Falls auch Eigenwerte auftreten, deren geometrische Vielfachheit kleiner ist als die algebraische, können wir nicht so argumentieren; dann können wir A nicht auf eine Diagonalmatrix transformieren. Standardbeispiel ist die Matrix

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

deren einziger Eigenwert eins zwar die algebraische Vielfachheit zwei, aber nur die geometrische Vielfachheit eins hat. Trotzdem läßt sich auch hier e^{At} leicht berechnen: Wir schreiben

$$A = E + N \quad \text{mit} \quad E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Da N^2 sowie auch alle höheren Potenzen gleich der Nullmatrix sind, ist $e^{Nt} = E + Nt$, und da die Einheitsmatrix mit N (wie mit jeder anderen Matrix) kommutiert, ist

$$e^{At} = e^{Et+Nt} = e^{Et}e^{Nt} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix}.$$

Diese Vorgehensweise läßt sich verallgemeinern: Wie wir aus **[HM1]**, §6*j)* wissen, gibt es zu jeder Matrix $n \times n$-Matrix A eine Basis von $\mathbb{C}^n$, bezüglich derer A als Summe einer Diagonalmatrix D und einer damit vertauschbaren oberen Dreiecksmatrix N mit Nullen in der Hauptdiagonalen geschrieben werden kann, d.h.

$$B^{-1}AB = D + N \quad \text{mit} \quad DN = ND,$$

wobei die invertierbare Matrix B den zugehörigen Basiswechsel beschreibt. Damit ist

$$e^A = e^{\left(B(D+N)B^{-1}\right)t} = Be^{(D+N)t}B^{-1} = Be^{Dt} \cdot e^{Nt}B^{-1}.$$

e^{Dt} ist eine Diagonalmatrix ist mit Einträgen $e^{\lambda_i t}$ in der Hauptdiagonalen, und e^{Nt} ist, da spätestens N^n gleich der Nullmatrix ist, ein Polynom in Nt, dessen Einträge somit Polynome in t sind. Die Einträge von $e^{Dt}e^{Nt}$ sind daher Produkte aus den Exponentialfunktionen $e^{\lambda_i t}$ und Polynomen in t. Wir können das noch genauer sagen: Nach **[HM1]**, §6*j)* hat $D + N$ die spezielle Form

$$D + N = \begin{pmatrix} \boxed{M_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boxed{M_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boxed{M_s} \end{pmatrix}$$

hat mit oberen Dreiecksmatrizen M_i der Form

$$\begin{pmatrix} \lambda_i & * & \dots & * \\ 0 & \lambda_i & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i & 0 & \dots & 0 \\ 0 & \lambda_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_i \end{pmatrix} + \begin{pmatrix} 0 & * & \dots & * \\ 0 & 0 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

was wir kurz als $M_i = D_i + N_i$ schreiben wollen. Damit ist

$$e^{(D+N)t} = \begin{pmatrix} \boxed{e^{D_1t}e^{N_1t}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boxed{e^{D_2t}e^{N_2t}} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boxed{e^{D_st}e^{N_st}} \end{pmatrix}.$$

Ist M_i eine $r_i \times r_i$-Matrix, so ist die r_i-te Potenz von N_i gleich null; somit ist e^{N_i} ein Polynom vom Grad höchstens $r_i - 1$ in N_i, und die Einträge von e^{N_it} sind entsprechend Polynome vom Grad höchstens $r_i - 1$ in t.

Wir können noch mehr sagen: Nach **[HM1]**, §6*k)* können wir die N_i in der speziellen Form

$$N_i = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

bei denen direkt oberhalb der Hauptdiagonale lauter Einsen stehen, während alle anderen Einträge verschwinden. Bei der sukzessiven Potenzierung von N verschiebt sich, wie wir dort gesehen haben, einfach die Reihe von Einsen jeweils um eins weiter nach außen, und damit ist

$$e^{N_i t} = \sum_{\nu=0}^{r_i-1} \frac{1}{\nu!} N^\nu t^n u = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{r_i-2}}{(r_i-2)!} & \frac{t^{r_i-1}}{(r_i-1)!} \\ 0 & 1 & t & \dots & \frac{t^{r_i-3}}{(r_i-3)!} & \frac{t^{r_i-2}}{(r_i-2)!} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

und

$$\begin{aligned} e^{M_i t} &= e^{D_i t} e^{N_i t} = e^{\lambda_i t} e^{N_i t} \\ &= \begin{pmatrix} e^{\lambda_i t} & t e^{\lambda_i t} & \frac{t^2}{2!} e^{\lambda_i t} & \dots & \frac{t^{r_i-2}}{(r_i-2)!} e^{\lambda_i t} & \frac{t^{r_i-1}}{(r_i-1)!} e^{\lambda_i t} \\ 0 & e^{\lambda_i t} & t e^{\lambda_i t} & \dots & \frac{t^{r_i-3}}{(r_i-3)!} e^{\lambda_i t} & \frac{t^{r_i-2}}{(r_i-2)!} e^{\lambda_i t} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & e^{\lambda_i t} & t e^{\lambda_i t} \\ 0 & 0 & 0 & \dots & 0 & e^{\lambda_i t} \end{pmatrix} . \end{aligned}$$

Also treten Einträge mit Polynomen vom Grad $r_i - 1$ auch tatsächlich auf, und die Zahl r_i ist dadurch gekennzeichnet, daß wir für eine Basis, bezüglich derer die Matrix Dreiecksgestalt hat, Hauptvektoren bis zur r_i-ten Stufe berücksichtigen müssen.

Die Matrix e^{At}, die uns eigentlich interessiert, ist

$$e^{At} = B e^{Dt} e^{Nt} B^{-1}$$

mit einer konstanten (komplexen) Matrix B; ihre Einträge sind also Linearkombinationen von Funktionen der Form $t^j e^{\lambda_i t}$, wobei $0 \le j \le r_i - 1$ ist.

Betrachten wir als Beispiel das Differentialgleichungssystem

$$\begin{aligned} \dot{x}(t) &= 2x(t) - y(t) - z(t) \\ \dot{y}(t) &= x(t) + 5y(t) + 2z(t) \\ \dot{z}(t) &= -x(t) - 2y(t) + z(t) . \end{aligned}$$

Hier ist

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 5 & 2 \\ -1 & -2 & 1 \end{pmatrix} ,$$

und das charakteristische Polynom von A ist

$$\det(A - \lambda E) = -\lambda^3 + 8\lambda^2 - 21\lambda + 18 = -(\lambda - 2)(\lambda - 3)^2 .$$

Wir haben also den Eigenwert zwei mit algebraischer und somit auch geometrischer Vielfachheit eins und den Eigenwert drei mit algebraischer Vielfachheit zwei. In der Matrix

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

ist die mittlere Spalte gleich der Summe der beiden äußeren,

$$\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

erzeugt also den Eigenraum. In

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

stimmen die zweite und die dritte Spalte miteinander überein,

$$\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

ist also ein Eigenvektor und erzeugt auch den Eigenraum, denn da die erste Spalte kein Vielfaches der zweiten ist, hat die Matrix den Rang zwei. Die geometrische Vielfachheit des Eigenwerts drei ist also nur eins: Um zu einer Dreiecksmatrix zu kommen, müssen wir einen Hauptvektor zweiter Stufe berechnen. Aus

$$(A - 3E)^2 = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

sieht man, daß sich

$$\vec{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

als von $\vec{v}_2$ linear unabhängiger Kandidat anbietet. Da

$$A\vec{v}_3 = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix} = 3\vec{v}_3 - \vec{v}_2$$

ist, hat A bezüglich der Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ die Form

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 3 \end{pmatrix} ,$$

der erste „Kasten“ ist also einfach eine 1×1-Matrix und der zweite ist

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

Da das Quadrat des zweiten Summanden verschwindet, ist

$$e^{\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} t} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} t = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}$$

und

$$e^{\begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix} t} = e^{3t} \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{3t} & -te^{3t} \\ 0 & e^{3t} \end{pmatrix} .$$

Damit ist

$$e^{Mt} = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{3t} & -te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix} .$$

Um daraus e^{At} zu berechnen, müssen wir die Standardbasis des $\mathbb{R}^3$ durch die Hauptvektoren ausdrücken; man überzeugt sich leicht, daß

$$\vec{e}_1 = \vec{v}_1 + \vec{v}_2, \quad \vec{e}_2 = \vec{v}_1 + \vec{v}_2 - \vec{v}_3 \quad \text{und} \quad \vec{e}_3 = \vec{v}_1 - \vec{v}_3$$

ist. Bezüglich der Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ ist also

$$e^{At}\vec{e}_1 = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{3t} & -te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{2t} \\ e^{3t} \\ 0 \end{pmatrix} ,$$

was bezüglich der Standardbasis der Vektor

$$e^{2t}\vec{v}_1 + e^{3t}\vec{v}_2 = \begin{pmatrix} e^{2t} \\ -e^{2t} \\ e^{2t} \end{pmatrix} + \begin{pmatrix} 0 \\ e^{3t} \\ -e^{3t} \end{pmatrix} = \begin{pmatrix} e^{2t} \\ e^{3t} - e^{2t} \\ e^{2t} - e^{3t} \end{pmatrix} .$$

Also ist, bezüglich der Standardbasis ausgedrückt,

$$e^{At} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{2t} \\ e^{3t} - e^{2t} \\ e^{2t} - e^{3t} \end{pmatrix} .$$

Genauso überlegt man sich, daß

$$e^{At} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e^{2t}\vec{v}_1 + \left(e^{3t} + te^{3t}\right)\vec{v}_2 - e^{3t}\vec{v}_3 = \begin{pmatrix} e^{2t} - e^{3t} \\ 2e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - e^{3t} - te^{3t} \end{pmatrix}$$

und

$$e^{At} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{2t} - e^{3t} \\ e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - te^{3t} \end{pmatrix}$$

ist. Da in den Spalten einer Matrix die Bilder der Basisvektoren stehen, ist somit

$$e^{At} = \begin{pmatrix} e^{2t} & e^{2t} - e^{3t} & e^{2t} - e^{3t} \\ e^{3t} - e^{2t} & 2e^{3t} - e^{2t} + te^{3t} & e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - e^{3t} & e^{2t} - e^{3t} - te^{3t} & e^{2t} - te^{3t} \end{pmatrix} .$$

Die Lösung, die den Anfangsbedingungen

$$x(0) = x_0, \quad y(0) = y_0 \quad \text{und} \quad z(0) = z_0$$

genügt, ist also

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} e^{2t} & e^{2t} - e^{3t} & e^{2t} - e^{3t} \\ e^{3t} - e^{2t} & 2e^{3t} - e^{2t} + te^{3t} & e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - e^{3t} & e^{2t} - e^{3t} - te^{3t} & e^{2t} - te^{3t} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

$$= \begin{pmatrix} (x_0 + y_0 + z_0)e^{2t} - (y_0 + z_0)e^{3t} \\ \left(x_0 + (t+2)y_0 + (t+1)z_0\right)e^{3t} - (x_0 + y_0 + z_0)e^{2t} \\ (x_0 + y_0 + z_0)e^{2t} - \left(x_0 + (t+1)y_0 + tz_0\right)e^{3t} \end{pmatrix} .$$

g) Lineare homogene Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten

Eine der wichtigsten Anwendungen der Theorie aus dem letzten Abschnitt sind lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten. Wir wollen in diesem Abschnitt sehen, daß in diesem Spezialfall schon ein sehr viel einfacherer Ansatz zum Ziel führt. Wir betrachten die Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0 \quad \text{mit} \quad a_i \in \mathbb{C}\,.$$

Um die Theorie aus dem vorigen Abschnitt anwenden zu können, schreiben wir sie in der üblichen Weise um als ein System

$$\begin{aligned} \dot{y}_0(t) &= y_1(t) \\ \dot{y}_1(t) &= y_2(t) \\ &\vdots \\ \dot{y}_{n-2}(t) &= y_{n-1}(t) \\ \dot{y}_{n-1}(t) &= -a_{n-1}y_{n-1}(t) - \cdots - a_1\dot{y}_1(t) - a_0 y_0(t) \end{aligned}$$

oder kurz $\dot{\vec{y}}(t) = A\vec{y}(t)$ mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

und

$$\vec{y}(t) = \begin{pmatrix} y_0(t) \\ \vdots \\ y_{n-1}(t) \end{pmatrix}.$$

Die Lösungen $\vec{y}(t)$ dieses Systems definieren Lösungen $y(t) = y_0(t)$ der obigen Gleichung, und umgekehrt ist auch für jede solche Lösung $y(t)$ der Vektor

$$\vec{y}(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \ddot{y}(t) \\ \vdots \\ y^{(n-2)}(t) \\ y^{(n-1)}(t) \end{pmatrix}$$

eine Lösung des Systems $\dot{\vec{y}}(t) = A\vec{y}(t)$.

Aus dem vorigen Abschnitt wissen wir, daß die Lösungen dieses Systems einen n-dimensionalen Vektorraum V bilden, aufgespannt von den Funktionen $t^j e^{\lambda_i t}$, wobei $\lambda_1, \dots, \lambda_r$ die verschiedenen Eigenwerte von A sind und $j \geq 0$ kleiner ist als die algebraische Vielfachheit des Eigenwerts λ_i. Insbesondere tritt der Vorfaktor t^j bei einfachen Eigenwerten erst gar nicht auf.

Außerdem wissen wir, daß jedes Anfangswertproblem eindeutig lösbar ist: Bei Anfangsbedingungen

$$y_0(t_0) = c_0, \quad \dots, \quad y_{n-1}(t_0) = c_{n-1}$$

ist

$$\vec{y}(t) = e^{A(t-t_0)}$$

die Lösung. Vornehm ausgedrückt können wir auch sagen, daß für jedes $t_0 \in \mathbb{R}$ die Abbildung

$$V \to \mathbb{C}^n; \quad \vec{y}(t) \mapsto \vec{y}(t_0)$$

ein Isomorphismus ist.

Damit bilden auch die Lösungsfunktionen $y(t)$ der obigen Gleichung einen n-dimensionalen Vektorraum W und sind Linearkombinationen dieser Funktionen; insbesondere ist jedes Anfangswertproblem

$$y(t_0) = c_0, \quad \dot{y}(t_0) = c_1, \quad \dots, \quad y^{(n-1)}(t_0) = c_{n-1}$$

eindeutig lösbar und die Abbildung

$$W \to \mathbb{C}^n; y(t) \mapsto \big(y(t), \dot{y}(t), \dots, y^{(n-1)}(t)\big)$$

ist ein Isomorphismus.

Da die Summe der algebraischen Vielfachheiten der (komplexen) Eigenwerte einer komplexen $n \times n$-Matrix gleich n ist, werden tatsächlich alle dieser Funktionen gebraucht. Damit kennen wir eine Basis des Lösungsraums: Sind $\lambda_1, \dots, \lambda_r$ die Eigenwerte von A ist α_i die algebraische Vielfachheit von λ_i, so ist sie

$$\left\{e^{\lambda_1 t}, te^{\lambda_1 t}, \dots, t^{\alpha_1 - 1}e^{\lambda_1 t}, \quad \dots \quad e^{\lambda_r t}, te^{\lambda_r t}, \dots, t^{\alpha_r - 1}e^{\lambda_r t}\right\}.$$

Die Zahlen λ_i lassen sich leicht direkt aus der Gleichung ablesen: Die Funktion

$$y(t) = e^{\lambda t} \quad \text{mit} \quad y^{(i)}(t) = \lambda^i e^{\lambda t}$$

ist genau dann eine Lösung der Differentialgleichung, wenn

$$\begin{aligned} &\lambda^n e^{\lambda t} + a_{n-1}\lambda^{n-1}e^{\lambda t} + \cdots + a_1\lambda e^{\lambda t} + a_0 e^{\lambda t} \\ &= (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)e^{\lambda t} \end{aligned}$$

verschwindet; da $e^{\lambda t}$ nirgends verschwindet, ist dies genau dann der Fall, wenn

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0 \qquad (*)$$

ist.

Definition: Die Gleichung $(*)$ heißt *charakteristische Gleichung* der Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0\,.$$

Die Eigenwerte λ_i von A sind also genau die Nullstellen der charakteristischen Gleichung. Da diese denselben Grad hat wie das charakteristische Polynom von A liegt die Vermutung nahe, daß sie (eventuell bis auf eine Konstante) damit übereinstimmt, und das ist auch tatsächlich der Fall:

Lemma: Die charakteristische Gleichung ist das mit $(-1)^n$ multiplizierte charakteristische Polynom von A.

Beweis: Für $n = 1$ ist die Differentialgleichung $\dot{y}(t) = ay(t)$ identisch mit dem zugehörigen System; die charakteristische Gleichung ist $\lambda - a$, und die „Matrix“ $A = (a)$ hat das charakteristische Polynom $a - \lambda$. Für $n = 1$ stimmt die Behauptung also.

Für $n > 1$ entwickeln wir das charakteristische Polynom

$$\det(A - \lambda E) = \begin{vmatrix} -\lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\lambda & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} - \lambda \end{vmatrix}$$

nach der ersten Spalte und erhalten

$$\begin{aligned} \det(A - \lambda E) = -\lambda &\begin{vmatrix} -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -\lambda & 1 \\ -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} - \lambda \end{vmatrix} \\ &+ (-1)^{n-1}(-a_0) \begin{vmatrix} 1 & 0 & \dots & 0 & 0 \\ -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -\lambda & 1 \end{vmatrix}. \end{aligned}$$

Die erste Determinate auf der rechten Seite ist von derselben Form wie das betrachtete charakteristische Polynom, hat aber nur die Größe $(n-1) \times (n-1)$. Nach Induktionsannahme wissen wir deshalb, daß sie gleich

$$(-1)^{n-1}\left(a_n\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \cdots + a_2\lambda + a_1\right)$$

ist. Die zweite Determinante läßt sich leicht direkt berechnen: Entweder durch Entwicklung nach der ersten Zeile und Induktion, oder aber, indem wir uns daran erinnern, daß nach **[HM1]**, §6*g),* die Determinante einer Dreiecksmatrix gleich dem Produkt ihrer Diagonalelemente ist, sehen wir, daß diese Determinante eins ist. Damit folgt

$$\begin{aligned}\det(A-\lambda E) &= (-\lambda)(-1)^{n-1}\big(a_n\lambda^{n-1}+a_{n-1}\lambda^{n-2}+\cdots+a_2\lambda+a_1\big)\\ &\quad +(-1)^{n-1}(-a_0)\\ &= (-1)^n\big(a_n\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_2\lambda^2+a_1\lambda+a_0\big)\,,\end{aligned}$$

wie behauptet. ■

Insbesondere haben also die charakteristische Gleichung einer Differentialgleichung und das charakteristische Polynom der zugehörigen Matrix nicht nur dieselben Nullstellen, sondern auch die Vielfachheiten dieser Nullstellen sind gleich.

Wir können den bisherigen Inhalt dieses Abschnitts zusammenfassen im

Satz: Die charakteristische Gleichung

$$\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0=0$$

der Differentialgleichung

$$y^{(n)}(t)+a_{n-1}y^{(n-1)}(t)+\cdots+a_1\dot y(t)+a_0y(t)=0$$

habe die Nullstellen $\lambda_1,\dots,\lambda_r$; die Vielfachheit der Nullstelle λ_i sei α_i. Dann bilden die Lösungen der Differentialgleichung einen n-dimensionalen Vektorraum V mit Basis

$$\big\{e^{\lambda_1 t},te^{\lambda_1 t},\dots,t^{\alpha_1-1}e^{\lambda_1 t},\quad\dots\quad e^{\lambda_r t},te^{\lambda_r t},\dots,t^{\alpha_r-1}e^{\lambda_r t}\big\}\,.$$

Für jedes $t_0\in\mathbb{R}$ ist die lineare Abbildung

$$V\to\mathbb{C}^n;\quad y\mapsto\big(y(t_0),\dot y(t_0),\ddot y(t_0),\dots,y^{(n-1)}(t_0)\big)$$

ein Isomorphismus; insbesondere ist jedes Anfangswertproblem

$$y(t_0)=c_0,\quad \dot y(t_0)=c_1,\quad\dots\quad y^{(n-1)}(t_0)=c_{n-1}$$

eindeutig lösbar. ■

Sofern wir uns für komplexwertige Lösungen interessieren, liefert uns dieser Satz alles was wir brauchen. Bei vielen Anwendungen entstehen aber Gleichungen mit reellen Koeffizienten, und von der Natur des Problems her interessieren nur reelle Lösungen. Mit unserem bisherigen Ansatz kommen wir auch bei diesen Problemen nicht um komplexe Lösungen herum, beispielsweise hat die wohlbekannte Differentialgleichung

$$\ddot y(t)+y(t)=0$$

die charakteristische Gleichung

$$\lambda^2+1=0$$

mit den beiden rein imaginären Lösungen $\lambda=\pm i$, die oben angegebene Basis des Lösungsraums ist also

$$\big\{e^{it},\ e^{-it}\big\}\,.$$

Falls wir noch nie etwas vom obigen Satz gehört hätten, würden wir stattdessen sagen, daß die Lösungen dieser Differentialgleichung genau die Linearkombinationen von Sinus und Cosinus sind, Basis des Lösungsraums ist also

$$\big\{\sin t,\ \cos t\big\}\,.$$

Über $\mathbb{C}$ sind beide Aussagen äquivalent, denn auf Grund der EULERschen Formeln

$$\cos t=\frac{1}{2}\big(e^{it}+e^{-it}\big)\quad\text{und}\quad \sin t=\frac{1}{2i}\big(e^{it}ie^{-it}\big)$$

bzw.

$$e^{it}=\cos t+i\sin t\quad\text{und}\quad e^{-it}=\cos t-i\sin t$$

erzeugen beide Basen denselben $\mathbb{C}$-Vektorraum. Wenn wir an reellen Lösungen interessiert sind, ist aber die zweite Basis erheblich nützlicher, denn sie erzeugt auch den $\mathbb{R}$-Vektorraum der reellen Lösungen.

Entsprechend hätten wie auch für allgemeinere Differentialgleichungen mit reellen Koeffizienten gerne eine $\mathbb{C}$-Basis des (komplexen) Lösungsraums, die gleichzeitig $\mathbb{R}$-Basis des reellen Lösungsraums ist. Im Rest dieses Abschnitts wollen wir uns überlegen, wie wir diese Basis konstruieren können.

Wir gehen also aus von einer Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0y(t) = 0 \quad \text{mit} \quad a_i \in \mathbb{R}$$

und betrachten auch hier wieder zunächst die charakteristische Gleichung

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0\,.$$

Diese *kann* reelle Lösungen haben; diese bezeichnen wir mit λ_1 bis λ_p, und falls es keine gibt, setzen wir $p = 0$.

Falls p gleich der Gesamtzahl r der Nullstellen der charakteristischen Gleichung ist, sind wir fertig: Die Funktionen $t^j e^{\lambda_i t}$ sind allesamt reell, und die, bei denen j kleiner als die Vielfachheit der Nullstelle λ_i ist, spannen sowohl den komplexen als auch den reellen Lösungsraum auf.

Andernfalls gibt es auch noch nichtreelle Nullstellen. Für jede solche Nullstelle λ ist auch die konjugiert komplexe Zahl $\overline{\lambda}$ eine Nullstelle, denn da die Koeffizienten a_i reell sind, ist $\overline{a_i} = a_i$ und damit

$$\overline{\lambda}^n + a_{n-1}\overline{\lambda}^{n-1} + \ldots + a_1\overline{\lambda} + a_0 = \overline{\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0} = 0\,.$$

Nichtreelle Nullstellen treten also immer als Paare konjugiert komplexer Zahlen auf; außerdem haben λ und $\overline{\lambda}$ dieselbe Vielfachheit, denn eine Nullstelle ist genau dann eine i-fache Nullstelle eines Polynoms P, wenn sie Nullstelle von P und allen seinen Ableitungen bis zur $(i-1)$-ten ist. Da auch alle diese Ableitungen Polynome mit reellen Koeffizienten sind, zeigt die gleiche Rechnung wie oben, daß $P^{(j)}(\lambda)$ genau dann verschwindet, wenn auch $P^{(j)}(\overline{\lambda})$ verschwindet.

Wir fassen daher die nichtreellen Nullstellen zu Paaren $\left(\lambda_{p+j}, \overline{\lambda_{p+j}}\right)$ zusammen, wobei j von 1 bis zur Anzahl q dieser Paare läuft; die Gesamtzahl der Nullstellen ist also $r = p + 2q$.

Die Vielfachheit der Nullstelle λ_k sei weiterhin mit α_k bezeichnet; für $k > p$ ist das gleichzeitig auch die Vielfachheit der Nullstelle $\overline{\lambda}_k$.

Für $k > p$ wenden wir, genau wie im obigen Beispiel, die EULERschen Formeln an: Für

$$\lambda_k = \mu_k + i\nu_k \quad \text{mit} \quad \mu_k, \nu_k \in \mathbb{R}$$

ist

$$e^{\lambda_k t} = e^{\mu_k t}\left(\cos\nu_k t + i\sin\nu_k t\right) \quad \text{und} \quad e^{\overline{\lambda_k} t} = e^{\mu_k t}\left(\cos\nu_k t - i\sin\nu_k t\right)\,,$$

genauso ist für jedes j

$$t^j e^{\lambda_k t} = t^j e^{\mu_k t}\left(\cos\nu_k t + i\sin\nu_k t\right)$$

und

$$t^j e^{\overline{\lambda_k} t} = t^j e^{\mu_k t}\left(\cos\nu_k t - i\sin\nu_k t\right)\,.$$

Also spannt

$$\left\{t^j e^{\lambda_k t},\; t^j e^{\overline{\lambda_k} t}\right\}$$

denselben $\mathbb{C}$-Vektorraum auf wie

$$\left\{t^j e^{\mu_k t}\cos\nu_k t,\; t^j e^{\mu_k t}\sin\nu_k t\right\}\,,$$

und letztere Basis spannt gleichzeitig den $\mathbb{R}$-Vektorraum aller reeller Funktionen aus diesem $\mathbb{C}$-Vektorraum auf.

Damit können wir auch eine Basis des reellen Lösungsraums angeben:

Satz: Die charakteristische Gleichung

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$$

der Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0y(t) = 0$$

habe die reellen Nullstellen $\lambda_1, \ldots, \lambda_p$ sowie die Paare konjugiert komplexer Nullstellen $\left(\lambda_{p+1}, \overline{\lambda_{p+1}}\right), \ldots, \left(\lambda_{p+q}, \overline{\lambda_{p+q}}\right)$ mit $\lambda_k = \mu_k + i\nu_k$; die Vielfachheit der Nullstelle λ_i sei α_i. Dann bilden die Lösungen der Differentialgleichung einen n-dimensionalen Vektorraum V, der aufgespannt wird von den Funktionen

$$t^j e^{\lambda_i t} \quad \text{für } i = 1, \ldots, p$$

und

$$t^j e^{\mu_i t}\cos\nu_i t, \quad t^j e^{\mu_i t}\sin\nu_i t \quad \text{für } i = p+1, \ldots, p+q\,,$$

wobei j jeweils von null bis $\alpha_i - 1$ läuft. Für jedes $t_0 \in \mathbb{R}$ ist die lineare Abbildung

$$V \to \mathbb{R}^n; \quad y \mapsto \big(y(t_0), \dot{y}(t_0), \ddot{y}(t_0), \dots, y^{(n-1)}(t_0)\big)$$

ein Isomorphismus; insbesondere ist jedes Anfangswertproblem

$$y(t_0) = c_0, \quad \dot{y}(t_0) = c_1, \quad \dots \quad y^{(n-1)}(t_0)$$

eindeutig lösbar. ■

h) Inhomogene Differentialgleichungen

Neben des bislang behandelten *homogenen* Differentialgleichungen interessieren uns auch inhomogene Systeme, z.B. die Differentialgleichung einer erzwungenen Schwingung. Nach dem Lemma aus Abschnitt *d)* genügt es für die Lösung einer solchen Gleichung, *eine* einzige Lösung zu finden; alle anderen sind da darstellbar als Summe aus dieser speziellen Lösung und einer Lösung der zugehörigen inhomogenen Gleichung.

Wir kennen bereits eine Methode, uns eine solche spezielle Lösung zu verschaffen: Sobald wir uns Anfangswerte vorgeben, können wir die entsprechende Lösung des inhomogenen Problems mittels LAPLACE-Transformation finden, sofern wir die LAPLACE-Transformierte des inhomogenen Anteils kennen und aus der LAPLACE-Transformierten der Lösungsfunktion diese Funktion selbst bestimmen können. In vielen praktisch relevanten Fällen wird dies mit Hilfe von Tabellen möglich sein, sofern man nur die grundlegenden Regeln für den Umgang mit LAPLACE-Transformationen kennt.

Eine zweite Methode ist das *Erraten* einer Lösung. In der Praxis betrachtet man Differentialgleichungen meist, um das künftige Verhalten eines Systems vorauszusagen, und oft wird man (keinesfalls immer richtige!) ungefähre Erwartungen haben, wie dieses Verhalten aussehen sollte. Falls man diese als Ansatz mit unbestimmten Parametern in die Differentialgleichung einsetzt und die Erwartungen richtig war, kann man die Parameter bestimmen und hat eine Lösung gefunden; andernfalls muß man sich etwas neues überlegen.

Als Beispiel betrachten wir nochmals die Differentialgleichung

$$\ddot{y}(t) + \rho\dot{y}(t) + \sigma y(t) = f(t)\,.$$

Wir kamen auf diese Differentialgleichung bei der betrachtung eines elektrischen Schwingkreises, an den eine externe Spannung proportional $f(t)$ angelegt war, und für vernünftiges $f(t)$ sollte man erwarten, daß es Lösungen gibt, bei denen sich $y(t)$ völlig von $f(t)$ bestimmen läßt und daher von ähnlicher Gestalt ist.

Im Falle einer Gleichstromquelle $f(t) = c$ vermuten wir, daß es vielleicht eine Lösung gibt, bei der nur ein Gleichstrom fließt. Der Ansatz $y(t) = a$ führt auf

$$\sigma a = c \quad \text{oder} \quad a = \frac{c}{\sigma}\,,$$

falls $\sigma \neq 0$ ist. Für $\sigma = 0$ legt die Mathematik einen Ansatz der Form $y(t) = bt + a$ nahe, und für $\rho \neq 0$ führt das zu

$$\rho b = c \quad \text{oder} \quad b = \frac{c}{\rho}\,.$$

Entsprechend können wir auch bei einer angelegten Wechselspannung argumentieren: Der Schwingkreis wird sicherlich die Amplitude und die Phase verändern, aber die permanent angelegte Wechselspannung sollte doch dem Schwingkreis ihre Frequenz aufzwingen, so daß es zumindest eine Lösung geben sollte, die eine reine Schwingung mit dieser Frequenz ist. Für

$$\ddot{y}(t) + \rho\dot{y}(t) + \sigma y(t) = A\cos\omega_0 t$$

machen wir also den Ansatz

$$y(t) = a\cos\omega_0 t + b\sin\omega_0 t$$

und berechnen die Ableitungen:

$$\dot{y}(t) = -a\omega_0\sin\omega_0 t + b\omega_0\cos\omega_0 t$$

und

$$\ddot{y}(t) = -a\omega_0^2\cos\omega_0 t - b\omega_0^2\sin\omega_0 t\,.$$

Einsetzen in die linke Seite der Differentialgleichung liefert

$$\begin{aligned}&\ddot{y}(t) + \rho\dot{y}(t) + \sigma y(t)\\ &= (-a\omega_0^2 + \rho b\omega_0 + \sigma a)\cos\omega_0 t + (-b\omega_0^2 - \rho a\omega_0 + \sigma b)\sin\omega_0 t\,,\end{aligned}$$

und das ist gleich $A\cos\omega_0 t$, falls

$$(\sigma-\omega_0^2)a+\rho\omega_0 b=A \quad\text{und}\quad (\sigma-\omega_0^2)b-\rho\omega_0 a=0$$

ist.

Für $\sigma=\omega_0^2$ ist dies problemlos zu lösen: Dann ist einfach

$$a=0 \quad\text{und}\quad b=\frac{A}{\rho\omega_0}\,.$$

Für $\sigma\neq\omega_0^2$ können wir die zweite Gleichung durch $(\sigma-\omega_0^2)$ dividieren und erhalten

$$b=\frac{\rho\omega_0}{\sigma-\omega_0^2}a\,.$$

Dies können wir in die erste Gleichung einsetzen:

$$(\sigma-\omega_0^2)a+\rho\omega_0\frac{\rho\omega_0}{\sigma-\omega_0^2}a=\frac{(\sigma-\omega_0^2)^2+\rho^2\omega_0^2}{\sigma-\omega_0^2}a=A$$

oder

$$a=\frac{A(\sigma-\omega_0^2)}{(\sigma-\omega_0^2)^2+\rho^2\omega_0^2} \quad\text{und}\quad b=\frac{A\rho\omega_0}{(\sigma-\omega_0^2)^2+\rho^2\omega_0^2}\,.$$

In beiden Fällen gibt es also in der Tat eine Lösung der postulierten Form, die wir hiermit explizit bestimmt haben.

Im Prinzip haben wir diese Lösung bereits in Abschnitt *c)* berechnet, mit dem Ansatz hier ist die Situation nun aber sehr viel übersichtlicher: Im häufigsten Falle, in dem die homogene Gleichung eine gedämpfte Schwingung beschreibt etwa wissen wir nun, daß es eine feste reine Schwingung mit der erregenden Frequenz gibt, gegen die *alle* Lösungen der Differentialgleichung langfristig konvergieren. Für solche allgemeinen Aussagen ist also dieser Ansatz besser geeignet; falls wir dagegen ein konkretes Anfangswertproblem lösen wollen, liefert die LAPLACE-Transformation direkt die Lösung, während wir bei diesem Ansatz noch ein lineares Gleichungssystem lösen müssen, um die beiden freien Parameter der allgemeinen Lösung an vorgegebene Anfangswerte anzupassen.

Falls man einigermaßen konkrete Erwartungen über das Verhalten zumindest einer Lösungsfunktion hat, ist das Erraten einer speziellen

Lösung also oft erfolgversprechend. Das im Abschnitt *c)* behandelte Beispiel der Resonanzkatastrophe zeigt aber auch, daß es Fälle gibt, wo man nur mit ziemlicher Erfahrung eine spezielle Lösung erraten kann: Die Gleichung

$$\ddot y(t)+\omega_0^2 y(t)=\cos\omega_0 t$$

hat schließlich *keine* Lösung, die eine reine Schwingung der Frequenz ω_0 ist.

Zum Glück gibt es außer LAPLACE-Transformation und Erraten noch ein drittes, rein formales Verfahren, das ebenfalls häufig zum Ziel führt, die *Variation der Konstanten.* Die Grundidee ist einfach: Die allgemeine Lösung der homogenen Gleichung hängt von Parametern c_i ab; man ersetze diese Parameter durch Funktionen $c_i(t)$ und setze dies ein in die inhomogene Gleichung. Dies stellt Bedingungen an die $c_i(t)$; falls man die entstehende Differentialgleichung lösen kann, hat man eine Lösung der inhomogenen Gleichung gefunden.

In Abschnitt *d)* wurde bereits erwähnt, daß diese Methode zur dort einfach vom Himmel gefallenen Lösung einer inhomogenen linearen Differentialgleichung erster Ordnung

$$\dot y(t)=a(t)y(t)+b(t)$$

führt: Die zugehörige *homogene Gleichung*

$$\dot y(t)=a(t)y(t)$$

hat, wie wir dort gesehen haben, die allgemeine Lösung

$$y(t)=c\,e^{\int_{t_0}^{t}a(\tau)\,d\tau}$$

mit einem $c\in\mathbb{R}$.

Zur Lösung der inhomogenen Gleichung variieren wir die Konstante c, d.h. wir versuchen unser Glück mit dem Ansatz

$$y(t)=c(t)\,e^{\int_{t_0}^{t}a(\tau)\,d\tau}\,.$$

Für eine solche Funktion ist

$$\dot y(t) = \dot c(t)\, e^{\int_{t_0}^{t} a(\tau)\,d\tau} + c(t)\,a(t)e^{\int_{t_0}^{t} a(\tau)\,d\tau} = \dot c(t)\, e^{\int_{t_0}^{t} a(\tau)\,d\tau} + a(t)y(t)\,.$$

$y(t)$ ist also genau dann eine Lösung der inhomogenen Differentialgleichung

$$\dot y(t) = a(t)y(t) + b(t)\,,$$

wenn

$$\dot c(t) = b(t)\, e^{-\int_{t_0}^{t} a(\tau)\,d\tau}$$

oder

$$c(t) = \int_{t_0}^{t} b(s)\, e^{-\int_{t_0}^{s} a(\tau)d\tau}\, ds$$

ist. Damit haben wir eine spezielle Lösung

$$y(t) = \int_{t_0}^{t} b(s)\, e^{-\int_{t_0}^{s} a(\tau)d\tau}\, ds \cdot e^{\int_{t_0}^{t} a(\tau)\,d\tau}$$

gefunden, und die allgemeine Lösung ist

$$y(t) = \int_{t_0}^{t} b(s)\, e^{-\int_{t_0}^{s} a(\tau)d\tau}\, ds \cdot e^{\int_{t_0}^{t} a(\tau)\,d\tau} + C e^{\int_{t_0}^{t} a(\tau)\,d\tau}$$

mit einer beliebigen Konstanten C.

In diesem einfachen Fall funktioniert der Ansatz also recht gut (wenn es auch nicht immer möglich sein wird, die hier stehenden Integrale in geschlossener Form auszudrücken), und er liefert uns genau die Lösung, von deren Richtigkeit wir uns schon im Abschnitt *d)* überzeugt hatten.

Betrachten wir als nächstes Beispiel die Differentialgleichung

$$\ddot y(t) - y(t) = 6\sinh t\,.$$

Erraten einer speziellen Lösung nach dem Motto, daß diese so ähnlich aussehen sollte wie die rechte Seite führt hier auf den Ansatz

$$y(t) = a\cosh t + b\sinh t\,,$$

und der führt definitiv nicht zum Erfolg, da für jede solche Funktion die linke Seite gleich null ist.

Probieren wir es also mit Variation der Konstanten: Der obige Ansatz ist gleichzeitig die allgemeine Lösung der homogenen Gleichung (und konnte genau deshalb auch unmöglich zu einer Lösung der inhomogenen Gleichung führen); durch Variation der Konstanten wird diese Lösung zu

$$y(t) = a(t)\cosh t + b(t)\sinh t\,.$$

Differenzieren führt auf

$$\dot y(t) = \dot a(t)\cosh t + a(t)\sinh t + \dot b(t)\sinh t + b(t)\cosh t$$

und

$$\begin{aligned}\ddot y(t) &= \ddot a(t)\cosh t + 2\dot a(t)\sinh t + a(t)\cosh t\\ &\quad + \ddot b(t)\sinh t + 2\dot b(t)\cosh t + b(t)\sinh t\,;\end{aligned}$$

einsetzen in die Differentialgleichung ergibt

$$\big(\ddot a(t) + 2\dot b(t)\big)\cosh t + \big(\ddot b(t) + 2\dot a(t)\big)\sinh t = 6\sinh t\,,$$

und damit können wir offensichtlich nicht sonderlich viel anfangen.

Trotzdem führt die Methode der Variation der Konstanten auch hier zum Erfolg, wir müssen nur die Differentialgleichung umschreiben in ein System von Differentialgleichungen erster Ordnung, also in

$$\begin{aligned}\dot y(t) &= z(t)\\ \dot z(t) &= y(t) + 6\sinh t\,.\end{aligned}$$

Dies ist ein System der allgemeinen Form

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

mit einer Matrix $A \in \mathbb{C}^{n\times n}$ und einem Vektor $\vec{b}(t)$ von Funktionen $b_i(t)$; bevor wir das spezielle Problem lösen, wollen uns überlegen, daß die Methode der Variation der Konstanten sogar das allgemeine Problem lösen kann.

Die Lösungsmenge des homogenen Systems $\dot{\vec{y}}(t) = A\vec{y}(t)$ besteht aus den Funktionen $e^{At}\vec{u}$ mit einem beliebigen Vektor $\vec{u} \in \mathbb{C}^n$; Variation der Konstanten bedeutet, daß dieser konstante Vektor durch einen Vektor $\vec{u}(t)$ von Funktionen ersetzt wird, d.h. wir machen den Ansatz

$$\vec{y}(t) = e^{At}\vec{u}(t)\,.$$

Da die LEIBNIZsche Produktregel auch für matrixwertige Funktionen gilt, ist

$$\dot{\vec{y}}(t) = Ae^{At}\vec{u}(t) + e^{At}\dot{\vec{u}}(t) = A\vec{y}(t) + e^{At}\dot{\vec{u}}(t)\,;$$

der Ansatz führt also genau dann auf eine Lösung der Differentialgleichung

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

wenn

$$\vec{b}(t) = e^{At}\dot{\vec{u}}(t) \quad \text{oder} \quad \dot{\vec{u}} = e^{-At}\vec{b}(t)$$

ist. Damit können wir die Komponenten von $\vec{u}(t)$ als Stammfunktionen der Komponenten des ganz rechts stehenden Vektors von Funktionen berechnen. Zur Vereinfachung der Schreibweise vereinbaren wir, daß für eine vektorwertige Funktion $\vec{v}(t)$ gelten soll

$$\int \vec{v}(t)\,dt = \begin{pmatrix} \int v_1(t)\,dt \\ \vdots \\ \int v_n(t)\,dt \end{pmatrix} \quad \text{falls} \quad \vec{v}(t) = \begin{pmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{pmatrix}.$$

Dann ist

$$\vec{u}(t) = \int e^{-At}\vec{b}(t)\,dt\,,$$

und wir haben eine spezielle Lösung gefunden. Die allgemeine Lösung des inhomogenen Systems ist daher

$$\vec{y}(t) = e^{At}\left(\int e^{-At}\vec{b}(t)\,dt\right) + e^{At}\vec{v}_0$$

mit einem beliebigen konstanten Vektor $\vec{v}_0$.

Mit dieser Methode läßt sich nun auch die oben betrachtete Differentialgleichung

$$\ddot{y}(t) - y(t) = 6\sinh t$$

lösen: Das dazu äquivalente System

$$\begin{aligned} \dot{y}(t) &= z(t) \\ \dot{z}(t) &= y(t) + 6\sinh t \end{aligned}$$

ist von der Form $\dot{\vec{y}}(t) = A\vec{y}(t) + \vec{b}(t)$ mit

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad \vec{b}(t) = \begin{pmatrix} 0 \\ 6\sinh t \end{pmatrix};$$

wir benötigen zunächst

$$\vec{u}(t) = \int e^{-At}\vec{b}(t)\,dt\,.$$

Wie wir bereits zu Beginn von Abschnitt *e)* gesehen haben, ist A^2 die Einheitsmatrix und somit

$$e^A = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix};$$

im wesentlichen die gleiche Rechnung zeigt, daß

$$e^{At} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$$

und

$$e^{-At} = \begin{pmatrix} \cosh(-t) & \sinh(-t) \\ \sinh(-t) & \cosh(-t) \end{pmatrix} = \begin{pmatrix} \cosh t & -\sinh t \\ -\sinh t & \cosh t \end{pmatrix}$$

ist. Wir brauchen die Funktion

$$e^{-At}\vec{b}(t) = \begin{pmatrix} \cosh t & -\sinh t \\ -\sinh t & \cosh t \end{pmatrix}\begin{pmatrix} 0 \\ 6\sinh t \end{pmatrix} = \begin{pmatrix} -6\sinh^2 t \\ 6\sinh t\cosh t \end{pmatrix}$$

und müssen diese integrieren.

Beginnen wir mit dem ersten Eintrag:

$$\begin{aligned}\int -6\sinh^2 t\,dt = -6\int\left(\frac{e^t-e^{-t}}{2}\right)^2 dt &= -\frac{3}{2}\int(e^{2t}+e^{-2t}-2)\,dt\\ &= -\frac{3}{2}\left(\frac{e^{2t}}{2}-\frac{e^{-2t}}{2}-2t\right)+C\\ &= -\frac{3}{4}(e^t-e^{-t})(e^t+e^{-t})+3t+C'\\ &= -3\sinh t\cosh t+3t+C'\,.\end{aligned}$$

Genauso finden wir auch eine Stammfunktion des zweiten Eintrags:

$$\begin{aligned}\int 6\sinh t\cosh t\,dt &= \frac{3}{2}\int(e^{2t}-e^{-2t})\,dt\\ &= \frac{3}{4}\left(e^{2t}+e^{-2t}\right)+C\\ &= \frac{3}{4}\left(e^{2t}+2+e^{-2t}\right)+C-\frac{3}{2}\\ &= 3\left(\frac{e^t+e^{-t}}{2}\right)^2+C' = 3\cosh^2 t+C'\,.\end{aligned}$$

Da wir nur eine spezielle Lösung brauchen, können wir die beiden Integrationskonstanten unbesorgt auf null setzen; jede andere Wahl würde nur bedeuten, daß wir eine Lösung der homogenen Gleichung dazuaddieren. Also arbeiten wir mit

$$\vec{u}(t) = \begin{pmatrix}-3\sinh t\cosh t+3t\\ 3\cosh^2 t\end{pmatrix}$$

und die gesuchte spezielle Lösung ist

$$\begin{aligned}e^{At}\vec{u}(t) &= \begin{pmatrix}\cosh t & \sinh t\\ \sinh t & \cosh t\end{pmatrix}\begin{pmatrix}-3\sinh t\cosh t+3t\\ 3\cosh^2 t\end{pmatrix}\\ &= \begin{pmatrix}-3\sinh t\cosh^2 t+3t\cosh t+3\cosh^2 t\sinh t\\ -3\sinh^2 t\cosh t+3t\cosh t+3\cosh^3 t\end{pmatrix}\,.\end{aligned}$$

In der ersten Zeile dieses Ergebnisses heben sich der erste und der dritte Term gegenseitig weg; in der zweiten ist

$$3\cosh^3 t-3\sinh^2 t\cosh t = 3\cosh t(\cosh^2 t-\sinh^2 t) = 3\cosh t\,.$$

Als Endergebnis erhalten wir somit die spezielle Lösung

$$\begin{pmatrix}y(t)\\ z(t)\end{pmatrix} = e^{At}\vec{u}(t) = \begin{pmatrix}3t\cosh t\\ 3t\sinh t+3\cosh t\end{pmatrix}\,.$$

Insbesondere ist

$$y(t) = 3t\cosh t$$

eine spezielle Lösung der Ausgangsgleichung

$$\ddot{y}(t)-y(t) = 6\sinh t\,,$$

und die allgemeine Lösung dieser Differentialgleichung ist

$$y(t) = 3t\cosh t+a\cosh t+b\sinh t$$

mit beliebigen Konstanten a, b aus $\mathbb{R}$ oder $\mathbb{C}$ – je nachdem, worüber wir das Problem betrachten.

Für lineare Differentialgleichungen mit konstanten Koeffizienten führt Variation der Konstanten also immer zu einer Lösung der inhomogenen Gleichung, allerdings enthält diese noch Integrale, und die Bestimmung dieser Stammfunktionen kann durchaus problematisch sein.

Auch bei Systemen von linearen Differentialgleichungen mit *variablen* Koeffizienten kann man mit Variation der Konstanten arbeiten, da wir für solche Systeme allerdings nicht wissen, wie man das homogene System löst, hat es keinen Sinn, hierauf genauer einzugehen. Merken sollte man sich nur, daß die Methode auch in vielen anderen Situationen einen Versuch wert ist.

i) Eindeutigkeitsfragen

Wir hatten Differentialgleichungen eingeführt, um aus dem gegenwärtigen Zustand eines Systems Folgerungen über die künftige Entwicklung zu ziehen. Dies ist natürlich nur dann möglich, wenn das zugehörige

Anfangswertproblem eindeutig lösbar ist. Bei linearen Differentialgleichungen mit konstanten Koeffizienten ist das, wie wir im vorigen Abschnitt gesehen haben, immer der Fall; in diesem Abschnitt wollen wir uns überlegen, welche Probleme im nichtlinearen Fall auftreten können.

Wir betrachten als erstes die Differentialgleichung

$$\dot{y}(t) = \frac{1}{2y(t)}\,,$$

wobei wir uns hier wie auch bei allen Beispielen in diesem Abschnitt der Einfachheit halber auf *reelle* Lösungsfunktionen beschränken wollen.

Multiplikation beider Seiten mit $2y(t)$ führt auf

$$2\dot{y}(t)y(t) = 1\,,$$

die Ableitung von $y(t)^2$ ist also gleich 1. Damit ist

$$y(t)^2 = t + C \quad \text{oder} \quad y(t) = \pm\sqrt{t+C}$$

mit einer beliebigen Konstanten $C \in \mathbb{R}$. Fordern wir nun, daß $y(0) = 0$ sein soll, so gibt es die beiden Lösungen

$$y(t) = \sqrt{t} \quad \text{und} \quad y(t) = -\sqrt{t}\,,$$

von denen die eine für $t \to \infty$ gegen $+\infty$ geht, die andere gegen $-\infty$.

Als zweites Beispiel betrachten wir das Anfangswertproblem

$$\dot{y}(t) = 2\sqrt{|y(t)|} \quad \text{mit} \quad y(0) = 0\,.$$

Es hat natürlich die Nullfunktion als Lösung, aber auch die stetig differenzierbare Funktion

$$y(t) = \begin{cases} t^2 & \text{für } t \geq 0 \\ -t^2 & \text{für } t \leq 0 \end{cases}.$$

($y(t) = t^2$ ist für $t < 0$ *keine* Lösung.) Da die Ableitung von $\pm t^2$ für $t = 0$ verschwindet, sind aber auch die beiden zusammengesetzten Funktionen

$$y(t) = \begin{cases} 0 & \text{für } t \geq 0 \\ -t^2 & \text{für } t \leq 0 \end{cases} \quad \text{und} \quad y(t) = \begin{cases} t^2 & \text{für } t \geq 0 \\ 0 & \text{für } t \leq 0 \end{cases}$$

stetig differenzierbar und Lösungen.

Es kommt noch schlimmer: Für jedes $a \in \mathbb{R}$ löst

$$y(t) = (t-a)^2$$

für $t \geq a$ (und nur dort) die Differentialgleichung, wenn auch nicht das Anfangswertproblem. Für $a \geq 0$ können wir dazu die zusammengesetzte Funktion

$$y(t) = \begin{cases} 0 & \text{für } t \leq a \\ (t-a)^2 & \text{für } t \geq a \end{cases}$$

betrachten; diese ist in jedem Punkt t stetig differenzierbar und erfüllt die Differentialgleichung samt Anfangsbedingung $y(0) = 0$.

Auf Grund der Differentialgleichung und der Anfangsbedingung wissen wir also nur, daß die Funktion entweder konstant gleich null ist oder aber nach einem Zeitraum $a \geq 0$ damit beginnt, quadratisch gegen unendlich zu gehen – nicht gerade viel Information.

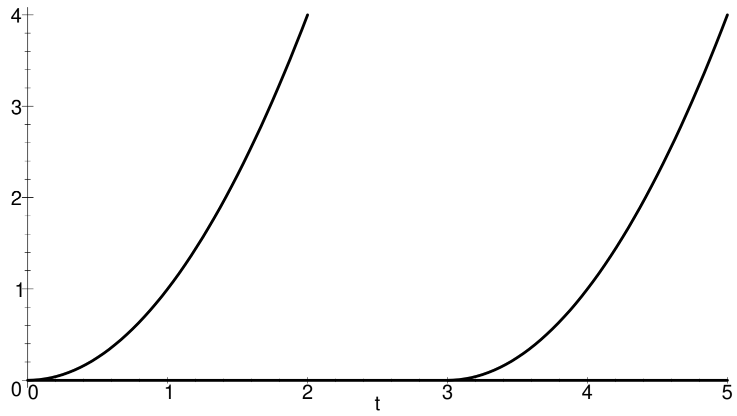

Abb. 10: Drei Lösungen von $\dot{y}(t) = \sqrt{|y(t)|}$ mit $y(0) = 0$

Bei der Differentialgleichung

$$\dot{y}(t) + \frac{y(t)}{2t^3} = 0$$

schließlich überzeugt man sich leicht, daß für jedes $\lambda \in \mathbb{R}$ die Funktion $y(t) = \lambda e^{-1/t^2}$ eine Lösung ist. Diese Funktion ist zwar *a priori* für $t = 0$

nicht definiert, sie kann aber stetig ergänzt werden durch die Vorschrift, daß ihr Wert dort gleich Null sein soll. Genauso verhält es sich mit ihrer Ableitung

$$\dot{y}(t) = -\frac{x(t)}{2t^3} = -\frac{\lambda}{2t^3} \cdot e^{-1/t^2},$$

denn e^{-1/t^2} geht schneller gegen Null als $\frac{\lambda}{2t^3}$ gegen Unendlich geht. Also ist die Funktion stetig differenzierbar (sogar beliebig oft, wie man sich ohne große Schwierigkeiten überlegen kann). Ihr Wert an der Stelle $t = 0$ ist unabhängig von λ immer gleich Null, d.h. das Anfangswertproblem

$$\dot{y}(t) + \frac{y(t)}{2t^3} = 0 \quad \text{und} \quad y(0) = 0$$

hat *jede* der Funktionen $y(t) = \lambda e^{-1/t^2}$ als Lösung, egal welchen Wert $\lambda \in \mathbb{R}$ man auch wählt. Über die Entwicklung eines Systems, das durch diese Gleichung beschrieben wird, läßt sich also *nichts* aussagen. Erst wenn $x(t_0)$ für einen Wert $t_0 \neq 0$ bekannt ist, lassen sich die Lösungen mit den verschiedenen Parameterwerten λ voneinander unterscheiden. Für $t \to \infty$ konvergiert die Lösungsfunktion gegen den Wert λ, der somit die entscheidende Größe für die Beschreibung des Langzeitverhaltens ist und trotzdem durch das Anfangswertproblem völlig unbestimmt gelassen wird.

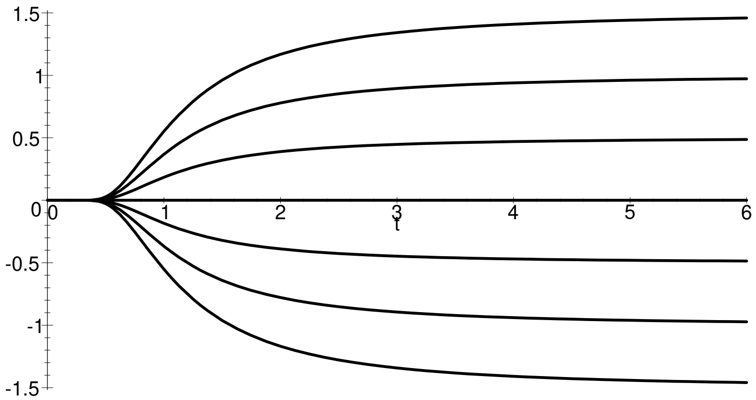

Abb. 11: Sieben Lösungen von $\dot{y}(t) + \frac{y(t)}{2t^3} = 0$ mit $y(0) = 0$

Falls man Differentialgleichungen zur Beschreibung und vor allem zur Vorhersage realer Phänomene einsetzt, ist es also definitiv kein überflüssiger Luxus, wenn man sich zunächst darum kümmert, ob die Differentialgleichung zusammen mit ihrer Anfangsbedingung überhaupt eine eindeutig bestimmte Lösungsfunktion festlegt. Ohne diese Vorüberlegung ist jede Lösungsfunktion wertlos.

j) Der Satz von Picard und Lindelöf

Wir betrachten in diesem Abschnitt das eindimensionale Anfangswertproblem

$$\dot{y}(t) = f\big(y(t), t\big) \quad \text{mit} \quad y(t_0) = c_0 .$$

Dabei sei

$$f \colon \begin{cases} \mathbb{R} \times [t_0, t_1] \to \mathbb{R} \\ \qquad (y, t) \mapsto f(y, t) \end{cases}$$

eine stetige Funktion, d.h. für jeden Punkt (y, t) aus dem Definitionsbereich und jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß

$$\left| f\left(\widetilde{y}, \widetilde{t}\right) - f(y, t) \right| < \varepsilon \quad \text{falls} \quad |\widetilde{y} - y| < \delta \text{ und } \left|\widetilde{t} - t\right| < \delta .$$

Indem wir beide Seiten der Differentialgleichung ab t_0 integrieren, erhalten wir unter Berücksichtigung der Anfangsbedingung $y(t_0) = c_0$ die neue Gleichung

$$y(t) = c_0 + \int_{t_0}^{t} f\big(y(\tau), \tau\big)\, d\tau ,$$

aus der durch Diffenzieren sofort folgt, daß jede ihrer Lösungen das obige Anfangswertproblem löst. Mit der Abkürzung

$$T\big(y(t)\big) \underset{\text{def}}{=} c_0 + \int_{t_0}^{t} f\big(y(\tau), \tau\big)\, d\tau$$

können wir auch kurz sagen, daß wir Lösungen der Gleichung

$$y(t) = T\big(y(t)\big)$$

oder, anders ausgedrückt, Fixpunkte des Operators T suchen.

Um einen solchen Fixpunkt zu erhalten starten wir mit einer beliebigen Funktion $y_0(t)$ auf dem Intervall $[t_0,\ t_1]$ und betrachten die Folge von Funktionen

$$y_n(t) = T^n\big(y(t)\big) ;$$

falls ein Grenzwert für $n \to \infty$ existiert, sollte das ein Fixpunkt sein.

Als Beispiel betrachten wir das (wohlbekannte) Anfangswertproblem

$$\dot{y}(t) = y(t) \quad \text{mit} \quad y(0) = 1 .$$

Hier ist

$$T\big(y(t)\big) = 1 + \int_0^t y(\tau)\, d\tau ;$$

falls wir also von der konstanten Funktion $y_0(t) = 1$ ausgehen, ist

$$y_1(t) = 1 + \int_0^t d\tau = 1 + t ,$$

$$y_2(t) = 1 + \int_0^t (1+\tau)\, d\tau = 1 + t + \frac{t^2}{2} ,$$

$$y_3(t) = 1 + \int_0^t \left(1 + \tau + \frac{\tau^2}{2}\right) d\tau = 1 + t + \frac{t^2}{2} + \frac{t^3}{6} ,$$

und so weiter. Ab hier wird man erraten, daß die Funktionen $y_n(t)$ gerade die TAYLOR-Polynome n-ten Grades der Exponentialfunktion sind, was man dann auch leicht durch vollständige Induktion beweisen kann. Somit ist

$$\lim_{n\to\infty} y_n(t) = e^t ,$$

wie es auch nicht anders zu erwarten war.

In diesem Beispiel geht also alles gut; für den allgemeinen Fall haben wir bislang noch zu wenig in der Hand und stellen deshalb die folgende

Lipschitz-Bedingung: Es gibt eine Konstante $L \in \mathbb{R}$, genannt LIPSCHITZ-Konstante, so daß für alle $t \in [t_0,\ t_1]$ und alle $y_1, y_2 \in \mathbb{R}$ gilt:

$$|f(y_2,t) - f(y_1,t)| \le L\,|y_2 - y_1| .$$

RUDOLF OTTO SIGISMUND LIPSCHITZ (1832–1903) wurde in Königsberg geboren und starb in Bonn. Am besten bekannt ist er durch die gerade definierte LIPSCHITZ-Bedingung für die Existenz und Eindeutigkeit der Lösungen von Differentialgleichungen. Sein Hauptarbeitsgebiet waren die Differentialgleichungen der mathematischen Physik, jedoch beschäftigte er sich auch mit weiteren Methoden der mathematischen Physik wie etwa Matrizengruppen. Seine Arbeiten über dynamische Systeme haben wichtige Anwendungen in der Himmelsmechanik.

Die LIPSCHITZ-Bedingung muß glücklicherweise in vielen Fällen nicht explizit durch Abschätzungen nachgeprüft werden, denn es gilt

Lemma: Falls die partielle Ableitung von f nach y existiert und betragsmäßig beschränkt ist durch eine Konstante L, genügt f einer LIPSCHITZ-Bedingung mit L als LIPSCHITZ-Konstante.

Beweis: Nach dem Mittelwertsatz der Differentialrechnung (**[Ba]**, §3.2, *Satz 2.4*) gibt es für jedes feste t (das wir bezüglich der partiellen Ableitung wie eine Konstante betrachten können) und für je zwei Werte $y_1, y_2 \in \mathbb{R}$ einen Punkt $z \in [y_1,\ y_2]$ mit der Eigenschaft, daß

$$\frac{f(y_2,t) - f(y_1,t)}{y_2 - y_1} = \frac{\partial f}{\partial y}(z,t)$$

ist. Also ist

$$|f(y_2,t) - f(y_1,t)| = \left|\frac{\partial f}{\partial y}(z,t)\right|\ |y_2 - y_1| ,$$

und da der Betrag der partiellen Ableitung überall höchstens L ist, erfüllt f eine LIPSCHITZ-Bedingung mit Konstante L.

■

Wir betrachten nun wieder den oben eingeführten Operator $T\big(y(t)\big)$, von dem uns im Augenblick nur interessieren soll, daß er stetige Funktionen auf dem Intervall $[t_0,\ t_1]$ in ebensolche Funktionen überführt. und wir suchen eine Funktion $y(t)$ mit $T\big(y(t)\big) = y(t)$. Dazu erinnern wir uns an die Supremumsnorm, die wir für auf $[t_0,\ t_1]$ stetige Funktionen f durch

$$\|f\| \underset{\text{def}}{=} \sup_{\tau\in[t_0,\,t_1]} |f(\tau)|$$

definieren. Unser wichtigstes Hilfsmittel ist der BANACHsche Fixpunktsatz, der in diesem Semester noch in der Numerik allgemein formuliert und bewiesen wird. Um keine neuen Begriffe einführen zu müssen, begnüge ich mich hier mit dem für unsere Zwecke notwendigen Spezialfall; der Beweis vereinfacht sich dadurch zwar nichts, aber die Formulierung des Satzes wird leichter verständlich.

Banachscher Fixpunktsatz: Für einen Operator T, der stetige Funktionen auf $[t_0,\ t_1]$ auf ebensolche Funktionen abbildet, gebe es eine Konstante $K < 1$ derart, daß für alle auf $[t_0,\ t_1]$ stetigen Funktionen $w(t), z(t)$ gilt

$$\Big\|T\big(z(t)\big) - T\big(w(t)\big)\Big\| \le K\,\|z(t) - w(t)\|\ .$$

Dann gibt es genau eine stetige Funktion $y\colon [t_0,\ t_1] \to \mathbb{R}$ mit der Eigenschaft

$$T\big(y(t)\big) = y(t)\,.$$

Beweis: Zunächst ist klar, daß es *höchstens* eine solche Funktion gibt, denn wären $y(t)$ und $z(t)$ zwei Lösungen, so wäre

$$\|z(t) - y(t)\| = \Big\|T\big(z(t)\big) - T\big(y(t)\big)\Big\| \le K\ \|z(t) - y(t)\|\ ,$$

was für $K < 1$ nur dann möglich ist, wenn $\|z(t) - y(t)\|$ verschwindet, wenn also überall $y(t) = z(t)$ ist.

Zum Nachweis der Existenz starten wir mit irgendeiner stetigen Funktion $y_0(t)$ und iterieren sie mit T, d.h. für $n > 0$ sei

$$y_n(t) = T\big(y_{n-1}(t)\big)\ .$$

Wir überlegen uns zunächst, daß für jedes feste $\tau \in [t_0,\ t_1]$ die Folge der reellen Zahlen $y_n(t)$ konvergiert. Nach dem CAUCHYschen Konvergenzkriterium (**[Ba]**, §2.1, *Satz 1.29*) müssen wir zeigen, daß es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß

$$|y_m(\tau) - y_n(\tau)| \le \varepsilon \quad \text{falls} \quad m > n > N$$

ist. Da wir mit der Supremumsnorm arbeiten, ist

$$|y_m(\tau) - y_n(\tau)| \le \|y_m(t) - y_n(t)\|\ ,$$

und nach Voraussetzung ist für jedes $r \in \mathbb{N}$

$$\|y_{r+2}(t) - y_{r+1}(t)\| \le K\,\|y_{r+1}(t) - y_r(t)\|\ .$$

Induktiv folgt weiter, daß

$$\|y_{r+1}(t) - y_r(t)\| \le K^r\,\|y_1(t) - y_0(t)\|$$

ist und damit nach der Dreiecksungleichung

$$\begin{aligned}
\|y_m(t) - y_n(t)\| &\le \big\|y_m(t) - y_{m-1}(t)\big\| + \cdots + \|y_{n+1}(t) - y_n(t)\| \\
&\le K^m\,\|y_1(t) - y_0(t)\| + \cdots + K^n\,\|y_1(t) - y_0(t)\| \\
&= (K^m + \cdots + K^n)\,\|y_1(t) - y_0(t)\| \\
&= K^n(1 + K + \cdots + K^{m-n})\,\|y_1(t) - y_0(t)\| \\
&\le K^n i\left(\sum_{i=0}^{\infty} K^i\right)\|y_1(t) - y_0(t)\| \\
&= \frac{K^n}{1-K}\,\|y_1(t) - y_0(t)\|\ .
\end{aligned}$$

Wählen wir also N so, daß

$$\frac{K^N}{1-K}\,\|y_1 - y_0(t)\| < \varepsilon$$

ist, gilt die Voraussetzung des CAUCHYschen Konvergenzkriteriums mit diesem N für alle $\tau \in [t_0,\ t_1]$. Somit konvergiert die Folge der Funktionen $y_n(t)$ *punktweise* gegen eine Funktion $y(t)$. Für jedes $\tau \in [t_0,\ t_1]$ ist dann

$$\begin{aligned}
|y(\tau) - y_n(\tau)| &= \lim_{m\to\infty} |y_m(\tau) - y_n(\tau)| \le \lim_{m\to\infty} \|y_m(t) - y_n(t)\| \\
&\le \frac{K^n}{1-K}\,\|y_1(t) - y_0(t)\|\ ,
\end{aligned}$$

die Folge konvergiert also in $[t_0,\, t_1]$ *gleichmäßig* gegen $y(t)$. (*vgl.* **[Ba]**, §5.1, *Def. 1.3*) Damit ist $y(t)$ nach **[Ba]**, §5.1, *Satz 1.5* eine stetige Funktion, und der Satz ist bewiesen. ∎

STEFAN BANACH (1892–1945) wurde in Krakau geboren und ausgebildet, promovierte und arbeitete dann aber an der Universität von Lvov in der Ukraine, wo er unter schwierigen Bedingungen unter deutscher Besatzung den zweiten Weltkrieg verbrachte. Durch seine Arbeiten über lineare Operatoren und über Vektorräume von Funktionen wurde er zum Begründer der modernen Funktionalanalysis. Nach dem Krieg wollte er auf einen Lehrstuhl an der Universität Krakau wechseln, starb aber 1945 an Lungenkrebs. Das wichtigste mathematische Forschungsinstitut Polens, das Banach-Zentrum in Warschau, ist nach ihm benannt.

Als Anwendung erhalten wir den

Satz von Picard-Lindelöf: Die stetige Funktion

$$f\colon \begin{cases} \mathbb{R}\times[t_0,\, t_1] \to \mathbb{R} \\ \qquad (y,t) \mapsto f(y,t) \end{cases}$$

genüge einer LIPSCHITZ-Bedingung mit irgendeiner Konstanten $L \in \mathbb{R}$. Dann hat das Anfangswertproblem

$$\dot y(t) = f\big(y(t),t\big) \quad \text{mit} \quad y(t_0) = c_0$$

für jedes $c_0 \in \mathbb{R}$ und $t_0 \in [t_0,\, t_1]$ eine eindeutig bestimmte Lösung.

Beweis: Wir müssen zeigen, daß für den oben definierten Operator

$$T\big(y(t)\big) \underset{\text{def}}{=} c_0 + \int_{t_0}^{t} f\big(y(\tau),\tau\big)\, d\tau$$

und zwei beliebige stetige Funktionen z, w auf $[t_0,\, t_1]$

$$\big\|T\big(z(t)\big) - T\big(w(t)\big)\big\| \le K\,\|z(t) - w(t)\|$$

ist mit einer Konstanten $K < 1$. Links steht die Norm von

$$T\big(z(t)\big) - T\big(w(t)\big) = \int_{t_0}^{t} \Big(f\big(z(\tau),\tau\big) - f\big(w(\tau),\tau\big)\Big)\, d\tau\,,$$

und auf Grund der LIPSCHITZ-Bedingung ist

$$\big|f\big(z(\tau),\tau\big) - f\big(w(\tau),\tau\big)\big| \le L\,|z(\tau) - w(\tau)| \le \|z(t) - w(t)\|\ .$$

Also ist für $t \in [t_0,\, t_1]$

$$\begin{aligned} \left|\int_{t_0}^{t} \Big(f\big(z(\tau),\tau\big) - f\big(w(\tau),\tau\big)\Big)\, d\tau\right| &\le \int_{t_0}^{t} L\,\|z(t) - w(t)\|\, d\tau \\ &= L\,\|z(t) - w(t)\|\,(t - t_0) \\ &\le L\cdot(t_1 - t_0)\,\|z(t) - w(t)\|\ . \end{aligned}$$

Falls $L\cdot(t_1 - t_0)$ kleiner als eins ist, haben wir eine Konstante $K < 1$ gefunden und können den BANACHschen Fixpunktsatz anwenden, der uns eine stetige Funktion $y(t)$ liefert, die der Bedingung

$$y(t) = c_0 + \int_{t_0}^{t} f\big(y(\tau),\tau\big)\, d\tau$$

genügt. Da die rechte Seite wegen der Stetigkeit von f differenzierbar ist, haben wir sogar eine differenzierbare Funktion, und

$$\dot y(t) = f\big(y(t),t\big)\ .$$

Außerdem ist $y(t_0) = c_0$, wir haben also eine Lösung des Anfangswertproblems gefunden. Sie ist die einzige Lösung, denn jede Lösung muß Fixpunkt von T sein, und nach dem BANACHschen Fixpunktsatz gibt es nur einen Fixpunkt.

Falls $L\cdot(t_1 - t_0)$ größer als eins ist, gibt es immerhin noch eine natürliche Zahl n, so daß $L\cdot(t_1 - t_0) < n$ ist. Wir unterteilen das Intervall $[t_0,\, t_1]$ in n gleichlange Teilintervalle $[\tau_i,\, \tau_{i+1}]$ mit Grenzen

$$t_0 = \tau_0 < \tau_1 < \cdots < \tau_{n-1} < \tau_n = t_1\,,$$

d.h.

$$\tau_i = t_0 + \frac{i}{n}(t_1 - t_0)\,.$$

Für jedes dieser Intervalle ist dann $L\cdot(\tau_{i+1} - \tau_i) < 1$, insbesondere ist also das Anfangswertproblem im Anfangsintervall $[\tau_0,\, \tau_1]$ eindeutig lösbar.

Für die entsprechende Lösung sei $y(\tau_1) = c_1$. Dann betrachten wir im Intervall $[\tau_1, \tau_2]$ das entsprechende Anfangswertproblem mit $y(\tau_1) = c_1$ und erhalten eine in diesem Intervall eindeutige Lösung und so weiter, bis wir die Lösung auf ganz $[t_0, t_1]$ ausgedehnt haben.

■

ÉMILE PICARD (1856–1941) wurde in Paris geboren und hatte an der dortigen Universität auch seine erste Stelle, bis er 1878 eine Professur in Toulouse bekam. 1898 wurde er Professor an der Sorbonne und kehrte er nach Paris zurück, wo er auch starb. Seine Arbeiten beschäftigen sich mit der Analysis (z.B. dem gerade bewiesenen Satz) und der Geometrie. Dort interessierte er sich für algebraische Flächen und die eng damit verbundenen algebraiscchen Funktionen zweier Veränderlichen sowie den zugehörigen Integralen. Weitere Arbeiten beschäftigen sich auch mit der Topologie von Flächen und wieder andere mit Anwendungen der Analysis in der Elektrodynamik, Wärmelehre und Elastizitätstheorie. Er war mit der Tochter von CHARLES HERMITE verheiratet.

ERNST LEONARD LINDELÖF (1870–1946) wurde in Helsinki geboren, als Finnland noch eine russische Provinz war. Sein Vater war Mathematikprofessor an der damals noch schwedischen (heute finnischen) Universität Helsingfors, an der später auch Lindelöf selbst sein Studium begann, unterbrochen durch Aufenthalte in Stockholm (1891), Paris (1893-1894) und Göttingen (1901). Der gerade bewiesene Satz stammt aus einer Arbeit von 1890; weitere Arbeiten beschäftigen sich mit analytischen Funktionen und Singularitäten. Nachdem er Professor in Helsingfors wurde, widmete er sich vor allem der Lehre und publizierte mehrere Bücher.

Die Anfangswertprobleme aus dem vorigen Abschnitt erfüllen offensichtlich allesamt *keine* LIPSCHITZ-Bedingung, denn sie sind ja nicht eindeutig lösbar. Im Falle von

$$\dot{y}(t) = \frac{1}{2y(t)} \quad \text{mit} \quad y(0) = 0$$

ist

$$f(y,t) = \frac{1}{2y},$$

also

$$|f(y_2,t) - f(y_1(t)| = \left|\frac{1}{2y_2} - \frac{1}{2y_1}\right| = \frac{1}{|2y_1y_2|}\,|y_2 - y_1| \;,$$

und das kann man nur dann durch eine Schranke der Form $L\,|y_2 - y_1|$ abschätzen, wenn $1/\,|2y_1y_2|$ kleiner ist als L. In der Nähe des Nullpunkts kann $1/\,|2y_1y_2|$ aber beliebig groß werden, und somit kann hier keine LIPSCHITZ-Bedingung gelten.

Ähnlich verhält es sich beim Anfangswertproblem

$$\dot{y}(t) = 2\sqrt{|y(t)|} \quad \text{mit} \quad y(0) = 0\,.$$

Hier ist

$$f(y,t) = 2\sqrt{|y(t)|}$$

zwar überall stetig, aber die Ableitung

$$\frac{\partial f}{\partial y}(y,t) = 2\frac{d}{dy}\sqrt{|y|} = \pm\frac{1}{\sqrt{|y|}}$$

mit Pluszeichen für $y > 0$ und Minuszeichen für $y < 0$ wächst bei Annäherung an den Nullpunkt betragsmäßig unbeschränkt, die Kurve wird also immer steiler, und damit zeigt der Mittelwertsatz der Differentialrechnung, daß es keine LIPSCHITZ-Konstante geben kann.

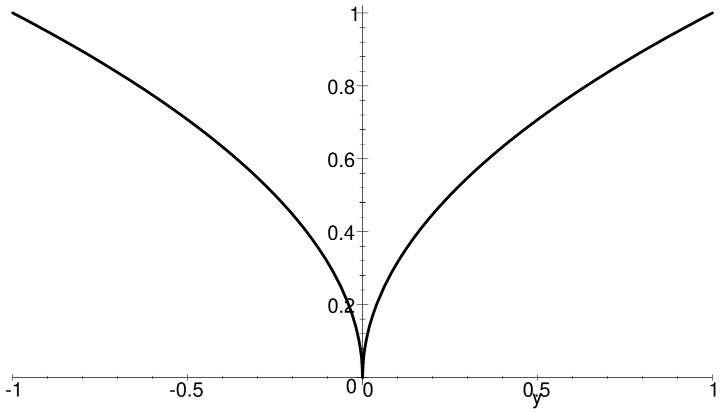

Abb. 12: Die Funktion $2\sqrt{|y|}$