y_1(t), \ldots, y_n(t)\big)\,.
\end{aligned}$$

Falls die Funktionen f_i nur von $y_1(t), \ldots, y_n(t)$ abhängen und nicht auch noch direkt von der Zeit t spricht man von einem *autonomen* System. Da Naturgesetze nicht von der Zeit abhängen, hat man es in wissenschaftlich-technischen Anwendungen meist mit autonomen Systemen zu tun; man kann allerdings auch den Einfluß von Umgebungsgrößen in einem zeitabhängigen Term zusammenfassen und so ein nichtautonomes System erhalten.

Falls wir, wie im obigen Beispiel die Werte der beteiligten Funktionen zu einem festen Zeitpunkt $t = t_0$ kennen, reden wir von einem *Anfangswertproblem;* die meisten Anwendungen gewöhnlicher Differentialgleichungen sind Anfangswertprobleme.

Auch Differentialgleichungen, in denen wie dort im Beispiel höhere Ableitungen vorkommen, lassen sich so interpretieren: Wenn wir dort die drei Komponenten des Geschwindigkeitsvektors als neue Funktionen

$$u(t) = \dot{x}(t), \quad v(t) = \dot{y}(t) \quad \text{und} \quad w(t) = \dot{z}(t)$$

eingeführt hätten, wäre das System

$$\begin{aligned}
\dot{x}(t) &= u(t) \\
\dot{y}(t) &= v(t) \\
\dot{z}(t) &= w(t) \\
\dot{u}(t) &= 0 \\
\dot{v}(t) &= 0 \\
\dot{w}(t) &= -g
\end{aligned}$$

von obiger Form, und wir hätten für jede der sechs beteiligten Funktionen den Wert an der Stelle $t = t_0$ gekannt.

Ein für die Informationstechnik wichtiger Spezialfall sind Gleichungen der Form

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0y(t) = b(t)\,,$$

die sogenannten linearen Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten. Auch diese Gleichungen lassen sich leicht auf die obige Form bringen: Wir betrachten n neue Funktionen

$$y_0(t), y_1(t), \ldots, y_{n-1}(t)$$

mit der Idde, daß sich $y_i(t)$ so verhalten soll wie die i-te Ableitung von $y(t)$. Dann betrachten wir das Differentialgleichungssystem

$$\begin{aligned}
\dot{y}_0(t) &= y_1(t) \\
\dot{y}_1(t) &= y_2(t) \\
\vdots \quad & \vdots \\
\dot{y}_{n-2}(t) &= y_{n-1}(t) \\
\dot{y}_{n-1}(t) &= b(t) - a_{n-1}y^{(n-1)}(t) - \cdots - a_1\dot{y}(t) - a_0y(t)\,.
\end{aligned}$$

Für jede Lösung $y(t)$ der obigen Gleichung ist dann das n-tupel

$$\big(y(t), \dot{y}(t), \ddot{y}(t), \ldots, y^{(n-1)}(t)\big)$$

eine Lösung des Systems, und für jede Lösung

$$\big(y_0(t), y_1(t), y_2(t), \ldots, y_{n-1}(t)\big)$$

des Differentialgleichungssystems ist $y_0(t)$ eine Lösung der obigen Gleichung.

Auch wenn das Differentialgleichungssystem als Anfangswertproblem gegeben ist, läßt sich das leicht in Anfangswerte für die Gleichung höherer Ordung umschreiben: Hier werden die Werte $y(t_0)$, $\dot{y}(t_0)$ *usw.* bis $y^{(n-1)}(t_0)$ vorgegeben.

d) Systeme linearer Differentialgleichungen

Im letzten Abschnitt haben wir gesehen, daß Differentialgleichungen höherer Ordnung interessante Anwendungen haben; wir wissen aber bereits, daß man solche Gleichungen stets umschreiben kann in Systeme, die nur erste Ableitungen enthalten, und zumindest in den gerade betrachteten Beispielen hatten die neu eingeführten Funktionen auch eine physikalische Interpretation als Geschwindigkeit *bzw.* Stromstärke. Die Auffassung als System hat außerdem den Vorteil, daß wir den Apparat der linearen Algebra anwenden können, der uns eine strukturelle Übersicht über die Lösungsmenge einer Differentialgleichung gibt – zumindest im linearen Fall.

Wir betrachten in diesem Abschnitt Systeme von Differentialgleichungen, wie sie zu Beginn dieses Paragraphen definiert wurden, unter der (sehr) einschränkenden Voraussetzungen, daß die rechten Seiten linear in den gesuchten Funktionen $y_1(t), \ldots, y_n(t)$ sind; wir betrachten also ein System

$$\begin{aligned}
\dot{y}_1(t) &= a_{11}(t)y_1(t) + a_{12}(t)y_2(t) + \cdots + a_{1n}(t)y_n(t) + b_1(t) \\
\dot{y}_2(t) &= a_{21}(t)y_1(t) + a_{22}(t)y_2(t) + \cdots + a_{2n}(t)y_n(t) + b_2(t) \\
\vdots \quad & \quad \vdots \\
\dot{y}_n(t) &= a_{n1}(t)y_1(t) + a_{n2}(t)y_2(t) + \cdots + a_{nn}(t)y_n(t) + b_n(t)
\end{aligned} \quad . \qquad (*)$$

Eventuell haben wir noch Anfangsbedingungen der Form

$$y_1(t_0) = c_0, \quad y_2(t_0) = c_2, \quad \ldots, \quad y_n(t_0) = c_n \qquad (**)$$

für ein festes $t_0 \in \mathbb{R}$.

Für das Beispiel des elektrischen Schwingkreises mit angelegter Wechselspannung etwa haben wir bei dieser Sicht der Dinge die beiden Funktionen $Q(t)$, die Ladung des Kondensators zum Zeitpunkt t, und $I(t) = \dot{Q}(t)$, die resultierende Stromstärke; das Differentialgleichungssystem $(*)$ wird zu

$$\begin{aligned} \dot{Q}(t) &= I(t) \\ \dot{I}(t) &= -\frac{R}{L} I(t) - \frac{Q(t)}{LC} + A_0 \cos \omega_0 t \end{aligned} .$$

Wir können die Funktionen $y_i(t)$, $b_i(t)$ und die Anfangswerte c_i jeweils zu Vektoren zusammenfassen und die Koeffizientenfunktionen $a_{ij}(t)$ zu einer Matrix; mit

$$\vec{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \qquad \vec{b}(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix}, \qquad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

und

$$A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \ldots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \ldots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \ldots & a_{nn}(t) \end{pmatrix},$$

bekommt das System die übersichtlichere Form

$$\dot{\vec{y}}(t) = A(t) \cdot \vec{y}(t) + \vec{b}(t),$$

wobei die Ableitung eines Vektors von Funktionen natürlich der Vektor der abgeleiteten Funktionen sein soll. Falls es Anfangsbedingungen gibt, haben sie nun die kompakte Form $\vec{y}(t_0) = \vec{c}$.

In Analogie zu den linearen Gleichungssystemen bezeichnen wir das System $(*)$ als *homogen,* wenn $\vec{b}(t)$ der Nullvektor ist, wenn also alle

Funktionen $b_i(t)$ verschwinden; andernfalls bezeichnen wir es als *inhomogen.* Das homogene System zu einem gegebenen inhomogenen System soll einfach dasjenige System sein, in dem alle $b_i(t)$ durch null ersetzt wurden.

Wie bei linearen Gleichungssystemen gilt auch hier

Lemma: *a)* Die Menge aller Lösungen eines *homogenen* Differentialgleichungssystems der Form $(*)$ ist ein $\mathbb{R}$-Vektorraum.
b) Ist das System nicht homogen und ist $\vec{y}(t)$ eine feste Lösung, so läßt sich jede andere Lösung $\vec{z}(t)$ schreiben als $\vec{z}(t) = \vec{y}(t) + \vec{x}(t)$ mit einer Lösung $\vec{x}(t)$ des zugehörigen homogenen Systems; die Lösungsmenge ist also ein affiner Raum.

Beweis: a) Wir müssen zeigen, daß für zwei Lösungen $\dot{\vec{x}}(t)$ und $\dot{\vec{y}}(t)$ eines homogenen Systems auch jede Linearkombination $\lambda\dot{\vec{x}}(t) + \mu\dot{\vec{y}}(t)$ mit $\lambda, \mu \in \mathbb{R}$ wieder eine Lösung ist. Das ist aber klar, denn wenn

$$\dot{\vec{x}}(t) = A(t) \cdot \vec{x}(t) \quad \text{und} \quad \dot{\vec{y}}(t) = A(t) \cdot \vec{y}(t)$$

ist, gilt auch

$$\begin{aligned} \frac{d}{dt}\big(\lambda\vec{x}(t) + \mu\vec{y}(t)\big) &= \lambda\dot{\vec{x}}(t) + \mu\dot{\vec{y}}(t) = \lambda A(t) \cdot \vec{x}(t) + \mu A(t) \cdot \vec{y}(t) \\ &= A(t) \cdot \big(\lambda\vec{x}(t) + \mu\vec{y}(t)\big). \end{aligned}$$

b) Sind $\vec{y}(t)$ und $\vec{z}(t)$ zwei Lösungen von $(*)$, so ist

$$\dot{\vec{y}}(t) = A(t) \cdot \vec{y}(t) + \vec{b}(t) \quad \text{und} \quad \dot{\vec{z}}(t) = A(t) \cdot \vec{z}(t) + \vec{b}(t);$$

die Differenz $x(t) = z(t) - y(t)$ hat somit die Ableitung

$$\begin{aligned} \dot{\vec{x}}(t) = \dot{\vec{z}}(t) - \dot{\vec{y}}(t) &= \Big(A(t) \cdot \vec{z}(t) + \vec{b}(t)\Big) - \Big(A(t) \cdot \vec{z}(t) + \vec{b}(t)\Big) \\ &= A(t) \cdot \vec{z}(t) - A(t) \cdot y(t) = A(t) \cdot \Big(\vec{z}(t) - \vec{y}(t)\Big) = A(t) \cdot \vec{x}(t) \end{aligned}$$

und $x(t)$ löst also in der Tat das zugehörige homogene System. ■

Um die Lösungsmenge des Differentialgleichungssystems $(*)$ zu verstehen, müssen wir nach diesem Lemma zwei Teilaufgaben lösen:

1.) Wir müssen den Vektorraum der Lösungen des homogenen Systems bestimmen.
2.) Wir müssen uns wenigstens eine Lösung des inhomogenen Systems verschaffen – oder zumindest wissen, daß eine existiert.

Um im einfachsten Fall zu sehen, wie so etwas funktionieren könnte, betrachten wir ein „System“ aus genau einer Gleichung

$$\dot{y}(t) = a(t) \cdot y(t) + b(t)\,;$$

dabei nehmen wir an, daß y eine differenzierbare Funktion von $\mathbb{R}$ nach $\mathbb{R}$ sei und $a, b\colon \mathbb{R} \to \mathbb{R}$ stetige Funktionen.

Wir beginnen mit der Lösung des homogenen Systems

$$\dot{y}(t) = a(t) \cdot y(t)\,.$$

Unter der Annahme, daß wir das dürfen, dividieren wir durch $y(t)$ und erhalten

$$\frac{\dot{y}(t)}{y(t)} = a(t)\,.$$

Der Quotient links ist bekanntlich die logarithmische Ableitung von $y(t)$; falls dies nicht mehr bekannt sein sollte, zeigt eine einfache Anwendung der Kettenregel, daß in einem Intervall, in dem $y(t)$ positiv ist,

$$\frac{d}{dt} \ln y(t) = \frac{\dot{y}(t)}{y(t)} = a(t)$$

ist. In einem Intervall, in dem $y(t)$ negativ ist, gilt entsprechend

$$\frac{d}{dt} \ln\bigl(-y(t)\bigr) = \frac{-\dot{y}(t)}{-y(t)} = \frac{\dot{y}(t)}{y(t)} = a(t)\,,$$

und allgemein haben wir somit

$$\frac{d}{dt} \ln |y(t)| = \frac{\dot{y}(t)}{y(t)} = a(t)$$

in jedem Intervall, in dem $y(t)$ nirgends verschwindet.

Integration beider Seiten führt auf

$$\ln |y(t)| = \int a(t)\,dt + C \quad \text{oder} \quad y(t) = e^{\int a(t)\,dt + C} = e^{C} \cdot e^{\int a(t)\,dt}$$

oder

$$y(t) = \pm e^{C} \cdot e^{\int a(t)\,dt}\,,$$

wobei das Vorzeichen wegen der Stetigkeit von y im gesamten Intervall konstant ist, da die Exponentialfunktion nie null wird.

Damit ist in diesem Fall das erste Problem auf eine einfache Integration zurückgeführt.

Bleibt noch die Frage, was passiert, wenn $y(t)$ an irgendeinem Punkt t_0 eine Nullstelle haben sollte. Wir wollen uns überlegen, daß $y(t)$ dann auch für jedes $t > t_0$ verschwinden muß. Falls nicht, gibt es einen Punkt $t_1 > t_0$, so daß $y(t_1) \neq 0$ ist. Wegen der Stetigkeit von $y(t)$ ist die Funktion dann auch in einer Umgebung von t_1 von Null verschieden, d.h. dort können wir die obigen Argumente anwenden und sehen, daß $y(t)$ dort die Form $e^{h(t)}$ hat mit irgendeiner Funktion h. Da y als differenzierbare Funktion insbesondere überall stetig sein muß und $e^{h(t)}$ nirgends verschwindet, ist das nicht möglich. Genauso überlegt man sich, daß $y(t)$ für jedes $t < t_0$ verschwinden muß, $y(t)$ ist also gleich der Nullfunktion. Diese ist somit die einzige Lösung, die noch zusätzlich betrachtet werden muß. Insbesondere folgt daraus auch, daß eine Lösungsfunktion, die in irgendeinem Punkt positiv *bzw.* negativ ist, überall positiv *bzw.* negativ sein muß, denn eine stetige Funktion kann ihr Vorzeichen nur wechseln, wenn sie in irgendeinem Punkt null wird; wie wir gerade gesehen haben, ist das genau dann der Fall, wenn sie überall verschwindet.

Somit hat jede Lösung die Form

$$y(t) = a e^{\int a(t)\,dt} \quad \text{mit einem } a \in \mathbb{R}\,.$$

Insbesondere bilden diese Lösungen einen eindimensionalen Vektorraum.

Auch eine spezielle Lösung der inhomogenen Gleichung läßt sich in diesem Fall angeben: Die Methode der *Variation der Konstanten,* mit der wir uns später noch genauer beschäftigen werden, führt auf den Kandidaten

$$y(t) = \left(\int b(t) \cdot e^{-\int a(t)\,dt}\,dt \right) \cdot e^{\int a(t)\,dt}\,,$$

und in der Tat rechnet man leicht nach, daß nach der Produktregel

$$\begin{aligned}\dot{y}(t) &= \left(b(t)\cdot e^{-\int a(t)\,dt}\right)\cdot e^{\int a(t)\,dt}\\ &\quad + \left(\int b(t)\cdot e^{-\int a(t)\,dt}\,dt\right)\cdot\left(a(t)\cdot e^{\int a(t)\,dt}\right)\\ &= b(t) + a(t)\cdot\left(\int b(t)\cdot e^{-\int a(t)\,dt}\,dt\right)\cdot e^{\int a(t)\,dt}\\ &= a(t)\cdot y(t) + b(t)\end{aligned}$$

ist.

Es wäre schön, wenn wir im mehrdimensionalen Fall genauso vorgehen könnten. Zumindest für homogene Systeme bietet sich an, formal genauso vorzugehen wie im eindimensionalen: Dort hatten wir

$$\dot{y}(t) = a(t)\cdot y(t) \Longrightarrow y(t) = ce^{\int a(t)\,dt} \quad \text{mit} \quad c\in\mathbb{R}\,.$$

Warum sollte also nicht gelten

$$\dot{\vec{y}}(t) = A(t)\cdot\vec{y}(t) \Longrightarrow y(t) = e^{\int A(t)\,dt}\cdot\vec{c} \quad \text{mit} \quad \vec{c}\in\mathbb{R}^n \quad ?$$

Das Problem dabei ist nur, daß wir nicht die geringste Ahnung haben, was die rechte Seite hier bedeuten soll; unser nächstes Ziel wird sein, ihr eine Bedeutung zu geben und uns dann zu überlegen, ob *bzw.* unter welchen Bedingungen die obige Formel korrekt ist.

e) Die Matrixexponentialfunktion

Wir orientieren uns wieder am Eindimensionalen: Für eine reelle Zahl x ist

$$e^x = \sum_{i=0}^{\infty}\frac{x^i}{i!}\,,$$

also setzen wir analog für eine $n\times n$-Matrix X

$$e^X \underset{\text{def}}{=} \sum_{i=0}^{\infty}\frac{1}{i!}\cdot X^i\,.$$

Damit ist klar, daß e^X eine $n\times n$-Matrix sein soll, und das erklärt auch, warum oben der Konstantenvektor $\vec{y}_0$ auf der rechten Seite steht. Was wir uns noch überlegen müssen, ist die Konvergenz der Reihe.

Dazu müssen wir die Größe der Einträge in den Matrizen X^i abschätzen: Sind allgemein A, B zwei $n\times n$-Matrizen und sind die Beträge aller Einträge von A kleiner oder gleich a und die von B kleiner oder gleich b, so kann es in AB offensichtlich keinen Eintrag geben, dessen Betrag größer ist als nab: Schließlich ist jeder Eintrag in der Produktmatrix eine Summe von n Summanden, deren jeder Produkt je eines Eintrags von A und von B ist.

Um diese Formel leichter anwenden zu können, machen wir sie mutwillig schlechter und begnügen uns damit, daß jeder Eintrag von AB höchstens den Betrag $(na)\cdot(nb)$ hat.

Ist nun x der Betrag des größten Eintrags in der Matrix X, so folgt induktiv sofort, daß in X^i höchstens Zahlen bis zum Betrag $(nx)^i$ stehen können; in der endlichen Teilsumme

$$\sum_{i=0}^{M}\frac{1}{i!}X^i$$

hat daher jeder Eintrag einen Betrag kleiner

$$\sum_{i=0}^{M}\frac{(nx)^i}{i!}\,.$$

Letztere Summe konvergiert für $M\to\infty$ gegen e^{nx}, und damit muß auch die Matrixsumme absolut konvergieren, denn die Reihe für e^{nx} ist konvergente Majorante des Betrags eines jeden Eintrags. Insbesondere hat jeder Eintrag von e^X höchstens den Betrag e^{nx}.

Damit wissen wir also, daß die Matrix e^X für jede $n\times n$-Matrix X existiert; somit ist die Funktion $t\mapsto e^{At}$ wohldefiniert.

f) Eigenschaften der Matrixexponentialfunktion

Wir können natürlich nicht erwarten, daß die Matrixexponentialfunktion alle schönen Eigenschaften der gewöhnlichen Exponentialfunktion erbt.

Beispielsweise ist nur schwer vorstellbar, daß für beliebige Matrizen A und B gelten sollte $e^{A+B} = e^A \cdot e^B$: Da für zwei Matrizen A und B stets $A + B = B + A$ ist, müßte dann auch $e^A \cdot e^B = e^B \cdot e^A$ sein, was zumindest unwahrscheinlich aussieht. In der Tat ist etwa für

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

sowohl A^2 als auch B^2 gleich der Nullmatrix, d.h.

$$e^A = E + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad e^B = E + B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

und

$$e^A \cdot e^B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Das Quadrat von $C = A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist aber gleich der Einheitsmatrix und daher ist

$$\begin{aligned} e^{A+B} &= E + C + \frac{1}{2!}E + \frac{1}{3!}C + \frac{1}{4!}E + \frac{1}{5!}C + \cdots \\ &= \left(\sum_{i=0}^{\infty} \frac{1}{(2i)!}\right) \cdot E + \left(\sum_{i=0}^{\infty} \frac{1}{(2i+1)!}\right) \cdot C \\ &= \cosh 1 \cdot E + \sinh 1 \cdot C = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e + e^{-1} & e - e^{-1} \\ e - e^{-1} & e + e^{-1} \end{pmatrix}. \end{aligned}$$

Allgemeiner ist

$$e^{At} = E + tA = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad e^{Bt} = E + tB = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$$

und

$$e^{Ct} = \sum_{i=0}^{\infty} \frac{1}{i!} C^i = \left(\sum_{i=0}^{\infty} \frac{1}{(2i)!}\right) \cdot E + \left(\sum_{i=0}^{\infty} \frac{1}{(2i+1)!}\right) \cdot C = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

Zum Glück gilt aber wenigstens

Lemma: Für zwei Matrizen $A, B \in \mathbb{R}^{n \times n}$ mit $AB = BA$ ist

$$e^{A+B} = e^A \cdot e^B = e^B \cdot e^A.$$

Insbesondere ist für $s, t \in \mathbb{R}$

$$e^{A(s+t)} = e^{As} \cdot e^{At}.$$

Beweis: Für zwei reelle Zahlen x, y ist $e^{x+y} = e^x e^y = e^y e^x$; damit gilt dieselbe Formel auch für zwei reellwertige Variablen x und y. Wenn wir in allen Potenzreihen alle x- und y-Potenzen oberhalb der N-ten ignorieren, sagt die Gleichung aus, daß drei Polynome in x und y *als Polynome* identisch sind.

Beim Rechnen mit Polynomen in x und y verwendet man keine speziellen Eigenschaften dieser Variablen *außer, daß sie kommutieren.* Damit kann man in so eine Polynomidentität auch kommutierende Matrizen A und B einsetzen: Beispielsweise führt die Polynomidentität

$$(x + y)^2 = x^2 + 2xy + y^2$$

zur Identität

$$(A + B)^2 = A^2 + 2AB + B^2,$$

die wegen der für beliebige Matrizen gültigen Gleichung

$$(A + B)^2 = A^2 + AB + BA + B^2$$

für kommutierende Matrizen in der Tat erfüllt ist.

Damit ist für kommutierende Matrizen A, B speziell stets

$$e^{A+B} = e^A \cdot e^B = e^B \cdot e^A.$$

Da zwei skalare Vielfache derselben Matrix stets miteinander kommutieren, folgt damit auch die letzte Aussage des Lemmas. ■

Die für uns wichtigste Anwendung hiervon ist

Satz: Für jede $n \times n$-Matrix A ist die Funktion

$$\begin{cases} \mathbb{R} \to \mathbb{R}^{n \times n} \\ t \mapsto e^{At} \end{cases}$$

stetig differenzierbar mit Ableitung $t \mapsto A \cdot e^{At} = e^{At} \cdot A$.

Beweis: Die Ableitung ist definiert als

$$\lim_{h\to 0} \frac{e^{A(t+h)} - e^{At}}{h} .$$

Da die Matrizen At und Ah miteinander vertauschbar sind, ist nach dem gerade bewiesenen Lemma $e^{A(t+h)} = e^{At} \cdot e^{Ah}$, also

$$\frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \cdot \frac{e^{Ah} - E}{h} .$$

Dabei ist

$$\frac{e^{Ah} - E}{h} = \frac{1}{h} \cdot \sum_{i=1}^{\infty} \frac{(Ah)^i}{i!} = A + A^2 h \sum_{i=2}^{\infty} \frac{(Ah)^{i-2}}{i!} = A + A^2 h \sum_{i=0}^{\infty} \frac{(Ah)^i}{(i+2)!} .$$

Nach der obigen Diskussion ist jeder Eintrag der Matrix in der rechtsstehenden Summenmatrix höchstens gleich

$$\sum_{i=0}^{\infty} \frac{(ah)^i}{(i+2)!} = \frac{e^{ah} - (1+ah)}{a^2 h^2} ,$$

bleibt also insbesondere beschränkt. Dies gilt auch für $h \to 0$, denn wie zweimalige Anwendung der DE L'HOSPITALschen Regel oder TAYLOR-Entwicklung zeigen, ist der Grenzwert dann $\frac{1}{2}$. Damit existiert

$$\lim_{h\to 0} \sum_{i=1}^{\infty} \frac{(Ah)^i}{(i+2)!} ,$$

und somit ist $\frac{d}{dt} e^{At} = \lim_{h\to 0} \frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \cdot A$, da der Vorfaktor $A^2 h$ gegen Null geht. Dies ist auch gleich $A \cdot e^{At}$, denn da A mit jeder seiner Potenzen vertauschbar ist, ist es auch mit jeder endlichen Teilsumme der Reihe von e^{At} vertauschbar, also auch mit e^{At} selbst. ■

Am Ende des vorigen Abschnitts hatten wir gehofft, daß vielleicht auch für jede matrixwertige Funktion $A(t)$ gelten könnte, daß

$$\frac{d}{dt} e^{A(t)} = \dot{A}(t) e^{A(t)}$$

ist; dies war offensichtlich zu optimistisch: Da

$$e^{A(t+h)} = e^{A(t) + h\dot{A}(t) + o(h)}$$

ist, bräuchten wir für einen Beweis nach obigem Vorbild, daß $A(t)$ und $\dot{A}(t)$ miteinander kommutieren; dies ist aber im allgemeinen nicht der Fall. Für

$$A(t) = \begin{pmatrix} 1 & t \\ 1 & 1 \end{pmatrix}$$

beispielsweise ist

$$\dot{A}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} ,$$

also

$$A(t) \cdot \dot{A}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} , \quad \text{aber} \quad \dot{A}(t) \cdot A(t) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} .$$

Dies ist natürlich kein Beweis dafür, daß die Ableitung von $e^{A(t)}$ ungleich $\dot{A}(t) \cdot e^{A(t)}$ ist, aber im vorliegenden Fall ist die Ableitung in der Tat verschieden sowohl von $\dot{A}(t) e^{A(t)}$ als auch von $e^{A(t)} \cdot A(t)$: Mit den Methoden, die wir im nächsten Abschnitt kennenlernen werden, können wir durch eine (alles andere als angenehme) Rechnung zeigen, daß

$$e^{A(t)} = \begin{pmatrix} \dfrac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{2} & \dfrac{\sqrt{t}\left(e^{1+\sqrt{t}} - e^{1-\sqrt{t}}\right)}{2} \\ \dfrac{e^{1+\sqrt{t}} - e^{1-\sqrt{t}}}{2\sqrt{t}} & \dfrac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{2} \end{pmatrix} ,$$

$$\frac{d}{dt} e^{A(t)} = \begin{pmatrix} \dfrac{\left(e^{2\sqrt{t}} - 1\right) e^{1-\sqrt{t}}}{4\sqrt{t}} & \dfrac{\left(-1 + \sqrt{t} + \sqrt{t} e^{2\sqrt{t}} + e^{2\sqrt{t}}\right) e^{1-\sqrt{t}}}{4\sqrt{t}} \\ \dfrac{\left(1 - e^{2\sqrt{t}} + \sqrt{t} + \sqrt{t} e^{2\sqrt{t}}\right) e^{1-\sqrt{t}}}{4t\sqrt{t}} & \dfrac{\left(e^{2\sqrt{t}} - 1\right) e^{1-\sqrt{t}}}{4\sqrt{t}} \end{pmatrix} ,$$

aber

$$\dot{A}(t) \cdot e^{A(t)} = \begin{pmatrix} \dfrac{e^{1+\sqrt{t}} - e^{1-\sqrt{t}}}{2\sqrt{t}} & \dfrac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{2} \\ 0 & 0 \end{pmatrix}$$

und

$$e^{A(t)} \cdot \dot{A}(t) = \begin{pmatrix} 0 & \dfrac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{2} \\ 0 & \dfrac{e^{1+\sqrt{t}} - e^{1-\sqrt{t}}}{2\sqrt{t}} \end{pmatrix}$$

ist. Wir müssen uns bei diesem Ansatz also begnügen mit linearen homogenen Differentialgleichungen mit *konstanten* Koeffizienten.

Alles was uns zu deren theoretischer Lösung jetzt noch fehlt sind Rechenregeln für den Umgang mit Ableitungen von Matrixfunktionen; für die praktische Lösung fehlen natürlich auch noch Verfahren zur effizienten Berechnung der Matrixexponentialfunktion.

Zumindest für Summen und Produkte gelten, wenn man von der Nichtkommutativität der Multiplikation absieht, für matrixwertige Funktionen die üblichen Regeln:

Lemma: *a)* $F, G\colon (a,b) \to \mathbb{R}^{n\times n}$ seien zwei matrixwertige Funktionen auf dem offenen Intervall (a,b). Dann ist

$$\frac{d}{dt}\big(F(t)+G(t)\big) = \dot{F}(t) + \dot{G}(t)\,.$$

b) Für $F\colon (a,b) \to \mathbb{R}^{n\times m}$ und $G\colon (a,b) \to \mathbb{R}^{m\times p}$ ist

$$\frac{d}{dt}\big((F(t)G(t)\big) = \dot{F}(t)\cdot G(t) + F(t)\cdot \dot{G}(t)\,.$$

c) Für einen konstanten Vektor $\vec{v} \in \mathbb{R}^n$ ist

$$\frac{d}{dt}\big(F(t)\cdot \vec{v}\big) = \dot{F}(t)\cdot \vec{v}\,.$$

Beweis: a) Sind $f_{ij}(t)$ und $g_{ij}(t)$ die Komponenten der Matrizen F und G, so sind die Summen $f_{ij}(t)+g_{ij}(t)$ die Komponenten von $F+G$, und deren Ableitung ist die Summe der Ableitungen.

b) Die (i,j)-Komponente von FG ist die Funktion $\sum_{\nu=1}^{m} f_{i\nu}(t)\cdot g_{\nu j}(t)$, und deren Ableitung ist

$$\sum_{\nu=1}^{m}\Big(\dot{f}_{i\nu}(t)\cdot g_{\nu j}(t) + f_{i\nu}(t)\cdot \dot{g}_{\nu j}(t)\Big)\,,$$

die (i,j)-Komponente von $\dot{F}(t)\cdot G(t) + F(t)\cdot \dot{G}(t)$.

c) Ist der Spezialfall $n+m$ und $p=1$ von *b)*, wobei zusätzlich noch G eine konstante Funktion ist, so daß alle Terme mit $\dot{G}(t)$ verschwinden. ■

Damit haben wir alles zusammen und können zeigen

Satz: Die sämtlichen Lösungen des homogenen Differentialgleichungssystems $\dot{y} = Ay$ sind genau die Funktionen $t \mapsto e^{At}\vec{y}_0$ mit $\vec{y}_0 \in \mathbb{R}^n$.

Beweis: Da e^{At} die Ableitung Ae^{At} hat, ist die Ableitung der vektorwertigen Funktion $\vec{f}(t) = e^{At}\cdot \vec{y}_0$ nach der zuletzt bewiesenen Formel gleich $A\cdot e^{At}\cdot \vec{y}_0$, also in der Tat gleich $A\cdot \vec{f}(t)$.

Nun sei $\vec{y}\colon (a,b) \to \mathbb{R}^n$ irgendeine differenzierbare vektorwertige Funktion mit der Eigenschaft, daß $\dot{\vec{y}}(t) = A\cdot \vec{y}(t)$ ist. Wir betrachten die Funktion $\vec{g}(t) = e^{-At}\cdot \vec{y}(t)$. Deren Ableitung ist

$$\dot{\vec{g}}(t) = -A\cdot e^{-At}\cdot \vec{y}(t) + e^{-At}\cdot \dot{\vec{y}}(t) = -A\cdot e^{-At}\cdot \vec{y}(t) + e^{-At}\cdot A\vec{y}(t) = \vec{0}\,,$$

denn die Matrix A ist mit e^{-At} vertauschbar. Damit haben alle Komponenten von $\vec{g}$ die Ableitung Null, sind also konstant, und somit ist $\vec{g}(t) \underset{\text{def}}{=} \vec{y}_0$ ein konstanter Vektor mit der Eigenschaft, daß

$$\vec{y}_0 = e^{-At}\cdot \vec{y}(t), \quad d.h. \quad \vec{y}(t) = e^{At}\cdot \vec{y}_0\,,$$

wie behauptet. ■

Korollar: Für jeden Vektor $\vec{y}_0 \in \mathbb{R}^n$ und jede reelle Zahl $t_0 \in \mathbb{R}$ gibt es genau eine differenzierbare Funktion $\vec{y}(t)$ mit den Eigenschaften, daß $\dot{\vec{y}}(t) = A\vec{y}(t)$ und $\vec{y}(t_0) = \vec{y}_0$ ist; dies ist $\vec{y}(t) = e^{A(t-t_0)}\cdot \vec{y}_0$. ■

Wir wußten bereits, daß die Lösungen einen Vektorraum bilden; obiges Korollar sagt uns, daß dieser Vektorraum die Dimension n hat und daß für jede reelle Zahl t_0 die Abbildung $\vec{y} \mapsto \vec{y}(t_0)$ ein Isomorphismus auf den $\mathbb{R}^n$ ist.

Als erstes Beispiel betrachten wir das Anfangswertproblem

$$\dot{x}(t) = y(t) \quad \text{und} \quad \dot{y}(t) = x(t) \quad \text{mit} \quad x(0) = a \quad \text{und} \quad y(0) = b$$

oder

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \quad \text{mit} \quad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Die Koeffizientenmatrix ist gleich der zu Beginn des Abschnitts betrachteten Beispielmatrix C, deren Exponentialfunktion wird dort berechnet haben; die Lösung des Anfangswertproblems ist also

$$e^{Ct}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a\cosh t + b\sinh t \\ a\sinh t + b\cosh t \end{pmatrix}.$$

§ 2: Eigenwerte, Eigenvektoren und Hauptvektoren

Die Matrixexponentialfunktion ist zwar wohldefiniert, aber eine matrixwertige Potenzreihe ist für allgemeine Matrizen A nicht gerade einfach zu berechnen. Wir brauchen daher alternative Rechenverfahren.

Zumindest ein Fall ist problemlos: Ist nämlich

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

eine Diagonalmatrix, ist offensichtlich

$$e^{Dt} = \begin{pmatrix} e^{d_1 t} & 0 & \dots & 0 \\ 0 & e^{d_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{d_n t} \end{pmatrix}$$

wieder eine Diagonalmatrix, wobei die Exponentialfunktion einfach komponentenweise auf die Diagonaleinträge ihres Arguments angewandt wird.

Noch ein weiterer Fall ist relativ unproblematisch: für eine obere (oder untere) Dreiecksmatrix N mit Nullen in der Hauptdiagonalen. Eine solche Matrix definiert eine lineare Abbildung $\mathbb{C}^n \to \mathbb{C}^n$, die den k-ten Einheitsvektor $\vec{e}_k$ in den von $\vec{e}_{k+1}$ bis $\vec{e}_n$ erzeugten Untervektorraum abbildet. Das Quadrat von N bildet ihn entsprechend in den von $\vec{e}_{k+2}$ bis $\vec{e}_n$ erzeugten Untervektorraum ab und so weiter, spätestens N^n ist also die Nullmatrix. Damit wird die Potenzreihe der Exponentialfunktion zu einer endlichen Summe, die, wir wir bereits in zwei Beispielen gesehen haben, zumindest für kleine n leicht berechnet werden kann.

In diesem Paragraphen wollen wir sehen, daß sich die Berechnung einer beliebigen Matrixexponentialfunktion auf diese beiden Spezialfälle zurückführen läßt.

Wir arbeiten dabei über einem beliebigen Körper k, denn auch wenn uns bei Differentialgleichungen nur die Fälle $k = \mathbb{R}$ und $k = \mathbb{C}$ interessieren, hat die hier entwickelte Theorie doch auch interessante Anwendungen über anderen Körpern: Eigenvektoren über endlichen Körpern spielen beispielsweise bei einigen Problem der Signalverarbeitung eine Rolle.

a) Eigenwerte und Eigenvektoren

Nicht nur bei der Berechnung der Matrixexponentialfunktion sind Diagonalmatrizen sehr viel angenehmer als allgemeine Matrizen. Eine Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \dots, \lambda_n$ hat die Eigenschaft, daß sie den i-ten Einehitsvektor $ve_i \in k^n$ auf dessen Vielfaches $\lambda_i \vec{e}_i$ abbildet, und diese Eigenschaft soll uns als Ausgangspunkt für die Definition von Eigenwerten und Eigenvektoren dienen:

Definition: *a)* V sei ein k-Vektorraum. Ein Vektor $\vec{v} \in V \setminus \{\vec{0}\}$ heißt *Eigenvektor* der linearen Abbildung $\varphi\colon V \to V$ zum *Eigenwert* $\lambda \in k$, wenn $\varphi(\vec{v}) = \lambda\vec{v}$ ist.

b) $\lambda \in k$ heißt *Eigenwert* von φ, falls φ einen Eigenvektor zum Eigenwert λ hat.

c) Eigenwerte und Eigenvektoren einer Matrix $A \in k^{n\times n}$ sind die Eigenwerte und Eigenvektoren der linearen Abbildung

$$\varphi\colon \begin{cases} k^n \to k^n \\ \vec{v} \mapsto A\vec{v} \end{cases}.$$

Offensichtlich ist mit einem Vektor $\vec{v}$ auch jedes Vielfache (außer dem nach Definition ausgeschlossenen Nullvektor) ein Eigenvektor zum selben Eigenwert; allgemeiner ist sogar jede Linearkombination (außer $\vec{0}$)

von Eigenvektoren zum Eigenwert λ wieder ein Eigenvektor zum Eigenwert λ, d.h. die Eigenvektoren zu einem festen Eigenwert λ bilden zusammen mit dem Nullvektor einen Untervektorraum von V, den sogenannten *Eigenraum* von λ.

Definition: Die Dimension des Eigenraums von λ heißt *geometrische Vielfachheit* des Eigenwerts λ.

Lemma: Sind $\vec{v}_1, \ldots, \vec{v}_r \in V$ Eigenvektoren der linearen Abbildung $\varphi: V \to V$ zu verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r$, so sind diese Vektoren linear unabhängig.

Beweis: Angenommen, $\vec{v}_1, \ldots, \vec{v}_r$ seien linear abhängig. Dann können wir eine Zahl $2 \leq s \leq r$ finden, so daß zwar $\vec{v}_1, \ldots, \vec{v}_s$ linear abhängig sind, nicht aber $\vec{v}_1, \ldots, \vec{v}_{s-1}$. Es gibt daher Skalare $\alpha_i \in k$, so daß

$$\alpha_1 \vec{v}_1 + \cdots + \alpha_s \vec{v}_s = \vec{0}$$

ist. Wenden wir auf beide Seiten dieser Gleichung die Abbildung φ an und beachten, daß $\varphi(\vec{v}_i) = \lambda_i \vec{v}_i$ ist, folgt, daß auch

$$\alpha_1 \lambda_1 \vec{v}_1 + \cdots + \alpha_s \lambda_s \vec{v}_s = \vec{0}$$

ist. Andererseits können wir obige Gleichung auch einfach mit λ_s multiplizieren mit dem Ergebnis, daß

$$\lambda_s \alpha_1 \vec{v}_1 + \cdots + \lambda_s \alpha_s \vec{v}_s = \vec{0}\,.$$

Durch Subtraktion der letzten beiden Gleichungen voneinander erhalten wir eine lineare Abhängigkeit

$$\alpha_1(\lambda_s - \lambda_1)\vec{v}_1 + \cdots + \alpha_{s-1}(\lambda_s - \lambda_{s-1})\vec{v}_{s-1} = \vec{0}$$

zwischen $\vec{v}_1, \ldots, \vec{v}_{s-1}$. Da diese Vektoren linear unabhängig sind, müssen alle Koeffizienten verschwinden. Da aber die Eigenwerte $\lambda_1, \cdots, \lambda_s$ allesamt verschieden sind, ist dies nur möglich, wenn $\alpha_1 = \cdots = \alpha_{s-1} = 0$ ist. Wegen $\vec{v}_s \neq \vec{0}$ muß dann aber auch α_s verschwinden, im Widerspruch zur angenommenen linearen Unabhängigkeit von $\vec{v}_1, \ldots, \vec{v}_s$.

Also sind die Vektoren $\vec{v}_1, \ldots, \vec{v}_r$ linear unabhängig. ∎

Eigenwerte und Eigenvektoren sind auch interessant für Selbstabbildungen eines unendlichdimensionalen Vektorraums: Ist $V = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ beispielsweise der Vektorraum aller beliebig oft stetig differenzierbarer reeller Funktionen, so sind $\sin \omega t$ und $\cos \omega t$ Eigenvektoren der linearen Abbildung

$$\varphi: \begin{cases} \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \\ f \mapsto \ddot{f} \end{cases}$$

zum Eigenwert $-\omega^2$; genauso sind $\sinh \omega t$ und $\cosh \omega t$ Eigenvektoren zum Eigenwert ω^2. Für $\omega = 0$ degenerieren diese beiden Eigenvektoren jeweils zu null und eins, wodurch ein Eigenwert verschwindet; dafür kommt die Identität als neuer Eigenvektor hinzu. Damit ist also jede reelle Zahl Eigenwert von φ mit einer geometrischen Vielfachheit von mindestens zwei. (Wir werden im nächsten Paragraphen sehen, daß die Vielfachheit immer gleich zwei ist.)

Eigenwertprobleme für lineare Abbildungen, die durch Differentialoperatoren gegeben sind, spielen in vielen Anwendungen eine wichtige Rolle; im Hinblick auf solche Anwendungen bezeichnet man die Menge aller Eigenwerte einer linearen Abbildung oder Matrix auch als deren *Spektrum.* Dieses Wort kommt daher, daß z.B. beim (mehrdimensionalen) Differentialoperator, der die Schwingungen des Fells einer Trommel beschreibt, die Eigenwerte gerade die Frequenzen sind, die die Trommel produzieren kann.

Uns interessieren hauptsächlich Eigenwerte und Eigenvektoren in endlichdimensionalen Vektorräumen. Dort können wir konkret mit Matrizen rechnen; ist A die Abbildungsmatrix zu $\varphi: V \to V$, so ist $\varphi(\vec{v}) = \lambda \vec{v}$ äquivalent dazu, daß $A\vec{v} = \lambda \vec{v}$ oder $(A - \lambda E)\vec{v} = \vec{0}$ ist, wobei E wie üblich die Einheitsmatrix bezeichnet.

In letzterer Form ist dies ein homogenes lineares Gleichungssystem für die Komponenten von $\vec{v}$; wie jedes homogene lineare Gleichungssystem hat es den Nullvektor als Lösung, der allerdings nach Definition *nie* ein Eigenvektor ist. Weitere Lösungen gibt es genau dann, wenn die Matrix $A - \lambda E$ des Gleichungssystems singulär ist, wenn also $\det(A - \lambda E)$ verschwindet. Somit ist $\lambda \in k$ genau dann ein Eigenwert, wenn $\det(A - \lambda E) = 0$ ist; die zugehörigen Eigenvektoren sind

die nichttrivialen Lösungen des homogenen linearen Gleichungssystems $(A - \lambda E)\vec{v} = \vec{0}$.

Damit ist klar, wie man Eigenwerte und Eigenvektoren berechnen kann: Man löse die Gleichung $\det(A - \lambda E) = 0$ und dann für jede Nullstelle λ_i dieser Gleichung das lineare Gleichungssystem $(A - \lambda E)\vec{v} = \vec{0}$. Dieses homogene lineare Gleichungssystem hat *nie* maximalen Rang, da es nach Definition eines Eigenwerts nichttriviale Lösungen geben muß; kommt man also auf ein eindeutig lösbares Gleichungssystem (und damit auf den Nullvektor als einzige Lösung), ist das immer ein Zeichen für einen Rechenfehler.

Dies zeigt auch, daß die numerische Bestimmung von Eigenvektoren nicht nach obigem Schema vorgehen kann: Falls man die Nullstellen der charakterischen Gleichung nur näherungsweise kennt, kann die Matrix des linearen Gleichungssystems eine leicht von null verschiedene Determinante haben, so daß das Gleichungssystem nur die triviale Lösung hat. In der Numerik werden Eigenwerte und Eigenvektoren daher simultan berechnet – sofern man beides braucht. Es gibt auch numerische Algorithmen, die nur die (angenäherten) Eigenwerte bestimmen ohne $\det(A - \lambda E)$ zu berechnen; dies ist vor allem nützlich für große n, wo die Berechnung einer Determinanten sehr aufwendig wäre. Für Genaueres sei auf die Numerikvorlesung verwiesen.

b) Ein erstes Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{pmatrix} .$$

Hier ist

$$A - \lambda E = \begin{pmatrix} 1-\lambda & 2 & 3 & 4 \\ 5 & 6-\lambda & 7 & 8 \\ 8 & 7 & 6-\lambda & 5 \\ 4 & 3 & 2 & 1-\lambda \end{pmatrix} ,$$

und nach dem Entwicklungssatz ist (bei Entwicklung nach der ersten Zeile)

$$\begin{aligned}
\det(A - \lambda E) &= (1-\lambda)\begin{vmatrix} 6-\lambda & 7 & 8 \\ 7 & 6-\lambda & 5 \\ 3 & 2 & 1-\lambda \end{vmatrix} \\
&\quad - 2\begin{vmatrix} 5 & 7 & 8 \\ 8 & 6-\lambda & 5 \\ 4 & 2 & 1-\lambda \end{vmatrix} \\
&\quad + 3\begin{vmatrix} 5 & 6-\lambda & 8 \\ 8 & 7 & 5 \\ 4 & 3 & 1-\lambda \end{vmatrix} - 4\begin{vmatrix} 5 & 6-\lambda & 7 \\ 8 & 7 & 6-\lambda \\ 4 & 3 & 2 \end{vmatrix} \\
&= (1-\lambda)(-\lambda^3 + 13\lambda^2 + 35\lambda) - 2(5\lambda^2 + 53\lambda) \\
&\quad + 3(-8\lambda^2 + \lambda) - 4(4\lambda^2 - 17\lambda) \\
&= \lambda^4 - 14\lambda^3 - 72\lambda^2 = \lambda^2(\lambda^2 - 14\lambda - 72) .
\end{aligned}$$

Dieser Ausdruck verschwindet genau dann, wenn entweder λ verschwindet oder der Klammerausdruck ganz hinten. Letzterer ist eine quadratische Gleichung für λ, die man entweder nach der üblichen Formel lösen kann, oder aber man beachtet den Wurzelsatz von VIÈTE, wonach die Summe der beiden Lösungen 14 und ihr Produkt -72 sein muß; da $72 = 4 \times 18$ ist, können die beiden Lösungen daher nur -4 und 18 sein. Die Nullstellen des Polynoms sind also $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = -4$ und $\lambda_4 = 18$.

Die Eigenwerte von A sind damit 0, -4 und 18; für jeden dieser Werte müssen wir über ein lineares Gleichungssystem die Eigenvektoren bestimmen.

Beginnen wir mit $\lambda_1 = \lambda_2 = 0$. Hier ist das Gleichungssystem einfach $A\vec{x} = \vec{0}$, d.h.

$$\begin{aligned}
x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\
5x_1 + 6x_2 + 7x_3 + 8x_4 &= 0 \\
8x_1 + 7x_2 + 6x_3 + 5x_4 &= 0 \\
4x_1 + 3x_2 + 2x_3 + x_4 &= 0 .
\end{aligned}$$

Aufgrund der speziellen Struktur der Matrix A ergänzen sich die erste und die vierte Gleichung zu

$$5(x_1 + x_2 + x_3 + x_4) = 0\,;$$

entsprechend addieren sich die beiden mittleren Gleichungen zu

$$13(x_1 + x_2 + x_3 + x_4) = 0\,.$$

Da außerdem noch die Differenz zwischen den ersten beiden Gleichungen auf

$$4(x_1 + x_2 + x_3 + x_4) = 0$$

führt, genügt es, außer der ersten Gleichung noch als einzige weitere Gleichung

$$x_1 + x_2 + x_3 + x_4 = 0 \quad \text{oder} \quad x_1 = -(x_2 + x_3 + x_4)$$

zu betrachten. Subtrahiert man diese Gleichung von der ersten oder, was dasselbe ist, setzt man sie ein, so bleibt

$$x_2 + 2x_3 + 3x_4 = 0$$

als einzige Relation zwischen x_2, x_3 und x_4. Das gegebene Gleichungssystem ist also äquivalent zu

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_2 + 2x_3 + 3x_4 &= 0\,; \end{aligned}$$

insbesondere hat es also den Rang zwei und damit einen Lösungsraum der Dimension $4 - 2 = 2$.

(Es ist klar, daß der Rang des Gleichungssystems kleiner als vier sein muß, da es nichttriviale Lösungen gibt. Bei der Berechnung von Eigenvektoren über lineare Gleichungssysteme gibt es also immer Abhängigkeiten zwischen den einzelnen Gleichungen.)

Im noch verbliebenen Gleichungssystem kann beispielsweise x_4 beliebig gewählt werden. Ist $x_4 = 0$, so bleibt noch die Relation $x_2 + 2x_3 = 0$ übrig, in der etwa x_3 beliebig gewählt werden kann; danach sind $x_2 = -2x_3$ und $x_1 = x_3$ eindeutig festgelegt; die entsprechenden Lösungen sind also die Vielfachen der Lösung $(1, -2, 1, 0)$.

Genausogut könnte etwa $x_3 = 0$ gesetzt werden; dann wäre $x_2 = -3x_4$ und $x_1 = 2x_4$; dies führt auf die Vielfachen der Lösung $(2, -3, 0, 1)$.

Damit haben wir zwei linear unabhängige Lösungen gefunden; da der Lösungsraum zweidimensional ist, sind die sämtlichen Lösungen des Gleichungssystems die Linearkombinationen

$$x_1 = -\lambda - 2\mu, \quad x_2 = 2\lambda - 3\mu, \quad x_3 = \lambda \quad \text{und} \quad x_4 = \mu$$

dieser beiden Lösungen; die beiden Vektoren

$$\begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

bilden also eine Basis des Eigenraums von A zum Eigenwert Null.

Genauso lassen sich auch die Eigenwerte -4 und 18 behandeln: -4 führt auf das lineare Gleichungssystem

$$\begin{aligned} 5x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ 5x_1 + 10x_2 + 7x_3 + 8x_4 &= 0 \\ 8x_1 + 7x_2 + 10x_3 + 5x_4 &= 0 \\ 4x_1 + 3x_2 + 2x_3 + 5x_1 &= 0\,, \end{aligned}$$

dessen sämtliche Lösungen die Vielfachen von $(-1, -1, 1, 1)$ sind, und 18 führt auf

$$\begin{aligned} -17x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ 5x_1 - 12x_2 + 7x_3 + 8x_4 &= 0 \\ 8x_1 + 7x_2 - 12x_3 + 5x_4 &= 0 \\ 4x_1 + 3x_2 + 2x_3 - 17x_4 &= 0 \end{aligned}$$

mit den Vielfachen von $(5, 13, 13, 5)$ als Lösungsraum.

Was haben wir nun erreicht? Betrachten wir die lineare Abbildung

$$\varphi\colon \mathbb{R}^4 \to \mathbb{R}^4; \quad \vec{v} \mapsto A \cdot \vec{v}\,.$$

Für die Vektoren

$$\vec{b}_1 = \lambda \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \ \vec{b}_2 = \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}, \ \vec{b}_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \ \text{und} \ \vec{b}_4 = \begin{pmatrix} 5 \\ 13 \\ 13 \\ 5 \end{pmatrix}$$

ist

$$\varphi(\vec{b}_1) = 0\,\vec{b}_1, \quad \varphi(\vec{b}_2) = 0\,\vec{b}_2, \quad \varphi(\vec{b}_3) = -4\,\vec{b}_3 \quad \text{und} \quad \varphi(\vec{b}_4) = 18\,\vec{b}_4\,;$$

bezüglich der neuen Basis $\{\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4\}$ hat φ daher die Abbildungsmatrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix},$$

bei der die Eigenwerte von A in der Hauptdiagonalen stehen und alle sonstigen Einträge verschwinden. Mit dieser Matrix läßt sich sehr viel angenehmer rechnen als mit der ursprünglichen Matrix A; insbesondere kann ihre Exponentialfunktion

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-4} & 0 \\ 0 & 0 & 0 & e^{18} \end{pmatrix}$$

sehr einfach berechnet werden.

Damit läßt sich auch e^A berechnen: Sind nämlich

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

die vier Einheitsvektoren des $\mathbb{R}^4$, so ist

$$\vec{b}_i = B\vec{e}_i \quad \text{mit} \quad B = \begin{pmatrix} 1 & 2 & -1 & 5 \\ -2 & -3 & -1 & 13 \\ 1 & 0 & 1 & 13 \\ 0 & 1 & 1 & 5 \end{pmatrix},$$

und die Gleichung

$$A\vec{b}_i = \lambda_i \vec{b}_i$$

mit $\lambda_{1/2} = 0$, $\lambda_3 = -4$ und $\lambda_4 = 18$ wird zu

$$AB\vec{e}_i = \lambda_i B\vec{b}_i \quad \text{oder} \quad B^{-1}AB\,\vec{e}_i = \lambda_i \vec{e}_i\,.$$

Also ist

$$B^{-1}AB = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix}$$

oder

$$A = B \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix} B^{-1}\,.$$

Für eine beliebige Matrix M ist

$$e^{BMB^{-1}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(BMB^{-1}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!} BM^nB^{-1} = Be^M B^{-1}\,,$$

also ist

$$e^A = B \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-4} & 0 \\ 0 & 0 & 0 & e^{18} \end{pmatrix} B^{-1}\,,$$

was sich zumindest im Prinzip berechnen läßt – auch wenn das Ergebnis

$$\frac{1}{72} \begin{pmatrix} 27e^{-4}+35+10e^{18} & 10e^{18}-19+9e^{-4} & 10e^{18}-1-9e^{-4} & 10e^{18}+17-27e^{-4} \\ -53+27e^{-4}+26e^{18} & 9e^{-4}+26e^{18}+37 & 26e^{18}-17-9e^{-4} & 26e^{18}+1-27e^{-4} \\ 26e^{18}+1-27e^{-4} & 26e^{18}-17-9e^{-4} & 9e^{-4}+26e^{18}+37 & -53+27e^{-4}+26e^{18} \\ 10e^{18}+17-27e^{-4} & 10e^{18}-1-9e^{-4} & 10e^{18}-19+9e^{-4} & 27e^{-4}+35+10e^{18} \end{pmatrix}$$

alles andere als angenehm ist. Dies zeigt wieder einmal, wieviel man sich ersparen kann, wenn man *vor* Beginn einer Rechnung eine gute Basis *bzw.* ein gutes Koordinatensystem wählt.

c) Das charakteristische Polynom und seine Nullstellen

Es ist kein Zufall, daß im obigen Beispiel die Gleichung $\det(A - \lambda E) = 0$ auf ein Polynom vierten Grades führte: Ist $A = (a_{ij})$ eine $n \times n$-Matrix, so hat die Matrix $A - \lambda E$ in der Diagonalen die Einträge $a_{ii} - \lambda$,

ansonsten stimmen alle Einträge mit denen von A überein. Berechnet man daher $\det(A - \lambda E)$ gemäß der definierenden Formel, so gibt es genau ein Produkt, in dem n mit λ behaftete Faktoren vorkommen, nämlich das Produkt

$$(a_{11} - \lambda) \cdots (a_{nn} - \lambda) = (-1)^n \lambda^n + \textit{Terme niedrigerer Ordnung}$$

der Diagonaleinträge. Die restlichen Produkte, die zur Determinante aufsummiert werden, enthalten zwischen null und $n-1$ mit λ behaftete Faktoren, die Summe ist also ein Polynom vom Grad n mit höchstem Term $(-1)^n \lambda^n$.

Definition: Das Polynom $\det(A - \lambda E)$ heißt *charakteristisches Polynom* der Matrix A.

Demgemäß sind also die Eigenwerte von A gleich den Nullstellen des charakteristischen Polynoms von A, und wir sollten uns wenigsten kurz überlegen, wie man die Nullstellen eines solchen Polynoms bestimmen kann.

Zur Bestimmung der Eigenwerte muß man somit die Nullstellen des charakteristischen Polynoms finden. Für ein Polynom vom Grad höchstens zwei (oder aber ein Polynom, das man als Produkt solcher Polynome schreiben kann) ist das nicht schwer: Nullstellen eines linearen Polynoms erhält man durch eine einfache Division, solche eines quadratischen durch quadratische Ergänzung: Da für $a \neq 0$

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}$$

ist, hat das linksstehende Polynom die Nullstellen

$$x_{1/2} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a} - c} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\,.$$

Bei Polynomen höherer Grade, die man nicht auf einfache Weise über binomische Formeln oder ähnliches in kleinere Faktoren zerlegen kann, ist es oft einen Versuch wert, einige der Lösungen zu *erraten,* um so den Grad des Polynoms zu reduzieren.

Bei Polynomen mit ganzzahligen (und eventuell auch rationalen) Nullstellen, ist dazu der Wurzelsatz von VIÈTE ein vielversprechender Ansatzpunkt: Angenommen, das Polynom n-ten Grades

$$f(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0$$

mit höchstem Koeffizient eins habe die Nullstellen $z_1, \ldots, z_n$. Dann ist

$$f(x) = (x - z_1)(x - z_2) \cdot \cdots \cdot (x - z_n)\,.$$

Dies läßt sich ausmultiplizieren und liefert dann einen Zusammenhang zwischen Nullstellen und Koeffizienten: Beispielsweise ist

$$a_0 = (-1)^n z_1 z_2 \cdots z_n \quad \text{und} \quad a_{n-1} = -(z_1 + z_2 + \cdots + z_n)\,,$$

und genauso zeigt man auch daß der allgemeine Koeffizient a_i die Summe aller möglicher Produkte aus $n-i$ Nullstellen z_j ist, multipliziert mit $(-1)^{n-i}$. Diese Aussage bezeichnet man als den Wurzelsatz von VIÈTE.

FRANÇOIS VIÈTE (1540–1603) studierte Jura an der Universität Poitiers, danach arbeitete er als Hauslehrer. 1573, ein Jahr nach dem Massaker an den Hugenotten, berief ihn CHARLES IX (obwohl VIÈTE Hugenotte war) in die Regierung der Bretagne; unter HENRI III wurde er geheimer Staatsrat. 1584 wurde er auf Druck der katholischen Liga vom Hofe verbannt und beschäftigte sich fünf Jahre lang mit Mathematik. Unter HENRI IV arbeitete er wieder am Hof und knackte u.a. verschlüsselte Botschaften an den spanischen König PHILIP II. In seinem Buch *In artem analyticam isagoge* rechnete er als erster systematisch mit symbolischen Größen.

Für das Erraten von Nullstellen einfacher Polynome, bei denen man (aus inhaltlichen Gründen oder aber weil so etwas in Übungs- und Klausuraufgaben fast die Regel ist) ganzzahlige Lösungen erwartet, ist vor allem die erstgenannte Beziehung wichtig: In der Form

$$(-1)^n a_0 = z_1 z_2 \cdots z_n$$

gibt sie das Produkt aller Nullstellen. Falls die a_i alle ganzzahlig sind, lohnt es sich also, die Teiler von a_0 zu testen. Ist beispielsweise

$$f(x) = x^4 + 14x^3 - 52x^2 - 14x + 51\,,$$

so ist

$$a_0 = 51 = 3 \cdot 17 .$$

Da das Produkt aller Nullstellen gleich diesem Wert sein muß, kommen – falls *alle* Nullstellen ganzzahlig sind – für diese nur die Werte $\pm 1, \pm 3$ und ± 17 in Frage. Da das Produkt aller vier Nullstellen gleich 51 ist, gibt es jeweils genau eine Nullstelle vom Betrag 3 *bzw.* 17, sowie zwei Nullstellen vom Betrag eins. Welche Vorzeichen wirklich auftreten, läßt sich durch Einsetzen festellen oder aber auch dadurch, daß nach VIÈTE die *Summe* aller Nullstellen gleich -14 sein muß. Das ist offenbar nur möglich, wenn sowohl +1 als auch -1 Nullstellen sind, sowie -17 und +3. In der Tat zeigt Einsetzen, daß dies auch tatsächlich Nullstellen sind. (Das Einsetzen ist notwendig, da wir nicht sicher sein können, daß wirklich alle Nullstellen ganzzahlig sind.)

In diesem extrem einfachen (und konstrierten) Fall führt also die Primfaktorzerlegung direkt zur Lösung; in komplizierteren Fällen, wenn a_0 mehr Primfaktoren hat, muß man zunächst alle Kombinationsmöglichkeiten, die zum Produkt a_0 führen können, in Betracht ziehen und davon dann durch Einsetzen potentieller Nullstellen alle bis auf die tatsächlichen Nullstellen eliminieren.

Beim Polynom

$$f(x) = x^6 + 27x^5 - 318x^4 - 5400x^3 - 10176x^2 + 27648x + 32768$$

etwa ist $a_0 = 32768 = 2^{15}$; hier wissen wir also nur, daß jede Nullstelle die Form $\pm 2^i$ haben muß, wobei die Summe aller Exponenten gleich 15 sein muß und die Anzahl der negativen Vorzeichen gerade. Einsetzen zeigt, daß

$$-1, \quad 2, \quad -4, \quad -8, \quad 16, \quad -32$$

die Nullstellen sind.

Man beachte, daß diese Vorgehensweise nur funktioniert, wenn das Polynom höchsten Koeffizienten eins hat; andernfalls ist das Produkt der Nullstellen gleich dem Quotienten aus konstantem Koeffizienten und führendem Koeffizienten mal $(-1)^{\text{Grad}}$.

Falls man nicht sicher sein kann, daß alle Nullstellen ganzzahlig sind, gibt es immer noch eine ganze Reihe von Methoden, um Nullstellen

exakt zu berechnen: Beispielsweise kennt die Computeralgebra Algorithmen, um ein Polynom (soweit dies möglich ist) in ein Produkt von Polynomen kleineren Grades zu zerlegen mit Koeffizienten aus einem vorgegebenen Körper, der (in einem hier nicht präzisierten) Sinne nicht *zu weit* vom Körper der rationalen Zahlen *bzw.* einem endlichen Körper entfernt ist, und es gibt auch, seit der ersten Hälfte des sechzehnten Jahrhunderts, allgemeine Formeln zur Lösung von Gleichungen dritten und vierten Grades. Diese Formeln spielen wegen ihrer Komplexität und numerischen Instabilität in der Praxis keine sonderlich große Rolle und sollen daher hier nur im Kleindruck behandelt werden:

Für die kubische Gleichung

$$ax^3 + bx^2 + cx + d = 0$$

wenden wir zunächst einen ähnlichen Trick an wie die quadratische Ergänzung beim Fall der quadratischen Gleichungen: Durch die Substitution

$$z = x + \frac{b}{3a}$$

wird die Gleichung zu

$$az^3 + \left(c - \frac{b^2}{3a}\right) z + \frac{2b^3}{27a^2} + d ,$$

was wir auch kurz als

$$z^3 + pz + q = 0$$

schreiben können. Zur Lösung dieser Gleichung ersetzen wir z durch die Summe

$$z = u + v$$

zweier Variablen und erhalten

$$u^3 + 3u^2v + 3uv^2 + v^3 = u^3 + v^3 + 3uv(u + v) + p(u + v) + q = 0 .$$

Da die Zerlegung von z in eine Summe äußerst willkürlich ist, können wir hoffen, daß diese Gleichung für die beiden Variablen u und v auch Lösungen hat, wenn wir zusätzliche Bedingungen stellen: Die obige Gleichung für z wird beispielsweise sicherlich dann gelöst, wenn

$$u^3 + v^3 = -q \quad \text{und} \quad 3uv = -p$$

ist. Dann ist

$$u^3 + v^3 = -q \quad \text{und} \quad u^3 \cdot v^3 = -\frac{p^3}{3} ,$$

wir kennen also Summe und Produkt von u^3 und v^3.

Sind aber Summe und Produkt zweier Zahlen r und s bekannt, so können wir leicht die Zahlen selbst bestimmen: Aus

$$r + s = c \quad \text{und} \quad rs = d$$

folgt, daß

$$r(c-r) = -r^2 + cr = d \quad \text{oder} \quad r^2 - cr + d = 0$$

ist; wir müssen also einfach eine quadratische Gleichung lösen und erhalten

$$r = \frac{c}{2} \pm \sqrt{\frac{c^2}{4} - d}\,.$$

Da die Summe dieser beiden Lösungen gleich c ist, muß also die eine gleich r und die andere gleich s sein.

Auf die kubische Gleichung angewandt heißt das, daß u^3 und v^3 die beiden Zahlen

$$-\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}$$

sind, also

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \quad \text{und} \quad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

oder umgekehrt.

Uns interessiert nur die Summe der beiden Zahlen, also

$$z = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}\,.$$

Damit sind wir fast fertig. Das verbleibende Problem ist, daß hier formal eine Lösung steht, wohingegen wir für eine kubische Gleichung *drei* Lösungen erwarten. Dieses Problem verkehrt sich sofort in sein Gegenteil, wenn wir beachten, daß genauso, wie die Quadratwurzel nur bis aufs Vorzeichen bestimmt ist, die Kubikwurzel nur bis dritte Einheitswurzeln bestimmt ist: Da die drei komplexen Zahlen

$$1, \quad \frac{-1+i\sqrt{3}}{2} \quad \text{und} \quad \frac{-1-i\sqrt{3}}{2}$$

alle dritte Potenz eins haben, ist mit jeder Kubikwurzel w einer Zahl y auch w mal einer dieser drei Zahlen Kubikwurzel; es gibt also (für $y \neq 0$) drei verschiedene Kubikwurzeln, und somit hat obige Formel für z gleich *neun* mögliche Interpretationen.

Daraus können wir die drei richtigen herausfiltern, wenn wir beachten, daß wir nicht nur das Produkt von u^3 und v^3 kennen, sondern auch das von u und v, nämlich $-p/3$. Damit ist der zweite Summand in der Formel für z eindeutig durch den ersten bestimmt, und es gibt nur die zu erwartenden drei Lösungen: Ist

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

irgendeiner der drei möglichen Werte der Wurzel, so ist

$$z = u - \frac{p}{3u}$$

eine Lösung der Gleichung $z^3 + pz + q = 0$ und

$$x = z - \frac{b}{3a}$$

eine Lösung der ursprünglichen Gleichung $ax^3 + bx^2 + cx + d = 0$.

Auch biquadratische Gleichungen lassen sich auflösen: Hier eliminiert man den kubischen Term von

$$ax^4 + bx^3 + cx^2 + dx + e = 0$$

durch die Substitution

$$z = x + \frac{b}{4a}\,;$$

dies führt auf eine Gleichung der Form

$$z^4 + pz^2 + qz + r = 0\,.$$

Für eine beliebige Zahl y folgt daraus für jede Nullstelle z dieser Gleichung die Beziehung

$$(z^2+y)^2 = z^4 + 2yz^2 + y^2 = (2y-p)z^2 - qz + y^2 - r\,.$$

Falls rechts das Quadrat eines linearen Polynoms $sz+t$ steht, ist

$$(z^2+y)^2 = (sz+t)^2 \Longrightarrow z = \pm\sqrt{-y \pm (sz+t)}\,,$$

wir können die Gleichung also auflösen.

Nun wird die rechte Seite

$$(2y-p)z^2 - qz + y^2 - r$$

im allgemeinen kein Quadrat eines linearen Polynoms in z sein, wir können aber hoffen, daß es zumindest für gewisse spezielle Werte der bislang noch willkürlichen Konstante y eines ist.

Ein quadratisches Polynom

$$\alpha z^2 + \beta z + \gamma$$

ist genau dann Quadrat eines linearen, wenn die beiden Nullstellen der quadratischen Gleichung

$$\alpha z^2 + \beta z + \gamma = 0$$

übereinstimmen. Nach der obigen Lösungsformel für quadratische Gleichungen ist dies genau dann der Fall, wenn dort der Ausdruck unter der Wurzel verschwindet, d.h. wenn

$$\beta^2 - 4\alpha\gamma = 0$$

ist. In unserem Fall muß also

$$q^2 - 4(2y-p)(y^2-r) = -8y^3 + 4py^2 + 8ry + q^2 - 4pr$$

verschwinden. Dies ist eine kubische Gleichung für y; indem wir diese Gleichung lösen und eine der Lösungen für y einsetzen, erhalten wir die vier Lösungen der biquadratischen Gleichung.

Die erste Lösung einer kubischen Gleichung geht wohl aus SCIPIONE DEL FERRO (1465–1526) zurück, der von 1496 bis zu seinem Tod an der Universität Bologna lehrte. 1515 fand er eine Methode, um die Nullstellen von $x^3+px=q$ für *positive* Werte von p und q zu bestimmen (Negative Zahlen waren damals in Europa noch nicht im Gebrauch). Er veröffentlichte diese jedoch nie, so daß NICCOLO FONTANA (1499–1557, oberes Bild), genannt TARTAGLIA (der Stotterer), dieselbe Methode 1535 noch einmal entdeckte und gleichzeitig auch noch eine Modifikation, um einen leicht verschiedenen Typ kubischer Gleichungen zu lösen. TARTAGLIA war mathematischer Autodidakt, war aber schnell als Fachmann anerkannt und konnte seinen Lebensunterhalt als Mathematiklehrer in Verona und Venedig verdienen.

Die Lösung allgemeiner kubischer Gleichungen geht auf den Mathematiker, Arzt und Naturforscher GIROLAMO CARDANO (1501–1576, unteres Bild) zurück, dem TARTAGLIA nach langem Drängen und unter dem Siegel der Verschwiegenheit seine Methode mitgeteilt hatte. LODOVICO FERRARI (1522–1565) kam 14-jährig als Diener zu CARDANO; als dieser merkte, daß FERRARI schreiben konnte, machte er ihn zu seinem Sekretär. 1540 fand er die Lösungsmethode für biquadratische Gleichungen; 1545 veröffentlichte CARDANO in seinem Buch *Ars magna* die Lösungsmethodem für kubische und biquadratische Gleichungen.

Nach der erfolgreichen Auflösung der kubischen und biquadratischen Gleichungen in der ersten Hälfte des sechzehnten Jahrhunderts beschäftigten sich natürlich viele Mathematiker mit dem nächsten Fall, der Gleichung fünften Grades. Hier gab es jedoch über 250 Jahre lang keinerlei Fortschritt, bis zu Beginn des neunzehnten Jahrhunderts ABEL glaubte, eine Lösung gefunden zu haben. Er entdeckte dann aber recht schnell seinen Fehler und bewies stattdessen 1824, daß es *unmöglich* ist, die Lösungen einer allgemeinen Gleichung fünften (oder höheren) Grades durch Grundrechenarten und Wurzeln auszudrücken.

Die Grundidee seines Beweises liegt in der Betrachtung von Symmetrien innerhalb der Lösungsmenge, ähnlich wie wir in einem späteren Abschnitt einige Differentialgleichungen durch Symmetriebetrachtungen lösen werden. Unmöglichkeitsbeweise sind allerdings deutlich aufwendiger als Lösungsversuche mit Hilfe von Symmetriebetrachtungen; daher kann über Einzelheiten des ABELschen Beweises hier nichts weiter gesagt werden. Interessenten finden ihn in fast jedem Algebralehrbuch im Kapitel über GALOIS-Theorie.

Der norwegische Mathematiker NILS HENRIK ABEL (1802–1829) ist trotz seinen frühen Todes (an Tuberkulose) Initiator vieler Entwicklungen der Mathematik des neunzehnten Jahrhunderts; Begriffe wie abelsche Gruppen, abelsche Integrale, abelsche Funktionen, abelsche Varietäten, die auch in der heutigen Mathematik noch allgegenwärtig sind, verdeutlichen seinen Einfluß. Im August 2001 kündigte der norwegische Ministerpräsident zu seinen Ehren die Stiftung eines ABEL-Preises für Mathematik an, der ab 2003 in ähnlicher Ausstattung und ähnlicher Weise wie die Nobelpreise verliehen werden soll.

Der ABELsche Satz besagt selbstverständlich nicht, daß Gleichungen höheren als vierten Grades *unlösbar* seien; er sagt nur, daß es *im allgemeinen* nicht möglich ist, die Lösungen durch Wurzelausdrücke in den Koeffizienten darzustellen: Für eine allgemeine Lösungsformel muß man also außer Wurzeln und Grundrechenarten noch weitere Funktionen zulassen. Beispielsweise fanden sowohl HERMITE als auch KRONECKER 1858 Lösungsformeln für Gleichungen fünften Grades mit sogenannten elliptischen Modulfunktionen; 1870 löste JORDAN damit Gleichungen beliebigen Grades.

Für die Berechnung von Eigenvektoren sind schon die Lösungen einer kubischen Gleichung nach CARDANOs Formel im allgemeinen zu kompliziert, als daß man ohne Computer damit rechnen könnte; dasselbe gilt erst für höhere Grade. Insbesondere sind die Formeln in vielen Fällen numerisch instabil, wenn annähern gleich große Zahlen voneinander subtrahiert werden. Die Numerik geht daher aus gutem Grund anders vor, wenn sie Nullstellen von Polynomen berechnet.

d) Vielfachheiten von Eigenwerten

Ist x eine Nullstelle eines Polynoms $f(X)$, so kann $f(X)$ bekanntlich durch $(X-x)$ geteilt werden, und x heißt *r-fache Nullstelle* von $f(X)$, wenn $f(X)$ durch $(X-x)^r$ teilbar ist, nicht aber durch $(X-x)^{r+1}$.

Definition: Wir sagen, der Eigenwert λ von φ *bzw.* A habe die *algebraische Vielfachheit* r, wenn λ eine r-fache Nullstelle des charakteristischen Polynoms ist.

Im obigen Beispiel hatte also der Eigenwert Null die algebraische Vielfachheit zwei, die anderen beiden hatten algebraische Vielfachheit eins.