

Kapitel 5

Gewöhnliche Differentialgleichungen

In diesem Kapitel werden wir uns mit Anfangswertaufgaben bei gewöhnlichen Differentialgleichungen beschäftigen. Statt am Anfang genau zu erklären, was das ist, geben wir einige Beispiele von gewöhnlichen Differentialgleichungen an. Die Visualisierungsmöglichkeiten, symbolische und numerische Löser von Maple werden ausgenutzt. Anschließend gehen wir auf wichtige theoretische Fragestellungen für Anfangswertaufgaben bei gewöhnlichen Differentialgleichungen (Existenz, Eindeutigkeit, lineare Systeme) ein und beschäftigen uns am Schluss sehr kurz mit ihrer numerischen Lösung durch die einfachste Klasse von Verfahren, nämlich Einschrittverfahren.

5.1 Beispiele

5.1.1 Populationsmodelle

In einem Populationsmodell bezeichne $p(t)$ die Population einer gewissen Spezies (etwa der Menschen auf der Erde oder der Lachse in der Weser) zur Zeit t . Geht man wie T. R. Malthus (1766-1834) von einer konstanten Geburtenrate γ und Sterberate δ pro Kopf der Bevölkerung und Zeiteinheit aus und nimmt man an, dass sich die Bevölkerungszahl innerhalb eines Zeitraum Δt von t bis $t + \Delta t$ gemäß der Formel

$$p(t + \Delta t) = p(t) + \lambda p(t) \Delta t$$

verändert, wobei wir

$$\lambda := \gamma - \delta$$

gesetzt haben, so erhält man aus

$$\frac{p(t + \Delta t) - p(t)}{\Delta t} = \lambda p(t)$$

durch den Grenzübergang $\Delta t \rightarrow 0$ die *Differentialgleichung*

$$\frac{dp}{dt}(t) = \lambda p(t)$$

mit der eindeutigen Lösung $p(t) = p_0 e^{\lambda(t-t_0)}$, wenn man noch $p(t_0) = p_0$ vorgibt¹. Wenn man die Zeit T kennt, in der sich die Bevölkerung verdoppelt, so kann man unter Zugrundelegung des Malthus-Wachstumsmodells auf die Wachstumsrate λ schließen. Denn aus

$$p_0 e^{\lambda(t+T-t_0)} = p(t+T) = 2p(t) = 2p_0 e^{\lambda(t-t_0)}$$

erhält man

$$\lambda = \frac{\log 2}{T}.$$

Beispiel: Bei M. Braun (1983)² wird bemerkt, dass die Weltbevölkerung 1961 mit 3 060 000 000 geschätzt wurde. Geht man von einer Wachstumsrate von 2%/Jahr (bzw. einer Verdoppelung der Weltbevölkerung alle 35 Jahre) aus, so erhalte man unter Zugrundelegung des Malthus-Ansatzes

$$p(t) = (3.06)10^9 e^{0.02(t-1961)}.$$

Dies scheint mit den Zahlen der Jahre 1700 bis etwa 1960 gut im Einklang zu sein. \square

In der fernen Zukunft ist das Malthus-Modell unrealistisch, da bei großen Populationen die Wachstumsrate von der Populationsgröße abhängen wird. Die nächst einfache Annahme ist, dass die Wachstumsrate linear von der Populationsgröße abhängt und kleiner wird, wenn die Population wächst. Dies führt auf das sogenannte Verhulst-Modell, genannt nach P. F. Verhulst (1804-1849). Hiernach berechnet sich die Population p als Lösung der Anfangswertaufgabe

$$p' = ap - bp^2, \quad p(t_0) = p_0.$$

Hierbei ist $p_0 > 0$ und $0 < b \ll a$, so dass bei kleinen Populationen bp^2 gegenüber ap vernachlässigt werden kann. Diese (nichtlineare) Differentialgleichung (man nennt sie auch Gleichung des beschränkten Wachstums oder logistische Differentialgleichung) mit Anfangsbedingung (man spricht dann, wie wir es schon getan haben, von einer *Anfangswertaufgabe*) kann man geschlossen lösen, und zwar ist durch

$$p(t) = \frac{ap_0}{bp_0 + (a - bp_0) \exp[-a(t - t_0)]}$$

¹Dass es sich bei $p(t) = p_0 e^{\lambda(t-t_0)}$ um eine Lösung der Differentialgleichung $p' = \lambda p$ mit der Anfangsbedingung $p(t_0) = p_0$ handelt, ist klar. Weshalb aber ist es die einzige? Sei hierzu p eine Lösung von $p' = \lambda p$. Dann ist

$$p'(t)e^{-\lambda(t-t_0)} = \lambda e^{-\lambda(t-t_0)} p(t)$$

bzw.

$$\frac{d}{dt}[e^{-\lambda(t-t_0)} p(t)] = 0.$$

Folglich ist $e^{-\lambda(t-t_0)} p(t)$ konstant. Für eine Lösung p von $p' = \lambda p$, welche der Anfangsbedingung $p(t_0) = p_0$ genügt, ist diese Konstante gleich p_0 , so dass also $p(t) = p_0 e^{\lambda(t-t_0)}$ notwendigerweise die einzige Lösung der gestellten Anfangswertaufgabe ist.

²M. BRAUN (1983) "Single species population models". In: *Differential Equation Models* (eds. M. Braun et al.), Springer-Verlag, New York-Heidelberg-Berlin.

die Lösung der gestellten Anfangswertaufgabe gegeben. Durch Einsetzen kann man leicht nachprüfen, dass dies wirklich eine Lösung ist.

Beispiel: Will man die Anfangswertaufgabe

$$p' = ap - bp^2, \quad p(t_0) = p_0$$

mit Maple lösen, zunächst für allgemeine Daten a, b, t_0, p_0 , danach für spezielle Daten, so kann dies mittels

```
> ode:=diff(p(t),t)=a*p(t)-b*p(t)^2;
```

$$ode := \frac{\partial}{\partial t} p(t) = a p(t) - b p(t)^2$$

```
> initial:=p(t_0)=p_0;
```

$$initial := p(t_0) = p_0$$

```
> sol:=dsolve({ode,initial},p(t));
```

$$sol := p(t) = \frac{a}{b + \frac{e^{(-at)}(a - p_0 b)}{p_0 e^{(-at_0)}}$$

```
> simplify(%);
```

$$p(t) = \frac{a p_0}{p_0 b + e^{(-a(t-t_0))} a - e^{(-a(t-t_0))} p_0 b}$$

```
> par:={a=100,b=0.1,t_0=0,p_0=10};
```

$$par := \{a = 100, b = .1, t_0 = 0, p_0 = 10\}$$

```
> solspecial:=dsolve(subs(par,{ode,initial}),p(t));
```

$$solspecial := p(t) = 1000 \frac{1}{1 + 99 e^{(-100t)}}$$

geschehen. □

Bemerkenswert ist die Folgerung, die wir aus der angegebenen Lösung der logistischen Differentialgleichung ziehen können. Es ist nämlich $\lim_{t \rightarrow \infty} p(t) = a/b$, die Gesamtpopulation strebt also mit wachsender Zeit der Grenzpopulation $\xi := a/b$ zu. Dies geschieht monoton fallend, wenn $p_0 > \xi$ und monoton wachsend, wenn $0 < p_0 < \xi$, wobei der letztere Fall sicher der interessantere ist. In Abbildung 5.1 zeigen wir die typische S-Form der Lösung der logistischen Differentialgleichung. Den Plot haben wir in Maple durch

```
plot(rhs(solspecial),t=0..0.1,title="Logistisches Wachstum");
```

erzeugt. subsectionDas Räuber-Beute-Modell Im letzten Unterabschnitt wurden mathematische Modelle zum Wachstum einer einzigen Population aufgestellt, Wechselwirkungen mit anderen Species wurden ebenso vernachlässigt wie “Ein- oder Auswanderungen”. Nun wollen wir das Wachstum von zwei Arten untersuchen, die sich gegenseitig beeinflussen und die wir Räuber und Beute (engl.: predator, prey) nennen wollen. Man stelle sich etwa Raub- und Beutfische in der Adria vor. Dies war der Ausgangspunkt für die Untersuchungen von V. Volterra (1860-1940) und A. J. Lotka (1880-1949). Sei $x(t)$ die Population der Beute, $y(t)$ die Population der Räuber zur Zeit t . Falls genügend Nahrung für die Beute vorhanden ist, so dass sich diese nicht gegenseitig das Futter wegzunehmen brauchen, und keine Räuber vorhanden sind, ist $x' = ax$ mit einer positiven Konstante a , wenn vom Malthus-Modell ausgegangen wird. Andererseits

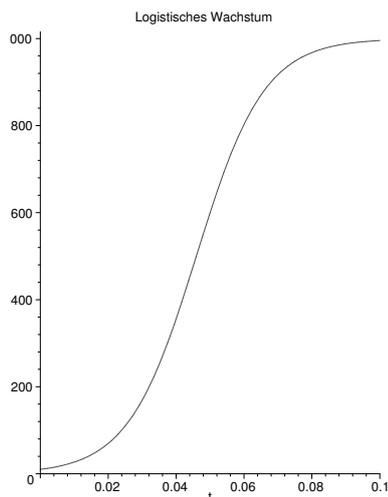


Abbildung 5.1: Eine Lösung der logistischen Differentialgleichung

ist die Anzahl der “Kontakte” zwischen Räuber und Beute proportional zu xy , so dass $x' = ax - bxy$ mit einer positiven Konstanten b . Ist dagegen keine Beute vorhanden, so sterben die Räuber aus: $y' = -cy$. Andererseits ist die Zuwachsrates proportional zu xy , so dass die zeitliche Änderung der Räuberpopulation durch $y' = -cy + dxy$ gegeben ist. Insgesamt erhält man ein System von zwei Differentialgleichungen erster Ordnung, das sogenannte Lotka-Volterra-System:

$$\begin{aligned}x' &= ax - bxy, \\y' &= -cy + dxy,\end{aligned}$$

wobei a, b, c, d positive Konstanten sind. Sind positive Anfangspopulationen x_0, y_0 zur Zeit $t = 0$ vorgegeben, so können die zukünftigen Räuber- bzw. Beute-Populationen $x(t)$ bzw. $y(t)$ durch (numerisches) Lösen der Anfangswertaufgabe

$$\begin{aligned}x' &= ax - bxy, & x(0) &= x_0, \\y' &= -cy + dxy, & y(0) &= y_0\end{aligned}$$

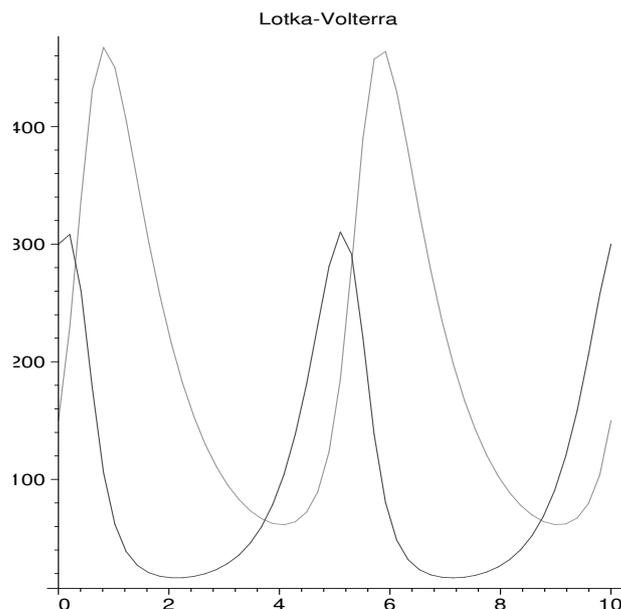
bestimmt werden.

Beispiel: Wir wollen die Lösung von

$$(*) \quad \begin{aligned}x' &= 2x - 0.01xy, & x(0) &= 300, \\y' &= -y + 0.01xy, & y(0) &= 150\end{aligned}$$

veranschaulichen. Diese Anfangswertaufgabe für ein Differentialgleichungssystem von zwei Differentialgleichungen erster Ordnung ist nicht geschlossen lösbar, so dass wir auf numerische Methoden angewiesen sind. Wir geben nach dem Maple-Prompt jeweils ein:

```
par:=a=2,b=0.01,c=1,d=0.01,x_0=300,y_0=150;
eqn:=diff(x(t),t)=a*x(t)-b*x(t)*y(t),diff(y(t),t)=-c*y(t)+d*x(t)*y(t);
initial:=x(0)=x_0,y(0)=y_0;
sol:=dsolve(subs(par,{eqn,initial}},{x(t),y(t)},type=numeric);
plots[odeplot](sol,[[t,x(t)],[t,y(t)]],0..10,title="Lotka-Volterra");
```

Abbildung 5.2: Die Population der Beute x und der Räuber y

und erhalten den Plot in Abbildung 5.2. Ganz erstaunlich ist nun, dass eine Konstante $T > 0$ existiert mit

$$(x(t), y(t)) = (x(t + T), y(t + T))$$

für alle t . D.h. die Populationen der Räuber und der Beute haben eine Periode T , so dass nach der Zeit T der Anfangszustand wieder erreicht wird. Anders gesagt: In der sogenannten Phasenebene, der (x, y) -Ebene, beschreibt $(x(\cdot), y(\cdot))$ eine geschlossene Bahn. Durch

```
with(plots);
odeplot(sol, [x(t), y(t)], 0..5, title="Geschlossene
Phasenbahn", labels=["Beute", "Raeuber"]);
```

wird der entsprechende Plot hergestellt, der in Abbildung 5.3 zu sehen ist. \square

Bemerkung: Sei $(x(t), y(t))$ eine T -periodische Lösung des Lotka-Volterra-Systems. Durch

$$\bar{x} := \frac{1}{T} \int_0^T x(t) dt, \quad \bar{y} := \frac{1}{T} \int_0^T y(t) dt$$

sind die mittleren Populationen der Beute bzw. der Räuber über das Periodenintervall $[0, T]$ gegeben. Es ist überraschend, dass man diese Mittelwerte berechnen kann, ohne die Populationen x bzw. y zu kennen. Denn wegen

$$\begin{aligned} 0 &= \frac{1}{T} [\log y(t) - \log y(0)] \\ &= \frac{1}{T} \int_0^T \frac{d}{dt} \log y(t) dt \\ &= \frac{1}{T} \int_0^T \frac{y'(t)}{y(t)} dt \end{aligned}$$

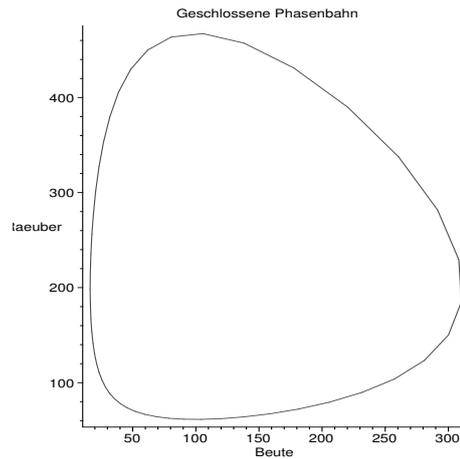


Abbildung 5.3: Beute und Räuber in der Phasenebene

$$\begin{aligned}
 &= \frac{1}{T} \int_0^T [-c + dx(t)] dt \\
 &= -c + d\bar{x},
 \end{aligned}$$

so dass $\bar{x} = c/d$. Entsprechend erhält man $\bar{y} = a/b$. Hieraus kann eine interessante Folgerung gezogen werden. Bei den Räubern und der Beute handele es sich jeweils um Fische. Es sollen Aussagen über die Auswirkung des Fischfangs auf die jeweiligen Populationen gemacht werden, wobei davon ausgegangen wird, dass die Fischer beim Fang zwischen Räubern und Beute machen können. Durch die Konstante $\epsilon \geq 0$ werde die "Intensität" des Fischfangs angegeben. Das modifizierte Differentialgleichungssystem lautet dann

$$\begin{aligned}
 x' &= (a - \epsilon)x - bxy, \\
 y' &= -(c + \epsilon)y + dxy.
 \end{aligned}$$

Ist nun $\epsilon \in (0, a)$, erhält man nach obiger Überlegung als mittlere Populationen

$$\bar{x} = \frac{c + \epsilon}{d}, \quad \bar{y} = \frac{a - \epsilon}{b}.$$

Ein nicht zu intensiver Fischfang erhöht also die mittlere Population der Beute und erniedrigt die der Räuber. \square

5.1.2 Das mathematische Pendel

Ein Massenpunkt³ M der Masse m sei durch eine masselose Stange der Länge l an einem festen Punkt drehbar aufgehängt. Von der Reibung im Aufhängepunkt und vom Luftwiderstand wird abgesehen. Auf M wirkt als bewegende Kraft also lediglich die Schwerkraft, genauer: ihre tangentielle Komponente $-mg \sin \phi$, wobei ϕ der Winkel zwischen der Pendelstange und der Vertikalen, g die Erdbeschleunigung ist. Eine solche

³Siehe z. B.

Vorrichtung heißt ein *mathematisches Pendel*. In Abbildung 5.4 haben wir versucht, dies zu verdeutlichen⁴. Da die Bogenlänge s vom Ruhe- oder Tiefstpunkt R bis zum

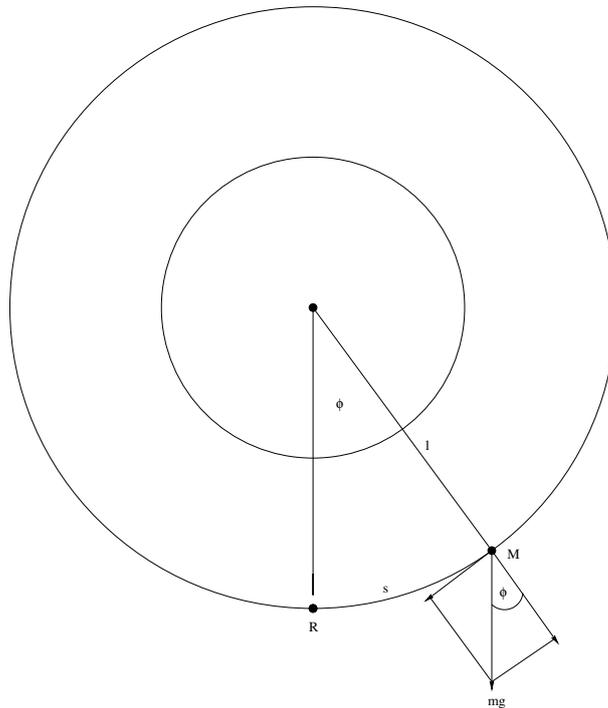


Abbildung 5.4: Das mathematische Pendel

Massenpunkt M gemessen gleich $l\phi$ ist, erhält man aus dem Newtonschen Bewegungsgesetz

$$\text{Kraft} = \text{Masse} \times \text{Beschleunigung},$$

das (zeitliche Ableitungen werden durch einen hochgestellten Punkt verdeutlicht)

$$m\ddot{s} = m l \ddot{\phi} = -m g \sin \phi.$$

Ist die Anfangsauslenkung ϕ_0 gegeben und befindet sich der Massenpunkt zur Anfangszeit im Ruhezustand, so hat man also die Anfangswertaufgabe

$$(*) \quad \ddot{\phi} + \frac{g}{l} \sin \phi = 0, \quad \phi(0) = \phi_0, \quad \dot{\phi}(0) = 0.$$

Für kleine Winkel ϕ ist $\sin \phi \approx \phi$, die exakte Pendelgleichung (*) geht also bei kleinen Ausschlägen über in die Näherungsgleichung (auch Gleichung des ungedämpften harmonischen Oszillators genannt)

$$(**) \quad \ddot{\phi} + \frac{g}{l} \phi = 0, \quad \phi(0) = \phi_0, \quad \dot{\phi}(0) = 0$$

⁴Die Zeichnung ist mit xfig hergestellt worden, was (zumindestens für mich noch) relativ kompliziert ist und nicht das gewünschte Resultat lieferte.

mit der Lösung $\phi(t) = \phi_0 \cos \omega_0 t$, wobei

$$\omega_0 := \sqrt{\frac{g}{l}}$$

gesetzt wurde. Diese Näherungslösung hat die (von der Pendelmass m und der Anfangsauslenkung ϕ_0 unabhängige) Periode $T_0 = 2\pi/\omega_0 = 2\pi\sqrt{l/g}$.

Beispiel: Für $\phi_0 := \pi/4$ (kein ganz kleiner Ausschlag mehr) und $\omega_0^2 = 4$ veranschaulichen wir die Lösungen von (*) und (**) in den Abbildungen 5.5 und 5.6. In 5.5 geben wir die Lösung ϕ und die zugehörige Bahn in der $(\phi, \dot{\phi})$ -Phasenebene an. Dagegen

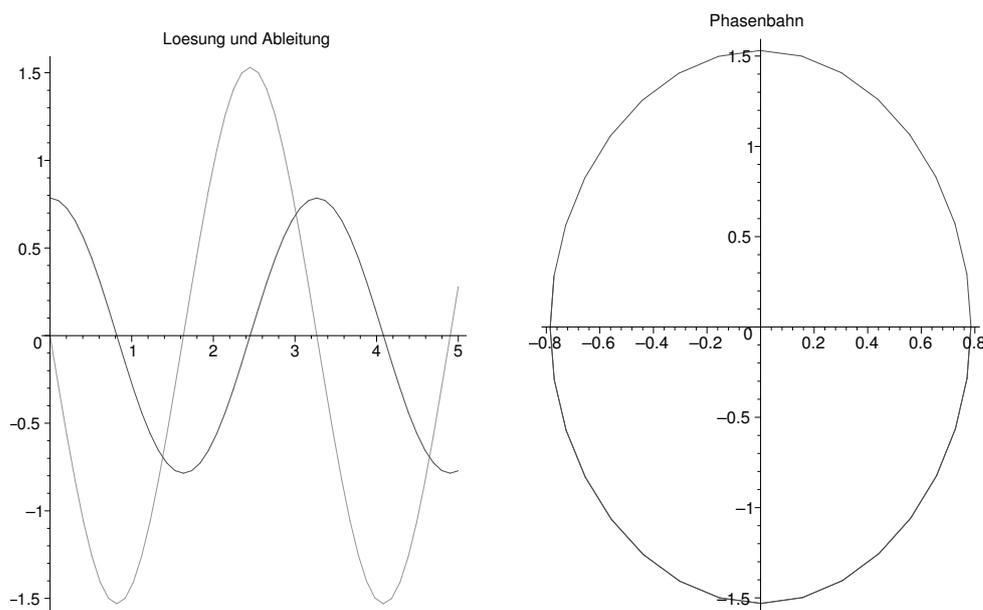


Abbildung 5.5: Mathematisches Pendel: Lösung und Bahn

findet man in Abbildung 5.6 die entsprechenden Plots für das linearisierte mathematische Pendel. Die Unterschiede sind offensichtlich nicht groß. Etwas überraschender (?) ist, dass auch das nichtlineare mathematische Pendel periodisch schwingt. In Kürze werden wir die Schwingungsdauer des mathematischen Pendels berechnen. Die beiden Abbildungen in 5.5 haben wir durch

```

eqn1:=diff(phi(t),t)=psi(t),diff(psi(t),t)=-4*sin(phi(t));
initial:=phi(0)=Pi/4;psi(0)=0;
sol1:=dsolve({eqn1,initial},{phi(t),psi(t)},type=numeric);
plots[odeplot](sol1,[[t,phi(t)],[t,psi(t)]],0..5,
    title="Loesung und Ableitung");
with(plots);
odeplot(sol1,[phi(t),psi(t)],0..5,title="Phasenbahn");

```

erhalten. Hierbei haben wir also die Anfangswertaufgabe zweiter Ordnung

$$(*) \quad \ddot{\phi} + \omega_0^2 \sin \phi = 0, \quad \phi(0) = \phi_0, \quad \dot{\phi}(0) = 0.$$

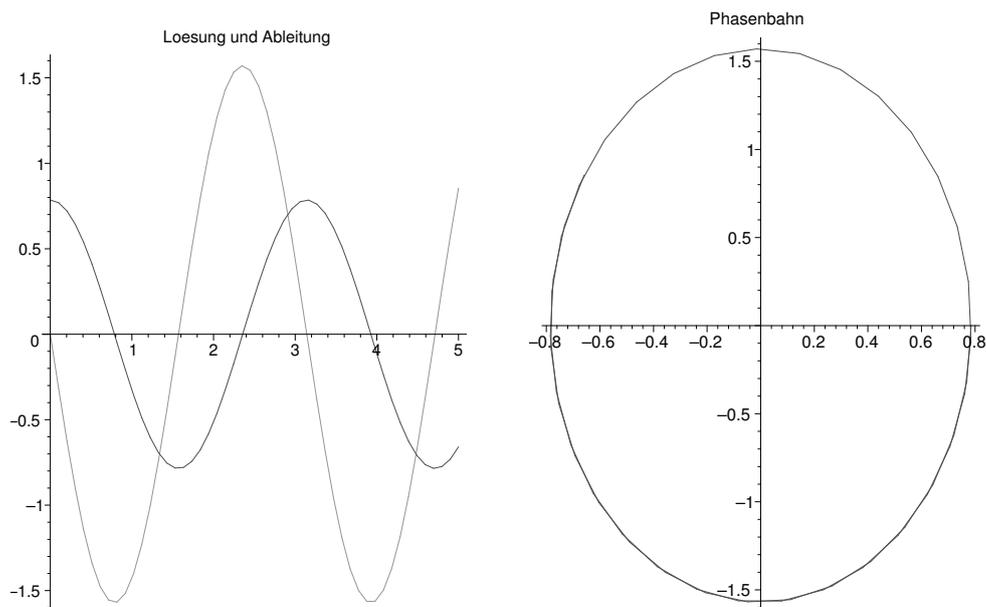


Abbildung 5.6: Linearisiertes mathematisches Pendel: Lösung und Bahn

in eine Anfangswertaufgabe für ein System von Differentialgleichungen erster Ordnung umformuliert:

$$\begin{aligned}\dot{\phi} &= \psi, & \phi(0) &= \phi_0, \\ \dot{\psi} &= -\omega_0^2 \sin \phi, & \psi(0) &= 0.\end{aligned}$$

Entsprechendes gilt natürlich für die Abbildung 5.6. □

Definiert man

$$E(\phi, \dot{\phi}) := \frac{1}{2} \dot{\phi}^2 + \omega_0^2 (1 - \cos \phi),$$

so stellt man mit einer Lösung ϕ von (*) fest, dass

$$\frac{d}{dt} E(\phi(t), \dot{\phi}(t)) = \dot{\phi}(t) [\ddot{\phi}(t) + \omega_0^2 \sin \phi(t)] = 0.$$

Die Niveaulinien $E(\phi, \dot{\phi}) = C$ findet man in Abbildung 5.7. Diese wurden durch

```
f:=phi->4*(1-cos(phi));g:=psi->(1/2)*psi^2;
contourplot(f(phi)+g(psi),phi=-2*Pi..2*Pi,psi=-Pi..Pi,filled=true);
contourplot(f(phi)+g(psi),phi=-2*Pi..2*Pi,psi=-Pi..Pi);
```

hergestellt. Daher ist

$$\frac{1}{2} \dot{\phi}(t)^2 + \omega_0^2 (1 - \cos \phi(t)) = \omega_0^2 (1 - \cos \phi_0)$$

bzw.

$$\dot{\phi}(t)^2 = 2\omega_0^2 [\cos \phi(t) - \cos \phi_0]$$

und daher

$$\dot{\phi}(t) = \pm \omega_0 \sqrt{2(\cos \phi(t) - \cos \phi_0)}$$

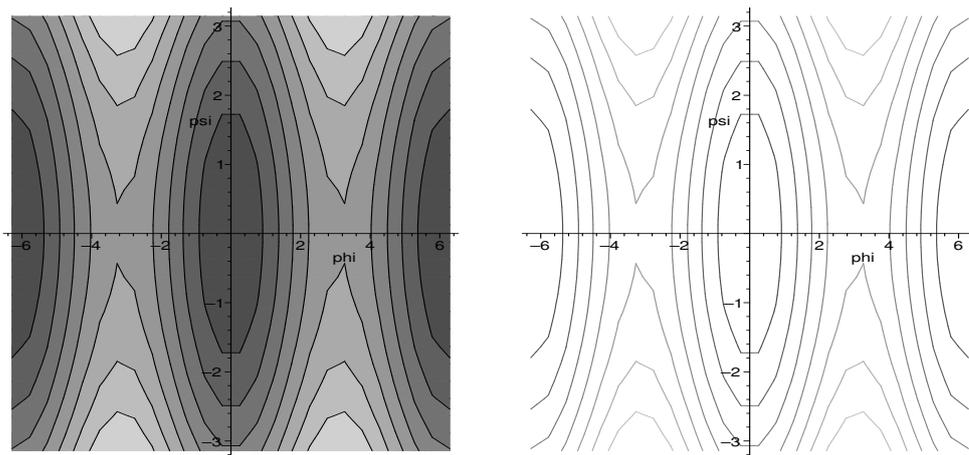


Abbildung 5.7: Niveaulinien zu $1 - \cos \phi + \frac{1}{2}\dot{\phi}^2 = C$

für alle t . Die Pendelschwingungsdauer des mathematischen Pendels (welche von dem Anfangsausschlag ϕ_0 abhängt) ist dann

$$T(\phi_0) = \frac{4}{\omega_0} \int_0^{\phi_0} \frac{d\phi}{\sqrt{2(\cos \phi - \cos \phi_0)}}.$$

Macht man hier die Substitution $\sin \frac{1}{2}\phi = k \sin \theta$ mit $k := \sin \frac{1}{2}\phi_0$, so erhält man unter Berücksichtigung von $\cos \phi = 1 - 2 \sin^2 \frac{1}{2}\phi$, dass

$$T(\phi_0) = \frac{4}{\omega_0} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = \frac{4}{\omega_0} K(k),$$

wobei

$$K(k) := \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2 t^2)}}$$

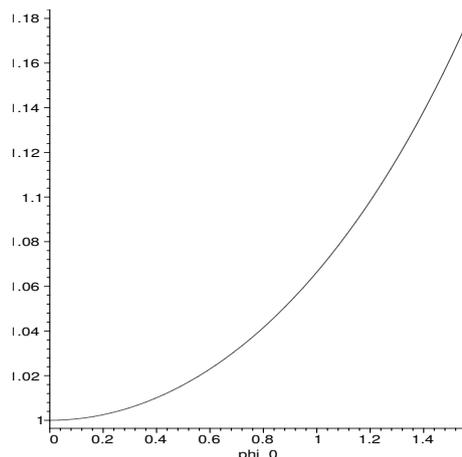
das *vollständige elliptische Integral erster Art* ist. Das Verhältnis zwischen der Schwingungsdauer des mathematischen Pendels und des linearisierten mathematischen Pendels ist also

$$\frac{T(\phi_0)}{T_0} = \frac{2K(\sin \frac{1}{2}\phi_0)}{\pi}.$$

Dieses Verhältnis wird in Abbildung 5.8 als Funktion von ϕ_0 aufgetragen. Den Plot haben wir hierbei durch

```
f:=phi_0->2*EllipticK(sin(0.5*phi_0))/Pi;
plot(f(phi_0),phi_0=0..Pi/2);
```

hergestellt.

Abbildung 5.8: Das Verhältnis $T(\phi_0)/T_0$

5.1.3 Planetenbahnen

Das Newtonsche Gravitationsgesetz sagt aus, dass zwei (punktförmige) Objekte aufeinander eine Kraft ausüben, deren Länge (die Kraft ist ein Vektor) linear jeweils von ihrer Masse und umgekehrt proportional vom Quadrat ihrer Entfernung abhängt. Im folgenden sei das eine Objekt die Sonne mit der Masse M , das andere Objekt ein gegebener Planet mit der Masse m . Der Planet bewegt sich in einer Ebene, als dessen Nullpunkt die (unbewegliche) Sonne genommen wird. Die Bewegungsgleichungen sind dann

$$(*) \quad m\ddot{x} = -\frac{\gamma mM}{\|x\|^3}x.$$

Hierbei ist $x(t) = (x_1(t), x_2(t))$ der Ort des Planeten zur Zeit t und γ die Gravitationskonstante ist. Die Bewegungsgleichung (*) ist also eigentlich ein System von zwei Differentialgleichungen zweiter Ordnung, nämlich

$$(*) \quad \begin{aligned} \ddot{x}_1 &= -\frac{\gamma M}{(x_1^2 + x_2^2)^{3/2}}x_1, \\ \ddot{x}_2 &= -\frac{\gamma M}{(x_1^2 + x_2^2)^{3/2}}x_2. \end{aligned}$$

Das Zweikörperproblem besteht darin, die Bahn des Planeten zu beschreiben. Die Lösung dieses Problems durch J. Kepler (1571-1630) gehört sicherlich zu den größten Leistungen in der Geschichte der Menschheit. Aus den Bewegungsgleichungen (*) wollen wir die drei Keplerschen Gesetze⁵ der Planetenbewegung ableiten⁶. Dies sind bekannt-

⁵Es gibt im Internet einige Seiten, auf denen animierte Erläuterungen der Keplerschen Gesetze zu sehen sind. Man sehe sich z. B. an:

<http://www.cvc.org/science/kepler.htm>

http://www.phy.syr.edu/courses/java/mc_html/kepler.html

<http://sunsite.ubc.ca/LivingMathematics/V001N01/UBCExamples/Kepler/kepler.html>

⁶Wir folgen

W. WALTER (1990) *Analysis II*. Springer-Verlag, Berlin-Heidelberg, New York.

lich:

(K1) Die Bahnen der Planeten sind Ellipsen, in deren einem Brennpunkt die Sonne steht.

(K2) Der von der Sonne zu einem Planeten weisende Radiusvektor überstreicht in gleichen Zeiten gleiche Flächen.

(K3) Das Verhältnis zwischen dem Quadrat der Umlaufzeit und dem Kubus der großen Achse (der Bahnellipse) ist für alle Planeten des Sonnensystems konstant.

Satz 3.1 Für Planeten mit der Bewegungsgleichung (*) gelten die drei Keplerschen Gesetze.

Beweis: Wir suchen eine Lösung der Anfangswertaufgabe

$$\begin{aligned} \ddot{x}_1 &= -\frac{\gamma M}{(x_1^2 + x_2^2)^{3/2}} x_1, & x_1(0) &= R, & \dot{x}_1(0) &= v_1, \\ \ddot{x}_2 &= -\frac{\gamma M}{(x_1^2 + x_2^2)^{3/2}} x_2, & x_2(0) &= 0, & \dot{x}_2(0) &= v_2, \end{aligned}$$

die sich darstellen lässt in der Form

$$x_1(t) = r(t) \cos \phi(t), \quad x_2(t) = r(t) \sin \phi(t),$$

wobei wir voraussetzen, dass $R > 0$ der Abstand des Planeten von der Sonne zur Zeit $t_0 = 0$ ist. Ferner wird $v_2 \neq 0$ vorausgesetzt (andernfalls wäre $x_2(t) = 0$, die Bewegung des Planeten würde auf einer Geraden erfolgen). Benutzt man die komplexe Schreibweise $z(t) = r(t)e^{i\phi(t)}$, so ist

$$\dot{z} = (\dot{r} + ir\dot{\phi})e^{i\phi}, \quad \ddot{z} = (\ddot{r} + 2i\dot{r}\dot{\phi} + ir\ddot{\phi} - r\dot{\phi}^2)e^{i\phi}.$$

Die Bewegungsgleichungen sind also äquivalent zu

$$\ddot{r} + 2i\dot{r}\dot{\phi} + ir\ddot{\phi} - r\dot{\phi}^2 = -\frac{\gamma M}{r^2}.$$

Zerlegt man in Real- und Imaginärteil, so ergeben sich die beiden folgenden, mit (*) äquivalenten Gleichungen:

$$(1) \quad \ddot{r} - r\dot{\phi}^2 + \frac{\gamma M}{r^2} = 0, \quad 2\dot{r}\dot{\phi} + r\ddot{\phi} = 0.$$

Die Anfangsbedingungen sind gegeben durch

$$(2) \quad r(0) = R, \quad \phi(0) = 0, \quad \dot{r}(0) = v_1, \quad \dot{\phi}(0) = \frac{v_2}{R}.$$

Also ist die gegebene Anfangswertaufgabe zu (1), (2) äquivalent.

Zum Nachweis des ersten Keplerschen Gesetzes gehen wir folgendermaßen vor. Zunächst definieren wir mit noch unbekanntenen Konstanten $\epsilon \geq 0$, $p > 0$ und $0 \leq \alpha < 2\pi$ die Funktion f durch

$$f(\phi) := \frac{p}{1 + \epsilon \cos(\phi - \alpha)}.$$

Anschließend sei

$$t(\phi) := \frac{1}{A} \int_0^\phi f^2(\phi) d\phi \quad \text{mit} \quad A := Rv_2,$$

also

$$\frac{dt(\phi)}{d\phi} = \frac{1}{A} f^2(\phi), \quad t(0) = 0.$$

Mit $\phi = \phi(t)$ sei die Umkehrfunktion zu $t = t(\phi)$ bezeichnet und $r(t) := f(\phi(t))$ gesetzt. Aus

$$t = t(\phi(t)) = \frac{1}{A} \int_0^{\phi(t)} f^2(\phi) d\phi$$

erhält man durch Differentiation nach t , dass

$$\dot{\phi}(t) = \frac{A}{f^2(\phi(t))}.$$

Wir wollen uns überlegen, dass bei geeigneter Wahl der noch freien Konstanten ϵ, p, α durch (r, ϕ) eine Lösung von (1) und (2) gegeben ist. Am einfachsten ist die zweite Gleichung in (1) einzusehen. Es ist nämlich

$$r^2(t) \dot{\phi}(t) = f^2(\phi(t)) \frac{A}{f^2(\phi(t))} = A,$$

insbesondere also auch

$$0 = \frac{d}{dt} [r^2(t) \dot{\phi}(t)] = r(t) [2\dot{r}(t) \dot{\phi}(t) + r(t) \ddot{\phi}(t)].$$

Wegen $r(t) > 0$ ist die zweite Gleichung in (1) erfüllt. Von den Anfangsbedingungen in (2) ist $\phi(0) = 0$ schon erfüllt (wegen $t(0) = 0$). Nun kommen wir zu der ersten Gleichung in (1). Zunächst folgt aus $r(t) := f(\phi(t))$, dass

$$\begin{aligned} \dot{r}(t) &= f'(\phi(t)) \dot{\phi}(t) \\ &= f'(\phi(t)) \frac{A}{f^2(\phi(t))} \\ &= \frac{p\epsilon \sin(\phi(t) - \alpha)}{(1 + \epsilon \cos(\phi(t) - \alpha))^2} \cdot \frac{A(1 + \epsilon \cos(\phi(t) - \alpha))^2}{p^2} \\ &= \frac{\epsilon A}{p} \sin(\phi(t) - \alpha), \end{aligned}$$

anschließend

$$\ddot{r}(t) = \frac{\epsilon A}{p} \cos(\phi(t) - \alpha) \dot{\phi}(t).$$

Folglich ist

$$\begin{aligned}
 r^2(t) \left[\ddot{r}(t) - r(t) \dot{\phi}^2(t) + \frac{\gamma M}{r^2(t)} \right] &= r^2(t) \left[\frac{\epsilon A}{p} \cos(\phi(t) - \alpha) \dot{\phi}(t) - \frac{A^2}{r^3(t)} + \frac{\gamma M}{r^2(t)} \right] \\
 &= \frac{\epsilon A^2}{p} \cos(\phi(t) - \alpha) - \frac{A^2}{r(t)} + \gamma M \\
 &= \frac{\epsilon A^2}{p} \cos(\phi(t) - \alpha) - \frac{A^2}{p} [1 + \epsilon \cos(\phi(t) - \alpha)] + \gamma M \\
 &= \gamma M - \frac{A^2}{p}.
 \end{aligned}$$

Setzt man also

$$p := \frac{A^2}{\gamma M} = \frac{R^2 v_2^2}{\gamma M},$$

so erfüllt (r, ϕ) auch die erste Gleichung in (1). Die Konstanten ϵ und α werden durch die Anfangsbedingungen $r(0) = R$ und $\dot{r}(0) = v_1$ festgelegt. Dies führt auf die Gleichungen

$$R = \frac{p}{1 + \epsilon \cos \alpha}, \quad v_1 = -\frac{\epsilon A}{p} \sin \alpha$$

bzw. nach Einsetzen von p auf

$$\epsilon \cos \alpha = \frac{Rv_2^2 - \gamma M}{\gamma M}, \quad \epsilon \sin \alpha = -\frac{Rv_1 v_2}{\gamma M}.$$

Diese beiden Gleichungen sind lösbar, da $\epsilon \geq 0$ und $\alpha \in [0, 2\pi)$ als Polarkoordinaten des Punktes

$$Q := \frac{1}{\gamma M} (Rv_2^2 - \gamma M, -Rv_1 v_2)$$

bestimmt werden können. Für $Q \neq 0$ sind $\epsilon \geq 0$ und $\alpha \in [0, 2\pi)$ sogar eindeutig festgelegt. Die letzte Anfangsbedingung ist ebenfalls erfüllt:

$$\dot{\phi}(0) = \frac{A}{f^2(\phi(0))} = \frac{A}{r^2(0)} = \frac{Rv_2}{R^2} = \frac{v_2}{R}.$$

Damit ist gezeigt: Die Anfangswertaufgabe (1), (2) besitzt eine Lösung (r, ϕ) mit

$$r(t) = \frac{p}{1 + \epsilon \cos(\phi(t) - \alpha)},$$

wobei $p > 0$, $\epsilon \geq 0$ und $\alpha \in [0, 2\pi)$ geeignete Konstanten sind.

Durch

$$K := \{(f(\phi) \cos \phi, f(\phi) \sin \phi) : \phi \in [0, 2\pi]\}$$

mit

$$f(\phi) := \frac{p}{1 + \epsilon \cos(\phi - \alpha)}$$

ist ein Kegelschnitt gegeben und zwar eine Ellipse ($\epsilon < 1$), eine Parabel ($\epsilon = 1$) oder eine Hyperbel ($\epsilon > 1$). Wir gehen jetzt davon aus, dass die Anfangsdaten so vernünftig

sind, dass es sich bei der Planetenbahn um eine Ellipse handelt, da sie ja geschlossen ist. Damit ist das erste Keplersche Gesetz bewiesen. Die Ellipse habe die Halbachsen $a \geq b$. Diese lassen sich aus ϵ und p berechnen und man erhält

$$a = \frac{p}{1 - \epsilon^2}, \quad b = \frac{p}{\sqrt{1 - \epsilon^2}}.$$

Nun zum zweiten Keplerschen Gesetz. Man bezeichne mit $F(t_1, t_2)$ die Größe der vom Fahrstrahl für $t_1 \leq t \leq t_2$ überstrichenen Fläche, also den Flächeninhalt des von den Strahlen $\phi = \phi(t_1)$, $\phi = \phi(t_2)$ und der Kurve $f(\phi)e^{i\phi}$, $\phi(t_1) \leq \phi \leq \phi(t_2)$, begrenzten Gebietes

$$S_{1,2} := \{(r \cos \phi, r \sin \phi) : 0 \leq r \leq f(\phi), \phi(t_1) \leq \phi \leq \phi(t_2)\}.$$

Dann ist

$$F(t_1, t_2) = \frac{1}{2} \int_{\phi(t_1)}^{\phi(t_2)} f^2(\phi) d\phi = \frac{1}{2} \int_{t_1}^{t_2} f^2(\phi(t)) \dot{\phi}(t) dt = \frac{1}{2} \int_{t_1}^{t_2} r^2(t) \dot{\phi}(t) dt = \frac{A}{2} (t_2 - t_1).$$

Hierbei erhält man die erste Gleichung (Leibnizsche Sektorformel⁷) z. B. aus der Transformationsformel für mehrfache Integrale. Damit ist auch das zweite Keplersche Gesetz bewiesen.

Nun sei T die Umlaufzeit des Planeten, also $\phi(T) = 2\pi$. Dann ist $F(0, T) = \frac{1}{2}AT$ die Fläche der Ellipse, die andererseits bekanntlich πab beträgt. Folglich ist $T = 2\pi ab/A$ und daher

$$T^2 = \frac{4\pi^2 a^2 b^2}{A^2} = \frac{4\pi^2 a^2 b^2}{p\gamma M} = \frac{4\pi^2 a^2 b^2}{(b^2/a)\gamma M} = \frac{4\pi^2}{\gamma M} a^3.$$

Damit ist auch das dritte Keplersche Gesetz bewiesen. □

Beispiel: Wir wollen einmal die Anfangswertaufgabe

$$\begin{aligned} \ddot{x}_1 &= -\frac{10}{(x_1^2 + x_2^2)^{3/2}} x_1, & x_1(0) &= 3, & \dot{x}_1(0) &= 1, \\ \ddot{x}_2 &= -\frac{10}{(x_1^2 + x_2^2)^{3/2}} x_2, & x_2(0) &= 0, & \dot{x}_2(0) &= 1 \end{aligned}$$

mit Maple lösen und die zugehörige Planetenbahn in einer (x_1, x_2) -Ebene plotten. Anschließend werden wir die Phasenbahn mit der im Beweis für die Keplerschen Gesetze erhaltenen vergleichen. Mit Hilfe der Eingabe

```
eqn1:=diff(x_1(t),t,t)=-10*x_1(t)/(x_1(t)^2+x_2(t)^2)^(3/2);
eqn2:=diff(x_2(t),t,t)=-10*x_2(t)/(x_1(t)^2+x_2(t)^2)^(3/2);
initial1:=x_1(0)=3,D(x_1)(0)=1;
initial2:=x_2(0)=0,D(x_2)(0)=1;
two_body:=dsolve({eqn1,eqn2,initial1,initial2},{x_1(t),x_2(t)},
  type=numeric);
with(plots);
odeplot(two_body,[x_1(t),x_2(t)],0..10,
  numpoints=200,title="Planetenbahn");
```

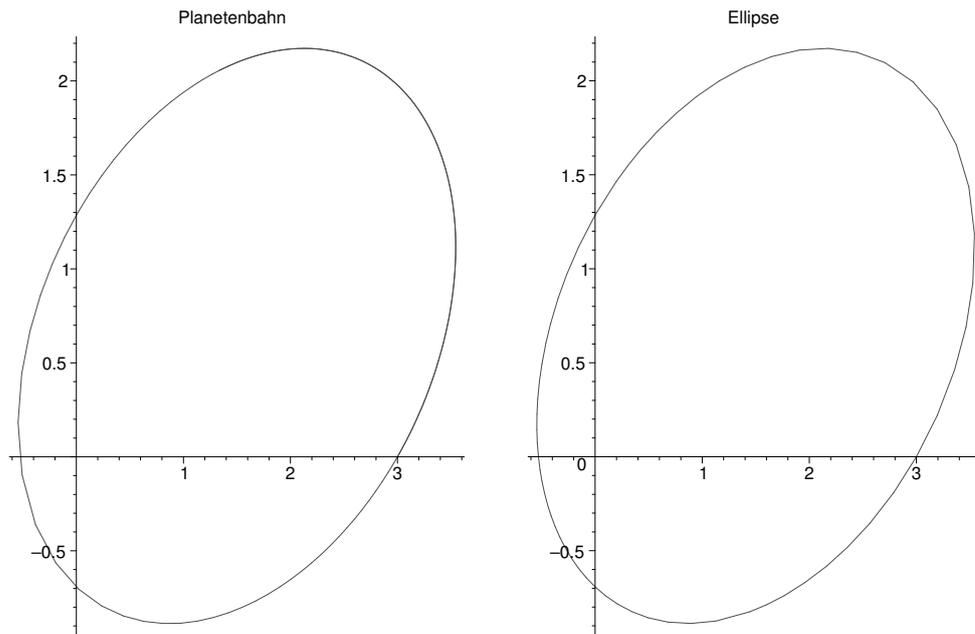


Abbildung 5.9: Eine auf zwei Arten berechnete Planetenbahn

erhalten wir den in der folgenden Abbildung 5.9 links angegebenen Plot. Natürlich hätten wir das System von zwei Differentialgleichungen zweiter Ordnung auch als System von vier Differentialgleichungen erster Ordnung schreiben können. In der Darstellung der Ellipse

$$K := \{(f(\phi) \cos \phi, f(\phi) \sin \phi) : 0 \leq \phi \leq 2\pi\}, \quad f(\phi) := \frac{p}{1 + \epsilon \cos(\phi - \alpha)},$$

erhalten wir $p = 0.9$, während $\epsilon \geq 0$ und $\alpha \in [0, 2\pi)$ aus

$$\epsilon \cos \alpha = -0.7, \quad \epsilon \sin \alpha = -0.3$$

zu berechnen sind. Dies ergibt mittels

```
fsolve({epsilon*cos(alpha)=-0.7,epsilon*sin(alpha)=-0.3},
      {epsilon,alpha},{epsilon=0..1,alpha=0..2*Pi});
```

die Werte

$$\epsilon = 0.7615773106, \quad \alpha = 3.546484440.$$

Mit

```
p:=0.9;epsilon:=0.7615773106;alpha:=3.546484440;
f:=phi->p/(1+epsilon*cos(phi-alpha));
plot([f(phi)*cos(phi),f(phi)*sin(phi),phi=0..2*Pi],title="Ellipse");
```

erhalten wir die Ellipse in Abbildung 5.9 rechts. Einen Unterschied zum linken Pendant wird man nicht feststellen. \square

⁷Siehe

W. WALTER (1990, S. 251) *Analysis II*. Springer-Verlag. Berlin-Heidelberg-New York.

5.1.4 Aufgaben

1. Sei⁸ p die Lösung der Anfangswertaufgabe für die logistische Differentialgleichung

$$p' = ap - bp^2, \quad p(t_0) = p_0,$$

wobei a, b, p_0 positive Konstanten mit $p_0 < \frac{1}{2}(a/b)$ sind.

- (a) Seien $t_1 < t_2$ mit $t_1 > t_0$ und $t_1 - t_0 = t_2 - t_1$ gegeben. Man zeige, dass a und b eindeutig durch $p_0 = p(t_0), p(t_1), p(t_2)$ bestimmt sind. Dies bedeutet: Legt man das logistische Wachstumsmodell zugrunde und sind die Populationen p_0, p_1, p_2 zu äquidistanten Zeiten t_0, t_1, t_2 bekannt, so sind hierdurch die Parameter a, b im Modell eindeutig festgelegt.
- (b) Man zeige, dass genau ein $t^* > t_0$ mit $p(t^*) = \frac{1}{2}(a/b)$ existiert und die Darstellung

$$p(t) = \frac{a/b}{1 + e^{-a(t-t^*)}}$$

gilt.

- (c) Aus

k	t_k	$p(t_k)$
0	1790	3 929 000
1	1850	23 192 000
2	1910	91 972 000

bestimme man a und b . Anschließend berechne man t^* mit $p(t^*) = \frac{1}{2}(a/b)$.

Hinweis: Es darf Maple eingesetzt werden.

2. Das vollständige elliptische Integral erster Ordnung ist mit dem Gaußschen arithmetisch-geometrischen Mittel (AGM) verwandt⁹. Insbesondere zeige man:

- (a) Gegeben seien Zahlen a, b mit $0 < b \leq a$. Auf die folgende Weise erzeuge man Folgen $\{a_k\}, \{b_k\}$.
- Setze $a_0 := a, b_0 := b$.
 - Für $k = 0, 1, \dots$:
 - Berechne $a_{k+1} := \frac{1}{2}(a_k + b_k)$.
 - Berechne $b_{k+1} := \sqrt{a_k b_k}$.

Man zeige: Die Folgen $\{a_k\}$ und $\{b_k\}$ konvergieren monoton nicht wachsend bzw. monoton nicht fallend gegen einen gemeinsamen Grenzwert $M(a, b)$, das sogenannte arithmetisch-geometrische Mittel von a und b .

⁸Diese Aufgabe findet man auf S. 88 von

M. BRAUN (1983) "Single species population models". In: *Differential Equation Models* (eds. M. Braun et al.), Springer-Verlag, New York-Heidelberg-Berlin.

⁹Hierüber kann man sich sehr gut auf den ersten Seiten von

J. M. BORWEIN, P. B. BORWEIN (1987) *Pi and the AGM*. J. Wiley, New York

informieren. Zu recht bezeichnen sie das arithmetisch-geometrische Mittel als eine der Juwelen der klassischen Analysis.

(b) Für $0 < b \leq a$ und $\lambda > 0$ ist $M(\lambda a, \lambda b) = \lambda M(a, b)$.

(c) Für $0 < b \leq a$ ist

$$M(a, b) = M\left(\frac{a+b}{2}, \sqrt{ab}\right).$$

(d) Für $0 < b \leq 1$ ist

$$M(1, b) = \frac{1+b}{2} M\left(1, \frac{2\sqrt{b}}{1+b}\right).$$

(e) Für $0 < x \leq 1$ ist

$$\frac{1}{M(1, x)} = \frac{2}{\pi} K(\sqrt{1-x^2}),$$

wobei mit

$$K(k) := \int_0^{\pi/2} \frac{d\theta}{\sqrt{1-k^2 \sin^2 \theta}}$$

das vollständige elliptische Integral erster Art bezeichnet wird¹⁰.

(f) Man zeige¹¹, dass

$$\frac{1}{M(\sqrt{2}, 1)} = \frac{2}{\pi} \int_0^1 \frac{dt}{\sqrt{1-t^4}}.$$

3. Wir betrachten den Fall eines Körpers der Masse m , der unter dem Einfluss der Schwerkraft sich senkrecht nach unten bewegt, wobei der zur Geschwindigkeit proportionale Luftwiderstand berücksichtigt werde. Mit einer Konstanten $\rho > 0$ und (konstantem) $g = 9.81 \text{ m/sec}^2$ hat man die Anfangswertaufgabe

$$m\ddot{x} = mg - \rho\dot{x}, \quad x(0) = 0, \quad \dot{x}(0) = v_0$$

zu lösen. Man zeige, dass $\lim_{t \rightarrow \infty} \dot{x}(t)$ existiert, der Körper also eine endliche Endgeschwindigkeit erreicht¹².

¹⁰Für einen Beweis kann man J. M. Borwein, P. B. Borwein (1987, S. 5) oder auch J. TODD (1979, S. 18) *Basic Numerical Mathematics*, vol. 1. Birkhäuser Verlag, Basel-Stuttgart konsultieren. Aber selbst dann wird der Beweis nicht ganz einfach sein

¹¹Bei J. Todd (1979, S. 17) kann man nachlesen:

As a teenager in 1791 Gauss, without computers, made extensive calculations of arithmetic-geometric means. In particular he found that

$$M(\sqrt{2}, 1) = 1.19814\ 02347\ 35592\ 20744.$$

It seems clear that he was searching for a formula for $M(a, b)$ It was not until 1799 that he made progress. At that time he computed the definite integral

$$A = \int_0^1 \frac{dt}{\sqrt{1-t^4}}.$$

He then recalled his value of $M(\sqrt{2}, 1)$ given above and observed that the product $AM(\sqrt{2}, 1)$ coincided to many decimal places with $\frac{1}{2}\pi$. In his diary, on 30 May 1799, Gauss wrote that if one could prove rigorously that $AM(\sqrt{2}, 1) = \frac{1}{2}\pi$, then new fields of mathematics would open. In his diary, on 23 December 1799, Gauss noted that he had proved this result, and more; in later years his prophesy was fulfilled.

¹²Bei H. HEUSER (1989, S. 30) findet man hierzu die Bemerkung: Von dieser Tatsache profitiert der Fallschirmspringer immer dann, wenn sein Schirm überhaupt aufgeht.

4. Bei W. Walter (1996, S. 5)¹³ findet man ein System von zwei Differentialgleichungen zweiter Ordnung, durch das die Bewegung eines Satelliten im Gravitationsfeld zweier Körper (z. B. Erde und Mond) modelliert wird, nämlich

$$\begin{aligned}\ddot{x} &= x + 2\dot{y} - \mu' \frac{x + \mu}{[(x + \mu)^2 + y^2]^{3/2}} - \mu \frac{x - \mu'}{[(x - \mu')^2 + y^2]^{3/2}}, \\ \ddot{y} &= y - 2\dot{x} - \mu' \frac{y}{[(x + \mu)^2 + y^2]^{3/2}} - \mu \frac{y}{[(x - \mu')^2 + y^2]^{3/2}}.\end{aligned}$$

Hierbei ist μ eine gegebene Konstante und $\mu' := 1 - \mu$. Für $\mu := 0.01213$ und die Anfangsbedingungen

$$x(0) = 1.2, \quad \dot{x}(0) = 0, \quad y(0) = 0, \quad \dot{y}(0) = -1.04936$$

plotte man die Bahn $\{(x(t), y(t)) : 0 \leq t \leq 10\}$.

5.2 Existenz und Eindeutigkeit beim Anfangswertproblem

In diesem Abschnitt sei die Anfangswertaufgabe

$$(P) \quad x' = f(t, x), \quad x(t_0) = x_0$$

gegeben, wobei die Anfangszeit $t_0 \in \mathbb{R}$, der Anfangszustand $x_0 \in \mathbb{R}^n$ und die auf einer Umgebung U von (t_0, x_0) definierte und dort zumindestens stetige Abbildung $f: U \rightarrow \mathbb{R}^n$ gegeben sind. Gefragt ist nach der Existenz einer Lösung x von (P), also einer auf einer (eventuell nur rechtsseitigen) Umgebung I von t_0 definierten und dort stetig differenzierbaren Funktion $x: I \rightarrow \mathbb{R}^n$ mit $x'(t) = f(t, x(t))$ für alle $t \in I$ und $x(t_0) = x_0$. Ebenso stellt sich die Frage nach der Eindeutigkeit. Die Existenzfrage wird i. Allg. dadurch behandelt, dass man (P) äquivalent als *Fixpunktaufgabe*

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds$$

schreibt und auf diese Fixpunktaufgabe den Kontraktionssatz (dies führt auf den Satz von Picard-Lindelöf) bzw. den Schauderschen Fixpunktsatz (dies liefert den Existenzsatz von Peano) anwendet. Man beachte, dass man eine (explizite) Differentialgleichung n -ter Ordnung, etwa

$$x^{(n)} = f(t, x, x', \dots, x^{(n-1)})$$

als ein System von n Differentialgleichungen erster Ordnung schreiben kann:

$$\begin{aligned}x_1' &= x_2 \\ x_2' &= x_3 \\ &\vdots \\ x_{n-1}' &= x_n \\ x_n' &= f(t, x_1, x_2, \dots, x_n).\end{aligned}$$

¹³W. WALTER (1996) *Gewöhnliche Differentialgleichungen. 6. Auflage.* Springer-Verlag, Berlin-Heidelberg-New York.

5.2.1 Banachscher, Schauderscher Fixpunktsatz

Wir nehmen an (ähnlich wie in Unterabschnitt 2.2.3, dass die folgenden Begriffe aus Analysis II bekannt sind:

- Linearer¹⁴ normierter Raum, Norm, Konvergenz und Cauchy-Folge in einem linearen normierten Raum, Banachraum,
- Abgeschlossene, kompakte, relativ kompakte Teilmengen eines linearen normierten Raumes,
- Stetige Abbildungen zwischen linearen normierten Räumen.

Beispiele: Der \mathbb{R}^n , versehen mit einer beliebigen Vektornorm $\|\cdot\|$, ist bekanntlich ein Banachraum. Die wichtigsten Normen im \mathbb{R}^n sind die euklidische Vektornorm, die Betragssummennorm und die Maximumnorm, welche für einen Vektor $x = (x_j) \in \mathbb{R}^n$ durch

$$\|x\|_2 := \left(\sum_{j=1}^n x_j^2 \right)^{1/2}, \quad \|x\|_1 := \sum_{j=1}^n |x_j|, \quad \|x\|_\infty := \max_{j=1, \dots, n} |x_j|$$

gegeben sind.

Mit $C_n[a, b]$ bezeichnen wir den linearen Raum der stetigen Abbildungen $x: [a, b] \rightarrow \mathbb{R}^n$, wobei bei dieser Schreibweise $[a, b]$ stets als ein kompaktes Intervall in \mathbb{R} verstanden wird und $C[a, b]$ statt $C_1[a, b]$ geschrieben wird. Definiert man auf $C_n[a, b]$ eine Norm durch

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

wobei auf der rechten Seite $\|\cdot\|$ eine beliebige Norm auf dem \mathbb{R}^n ist, so ist $C_n[a, b]$ versehen mit dieser Norm ein Banachraum (siehe Aufgabe 1). Man beachte, dass aus dem Zusammenhang immer hervorgeht, ob es sich bei $\|\cdot\|$ um eine Norm auf dem \mathbb{R}^n oder auf $C_n[a, b]$ handelt. \square

Ein häufig benutztes Hilfsmittel ist in dem folgenden Lemma angegeben (siehe auch Lemma 2.4 in Abschnitt 2.2).

Lemma 4.1 *Ist $x \in C_n[a, b]$ und $\|\cdot\|$ eine Norm auf \mathbb{R}^n , so ist*

$$\left\| \int_a^b x(t) dt \right\| \leq \int_a^b \|x(t)\| dt.$$

Jetzt geben wir noch einmal den Banachschen Fixpunktsatz an (siehe Satz 2.3 in Abschnitt 2.2).

Satz 4.2 (Banach) *Sei $(X, \|\cdot\|)$ ein Banachraum, $K \subset X$ abgeschlossen und F eine Abbildung mit $F(K) \subset K$, die also K in sich abbildet, und die auf K kontrahierend ist, zu der also eine Konstante $q \in (0, 1)$ mit*

$$\|F(x) - F(y)\| \leq q \|x - y\| \quad \text{für alle } x, y \in K$$

¹⁴Der zugrunde gelegte Skalarkörper ist bei uns grundsätzlich der Körper \mathbb{R} der reellen Zahlen.

existiert. Dann besitzt F genau einen Fixpunkt x^* in K und es gilt die Fehlerabschätzung

$$\|x^* - F(x)\| \leq \frac{q}{1-q} \|x - F(x)\| \quad \text{für alle } x \in K.$$

Die beiden folgenden Fixpunktsätze werden wir nicht beweisen¹⁵. Der erste ist der Brouwersche Fixpunktsatz. Von ihm sagt J. Franklin (1980, S. 232):

- This is what the Brouwer theorem says in everyday terms: Sit down with a cup of coffee. Gently and continuously swirl the coffee about in the cup. Put the cup down, and let the motion subside. When the coffee is still, Brouwer says there is at least one point in the coffee that has returned to the exact spot where it was when you first sat down.

Satz 4.3 (Brouwer) Sei $K \subset \mathbb{R}^n$ kompakt und konvex. Ist $F: K \rightarrow \mathbb{R}^n$ eine stetige Abbildung mit $F(K) \subset K$, die K also in sich abbildet, so besitzt F einen Fixpunkt in K .

Diesen Brouwerschen Fixpunktsatz haben wir wegen seiner leichten Formulierbarkeit aufgeführt, weil seine Aussage zum mathematischen Allgemeinwissen gehören sollte und weil er i. Allg. zum Beweis des jetzt folgenden Schauderschen Fixpunktsatzes herangezogen wird.

Satz 4.4 (Schauder) Sei X ein Banachraum und $K \subset X$ abgeschlossen und konvex. Die Abbildung $F: K \rightarrow X$ sei stetig, es sei $F(K) \subset K$ und $F(K)$ relativ kompakt. Dann besitzt F einen Fixpunkt in F .

Der letzte Satz in diesem Unterabschnitt dient dazu, die relative Kompaktheit einer Teilmenge des Banachraumes $C_n[a, b]$ (wie stets versehen mit einer Norm $\|x\| := \max_{t \in [a, b]} \|x(t)\|$) nachzuweisen. Er stammt von Arzela-Ascoli, auch ihn geben wir ohne Beweis an.

Satz 4.5 (Arzela-Ascoli) Die Menge $K \subset C_n[a, b]$ habe die folgenden beiden Eigenschaften:

1. Die Menge K ist beschränkt, d. h. es existiert eine Konstante $C > 0$ mit $\|x\| \leq C$ für alle $x \in K$.
2. Die Menge K ist gleichgradig stetig, d. h. zu jedem $\epsilon > 0$ gibt es ein $\delta = \delta(\epsilon) > 0$ mit

$$t, s \in [a, b], \quad |t - s| \leq \delta, \quad x \in K \implies \|x(t) - x(s)\| \leq \epsilon.$$

Dann ist K relativ kompakt in $C_n[a, b]$.

¹⁵Für Beweise siehe z. B.

J. FRANKLIN (1980) *Methods of Mathematical Economics. Linear and Nonlinear Programming, Fixed-Point Theorems*. Springer-Verlag, New York-Heidelberg-Berlin.

5.2.2 Die Sätze von Picard-Lindelöf und Peano

Wir beginnen mit dem Satz von Picard-Lindelöf, in dem im wesentlichen ausgesagt wird, dass unter der Voraussetzung der lokalen Lipschitzstetigkeit an die rechte Seite der gegebenen Anfangswertaufgabe die lokale eindeutige Lösbarkeit folgt.

Satz 4.6 (Picard-Lindelöf) Sei $D \subset \mathbb{R}^{n+1}$ offen und $f: D \rightarrow \mathbb{R}^n$ stetig und bezüglich der letzten n Variablen lokal Lipschitzstetig, d. h. zu jedem $(t_0, x_0) \in D$ existiert eine Umgebung $U = U(t_0, x_0)$ und eine Konstante $L = L(t_0, x_0)$ derart, dass

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{für alle } (t, x), (t, y) \in D \cap U.$$

Dann existiert zu jedem $(t_0, x_0) \in D$ ein $\alpha^* > 0$ derart, dass die Anfangswertaufgabe

$$x' = f(t, x), \quad x(t_0) = x_0$$

auf $I^* := \{t \in \mathbb{R} : |t - t_0| \leq \alpha^*\}$ eindeutig lösbar ist.

Beweis: Sei $(t_0, x_0) \in D$ gegeben. Man wähle zunächst positive Zahlen α, β derart, dass $I_\alpha \times B_\beta \subset D \cap U$, wobei

$$I_\alpha := \{t \in \mathbb{R} : |t - t_0| \leq \alpha\}, \quad B_\beta := \{x \in \mathbb{R}^n : \|x - x_0\| \leq \beta\}.$$

Dann definiere man

$$M := \max_{(t,x) \in I_\alpha \times B_\beta} \|f(t, x)\|$$

und bestimme $\alpha^* \in (0, \alpha]$, $\beta^* \in (0, \beta]$ mit $M\alpha^* \leq \beta^*$. Z. B. setze man $\beta^* := \beta$ und $\alpha^* := \min(\alpha, \beta/M)$. Anschließend definiere man $I^* := \{t \in \mathbb{R} : |t - t_0| \leq \alpha^*\}$ und die nichtleere Menge

$$K := \left\{ x \in C_n(I^*) : \max_{t \in I^*} \|x(t) - x_0\| \leq \beta^* \right\}.$$

Dann bildet die Abbildung $F: C_n(I^*) \rightarrow C_n(I^*)$, definiert durch

$$F(x)(t) := x_0 + \int_{t_0}^t f(s, x(s)) ds$$

die Menge K in sich ab. Denn ist $x \in K$ und $t \in I^*$, so ist

$$\|x(t) - x_0\| = \left\| \int_{t_0}^t f(s, x(s)) ds \right\| \leq |t - t_0| M \leq \alpha^* M \leq \beta^*$$

und daher auch $F(x) \in K$. Definiert man auf $C_n(I^*)$ die gewichtete Maximumnorm $\|x\|_* := \max_{t \in I^*} e^{-2L|t-t_0|} \|x(t)\|$, wobei L die Lipschitzkonstante von f bezüglich des zweiten Argumentes auf $D \cap U$ ist, so ist $\|F(x) - F(y)\|_* \leq \frac{1}{2} \|x - y\|_*$ für alle $x, y \in K$,

insbesondere bildet F also die Menge K bezüglich der Norm $\|\cdot\|_*$ kontrahierend in sich ab. Denn sind $x, y \in K$ und $t \in I^*$, so ist

$$\begin{aligned}
e^{-2L|t-t_0|} \|F(x)(t) - F(y)(t)\| &= e^{-2L|t-t_0|} \left\| \int_{t_0}^t [f(s, x(s)) - f(s, y(s))] ds \right\| \\
&\leq e^{-2L|t-t_0|} \operatorname{sign}(t-t_0) \int_{t_0}^t \|f(s, x(s)) - f(s, y(s))\| ds \\
&\leq L e^{-2L|t-t_0|} \operatorname{sign}(t-t_0) \int_{t_0}^t \|x(s) - y(s)\| ds \\
&\leq L e^{-2L|t-t_0|} \operatorname{sign}(t-t_0) \int_{t_0}^t e^{2L|s-t_0|} ds \|x - y\|_* \\
&= L e^{-2L|t-t_0|} \frac{1}{2L} (e^{2L|t-t_0|} - 1) \|x - y\|_* \\
&\leq \frac{1}{2} \|x - y\|_*.
\end{aligned}$$

Der Kontraktionssatz (angewandt auf den mit der Norm $\|\cdot\|_*$ versehenen Banachraum $C_n(I^*)$ und die Abbildung F , welche die abgeschlossene Menge K kontrahierend in sich abbildet) liefert, dass F in K genau einen Fixpunkt x besitzt bzw. die Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$, auf dem Intervall I^* genau eine Lösung x in K besitzt. Um die Eindeutigkeit einer Lösung in I^* zu beweisen, nehmen wir an, y sei eine weitere Lösung der Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$, auf dem Intervall I^* . Sei

$$\hat{\alpha} := \sup\{\alpha \in [0, \alpha^*] : \|y(t) - x_0\| \leq \beta^* \text{ für alle } t \in I_\alpha\}.$$

Ist $\hat{\alpha} = \alpha^*$, so ist auch $y \in K$ und folglich $x = y$. Wir nehmen daher an, es sei $\hat{\alpha} < \alpha^*$ und führen dies zum Widerspruch. Für $t \in I_{\hat{\alpha}}$ ist

$$\|y(t) - x_0\| = \|F(y)(t) - x_0\| \leq M|t - t_0| \leq M\hat{\alpha} < M\alpha^* \leq \beta^*,$$

was ein Widerspruch zur Definition von $\hat{\alpha}$ ist. Insgesamt ist der Satz von Picard-Lindelöf bewiesen. \square

Bemerkung: Die lokale Lipschitzstetigkeit von f bezüglich der letzten n Komponenten ist eine schwache Voraussetzung. Ist z. B. $D \subset \mathbb{R}^n$ konvex und sind in D die Ableitungen $\partial f_i / \partial x_j$ stetig und beschränkt ($i, j = 1, \dots, n$), so genügt f in D einer Lipschitzbedingung. Denn wegen des Mittelwertsatzes existiert zu $(t, x), (t, y) \in D$ und $i \in \{1, \dots, n\}$ ein $z \in \mathbb{R}^n$ mit $(t, z) \in D$ (nämlich auf der Verbindungsstrecke zwischen (t, x) und (t, y)) und

$$f_i(t, x) - f_i(t, y) = \sum_{j=1}^n \frac{\partial f_i(t, z)}{\partial x_j} (x_j - y_j),$$

woraus die Behauptung sofort folgt. \square

Etlliche Varianten zum obigen Satz von Picard-Lindelöf sind denkbar.

Korollar 4.7 Sei $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$. Mit positiven Zahlen α^*, β^* sei

$$I^* := \{t \in \mathbb{R} : |t - t_0| \leq \alpha^*\}, \quad B^* := \{x \in \mathbb{R}^n : \|x - x_0\| \leq \beta^*\}.$$

Die Funktion $f: I^* \times B^* \rightarrow \mathbb{R}^n$ sei stetig und bezüglich der zweiten Variablen lipschitzstetig auf $I^* \times B^*$. Sei $M := \max_{(t,x) \in I^* \times B^*} \|f(t, x)\|$ und $M\alpha^* \leq \beta^*$. Dann besitzt die Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$, genau eine Lösung x auf I^* mit $x(t) \in B^*$ für alle $t \in I^*$.

Die Bedingung $M\alpha^* \leq \beta^*$ dient jeweils dazu, zu sichern, dass die Abbildung F die abgeschlossene Kugel $K \subset C_n(I^*)$ in sich abbildet. Ist $\beta^* = \infty$ im obigen Korollar, die rechte Seite f der Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$, also bezüglich des zweiten Argumentes global lipschitzstetig, so kann man auf diese Bedingung verzichten und erhält:

Korollar 4.8 Sei $I \subset \mathbb{R}$ ein kompaktes Intervall. Die Funktion $f: I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei stetig und bezüglich der zweiten Variablen lipschitzstetig auf $I \times \mathbb{R}^n$. Dann besitzt die Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$, für jedes $(t_0, x_0) \in I \times \mathbb{R}^n$ genau eine Lösung auf dem Intervall I .

Beispiel: Der wichtigste Spezialfall des letzten Korollars besteht sicherlich darin, dass $f(t, x) = A(t)x + b(t)$, wobei $A: I \rightarrow \mathbb{R}^{n \times n}$ und $b: I \rightarrow \mathbb{R}^n$ stetig sind. Denn für beliebige $(t, x), (t, y) \in I \times \mathbb{R}^n$ ist dann

$$\|f(t, x) - f(t, y)\| = \|A(t)(x - y)\| \leq \|A(t)\| \|x - y\|.$$

Hierbei bezeichnen wir die der gegebenen Vektornorm $\|\cdot\|$ zugeordnete Matrixnorm ebenfalls mit $\|\cdot\|$. Genauer:

- Ist $\|\cdot\|$ eine Norm auf \mathbb{R}^n , so wird auf $\mathbb{R}^{n \times n}$ durch

$$\|A\| := \max_{\|x\|=1} \|Ax\|, \quad A \in \mathbb{R}^{n \times n},$$

die sogenannte *zugeordnete Matrixnorm* definiert.

Die Abbildung $t \mapsto \|A(t)\|$ ist stetig. Ist also I ein kompaktes Intervall, so ist f global lipschitzstetig mit der Lipschitzkonstanten $L := \max_{t \in I} \|A(t)\|$. \square

Beispiel: Eine direkte Anwendung des Kontraktionssatzes auf spezielle Anfangswertaufgaben kann auch numerisch befriedigende Ergebnisse liefern. Dies wollen wir an einem Beispiel demonstrieren.

Gesucht sei die Lösung der Anfangswertaufgabe

$$x' = t^2 + x^2, \quad x(0) = 0$$

auf dem Intervall $I := [-\frac{1}{2}, \frac{1}{2}]$. Als Näherung nehmen wir $x_0(t) := \frac{1}{3}t^3$. Mit

$$F(x)(t) := \int_0^t [s^2 + x(s)^2] ds$$

wird

$$x_1(t) := F(x_0)(t) = \frac{1}{3}t^3 + \frac{1}{63}t^7.$$

Wenn wir hier, was natürlich nicht unbedingt nötig ist, Maple anwenden wollen, so würden wir dies als Ergebnis von

```
int(s^2+(1/3*s^3)^2,s=0..t);
```

erhalten. Will (oder muss) man aber weitere Iterationen durchführen, so lernt man Maple sehr schnell schätzen. So erhält man

$$x_2(t) = \frac{t^3}{3} + \frac{t^7}{63} + \frac{2t^{11}}{2079} + \frac{t^{15}}{59535}$$

und (das wird keiner mehr zu Fuß ausrechnen wollen)

$$x_3(t) = \frac{t^3}{3} + \frac{t^7}{63} + \frac{2t^{11}}{2079} + \frac{13t^{15}}{218295} + \frac{82t^{19}}{37328445} + \frac{662t^{23}}{10438212015} \\ + \frac{4t^{27}}{3341878155} + \frac{t^{31}}{109876902975}.$$

Übrigens gibt es in Maple bei der Lösung von Anfangswertaufgaben mit `dsolve` die Option `series`, was wir durch den folgenden Ausschnitt illustrieren:

```
> restart;
> Order:=10;dsolve({diff(x(t),t)=t^2+x(t)^2,x(0)=0},x(t),series);
      Order := 10
      x(t) = 1/3 t^3 + 1/63 t^7 + O(t^10)
> x_1:=convert(rhs(%),polynom);
      x_1 := 1/3 t^3 + 1/63 t^7
```

Als Norm auf $C(I)$ wählen wir wie üblich die Maximumnorm, also $\|x\| := \max_{t \in I} |x(t)|$. Als Menge K wählen wir $K := \{x \in C(I) : \|x - x_0\| \leq \beta\}$ und bestimmen $\beta > 0$ so, dass einerseits $F(K) \subset K$ und andererseits F auf K kontrahiert. Ersteres ist erfüllt, wenn $\frac{1}{2}(\frac{1}{24} + \beta)^2 \leq \beta$. Andererseits ist für $x, y \in K$ und $t \in I$ offenbar

$$|F(x)(t) - F(y)(t)| = \left| \int_0^t [x(s) + y(s)][x(s) - y(s)] ds \right| \\ \leq \frac{1}{2} 2 \left(\frac{1}{24} + \beta \right) \|x - y\|$$

und daher $\|F(x) - F(y)\| \leq (\frac{1}{24} + \beta) \|x - y\|$. Daher kontrahiert die Abbildung F auf K , wenn $\frac{1}{24} + \beta < 1$. Z. B. kann man $\beta = 0.001$ wählen und hat als Lipschitzkonstante $q := \frac{1}{24} + 0.001 \leq 0.042667$. Wegen $\|x_1 - x_0\| = \frac{1}{63}(\frac{1}{2})^7 \leq 0.000125$ erhält man aus dem Kontraktionssatz für die Lösung x die Fehlerabschätzung

$$\|x - x_1\| \leq \frac{q}{1 - q} \|x_1 - x_0\| \leq 6 \cdot 10^{-6},$$

was schon ganz befriedigend ist. \square

Ohne die lokale Lipschitzbedingung an die rechte Seite der gegebenen Differentialgleichung ist die Eindeutigkeit einer Lösung i. Allg. nicht gegeben, wie das folgende Beispiel zeigt.

Beispiel: Gegeben sei die Anfangswertaufgabe

$$x' = \sqrt{|x|}, \quad x(0) = 0.$$

Natürlich ist $x := 0$ eine Lösung. Weitere Lösungen erhält man, wenn man mit beliebigem $a \leq 0$ definiert:

$$x(t) := \frac{1}{4} \begin{cases} t^2, & \text{für } t > 0, \\ 0, & \text{für } a \leq t \leq 0, \\ -(t-a)^2, & \text{für } t < a. \end{cases}$$

Dies liegt natürlich daran, dass für $f(t, x) := \sqrt{|t|}$ die lokale Lipschitzbedingung auf einer Umgebung von $(0, 0)$ nicht erfüllt ist. \square

Nun kommen wir zum Existenzsatz von Peano. Grob sagt dieser aus: Ist $D \subset \mathbb{R}^{n+1}$ offen, und $f: D \rightarrow \mathbb{R}^n$ stetig, so "geht durch jeden Punkt $(t_0, x_0) \in D$ eine Lösung von $x' = f(t, x)$ ".

Satz 4.9 (Peano) Sei $D \subset \mathbb{R}^{n+1}$ offen und $f: D \rightarrow \mathbb{R}^n$ stetig. Dann existieren zu jedem $(t_0, x_0) \in D$ positive Zahlen α^*, β^* derart, dass mit

$$I^* := \{t \in \mathbb{R} : |t - t_0| \leq \alpha^*\}, \quad B^* := \{x \in \mathbb{R}^n : \|x - x_0\| \leq \beta^*\}$$

gilt: Es existiert mindestens ein $x \in C_n(I^*)$ mit

(a) Es ist $x(t) \in B^*$ für alle $t \in I^*$.

(b) Auf I^* ist x eine Lösung der Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$.

Beweis: Sei $(t_0, x_0) \in D$. Man wähle positive Zahlen α, β derart, dass $I_\alpha \times B_\beta \subset D$ mit

$$I_\alpha := \{t \in \mathbb{R} : |t - t_0| \leq \alpha\}, \quad B_\beta := \{x \in \mathbb{R}^n : \|x - x_0\| \leq \beta\},$$

was wegen der vorausgesetzten Offenheit von D möglich ist. Anschließend definiere man

$$M := \max_{(t,x) \in I_\alpha \times B_\beta} \|f(t, x)\|,$$

bestimme $\alpha^* \in (0, \alpha]$ und $\beta^* \in (0, \beta]$ mit $M\alpha^* \leq \beta^*$ und setze I^* und B^* wie oben angegeben. Nun wende man den Schauderschen Fixpunktsatz mit folgenden Daten an: Als linearen normierten Raum $(X, \|\cdot\|)$ nehme man $X := C_n(I^*)$ mit der Norm $\|x\| := \max_{t \in I^*} \|x(t)\|$ (wobei rechts $\|\cdot\|$ eine gegebene Norm auf dem \mathbb{R}^n ist), die nichtleere, abgeschlossene und konvexe Menge K sei durch

$$K := \{x \in C_n(I^*) \in B^* \text{ für alle } t \in I^*\}$$

gegeben und schließlich sei die Abbildung $F: C_n(I^*) \rightarrow C_n(I^*)$ definiert durch

$$F(x)(t) := x_0 + \int_{t_0}^t f(s, x(s)) ds.$$

Die Voraussetzungen des Schauderschen Fixpunktsatzes sind leicht nachgeprüft:

1. Die Abbildung $F: K \rightarrow C_n(I^*)$ ist stetig.

Denn: Sei $\epsilon > 0$ beliebig vorgegeben. Die Menge $I^* \times B^*$ ist kompakt und daher f auf $I^* \times B^*$ gleichmäßig stetig. Folglich existiert ein $\delta = \delta(\epsilon) > 0$ mit

$$(t, x), (t, y) \in I^* \times B^*, \quad \|x - y\| \leq \delta \implies \|f(t, x) - f(t, y)\| \leq \frac{\epsilon}{\alpha^*}.$$

Sind nun $x, y \in K$ mit $\|x - y\| \leq \delta$ gegeben, so ist $(s, x(s)), (s, y(s)) \in I^* \times B^*$ für alle $s \in I^*$ und daher

$$\begin{aligned} \|F(x)(t) - F(y)(t)\| &= \left\| \int_{t_0}^t [f(s, x(s)) - f(s, y(s))] ds \right\| \\ &\leq \text{sign}(t - t_0) \int_{t_0}^t \|f(s, x(s)) - f(s, y(s))\| ds \\ &\leq |t - t_0| \frac{\epsilon}{\alpha^*} \\ &\leq \epsilon. \end{aligned}$$

Also gilt die Implikation

$$x, y \in K, \quad \|x - y\| \leq \delta \implies \|F(x) - F(y)\| \leq \epsilon,$$

womit die Stetigkeit von F auf K bewiesen ist.

2. Es ist $F(K) \subset K$.

Sind $x \in K$ und $t \in I^*$ beliebig, so ist

$$\|F(x)(t) - x_0\| \leq M |t - t_0| \leq M\alpha^* \leq \beta^*$$

und damit $F(x) \in K$.

3. K ist nichtleer, abgeschlossen und konvex, $F(K)$ ist relativ kompakt.

Die erste Aussage ist trivial. Die relative Kompaktheit von $F(K)$ folgt aus dem Satz von Arzela-Ascoli, indem man zeigt, dass $F(K)$ beschränkt und gleichgradig stetig ist.

- (a) $F(K)$ ist beschränkt.

Denn: Sind $x \in K$ und $t \in I^*$ beliebig, so ist

$$\|F(x)(t)\| \leq \|x_0\| + \text{sign}(t - t_0) \int_{t_0}^t \|f(s, x(s))\| ds \leq \|x_0\| + M\alpha^*,$$

also $\|F(x)\| \leq C := \|x_0\| + M\alpha^*$ für alle $x \in K$, womit die Beschränktheit von $F(K)$ nachgewiesen ist.

(b) $F(K)$ ist gleichgradig stetig.

Denn: Sei $x \in K$ beliebig vorgegeben. Für beliebige $s, t \in I^*$ ist

$$\|F(x)(s) - F(x)(t)\| = \left\| \int_t^s f(\tau, x(\tau)) d\tau \right\| \leq M |s - t|,$$

woraus man die gleichgradige Stetigkeit von $F(K)$ abliest. Denn zu vorgegebenem $\epsilon > 0$ setze man $\delta := \epsilon/M$ und erhält die Implikation

$$s, t \in I^*, \quad |s - t| \leq \delta, \quad x \in K \implies \|F(x)(s) - F(x)(t)\| \leq \epsilon,$$

was die gleichgradige Stetigkeit von $F(K)$ nach sich zieht.

Der Schaudersche Fixpunktsatz liefert die Behauptung. □

5.2.3 Aufgaben

1. Der lineare Raum $C_n[a, b]$ aller stetigen Abbildungen $x : [a, b] \rightarrow \mathbb{R}^n$, versehen mit der Norm

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

wobei auf der rechten Seite $\|\cdot\|$ eine beliebige Norm auf dem \mathbb{R}^n ist, ist ein Banachraum.

2. Der lineare Raum $C_n^1[a, b]$ aller stetig differenzierbaren Abbildungen $x : [a, b] \rightarrow \mathbb{R}^n$, versehen mit der Norm

$$\|x\| := \max\left(\max_{t \in [a, b]} \|x(t)\|, \max_{t \in [a, b]} \|x'(t)\|\right),$$

wobei auf der rechten Seite $\|\cdot\|$ eine beliebige Norm auf dem \mathbb{R}^n ist, ist ein Banachraum.

3. Die lineare Anfangswertaufgabe erster Ordnung

$$x' = 2tx + t, \quad x(0) = x_0$$

besitzt die Lösung

$$x(t) = x_0 e^{t^2} + \frac{1}{2}(e^{t^2} - 1),$$

wie man durch `dsolve({diff(x(t),t)=2*t*x(t)+t,x(0)=x_0},x(t));` oder eigene Rechnung feststellt. Mit $x_0(t) := x_0$ sei

$$x_{k+1}(t) := x_0 + \int_0^t (2sx_k(s) + s) ds.$$

Man stelle die Iterierten x_k geschlossen dar und begründe, weshalb die Folge $\{x_k\}$ auf jedem kompakten Intervall in \mathbb{R} gleichmäßig gegen die Lösung der gegebenen Anfangswertaufgabe konvergiert.

4. Gegeben sei die Anfangswertaufgabe

$$x' = tx^2, \quad x(0) = 1.$$

Mit $x_0 := 1$ und

$$x_{k+1}(t) := 1 + \int_0^t s x_k(s)^2 ds$$

berechne man x_1, x_2, x_3 . Man bestimme ein Intervall $[0, \alpha]$ mit $\alpha > 0$, auf dem eine Lösung eindeutig existiert und mache eine Fehlerabschätzung.

5. Man zeige, dass die Anfangswertaufgabe für das mathematische Pendel, also

$$x'' + \omega_0^2 \sin x = 0, \quad x(0) = x_0, \quad x'(0) = 0,$$

für beliebige ω_0 und x_0 genau eine Lösung besitzt. Diese existiert auf ganz \mathbb{R} und ist gerade, also $x(t) = x(-t)$ für alle t . Für $\omega_0 := 2$ und $x_0 := 1$ berechne man mit Hilfe des Gaußschen Verfahrens vom arithmetisch-geometrischen Mittel die Periodenlänge $T = (4/\omega_0)K(\sin \frac{1}{2}x_0)$. Schließlich plote man die Lösung auf $[0, 2T]$.

6. Gegeben sei die Anfangswertaufgabe

$$(P) \quad x' = t + \sin x, \quad x(0) = 0.$$

- (a) Man zeige, dass (P) auf jedem kompakten Teilintervall I von \mathbb{R} mit $0 \in I$ genau eine Lösung besitzt.
- (b) Mit Hilfe von Maple-Befehlen plote man die Lösung von (P) auf dem Intervall $[-1, 1]$.

5.3 Lineare Differentialgleichungssysteme

Nur sehr kurz wollen wir auf lineare Systeme von Differentialgleichungen erster Ordnung eingehen.

5.3.1 Lineare Systeme mit variablen Koeffizienten

In diesem Unterabschnitt betrachten wir lineare Differentialgleichungssysteme der Form

$$(*) \quad x' = A(t)x + b(t).$$

Es werde vorausgesetzt, dass $A : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ und $b : \mathbb{R} \rightarrow \mathbb{R}^n$ stetig sind. Die Modifikationen für den Fall, dass A und b nicht auf ganz \mathbb{R} , sondern nur einem Teilintervall $I \subset \mathbb{R}$ erklärt und stetig sind, werden offensichtlich sein. Alle Lösungen von (*) erhält man natürlich dadurch, dass man zu einer speziellen Lösung von (*) alle Lösungen der zugehörigen homogenen Aufgabe $x' = A(t)x$ addiert.

Bei gegebenem $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$ ist die Anfangswertaufgabe

$$x' = A(t)x, \quad x(t_0) = x_0$$

eindeutig lösbar, ferner existiert die Lösung, die wir mit $x(\cdot; t_0, x_0)$ bezeichnen wollen, auf ganz \mathbb{R} . Dies hatten wir uns im Anschluss an den Existenz- und Eindeutigkeitsatz von Picard-Lindelöf überlegt, siehe Korollar 4.8. Die Abbildung $x_0 \mapsto x(\cdot; t_0, x_0)$ vom

\mathbb{R}^n in den Lösungsraum der homogenen Aufgabe $x' = A(t)x$ ist linear und bijektiv. Daher hat der Lösungsraum der homogenen Aufgabe die Dimension n .

Seien x_1, \dots, x_n (nicht notwendig linear unabhängige) Lösungen von $x' = A(t)x$ und $X(t) = \begin{pmatrix} x_1(t) & \cdots & x_n(t) \end{pmatrix}$ diejenige Matrix, die $x_1(t), \dots, x_n(t)$ als Spalten besitzt. Dann ist $X'(t) = A(t)X(t)$ und es gilt:

$$\det X(\tau) \neq 0 \text{ für ein } \tau \in \mathbb{R} \iff \det X(t) \neq 0 \text{ für alle } t \in \mathbb{R}.$$

Denn: Angenommen, es ist $\det X(\tau) = 0$, also $X(\tau)$ singular. Dann existiert ein $c \neq 0$ mit $X(\tau)c = 0$. Dann ist aber $X(\cdot)c = \sum_{j=1}^n c_j x_j(\cdot)$ eine Lösung von $x' = A(t)x$, die zur Zeit τ verschwindet. Da die triviale Lösung $x \equiv 0$ eine ebensolche ist und eine Lösung linearer Anfangswertaufgaben eindeutig bestimmt ist, ist $X(t)c = 0$ für alle t und daher $\det X(t) = 0$ für alle $t \in \mathbb{R}$.

Ist $X'(t) = A(t)X(t)$ für alle t und $\det X(t) \neq 0$ (für ein oder für alle t , das bleibt sich gleich), so nennt man X ein *Fundamentalsystem* von $x' = A(t)x$. Ist X ein Fundamentalsystem, so erhält man *sämtliche* Lösungen des homogenen Systems in der Form $x(t) = X(t)c$ mit $c \in \mathbb{R}^n$.

Ist X ein beliebiges Fundamentalsystem von $x' = A(t)x$, so ist durch $x(t) := X(t) \int_{t_0}^t X^{-1}(s)b(s) ds$ eine spezielle Lösung von (*) gegeben, wie man sofort nachrechnet. Daher ist

$$x(t) := X(t)c + X(t) \int_{t_0}^t X^{-1}(s)b(s) ds$$

mit beliebigem $c \in \mathbb{R}^n$ die allgemeine Lösung von (*), während durch

$$x(t) := X(t) \left[X^{-1}(t_0)x_0 + \int_{t_0}^t X^{-1}(s)b(s) ds \right]$$

die (eindeutige) Lösung der Anfangswertaufgabe

$$x' = A(t)x + b(t), \quad x(t_0) = x_0$$

gegeben ist.

Beispiel: Nur in Ausnahmefällen kann man bei einem linearen Differentialgleichungssystem mit variablen Koeffizienten ein Fundamentalsystem geschlossen angeben. Wir reproduzieren ein Beispiel bei W. Walter (1996, S. 144), in welchem ein Fundamentalsystem angegeben werden kann. Und zwar betrachten wir die inhomogene Aufgabe

$$(*) \quad \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \frac{1}{t} & -1 \\ \frac{1}{t^2} & \frac{2}{t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} t \\ -t^2 \end{pmatrix}, \quad \begin{pmatrix} x_1(1) \\ x_2(1) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Mit Hilfe von

```
dsolve({diff(x_1(t),t)=(1/t)*x_1(t)-x_2(t),
diff(x_2(t),t)=(1/t^2)*x_1(t)+(2/t)*x_2(t),
x_1(1)=1, x_2(1)=0},{x_1(t),x_2(t)});
```

und

```
dsolve({diff(x_1(t),t)=(1/t)*x_1(t)-x_2(t),
diff(x_2(t),t)=(1/t^2)*x_1(t)+(2/t)*x_2(t),
x_1(1)=0, x_2(1)=1},{x_1(t),x_2(t)});
```

erhält man das Fundamentalsystem

$$X(t) = \begin{pmatrix} (-\ln t + 1)t^2 & -t^2 \ln t \\ t \ln t & t(1 + \ln t) \end{pmatrix}.$$

Die Berechnung der Lösung von (*) mittels der Formel

$$x(t) = X(t) \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_1^t X(s)^{-1} \begin{pmatrix} s \\ -s^2 \end{pmatrix} ds \right)$$

scheint (wer kennt eine bessere Lösung?) komplizierter zu sein als z. B. in *Mathematica*. Dies liegt daran, dass `int` offenbar nicht auf Vektoren komponentenweise angewandt werden kann. Die gesuchte Lösung erhält man etwa folgendermaßen:

```
> with(LinearAlgebra):
> b:=t-><t,-t^2>;
> X:=t->Matrix([[(-ln(t)+1)*t^2,-t^2*ln(t)],[t*ln(t),t*(1+ln(t))]]):
> x:=t->X(t).(<1,0>+<seq(int((MatrixInverse(X(s)).b(s))[i],s=1..t),i=1..2)>):
> <seq(expand(x(t)[i]),i=1..2)>;
```

$$\begin{bmatrix} -\frac{1}{2}t^2 \ln(t) - \frac{1}{2}t^2 \ln(t)^2 + \frac{3}{4}t^2 + \frac{1}{4}t^4 \\ \frac{3}{2}t \ln(t) + \frac{1}{2}t \ln(t)^2 - \frac{3}{4}t^3 + \frac{3}{4}t \end{bmatrix}$$

Diese kann man natürlich auch versuchen, direkt durch

```
dsolve({diff(x_1(t),t)=(1/t)*x_1(t)-x_2(t)+t,
diff(x_2(t),t)=(1/t^2)*x_1(t)+(2/t)*x_2(t)-t^2,
x_1(1)=1,x_2(1)=0},{x_1(t),x_2(t)});
```

zu berechnen. Genau wie bei *Mathematica* ist man hiermit nicht erfolgreich. \square

5.3.2 Lineare Systeme mit konstanten Koeffizienten

In diesem Unterabschnitt kommt es vor allem darauf an, ein Fundamentalsystem zu $x' = Ax$ zu bestimmen, wobei $A \in \mathbb{R}^{n \times n}$ eine konstante Matrix ist. Ist dies gelungen, so können inhomogene Aufgaben wie im letzten Unterabschnitt behandelt werden.

Zunächst machen wir für eine Lösung von $x' = Ax$ aber den Ansatz $x(t) = ce^{\lambda t}$, wobei der Vektor $c \in \mathbb{R}^n$ vom Nullvektor verschieden sein sollte, damit der Lösungskandidat nichttrivial ist. Wie man sehr leicht nachrechnet ist x genau dann eine Lösung, wenn $Ac = \lambda c$, also λ ein Eigenwert von A mit zugehörigem Eigenvektor c ist. Man erkennt hieran eine (kleine) Schwierigkeit: Eine reelle Matrix hat i. allg. nicht nur reelle, sondern auch komplexe Eigenwerte, die dann in konjugiert komplexen Paaren auftreten. Dann sind auch die Eigenvektoren komplex, sie treten in konjugiert komplexen

Paaren auf. Damit ist $x(t) = ce^{\lambda t}$ in diesem Falle komplexwertig, womit man bei reellen Ausgangsdaten nicht immer zufrieden ist. Dies ist aber wirklich nur eine kleine Schwierigkeit. Denn ist $A \in \mathbb{R}^{n \times n}$ eine reelle $n \times n$ -Matrix mit dem komplexen Eigenwert $\lambda = \mu + i\nu$ und dem zugehörigen Eigenvektor $c = a + ib$, so ist auch $\bar{\lambda} = \mu - i\nu$ ein Eigenwert von A mit dem zugehörigen Eigenvektor $\bar{c} = a - ib$. Dann ist

$$ce^{\lambda t} = (a + ib)e^{(\mu + i\nu)t} = e^{\mu t}(a \cos \nu t - b \sin \nu t) + ie^{\mu t}(b \cos \nu t + a \sin \nu t).$$

Sowohl der Real- als auch der Imaginärteil hiervon sind Lösungen. Wichtig ist die Bemerkung: Besitzt $A \in \mathbb{R}^{n \times n}$ ein System von n linear unabhängigen Eigenvektoren, so gewinnt man auf diese Weise n linear unabhängige Lösungen von $x' = Ax$ bzw. ein Fundamentalsystem von $x' = Ax$. Insbesondere ist das der Fall, wenn A symmetrisch ist oder die Eigenwerte von A paarweise von einander verschieden sind.

Beispiel: Man bestimme die allgemeine Lösung von

$$x''' + 6x'' + 11x' + 6x = 0.$$

Dies ist äquivalent zu

$$x' = Ax \quad \text{mit} \quad A := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{pmatrix}.$$

Als charakteristisches Polynom erhält man

$$p_3(\lambda) := \det(A - \lambda I) = -\lambda^3 - 6\lambda^2 - 11\lambda - 6$$

mit den Wurzeln

$$\lambda_1 = -1, \quad \lambda_2 = -2, \quad \lambda_3 = -3$$

und den zugehörigen Eigenvektoren

$$c_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}, \quad c_3 = \begin{pmatrix} 1 \\ -3 \\ 9 \end{pmatrix}.$$

Als allgemeine Lösung der gegebenen Differentialgleichung erster Ordnung hat man also

$$x(t) = ae^{-t} + be^{-2t} + ce^{-3t}.$$

Will man z. B. die Lösung bestimmen, die den Anfangsbedingungen $x(0) = 1$, $x'(0) = x''(0) = 0$ genügt, so hat man das Gleichungssystem

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

zu lösen, woraus man $(a, b, c) = (3, -3, 1)$ erhält. Bei diesen Rechnungen (die hier noch leicht "per hand" durchgeführt werden können) kann wieder Maple helfen, da

es im package `LinearAlgebra` u. a. Funktionen zur Berechnung des charakterischen Polynoms sowie der Eigenwerte und Eigenvektoren einer Matrix gibt. \square

Die Matrix $A \in \mathbb{C}^{n \times n}$ (es ist zweckmäßig, hier gleich den komplexen Fall zuzulassen) besitzt genau dann n linear unabhängige Eigenvektoren, wenn A *diagonalisierbar* ist, d. h. durch eine Ähnlichkeitstransformation auf Diagonalgestalt transformiert werden kann. Ist also $C = (c_1 \ \cdots \ c_n) \in \mathbb{C}^{n \times n}$ eine nichtsinguläre Matrix mit $C^{-1}AC = \text{diag}(\lambda_1, \dots, \lambda_n)$, so ist durch $X(t) := (e^{\lambda_1 t} c_1 \ \cdots \ e^{\lambda_n t} c_n)$ ein Fundamentalsystem zu $x' = Ax$ gegeben. Nun ist aber bekanntlich nicht jede Matrix $A \in \mathbb{C}^{n \times n}$ diagonalisierbar, die Matrix

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

ist hierfür ein Beispiel. Aber jede Matrix kann durch eine Ähnlichkeitstransformation auf Jordansche Normalform transformiert werden. Dies ist der Schlüssel, um auch für solche Matrizen ein Fundamentalsystem berechnen zu können. Wir verweisen hierzu nur auf die Literatur, z. B. W. Walter (1996, S. 151 ff.), und auf Aufgabe 1.

5.3.3 Aufgaben

1. Für $A \in \mathbb{C}^{n \times n}$ bezeichne e^{At} das durch $X(0) = I$ normierte Fundamentalsystem $X(t)$ zu $x' = Ax$. Man zeige:

- (a) Es ist

$$\frac{d}{dt} e^{At} = A e^{At} = e^{At} A, \quad e^{A(t+s)} = e^{At} e^{As}, \quad (e^{At})^{-1} = e^{(-A)t} = e^{A(-t)}$$

für alle $t, s \in \mathbb{R}$.

- (b) Ist $C \in \mathbb{C}^{n \times n}$ nichtsingulär und $J := C^{-1}AC$, so ist $e^{At} = C e^{Jt} C^{-1}$.

- (c) Mit $\lambda \in \mathbb{C}$ sei $J \in \mathbb{C}^{n \times n}$ definiert durch

$$J := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & \lambda \end{pmatrix}.$$

Dann ist

$$e^{Jt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n-2}}{(n-2)!} & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{n-3}}{(n-3)!} & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & t \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

(d) Die inhomogene Anfangswertaufgabe

$$x' = Ax + b(t), \quad x(t_0) = x_0$$

besitzt die eindeutige Lösung

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^t e^{A(t-s)}b(s) ds.$$

(e) Sind $A, B \in \mathbb{C}^{n \times n}$ mit $AB = BA$, so ist $e^{(A+B)t} = e^{At}e^{Bt}$ für alle t .

(f) Es ist

$$e^{At} = \sum_{j=0}^{\infty} \frac{A^j t^j}{j!}.$$

Genauer ist

$$\lim_{k \rightarrow \infty} \sum_{j=0}^k \frac{A^j t^j}{j!} = e^{At} \quad \text{für jedes } t \in \mathbb{R},$$

wobei diese Konvergenz auf kompakten Teilmengen von \mathbb{R} gleichmäßig ist.

(g) Sei

$$A := \begin{pmatrix} 0 & 1 & 0 \\ 4 & 3 & -4 \\ 1 & 2 & -1 \end{pmatrix}.$$

Man zeige, dass

$$e^{At} = \begin{pmatrix} 1 & 4 & 0 \\ 0 & 4 & 4 \\ 1 & 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^t & te^t \\ 0 & 0 & e^t \end{pmatrix} \begin{pmatrix} 1 & 4 & 0 \\ 0 & 4 & 4 \\ 1 & 6 & 1 \end{pmatrix}^{-1}.$$

Man vergleiche die hierdurch gewonnene Matrix e^A mit $\sum_{j=0}^{10} A^j/j!$.

2. Für zwei durch eine Feder gekoppelte Pendel gleicher Masse $m = 1$ und gleicher Länge l lauten die Bewegungsgleichungen

$$\begin{aligned} \ddot{x} &= -\alpha x - k(x - y) \\ \ddot{y} &= -\alpha y - k(y - x), \end{aligned}$$

wobei g die Erdbeschleunigung und k die (positive) Federkonstante bedeuten und $\alpha := g/l$ gesetzt ist. Schreibt man die beiden Differentialgleichungen zweiter Ordnung als ein System von vier Differentialgleichungen erster Ordnung, so erhält man ein homogenes System mit der Koeffizientenmatrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -(\alpha + k) & 0 & k & 0 \\ 0 & 0 & 0 & 1 \\ k & 0 & -(\alpha + k) & 0 \end{pmatrix}.$$

Man bestimme ein zu $x' = Ax$ gehörendes (nicht notwendig normiertes) Fundamentalsystem. Ferner löse man die Anfangswertaufgabe für den Fall, dass zur Zeit $t = 0$ ein Pendel angestoßen wird bzw. die Anfangswerte $x(0) = y(0) = y'(0) = 0$, $x'(0) = 1$ vorgegeben werden.

3. Man bestimme (wie auch immer)¹⁶ ein reelles Fundamentalsystem von Lösungen der Differentialgleichungssysteme $x' = Ax$ mit

$$A := \begin{pmatrix} 3 & 6 \\ -2 & -3 \end{pmatrix}, \quad A := \begin{pmatrix} 8 & 1 \\ -4 & 4 \end{pmatrix}.$$

4. Man bestimme die allgemeine Lösung von

$$x'' - 6x' + 25x = e^{2t}.$$

Anschließend bestimme man die Lösung zu den Anfangswerten $x(0) = 1$, $x'(0) = 0$.

5.4 Einschrittverfahren bei Anfangswertaufgaben

Zwar sind in Maple, Mathematica, MATLAB und anderen mathematischen Anwendersystemen Löser für Anfangswertaufgaben bei gewöhnlichen Differentialgleichungen enthalten, aber man möchte natürlich wenigstens annähernd verstehen, auf welchen Prinzipien diese basieren. Daher wollen wir in diesem Abschnitt eine ganz kurze Einführung in die einfachste Klasse von Verfahren zur numerischen Behandlung von Anfangswertaufgaben bei gewöhnlichen Differentialgleichungssystemen geben, nämlich Einschrittverfahren. Wir betrachten Anfangswertaufgaben der Form

$$(P) \quad x' = f(t, x), \quad x(t_0) = x_0$$

und setzen der Einfachheit halber generell voraus, dass:

- (V) Es ist $f: [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig und es existiert eine Konstante $L > 0$ mit

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{für alle } (t, x), (t, y) \in [t_0, T] \times \mathbb{R}^n.$$

Dies impliziert die Existenz genau einer Lösung von (P) auf dem Intervall $[t_0, T]$. Wir brauchen uns also keine Gedanken darüber zu machen, dass wir vielleicht versuchen etwas zu berechnen, was es gar nicht gibt.

5.4.1 Beispiele von Einschrittverfahren

Sei x die Lösung von (P), also $x(t)$ der Wert der Lösung zur Zeit t , und $h > 0$ eine gewisse Schrittweite. Bei einem Einschrittverfahren ist eine *Näherung* $u(t)$ für $x(t)$ bekannt und es wird eine Vorschrift angegeben, wie $u(t+h)$, eine Näherung für $x(t+h)$, zu berechnen ist. Ausgangspunkt ist stets die Identität

$$x(t+h) = x(t) + \int_t^{t+h} f(s, x(s)) ds.$$

Indem man rechts das Integral durch einen durch eine *Quadraturformel* ermittelten Wert ersetzt, erhält man geeignete Verfahren.

¹⁶Die Aufgabe ist W. Walter (1996, S. 159) entnommen.

(a) Das Eulersche Polygonzugverfahren.

Wendet man die sehr simple Quadraturformel

$$\int_a^b f(s) ds \approx (b-a)f(a)$$

an, so erhält man $x(t+h) \approx x(t) + hf(t, x(t))$. Ersetzt man hier auf der rechten Seite die unbekannte exakte Lösung $x(t)$ zur Zeit t durch $u(t)$, so erhält man das Euler'sche Polygonzugverfahrens, nämlich

$$u(t+h) := u(t) + hf(t, u(t)).$$

(b) Verfahren von Heun (auch Verfahren von Euler-Cauchy genannt).

Es ist

$$x(t+h) \approx x(t) + \frac{1}{2} h [f(t, x(t)) + f(t+h, x(t+h))],$$

was einer Anwendung der Trapezregel

$$\int_a^b f(s) ds \approx \frac{1}{2} [f(a) + f(b)]$$

entspricht. Hieraus könnte man natürlich das *implizite* Einschrittverfahren

$$u(t+h) := u(t) + \frac{1}{2} h [f(t, u(t)) + f(t+h, u(t+h))]$$

gewinnen, das sogenannte verbesserte Euler-Verfahren oder Euler-Heun-Verfahren. Zur Bestimmung von $u(t+h)$ muss hier also ein nichtlineares Gleichungssystem (bzw. eine Fixpunktaufgabe) gelöst werden. Unter der Voraussetzung (V) hat diese Fixpunktaufgabe wegen des Kontraktionssatzes für $\frac{1}{2} hL < 1$, also für eine hinreichend kleine Schrittweite h , eine eindeutige Lösung. Man kann aber auch aus

$$x(t+h) \approx x(t) + hx'(t) = x(t) + hf(t, x(t))$$

die Verfahrensvorschrift

$$u(t+h) := u(t) + \frac{1}{2} h [f(t, u(t)) + f(t+h, u(t) + hf(t, u(t)))]$$

erhalten, das Verfahren von Heun.

Die bisher angegebenen Verfahren spielen i. Allg. in der Praxis keine Rolle. Wir haben sie trotzdem angegeben und hergeleitet, weil man an ihnen einige theoretische Ergebnisse besonders einfach erläutern kann. Die bisher benutzten Quadraturformeln zur näherungsweisen Berechnung von $I(f) := \int_a^b f(s) ds$ haben die Form

$$Q(f) := (b-a)f(a), \quad Q(f) := \frac{1}{2} (b-a) [f(a) + f(b)].$$

Bezeichnet man mit \mathcal{P}_k die Menge der Polynome vom Grad $\leq k$, so stellt man sofort fest, dass die erste Quadraturformel auf \mathcal{P}_0 (der Menge der Konstanten) exakt ist, während die zweite auf \mathcal{P}_1 exakt ist.

(c) Verfahren von Runge-Kutta (vierter Ordnung).

Ausgangspunkt ist

$$x(t+h) \approx x(t) + \frac{1}{6} h [f(t, x(t)) + 4f(t + \frac{1}{2}h, x(t + \frac{1}{2}h)) + f(t+h, x(t+h))].$$

Dies entspricht der Anwendung der Simpson-Regel (siehe Aufgabe 1)

$$\int_a^b f(s) ds \approx \frac{b-a}{6} [f(a) + 4f(\frac{1}{2}(a+b)) + f(b)].$$

Es ist

$$x(t + \frac{1}{2}h) \approx x(t) + \frac{1}{2} h x'(t) = x(t) + \frac{1}{2} h k_1(t)$$

mit

$$k_1(t) := f(t, x(t)).$$

Entsprechend ist

$$\begin{aligned} x(t + \frac{1}{2}h) &\approx x(t) + \frac{1}{2} h x'(t + \frac{1}{2}h) \\ &= x(t) + \frac{1}{2} h f(t + \frac{1}{2}h, x(t + \frac{1}{2}h)) \\ &\approx x(t) + \frac{1}{2} h f(t + \frac{1}{2}h, x(t) + \frac{1}{2} h k_1(t)) \\ &= x(t) + \frac{1}{2} h k_2(t) \end{aligned}$$

mit

$$k_2(t) := f(t + \frac{1}{2}h, x(t) + \frac{1}{2} h k_1(t)).$$

Ferner ist

$$\begin{aligned} x(t+h) &\approx x(t) + h x'(t + \frac{1}{2}h) \\ &= x(t) + h f(t + \frac{1}{2}h, x(t + \frac{1}{2}h)) \\ &\approx x(t) + h f(t + \frac{1}{2}h, x(t) + \frac{1}{2} h k_2(t)) \\ &= x(t) + h k_3(t) \end{aligned}$$

mit

$$k_3(t) := f(t + \frac{1}{2}h, x(t) + \frac{1}{2} h k_2(t)).$$

Schließlich ist

$$f(t+h, x(t+h)) \approx f(t+h, x(t) + h k_3(t)) =: k_4(t).$$

Damit wird

$$\begin{aligned} x(t+h) &\approx x(t) + \frac{1}{6} h [f(t, x(t)) + 2f(t + \frac{1}{2}h, x(t + \frac{1}{2}h)) \\ &\quad + 2f(t + \frac{1}{2}h, x(t + \frac{1}{2}h)) + f(t+h, x(t+h))] \\ &\approx x(t) + \frac{1}{6} h [k_1(t) + 2k_2(t) + 2k_3(t) + k_4(t)] \end{aligned}$$

Das (klassische) Runge-Kutta Verfahren ist daher durch die folgende Vorschrift gegeben:

$$u(t+h) := u(t) + \frac{h}{6} [k_1(t) + 2k_2(t) + 2k_3(t) + k_4(t)]$$

mit

$$\begin{aligned}k_1(t) &:= f(t, u(t)), \\k_2(t) &:= f\left(t + \frac{1}{2}h, u(t) + \frac{1}{2}hk_1(t)\right), \\k_3(t) &:= f\left(t + \frac{1}{2}h, u(t) + \frac{1}{2}hk_2(t)\right), \\k_4(t) &:= f(t + h, u(t) + hk_3(t)).\end{aligned}$$

Beispiel: Wir geben ein MATLAB-Programm zum Runge-Kutta-Verfahren an¹⁷. Das Intervall $[t_0, t_{\max}]$, auf dem die Lösung der gegebenen Anfangswertaufgabe zu berechnen ist, sei äquidistant unterteilt.

```
function [tvals,xvals]=FixedRK(fname,x_0,t_0,t_max,m);
%*****
%Berechnet wird mit dem Runge-Kutta-Verfahren eine Naehungsl"osung fuer
%die Anfangswertaufgabe x'=f(t,x), x(t_0)=x_0 mit der festen Schrittweite
%h=(t_max-t_0)/(m-1).
%Input:  fname=string fuer die rechte Seite f (siehe folgende Beispiele).
%        x_0=Vektor des Anfangszustandes.
%        t_0=Anfangszeit
%        t_max=Endzeit.
%        m=Zahl der Schritte, es wird mit der konstanten Schrittweite
%          h=(t_max-t_0)/(m-1) gerechnet.
%Output: tvals(k)=t_0+(k-1)h, k=1:m
%        xvals(:,k)=Naeherungsloesung in t=tvals(k), k=1..m
%*****
t=t_0; x=x_0; tvals=t; xvals=x; f=feval(fname,t,x);
h=(t_max-t_0)/(m-1);
for k=1:m-1
    k_1=f;
    k_2=feval(fname,t+(h/2),x+(h/2)*k_1);
    k_3=feval(fname,t+(h/2),x+(h/2)*k_2);
    k_4=feval(fname,t+h,x+h*k_3);
    x=x+(h/6)*(k_1+2*k_2+2*k_3+k_4);
    t=t+h;
    f=feval(fname,t,x);
    xvals=[xvals x];
    tvals=[tvals t];
end;
```

Als Beispiel wenden wir die obige Funktion auf die Anfangswertaufgabe (Lotka-Volterra-Modell)

$$\begin{aligned}x' &= 2x - 0.01xy, & x(0) &= 300, \\y' &= -y + 0.01xy, & y(0) &= 150\end{aligned}$$

¹⁷Siehe auch

an. Wir definieren eine Funktion, die die rechte Seite des Differentialgleichungssystems liefert (dies kann ein eigenes Function-File sein oder der Funktion `FixedRK` angehängt sein), etwa

```
function out=Lotka(t,x);
    out=[2*x(1)-0.01*x(1)*x(2);-x(2)+0.01*x(1)*x(2)];
```

Anschließend gibt man z. B. ein:

```
x_0=[300;150];
[t,x]=FixedRK('Lotka',x_0,0,10,200);
plot(x(1,:),x(2,:))
title('Phasenbahn im Lotka-Volterra-Modell')
xlabel('Beute')
ylabel('Raeuber')
```

Wir erhalten das in Abbildung 5.10 angegebene Bild für die Phasenbahn. □

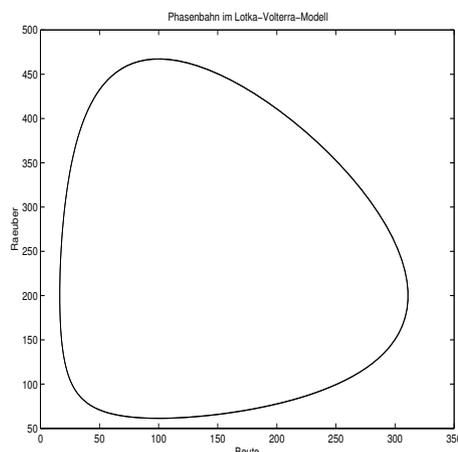


Abbildung 5.10: Phasenbahn beim Lotka-Volterra-Modell

5.4.2 Konsistenz von Einschrittverfahren

Nun kommen wir zur Definition von Einschrittverfahren und überlegen uns dazu, was die im letzten Unterabschnitt angegebenen Verfahren gemeinsam haben. Gemeinsam ist, daß aus der Kenntnis einer Näherung $u(t)$ für die Lösung $x(t)$ zur Zeit t bei vorgegebener Schrittweite $h > 0$ und rechter Seite f eine Näherung $u(t+h)$ für $x(t+h)$ aus der Vorschrift

$$u(t+h) := u(t) + h\Phi(h, f)(t, u(t))$$

berechnet wird. Hierbei heißt Φ die *Verfahrensfunktion* des zugehörigen *Einschrittverfahrens*. In den bisher angegebenen Beispielen ist die Verfahrensfunktion gegeben durch:

- (a) Euler: $\Phi(h, f)(t, u) := f(t, u)$.

- (b) Heun: $\Phi(h, f)(t, u) := \frac{1}{2} [f(t, u) + f(t + h, u + hf(t, u))]$.
- (c) Runge-Kutta: $\Phi(h, f)(t, u) := \frac{1}{6} [k_1 + 2k_2 + 2k_3 + k_4]$ mit $k_1 := f(t, u)$, $k_2 := f(t + \frac{1}{2}h, f(u + \frac{1}{2}hk_1))$, $k_3 := f(t + \frac{1}{2}h, u + \frac{1}{2}hk_2)$, $k_4 := f(t + h, u + hk_3)$.

Im folgenden sei $\|\cdot\|$ eine beliebig gegebene Norm auf dem \mathbb{R}^n . Die folgenden Funktionenräume werden in der Konvergenztheorie eine Rolle spielen:

- (a) Sei $\text{Lip}[t_0, T]$ die Menge aller stetigen Abbildungen $f: [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, zu denen eine Konstante $L > 0$ mit

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{für alle } (t, x), (t, y) \in [t_0, T] \times \mathbb{R}^n$$

existiert.

- (b) Sei $F_N[t_0, T]$ die Menge aller stetigen Abbildungen $f: [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, deren partielle Ableitungen bis zur Ordnung N auf $[t_0, T] \times \mathbb{R}^n$ existieren und dort stetig und beschränkt sind.

Dann ist offenbar $F_1[t_0, T] \subset \text{Lip}[t_0, T]$. Natürlich ist die Voraussetzung, dass die rechte Seite f einer Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$, zu einer Funktionenklasse $F_N[t_0, T]$ gehört, ziemlich einschränkend. "Lokale Versionen" (lokal bezieht sich hier auf die Variable x) der obigen Funktionenräume wären angebracht. Da die Argumentation aber nicht wesentlich anders würde, bleiben wir bei obigen Funktionsklassen.

Definition 6.1 Für $(t, u) \in [t_0, T] \times \mathbb{R}^n$ sei $z = z(s)$ die Lösung der Anfangswertaufgabe $z' = f(s, z)$, $z(t) = u$. Dann heißt

$$\Delta(h, f)(t, u) := \frac{z(t+h) - z(t)}{h} - \Phi(h, f)(t, u)$$

der *lokale Diskretisierungsfehler*. Das durch die Verfahrensfunktion Φ definierte (explizite) Einschrittverfahren heißt *konsistent* (zur gegebenen Anfangswertaufgabe), wenn

$$\lim_{h \rightarrow 0^+} \Delta(h, f)(t, u) = 0$$

für alle $(t, u) \in [t_0, T] \times \mathbb{R}^n$ und alle $f \in F_1[t_0, T]$. Das Einschrittverfahren hat die *Konsistenzordnung* p , falls $\Delta(h, f)(t, u) = O(h^p)$ für alle $(t, u) \in [t_0, T] \times \mathbb{R}^n$ und alle $f \in F_p[t_0, T]$.

Es folgt eine einfach nachprüfbare (notwendige und hinreichende) Bedingung für Konsistenz.

Satz 6.2 Das Einschrittverfahren mit der Verfahrensfunktion Φ ist genau dann konsistent zu der gegebenen Anfangswertaufgabe, wenn

$$\lim_{h \rightarrow 0^+} \Phi(h, f)(t, u) = f(t, u)$$

für alle $(t, u) \in [t_0, T] \times \mathbb{R}^n$ und alle $f \in F_1[t_0, T]$.

Beweis: Sei z die Lösung der Anfangswertaufgabe $z' = f(s, z)$, $z(t) = u$. Es ist

$$\begin{aligned} \|\Delta(h, f)(t, u) + \Phi(h, f)(t, u) - f(t, u)\| &= \left\| \frac{z(t+h) - z(t)}{h} - z'(t) \right\| \\ &= \frac{1}{h} \left\| \int_0^h [z'(t+\tau) - z'(t)] d\tau \right\| \\ &= \frac{1}{h} \left\| \int_0^h [f(t+\tau, z(t+\tau)) - f(t, z(t))] d\tau \right\| \\ &\leq \max_{\tau \in [0, h]} \|f(t+\tau, z(t+\tau)) - f(t, z(t))\| \\ &\rightarrow 0. \end{aligned}$$

Hieraus liest man die Behauptung ab. \square

Die bisher angegebenen Einschrittverfahren sind wegen des eben angegebenen Satzes offenbar sämtlich konsistent.

Beispiele: Wir wollen für die bisher angegebenen Einschrittverfahren die Konsistenzordnung berechnen. Bei vorgegebenem $(t, u) \in [t_0, T] \times \mathbb{R}^n$ sei z jeweils die Lösung der Anfangswertaufgabe $z' = f(s, z)$, $z(t) = u$. Beim Eulerschen Polygonzugverfahren ist der lokale Diskretisierungsfehler für $f \in F_1[t_0, T]$ durch

$$\begin{aligned} \Delta(h, f)(t, u) &= \frac{z(t+h) - z(t)}{h} - f(t, u) \\ &= z'(t) + O(h) - f(t, u) \\ &= O(h) \end{aligned}$$

gegeben. Das Euler'sche Polygonzugverfahren ist also ein Verfahren erster Ordnung.

Für $f \in F_2[t_0, T]$ ist beim Verfahren von Heun

$$\begin{aligned} \Delta(h, f)(t, u) &= \frac{z(t+h) - z(t)}{h} - \frac{1}{2}[f(t, u) + f(t+h, u + hf(t, u))] \\ &= z'(t) + \frac{h}{2}z''(t) + O(h^2) - \frac{1}{2}[f(t, u) + f(t+h, u + hf(t, u))] \\ &= f(t, u) + \frac{h}{2} \frac{d}{ds} f(s, z(s)) \Big|_{s=t} + O(h^2) \\ &\quad - \frac{1}{2}[f(t, u) + f(t, u) + f_t(t, u)h + f_x(t, u)hf(t, u)] + O(h^2) \\ &= f(t, u) + \frac{h}{2}[f_t(t, u) + f_x(t, u) \underbrace{z'(t)}_{=f(t, u)}] + O(h^2) \\ &\quad - \frac{1}{2}[f(t, u) + f(t, u) + f_t(t, u)h + f_x(t, u)hf(t, u)] + O(h^2) \\ &= O(h^2). \end{aligned}$$

Das Verfahren von Heun ist also ein Verfahren zweiter Ordnung.

Man kann (mit ziemlich mühsamer Rechnung) zeigen, dass das klassische Runge-Kutta-Verfahren die Ordnung vier hat. Ein Verfahren der Ordnung drei ist in Aufgabe 3 angegeben.

Natürlich stellt sich die Frage, ob die “ziemlich mühsame Rechnung” von Maple übernommen werden kann. Dies ist in der Tat der Fall. Wir beginnen mit dem Eulerschen-Polygonzugverfahren, bei dem wir folgendermaßen vorgehen können:

```
> g:=t->f(t,z(t));
                                g := t → f(t, z(t))
> Delta:=h->(z(t+h)-z(t))/h-f(t,z(t));
                                Δ := h →  $\frac{z(t+h) - z(t)}{h} - f(t, z(t))$ 
> s:=series(Delta(h),h,2);
                                s := (D(z)(t) - f(t, z(t))) + O(h)
> s1:=subs(D(z)(t)=g(t),s);
                                s1 := O(h)
```

Durch diese Rechnung erhalten wir, dass

$$\Delta(h, f)(t, u) = O(h),$$

das Euler-Verfahren also ein Verfahren erster Ordnung ist. Entsprechend gehen wir beim Heun-Verfahren vor (wir verzichten diesmal auf das Echo):

```
> g:=t->f(t,z(t));
> k_1:=h->f(t,z(t));
> k_2:=h->f(t+h,z(t)+h*k_1(h));
> Delta:=h->(z(t+h)-z(t))/h-(1/2)*(k_1(h)+k_2(h));
> s:=series(Delta(h),h,3);
> s1:=subs((D@@2)(z)(t)=D(g)(t),s);
> s2:=subs(D(z)(t)=g(t),s1);
                                s2 := O(h^2)
```

Das Heun-Verfahren ist also ein Verfahren der Ordnung 2. Dass das klassische Runge-Kutta-Verfahren die Ordnung 4 hat, erhält man durch:

```
> restart;
> g:=t->f(t,z(t));
> k_1:=h->f(t,z(t));
> k_2:=h->f(t+(h/2),z(t)+(h/2)*k_1(h));
> k_3:=h->f(t+(h/2),z(t)+(h/2)*k_2(h));
> k_4:=h->f(t+h,z(t)+h*k_3(h));
> Delta:=h->(z(t+h)-z(t))/h-(1/6)*(k_1(h)+2*k_2(h)+2*k_3(h)+k_4(h));
> s:=series(Delta(h),h,5);
> s1:=subs((D@@4)(z)(t)=(D@@3)(g)(t),s);
> s2:=subs((D@@3)(z)(t)=(D@@2)(g)(t),s1);
> s3:=subs((D@@2)(z)(t)=D(g)(t),s2);
> s4:=subs(D(z)(t)=g(t),s3);
> simplify(%);
                                O(h^4)
```

In Aufgabe 3 kann “zu Fuß” und mit Hilfe von Maple nachgewiesen werden, dass ein gewisses Einschrittverfahren die Konsistenzordnung 3 besitzt. \square

5.4.3 Konvergenz von Einschrittverfahren

Nun wollen wir uns mit der *Konvergenz* von Einschrittverfahren zur Lösung der Anfangswertaufgabe

$$x' = f(t, x), \quad x(t_0) = x_0$$

beschäftigen, wobei wir uns zunächst darüber klar werden müssen, was unter Konvergenz zu verstehen ist. Hierzu geben wir uns ein beliebiges $t \in [t_0, T]$ vor. Man erreicht t von t_0 ausgehend in m äquidistanten Schritten der Länge $h_m := (t - t_0)/m$, erhält also eine Näherung u_m für die Lösung $x(t)$ zur Zeit t durch

$$u_0 := x_0, \quad u_{i+1} := u_i + h_m \Phi(h_m, f)(t_0 + ih_m, u_i) \quad (i = 0, \dots, m-1).$$

Es werden daher die Schrittweiten

$$H_t := \left\{ \frac{t - t_0}{m} : m \in \mathbb{N} \right\}$$

eine besondere Rolle spielen. Mit einem durch die Verfahrensfunktion Φ definierten Einschrittverfahren kann mit einer Maschenweite $h \in H_t$ die Näherung $u(t; h)$ für die Lösung $x(t)$ berechnet werden. Der Fehler

$$e(t; h) := u(t; h) - x(t)$$

heißt *globaler Diskretisierungsfehler*.

Die folgende Definition ist nun naheliegend.

Definition 6.3 Sei $x(\cdot)$ die Lösung der Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$. Ein durch eine Verfahrensfunktion Φ definiertes Einschrittverfahren heißt *konvergent*, falls

$$\lim_{m \rightarrow \infty} u(t; h_m) = x(t) \quad \text{für alle } t \in [t_0, T] \text{ und alle } f \in F_1[t_0, T].$$

Hierbei ist $h_m := (t - t_0)/m$, während $u(t; h_m)$ aus

- $u_0 := x_0$,
- Für $i = 0, \dots, m-1$:

$$u_{i+1} := u_i + h_m \Phi(h_m, f)(t_i, u_i), \quad t_{i+1} := t_i + h_m,$$

- $u(t; h_m) := u_m$

berechnet wird.

Nun kommen wir zum Konvergenzsatz für Einschrittverfahren. In seiner einfachsten Version sagt er aus, dass ein mit der gegebenen Anfangswertaufgabe konsistentes Einschrittverfahren konvergent ist¹⁸. Diese Version geben wir zunächst an.

¹⁸Es gilt hier auch die Umkehrung, worauf wir aber nicht eingehen wollen.

Satz 6.4 Mit $f \in F_1[t_0, T]$ und $x_0 \in \mathbb{R}^n$ sei die Anfangswertaufgabe

$$(P) \quad x' = f(t, x), \quad x(t_0) = x_0$$

gegeben, $x(\cdot)$ sei die eindeutige Lösung. Die Verfahrensfunktion $\Phi(h; t, u)$ (wir werden die Abhängigkeit von f unterdrücken) des betrachteten Einschrittverfahrens sei stetig auf $G := [0, h_0] \times [t_0, T] \times \mathbb{R}^n$, wobei $h_0 > 0$, und lipschitzstetig bezüglich der letzten Variablen, es existiere also eine Konstante $M > 0$ mit

$$\|\Phi(h; t, u) - \Phi(h; t, v)\| \leq M \|u - v\| \quad \text{für alle } (h, t, u), (h, t, v) \in G.$$

Dann gilt: Ist das Einschrittverfahren konsistent, so ist es auch konvergent.

Beweis: Sei $t \in (t_0, T]$ fest gewählt und $h := (t - t_0)/m$ mit einem so großen $m \in \mathbb{N}$, dass $h \leq h_0$. Sei $t_i := t_0 + ih$, $i = 0, \dots, m$. Wie schon mehrfach beschrieben, erhält man $u(t; h)$ aus

- $u_0 := x_0$,
- Für $i = 0, \dots, m - 1$:

$$u_{i+1} := u_i + h\Phi(h; t_i, u_i)$$
- $u(t; h) := u_m$

Zur Abkürzung setze man $x_i := x(t_i)$ und $e_i := u_i - x_i$. Wegen

$$\Delta(h; t_i, x_i) = \frac{x_{i+1} - x_i}{h} - \Phi(h; t_i, x_i)$$

ist

$$x_{i+1} = x_i + h[\Phi(h; t_i, x_i) + \Delta(h; t_i, x_i)]$$

und folglich

$$e_{i+1} = u_{i+1} - x_{i+1} = e_i + h[\Phi(h; t_i, u_i) - \Phi(h; t_i, x_i)] - h\Delta(h; t_i, x_i).$$

Mit

$$\sigma(h) := \max_{t \in [t_0, T]} \|\Delta(h; t, x(t))\|$$

ist dann

$$\|e_{i+1}\| \leq (1 + hM) \|e_i\| + h\sigma(h), \quad i = 0, \dots$$

Wegen $e_0 = 0$ erhält man durch Zurückspulen

$$\begin{aligned} \|u(t; h) - x(t)\| &= \|e_m\| \\ &\leq h\sigma(h) \sum_{i=0}^{m-1} (1 + hM)^i \\ &= h\sigma(h) \frac{(1 + hM)^m - 1}{hM} \\ &\leq h\sigma(h) \frac{e^{mhM} - 1}{hM} \\ &= \sigma(h) \frac{e^{M(t-t_0)} - 1}{M}. \end{aligned}$$

Wegen der vorausgesetzten Konsistenz ist $\lim_{h \rightarrow 0^+} \sigma(h) = 0$, aus der gerade eben bewiesenen Ungleichungskette folgt die Konvergenz des Einschrittverfahrens. \square

Korollar 6.5 Mit $f \in F_p[t_0, T]$ (mit $p \in \mathbb{N}$) und $x_0 \in \mathbb{R}^n$ sei die Anfangswertaufgabe

$$(P) \quad x' = f(t, x), \quad x(t_0) = x_0$$

gegeben, $x(\cdot)$ sei die eindeutige Lösung. Die Verfahrensfunktion $\Phi(h; t, u)$ (wir werden die Abhängigkeit von f unterdrücken) des betrachteten Einschrittverfahrens sei stetig auf $G := [0, h_0] \times [t_0, T] \times \mathbb{R}^n$, wobei $h_0 > 0$. Es gebe positive Konstanten M und N derart, dass

(a) Es ist

$$\|\Phi(h; t, u) - \Phi(h; t, v)\| \leq M \|u - v\| \quad \text{für alle } (h, t, u), (h, t, v) \in G.$$

(b) Für den lokalen Diskretisierungsfehler (auch hier lassen wir die Abhängigkeit von f fort) gilt die Abschätzung

$$\|\Delta(h; t, x(t))\| \leq Nh^p \quad \text{für alle } t \in [t_0, T], h \in [0, h_0].$$

Dann läßt sich der globale Diskretisierungsfehler $e(t; h) := u(t; h) - x(t)$ abschätzen durch

$$\|e(t; h)\| \leq Nh^p \frac{e^{M(t-t_0)} - 1}{M}$$

für alle $t \in [t_0, T]$ und $h = (t - t_0)/m$, wobei $m \in \mathbb{N}$ so groß sei, dass $h \leq h_0$.

Beweis: Die Aussage folgt offenbar aus dem Beweis des letzten Satzes, da $\sigma(h) \leq Nh^p$ vorausgesetzt wird. \square

Theoretisch gibt der letzte Satz die Möglichkeit, nicht nur die qualitative Aussage zu machen, dass der globale Diskretisierungsfehler die gleiche Ordnung wie der lokale Diskretisierungsfehler hat, sondern sogar eine quantitative Fehlerabschätzung anzugeben, zumindestens dann, wenn man geeignete Konstanten M und N kennt, was aber sehr selten der Fall ist.

5.4.4 Einschrittverfahren und Extrapolation

Statt vom *Lösen* einer Differentialgleichung spricht man häufig auch von ihrer *Integration*, was natürlich ist, denn die Lösung x von $x' = f(t)$, $x(t_0) = x_0$ erhält man durch das Integral $x(t) = x_0 + \int_{t_0}^t f(s) ds$. Durch jedes der vorgestellten Verfahren kann man also auch bestimmte Integrale berechnen. Aus der numerischen Integration wissen wir (eventuell), dass man die Ordnung der (zusammengesetzten) Trapezregel und der Simpson-Regel sukzessive durch *Extrapolation* erhöhen kann, was z. B. auf das Romberg-Verfahren führt. Dass ein solches Vorgehen auch bei der numerischen Behandlung von Anfangswertaufgaben möglich ist, folgt aus dem nächsten Satz. Dieser stammt

von W. Gragg (1963), sein sehr hübscher Beweis von E. Hairer, Ch. Lubich (1984)¹⁹. In ihm wird die Existenz einer asymptotischen Entwicklung des lokalen Diskretisierungsfehlers vorausgesetzt und auf eine des globalen Diskretisierungsfehlers geschlossen.

Satz 6.6 Sei $f \in F_{N+2}[t_0, T]$ und $u(t; h)$ die von dem Einschrittverfahren mit der (hinreichend glatten) Verfahrensfunktion Φ gelieferte Näherungslösung für $x(t)$, wobei $x(\cdot)$ die Lösung der Anfangswertaufgabe

$$(P) \quad x' = f(t, x), \quad x(t_0) = x_0$$

ist. Es sei $\Phi(0; t, u) = f(t, u)$ (Konsistenzbedingung), ferner gelte mit einem $p \geq 1$ die folgende Entwicklung des lokalen Diskretisierungsfehlers:

$$(*) \quad \Delta(h, t, x(t)) = \frac{x(t+h) - x(t)}{h} - \Phi(h; t, x(t)) = d_p(t)h^p + \dots + d_N(t)h^N + O(h^{N+1}),$$

insbesondere habe das Verfahren also die Ordnung p . Dann besitzt $u(t; h)$ eine asymptotische Entwicklung der Form

$$u(t; h) = x(t) + e_p(t)h^p + e_{p+1}(t)h^{p+1} + \dots + e_N(t)h^N + E_{N+1}(t; h)h^{N+1}$$

für alle $t \in [t_0, T]$, $h = h_m = (t - t_0)/m$, $m \in \mathbb{N}$. Dabei sind die Funktionen e_i von h unabhängig und das Restglied $E_{N+1}(t; h)$ ist bei festem t für alle $h = h_m$, $m \in \mathbb{N}$, beschränkt.

Beweis: Zunächst wird nur ausgenutzt, dass wegen (*) insbesondere

$$\Delta(h; t, x(t)) = \frac{x(t+h) - x(t)}{h} - \Phi(h; t, x(t)) = d_p(t)h^p + O(h^{p+1}).$$

Wir zeigen, dass eine stetig differenzierbare Funktion e_p existiert mit

$$u(t; h) - x(t) = e_p(t)h^p + O(h^{p+1}).$$

Hierzu definieren wir

$$\hat{u}(t; h) := u(t; h) - e_p(t)h^p$$

mit noch unbestimmter Funktion e_p . Dann ist

$$\begin{aligned} \hat{u}(t+h; h) &= u(t+h; h) - e_p(t+h)h^p \\ &= u(t; h) + h\Phi(h; t, u(t; h)) - e_p(t+h)h^p \\ &= \hat{u}(t; h) + e_p(t)h^p + h\Phi(h; t, \hat{u}(t; h) + e_p(t)h^p) - e_p(t+h)h^p \\ &= \hat{u}(t; h) + h\hat{\Phi}(h; t, \hat{u}(t; h)) \end{aligned}$$

mit

$$\hat{\Phi}(h; t, u) := \Phi(h; t, u + e_p(t)h^p) - [e_p(t+h) - e_p(t)]h^{p-1}.$$

¹⁹HAIRER, E., C. LUBICH (1984) "Asymptotic expansions of the global error of fixed-stepsize methods." Numer. Math. 45, 345–360.

Also ist $\hat{u}(t; h)$ Resultat eines Schrittes eines Einschrittverfahrens mit der Verfahrensfunktion $\hat{\Phi}$. Nun entwickeln wir den lokalen Diskretisierungsfehler des zu $\hat{\Phi}$ gehörenden Einschrittverfahrens:

$$\begin{aligned}\hat{\Delta}(h; t, x(t)) &= \frac{x(t+h) - x(t)}{h} - \hat{\Phi}(h; t, x(t)) \\ &= \Delta(h; t, x(t)) + [\Phi(h; t, x(t)) - \hat{\Phi}(h; t, x(t))] \\ &= d_p(t)h^p + O(h^{p+1}) + [\Phi(h; t, x(t)) - \hat{\Phi}(h; t, x(t))] \\ &= [\Phi(h; t, x(t)) - \Phi(h; t, x(t) + e_p(t)h^p)] \\ &\quad + [d_p(t) + e'_p(t)]h^p + O(h^{p+1}) \\ &= [d_p(t) - f_x(t, x(t))e_p(t) + e'_p(t)]h^p + O(h^{p+1}).\end{aligned}$$

Hierbei haben wir ausgenutzt, daß

$$\begin{aligned}\Phi(h; t, x(t) + e_p(t)h^p) &= \Phi(h; t, x(t)) + \Phi_x(h; t, x(t))e_p(t)h^p + O(h^{p+1}) \\ &= \Phi(h; t, x(t)) + \Phi_x(0; t, x(t))e_p(t)h^p + O(h^{p+1}) \\ &= \Phi(h; t, x(t)) + f_x(t, x(t))e_p(t)h^p + O(h^{p+1}).\end{aligned}$$

Daher bestimme man e_p als Lösung der linearen Anfangswertaufgabe

$$e'_p = f_x(t, x(t))e_p - d_p(t), \quad e_p(t_0) = 0.$$

Dann ist durch die Verfahrensfunktion $\hat{\Phi}$ ein Einschrittverfahren der Ordnung $p+1$ gegeben, so dass die zugehörige Konvergenzordnung auch mindestens $p+1$ ist. Also ist

$$u(t; h) - x(t) = e_p(t)h^p + \hat{u}(t; h) - x(t) = e_p(t)h^p + O(h^{p+1}).$$

Ist $N = p$, so ist der Beweis abgeschlossen, andernfalls kann die obige Konstruktion fortgesetzt werden, wobei Φ durch $\hat{\Phi}$ und p durch $p+1$ zu ersetzen ist. \square

Eine asymptotische Entwicklung des globalen Diskretisierungsfehlers ist aus mindestens zwei Gründen wichtig. Zum einen kann man in diesem Falle den globalen Diskretisierungsfehler abschätzen bzw. genauer schätzen. Es gelte also etwa

$$u(t; h) - x(t) = e_p(t)h^p + O(h^{p+1}).$$

Hat man mit der Schrittweite h den Näherungswert $u(t; h)$ für $x(t)$ bestimmt, so berechne man anschließend mit einer anderen Schrittweite, z. B. mit der Schrittweite $\frac{1}{2}h$, für dasselbe t den Näherungswert $u(t; \frac{1}{2}h)$. Aus

$$\begin{aligned}u(t; h) - x(t) &= e_p(t)h^p + O(h^{p+1}), \\ u(t; \frac{1}{2}h) - x(t) &= e_p(t)\left(\frac{h}{2}\right)^p + O(h^{p+1})\end{aligned}$$

folgt durch Subtraktion

$$u(t; h) - u(t; \frac{1}{2}h) = e_p(t)\left(\frac{h}{2}\right)^p (2^p - 1) + O(h^{p+1})$$

und anschließend

$$e_p(t) \left(\frac{h}{2}\right)^p = \frac{u(t; h) - u(t; \frac{1}{2}h)}{2^p - 1} + O(h^{p+1}).$$

Dann ist schließlich durch

$$u(t; \frac{1}{2}h) - x(t) = \frac{u(t; h) - u(t; \frac{1}{2}h)}{2^p - 1} + O(h^{p+1})$$

eine Schätzung des Fehlers $e(t; \frac{1}{2}h)$ gelungen. Beim Runge-Kutta Verfahren ist $p = 4$, hier lautet die entsprechende Schätzung

$$u(t; \frac{1}{2}h) - x(t) \approx \frac{u(t; h) - u(t; \frac{1}{2}h)}{15}.$$

Beispiel: Wir wollen die Qualität des gerade eben vorgestellten Fehlerschätzers an einem Beispiel testen. Hierzu lösen wir die Anfangswertaufgabe $x' = x - t^2 + 1$, $x(0) = 0.5$, deren Lösung $x(t) = (1+t)^2 - \frac{1}{2}e^t$ ist, mit dem Runge-Kutta-Verfahren mit $h := 0.2$ und $\frac{1}{2}h = 0.1$ auf dem Intervall $[0, 2]$. Wir erhalten die folgenden Werte für $t = 2$:

$u(2; h)$	$u(2; \frac{1}{2}h)$	$x(2)$
5.30536300069265	5.30546496022735	5.30547195053467

und damit

$$\begin{aligned} \frac{u(2; h) - u(2; \frac{1}{2}h)}{15} &= -6.797302313129213e - 06, \\ u(2; \frac{1}{2}h) - x(2) &= -6.990307323206935e - 06. \end{aligned}$$

Das ist also schon eine verblüffend gute Schätzung. □

Die Idee der Extrapolation wollen wir nur ganz kurz beschreiben. Angenommen, mit dem Runge-Kutta-Verfahren seien $u(t; h)$ und $u(t; \frac{1}{2}h)$ berechnet worden. Dann ist

$$\begin{aligned} u(t; h) &= x(t) + e_4(t)h^4 + e_5(t)h^5 + \dots \\ u(t; \frac{1}{2}h) &= x(t) + \frac{1}{16}e_4(t)h^4 + \frac{1}{32}e_5(t)h^5 + \dots, \end{aligned}$$

woraus man den (hoffentlich) verbesserten Wert

$$u^*(t; \frac{1}{2}h) := \frac{16u(t; \frac{1}{2}h) - u(t; h)}{15} = x(t) - \frac{1}{30}e_5(t)h^5 + \dots$$

erhält.

Beispiel: Wir setzen das eben angegebene Beispiel fort und geben auch noch den extrapolierten Wert $u^*(2; \frac{1}{2}h)$ an:

$u(2; h)$	$u(2; \frac{1}{2}h)$	$u^*(2; \frac{1}{2}h)$	$x(2)$
5.30536300069265	5.30546496022735	5.30547175752966	5.30547195053467

Der Erfolg ist offensichtlich. □

5.4.5 Schrittweitensteuerung

Nun wollen wir einiges zu dem wichtigen Problem der Schrittweitensteuerung sagen, natürlich nach wie vor bei Einschrittverfahren. Einer der Vorteile von Einschrittverfahren gegenüber sogenannten Mehrschrittverfahren, auf die wir nicht mehr eingehen werden, ist, dass eine Schrittweitensteuerung im Prinzip sehr einfach durchführbar ist. Das Problem besteht darin, dass man zwei Ziele verfolgen will, die sich gegenseitig ausschließen:

- Zum einen will man die Schrittweite möglichst groß wählen, um den Arbeitsaufwand zur Integration einer Anfangswertaufgabe über ein vorgegebenes Zeitintervall möglichst klein zu halten,
- andererseits will man die Lösung möglichst genau berechnen, wozu i. allg. eine kleine Schrittweite nötig ist,
- außerdem will man die Schrittweite den jeweiligen Verhältnissen anpassen, also dort feiner diskretisieren, wo es nötig ist, sonst aber mit einer größeren Schrittweite arbeiten.

Wir beschreiben nun eine mögliche Form der Schrittweitensteuerung (und folgen hier J. Stoer, R. Bulirsch (1990, S. 127 ff.)).

Gegeben seien (t_0, x_0) und $\epsilon > 0$, gesucht ist eine möglichst große Schrittweite $h > 0$ mit $\|e(t_0 + h; h)\| \leq \epsilon$ oder wenigstens $\|e(t_0 + h; h)\| \approx \epsilon$. Hierbei bedeutet $e(t; h)$ wieder den globalen Diskretisierungsfehler in t bei Verwendung eines Einschrittverfahrens mit der Schrittweite h . Hierbei sollte ϵ nicht zu klein gewählt werden, etwa

$$\epsilon \approx \text{eps } K \quad \text{mit} \quad K := \max_{t \in [t_0, t_0 + h]} \|x(t)\|,$$

wobei eps die Maschinengenauigkeit bedeutet. Die Maschinengenauigkeit eps (kleinste positive Zahl mit $1 + \text{eps} \neq 1$) wird häufig durch das folgende (MATLAB-) Programm berechnet:

```
u=1;
while 1+u~=1
    u=u/2;
end;
u=2*u;
```

Als Resultat erhalten wir (mit `format long g`) `u=2.22044604925031e-16`, was genau mit dem Resultat der MATLAB-Funktion `eps` übereinstimmt. Wir nehmen an, dass ein Verfahren der Ordnung p benutzt wird, so dass $e(t; h) = e_p(t)h^p + O(h^{p+1})$. Wegen $e_p(t_0) = 0$ ist $e_p(t) \approx (t - t_0)e'_p(t_0)$. Es ist

$$e(t_0 + h; h) \approx e_p(t_0 + h)h^p \approx e'_p(t_0)h^{p+1},$$

wenn also $\|e(t_0 + h; h)\| \approx \epsilon$ sein soll, hat man h so zu wählen, dass $\|e'_p(t_0)\| h^{p+1} \approx \epsilon$. Um hieraus h zu berechnen, müßte man $\|e'_p(t_0)\|$ kennen, was aber i. allg. nicht der Fall ist. Daher versuchen wir, $\|e'_p(t_0)\|$ zu schätzen.

Man wähle eine Schrittweite $H > 0$ und berechne $u(t_0 + H; H)$ und $u(t_0 + H; \frac{1}{2}H)$. Dann ist (siehe oben)

$$e(t_0 + H; \frac{1}{2}H) \approx \frac{u(t_0 + H; H) - u(t_0 + H; \frac{1}{2}H)}{2^p - 1}.$$

Andererseits ist

$$e(t_0 + H; \frac{1}{2}H) \approx e_p(t_0 + H) \left(\frac{H}{2}\right)^p \approx e'_p(t_0) H \left(\frac{H}{2}\right)^p$$

und daher

$$\|e'_p(t_0)\| \approx \frac{1}{H^{p+1}} \frac{2^p}{2^p - 1} \|u(t_0 + H; H) - u(t_0 + H; \frac{1}{2}H)\|.$$

Aus $\|e'_p(t_0)\| h^{p+1} \approx \epsilon$ erhält man also h aus

$$\frac{H}{h} \approx \sqrt[p+1]{\frac{2^p}{2^p - 1} \frac{\|u(t_0 + H; H) - u(t_0 + H; \frac{1}{2}H)\|}{\epsilon}}.$$

Ist $H/h \gg 2$, so ist $\|e(t_0 + H; \frac{1}{2}H)\| \gg 2\epsilon$, da

$$\frac{H}{h} \approx \sqrt[p+1]{\frac{2^p}{\epsilon} \|e(t_0 + H; \frac{1}{2}H)\|}.$$

Dann ersetzt man H durch den (wesentlich) kleineren Wert $2h$ und beginnt die Rechnung noch einmal. Dies sieht dann also folgendermaßen aus:

- Gegeben (t_0, x_0) (Anfangszeit, Anfangszustand), $H > 0$ (Anfangsschrittweite), $\epsilon > 0$ (gewünschte Genauigkeit), $T > t_0$ (Endzeit, die Lösung soll auf $[t_0, T]$ berechnet werden).

(1) Berechne $u(t_0 + H; H)$ und $u(t_0 + H; \frac{1}{2}H)$. Anschließend berechne man

$$h := H \sqrt[p+1]{\frac{2^p - 1}{2^p} \frac{\epsilon}{\|u(t_0 + H; H) - u(t_0 + H; \frac{1}{2}H)\|}}.$$

(2) Falls $H/h \gg 2$, dann setze $H := 2h$ und gehe zu (1).

(3) Setze $(t_0, x_0) := (t_0 + H, u(t_0 + H; \frac{1}{2}H))$, $H := 2h$.

(4) Falls $t_0 \geq T$, dann: STOP, die Anfangswertaufgabe ist auf $[t_0, T]$ numerisch gelöst. Andernfalls gehe man nach (1).

5.4.6 Aufgaben

1. Man zeige: Ist $p \in \mathcal{P}_3$ (Menge der kubischen Polynome), so ist

$$\int_a^b p(t) dt = \frac{b-a}{6} [p(a) + 4p(\frac{1}{2}(a+b)) + p(b)].$$

2. Man betrachte ein Einschrittverfahren mit der Verfahrensfunktion

$$\Phi(h, f)(t, u) := a_1 f(t, u) + a_2 f(t + b_1 h, u + b_2 h f(t, u))$$

und zeige, dass dieses die Ordnung 2 besitzt, falls

$$a_1 + a_2 = 1, \quad a_2 b_1 = \frac{1}{2}, \quad a_2 b_2 = \frac{1}{2}.$$

Spezialfälle erhält man übrigens für $a_1 = 0, a_2 = 1, b_1 = b_2 = \frac{1}{2}$ (modifiziertes Euler-Verfahren) und für $a_1 = a_2 = \frac{1}{2}, b_1 = b_2 = 1$ (Heun-Verfahren).

3. Man betrachte ein Einschrittverfahren mit der Verfahrensfunktion

$$\Phi(h, f)(t, u) := \frac{1}{4} k_1 + \frac{3}{4} k_3,$$

wobei

$$k_1 := f(t, u), \quad k_2 := f(t + \frac{1}{3}h, u + \frac{1}{3}h k_1), \quad k_3 := f(t + \frac{2}{3}h, u + \frac{2}{3}h k_2).$$

Man zeige, dass dies ein Verfahren der Ordnung 3 ist. Hierbei darf man sich auf den Fall einer Differentialgleichung erster Ordnung, also $n = 1$, beschränken. Anschließend löse man diese Aufgabe mit Maple.

4. Man betrachte ein Einschrittverfahren mit der Verfahrensfunktion

$$\Phi(h, f)(t, u) := \frac{1}{6}(k_1 + 4k_2 + k_3),$$

wobei

$$k_1 := f(t, u), \quad k_2 := f(t + \frac{1}{2}h, u + \frac{1}{2}h k_1), \quad k_3 := f(t + h, u - h k_1 + 2h k_2).$$

Man zeige, dass dies ein Verfahren der Ordnung 3 ist. Hierbei darf man sich auf den Fall einer Differentialgleichung erster Ordnung, also $n = 1$, beschränken und die Aufgabe mit Maple lösen.

5. Man bestimme²⁰ die exakte Lösung der Anfangswertaufgabe $x' = (2/t)x, x(1) = 1$. Anschließend bestimme man einen analytischen Ausdruck für die durch das Eulersche Polygonzugverfahren erhaltene Näherung und gebe den globalen und den lokalen Diskretisierungsfehler an.
6. Man schreibe ein MATLAB-Programm für das klassische Runge-Kutta-Verfahren mit automatischer Schrittweitensteuerung. Anschließend teste man das Programm an der Anfangswertaufgabe (siehe Stoer-Bulirsch)

$$x' = -200tx^2, \quad x(-3) = \frac{1}{901},$$

welche auf $[-3, 0]$ zu lösen sei.

²⁰Diese Aufgabe ist dem Lehrbuch

R. KRESS (1998) *Numerical Analysis*. Springer-Verlag, New York-Berlin-Heidelberg. entnommen.

