

# Diskretisierung und numerische Optimierung

Sommersemester 2022

**Martin Burger**  
Department Mathematik, AMMN

Version vom 27. April 2022

---

# Inhaltsverzeichnis

---

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Numerische Lösung von Anfangswertproblemen</b>	<b>3</b>
2.1	Theorie von Anfangswertproblemen für gewöhnliche Differentialgleichungen	4

---

# Abbildungsverzeichnis

---

---

# Kapitel 1

## Einleitung

---

In dieser Vorlesung werden wir einige weiterführende Aspekte der numerischen Mathematik diskutieren, nämlich numerische Verfahren zur Lösung von Optimierungsproblemen und von (gewöhnlichen) Differentialgleichungen.

Der erste Teil der Vorlesung beschäftigt sich mit der Lösung von Differentialgleichungen, insbesondere mit gewöhnlichen Differentialgleichungen. Wir beginnen mit Anfangswertproblemen der Form

$$u'(t) = F(u(t), t), \quad u(0) = u_0,$$

die nur für spezielle Formen von  $F$  gelöst werden können. Allgemeinere Funktionen  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  treten aber in einer Vielzahl von Anwendungen auf, z.B. bei den Newtonschen Gesetzen für die Dynamik von Teilchen. Die entstehenden Systeme sind dann auch beliebig groß, z.B. in der Molekulardynamik, wo  $u(t)$  die Koordinaten  $M$  verschiedener Teilchen im Zeitverlauf beschreibt ( $N = 3M$ ). Andere klassische Anwendungsgebiete gewöhnlicher Differentialgleichungen sind die Populationsdynamik oder auch die Modellierung von Aktienmärkten, wo meist noch eine zufällige Komponente hinzugefügt wird (stochastische Differentialgleichungen). Die numerischen Verfahren zur Lösung von Anfangswertproblemen sind einerseits ähnlich zu Iterationsverfahren, wenn die Ableitung durch Differenzenquotienten auf einem Gitter approximiert wird, andererseits zu numerischer Integration, wenn man die äquivalente Formulierung als Volterra-Integralgleichung

$$u(t) = u_0 + \int_0^t F(u(s), s) \, ds$$

benutzt und Quadraturformeln auf das Integral anwendet. Ein wichtiger Aspekt ist dabei die Diskretisierung, d.h. wir approximieren das Problem in einem endlich-dimensionalen Lösungsraum, z.B. durch Werte auf einem Gitter. Mathematisch stellt sich dann natürlich die Frage ob und in welchem Sinne das diskretisierte Problem gegen das ursprüngliche konvergiert.

Weiterhin werden wir auch Randwertprobleme betrachten, die zu partiellen Differentialgleichungen führen. Ein einfaches Beispiel ist die Lösung von

$$-(a(x)u'(x))' + c(x)u(x) = f(x), \quad x \in (0, 1)$$

mit Randwerten  $u(0) = u(1) = 0$ . Hier müssen wir die Diskretisierung auf einmal im ganzen Intervall  $(0, 1)$  durchführen und nicht von einem Schritt zum Nächsten wie bei Anfangswertproblemen. Die Diskretisierung liefert hier ein lineares Gleichungssystem, das wir anschließend lösen müssen. Die Abschätzung des Diskretisierungsfehlers erfordert dann weiterführende Methoden.

Abschließend beschäftigen wir uns mit Optimierungsproblemen, welche in vielen mathematischen Anwendungsbereichen auftreten, von klassischen ökonomischen Problemen über Materialoptimierung bis hin zu modernen Problemen in der Bildverarbeitung und im maschinellen Lernen. Methodisch knüpfen wir im Teil zur Optimierung an die iterativen Methoden zur Lösung von Gleichungssystemen an, allerdings kommen hier noch einige Aspekte dazu: Mit einem Optimierungsproblem im Hintergrund können wir die Iterationsverfahren anpassen um tatsächlich durch die Iteration die Funktionswerte zu verkleinern. Darüber hinaus werden wir geeignete Wahlen der Schrittweite kennenlernen, um besser die Konvergenz gegen Minimierer oder zumindest stationäre Punkte gewährleisten zu können. Ein weiterer Aspekt ist die Optimierung unter Nebenbedingung und die Minimierung konvexer nicht-differenzierbarer Probleme, die in vielen modernen Anwendungen auftreten. Dazu werden wir exemplarisch ein Verfahren kennenlernen.

---

## Kapitel 2

# Numerische Lösung von Anfangswertproblemen

---

Im Folgenden wollen wir uns mit der numerischen Lösungen von Anfangswertproblemen für gewöhnliche Differentialgleichungen der Form

$$u'(t) = \frac{du}{dt} = F(t, u(t)), \quad u(0) = u_0 \quad (2.1)$$

beschäftigen, wobei  $F : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine gegebene stetige Funktion ist. Die Modellierung mit Differentialgleichungen ist oft kanonisch, da man nur verstehen muss wie sich eine Größe in hinreichend kleiner Zeit ändern wird, d.h. wir schreiben

$$u(t + \Delta t) \approx u(t) + \Delta t F(t, u(t)),$$

wobei der Unterschied zur Exaktheit in dieser Relation dann von höherer Ordnung in  $\Delta t$  sein kann. Im Grenzwert  $\Delta t \rightarrow 0$  erhalten wir dann die Differentialgleichungen. Einfache Heuristiken für Differentialgleichungen kennen wir beispielsweise aus der Physik,

- Geschwindigkeit ist Änderung des Orts pro Zeit,
- Beschleunigung ist Änderung der Geschwindigkeit pro Zeit,
- Kraft ist Masse mal Beschleunigung.

Beschreibt  $x(t)$  den Ort eines Teilchens zur Zeit  $t$ ,  $v(t)$  seine Geschwindigkeit und  $a(t)$  die Beschleunigung, dann gilt

$$v(t) = \frac{dx}{dt}(t), \quad a(t) = \frac{dv}{dt}(t), \quad ma(t) = F(x(t), v(t), t),$$

wobei  $F$  ein Kraftfeld ist. Diese Gesetze können wir auch als Differentialgleichungen für  $x$  und  $v$  (oder den Impuls  $p(t) = mv(t)$ ) lesen, wenn wir  $a$  eliminieren, es gilt dann

$$\frac{d}{dt}(x(t), v(t)) = (v(t), F(x(t), v(t), t)).$$

Kennen wir den Anfangsort  $x(0)$  und die Anfangsgeschwindigkeit  $v(0)$ , so haben wir ein Anfangswert für ein System von Differentialgleichungen (im  $\mathbb{R}^3$  insgesamt sechs Gleichungen für sechs Unbekannte).

In großen Systemen von  $N$  Teilchen  $x_1, \dots, x_N$  hat man dann Gleichungen für jede Position und die Geschwindigkeiten  $v_1, \dots, v_N$ , mit

$$\frac{d}{dt}(x_i(t), v_i(t))_{i=1, \dots, N} = (v_i(t), F(x_1(t), v_1(t), \dots, x_N(t), v_N(t), t)).$$

Zur Beschreibung realer Vorgänge mit vielen Teilchen (Moleküle, Zellen, Tiere in Herden, Fußgänger, Autos ...) erhält man also schnell beliebig komplexe Systeme von Differentialgleichungen. Unser Ziel in diesem Kapitel ist es die numerische Lösung solcher Differentialgleichungen zu untersuchen, zuvor klären wir aber noch einige Grundlagen der gewöhnlichen Differentialgleichungen um diese auch vernünftig verstehen zu können.

## 2.1 Theorie von Anfangswertproblemen für gewöhnliche Differentialgleichungen

Wir diskutieren im Folgenden kurz einige theoretische Aspekte gewöhnlicher Differentialgleichungen. Wir beginnen mit einer allgemeinen Theorie der Existenz und Eindeutigkeit. Grundlage dafür ist eine Umformulierung in eine Fixpunktform, sodass wir dann einfach einen passenden Fixpunktsatz anwenden können. Ist  $u \in C^1([0, T])$  eine Lösung des Anfangswertproblems (2.1), dann gilt aus dem Hauptsatz der Integralrechnung auch

$$u(t) = u_0 + \int_0^t F(s, u(s)) \, ds, \quad 0 \leq s \leq T.$$

Dies können wir als Fixpunktgleichung  $u = \mathcal{F}(u)$  in einem Banachraum interpretieren. Dafür können wir zwei Arten von Fixpunktsätzen anwenden: die erste Art (Satz von Browder, Schauder oder andere Varianten) basiert auf Kompaktheit, d.h. wenn man Funktionen  $u$  in den Operator reinsteckt sind diese topologisch danach in einer schöneren Menge. Insbesondere hat diese Menge dann einen Fixpunkt. In unserem Fall entsteht die Kompaktheit aus dem Satz von Arzela–Ascoli, man kann zeigen, dass für  $u_n$  beschränkt die Folge  $\mathcal{F}(u_n)$  immer eine konvergente Teilfolge hat. Dies liefert den sogenannten Satz von Peano, der die Existenz einer Lösung für stetiges  $F$  garantiert. Da man hier recht abstrakt und über Teilfolgen argumentiert, hat man keine Chance die Eindeutigkeit eines Fixpunkts nachzuweisen.

Die zweite Art an Beweisen basiert eigentlich immer auf dem Banach'schen Fixpunktsatz, den wir hier näher diskutieren wollen. Dazu beachten wir, dass falls  $F$  im zweiten Argument Lipschitz-stetig ist (mit Modul  $L$ ), folgendes gilt

$$\begin{aligned}
 \|\mathcal{F}(u_1) - \mathcal{F}(u_2)\|_\infty &= \max_{0 \leq t \leq T} \left| \int_0^t F(s, u_1(s)) - F(s, u_2(s)) \, ds \right| \\
 &\leq \max_{0 \leq t \leq T} \int_0^t |F(s, u_1(s)) - F(s, u_2(s))| \, ds \\
 &\leq LT \max_{0 \leq s \leq T} |u_1(s) - u_2(s)| = LT \|u_1 - u_2\|_\infty.
 \end{aligned}$$

Wir erkennen daraus, dass die Abbildung  $\mathcal{F} : C([0, T]) \rightarrow C([0, T])$  kontraktiv ist, wenn  $T$  klein genug ist, da dann immer  $LT < 1$  ist. Damit liefert der Banach'sche Fixpunktsatz die Existenz und Eindeutigkeit einer Lösung in  $C([0, T])$  für  $T$  hinreichend klein. Da  $u$  dann die Stammfunktion der stetigen Funktion  $t \mapsto F(t, u(t))$  ist, gilt auch  $u \in C^1([0, T])$ . Dies ist die erste Version des Satzes von Picard-Lindelöf, die uns die Existenz und Eindeutigkeit für kleine Zeiten liefert. Wir sehen dabei schon einige typische Techniken bei der Behandlung von gewöhnlichen und partiellen Differentialgleichungen: erst formulieren wir das Problem um, sodass weniger oder keine Ableitungen mehr vorkommen und zeigen die Existenz / Eindeutigkeit einer Lösung in einem größeren Raum (hier die stetigen Funktionen). Danach beweisen wir zusätzliche Regularität der Lösung, in unserem Fall stetige Differenzierbarkeit. Wir beachten, dass wenn  $F$   $k$ -mal stetig differenzierbar in beiden Variablen ist, eine Iteration des obigen Arguments sogar liefert, dass  $u$   $k + 1$ -mal stetig differenzierbar ist. Auch die Analyse in kleinen Zeiten ist ein typisches Vorgehen für Anfangswertprobleme.

In unserem Fall können wir aber ein besseres Resultat erreichen, in dem wir einfach die Norm passend wählen. Die Idee dazu liefert zunächst die einfache Gleichung

$$u'(t) = Lu(t)$$

mit  $L > 0$ . Ist  $u$  eine positive Lösung, dann folgt mit der Kettenregel

$$\frac{d}{dt} \log u(t) = L.$$

Dies können wir integrieren zu

$$\log u(t) - \log u(0) = Lt$$

und auflösen als

$$u(t) = u_0 e^{Lt}.$$

In diesem Fall ist trivialerweise  $L$  die Lipschitzkonstante von  $F$  und wir sehen, dass wir dann ein exponentielles Wachstum mit  $e^{Lt}$  erwarten müssen. Dies ist auch allgemein der Fall, wie das folgende Lemma von Gronwall zeigt:



**LEMMA 2.1.**

Sei  $v(t)$  eine nichtnegative stetige Funktion, die

$$v(t) \leq a + \int_0^t bv(s) \, ds, \quad \forall 0 \leq t \leq T$$

mit  $a > 0$  und  $b \in \mathbb{R}$  erfüllt. Dann gilt

$$v(t) \leq ae^{bt}, \quad \forall 0 \leq t \leq T.$$

*Beweis.* Wir definieren  $w(t) = e^{-bt}v(t) - a$ , dann gilt

$$w(t) \leq a(e^{-bt} - 1) + \int_0^t e^{b(s-t)}b(w(s) + a) \, ds = \int_0^t bw(s) \, ds.$$

Aus der Ungleichung bei  $t = 0$  folgt  $w(0) \leq 0$ . Sei  $T_0$  die maximale Zeit zu der  $w(t) \leq 0$  für alle  $t \leq T_0$  gilt. Ist  $T_0 = T$ , so sind wir fertig. Ist  $T_0 < T$ , so gibt es ein hinreichend kleines Zeitintervall  $(T_0, T_0 + \delta)$ , in dem  $w$  positiv ist. Dann folgt für  $t$  in diesem Intervall

$$w(t) \leq \int_{T_0}^t bw(s) \, ds \leq \delta b \max_{T_0 \leq s \leq T_0 + \delta} w(s)$$

und damit auch

$$\max_{T_0 \leq t \leq T_0 + \delta} w(t) \leq \delta b \max_{T_0 \leq s \leq T_0 + \delta} w(s).$$

Für  $\delta b < 1$  ist dies aber ein Widerspruch zur Positivität von  $w$ . Also muss  $w(t) \leq 0$  und damit  $u(t) \leq ae^{bt}$  für alle  $t$  gelten.  $\square$

Wenden wir das Ergebnis auf eine Differentialgleichung mit Lipschitz-stetigem  $F$  an, so folgt

$$|u(t) - u_0| \leq \int_0^t |F(t, u(t)) - F(t, u_0)| \, dt + T \max_{0 \leq t \leq T} |F(t, u_0)| \leq L \int_0^t |u(t) - u_0| \, dt + a.$$

Aus dem Lemma von Gronwall angewandt auf  $v(t) = |u(t) - u_0|$  und der Dreiecksungleichung sehen wir, dass  $u(t)$  höchstens wie  $e^{Lt}$  wächst. Dies legt nahe, die folgende gewichtete Norm zu wählen:

$$\|u\|_{\infty, L} := \max_{0 \leq t \leq T} e^{-Lt} |u(t)|.$$

Da  $e^{-Lt}$  nach oben durch eins und nach unten durch  $e^{-LT}$  beschränkt ist, ist dies eine äquivalente Norm im Raum der stetigen Funktionen. Wir wiederholen also unsere Abschätzung an den Fixpunktoperator in dieser Norm

$$\begin{aligned}
 \|\mathcal{F}(u_1) - \mathcal{F}(u_2)\|_{L,\infty} &= \max_{0 \leq t \leq T} e^{-Lt} \left| \int_0^t F(s, u_1(s)) - F(s, u_2(s)) \, ds \right| \\
 &\leq \max_{0 \leq t \leq T} \int_0^t e^{L(s-t)} e^{-Ls} |F(s, u_1(s)) - F(s, u_2(s))| \, ds \\
 &\leq \int_0^T L e^{-L\tau} \, d\tau \max_{0 \leq s \leq T} e^{-Ls} |u_1(s) - u_2(s)| = (1 - e^{-LT}) \|u_1 - u_2\|_{L,\infty}.
 \end{aligned}$$

Der Operator ist nun also kontraktiv für beliebiges  $T$ , da  $1 - e^{-LT}$  gilt. Wir haben mit dem Banach'schen Fixpunktsatz also die folgende Version des Satzes von Picard-Lindelöf bewiesen:

**THEOREM 2.2.**

Sei  $F$  stetig und Lipschitzstetig bezüglich der zweiten Variable, dann besitzt das Anfangswertproblem (2.1) genau eine Lösung in  $C^1([0, T])$ .

Wir werden sehen, dass wir auch bei numerischen Verfahren ähnliche Aussagen und insbesondere ein Version des Lemma von Gronwall benötigen werden, um die Stabilität der Verfahren garantieren zu können. Abstrakt gesehen liefert das Lemma von Gronwall eine Stabilitätsaussage für die Differentialgleichung, in endlicher Zeit kann die Norm der Lösung nicht beliebig schnell wachsen.

Wir betrachten zum Abschluss noch einige spezielle Fälle von gewöhnlichen Differentialgleichungen in denen wir eine explizitere Form der Lösung berechnen können. Die Integration der einfachen linearen Gleichung oben ist ein Spezialfall sogenannter separabler Gleichungen der Form

$$u'(t) = G(u(t))H(t).$$

Für skalare Gleichungen, d.h.  $u(t) \in \mathbb{R}$ , können wir durch  $G$  dividieren (vorausgesetzt dieser Term ist ungleich null) und es gilt

$$\frac{u'}{G(u)} = H(t).$$

Ist  $g$  eine Stammfunktion von  $\frac{1}{G}$  und  $h$  eine Stammfunktion von  $H$ , so schreiben wir das als  $g'(u)u' = h'$  und integrieren zu  $g(u) = h(t) + c$ . Die Integrationskonstante  $c$  können wir aus dem Anfangswert mit  $c = g(u_0) - h(0)$  berechnen. Damit gilt, vorausgesetzt  $g$  ist invertierbar

$$u(t) = g^{-1}(h(t) - h(0) + g(u_0)).$$

Ein wichtiger Fall, der uns auch kanonische Beispiele für numerische Verfahren liefert, sind lineare Gleichungen mit konstanten Koeffizienten, d.h.

$$u'(t) = Au(t) + b(t)$$

mit einer gegebenen Matrix  $A \in \mathbb{R}^{n \times n}$ . Wir betrachten zunächst den homogenen Fall  $b = 0$ . Ist  $A$  diagonalisierbar als  $A = B^{-1}DB$  mit Diagonalmatrix  $D$ , so können wir analog eine Gleichung für  $v = Bu$  betrachten, die dann von der Form  $v'(t) = Dv(t)$ , d.h. jeder Eintrag erfüllt  $v_i'(t) = D_{ii}v_i(t)$  und wir erhalten daraus  $v_i(t) = v_i(0)e^{D_{ii}t}$ . Die Lösung  $u$  erhalten wir wieder durch Multiplikation mit  $B^{-1}$ . Im Fall einer Diagonalmatrix ist es naheliegend die Exponentialfunktion einer Matrix  $e^{Dt}$  als die Diagonalmatrix mit den Einträgen  $e^{D_{ii}t}$  zu definieren. Wir haben dann

$$u(t) = B^{-1}e^{Dt}v(0) = B^{-1}e^{Dt}Bu(0) =: e^{At}u(t).$$

Wir bekommen durch Diagonalisieren der Matrix also eine Definition des Matrixexponentials, die dann eine Lösung des Anfangswertproblems liefert. Im Fall einer nicht diagonalisierbaren Matrix ist ein ähnliches Vorgehen über die Jordan'sche Normalform möglich, was hier aber zu weit führen würde.

Ist  $b \neq 0$ , so können wir die sogenannte Variation der Konstanten benutzen um eine Lösung auszurechnen. Die Idee dabei ist ein Produktansatz  $u(t) = e^{At}w(t)$ , anstatt der Konstanten in der homogenen Lösung haben wir also jetzt eine Funktion. Dann gilt mit Produktregel

$$u'(t) = Ae^{At}w(t) + e^{At}w'(t) = Au(t) + e^{At}w'(t),$$

der Vergleich mit der rechten Seite der Differentialgleichung liefert  $e^{At}w'(t) = b(t)$  bzw.  $w'(t) = e^{-At}b(t)$ . Um die Lösung zu erhalten müssen wir also nur  $e^{-At}b(t)$  aufintegrieren.

Zum Abschluss betrachten wir noch eine Klasse von Gleichungen, die wir aus dem Gradientenverfahren der Optimierung erhalten, wenn die Schrittweite gegen Null geht. Interpretieren wir die Iterierten  $x^k$  als Wert einer Funktion  $u$  zur Zeit  $t$ , dann schreiben wir das Verfahren als

$$u(t + \alpha^k) = u(t) - \alpha^k \nabla F(u(t)),$$

die Differentialgleichung im Grenzwert ist der sogenannte Gradientenfluss

$$u'(t) = -\nabla G(u(t)).$$

Dieser hat natürlich immer noch eine Abstiegseigenschaft, es gilt

$$G(u)' = \nabla G(u)u' = -\|\nabla G(u)\|^2 = -\|u'\|^2.$$

Damit haben wir immer eine Funktion von  $u$ , die sogar gleichmäßig in der Zeit beschränkt ist, und nicht exponentiell wächst wie die Norm im schlimmsten Fall. Ist  $G$  konvex, dann gilt für einen Minimierer  $u^*$  sogar

$$\frac{d}{dt}\|u - u^*\|^2 = -2\langle \nabla G(u) - \nabla G(u^*), u - u^* \rangle \leq 0,$$

d.h. die Norm von  $u$  ist gleichmäßig beschränkt.

---

# Literatur

---