

Modellierung mit Differentialgleichungen

Lars Grüne
Mathematisches Institut
Fakultät für Mathematik und Physik
Universität Bayreuth
95440 Bayreuth
lars.gruene@uni-bayreuth.de
www.uni-bayreuth.de/departments/math/~lgruene/

Vorlesungsskript
Sommersemester 2003

Vorwort

Dieses Skript ist im Rahmen einer gleichnamigen Vorlesung entstanden, die ich im Sommersemester 2003 an der Universität Bayreuth gehalten habe.

Die einzelnen Kapitel des Skriptes wurden auf Basis verschiedener Lehrbücher und Monographien erstellt, Hinweise dazu finden sich am Ende jedes Kapitels. Besonders bedanken möchte ich mich bei Fabian Wirth (<http://www.math.uni-bremen.de/~fabian/>) für seine sehr hilfreichen Notizen zu einer im Wintersemester 2002/2003 an der J.W. Goethe-Universität Frankfurt am Main gehaltenen Vorlesung über mathematische Modellierung. Ebenfalls möchte ich mich bei den aufmerksamen StudentInnen bedanken, die viele kleinere Fehler gefunden haben, die in dieser Version korrigiert werden konnten.

Eine elektronische Version dieses Skripts sowie die zu dieser Vorlesung gehörigen Übungsaufgaben finden sich im WWW unter dem Link “Lehrveranstaltungen” auf der Seite <http://www.uni-bayreuth.de/departments/math/~lgruene/>.

Bayreuth, Juli 2003

LARS GRÜNE

Inhaltsverzeichnis

Vorwort	i
1 Einführung	1
1.1 Überblick	1
1.2 Gewöhnliche Differentialgleichungen	2
1.2.1 Definition	2
1.2.2 Anfangswertprobleme	3
1.2.3 Ein Existenz- und Eindeutigkeitssatz	4
1.2.4 Grafische Darstellung der Lösungen	7
1.3 Literaturhinweise	8
2 Biologie	9
2.1 Populationsdynamik für eine Art	9
2.1.1 Differenzen- und Differentialgleichungen	9
2.1.2 Einfache Modelle	11
2.1.3 Eine Anwendung des Modells	16
2.1.4 Abschließende Diskussion	19
2.2 Populationsdynamik für mehrere Arten	19
2.2.1 Das Räuber-Beute Modell mit unbeschränkten Ressourcen	20
2.2.2 Das Räuber-Beute Modell mit beschränkten Ressourcen	24
2.2.3 Verallgemeinerung auf n Arten	27
2.3 Anwendungen der Populationsdynamik	28
2.3.1 Auswirkungen der Befischung	28
2.3.2 Selektion gleichartiger Spezies	29
2.4 Ausbreitung von Epidemien	32
2.5 Literaturhinweise	35

3	Mechanik	37
3.1	Mechanisch–technischer Ansatz	37
3.1.1	Translationale Bewegungselemente	37
3.1.2	Einfache translationale Modelle	40
3.1.3	Rotationale Bewegungselemente	45
3.1.4	Das Pendel	48
3.2	Lagrange–Gleichungen und Hamilton–Formalismus	54
3.2.1	Lagrange–Gleichungen	54
3.2.2	Dissipative Systeme	57
3.2.3	Die Hamilton’sche Methode	58
4	Finanzmathematik	61
4.1	Stochastische Differentialgleichungen	62
4.1.1	Einige stochastische Grundbegriffe	62
4.1.2	Der Wiener Prozess	67
4.1.3	Das Ito–Integral	68
4.2	Die Bewertung von Optionen	72
4.2.1	Optionen	72
4.2.2	Das Ito–Lemma	74
4.2.3	SDG–Modelle für Kursverläufe	76
4.2.4	Die Monte–Carlo Methode	78
4.2.5	Die Black–Scholes–Gleichung	79
4.3	Literaturhinweise	82
	Index	83

Kapitel 1

Einführung

1.1 Überblick

In dieser Vorlesung werden wir uns mit Modellen realer Phänomene beschäftigen, die auf Differentialgleichungen beruhen. Wir wollen für eine Reihe von Anwendungen Modelle entwickeln, die zu Grunde liegenden Mechanismen untersuchen und Modellannahmen diskutieren. Zudem werden wir einige Aspekte der qualitativen Analyse von Differentialgleichungen kennen lernen, die zur Untersuchung des Verhaltens der Modelle nützlich sind. Diese werden dabei in eher informeller Weise eingeführt, d.h. wir werden auf Beweise ganz verzichten oder nur die Beweisideen geben. Interessierten TeilnehmerInnen sei zur Vertiefung der Besuch von Vorlesungen über Differentialgleichungen empfohlen.

Speziell werden wir in dieser Vorlesung auf drei Anwendungsgebiete eingehen

- Biologie
- Mechanik
- Finanzmathematik

Jedes dieser Teilgebiete ist dabei natürlich so groß, dass man es in dieser zweistündigen Vorlesung nur auszugsweise betrachten kann. Wir werden uns deshalb auf spezielle Aspekte aus den Gebieten beschränken. In der Biologie werden dies Wachstumsmodelle sein, mit denen man das Wachstum von einer einzelnen Art oder von mehreren konkurrierenden Arten beschreiben kann. In der Mechanik werden wir aus den Bewegungsgesetzen der klassischen Mechanik modulare Modelle für mechanische Systeme herleiten. In der Finanzmathematik schließlich werden uns mit der Bewertung von Optionen beschäftigen und unter anderem die Black–Scholes Gleichung herleiten.

Die verschiedenen Teilgebiete bieten hierbei die Möglichkeit, unterschiedliche Techniken und Möglichkeiten aber auch Beschränkungen der Modellierung kennen zu lernen. In der Biologie werden wir sehen, dass zwar die einzelnen Elemente der Modelle durchaus mathematisch stringent begründet werden können, die Aussagekraft der Gesamtmodelle aber dadurch stark beeinflusst wird, dass viele Einwirkungen von außen in der Modellierung nicht berücksichtigt werden können, wenn das mathematische Modell noch handhabbar sein soll.

Man muss also *Modellannahmen* einführen, z.B. dass ein Ökosystem abgeschlossen ist und nicht von außen beeinflusst wird oder dass die Einflüsse der wechselnden Jahreszeiten vernachlässigt werden. Auch in der Mechanik kommt man um Modellannahmen nicht herum, allerdings können hier die Situationen, in denen die vernachlässigten Effekte überhaupt einen Einfluss haben i.A. besser eingeschätzt werden. Ein Beispiel dafür ist die Haftreibung, die oft vernachlässigt wird, da sie sich mit Differentialgleichungen nur schlecht modellieren lässt. Wenn ein mechanisches System aber z.B. immer in hinreichend schneller Bewegung ist, so tritt dieser Effekt gar nicht auf. Das Kapitel über die Mechanik ist zudem sicherlich der mathematisch “sauberste” Teil der Vorlesung, da sich die mechanischen Gesetze in sehr direkter Weise in die mathematische Formelsprache übersetzen lassen. Ganz anders ist es in der Finanzmathematik. Hier werden Differentialgleichungen verwendet, die die zu Grunde liegenden Kursverläufe im Wesentlichen phänomenologisch modellieren: man verwendet Gleichungen, deren Lösungen sich in etwa wie reale Kursverläufe verhalten, ohne dass man tiefere Gesetzmäßigkeiten (die man hier ja oft gar nicht kennt) in der Modellierung berücksichtigt. In diesem Kapitel werden wir auch weniger als bei den vorhergehenden Themen auf die eigentliche Modellbildung eingehen und statt dessen die Anwendungen der Modelle auf finanzmathematische Probleme — hier speziell die Bewertung von Optionen — genauer betrachten.

Im größeren Teil dieser Vorlesung werden wir uns in der Modellierung auf *gewöhnliche* Differentialgleichungen im klassischen Sinne beschränken. Eine Ausnahme bildet das Kapitel 4 über Finanzmathematik, in dem *stochastische* gewöhnliche Differentialgleichungen im Mittelpunkt stehen; eine Einführung in die Theorie dieser Gleichungen werden wir in Kapitel 4 geben. Gelegentlich werden uns bei unseren Überlegungen auch noch weitere Typen von Differentialgleichungen begegnen, z.B. *Delay-Differentialgleichungen* in Abschnitt 2.4 und *partielle Differentialgleichungen* in Abschnitt 4.2.5.

Im Rahmen unserer Betrachtungen werden wir die benötigten theoretischen Resultate dort zitieren (manchmal auch beweisen), wo wir sie verwenden. Bevor wir mit der Modellierung beginnen können, benötigen wir aber zumindest ein theoretisches Grundgerüst mit einigen Basisdefinitionen und Resultaten zu den gewöhnlichen Differentialgleichungen, das der nun folgende Abschnitt bereit stellt.

1.2 Gewöhnliche Differentialgleichungen

In diesem Abschnitt werden wir die grundlegenden Gleichungen definieren, mit denen wir uns in dieser Vorlesung beschäftigen wollen. Als Hauptresultat dieses Abschnittes werden wir einen Existenz- und Eindeutigkeitssatz formulieren und beweisen. Zudem werden wir zwei verschiedene grafische Darstellungsmöglichkeiten für die Lösungen kennen lernen.

1.2.1 Definition

Eine gewöhnliche Differentialgleichung setzt die Ableitung einer Funktion $x : \mathbb{R} \rightarrow \mathbb{R}^n$ nach ihrem (eindimensionalen) Argument mit der Funktion selbst in Beziehung. Formal beschreibt dies die folgende Definition.

Definition 1.1 Ein *gewöhnliche Differentialgleichung* (DGL) im \mathbb{R}^n , $n \in \mathbb{N}$, ist gegeben durch die Gleichung

$$\frac{d}{dt}x(t) = f(t, x(t)), \quad (1.1)$$

wobei $f : D \rightarrow \mathbb{R}^n$ eine stetige Funktion ist und *Vektorfeld* genannt wird, deren Definitionsbereich D eine offene Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ ist.

Eine *Lösung* von (1.1) ist eine stetig differenzierbare Funktion $x : \mathbb{R} \rightarrow \mathbb{R}^n$, die (1.1) erfüllt. \square

Einige Anmerkungen zur Notation bzw. Sprechweise:

- Die unabhängige Variable t werden wir üblicherweise als Zeit interpretieren, obwohl (abhängig vom modellierten Sachverhalt) gelegentlich auch andere Interpretationen möglich sind.
- Statt $\frac{d}{dt}x(t)$ schreiben wir oft kurz $\dot{x}(t)$.
- Die Lösungsfunktion $x(t)$ nennen wir auch *Lösungskurve* oder (*Lösungs-*)*Trajektorie*.
- Falls das Vektorfeld f nicht von t abhängt, also $\dot{x}(t) = f(x(t))$ ist, nennen wir die Differentialgleichung *autonom*.

1.2.2 Anfangswertprobleme

Eine gewöhnliche Differentialgleichung besitzt im Allgemeinen unendlich viele Lösungen. Als Beispiel betrachte die (sehr einfache) eindimensionale DGL mit $f(x, t) = x$, also

$$\dot{x}(t) = x(t)$$

mit $x(t) \in \mathbb{R}$. Betrachte die Funktion $x(t) = Ce^t$ mit beliebigem $C \in \mathbb{R}$. Dann gilt

$$\dot{x}(t) = \frac{d}{dt}Ce^t = Ce^t = x(t).$$

Für jedes feste C löst Ce^t die obige DGL, es gibt also unendlich viele Lösungen.

Um *eindeutige* Lösungen zu erhalten, müssen wir eine weitere Bedingung festlegen. Dies geschieht in der folgenden Definition.

Definition 1.2 Ein *Anfangswertproblem* für die gewöhnliche Differentialgleichung (1.1) besteht darin, zu gegebenem $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}^n$ eine Lösungsfunktion $x(t)$ zu finden, die (1.1) erfüllt und für die darüberhinaus die Gleichung

$$x(t_0) = x_0 \quad (1.2)$$

gilt. \square

Notation und Sprechweisen:

- Für die Lösung $x(t)$, die (1.1) und (1.2) erfüllt, schreiben wir $x(t; t_0, x_0)$. Im Spezialfall $t_0 = 0$ werden wir oft kurz $x(t; x_0)$ schreiben.
- Die Zeit $t_0 \in \mathbb{R}$ bezeichnen wir als *Anfangszeit*, den Wert $x_0 \in \mathbb{R}^n$ als *Anfangswert*. Das Paar (t_0, x_0) bezeichnen wir als *Anfangsbedingung*, ebenso nennen wir die Gleichung (1.2) *Anfangsbedingung*.

Bemerkung 1.3 Eine stetig differenzierbare Funktion $x : I \rightarrow \mathbb{R}^n$ löst das Anfangswertproblem (1.1), (1.2) für ein $t_0 \in I$ und ein $x_0 \in \mathbb{R}^n$ genau dann, wenn sie für alle $t \in I$ die *Integralgleichung*

$$x(t) = x_0 + \int_{t_0}^t f(\tau, x(\tau)) d\tau \quad (1.3)$$

erfüllt. Dies folgt sofort durch Integrieren von (1.1) bzgl. t bzw. durch Differenzieren von (1.3) nach t unter Verwendung des Hauptsatzes der Differential- und Integralrechnung. \square

1.2.3 Ein Existenz- und Eindeutigkeitsatz

Unter geeigneten Bedingungen an f können wir einen Existenz- und Eindeutigkeitsatz für Anfangswertprobleme der Form (1.1), (1.2) erhalten.

Satz 1.4 Betrachte die gewöhnliche Differentialgleichung (1.1) für ein $f : D \rightarrow \mathbb{R}^n$ mit $D \subset \mathbb{R} \times \mathbb{R}^n$. Das Vektorfeld f sei stetig, darüberhinaus sei f Lipschitz-stetig im zweiten Argument im folgenden Sinne: Für jede kompakte Teilmenge $K \subset D$ existiere eine Konstante $L > 0$, so dass die Ungleichung

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\|$$

gilt für alle $t \in \mathbb{R}$ und $x, y \in K$ mit $(t, x), (t, y) \in K$.

Dann gibt es für jede Anfangsbedingung $(t_0, x_0) \in D$ genau eine Lösung $x(t; t_0, x_0)$ des Anfangswertproblems (1.1), (1.2). Diese ist definiert für alle t aus einem offenen *maximalen Existenzintervall* $I_{t_0, x_0} \subseteq \mathbb{R}$ mit $t_0 \in I_{t_0, x_0}$.

Beweis: Da dieser Satz grundlegend für die gesamte Theorie ist, werden wir ihn (im Gegensatz zu den meisten anderen Sätzen in dieser Vorlesung) ausführlich beweisen.

Wir zeigen zunächst, dass für jede Anfangsbedingung $(t_0, x_0) \in D$ ein offenes Intervall J um t_0 existiert, auf dem die Lösung eindeutig ist. Dieses Intervall wählen wir wie folgt: Zunächst wählen wir eine beschränkte offene Umgebung $U = I \times B_\delta(x_0)$ von (t_0, x_0) in D deren (kompakter) Abschluss \overline{U} ebenfalls in D liegt, und für die I ein offenes Intervall und $B_\delta(x_0)$ der offene Ball mit Radius δ um x_0 im \mathbb{R}^n ist (dies ist möglich, da D eine offene Menge ist). Da f stetig ist und \overline{U} kompakt ist, existiert eine Konstante M , so dass $\|f(t, x)\| \leq M$ für alle $(t, x) \in \overline{U}$ gilt. Wir wählen nun $J = (t_0 - a, t_0 + a)$ wobei $a > 0$ so gewählt ist, dass $J \subseteq I$ gilt und $La < 1$ sowie $Ma < \delta$ erfüllt ist, wobei L die Lipschitz-Konstante aus der Annahme für $K = \overline{U}$ ist.

Nun verwenden wir zum Beweis der Existenz und Eindeutigkeit der Lösung auf J den Banach'schen Fixpunktsatz, den wir auf den Banachraum $\mathcal{B} = C(J, \overline{B}_\delta(x_0))$ der stetigen Funktionen von J nach $\overline{B}_\delta(x_0) \subseteq \mathbb{R}^n$ anwenden. Auf \mathcal{B} definieren wir die Abbildung

$$T : \mathcal{B} \rightarrow \mathcal{B}, \quad T(x)(t) := x_0 + \int_{t_0}^t f(\tau, x(\tau)) d\tau.$$

Beachte, dass für jedes $t \in J$ und jedes $x \in \mathcal{B}$ die Ungleichung

$$\|T(x)(t) - x_0\| = \left\| \int_{t_0}^t f(\tau, x(\tau)) d\tau \right\| \leq \left| \int_{t_0}^t \underbrace{\|f(\tau, x(\tau))\|}_{\leq M, \text{ weil } (\tau, x(\tau)) \in \overline{U}} d\tau \right| \leq aM \leq \delta$$

gilt, weswegen $T(x)(t) \in \overline{B}_\delta(x_0)$ und damit $T(x) \in \mathcal{B}$ liegt.

Um den Banach'schen Fixpunktsatz anzuwenden, müssen wir zum einen nachweisen, dass \mathcal{B} ein Banachraum ist. Dies ist mit der Wahl der Norm $\|x\|_\infty := \sup_{t \in J} \|x(t)\|$ tatsächlich der Fall, wir werden es hier aber nicht explizit nachrechnen. Zum anderen müssen wir zeigen, dass T eine Kontraktion auf \mathcal{B} ist, also

$$\|T(x) - T(y)\|_\infty \leq k \|x - y\|_\infty$$

gilt für alle $x, y \in \mathcal{B}$ und ein $k < 1$. Diese Eigenschaft folgt für $k = La < 1$ aus

$$\begin{aligned} \|T(x) - T(y)\|_\infty &= \sup_{t \in J} \left\| \int_{t_0}^t f(\tau, x(\tau)) d\tau - \int_{t_0}^t f(\tau, y(\tau)) d\tau \right\| \\ &\leq \sup_{t \in J} \left| \int_{t_0}^t \underbrace{\|f(\tau, x(\tau)) - f(\tau, y(\tau))\|}_{\leq L\|x(\tau) - y(\tau)\| \leq L\|x - y\|_\infty} d\tau \right| \\ &\leq \sup_{t \in J} |t - t_0| L \|x - y\|_\infty = aL \|x - y\|_\infty. \end{aligned}$$

Also sind alle Voraussetzungen des Banach'schen Fixpunktsatzes erfüllt, weswegen T einen eindeutigen Fixpunkt $x \in \mathcal{B}$ (also eine "Fixpunktfunktion") besitzt. Da dieser Fixpunkt x offenbar die Integralgleichung (1.3) erfüllt, ist er eine Lösung des Anfangswertproblems. Umgekehrt ist jede weitere Lösung des Anfangswertproblems ein Fixpunkt von T ; da T aber einen eindeutigen Fixpunkt besitzt, folgt die Eindeutigkeit der Lösung.

Wir wissen also, dass eine Lösung $x(t) = x(t; t_0, x_0)$ existiert, die auf J eindeutig ist. Betrachte nun das maximale Existenzintervall I der Lösung $x(t)$, d.h. I sei so gewählt, dass sich die Lösung über den Rand von I hinaus nicht mehr fortsetzen lässt. Es bleibt zu zeigen, dass x auf ganz I eindeutig ist. Sei dazu \tilde{x} eine weitere Lösung des Anfangswertproblems, die auf einem Intervall \tilde{I} existiert und die für ein $t \in I \cap \tilde{I}$ nicht mit x übereinstimmt, also $x(t) \neq \tilde{x}(t)$. O.b.d.A. sei $t > t_0$. Da beide Lösungen auf $(t_0 - a, t_0 + a)$ übereinstimmen müssen und stetig sind, existieren $t_2 > t_1 \geq t_0 + a$, so dass

$$x(t_1) = \tilde{x}(t_1) \quad \text{und} \quad x(t) \neq \tilde{x}(t) \quad \text{für alle } t \in (t_1, t_2) \quad (1.4)$$

gilt. Nun kann man leicht nachrechnen, dass beide Funktionen das Anfangswertproblem (1.1), (1.2) mit Anfangsbedingung $(t_1, x(t_1)) \in D$ lösen. Aus dem ersten Teil des Beweises folgt Eindeutigkeit der Lösungen dieses Problems auf einem offenen Intervall $\tilde{J} = (t_1 - \tilde{a}, t_1 + \tilde{a})$, also insbesondere

$$x(t) = \tilde{x}(t) \text{ für alle } t \in (t_1, t_1 + \tilde{a}).$$

Dies widerspricht (1.4), weswegen x und \tilde{x} für alle $t \in I \cap \tilde{I}$ übereinstimmen müssen. Damit müssen auch I und \tilde{I} übereinstimmen, da ansonsten x außerhalb von I durch \tilde{x} bzw. \tilde{x} außerhalb von \tilde{I} durch x fortgesetzt werden könnte, was der Maximalität von I und \tilde{I} widerspricht. \square

Das Verhalten am Rand von I_{t_0, x_0} kann für die Lösungen wie folgt beschrieben werden: Falls $t^* = \sup I_{t_0, x_0} < \infty$ ist, so divergiert die Lösung $x(t; t_0, x_0)$ für $t \nearrow t^*$ oder sie konvergiert gegen ein $x_1 \in \mathbb{R}^n$ mit $(t^*, x_1) \notin D$; falls $t_* = \inf I_{t_0, x_0} > -\infty$ ist, so divergiert die Lösung $x(t; t_0, x_0)$ für $t \searrow t_*$ oder sie konvergiert gegen ein $x_1 \in \mathbb{R}^n$ mit $(t_*, x_1) \notin D$. Die Begründung hierfür ist die folgende: Wenn $x(t; t_0, x_0)$ z.B. für $t \nearrow t^*$ gegen $x_1 \in \mathbb{R}^n$ mit $(t^*, x_1) \in D$ konvergieren würde, so könnte man die auf einem offenen Intervall $(t^* - a, t^* + a)$ definierte Lösung $x(t; b, x_1)$ betrachten, die auf $(t^* - a, t^*]$ mit $x(t; t_0, x_0)$ übereinstimmen müsste und damit eine Fortsetzung von $x(t; t_0, x_0)$ über t^* hinaus wäre. Dies widerspricht der Maximalität von I_{t_0, x_0} . Im Fall $D = \mathbb{R} \times \mathbb{R}^n$ gilt insbesondere $\|x(t; t_0, x_0)\| \rightarrow \infty$ für $t \nearrow t^*$ oder $t \searrow t_*$, falls t^* oder t_* endlich sind.

Wir werden im Folgenden zumeist annehmen, dass die Annahmen von Satz 1.4 erfüllt sind, auch ohne dies explizit zu erwähnen. Lediglich auf Ausnahmen von dieser Regel werden wir ausdrücklich hinweisen.

Eine einfache Konsequenz aus Satz 1.4 ist die sogenannte *Kozykluseigenschaft* der Lösungen, die für $(t_0, x_0) \in D$ und zwei Zeiten $t_1, t \in \mathbb{R}$ gegeben ist durch

$$x(t; t_0, x_0) = x(t; t_1, x(t_1; t_0, x_0)), \quad (1.5)$$

vorausgesetzt natürlich, dass alle hier verwendeten Lösungen zu den angegebenen Zeiten auch existieren. Zum Beweis rechnet man nach, dass linke Ausdruck in (1.5) das Anfangswertproblem (1.1), (1.2) zur Anfangsbedingung $(t_1, x(t_1; t_0, x_0))$ löst. Da der rechte dies ebenfalls tut, müssen beide übereinstimmen.

Eine weitere Konsequenz aus Satz 1.4 ist, dass sich Lösungen nicht schneiden können. Falls für zwei Lösungen $x_1(t)$ und $x_2(t)$ (zu beliebigen Anfangswerten und Anfangszeiten) eine Zeit t_1 mit $x_1(t_1) = x_2(t_1)$ existiert, so müssen diese Lösungen bereits übereinstimmen, da sie beide Lösungen zur Anfangsbedingung $(t_1, x_1(t_1))$ sind. (Für autonome DGL gilt sogar noch mehr, vgl. das Ende dieses Abschnittes).

Einige gewöhnliche Differentialgleichungen können explizit analytisch gelöst werden, für diesen Zweck gibt es eine Reihe von Methoden, mit denen die Lösungen berechnet werden können. Im Allgemeinen stehen explizite Lösungen nicht zur Verfügung, so dass allein die numerische Lösung Aufschluss über die Lösungen erlaubt. Auch in diesem Fall verlässt man sich aber nicht allein auf die Numerik, sondern versucht, gewisse qualitative Eigenschaften der Lösungen mit analytischen Methoden zu beweisen. Wir werden im Rahmen dieser Vorlesung Beispiele für diese Vorgehensweise kennen lernen.

1.2.4 Grafische Darstellung der Lösungen

Zur grafischen Darstellung von Lösungen verwenden wir zwei verschiedene Methoden, die wir hier an der zweidimensionalen DGL

$$\dot{x}(t) = \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} x(t)$$

mit $x(t) = (x_1(t), x_2(t))^T$ und Anfangsbedingung $x(0) = (1, 1)^T$ illustrieren wollen. Da jede Lösung einer Differentialgleichung eine Funktion von \mathbb{R} nach \mathbb{R}^n darstellt, kann man die Graphen der einzelnen Komponenten $x_i(t)$ der Lösung in Abhängigkeit von t darstellen. Für die obige DGL ist dies in Abbildung 1.1 dargestellt. Die durchgezogene Linie zeigt $x_1(t)$ während die gestrichelte Linie $x_2(t)$ darstellt.

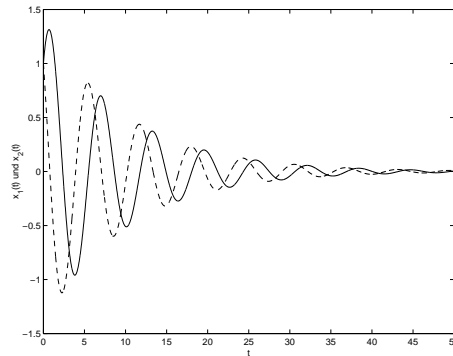


Abbildung 1.1: Darstellung von $x(t)$ mittels Graphen ($x_1(t)$ durchgezogen, $x_2(t)$ gestrichelt)

Eine alternative Darstellung, die speziell für zwei- und dreidimensionale Differentialgleichungen geeignet ist, ergibt sich, wenn man statt der Funktionsgraphen der Komponenten x_i die Kurve $\{x(t) \mid t \in [0, T]\} \subset \mathbb{R}^n$ darstellt. Hier geht in der Grafik die Information über die Zeit (sowohl über die Anfangszeit t_0 als auch über die laufende Zeit t) verloren. Letzteres kann zumindest teilweise durch das Anbringen von Pfeilen, die die Zeitrichtung symbolisieren, ausgeglichen werden. Ein Beispiel für diese Darstellung zeigt Abbildung 1.2.

Für autonome Differentialgleichungen ist der Verlust der Anfangszeit in der Grafik nicht weiter schlimm, da die Lösungen nicht wirklich von der Anfangszeit abhängen: man rechnet leicht nach, dass hier für zwei verschiedene Anfangszeiten $t_0 \neq t_1$ die Beziehung

$$x(t; t_0, x_0) = x(t - t_0 + t_1; t_1, x_0) \quad (1.6)$$

gilt. Die Lösung verschiebt sich also auf der t -Achse, verändert sich aber ansonsten nicht. Insbesondere ist die in Abbildung 1.2 dargestellte Kurve für autonome DGL für alle Anfangszeiten gleich. Daraus folgt für autonome DGL, dass sich auch in der Kurvendarstellung zwei Lösungen nicht schneiden können, es sei denn, sie stimmen bereits vollständig überein.

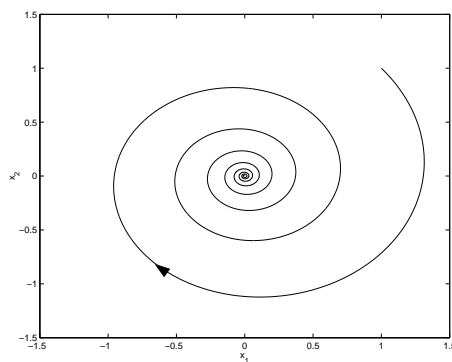


Abbildung 1.2: Darstellung von $x(t)$ als Kurve

1.3 Literaturhinweise

Eine schöne Einführung in die Theorie der gewöhnlichen Differentialgleichungen liefert z.B. das Lehrbuch

B. Aulbach, Gewöhnliche Differentialgleichungen, Spektrum-Verlag, Heidelberg, 1997

Kapitel 2

Biologie

Mathematische Modelle werden in vielen verschiedenen Bereichen der Biologie verwendet. Klassische Anwendungen sind z.B. die Untersuchung von Wachstumsprozessen und biochemischen Reaktionen oder die Ausbreitung von Epidemien, neuere Anwendungen finden sich z.B. in vielen Teilgebieten der Gentechnik oder in der Immunologie. Wir werden hier bei den klassischen Bereichen bleiben und uns (ausführlich) mit der Populationsdynamik und (kürzer) mit Epidemien beschäftigen.

2.1 Populationsdynamik für eine Art

Populationsdynamik bezeichnet die Analyse des Wachstums einer oder mehrerer Arten oder Spezies in einem (meist sehr einfach modellierten) Ökosystem. In diesem Abschnitt wollen wir mit Modellen für eine Art beginnen.

2.1.1 Differenzen- und Differentialgleichungen

In der mathematischen Modellierung von Wachstumsprozessen stellt sich zunächst die Frage, ob gewöhnliche Differentialgleichungen überhaupt das richtige mathematische Modellierungswerkzeug sind. Tatsächlich “lebt” eine Differentialgleichung immer auf kontinuierlichen Räumen, während die in der Populationsdynamik auftretenden Größen zunächst einmal diskreter Natur sind: Die Größe einer Population wird üblicherweise in der Anzahl der Individuen gemessen, die selbstverständlich eine natürliche Zahl ist. Dieses Problem wird in praktisch allen Modellen dadurch gelöst, dass man die Größe der Population nicht anhand der diskreten Anzahl der Individuen sondern anhand ihrer Biomasse x misst, und genau so wollen wir es hier halten. Die Biomasse x ist eine (nichtnegative) reelle Zahl, deren zeitliche Entwicklung man durch eine Differentialgleichung modellieren kann.

Das nächste Problem ist die richtige Wahl der Zeitachse. Biologische Messungen (z.B. zum Bestand einer Population) werden niemals kontinuierlich für $t \in [t_0, t_1]$ durchgeführt, sondern zu diskreten Zeiten $t_1 < t_2 < t_3 < \dots$. Der Zuwachs oder die Abnahme einer Population wird dementsprechend oft bezüglich diskreter Zeitpunkte ausgedrückt. Ein allgemeines diskretes Modell einer Populationsdynamik für x in einem festgelegten Gebiet A

ist gegeben durch

$$x(t_{i+1}) = x(t_i) + \Delta G(t_i) - \Delta S(t_i) + \Delta M(t_i). \quad (2.1)$$

Hierbei bezeichnet

$$\Delta G(t_i): \text{Anzahl der Geburten im Intervall } [t_i, t_{i+1}] \quad (\geq 0)$$

$$\Delta S(t_i): \text{Anzahl der Sterbefälle im Intervall } [t_i, t_{i+1}] \quad (\geq 0)$$

$$\Delta M(t_i): \text{Migration (Zu- und Abwanderung) im Intervall } [t_i, t_{i+1}] \quad (\geq 0 \text{ oder } \leq 0)$$

Gleichungen von Typ (2.1) nennt man *Differenzgleichungen* und tatsächlich kann man mit solchen zeitdiskreten Modellen arbeiten und es gibt viele (auch aktuelle) Forschungsarbeiten, die sich mit der Theorie von Differenzgleichungen beschäftigen.

Wir werden hier nicht mit Differenzgleichungen arbeiten, sondern statt dessen ein Differentialgleichungsmodell herleiten. Der Grund dafür, Differentialgleichungen vorzuziehen, liegt im Wesentlichen darin, dass es für Differentialgleichungen viele mathematische Analysemethoden gibt, die für Differenzgleichungen entweder komplizierter sind oder gar nicht zur Verfügung stehen, zum Teil aus prinzipiellen mathematischen Gründen (da Lösungen von Differenzgleichungen i.A. ein sehr viel komplexeres Verhalten aufweisen als Lösungen von Differentialgleichungen) oder einfach, weil noch niemand versucht hat sie herzuleiten und zu beweisen. Als Modellierungswerkzeug sind Differenzgleichungen den Differentialgleichungen sicherlich ebenbürtig.

Wie kommt man nun von (2.1) zu einer Differentialgleichungsformulierung? Nehmen wir an, dass die Zeitpunkte t_i äquidistant verteilt sind, dass also $t_{i+1} - t_i =: \Delta t$ für ein von i unabhängiges Δt gilt. Dann kann man (2.1) für $t = t_i$ umformulieren als

$$\frac{x(t + \Delta t) - x(t)}{\Delta t} = \frac{\Delta G(t)}{\Delta t} - \frac{\Delta S(t)}{\Delta t} + \frac{\Delta M(t)}{\Delta t}.$$

Beachte, dass ΔG , ΔS und ΔM von Δt abhängen, auch wenn dies in der Notation nicht explizit klar wird. Für $\Delta t \rightarrow 0$ erhält man so

$$\frac{d}{dt}x(t) = g(t) - s(t) + m(t).$$

Die Funktionen g , s und m heißen *Geburten-*, *Sterbe-* und *Migrationsrate*. Man könnte versuchen, sie mittels

$$g(t) = \lim_{\Delta t \rightarrow 0} \frac{\Delta G(t)}{\Delta t}, \quad s(t) = \lim_{\Delta t \rightarrow 0} \frac{\Delta S(t)}{\Delta t} \quad \text{und} \quad M(t) = \lim_{\Delta t \rightarrow 0} \frac{\Delta M(t)}{\Delta t}$$

aus ΔG , ΔS und ΔM zu bestimmen, was sinnvoll wäre, wenn wir ΔG , ΔS und ΔM definiert hätten. Diesen Umweg wollen wir nicht gehen, statt dessen werden wir g und s direkt aus geeigneten Modellannahmen ableiten. Migration werden wir in unseren Modellen nicht betrachten, weswegen m immer gleich 0 sein wird.

2.1.2 Einfache Modelle

Das einfachste Modell der Populationsdynamik macht die folgenden Annahmen:

- (i) Die Geburtenrate $g(t)$ ist linear proportional zum aktuellen Bestand der Population:

$$g(t) = \gamma x(t) \text{ für ein } \gamma \in \mathbb{R}$$

- (ii) Die Sterberate $s(t)$ ist linear proportional zum aktuellen Bestand der Population:

$$s(t) = \sigma x(t) \text{ für ein } \sigma \in \mathbb{R}$$

- (iii) Migration findet nicht statt: $m(t) \equiv 0$

Dies führt auf die Differentialgleichung

$$\dot{x}(t) = \lambda x(t) \tag{2.2}$$

wobei $\lambda x(t)$ mit $\lambda = \gamma - \sigma \in \mathbb{R}$ die *Wachstumsrate* genannt wird.¹ Man rechnet leicht nach, dass die Lösungen von (2.2) mit Anfangsbedingung $x(t_0) = x_0$ durch

$$x(t; x_0) = x_0 e^{\lambda(t-t_0)}$$

gegeben sind. Beachte, dass $x(t)$ hier — wie in allen Wachstumsmodellen — die Größe einer Population beschreibt, so dass in diesen Modellen nur $x \geq 0$ und damit insbesondere $x_0 \geq 0$ sinnvoll ist. Wir schreiben hier $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$ und $\mathbb{R}_0^+ = \mathbb{R}^+ \cup \{0\}$.

Auch wenn dies ein sehr einfaches Modell ist, so beschreibt es doch manche realen Wachstumsphänomene relativ gut. Abbildung 2.1 zeigt z.B. die Größe der Weltbevölkerung zwischen 1950 und 2000 (in Milliarden Menschen) mit einer Lösung von (2.2). Die Werte x_0 und λ wurden hier über ein nichtlineares Ausgleichsproblem geschätzt, vgl. Abschnitt 6.6 des Skripts zur Vorlesung “Numerische Mathematik 1”²: Mit $t_0 = 1950$ erhalten wir hier $\lambda = 0.0173456$ und $x_0 = 2.605331$. Es sollte erwähnt werden, dass die Ermittlung geeigneter Parameter für ein Modell (man spricht von *Parameterschätzung*) ein eigenständiges anspruchsvolles mathematisches (oft auch numerisches) Problem ist, das wir in dieser Vorlesung nicht betrachten können. Beachte auch, dass die Parameter von den verwendeten Einheiten abhängen. Hier haben wir x in Milliarden Menschen und t in Jahren angegeben.

Das derzeitige Weltbevölkerungswachstum wird also durch (2.2) offenbar recht gut beschrieben, andere reale Wachstumsprozesse hingegen werden durch dieses Modell überhaupt nicht gut beschrieben, beispielsweise das Wachstum der Bevölkerung in Europa, welches in den letzten Jahrzehnten praktisch zum Stillstand gekommen ist. Der Grund dafür ist aus der Struktur der Lösungen sofort ersichtlich: Aus $\lambda > 0$ folgt $e^{\lambda t} \rightarrow \infty$ für $t \rightarrow \infty$, für $x_0 > 0$ wächst die modellierte Population also über alle Grenzen; die Wahl $\lambda < 0$ (d.h., die Sterberate ist größer als die Geburtenrate) schafft hier keine brauchbare Abhilfe: in diesem Fall folgt $e^{\lambda t} \rightarrow 0$ für $t \rightarrow \infty$, was das reale Verhalten sicherlich auch nicht korrekt widerspiegelt — zumindest derzeit nicht.

¹Oft wird auch der reelle Parameter λ als Wachstumsrate bezeichnet.

²<http://www.uni-bayreuth.de/departments/math/~lgruene/numerik0203/>

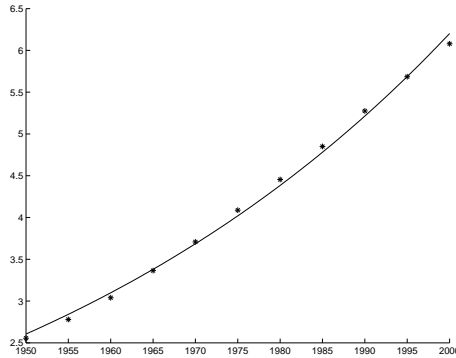


Abbildung 2.1: Wachstum der Weltbevölkerung und Lösung von (2.2)

Um sich verlangsames Wachstum modellieren zu können, werden wir (2.2) um eine “Wachstumsgrenze” erweitern, die wir hier durch eine obere Schranke $K > 0$ für die Größe der Population modellieren; K steht für die *Kapazität* des Lebensraums. Diese ergibt sich aus den zur Verfügung stehenden Ressourcen, wie z.B. Nahrung, Trinkwasser etc. Wir fügen dazu einen Faktor $w(x)$ mit den folgenden Eigenschaften in die Gleichung (2.2) ein.

- (i) Falls $x < K$ ist, soll $w(x) > 0$ sein, da noch “Platz” für Wachstum vorhanden ist.
- (ii) Falls $x > K$ ist, soll $w(x) < 0$ sein, um “negatives Wachstum” zu erzwingen.

Die einfachste Funktion, die dieses leistet, ist die lineare Funktion $w(x) = K - x$. Wir erhalten damit die Gleichung

$$\dot{x}(t) = \lambda(K - x(t))x(t), \quad (2.3)$$

die als *logistisches Wachstum* bezeichnet wird. Der Ausdruck $\lambda(K - x(t))x(t)$ ist hier die — nun nichtlineare — Wachstumsrate. Auch für diese DGL ist die explizite Lösung bekannt, sie ist gegeben durch

$$x(t; t_0, x_0) = \frac{K}{1 + \left(\frac{K}{x_0} - 1\right) e^{-\lambda K(t-t_0)}}.$$

Man kann das Verhalten der Lösung nun an diesem expliziten Ausdruck untersuchen. Wir wollen hier aber — zur Einübung — einen anderen Weg gehen und die dadurch erhaltenen Resultate an der expliziten Lösung überprüfen.

Hierzu definieren und betrachten wir zunächst einige wichtige Begriffe für Differentialgleichungen, und zwar allgemein im \mathbb{R}^n .

Definition 2.1 Ein Punkt $x^* \in \mathbb{R}^n$ heißt *Gleichgewicht* (auch *Ruhelage*, *Fixpunkt* oder *Equilibrium*) für eine DGL (1.1), falls $x(t; t_0, x^*) = x^*$ ist für alle $t, t_0 \in \mathbb{R}$. \square

Man sieht leicht, dass ein Punkt x^* genau dann ein Gleichgewicht ist, wenn $f(t, x^*) = 0$ ist für alle $t \in \mathbb{R}$. Gleichgewichte sind vor allem deswegen interessant, weil sie Aufschluss über das Langzeitverhalten der Lösungen geben können.

Satz 2.2 Betrachte eine DGL (1.1) mit autonomem f . Sei $x(t; t_0, x_0)$ eine Lösung, die für $t \rightarrow \infty$ oder $t \rightarrow -\infty$ gegen einen Punkt $x^* \in \mathbb{R}^n$ konvergiert. Dann ist x^* ein Gleichgewicht.

Beweis: Wir beweisen den Fall $t \rightarrow \infty$, der Fall $t \rightarrow -\infty$ folgt analog. Betrachte dazu die Lösung $x(t) = x(t; t_0, x_0)$. Da diese Lösung gegen x^* konvergiert, folgt aus der Stetigkeit von f die Konvergenz $f(x(t)) \rightarrow f(x^*)$. Sei nun für ein gegebenes $\varepsilon > 0$ die Zeit $t^* > 0$ so groß gewählt, dass die Ungleichungen

$$\|x(t) - x^*\| \leq \varepsilon \quad \text{und} \quad \|f(x(t)) - f(x^*)\| \leq \varepsilon$$

für alle $t \geq t^*$ gelten. Dann folgt für alle $t \geq t^*$ aus (1.3) die Ungleichung

$$\|x(t) - x(t^*)\| = \left\| \int_{t^*}^t f(x(\tau)) d\tau \right\| \geq \left\| \int_{t^*}^t f(x^*) d\tau \right\| - \left\| \int_{t^*}^t f(x(\tau)) - f(x^*) d\tau \right\|$$

und daraus

$$\begin{aligned} (t - t^*)\|f(x^*)\| &= \left\| \int_{t^*}^t f(x^*) d\tau \right\| \\ &\leq \|x(t) - x(t^*)\| + \left\| \int_{t^*}^t f(x(\tau)) - f(x^*) d\tau \right\| \\ &\leq \underbrace{\|x(t) - x(t^*)\|}_{\leq \|x(t) - x^*\| + \|x^* - x(t^*)\|} + \int_{t^*}^t \|f(x(\tau)) - f(x^*)\| d\tau \leq 2\varepsilon + (t - t^*)\varepsilon. \end{aligned}$$

Diese Ungleichung gilt für alle $t > t^*$, insbesondere also für $t = t^* + 1$. Mit dieser Wahl folgt

$$\|f(x^*)\| \leq 3\varepsilon,$$

also, da $\varepsilon > 0$ beliebig war, $\|f(x^*)\| = 0$ und damit $f(x^*) = 0$. Folglich ist x^* ein Gleichgewicht der DGL. \square

Satz 2.2 hat eine wichtige Konsequenz für die Analyse von Differentialgleichungen. Er besagt nämlich, dass wir mit den Gleichgewichten im autonomen Fall bereits alle möglichen Grenzwerte von Lösungstrajektorien kennen. Nicht jedes Gleichgewicht ist allerdings ein möglicher Grenzwert. Die folgende Definition beschreibt mögliche Situationen in der Umgebung eines Gleichgewichtes etwas genauer.

Definition 2.3 (i) Ein Gleichgewicht x^* einer DGL (1.1) heißt *(lokal) exponentiell stabil*, falls eine Umgebung N von x^* sowie Parameter $\sigma, \lambda > 0$ existieren, so dass für alle $x_0 \in N$, alle $t_0 \in \mathbb{R}$ und alle $t \geq t_0$ die Ungleichung

$$\|x(t; t_0, x_0) - x^*\| \leq \sigma e^{-\lambda t} \|x_0 - x^*\|$$

gilt.

(ii) Ein Gleichgewicht x^* einer DGL (1.1) heißt *exponentiell instabil*, falls Parameter $\sigma, \lambda > 0$ und eine Umgebung N von x^* existieren, so dass in jeder Umgebung N_0 von x^* ein Punkt $x_0 \in N_0$ existiert, für den für alle $t_0 \in \mathbb{R}$ die Ungleichung

$$\|x(t; t_0, x_0) - x^*\| \geq \sigma e^{-\lambda t} \|x_0 - x^*\|$$

gilt für alle $t \geq t_0$ für die $x(t; t_0, x_0) \in N$ gilt.

(iii) Ein Gleichgewicht x^* einer DGL (1.1) heißt *exponentiell antistabil*, falls Parameter $\sigma, \lambda > 0$ und eine Umgebung N von x^* existieren, so dass für alle $x_0 \in N$ mit $x_0 \neq x^*$ und alle $t_0 \in \mathbb{R}$ die Ungleichung

$$\|x(t; t_0, x_0) - x^*\| \geq \sigma e^{-\lambda t} \|x_0 - x^*\|$$

gilt für alle $t \geq t_0$ für die $x(t; t_0, x_0) \in N$ gilt. □

Für $t \rightarrow \infty$ konvergieren also im Fall (i) alle Lösungen aus einer Umgebung des Gleichgewichtes gegen das Gleichgewicht x^* . Im Fall (iii) laufen alle Lösungen für wachsendes t weg von x^* , Konvergenz gegen x^* ist nicht möglich. Im Fall (ii) gibt es beliebig nahe and x^* startende Lösungen die von x^* weg laufen, es ist aber nicht ausgeschlossen, dass ein Anfangswert $x_0 \neq x^*$ existiert, für den $x(t; t_0, x_0)$ gegen x^* konvergiert. Wir werden später Beispiele dafür kennen lernen.

Beachte, dass (i)–(iii) keineswegs alle möglichen Szenarien beschreiben. So könnte z.B. eine Funktion $\beta(\|x_0 - x^*\|, t)$ existieren, die langsamer als $\sigma e^{-\lambda t} \|x_0 - x^*\|$ gegen Null konvergiert und für die statt (i) die Ungleichung

$$\|x(t; t_0, x_0) - x^*\| \leq \beta(\|x_0 - x^*\|, t)$$

gilt.

Der Grund dafür, in diesen Definitionen die (doch recht speziellen) exponentiellen Abschätzungen zu verwenden, liegt darin, dass sich für diese Definitionen einfache nachprüfbarere Kriterien beweisen lassen — zumindest falls die DGL autonom ist.

Satz 2.4 Sei x^* ein Gleichgewicht einer DGL (1.1) mit autonomem Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Sei f in einer Umgebung von x^* stetig differenzierbar und sei $Df(x^*) \in \mathbb{R}^{n \times n}$ die Ableitung (also die Jacobi-Matrix) von f an der Stelle x^* . Dann gilt:

(i) Das Gleichgewicht x^* ist genau dann (lokal) exponentiell stabil, wenn alle Eigenwerte $\lambda_i \in \mathbb{C}$ von $Df(x^*)$ negativem Realteil haben.

(ii) Das Gleichgewicht x^* ist genau dann exponentiell instabil, wenn es einen Eigenwert $\lambda_i \in \mathbb{C}$ von $Df(x^*)$ gibt, der positiven Realteil besitzt.

(iii) Das Gleichgewicht x^* ist genau dann exponentiell antistabil, wenn alle Eigenwerte $\lambda_i \in \mathbb{C}$ von $Df(x^*)$ positivem Realteil haben.

Ein Beweis für (i) findet sich z.B. als Korollar 7.6 in meinem Skript zur Vorlesung “Stabilität und Stabilisierung linearer Systeme”³. Beweise für (ii) und (iii) finden sich in Büchern über

³<http://www.uni-bayreuth.de/departments/math/~lgruene/linstab0203/>

gewöhnliche Differentialgleichungen. Die Jacobi-Matrix $Df(x^*)$ wird oft *Linearisierung* von (1.1) in x^* genannt.

Wir wollen die theoretischen Resultate nun zur Analyse unseres Modells (2.3) verwenden. Hier gilt

$$f(x) = \lambda(K - x)x.$$

Die Gleichgewichte von (2.3) sind also gegeben durch

$$f(x) = 0 \quad \Leftrightarrow \quad x^* = 0 \text{ oder } x^{**} = K.$$

Für jeden Anfangswert $x_0 \neq x^*, x^{**}$ muss die Lösung nun entweder divergieren oder gegen eines der zwei Gleichgewichte konvergieren. Betrachten wir also nun zunächst das lokale Stabilitätsverhalten der Gleichgewichte, um mögliche Grenzwerte zu ermitteln. Da die DGL eindimensional ist, ist die Ableitung Df von f reellwertig. Nach Produktregel gilt

$$Df(x) = f'(x) = \lambda(K - x) - \lambda x \quad \Rightarrow \quad Df(x^*) = \lambda K \text{ und } Df(x^{**}) = -\lambda K.$$

Die Eigenwerte dieser "1 × 1-Matrizen" sind natürlich gerade $K > 0$ für $x^* = 0$ und $-\lambda K < 0$ für $x^{**} = K$. Das Gleichgewicht x^* ist also exponentiell antistabil, während x^{**} exponentiell stabil ist. Nur x^{**} kommt also als Limes von Lösungen für $t \rightarrow \infty$ in Frage.

Aus der lokalen exponentiellen Stabilität wissen wir, dass es eine Umgebung N von x^{**} gibt, aus der alle Lösungen gegen x^{**} konvergieren. Beachte, dass nicht nur alle Lösungen, die in N starten, konvergieren müssen, sondern wegen (1.5) auch alle Lösungen, die außerhalb von N starten und irgendwann in diese Umgebung N hineinlaufen. Der nächste Schritt in der Analyse von (2.3) besteht nun daraus, zu ermitteln, für welche Anfangswerte die Lösungen gegen x^{**} konvergieren. Dies ist die Frage nach dem *Einzugsbereich* des Gleichgewichtes x^{**} . Allgemein ist der Einzugsbereich eines lokal exponentiell stabilen Gleichgewichtes x^* für eine autonome DGL gegeben als

$$\mathcal{D}(x^*) := \{x_0 \in \mathbb{R}^n \mid \lim_{t \rightarrow \infty} x(t; x_0) = x^*\}$$

und für die Umgebung N aus Definition 2.3(i) gilt

$$\mathcal{D}(x^*) = \{x_0 \in \mathbb{R}^n \mid x(t; x_0) \in N \text{ für ein } t \geq 0\},$$

da alle Lösungen, die nach N laufen wegen (1.5) gegen x^* konvergieren müssen und umgekehrt alle Lösungen, die gegen x^* konvergieren, durch N laufen müssen.

Im \mathbb{R}^n ist die Ermittlung von \mathcal{D} eine schwierige, oft unlösbare Aufgabe. Im eindimensionalen Fall ist die Sache einfacher, da man mit der Monotonie der Lösungen argumentieren kann. Man sieht leicht, dass $f(x) > 0$ gilt für $x \in (x^*, x^{**})$. Die Lösungen in diesem Intervall wachsen also streng monoton. Da sie nicht gegen einen Wert $\tilde{x} \in (x^*, x^{**})$ konvergieren können (dieses \tilde{x} wäre nach Satz 2.2 ein weiteres Gleichgewicht, ein solches existiert aber nicht) müssen sie wachsen, bis sie die exponentielle Stabilitätsumgebung N von x^{**} erreichen und dann gegen x^{**} konvergieren. Es gilt also $(x^*, x^{**}) = (0, K) \subset \mathcal{D}(x^{**})$. Für $x \in (x^{**}, \infty)$ ist f negativ, die Lösungen fallen in diesem Bereich also monoton. Wiederum können sie gegen kein $\tilde{x} \in (x^{**}, \infty)$ konvergieren, sie müssen also irgendwann in N landen und dann gegen x^{**} konvergieren. Wir erhalten also $\mathcal{D}(x^{**}) = (x^*, \infty) = (0, \infty)$, die Gleichheit folgt

hierbei daraus, dass die Lösungen mit $x_0 = x^*$ sicherlich nicht gegen x^{**} konvergieren, für $x < x^* = 0$ ist $f(x) < 0$, weswegen die Lösungen fallen und niemals in die Umgebung N gelangen können.

Unsere Analyse ergibt also zusammenfassend: (1) Es gibt zwei Gleichgewichte, $x^* = 0$ und $x^{**} = K$. (2) Alle Lösungen mit Anfangswert $x_0 \in (0, \infty)$ konvergieren gegen x^{**} . In Abbildung 2.2 sind die oben angegebenen expliziten Lösungen mit $K = \lambda = 1$ für die Anfangswerte $x_0 = 0, 1/100, 1$ und 2 dargestellt. Es ergibt sich genau das beschriebene Verhalten.

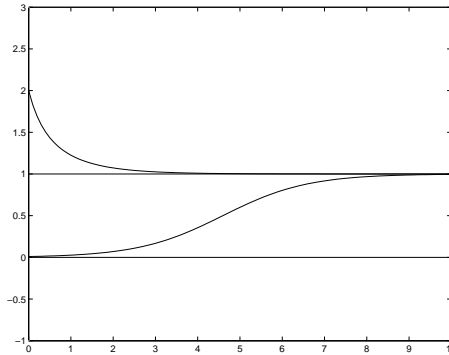


Abbildung 2.2: Lösungen der logistischen Wachstumsgleichung (2.3)

Bemerkung 2.5 Das logistische Wachstum (2.3) ist nicht das einzige Modell für beschränktes Wachstum. Zur Modellierung von Zellwachstum z.B. wird oft die DGL

$$\dot{x}(t) = \lambda x(t) \log \left(\frac{K}{x(t)} \right) \quad (2.4)$$

verwendet, die als *Gompertz-Wachstum* bezeichnet wird und mit deren Lösungen sich klinische Ergebnisse gut nachvollziehen lassen. Hier sind die expliziten Lösungen unbekannt; mit ähnlichen Methoden wie oben kann man aber nachweisen, dass das qualitative Lösungsverhalten dem von (2.3) entspricht. \square

2.1.3 Eine Anwendung des Modells

Mathematische Modelle werden oft zur Beschreibung und Erklärung realer Situationen eingesetzt. Sie dienen aber auch als Teilsysteme in der mathematischen Untersuchung komplexerer Phänomene. Wir wollen dies Prinzip an einem Beispiel illustrieren.

Wir wollen ein Modell für den Fischfang aufstellen, bei dem $x(t)$ eine Fischpopulation beschreibt. Analog zu den oben betrachteten Geburts- und Sterberaten führen wir eine Fangrate $u(t)$ zur Zeit t ein. Als Modellannahme verwenden wir, dass sich das Fischwachstum durch das logistische Wachstum (2.3) beschreiben lässt, falls keine Fische gefangen werden.

Die sich daraus ergebende DGL

$$\dot{x}(t) = \lambda(K - x(t))x(t) - u(t). \quad (2.5)$$

wird *Schäfers Modell* genannt.

Beachte, dass dies nun eine nichtautonome DGL ist. Zudem ist es — je nach Wahl von $u(t)$ — möglich, dass die Lösungen mit positiven Anfangswert x_0 negativ werden, was aber vom Modell her nicht sinnvoll ist, weswegen wir $x(t) > 0$ durch die Wahl von $D = \mathbb{R} \times \mathbb{R}^+$ als Definitionsbereich sicher stellen. Die Wahl von D ist also nicht mathematisch sondern aus Modellgesichtspunkten motiviert.

Eine Fangstrategie auf einem Intervall $[t_0, t_1]$ ist nun einfach eine stetige Funktion $u : [t_0, t_1] \rightarrow \mathbb{R}_0^+$. Für einen Anfangswert $x_0 > 0$ nennen wir u *zulässig*, falls die zugehörige Lösung $x(t; t_0, x_0, u)$ auf dem ganzen Intervall $[t_0, t_1]$ existiert. Wir schreiben die Funktion u hier als zusätzlichen Parameter in die Lösung, um die Abhängigkeit der Lösungen von u zu betonen.

Die Anzahl M der gefangenen Fische (natürlich wieder als Biomasse ausgedrückt) ergibt sich nun als Integral über $u(t)$, also

$$M = \int_{t_0}^{t_1} u(t) dt.$$

Ziel des Fischers könnte es nun sein, diese Größe M zu maximieren. Dies würde jedoch unausweichlich zur Ausrottung der Fische führen: Wären zum Zeitpunkt t_1 noch Fische da, so könnte man $u(t)$ erhöhen und würde trotzdem noch eine zulässige Fangstrategie erhalten. Dies wäre zwar auf dem betrachteten Intervall optimal, nach der Zeit t_1 wäre der Fischer aber arbeitslos, weswegen dies auf lange Sicht keine gute Strategie ist. Selbst wenn der Fischer sich nach der Zeit t_1 zur Ruhe setzen will, wäre dies keine gute Lösung, in jedem Fall aus ökologischer Sicht aber auch aus ökonomischer Sicht, da dies zur Vernichtung der Bestände führen würde.

Man muss also das Überleben der Fische in die Optimierung einbeziehen. Dies führt auf ein Optimierungsproblem unter Nebenbedingungen:

$$\text{maximiere } \int_{t_0}^{t_1} u(t) dt$$

unter den Nebenbedingungen

- (i) u ist zulässig für den Anfangswert x_0
- (ii) $x(t_1; t_0, x_0, u) \geq x_1$ für einen vorgegebenen Wert $x_1 > 0$

Dies ist ein sogenanntes *optimales Steuerungsproblem*, für dessen Lösung es eine Vielzahl von analytischen und numerischen Techniken gibt. (Weiterführende Vorlesungen in diesem Gebiet werden an der Uni Bayreuth regelmäßig angeboten.)

Hier können wir dieses Problem nicht lösen, statt dessen betrachten wir einen alternativen Ansatz, den wir mit unseren Methoden behandeln können. Wir wählen die Fangrate $u(t)$ proportional zur Menge der vorhandenen Fische: $u(t) = cx(t)$ für ein $c > 0$. Dies vereinfacht nicht nur die Analyse, sondern liefert auch ein Modell für die Tatsache, dass man bei gleichbleibender Befischung (z.B. durch Auslegen von Netzen) in der Regel immer eine zu $x(t)$ proportionale Menge von Fischen fangen wird. Der Parameter c ergibt sich dabei z.B.

aus der Größe der Netze und der Dauer des Auslegens. Mit dieser Wahl von u ergibt sich (2.5) zu

$$\dot{x}(t) = \lambda \left(K - \frac{c}{\lambda} - x(t) \right) x(t).$$

Dadurch verschiebt sich das Gleichgewicht x^{**} aus der obigen Analyse, genauer kann man leicht die Gleichgewichte

$$x^* = 0 \quad \text{und} \quad x^{**} = K - \frac{c}{\lambda}$$

berechnen. Für die Ableitung gilt

$$Df(x^*) = \lambda K - c = \lambda x^{**} \quad \text{und} \quad Df(x^{**}) = c - \lambda K = -\lambda x^{**}.$$

Jetzt muss man drei Fälle unterscheiden.

1. Fall: $x^{**} = K - \frac{c}{\lambda} > 0$. In diesem Fall bleibt alles wie oben, x^{**} ist lokal exponentiell stabil und jede Lösung mit Anfangswert $x_0 > 0$ konvergiert gegen x^{**} .

2. Fall: $x^{**} = K - \frac{c}{\lambda} < 0$. In diesem Fall wird $x^* = 0$ lokal exponentiell stabil und es gilt $f(x) < 0$ für alle $x > 0$. Alle Lösungen fallen also monoton und konvergieren schließlich gegen $x^* = 0$, für $t \rightarrow \infty$ werden die Fische also ausgerottet.

3. Fall: $x^{**} = K - \frac{c}{\lambda} = 0$. In diesem Fall vereinfacht sich die DGL zu $\dot{x}(t) = -\lambda(x(t))^2$, also ist jede Lösung monoton fallend. Zudem gilt $x^* = x^{**} = 0$. Alle Lösungen mit $x_0 > 0$ konvergieren gegen $x^* = 0$: sie können gegen keinen größeren Wert konvergieren, da kein größeres Gleichgewicht existiert; andererseits können sie die konstante Lösung $x(t; x^*) \equiv 0$ aber auch nicht schneiden. Beachte, dass das Gleichgewicht $x^* = 0$ weder lokal exponentiell stabil noch exponentiell instabil ist. Wie in Fall 2 werden die Fische für $t \rightarrow \infty$ ausgerottet.

Aus dieser Analyse kann man nun versuchen zu berechnen, wie $c > 0$ gewählt werden muss, damit der Ertrag maximiert wird. Auf beliebigen endlichen Intervallen ist das nicht so einfach, da aber alle Lösungen gegen eines der Gleichgewichte konvergieren, können wir zumindest approximativ den Ertrag für die Zeiten bestimmen, in denen die Lösung bereits nahe am Gleichgewicht liegt. Wir betrachten den Ertrag in einem Zeitintervall $[t_1, t_1 + 1]$ der Länge 1, wobei wir annehmen, dass t_1 so groß ist, dass wir uns bereits in der Nähe des Gleichgewichtes befinden. Im Fall 1 erhalten wir so

$$M = \int_{t_1}^{t_1+1} cx(t)dt \approx \int_{t_1}^{t_1+1} cx^{**}dt = cx^{**} = cK - \frac{c^2}{\lambda} > 0$$

und im Fall 2 und 3 ergibt sich analog

$$M \approx cx^* = 0.$$

Offensichtlich ist Fall 1 vorzuziehen, da nur dort (auf lange Sicht) ein positiver Ertrag erzielt wird. Zur Maximierung des Fangergebnisses muss man nun den Ausdruck $M(c) = cK - \frac{c^2}{\lambda}$ in c maximieren. Ableiten liefert die notwendige Bedingung

$$M'(c^*) = K - 2\frac{c^*}{\lambda} = 0 \quad \Leftrightarrow \quad c^* = \lambda K/2,$$

und da die zweite Ableitung $M''(c) = -2/\lambda < 0$ ist, ist dies tatsächlich ein lokales Maximum, sogar ein globales, da es das einzige ist. Der maximale Ertrag ergibt sich also zu

$$M(c^*) = cK - \frac{c^2}{\lambda} = \frac{\lambda K^2}{2} - \frac{\lambda^2 K^2}{4\lambda} = \frac{K^2 \lambda}{4}.$$

Welchen Wert haben solche Folgerungen aus einem Modell? Zunächst einmal muss man sich die möglichen Unzulänglichkeiten des Basismodells vergegenwärtigen; für das hier zu Grunde liegende Modell (2.3) machen wir dies im folgenden Abschnitt. Wenn man nun annimmt (oder experimentell belegen kann), dass das Modell Aussagekraft besitzt, so erlauben solche Rechnungen Einsicht in die Struktur des modellierten Phänomens. Hier zum Beispiel beobachtet man, dass man auf lange Sicht den maximalen Ertrag nicht erzielt, indem man die Fangrate beliebig erhöht, denn oberhalb des Wertes c^* wird der langfristig erzielbare Ertrag wieder sinken. In unserem Fall erlaubt dies durchaus gerechtfertigte *qualitative* Folgerungen für das modellierte Fischfangproblem. Eine zuverlässige *quantitative* Berechnung der realen optimalen Fangrate dürfte auf Basis eines so einfachen Modells allerdings nahezu unmöglich sein.

2.1.4 Abschließende Diskussion

Wir wollen das Modell (2.3) noch einmal abschließend diskutieren:

- Das Modell eignet sich gut zur Beschreibung von Wachstum unter idealen Bedingungen; die Ergebnisse von Laborversuchen lassen sich damit gut reproduzieren

In der realen Anwendung gibt es allerdings eine Reihe von weiteren Einflüssen, die hier nicht berücksichtigt werden:

- Naturbedingungen sind in der Regel variabel, z.B. durch Jahreszeiten bedingt. Im Modell ist alles konstant (realistischere Modelle verwenden hier zeitabhängige bzw. stochastische Parameter, wie wir sie im Kapitel über Finanzmathematik kennen lernen werden).
- Die räumliche Verteilung sowohl der Population als auch der Ressourcen wird nicht modelliert (dies könnte z.B. eine partielle Differentialgleichung leisten, mit der vom Ort abhängige Populationsdichten modelliert werden können).
- Die Geburts- und Sterberate hängen unmittelbar von der Größe der Population ab. Faktoren wie z.B. die Altersverteilung werden nicht berücksichtigt (hier können Delay-Differentialgleichungen Abhilfe schaffen, die wir später betrachten werden).
- Der Einfluss anderer Arten ist nicht im Modell enthalten.

Im nächsten Abschnitt werden wir uns mit Modellen beschäftigen, in denen der letzte Punkt berücksichtigt wird.

2.2 Populationsdynamik für mehrere Arten

In diesem Abschnitt werden wir die Modelle (2.2) und (2.3) auf den Fall mehrerer Arten verallgemeinern. Wir werden dabei zunächst auf den Fall von zwei Arten eingehen, wobei die erste Art (Beute) die Nahrung der zweiten Art (Räuber) darstellt.

2.2.1 Das Räuber–Beute Modell mit unbeschränkten Ressourcen

Dieser Abschnitt behandelt die Erweiterung des sehr einfachen Modells 2.2 auf den Fall von zwei Arten, und zwar Beutetiere (z.B. Ziegen) und Räubertiere (z.B. Wölfe).

Es bezeichne also x_1 die Größe der Beutepopulation und x_2 die Größe der Räuberpopulation. Wir machen die folgenden Modellannahmen:

- (i) Die Beutepopulation x_1 verhält sich gemäß (2.2) mit $\lambda = \gamma - \sigma$, wobei γ konstant ist und $\sigma = \tilde{\sigma} + bx_2$. Für die Beutetiere gibt es also unbegrenzte Ressourcen und die Sterbequote σ besteht aus einem konstanten Term $\tilde{\sigma} \in (0, \gamma)$ (natürlicher Tod) und einem zu x_2 proportionalen Term bx_2 (Tod durch Räuber). Für $x_2 = 0$ wächst die Population exponentiell. Wir setzen $a = \gamma - \tilde{\sigma} > 0$.
- (ii) Die Räuberpopulation verhält sich ebenfalls gemäß (2.2) mit $\lambda = \gamma - \sigma$. Hier ist die Sterbequote σ konstant und $\gamma = \tilde{\gamma} + dx_1$ für $\tilde{\gamma} \in (0, \sigma)$ und $d > 0$, d.h. die Geburtenquote hängt affin linear von der Anzahl der zur Verfügung stehenden Beute x_1 ab; für $x_1 = 0$ stirbt die Räuberpopulation wegen $\sigma > \tilde{\gamma}$ aus. Wir setzen $c = \sigma - \tilde{\gamma} > 0$.

Zusammen erhalten wir so die zweidimensionale Differentialgleichung

$$\begin{aligned} \dot{x}_1(t) &= ax_1(t) - bx_1(t)x_2(t) \\ \dot{x}_2(t) &= -cx_2(t) + dx_1(t)x_2(t) \end{aligned} \quad (2.6)$$

mit den Parametern $a, b, c, d > 0$. Dieses Modell wird als *Lotka–Volterra Modell* bezeichnet. V. Volterra⁴ hat dieses Modell im biologischen Kontext eingeführt (vgl. dazu Abschnitt 2.3.1), A.J. Lotka⁵ hat das Modell unabhängig von Volterra zur Beschreibung einer hypothetischen chemischen Reaktion entwickelt.

Um die Analyse von (2.6) zu vereinfachen wollen wir die Zahl der Parameter zu reduzieren. Dazu führt man die Koordinatentransformation $x_1 \rightarrow \frac{d}{c}x_1$ und $x_2 \rightarrow \frac{b}{a}x_2$ durch. Dies führt auf die neuen Gleichungen

$$\begin{aligned} \dot{x}_1(t) &= ax_1(t) - ax_1(t)x_2(t) = ax_1(t)(1 - x_2(t)) \\ \dot{x}_2(t) &= -cx_2(t) + cx_1(t)x_2(t) = -cx_2(t)(1 - x_1(t)) \end{aligned} \quad (2.7)$$

Beachte, dass die Lösungen $\tilde{x}(t; t_0, \tilde{x}_0)$ von (2.6) und $x(t; t_0, x_0)$ von (2.7) mittels

$$x(t; t_0, x_0) = A\tilde{x}(t; t_0, A^{-1}x_0) \quad \text{und} \quad \tilde{x}(t; t_0, \tilde{x}_0) = A^{-1}x(t; t_0, A\tilde{x}_0) \quad \text{für} \quad A = \begin{pmatrix} \frac{d}{c} & 0 \\ 0 & \frac{b}{a} \end{pmatrix}$$

zusammenhängen; alle Lösungen von (2.6) lassen sich also aus (2.7) berechnen und umgekehrt. Man nennt die zwei Gleichungen auch *äquivalent*.

Wir bestimmen zunächst die Gleichgewichte von (2.7), also die Nullstellen des Vektorfeldes

$$f(x) = \begin{pmatrix} ax_1(1 - x_2) \\ -cx_2(1 - x_1) \end{pmatrix}.$$

⁴italienischer Physiker und Mathematiker, 1860–1940

⁵US–amerikanischer Chemiker und Mathematiker, 1880–1949

Hier sieht man leicht, dass die Punkte

$$x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{und} \quad x^+ = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

die einzigen Gleichgewichtspunkte sind. Zur Bestimmung der Stabilität dieser Gleichgewichte berechnen wir

$$Df(x^*) = \begin{pmatrix} a(1-x_2^*) & -ax_1^* \\ cx_2^* & -c(1-x_1^*) \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & -c \end{pmatrix} \quad \text{und} \quad Df(x^+) = \begin{pmatrix} 0 & -a \\ c & 0 \end{pmatrix}$$

Als Eigenwerte dieser Matrizen ergeben sich a und $-c$ in x^* sowie $\pm\sqrt{-ca}$ in x^+ . Aus Satz 2.4 folgt damit exponentielle Instabilität (aber nicht Antistabilität) von x^* . Dies ist gut zu erklären: Für Anfangswerte der Form $x_0 = (x_1, 0)^T$ (also keine Räuber) mit $x_1 \neq 0$ wächst der Betrag der Lösung exponentiell, sie läuft also exponentiell von $x^* = 0$ weg. Die Menge aller Punkte, die exponentiell wegläuft, heißt *instabile Mannigfaltigkeit* $M_i(x^*)$ von x^* , hier ist das einfach der Unterraum $M_i(x^*) = \langle (1, 0)^T \rangle$. Umgekehrt konvergieren alle Lösungen zu Anfangswerten der Form $x_0 = (0, x_2)^T$ (also keine Beute) mit $x_2 \in \mathbb{R}$ exponentiell gegen $x^* = 0$, dies ist die sogenannte *stabile Mannigfaltigkeit* $M_s(x^*)$, hier wiederum ein Unterraum, nämlich $M_s(x^*) = \langle (0, 1)^T \rangle$.

Auf x^+ trifft keiner der Fälle in Satz 2.4 zu, da hier beide Eigenwerte wegen $ca > 0$ offenbar die Realteile 0 besitzen. Wir wissen also, dass Lösungen weder exponentiell konvergieren noch wegläuft können. Was aber passiert statt dessen? Um sich einen Überblick über das Verhalten dieses Systems zu verschaffen, empfiehlt sich hier die numerische Lösung und Darstellung in Kurvenform, die in Abbildung 2.3 mit $a = c = 1$ zu sehen ist.

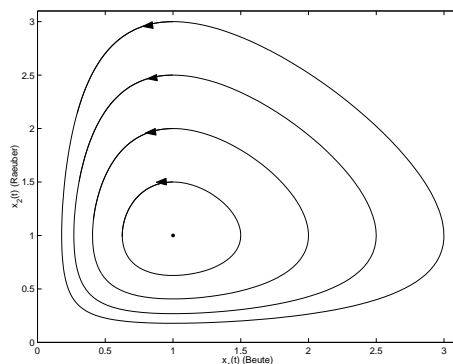


Abbildung 2.3: Lösungen der Räuber–Beute Modells (2.7) mit $a = c = 1$

Man erkennt in dieser Grafik gut, warum die Ruhelage $x^+ = (1, 1)^T$ weder exponentiell stabil noch exponentiell instabil ist: Alle Lösungen, die nicht auf $M_s(x^*)$ oder $M_i(x^*)$ liegen, laufen auf periodischen Bahnen um dieses x^+ herum, weder konvergieren sie noch laufen sie weg. Formal nennt man eine Lösung $x(t; t_0, x_0)$ *periodisch*, falls ein $T > 0$ existiert, so dass

$$x(t; t_0, x_0) = x(t + T; t_0, x_0)$$

gilt für alle $t \in \mathbb{R}$. Die Zeit $T > 0$ heißt die *Periode* der Lösung. (Wir verlangen hier i.A. nicht, dass T die minimale Periode ist.) Beachte, dass eine Lösung $x(t)$ einer autonomen

Gleichung genau dann periodisch ist, wenn es zwei Zeiten $t_1 < t_2 \in \mathbb{R}$ gibt, so dass $x(t_1) = x(t_2) =: x_P$ gilt. Aus dieser Gleichheit folgt nämlich sowohl $x(t) = x(t; t_1, x_P)$ als auch $x(t) = x(t; t_2, x_P)$. Aus (1.6) folgt damit $x(t+t_2-t_1) = x(t)$ für alle $t \in \mathbb{R}$, also Periodizität für $T = t_2 - t_1$.

Wir wollen diese numerische Erkenntnis nun mathematisch rigoros beweisen. Dazu betrachten wir den Quotienten

$$\frac{\dot{x}_2(t)}{\dot{x}_1(t)} = \frac{-cx_2(t)(1-x_1(t))}{ax_1(t)(1-x_2(t))}.$$

Aus dieser Gleichung folgt

$$ax_1(t)\dot{x}_2(t) - ax_1(t)x_2(t)\dot{x}_2(t) = -cx_2(t)\dot{x}_1(t) + cx_2(t)x_1(t)\dot{x}_1(t)$$

und damit

$$c\dot{x}_1(t) - c\frac{\dot{x}_1(t)}{x_1(t)} + a\dot{x}_2(t) - a\frac{\dot{x}_2(t)}{x_2(t)} = 0.$$

Beachte, dass alle diese Gleichungen nur gelten, wenn alle Nenner ungleich Null sind, also nur für Lösungen $x(t) = (x_1(t), x_2(t))$, die sich in $\mathbb{R}^+ \times \mathbb{R}^+$ befinden und keine Gleichgewichte sind.

Integrieren wir diese Gleichung nun von 0 bis t , so erhalten wir

$$cx_1(t) - c \log x_1(t) + ax_2(t) - a \log x_2(t) = k(x(0))$$

mit $k(x(0)) = cx_1(0) - c \log x_1(0) + ax_2(0) - a \log x_2(0)$. Die auf $D_V = \mathbb{R}^+ \times \mathbb{R}^+$ definierte Funktion

$$V(x) = cx_1 - c \log x_1 + ax_2 - a \log x_2 \tag{2.8}$$

ist also konstant entlang von Lösungen; es gilt

$$V(x(t; x_0)) = V(x_0) \text{ für alle } t \geq 0$$

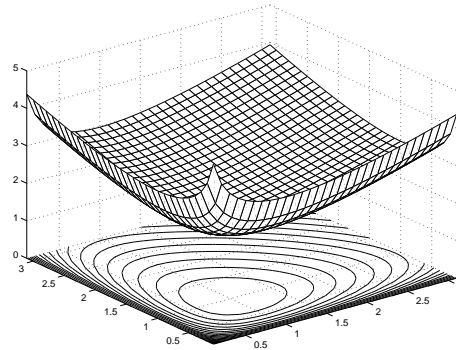
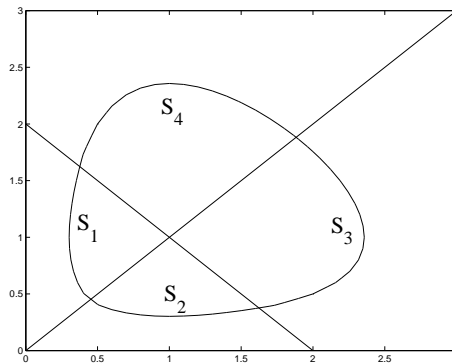
bzw.

$$\frac{d}{dt}V(x(t; x_0)) = 0.$$

V heißt *Konstante der Bewegung* für (2.7). Die Lösungen von (2.7) mit Anfangswert $x_0 \in D_V$ laufen also entlang der Höhenlinien $V^{-1}(l) := \{x \in D_V \mid V(x) = l\}$ von V , die in Abbildung 2.4 gemeinsam mit dem Graphen von V skizziert sind. Man sagt, die Höhenlinien $V^{-1}(l)$ sind *invariante Mengen* bezüglich (2.7). Beachte, dass V ein globales Minimum in x^+ mit $V(x^+) = c + a$ besitzt.

Dass die Lösungen tatsächlich periodisch sind, folgt aus der Analyse des Vektorfeldes auf den Höhenlinien. Wir betrachten eine Höhenlinie $V^{-1}(l)$ für ein $l > V(x^+)$ und teilen $V^{-1}(l)$ in die vier Segmente

$$\begin{aligned} S_1 &= \{x \in V^{-1}(l) \mid x_1 \leq x_2 \leq 2 - x_1\} \\ S_2 &= \{x \in V^{-1}(l) \mid x_2 \leq x_1 \leq 2 - x_2\} \\ S_3 &= \{x \in V^{-1}(l) \mid x_1 \geq x_2 \geq 2 - x_1\} \\ S_4 &= \{x \in V^{-1}(l) \mid x_2 \geq x_1 \geq 2 - x_2\} \end{aligned}$$

Abbildung 2.4: Graph und Höhenlinien von V aus (2.8) mit $a = c = 1$ Abbildung 2.5: Segmente S_1 , S_2 , S_3 und S_4

ein, vgl. Abbildung 2.5.

Aus der Form der Höhenlinien folgt jetzt, dass ein $\alpha > 0$ existiert, so dass $|x_1 - 1| \geq \alpha$ gilt für alle $x \in S_1$ und $x \in S_3$ und $|x_2 - 1| \geq \alpha$ gilt für alle $x \in S_2$ und alle $x \in S_4$. Desweiteren existiert ein $\beta > 0$ mit $x_1 > \beta$ und $x_2 > \beta$ für alle $x \in V^{-1}(l)$. Aus der Gleichung 2.7 folgert man damit die Ungleichungen

$$\begin{aligned} \dot{x}_2(t) &< -c\beta\alpha, & \text{falls } x(t) \in S_1 \\ \dot{x}_1(t) &> a\beta\alpha, & \text{falls } x(t) \in S_2 \\ \dot{x}_2(t) &> c\beta\alpha, & \text{falls } x(t) \in S_3 \\ \dot{x}_1(t) &< -a\beta\alpha, & \text{falls } x(t) \in S_4 \end{aligned}$$

In jedem Sektor ist also eine der beiden Komponenten $x_1(t)$ oder $x_2(t)$ streng monoton wachsend oder fallend mit von 0 (gleichmäßig in t) verschiedener Steigung. Deswegen muss jeder Sektor nach einer endlichen Zeit verlassen werden, und zwar in der Reihenfolge $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4 \rightarrow S_1$. Die Lösung ist also tatsächlich periodisch.

Um die Aussagen des Modells für die modellierten Populationsgrößen zu interpretieren, ist es sinnvoll, eine beispielhafte Lösung in Abhängigkeit von t darzustellen. In Abbildung 2.6 ist dies gemacht.

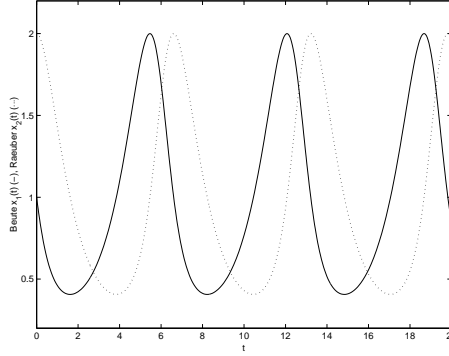


Abbildung 2.6: Lösung von (2.7) mit $x_0 = (1, 2)^T$ und $a = c = 1$

Beide Populationen schwanken also periodisch. Wenn (wie am Anfang) viele Räuber und wenig Beute vorhanden sind, nehmen beide Populationen ab. Wenn die Zahl der Räuber unter einer gewissen Schwelle liegt, nimmt die Beutepopulation wieder zu. Wenn genügend Beute vorhanden ist, beginnt auch die Räuberpopulation wieder zuzunehmen und wenn diese eine kritische Marke überschritten hat, nimmt die Zahl der Beute wieder ab, usw. Ein solches Verhalten ist in der Natur durchaus zu beobachten.

2.2.2 Das Räuber–Beute Modell mit beschränkten Ressourcen

Im Modell (2.6) hat die (unrealistische) Eigenschaft, dass sich die Beutepopulation in Abwesenheit der Räuber gemäß (2.2) verhält, also unbeschränkt wächst. Wir wollen dies durch das realistischere Modell (2.3) ersetzen, das wir hier mit $\mu = \lambda K$ und $e = \lambda$ als

$$\dot{x}(t) = \mu x_1(t) - e x_1(t)^2 \quad (2.9)$$

schreiben. Wir ändern damit die Modellannahme (i) wie folgt ab.

- (i') Die Beutepopulation x_2 verhält sich gemäß (2.9) mit $\mu = \gamma - \sigma$ und $e > 0$, wobei e und γ konstant sind und $\sigma = \tilde{\sigma} + b x_2$. Für die Beutetiere gibt es also begrenzte Ressourcen und die Sterbequote σ besteht aus einem konstanten Term $\tilde{\sigma} \in (0, \gamma)$ (natürlicher Tod) und einem zu x_2 proportionalen Term $b x_2$ (Tod durch Räuber). Für $x_2 = 0$ konvergiert die Populationsgröße gegen $K = a/e$ mit $a = \gamma - \tilde{\sigma} > 0$.

Damit erhalten wir die Gleichung

$$\begin{aligned} \dot{x}_1(t) &= a x_1(t) - b x_1(t) x_2(t) - e x_1(t)^2 \\ \dot{x}_2(t) &= -c x_2(t) + d x_1(t) x_2(t) \end{aligned} \quad (2.10)$$

Analog zu (2.6) können wir diese Gleichung durch eine lineare Koordinatentransformation vereinfachen. Hier transformieren wir $x_1 \rightarrow \frac{d}{c} x_1$, $x_2 \rightarrow \frac{bd}{da - ec} x_2$ und erhalten so

$$\begin{aligned} \dot{x}_1(t) &= \alpha x_1(t)(1 - x_2(t)) + \beta x_1(t)(1 - x_1(t)) \\ \dot{x}_2(t) &= -c x_2(t)(1 - x_1(t)) \end{aligned} \quad (2.11)$$

mit $\alpha = a - ec/d$ und $\beta = ec/d$. Hier muss man aufpassen, dass bei dieser Transformation positive x_1, x_2 wieder auf positive x_1, x_2 abgebildet werden. Da $a, b, c, d, e > 0$ sind, ist dies genau dann der Fall, wenn $\frac{bd}{da-ec} > 0$ ist, also wenn $ad > ec$ gilt. Wir wollen uns auf diesen Fall einschränken, nicht nur aus formalen Gründen, sondern auch aus Modellierungsgründen: Für $ad \leq ec$ kann man zeigen, dass die Räuberpopulation für $t \rightarrow \infty$ für alle Anfangswerte ausstirbt, wir wollen hier aber den Fall der langfristigen Koexistenz beider Arten betrachten, für den $ad > ec$ eine notwendige Bedingung ist.

Als Gleichgewichte erhält man hier $x^* = (0, 0)^T$, $x^{**} = ((\alpha + \beta)/\beta, 0)^T$ und $x^+ = (1, 1)^T$. Nur x^+ liegt in $\mathbb{R}^+ \times \mathbb{R}^+$, weswegen wir dieses Gleichgewicht genauer untersuchen wollen.

Die Linearisierung ergibt sich zu

$$Df(x) = \begin{pmatrix} \alpha(1-x_2) + \beta(1-2x_1) & -\alpha x_1 \\ cx_2 & -c(1-x_1) \end{pmatrix}$$

also

$$Df(x^+) = \begin{pmatrix} -\beta & -\alpha \\ c & 0 \end{pmatrix}.$$

Die Eigenwerte dieser Matrix sind

$$\lambda_{1/2} = -\frac{\beta}{2} \pm \sqrt{\frac{\beta^2}{4} - \alpha c}.$$

Falls die Wurzel komplex ist, sind die Realteile $-\beta/2$ negativ, falls die Wurzel reell ist, sind auch $\lambda_{1/2}$ reell und es gilt

$$\lambda_{1/2} \leq -\frac{\beta}{2} + \sqrt{\frac{\beta^2}{4} - \alpha c} < -\frac{\beta}{2} + \sqrt{\frac{\beta^2}{4}} = 0,$$

also erhalten wir in beiden Fällen negative Realteile, weswegen x^+ lokal exponentiell stabil ist. Wir wissen also insbesondere, dass es eine Umgebung von x^+ gibt, so dass alle Lösungen in dieser Umgebung gegen x^+ konvergieren. Was ist aber nun der Einzugsbereich $\mathcal{D}(x^+)$?

Dieser lässt sich hier analytisch ermitteln: Betrachte dazu die Konstante der Bewegung (2.8)

$$V(x) = cx_1 - c \log x_1 + \alpha x_2 - \alpha \log x_2.$$

Im Gegensatz zu (2.7) ist diese Funktion für (2.11) nicht mehr konstant entlang von Lösungen, statt dessen gilt für jede Lösung $x(t)$ in D_V die Gleichung

$$\begin{aligned} \frac{d}{dt}V(x(t)) &= c\dot{x}_1(t) - c\frac{\dot{x}_1(t)}{x_1(t)} + \alpha\dot{x}_2(t) - \alpha\frac{\dot{x}_2(t)}{x_2(t)} \\ &= \left(c\alpha x_1(t)(1-x_2(t)) + c\beta x_1(t)(1-x_1(t)) \right) \left(1 - \frac{1}{x_1(t)} \right) \\ &\quad - \alpha c x_2(t)(1-x_1(t)) \left(1 - \frac{1}{x_2(t)} \right) \\ &= c\beta(1-x_1(t))(x_1(t)-1) = -c\beta(x_1(t)-1)^2 \end{aligned}$$

Die Funktion $V(x(t))$ fällt also monoton in t , für $x_1(t) \neq 1$ sogar streng monoton. Beachte, dass $V(x)$ in $x = x^+$ ein globales Minimum besitzt; weitere lokale Minima existieren nicht.

Wir beweisen nun $x(t) \rightarrow x^+$ für $t \rightarrow \infty$. Da $V(x(t))$ monoton fällt und nach unten beschränkt ist, konvergiert $V(x(t))$ gegen einen Wert V_∞ . Ähnlich wie im Beweis von Satz 2.2 sieht man nun, dass $\frac{d}{dt}V(x(t)) \rightarrow 0$ für $t \rightarrow \infty$ gilt, also muss $x_1(t) \rightarrow 1$ für $t \rightarrow \infty$ gelten. Dies ist aber nur dann möglich, falls $x_2(t) \rightarrow 1$ konvergiert: Wäre $|x_2(t) - 1| \geq \delta$ so würde aus (2.11) für $x_1(t)$ in einer Umgebung der 1 entweder $\dot{x}_1(t) > \varepsilon$ oder $\dot{x}_1(t) < -\varepsilon$ folgen, was der Konvergenz $x_1 \rightarrow 1$ widersprechen würde. Also gilt $x_2(t) \rightarrow 1$ und damit $x(t) \rightarrow x^+$. Alle Lösungen mit Anfangswerten in D_V konvergieren also gegen x^+ , weswegen $\mathcal{D}(x^+) = \mathbb{R}^+ \times \mathbb{R}^+$ ist. Die numerischen Ergebnisse in Abbildung 2.7 bestätigen dieses Ergebnis.

Die Argumentation, die wir hier verwendet haben, ist als *Lasalles Invarianzprinzip* bekannt und lässt sich auch allgemein als Satz formulieren, was wir hier aber nicht vertiefen wollen.

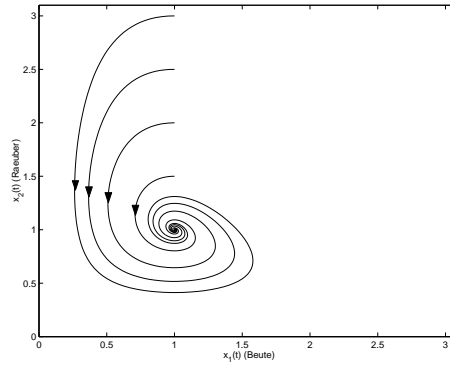


Abbildung 2.7: Lösungen des Räuber–Beute Modells (2.11) mit $a = c = 1$, $\beta = 0.5$

Zur Interpretation des Modells ist wieder die Darstellung einer Lösung in Abhängigkeit von t nützlich, wie sie in Abbildung 2.8 gegeben ist.

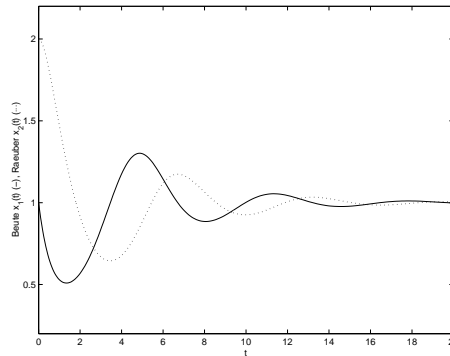


Abbildung 2.8: Lösung von (2.11) mit $x_0 = (1, 2)^T$ und $a = c = 1$, $\beta = 0.5$

Die Lösung zeigt zwar ein ähnliches Schwankungsverhalten wie in Abbildung 2.6, konvergiert aber mit wachsendem t gegen ein Gleichgewicht. Solche Gleichgewichte verschiedener gemeinsam existierender Arten sind in realen Ökosystemen oft zu beobachten, ebenso wie die charakteristischen Schwankungen, die auftreten, wenn das System durch äußere Ein-

flüsse “aus dem Gleichgewicht” gebracht wurde.

Auf eine abschließende Diskussion dieses Modells können wir hier verzichten, da hier exakt die gleichen Unzulänglichkeiten wie in der Diskussion im Abschnitt 2.1.4 bestehen, mit Ausnahme des letzten Punktes natürlich.

2.2.3 Verallgemeinerung auf n Arten

Wir wollen in diesem Abschnitt abschließend auf die Verallgemeinerung des Modells (2.10) auf n verschiedene Arten x_1, \dots, x_n eingehen. Wenn wir für alle Arten die gleichen Modellannahmen treffen, nämlich, dass die Dynamik durch (2.9) gegeben ist, wobei die jeweilige Wachstumsrate μ affin linear von allen anderen Arten abhängt, so erhalten wir das Modell

$$\dot{x}_i(t) = k_i x_i(t) + b_i^{-1} \sum_{j=1}^n a_{ij} x_i(t) x_j(t), \quad i = 1, \dots, n. \quad (2.12)$$

mit $k_i \neq 0$, $a_{ii} \leq 0$, $a_{ij} \in \mathbb{R}$ und $b_i > 0$. Wir definieren mittels der a_{ij} die Matrix $A = (a_{ij})$. Der Koeffizient a_{ii} entspricht für jede Art gerade dem e aus (2.9), er modelliert also die Ressourcenbeschränkung, während die a_{ij} für $i \neq j$ die Interaktion der Arten beschreibt. Für Beute x_i und Räuber x_j muss die Bedingung $a_{ij} < 0$ und $a_{ji} > 0$ gelten. Die etwas seltsam anmutende Notation mit b_i^{-1} ergibt sich aus der ursprünglichen, etwas anderen Schreibweise des Modells. Beachte, dass die Modelle (2.6) und (2.10) Spezialfälle dieses Modells sind.

Der Spezialfall $a_{ii} = 0$ und $a_{ij} = -a_{ji}$ wird als *Volterra-Ökologie* bezeichnet. In diesem Fall ist die Matrix $A = (a_{ij})$ antisymmetrisch, d.h. $x^T A x = 0$ für alle $x \in \mathbb{R}^n$.

Wenn wir nach Gleichgewichten x^+ suchen, für die alle Arten koexistieren, so gilt $x_i^+ > 0$, also

$$k_i x_i^+ + b_i^{-1} \sum_{j=1}^n a_{ij} x_i^+ x_j^+ = 0 \quad \Rightarrow \quad b_i k_i + \sum_{j=1}^n a_{ij} x_j^+ = 0, \quad (2.13)$$

diese Gleichgewichte sind also als Lösungen eines linearen Gleichungssystems gegeben. Wenn A invertierbar ist, existiert also höchstens ein solches Gleichgewicht: es gibt genau eine Lösung x^* des linearen Gleichungssystems, für die aber nicht $x_i^* > 0$ gelten muss.

Die Konstruktion der Konstanten der Bewegung V lässt sich auf dieses Modell verallgemeinern. Wenn ein Gleichgewicht x^+ mit $x_i^+ > 0$ für $i = 1, \dots, n$ existiert, so kann man nachrechnen, dass die Funktion

$$V(x) = \sum_{i=1}^n b_i (x_i - x_i^+ \log x_i) \quad (2.14)$$

die Gleichung

$$\frac{d}{dt} V(x(t)) = (x(t) - x^+)^T A (x(t) - x^+)$$

erfüllt. Falls A negativ semidefinit ist, so ist diese Ableitung ≤ 0 und wir können die obige Argumentation auf das n -dimensionale Modell übertragen. Im Falle einer Volterra-Ökologie ist A antisymmetrisch, weswegen $\frac{d}{dt} V(x(t)) = 0$ ist. Hier erhalten wir also wieder das Phänomen periodischer Lösungen.

2.3 Anwendungen der Populationsdynamik

2.3.1 Auswirkungen der Befischung

Die Volterra-Ökologie und speziell das Lotka-Volterra Modell gilt in der Biologie i.A. als zu stark vereinfacht, da hierbei in dem sowieso schon einfachen Modell (2.12) noch weitere strukturelle Vereinfachungen gemacht werden. Man muss aber berücksichtigen, dass dieses Modell zur Erklärung eines speziellen Sachverhaltes entwickelt wurde, für den es tatsächlich gut funktioniert. Wir wollen diese Anwendung nun erläutern.

In den 1920er Jahren beobachtete der italienische Biologe D'Ancona, dass der Anteil der Raubfische (Haie, Rochen, ...) am Gesamtfischfang während des 1. Weltkrieges im Mittelmeer deutlich höher als vorher und nachher war. Im Hafen Fiume in Italien wurden die folgenden Anteile der Raubfische am Gesamtfang festgestellt:

Jahr	1914	1915	1916	1917	1918
Raubfischanteil	11,9%	21,4%	22,1%	21,2%	36,4%
Jahr	1919	1920	1921	1922	1923
Raubfischanteil	27,3%	16,0%	15,9%	14,8%	10,7%

Natürlich war D'Ancona klar, dass während des Krieges weniger gefischt wurde, aber warum sollte dies die Raubfische mehr begünstigen?

Das Volterra-Modell wurde zur Erklärung dieses Phänomens entwickelt. Tatsächlich handelt es sich hier nur um zwei (Gruppen von) Arten, so dass sich (2.12) zu (2.6) vereinfacht, wenn man $a = k_1$, $c = -k_2$, $b = -b_1^{-1}a_{12}$, $d = b_2^{-1}a_{21} = -b_2^{-1}a_{12}$ setzt (dies zeigt insbesondere, dass (2.6) ein Spezialfall der Volterra-Ökologie ist). Wie kann dieses Modell mit den bekannten periodischen Lösungen aus den Abbildungen 2.3 und 2.6 das Phänomen beschreiben? Die Werte in der obigen Tabelle sind Jahresmittelwerte, weswegen es sich anbietet, auch die vom Modell gegebenen Werte zu mitteln. Hier gilt das folgende Lemma.

Lemma 2.6 Sei $x(t)$ eine periodische Lösung von (2.6) mit Periode T . Dann gilt

$$\bar{x}_1 := \frac{1}{T} \int_0^T x_1(t) dt = \frac{c}{d} \quad \text{und} \quad \bar{x}_2 := \frac{1}{T} \int_0^T x_2(t) dt = \frac{a}{b}.$$

Beweis: Es gilt

$$\frac{\dot{x}_1(t)}{x_1(t)} = a - bx_2(t).$$

Für diesen Ausdruck gilt

$$\frac{1}{T} \int_0^T \frac{\dot{x}_1(t)}{x_1(t)} dt = \frac{1}{T} \int_0^T a - bx_2(t) dt.$$

Andererseits gilt

$$\begin{aligned} \frac{1}{T} \int_0^T \underbrace{\frac{\dot{x}_1(t)}{x_1(t)}}_{= \frac{d}{dt} \log x_1(t)} dt &= \frac{1}{T} (\log x(T) - \log x(0)) = 0, \\ &= \frac{d}{dt} \log x_1(t) \end{aligned}$$

da die Lösung periodisch mit $x(T) = x(0)$ ist. Also folgt

$$0 = \frac{1}{T} \int_0^T a - bx_2(t) dt = a - b\bar{x}_2$$

und damit die Behauptung für \bar{x}_2 . Analog berechnet man den Wert für \bar{x}_1 . \square

Die zunächst vielleicht etwas überraschende Erkenntnis dieses Lemmas ist, dass die Mittelwerte über eine Periode nicht vom Anfangswert abhängen. Der Anteil der Raubfische an der Gesamtmenge ist im Mittel also gegeben durch

$$\bar{x}_2^A = \frac{\bar{x}_2}{\bar{x}_1 + \bar{x}_2} = \frac{\frac{a}{b}}{\frac{a}{b} + \frac{c}{d}} = \frac{ad}{ad + cb}$$

Um die veränderten Anteile während des 1. Weltkrieges zu erklären, müssen wir den Fischfang in (2.6) berücksichtigen. Nimmt man hier proportionale Fangraten px_1 und px_2 an, so ergibt sich das Modell mit Fischfang zu

$$\begin{aligned} \dot{x}_1(t) &= (a - p)x_1(t) - bx_1(t)x_2(t) \\ \dot{x}_2(t) &= -(c + p)x_2(t) + dx_1(t)x_2(t) \end{aligned}$$

Der mittlere Raubfischanteil bei Fangrate p ist demnach

$$\bar{x}_2^A(p) = \frac{(a - p)d}{(a - p)d + (c + p)b}$$

oder als Kehrwert ausgedrückt

$$\bar{x}_2^A(p)^{-1} = \frac{(a - p)d + (c + p)b}{(a - p)d} = \frac{(c + p)b}{(a - p)d} + 1.$$

Wenn also die Fangrate p abnimmt, so verringert sich der Bruch $\frac{(c+p)b}{(a-p)d}$ ebenfalls (der Zähler wird kleiner und der Nenner größer), womit auch $\bar{x}_2^A(p)^{-1}$ kleiner wird, weswegen der Raubfischanteil $\bar{x}_2^A(p)$ zunimmt. Das Modell liefert also eine Erklärung dafür, warum bei geringerer Befischung der Raubfischanteil zunimmt.

2.3.2 Selektion gleichartiger Spezies (in den Übungen behandelt)

Ähnlich wie wir das beim Räuber–Beute Modell (2.10) gemacht haben, wollen wir hier wieder die Interaktion zweier Populationen beschreiben, die dem logistischen Wachstum (2.3) unterliegen. Diesmal wollen wir aber nicht Räuber und Beute beschreiben, sondern zwei friedlich koexistierende Arten x_1 und x_2 modellieren, die teilweise um die gleichen Ressourcen konkurrieren.

Beide Arten sollen also durch die Gleichung

$$\dot{x}_i(t) = \lambda_i x_i(t)(K_i - x_i(t))$$

beschrieben werden. Um die Konkurrenz zu modellieren, ersetzen wir K_i durch $K_i - m_i$ und machen dazu die Modellannahme, dass m_1 der Anteil der Ressourcen K_1 von x_1 ist, der von

x_2 in Anspruch genommen wird, und umgekehrt. Die einfachste Art, dies zu modellieren, ist die Wahl

$$m_1 = \alpha x_2 \quad \text{und} \quad m_2 = \beta x_1.$$

Das Gesamtmodell ergibt sich so zu

$$\begin{aligned} \dot{x}_1(t) &= \lambda_1 x_1(t)(K_1 - x_1(t) - \alpha x_2(t)) \\ \dot{x}_2(t) &= \lambda_2 x_2(t)(K_2 - x_2(t) - \beta x_1(t)) \end{aligned} \quad (2.15)$$

Dieses Modell ist wieder ein Spezialfall des Modells (2.12), mit $k_i = \lambda_i K_i$, $b_i = 1$, $a_{ii} = -\lambda_i$, und $a_{12} = -\lambda_1 \alpha$ und $a_{21} = -\lambda_2 \beta$.

Wir analysieren hier zunächst den Fall $\alpha = \beta = 1$ und betrachten, wie das Verhalten der Lösungen von K_1 und K_2 abhängt. Beachte zunächst, dass Lösungen, die auf der x_1 - bzw. x_2 -Achse starten, für alle Zeiten dort bleiben. Umgekehrt bedeutet dies, da sich Lösungen nicht schneiden können, dass Lösungen in $\mathbb{R}^+ \times \mathbb{R}^+$ für alle Zeiten in $\mathbb{R}^+ \times \mathbb{R}^+$ bleiben. Eine Art kann also nicht in endlicher Zeit aussterben, dies kann aber durchaus für $t \rightarrow \infty$ passieren. Genau dieser Fall tritt hier ein; es gilt:

Falls $\alpha = \beta = 1$ und $K_1 > K_2$ ist, so konvergiert jede Lösung $x(t; t_0, x_0)$ mit $x_0 \in \mathbb{R}^+ \times \mathbb{R}^+$ gegen $x^* = (K_1, 0)^T$ für $t \rightarrow \infty$. Mit anderen Worten: Unabhängig von den Wachstumsraten λ_i überlebt nur die Art mit den größeren Ressourcen K_i (denn für $K_2 > K_1$ ergibt sich wegen der Symmetrie des Modells gerade das umgekehrte Verhalten).

Wir wollen den Beweis dieser Behauptung skizzieren. Zunächst rechnet man nach, dass die Gleichung (2.15) genau die Gleichgewichte $(0, 0)^T$, $(K_1, 0)^T$ und $(0, K_2)^T$ besitzt. Da kein Gleichgewicht $x^+ \in \mathbb{R}^+ \times \mathbb{R}^+$ existiert, können wir V aus (2.14) nicht zur Analyse verwenden, wir müssen das Modell also direkt analysieren.

Dazu teilt man den positiven Quadranten $\mathbb{R}^+ \times \mathbb{R}^+$ auf in die drei Bereiche

$$\begin{aligned} A &:= \{(x_1, x_2) \mid x_1 > 0, x_2 > 0, K_1 \geq x_1 + x_2\} \\ B &:= \{(x_1, x_2) \mid x_1 > 0, x_2 > 0, K_1 \leq x_1 + x_2 \leq K_2\} \\ C &:= \{(x_1, x_2) \mid x_1 > 0, x_2 > 0, K_2 \leq x_1 + x_2\}, \end{aligned}$$

vgl. Abbildung 2.9.

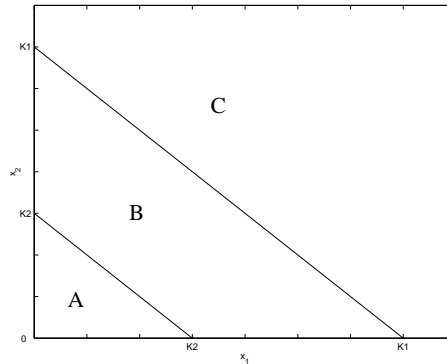


Abbildung 2.9: Bereiche von $\mathbb{R}^+ \times \mathbb{R}^+$

Nun unterscheidet man die folgenden Fälle:

(1) In $\text{int } A$ ist $\dot{x}_1(t) > 0$ und $\dot{x}_2(t) > 0$, die Lösungen wachsen also in beiden Komponenten streng monoton. Für $x_0 \in \mathbb{R}^+ \times \mathbb{R}^+$ muss die Lösung also entweder nach $\text{int } B$ laufen, oder gegen ein Gleichgewicht $x^* \in A$ konvergieren, für das $x_1^* > x_{0,1} > 0$ und $x_2^* > x_{0,1} > 0$ gilt. Da ein solches Gleichgewicht nicht existiert, müssen die Lösungen also nach $\text{int } B$ laufen.

(2) In $\text{int } C$ gilt $\dot{x}_1(t) < 0$ und $\dot{x}_2(t) < 0$, die Lösungen fallen also in beiden Komponenten streng monoton. Folglich muss die Lösung hier entweder nach $\text{int } B$ laufen oder gegen ein Gleichgewicht $x^* \in C$ konvergieren. Dies kann nur $x^* = (K_1, 0)^T$ sein, so dass in diesem Fall die Behauptung gezeigt ist.

(3) In $\text{int } B$ gilt $\dot{x}_1(t) > 0$ und $\dot{x}_2(t) < 0$, $x_1(t)$ wächst und $x_2(t)$ fällt also streng monoton. Man rechnet nach, dass eine Lösung $x(t)$, die für ein t^* in $\text{int } B$ liegt, für alle zukünftigen Zeiten $t \geq t^*$ auch in $\text{int } B$ liegt (dies leitet man aus den Nullstellen der Ableitungen \dot{x}_1 bzw. \dot{x}_2 an den Übergängen zwischen den Mengen ab). Also müssen beide Komponenten $x_1(t)$ und $x_2(t)$ konvergieren, weswegen die Gesamtlösung auch konvergieren muss, und zwar gegen ein Gleichgewicht $x^* \in B$. Da $x_1(t)$ wächst und $x_2(t)$ fällt, kann dies nur $x^* = (K_1, 0)^T$ sein.

Abbildung 2.10 zeigt verschiedene numerisch berechnete Trajektorien von (2.15), die diese Analyse bestätigen.

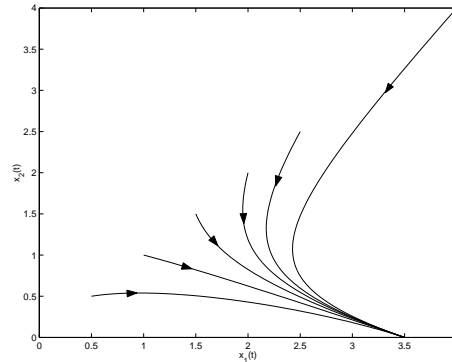


Abbildung 2.10: Lösungen von (2.15) mit $\alpha = \beta = 1$, $\lambda_1 = \lambda_2 = 1$, $K_1 = 3.5$, $K_2 = 1.5$

Die Situation ändert sich deutlich, wenn wir die Annahme $\alpha = \beta = 1$ wegfallen lassen. Wenn also die Art x_1 z.B. auf andere Nahrungsressourcen ausweichen kann, die von x_2 nicht beansprucht werden, so würde sich K_1 vergrößern und β verkleinern. Tatsächlich reicht es aus, β zu verkleinern, um eine langfristige Koexistenz der Arten zu erreichen. Für $\beta < K_2/K_1 < 1$ liegt das mittels (2.13) errechnete Gleichgewicht

$$x^+ = \begin{pmatrix} \frac{K_1 - \alpha K_2}{1 - \alpha\beta} \\ \frac{K_2 - \beta K_1}{1 - \alpha\beta} \end{pmatrix}$$

in $\mathbb{R}^+ \times \mathbb{R}^+$. Zudem ist $A = (a_{ij})$ negativ definit, so dass V aus (2.14) entlang aller Lösungen streng monoton fällt, weswegen alle Lösungen in $\mathbb{R}^+ \times \mathbb{R}^+$ gegen x^+ konvergieren müssen. Mit $\beta = 0.1$ ergeben sich die in Abbildung 2.11 dargestellten Lösungen, die dieses Verhalten bestätigen.

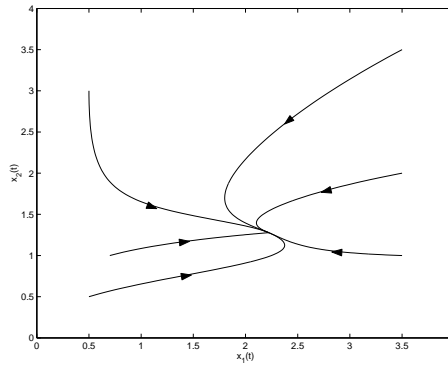


Abbildung 2.11: Lösungen von (2.15) mit $\alpha = 1$, $\beta = 0.1$, $\lambda_1 = \lambda_2 = 1$, $K_1 = 3.5$, $K_2 = 1.5$

2.4 Ausbreitung von Epidemien

Zum Abschluss dieses Kapitels wollen wir ein Modell für die Ausbreitung von Epidemien betrachten, das zu Modellierung eine andere Art von Differentialgleichungen verwendet. Als Beispiel für eine Epidemie betrachten wir hier eine Pflanzenkrankheit, nämlich die Kartoffelfäule.

Wir machen zunächst die Modellannahme, dass sich die Masse x der infizierten Pflanzen gemäß dem logistischen Wachstum mit $K = 1$ verhält, also

$$\dot{x}(t) = \lambda x(t)(1 - x(t)). \quad (2.16)$$

Die Kapazität $K = 1$ ist hierbei gerade die normierte Größe des Gesamtbestandes der Pflanzen, die befallen werden können. Natürlich muss man hierbei annehmen, dass alle diese Pflanzen so gleichmäßig stehen, dass der Erreger sich konstant mit Infektionsrate λ ausbreiten kann. Der Faktor $(1 - x(t))$ modelliert in (2.16) die Kapazität des Lebensraumes, während der Faktor $\lambda x(t)$ die Ausbreitung der Infektion bestimmt.

Für Epidemien ist dies aber ein zu einfaches Modell, da wir aus der Analyse des Modells ja bereits wissen, dass die Lösungen gegen 1 konvergieren. Insbesondere würde eine mittels (2.16) modellierte Epidemie immer den gesamten Bestand befallen. In einem realistischeren Modell sollten also weitere aus der Biologie bekannte Tatsachen einbezogen werden. Wir werden hier nun den zeitlichen Verlauf einer Infektion berücksichtigen. Für die Kartoffelfäule ist bekannt, dass sich die Krankheit nach erfolgter Infektion zum Zeitpunkt t^* in zwei Stadien entwickelt:

- Das Latenzstadium $[t^*, t^* + p]$, in dem sich der Erreger nicht ausbreiten kann
- Das Infektionsstadium $[t^* + p, t^* + p + q]$, in dem sich der Erreger verbreiten kann

Nach der Zeit $t^* + p + q$ ist eine weitere Ausbreitung nicht möglich.

Mit $x(t)$ bezeichnen wir weiterhin die Masse der infizierten Pflanzen. Wir wollen nun eine Differentialgleichung für $x(t)$ aufstellen. Wir nehmen an, dass die Kapazität des "Lebensraumes" der Infektion von den Stadien der Krankheit nicht abhängt, so dass der Faktor

$(1 - x(t))$ in (2.16) unverändert bleibt. Der Wachstumsfaktor $\lambda x(t)$ muss aber geändert werden: Die Zunahme der Infektion ist nun proportional zur Größe des Anteils der infizierten Population, die sich zur Zeit t im Infektionsstadium befindet. Diese Größe ist gegeben durch die Menge aller Infektionen, die älter als p sind, also $x(t - p)$, minus der Anzahl der Infektionen, die älter als $p + q$ sind, also $x(t - p - q)$. Wir ersetzen $\lambda x(t)$ also durch $\lambda(x(t - p) - x(t - p - q))$ und erhalten so die Gleichung

$$\dot{x}(t) = \lambda(1 - x(t))(x(t - p) - x(t - p - q)). \quad (2.17)$$

Dies ist jetzt keine gewöhnliche Differentialgleichung im üblichen Sinne mehr, da die rechte Seite nicht nur von $x(t)$ sondern auch von $x(t - p)$ und $x(t - p - q)$ abhängt. Eine solche Gleichung nennt man *Delay-Differentialgleichung*, auf deutsch auch *verzögerte Differentialgleichung*.

Allgemein kann man diese Gleichungen in der Form

$$\dot{x}(t) = f(x(t), x(t - \tau_1), \dots, x(t - \tau_k))$$

für ein $f : (\mathbb{R}^n)^{k+1} \rightarrow \mathbb{R}^n$ schreiben, wobei wir $\tau_k > \tau_{k-1} > \dots > \tau_1$ annehmen. Auch für diese Gleichungen gibt es einen Existenz- und Eindeutigkeitsatz, der dem Satz 1.4 sehr ähnlich ist (man benötigt wieder eine Lipschitz-Bedingung etc.). Ein wesentlicher Unterschied besteht aber bei der Wahl der Anfangsbedingung: Es genügt hier nicht, nur die Zeit t_0 und den Wert $x(t_0)$ festzulegen. Tatsächlich reicht es auch nicht, zusätzlich die Werte $x(t_0 - \tau_i)$ für $i = 1, \dots, k$ festzulegen, denn für jeden Zeitpunkt $t > t_0$ benötigt man zur Berechnung von $\dot{x}(t)$ ja insbesondere die Werte $x(t - \tau_k)$. Da $t - \tau_k$ das gesamte Intervall $[t_0 - \tau_k, t_0]$ durchläuft, müssen wir als Anfangs"wert" also zusätzlich zu $x(t_0) = x_0$ noch eine Funktion $\Psi : [t_0 - \tau_k, t_0] \rightarrow \mathbb{R}^n$ festlegen. Für die Existenz- und Eindeutigkeitsaussage reicht es dabei aus, Ψ als stetige Funktion zu wählen, wobei es nicht nötig ist, dass Ψ in t_0 durch x_0 stetig fortgesetzt wird.

Wir wollen nun das Langzeitverhalten der Lösungen von (2.17) untersuchen um damit zu ermitteln, wie groß der für $t \rightarrow \infty$ befallene Pflanzenbestand bei der durch (2.17) modellierten Epidemie ist und wie dieser Wert von p und q abhängt. Wir wählen dabei die Anfangsfunktion $\Psi \equiv 0$ und einen Anfangswert $x(0) = x_0 \in (0, 1)$; die Infektion gelangt also zum Zeitpunkt $t_0 = 0$ von außen in den Pflanzenbestand.

Unter dieser Annahme sieht man per Induktion über $n = 1, 2, 3, \dots$ aus (2.17), dass auf jedem Intervall $[(n-1)(p+q), n(p+q)]$ die Ungleichungen $\dot{x}(t) \geq 0$ und $x(t) \in [0, 1]$ gelten. Also ist $x(t)$ monoton wachsend und durch 1 beschränkt und konvergiert damit gegen einen Wert $\beta \in (0, 1]$. Mit $g(t) := x(t - p) - x(t - p - q)$ können wir (2.17) als

$$\dot{x}(t) = \lambda g(t)(1 - x(t))$$

schreiben. Dies ist nun wieder eine klassische gewöhnliche Differentialgleichung, für die man (mit einer Technik, die in Lehrbüchern unter dem Namen *Trennung der Variablen* zu finden ist) die explizite Lösung

$$x(t) = 1 - (1 - x_0) \exp\left(-\lambda \int_0^t g(\tau) d\tau\right)$$

berechnen kann. Durch Ableiten nach t prüft man leicht nach, dass dies tatsächlich die Lösung ist. Für $g(t)$ gilt nun wegen $x(\sigma) = \Psi(\sigma) = 0$ für $\sigma < 0$ die Gleichung

$$\begin{aligned} \int_0^t g(\tau) d\tau &= \int_0^t x(\tau - p) d\tau - \int_0^t x(\tau - p - q) d\tau \\ &= \int_0^{t-p} x(\sigma) d\sigma - \int_0^{t-p-q} x(\sigma) d\sigma = \int_{t-p-q}^{t-p} x(\sigma) d\sigma \end{aligned}$$

für alle $t \geq 0$. Wir erhalten somit

$$x(t) = 1 - (1 - x_0) \exp\left(-\lambda \int_{t-p-q}^{t-p} x(\sigma) d\sigma\right),$$

(was man wiederum durch Ableiten nach t überprüfen kann) und damit

$$\beta = \lim_{t \rightarrow \infty} x(t) = 1 - (1 - x_0) \exp\left(-\lambda \lim_{t \rightarrow \infty} \int_{t-p-q}^{t-p} x(\sigma) d\sigma\right).$$

Mit dem Mittelwertsatz der Integralrechnung und wegen $\lim_{t \rightarrow \infty} x(t) = \beta$ folgt

$$\lim_{t \rightarrow \infty} \int_{t-p-q}^{t-p} x(\sigma) d\sigma = \lim_{t \rightarrow \infty} qx(t - p - \theta q) = q\beta.$$

Der Limes β ist also bestimmt durch die Gleichung

$$\beta = 1 - (1 - x_0)e^{-q\lambda\beta}.$$

Leider erlaubt diese Gleichung keine explizite Lösung. In Abbildung 2.12 ist $\beta(q, x_0)$ für $\lambda = 1$ und $x_0 = 0.1, 0.2, \dots, 0.9$ in Abhängigkeit von q dargestellt. Die Graphen wurden numerisch berechnet. Beachte, dass $\beta(0, x_0) = x_0$ gilt.

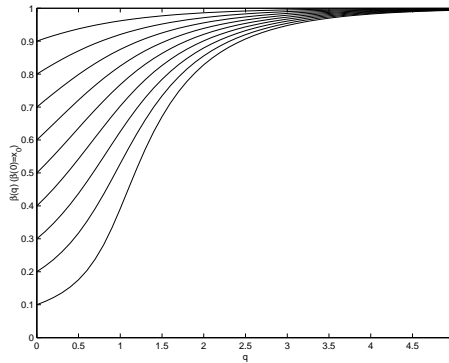


Abbildung 2.12: Die Funktion $\beta(q, x_0)$ für verschiedene x_0 in Abhängigkeit von q

Die Graphen geben also den aus dem Modell (2.17) berechneten Anteil befallener Pflanzen für $t \rightarrow \infty$ in Abhängigkeit vom Anfangsbefall x_0 und der Infektionszeit q an. Für wachsendes q nähert sich die Größe der 1 an, d.h. fast der gesamte Bestand wird befallen. Für

kleinere Infektionszeiten q hingegen wird nur ein Teil des Bestandes befallen. Für q gegen 0 nähert sich dieser Wert dem Anfangsbefall x_0 an. Der Grund für dieses Verhalten liegt darin, dass zur Ausbreitung der Krankheit eine gewisse Anzahl (relativ) frischer Infektionen vorliegen muss. Wenn die Wachstumsrate $\dot{x}(t)$ abnimmt, so fehlt der "Nachschub" an frischen Infektionen, dadurch nimmt $\dot{x}(t)$ weiter ab usw., weswegen die Lösung gegen $\beta < 1$ konvergiert.

Auch dieses Modell ist sicherlich für viele praktische Zwecke zu einfach, weil viele wichtige Einflüsse unberücksichtigt bleiben, z.B. Resistenzen gegen die Krankheit oder die räumliche Verteilung der Pflanzen. Trotzdem kann es zum Verständnis der Abhängigkeiten zwischen Infektionszeiten und Ausbreitungen von Epidemien beitragen und hierbei insbesondere die Komplexität der möglichen Abhängigkeiten illustrieren.

2.5 Literaturhinweise

Eine umfassende Einführung in die mathematische Biologie bietet das (in der ersten und zweiten Auflage einbändige, in der dritten Auflage zweibändige) Buch

J.D. Murray, Mathematical Biology, Springer-Verlag, 2002 (dritte Auflage).

Kapitel 3

Mechanik

Die mathematische Modellierung der klassischen Mechanik geht im Wesentlichen auf die Arbeiten von Isaac Newton¹, Joseph–Louis Lagrange² und William R. Hamilton³ zurück. Newton entwickelte die elementaren Bewegungsgleichungen (und nebenbei die Differentialrechnung), während Lagrange und Hamilton weiterführende mathematische Modellierungs- und Analysemethoden entwickelten, die wir im zweiten Abschnitt dieses Kapitels kennen lernen werden.

3.1 Mechanisch–technischer Ansatz

In diesem ersten Abschnitt wollen wir uns zunächst mit einem “modularen” Ansatz zur Modellierung mechanischer Systeme kennen beschäftigen, der auf der Kombination verschiedener Elemente und der dazu gehörigen DGLs beruht. Für jedes Element werden wir ein grafisches Symbol und die dazugehörige Bewegungsgleichung (die nicht immer eine Differentialgleichung ist) betrachten. Man unterscheidet dabei zwischen verschiedenen Arten von Bewegungen, die wir der Reihe nach einführen wollen. Der Ansatz ist konstruktiv und erlaubt mit sehr wenig mathematischem Aufwand die Modellierung von (im Prinzip) beliebig komplizierten mechanischen Systemen. Wir werden allerdings auch sehen, dass der Ansatz für komplizierte Systeme unpraktikabel wird, was die Einführung mathematisch anspruchsvollerer Techniken im nachfolgenden zweiten Abschnitt rechtfertigt.

3.1.1 Translationale Bewegungselemente

Wir betrachten in diesem Teilabschnitt Bewegungselemente, die sich in eine Richtung auf einer Geraden bewegen können, also 1–dimensionale Bewegungen.

Wir verwenden dabei die folgenden Bezeichnungen:

¹englischer Mathematiker und Physiker, 1642–1727

²französischer Mathematiker, 1736–1813 (geboren als Giuseppe Lodovico Lagrangia in Turin, deshalb manchmal — vor allem in italienischen Büchern — auch als italienischer Mathematiker bezeichnet)

³irischer Mathematiker, 1805–1865

Variable	Bedeutung	Maßeinheit	
y	Ort, Ausdehnung	m	[Meter]
$v = \dot{y}$	Geschwindigkeit	m/s	[Meter pro Sekunde]
$a = \ddot{y}$	Beschleunigung	m/s^2	[Meter pro Sekunde zum Quadrat]
F	Kraft	$N = kg\,m/s^2$	[Newton]

a) Das Trägheitselement (Masse)

Das Trägheitselement besteht aus einer (zeitlich konstanten) Masse m auf die eine Kraft F wirkt und die sich mit einer Geschwindigkeit v bewegt. Das Symbol für das Trägheitselement ist in Abbildung 3.1 dargestellt.

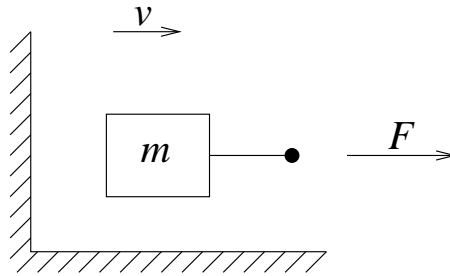


Abbildung 3.1: Symbol für das Trägheitselement

Die Differentialgleichung für das Trägheitselement ist nach dem 2. Newton'schen Gesetz gegeben durch

$$F(t) = ma(t) = m\dot{v}(t) \quad (3.1)$$

Beachte, dass hier die Kraft F und die Geschwindigkeit v in die selbe Richtung zeigen müssen, ansonsten muss F durch $-F$ ersetzt werden; dies ist eine beliebte Quelle für Vorzeichenfehler! Die "Wände" im Symbol symbolisieren das verwendete Koordinatensystem, was wichtig sein kann, wenn mehrere Massen in einem System verbunden werden.

Eine Masse speichert *Energie*: Wenn die Masse in Bewegung ist, so besitzt sie die kinetische Energie

$$E_k(t) = \frac{m}{2}v(t)^2$$

und wenn sie sich in einem Schwerfeld befindet, so besitzt sie potentielle Energie, auf der Erde nahe der Erdoberfläche gerade

$$E_p(t) = mgh(t),$$

wobei $g \approx 9,80665m/s^2$ die Erdbeschleunigung und h die Höhe über der Erdoberfläche bezeichnet.

b) Das Elastizitätselement (Feder)

Das Elastizitätselement wird ganz allgemein als deformierbares Objekt definiert, bei dem die Größe y der Deformation eine Funktion der einwirkenden Kraft F ist. Zwei gebräuchliche Symbole sind in Abbildung 3.2 dargestellt.

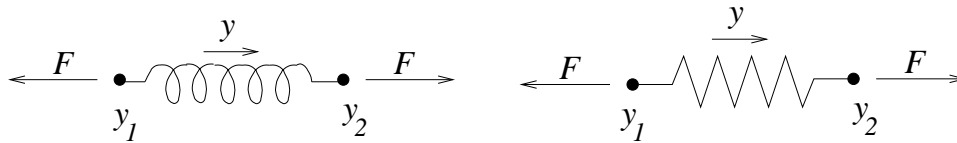


Abbildung 3.2: Symbole für das Elastizitätselement

Beim *linearen Modellansatz* wird zur Beschreibung des Elastizitätselementes das *Hook'sche Gesetz* verwendet. Mit $y = y_2 - y_1$ ist dies durch

$$ky(t) = F(t) \quad (3.2)$$

gegeben, wobei $k > 0$ die *Federkonstante* ist. Per Konvention ist y_2 der Angriffspunkt in positiver Koordinatenrichtung und y_1 der Angriffspunkt in negativer Koordinatenrichtung.

Dieses Modell beschreibt eine reale Feder bei kleinen Auslenkungen i.A. hinreichend gut. Realistischere Ansätze verwenden einen nichtlinearen Zusammenhang zwischen y und F , worauf wir hier aber nicht näher eingehen wollen. Unabhängig von der Modellierung dieses Zusammenhangs sind reine Elastizitätselemente prinzipiell eine Idealisierung, da in der Realität keine Federn ohne Masse und Dämpfung (s.u.) existieren. Beachte, dass bei der obigen Wahl der Punkt $y = 0$ gerade die Feder in Ruhelage bezeichnet, die Ausdehnung kann in diesem mathematischen Modell also positiv (gedehnte Feder) oder negativ (gestauchte Feder) sein.

Auch Federn speichern potentielle Energie, falls man (3.2) annimmt, ist diese durch

$$E_p(t) = \frac{k}{2}y(t)^2$$

gegeben.

c) Das Dämpfungselement (Dämpfer, Reibung)

Die allgemeine Definition ist hier gegeben durch ein mechanisches Element, das keinerlei Energie speichert, sondern die aufgenommene Energie in Wärme umwandelt und abgibt (Dissipator). Das Symbol für das Dämpfungselement ist in Abbildung 3.3 angegeben.

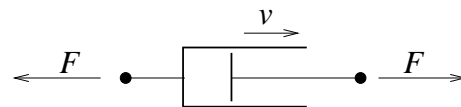


Abbildung 3.3: Symbol für das Dämpfungselement

Wiederum betrachten wir hier nur das lineare Modell, das durch

$$F(t) = cv(t) \quad (3.3)$$

gegeben ist, wobei v die relative Geschwindigkeit zweier (in Abbildung 3.3 durch Kolben und Zylinder symbolisierten) Körper ist, F bezeichnet wirkende Kraft und $c > 0$ eine

Dämpfungskonstante ist: wenn die Kraft F wirkt, so wird die Geschwindigkeit cv erreicht. Die Relativgeschwindigkeit v berechnet sich hierbei als $v = v_+ - v_-$, wobei v_+ die Geschwindigkeit des Endpunktes in positiver und v_- die Geschwindigkeit des Endpunktes in negativer Koordinatenrichtung ist.

Dieses Modell nennt man *viskose Reibung*. Andere Modelle sind z.B. die *trockene Reibung*, bei der die Kraft F bei niedriger Geschwindigkeit größer wird (diese Funktion ist unstetig in $v = 0$) oder die *Strömungsreibung* (z.B. der Luftwiderstand), bei der $F = c|v|v$ ist, also quadratisch von v abhängt. Noch komplizierter ist die Haftreibung, die sich als klassische Funktion zwischen F und v nicht modellieren lässt, sondern nur mit sogenannten *Hysteresemodellen* beschrieben werden kann.

Die von dem Dämpfungselement zur Zeit t absorbierte Leistung ist gerade das Produkt $F(t)v(t)$, die im Intervall $[t_0, t_1]$ absorbierte Energie ergibt sich als Integral über die Leistung, also

$$E_{\text{abs}} = \int_{t_0}^{t_1} F(t)v(t)dt.$$

3.1.2 Einfache translationale Modelle

Die im letzten Abschnitt beschriebenen drei Elemente bilden die Grundbausteine für translationale mechanische Systeme. Der Ansatz, um kompliziertere Systeme beschreiben zu können funktioniert nun wie folgt.

- (1) Modelliere ein translationales mechanisches System als Verbindung von Massen, Federn und Dämpfern.
- (2) Stelle die zugehörigen Bewegungsgleichungen auf.
- (3) Formuliere die Verbindungsgesetze (oder Kontaktkräfte).

Grundlegend für (3) ist Newtons 3. Gesetz *actio = reactio*: In jeder Masse ist die Summe der Kräfte (mit Berücksichtigung ihrer Vorzeichen) gleich Null. Falls zusätzlich eine externe Kraft wirkt, so ist die Summe der (systeminternen) Kräfte gleich der externen Kraft, wobei auch hier die Richtung der Kraft über das Vorzeichen berücksichtigt werden muss.

Wir illustrieren dies an zwei Beispielen.

Beispiel 3.1 Mechanischer Oszillator (oder Schwinger)

Eine Masse ist an einer (realen) Feder an der Decke aufgehängt, auf die Masse wirkt eine Kraft $F(t)$, die nach unten gerichtet ist und die aus der Schwerkraft und einer externen Kraft zusammen gesetzt ist. Da eine reale Feder immer auch Dämpfung bewirkt, modellieren wir sie durch eine Kombination aus Elastizitäts- und Dämpfungselement. Abbildung 3.4 stellt diese Kombination dar.

Mit $y(t)$ und $v(t)$ bezeichnen wir Position und Geschwindigkeit der Masse. Für jedes der drei Elemente kennen wir die Gleichungen, es gilt

$$F_1(t) = m\dot{v}(t), \quad F_2(t) = -cv(t), \quad F_3(t) = -ky(t).$$

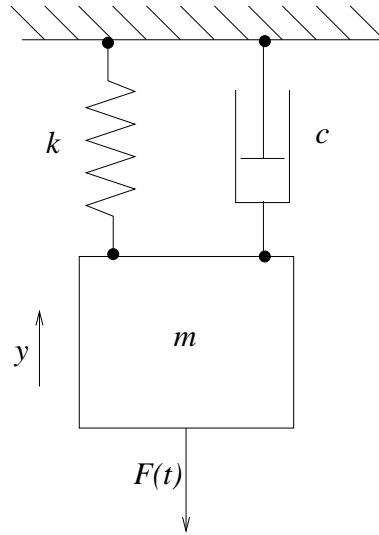


Abbildung 3.4: Mechanischer Oszillator

F_1 beschreibt die Kraft im Masselement, F_2 die im Dämpfungselement und F_3 die im Elastizitätselement. Die umgekehrten Vorzeichen ergeben sich aus der gemäß $y_k = y_2 - y_1$ und $v_k = v_+ - v_-$ berechneten Ausdehnung und Relativgeschwindigkeit, wobei hier $y_1 = y$, $v_+ = 0$ und $v_- = v$ gelten und wir $y_2 = 0$ setzen.

Wir müssen nun die Richtungen der Kräfte F_1 , F_2 , F_3 in den Elementen bestimmen. Da y durch den Pfeil angedeutet nach oben zunimmt, zeigt auch F_1 nach oben. Gemäß den Elementarmodellen zeigen F_2 und F_3 nach unten, da wir uns am unteren Ende der Elemente befinden. Ebenfalls zeigt die Kraft F gemäß dem eingezeichneten Pfeil nach unten; alle Kräfte sind also F_1 entgegengerichtet. Nach Newtons drittem Gesetz muss die Summe aller Kräfte (versehen mit Vorzeichen gemäß ihrer Richtungen) daher gleich $-F$ sein, also $F_1 - F_2 - F_3 = -F$. Daraus erhalten wir die Gesamtgleichung des Systems:

$$-F(t) = F_1(t) - F_2(t) - F_3(t) = m\dot{v} + cv(t) + ky(t) = m\ddot{y} + c\dot{y}(t) + ky(t).$$

Dies ist noch keine Differentialgleichung in der Form (1.1), denn hier treten höhere Ableitungen der unbekanntten Funktion $y(t)$ auf. Man nennt diese Gleichungen *gewöhnliche Differentialgleichungen höherer Ordnung*, hier haben wir eine DGL zweiter Ordnung. In der Mechanik lässt man die Gleichungen meist in der obigen Form stehen, wir wollen sie hier in die Form (1.1) bringen, um sie in unseren abstrakten Rahmen einzupassen. Man kann jede DGL höherer Ordnung formal in eine Gleichung erster Ordnung der Form (1.1) bringen, indem man Hilfsgrößen für die Ableitungen einführt. Dazu definiert man $x_1(t) := y(t)$ und $x_2(t) := \dot{y}(t)$ (und so weiter, falls nötig) und erhält so das System

$$\begin{aligned}\dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -\frac{c}{m}x_2(t) - \frac{k}{m}x_1(t) - \frac{1}{m}F(t)\end{aligned}$$

in Form (1.1). In unserem Fall haben wir bereits die Symbole $y(t) = x_1(t)$ und $v(t) = x_2(t)$,

so dass die Schreibweise

$$\begin{aligned}\dot{y}(t) &= v(t) \\ \dot{v}(t) &= -\frac{c}{m}v(t) - \frac{k}{m}y(t) - \frac{1}{m}F(t)\end{aligned}$$

mit der zweidimensionalen unbekanntem Funktion $(y(t), v(t))^T$ aussagekräftiger ist.

Der Nullpunkt $y = v = 0$ ist gerade das einzige Gleichgewicht dieser Gleichung und entspricht dem Gleichgewichtspunkt der Masse wenn keine Kraft wirkt, also $F(t) \equiv 0$ ist. Dies ist i.A. keine gute Wahl für unser Modell, da (unabhängig von weiteren äußeren Kräften) immer die Schwerkraft $F_G = mg$ auf die Masse wirkt. Man kann die Schwerkraft aber aus dem Modell eliminieren, wenn man den Nullpunkt für y anders wählt. Wir zerlegen dazu $F(t) = F_G + F_e(t)$ in die Schwerkraft und eine (gegebenfalls wirkende) weitere externe Kraft $F_e(t)$. Sicherlich besitzt die Gleichung ein Gleichgewicht $(y_G, 0)^T$ für $F_e(t) \equiv 0$, also wenn nur die Schwerkraft F_G wirkt. Für dieses gilt

$$0 = -\frac{k}{m}y_G - \frac{1}{m}F_G \Leftrightarrow y_G = -\frac{F_G}{k} = -\frac{mg}{k}.$$

Mit der Koordinatentransformation $\tilde{y} = y - y_G$ wird dies der neue Nullpunkt, und in den neuen Koordinaten gilt nun

$$\dot{\tilde{y}}(t) = \dot{y}(t) = v(t)$$

und

$$\begin{aligned}\dot{v}(t) &= -\frac{c}{m}v(t) - \frac{k}{m}y(t) + \frac{1}{m}F(t) \\ &= -\frac{c}{m}v(t) - \frac{k}{m}\tilde{y}(t) - \underbrace{\frac{k}{m}y_G - \frac{1}{m}F_G}_{=0} - \frac{1}{m}F_e(t) \\ &= -\frac{c}{m}v(t) - \frac{k}{m}\tilde{y}(t) - \frac{1}{m}F_e(t)\end{aligned}$$

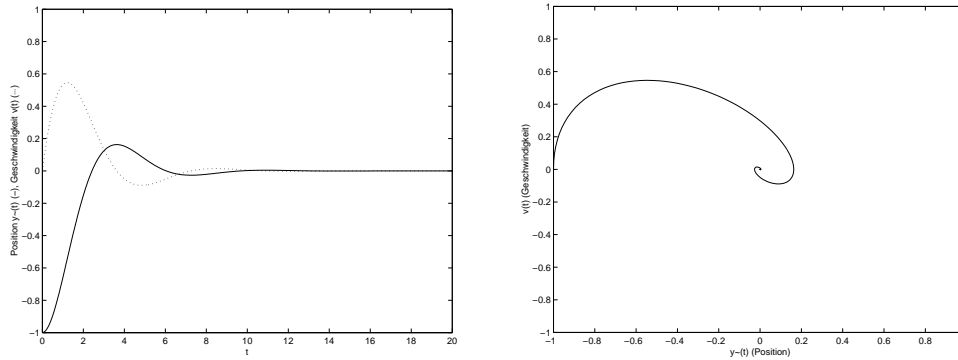
Also ergibt sich die Gleichung

$$\begin{aligned}\dot{\tilde{y}}(t) &= v(t) \\ \dot{v}(t) &= -\frac{c}{m}v(t) - \frac{k}{m}\tilde{y}(t) - \frac{1}{m}F_e(t)\end{aligned}\tag{3.4}$$

in der die Schwerkraft nicht mehr auftaucht. Wenn man den Nullpunkt für y also von vornherein in das Gleichgewicht unter Schwerkraft legt, so braucht man F_G gar nicht berücksichtigen oder mit anderen Worten: Wenn wir F_G nicht ins Modell aufnehmen, so beschreibt der Nullpunkt "automatisch" das Gleichgewicht unter Schwerkraft.

Abbildung 3.5 zeigt eine Lösung dieser Gleichung mit $F_e(t) \equiv 0$ und Anfangswert $(y(0), v(0)) = (-1, 0)$. Die Masse wird also nach unten in den Punkt $y = -1$ "gezogen" und zum Zeitpunkt $t = 0$ losgelassen.

Die Grafik zeigt das zu erwartende Verhalten: Die Masse schwingt etwas und pendelt sich, bedingt durch die Reibung, auf das Gleichgewicht $(0, 0)$ ein. \square

Abbildung 3.5: Lösungen von (3.4) mit $m = k = c = 1$

Beispiel 3.2 Ein einfaches Fahrzeug-Federungsmodell

Abbildung 3.6 zeigt ein einfaches Modell für eine Fahrzeugfederung. Die Modellannahmen hier sind:

- Es werden nur vertikale Bewegungen berücksichtigt (keine Drehungen)
- Es wird nur ein Rad modelliert
- Die Karosserie wird als Masse m_1 mit Position y_1 modelliert, die Federung mittels Elastizitäts- und Dämpfungselement k_1, c_1
- Rad und Achse wird als Masse m_2 mit Position y_2 modelliert, der Reifen mittels Elastizitäts- und Dämpfungselement k_2, c_2
- Die Straßenunebenheiten werden durch eine “Straßenhöhenfunktion” $u(t)$ modelliert

Aus den Bewegungsgleichungen erhalten wir die Gleichungen für die Einzelkräfte

$$m_i \dot{v}_i^m(t) = F_i^m(t), \quad c_i v_i^c(t) = F_i^c(t), \quad k_i y_i^k(t) = F_i^k(t)$$

für $i = 1, 2$, wobei die Indizes angeben, für welches Element die Kräfte bzw. Positionen gelten. Hierbei gelten die Beziehungen

$$\begin{aligned} v_1^m(t) &= \dot{y}_1(t), & v_1^c(t) &= \dot{y}_1(t) - \dot{y}_2(t), & y_1^k(t) &= y_1(t) - y_2(t), \\ v_2^m(t) &= \dot{y}_2(t), & v_2^c(t) &= \dot{y}_2(t) - \dot{u}(t), & y_2^k(t) &= y_2(t) - u(t). \end{aligned}$$

Um die obigen Gleichungen zusammensetzen müssen wir nun die Einzelkräfte an den Massen bestimmen. Hierbei muss man wieder auf die Richtungen aufpassen, die von der betrachteten Masse abhängen. In m_1 ist die Kraft F_1^m (entsprechend der Richtung von y_1) nach oben gerichtet, ebenso zeigen F_1^c und F_1^k nach oben, da dies das obere Ende der Elemente ist. In m_1 ergibt sich so

$$F_1^m + F_1^c + F_1^k = 0.$$

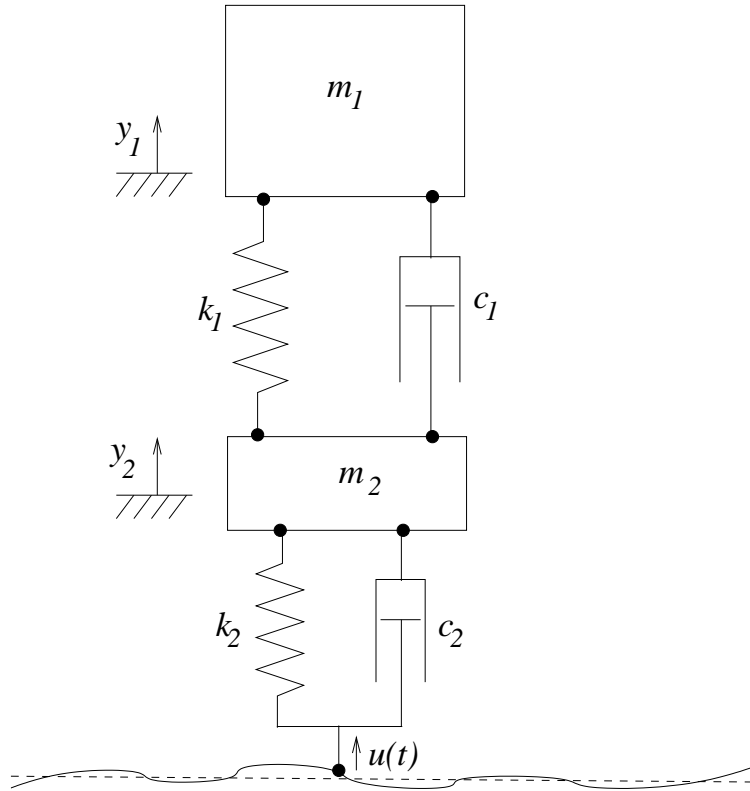


Abbildung 3.6: Federungsmodell

In m_2 zeigen F_1^k und F_1^c nach unten und alle anderen Kräfte nach oben, also erhalten wir in m_2 die Gleichung

$$F_2^m - F_1^k - F_1^c + F_2^k + F_2^c = 0.$$

Zusammen ergibt dies die Gleichungen

$$\begin{aligned} 0 &= F_1^m(t) + F_1^c(t) + F_1^k(t) \\ &= m_1 \dot{v}_1^m(t) + c_1 v_1^c(t) + k_1 y_1^k(t) \\ &= m_1 \ddot{y}_1(t) + c_1 (\dot{y}_1(t) - \dot{y}_2(t)) + k_1 (y_1(t) - y_2(t)) \end{aligned}$$

und

$$\begin{aligned} 0 &= F_2^m(t) - F_1^c(t) - F_1^k(t) + F_2^c(t) + F_2^k(t) \\ &= m_2 \dot{v}_2^m(t) - c_1 v_1^c(t) - k_1 y_1^k(t) + c_2 v_2^c(t) + k_2 y_2^k(t) \\ &= m_2 \ddot{y}_2(t) - c_1 (\dot{y}_1(t) - \dot{y}_2(t)) - k_1 (y_1(t) - y_2(t)) + c_2 (\dot{y}_2(t) - \dot{u}(t)) + k_2 (y_2(t) - u(t)) \end{aligned}$$

Analog zum vorherigen Beispiel kann diese Gleichung nun in die Form (1.1) mit der vierdimensionalen unbekanntem Funktion $(y_1(t), v_1(t), y_2(t), v_2(t))^T$ umgeformt werden. \square

3.1.3 Rotationale Bewegungselemente

Bisher können wir nur Bewegungen in gerader Richtung beschreiben. Nun lernen wir drei analoge Elemente kennen, die Drehbewegungen darstellen können.

Wir verwenden dabei die folgenden Bezeichnungen:

Variable	Bedeutung	Maßeinheit	
θ	Winkel	rad	[Radiant]
$\omega = \dot{\theta}$	Winkelgeschwindigkeit	rad/s	[Radiant pro Sekunde]
$\alpha = \ddot{\theta}$	Winkelbeschleunigung	rad/s^2	[Radiant pro Sek. zum Quadrat]
τ	Drehmoment	$Nm = kg\,m^2/s^2$	[Newtonmeter]

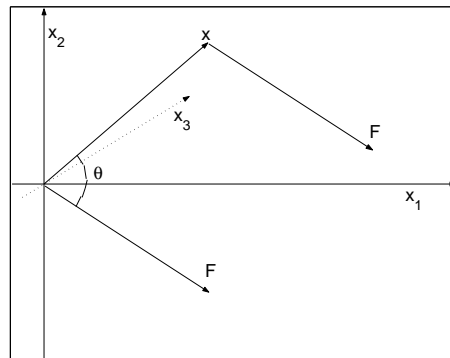


Abbildung 3.7: Schematische Darstellung des Drehmoments

Das Drehmoment beschreibt dabei die auf einen rotierenden Körper wirkenden Kraft: Sei $F = (F_1, F_2, 0)$ eine gerichtete Kraft die im Punkt $x = (x_1, x_2, 0)$ an einem um die x_3 -Achse rotierenden Körper angreift, vgl. Abbildung 3.7: Man kann sich den Vektor x als Hebel an dem (nicht dargestellten) Körper vorstellen. Dann ist das resultierende Drehmoment gegeben durch

$$\tau = x_1 F_2 - x_2 F_1 = \|x\| \|F\| \sin \theta, \quad (3.5)$$

wobei θ den Winkel zwischen x und F beschreibt.

Zu beachten ist, dass wir die Kraft F nun als Kraftvektor im Koordinatensystem schreiben, die Richtungsinformation ist — im Gegensatz zu den translationalen Modellen — hier also bereits in F enthalten, so dass wir beim Bestimmen der Kontaktkräfte keine Richtungen mehr berücksichtigen müssen.

a) Das Trägheitselement (Trägheitsmoment)

Das Trägheitselement der Rotation besteht aus einer Masse, die um eine Achse rotiert. Die Gleichung lautet

$$\tau(t) = J\alpha(t) = J\dot{\omega}(t). \quad (3.6)$$

Das Trägheitsmoment J des Elementes wird durch die Masse des Elementes und ihre Verteilung bzgl. der Rotationsachse bestimmt.

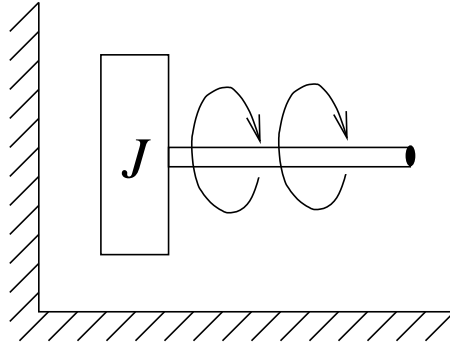


Abbildung 3.8: Symbol für das Rotations-Trägheitselement

Für einen rotierenden Körper $B \subset \mathbb{R}^3$ mit Masse m und Massendichte $\rho : B \rightarrow \mathbb{R}_0^+$ gilt allgemein

$$J = \int_B r^2(x) \rho(x) dx.$$

In Spezialfällen kann man explizite Formeln angeben:

Ein rotierender Massepunkt mit Masse m und Abstand r von der Rotationsachse besitzt gerade das Trägheitsmoment

$$J = mr^2.$$

Für eine Menge von N Massepunkten x^i mit Koordinaten (x_1^i, x_2^i, x_3^i) , die um die x_3 -Achse rotieren, gilt

$$J = \sum_{i=1}^N m_i r_i^2 = \sum_{i=1}^N m_i ((x_1^i)^2 + (x_2^i)^2).$$

Für einen dünnen Stab mit Länge l und einer eindimensionalen Massenverteilung $\rho : [0, l] \rightarrow \mathbb{R}_0^+$ gilt bei Rotation um den Endpunkt 0

$$J = \int_0^l x^2 \rho(x) dx$$

und bei Rotation um den Mittelpunkt $l/2$

$$J = \int_0^l (x - l/2)^2 \rho(x) dx.$$

Falls die Masse m im Stab gleichverteilt ist, gilt $\rho(x) = m/l$ und damit

$$J = \int_0^l x^2 m/l dx = \frac{mL^2}{3}$$

bei Rotation um den Endpunkt und

$$J = \int_0^l (x - l/2)^2 m/l dx = \int_{-l/2}^{l/2} x^2 m/L dx = \frac{ml^2}{12}$$

bei Rotation um den Mittelpunkt.

Allgemein gilt der *Parallelachsensatz*, auch *Steiner'scher Satz* genannt.

Satz 3.3 Gegeben ein Körper $B \subset \mathbb{R}^3$ der Masse m mit Massenverteilung $\rho : B \rightarrow \mathbb{R}_0^+$ und Schwerpunkt

$$\bar{x} = \frac{1}{m} \int_B x \rho(x) dx \in \mathbb{R}^3.$$

Sei J das Trägheitsmoment des Körpers um eine beliebige Achse, J' das Trägheitsmoment des Körpers um die durch den Schwerpunkt verlaufende parallele Achse. Dann gilt

$$J = J' + mR^2,$$

wobei R den Abstand der beiden Achsen bezeichnet.

b) Das Elastizitätselement (Torsionsfeder)

Das Elastizitätselement und das nachfolgende Dämpfungselement sind völlig analog zu ihren translationalen Gegenstücken. Wir schon bei diesen betrachten wir hier nur die linearen Bewegungsmodelle. Für das rotationale Elastizitätselement its die entsprechende Gleichung durch

$$k\theta(t) = \tau(t) \tag{3.7}$$

gegeben.

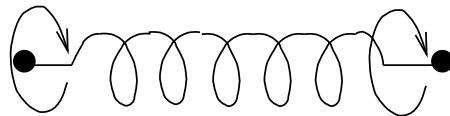


Abbildung 3.9: Symbol für das Rotations-Elastizitätselement

c) Das Dämpfungselement (Rotationsdämpfer)

Wiederum analog zum translationalen Fall gibt es das rotationale Dämpfungselement, dessen Gleichung

$$c\omega(t) = \tau(t) \tag{3.8}$$

lautet.

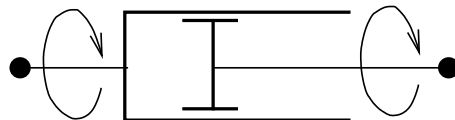


Abbildung 3.10: Symbol für das Rotations-Dämpfungselement

3.1.4 Das Pendel

Wir wollen die besprochenen Elemente nun zu einem Modell eines Pendels zusammensetzen. Wir machen zuerst die folgenden vereinfachenden Modellannahmen

- Das Pendel ist eine Punktmasse m , die an einem masselosen Stab der Länge l befestigt ist
- Es gibt keine Reibung

Das Modell ist in Abbildung 3.11 schematisch dargestellt.

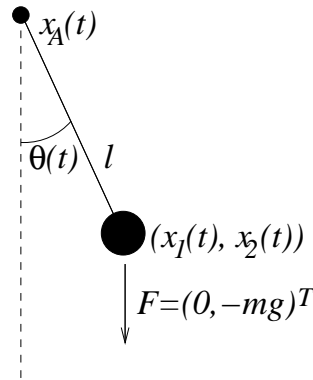


Abbildung 3.11: Pendelmodell

Es sei $x(t) = (x_1(t), x_2(t))^T$ der Endpunkt des Pendels. Der Aufhängepunkt bzw. die Position der Drehachse $x_A(t)$ sei zunächst konstant gleich 0. Wie üblich im Koordinatensystem nehmen x_1 und x_2 nach rechts bzw. oben zu. Der Punkt $x(t)$ lässt sich mittels der Länge l und des Winkels $\theta(t)$ als

$$x(t) = (l \sin \theta(t), -l \cos \theta(t))^T$$

schreiben.

Der in x angreifende Kraftvektor F ist durch die Erdbeschleunigung gegeben als $F = (0, -mg)^T$ (er zeigt nach unten, deswegen '-'). Gemäß (3.5) gilt für das erzeugte Drehmoment also

$$\tau_F(t) = x_1(t) \cdot (-mg) - x_2(t) \cdot 0 = -mgx_1(t) = -mgl \sin \theta(t).$$

Andererseits gilt für das Trägheitselement die Gleichung

$$\tau_J(t) = J\ddot{\theta}(t) = ml^2\ddot{\theta}(t),$$

da wir das Pendel ja als Punktmasse modelliert haben. Gleichsetzen des externen Drehmoments τ_F mit τ_J ergibt

$$\tau_J(t) = \tau_F(t),$$

also

$$ml^2\ddot{\theta}(t) = -mgl \sin \theta(t).$$

Wiederum erhalten wir eine DGL zweiter Ordnung, die wir über die Gleichung $\omega(t) = \dot{\theta}(t)$ als

$$\begin{aligned}\dot{\theta}(t) &= \omega(t) \\ \dot{\omega}(t) &= -\frac{g}{l} \sin \theta(t)\end{aligned}\tag{3.9}$$

in der Form (1.1) schreiben können. Abbildung 3.12 zeigt einige Lösungen dieser Gleichung.

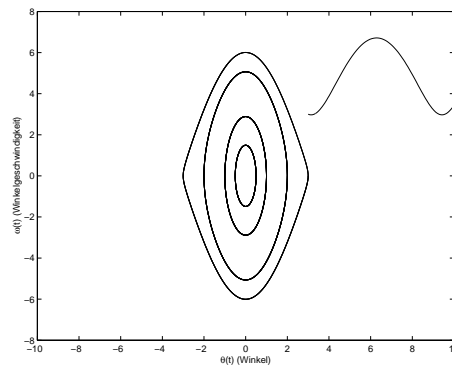


Abbildung 3.12: Lösungen des Pendelmodells (3.9) mit $l = 1$

Die periodischen Lösungen gehören hierbei zu Anfangswerten, für die das Pendel hin- und her schwingt. Da keine Reibung modelliert wurde, wird das Pendel nicht gebremst und die Pendel schwingt für alle Zeiten mit der gleichen Bewegung, daher die Periodizität. Die Lösung, die rechts aus dem Bild hinausläuft wurde mit größerer Anfangsgeschwindigkeit gestartet. Hier überschlägt sich das Pendel, und zwar — da keine Reibung vorhanden ist — nicht nur einmal sondern immer wieder. Beachte, dass die Winkel θ und $\theta + k2\pi$ für alle $k \in \mathbb{Z}$ die gleiche Pendelposition bedeuten, aber in in unserem Modell unterschieden werden. Die Gleichung besitzt übrigens genau die Gleichgewichte $(\theta_k^*, \omega_k^*) = (k\pi, 0)$ für $k \in \mathbb{Z}$. Für gerades k ist dies gerade das herunterhängende Pendel, für ungerades k ist dies das aufrecht stehende Pendel. Die aufrechten Gleichgewichte sind exponentiell instabil (aber nicht antistabil), die herabhängenden sind weder exponentiell stabil noch instabil, denn die Realteile der Eigenwerte der Linearisierung sind gleich 0.

Wir wollen unser Modell nun realistischer machen. Zunächst fügen wir Reibung hinzu, Wir machen die Annahme:

- Es gibt viskose (d.h. lineare) Reibung mit Koeffizient c .

Wenn wir ein Dämpfungselement an der Achse hinzunehmen, so erhalten wir für das zusätzliche Drehmoment $\tau_c(t)$ die Gleichung

$$\tau_c(t) = c\omega_c(t) = c\dot{\theta}(t).$$

Wie bei den Translationselementen müssen wir nun die Drehmomente addieren und gleich der externen Kraft setzen. Da $\omega_c = \omega$ ist, erhalten wir also

$$\tau_J(t) + \tau_c(t) = \tau_F(t) \quad (3.10)$$

und damit die neue Gleichung

$$ml^2\ddot{\theta}(t) = -c\dot{\theta}(t) - mgl \sin \theta(t).$$

Nun wollen wir noch die Punktmassenannahme verallgemeinern. Wir nehmen nun an:

- Das Pendel ist ein starrer Körper mit Masseverteilung $\rho(x_1, x_2)$ und Trägheitsmoment J .

Die Frage ist jetzt: Wie berechnen wir das durch F erzeugte Drehmoment $\tau_F(t)$? Wir leiten die Lösung heuristisch her. In jedem Rechteck der Form $[x_1(t), x_1(t) + \Delta x_1] \times [x_2(t), x_2(t) + \Delta x_2]$ mit Masse Δm erzeugt F das Drehmoment

$$\Delta\tau_F(t) = -x_1(t)\Delta mg \approx -x_1(t)\rho(x(t))\Delta x_1\Delta x_2g.$$

Aufsummieren über alle Rechtecke und Grenzübergang $\Delta x_i \rightarrow 0$ liefert dann

$$\tau_F(t) = - \int_{B(t)} x_1\rho(x)gdx = -mg\bar{x}_1(t),$$

wobei $\bar{x}(t) = (\bar{x}_1(t), \bar{x}_2(t))^T$ die Position des Schwerpunktes des Körpers zur Zeit t bezeichnet.

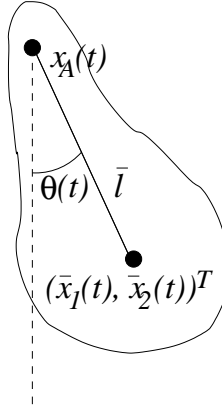


Abbildung 3.13: Pendelmodell mit starrem Körper

Wenn wir mit $\bar{l} = \|\bar{x}(t)\|$ den Abstand des Schwerpunktes zur Rotationsachse (die zunächst weiterhin im Nullpunkt liegt) bezeichnen, und mit $\theta(t)$ den Winkel von $\bar{x}(t)$ mit der x_2 -Achse bezeichnen (siehe Abbildung 3.13), so gilt

$$\bar{x}(t) = (\bar{l} \sin \theta(t), -\bar{l} \cos \theta(t)).$$

Aus der schon bekannten Gleichung (3.10)

$$\tau_J(t) + \tau_c(t) - \tau_F(t) = 0$$

erhalten wir damit

$$J\ddot{\theta}(t) = -c\dot{\theta}(t) - mg\bar{l} \sin \theta(t)$$

bzw. in der Form (1.1) die Gleichung

$$\begin{aligned} \dot{\theta}(t) &= \omega(t) \\ \dot{\omega}(t) &= -\frac{c}{J}\omega(t) - \frac{mg\bar{l}}{J} \sin \theta(t) \end{aligned} \quad (3.11)$$

Abbildung 3.14 zeigt die Lösungen dieser Gleichung mit den gleichen Anfangswerten wie in Abbildung 3.12.

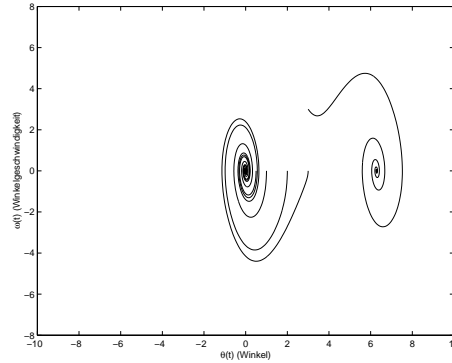


Abbildung 3.14: Lösungen des Pendelmodells (3.11) mit $c/J = 1$ und $m\bar{l}/J = 1$

Hier streben alle Lösungen dem Gleichgewicht $(0, 0)$, bzw. nach einmaligem Überschlagen dem Gleichgewicht $(2\pi, 0)$ zu. Tatsächlich kann man zeigen, dass die Gleichgewichte der Form $(2k\pi, 0)$ mit $k \in \mathbb{Z}$ (herabhängendes Pendel) nun lokal exponentiell stabil sind. Die Gleichgewichte der Form $((2k+1)\pi, 0)$ mit $k \in \mathbb{Z}$ (aufrechtes Pendel) bleiben exponentiell instabil. Beachte, dass sich (3.9) und (3.11) in diesen Simulationen wegen der Wahl der Parameter nur durch den Reibungsterm unterscheiden.

Bemerkung 3.4 Das oben heuristisch hergeleitete Prinzip gilt allgemein: Wenn Kräfte auf einen starren Körper wirken, so bewegt sich dessen Schwerpunkt genau so, wie sich die im Schwerpunkt konzentrierte Gesamtmasse unter Einfluss der Summe der Kräfte bewegen würde (ohne Beweis). \square

Wir wollen nun untersuchen, wie sich das Pendel bei Bewegung des Aufhängepunktes x_A verhält. Wir betrachten dabei horizontale und vertikale Bewegung. Wenn der Aufhängepunkt $x_A(t)$ horizontal und vertikal bewegt wird, so gilt für die Kraft $F_A = (F_{A,1}, F_{A,2})^T$ nach (3.1)

$$F_{A,1}(t) = m\ddot{x}_{A,1}(t), \quad F_{A,2}(t) = m\ddot{x}_{A,2}(t)$$

Diese Kraft greift im Punkt x_A an und erzeugt gemäß Bemerkung 3.4 und Gleichung (3.5) (beachte, dass der Vektor x in (3.5) hier gerade gleich $\bar{x} - x_A$ ist) nun das Drehmoment

$$\tau_A(t) = (\bar{x}_1(t) - x_{A,1}(t))F_{A,2}(t) - (\bar{x}_2(t) - x_{A,2}(t))F_{A,1}(t).$$

Für das Pendel gilt nun als Erweiterung der Gleichung (3.10)

$$\tau_J(t) + \tau_c(t) + \tau_A(t) = \tau_F(t).$$

Wegen

$$\bar{x}(t) - x_A(t) = (\bar{l} \sin \theta(t), -\bar{l} \cos \theta(t))^T$$

erhalten wir

$$\tau_A(t) = F_{A,2}(t)(\bar{l} \sin \theta(t)) - F_{A,1}(t)(-\bar{l} \cos \theta(t)) = m(\ddot{x}_{A,2}(t)\bar{l} \sin \theta(t) + \ddot{x}_{A,1}(t)\bar{l} \cos \theta(t))$$

und damit

$$J\ddot{\theta}(t) = -c\dot{\theta}(t) - mg\bar{l} \sin \theta(t) - m\ddot{x}_{A,2}(t)\bar{l} \sin \theta(t) - m\ddot{x}_{A,1}(t)\bar{l} \cos \theta(t)$$

bzw. wiederum in der Form (1.1) die Gleichung

$$\begin{aligned} \dot{\theta}(t) &= \omega(t) \\ \dot{\omega}(t) &= -\frac{c}{J}\omega(t) - \frac{m\bar{l}}{J}\left(g \sin \theta(t) + \ddot{x}_{A,2}(t) \sin \theta(t) + \ddot{x}_{A,1}(t) \cos \theta(t)\right) \end{aligned} \quad (3.12)$$

Wir wollen die Lösungen für eine spezielle Wahl der Beschleunigungen \ddot{x}_A veranschaulichen. Nehmen wir an, dass nur vertikale Bewegungen vorliegen, wobei der Aufhängepunkt mit konstanter Frequenz Ω und Amplitude a cosinusförmig auf- und abschwingt, also $x_{A,2}(t) = a \cos \Omega t$ gilt. Dann folgt

$$\ddot{x}_{A,2}(t) = -a\Omega^2 \cos t,$$

also

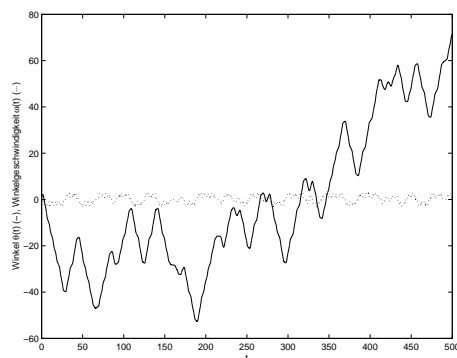
$$\begin{aligned} \dot{\theta}(t) &= \omega(t) \\ \dot{\omega}(t) &= -\frac{c}{J}\omega(t) - \frac{m\bar{l}}{J}\left(g - a\Omega^2 \cos \Omega t\right) \sin \theta(t) \end{aligned} \quad (3.13)$$

Mit den Parametern

$$\Omega = 1.57, \quad \frac{c}{J} = 0.2, \quad \frac{m\bar{l}g}{J} = 1 \quad \text{und} \quad \frac{m\bar{l}a\Omega^2}{J} = 1.42$$

erhält man das in Abbildung 3.15 dargestellte Verhalten.

Das Ganze sieht recht "chaotisch" aus und tatsächlich ist dies ein Beispiel für eine Differentialgleichung mit sogenanntem *chaotischen* Verhalten. Die Lösungen der Gleichung zeigen einen quasi zufälligen Verlauf: es ist nicht vorhersagbar, wann sich das Pendel überschlägt, die Richtung ändert etc. Trotzdem lassen sich in diesem chaotischen Verhalten Gesetzmäßigkeiten erkennen. Hierzu muss man zunächst die vorhandenen Periodizitäten berücksichtigen: Man kann in der Gleichung alle Winkel θ_1, θ_2 mit $\theta_1 = \theta_2 + 2k\pi$ für ein $k \in \mathbb{Z}$ identifizieren, da diese Werte die gleichen Positionen darstellen. Mittels $\tilde{\theta}(t) = \theta(t) - 2k(t)\pi$ (für das richtige $k(t)$) kann man die θ -Komponente der Lösung in das

Abbildung 3.15: Lösungen des Pendelmodells (3.13) in Abhängigkeit von t

Intervall $[-\pi, \pi]$ “projizieren”. Zusätzlich kann man die Periodizität der Beschleunigung berücksichtigen: Für alle $t \in \mathbb{R}$ gilt $\cos \Omega t = \cos \Omega(t + kT)$ für $T = 2\pi/\Omega$ und alle $k \in \mathbb{Z}$. Aufgrund dieser Beobachtung ist es sinnvoll, die Lösungen jeweils nach einer Periode der Beschleunigung darzustellen. Es zeigt sich, dass für jedes $t_0 \in \mathbb{R}$ eine (recht komplizierte) Menge A_{t_0} existiert, so dass jede “periodisch ausgewertete” Lösung, also jede Folge der Form

$$x(t_0 + kT; t_0, x_0) \text{ für } k = 0, 1, 2, \dots$$

gegen diese Menge konvergiert. Diese Menge A_{t_0} heißt *Attraktor*. Durch Darstellung der Punkte $x(t_0 + kT; t_0, x_0)$ für $k = k_0, \dots, k_1$ mit hinreichend großen $k_1 \gg k_0 \gg 0$ kann man einen Eindruck von dieser Menge gewinnen⁴. Abbildung 3.16 zeigt die Punkte $x(t_0 + kT; t_0, x_0)$ für $x_0 = (1, 1)^T$, $t_0 = T/2$ und $k = 101, \dots, 10000$.

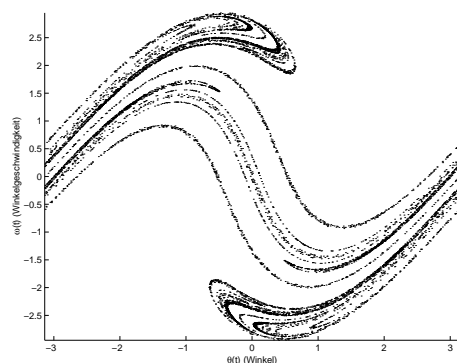


Abbildung 3.16: Approximation des Attraktors des Pendelmodells (3.13)

Diese numerischen Ergebnisse lassen sich durch reale Experimente bestätigen. Im Buch *R.W. Leven, B. Koch und B. Pompe, “Chaos in dissipativen Systemen”, Akademie Verlag 1989 (1. Auflage) und 1994 (2. Auflage)* finden sich in Kapitel 1 der Versuchsaufbau und experimentelle Resultate.

⁴Es gibt auch spezielle Algorithmen zur genaueren Berechnung von Attraktoren, diese werden in der Vorlesung “Numerik dynamischer Systeme” im kommenden Wintersemester behandelt.

3.2 Lagrange–Gleichungen und Hamilton–Formalismus

Die im vorherigen Abschnitt vorgestellte Methode hat den Nachteil, dass das Zusammen­setzen der elementaren Gleichungen für große Systeme sehr kompliziert wird. Man muss für große Systeme und die Verbindungsgesetze bzw. Kontaktkräfte vieler Einzelgleichungen berücksichtigen, was zu sehr großen Gleichungssystemen führt, die dann nur schwer aufzulösen sind.

Die Alternative, die wir in diesem Abschnitt (aus Zeitgründen nur skizziert) vorstellen werden, ist die *Energie­basierte* Methode mit Hilfe der *Lagrange­Gleichungen*.

3.2.1 Lagrange–Gleichungen

Die Idee der Lagrange–Gleichungen basiert auf der Betrachtung der Energie des Systems. Wir beschränken uns hierbei auf ein System s von N Massepunkten mit Positionen $r_i = (x_i, y_i, z_i)^T$ und Massen m_i , $i = 1, \dots, N$, dessen kinetische Energie gerade durch

$$E_{kin} = \sum_{i=1}^N \frac{m_i}{2} \|v_i\|^2$$

gegeben ist.

Zur Modellierung der Tatsache, dass sich ein mechanisches System — bedingt durch die mechanische Struktur — nur auf vorgegebenen Bahnen bewegen kann, verwenden wir Nebenbedingungen der Form

$$c_j(r_1, \dots, r_n, t) = 0, \quad \text{für } j = 1, \dots, J,$$

wobei die $r_i = (x_i, y_i, z_i)^T \in \mathbb{R}^3$ die Positionen der Massepunkte beschreiben.

Beispiel: Wir betrachten ein im Nullpunkt aufgehängtes und in der xy -Ebene schwingendes starres Pendel mit Punktmasse m im Punkt $r(t) = (x(t), y(t), z(t))^T$ und Länge ρ . Die möglichen Positionen von $r(t)$ werden dann genau durch die Gleichungen

$$c_1(r) = \|r\|^2 - \rho \quad \text{und} \quad c_2(r) = z$$

beschrieben.

Wir nehmen nun an, dass die durch

$$M = \{(r_1, \dots, r_N)^T \mid c_j(r_1, \dots, r_n, t) = 0, \quad \text{für } j = 1, \dots, J\}$$

implizit definierte *Mannigfaltigkeit der verträglichen Konfigurationen* durch Koordinaten $q = (q_1, \dots, q_l) \in Q$ mit einer offenen Menge $Q \subset \mathbb{R}^l$ parametrisieren lässt, d.h. dass stetig differenzierbare Funktionen $r_i(q, t)$ existieren mit

$$M = \{(r_1(q, t), \dots, r_N(q, t))^T \mid q \in Q\}.$$

Wir nehmen weiterhin an, dass die partiellen Ableitungen

$$\frac{\partial r}{\partial q_i}(q, t) \in \mathbb{R}^{3N}$$

für $i = 1, \dots, l$ linear unabhängig sind. Die Größen q_1, \dots, q_l heißen *verallgemeinerte Koordinaten*.

Beispiel: Für das Pendel gilt

$$r(q) = \begin{pmatrix} \rho \sin q \\ -\rho \cos q \\ 0 \end{pmatrix}$$

mit $q = q_1 \in Q = (-\varepsilon, 2\pi) \subset \mathbb{R}$ für beliebiges $\varepsilon > 0$. Beachte, dass q hier gerade den Winkel des Pendels beschreibt, also gerade gleich dem θ in unserem Pendelmodell aus Abschnitt 3.1.4 ist.

Wir können das System nun vollständig mittels $q(t)$ beschreiben. Mittels der Kettenregel kann man die Geschwindigkeit über $q(t)$ ausdrücken. Es gilt

$$v_i(t) = \frac{d}{dt} r_i(q(t), t) = \sum_{j=1}^l \frac{\partial r_i}{\partial q_j}(q(t), t) \dot{q}_j(t) + \frac{\partial r_i}{\partial t}(q(t), t), \quad i = 1, \dots, N.$$

Diese Gleichung kann wegen der linearen Unabhängigkeit der partiellen Ableitungen nach \dot{q}_j aufgelöst werden, was i.A. aber nicht explizit durchgeführt werden muss. Die Größen $\dot{q}_1, \dots, \dot{q}_l$ heißen *verallgemeinerte Geschwindigkeiten*.

Beispiel: Für das Pendel gilt

$$v(t) = \begin{pmatrix} \rho \cos q \\ \rho \sin q \\ 0 \end{pmatrix} \dot{q}(t),$$

Ebenso kann die kinetische Energie mittels q und \dot{q} als

$$E_{kin} = \sum_{i=1}^N \frac{m_i}{2} \|v_i\|^2 = \sum_{i=1}^N \frac{m_i}{2} \left\| \sum_{j=1}^l \frac{\partial r_i}{\partial q_j}(q(t), t) \dot{q}_j(t) + \frac{\partial r_i}{\partial t}(q(t), t) \right\|^2 =: \mathcal{T}(q, \dot{q}, t)$$

geschrieben werden.

Beispiel: Für das Pendel gilt

$$\mathcal{T}(q, \dot{q}, t) = \frac{m}{2} \rho^2 \dot{q}(t)^2$$

Für Kräfte $f^i \in \mathbb{R}^3$, $i = 1, \dots, N$, die jeweils auf den i -ten Massepunkt wirken, definiert man die *verallgemeinerten Kräfte*

$$F_j = \sum_{i=1}^N \left\langle f^i, \frac{\partial r_i}{\partial q_j} \right\rangle, \quad j = 1, \dots, l.$$

Wir nennen das mechanische System *konservativ*, falls eine reelle Funktion $W(r_1, \dots, r_N, t)$ existiert, so dass

$$f^i = -\frac{\partial W}{\partial r_i} =: -\nabla_i W$$

gilt. Für die verallgemeinerten Kräfte berechnet man dann

$$F_j = -\frac{\partial W(q, t)}{\partial q_j}$$

mit $W(q, t) = W(r(q), t)$. In Vektorform schreiben wir $F = -\nabla_q W(q, t)$. Die Funktion W kann physikalisch als die potentielle Energie des Systems interpretiert werden, weswegen man üblicherweise durch Addition einer geeigneten Konstanten die Bedingung $\min_q W(q, t) = 0$ sicher stellt. Beachte, dass die Addition einer Konstanten an $\nabla_q W$ nichts ändert.

Beispiel: Beim Pendel ohne Reibung wirkt auf den Massenpunkt die Kraft $f = (0, -mg, 0)^T$, die sich als $f = -\nabla W$ mit $W(r) = mgy$ schreiben lässt. Mit der oben eingeführten Darstellung $r(q) = (\rho \sin q, -\rho \cos q, 0)^T$ gilt $W(q, t) = -mg\rho \cos q$. Um $\min_q W(q, t) = 0$ zu gewährleisten, addieren wir $mg\rho$, d.h. wir setzen $W(q, t) = -mg\rho \cos q + mg\rho$.

Definition 3.5 Die Funktion

$$L(q, \dot{q}, t) = \mathcal{T}(q, \dot{q}, t) - W(q, t)$$

heißt *Lagrange-Funktion* des konservativen mechanischen Systems. \square

Die Variablen q und \dot{q} werden hier als (formal) unabhängige Variablen aufgefasst.

Aus der Lagrange-Funktion kann man nun die Bewegungsgleichungen des Systems herleiten: Es gelten die *Lagrange-Gleichungen*

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j}(q(t), \dot{q}(t), t) \right) - \frac{\partial L}{\partial q_j}(q(t), \dot{q}(t), t) = 0, \quad j = 1, \dots, l. \quad (3.14)$$

Die Herleitung dieser Gleichungen ergibt sich aus der physikalischen Bedingung, dass das *Wirkungsfunktional*

$$I(q) = \int_{t_0}^{t_1} L(q(t), \dot{q}(t), t) dt$$

entlang von Lösungen q minimal sein muss bzgl. aller differenzierbarer Funktionen, die die Punkte $(t_0, q(t_0))$ und $(t_1, q(t_1))$ verbinden. Setzt man $g(\alpha) = I(q + \alpha z)$ für $\alpha \in \mathbb{R}$ und eine beliebige differenzierbare Funktion z mit $z(t_0) = z(t_1) = 0$, so muss $g'(0) = 0$ gelten. Mit etwas Rechnung sieht man, dass

$$g'(0) = \sum_{j=1}^l \int_{t_0}^{t_1} \left(\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j}(q(t), \dot{q}(t), t) \right) - \frac{\partial L}{\partial q_j}(q(t), \dot{q}(t), t) \right) z(t) dt$$

ist, was schließlich auf die Lagrange-Gleichungen führt.

Beispiel: Für das Pendel ohne Reibung erhalten wir aus den obigen Überlegungen

$$L(q, \dot{q}, t) = \frac{m}{2} \rho^2 \dot{q}^2 + mg\rho \cos q + mg\rho,$$

also

$$\frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t) = \frac{m}{2} \rho^2 2\dot{q} = m\rho^2 \dot{q}$$

und

$$\frac{\partial L}{\partial q}(q, \dot{q}, t) = -mg\rho \sin q.$$

Damit erhalten wir die Bewegungsgleichung

$$\begin{aligned} 0 &= \frac{d}{dt}(m\rho^2\dot{q}(t)) + mg\rho \sin q(t) \\ &= m\rho^2\ddot{q}(t) + mg\rho \sin q(t). \end{aligned}$$

Da $\rho > 0$ und $m > 0$ ist, vereinfacht sich diese zu

$$0 = \rho\ddot{q}(t) + g \sin q(t),$$

was gerade die in Abschnitt 3.1.4 hergeleitete Gleichung (mit $q = \theta$) ist.

3.2.2 Dissipative Systeme

Wir haben die Lagrange–Gleichungen unter der Annahme hergeleitet, dass das mechanische System konservativ ist. Tatsächlich bedeutet dies, dass die wirkenden Kräfte nur von den Positionen r_i abhängen; weder externe noch geschwindigkeitsabhängige Kräfte (wie die Reibung) können hiermit modelliert werden.

Externe Kräfte können — wenn sie in verallgemeinerter Form F_e^j also mittels Ihrer Wirkung auf die q ausgedrückt sind — einfach durch Ersetzen der “0” durch F_e^j auf der rechten Seite von (3.14) eingeführt werden. Die Umrechnung von “realen” Kräften f^i auf verallgemeinerte Kräfte F_i kann dabei gemäß der Kettenregel mittels

$$F_i(q, t) = f^i(r(q), t)^T r'(q)$$

geschehen.

Bei der Reibung beschränken wir uns auf den einfachen Fall viskoser Reibung und bezeichnen mit $f^i \in \mathbb{R}^3$ nun die Reibungskräfte, die durch $f^i = -C^i v_i$ mit einer Diagonalmatrix $C^i = \text{diag}(c_x^i, c_y^i, c_z^i) \in \mathbb{R}^{3 \times 3}$ und den Geschwindigkeitsvektoren $v_i = (v_x^i, v_y^i, v_z^i)^T$ allgemein beschrieben werden können.

Wir wollen die Reibungskräfte analog zu den konservativen Kräften als Ableitung einer reellwertigen Funktion darstellen. Dazu definiert man die sogenannte *Ragleigh’sche Dissipationsfunktion*

$$\mathcal{D}(v_1, \dots, v_N) = \frac{1}{2} \sum_{i=1}^N (c_x^i v_x^i{}^2 + c_z^i v_z^i{}^2 + c_z^i v_z^i{}^2)$$

und definiert die Reibungskräfte f^i als

$$f^i = -\nabla_i \mathcal{D},$$

wobei $\nabla_i \mathcal{D}(v) \in \mathbb{R}^3$ den Gradienten von \mathcal{D} nach v^i bezeichnet.

Die Ragleigh–Funktion lässt sich als infinitesimale Arbeit des i -ten Partikels gegen die Reibungskraft interpretieren. Die von den Reibungskräften absorbierte Leistung ist gerade $2\mathcal{D}$ und wird als *Dissipationsrate* bezeichnet. Analog zu den verallgemeinerten Kräften lassen sich die *verallgemeinerten Reibungskräfte* als

$$F_j = -\frac{\partial \mathcal{D}(q, \dot{q}, t)}{\partial \dot{q}(t)}$$

mit $\mathcal{D}(q, \dot{q}, t) = \mathcal{D}(v(q, \dot{q}, t))$ berechnen. In Kurzform schreiben wir

$$F = -\nabla_{\dot{q}}\mathcal{D}(q, \dot{q}, t).$$

Die unter der Berücksichtigung der externen und Reibungskräfte erhaltenen *verallgemeinerten Lagrange-Gleichungen* lauten damit

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j}(q(t), \dot{q}(t), t) \right) - \frac{\partial L}{\partial q_j}(q(t), \dot{q}(t), t) + \frac{\partial \mathcal{D}}{\partial \dot{q}_j}(q(t), \dot{q}(t), t) = F_e^j(t), \quad j = 1, \dots, l \quad (3.15)$$

Beispiel: Wir wollen unser Pendelmodell um einen Reibungsterm ergänzen und setzen $\mathcal{D}(v) = \frac{1}{2}(cv_x^2 + cv_y^2)$. Damit erhalten wir

$$\mathcal{D}(q, \dot{q}, t) = \frac{c}{2}\rho^2(\cos^2 q + \sin^2 q)\dot{q}^2 = \frac{c}{2}\rho^2\dot{q}^2$$

und folglich

$$\frac{\partial \mathcal{D}}{\partial \dot{q}}(q(t), \dot{q}(t), t) = c\rho^2\dot{q}.$$

Die Bewegungsgleichung ergibt sich damit nach Division durch ρ zu

$$0 = m\rho^2\ddot{q}(t) + mg\rho \sin q(t) + c\rho\dot{q}.$$

Wir erhalten also wieder das bereits bekannte Modell, bei dem die Reibungskonstante c nun allerdings mit der Länge ρ multipliziert ist. Dies liegt daran, dass die Reibung hier an der Punktmasse wirkt, während sie im früheren Modell an der Drehachse wirkt.

3.2.3 Die Hamilton'sche Methode

Die Lagrange-Gleichungen führen in natürlicher Weise auf eine Differentialgleichung zweiter Ordnung, also eine Gleichung, in der q und \dot{q} auftreten. Der Hamilton-Formalismus, den wir abschließend kurz behandeln wollen, ermöglicht es, für konservative mechanische Systeme direkt ein System erster Ordnung herzuleiten, in dem die dabei verwendete "Hilfsfunktion" H eine wohldefinierte physikalische Interpretation besitzt.

Hierzu definieren wir das *verallgemeinerte Moment* als

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t) \in \mathbb{R}^l$$

und nehmen an, dass eine stetig differenzierbare Funktion $\dot{q}(q, p, t)$ existiert, so dass die Gleichung

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}(q, p, t), t)$$

gilt. Die Abbildung

$$(q, p, t) \mapsto (q, \dot{q}(q, p, t), t)$$

heißt dabei *Legendre-Transformation*.

Definition 3.6 Die reellwertige Funktion

$$H(q, p, t) = p^T \dot{q}(q, p, t) - L(q, \dot{q}(q, p, t), t)$$

heißt *Hamilton-Funktion* eines konservativen mechanischen Systems. \square

Beispiel: Für das Pendel gilt

$$p = \frac{\partial L}{\partial \dot{q}} L(q, \dot{q}, t) = m\rho^2 \dot{q},$$

also ist

$$\dot{q}(q, p, t) = \frac{p}{m\rho^2}$$

wegen

$$\frac{\partial L}{\partial \dot{q}} L(q, \dot{q}(q, p, t), t) = m\rho^2 \dot{q}(q, p, t) = m\rho^2 \frac{p}{m\rho^2} = p$$

die gesuchte Abbildung mit Legendre-Transformation

$$(q, p, t) \mapsto (q, p/(\rho^2 m), t).$$

Die Hamilton-Funktion unseres Pendels lautet demnach

$$H(q, p, t) = \frac{p^2}{m\rho^2} - \frac{m}{2}\rho^2 \left(\frac{p}{m\rho^2} \right)^2 - mg\rho \cos q + mg\rho = \frac{1}{2} \frac{p^2}{m\rho^2} - mg\rho \cos q + mg\rho.$$

Sei nun $q(t)$ eine Lösung von (3.14) mit

$$p(t) = \frac{\partial L}{\partial \dot{q}}(q(t), \dot{q}(t), t).$$

Dann folgt aus der Definition von $\dot{q}(q, p, t)$ die Identität

$$\dot{q}(t)(q(t), p(t), t) = \dot{q}(t).$$

Für H gilt nun

$$\begin{aligned} \frac{\partial H}{\partial p_j}(q(t), p(t), t) &= \dot{q}_j(t) + p^T \frac{\partial \dot{q}}{\partial p_j}(q(t), p(t), t) - \underbrace{\sum_{i=1}^l \frac{\partial L}{\partial \dot{q}_i}(q(t), \dot{q}(t), t)}_{=p_i(t)} \frac{\partial \dot{q}_i}{\partial p_j}(q(t), p(t), t) \\ &= \dot{q}_j(t) \end{aligned}$$

für $j = 1, \dots, l$ und

$$\begin{aligned} \frac{\partial H}{\partial q_j}(q(t), p(t), t) &= p^T \frac{\partial \dot{q}}{\partial q_j}(q(t), p(t), t) - \frac{\partial L}{\partial q_j}(q(t), p(t), t) \\ &\quad - \underbrace{\sum_{i=1}^l \frac{\partial L}{\partial \dot{q}_i}(q(t), \dot{q}(t), t)}_{=p_i(t)} \frac{\partial \dot{q}_i}{\partial q_j}(q(t), p(t), t) \\ &= -\frac{\partial L}{\partial q_j}(q(t), \dot{q}(t), t) = -\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j}(q(t), \dot{q}(t), t) \right) \\ &= -\dot{p}(t) \end{aligned}$$

Also erfüllt die Funktion $(q(t), p(t))$ die DGL erster Ordnung

$$\dot{q}(t) = \frac{\partial H}{\partial p}(q(t), p(t), t)$$

$$\dot{p}(t) = -\frac{\partial H}{\partial q}(q(t), p(t), t)$$

das sogenannte *Hamilton-System*.

Umgekehrt kann man nachweisen, dass jede Lösung des Hamilton-Systems eine Lösung der Lagrange-Gleichungen induziert. Die zwei Systeme sind also äquivalent.

Die Hamilton-Funktion ist deswegen eine schöne Form der Gleichung, da sie (in vielen Fällen) eine explizite physikalische Interpretation besitzt: $H(q, p, t)$ ist gerade die Gesamtenergie des Systems, die — aufgrund der Konservativität des Systems — entlang von Lösungen konstant ist.

Beispiel: Für das Pendel gilt

$$\frac{\partial H}{\partial p}(q, p, t) = \frac{p}{m\rho^2}$$

und

$$\frac{\partial H}{\partial q}(q, p, t) = mg\rho \sin q.$$

Wir erhalten also das Hamilton-System

$$\dot{q}(t) = \frac{p(t)}{m\rho^2}$$

$$\dot{p}(t) = -mg\rho \sin q(t)$$

Dies ist genau das bekannte Modell (3.9), wenn wir $\theta = q$ und $\omega = p/(m\rho^2)$ setzen. Der Vorteil der hier erhaltenen Skalierung liegt darin, dass die Hamilton-Funktion

$$H(q, p, t) = \frac{1}{2} \frac{p^2}{m\rho^2} - mg\rho \cos q + mg\rho$$

hier tatsächlich die Gesamtenergie des Systems beschreibt, die gerade die Summe der kinetischen Energie $\frac{1}{2} \frac{p^2}{m\rho^2}$ und der potentiellen Energie $-mg\rho \cos q + mg\rho$ ist.

Kapitel 4

Finanzmathematik

Die Finanzmathematik ist ein relativ junges und in den letzten Jahren stark gewachsenes Teilgebiet der angewandten Mathematik. Auch wenn die Börseneuphorie durch die anhaltende Baisse in den letzten Jahren deutlich nachgelassen hat, so hat die Finanzmathematik doch weiterhin ihre Berechtigung, da der Handel mit komplexen Finanzinstrumenten zwar etwas zurück gegangen ist, aber sicherlich auch in den kommenden Jahrzehnten von Bedeutung sein wird. Zudem tauchen komplexe Finanzinstrumente nicht nur im Zusammenhang mit Börsenspekulationen auf, sondern auch wenn es z.B. um die Absicherung gegen Wechselkursschwankungen geht, die für international agierende Firmen zum Alltagsgeschäft gehört. Dies ist auch der Bereich der Finanzmathematik, den wir hier näher betrachten wollen: Moderne Finanzderivate (deren einfachste Vertreter, die europäischen Optionen, wir hier genauer untersuchen werden) hängen in komplexer Weise von den zu Grunde liegenden Kursen (z.B. Aktien- oder Wechselkurse) ab. Eines der Ziele der Finanzmathematik ist es nun, den Wert eines solchen Derivates zu einer gegebenen Zeit zu ermitteln. Da der Wert von der (natürlich unbekannt) zukünftigen Entwicklung des zu Grunde liegenden Kurses abhängt, muss dafür zunächst ein Modell für den zukünftigen Kursverlauf entwickelt werden.

Da niemand ernsthaft behaupten kann, dass man mit mathematischen Methoden den Verlauf von Aktienkursen, Wechselkursen etc. vorhersagen kann, verwendet man in der Finanzmathematik üblicherweise keine *deterministischen* Differentialgleichungen, wie wir sie in den vorhergehenden Kapiteln kennen gelernt haben. Statt dessen benutzt man hier sogenannte *stochastische Differentialgleichungen*, die für jeden Anfangswert nicht nur eine, sondern viele Lösungstrajektorien besitzen, welche vom Zufall abhängen. Die Idee dabei ist, dass die verschiedenen Lösungen dieser Gleichung die möglichen zukünftigen Kursverläufe so approximieren, dass gewisse aus der Vergangenheit des Kursverlaufs ermittelte statistische Werte, wie z.B. die Varianz (also die Streuung um den Mittelwert) des Kurses gut modelliert werden.

In dem nun folgenden Abschnitt 4.1 werden wir die in der Finanzmathematik gebräuchlichste Klasse stochastischer Differentialgleichungen einführen. Wir versuchen dabei, auf stochastische Vorkenntnisse so weit wie möglich zu verzichten und werden statt dessen oftmals intuitive oder heuristische Erklärungen geben und Beweise weitgehend vermeiden. Eine tiefere Beschäftigung mit dieser Materie ist natürlich nur mit (zumindest grundlegenden) Stochastikkenntnissen möglich.

Auf Basis dieser Gleichungen werden wir dann im Abschnitt 4.2 ihre Anwendung in der Optionspreisbewertung erläutern.

4.1 Stochastische Differentialgleichungen

Um eine stochastische Differentialgleichung zu formulieren, müssen wir zunächst erklären, auf welche Weise der “Zufall” in unsere Differentialgleichung eingeht.

Vereinfacht kann man eine stochastische Differentialgleichung zunächst als eine Differentialgleichung der Form (1.1) mit der speziellen Wahl

$$\dot{x}(t) = f(t, x(t)) := a(t, x(t)) + b(t, x(t))g(t) \quad (4.1)$$

schreiben¹. Hierbei sind $a, b : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ gewöhnliche Funktionen, der Zufall kommt durch die Funktion g ins Spiel, die nämlich keine “feste” Funktion ist, sondern eine Funktion, die abhängig vom Zufall verschiedene Gestalt annehmen kann. Für jedes mögliche g erhalten wir für einen festen Anfangswert weiterhin eine eindeutige Lösung. Da diese aber nun von g abhängt und g zufällig ist, ist in diesem Sinne auch die Lösungstrajektorie zufällig.

Natürlich ist g nicht “irgendwie” vom Zufall abhängig, sondern genügt gewissen statistischen Regeln. Um diese zu definieren, werden wir zuerst einige Grundbegriffe einführen.

4.1.1 Einige stochastische Grundbegriffe

Wir beginnen mit dem Konzept des *Wahrscheinlichkeitsraumes*. Dieser besteht aus einer Menge Ω von *Elementarereignissen*, einer Menge von *Ereignissen* \mathcal{A} , deren Elemente $A \in \mathcal{A}$ Teilmengen $A \subseteq \Omega$ sind und einem *Wahrscheinlichkeitsmaß* $\mathbb{P} : \mathcal{A} \rightarrow \mathbb{R}$, das jedem Ereignis $A \in \mathcal{A}$ eine Wahrscheinlichkeit zuordnet. Man schreibt einen Wahrscheinlichkeitsraum als Tripel $(\Omega, \mathcal{A}, \mathbb{P})$ (\mathbb{P} steht für “Probability”=Wahrscheinlichkeit). Am Einfachsten lässt sich dieses abstrakte Konzept durch ein einfaches Beispiel erläutern:

Wenn der Raum z.B. das mathematische Modell eines Würfels ist, so kann das Ergebnis eines jeden Wurfs einen der Werte $\Omega = \{1, 2, 3, 4, 5, 6\}$ annehmen. Bei einem idealen Würfel ist die Wahrscheinlichkeit für jeden Wert gleich groß. Da man in der Mathematik die Konvention macht, dass sich alle Einzelwahrscheinlichkeiten zu 1 aufsummieren (formal: $\mathbb{P}(\Omega) = 1$) und gerade 6 Werte existieren deren Wahrscheinlichkeiten gleich groß sind, müssen diese also gleich $1/6$ sein. Man schreibt $\mathbb{P}(\omega) = 1/6$ für jedes $\omega \in \Omega$. Durch die Funktion $\mathbb{P}(\omega)$ für alle $\omega \in \Omega$ ist die Abbildung \mathbb{P} vollständig beschrieben. Hieraus kann man z.B. leicht die Wahrscheinlichkeit ausrechnen, dass man eine gerade Zahl würfelt, da sich diese Wahrscheinlichkeit als Summe der Einzelwahrscheinlichkeiten ergibt; mit $A = \{2, 4, 6\}$ ist nämlich

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \mathbb{P}(2) + \mathbb{P}(4) + \mathbb{P}(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

¹Dies ist eine vorläufige Schreibweise, siehe (4.13) für die mathematisch gebräuchliche Formulierung.

Im endlichen Fall gilt die allgemein die einfache Formel

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega), \quad (4.2)$$

d.h. aus der Wahrscheinlichkeit der Elementarereignisse kann man die Wahrscheinlichkeit beliebiger Ereignisse $A \in \mathcal{A}$ durch Aufsummieren berechnen. Welche Ereignisse man in die Ereignismenge \mathcal{A} aufnimmt, hängt davon ab, welche Ereignisse man modellieren will. In jedem Fall muss Ω und \emptyset in \mathcal{A} enthalten sein und die Menge \mathcal{A} muss abgeschlossen bezüglich Vereinigung und Durchschnittbildung sein. Diese Details sind für eine rigorose mathematische Betrachtung zwar wichtig, für uns hier aber eher nebensächlich, da wir sie nicht direkt verwenden werden.

Stattdessen benutzen wir ein weiteres Konzept, nämlich das der *Zufallsvariablen*. Eine Zufallsvariable X ist keine Variable im klassischen Sinne, sondern eine Abbildung $X : \Omega \rightarrow \mathbb{R}$; jedem Element aus Ω wird durch X eine reelle Zahl zugeordnet. Wenn wie im obigen Würfelbeispiel $\Omega \subset \mathbb{R}$ gilt, kann man X im einfachsten Fall als Identität definieren. Der große Vorteil bei der Verwendung von Zufallsvariablen liegt darin, dass man auf den reellen Zahlen alle Techniken der Analysis zur Verfügung hat, was das Rechnen mit Wahrscheinlichkeiten deutlich vereinfacht. Wenn man zufällig einen Wert ω aus Ω “zieht”, so nennt man den Wert $X(\omega)$ eine *Realisierung* von X . Anders gesagt, ist eine Realisierung $X(\omega)$ von X gerade das Ergebnis eines mittels X ausgewerteten Zufallsexperimentes.

Um zu beschreiben, wie sich die zufälligen Realisierungen von X statistisch verhalten, benötigt man das Konzept der *Wahrscheinlichkeitsverteilung* \mathbb{P}_X von X . Für eine Menge $B \subset \mathbb{R}$ gibt $\mathbb{P}_X(B)$ die Wahrscheinlichkeit an, mit der eine Realisierung $X(\omega)$ in B liegt. Man kann \mathbb{P}_X ausgehend vom Wahrscheinlichkeitsmaß \mathbb{P} auf Ω definieren, indem man einfach

$$\mathbb{P}_X(B) := \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\})$$

setzt. Diese Definition setzt natürlich die Kenntnis von \mathbb{P} voraus. (Wir nehmen hier zudem implizit an, dass $A = \{\omega \in \Omega \mid X(\omega) \in B\}$ in \mathcal{A} liegt, so dass wir \mathbb{P} darauf anwenden können.) Wir werden im Folgenden aber meist andersherum vorgehen, indem wir uns nämlich für X ein \mathbb{P}_X vorgeben, ohne dass wir das zugrunde liegende Wahrscheinlichkeitsmaß \mathbb{P} auf Ω näher betrachten oder auch nur kennen. Falls X nur endlich viele Werte annimmt, also $X(\Omega) = \{x_1, \dots, x_N\}$ ist (man sagt, X ist eine *diskrete Zufallsvariable*), so reicht es dazu wiederum aus, nur $\mathbb{P}_X(x_i)$ festzulegen und $\mathbb{P}_X(x) = 0$ für $x \notin X(\Omega)$ zu setzen. Analog zu (4.2) gilt dann

$$\mathbb{P}_X(B) = \sum_{x \in B} \mathbb{P}_X(x), \quad (4.3)$$

Falls Ω eine unendliche Menge ist und X unendlich viele verschiedene Werte annimmt, so ist die Sache etwas komplizierter. Wir können uns für unsere Zwecke aber auf den Spezialfall einschränken, in dem $X(\Omega) = \mathbb{R}$ ist, also alle möglichen Werte in \mathbb{R} angenommen werden (man sagt, X ist eine *kontinuierliche Zufallsvariable*) und wir die Verteilung von X durch eine stetige Funktion $p : \mathbb{R} \rightarrow \mathbb{R}$, die sogenannte *Dichtefunktion* beschreiben können. Diese Funktion p spielt in gewissem Sinne die Rolle der Wahrscheinlichkeiten $\mathbb{P}_X(x_i)$ für Elementarereignisse im endlichen Fall. Zwar gibt $p(x)$ hier nicht die Wahrscheinlichkeit des

Elementarereignisses $X(\omega) = x$ an, man kann aber analog zu (4.3) eine Formel für die Wahrscheinlichkeit $\mathbb{P}_X(B)$ von Ereignissen $B \subseteq \mathbb{R}$ angeben. Diese ist gegeben durch

$$\mathbb{P}_X(B) = \int_{-\infty}^{\infty} \chi_B(x)p(x)dx, \quad (4.4)$$

wobei χ_B die sogenannte *charakteristische Funktion* der Menge B ist, die durch

$$\chi_B(x) = \begin{cases} 1, & \text{falls } x \in B \\ 0, & \text{falls } x \notin B \end{cases}$$

gegeben ist. Die spezielle Wahrscheinlichkeitsfunktion

$$P(a) = \mathbb{P}_X((-\infty, a]) = \int_{-\infty}^a p(x)dx$$

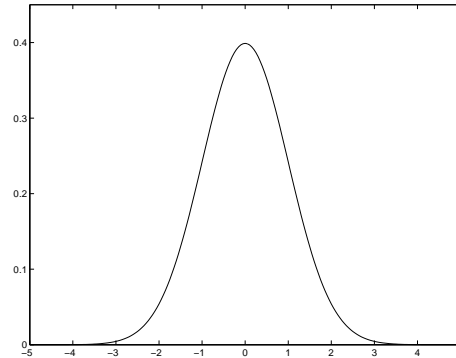
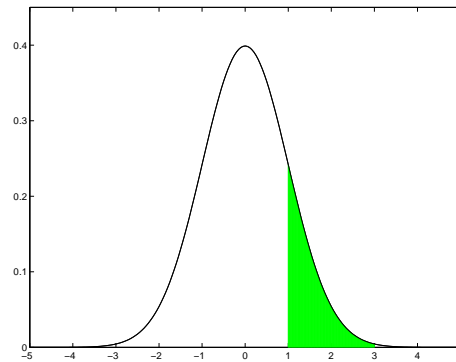
heißt (zu p gehörige) *Verteilungsfunktion*. An diesem Punkt stellen sich eine ganze Reihe von mathematischen Fragen: Wie ist das Integral (4.4) definiert (Antwort: als Lebesgue–Integral), für welche Mengen $B \subseteq \Omega$ existiert (4.4) (Antwort: gerade die Mengen, für die χ_B eine Lebesgue–messbare Funktion ist), etc. Diese Fragen werden wir hier nicht im Detail betrachten können; wir verweisen dazu auf die Einführung in die Stochastik.

Ereignisse $B \subset \mathbb{R}$, für die $\mathbb{P}_X(B) = 1$ gilt, nennt man *fast sicher*. Das Adjektiv “fast” ergibt sich hierbei aus einer Eigenschaft kontinuierlicher Zufallsvariablen, die am Anfang paradox erscheint: $\mathbb{P}_X(B) = 1$ garantiert nicht, dass $X(\omega)$ für *jedes* $\omega \in \Omega$ in B liegt. Die tiefere mathematische Begründung liegt in der Lebesgue–Integrationstheorie; intuitiv kann man dies aber auch schon auf den natürlichen Zahlen erläutern: Nehmen wir an, dass der Wahrscheinlichkeitsraum einen “Würfel” mit N verschiedenen Seiten $\Omega = \{1, \dots, N\}$ beschreibt und X die Identität ist. Die Wahrscheinlichkeit des Ereignisses $X \in B = \{2, \dots, N\}$ beträgt offenbar $\mathbb{P}_X(B) = (N - 1)/N = 1 - 1/N$. Je größer die Seitenanzahl N wird, desto näher ist $\mathbb{P}_X(B)$ an 1. Für einen hypothetischen Würfel mit $N = \infty$ Seiten gilt demnach $\mathbb{P}_X(B) = 1$, trotzdem ist es nicht unmöglich, dass eine 1 geworfen wird, also $X(\omega) = 1$ ist. Dieses Ereignis ist nur extrem unwahrscheinlich in dem Sinne, dass es unendlich viele andere mögliche Ergebnisse gibt. Beim Rechnen mit kontinuierlichen Zufallsvariablen muss man zwischen “fast sicher” und “sicher” nicht unterscheiden, da man die sehr unwahrscheinlichen anderen Ereignisse einfach vernachlässigen kann, was sich wiederum mathematisch rigoros aus der Lebesgue–Integrationstheorie ergibt.

Wir werden uns im Folgenden nur mit einer ganz speziellen kontinuierlichen Zufallsvariablen X beschäftigen, nämlich der sogenannten *Gauß–verteilten Zufallsvariablen*. Für diese Zufallsvariable hängt die Dichte p von zwei (konstanten) Parametern $\mu, \sigma \in \mathbb{R}$ mit $\sigma > 0$ ab und ist gegeben durch

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x - \mu)^2}{2\sigma^2}\right). \quad (4.5)$$

Für Gauß–verteilte (auch *normalverteilte* oder kurz *Gauß’sche*) Zufallsvariablen schreibt man oft $X \sim N(\mu, \sigma^2)$. Eine Zufallsvariable $X \sim N(0, 1)$ heißt *standard–normalverteilt*. Abbildung 4.1 zeigt die Dichtefunktion p für die Standard–Normalverteilung.

Abbildung 4.1: Dichtefunktion der Gauß'schen Zufallsvariablen für $\mu = 0$ und $\sigma = 1$ Abbildung 4.2: Wahrscheinlichkeit von $B = [1, 3]$ als Fläche unter p

Grafisch kann man sich (4.4) so vorstellen: Falls $B = [a, b]$ ein Intervall ist, so ist die Wahrscheinlichkeit, dass eine Realisierung $X(\omega)$ von X in B liegt, gerade gleich der Fläche unter dem Graphen von p . In Abbildung 4.2 ist dies für $B = [1, 3]$ dargestellt.

Die Gauß'sche Zufallsvariable kann also Werte in ganz \mathbb{R} annehmen (p ist nirgends gleich 0), allerdings ist die Wahrscheinlichkeit betragsmäßig großer Werte sehr gering, da p für große $|x|$ sehr klein wird.

Ausgehend von der eben betrachteten Verteilung kann man eine Reihe von Kenngrößen für X definieren; zwei davon sind für uns wichtig. Wir beschränken uns hierbei auf kontinuierliche Zufallsvariablen.

Die erste wichtige Größe ist der *Erwartungswert* von X , der mit $\mathbb{E}(X)$ bezeichnet wird. Anschaulich ist dies nichts anderes als der Mittelwert, den man erhält, wenn man über alle möglichen Realisierungen $X(\omega)$ von X bezüglich ihrer Wahrscheinlichkeit mittelt, für eine Folge $X(\omega_1), X(\omega_2), \dots$ von Realisierungen also

$$\mathbb{E}(X) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X(\omega_i). \quad (4.6)$$

Formal definiert man den Erwartungswert geeigneter über die Dichtefunktion, nämlich als

$$\mathbb{E}(X) := \sum_{x_i \in X(\Omega)} x \mathbb{P}_X(x) \quad (4.7)$$

für diskrete und, analog,

$$\mathbb{E}(X) := \int_{-\infty}^{\infty} xp(x)dx, \quad (4.8)$$

für kontinuierliche Zufallsvariablen, wobei wir voraussetzen, dass dieses Integral existiert. Für die Gauß-Verteilung errechnet man $\mathbb{E}(X) = \mu$. Es ist im übrigen nicht trivial zu beweisen, dass diese beiden Definitionen von $\mathbb{E}(X)$ übereinstimmen.

Man rechnet leicht nach, dass der Erwartungswert linear in X ist, d.h. für zwei Zufallsvariablen X_1 und X_2 und $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt $\mathbb{E}(\alpha_1 X_1 + \alpha_2 X_2) = \alpha_1 \mathbb{E}(X_1) + \alpha_2 \mathbb{E}(X_2)$. Beachte hierbei, dass Summen, Differenzen und reelle Vielfache von Zufallsvariablen wieder Zufallsvariablen sind.

Die zweite wichtige Größe ist die *Varianz* von X , geschrieben $\text{Var}(X)$. Sie gibt an, wie weit sich die einzelnen Realisierungen im Mittel vom Erwartungswert entfernen können, man sagt auch, wie weit die Werte “streuen”. Sie ist definiert durch

$$\text{Var}(X) = \mathbb{E}((X - \mu)^2).$$

Eine einfache Rechnung ergibt wegen $\mathbb{E}(X) = \mu$

$$\text{Var}(X) = \mathbb{E}((X - \mu)^2) = \mathbb{E}(X^2 - 2\mu X + \mu^2) = \mathbb{E}(X^2) - 2\mu \mathbb{E}(X) + \mu^2 = \mathbb{E}(X^2) - \mu^2.$$

Für die Gauß-Verteilung errechnet man $\text{Var}(X) = \sigma^2$.

Zuletzt wollen wir noch kurz das Prinzip der *Unabhängigkeit* von Zufallsvariablen betrachten. Zwei auf dem gleichen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ definierte Zufallsvariablen $X_1 : \Omega \rightarrow \mathbb{R}$ und $X_2 : \Omega \rightarrow \mathbb{R}$ heißen *unabhängig*, wenn eine Realisierung $X_1(\omega)$ von X_1 keine Informationen über die Realisierung $X_2(\omega)$ von X_2 (für das gleiche ω) enthält, und umgekehrt. Auf die mathematisch präzise Beschreibung dieser informellen Definition wollen wir verzichten; wir wollen sie aber an einem Beispiel erläutern.

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ der Wahrscheinlichkeitsraum, der das (gleichzeitige) Würfeln mit zwei Würfeln beschreibt. Jedes Elementarereignis ist also ein Paar $\omega = (\omega_1, \omega_2)$ von Würfelwerten, wobei die erste Komponente ω_1 das Ergebnis des ersten Würfels und ω_2 das Ergebnis des zweiten Würfels beschreibt (wir nehmen an, dass die zwei Würfel unterscheidbar sind). Dies führt zu $\Omega = \{(1, 1), (1, 2), \dots, (6, 6)\} \subset \mathbb{N}^2$. Als \mathcal{A} wählen wir alle möglichen Teilmengen von Ω . Offenbar ist (für ideale Würfel) $\mathbb{P}(\omega) = 1/36$ für jedes Element $\omega \in \Omega$, da Ω gerade 36 = 6 × 6 Elementarereignisse enthält, die alle gleich wahrscheinlich sind. Wir betrachten nun drei Zufallsvariablen, die als Wert das Ergebnis des ersten Würfels, des zweiten Würfels und die Summe der Würfelwerte ausgeben, also $X_1(\omega) = \omega_1$, $X_2(\omega) = \omega_2$ und $X_3(\omega) = \omega_1 + \omega_2$. Jede Realisierung dieser Variablen entspricht nun einem (simultanen) Wurf der zwei Würfel. Offenbar kann man aus dem Ergebnis des ersten Würfels keinerlei Rückschlüsse auf das Ergebnis des zweiten Würfels ziehen: Unabhängig davon, was beim ersten Würfel herauskam, ist beim zweiten Würfel jeder Wert 1, . . . , 6 weiterhin gleich wahrscheinlich. Die Zufallsvariablen X_1 und X_2 sind also *unabhängig*. Wenn wir aber X_1 und X_3

betrachten, so ändert sich die Situation: Die Wahrscheinlichkeit, dass $X_3(\omega)$ den Wert 12 annimmt, ist gerade $1/36$ (von den 36 möglichen Würfelkombinationen liefert gerade eine die Summe 12). Wenn wir aber wissen, dass die Realisierung $X_1(\omega) = \omega_1 = 1$ ist, so kann die Summe $X_3(\omega) = \omega_1 + \omega_2 = 1 + \omega_2$ höchstens den Wert 7 annehmen; die Wahrscheinlichkeit für $X_3(\omega) = 12$ ist unter dieser Vorinformation also gleich 0. Diese Wahrscheinlichkeit unter zusätzlicher Vorinformation nennt man *bedingte Wahrscheinlichkeit*. Da sich nun die bedingte Wahrscheinlichkeit der Zufallsvariablen X_3 unter Vorinformation aus X_1 von der unbedingten Wahrscheinlichkeit unterscheidet, ist hier gerade der Fall *nicht unabhängiger* Zufallsvariablen gegeben.

Für unabhängige Zufallsvariablen X_1 und X_2 gelten eine Reihe von Eigenschaften, z.B. gilt

$$\mathbb{E}(X_1 X_2) = \mathbb{E}(X_1) \mathbb{E}(X_2), \quad (4.9)$$

was für beliebige Zufallsvariablen im Allgemeinen nicht gilt.

4.1.2 Der Wiener Prozess

Zur Definition der stochastischen Differentialgleichung benötigen wir nun einen Begriff für zufällige Funktionen. Eine “zufällige Funktion” ist dabei nichts anderes als eine Funktion $X : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$, die für jedes feste t gerade eine Zufallsvariable $X(t, \cdot) : \Omega \rightarrow \mathbb{R}$ ist, wobei für alle t der gleiche Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ zu Grunde liegt. Solch eine Funktion nennt man *stochastischen Prozess*. Beachte, dass wir keinerlei Annahmen über den Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ voraussetzen; wichtig ist nur, dass alle Zufallsvariablen X auf dem selben Raum definiert sind. Für jedes feste $\omega \in \Omega$ ist die Realisierung $X(\cdot, \omega) : \mathbb{R} \rightarrow \mathbb{R}$ eine “normale” reelle Funktion. Statt Realisierung spricht man bei stochastischen Prozessen vom *Pfad*.

Der stochastische Prozess, den wir hier näher betrachten wollen und aus dem das obige g später hervorgehen wird, ist der sogenannte *Wiener Prozess*, der üblicherweise mit W bezeichnet wird und für $t \geq 0$ definiert ist. Die Definition des Wiener Prozesses ergibt sich aus der von N. Wiener² eingeführten mathematischen Beschreibung der *Brown'schen Bewegung*, die in der Physik die zufällige Bewegung eines auf einer Wasseroberfläche schwimmenden Teilchens beschreibt. Formal verlangt man die folgenden Bedingungen:

- (i) $W(t)$ ist eine Gauß-verteilte Zufallsvariable mit $\mathbb{E}(W(t)) = 0$ und $\text{Var}(W(t)) = t$
- (ii) Für $t_1 \geq t_0 \geq 0$ sind die *Inkmente* $W(t_1) - W(t_0)$ Gauß-verteilte Zufallsvariablen mit $\mathbb{E}(W(t_1) - W(t_0)) = 0$ und $\text{Var}(W(t_1) - W(t_0)) = t_1 - t_0$
- (iii) Für $s_1 \geq s_0 \geq t_1 \geq t_0 \geq 0$ sind die Inkmente $W(t_1) - W(t_0)$ und $W(s_1) - W(s_0)$ unabhängige Zufallsvariablen.

Beachte, dass wir oben die Gauß-Verteilung nur für $\sigma > 0$ definiert haben. Hier erhalten wir für $t = 0$ die Bedingung $\text{Var}(W(0)) = 0$, womit einfach $\mathbb{P}_{W(0)}(\mu) = 1$ gemeint ist. Mit anderen Worten nimmt die Zufallsvariable $W(0)$ also fast sicher ihren Erwartungswert μ (hier $\mu = 0$) an.

²US-amerikanischer Mathematiker, 1894–1964

Ein Pfad $W(t, \omega)$ von W ist also gerade eine der vielen möglichen zufälligen Funktionen, die (in ihrer Gesamtheit) den obigen Bedingungen genügen. Für diese Pfade kann man beweisen, dass sie fast sicher stetig in t sind (zur Erinnerung: dies bedeutet, dass das Ereignis $A = \{\omega \in \Omega \mid X(t, \omega) \text{ ist stetig in } t\}$ die Wahrscheinlichkeit $\mathbb{P}(A) = 1$ besitzt) aber auch fast sicher nirgends differenzierbar. Abbildung 4.3 zeigt drei verschiedene Pfade des Wiener Prozesses.

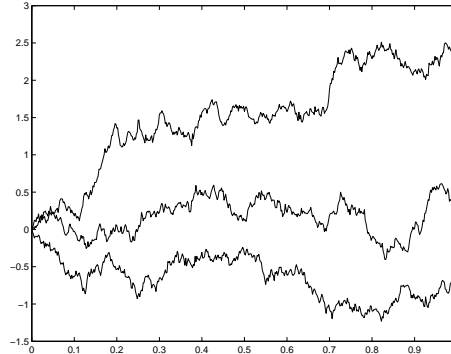


Abbildung 4.3: Verschiedene Pfade des Wiener Prozesses

Die Bedingung (i) besagt gerade, dass die Streuung der Werte der Pfade $W(t, \omega)$ um so größer wird, je größer t wird. Der Mittelwert ist dabei zu jedem Zeitpunkt gleich 0. Die Bedingung (ii) bedeutet, dass der stochastische Prozess $\widetilde{W}(t) = W(t + t_0) - W(t_0)$ wieder ein Wiener Prozess ist. Wenn wir also ein ‘‘Endstück’’ eines Pfades abschneiden und dessen Anfangspunkt in den Nullpunkt verschieben, so erhalten wir wieder einen Pfad des Wiener Prozesses. Diese zwei Eigenschaften kann man anhand der Grafik gut nachvollziehen. Bedingung (iii) schließlich besagt, dass man aus der Kenntnis eines Teilstücks eines Pfades auf einem Teilintervall $[t_0, t_1]$ keinerlei Vorhersagen über den Verlauf auf einem anderen Teilintervall $[s_0, s_1]$ mit $s_0 \geq t_1$ machen kann, wenn man nur die Informationen über die Anfangswerte durch Abziehen der Werte $W(s_0)$ und $W(t_0)$ ‘‘löscht’’. Der Pfad kann sich also zu jedem Zeitpunkt mit exakt der gleichen Wahrscheinlichkeit nach oben oder unten bewegen, egal welchen Verlauf er bis zu diesem Zeitpunkt genommen hat.

4.1.3 Das Ito-Integral

Betrachtet man die Pfade in Abbildung 4.3, so kann man durchaus eine vage Ähnlichkeit mit den Kursdiagrammen z.B. von Aktienkursen erkennen. Selbst für eine sehr grobe Modellierung von solchen Kurswerten ist der Wiener Prozess aber doch noch zu einfach, denn es gibt keine Parameter, an denen man ‘‘drehen’’ könnte, um den modellierten Verlauf einem echten Kursverlauf anzupassen: Der Wiener Prozess besitzt noch zu wenig Struktur.

Er ist aber bestens geeignet, um als ‘‘Zutat’’ in der Definition stochastischer Differentialgleichungen eine wichtige Rolle zu spielen; vereinfachend gesagt soll die zufällige Funktion g in (4.1) gerade die Ableitung des Wiener Prozesses sein. Um die dabei auftretenden mathematischen Probleme und ihre Lösung zu erläutern, wollen wir hier zunächst die denkbar

einfachste stochastische Differentialgleichung betrachten und dann zur allgemeinen Formulierung übergehen.

Zunächst einige Anmerkungen zur Notation: Da die Lösung einer SDG über den eingehenden Wiener Prozess wieder eine zufällige Funktion — also ein stochastischer Prozess — ist, verwenden wir hier für den gesuchten unbekanntenen Prozess die groß geschriebene Bezeichnung $X(t)$, bzw. mit Anfangswert $X_0 \in \mathbb{R}^n$ und Anfangszeit $t_0 \in \mathbb{R}$ die Schreibweise $X(t; t_0, X_0)$. Wir erlauben hierbei, dass X vektorwertig, also aus dem \mathbb{R}^n ist, was einfach bedeutet, dass $X = (X_1, X_2, \dots, X_n)^T$ ist, wobei die X_i reellwertige stochastische Prozesse im oben eingeführten Sinne sind. Zu jedem Pfad $W(t, \omega)$ des eingehenden Wiener Prozesses gehört dann ein Lösungspfad des X -Prozesses, den wir mit $X(t; t_0, X_0, \omega)$ bezeichnen.

Die technische Hauptschwierigkeit in der mathematischen Formulierung stochastischer Differentialgleichungen zeigt sich nun bereits bei einer scheinbar trivialen Aufgabe, nämlich dem Problem, eine stochastische Differentialgleichung aufzustellen, deren Lösung gerade der Wiener Prozess ist. Scheinbar trivial ist die Aufgabe deswegen, weil wir ja den Wiener Prozess als gegeben voraussetzen und in der Formulierung verwenden dürfen, weswegen es nahe liegt, einfach die Differentialgleichung

$$\frac{d}{dt}X(t) = \frac{d}{dt}W(t) \quad (4.10)$$

mit Anfangsbedingung $X_0 = W(0)$ zur Anfangszeit $t_0 = 0$ zu verwenden. Das Problem ist jetzt aber: Was verstehen wir unter “ $\frac{d}{dt}W(t)$ ”? Man würde vielleicht versuchen, die Ableitung pfadweise auffassen, d.h., wir berechnen die Ableitung für jeden Pfad $W(t, \omega)$. Nur ist ein typischer Pfad $W(t, \omega)$, wie oben erwähnt, nirgends differenzierbar.

Zunächst einmal bietet es sich an, die Gleichung (4.10) analog zu (1.3) in Integralform

$$X(t) = X_0 + \int_0^t \frac{d}{d\tau}W(\tau)d\tau$$

zu schreiben. Jetzt könnten wir formal integrieren, was uns aber bei der Frage “was ist $\frac{d}{dt}W(t)$?” nicht weiter bringt. Für das obige sogenannte *stochastische Integral* hat sich in der Literatur die kürzere Schreibweise

$$\int_0^t dW_\tau$$

eingebürgert, die wir hier übernehmen wollen. Dies zeigt die Richtung auf, die wir zur Lösung unseres Problems einschlagen wollen: Anstatt die Ableitung $\frac{d}{dt}W(t)$ zu betrachten, werden wir versuchen, diesem stochastischen Integral eine mathematische Definition zu geben, die

- (i) wohldefiniert ist, obwohl $\frac{d}{dt}W(t)$ nicht existiert
- (ii) das gewünschte Ergebnis, nämlich $X(t) = W(t)$, liefert
- (iii) sich auf allgemeinere Integrale der Form

$$I(F) := \int_{t_0}^{t_1} F(t)dW_t \quad (4.11)$$

verallgemeinern lässt, damit wir auch kompliziertere SDGs formulieren können. Hierbei ist F wiederum ein stochastischer Prozess, der auf dem selben Wahrscheinlichkeitsraum wie W definiert ist.

Wir wollen dieses Konzept nun für Integrale der Form (4.11) angeben. Die hier vorgestellte Lösung geht auf Kiyosi Ito³ zurück und wurde in den 1940er Jahren entwickelt. Die Idee besteht darin, das Integral (4.11) für jedes Paar von Pfaden $F(t, \omega)$ und $W(t, \omega)$ durch den Limes einer geeigneten Summe zu approximieren. Wir wählen dazu ein $N \in \mathbb{N}$ und eine Folge von Zeiten $\tau_i^{(N)}$, $i = 0, 1, \dots, N$ mit

$$t_0 = \tau_0^{(N)} < \tau_1^{(N)} < \dots < \tau_N^{(N)} = t_1$$

und definieren für jedes $\omega \in \Omega$

$$I^{(N)}(F)(\omega) := \sum_{i=0}^{N-1} F(\tau_i^{(N)}, \omega)(W(\tau_{i+1}^{(N)}, \omega) - W(\tau_i^{(N)}, \omega)).$$

Das Integral (4.11) wird nun über den Limes dieser Summe definiert. Betrachte eine Familie von Folgen $\tau_i^{(N)}$ für $N \in \mathbb{N}$ mit $\lim_{N \rightarrow \infty} \delta(N) = 0$, wobei $\delta(N) := \max_{i=1, \dots, N} \tau_i^{(N)} - \tau_{i-1}^{(N)}$ ist. Dann definieren wir

$$I(F) := \lim_{N \rightarrow \infty} I^{(N)}(F). \quad (4.12)$$

Diese Definition wirft zunächst eine Reihe von Fragen auf, denn da die Pfade des Wiener Prozesses sehr unangenehme Funktionen sein können, ist nicht garantiert, dass dieser Limes für jedes ω überhaupt existiert. Tatsächlich lag der Haupttrick von Ito darin, zu definieren, was der Limes in (4.12) eigentlich bedeuten soll. Dieser Limes ist nämlich *nicht* pfadweise zu verstehen (in dem Sinne, dass wir $\lim_{N \rightarrow \infty} I^{(N)}(F)(\omega)$ für jedes $\omega \in \Omega$ bilden), sondern man muss die Werte $I^{(N)}(F)$ ebenso wie das Integral $I(F)$ wieder als Zufallsvariablen $I^{(N)}(F) : \Omega \rightarrow \mathbb{R}$ bzw. $I(F) : \Omega \rightarrow \mathbb{R}$ auffassen. Für Zufallsvariablen gibt es verschiedene Konvergenzbegriffe und der hier geeignete ist der Begriff der *Quadrat-Mittel-Konvergenz*, der wie folgt definiert ist: Eine Folge von Zufallsvariablen $X_N : \Omega \rightarrow \mathbb{R}$ konvergiert im Quadrat-Mittel-Sinne gegen eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$, falls

$$\lim_{N \rightarrow \infty} \mathbb{E}(|X_N - X|^2) = 0$$

gilt. Mit diesem Konvergenzbegriff kann man zeigen, dass die Folge $I^{(N)}(F)$ (unter geeigneten Bedingungen an F) tatsächlich konvergiert und (4.12) also wohldefiniert ist. Das resultierende Integral wird *Ito-Integral* genannt und es besitzt tatsächlich die oben aufgeführten gewünschten Eigenschaften (i)–(iii).

Mit Hilfe des Ito-Integrals können wir die informelle Schreibweise (4.1) mathematisch präzise formulieren. Statt der üblichen Differentialgleichungsschreibweise schreibt man *Ito-stochastische Differentialgleichungen* nämlich als

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW_t. \quad (4.13)$$

³japanischer Mathematiker, *1915, oft auch Itô geschrieben

Dies ist nur eine symbolische Schreibweise; was mit (4.13) tatsächlich gemeint ist, ist die längere Integralschreibweise

$$X(t) = X(t_0) + \int_{t_0}^t a(\tau, X(\tau))d\tau + \int_{t_0}^t b(\tau, X(\tau))dW_\tau,$$

bei der das zweite Integral gerade das Ito-Integral ist. Dies liefert eine mathematisch fundierte und brauchbare Definition stochastischer Differentialgleichungen. Falls $b(t, x) \equiv 0$ ist, also kein stochastischer Anteil vorhanden ist, reduziert sich (4.13) auf

$$X(t) = X(t_0) + \int_{t_0}^t a(\tau, X(\tau))dt \iff \frac{d}{dt}X(t) = a(t, X(t)),$$

also auf die wohlbekannteste deterministische gewöhnliche Differentialgleichung. Der deterministische Anteil der Gleichung $a(t, X(t))$ wird oft *Drift* genannt, der stochastische Anteil $b(t, X(t))$ oft als *Diffusion* bezeichnet.

Natürlich lässt sich (4.13) in vielfacher Hinsicht erweitern, z.B. kann man statt nur einem W mehrere unabhängige Wiener Prozesse W^1, \dots, W^m eingehen lassen, was zur Gleichung

$$dX(t) = a(t, X(t))dt + \sum_{j=1}^m b_j(t, X(t))dW_t^j$$

führt.

Für das Ito-Integral lassen sich eine Reihe von Eigenschaften beweisen; auf die meisten können wir hier nicht näher eingehen. In Lemma 4.2 werden wir eine wichtige Eigenschaft kennen lernen, eine weitere Eigenschaft ist die Tatsache, dass

$$\mathbb{E} \left(\int_{t_0}^{t_1} F(t)dW_t \right) = 0 \tag{4.14}$$

ist, falls die Zufallsvariablen $F(t)$ und $W(s) - W(t)$ unabhängig sind für $s > t \geq 0$ (dies gilt immer, falls F die Lösung einer SDE bzgl. W ist). Eigenschaft (4.14) gilt zunächst für die Approximationen $I^{(N)}(F)$, weil nach Annahme an W und wegen der Unabhängigkeit und (4.9) die Gleichung

$$\mathbb{E}(F(\tau_i)(W(\tau_{i+1}) - W(\tau_i))) = \mathbb{E}(F(\tau_i)) \underbrace{\mathbb{E}(W(\tau_{i+1}) - W(\tau_i))}_{=0} = 0 \tag{4.15}$$

gilt, und überträgt sich durch die Limesbildung auf $I(F)$.

Bemerkung 4.1 Es sollte hier erwähnt werden, dass es eine weitere sinnvolle stochastische Integraldefinition gibt, die auf R. Stratonovich⁴ zurück geht. Das Stratonovich-Integral

$$\int_{t_0}^{t_1} F(t) \circ dW_t$$

⁴russischer Mathematiker, 1930–1997

wird über eine ähnliche Limes-Bildung wie das Ito-Integral definiert und liefert ebenfalls eine mathematisch fundierte Definition stochastischer DGLs. Die beiden Integrale unterscheiden sich allerdings in den Rechenregeln ebenso wie in der Form der Lösungen. Die zugehörigen *Stratonovich-SDGs* werden in der Form

$$dX(t) = a(t, X(t))dt + b(t, X(t)) \circ dW_t$$

geschrieben. Aus Zeitgründen können wir auf diese zweite Definition und auf die Gemeinsamkeiten und Unterschiede zum Ito-Integral hier nicht näher eingehen. \square

4.2 Die Bewertung von Optionen

Die Optionsbewertung ist eine der einfachsten Aufgabenstellung der Finanzmathematik. Trotzdem ist sie alles andere als trivial, weswegen wir sie hier beispielhaft als mathematische Anwendung betrachten werden. Wir werden zunächst definieren, was eine Option überhaupt ist, werden dann kurz auf eine wichtige Formel für die stochastische Analysis eingehen und dann SDG-Modelle zur Modellierung von Kursverläufen betrachten. Schließlich werden wir zwei Methoden kennen lernen, mit denen die Preise von Optionen dann in der Praxis wirklich berechnet werden können.

4.2.1 Optionen

Ganz allgemein ist eine Option ein Vertrag, der dem Inhaber die Möglichkeit (aber nicht die Verpflichtung) gibt, ein Wertpapier zu einem bestimmten Zeitpunkt zu einem vorher vereinbarten *Ausübungspreis* zu kaufen (*Call-Option*) oder zu verkaufen (*Put-Option*). Das zugrunde liegende Wertpapier heißt *Basiswert* und ist dabei z.B. ein Aktienpaket eines Unternehmens oder eine Währung. Optionen und andere ähnliche Finanzinstrumente werden *Derivate* genannt, da sich ihr Wert von dem Basiswert ableitet.

Wir wollen hier als Beispiel die sogenannten *europäischen Optionen* betrachten, bei denen der Kauf (bzw. Verkauf) zu einem von vornherein festgelegten *Ausübungszeitpunkt* T statt finden muss. Ist dieser Ausübungszeitpunkt T erreicht, so kann der Inhaber die Option ausüben, also den Basiswert zu dem vorher festgelegten Ausübungspreis kaufen (bzw. verkaufen), oder er kann die Option verfallen lassen, den Basiswert also nicht kaufen (bzw. verkaufen). (Alternativ gibt es die *amerikanischen Optionen*, bei denen die Ausübung zu jedem beliebigen Zeitpunkt $t \in [0, T]$ durchgeführt werden kann und die wir hier aus Zeitgründen nicht betrachten werden.)

Das Problem der Optionspreisbewertung ist nun: Welchen Wert besitzt diese Option selbst zu einem Zeitpunkt $t < T$? Diese Frage ist wichtig, wenn z.B. eine Bank eine solche Option ausgibt, oder wenn ein Inhaber einer Option diese auf dem Markt anbieten will. Wir betrachten zunächst die Call-Option. Wir bezeichnen mit S den (bekannten) Basiswert zu einem beliebigen aber festen Zeitpunkt $t \in [0, T]$ und mit $V(S, t)$ den (zu bestimmenden) Wert der Option zu diesem Zeitpunkt t . Es ist relativ einfach, den Wert der Option für $t = T$ zu ermitteln: Sei K der festgelegte Ausübungspreis. Falls $S > K$ ist, so erzielt man durch Ausübung der Option offenbar den Gewinn $V(S, T) = S - K > 0$. Falls $S \leq K$, so würde man durch Ausübung der Option den Basiswert zu einem höheren Preis als dem

aktuellen Marktpreis kaufen, folglich würde man die Option nicht ausüben, der Wert ist also $V(S, T) = 0$. Zusammen erhält man so den Wert

$$V(S, T) = \max\{S - K, 0\} =: (S - K)^+.$$

Gerade umgekehrt verhält es sich bei einer Put-Option. Hier wird man die Option nur dann ausüben, wenn der Marktpreis S unterhalb des Ausübungspreises K liegt. Man erhält so

$$V(S, T) = \max\{K - S, 0\} =: (K - S)^+.$$

Zur Berechnung des Optionswerts zu einem Zeitpunkt $t < T$ benötigt man nun zweierlei:

- (1) Eine Regel zur Berechnung von $V(S, t)$ aus $V(S(T), T)$, falls $S(T)$ bekannt ist
- (2) Eine Schätzung des Basiswertes $S(T)$ zur Zeit T abhängig vom Basiswert S zur Zeit t

Hat man (1) und (2) zur Verfügung, so kann man $V(S(T), T)$ mittels (2) schätzen und die Regel aus (1) auf diesen Schätzwert anwenden.

Zu (1): Wir nehmen an, dass $S(T)$ und damit — gemäß der obigen Formeln — auch $V(S(T), T)$ bekannt ist. Wir schreiben kurz $V(t) = V(S, t)$ und $V(T) = V(S(T), T)$. Dann könnte man einfach $V(t) = V(T)$ setzen. Dies entspricht aber nicht der wirtschaftlichen Sichtweise. Man muss den Wert $V(T)$ nämlich noch *diskontieren*, d.h. mit $e^{-r(T-t)}$ multiplizieren, wobei r die (hier vereinfachend als konstant angenommene) Zinsrate für risikofreie Anlagen ist. Diese Diskontierung ergibt sich aus einer Grundannahme in der mathematischen Modellierung von Finanzmärkten, die *Arbitrage-Freiheit* genannt wird. Arbitrage bezeichnet einen am Finanzmarkt erzielbaren risikofreien Gewinn. Beispielsweise könnte ein Produkt auf zwei verschiedenen Märkten zu verschiedenen Preisen gehandelt werden. Indem man das Produkt auf dem "billigeren" Markt kauft und auf dem "teureren" Markt verkauft, könnte man einen risikofreien Profit erzielen. Das Postulat der Arbitrage-Freiheit beruht nun auf der Annahme, dass sich in einer solchen Situation die Preise auf den beiden Märkten sofort angleichen würden, was den risikofreien Gewinn unmöglich macht. In der Praxis ist die Arbitrage-Freiheit in dieser idealisierten Form natürlich nicht gegeben; sie ist aber als Modellannahme für Finanzmärkte weitgehend akzeptiert. Die oben erwähnte Diskontierung leitet sich nun aus der Arbitrage-Freiheit ab: Man vergleicht den Gewinn aus einer Option mit dem Gewinn, den man bei einer risikofreien Anlage des Kapitals erzielen würde. Würde man das dem Wert $V(t)$ entsprechende Kapital in dieser Form anlegen, so würde man zur Zeit T den Betrag $B(T) = V(t)e^{r(T-t)}$ erhalten. Betrachten wir nun den als bekannt angenommenen Wert $V(T)$ der Option zum Zeitpunkt T . Wäre $V(t) > e^{-r(T-t)}V(T)$, so könnte man die Option sofort zu diesem Preis verkaufen und das Geld risikofrei anlegen; damit würde man zur Zeit T den Betrag

$$B(T) = V(t)e^{r(T-t)} > e^{-r(T-t)}V(T)e^{r(T-t)} = V(T)$$

erhalten und hätte einen risikofreien Gewinn $B(T) - V(T) > 0$ erzielt. Wäre umgekehrt $V(t) < e^{-r(T-t)}V(T)$, so würde man einen Betrag $B(t) = V(t)$ aus dem risikofrei angelegten Geld nehmen, die Option kaufen und zur Zeit T den Betrag

$$B(T) = V(T) > e^{r(T-t)}V(t) = e^{r(T-t)}B(t)$$

erhalten. Wiederum hätte man einen risikofreien Gewinn $B(T) - e^{r(T-t)}B(t) > 0$ erzielt. Da das Postulat der Arbitrage-Freiheit aber risikofreien Gewinn gerade ausschließt, muss also auf Basis dieser Annahme

$$V(t) = e^{-r(T-t)}V(T)$$

gelten.

Zu (2): Hier kommen nun die stochastischen Differentialgleichungen ins Spiel. Man modelliert den typischen Kursverlauf durch eine SDG (4.13) mit Lösung $X(t; t_0, X_0)$ und setzt $S(T) = X(T; t, S)$. Geeignete SDG Modelle werden wir im übernächsten Abschnitt betrachten. Beachte, dass $S(T)$ damit kein fester Wert sondern eine Zufallsvariable ist, der hieraus errechnete ‘‘Schätzwert’’ für $V(S(T), T)$ muss also als Erwartungswert $\mathbb{E}(V(X(T; t, S), T))$ ausgedrückt werden.

Indem wir (1) und (2) zusammenfassen, erhalten wir schließlich die Gleichung

$$V(S, t) = e^{-r(T-t)}\mathbb{E}(V(X(T; t, S), T)), \quad (4.16)$$

mit der der Wert der Option zur Zeit t abhängig vom Basiswert S zur Zeit t berechnet werden kann.

In Verlauf dieses Kapitels werden wir Techniken kennen lernen, um diese Größe zu berechnen. Wir betrachten jetzt zunächst eine Rechenregel für Ito-Integrale und gehen dann kurz auf (sehr einfache) Modelle zur Simulation von Kursverläufen ein. Schließlich kommen wir zu den eigentlichen Berechnungsmethoden, bei denen wir zwei Methoden, nämlich Monte-Carlo-Verfahren und die Black-Scholes-Gleichung, besprechen werden.

4.2.2 Das Ito-Lemma

Zur Bewertung von Optionen, aber auch, um ganz allgemein mit dem Ito-Integral zu rechnen, benötigt man geeignete Rechenregeln. Falls $x(t)$ eine Lösung einer eindimensionalen deterministischen DGL $\dot{x}(t) = f(t, x(t))$ ist, so rechnet man für eine beliebige Funktion $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mittels der Kettenregel leicht nach, dass die Funktion $y(t) = g(t, x(t))$ die Gleichung

$$\dot{y}(t) = \frac{\partial g}{\partial t}(t, x(t)) + \frac{\partial g}{\partial x}(t, x(t))f(t, x(t))$$

erfüllt.

Das Ito-Lemma liefert nun eine Verallgemeinerung dieser Gleichung für die Lösung Ito-stochastischer DGL, und wird daher oft auch als stochastische Verallgemeinerung der Kettenregel bezeichnet.

Lemma 4.2 Sei $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ eine zwei mal stetig differenzierbare Funktion und sei $X(t)$ die Lösung einer reellwertigen Ito-SDG vom Typ (4.13). Dann erfüllt $Y(t) = g(t, X(t))$ die Gleichung

$$\begin{aligned} dY(t) &= \left(\frac{\partial g}{\partial t}(t, X(t)) + \frac{\partial g}{\partial x}(t, X(t))a(t, X(t)) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X(t))b(t, X(t))^2 \right) dt \\ &+ \frac{\partial g}{\partial x}(t, X(t))b(t, X(t))dW_t, \end{aligned}$$

wobei W hier gerade der Wiener Prozess aus der SDG ist, die $X(t)$ erfüllt. Diese Formel wird auch *Ito-Formel* genannt.

Beweisidee: Wir werden hier keinen vollständigen Beweis betrachten, wollen aber begründen, warum die (im Vergleich zur deterministischen Formel ungewöhnliche) zweite Ableitung von g nach x hier auftritt. Wir betrachten dazu zwei Zeitpunkte t und $t + \Delta t$. Dann gilt $X(t + \Delta t) = X(t) + \Delta X(t)$ mit

$$\begin{aligned}\Delta X(t) &= \int_t^{t+\Delta t} a(s, X(s))ds + \int_t^{t+\Delta t} b(s, X(s))dW_s \\ &\approx a(t, X(t))\Delta t + b(t, X(t))\Delta W_t,\end{aligned}\tag{4.17}$$

wobei $\Delta W_t = W(t + \Delta t) - W(t)$ ist. Die Approximation (4.17) folgt hierbei aus der Limes-Definition des Ito-Integrals, wenn wir $\tau_{i+1} - \tau_i = \Delta t$ setzen. Wir betrachten nun die Größe

$$\Delta Y := g(t + \Delta t, X(t) + \Delta X(t)) - g(t, X(t)).$$

Aus der Taylor-Entwicklung von g folgt

$$\begin{aligned}\Delta Y &\approx \frac{\partial g}{\partial x} \Delta X + \frac{\partial g}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (\Delta X)^2 \\ &\approx \frac{\partial g}{\partial x} \left(a(t, X(t))\Delta t + b(t, X(t))\Delta W_t \right) + \frac{\partial g}{\partial t} \Delta t \\ &\quad + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \left(a(t, X(t))^2 (\Delta t)^2 + 2a(t, X(t))b(t, X(t))\Delta t \Delta W_t + (b(t, X(t))\Delta W_t)^2 \right),\end{aligned}$$

wobei alle Ableitungen in $(t, X(t))$ ausgewertet werden. Um aus dieser Approximation die ‘‘Ableitung’’

$$dY(t) = \lim_{\Delta t \rightarrow 0} \Delta Y$$

zu berechnen, müssen wir alle Terme der Ordnung $O(\Delta t)$ berücksichtigen. Wären alle Größen deterministisch, so blieben hier gerade die Terme mit den ersten Ableitungen von g stehen. Im stochastischen Fall ist aber auch $(\Delta W_t)^2$ ein Term der Ordnung $O(\Delta t)$, denn es gilt

$$\mathbb{E}((\Delta W_t)^2) = \mathbb{E}(\underbrace{\mathbb{E}(\Delta W_t)}_{=0} - \Delta W_t)^2 = \text{Var}(W_t) = \text{Var}(W(t + \Delta t) - W(t)) = \Delta t.$$

Der Term

$$\frac{1}{2} \frac{\partial^2 g}{\partial x^2} b(t, X(t))^2 (\Delta W_t)^2$$

muss also mit berücksichtigt werden. Führt man nun den Grenzübergang für $\Delta t \rightarrow 0$ im richtigen stochastischen Sinne durch, so erhält man gerade die behauptete Formel. \square

Bemerkung 4.3 Es existiert auch eine Version des Ito-Lemmas für mehrdimensionale SDGs. In diesem Fall gehen — durch die im Beweis verwendete Taylor-Entwicklung für g — alle gemischten zweiten Ableitungen der Form $\frac{\partial^2 g}{\partial x_i \partial x_j}$ in die Gleichung ein. \square

4.2.3 SDG–Modelle für Kursverläufe

Wir werden in diesem Abschnitt zwei Modelle aus der Finanzmathematik vorstellen, die zur Modellierung von Kursverläufen verwendet werden.

Die absolute Minimalanforderung, die man in dieser Modellierung zur Verfügung haben möchte, sind zwei Parameter: Einen Parameter μ , mit dem ein “Trend” des Kurses nach oben oder unten erzielt werden kann und einen weiteren Parameter σ , mit dem man die stochastische Streuung, also die Varianz des Kurses beeinflussen kann.

Das denkbar einfachste SDG–Modell, das diese Anforderungen erfüllt ist gegeben durch

$$dX(t) = \mu X(t)dt + \sigma X(t)dW_t. \quad (4.18)$$

Die Lösungen dieser Gleichung werden *geometrische Brown'sche Bewegung* genannt. In der Finanzmathematik wird der Parameter $\mu \in \mathbb{R}$ als *Rücklaufquote* oder auch *Rendite* bezeichnet, während $\sigma > 0$ die *Volatilität* genannt wird.

Trotz seiner Einfachheit ist dieses Modell Grundlage vieler Finanzmathematischer Anwendungen in der Modellierung von Kurverläufen, auch wenn es heutzutage immer mehr durch komplexere Modelle ersetzt wird.

Das Schöne an dieser einfachen Gleichung ist die Tatsache, dass man die Lösungen noch per Hand angeben kann. Mit Hilfe des Ito–Lemmas rechnet man nach, dass

$$X(t; X_0) = X_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right) \quad (4.19)$$

die Lösung von (4.18) ist: Wir betrachten die SDG $dZ_t = dW_t$ (also $a \equiv 0$ und $b \equiv 1$) und wenden das Ito–Lemma an auf $Y(t) = g(t, Z(t))$ mit

$$g(t, x) = X_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma x\right).$$

Wegen

$$\frac{\partial g}{\partial t}(t, x) = g(t, x)\left(\mu - \frac{1}{2}\sigma^2\right), \quad \frac{\partial g}{\partial x}(t, x) = g(t, x)\sigma \quad \text{und} \quad \frac{\partial^2 g}{\partial x^2}(t, x) = g(t, x)\sigma^2$$

und $Y(t) = g(t, Z(t)) = X(t)$ (weil $Z(t) = W(t)$ ist) folgt aus dem Ito–Lemma die SDG

$$dX(t) = dY(t) = \left(X(t)\left(\mu - \frac{1}{2}\sigma^2\right) + \frac{1}{2}\sigma^2 X(t)\right) dt + X(t)\sigma dW_t = \mu X(t)dt + \sigma X(t)dW_t,$$

was zeigt, dass (4.19) tatsächlich eine Lösung der SDG (4.18) ist, wegen der Eindeutigkeit der Lösung also die einzige.

Für $\sigma = 0$ erhalten wir aus (4.18) gerade die bekannte lineare DGL $\dot{x}(t) = \mu x(t)$ und ihre zugehörige Lösung $x(t) = x(0)e^{\mu t}$. Durch diese Gleichung ist der Erwartungswert $\mathbb{E}(X(t; X_0))$ bestimmt. Für die Lösungen von (4.18) gilt nämlich

$$\begin{aligned} \mathbb{E}(X(t; X_0)) &= \mathbb{E}(X(0)) + \mathbb{E}\left(\int_0^t \mu X(\tau; X_0)d\tau\right) + \underbrace{\mathbb{E}\left(\int_0^t \sigma X(\tau; X_0)dW_\tau\right)}_{=0, \text{ wegen (4.14)}} \\ &= \mathbb{E}(X(0)) + \int_0^t \mu \mathbb{E}(X(\tau; X_0))d\tau. \end{aligned}$$

Die Funktion $e(t) = \mathbb{E}(X(t, x_0))$ erfüllt also gerade die lineare DGL $\dot{e}(t) = \mu e(t)$ mit Anfangswert $e(0) = \mathbb{E}(X_0) = X_0$, weswegen $\mathbb{E}(X(t, x_0)) = X_0 e^{\mu t}$ ist.

Ebenso wie den Erwartungswert kann man auch die Varianz der Lösungen von (4.18) explizit berechnen: Es gilt $\text{Var}(X(t)) = X_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1)$. Für einen zu modellierenden Kurs können die zwei Parameter üblicherweise aus vergangenen Kursverläufen geschätzt werden, vereinfacht gesagt gibt μ den Trend des bisherigen Kursverlaufs an, während sich σ aus der Streuung der vergangenen Werte ergibt. Dies zeigt wiederum, dass diese Art der Modellierung keine echten Voraussagen über den Kurs zulässt; sie erlaubt nur eine stochastische Modellierung des Verlaufes auf Basis vergangener Werte.

Falls man dieses Modell in der Optionsbewertung einsetzt (also als Modell für (2) in Abschnitt 4.2.1 verwendet) und eine sogenannte *risikoneutrale* Optionsbewertung durchführen will, so setzt man $\mu = r$, wobei r die risikofreie Zinsrate ist. In der Praxis würde man für risikoreiche Anlagen natürlich $\mu > r$ erwarten, damit das Anlagerisiko durch einen höheren Erwartungswert des Gewinns ausgeglichen wird. Trotzdem wird in der risikoneutralen Optionspreisbewertung $\mu = r$ gesetzt; der Grund dafür ist eine weitere Modellannahme, die an die Anlagestrategie des Emittenten (Ausgebers) der Option gemacht wird, der bisher in unserem Modell nicht berücksichtigt wurde: Bisher haben wir angenommen, dass der Emittent mehr oder weniger tatenlos abwartet, bis die Option ausläuft. Der Preis der Option ist aber gerade durch den erwarteten Verlust des Emittenten gegeben. Wenn dieser also seinen Verlust durch eine geeignete Anlagestrategie beeinflussen kann, so beeinflusst dies wiederum den Optionspreis. Wir werden dies später im Rahmen der Black-Scholes Gleichung genauer erläutern, und sehen, warum dies auf die Bedingung $\mu = r$ führt, vgl. Bemerkung 4.4.

Wir wollen uns hier nur kurz auf die Modellierung von Kursverläufen eingehen, da wir noch zur Anwendung dieser Modelle in der Optionspreisbewertung kommen wollen. Um zu zeigen, dass es durchaus noch kompliziertere Modelle als (4.18) gibt, wollen wir hier nur ein mehrdimensionales Modell angeben. Dieses ist gegeben durch den dreidimensionalen Prozess $X(t) = (S(t), \sigma(t), \zeta(t))^T$, der durch die Gleichungen

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sigma(t)S(t)dW_t^1 \\ d\sigma(t) &= -(\sigma - \zeta)dt + \alpha\sigma(t)dW_t^2 \\ d\zeta(t) &= \beta(\sigma(t) - \zeta(t))dt \end{aligned}$$

mit Parametern $\mu \in \mathbb{R}$, $\alpha > 0$ und $\beta \geq 0$ gegeben ist. In der ersten Komponente erkennen wir (4.18) wieder, allerdings ist σ nun kein fester Parameter sondern ebenfalls ein stochastischer Prozess, was die Realität sicherlich deutlich angemessener widerspiegelt. Um zu verstehen, wie sich σ hier in der Zeit verändert, ist es sinnvoll, zunächst die Gleichung für ζ anzuschauen. Für jeden Pfad $\sigma(t, \omega)$ ist dies eine einfach inhomogene lineare DGL, für sie man die exakte Lösung als

$$\zeta(t) = e^{-\beta t} \zeta(0) + \int_0^t e^{-\beta(t-s)} \sigma(s, \omega) ds$$

angeben kann. Der Wert ζ kann also als (durch $e^{-\beta(t-s)}$ gewichteter) Mittelwert für σ aufgefasst werden. Die Gleichung für σ bewirkt dann, dass die Volatilität σ ihrem Mittelwert folgt. Wäre dies eine deterministische Gleichung, so wäre das Verhalten langweilig da (für

$\beta > 0$) sowohl ζ als auch σ gegen Null konvergieren würden. Da die Gleichung für σ aber durch einen zweiten Wiener Prozess W^2 wiederum stochastisch gestört wird, ergibt sich ein durchaus interessantes Verhalten der Gleichung.

4.2.4 Die Monte–Carlo Methode

Wie oben beschrieben, lässt sich das Problem der Optionspreisbewertung auf die Berechnung von $V(t, S)$ mittels (4.16) zurückführen, wobei die Hauptschwierigkeit in der Berechnung des Erwartungswertes $\mathbb{E}(V(X(T; t, S), T))$ steckt.

Eine direkte und sehr vielseitige Methode, um Erwartungswerte komplizierter Ausdrücke zu berechnen ist die Monte–Carlo Methode. Der Name leitet sich (angeblich) von der gleichnamigen Hauptstadt des Fürstentums Monaco ab, genau genommen von dem dortigen Kasino. Ähnlich wie beim Glücksspiel führt man bei der Monte–Carlo Methode Zufallsexperimente durch; statt aber auf Gewinn zu hoffen, berechnet man aus den einzelnen Ergebnissen der Experimente hier eine Schätzung des Erwartungswertes. Die Zufallsexperimente werden hierbei mittels Computersimulationen durchgeführt.

Im Einzelnen läuft die Monte–Carlo Methode wie folgt ab: Gegeben seien eine dem Kursverlauf zugrunde liegende SDG (z.B. (4.18)), eine Anfangszeit t , eine Endzeit T , die risikofreie Zinsrate r und die Funktion $V(S(T), T)$ aus Abschnitt 4.2.1.

- (i) Mittels eines Zufallsgenerators erzeugt man zufällig Pfade $W(t, \omega_i)$, $i = 1, 2, \dots, N$ des Wiener Prozesses W . Genauer erzeugt man Approximationen dieser Pfade, die so konstruiert werden, dass sie den geforderten Bedingungen an W zumindest approximativ genügen.
- (ii) Mittels eines numerischen Verfahrens zur Lösung stochastischer Differentialgleichungen werden dann die zu $W(t, \omega_i)$ gehörigen Pfade $X(\tau, t; S, \omega_i)$ der SDG–Lösung — wiederum approximativ — berechnet⁵. Dies liefert eine Menge von zufälligen Approximationen $\tilde{S}_i(T)$ von $S(T, \omega_i) = X(T, t, S), \omega_i$, $i = 1, 2, \dots, N$.
- (iii) Zur Approximation von $\mathbb{E}(V(S(T), T))$ verwendet man (4.6), d.h. man berechnet

$$\tilde{\mathbb{E}}(V(S(T), T)) = \frac{1}{N} \sum_{i=1}^N V(\tilde{S}_i, T).$$

Dieser Wert $\tilde{\mathbb{E}}(V(S(T), T))$ liefert dann eine Approximation von $\mathbb{E}(V(S(T), T))$.

- (iv) Schließlich berechnet man die Approximation $\tilde{V}(S, t) = e^{-r(T-t)} \tilde{\mathbb{E}}(V(S(T), T))$ von $V(S, t)$.

Wenn man das einfache Modell (4.18) verwendet, so kann man sich die numerische Lösung der SDE sparen und direkt die Formel (4.19) verwenden. Hierfür müssen auch keine kompletten Pfade des Wiener–Prozesses simuliert werden, sondern nur die Werte $W(T, \omega_i)$, also — gemäß der Definition im Abschnitt 4.1.2 — nur $N(0, T)$ –verteilte Zufallszahlen.

⁵Details zur Approximation von W und zur numerischen Lösung von SDGs werden in der parallel stattfindenden Vorlesung “Numerische Mathematik II” behandelt.

Die Monte–Carlo Methode ist aus verschiedenen Gründen in der Anwendung beliebt: Sie ist intuitiv und damit — auch für Nichtmathematiker — sehr einfach zu verstehen, sie erlaubt die Verwendung komplizierter SDGs, allgemeiner “Endwertfunktionen” $V(S(T), T)$ oder zeitvarianter Zinsrate $r(t)$, und die Konvergenz der Methode ist relativ allgemein bewiesen.

Sie hat aber durchaus auch Nachteile: Zum einen konvergiert sie sehr langsam, da für eine genaue Approximation eine große Anzahl N von Simulationen durchgeführt werden muss. Für dieses Problem gibt es Ansätze zur Konvergenzbeschleunigung, die aber das Grundproblem, dass nämlich viele Pfade simuliert werden müssen, nicht prinzipiell beseitigen. Zum anderen liefert sie nur den Optionswert für einen festen Basiswert S und eine feste Zeit t . Will man $V(S, t)$ als Funktion in S und t errechnen, so muss man viele verschiedene Durchläufe der Methode durchführen.

Eine ganz andere Variante zur Bewertung von Optionen stellen Methoden dar, die den Wert mittels partieller Differentialgleichungen berechnen. Eine solche Methode wollen wir im folgenden letzten Abschnitt vorstellen.

4.2.5 Die Black–Scholes–Gleichung

Fisher Black⁶ und Myron Scholes⁷ sowie Robert Merton⁸ haben die nach den ersten beiden Autoren benannte partielle Differentialgleichung gleichzeitig in der ersten Hälfte der 1970er Jahre entwickelt. Merton und Scholes bekamen für ihre Arbeiten auf dem Gebiet der Finanzmathematik 1973 den Nobel–Preis für Wirtschaftswissenschaften (Black war 1995 gestorben).

Wir wollen hier sowohl die Herleitung dieser Gleichung als auch ihre explizite Lösung angeben. Ziel ist die Berechnung der Funktion $V(S, t)$, die hier als Funktion in den zwei reellen Variablen S und t aufgefasst wird, also $V : \mathbb{R}_0^+ \times [0, T] \rightarrow \mathbb{R}$. Wir erinnern daran, dass der Wert $V(S, T)$ dieser Funktion aus Abschnitt 4.2.1 bekannt ist.

Wir fassen zunächst die Grundannahmen für die Black–Scholes Gleichung zusammen:

- (a) Es gibt keine Transaktionskosten und Gebühren, ebenso hat das eigene Kaufverhalten keinen Einfluss auf den Markt.
- (b) Der Markt ist Arbitrage–frei.
- (c) Der Kursverlauf ist durch (4.18) beschrieben.
- (d) Die risikofreie Zinsrate r ist konstant und im betrachteten Zeitraum werden keine Dividenden gezahlt.
- (e) Es wird eine risikoneutrale Optionsbewertung durchgeführt.

Beachte, dass wir die meisten dieser Annahmen bereits bisher — explizit oder implizit — gemacht haben. Ausnahmen bilden die Annahmen (c) und (e), die wir zwar diskutiert aber bisher nicht zwingend vorausgesetzt hatten (was Annahme (e) bedeutet werden wir im

⁶US–amerikanischer Wirtschaftswissenschaftler, 1938–1995

⁷kanadischer Wirtschaftswissenschaftler, *1941

⁸US–amerikanischer Mathematiker und Wirtschaftswissenschaftler, *1944

Laufe dieses Abschnitts erläutern). Viele dieser Annahmen können verallgemeinert werden, führen dann aber auf (zum Teil deutlich) kompliziertere Gleichungen.

Die Idee der Black–Scholes Gleichung besteht darin, eine Gleichung für V herleiten, die nicht explizit von den Lösungen $X(t)$ von 4.18 abhängt. Diese Herleitung werden wir nun durchführen.

Zunächst wenden wir das Ito–Lemma 4.2 auf $V(X(t), t)$ an. Um die Schreibweise zu verkürzen, lassen wir die Argumente der Funktionen weg. Damit erhalten wir

$$dV = \left(\mu X \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 X^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma X \frac{\partial V}{\partial S} dW.$$

$V(S, t)$ ist hier eine deterministische Funktion in den unabhängigen Variablen S und t , W und X hingegen sind stochastische Prozesse. Um die stochastischen Terme zu eliminieren, definieren wir die Größe

$$\Pi(V, S) := -V + \Delta \cdot S.$$

Die Größe Π stellt den Wert eines *Portfolios* dar, d.h., den Wert einer Kombination verschiedener Wertpapiere. Hierbei geht V negativ ein, Π stellt also den Wert eines Portfolios des *Emittenten* (*Ausgebers*) der Option dar, der Wert $-V$ wird auch als Wert der *emittierten* Option bezeichnet. Zudem enthält das Portfolio Δ Einheiten des Basiswertes, wobei Δ eine konstante reelle Zahl ist. Da Π eine Linearkombination aus V und S ist, erhalten wir für $\Pi(V, X)$ die Gleichung $d\Pi = -dV + \Delta dX$, also

$$d\Pi = - \left(\mu X \left(\frac{\partial V}{\partial S} - \Delta \right) + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 X^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma X \left(\frac{\partial V}{\partial S} - \Delta \right) dW.$$

Wir betrachten nun einen festen Punkt $(S, t) \in \mathbb{R} \times \mathbb{R}$, an dem wir V charakterisieren wollen und machen die folgende weitere Modellannahme: Ziel des Emittenten ist es nun, gerade eine so große Anzahl Δ von Anteilen zu halten, dass das stochastische Risiko in der Gleichung für $d\Pi$ eliminiert wird. Man sagt, der Emittent verfolgt eine *risikoneutrale* Anlagestrategie

Dazu wählen wir für den Prozess X die Anfangszeit t und den Anfangswert S , so dass sich in der obigen Gleichung $X = X(t; t, S) = S$ ergibt. Wenn wir nun Δ als

$$\Delta = \frac{\partial V}{\partial S}$$

wählen, so erhalten wir

$$d\Pi = - \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \tag{4.20}$$

also eine deterministische Gleichung. Wir haben also die stochastischen Variablen gegen die unbekante Größe $d\Pi$ vertauscht.

Um dieses $d\Pi$ aus (4.20) wieder zu eliminieren, müssen wir die Arbitrage–Freiheit verwenden. Die Größe $d\Pi$ stellt gerade die Zunahme des Wertes des Portfolios, also den “infinitesimalen Gewinn” zur Zeit t dar. Wenn der Wert Π des Portfolios risikofrei angelegt wäre, so würde es sich für $\Delta t > 0$ wie $\Pi(t + \Delta t) = \Pi(t)e^{r\Delta t}$ verhalten. In t genügt Π also der

Differentialgleichung $\dot{\Pi}(t) = r\Pi(t)$. Wenn wir dies in SDG-Form schreiben und die Wahl von Δ sowie die Definition von Π einsetzen, so erhalten wir

$$d\Pi = r\Pi dt = \left(-rV + rS \frac{\partial V}{\partial S} \right) dt. \quad (4.21)$$

Nun haben wir zwei Gleichungen für $d\Pi$, nämlich gerade (4.20) und (4.21), die beide *nicht* vom Zufall abhängen, also einen garantierten Gewinn beschreiben. Aufgrund der Arbitrage-Freiheit müssen die beiden Gewinnzunahmen $d\Pi$, die durch (4.20) und (4.21) beschrieben werden also gleich sein, denn sonst könnte man ganz analog zu der Argumentation auf Seite 73 (dort für einen längeren Zeithorizont, hier nun für einen infinitesimal kleinen Zeithorizont) durch entsprechenden Wechsel der Anlagestrategie einen risikofreien Gewinn erzielen. Die rechten Seiten der Gleichungen (4.20) und (4.21) müssen also gleich sein, woraus die Gleichung

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV + rS \frac{\partial V}{\partial S} = 0 \quad (4.22)$$

folgt. Dies ist gerade die Black-Scholes Gleichung.

Die Black-Scholes Gleichung (4.22) ist also eine partielle Differentialgleichung, die von V erfüllt wird. Die Gleichung (4.22) hat (genau wie eine gewöhnliche Differentialgleichung ohne Anfangsbedingung) unendlich viele Lösungen, wenn wir aber geeignete *Randwerte* $V(S, T)$, für ein $T > 0$ und alle $S \in [0, \infty)$ festlegen, so erhalten wir eine eindeutige Lösung. Dies ist hier einfach, da sich die Randwerte in natürlicher Weise aus den Optionswerten zum Fälligkeitszeitpunkt $t = T$ ergeben.

Tatsächlich ist Gleichung (4.22) explizit lösbar. Für eine europäische Call-Option gilt

$$V(S, t) = SP(a) - Ke^{-r(T-t)}P(b),$$

während sich für die europäische Put-Option

$$V(S, t) = S(P(a) - 1) - Ke^{-r(T-t)}(P(b) - 1)$$

ergibt. Hierbei sind in beiden Fällen

$$a = \frac{\log \frac{S}{K} + \left(r + \frac{\sigma^2}{2} \right) (T - t)}{\sigma \sqrt{T - t}}, \quad b = a - \sigma \sqrt{T - t}$$

und die Funktion P ist die Verteilungsfunktion der Standard-Normalverteilung gegeben durch

$$P(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx,$$

vgl. (4.5).

Abbildung 4.4 zeigt Lösungen für den europäischen Call mit $r = 0.06$, $K = 1$, $T = 1$ und $\sigma = 0.3$ bzw. $\sigma = 1$.

Wenn man die vereinfachenden Annahmen abschwächt oder wenn man amerikanische an Stelle von europäischen Optionen betrachtet, so erhält man deutlich kompliziertere partielle Differentialgleichungen von ähnlichem Typ wie (4.22), die dann nur noch numerisch gelöst werden können.

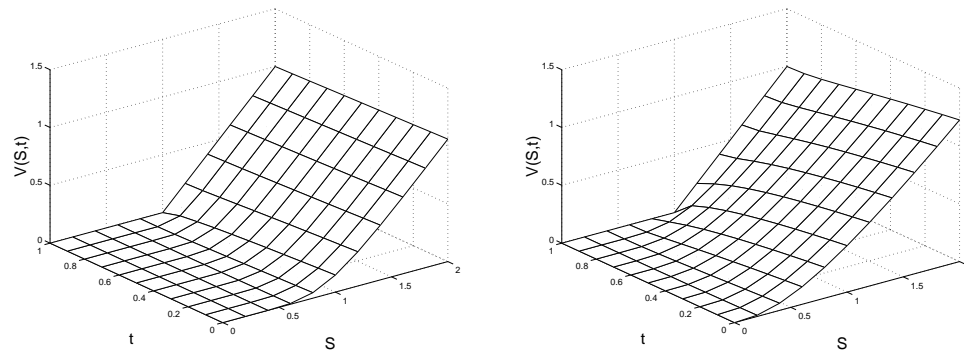


Abbildung 4.4: Lösung der Black–Scholes Gleichung für den europäischen Call mit $r = 0.06$, $K = 1$, $T = 1$ und $\sigma = 0.3$ (links) bzw. $\sigma = 1$ (rechts)

Bemerkung 4.4 Die Annahme, dass der Emittent der Option eine risikoneutrale Anlagestrategie verfolgt ist wesentlich für den resultierenden Optionspreis. Der so hergeleitete Preisberechnung wird deswegen auch *risikoneutrale Optionsbewertung* genannt. Dies erklärt auch die Wahl $\mu = r$ in dem SDE–Kursmodell (4.18), falls dieses (z.B. mittels Monte–Carlo–Simulationen) direkt zur risikoneutralen Optionspreisbewertung verwendet wird: Mit dieser Wahl ergibt die Bewertung mittels Simulation von (4.18) genau die gleichen Werte wie die Black–Scholes Gleichung. Im Modell entspricht die Lösung der SDG dann tatsächlich nicht mehr dem erwarteten Verlauf des Kurses sondern steht statt dessen in Beziehung zum Wert des Portfolios des Emittenten.

Dies bedeutet insbesondere, dass die Black–Scholes Gleichung tatsächlich viel mehr als nur den Optionspreis liefert: Sie liefert auch eine Regel für die risikoneutrale Anlagestrategie (auch “Hedging–Strategie” genannt) des Emittenten. \square

4.3 Literaturhinweise

Dieses Kapitel wurde zusammengestellt auf Basis der Bücher

P. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations, Springer–Verlag, 1999 (3. Auflage)

und

R. Seydel, Tools for Computational Finance, Springer–Verlag, 2002,

von dem es auch die deutsche Ausgabe

R. Seydel, Einführung in die numerische Berechnung von Finanz–Derivaten, Springer–Verlag, 2000

gibt (die englische Ausgabe ist eine etwas erweiterte Übersetzung der deutschen Ausgabe). Diese Bücher sind — wie die Titel bereits andeuten — numerisch orientiert, liefern aber auch eine knappe Darstellungen der zu Grunde liegenden Theorie, auf die hier zurück gegriffen wurden.

Index

- Anfangsbedingung, 4
- Anfangswert, 4
- Anfangswertproblem, 3
- Anfangszeit, 4
- angeregtes Pendel, 51
- Antistabilität, 13
- Antisymmetrie, 27
- Arbitrage, 73
- Attraktor, 53
- autonome Differentialgleichung, 3

- Banach'scher Fixpunktsatz, 5
- Bewegungselement, 37
- Bewegungsgleichung, 37
- Black–Scholes Gleichung, 81
- Brown'sche Bewegung, 67
 - geometrische, 76

- Call–Option, 72
- charakteristische Funktion, 64

- Dämpfer, 39
- Dämpfungselement, 39, 47
- Delay–Differentialgleichung, 33
- Dichtefunktion, 63
- Differentialgleichung, 2
- Differenzgleichung, 9
- Diffusion, 71
- diskontieren, 73
- Dissipation, 57
- Dissipator, 39
- Drehmoment, 45
- Drift, 71
- drittes Newton'sches Gesetz, 40

- Eigenwert, 14
- Eindeutigkeitssatz, 4
- Einzugsbereich, 15
- Elastizitätselement, 38, 47
- Elementarereignis, 62

- Energie, 38, 39
- Epidemien, 32
- Equilibrium, 12
- Erwartungswert, 65
- Existenzintervall, 4
- Existenzsatz, 4
- exponentielle Antistabilität, 13
- exponentielle Instabilität, 13
- exponentielle Stabilität, 13
- exponentielles Wachstum, 11

- Fahrzeugfederung, 43
- Fangstrategie, 17
- fast sicher, 64
- Feder, 38
- Federung, 43
- Fischfang, 16, 28

- Gauß–Verteilung, 64
- Geburtenrate, 11
- geometrische Brown'sche Bewegung, 76
- gewöhnliche Differentialgleichung, 2
- Gleichgewicht, 12
- Gompertz–Wachstum, 16
- grafische Darstellung
 - als Graph, 7
 - als Kurve, 7

- Hamilton–Funktion, 58
- Hamilton–System, 60
- Hook'sches Gesetz, 38

- Infektionsstadium, 32
- Inkrement des Wiener Prozesses, 67
- instabile Mannigfaltigkeit, 21
- Instabilität, 13
- invariante Menge, 22
- Invarianzprinzip von Lasalle, 26
- Ito–Integral, 70
- Ito–Lemma, 74

- Jacobi-Matrix, 14
- Kapazität eines Lebensraums, 11
- Kartoffelfäule, 32
- Konstante der Bewegung, 22
- Koordinatentransformation, 20, 24
- Kozykluseigenschaft, 6
- Kraftvektor, 45
- Kursverlaufsmodelle, 76
- Lagrange-Gleichungen, 56
 - verallgemeinert, 58
- Lasalles Invarianzprinzip, 26
- Latenzstadium, 32
- Legendre-Transformation, 58
- Linearisierung, 14
- logistisches Wachstum, 12
- Lösungskurve, 3
- Lösungstrajektorie, 3
- Lotka-Volterra Modell, 20
- Mannigfaltigkeit
 - der verträglichen Konfigurationen, 54
 - instabile, 21
 - stabile, 21
- Masse, 38
- Migrationsrate, 11
- Monte-Carlo Methode, 78
- n-Arten Modell, 27
- Newton'sches Gesetz
 - drittes, 40
 - zweites, 38
- Normalverteilung, 64
- optimale Fangstrategie, 17
- Option, 72
 - amerikanische, 72
 - europäischer Call, 72
 - europäischer Put, 72
- Optionsbewertung, 72
 - risikoneutral, 77, 79, 82
- Oszillator, 40
- Parallelachsensatz, 46
- Pendel, 48
 - angeregt, 51
- periodische Lösungen, 22
- Pfad eines stochastischen Prozesses, 67
- Populationsdynamik, 9
- Portfolio, 80
- Put-Option, 72
- Ragleigh'sche Dissipationsfunktion, 57
- Randwert, 81
- Räuber-Beute Modell
 - beschränkte Ressourcen, 24
 - unbeschränkte Ressourcen, 20
- Raubfischanteil, 28
- Realisierung, 63
- Reibung, 39
- Rotationsdämpfer, 47
- Rücklaufquote, 76
- Ruhelage, 12
- Schäfers Modell, 16
- Schneiden von Lösungen, 6, 7
- Schwerkraft, 42
- Schwinger, 40
- Selektion gleichartiger Spezies, 29
- stabile Mannigfaltigkeit, 21
- Stabilität, 13
- Standard-Normalverteilung, 64
- Steiner'scher Satz, 46
- Sterberate, 11
- stochastische Differentialgleichung, 62
 - Ito, 70
 - Stratonovich, 71
- stochastischer Prozess, 67
- Strömungsreibung, 40
- Torsionsfeder, 47
- Trägheitselement, 38, 45
- Trägheitsmoment, 45
- Trajektorie, 3
- translationale Modelle, 40
- trockene Reibung, 40
- Unabhängigkeit von Zufallsvariablen, 66
- Varianz, 66
- Vektorfeld, 3
- verallgemeinerte Geschwindigkeiten, 55
- verallgemeinerte Koordinaten, 55
- verallgemeinerte Kräfte, 55
- verallgemeinertes Moment, 58

- Verteilungsfunktion, 64
- verträgliche Konfigurationen, 54
- verzögerte Differentialgleichung, 33
- viskose Reibung, 40
- Volatilität, 76
- Volterra-Ökologie, 27

- Wachstum
 - exponentiell, 11
 - Gompertz, 16
 - logistisch, 12
- Wachstumsrate, 11
- Wahrscheinlichkeitsmaß, 62
- Wahrscheinlichkeitsraum, 62
- Wahrscheinlichkeitsverteilung, 63
- Weltbevölkerungswachstum, 11
- Wiener Prozess, 67
- Wirkungsfunktion, 56
- Würfelmodell
 - ein Würfel, 62
 - zwei Würfel, 66

- Zufallsvariable, 63
 - diskret, 63
 - Gauß-verteilt, 64
 - kontinuierlich, 63
 - normalverteilt, 64
- zweites Newton'sches Gesetz, 38