

6 Differentialgleichungen – eine kurze Wiederholung

Viele Zusammenhänge in der Natur lassen sich durch Differentialgleichungen beschreiben. Dazu gehören Vorgänge wie die Wärmeleitung, die Diffusion in Flüssigkeiten, Wellenphänomene aller Art, oder auch die zeitliche Entwicklung von Tierpopulationen. Eine Differentialgleichung beschreibt dabei, wie Ableitungen einer Funktion (etwa nach der Zeit oder nach räumlichen Komponenten) wieder von der ursprünglichen Funktion und evtl. der Zeit und den Koordinaten abhängig sind. Hier wollen wir nur gewöhnliche Differentialgleichungen ansehen, es kommen also keine partiellen Ableitungen vor. Allerdings wollen wir auch Differentialgleichungen höherer Ordnung ansehen, wie sie etwa bei Schwingungsphänomenen vorkommen.

Eine allgemeine Differentialgleichung (DGL) 1. Ordnung besitzt die Form

$$\dot{x}(t) := \frac{d}{dt} x(t) = f(x(t), t), \quad (31)$$

wobei wir hier die Zeit als Parameter nehmen. Oft schreibt man auch kurz $\dot{x} = f(x, t)$. In der Regel liegt weiter ein *Anfangswertproblem* (oder auch *Cauchy-Problem*) vor, d.h. es ist die Lösung dieser DGL gesucht mit einem gegebenen Anfangswert, $x(t_0) = x_0$. In einfachen Fällen kann man ein solches Problem explizit lösen. Falls nicht, gibt es verschiedene Verfahren aus der Numerik für eine näherungsweise (numerische) Lösung. An dieser Stelle wird ein iteratives Verfahren nach Volterra vorgestellt.

6.1 Volterra-Integration

Die Volterra-Integralgleichung entsteht durch Integration von Gl. (31) und lautet

$$x(t) = x_0 + \int_{t_0}^t f(x(\tau), \tau) d\tau.$$

Diese Gleichung, deren Lösungen im Wesentlichen dieselben sind wie diejenigen des obigen Anfangswertproblems, kann iterativ eingesetzt werden. Mit $x_0(t) \equiv x_0$ als Startwert setzen wir dazu

$$x_{n+1}(t) := x_0 + \underbrace{\int_{t_0}^t f(x_n(\tau), \tau) d\tau}_{=:(\Phi(x_n))(t)}$$

für $n \in \mathbb{N}_0$. Dabei ist $\Phi: C^0[a, b] \rightarrow C^0[a, b]$ mit $t_0 \in [a, b]$ eine Abbildung, die wir noch genauer ansehen werden. Die Iteration wird — je nach Text — *Picard-Iteration* oder *Volterra-Iteration* genannt.

Beispiel 6.1. Die DGL $\dot{x}(t) = f(t)$ mit Anfangswert $x(t_0) = x_0$ kann einfach durch Integration gelöst werden,

$$x(t) = x_0 + \int_{t_0}^t f(\tau) d\tau.$$

Etwas schwieriger ist die DGL $\dot{x} = f(x)$, deren Lösung mit der Methode der Trennung der Variablen berechnet werden kann. Durch Integration bekommt man

$$\int_{x_0}^{x(t)} \frac{d\xi}{f(\xi)} = \int_{t_0}^t \frac{\dot{x}(\tau)}{f(x(\tau))} d\tau = \int_{t_0}^t 1 d\tau = (t - t_0)$$

wobei angenommen sei, dass der Nenner nirgends verschwindet (sonst muss man natürlich anders vorgehen). Wenn man nun das linke Integral ausrechnen kann, löst man danach nach $x(t)$ auf. Diese Methode kommt etwa bei der logistischen Gleichung zum Einsatz; vgl. [12] für Details.

Wichtiger als die bisherigen Beispiele ist die folgende DGL, die wir später noch vektorwertig verallgemeinern werden.

Beispiel 6.2. Betrachten wir $\dot{x} = ax$ mit $x(0) = 1$, wobei $a \in \mathbb{R}$ beliebig, aber fest sei. Dann kann man induktiv nachrechnen, dass die Picard-Iteration auf

$$x_n(t) = \sum_{\ell=0}^n \frac{(at)^\ell}{\ell!}$$

für alle $n \in \mathbb{N}_0$ führt (Übung: nachrechnen!). Aus der Analysis wissen wir, dass dann $\lim_{n \rightarrow \infty} x_n(t) = e^{at}$ gilt, mit punktwiser Konvergenz für jedes $t \in \mathbb{R}$ und gleichmäßiger Konvergenz auf jedem kompakten Intervall. In der Tat ist e^{at} die *eindeutige* Lösung unseres Anfangswertproblems.

Die Situation im letzten Beispiel hat eine wichtige Verallgemeinerung, den Existenz- und Eindeutigkeitssatz für gewöhnliche DGLen.

Satz 6.3. (*Picard–Lindelöf*)

Gegeben sei eine Funktion $f: \mathbb{R} \times [a, b] \rightarrow \mathbb{R}$. Sei f stetig und zudem Lipschitz bzgl. der x -Koordinate, mit Lipschitz-Konstante L . Sei außerdem $t_0 \in [a, b]$. Dann besitzt das Anfangswertproblem $\dot{x} = f(x, t)$ mit $x(t_0) = x_0$ eine eindeutige Lösung, die in t bis zum Rand des Intervalls fortgesetzt werden kann. Diese Lösung kann durch die Picard-Iteration beliebig genau approximiert werden.

Beweisskizze. Der klassische Beweis beruht auf einer Reduktion auf den Banach'schen Fixpunktsatz. Um ihn anwenden zu können, definieren wir uns zunächst eine geeignete Norm,

$$\|x\| := \max_{t \in [a, b]} |x(t)| e^{-\alpha t}$$

für ein (noch geschickt zu wählendes) $\alpha > 0$. Nun rechnet man nach, dass für die Abbildung Φ aus der Picard-Iteration die Ungleichung

$$\|\Phi(x) - \Phi(y)\| \leq \frac{L}{\alpha} \|x - y\|$$

gilt, wobei L die Lipschitz-Konstante von f ist. Wählt man nun $\alpha > L$, so ist Φ eine Kontraktion, mit Kontraktionskonstante $\frac{L}{\alpha}$. Satz 2.13 liefert nun sowohl die Existenz als auch die Eindeutigkeit der Lösung, und zudem die exponentiell schnelle Konvergenz der Picard-Iteration gegen diese Lösung. $\square$

Man kann eine ganze Reihe von Verallgemeinerungen dieses Satzes formulieren. So muss f i.A. in der x -Koordinate nicht auf ganz $\mathbb{R}$ definiert sein, oder die Lipschitz-Eigenschaft wird nur lokal gefordert — dementsprechend ändert sich dann immer ein wenig am Resultat. Dies ist sehr schön im Buch von Walter [12] ausgeführt. Außerdem kann man Gl. (31) auch als Vektor-Gleichung lesen — und alles Weitere danach lässt sich analog entwickeln. Man spricht dann von *Systemen* von DGLen erster Ordnung.

Beispiel 6.4. Das für uns wichtigste System ist das lineare, mit konstanten Koeffizienten. Dies schreiben wir als $\dot{\mathbf{x}} = A\mathbf{x}$ mit einer konstanten Matrix A , die i.A. komplexe Einträge haben darf. Explizit geschrieben sieht das so aus:

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A\mathbf{x},$$

und die Anfangsbedingung lautet wieder $\mathbf{x}(t_0) = \mathbf{x}_0$.

Die eindeutige Lösung lautet in diesem Fall $\mathbf{x}(t) = \exp((t - t_0)A)\mathbf{x}_0$, und man kann auch diese explizit aus der Picard-Iteration herleiten. Dabei verwendet man die Reihendefinition für e^A , die wir im nächsten Kapitel noch genauer besprechen werden. Die Eindeutigkeit kann man auch beweisen, indem man annimmt, dass $\mathbf{y}(t)$ eine Lösung ist und dann $\mathbf{z}(t) := \exp(-(t - t_0)A)\mathbf{y}(t)$ setzt. Nun zeigt man, dass $\dot{\mathbf{z}}(t) \equiv \mathbf{0}$ gilt, also $\mathbf{z}(t) = \text{const.}$ An der Stelle t_0 kennen wir den Wert aber, das ist $\mathbf{z}(t_0) = \mathbf{x}_0$, und dann folgt $\mathbf{y}(t) = \exp((t - t_0)A)\mathbf{x}_0$.

6.2 Lineare Differentialgleichungen höherer Ordnung

Differentialgleichungen höherer Ordnung können behandelt werden, indem man sie auf DGL-Systeme 1. Ordnung zurückführt.

Für die Lösung von DGLs 2. Ordnung soll dies zunächst an einem Beispiel erläutert werden, nämlich anhand der Schwingungsgleichung $\ddot{x} + x = 0$. Diese besitzt die allgemeine Lösung $x(t) = \alpha \cdot \cos(t) + \beta \cdot \sin(t)$. Die Parameter α und β werden dann eindeutig festgelegt durch die Vorgabe von

$$\begin{aligned} x(t_0) &= \alpha \cos(t_0) + \beta \sin(t_0), \\ \dot{x}(t_0) &= -\alpha \sin(t_0) + \beta \cos(t_0). \end{aligned}$$

Durch Auflösen bekommt man

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \cos(t_0) & -\sin(t_0) \\ \sin(t_0) & \cos(t_0) \end{pmatrix} \begin{pmatrix} x(t_0) \\ \dot{x}(t_0) \end{pmatrix}.$$

Damit ist die Lösung dann eindeutig bestimmt.

Dieses Ergebnis erhält man durch Rückführung auf ein Paar an Differentialgleichungen 1. Ordnung mittels

$$x_1 := \dot{x} \quad \text{und} \quad x_2 := x.$$

Damit bekommt man dann das System

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{=: A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (32)$$

mit den Anfangswerten $x_1(t_0) = \dot{x}(t_0)$ und $x_2(t_0) = x(t_0)$. Für eine Differentialgleichung n -ter Ordnung funktioniert diese Rückführung auf ein n -Tupel von Differentialgleichungen 1. Ordnung analog.

Die Lösung des Systems (32) kann man nun folgendermaßen berechnen. Zunächst zeigt man, dass $A^{2m} = (-1)^m \mathbb{1}$ und $A^{2m+1} = (-1)^m A$ gilt, für alle $m \in \mathbb{N}_0$. Damit können wir $\exp(tA)$ berechnen, und bekommen durch Aufspalten der Summe nach geraden und ungeraden Exponenten

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n = \underbrace{\sum_{m \geq 0} (-1)^m \frac{t^{2m}}{(2m)!} \mathbb{1}}_{=\cos(t)} + \underbrace{\sum_{\ell \geq 0} (-1)^\ell \frac{t^{2\ell+1}}{(2\ell+1)!} A}_{=\sin(t)} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Daraus folgt nun die zuvor bereits angegebene Lösung der DGL 2. Ordnung (mit Anfangswerten):

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \exp((t - t_0)A) \begin{pmatrix} x_1(t_0) \\ x_2(t_0) \end{pmatrix} = \begin{pmatrix} \cos(t - t_0) & -\sin(t - t_0) \\ \sin(t - t_0) & \cos(t - t_0) \end{pmatrix} \begin{pmatrix} x_1(t_0) \\ x_2(t_0) \end{pmatrix}.$$

Dies stimmt mit der obigen Lösung überein.

Eine lineare Differentialgleichung höherer Ordnung mit konstanten Koeffizienten besitzt die allgemeine Form

$$x^{(n)} + a_1 x^{(n-1)} + a_2 x^{(n-2)} + \dots + a_{n-1} \dot{x} + a_n x = 0. \quad (33)$$

Dies kann man nun wie oben erwähnt umformen in ein System n -ter Ordnung, und erhält dann eine $n \times n$ -Matrix A , deren Eigenwerte und ggf. Eigenvektoren zur Lösung herangezogen werden.

Das kann man aber auch abkürzen, mit einem Ansatz von Euler. Dabei wird ein Polynom n -ter Ordnung aufgestellt, welches nach dem Fundamentalsatz der Algebra genau n Nullstellen $\lambda_1, \dots, \lambda_n$ in $\mathbb{C}$ besitzt. Mit $x(t) := e^{\lambda t}$ bekommen wir durch Einsetzen in Gl. (33) in der Tat

$$\underbrace{(\lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n)}_{=: p(\lambda)} \underbrace{e^{\lambda t}}_{\neq 0 \text{ auf ganz } \mathbb{C}} = 0.$$

Diese Gleichung kann nur stimmen, wenn $p(\lambda) = 0$ erfüllt ist, was uns die n Nullstellen (mit Vielfachheiten gezählt) verschafft.

Falls nun die n Nullstellen paarweise verschieden sind (was der generische Fall ist), so lautet die allgemeine Lösung der DGL (33) einfach

$$x(t) = \sum_{\ell=1}^n c_{\ell} e^{\lambda_{\ell} t}.$$

Mit Hilfe der Anfangsbedingungen $x(t_0), \dot{x}(t_0), \dots, x^{(n-1)}(t_0)$ lässt sich ein lineares Gleichungssystem aufstellen, das die Koeffizienten c_{ℓ} eindeutig festlegt.

Wenn Multiplizitäten bei den Nullstellen auftreten, so hängt die Form der allgemeinen Lösung davon ab, ob die Matrix A , die wir hier nicht explizit hingeschrieben haben, diagonalisierbar ist oder nicht. Für Details hierzu sei auf [12] verwiesen.

7 Markov-Ketten

Um das Matrix-Exponential im nächsten Kapitel einsetzen zu können, sind zunächst einige Vorüberlegungen notwendig.

7.1 Matrix-Normen und Matrix-Exponential

Für den Vektorraum $V = \mathbb{R}^n$ (oder $\mathbb{C}^n$) mit einer beliebigen Norm $\|\cdot\|$ und einer linearen Abbildung $A: V \rightarrow V$ ist

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

die zugehörige *Operatornorm*. Dabei gilt

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \max_{\|x\|=1} \|Ax\|.$$

In der Gleichung ist das Supremum gleich dem Maximum, da die Einheitskugel $\{\|x\| \leq 1\}$ im $\mathbb{R}^n$ (bzw. im $\mathbb{C}^n$) kompakt ist, und daher ihr Rand, der eine abgeschlossene Teilmenge ist, auch. Dieser Zusammenhang ist nicht trivial und bedeutet, dass der Extremwert für ein x mit $\|x\| = 1$ wirklich angenommen wird. Es gelten weiter folgende Eigenschaften:

$$\|A\| \geq 0, \text{ mit } \|A\| = 0 \Leftrightarrow A = \mathbf{0} \quad (34)$$

$$\|cA\| = |c| \|A\| \text{ für alle } c \in \mathbb{R} \text{ (} c \in \mathbb{C} \text{)} \quad (35)$$

$$\|A + B\| \leq \|A\| + \|B\| \quad (36)$$

$$\|Ax\| \leq \|A\| \|x\| \quad (37)$$

$$\|AB\| \leq \|A\| \|B\| \quad (38)$$

Dabei beschreiben die Eigenschaften (34), (35) und (36) die üblichen Norm-Axiome und (37) ist eine Verträglichkeits-Bedingung. Schließlich ist (38) eine wichtige neue Eigenschaft. Man überlege sich, warum hier i.A. keine Gleichheit gelten kann!

Als Beispiele werden an dieser Stelle die Matrixnormen zu bereits bekannten Normen aufgestellt für eine $n \times n$ -Matrix A mit Elementen in $\mathbb{C}$ (kurz: $A \in \text{Mat}(n, \mathbb{C})$). Die wichtigsten Beispiele sind in Tabelle 2 zusammengefasst. Für eine dieser Normen wird noch ein neuer Begriff benötigt.

Norm $\ \cdot\ $	Operatornorm	$\ A\ $
$\ \cdot\ _1$	Spaltensummennorm:	$\ A\ _1 = \max_{1 \leq j \leq n} \sum_{i=1}^n a_{ij} $
$\ \cdot\ _2$	Spektralnrm:	$\ A\ _2 = \sqrt{\varphi(A^+ A)}$, mit $A^+ = \overline{A}^t$
$\ \cdot\ _\infty$	Zeilensummennorm:	$\ A\ _\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} $

Tabelle 2: Einige gängige Matrixnormen

Definition 7.1. Das *Spektrum* einer Matrix $A \in \text{Mat}(n, \mathbb{C})$ ist die Menge ihrer Eigenwerte, also

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid \lambda \text{ ist EW von } A\}.$$

Dann heißt

$$\varphi(A) = \max_{\lambda \in \sigma(A)} |\lambda|$$

der *Spektralradius* von A und definiert den kleinsten Kreis um 0 in $\mathbb{C}$, in dem alle Eigenwerte der Matrix A liegen.

Es gibt verschiedene Ansätze, das Matrix-Exponential e^A zu betrachten. Wir beginnen mit der Definition mittels der Potenzreihe der Exponential-Funktion, und erhalten (zunächst formal)

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \quad (39)$$

Wie auch bei der Exponential-Funktion aus Analysis I ist die Konvergenz der Reihe von großer Bedeutung. Daher wird zunächst das Konvergenzverhalten geprüft.

$$\left\| \sum_{m=0}^N \frac{A^m}{m!} \right\| \leq \sum_{m=0}^N \frac{\|A^m\|}{m!} \leq \sum_{m=0}^N \frac{\|A\|^m}{m!} \leq \sum_{m=0}^{\infty} \frac{\|A\|^m}{m!} = e^{\|A\|} < \infty.$$

Daraus folgt die Konvergenz der Reihe für jede gegebene Matrix A , weil wir hier eine Majorante für jedes Element der durch Gl. (39) definierten Matrix berechnet haben (man mache sich die Details als Übung klar).

Beispiel 7.2. Wir berechnen e^{tA} für $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Da $A^2 = 0$, wird die Reihe einfach:

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

Für $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist die Rechnung etwas komplizierter. Da $A^2 = \mathbb{1}$, können wir die Summe nach geraden und ungeraden Indices aufteilen, und bekommen (Details: Übung):

$$e^{tA} = \cosh(t) \mathbb{1} + \sinh(t) A = \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}.$$

Aus unseren obigen Überlegungen folgt auch, dass e^{tA} für jedes $t \in \mathbb{R}$ wohldefiniert ist. Dabei gilt sogar gleichmäßige Konvergenz auf kompakten Intervallen für den Parameter t . Insbesondere gilt $e^{0A} = e^0 = \mathbb{1}$.

Nun betrachte wir die Ableitung:

$$\begin{aligned} \frac{d}{dt} e^{tA} &= \frac{d}{dt} \sum_{m \geq 0} \frac{t^m A^m}{m!} = \sum_{m \geq 1} \frac{t^{m-1}}{(m-1)!} A^m \\ &= A \underbrace{\sum_{m \geq 1} \frac{(tA)^{m-1}}{(m-1)!}}_{=e^{tA}} = A e^{tA} = e^{tA} A. \end{aligned}$$