

Grundlagen physikalischer Prozesse

Vorlesung WS 2010/2011

Bernd Thaller

Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstraße 36
A-8010 Graz
mailto:bernd.thaller@uni-graz.at
<http://math.uni-graz.at/thaller/>

Download Link: <http://math.uni-graz.at/thaller/lehre/mpt/skriptum.pdf>

12. Januar 2011

Inhaltsverzeichnis

Kapitel 1. Einleitung	1
1.1. Mathematische Modellierung	1
1.2. Ein Beispiel	3
Kapitel 2. Dimensionsanalyse	7
2.1. Maßzahlen und Dimensionen	7
2.2. Dimensionsanalyse physikalischer Zusammenhänge	13
Kapitel 3. Einige Grundsätze der Mechanik	29
3.1. Kraft und Energie	29
3.2. Die Newton'schen Bewegungsgleichungen	30
3.3. Wie man physikalische Konstanten los wird	31
3.4. Potenzielle Energie und Arbeit im Kraftfeld	33
Kapitel 4. Mechanische Systeme und Gleichgewichtslagen	37
4.1. Durch Federkräfte verbundene Massen	37
4.2. Die potenzielle Energie des Federsystems	40
4.3. Das Minimum der potenziellen Energie	42
4.4. Gleichgewicht kontinuierlicher Systeme - ein elastischer Stab	44
Kapitel 5. Lineare Widerstandsnetzwerke	53
5.1. Grundgesetze	53
5.2. Beispiel und Theorie	54
5.3. Ein äquivalentes Minimierungsproblem	59
5.4. Minimierung unter Nebenbedingungen	62
5.5. Das duale Problem	64
Kapitel 6. Differentialgleichungen und Eigenwertproblem	67
6.1. Einleitung	67
6.2. Die einfachste Gleichung	68
6.3. Die logistische Gleichung	69
6.4. Zinsrechnung	70
6.5. Ein Gleichungssystem	72
6.6. Systeme zweiter Ordnung und Schwingungen	76

6.7. Gedämpfte Schwingungen	80
-----------------------------	----

Kapitel 1

Einleitung

Diese Vorlesung behandelt die physikalischen Grundlagen mathematischer Modellbildung. Sie ist gedacht für Studierende der Mathematik etwa ab dem dritten Semester. Das erste Kapitel sollte allerdings auch auf den Niveau der Sekundarstufe 2 zugänglich sein. Vom Umfang her entspricht der Text etwa dem, was in einer dreistündigen Vorlesung in einem Semester behandelt werden kann.

1.1. Mathematische Modellierung

1.1.1. Was ist ein Modell?

Ganz allgemein ist ein *Modell* ein vereinfachtes und idealisiertes Abbild eines realen Systems, das mit diesem wichtige Eigenschaften gemeinsam hat, und zu einem bestimmten *Zweck* entworfen wurde. Der Zweck des Modells ist meist die Beantwortung einer konkreten Frage über das reale System. Man hofft, dass man durch die Untersuchung des Modells Aussagen über Eigenschaften des realen Systems bekommt und sich so oft teure und schwierige Experimente und Messungen am realen System ersparen kann.

Der Modellierungszweck bestimmt unter anderem die Komplexität des Modells, wobei folgende Prinzipien zu beachten sind:

- Das Modell muss alle Eigenschaften, die zur Klärung der Frage notwendig sind, einschließen.
- Das Modell soll keine zur Beantwortung der Frage unnötigen Eigenschaften enthalten.

Ein zu einfaches Modell verfügt möglicherweise nicht über die zur Erklärung eines Effektes nötigen Eigenschaften. Ein unnötig kompliziertes Modell ist möglicherweise zu schwierig zu handhaben oder verschleiert gar die eigentlich interessierenden Zusammenhänge.

Manchmal wird der Modellbegriff sehr weit gefasst und umfasst zum Beispiel *Verbalmodelle*, wie einen Zeitungsartikel, der ein Ereignis beschreibt, oder Landkarten, Fotos, und Listen von Messresultaten. Eine Crash-Test Puppe oder ein Modellflugzeug sind Beispiele für sogenannte *Realmodelle*. Uns interessieren hier aber ausschließlich *mathematische Modelle*.

Mathematische Modelle klassifiziert man oft in

- Beschreibende Modelle (= deskriptive Modelle, z.B. das Fitten von Messwerten durch eine Kurve)
- Erklärende Modelle (beschreiben kausale Zusammenhänge)

oder in

- Qualitative Modelle (beschreiben das Verhalten und die Eigenschaften ohne numerisch nachprüfbare Resultate)
- Quantitative Modelle (sagen die Werte von Messgrößen vorher, machen numerische Angaben)

Sehr oft besteht ein mathematisches Modell aus einem mathematischen Gleichungssystem, dessen Lösungen das Verhalten des realen Systems beschreiben sollen. Nicht nur physikalisch-technische Systeme sind auf diese Art der mathematischen Modellierung zugänglich, sondern auch biologische, physiologische, und ökologische Systeme, sowie Systeme der Ökonomie — um nur einige Bereiche zu nennen, in denen heutzutage die mathematische Modellierung eine große Rolle spielt.

1.1.2. Der Modellierungszyklus

Der Ausgangspunkt für die mathematische Modellierung ist oft nicht ein bestimmtes konkretes System sondern ein “typisches System”, bei dem man bereits von einigen konkreten Eigenschaften, die im Moment nicht interessant sind, absieht. Will man untersuchen, wie lange es braucht, bis ein aus dem Fenster geworfener Stein unten aufschlägt, ist es ziemlich uninteressant, um welchen konkreten Stein es sich handelt und wie seine genaue Zusammensetzung ist. Wir werden da eher einen “typischen Stein” im Sinn haben. Wir bezeichnen so ein zu beschreibendes typisches System als *Prototyp*.

Wenn wir so ein mathematisches Modell haben, können wir Fragen über den Prototyp in eine mathematische Frage betreffend das Gleichungssystem übersetzen. Die Berechnung (Lösung der Gleichungen) liefert die Antwort auf die mathematische Frage. Diese Antwort muss in der Folge interpretiert werden (“Was sagt die mathematische Lösung über die Wirklichkeit aus?”). Dann können wir die vorgeschlagene Lösung des Ausgangsproblems am Prototypen (dem realen System) testen, um das Modell zu verifizieren oder zu falsifizieren. Eventuelle Wiederholungen dieses Prozesses dienen zur Anpassung und Verbesserung des Modells. Diese hier stark vereinfachend dargestellte Prozedur nennt man den *Modellierungszyklus*. Er ist schematisch in der Abbildung 1.1 dargestellt.

Ein ganz ähnliches Diagramm könnte man auch für das Aufstellen einer physikalische Theorie (wie zB Mechanik, Elektrodynamik, Relativitätstheorie, ...) verwenden. Als Prototyp einer physikalischen Theorie dient ganz allgemein ein Teilbereich der beobachtbaren Wirklichkeit, repräsentiert durch

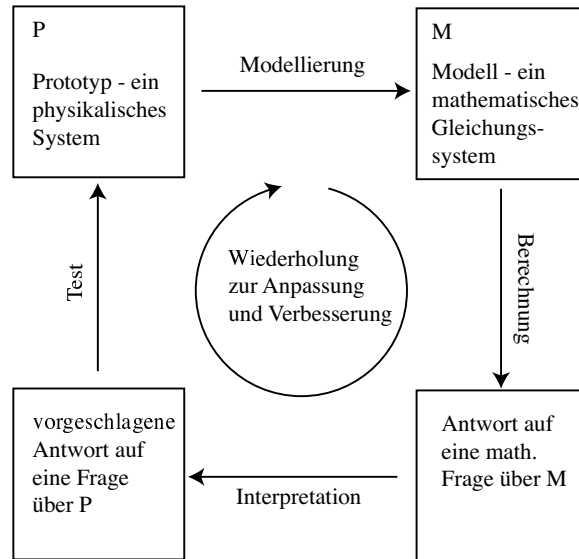


ABBILDUNG 1.1. Der Modellierungszyklus (schematisch)

eine umfangreiche Sammlung von genauen Beobachtungen, experimentellen Ergebnissen und Messresultaten. Das mathematische Modell ist meist kein einzelnes Gleichungssystem, sondern ein ganzes Gebäude aus mathematischen Definitionen und Sätzen. Im Rahmen mathematischer Modellierung dienen physikalische Theorien vor allem dazu, die Vielfalt denk möglicher Modelle für ein Problem einzuschränken. Die physikalische Theorie ist somit eine Richtlinie zur Aufstellung und Interpretation einer konkreten mathematischen Modellgleichung.

1.2. Ein Beispiel

Um den abstrakten Modellierungszyklus durch ein konkretes Beispiel zu erläutern, betrachten wir zunächst ein ganz einfaches System, das wir alle kennen.

Prototyp: Ein fallender Stein.

Frage: Wie lange dauert der Fall aus einer bestimmten Höhe h ?

Modell: (Es stammt im Prinzip von Galileo Galilei). Die Höhe $z(t)$ über dem Boden zur Zeit t wird durch eine Differentialgleichung beschrieben:

$$\frac{d^2 z(t)}{dt^2} = -g. \quad (1.1)$$

Hier ist t die Zeit seit dem Loslassen, g eine Konstante. Dazu gehören die Anfangsbedingungen:

$$z(0) = h, \quad \left. \frac{dz(t)}{dt} \right|_{t=0} = z'(0) = 0. \quad (1.2)$$

(Anfangshöhe = h , Anfangsgeschwindigkeit = 0).

Dieses Modell beinhaltet natürlich viele Vereinfachungen (idealisierende Modellannahmen):

- genau senkrechter Fall, eine ev. Seitwärtsbewegung wird vernachlässigt,
- kein Luftwiderstand, Gravitation ist einzige Kraft auf den Stein,
- keine Höhenabhängigkeit der Gravitation,
- kein Einfluss des Gebäudes oder des Mondes,
- möglicherweise wurden bis dato unbekannte Einflüsse vernachlässigt.

Diese Vereinfachungen sind durchaus praktisch. Wenn das Modell genauso kompliziert ist, wie die Wirklichkeit, ist es nicht nützlich. Ein Modell soll ja nur so kompliziert sein, wie es zur Beantwortung der Frage sein muss.

Mathematische Frage an das Modell:

Bei welchem Wert von t ist $z(t) = 0$?

Lösung: Durch zweimalige Integration:

$$\begin{aligned} \frac{d^2 z(t)}{dt^2} &= -g, \\ \frac{dz(t)}{dt} &= -gt + z'(0), \\ z(t) &= -\frac{1}{2} g t^2 + z'(0) t + z(0). \end{aligned} \quad (1.3)$$

Berücksichtigung der Anfangsbedingungen führt auf

$$z(t) = -\frac{1}{2} g t^2 + h. \quad (1.4)$$

Die Antwort auf unsere mathematische Frage lautet also:

$z(t) = 0$, falls $\frac{1}{2} g t^2 = h$, oder falls

$$t = \pm \sqrt{2 \frac{h}{g}}. \quad (1.5)$$

Interpretation: Wir können das Minus in (1.5) ausschließen (der Stein kommt später am Boden an, als er ausgelassen wird). Das Minus ist deshalb nicht sinnlos, wie sich durch Betrachtung anderer Anfangsbedingungen oder Fragestellungen zeigt. Zur Beantwortung unserer Frage über den fallenden Stein

wird aber die Lösung mit dem Minus nicht gebraucht. Die genauere Frage für unser Problem lautet also: Bei welchem positiven Wert von t ist $z(t) = 0$, und die mathematische Antwort lautet $t = \sqrt{2h/g}$.

Test: Messungen der Fallzeit zeigen, dass das Modell für Steine sehr gut ist, wenn man $g = 9,81\text{m/s}^2$ setzt. Für Herbstblätter, die vom Baum fallen, ist das Modell aber sehr schlecht.

AUFGABE 1.1. Mit welcher Geschwindigkeit muss man einen Stein senkrecht in die Luft werfen, damit er eine maximale Höhe von 20 m erreicht?

Kapitel 2

Dimensionsanalyse

2.1. Maßzahlen und Dimensionen

Die angewandte Mathematik unterscheidet sich von der reinen Mathematik vor allem dadurch, dass die betrachteten Größen durch Einheiten charakterisiert werden. Diese Einheiten beschreiben die Beziehung zwischen den Zahlenwerten, die man durch Rechnung bekommt, und den beobachtbaren Größen in der Wirklichkeit.

2.1.1. Physikalische Größen, Maßzahl, Maßeinheit

Eine physikalische Messgröße A besitzt (zumindest im Rahmen der klassischen Physik) einen Zahlenwert $m(A)$. Das ist eine in der Regel als reell angenommene Zahl, die man durch Ablesung eines Messgerätes erhält. Wir nennen $m(A)$ daher die Maßzahl oder den Messwert der Größe A . Zum Unterschied von einer rein mathematischen Rechengröße gehört zur vollständigen Angabe einer physikalischen Größe noch eine Maßeinheit, die wir hier mit dem Symbol $[A]$ bezeichnen wollen. Wir notieren eine physikalische Größe in der Form „Maßzahl mal Maßeinheit“, also

$$A = m(A) [A]. \quad (2.1)$$

Die Maßeinheit $[A]$ ist ein Symbol, das dem Messwert $m(A)$ für die Größe A angeheftet wird, um damit das Messverfahren zu beschreiben, mit dem die Maßzahl gewonnen wurde. Erst diese Angabe ermöglicht den Vergleich der Zahlenwerte physikalischer Größen mit Erscheinungen in der Natur.

Für den Umgang mit Einheiten gilt folgende Grundregel: Wenn sich eine physikalische Größe als Produkt anderer physikalischer Größen schreiben lässt (zum Beispiel definiert man die Kraft als Masse mal Beschleunigung), so ist die Einheit der Produktgröße das Produkt der einzelnen Maßeinheiten. Allgemeiner seien A_i physikalische Größen und λ_i Zahlen. Dann ist auch das Potenzprodukt

$$A = \prod_{i=1}^n A_i^{\lambda_i}$$

eine physikalische Größe und ihre Einheit ist

$$[A] = \prod_{i=1}^n [A]^{\lambda_i}.$$

Für die Maßzahl gilt natürlich derselbe Zusammenhang:

$$m(A) = \prod_{i=1}^n m(A)^{\lambda_i}.$$

Wir wollen die Zahlen λ_i nicht von vornherein einschränken und erlauben beliebige reelle Zahlen, auch wenn in der Praxis nur rationale Zahlen als Potenzen physikalischer Größen vorkommen. Auch muss man zum Beispiel darauf achten, dass Wurzeln ($\lambda = 1/2$) nur aus Größen gezogen werden, deren Maßzahlen nichtnegativ sind.

Wie gewohnt gilt $[A^0] = [A]^0 = 1$.

Man kann sich denken, dass physikalische Größen auch durch Addition kombiniert werden können. Allerdings muss man beachten:

Die Ausdrücke $A_1 + A_2$ oder $A_1 - A_2$ können nur dann physikalische Größen darstellen, wenn $[A_1] = [A_2]$ gilt.

Summen von Größen mit unterschiedlichen Einheiten haben keine sinnvolle Maßeinheit und werden daher aus physikalischen Betrachtungen ausgeschlossen. Denn was soll es zB. bedeuten, wenn man eine Größe mit der Einheit Meter (m) zu einer Größe mit der Einheit Sekunde s addiert? Aus demselben Grund kann man keine Polynome oder Potenzreihen akzeptieren: Was ist $m + m^2 + m^3 \dots$? (Was soll Länge + Fläche + Volumen +.. darstellen?) Aus diesem Grund kann man auch Funktionen wie zum Beispiel $\sin(1\text{kg})$ keine Einheit auf konsistente Weise zuordnen. Argumente von Funktionen, müssen daher einheitenfrei sein.

Beispiel: Eine Sinuswelle wird durch den mathematischen Ausdruck

$$A \sin(kx - \omega t)$$

beschrieben. Dabei ist die Amplitude A eine Größe mit Einheit (zum Beispiel ein Druck, wenn es sich um eine Druckwelle handelt, oder eine Länge, wenn die Höhe einer Wasserwelle beschrieben werden soll). Die Größen im Argument der Sinusfunktion bedeuten folgendes:

- x ist die Ortskoordinate, ihre Einheit ist Meter m.
- k ist die Wellenzahl, die Wellenlänge ist durch $\lambda = 2\pi/k$ gegeben. Da die Wellenlänge die Einheit m hat, hat die Wellenzahl die Einheit m^{-1} .
- t ist die Zeitkoordinate, ihre Einheit ist Sekunde s.

- ω nennt man Kreisfrequenz. Die Schwingungsdauer der Welle ist durch $T = 2\pi/\omega$ gegeben. Die Kreisfrequenz hat also die Einheit s^{-1} .

Die Kombination kx ist einheitenlos, denn $[kx] = \text{m}^{-1}\text{m} = \text{m}^0 = 1$. Analog dazu ist ωt einheitenlos. Das Argument der Sinusfunktion ist also die Summe von einheitenlosen Größen und selbst einheitenlos.

Zusammenfassung der Rechenregeln für Größen mit Einheiten

- In einer Gleichung müssen links und rechts Größen mit der gleichen Einheit stehen
- Nur Größen mit gleicher Einheit dürfen addiert (oder subtrahiert) werden. Die Summe hat die gleiche Einheit wie jeder Summand.
- Größen dürfen uneingeschränkt miteinander multipliziert oder durcheinander dividiert werden. Das Produkt der Größen hat als Einheit das Produkt der Einheiten der einzelnen Faktoren.
- Differenziert man eine Größe nach einer Variablen, so ist die Einheit der Ableitung der Quotient der Einheiten der abgeleiteten Größe und der Variablen.
- Integriert man eine Größe über eine Variable, so ist die Einheit des Integrals das Produkt der Einheiten des Integranden und der Variablen.
- Nur einheitenlose Größen dürfen als Argumente von Exponentialfunktion, Logarithmus, Winkelfunktionen, etc., eingesetzt werden.

AUFGABE 2.1. *Bei einer Welle gibt es einen Zusammenhang zwischen der Wellengeschwindigkeit c , der Wellenlänge λ und der Frequenz ν . Welche der folgenden Formeln ist richtig?*

- (1) $c = \lambda/\nu$, (2) $c = \nu/\lambda$, oder (3) $c = \lambda\nu$.

2.1.2. Grundgrößen und Dimension

Innerhalb eines bestimmten Teilgebiets der Physik zeichnet man einige (willkürlich gewählte) physikalische Größen G_1, G_2, \dots, G_r als *Grundgrößen* aus. Die Grundgrößen müssen so gewählt werden, dass kein nichttriviales Potenzprodukt der Grundgrößen einheitenlos ist:

$$[G_1^{\alpha_1} G_2^{\alpha_2} \cdots G_r^{\alpha_r}] = 1 \quad \Leftrightarrow \quad \alpha_1 = \alpha_2 = \dots = \alpha_r = 0. \quad (2.2)$$

(Insbesondere darf keine der Grundgrößen selbst einheitenlos sein und keine zwei Grundgrößen dürfen die selbe Einheit haben).

Die Einheiten $g_j = [G_j]$ der Grundgrößen bezeichnet man als *Grundeinheiten*.

BEISPIEL 2.1. *Die Grundgrößen der Mechanik sind Länge L , Zeit T , Masse M , mit den Einheiten Meter m, Sekunde s, Masse kg.*

In der *Geometrie* kommt man sogar nur mit der Länge L und ihrer Maßeinheit aus. In der *Kinematik* benötigt man zur Beschreibung von Vorgängen zusätzlich die Zeit T . Die Verwendung von Länge, Zeit und Masse charakterisiert die *Dynamik*, wo man sich um das Verhalten von Gegenständen unter dem Einfluss von Kräften kümmert.

Man wählt die Grundgrößen und Grundeinheiten anhand theoretischer und praktischer Überlegungen. Es müssen insbesondere gute Verfahren zur Messung der Grundgrößen zur Verfügung stehen, damit die Messgeräte leicht geeicht werden können. Verschiedene Maßsysteme können sich sowohl durch die Wahl der Grundeinheiten (zum Beispiel cm, g, s, statt m, kg, s) als auch durch die Wahl der Grundgrößen (zB. Ladung oder Stromstärke als Grundgröße in der Elektrodynamik) unterscheiden.

Die Grundeinheiten werden so gewählt, dass die Einheit einer jeden physikalischer Größe A als Potenzprodukte der Grundeinheiten dargestellt werden kann:

$$[A] = \prod_{i=1}^r g_i^{\alpha_i} = g_1^{\alpha_1} g_2^{\alpha_2} \cdots g_r^{\alpha_r}, \quad \alpha_i \in \mathbb{R}. \quad (2.3)$$

(Eigentlich kommen nur rationale Exponenten tatsächlich vor, wir werden uns um diese Einschränkung allerdings nicht kümmern.)

DEFINITION 2.1. Die Dimension $\dim(A)$ einer Größe A , deren Einheit durch (2.3) gegeben ist, ist definiert durch

$$\dim(A) = \prod_{i=1}^r G_i^{\alpha_i} = G_1^{\alpha_1} G_2^{\alpha_2} \cdots G_r^{\alpha_r}. \quad (2.4)$$

Die Dimension ist also eine Bezeichnung, die angibt, welche Grundgrößen mit welchen Potenzen kombiniert werden müssen, damit man eine Größe mit derselben Einheit wie A erhält. Zum Beispiel ist die Dimension einer Geschwindigkeit v

$$\dim(v) = LT^{-1},$$

da die Einheit der Geschwindigkeit in jedem Maßsystem die Einheit der Länge dividiert durch die Einheit der Zeit ist. Die Dimension einer jeden mechanischen Größe ist von der Form $L^\alpha T^\beta M^\gamma$ mit drei (rationalen) Zahlen α , β , und γ .

In der Folge geben wir die Grundgrößen immer als geordnetes r -Tupel (G_1, \dots, G_r) vor. Die Einheit (und die Dimension) einer beliebigen Größe A ist dann durch den sogenannten *Dimensionsvektor*

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_r)$$

gemäß (2.3) (bzw. gemäß (2.4)) eindeutig bestimmt.

Die Grundgrößen selbst haben die kanonischen Einheitsvektoren als Dimensionsvektoren.

2.1.3. Dimensionslose Größen

Einheitenlose Größen nennt man auch dimensionslos. Für eine dimensionslose Größe A gilt $[A] = 1$ und $\dim(A) = 1$. Wegen (2.2) ist das genau dann der Fall, wenn in (2.3) alle Potenzen 0 sind:

$$[A] = \prod_{i=1}^r g_i^{\alpha_i} = 1 \quad \text{genau dann, wenn} \quad \alpha_1 = \alpha_2 = \dots = \alpha_r = 0.$$

Der Dimensionsvektor einer dimensionslosen Größe ist der Nullvektor.

BEISPIEL 2.2. *Ein Beispiel für eine dimensionslose Größe ist der Winkel im Bogenmaß, der als Verhältnis von Bogenlänge zu Radius definiert ist:*

$$\text{Winkel} = \frac{\text{Bogenlänge}}{\text{Radius}}, \quad [\text{Winkel}] = \frac{L}{L} = 1. \quad (2.5)$$

Der Zahlenwert einer dimensionslosen Größe ist in jedem Einheitensystem derselbe. Der rechte Winkel ist $\pi/2$, egal, ob man Bogenlänge und Radius in Meter, Zentimeter, oder Inch mißt.

Dennoch sagt man zum Einheitswinkel oft "ein Radiant", das ist aber nur eine Bezeichnung für die Zahl 1. (Es wäre allerdings durchaus denkbar, den Winkel als neue Grundgröße mit der Grundeinheit Radiant einzuführen. Das wird aber normalerweise nicht gemacht).

2.1.4. Dimensionale Unabhängigkeit

DEFINITION 2.2. *Seien n physikalische Größen A_1, A_2, \dots, A_n mit Einheiten $[A_1], [A_2], \dots, [A_n]$ gegeben. Das n -Tupel (A_1, \dots, A_n) heißt dimensional unabhängig, wenn folgende Aussage gilt:*

$$\prod_{i=1}^n [A_i]^{\lambda_i} = 1 \quad \iff \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0. \quad (2.6)$$

Ein Tupel, das nicht dimensional unabhängig ist, heißt dimensional abhängig.

SATZ 2.1. *Die physikalischen Größen A_1, A_2, \dots, A_n sind genau dann dimensional unabhängig, wenn ihre Dimensionsvektoren $\alpha_1, \alpha_2, \dots, \alpha_n$ ein linear unabhängiges n -Tupel von Vektoren bilden.*

Die r Grundgrößen (G_1, \dots, G_r) bilden per definitionem ein dimensional unabhängiges r -Tupel.

BEISPIEL 2.3. Seien m , s und kg die Einheiten der Grundgrößen Länge, Zeit und Masse. Betrachte folgende physikalische Größen:

$$A_1 = \text{Geschwindigkeit}, \quad [A_1] = m s^{-1}, \quad (2.7)$$

$$A_2 = \text{Kraft}, \quad [A_2] = kg m s^{-2}, \quad (2.8)$$

$$A_3 = \text{Leistung}, \quad [A_3] = kg m^2 s^{-3}. \quad (2.9)$$

Die Grundgrößen sind offenbar unabhängig, aber die drei Größen A_1 , A_2 , und A_3 sind dimensional abhängig, denn es gilt

$$[A_1 A_2 A_3^{-1}] = m s^{-1} kg m s^{-2} kg^{-1} m^{-2} s^3 = 1. \quad (2.10)$$

2.1.5. Vergleich von Einheiten

Die Auswahl von Maßeinheiten zu gegebenen Größen ist völlig willkürlich. Anstelle von Meter, Sekunde, Kilogramm könnte man ja auch mit Meile, Stunde, Pfund rechnen. Entsprechend wird dann die Geschwindigkeitseinheit zu Meilen/Stunde. Amerikanern könnte man sogar 1 "Furlong per Fortnight" als Geschwindigkeitseinheit zutrauen. Um Messungen vergleichen zu können, muß man Maßzahlen physikalischer Größen von einem Maßsystem ins andere übersetzen können.

Betrachten wir allgemein den Übergang $g_i \rightarrow \hat{g}_i$ zu anderen Grundeinheiten,

$$g_i = k_i \hat{g}_i \quad (2.11)$$

(z.B. $m = 39,3701$ inch). Wenn der Konversionsfaktor k_i , wie es oft der Fall ist, eine dimensionslose Zahl ist, dann hat die betrachtete Größe in beiden Einheitensystemen dieselbe Dimension.

Für eine physikalische Größe A erhalten wir also

$$A = m(A) [A] = m(A) \prod_{i=1}^r k_i^{\alpha_i} \hat{g}_i^{\alpha_i} = \left(\prod_{i=1}^r k_i^{\alpha_i} m(A) \right) \prod_{i=1}^r \hat{g}_i^{\alpha_i}. \quad (2.12)$$

Die Maßzahl in den neuen Einheiten ist also

$$\hat{m}(A) = \prod_{i=1}^r k_i^{\alpha_i} m(A) \quad (2.13)$$

und A in den neuen Einheiten ist

$$A = \hat{m}(A) \prod_{i=1}^r \hat{g}_i^{\alpha_i} = \hat{m}(A) [\hat{A}]. \quad (2.14)$$

Dieselbe physikalische Größe A hat also in jedem Einheitensystem eine andere Darstellung

$$A = m(A) [A] = \hat{m}(A) [\hat{A}].$$

Offensichtlich sind die Begriffe Dimension und Einheit verschieden, denn eine Größe mit bestimmter Dimension (zB. Länge) kann verschiedene Einheiten haben (zB. Meter oder Inch). Dennoch werden diese Worte in Dimensionsbetrachtungen oft synonym gebraucht. Betrachtungen über Einheiten lassen sich tatsächlich meistens auch auf Dimensionen übertragen.

2.2. Dimensionsanalyse physikalischer Zusammenhänge

Vieles in diesem Kapitel beruht auf privater Kommunikation mit Dr. Michael A. B. Deakin, Honorary Research Fellow, Monash University, Australien.

2.2.1. Physikalische Gesetze entdecken

Wir betrachten als Beispiel wieder den Fall eines Steines aus gegebener Höhe h . Wir fragen nach einer Formel, einem Naturgesetz, das beschreibt, mit welcher Geschwindigkeit der Stein am Boden auftrifft.

Machen wir uns eine Liste der relevanten Größen und ihrer Dimensionen:

v	Aufprallgeschwindigkeit:	LT^{-1}
h	Anfangshöhe:	L
m	Masse des Steins:	M
g	Erdbeschleunigung:	LT^{-2}
t	Fallzeit:	T

Wir suchen ein Gesetz der Form

$$v = f(h, m, g, t). \quad (2.15)$$

Dabei muss f eine geeignete Funktion ihrer Argumente sein. Geeignet bedeutet hier, dass der Größe $f(h, m, g, t)$ eine eindeutige physikalische Einheit bzw. Dimension zugeordnet werden kann. Geeignete Funktionen sind also zum Beispiel Potenzprodukte ihrer Argumente, sie können aber auch numerische (also dimensionslose) Konstanten enthalten. Hier muss f die Dimension LT^{-1} haben. Wir schließen daraus sofort, daß m auf der rechten Seite gar nicht vorkommen darf.

Aus den Größen g und h können wir bereits so eine Kombination bilden: \sqrt{gh} . Die Dimension ist $\sqrt{LT^{-2}L} = LT^{-1}$. Wir erhalten also folgenden Zusammenhang:

$$v = k\sqrt{gh} \quad (2.16)$$

wobei k eine numerische Konstante ist. Wir können k entweder durch Experimente messen, oder durch die Theorie bestimmen. Die Theorie des freien Falls zeigt: $k = \sqrt{2}$. Durch Dimensionsbetrachtungen allein kann man natürlich die Zahlenwerte dimensionsloser Konstanten nicht bestimmen.

Es gibt noch weitere Möglichkeiten. Zum Beispiel können wir die Erdbeschleunigung und die Fallzeit kombinieren:

$$v = k'gt \quad (\text{Theorie impliziert: } k' = 1). \quad (2.17)$$

Dieses Gesetz besagt, dass die Endgeschwindigkeit proportional der Fallzeit ist. Das ist ein Charakteristikum der gleichmäßig beschleunigten Bewegung.

Oder:

$$v = k''\frac{h}{t} \quad (\text{Theorie impliziert: } k'' = 2). \quad (2.18)$$

Die Endgeschwindigkeit ist das Doppelte der Durchschnittsgeschwindigkeit. Man verwendet Dimensionsbetrachtungen um

- (1) Probleme im vorhinein zu vereinfachen. In $v = f(h, m, g, t)$ müßte man (wenn wir keine Theorie hätten) durch Messungen eine unbekannte Funktion von vier Variablen bestimmen, in $v = k\sqrt{gh}$ ist nur eine einzige Konstante unbekannt.
- (2) Rechenergebnisse zu überprüfen.

Zur Übung überlege man sich die möglichen Gesetze, die die Fallhöhe als Funktion der Zeit darstellen.

Dimensionsüberlegungen dieser Art sind durch Plausibilitätsbetrachtungen mitbestimmt. Schon die Aufstellung der Liste der involvierten Größen verlangt Geschick. Einen gewissen Beitrag zur Systematisierung derartiger Überlegungen leistet das Buckingham-Pi Theorem.

AUFGABE 2.2. Welche Beziehung herrscht bei der gleichförmigen Kreisbewegung zwischen der Fliehkraft F und der Geschwindigkeit v des Körpers? Kombiniere die beteiligten Größen F , m (Masse des Körpers), r (Radius der Kreisbahn) und v (Bahngeschwindigkeit) zu einer dimensional sinnvollen Formel.

2.2.2. Das Buckingham-Pi Theorem

Gegeben seien r Grundgrößen (G_1, \dots, G_r) mit den Einheiten $[G_i] = g_i$. Wir betrachten nun ein physikalisches Problem, für dessen Beschreibung insgesamt n Größen A_1, \dots, A_n eine Rolle spielen. Im allgemeinen ist $r \leq n$, d.h., wir haben mehr Größen A_j als Grundgrößen, und wir können nicht erwarten, dass das n -Tupel (A_1, \dots, A_n) dimensional unabhängig ist.

Die Einheit einer jeden Größe A_j ist ein Potenzprodukt der Grundeinheiten:

$$[A_j] = \prod_{i=1}^r g_i^{\alpha_{ij}}, \quad j = 1, 2, \dots, n. \quad (2.19)$$

(Das häufige Auftreten des Produktzeichens Π verleiht angeblich dem Buckingham Pi Theorem seinen Namen). Die $r \times n$ -Matrix (r Zeilen und n Spalten)

$$M = (\alpha_{ij})_{\substack{i=1,\dots,r \\ j=1,\dots,n}} \quad (2.20)$$

heißt die *Dimensionsmatrix* der n Größen A_1, \dots, A_n . Die k -te Spalte der Dimensionsmatrix ist der Dimensionsvektor der Größe A_k .

Wir untersuchen nun eine *physikalisch sinnvolle Kombination* der Größen A_1, \dots, A_n zu einer neuen physikalischen Größe Π . Damit die Größe Π physikalisch sinnvoll ist, muss ihre Einheit aus den Einheiten der A_j berechenbar sein. Also muss die Einheit $[\Pi]$ ein Potenzprodukt der Einheiten der A_j sein:

$$[\Pi] = \prod_{j=1}^n [A_j]^{\lambda_j}, \quad \text{mit } \lambda_j \in \mathbb{R}. \quad (2.21)$$

Zu jeder solchen Kombination Π gehört also ein Vektor

$$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n, \quad (2.22)$$

der die Einheit von Π durch die Einheiten des n -Tupels (A_1, \dots, A_n) beschreibt. (Ebenso wird die Dimension von Π durch ein Potenzprodukt der Dimensionen $\dim(A_j)$ beschrieben).

Mehrere solche Größen $\Pi_1, \Pi_2, \dots, \Pi_m$ bezeichnen wir als *unabhängige Kombinationen* der gegebenen Größen A_1, \dots, A_n , falls die dazugehörigen Vektoren $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}, \dots, \boldsymbol{\lambda}^{(m)}$ linear unabhängig sind.

Die Einheit einer physikalisch sinnvollen Kombination Π der Größen A_1, \dots, A_n berechnet sich wie folgt:

$$\begin{aligned} [\Pi] &= [A_1]^{\lambda_1} [A_2]^{\lambda_2} \dots [A_n]^{\lambda_n} \\ &= \prod_{i=1}^r g_i^{\alpha_{i1}\lambda_1} \prod_{i=1}^r g_i^{\alpha_{i2}\lambda_2} \dots \prod_{i=1}^r g_i^{\alpha_{in}\lambda_n} \\ &= \prod_{i=1}^r g_i^{\alpha_{i1}\lambda_1 + \alpha_{i2}\lambda_2 + \dots + \alpha_{in}\lambda_n}. \end{aligned}$$

Schreiben wir

$$\boldsymbol{\mu} = A \boldsymbol{\lambda} \quad (2.23)$$

so ist die i -te Komponente des Vektors $\boldsymbol{\mu}$ gleich

$$\mu_i = (A \boldsymbol{\lambda})_i = \alpha_{i1}\lambda_1 + \alpha_{i2}\lambda_2 + \dots + \alpha_{in}\lambda_n \quad (2.24)$$

und wir erhalten aus obiger Rechnung

$$[\Pi] = \prod_{i=1}^r g_i^{\mu_i}. \quad (2.25)$$

Der Vektor $\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\lambda}$ ist also der Dimensionsvektor von Π . Die Größe Π ist genau dann dimensionslos, wenn

$$\mu_1 = \mu_2 = \dots = \mu_r = 0, \quad \text{d.h.} \quad \boldsymbol{\mu} = \mathbf{A}\boldsymbol{\lambda} = \mathbf{0}. \quad (2.26)$$

Wir haben also das folgende Resultat:

Dimensionslose Kombinationen:

Gegeben seien n physikalische Größen A_1, \dots, A_n mit Dimensionsmatrix \mathbf{M} . Eine physikalisch sinnvolle Kombination Π dieser Größen hat die Einheit

$$[\Pi] = \prod_{j=1}^n [A_j]^{\lambda_j} = \prod_{i=1}^n g_i^{\mu_i} \quad (\text{mit } \boldsymbol{\mu} = \mathbf{M}\boldsymbol{\lambda})$$

und ist genau dann dimensionslos (dh. $[\Pi] = 1$), wenn

$$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n) \in \text{Ker } \mathbf{M}$$

Für $n > r$ Größen A_1, \dots, A_n können in der $r \times n$ -Matrix \mathbf{M} höchstens r Spaltenvektoren linear unabhängig sein. Die Dimensionsmatrix hat also höchstens den Rang r . Also ist ihr Kern mindestens $n - r$ -dimensional.

Wir untersuchen nun Dimensionen im Zusammenhang mit physikalischen Gesetzen. Ein physikalisches Gesetz ist eine Beziehung zwischen den für ein gegebenes Problem relevanten physikalischen Größen A_1, \dots, A_n . So eine Beziehung hat zum Beispiel die Form

$$A_1 = F(A_2, \dots, A_n). \quad (2.27)$$

Dabei ist F eine ziemlich beliebige Funktion, der man allerdings eine Dimension (Einheit) zuordnen können muss. (Diese Einheit muss natürlich $[A_1]$ sein). Ein Beispiel für eine solche Funktion wäre ein Potenzprodukt der Größen A_1, \dots, A_n . Allerdings dürfen in F kompliziertere Abhängigkeiten von dimensionslosen Größen vorkommen. Zum Beispiel ist, vom Standpunkt der Dimensionsanalyse aus betrachtet,

$$A_1 = \sin(\Pi) A_2/A_3 \quad (2.28)$$

ein physikalisch sinnvolles Gesetz, falls die Einheit von A_1 gleich der Einheit von A_2/A_3 ist, und falls Π eine dimensionslose Kombination der gegebenen Größen (A_1, A_2, A_3) ist.

Wir formulieren nun das Buckingham Pi Theorem in der Form, wie wir es in den folgenden Kapiteln verwenden werden. Um nicht eine der physikalischen Größen willkürlich auszuzeichnen, betrachten anstelle von (2.27) ein physikalisches Gesetz in der Form

$$\Phi(A_1, \dots, A_n) = 0.$$

Also statt (2.28) betrachten wir $A_1 - F(A_2, \dots, A_n) = 0$. Man beachte, dass $\Phi(G_1, \dots, G_n)$ als Maßzahl 0 hat, aber durchaus eine Einheit haben könnte (im Beispiel die Einheit $[A_1]$). Durch multiplizieren mit einem geeigneten Potenzprodukt von A_1, \dots, A_n können wir aber das Gesetz auch durch eine dimensionslose Funktion beschreiben (im Beispiel: $\tilde{\Phi} = A_1^{-1}\Phi$ ist dimensionslos und $\tilde{\Phi}(A_1, \dots, A_n) = 0$).

Buckingham Pi Theorem:

Seien r Grundgrößen mit Einheiten g_1, g_2, \dots, g_r gegeben und n physikalische Größen A_1, \dots, A_n , deren Dimensionsmatrix den (maximalen) Rang r hat. Dann kann man genau $n - r$ dimensionslose unabhängige Kombinationen $\Pi_1, \Pi_2, \dots, \Pi_{n-r}$ bilden. Ein physikalisches Gesetz der Form

$$\Phi(A_1, \dots, A_n) = 0 \quad (2.29)$$

läßt sich äquivalent dazu in der Form

$$\Psi(\Pi_1, \dots, \Pi_{n-r}) = 0 \quad (2.30)$$

schreiben, also als eine Beziehung von nur $n - r$ dimensionslosen Größen.

BEWEISSKIZZE: Die Dimensionsmatrix der Größen A_1, \dots, A_n hat Rang r , also ist der Kern genau $n - r$ -dimensional. Wir wählen eine Basis im Kern der Dimensionsmatrix und finden so genau $n - r$ unabhängige und dimensionslose Kombinationen Π_1, \dots, Π_{n-r} der Größen A_1, \dots, A_n .

Wir nehmen an, dass in der Dimensionsmatrix von A_1, \dots, A_n gerade die ersten r Spaltenvektoren linear unabhängig sind (evt. muss man dazu umnummerieren). Dann sind die Größen A_1, \dots, A_r dimensional unabhängig, dh., die einzige dimensionslose Kombination dieser Größen ist die triviale Kombination,

$$\prod_{i=1}^r [A_i]^{\lambda_i} = 1 \quad \text{genau dann, wenn} \quad \lambda_1 = \dots = \lambda_r = 0. \quad (2.31)$$

Die Größe A_k entspricht der k -ten Spalte der Dimensionsmatrix. Für $k > r$ ist diese eine Linearkombination der ersten r linear unabhängigen Spalten. Es kann also die Einheit $[A_k]$ mit $k > r$ durch die Einheiten $[A_1], \dots, [A_r]$ ausgedrückt werden:

$$[A_k] = [A_1]^{\beta_1} [A_2]^{\beta_2} \dots [A_r]^{\beta_r}. \quad (2.32)$$

Somit muss

$$A_k = f(\Pi_1, \dots, \Pi_{n-r}) A_1^{\beta_1} A_2^{\beta_2} \dots A_r^{\beta_r} \quad (2.33)$$

sein.

Wir schließen daraus: Jede Funktion der n Größen A_1, \dots, A_n kann auch als Funktion der n Größen $A_1, \dots, A_r, \Pi_1, \dots, \Pi_{n-r}$ geschrieben werden. Also gilt

$$\Phi(A_1, \dots, A_n) = f(A_1, \dots, A_r, \Pi_1, \dots, \Pi_{n-r}) = 0. \quad (2.34)$$

Dabei sind die Größen A_1 bis A_r dimensionsbehaftet und dimensional unabhängig. Bei einer beliebigen Änderung der Grundeinheiten $g_i \rightarrow \hat{g}_i$ werden die Werte der Größen A_1, \dots, A_r gemäß (2.13) mit einem beliebigen Faktor multipliziert, während die dimensionslosen Größen Π_1, \dots, Π_{n-r} ihre Werte unverändert halten. Auch die Funktion f kann wie Φ als dimensionslos angenommen werden, und da sie ihren Wert nicht ändert, egal, welche Werte man für die ersten r Argumente einsetzt, kann sie von diesen ersten Argumenten nicht abhängen:

$$f(A_1, \dots, A_r, \Pi_1, \dots, \Pi_{n-r}) = \Psi(\Pi_1, \dots, \Pi_{n-r}).$$

Aus (2.34) (für jede Wahl der Einheiten) folgt also

$$\Psi(\Pi_1, \dots, \Pi_{n-r}) = 0.$$

□

2.2.3. Das dritte Keplersche Gesetz

“Die Quadrate der Umlaufzeiten zweier Planeten verhalten sich wie die Kuben ihrer großen Halbachsen.” Wir wollen eine angenäherte Version dieses Resultates durch Dimensionsbetrachtungen finden. Zunächst stellen wir wieder die Liste der involvierten Größen auf:

τ	Umlaufzeit:	T
G	Gravitationskonstante:	$M^{-1}L^3T^{-2}$
m_s	Masse der Sonne:	M
m_p	Masse des Planeten:	M
r	Länge der großen Halbachse:	L
e	Exzentrizität:	1

Nach Kepler's erstem Gesetz sind die Planetenbahnen Ellipsen. Daher haben wir die Exzentrizität in die Liste aufgenommen. Die Exzentrizität e einer Ellipse ist der Abstand Zentrum-Brennpunkt geteilt durch die Länge der großen Halbachse, also dimensionslos.

Die oben aufgelisteten 6 Größen werden durch 3 Grundgrößen ausgedrückt. Wir erwarten uns also 3 dimensionslose Kombinationen (Buckingham Pi). Ganz einfach finden wir

$$\Pi_2 = e, \quad \Pi_3 = \frac{m_p}{m_s}. \quad (2.35)$$

Wir brauchen noch eine weitere, davon unabhängige. Am besten findet man diese durch Probieren. Schreiben wir einfach einmal G hin und ergänzen durch die anderen Größen solange, bis wir etwas Dimensionsloses haben:

$$\Pi_1 = Gm_s\tau^2r^{-3}. \quad (2.36)$$

Die Wahl von Π_1 ist natürlich nicht eindeutig, aber unsere Wahl wird sich als günstig für die nun folgende qualitative Betrachtung herausstellen.

Die gesuchte physikalische Beziehung soll die Umlaufzeit beschreiben. Sie ist von der Form

$$\tau = F(G, m_p, m_s, r, e). \quad (2.37)$$

Nach Buckingham Pi können wir das ersetzen durch eine Beziehung der Form

$$\Pi_1 = f(\Pi_2, \Pi_3). \quad (2.38)$$

Dh.,

$$Gm_s\tau^2r^{-3} = f\left(e, \frac{m_p}{m_s}\right) \quad (2.39)$$

oder

$$\tau^2 = \frac{r^3}{Gm_s} f\left(e, \frac{m_p}{m_s}\right) \quad (2.40)$$

Weiter kommen wir ohne genauere Informationen nicht. Beachten wir aber, daß im realen Planetensystem die Exzentrizität der Bahnen sehr klein ist, und die Masse der Sonne sehr groß ist, so können wir näherungsweise

$$\Pi_2 = e \approx 0, \quad \Pi_3 = \frac{m_p}{m_s} \approx 0 \quad (2.41)$$

setzen. Weiters nehmen wir an:

$$f\left(e, \frac{m_p}{m_s}\right) \approx f(0, 0) = k \quad (\text{eine Konstante}). \quad (2.42)$$

Diese Annahme ist nicht selbstverständlich. Eine beliebige Funktion f muss ja für kleine Argumentwerte nicht konvergieren oder beschränkt bleiben. In der Physik probiert man aber oft einfach aus, wie weit man mit einfachen Annahmen kommt.

Aus (2.40) und (2.42) folgt

$$\tau^2 = \frac{k}{Gm_s} r^3. \quad (2.43)$$

Das Quadrat der Umlaufzeit ist also zur dritten Potenz des Sonnenabstandes proportional. Da die Proportionalitätskonstante $k/(Gm_s)$ als vom betrachteten Planeten unabhängig angenommen werden kann, erhalten wir das dritte Kepler'sche Gesetz, wenn wir das Verhältnis τ_1^2/τ_2^2 für zwei Planeten bilden.

Wenn man die Newton-Gleichungen für das Planetensystem löst, erhält man für die Funktion f das Resultat

$$f\left(e, \frac{m_p}{m_s}\right) = 4\pi^2. \quad (2.44)$$

Die Exzentrizität und die Planetenmasse ist also gar nicht involviert und die richtige Formel lautet:

$$\tau^2 = \frac{4\pi^2}{Gm_s} r^3. \quad (2.45)$$

Es ist bemerkenswert, daß wir die Formel (2.43) praktisch ohne Rechenaufwand gefunden haben. Die Methode braucht aber Geschick und physikalisches Verständnis, da sie voraussetzt, daß man die physikalisch relevanten Größen korrekt identifiziert. Wenn man zB. mit m_s die Masse des Sirius meint und mit r den Abstand des Planeten vom Zentrum der Galaxis kann man zwar dimensionsmäßig richtige aber dennoch sinnlose und physikalisch falsche Formeln erhalten.

2.2.4. Wie man Staatsgeheimnisse entdeckt

Mit der Dimensionsanalyse hat der englische Physiker Geoffrey Ingram Taylor (1886-1975) die Energie der ersten Atombombe herausgefunden, zu einer Zeit, als diese Zahl noch top secret war¹. Diese erste Atombombe explodierte in der Wüste von New Mexico im Jahr 1944. Die Explosion fand in der Atmosphäre statt. Der Feuerball war anfänglich in guter Näherung halbkugelförmig. Das US Militär hat den Vorgang gefilmt (siehe Abbildung 2.2.4). Durch die militärische Zusammenarbeit mit Großbritannien kam Taylor auch zum Manhattan Projekt nach Los Alamos. Dort überraschte Taylor die US-Militärs, indem er ihnen die freigesetzte Energie nannte — eine Zahl zu der er keinen Zugang hatte. Wir können hier seine Argumentation sinngemäß nachvollziehen.

Wir betrachten den Feuerball als einen Bereich von Luft mit extrem hoher Dichte und Druck innerhalb, einer Schockwelle am Rand, und normalen Verhältnissen außerhalb. Machen wir uns wieder eine Liste der involvierten Größen und ihrer Dimensionen.

E	Energie der Explosion (gesucht):	ML^2T^{-2}
t	Zeit seit der Explosion:	T
R	Radius des Feuerballs:	L
ρ_A	Massendichte der Luft außen:	ML^{-3}
ρ_I	Massendichte der Luft innen:	ML^{-3}
P_A	Druck der Luft außen:	Kraft/Fläche = $ML^{-1}T^{-2}$
P_I	Druck der Luft innen:	$ML^{-1}T^{-2}$

Es ist vielleicht bemerkenswert, daß wir die Temperatur nicht in die obige Liste aufgenommen haben. Wir können aber davon ausgehen, daß die Temperatur sich aus den anderen Bestimmungsgrößen über eine Gasgleichung ableiten läßt.

¹Siehe: G.I.Taylor: Proc.Roy.Soc. **A201** (1950), 159–196.

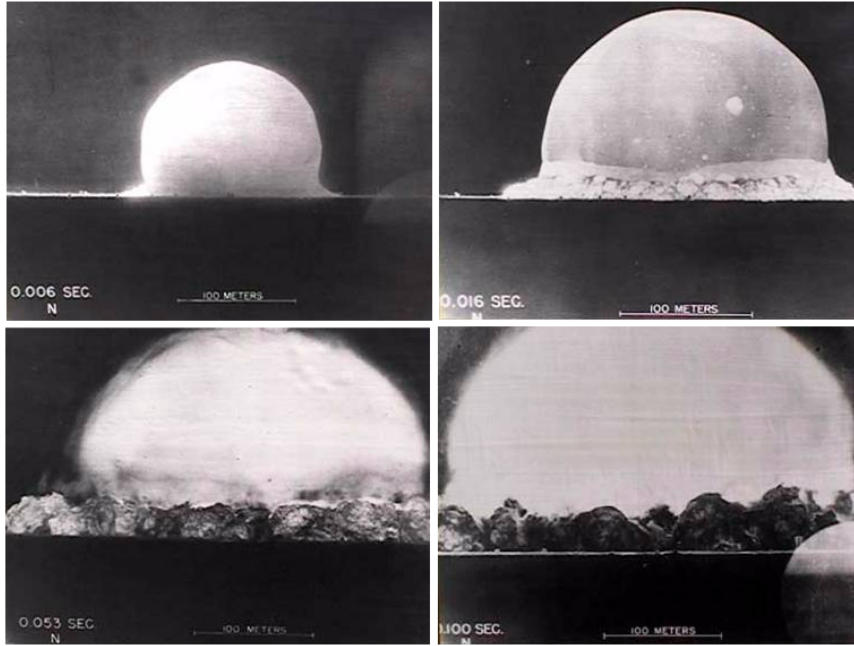


ABBILDUNG 2.1. Einige Bilder des Films von der ersten Atombombenexplosion in New Mexico

Wir haben in unserer Tabelle 7 Größen durch 3 Grundgrößen ausgedrückt, es sollten also 4 dimensionslose Kombinationen involviert sein. (Buckingham Pi Theorem). So einen Satz dimensionsloser Größen findet man am leichtesten durch Probieren. Zum Beispiel schreiben wir einfach einmal die gesuchte Größe E hin, und ergänzen mit den anderen Größen so lange, bis wir etwas Dimensionsloses erhalten:

$$\Pi_1 = E \rho_A^{-1} R^{-5} t^2 \quad (2.46)$$

Weiters finden wir

$$\Pi_2 = P_A^5 t^6 E^{-2} \rho_I^{-3} \quad (2.47)$$

und die trivialen Kombinationen

$$\Pi_3 = \frac{\rho_A}{\rho_I}, \quad \Pi_4 = \frac{P_A}{P_I}. \quad (2.48)$$

Wieder ist die Wahl der Größen nicht eindeutig und von Zusatzüberlegungen bestimmt. Die Größen Π_2 , Π_3 , und Π_4 sind nämlich so gewählt, daß ihre numerischen Meßwerte beim betrachteten Vorgang (und allen ähnlichen

Vorgängen) sehr klein sind. Es handelt sich ja um Verhältnisse von “normalen” Größen (die die Situation außerhalb beschreiben), zu Größen, die die Situation einer Atomexplosion beschreiben.

Die gesuchte Beziehung, die die Energie mit den anderen Größen verknüpft,

$$E = F(t, R, \rho_A, \rho_I, P_A, P_I) \quad (2.49)$$

läßt sich nach Buckingham Pi äquivalent dazu als Beziehung der dimensionslosen Größen

$$\Pi_1 = f(\Pi_2, \Pi_3, \Pi_4) \quad (2.50)$$

schreiben. Nun sind Π_2 , Π_3 , und Π_4 sehr klein. Setzen wir sie näherungsweise gleich Null, erhalten wir

$$\Pi_1 = f(0, 0, 0) = \text{const.} \quad (2.51)$$

oder

$$E = \rho_A \frac{R^5}{t^2} \text{const.} \quad (2.52)$$

Dabei haben wir wieder angenommen, daß die (unbekannte) Funktion f in einer Umgebung von $(0, 0, 0)$ stetig ist. Die auf diese Weise erhaltene Konstante sollte also näherungsweise unabhängig von den Details des Vorgangs sein, also für alle heftigen Explosionen ungefähr denselben Wert haben.

Diese Annahmen müssen natürlich nicht zutreffen. Man kann aber den gewonnenen Zusammenhang (2.52) am Film testen. Er besagt ja, daß der Radius des Feuerballs auf folgende Weise mit der Zeit zusammenhängt:

$$R = A t^{2/5} \quad \text{mit} \quad A = \left(\frac{E}{\rho_A \text{const.}} \right)^{1/5}. \quad (2.53)$$

Logarithmieren wird diese Beziehung,

$$\ln R = \ln A + \frac{2}{5} \ln t. \quad (2.54)$$

Wir können diese Beziehung am Film nachprüfen und daraus A bekommen (siehe Abbildung 2.2). Da ρ_A bekannt ist und die Konstante A aus dem Film bestimmt werden kann, fehlt zur Bestimmung von E nur noch $f(0, 0, 0)$. Diesen Wert kann man im Prinzip aus einer Vergleichsexplosion mit konventionellem Sprengstoff bestimmen. Die Funktion f ist ja eine ganz allgemein für energiereiche Explosionen vorkommende Funktion, die den gesetzmäßigen Zusammenhang zwischen den physikalischen Größen beschreibt. Die Explosion muß nur stark genug sein, damit die Näherung $f(\Pi_2, \Pi_3, \Pi_4) \approx f(0, 0, 0)$ gerechtfertigt ist. Man kann zB nach folgender Methode vorgehen. Wir bestimmen A aus dem Film wie beschrieben und ein A_0 aus einer Vergleichsexplosion einer bestimmten Menge (zB 1 kg) TNT. Dann bildet man das

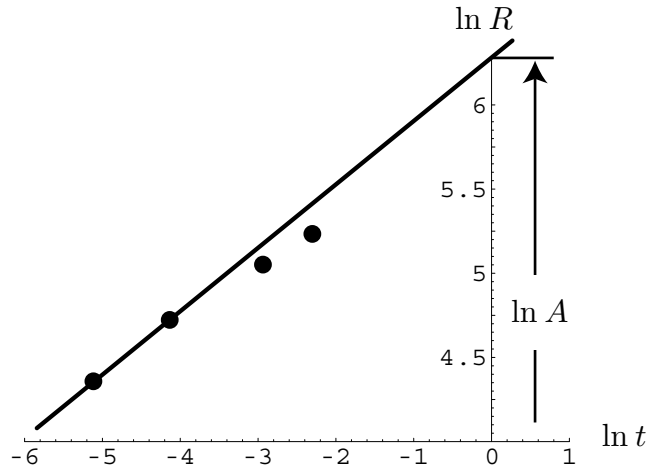


ABBILDUNG 2.2. Der Zusammenhang zwischen $\ln R$ und $\ln t$, empirisch aus Abb. 2.2.4 ermittelt. Der erwartete Zusammenhang (2.54) ist in der Anfangsphase der Explosion gut erfüllt.

Verhältnis

$$\frac{A}{A_0} = \frac{\left(\frac{E}{\rho_A \text{ const.}}\right)^{1/5}}{\left(\frac{E_0}{\rho_A \text{ const.}}\right)^{1/5}} = \left(\frac{E}{E_0}\right)^{1/5} \quad (2.55)$$

wobei E_0 die Energie des konventionellen Sprengstoffs ist. Wir erhalten also

$$\frac{E}{E_0} = \left(\frac{A}{A_0}\right)^5. \quad (2.56)$$

Dieses Verhältnis sagt uns, wieviel kg TNT man theoretisch braucht, um dieselbe Energie freizusetzen, wie bei der Atombombenexplosion. Tatsächliche Zahlen sind Hiroshima: 13 kt TNT, Nagasaki: 20 kt TNT. Heute sind bis zu einigen Megatonnen TNT möglich.

2.2.5. Wie klingen Insekten?

Dimensionsanalyse findet vor allem als begleitende Analyse von Experimenten Anwendung. Als Beispiel betrachten wir den Insektenflug. Wir fragen, wie die Frequenz des Flügelschlags eines Insekts von anderen relevanten Größen, wie Masse des Insekts und Flügelfläche, abhängt.

Liste der relevanten Größen und ihrer Dimensionen:

ν	Frequenz des Flügelschlags:	T^{-1}
A	Fläche der Flügel:	L^2
m	Masse des Insekts:	M
g	Erdbeschleunigung:	LT^{-2}
ρ	Dichte der Luft:	ML^{-3}

Nach Buckingham Pi erwarten wir $5 - 3 = 2$ dimensionslose Verhältnisse, zB

$$\frac{\rho A^{3/2}}{m}, \quad \frac{\nu^2 A^{1/2}}{g}. \quad (2.57)$$

Wir vermuten, dass diese beiden Größen durch irgendeine Beziehung (physikalisches Gesetz) verknüpft sind.

$$\frac{\nu^2 A^{1/2}}{g} = f\left(\frac{\rho A^{3/2}}{m}\right). \quad (2.58)$$

Wir denken uns diese Beziehung nach ν aufgelöst

$$\nu = g^{1/2} A^{-1/4} \psi\left(\frac{\rho A^{3/2}}{m}\right). \quad (2.59)$$

Die Größe im Argument der unbekanntten Funktion ist

$$x = \frac{\rho A^{3/2}}{m} \approx \frac{\text{Masse der Luft im Volumen } (\sqrt{A})^3}{\text{Masse des Insekts}} \quad (2.60)$$

Da das Insekt hauptsächlich aus Wasser besteht, welches viel schwerer als Luft ist, ist das eine sehr kleine Zahl. Wir könnten daraus also schließen,

$$\nu = g^{1/2} A^{-1/4} \psi(0), \quad (?) \quad (2.61)$$

wobei — analog zu den Beispielen im vorherigen Kapitel — $\psi(0)$ als Konstante angenommen wird. Tatsächlich ist dieses Resultat aber nicht sehr gut. Zum Beispiel kommt die Masse des Insekts nicht mehr vor — wir erwarten aber, dass ein schwereres Insekt bei gleicher Flügelfläche eine höhere Schlagfrequenz braucht, um in der Luft zu bleiben. Wir erwarten also, dass mit kleinem Argument x (also großer Masse und/oder kleine Abmessungen) die Schlagfrequenz groß wird, was man auch durch Messung der Frequenz ν an mehreren Insekten bestätigen kann.

Die Annahme, dass man $\psi(x)$ für kleine x durch eine Konstante ersetzen kann, die vom betrachteten Insekt weitgehend unabhängig ist, lässt sich also nicht aufrecht erhalten. Im Gegenteil, wir müssen annehmen, dass ψ eine Funktion ist, die für kleine Argumente sehr groß wird. Hat man dies erkannt, wird man folgenden Ansatz machen:

$$\psi(x) = k x^\alpha (1 + a_1 x + a_2 x^2 + \dots), \quad (2.62)$$

wobei α negativ ist. Dabei sind k und a_n dimensionslose Konstanten (die wir als unabhängig vom betrachteten Insekt annehmen). Die Funktion hat also eine Singularität für $x \rightarrow 0$. Für kleine x können wir also approximieren:

$$\psi(x) \approx k x^\alpha \quad (\text{für kleine } x). \quad (2.63)$$

Aus der Beziehung (2.59) wird dann

$$\nu = g^{1/2} A^{-1/4} k \left(\frac{\rho A^{3/2}}{m} \right)^\alpha = k g^{1/2} \rho^\alpha m^{-\alpha} A^{3\alpha/2-1/4}. \quad (2.64)$$

Den Exponenten α kann man letztlich nur aus dem Experiment bestimmen. Der Vergleich der Flügelschlagfrequenzen von Insekten gleicher Masse m zeigt, daß näherungsweise $\nu \approx 1/A$ ist. Damit wir das aus (2.64) erhalten, müssen wir annehmen:

$$\alpha \approx -1/2. \quad (2.65)$$

Dann erhalten wir:

$$\nu = k g^{1/2} \rho^{-1/2} m^{1/2} A^{-1}. \quad (2.66)$$

Aus den Messungen ergibt sich auch der Zahlenwert der Konstanten k :

$$k g^{1/2} \rho^{-1/2} \approx 0,1. \quad (2.67)$$

Für eine Hummel ist $m \approx 1,5 \times 10^{-3}$ kg und $A \approx 2 \times 10^{-5}$ m². Setzt man diese Zahlenwerte ein, erhält man

$$\nu = 0,1 \times \frac{\sqrt{1,5} \times 10^{-3/2}}{2 \times 10^{-5}} \text{ s}^{-1} = 193,6 \text{ Hz} \approx 200 \text{ Hz}. \quad (2.68)$$

Übrigens können solche hohen Bewegungsfrequenzen nur durch ungewöhnliche Muskeltypen erreicht werden. Bei diesen hängt die Kontraktionsgeschwindigkeit nicht vom stimulierenden Nervenimpuls ab, sondern von der Last, die auf den Muskel einwirkt.

Anhang: Überblick über die Einheiten des SI

Im folgenden beschreiben wir die Grundeinheiten im SI. Das SI (système international d'unités, internationales Einheitensystem) ist 1960 von der 11. Generalkonferenz für Maß und Gewicht (CGPM) eingeführt worden². Die nachfolgend aufgeführten sieben Grundeinheiten sind hinsichtlich ihrer Definition unabhängig voneinander:

Das **Meter** (m) ist die Länge der Strecke, die Licht im Vakuum während der Dauer von (1/299 792 458) Sekunden durchläuft.

Das **Kilogramm** (kg) ist die Einheit der Masse; es ist gleich der Masse des Internationalen Kilogrammprototyps ("Pariser Urkilogramm").

²Siehe die Web-Seite <http://physics.nist.gov/cuu/Units/units.html> für detailliertere Informationen.

Die **Sekunde** (s) ist das 9 192 631 770-fache der Periodendauer der dem Übergang zwischen den beiden Hyperfeinstruktur-niveaus des Grundzustandes von Atomen des Nuklids ^{133}Cs entsprechenden Strahlung.

Das **Ampere** (A) ist die Stärke eines konstanten elektrischen Stromes, der, durch zwei parallele, geradlinige, unendlich lange und im Vakuum im Abstand von einem Meter voneinander angeordnete Leiter von vernachlässigbar kleinem, kreisförmigen Querschnitt fließend zwischen diesen Leitern je einem Meter Leiterlänge die Kraft 2×10^{-7} Newton hervorrufen würde.

Das **Kelvin** (K), die Einheit der thermodynamischen Temperatur, ist definiert als der Bruchteil $1/273,16$ der thermodynamischen Temperatur des Tripelpunktes des Wassers.

Das **Mol** (mol) ist die Stoffmenge eines Systems, das aus ebensoviel Einzelteilchen besteht, wie Atome in 0,012 Kilogramm des Kohlenstoffnuklids ^{12}C enthalten sind. Bei Benutzung des Mol müssen die Einzelteilchen spezifiziert sein und können Atome, Moleküle, Ionen, Elektronen sowie andere Teilchen oder Gruppen solcher Teilchen genau angegebener Zusammensetzung sein.

Die **Candela** (cd) ist die Lichtstärke in einer bestimmten Richtung einer Strahlungsquelle, die monochromatische Strahlung der Frequenz 540×10^{12} Hertz aussendet und deren Strahlstärke in dieser Richtung $(1/683)$ Watt durch Steradian beträgt.

Die Einheiten von physikalischen Größen, die als Potenzprodukte aus den Grundeinheiten berechenbar sind, bekommen manchmal einen eigenen Namen und ein eigenes Einheitensymbol. Diese Einheiten heißen dann *abgeleiteten Einheiten*.

Die Dimension der Kraft F ist zum Beispiel (wegen Kraft = Masse mal Beschleunigung)

$$[F] = [m] [a] = LT^{-2}M. \quad (2.69)$$

Die dazugehörige abgeleitete Einheit nennt man ein Newton (N). Eine Kraft von einem Newton verleiht also der Masseneinheit 1 kg die Beschleunigung von 1m/s^2 :

$$1\text{N} = 1\text{kg m s}^{-2}. \quad (2.70)$$

Die Schwerkraft an der Erdoberfläche ist ungefähr zehnmal so stark, ein Kilogrammstück wird mit etwa 10 Newton nach unten gezogen.

Die Masse m ist als Grundgröße definiert, durch Vergleich mit einer Normmasse. (Man könnte die Masse auch über den Impulserhaltungssatz durch Betrachtung elastischer Stoßvorgänge einführen).

Es ist aber zu beachten, dass die Anzahl und Auswahl von Grundgrößen ebenfalls willkürlich ist. Man könnte z.B. auf die Stromstärke als Grundgröße

ganz verzichten, wenn das Coulombgesetz (Kraft zwischen zwei Ladungen q_1 und q_2) so geschrieben wird:

$$F = \frac{q_1 q_2}{r^2} \quad (2.71)$$

Die Dimension der Ladung ergibt sich daraus zu

$$[q] = \sqrt{[F][r]^2} = M^{1/2} L^{3/2} T^{-1}. \quad (2.72)$$

Die Strömstärke ist dann eine abgeleitete Größe, nämlich die pro Zeiteinheit transportierte Ladung mit der Dimension Ladung/Zeit. Statt dessen definiert man im SI das Ampere als Grundgröße durch ein eigenes Messverfahren. Dann ist die Ladung eine abgeleitete Größe mit der Einheit As (= 1 Amperesekunde = 1 Coulomb). Im SI wird das Coulombgesetz zu

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2} \quad (2.73)$$

mit einer dimensionsbehafteten Konstante ϵ_0 (Dielektrizitätskonstante des Vakuums).

Kapitel 3

Einige Grundsätze der Mechanik

3.1. Kraft und Energie

Wir erinnern (ganz kurz) an einige wichtige physikalische Grundbegriffe.

Kraft: Kräfte können auf bewegliche Körper einwirken und bewirken dann eine Beschleunigung. Die Form des Körpers wird durch eine geeignete Teilmenge K des \mathbb{R}^3 beschrieben. Ausgedehnte Körper haben oft eine komplizierte Massenverteilung und andere Eigenschaften (Elastizität), die aufwändig zu beschreiben sind. Sehr oft sieht man daher von der räumlichen Ausdehnung und Form der Körper ab und betrachtet *Massenpunkte*, die zur Gänze durch ihre Masse m charakterisiert werden. Der Zustand eines Massenpunktes wird durch seine Position im Raum und seine Geschwindigkeit angegeben (typischerweise also durch einen 6-dimensionalen Vektor).

Die Kraft ist eine vektorielle Größe und proportional zur erzielten Beschleunigung. Mehrere Kräfte können gleichzeitig auf einen Punkt des Körper einwirken. Dann werden die Kräfte vektoriell addiert und die Beschleunigung des Körpers ist proportional zur Summenkraft. Kräfte, die an verschiedenen Punkten eines ausgedehnten Körpers angreifen haben nichts miteinander zu tun. Sie können insbesondere nicht addiert werden. Das Bild eines frei verschiebbaren Pfeils eignet sich ganz und gar nicht als Visualisierung eines Kraftvektors.

In einem starren Festkörper tritt die besondere Situation auf, dass eine Kraft “entlang ihrer Wirkungslinie” verschoben werden kann. Diese Eigenschaft definiert die Starrheit des Körpers und ist keine generelle Eigenschaft von Kräften.

Die Kraft auf einen Massenpunkt kann durch die erzielte Beschleunigung definiert werden:

$$\mathbf{F} = m \mathbf{a} \quad (\text{Newton'sches Gesetz}). \quad (3.1)$$

Gleichgewicht: Ein Massenpunkt kann nur dann in einem *stationären Zustand* (in einer *Gleichgewichtslage*) sein, wenn die (vektorielle) Summe aller

angreifenden Kräfte Null ist (*Gleichgewichtsbedingung*), wenn sich also alle am Massenpunkt angreifenden Kräfte gegenseitig aufheben.

Kraftfeld: Nur in den einfachsten Fällen können Kräfte durch einzelne, an gegebenen Punkten “angeheftete” Vektoren beschrieben werden. In vielen Problemen tritt ein *Kraftfeld* auf. Dieses wird durch ein Vektorfeld beschrieben, dh. durch eine vektorwertige Funktion der Ortskoordinaten:

$$\mathbf{F} : \mathbf{x} \rightarrow \mathbf{F}(\mathbf{x}). \quad (3.2)$$

Wenn sich ein Massenpunkt am Ort \mathbf{x} befindet, dann wirkt auf ihn die Kraft $\mathbf{F}(\mathbf{x})$.

3.2. Die Newton’schen Bewegungsgleichungen

Wir betrachten zunächst einen Massenpunkt (ein Teilchen), daß sich unter dem Einfluß eines zeitunabhängigen Kraftfeldes bewegt. Seine Bewegung wird durch einen differenzierbaren Weg

$$\mathbf{x} : t \rightarrow \mathbf{x}(t) \quad (3.3)$$

beschrieben. Die Ableitung

$$\mathbf{v} : t \rightarrow \mathbf{v}(t) = \left. \frac{d\mathbf{x}}{dt} \right|_t \equiv \frac{d\mathbf{x}(t)}{dt} \quad (3.4)$$

ist natürlich die Geschwindigkeit.

Aus dem Newton’schen Bewegungsgesetz (Masse \times Beschleunigung = Kraft) erhalten wir nun sofort eine Bewegungsgleichung:

$$m \frac{d^2\mathbf{x}(t)}{dt^2} = F(\mathbf{x}(t)). \quad (3.5)$$

Eine eindeutige Lösung dieses Differentialgleichungssystems 2. Ordnung benötigt (außer geeigneten Stetigkeitsbedingungen an die Kraft) noch die Angabe von Anfangsort und Anfangsgeschwindigkeit (also die Angabe des Anfangszustandes):

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{v}(0) = \mathbf{v}_0 \quad (\text{Anfangsbedingungen}). \quad (3.6)$$

Kennt man den Anfangszustand, dann liefert die Lösung der Bewegungsgleichung den Zustand $(\mathbf{x}(t), \mathbf{v}(t))$ zur Zeit t . Der Ausdruck, der für die Kraft einzusetzen ist, hängt vom betrachteten physikalischen System ab. Zum Beispiel führt die Federkraft

$$\mathbf{F}(\mathbf{x}) = -k \mathbf{x} \quad (k = \text{Federkonstante}). \quad (3.7)$$

auf die Gleichung

$$m \frac{d^2\mathbf{x}(t)}{dt^2} = -k \mathbf{x}(t). \quad (3.8)$$

Für ein System aus n Massenpunkten wird das entsprechend komplizierter. Die Kräfte sind nun Funktionen der Koordinaten und evt. auch Geschwindigkeiten der einzelnen Körper. $\mathbf{F}^{(i)}$, die Kraft auf den i -ten Massenpunkt ist dann eine Abbildung von \mathbb{R}^{6n} in den \mathbb{R}^3 (vorausgesetzt, das System bewegt sich im dreidimensionalen Raum):

$$\mathbf{F}^{(i)}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}, \mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}) \quad (3.9)$$

Beispiel: Das Planetensystem. Eine im Weltall am Punkt $\mathbf{x}^{(0)}$ befindliche Masse $m^{(0)}$ ("Sonne") wirkt auf eine zweite Masse m am Punkt \mathbf{x} ("Erde") mit der anziehenden *Gravitationskraft*:

$$\mathbf{F}(\mathbf{x}) = G m^{(0)} m \frac{\mathbf{x}^{(0)} - \mathbf{x}}{|\mathbf{x}^{(0)} - \mathbf{x}|^3}. \quad (3.10)$$

(Die Kraft zeigt in Richtung des Einheitsvektors vom Punkt \mathbf{x} zum Punkt $\mathbf{x}^{(0)}$ und ist invers proportional zum Quadrat des Abstandes. G ist die universelle Gravitationskonstante.)

Ein System von n Massen $m^{(1)}, \dots, m^{(n)}$ muß also unter dem alleinigen Einfluß der Gravitationskraft folgende Bewegungsgleichungen erfüllen:

$$\frac{d^2 \mathbf{x}^{(i)}}{dt^2} = G \sum_{j \neq i} m^{(j)} \frac{\mathbf{x}^{(j)} - \mathbf{x}^{(i)}}{|\mathbf{x}^{(j)} - \mathbf{x}^{(i)}|^3}, \quad i = 1, \dots, n. \quad (3.11)$$

(Die Masse $m^{(i)}$ kürzt sich heraus.) Das ist ein System von $3n$ verkoppelten nichtlinearen Gleichungen zweiter Ordnung.

Beispiel: Lorentzkraft im Magnetfeld. Für die Bewegung eines geladenen Teilchens im Magnetfeld gilt eine andersartige Gleichung, denn hier hängt die Kraft von der Geschwindigkeit ab:

$$m \frac{d^2 \mathbf{x}(t)}{dt^2} = q \frac{d\mathbf{x}(t)}{dt} \times \mathbf{B}(\mathbf{x}(t)). \quad (3.12)$$

(Hier ist q die Ladung des Teilchens, und \mathbf{B} die magnetische Induktion). Die Beschleunigung ist hier immer senkrecht auf die Geschwindigkeit, ändert also nur deren Richtung und nicht deren Betrag.

3.3. Wie man physikalische Konstanten los wird

Die Zahlenwerte der in den Bewegungsgleichungen vorkommenden physikalischen Konstanten hängt von der Wahl der Grundeinheiten ab. Eine geeignete Änderung der Einheiten (Skalentransformation) kann die Bewegungsgleichungen vereinfachen.

Als Beispiel betrachten wir die Gleichung für den Massenpunkt an einer Feder (harmonischer Oszillator):

$$m \frac{d^2 \mathbf{x}(t)}{dt^2} = -k \mathbf{x}(t). \quad (3.13)$$

Hier ist k die Federkonstante, mit der Einheit kg s^{-2} .

Wir ändern die Zeiteinheit mit einem Konversionsfaktor k_t und schreiben:

$$\hat{t} = k_t t, \quad (3.14)$$

wobei t die Zeit in Sekunden ist, und \hat{t} die Zeit in den neuen Einheiten. Die Funktion, die den Ort des Massenpunkts als Funktion der Zeit in den neuen Einheiten beschreibt, schreiben wir $\hat{\mathbf{x}}$, also

$$\mathbf{x}(t) = \hat{\mathbf{x}}(\hat{t}) = \hat{\mathbf{x}}(k_t t). \quad (3.15)$$

Daher erhalten wir

$$\frac{d\mathbf{x}(t)}{dt} = \frac{d\hat{\mathbf{x}}(\hat{t})}{d\hat{t}} \frac{d\hat{t}}{dt} = k_t \frac{d\hat{\mathbf{x}}(\hat{t})}{d\hat{t}}, \quad (3.16)$$

$$\frac{d^2 \mathbf{x}(t)}{dt^2} = k_t^2 \frac{d^2 \hat{\mathbf{x}}(\hat{t})}{d\hat{t}^2}. \quad (3.17)$$

Die Bewegungsgleichung wird daher

$$m k_t^2 \frac{d^2 \hat{\mathbf{x}}(\hat{t})}{d\hat{t}^2} = -k \hat{\mathbf{x}}(\hat{t}). \quad (3.18)$$

Wählen wir

$$k_t = \sqrt{\frac{k}{m}}, \quad (3.19)$$

so erhalten wir die Gleichung in skalierten Einheiten

$$\frac{d^2 \hat{\mathbf{x}}(\hat{t})}{d\hat{t}^2} = -\hat{\mathbf{x}}(\hat{t}). \quad (3.20)$$

Der Konversionsfaktor k_t hat übrigens die Einheit $1/\text{s}$, sodass die neue Zeit $\hat{t} = k_t t$ dimensionslos ist.

Lösungen der Gleichung (3.20) sind proportional zu $\sin \hat{t}$ und $\cos \hat{t}$ oder Linearkombinationen dieser Funktionen:

$$\hat{\mathbf{x}}(\hat{t}) = a \cos \hat{t} + b \sin \hat{t} \quad (3.21)$$

$$= c \cos(\hat{t} + d) \quad (\text{harmonische Schwingung}). \quad (3.22)$$

Die Konstanten a und b (bzw. c und d) können nun einer Anfangsbedingung angepaßt werden. (a , b und c haben die Dimension einer Länge).

Weitere Vereinfachungen, die allerdings von der gewählten Anfangsbedingung abhängen: Die Längenskala kann so gewählt werden, daß c den Zahlenwert 1 hat. Der Anfangspunkt der Zeitskala kann so gewählt werden, daß $d = 0$ ist.

3.4. Potenzielle Energie und Arbeit im Kraftfeld

Sei \mathbf{F} ein Kraftfeld, dh. eine stetig differenzierbare Funktion

$$\mathbf{F} : \begin{cases} U \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3 \\ \mathbf{x} \rightarrow \mathbf{F}(\mathbf{x}) \end{cases} \quad (3.23)$$

definiert auf einer offenen Teilmenge U des Raumes. Weiters sei eine stetig differenzierbare, parametrisierte Kurve \mathbf{x} in U gegeben.

DEFINITION 3.1. *Eine parametrisierte Kurve \mathbf{x} in \mathbb{R}^n ist eine stetige Abbildung*

$$\mathbf{x} : \begin{cases} [a, b] \subset \mathbb{R} \rightarrow U \subset \mathbb{R}^n \\ t \mapsto \mathbf{x}(t) \in U \end{cases} \quad (3.24)$$

Abbildung 3.1 zeigt eine differenzierbar(!) parametrisierte Kurve.

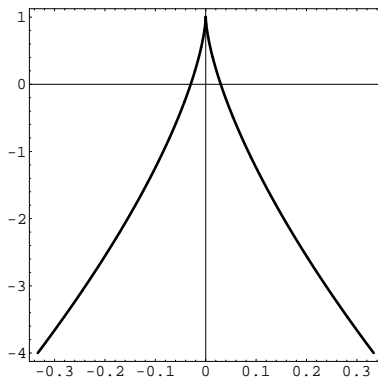


ABBILDUNG 3.1. Plot der differenzierbaren Kurve $\mathbf{x}(t) = (t^3/3, 1 - 5t^2)$ für $t \in [-1, 1]$.

Arbeit = Energie: In Gegenwart einer ortsabhängigen Kraft $\mathbf{F}(\mathbf{x})$ werde ein Körper entlang des (stetig differenzierbaren) Wegs $\gamma : t \rightarrow \mathbf{x}(t)$ vom Punkt $\mathbf{x}_1 = \mathbf{x}(a)$ zum Punkt $\mathbf{x}_2 = \mathbf{x}(b)$ geführt. Das Wegintegral

$$W_\gamma(\mathbf{x}_1, \mathbf{x}_2) = \int_\gamma \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} := \int_a^b \left(\mathbf{F}(\mathbf{x}(t)) \cdot \frac{d\mathbf{x}(t)}{dt} \right) dt \quad (3.25)$$

heißt die *Arbeit* oder *Energie*, die bei dieser Bewegung umgesetzt wird.

Wenn W positiv ist, kommt Energie “aus dem System heraus”, das Kraftfeld wirkt “in Bewegungsrichtung”. Wenn W negativ ist, ist die Bewegungsrichtung des Körpers entgegengesetzt zur Krafrichtung. Man muß also von außen Energie in den Körper hineinstecken, damit diese Bewegung stattfindet. Zugeführte Energie kann im System als *potenzielle Energie* gespeichert werden. Eine Erhöhung der potenziellen Energie erfolgt bei *reversiblen Vorgängen*. Überläßt man das System dann sich selbst, kann die potenzielle Energie wieder in Form von Arbeit entnommen werden. Bei sogenannten *irreversiblen Vorgängen* wird Energie in Form von Wärme dissipiert.

Potenzielle Energie: Sehr oft ist ein Kraftfeld der Gradient eines skalaren Feldes:

$$\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x}). \quad (3.26)$$

Der Gradient ist der aus den partiellen Ableitungen gebildete Vektor. Im \mathbb{R}^3 ,

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \quad (3.27)$$

Zum Beispiel ist die Federkraft der Gradient von

$$V(\mathbf{x}) = \frac{k}{2} x^2, \quad -\nabla V(\mathbf{x}) = -k\mathbf{x}, \quad (3.28)$$

und die Gravitationskraft (3.10) ist der Gradient von

$$V(\mathbf{x}) = -\frac{G m^{(0)} m}{|\mathbf{x}^0 - \mathbf{x}|}. \quad (3.29)$$

Die physikalische Bedeutung von V ist die einer potenziellen Energie.

Gleichgewichtslagen des Systems ($\mathbf{F} = -\nabla V = 0$) sind dann die kritischen Punkte von V . Die Minima von V sind stabile Gleichgewichtslagen.

SATZ 3.1. Sei $\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x})$ überall in einem offenen Gebiet U des \mathbb{R}^3 , und \mathbf{x}_1 und \mathbf{x}_2 Punkte in U . Dann ist die Arbeit zur Bewegung eines Körpers von \mathbf{x}_1 nach \mathbf{x}_2 unabhängig vom Weg in U , der diese Punkte verbindet, gleich der Differenz der potenziellen Energien

$$W(\mathbf{x}_1, \mathbf{x}_2) = V(\mathbf{x}_1) - V(\mathbf{x}_2). \quad (3.30)$$

Zum Beweis integrieren wir entlang einer beliebigen (stetig differenzierbaren) Kurve mit Anfangspunkt $\mathbf{x}(a) = \mathbf{x}_1$ und Endpunkt $\mathbf{x}(b) = \mathbf{x}_2$:

$$\begin{aligned}
 \int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} &= \int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \frac{d\mathbf{x}(t)}{dt} dt \\
 &= - \int_a^b \nabla V(\mathbf{x}(t)) \cdot \frac{d\mathbf{x}(t)}{dt} dt \quad (\text{Kettenregel}) \\
 &= - \int_a^b \frac{d}{dt} V(\mathbf{x}(t)) dt \\
 &= V(\mathbf{x}(a)) - V(\mathbf{x}(b)).
 \end{aligned} \tag{3.31}$$

□

Beachte: Wenn W positiv ist, so ist die potenzielle Energie am Ende kleiner: $V(\mathbf{x}_2) < V(\mathbf{x}_1)$.

Die Funktion

$$H(\mathbf{x}, \mathbf{v}) = \frac{mv^2}{2} + V(\mathbf{x}) \tag{3.32}$$

(mit $v^2 = \mathbf{v} \cdot \mathbf{v}$) beschreibt die Gesamtenergie eines Massenpunkts im Potenzialfeld V . Bei einer mechanischen Bewegung $t \rightarrow \mathbf{x}(t)$ bleibt die Energie konstant:

$$\frac{d}{dt} H(\mathbf{x}(t), \mathbf{v}(t)) = 0. \tag{3.33}$$

(Nachrechnen!)

Mechanische Systeme und Gleichgewichtslagen

Nachdem wir im vorherigen Kapitel einige Grundlagen der Mechanik eines Massenpunktes betrachtet haben, wenden wir uns nun den Erscheinungen zu, die bei einem aus mehreren Teilen zusammengesetzten Systems zu beobachten sind. Wir beginnen mit der Untersuchung von stationären Gleichgewichtslagen.

4.1. Durch Federkräfte verbundene Massen

Wir suchen ein Modell zur Beschreibung der Gleichgewichtslagen eines Systems von Körpern (Massen), die durch elastische Elemente verbunden sind und dem Einfluss äußerer Kräfte ausgesetzt sind. Die elastischen Elemente werden wir durch Federkräfte beschreiben.

Der Einfachheit halber betrachten wir ein eindimensionales System aus n Massen m_1, m_2, \dots, m_n und $n + 1$ Federn, wie es in Abbildung 4.1 dargestellt ist.

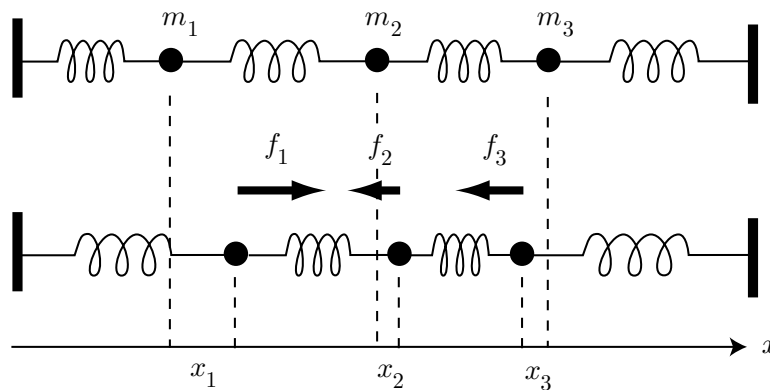


ABBILDUNG 4.1. Durch Federkräfte verbundene Massen in Ausgangslage (oben) und in einer neuen Gleichgewichtslage aufgrund angreifender äußerer Kräfte.

Wir nehmen an, daß sich die Massen nur in einer Dimension bewegen können (der x -Richtung). Anfänglich bilden wir eine Kette aus den Massen und Federn und legen diese einfach aus, sodaß das System kräftefrei ist und sich daher in einer Gleichgewichtslage befindet (oberer Teil der Abbildung 4.1). Wir denken uns die freien Enden der äußeren Federn fixiert (in der Abbildung durch dicke senkrechte Striche symbolisiert). Zunächst sind also alle Federn entspannt (weder gedehnt noch gestaucht).

Eine neue Gleichgewichtslage entsteht, wenn äußere Kräfte an den einzelnen Massen angreifen, also zusätzliche Kräfte, die nicht in den Federn erzeugt werden. Diese Kräfte beschreiben wir durch einen Vektor

$$\mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}, \quad \text{wobei } f_i = \text{äußere Kraft auf } m_i. \quad (4.1)$$

In der Abbildung sind die Komponenten von \mathbf{f} durch Pfeile symbolisiert. Hier ist f_1 positiv, f_2 und f_3 sind negativ. Eine positive Kraft (Pfeil nach rechts) will die Masse, an der sie angreift, nach rechts bewegen, eine negative Kraft will die Masse nach links ziehen.

Aufgrund der angreifenden Kräfte werden sich die Massen verschieben. Dabei werden die Federn gedehnt oder gestaucht, sodass Federkräfte auftreten werden. Wenn die Federkräfte und die äußeren Kräfte einander kompensieren, haben wir eine neue Gleichgewichtslage erreicht. Die Bedingung für die Gleichgewichtslage ist, dass für jede einzelne Masse die Summe aller an ihr angreifenden Kräfte Null ist.

Wir beschreiben eine beliebige *Systemkonfiguration* durch den Vektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{wobei } x_i = \text{Auslenkung von } m_i \text{ aus der Ruhelage.} \quad (4.2)$$

Die ursprüngliche Gleichgewichtslage ohne äußere Kräfte ist also durch den Nullvektor $\mathbf{x} = \mathbf{0}$ gegeben.

Bei einer Konfiguration $\mathbf{x} \neq \mathbf{0}$ sind im allgemeinen die Abstände zwischen den Massen gegenüber der Ausgangslage $\mathbf{x} = \mathbf{0}$ verändert, wodurch auch die einzelnen Federn deformiert werden. Wenn wir beachten, dass die Randpunkte des Systems fixiert sind, können wir schreiben

$$x_0 = x_{n+1} = 0. \quad (4.3)$$

Mit dieser Festsetzung kann die Längenänderung (*Elongation*) der i -ten Feder offenbar als

$$e_i = x_i - x_{i-1} \quad \text{für alle } i = 1, \dots, n+1 \quad (4.4)$$

geschrieben werden. Dieser Ausdruck kann positiv (Dehnung) oder negativ (Stauchung) sein.

Wir können nun die Elongationen e_i der Federn zu einem $n+1$ -dimensionalen Vektor zusammenfassen und schreiben

$$\mathbf{e} = \mathbf{A} \mathbf{x}, \quad (4.5)$$

oder

$$\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \\ e_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & 0 & \vdots \\ 0 & -1 & 1 & \vdots \\ \vdots & & \ddots & \ddots \\ \vdots & & & -1 & 1 \\ 0 & \dots & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \quad (4.6)$$

Die hier definierte Matrix \mathbf{A} ist offenbar eine $(n+1) \times n$ -Matrix.

Die Längenänderung e_i bewirkt in der i -ten Feder eine dazu proportionale Kraft y_i . Wenn c_i die Federkonstante der i -ten Feder ist (eine positive reelle Größe), dann gilt

$$y_i = c_i e_i \quad \text{für } i = 1, \dots, n+1.$$

Das ist das *Hook'sche Gesetz* für linear-elastische Materialien. Wir können diese $n+1$ Gleichungen zu einer Matrixgleichung zusammenfassen:

$$\mathbf{y} = \mathbf{C} \mathbf{e}. \quad (4.7)$$

Hier ist \mathbf{C} eine positiv definite $(n+1) \times (n+1)$ -dimensionale Diagonalmatrix, und \mathbf{y} ein $(n+1)$ -dimensionaler Vektor, der die in den Federn auftretenden Kräfte beschreibt.

Dabei ist es wichtig, sich der Bedeutung des Vorzeichens einer Federkraft bewusst zu sein. In der Gleichung (4.7) ist y_i positiv, wenn die Feder i gedehnt wird ($e_i > 0$), und negativ, wenn sie gestaucht wird ($e_i < 0$). Diesen Umstand müssen wir nämlich beachten, wenn wir die auf die Masse m_i einwirkenden Kräfte beschreiben wollen:

Falls die Feder i gedehnt ist, ($y_i > 0$), möchte sie sich zusammenziehen. Auf die rechts davon montierte Masse m_i wirkt diese Feder daher mit einer negativen (in negative x -Richtung wirkenden) Kraft $-y_i$. Analog: Wenn die i -te Feder gestaucht ist ($e_i < 0$ und $y_i < 0$) ist die Kraft auf m_i positiv und wird auch durch $-y_i$ beschrieben. Die $i+1$ -ste Feder hat die Masse m_i hingegen zu ihrer Linken und wirkt auf diese mit der Kraft $+y_{i+1}$.

Die Gesamtkraft auf die Masse m_i ist also

$$\text{Kraft auf } m_i = f_i - y_i + y_{i+1} \quad (4.8)$$

Gleichgewichtsbedingung: Die Ruhelage des Systems mit äußeren Kräften ist definiert durch die Bedingung, dass sich die an den einzelnen Massen angreifenden Kräfte gegenseitig aufheben. Die von außen angreifenden Kräfte und die inneren Federkräfte des Systems halten sich die Waage:

$$f_i = y_i - y_{i+1}, \quad i = 1, \dots, n, \quad (\text{Gleichgewichtsbedingung}). \quad (4.9)$$

Wir können diese n Gleichungen wieder zu einer Matrixgleichung zusammenfassen und erhalten

$$\mathbf{f} = \mathbf{A}^\top \mathbf{y}. \quad (4.10)$$

Zusammenfassung:

$$\begin{array}{ll} \mathbf{e} = \mathbf{A} \mathbf{x} & \text{Elongationen} \\ \mathbf{y} = \mathbf{C} \mathbf{e} & \text{innere Kräfte} \\ \mathbf{f} = \mathbf{A}^\top \mathbf{y} & \text{äußere Kräfte} \end{array}$$

also gilt

$$\mathbf{f} = \mathbf{K} \mathbf{x} \quad \text{mit} \quad \mathbf{K} = \mathbf{A}^\top \mathbf{C} \mathbf{A}. \quad (4.11)$$

Die Matrix \mathbf{K} heißt *Steifigkeits-Matrix*. Sie ist symmetrisch, dh. $\mathbf{K}^\top = \mathbf{K}$. Außerdem ist \mathbf{K} positiv definit, dh. alle Eigenwerte von \mathbf{K} sind reell und strikt positiv (denn die Matrix \mathbf{C} hat diese Eigenschaft und $\mathbf{A}^\top \mathbf{C} \mathbf{A}$ hat dieselben Eigenwerte wie \mathbf{C}).

Bei gegebenen Kräften führt die Lösung der Gleichung $\mathbf{f} = \mathbf{K} \mathbf{x}$ auf die Auslenkungen \mathbf{x} , die die neue Gleichgewichtslage definieren:

$$\mathbf{x} = \mathbf{K}^{-1} \mathbf{f}. \quad (4.12)$$

AUFGABE 4.1. *Wie ändert sich die Beschreibung, wenn in der Ausgangslage $\mathbf{x} = \mathbf{0}$ (Gleichgewichtslage ohne äußere Kräfte) die Federn bereits als gespannt angenommen werden?*

AUFGABE 4.2. *Beschreibe die Gleichgewichtslagen des in Abbildung 4.2 dargestellten Systems mit der in diesem Kapitel beschriebenen Technik. Die äußeren Kräfte werden dabei durch die Schwerkraft erzeugt.*

4.2. Die potenzielle Energie des Federsystems

Wir betrachten nun das Federsystem in einer beliebigen Lage \mathbf{x} (die nicht der Gleichgewichtslage entsprechen muß). Die gegebenen konstanten Kräfte \mathbf{f}_i haben bei der Auslenkung der Massen m_i aus der ursprünglichen Position um die Strecke x_i jeweils die Arbeit $f_i x_i$ (Kraft mal Weg) bereitgestellt. Insgesamt steckt also in der Positionsänderung aller Körper die Energiemenge

$$\mathbf{x}^\top \mathbf{f} = \sum_{i=1}^n x_i f_i = \mathbf{f}^\top \mathbf{x} \quad (4.13)$$

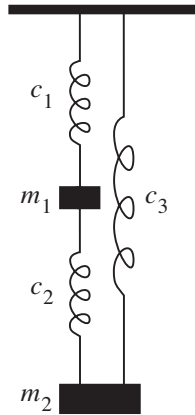


ABBILDUNG 4.2.

Die Kurzschreibweise verwendet die Matrix-Multiplikation: Zeilenvektor mal Spaltenvektor.

Der Beitrag dieser Energie zur potenziellen Energie des Systems wird negativ gerechnet. Begründung: Wenn x_i und f_i beide positiv sind, ist die neue Lage entstanden, indem das Teilchen der Krafrichtung folgte. Man müsste also Arbeit ins System hineinstecken, wollte man gegen die Kraft f_i die Ausgangslage $x_i = 0$ wiederherstellen. Die potenzielle Energie in der Lage x_i ist daher sicher geringer als in der Ausgangslage. Daher gilt für das System in der Lage \mathbf{x} ,

$$-\mathbf{x}^\top \mathbf{f} = \text{Beitrag der äußeren Kräfte zur pot. Energie des Systems.} \quad (4.14)$$

Die Federkräfte lieferten ebenfalls einen Beitrag zur potenziellen Energie. Der Betrag der Kraft der i -ten Feder bei einer Elongation s ist

$$\text{Kraft} = c_i s = \text{Federkonstante} \times \text{Elongation} \quad (4.15)$$

daher gilt

$$\text{Arbeit} = \int_{s=0}^{s=e_i} c_i s ds = \frac{c_i e_i^2}{2}. \quad (4.16)$$

Diese gegen die Federkräfte ins System gesteckte Arbeit bleibt als positive potenzielle Energie im System. Durch Rückführung in die Ausgangslage $x = 0$ erhält man diese Energie zurück. Es ist also

$$\sum_{i=1}^{n+1} \frac{c_i e_i^2}{2} = \frac{1}{2} \mathbf{e}^\top \mathbf{C} \mathbf{e} = \frac{1}{2} \mathbf{x}^\top \mathbf{A}^\top \mathbf{C} \mathbf{A} \mathbf{x} \quad (4.17)$$

die gesamte potenzielle Energie, die aufgrund der Dehnung der $n + 1$ Federn bei der Position \mathbf{x} im System gespeichert ist. $\mathbf{K} = \mathbf{A}^\top \mathbf{C} \mathbf{A}$ ist wieder die Steifigkeits-Matrix.

Zusammenfassung:

Die gesamte potenzielle Energie im Federsystem aufgrund der Verschiebung \mathbf{x} ist

$$P(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x} - \mathbf{x}^\top \mathbf{f}. \quad (4.18)$$

Dabei ist \mathbf{K} die Steifigkeits-Matrix und \mathbf{f} der Vektor der äußeren Kräfte.

Es ist naheliegend, dass wir wieder die Gleichgewichtslage erhalten, wenn wir diejenige Auslenkung \mathbf{x} suchen, für die die potenzielle Energie des Federsystems ein Minimum hat. Das beweisen wir im nächsten Abschnitt.

4.3. Das Minimum der potenziellen Energie

SATZ 4.1. *Sei \mathbf{C} eine symmetrische, positiv definite $m \times m$ -Matrix. Sei \mathbf{A} eine $m \times n$ -Matrix mit n linear unabhängigen Spalten.*

Dann ist die $n \times n$ -Matrix $\mathbf{K} = \mathbf{A}^\top \mathbf{C} \mathbf{A}$ symmetrisch und positiv definit.

Bemerkungen:

- (1) \mathbf{C} symmetrisch bedeutet $\mathbf{C}^\top = \mathbf{C}$.
- (2) Eine symmetrische Matrix läßt sich immer diagonalisieren. Alle Eigenwerte sind reell.
- (3) Eine symmetrische Matrix \mathbf{C} heißt positiv definit, wenn alle Eigenwerte > 0 sind. Es gilt dann $\mathbf{x}^\top \mathbf{C} \mathbf{x} > 0$ für alle $\mathbf{x} \neq 0$.
- (4) Eine $m \times n$ -Matrix hat m Zeilen und n Spalten. Sollen die Spalten linear unabhängig sein, muß $n \leq m$ gelten. Die Spaltenvektoren sind ja alle m -dimensional und spannen daher höchstens einen m -dimensionalen Raum auf. Allgemein gilt für eine Matrix: Anzahl der l.u. Spalten = Anzahl der l.u. Zeilen.

BEWEIS. Symmetrie: Wegen $\mathbf{C}^\top = \mathbf{C}$ und $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$ gilt

$$\mathbf{K}^\top = (\mathbf{A}^\top \mathbf{C} \mathbf{A})^\top = \mathbf{A}^\top \mathbf{C}^\top \mathbf{A}^{\top\top} = \mathbf{A}^\top \mathbf{C} \mathbf{A} = \mathbf{K}. \quad (4.19)$$

Positive Definitheit:

$$\mathbf{x}^\top \mathbf{K} \mathbf{x} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{C} \mathbf{A} \mathbf{x} = (\mathbf{A} \mathbf{x})^\top \mathbf{C} (\mathbf{A} \mathbf{x}) = \mathbf{y}^\top \mathbf{C} \mathbf{y} > 0, \quad (4.20)$$

für alle $\mathbf{y} \neq 0$, da ja laut Voraussetzung die Matrix \mathbf{C} positiv definit ist. Weiters ist $\mathbf{y} = \mathbf{A} \mathbf{x} = 0$ nur für $\mathbf{x} = 0$, denn:

$$\mathbf{A} \mathbf{x} = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \vec{v}_1 + \cdots + x_n \vec{v}_n, \quad (4.21)$$

was nur dann Null ist, wenn alle x_i Null sind, denn die Spalten \vec{v}_i der Matrix A wurden als linear unabhängig vorausgesetzt. Wir haben also $\mathbf{x}^\top \mathbf{K} \mathbf{x} > 0$ immer dann, wenn $\mathbf{x} \neq 0$ ist, also die positive Definitheit. \square

SATZ 4.2. Sei \mathbf{K} eine positiv definite $n \times n$ -Matrix, und $\mathbf{f} \in \mathbb{R}^n$. Dann hat die Funktion $P : \mathbb{R}^n \rightarrow \mathbb{R}$, definiert durch

$$P(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x} - \mathbf{x}^\top \mathbf{f} \quad (4.22)$$

das (eindeutige, globale) Minimum an der Stelle $\mathbf{x} = \mathbf{K}^{-1}\mathbf{f} \in \mathbb{R}^n$. Das Minimum ist also die (eindeutige) Lösung von

$$\mathbf{K} \mathbf{x} = \mathbf{f}. \quad (4.23)$$

Der Minimalwert von P ist

$$P(\mathbf{K}^{-1}\mathbf{f}) = -\frac{1}{2} \mathbf{f}^\top \mathbf{K}^{-1} \mathbf{f}. \quad (4.24)$$

BEWEIS. Sei \mathbf{x} die Lösung von $\mathbf{K} \mathbf{x} = \mathbf{f}$. Da \mathbf{K} positiv definit ist, ist \mathbf{K} invertierbar und die Lösung \mathbf{x} ist eindeutig durch $\mathbf{x} = \mathbf{K}^{-1}\mathbf{f}$ gegeben. Wir müssen zeigen:

$$P(\mathbf{y}) > P(\mathbf{x}) \quad \text{für } \mathbf{y} \neq \mathbf{x}. \quad (4.25)$$

Dazu berechnen wir

$$P(\mathbf{y}) - P(\mathbf{x}) = \frac{1}{2} \mathbf{y}^\top \mathbf{K} \mathbf{y} - \mathbf{y}^\top \mathbf{f} - \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x} + \mathbf{x}^\top \mathbf{f} \quad (4.26)$$

$$= \frac{1}{2} \mathbf{y}^\top \mathbf{K} \mathbf{y} - \mathbf{y}^\top \mathbf{K} \mathbf{x} - \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x} + \mathbf{x}^\top \mathbf{K} \mathbf{x} \quad (4.27)$$

$$= \frac{1}{2} \mathbf{y}^\top \mathbf{K} \mathbf{y} - \mathbf{y}^\top \mathbf{K} \mathbf{x} + \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x}. \quad (4.28)$$

Andererseits ist

$$\frac{1}{2} (\mathbf{y}^\top - \mathbf{x}^\top) \mathbf{K} (\mathbf{y} - \mathbf{x}) = \frac{1}{2} \mathbf{y}^\top \mathbf{K} \mathbf{y} - \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{y} - \frac{1}{2} \mathbf{y}^\top \mathbf{K} \mathbf{x} + \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x} \quad (4.29)$$

$$= \frac{1}{2} \mathbf{y}^\top \mathbf{K} \mathbf{y} - \mathbf{y}^\top \mathbf{K} \mathbf{x} + \frac{1}{2} \mathbf{x}^\top \mathbf{K} \mathbf{x}, \quad (4.30)$$

denn für eine symmetrische Matrix \mathbf{K} gilt

$$\mathbf{y}^\top \mathbf{K} \mathbf{x} = (\mathbf{K}^\top \mathbf{y})^\top \mathbf{x} = (\mathbf{K} \mathbf{y})^\top \mathbf{x} = \mathbf{x}^\top \mathbf{K} \mathbf{y}. \quad (4.31)$$

Wir haben also

$$P(\mathbf{y}) - P(\mathbf{x}) = \frac{1}{2} (\mathbf{y}^\top - \mathbf{x}^\top) \mathbf{K} (\mathbf{y} - \mathbf{x}) > 0, \quad (4.32)$$

außer für $\mathbf{y} = \mathbf{x}$, da \mathbf{K} als positiv definit vorausgesetzt war. Daraus folgt (4.25) und $\mathbf{x} = \mathbf{K}^{-1}\mathbf{f}$ ist tatsächlich Minimum von P .

Der Wert $P(\mathbf{x})$ an der Stelle $\mathbf{x} = \mathbf{K}^{-1}\mathbf{f}$ ist

$$P_{\min} = \frac{1}{2} (\mathbf{K}^{-1}\mathbf{f})^\top \mathbf{K} (\mathbf{K}^{-1}\mathbf{f}) - (\mathbf{K}^{-1}\mathbf{f})^\top \mathbf{f} \quad (4.33)$$

$$= \frac{1}{2} \mathbf{f}^\top (\mathbf{K}^{-1})^\top \mathbf{K} \mathbf{K}^{-1} \mathbf{f} - \mathbf{f}^\top (\mathbf{K}^{-1})^\top \mathbf{f} \quad (4.34)$$

$$= \frac{1}{2} \mathbf{f}^\top \mathbf{K}^{-1} \mathbf{K} \mathbf{K}^{-1} \mathbf{f} - \mathbf{f}^\top \mathbf{K}^{-1} \mathbf{f} \quad (4.35)$$

$$= \frac{1}{2} \mathbf{f}^\top \mathbf{K}^{-1} \mathbf{f} - \mathbf{f}^\top \mathbf{K}^{-1} \mathbf{f} \quad (4.36)$$

$$= -\frac{1}{2} \mathbf{f}^\top \mathbf{K}^{-1} \mathbf{f} \quad (4.37)$$

Dabei haben wir benützt dass für symmetrische Matrizen \mathbf{K} auch die Matrix \mathbf{K}^{-1} symmetrisch ist, also $(\mathbf{K}^{-1})^\top = \mathbf{K}^{-1}$. \square

Beachte: In einer Dimension ist

$$P(x) = \frac{1}{2} k x^2 - f x \quad (4.38)$$

eine Parabel. In mehreren Dimensionen beschreibt die Funktion $\mathbf{x} \rightarrow P(\mathbf{x})$ ein Paraboloid. Die Lösung des linearen Gleichungssystems $\mathbf{K}\mathbf{x} = \mathbf{f}$ ist äquivalent mit dem Auffinden des Minimums von P (Optimierungsaufgabe, Extremwertaufgabe). In unserer Anwendung auf das Federsystem ist das lineare Gleichungssystem eine Kräftebilanzgleichung, die quadratische Funktion P ist die potenzielle Energie der Konfiguration. Die Gleichgewichtslage des Federsystems ist daher genau jene Konfiguration, in der die potenzielle Energie P ihr Minimum hat.

4.4. Gleichgewicht kontinuierlicher Systeme - ein elastischer Stab

4.4.1. Problemstellung

Wir betrachten einen dünnen, elastischen Stab. Er sei mit einem Ende an der Decke befestigt und hängt senkrecht herab. Unter Einfluß der Schwerkraft wird der Stab gedehnt. Wir wollen wissen, wie stark diese Dehnung an verschiedenen Stellen des Stabes ist.

Um die Dehnung des Stabes zu beschreiben, verwenden wir ein eindimensionales Modell. Wir vernachlässigen somit alle Effekte, die mit dem Querschnitt des Stabes zu tun haben (z.B. die Änderung des Querschnittes bei Längsdehnung). Die Koordinate z beschreibe die Entfernung von der (bei $z = 0$ befestigten) oberen Stabspitze.

Ein Punkt des Stabes, der ohne Krafteinwirkung an der Stelle z wäre, befinde sich unter Einfluß der Schwerkraft an der Stelle $z + u(z)$. Die Größe

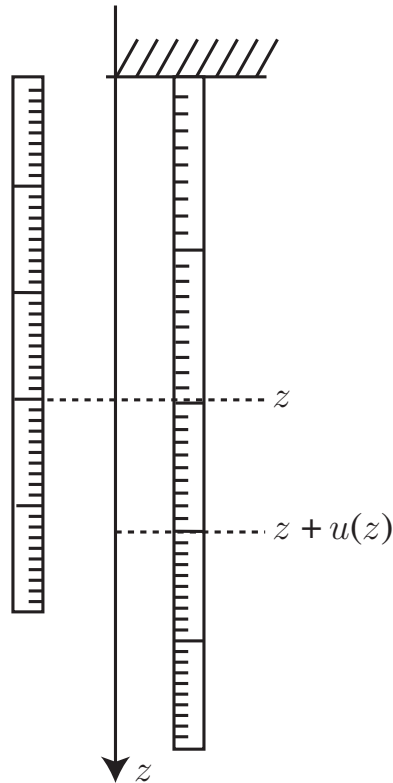


ABBILDUNG 4.3. Durch Schwerkraft gedehnter elastischer Stab im Vergleich zum ungedehnten Stab.

$u(z)$ beschreibt die Verschiebung an der Stelle z . Siehe Abbildung 4.3. Unser Modell soll die Bestimmung dieser Größe erlauben.

4.4.2. Modellbildung

Zur Beschreibung von kontinuierlich verteilten physikalischen Größen, verwenden wir Dichten. Zum Beispiel erhält man eine Massendichte, wenn man eine in einem Volumen enthaltene Masse durch dieses Volumen dividiert. In einer Dimension ist die Massendichte Masse/Länge.

Analog definiert man (in einer Dimension) die Kraftdichte als Kraft pro Länge. Diese Größe wird oft auch als (mechanische) Spannung bezeichnet. Die zweidimensionale Kraftdichte, also Kraft/Fläche bezeichnet man auch als Druck.

Beim hängenden Stab beschreiben wir den Einfluß der Gravitation durch eine Kraftdichte f . Die an einem kleinen Längenstück $[z, z + \Delta z]$ angreifende

Gewichtskraft ist dann ungefähr $f(z) \Delta z$, genauer:

$$F = \int_z^{z+\Delta z} f(\zeta) d\zeta. \quad (4.39)$$

Sie hängt davon ab, wieviel Stabmasse in diesem Längenstück vorhanden ist. Wenn der Stab an der Stelle z die Massendichte (Masse/Länge) $\mu(z)$ hat, ist die dort wirkende Dichte der Gewichtskraft

$$f(z) = \mu(z) g \quad (4.40)$$

und somit erhalten wir

$$F = \left(\int_z^{z+\Delta z} \mu(\zeta) d\zeta \right) g = \text{Gesamtmasse im Intervall} \times \text{Erdbeschleunigung}.$$

Die Dehnung des Stabes an einem Punkt z wird durch die Ableitung der Verschiebungsfunktion $u(z)$ gemessen:

$$\frac{du(z)}{dz} \approx \frac{u(z + \Delta z) - u(z)}{\Delta z} \quad (4.41)$$

Die Dehnung ist eine Elongation pro Länge. Dort, wo u konstant ist, ist das Material ungedehnt und $du/dz = 0$.

In einem elastischen Material entsteht aufgrund der Dehnung an der Stelle z eine Kraft. Diese Kraft erhält man aus

$$w(z) = c(z) \frac{du(z)}{dz} \quad (4.42)$$

wobei $c(z)$ eine Größe ist, die die Elastizität des Stabmaterials an der Stelle z beschreibt. Sie wird Elastizitätskoeffizient genannt. (Beachte: $w(z)$ ist eine Kraft, keine Kraftdichte). Aufgrund der Physik ist es vernünftig, anzunehmen, dass $c(z) > 0$ für alle z mit $0 \leq z \leq L$ gilt (L ist die Länge des Stabes).

Wir betrachten nun die Kräftebilanz eines sehr kurzen Stückes des elastischen Stabes, zB für das Stück zwischen zwei Markierungen an den Stellen z und $z + \Delta z$. Das Stück hat die (ungedehnte) Länge Δz und die dort angreifende Gewichtskraft ist (unabhängig von der Dehnung) gleich Masse zwischen z und $z + \Delta z$ mal Erdbeschleunigung g , also ungefähr gleich

$$F = \int_z^{z+\Delta z} f(z) dz \approx f(z) \Delta z = \mu(z) g \Delta z. \quad (4.43)$$

Diese Kraft zieht nach unten, in die positive z -Richtung, ist also positiv.

Im Gleichgewicht (also wenn das betrachtete Stück des Stabes in Ruhe ist) wird die Gewichtskraft durch die elastischen Kräfte an den Rändern

des Stückes kompensiert. Die an der Stelle z nach oben (also zu kleineren z -Werten hin) ziehende Kraft ist

$$-w(z) = -c(z) \frac{du(z)}{dz}. \quad (4.44)$$

Die an der Stelle $z + \Delta z$ nach unten (also zu größeren z -Werten hin) ziehende Kraft ist hingegen positiv zu rechnen:

$$w(z + \Delta z) = c(z + \Delta z) \left. \frac{du}{dz} \right|_{z+\Delta z}. \quad (4.45)$$

Die Kräftebilanz für das Stück des Stabes lautet somit:

$$w(z + \Delta z) - w(z) + f(z) \Delta z = 0, \quad (4.46)$$

oder

$$\frac{w(z + \Delta z) - w(z)}{\Delta z} + f(z) = 0, \quad (4.47)$$

und wenn wir Δz gegen Null gehen lassen,

$$\frac{dw(z)}{dz} + f(z) = 0. \quad (4.48)$$

Verwenden wir die Definition von $w(z)$, erhalten wir die folgende Differentialgleichung für u :

$$-\frac{d}{dz} \left(c(z) \frac{d}{dz} u(z) \right) = f(z). \quad (4.49)$$

Wir sollten noch die Randbedingungen notieren, die sich aus der Problemstellung ergeben. Der Stab ist oben angeklebt, also ist $u(0) = 0$. Am unteren Ende hängt kein weiteres Stück mehr, das dort noch eine Dehnung verursachen könnte, daher ist $w(L) = 0$, wenn L die Länge des (ungedehnten) Stabes ist.

Diese Differentialgleichung und die dazugehörigen Randbedingungen sind also das mathematische Modell, das den von der Decke hängenden elastischen Stab beschreibt. Wir können diese Differentialgleichung durch Integration lösen.

Integrieren wir beide Seiten von (4.49) von $z = x$ bis $z = L$, erhalten wir

$$\int_x^L -\frac{dw(z)}{dz} dz = -(w(L) - w(x)) = \int_x^L f(z) dz. \quad (4.50)$$

Wegen der Randbedingung $w(L) = 0$ erhalten wir

$$w(x) = c(z) \left. \frac{du(z)}{dz} \right|_{z=x} \equiv c(x) \frac{du(x)}{dx} = \int_x^L f(z) dz \quad (4.51)$$

und da $c(x) > 0$ ist,

$$\frac{du(x)}{dx} = \frac{1}{c(x)} \int_x^L f(z) dz. \quad (4.52)$$

Somit erhalten wir die Formel (unter Berücksichtigung von $u(0) = 0$)

$$u(x) = u(0) + \int_0^x \frac{du(y)}{dy} dy = \int_0^x \frac{1}{c(y)} \left(\int_y^L f(z) dz \right) dy. \quad (4.53)$$

4.4.3. Beispiel

Balken mit Einheitslänge $L = 1$, und $f(z) = 1$, $c(z) = 1$ für $0 \leq z \leq 1$:

$$u(x) = \int_0^x \frac{1}{c(y)} \left(\int_y^1 dz \right) dy = \int_0^x (1 - y) dy = x - x^2/2 \quad (4.54)$$

und das ist dann auch tatsächlich die eindeutige Lösung von

$$-u''(x) = 1, \quad u(0) = 0, \quad \left. \frac{du(x)}{dx} \right|_{x=1} = 0. \quad (4.55)$$

4.4.4. Modellvarianten

Wenn beide Enden des Stabes fixiert sind, kann man das offenbar durch die Randbedingungen $u(0) = 0$ und $u(L) = 0$ beschreiben. Die Randbedingungen beschreiben die physikalische Situation:

$$\begin{aligned} u(0) = 0 & \quad \text{Befestigung des oberen Stabendes} \\ u(L) = b & \quad \text{Befestigung unten an der Stelle } z = L + b \\ c(L) \left. \frac{du(z)}{dz} \right|_{z=L} = G & \quad \text{Gewicht } G \text{ ist am unteren Ende befestigt} \end{aligned}$$

In der obigen Rechnung haben wir bei der Integration die Randbedingungen

$$u(0) = 0 \quad \text{und} \quad w(L) = 0, \quad \text{also} \quad \left. \frac{du(z)}{dz} \right|_{z=L} = 0$$

verwendet. Andere Randbedingungen führen zu anderen Lösungsformeln.

4.4.5. Vergleich mit der Federkette

Die Gleichgewichtsbedingung für die Federkette

$$\mathbf{A}^\top \mathbf{C} \mathbf{A} \mathbf{x} = \mathbf{f} \quad (4.56)$$

ist sehr ähnlich zu der Gleichung

$$-\frac{d}{dz} \left(c(z) \frac{d}{dz} u(z) \right) = f(z). \quad (4.57)$$

Wir können folgende Gegenüberstellung machen:

diskretes System	\longleftrightarrow	kontinuierliches System
Vektor \mathbf{x}	\longleftrightarrow	Funktion u
Elongationen $\mathbf{e} = \mathbf{A}\mathbf{x}$	\longleftrightarrow	Dehnung $\frac{d}{dz}u(z)$
Diagonalmatrix \mathbf{C}	\longleftrightarrow	Multiplikation mit $c(z)$
innere Kräfte $\mathbf{y} = \mathbf{C}\mathbf{e}$	\longleftrightarrow	Dehnungskraft $w(z) = c(z)\frac{d}{dz}u(z)$
Matrix \mathbf{A}	\longleftrightarrow	Ableitungsoperator $\frac{d}{dz}$
transponierte Matrix \mathbf{A}^\top	\longleftrightarrow	Ableitungsoperator $-\frac{d}{dz}$
äußere Kräfte $\mathbf{A}^\top\mathbf{y} = \mathbf{f}$	\longleftrightarrow	Kraftdichte $-\frac{d}{dz}w(z) = f(z)$

4.4.6. Inneres Produkt im kontinuierlichen System

Warum entspricht der Matrix \mathbf{A} der Ableitungsoperator, aber der transponierten Matrix \mathbf{A}^\top das Negative des Ableitungsoperators?

Die Rolle der transponierten Matrix ist immer im Zusammenhang mit einem inneren Produkt zu sehen. Das innere Produkt zweier (n -dimensionaler) Spaltenvektoren ist

$$\mathbf{y}^\top \mathbf{x} = \sum_{i=1}^n y_i x_i.$$

Eine Matrix \mathbf{A} beschreibt eine lineare Abbildung, die Vektoren \mathbf{x} auf Vektoren $\mathbf{e} = \mathbf{A}\mathbf{x}$ abbildet. Die transponierte Matrix ermöglicht die Beschreibung der linearen Abbildung "auf der anderen Seite des inneren Produkts", denn

$$\mathbf{y}^\top (\mathbf{A}\mathbf{x}) = (\mathbf{A}^\top \mathbf{y})^\top \mathbf{x}.$$

In unserem kontinuierlichen Beispiel entspricht dem Vektor \mathbf{x} eine Funktion u , die auf dem Intervall $[0, L]$ definiert ist. Dem diskreten Index i , der die Komponenten des Vektors \mathbf{x} durchnummeriert, entspricht nun die kontinuierliche Variable x , von der die Funktionswerte abhängen. Dementsprechend ist das natürliche Analogon zur Summe über den Index das Integral über x : Wir können folgende Gegenüberstellung machen:

diskretes System	\longleftrightarrow	kontinuierliches System
Inneres Produkt $\sum_{i=1}^n y_i x_i$	\longleftrightarrow	$\int_0^L u(x) w(x) dx$

Im kontinuierlichen Fall bezeichnen wir das innere Produkt mit Klammern:

$$\langle u, w \rangle = \int_0^L u(x) w(x) dx. \quad (4.58)$$

Der Vektorraum besteht nun aus Funktionen, die auf dem Intervall $[0, L]$ stetig sind (und damit integrierbar). Die Summe zweier Funktionen und die Multiplikation einer Funktion mit einem Skalar sind wie üblich definiert. Wir bezeichnen diesen Vektorraum mit $\mathcal{C}([0, L])$. Unter einem linearen Operator versteht man nun eine lineare Abbildung, die auf einem linearen Teilraum $\mathcal{D}(T) \subset \mathcal{C}([0, L])$ des Funktionenraums definiert ist und in den Funktionenraum hinein abbildet:

$$T : \mathcal{D}(T) \subset \mathcal{C}([0, L]) \longrightarrow \mathcal{C}([0, L])$$

Wir können uns leicht davon überzeugen, dass der Ableitungsoperator in diesem Sinn ein linearer Operator ist:

$$T = \frac{d}{dx} \quad \text{auf dem Definitionsbereich} \quad (4.59)$$

$$\mathcal{D}(T) = \{u \in C([0, L]) \mid u \text{ ist stetig differenzierbar auf } [0, L]\}$$

Dabei verstehen wir unter stetiger Differenzierbarkeit auf dem abgeschlossenen Intervall, dass die Funktion auf dem offenen Intervall differenzierbar ist, am Rand die einseitigen Ableitungen existieren und die Ableitungsfunktion stetig ist.

Nun wollen wir untersuchen, was der zu T transponierte Operator sein könnte. Dazu müssen wir die Ableitung auf die andere Seite im Skalarprodukt bringen. Das gelingt mit einer partiellen Integration. Für alle stetig differenzierbaren Funktionen u und w erhalten wir:

$$\begin{aligned} \left\langle \frac{d}{dx} u, w \right\rangle &= \int_0^L u'(x) w(x) dx = u(x) w(x) \Big|_0^L - \int_0^L u(x) w'(x) dx \\ &= u(L) w(L) - u(0) w(0) + \left\langle u, \left(-\frac{d}{dx}\right) w \right\rangle. \end{aligned}$$

Die störenden Randterme werden wir los, indem wir die Randbedingungen unseres Systems berücksichtigen. In der Anwendung ist die Funktion u die Auslenkung, die am Befestigungspunkt 0 ist. Wir schränken daher den Definitionsbereich von T ein, sodass $u(0) = 0$ ist. Wir betrachten also anstelle des Differentialoperators T aus (4.59) den folgenden Operator A :

$$A = \frac{d}{dx} \quad \text{auf dem Definitionsbereich}$$

$$\mathcal{D}(A) = \{u \in C([0, L]) \mid u \text{ ist stetig differenzierbar und } u(0) = 0\}.$$

A unterscheidet sich von T nur durch die Wahl des Definitionsbereiches. Man beachte, dass auch der Definitionsbereich $\mathcal{D}(A)$ ein linearer Teilraum

von $C([0, L])$ ist. Definieren wir nun den linearen Operator

$$A^\top = -\frac{d}{dx} \quad \text{auf dem Definitionsbereich}$$

$$\mathcal{D}(A^\top) = \{w \in C([0, L]) \mid w \text{ ist stetig differenzierbar und } w(L) = 0\},$$

so gilt für alle $u \in \mathcal{D}(A)$ und alle $w \in \mathcal{D}(A^\top)$ die Beziehung

$$\langle Au, w \rangle = \langle u, A^\top w \rangle,$$

da nun bei der partiellen Integration alle Randterme wegfallen. Der zu d/dx (mit der Zusatzbedingung $u(0) = 0$) transponierte Operator ist also $-d/dx$ auf einem Definitionsbereich mit der Zusatzbedingung $w(L) = 0$.

Achtung: Bei linearen Operatoren in Funktionenräumen muss man immer die Definitionsbereiche mitberücksichtigen.

Lineare Widerstandsnetzwerke

5.1. Grundgesetze

Die Grundelemente der Theorie elektrischer Widerstandsnetzwerke sind die physikalischen Größen Strom und Spannung, und als Bauteile Spannungsquellen, Stromquellen und Widerstände (siehe Abbildung 5.1). Das Verhalten eines Widerstandes wird durch das Ohm'sche Gesetz beschrieben (siehe Abbildung 5.2):

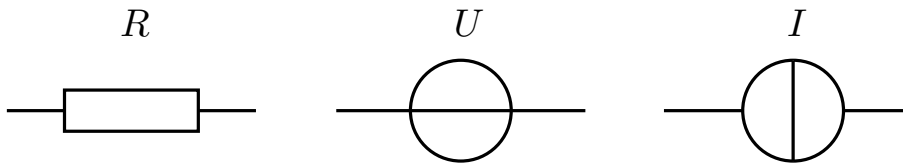


ABBILDUNG 5.1. Schaltungssymbole (europäisch) für Widerstand, Spannungsquelle, Stromquelle.

Ohm'sches Gesetz: An einem Widerstand stehen Strom und Spannung in einem linearen Zusammenhang:

$$U = R I \quad \text{oder} \quad I = G U. \quad (5.1)$$

Die Proportionalitätskonstante R heißt Ohm'scher Widerstand, die Konstante $G = 1/R$ heißt Leitwert.

Widerstände, Strom- und Spannungsquellen können zu Netzwerken zusammengeschlossen werden. Gesucht sind die Spannungen und Stromstärken an den einzelnen Widerständen, die sich im stationären Fall einstellen (nach Abklingen aller "Einschalteffekte"). Das Verhalten von elektrischen Netzwerken im stationären Fall wird durch die beiden Kirchhoff'schen Gesetze geregelt:

Kirchhoff'sches Spannungsgesetz: Die Summe der Spannungen in einer geschlossenen Schleife ("loop") des Netzwerks ist Null.

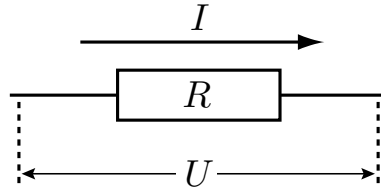


ABBILDUNG 5.2. Das Ohm'sche Gesetz.

Kirchhoff'sches Stromgesetz: Die Summe der Ströme, die in einen Knoten des Netzwerks fließen, ist Null.

Zur Erklärung: Man stelle sich vor, daß in einer zunächst noch offenen Leiterstrecke eine nichtverschwindende Spannung zwischen Anfangs- und Endpunkt herrscht. Wenn man Anfang und Ende nun zu einem Knoten zusammenfügt, und damit das Leiterstück zu einer geschlossenen Schleife macht, zwingt man diese Spannung zu Null (Kurzschluß).

Das Kirchhoff'sche Stromgesetz muss gelten, weil sich durch einen nichtverschwindenden Strom in einen Knoten dort eine Ladung aufbauen würde, was in einem stationären Netzwerk nicht sein kann.

5.2. Beispiel und Theorie

5.2.1. Beschreibung der geometrischen Struktur

Wir betrachten als Beispiel das folgende Netzwerk (Abbildung 5.3):

Ein Netzwerk besteht aus Eckpunkten (Knoten = nodes) und Verbindungslinien (Kanten = edges). In der Abbildung haben wir die vier Knoten und sechs Kanten durchnummeriert. In den Kanten sitzen Widerstände und Spannungsquellen. (In die Knoten werden Ströme auch noch durch Stromquellen eingespeist, siehe weiter unten.) Unsere Aufgabe ist es, die sich einstellenden Spannungen und Ströme zu berechnen.

Der erste Schritt besteht darin, jeder Kante eine (willkürliche) Richtung zuzuordnen. Die topologische Struktur des Netzwerks wird nun durch eine sogenannte "Inzidenzmatrix" A_0 beschrieben. Sie hat eine Zeile für jede Kante und eine Spalte für jeden Knoten. Die Elemente der Matrix A_0 sind die Zahlen 0, +1, und -1. Die Zahl a_{ij} in der Zeile i und Spalte j bestimmt sich daraus, ob die Kante i

- (a) nichts mit dem Knoten j zu tun hat ($a_{ij} = 0$)
- (b) vom Knoten j wegführt ($a_{ij} = -1$)
- (c) zum Knoten j hinführt ($a_{ij} = +1$)

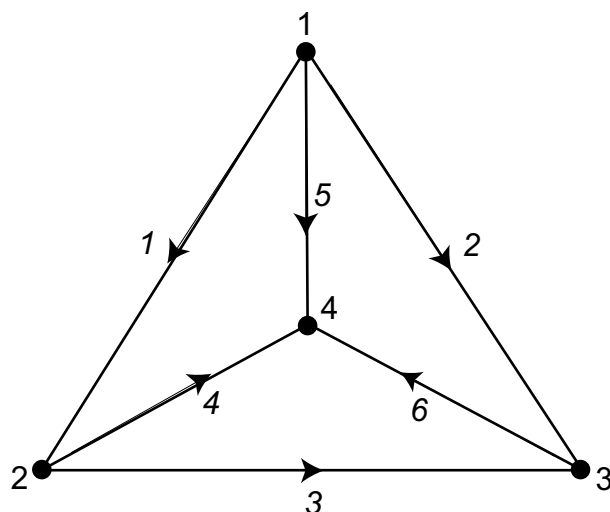


ABBILDUNG 5.3. Ein Netzwerk als gerichteter Graph.

Im durch Abbildung 5.3 definierten Beispiel ist also

$$A_0 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{array}{l} \leftarrow \\ \leftarrow \\ \leftarrow \\ \leftarrow \\ \leftarrow \\ \leftarrow \end{array} \quad \text{Kanten} \tag{5.2}$$

$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ \text{Knoten} & & & \end{array}$$

5.2.2. Berücksichtigung des Spannungsgesetzes

Wir nehmen an, daß die Spannungen an den Widerständen in den Kanten durch Potenzialdifferenzen gegeben sind. Wir ordnen also jedem Knoten i ein (unbekanntes) Potenzial x_i zu. Die Spannung entlang der gerichteten Kante von i zu j sei durch $U_{ij} = x_i - x_j$ gegeben. Dadurch gilt das Kirchhoff'sche Spannungsgesetz automatisch. Betrachte zB die durch die Knoten 1, 2 und 4 gebildete geschlossene Schleife in Abbildung 5.3: Die Summe der Spannungen entlang dieser Schleife ist

$$U_{12} + U_{24} + U_{41} = (x_1 - x_2) + (x_2 - x_4) + (x_4 - x_1) = 0. \tag{5.3}$$

Sei $\mathbf{x} = (x_1, x_2, x_3, x_4)^\top$ der aus den Potenzialwerten gebildete Vektor. Dann beschreibt offenbar

$$\mathbf{A}_0 \mathbf{x} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ x_3 - x_1 \\ x_3 - x_2 \\ x_4 - x_2 \\ x_4 - x_1 \\ x_4 - x_3 \end{pmatrix} \quad (5.4)$$

die Potentialdifferenzen (also die gesuchten Spannungen) entlang der Kanten 1, 2, ..., 6. Genauer: Die Spannungen in den Kanten sind durch die Komponenten des Vektors $-\mathbf{A}_0 \mathbf{x}$ gegeben.

Die Matrix \mathbf{A}_0 hat aber linear abhängige Spalten (die Summe der Spaltenvektoren ist Null). Daher ist die Gleichung $\mathbf{A}_0 \mathbf{x} = \mathbf{b}$ nicht eindeutig nach \mathbf{x} lösbar. Selbst bei gegebenen Spannungen kann man daher die Potenzialwerte an den Knoten nicht berechnen. Hat man eine Lösung (x_1, \dots, x_4) , so ist $(x_1 + c, \dots, x_4 + c)$ eine andere.

Wir können aber einen der Knoten „erden“. Dh., wir setzen das entsprechende Potenzial Null. Damit entfernt man einen nutzlosen Freiheitsgrad (Addition einer Konstanten zu allen x_i). In unserem Beispiel setzen wir (willkürlich)

$$x_4 = 0. \quad (5.5)$$

Wir haben dann

$$\mathbf{A}_0 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \end{pmatrix} = \mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (5.6)$$

mit einer Matrix \mathbf{A} , die wir aus \mathbf{A}_0 erhalten, indem wir die zum geerdeten Knoten gehörige Spalte weglassen:

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (5.7)$$

Die Spalten von \mathbf{A} sind linear unabhängig.

Wir fassen zusammen: Wenn die Kante k vom Knoten i zum Knoten j führt, dann ist die Spannung $U_{ij} = x_i - x_j$ durch die k -te Komponente des Vektors $-\mathbf{A}\mathbf{x}$ gegeben:

$$-(\mathbf{A}\mathbf{x})_k = x_i - x_j.$$

5.2.3. Berechnung der Ströme

In einem Netzwerk mit m Kanten und n (ungeerdeten) Knoten betrachten wir eine gerichtete Kante vom Knoten i zum Knoten j . Der Strom fließt vom höheren Potenzial zum tieferen. Eine positive Spannung $U_{ij} = x_i - x_j$ bewirkt dann einen positiven Strom in dieser Kante. Die Pfeilrichtung gibt also die Richtung an, in der der Strom positiv gezählt wird. Der Strom von i nach j berechnet sich über das Ohm'sche Gesetz: Sei R der Widerstand in dieser Kante, dann ist

$$I = \frac{1}{R}(x_i - x_j). \quad (5.8)$$

In einer Kante erlauben wir außer einem Widerstand noch eine Spannungsquelle (Abbildung 5.4). Wird die Spannungsquelle mit der Stärke U ihrem Minuspol an einen Punkt mit dem Potential x angeschlossen, so herrscht am anderen Pol das Potential $x + U$ (bzw. $x - U$, wenn die Spannungsquelle anders herum angeschlossen wird).

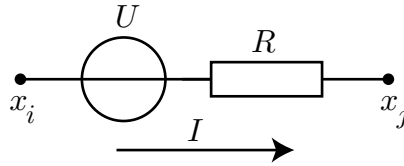


ABBILDUNG 5.4. Strom durch eine Kante mit Spannungsquelle und Widerstand.

In der Abbildung 5.4 herrscht an dem x_i nähergelegenen Anschluß des Widerstandes nicht das Potenzial x_i , sondern das Potenzial $x_i + U$. Die den Strom bestimmende Spannung am Widerstand R ist also $U + x_i - x_j$. Wir erhalten:

$$I = \frac{1}{R}(U + x_i - x_j). \quad (5.9)$$

Diese Stromstärke ändert sich übrigens nicht, wenn man die Reihenfolge von Spannungsquelle und Widerstand in Abbildung 5.4 vertauscht.

Wir bezeichnen mit $\mathbf{b} = (b_1, \dots, b_m)$ den aus den Spannungen der Spannungsquellen gebildete Vektor. b_k ist also die Spannungsquelle in der k -ten Kante (von insgesamt m Kanten). Sei y_k der Strom in der Kante k , die vom Knoten i zum Knoten j führt. In jeder Kante gilt dann

$$y_k = \frac{1}{R_k}(b_k + x_i - x_j) = \frac{1}{R_k}(b_k - (A\mathbf{x})_k). \quad (5.10)$$

Denn die Spannung in der Kante k ist die k -te Komponente des Vektors $-A\mathbf{x}$. Es ist dann der Vektor \mathbf{y} der Stromstärken in den Kanten durch

$$\mathbf{y} = C(\mathbf{b} - A\mathbf{x}) \quad (5.11)$$

gegeben, wobei C die $m \times m$ -Matrix der Leitwerte ist:

$$C = \begin{pmatrix} \frac{1}{R_1} & 0 & & 0 \\ 0 & \frac{1}{R_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{R_m} \end{pmatrix}. \quad (5.12)$$

5.2.4. Berücksichtigung des Stromgesetzes

Betrachten wir nun wieder unser Beispiel, Abbildung 5.3. Berechnen wir

$$A^T \mathbf{y} = \begin{pmatrix} -1 & -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_6 \end{pmatrix}. \quad (5.13)$$

Das ergibt

$$A^T \mathbf{y} = \begin{pmatrix} -y_1 - y_2 - y_5 \\ y_1 - y_3 - y_4 \\ y_2 + y_3 - y_6 \end{pmatrix} \begin{array}{l} \leftarrow \text{Strom in den Knoten 1,} \\ \leftarrow \text{Strom in den Knoten 2,} \\ \leftarrow \text{Strom in den Knoten 3.} \end{array} \quad (5.14)$$

Der Strom in den Knoten 4 ist irrelevant, denn die Erdung kann jeden Strom aufnehmen, ohne dass das restliche Netzwerk beeinflusst wird. In den anderen Knoten aber ist aber das Stromgesetz zu beachten.

Wir erlauben nun, dass in jeden Knoten des Netzwerks Ladung durch eine Stromquelle zugeführt wird. Diese zusätzlichen Ströme müssen in die Bilanz des Kirchhoff'schen Stromgesetzes aufgenommen werden. Sei $\mathbf{f} = (f_1, \dots, f_n)^T$ ein n -dimensionaler Vektor ($n = \text{Anzahl der nicht-geerdeten Knoten}$), der die Stromstärken der Stromquellen beschreibt. Per Konvention zählen wir f_j positiv, wenn der Strom aus dem Knoten j wegfließt. Die Strombilanz lautet dann:

$$A^T \mathbf{y} - \mathbf{f} = 0 \quad \text{oder} \quad A^T \mathbf{y} = \mathbf{f}. \quad (5.15)$$

5.2.5. Zusammenfassung

Ein elektrisches Widerstandsnetzwerk sei durch folgende Bestimmungsstücke mathematisch charakterisiert:

- (1) Eine Inzidenzmatrix A , die die topologische Struktur beschreibt. Es ist eine $m \times n$ -Matrix, wobei m die Anzahl der Kanten ist, und n die Anzahl der nicht geerdeten Knoten.
- (2) Eine Matrix C , die die Leitwerte der Widerstände angibt. Es ist eine $m \times m$ -Diagonalmatrix mit lauter positiven Eigenwerten (in jeder Kante ist ein positiver Widerstand).

- (3) Einen Vektor $\mathbf{b} \in \mathbb{R}^m$, der die Spannungsquellen in den Kanten beschreibt.
- (4) Einen Vektor $\mathbf{f} \in \mathbb{R}^n$, der die an den Knoten angeschlossenen Stromquellen beschreibt.

Dann erfüllen die Ströme $\mathbf{y} = (y_1, \dots, y_m)^\top$ in den Kanten und die Potenziale $\mathbf{x} = (x_1, \dots, x_n)$ an den Knoten die Gleichgewichtsbedingungen

$$\mathbf{y} = \mathbf{C}(\mathbf{b} - \mathbf{A}\mathbf{x}), \quad \mathbf{A}^\top \mathbf{y} = \mathbf{f}. \quad (5.16)$$

Wir erhalten daraus die folgende Gleichung für \mathbf{x} ,

$$\mathbf{A}^\top \mathbf{C} \mathbf{A} \mathbf{x} = \mathbf{A}^\top \mathbf{C} \mathbf{b} - \mathbf{f} \quad (5.17)$$

mit gegebener Inhomogenität und positiv definiten Matrix $\mathbf{A}^\top \mathbf{C} \mathbf{A}$. Berechnet man daraus \mathbf{x} , erhält man \mathbf{y} aus der ersten Gleichung (5.16).

Schreiben wir $\mathbf{C}^{-1}\mathbf{y} = \mathbf{b} - \mathbf{A}\mathbf{x}$, können wir die Gleichungen (5.16) zu dem folgenden $n + m$ -dimensionalen Gleichungssystem zusammenfassen:

$$\begin{pmatrix} \mathbf{C}^{-1} & \mathbf{A} \\ \mathbf{A}^\top & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{f} \end{pmatrix}. \quad (5.18)$$

Die Matrix hier ist zwar nicht positiv definit, das System ist dennoch eindeutig lösbar.

5.3. Ein äquivalentes Minimierungsproblem

In Kapitel 4.3 haben wir gesehen, dass ein inhomogenes lineares Gleichungssystem mit einer positiv definiten Matrix \mathbf{K} auch als eine Extremwertaufgabe geschrieben werden kann. Auch die beiden Gleichungen (5.16) können als eine Optimierungsaufgabe formuliert werden. Dazu müssen wir aber die Optimierung unter Nebenbedingungen untersuchen.

Sei wieder \mathbf{C} eine symmetrische, positiv definite $m \times m$ -Matrix. Sei \mathbf{A} eine $m \times n$ -Matrix mit n linear unabhängigen Spalten (also gilt $n \leq m$). Schließlich seien $\mathbf{f} \in \mathbb{R}^n$ und $\mathbf{b} \in \mathbb{R}^m$ beliebige Vektoren.

Problem: Wir suchen ein Minimum der quadratische Funktion $Q : \mathbb{R}^m \rightarrow \mathbb{R}$,

$$Q(\mathbf{y}) = \frac{1}{2} \mathbf{y}^\top \mathbf{C}^{-1} \mathbf{y} - \mathbf{b}^\top \mathbf{y}, \quad (5.19)$$

unter all jenen $\mathbf{y} \in \mathbb{R}^m$, für die gilt

$$\mathbf{A}^\top \mathbf{y} = \mathbf{f} \quad \text{“Nebenbedingung“ (engl. “constraint”)}. \quad (5.20)$$

Beachte: Die Nebenbedingung ist ein System von n linearen Gleichungen in m Unbekannten y_j (wobei $m \geq n$). Ohne diese Nebenbedingungen wäre das Minimum von $Q(\mathbf{y})$ durch $\mathbf{y} = \mathbf{C} \mathbf{b}$ gegeben, siehe Satz 4.2. Dieses “absolute

Minimum" muß aber die Zusatzbedingung nicht erfüllen, im allgemeinen ist ja $\mathbf{A}^\top \mathbf{C} \mathbf{b} \neq \mathbf{f}$.

Im Hinblick auf die Anwendung bei elektrischen Widerstandsnetzwerken ist noch bemerkenswert, dass die Aufgabe ausschließlich die Stromstärken y_j involviert und nicht die Potentiale x_i (welche mathematische Hilfsgrößen und nicht direkt beobachtbare sind).

Ein Minimierungsproblem mit Nebenbedingungen löst man oft mit der "Methode der Lagrange'schen Multiplikatoren", die wir an diesem Beispiel illustrieren wollen. Dazu führt man für die n Nebenbedingungen $\mathbf{A}^\top \mathbf{y} - \mathbf{f} = 0$ zusätzliche Unbekannte ein, nämlich die n Lagrange'schen Multiplikatoren x_1, x_2, \dots, x_n (hier fassen wir diese zu einem Vektor $\mathbf{x} \in \mathbb{R}^n$ zusammen). Daraus und aus der ursprünglichen Funktion Q bildet man nun eine neue Funktion

$$L(\mathbf{x}, \mathbf{y}) = Q(\mathbf{y}) + \mathbf{x}^\top (\mathbf{A}^\top \mathbf{y} - \mathbf{f}), \quad (5.21)$$

und sucht deren *kritische Punkte*, das sind jene Punkte (\mathbf{x}, \mathbf{y}) , an denen alle partiellen Ableitungen erster Ordnung von L verschwinden:

Ein kritischer Punkt $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)})$ von L ist eine Lösung des Gleichungssystems

$$\frac{\partial}{\partial x_i} L(\mathbf{x}, \mathbf{y}) = 0, \quad i = 1, \dots, n, \quad (5.22)$$

$$\frac{\partial}{\partial y_j} L(\mathbf{x}, \mathbf{y}) = 0, \quad j = 1, \dots, m. \quad (5.23)$$

(Achtung: Ein kritischer Punkt ist nicht notwendigerweise eine Extremstelle!)

Wegen

$$\frac{\partial}{\partial x_i} L(\mathbf{x}, \mathbf{y}) = \frac{\partial}{\partial x_i} \sum_j x_j (\mathbf{A}^\top \mathbf{y} - \mathbf{f})_j = (\mathbf{A}^\top \mathbf{y} - \mathbf{f})_i \quad (5.24)$$

ist die Gleichung (5.22) äquivalent mit (5.20):

$$\frac{\partial}{\partial x_i} L(\mathbf{x}, \mathbf{y}) = 0 \quad (\text{alle } i) \iff \mathbf{A}^\top \mathbf{y} - \mathbf{f} = 0. \quad (5.25)$$

Beachte: Die Nebenbedingung ist genau die zweite der Gleichgewichtsbedingungen (5.16).

Die Ableitungen nach y_j ergeben:

$$\begin{aligned} \frac{\partial}{\partial y_j} L(\mathbf{x}, \mathbf{y}) &= \frac{\partial}{\partial y_j} \left(\frac{1}{2} \sum_{i,k} (\mathbf{C}^{-1})_{ik} y_i y_k - \sum_i b_i y_i + \sum_i (\mathbf{A}\mathbf{x})_i y_i \right) \\ &= \frac{1}{2} \sum_k \mathbf{C}_{jk}^{-1} y_k + \frac{1}{2} \sum_i \mathbf{C}_{ij}^{-1} y_i - b_j + (\mathbf{A}\mathbf{x})_j \\ &= (\mathbf{C}^{-1}\mathbf{y})_j - b_j + (\mathbf{A}\mathbf{x})_j. \end{aligned}$$

Daher ist

$$\frac{\partial}{\partial y_j} L(\mathbf{x}, \mathbf{y}) = 0 \quad (\text{alle } j) \quad \iff \quad \mathbf{C}^{-1}\mathbf{y} - \mathbf{b} + \mathbf{A}\mathbf{x} = 0. \quad (5.26)$$

Das ist genau die erste der Gleichgewichtsbedingungen (5.16). Die kritischen Punkte von L sind also genau die Lösungen des Gleichungssystems (5.16).

Nach der Theorie der Lagrange-Multiplikatoren, die wir im Abschnitt 5.4 behandeln werden, gilt folgender Zusammenhang:

Falls $\mathbf{y}^{(0)}$ ein Minimum von Q unter der Nebenbedingung $\mathbf{A}^\top \mathbf{y} - \mathbf{f} = 0$ ist, so gibt es ein $\mathbf{x}^{(0)}$, sodass $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)})$ ein kritischer Punkt von L ist.

Das besagt, dass man die Lösungen des Minimierungsproblems mit Nebenbedingung unter den kritischen Punkten der Funktion L suchen muss. Hat man so einen kritischen Punkt gefunden (in unserem Beispiel ist die Lösung des Gleichungssystems (5.16) so ein kritischer Punkt), dann muss man noch nachprüfen, ob es sich dabei tatsächlich um ein Minimum von Q handelt.

Sei also ein kritischer Punkt $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)}) \in \mathbb{R}^{m+n}$ von L als Lösung des Gleichungssystems

$$\mathbf{y} = \mathbf{C}(\mathbf{b} - \mathbf{A}\mathbf{x}), \quad \mathbf{A}^\top \mathbf{y} = \mathbf{f} \quad (5.27)$$

gegeben. Wir wollen zeigen, dass $Q(\mathbf{y}^{(0)})$ wirklich der Minimalwert von Q unter der gegebenen Nebenbedingung ist. Dazu schreiben wir $\mathbf{y} = \mathbf{y}^{(0)} + \mathbf{h}$ und berechnen

$$\begin{aligned} Q(\mathbf{y}) &= \frac{1}{2}(\mathbf{y}^{(0)} + \mathbf{h})^\top \mathbf{C}^{-1}(\mathbf{y}^{(0)} + \mathbf{h}) - \mathbf{b}^\top \mathbf{y}^{(0)} - \mathbf{b}^\top \mathbf{h} \\ &= Q(\mathbf{y}^{(0)}) + \mathbf{h}^\top \mathbf{C}^{-1} \mathbf{y}^{(0)} + \frac{1}{2} \mathbf{h}^\top \mathbf{C}^{-1} \mathbf{h} - \mathbf{b}^\top \mathbf{h}. \end{aligned}$$

Hier haben wir die Symmetrie von \mathbf{C}^{-1} benützt. Aus der ersten Gleichung von (5.27) ersetzen wir $\mathbf{C}^{-1} \mathbf{y}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}$ und erhalten

$$\begin{aligned} Q(\mathbf{y}) &= Q(\mathbf{y}^{(0)}) + \mathbf{h}^\top (\mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}) + \frac{1}{2} \mathbf{h}^\top \mathbf{C}^{-1} \mathbf{h} - \mathbf{b}^\top \mathbf{h} \\ &= Q(\mathbf{y}^{(0)}) - \mathbf{h}^\top \mathbf{A}\mathbf{x}^{(0)} + \frac{1}{2} \mathbf{h}^\top \mathbf{C}^{-1} \mathbf{h} \\ &= Q(\mathbf{y}^{(0)}) - (\mathbf{A}^\top \mathbf{h})^\top \mathbf{x}^{(0)} + \frac{1}{2} \mathbf{h}^\top \mathbf{C}^{-1} \mathbf{h}. \end{aligned}$$

Nun ist

$$\mathbf{A}^\top \mathbf{h} = \mathbf{A}^\top \mathbf{y} - \mathbf{A}^\top \mathbf{y}^{(0)} = \mathbf{f} - \mathbf{f} = 0,$$

wenn wir nur solche \mathbf{y} betrachten, die die Nebenbedingung erfüllen. Daher gilt

$$Q(\mathbf{y}) = Q(\mathbf{y}^{(0)}) + \frac{1}{2} \mathbf{h}^\top \mathbf{C}^{-1} \mathbf{h}$$

und da C^{-1} laut Annahme eine positiv definite Matrix ist, gilt :

$$Q(\mathbf{y}) > Q(\mathbf{y}^{(0)}) \quad \text{für } \mathbf{y} \neq \mathbf{y}^{(0)}.$$

$Q(\mathbf{y}^{(0)})$ ist also tatsächlich der minimale Wert von $Q(\mathbf{y})$ unter all jenen \mathbf{y} , die der Nebenbedingung $A^T \mathbf{y} = \mathbf{f}$ genügen. Wir haben also den folgenden Satz bewiesen:

SATZ 5.1. $\mathbf{y}^{(0)} \in \mathbb{R}^m$ ist ein Minimum von

$$Q(\mathbf{y}) = \frac{1}{2} \mathbf{y}^T C^{-1} \mathbf{y} - \mathbf{b}^T \mathbf{y} \quad (5.28)$$

unter der Nebenbedingung $A^T \mathbf{y} - \mathbf{f} = 0$, genau dann, wenn es ein $\mathbf{x}^{(0)} \in \mathbb{R}^n$ gibt, sodass $\mathbf{x}^{(0)}$ und $\mathbf{y}^{(0)}$ das Gleichungssystem

$$C^{-1} \mathbf{y} + A \mathbf{x} = \mathbf{b}, \quad A^T \mathbf{y} - \mathbf{f} = 0 \quad (5.29)$$

lösen.

5.4. Minimierung unter Nebenbedingungen

Hier betrachten wir das Problem der Optimierung unter Nebenbedingungen ein wenig allgemeiner. Wir nehmen keine spezielle Form von Q an und nehmen auch nicht an, daß die Nebenbedingung linear in \mathbf{y} sei. Wir sind aber andererseits spezieller, weil wir das Problem nur in zwei Dimensionen behandeln. Die Verallgemeinerung auf m Dimensionen und n Nebenbedingungen ist jedoch lediglich eine Übung in mathematischer Buchhaltung ohne wesentlich neue Ideen.

Seien Q und A zweimal stetig differenzierbare Funktionen von \mathbb{R}^2 nach \mathbb{R} . Wir nehmen an, daß

$$\frac{\partial}{\partial y_2} A(y_1, y_2) \neq 0 \quad (5.30)$$

für alle $\mathbf{y} = (y_1, y_2)$ in einem offenen Bereich des \mathbb{R}^2 gilt. Sei f eine Konstante. Unter diesen Bedingungen gilt folgende Aussage:

Falls $\mathbf{y}^{(0)} = (y_1^{(0)}, y_2^{(0)})$ ein Extremum von $Q(y_1, y_2)$ unter der Nebenbedingung

$$A(y_1, y_2) - f = 0 \quad (5.31)$$

ist, so gibt es eine Zahl $x^{(0)} \in \mathbb{R}$, sodass die Funktion

$$L(x, y_1, y_2) = Q(y_1, y_2) + x(A(y_1, y_2) - f) \quad (5.32)$$

an der Stelle $(x^{(0)}, y_1^{(0)}, y_2^{(0)})$ einen kritischen Punkt hat. Das heißt, die Zahlen $x^{(0)}, y_1^{(0)}, y_2^{(0)}$ sind eine Lösung des Gleichungssystems

$$\frac{\partial}{\partial x} L(x, y_1, y_2) = 0, \quad \frac{\partial}{\partial y_i} L(x, y_1, y_2) = 0, \quad i = 1, 2. \quad (5.33)$$

Um die Extrema von $Q(y_1, y_2)$ unter der Nebenbedingung zu finden, geht man also so vor: Man sucht die Lösungen der Gleichungen (5.33). Unter diesen Lösungen muss sich das gesuchte Extremum befinden. Ob bei einer gegebenen Lösung $x^{(0)}, y_1^{(0)}, y_2^{(0)}$, der Wert $Q(y_1^{(0)}, y_2^{(0)})$ auch tatsächlich extremal ist, muss gesondert geprüft werden. Die Methode ist also besonders dann nützlich, wenn die Existenz des Extremums schon bekannt ist.

BEWEIS. Um eine notwendige Bedingung für das Extremum zu finden, kann man zuerst aus der Nebenbedingung y_2 ausrechnen, und in $Q(y_1, y_2)$ einsetzen. Die resultierende Funktion hängt nur noch von y_1 ab. Die Bedingung erhält man durch Nullsetzen ihrer Ableitung nach y_1 .

Um die Gleichung $A(y_1, y_2) - f = 0$ nach y_2 aufzulösen, brauchen wir den Satz von der impliziten Funktion. Danach ist die Bedingung für die Auflösbarkeit nach y_2 durch (5.30) gegeben. Es gibt dann lokal so eine Lösung

$$y_2 = g(y_1) \quad (5.34)$$

und g ist stetig differenzierbar. Im allgemeinen kann g allerdings nicht explizit angegeben werden, z.B. falls $A(y_1, y_2) = y_2 \exp(y_1 y_2)$. Setzen wir g in Q ein, wird Q zu einer differenzierbaren Funktion von y_1 alleine:

$$Q(y_1, g(y_1)) = R(y_1) \quad (5.35)$$

Das Extremum von R finden wir nun durch Nullsetzen der Ableitung

$$\frac{dR(y_1)}{dy_1} = \frac{\partial Q}{\partial y_1} + \frac{\partial Q}{\partial y_2} \frac{dg}{dy_1} = 0. \quad (5.36)$$

Unter den Lösungen dieser Gleichung finden wir auch die Extremwerte von Q unter der Nebenbedingung. Das Problem ist aber, dass wir g und somit R im allgemeinen nicht kennen.

Andererseits ist, nach der Definition von g ,

$$A(y_1, g(y_1)) = f \quad (5.37)$$

(eine Konstante), also verschwindet deren Ableitung nach y_1 :

$$\frac{\partial A}{\partial y_1} + \frac{\partial A}{\partial y_2} \frac{dg}{dy_1} = 0. \quad (5.38)$$

Multipliziert man (5.36) mit $\partial A / \partial y_2$, erhalten wir

$$\frac{\partial Q}{\partial y_1} \frac{\partial A}{\partial y_2} + \frac{\partial Q}{\partial y_2} \frac{\partial A}{\partial y_2} \frac{dg}{dy_1} = 0,$$

und unter Verwendung von (5.38) ergibt das

$$\frac{\partial Q}{\partial y_1} \frac{\partial A}{\partial y_2} - \frac{\partial Q}{\partial y_2} \frac{\partial A}{\partial y_1} = 0. \quad (5.39)$$

Zusammen mit der Nebenbedingung (5.37) beschreibt also (5.39) die möglichen Extrema: Dort, wo ein Extremum von Q unter der gegebenen Nebenbedingung ist, muss auch die Gleichung (5.39) erfüllt sein. Hat man umgekehrt einen Punkt $(y_1^{(0)}, y_2^{(0)})$, wo die Gleichungen (5.37) und (5.39) erfüllt sind, muss man aber noch überprüfen, ob tatsächlich ein Maximum oder Minimum vorliegt (indem man dort zB das Vorzeichen der zweiten Ableitung von R bestimmt).

Nun untersuchen wir den Zusammenhang mit den kritischen Punkten der Funktion

$$L(x, y_1, y_2) = Q(y_1, y_2) + x(A(y_1, y_2) - f) \quad (5.40)$$

Wir zeigen, dass die Punkte, an denen die Gleichungen (5.37) und (5.39) erfüllt sind, genau die kritischen Punkte dieser Funktion sind. Setzen wir also die Ableitungen dieser Funktion gleich Null:

$$\frac{\partial L}{\partial x} = 0 \quad \Leftrightarrow \quad A(y_1, y_2) - f = 0 \quad (5.41)$$

$$\frac{\partial L}{\partial y_1} = 0 \quad \Leftrightarrow \quad \frac{\partial Q}{\partial y_1} + x \frac{\partial A}{\partial y_1} = 0 \quad (5.42)$$

$$\frac{\partial L}{\partial y_2} = 0 \quad \Leftrightarrow \quad \frac{\partial Q}{\partial y_2} + x \frac{\partial A}{\partial y_2} = 0 \quad (5.43)$$

Die erste dieser Gleichungen ist bereits die Nebenbedingung (5.37). Berechnen wir x aus der letzten Gleichung (beachte die Bedingung (5.30)) und setzen wir es in die vorletzte Gleichung ein, erhalten wir wieder (5.39). Damit haben wir aber wieder genau die oben erhaltenen notwendigen Bedingungen für die möglichen Extrema. \square

Die Methode mit den Lagrange-Multiplikatoren ist nützlich, da sie die (im allgemeinen unmöglich explizit zu machende) Auflösung der Gleichung (5.31) nach y_2 vermeidet.

5.5. Das duale Problem

Primäres Problem: Minimiere das Funktional $Q : \mathbb{R}^m \rightarrow \mathbb{R}$

$$Q(\mathbf{y}) = \frac{1}{2} \mathbf{y}^\top \mathbf{C}^{-1} \mathbf{y} - \mathbf{b}^\top \mathbf{y} \quad (5.44)$$

unter der Nebenbedingung

$$\mathbf{A}^\top \mathbf{y} - \mathbf{f} = 0. \quad (5.45)$$

Dabei gelten die Standardvoraussetzungen:

- C sei eine $n \times n$ -Matrix, symmetrisch und positiv definit,
- A sei eine $m \times n$ -Matrix mit n linear unabhängigen Spalten,
- $\mathbf{f} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$.

Mit $\mathbf{x} = (x_1, \dots, x_n)^\top$ haben wir folgende Funktion definiert:

$$\begin{aligned} L(\mathbf{x}, \mathbf{y}) &= Q(\mathbf{y}) + \mathbf{x}^\top (\mathbf{A}^\top \mathbf{y} - \mathbf{f}) \\ &= \frac{1}{2} \mathbf{y}^\top \mathbf{C}^{-1} \mathbf{y} - (\mathbf{b} - \mathbf{A} \mathbf{x})^\top \mathbf{y} - \mathbf{x}^\top \mathbf{f}. \end{aligned} \quad (5.46)$$

Wir minimieren nun $L(\mathbf{x}, \mathbf{y})$ bzgl. \mathbf{y} bei festgehaltenem \mathbf{x} . Das ist ein ganz gewöhnliches Minimierungsproblem (ohne Nebenbedingungen), das wir schon gelöst haben (siehe Satz 4.2). Demnach ist das Minimum an der Stelle

$$\mathbf{y} = \mathbf{C} (\mathbf{b} - \mathbf{A} \mathbf{x}) \quad (5.47)$$

und der Minimalwert ergibt sich zu

$$\min_{\mathbf{y}} L(\mathbf{x}, \mathbf{y}) = -\frac{1}{2} (\mathbf{b} - \mathbf{A} \mathbf{x})^\top \mathbf{C} (\mathbf{b} - \mathbf{A} \mathbf{x}) - \mathbf{x}^\top \mathbf{f} =: R(\mathbf{x}) \quad (5.48)$$

Dabei haben wir ein neues Funktional R eingeführt. Wir haben also von $L(\mathbf{x}, \mathbf{y})$ als Funktion von \mathbf{y} den Minimalwert $R(\mathbf{x})$ in Abhängigkeit von den Lagrange-Multiplikatoren \mathbf{x} bestimmt.

Für alle \mathbf{x} ist $R(\mathbf{x})$ nach Definition der Minimalwert von $L(\mathbf{x}, \mathbf{y})$ bzgl. \mathbf{y} . Also gilt natürlich $L(\mathbf{x}, \mathbf{y}) \geq R(\mathbf{x})$ für alle \mathbf{y} und alle \mathbf{x} . Falls aber $\mathbf{A}^\top \mathbf{y} = \mathbf{f}$, gilt $L(\mathbf{x}, \mathbf{y}) = Q(\mathbf{y})$, also gilt für diese \mathbf{y} die Ungleichung $Q(\mathbf{y}) \geq R(\mathbf{x})$. Wir fassen zusammen:

$$Q(\mathbf{y}) \geq R(\mathbf{x}) \quad \text{für alle } \mathbf{x} \text{ und alle } \mathbf{y} \text{ mit } \mathbf{A}^\top \mathbf{y} = \mathbf{f}. \quad (5.49)$$

Gleichheit kann nur herrschen beim Minimum von $Q(\mathbf{y})$ (unter der gegebenen Nebenbedingung) und beim Maximum von $R(\mathbf{x})$.

Das führt uns direkt zum sogenannten dualen Problem:

Duales Problem: Maximiere das Funktional $\mathbf{x} \rightarrow R(\mathbf{x})$, wobei

$$R(\mathbf{x}) = -\frac{1}{2} (\mathbf{b} - \mathbf{A} \mathbf{x})^\top \mathbf{C} (\mathbf{b} - \mathbf{A} \mathbf{x}) - \mathbf{x}^\top \mathbf{f} \quad (5.50)$$

(ohne irgendwelche Nebenbedingungen).

Wir haben jetzt folgendes Verfahren: Bestimme für $L(\mathbf{x}, \mathbf{y})$ zuerst den Minimalwert $R(\mathbf{x})$ in Abhängigkeit von \mathbf{x} und dann das Maximum dieser Minimalwerte. Wenn man bei diesem Verfahren ein eindeutiges Resultat erhält, hat man zwangsläufig einen Sattelpunkt von L gefunden (Minimalwert in \mathbf{y} -Richtung, Maximalwert in \mathbf{x} -Richtung).

Wir suchen also einen Punkt $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)})$, wo

$$Q(\mathbf{y}^{(0)}) = R(\mathbf{x}^{(0)}) \quad \text{und} \quad \mathbf{A}^\top \mathbf{y}^{(0)} = \mathbf{f}. \quad (5.51)$$

Das heißt, wir suchen eine Nullstelle der Funktion

$$\begin{aligned} Q(\mathbf{y}) - R(\mathbf{x}) &= \frac{1}{2} \mathbf{y}^\top \mathbf{C}^{-1} \mathbf{y} - \mathbf{b}^\top \mathbf{y} \\ &\quad + \frac{1}{2} (\mathbf{b} - \mathbf{A} \mathbf{x})^\top \mathbf{C} (\mathbf{b} - \mathbf{A} \mathbf{x}) + \mathbf{x}^\top \mathbf{A}^\top \mathbf{y} \\ &= \frac{1}{2} (\mathbf{C}^{-1} \mathbf{y} + \mathbf{A} \mathbf{x} - \mathbf{b})^\top \mathbf{C} (\mathbf{C}^{-1} \mathbf{y} + \mathbf{A} \mathbf{x} - \mathbf{b}). \end{aligned}$$

Wegen der positiven Definitheit von \mathbf{C} ist dieser Ausdruck genau dann gleich Null, wenn

$$\mathbf{C}^{-1} \mathbf{y} + \mathbf{A} \mathbf{x} - \mathbf{b} = 0. \quad (5.52)$$

Das ist wieder die bekannte Gleichgewichtsbedingung (5.29).

Primäres und duales Problem:

Der Minimalwert $Q(\mathbf{y}^{(0)})$ von $Q(\mathbf{y})$ mit Nebenbedingung $\mathbf{A}^\top \mathbf{y} = \mathbf{f}$ ist gleich dem Maximalwert $R(\mathbf{x}^{(0)})$ von $R(\mathbf{x})$ (ohne Nebenbedingung):

$$Q(\mathbf{y}^{(0)}) = R(\mathbf{x}^{(0)}). \quad (5.53)$$

Der Punkt $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)})$ ist ein Sattelpunkt von $L(\mathbf{x}, \mathbf{y})$. Dort gilt

$$\mathbf{C}^{-1} \mathbf{y}^{(0)} + \mathbf{A} \mathbf{x}^{(0)} - \mathbf{b} = 0, \quad \mathbf{A}^\top \mathbf{y}^{(0)} = \mathbf{f}. \quad (5.54)$$

Differentialgleichungen und Eigenwertproblem

6.1. Einleitung

Gesetze, die den zeitlichen Verlauf bestimmter Vorgänge bestimmen, lassen sich in der Regel durch Differentialgleichungen beschreiben. Das sind Beziehungen, die eine (unbekannte) Größe $\mathbf{x}(t)$ mit ihren Ableitungen nach t verknüpfen.

Die gesuchte Größe im Falle einer Differentialgleichung ist eine Funktion, möglicherweise eine vektorwertige Funktion $t \rightarrow \mathbf{x}(t)$. Wenn die gesuchte Funktion nur von einer Variablen abhängt (z.B. von der Zeit t), spricht man von einer gewöhnlichen Differentialgleichung.

Im allgemeinsten Fall ist eine gewöhnliche Differentialgleichung von der Form

$$F(\mathbf{x}^{(m)}(t), \mathbf{x}^{(m-1)}(t), \dots, \mathbf{x}'(t), \mathbf{x}(t), t) = 0 \quad (6.1)$$

mit einer gegebenen Funktion F von $n(m+1)$ Variablen. Hier bedeutet natürlich

$$\mathbf{x}^{(k)}(t) = \frac{d^k}{dt^k} \mathbf{x}(t) \quad (6.2)$$

die k -te Ableitung der Funktion $\mathbf{x} : t \rightarrow \mathbf{x}(t)$ an der Stelle t . Wenn t die Zeit bedeutet wird man in der Literatur auch oft die Schreibweise $\dot{\mathbf{x}}(t)$ für die erste Ableitung $\mathbf{x}'(t)$ und $\ddot{\mathbf{x}}(t)$ für die zweite Ableitung $\mathbf{x}''(t)$ an der Stelle t finden.

Die Werte der Funktion \mathbf{x} können in einem n -dimensionalen reellen oder komplexen Vektorraum liegen. Abhängig von der gegebenen, die Differentialgleichung definierenden Funktion F können die Werte der Variablen t in ganz \mathbb{R} oder nur in einem Intervall liegen.

Es kann sein, daß die Funktion F es zuläßt, daß man (6.1) nach der höchsten Ableitung auflöst, dann bekommt man eine Gleichung der Form

$$\mathbf{x}^{(n)}(t) = f(\mathbf{x}^{(n-1)}(t), \dots, \mathbf{x}'(t), \mathbf{x}(t), t). \quad (6.3)$$

Das nennt man dann eine explizite gewöhnliche Differentialgleichung n -ter Ordnung für die gesuchte Funktion \mathbf{x} . Bevor wir uns aber in Klassifizierungsfragen verlieren, werden wir uns einem möglichst einfachen Spezialfall widmen, um ein Gefühl für Differentialgleichungen zu entwickeln.

6.2. Die einfachste Gleichung

Als erstes, Ihnen vielleicht schon bekanntes Beispiel betrachten wir die Differentialgleichung

$$\frac{d}{dt} x(t) = a x(t). \quad (6.4)$$

Hier ist a eine reelle Konstante. Diese Gleichung beschreibt eine Größe, deren zeitliche Änderungsgeschwindigkeit proportional zur Größe selbst ist. Diese Gleichung charakterisiert die Exponentialfunktion. Jede Funktion

$$x(t) = C e^{at} \quad \text{mit beliebigem } C \in \mathbb{R} \quad (6.5)$$

ist eine Lösung von (6.4). Wir sehen, dass die Lösung nicht eindeutig ist. Wenn wir einen Anfangswert vorgeben, $x(t_0) = x_0$, dann kann man C in (6.5) sogar ausrechnen, dh. die Lösung wird eindeutig. Es ergibt sich $C = x_0 \exp(-a t_0)$ und daher (bitte nachrechnen!)

$$x(t) = x_0 e^{a(t-t_0)}. \quad (6.6)$$

Für $a > 0$ beschreibt die Gleichung (6.4) daher ein exponentielles Anwachsen der Größe x . Falls $a < 0$ liegt ein exponentieller Abfall (Dämpfung, Zerfall) vor.

Für $a > 0$ heißt (6.4) die Gleichung des exponentiellen Wachstums. Sie modelliert ganz gut das Wachstum einer Population. Hier ist $u(t)$ proportional zur Anzahl der Individuen zur Zeit t . Natürlich ist das "in Wirklichkeit" eine Stufenfunktion, die nur Werte in den natürlichen Zahlen annimmt. Wenn wir aber von sehr großen Anzahlen sprechen, dann sind die Stufen im Vergleich zum Funktionswert winzig, und es ist eine ganz gute Näherung, wenn wir $x(t)$ als von t stetig abhängige, ja sogar differenzierbare Größe auffassen. In der Praxis ist ja auch die Anzahl der Individuen zu jeder Zeit nur mit einer gewissen Ungenauigkeit bekannt, die in der Regel viel größer als eins ist. Dann gehen die ganzzahligen Änderungen der Individuenzahl in der Unschärfe der Datenerfassung einfach unter.

Nehmen wir an, jedes Individuum der Population hat eine gewisse Wahrscheinlichkeit, sich pro Zeiteinheit zu reproduzieren. Dann wird die Änderung der Anzahl pro Zeiteinheit proportional zur Anzahl der Individuen sein. (Bei doppelt so vielen Individuen ist die Zahl der Geburten auch doppelt so hoch). Das wird genau durch (6.4) mit positivem a ausgedrückt.

Wichtig bei der Anwendung dieses Wachstumsmodells sind folgende Voraussetzungen: Die Änderung der Population geschieht nur durch innere Mechanismen. Es findet keine Immigration oder Emigration statt. Der Vermehrungsmechanismus muß unabhängig von Umwelteinflüssen sein, die Individuen müssen nicht um begrenzte Ressourcen wetteifern.

Konkrete Beispiele:

- $x(t)$ = Anzahl der Bakterien in einem Reagenzglas.
- $x(t)$ = Anzahl der Menschen auf der Erde ($a = 0.02/\text{Jahr}$). Die Gleichung wird aber bald ihre Gültigkeit verlieren. Das können Sie selbst sehen, wenn Sie zum Beispiel nach dieser Gleichung die Bevölkerungsdichte der Erde im Jahr 3000 berechnen.
- $x(t)$ = Anzahl der radioaktiven Teilchen in einer Substanz. (Hier ist a negativ – Exponentieller Zerfall).

6.3. Die logistische Gleichung

Wieder betrachten wir ein Wachstumsmodell (Populationsmodell). Sei $x(t)$ die Anzahl der Individuen einer Population. Die Größe

$$\frac{1}{x(t)} \frac{dx(t)}{dt} \quad (6.7)$$

beschreibt das *Pro-Kopf-Wachstum* der Bevölkerung. Nehmen wir an, dass dieses Pro-Kopf Wachstum durch Überbevölkerung oder Wettkampf um Ressourcen gebremst wird, und zwar um einen Betrag, der proportional zur Individuenanzahl ist:

$$\frac{1}{x(t)} \frac{dx(t)}{dt} = a - b x(t). \quad (6.8)$$

Wir erhalten dann die *logistische Differentialgleichung*

$$\frac{dx(t)}{dt} = a x(t) - b x(t)^2. \quad (6.9)$$

Es gilt $ax - bx^2 = x(a - bx) = 0$ für $x = 0$ und $x = a/b$. Wenn die Lösung $x(t)$ zu einem Anfangszeitpunkt zufällig bei einem dieser Werte ist, dann ist zu diesem Zeitpunkt nach der Differentialgleichung auch die Ableitung der Lösung Null, dh., es existieren die folgenden konstanten Lösungen:

$$\begin{aligned} x(t) &= 0 && \text{für alle } t \text{ — keine Bevölkerung,} \\ x(t) &= a/b && \text{für alle } t \text{ — stabile Verhältnisse.} \end{aligned} \quad (6.10)$$

Für $0 \neq x \neq a/b$ haben wir

$$\int \frac{dx}{ax - bx^2} = \int dt, \quad (6.11)$$

also (mit Partialbruchzerlegung)

$$\frac{1}{a} \ln \left| \frac{x}{a - bx} \right| = t + C', \quad (6.12)$$

mit einer beliebigen Integrationskonstanten C' . Durch Auflösung nach x erhalten wir mit $C = \exp(-aC')$ die Lösung

$$x(t) = \begin{cases} \frac{a e^{at}}{b e^{at} - C} & \text{für } x < 0 \text{ und } x > a/b, \\ \frac{a e^{at}}{b e^{at} + C} & \text{für } 0 < x < a/b. \end{cases} \quad (6.13)$$

Übung: Skizziere und interpretiere die Lösung für ausgewählte Anfangsbedingungen.

6.4. Zinsrechnung

Angenommen, man gibt einer Bank eine einmalige Einlage der Höhe x_0 zur Zeit t_0 . Die Verzinsung erfolge zu folgenden Konditionen: Man darf das Geld jederzeit beheben und erhält zur Einlage zusätzlich eine Prämie, die proportional zum Betrag x_0 und proportional zur Einlagedauer $t - t_0$ ist. Also

$$\text{Prämie} = a x_0 (t - t_0). \quad (6.14)$$

Bei 5% Zinsen pro Jahr rechnet man also $t - t_0 = 1$ Jahr und $a = 0.05$. Wenn man dann die Zinsen zu den selben Konditionen veranlagt, hat man ein neues Anfangskapital von

$$x_1 = x_0 + a x_0 (t - t_0). \quad (6.15)$$

auf dem Sparkonto, und von da an werden die Zinsen mitverzinst.

Man könnte nun auf die Idee kommen, schon nach der halben Zeit das Geld abzuheben und samt den bis dahin angefallenen Zinsen wieder zu veranlagen. Dabei macht man sogar ein Geschäft: Das neue Anfangskapital nach $(t - t_0)/2$ ist

$$x'_0 = x_0 + a x_0 \frac{t - t_0}{2} = x_0 \left(1 + \frac{a(t - t_0)}{2} \right). \quad (6.16)$$

Nach der Gesamtzeit $t - t_0$ hat man dann insgesamt

$$x'_1 = x'_0 \left(1 + \frac{a(t - t_0)}{2} \right) = x_0 \left(1 + \frac{a(t - t_0)}{2} \right)^2 \quad (6.17)$$

und das ist mehr als $x_1 = x_0 (1 + a(t - t_0))$. Am meisten Gewinn macht man, wenn man in ganz kurzen Zeitintervallen das Geld behebt und wieder

einlegt. Das führt auf die Idee der kontinuierlichen Verzinsung, bei der das Vermögen zur Zeit t wie folgt gegeben ist:

$$x(t) = x_0 \lim_{n \rightarrow \infty} \left(1 + \frac{a(t-t_0)}{n} \right)^n = x_0 e^{a(t-t_0)}. \quad (6.18)$$

Das Vermögen entwickelt sich dabei exponentiell! Bei kontinuierlicher Verzinsung gehorcht das Vermögen also der Differentialgleichung

$$\frac{dx(t)}{dt} = a x(t). \quad (6.19)$$

In der Finanzwelt ändern sich Zinsen aber oft. Wir können dann ein Modell mit variabler Verzinsung $a(t)$ verwenden:

$$\frac{dx(t)}{dt} = a(t) x(t). \quad (6.20)$$

Ausserdem will man vielleicht pro Zeiteinheit einen gewissen Geldbetrag $b(t)$ einzahlen (z.B. per Dauerauftrag), bzw. abheben (dann ist $b(t)$ negativ). Dann folgt die Vermögensentwicklung der Gleichung

$$\frac{dx(t)}{dt} = a(t) x(t) + b(t). \quad (6.21)$$

Dabei beschreibt $b(t)$ die von aussen beigesteuerte zeitliche Änderungsrate der GröÙe x . Man bezeichnet (6.20) als homogene Gleichung, und (6.21) als inhomogene Gleichung mit der Inhomogenität $b(t)$.

Natürlich sind die Anwendungen nicht auf die Finanzwelt beschränkt. In der Populationsdynamik etwa würde die GröÙe $b(t)$ die Anzahl der pro Zeiteinheit zuwandernden oder abwandernden Individuen beschreiben.

Die Lösung der homogenen Gleichung ist, wie man durch Einsetzen in (6.20) nachprüft,

$$x_{\text{hom}}(t) = x_0 \exp\left(\int_{t_0}^t a(\tau) d\tau\right) = x_0 w(t_0, t). \quad (6.22)$$

Hier ist $w(t_0, t)$ der Betrag zur Zeit t , den eine zur Zeit t_0 vorhandene Einheitsmenge ergibt.

Nun berücksichtigen wir die Ein/Auszahlungen. Im Zeitintervall zwischen τ und $\tau + \Delta\tau$ kommt $b(\tau) \Delta\tau$ hinzu (dabei nehmen wir an, dass $\Delta\tau$ so kurz ist, dass sich $b(\tau)$ während dieser Zeit nicht merklich ändert). Dann haben wir zur Zeit t den Beitrag

$$b(\tau) \Delta\tau w(\tau, t) \quad (6.23)$$

Insgesamt müssen wir die Beiträge von allen Zeitintervallen zwischen dem Anfangszeitpunkt t_0 und dem Zeitpunkt t aufsummieren:

$$x(t) = x_0 w(t_0, t) + \sum b(\tau) w(\tau, t) \Delta\tau. \quad (6.24)$$

Lassen wir die Länge der Zeitintervalle $\Delta\tau$ gegen Null gehen, wird die Summe offenbar zum Riemann-Integral,

$$x(t) = x_0 w(t_0, t) + \int_{t_0}^t b(\tau) w(\tau, t) d\tau. \quad (6.25)$$

Diese Formel nennt man auch die "Variation-der-Konstanten Formel". Man kann wiederum durch Einsetzen in die Differentialgleichung (6.21) verifizieren, dass dieses $x(t)$ die Lösung zum Anfangswert x_0 darstellt.

Die Funktion $w(s, t) = \exp(\int_s^t a(\tau) d\tau)$ hat offenbar folgende Eigenschaften:

$$w(s, \tau) w(\tau, t) = w(s, t), \quad w(s, s) = 1, \quad \text{für alle reellen } s \text{ und } t. \quad (6.26)$$

Daraus folgt auch $w(s, t)w(t, s) = 1$. Vergleichen wir das mit der Lösung der homogenen Gleichung (6.20), die zum Zeitpunkt t gleich 1 ist (also eine Geldeinheit)

$$1 = x_{\text{hom}}(t) = w(t_0, t) x_0 = w(t_0, t) w(t, t_0), \quad (6.27)$$

so sehen wir, wie die Größe $w(t, t_0)$ für t größer als t_0 zu interpretieren ist: Das ist der Betrag, den man zur Zeit t_0 einlegen muss, um zur Zeit t genau eine Geldeinheit zu haben.

Schreiben wir also im Integral (6.25) $w(\tau, t) = w(\tau, t_0) w(t_0, t)$, so erhalten wir die Lösung in der Form

$$x(t) = w(t_0, t) \left(x_0 + \int_{t_0}^t b(\tau) w(\tau, t_0) d\tau \right) = w(t_0, t) y(t_0, t). \quad (6.28)$$

$y(t_0, t)$ beschreibt den Anfangswert für die homogenen Gleichung, der auf den gleichen Endbetrag $x(t)$ führt, wie die inhomogene Gleichung mit Anfangswert x_0 . Zur Zeit t_0 hätten wir den Geldbetrag $y(t_0, t)$ einlegen müssen, um zur Zeit t ohne weitere Einzahlungen den Betrag $x(t)$ zu haben.

6.5. Ein Gleichungssystem

Hier betrachten wir zwei Größen $x_1(t)$ und $x_2(t)$, deren zeitliche Änderungen von der jeweils anderen Größe abhängig sind. Als ein einfaches mathematisches Beispiel betrachten wir das *Differentialgleichungssystem*

$$\frac{dx_1(t)}{dt} = 2x_1(t) + x_2(t), \quad (6.29)$$

$$\frac{dx_2(t)}{dt} = x_1(t) + 2x_2(t). \quad (6.30)$$

Wir fassen die beiden Größen zu einem Vektor $\mathbf{x}(t)$ zusammen und schreiben

$$\frac{d\mathbf{x}(t)}{dt} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \mathbf{x}(t), \quad \text{wobei } \mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}. \quad (6.31)$$

Die gesuchte Lösung ist also eine Kurve im \mathbb{R}^2 :

$$\mathbf{x} : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R}^2 \\ t & \mapsto & \mathbf{x}(t) \end{cases} \quad (6.32)$$

Das System (6.31) ist ein Beispiel von *Kooperation*: Das Anwachsen von x_2 fördert die Wachstumsgeschwindigkeit von x_1 und umgekehrt.

Ein Beispiel von *Konkurrenz* wäre hingegen durch die Matrix

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad (6.33)$$

gegeben. Hier hemmt das Wachstum von x_2 die Größe x_1 .

Um die Lösung eindeutig zu machen, braucht man noch eine *Anfangsbedingung*. Als Beispiel betrachten wir (6.31) mit

$$\mathbf{x}(t=0) = \mathbf{x}_0 = \begin{pmatrix} 5 \\ 0 \end{pmatrix}. \quad (6.34)$$

Lösungsmethode: Falls möglich, finde die Eigenwerte der Matrix A des Gleichungssystems. Hier ist

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad (6.35)$$

eine symmetrische Matrix und kann daher diagonalisiert werden. Die Eigenwerte und Eigenvektoren sind

$$\lambda_1 = 1, \quad \mathbf{x}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (6.36)$$

$$\lambda_2 = 3, \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (6.37)$$

Die aus den Eigenvektoren gebildete Matrix S bringt A in Diagonalform:

$$S = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \quad (6.38)$$

$$\Lambda = S^{-1} A S = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}. \quad (6.39)$$

Die beiden Eigenvektoren bilden eine Basis im \mathbb{R}^2 . Insbesondere ist der Anfangsvektor eine Linearkombination der Eigenvektoren von A:

$$\mathbf{x}_0 = a \mathbf{x}_1 + b \mathbf{x}_2 = \frac{5}{2} \mathbf{x}_1 + \frac{5}{2} \mathbf{x}_2. \quad (6.40)$$

Man verifiziert nun leicht, daß

$$\mathbf{x}(t) = a \mathbf{x}_1 e^{\lambda_1 t} + b \mathbf{x}_2 e^{\lambda_2 t} \quad (6.41)$$

eine Lösung des Differentialgleichungssystems ist, denn es ist $\mathbf{x}(t=0) = \mathbf{x}_0$ und

$$\frac{d}{dt} \mathbf{x}(t) = a \mathbf{x}_1 \frac{d}{dt} e^{\lambda_1 t} + b \mathbf{x}_2 \frac{d}{dt} e^{\lambda_2 t} \quad (6.42)$$

$$= a \mathbf{x}_1 \lambda_1 e^{\lambda_1 t} + b \mathbf{x}_2 \lambda_2 e^{\lambda_2 t} \quad (6.43)$$

$$= a A \mathbf{x}_1 e^{\lambda_1 t} + b A \mathbf{x}_2 e^{\lambda_2 t} \quad (6.44)$$

$$= A (a \mathbf{x}_1 e^{\lambda_1 t} + b \mathbf{x}_2 e^{\lambda_2 t}) \quad (6.45)$$

$$= A \mathbf{x}(t). \quad (6.46)$$

$$(6.47)$$

Mit den konkreten Zahlen des Beispiels ist

$$\mathbf{x}(t) = \frac{5}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^t + \frac{5}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{3t} = \frac{5}{2} \begin{pmatrix} e^t + e^{3t} \\ -e^t + e^{3t} \end{pmatrix}. \quad (6.48)$$

Die Kooperation bewirkt also ein Wachstum vom Typ e^{3t} , wo alleine (ohne Kooperation also ohne die Nebendiagonalterme in A) nur ein Wachstum wie e^{2t} erreichbar wäre. Das Wachstum von x_2 ist etwas langsamer im Vergleich zu x_1 , was an der schlechteren Ausgangssituation $x_2(0) = 0$ liegt.

Allgemein: Sei A eine $n \times n$ -Matrix mit n linear unabhängigen Eigenvektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$ mit dazugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$. Dann gilt:

Die Lösung des Anfangswertproblems

$$\frac{d}{dt} \mathbf{x}(t) = A \mathbf{x}(t) \quad \text{mit } \mathbf{x}(0) = \mathbf{x}_0 \quad (6.49)$$

ist

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + c_2 e^{\lambda_2 t} \mathbf{x}_2 + \dots + c_n e^{\lambda_n t} \mathbf{x}_n. \quad (6.50)$$

Die Konstanten c_j erhält man aus dem Gleichungssystem

$$\mathbf{x}_0 = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_n \mathbf{x}_n. \quad (6.51)$$

Bemerkungen

1. Die Eigenwerte λ_j müssen nicht alle verschieden sein. Außerdem funktioniert das ganze genauso, wenn die Matrixelemente, Eigenwerte, etc. komplex sind.

2. Es gilt

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = S^{-1} A S \quad \text{mit} \quad S = \begin{pmatrix} | & \cdots & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & \cdots & | \end{pmatrix}. \quad (6.52)$$

Die Gleichung (6.51) kann man auch als

$$\mathbf{x}_0 = \mathbf{S} \mathbf{c}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \quad (6.53)$$

schreiben, also ist

$$\mathbf{c} = \mathbf{S}^{-1} \mathbf{x}_0. \quad (6.54)$$

\mathbf{S}^{-1} ist also die Transformation, die in die Eigenvektor-Basisdarstellung überführt.

3. Die Variablentransformation $\mathbf{x} \rightarrow \mathbf{y} = \mathbf{S}^{-1} \mathbf{x}$ entkoppelt das Gleichungssystem:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A} \mathbf{x} \quad \Leftrightarrow \quad \mathbf{S}^{-1} \frac{d\mathbf{x}}{dt} = \mathbf{S}^{-1} \mathbf{A} \mathbf{x} \quad \Leftrightarrow \quad \frac{d\mathbf{y}}{dt} = \mathbf{S}^{-1} \mathbf{A} \mathbf{S} \mathbf{y} = \mathbf{\Lambda} \mathbf{y}. \quad (6.55)$$

Die letzte Gleichung ist äquivalent mit dem System

$$\frac{dy_1}{dt} = \lambda_1 y_1, \quad \dots \quad \frac{dy_n}{dt} = \lambda_n y_n. \quad (6.56)$$

In dieser entkoppelten Form löst man das System leicht:

$$y_1(t) = c_1 e^{\lambda_1 t}, \quad \dots \quad y_n(t) = c_n e^{\lambda_n t} \quad (6.57)$$

oder

$$\mathbf{y}(t) = e^{\mathbf{\Lambda} t} \mathbf{c}, \quad \text{mit} \quad e^{\mathbf{\Lambda} t} = \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix}. \quad (6.58)$$

Dabei haben wir auf naheliegende Weise die Exponentialfunktion einer Diagonalmatrix definiert.

4. Für die Lösung \mathbf{x} erhält man

$$\mathbf{x}(t) = \mathbf{S} \mathbf{y}(t) = \mathbf{S} e^{\mathbf{\Lambda} t} \mathbf{c} = \mathbf{S} e^{\mathbf{\Lambda} t} \mathbf{S}^{-1} \mathbf{x}_0 = e^{\mathbf{A} t} \mathbf{x}_0. \quad (6.59)$$

Dabei haben wir die Exponentialfunktion einer beliebigen diagonalisierbaren Matrix eingeführt:

$$e^{\mathbf{A} t} = \mathbf{S} e^{\mathbf{\Lambda} t} \mathbf{S}^{-1}, \quad \text{wobei} \quad \mathbf{\Lambda} = \mathbf{S}^{-1} \mathbf{A} \mathbf{S} \quad \text{diagonal.} \quad (6.60)$$

Allgemein definiert man die Exponentialfunktion einer Matrix durch

$$e^{\mathbf{A} t} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{A}^k \quad (6.61)$$

und hat

$$\begin{aligned} \frac{d}{dt} \mathbf{x}(t) &= \mathbf{A} \mathbf{x}(t), & \Leftrightarrow & \quad \mathbf{x}(t) = e^{\mathbf{A} t} \mathbf{x}_0. \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned} \quad (6.62)$$

Für Details verweisen wir auf Vorlesungen über gewöhnliche Differentialgleichungen.

6.6. Systeme zweiter Ordnung und Schwingungen

6.6.1. Harmonische Schwingungen in einer Dimension

Wir betrachten die folgende einfache Gleichung:

$$\frac{d^2 u(t)}{dt^2} + \lambda u(t) = 0, \quad \text{mit } \lambda > 0, \quad (6.63)$$

Anfangsbedingungen:

$$u(t=0) = u_0, \quad \left. \frac{du(t)}{dt} \right|_{t=0} = v_0. \quad (6.64)$$

Für eine Lösung dieser Gleichung gilt: Die zweite Ableitung der Funktion ist proportional der Funktion selbst. Das ist genau der Fall für Linearkombinationen der trigonometrischen Funktionen \sin und \cos . Der schnellste Weg zu einer Lösung führt also über den Ansatz

$$u(t) = a \cos \omega t + b \sin \omega t \quad (6.65)$$

mit noch zu bestimmenden Konstanten a , b und ω . Die *Kreisfrequenz* ω bestimmen wir durch Einsetzen in die Differentialgleichung. Die Funktion (6.65) erfüllt

$$\frac{d^2 u(t)}{dt^2} = -\omega^2 u(t). \quad (6.66)$$

Aus (6.63) folgt daher

$$\omega^2 = \lambda. \quad (6.67)$$

Wir wählen $\omega = \sqrt{\lambda}$ positiv. (Ein negatives ω bewirkt lediglich, dass in (6.65) die Konstante b durch $-b$ ersetzt wird).

Die Konstanten a und b bestimmt man aus den Anfangsbedingungen (6.64). Man erhält:

$$u(0) = u_0 = a, \quad u'(0) = v_0 = (-a\omega \sin \omega t + b\omega \cos \omega t)|_{t=0} = b\omega. \quad (6.68)$$

Die eindeutige Lösung, die die Anfangsbedingungen erfüllt, ist daher

$$u(t) = u_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t. \quad (6.69)$$

Diese Funktion beschreibt eine eindimensionale harmonische Schwingung mit der Kreisfrequenz ω . Die zeitliche Periode der Bewegung ist

$$T = \frac{2\pi}{\omega}. \quad (6.70)$$

6.6.2. Schwingende Systeme

Wir betrachten nun das Differentialgleichungssystem

$$\frac{d^2}{dt^2} \mathbf{u}(t) + \mathbf{K} \mathbf{u}(t) = 0, \quad (6.71)$$

wobei \mathbf{K} eine symmetrische, positiv definite $n \times n$ -Matrix ist. Entsprechend ist die Lösung eine vektorwertige Funktion der Zeit,

$$\mathbf{u} : \begin{cases} \mathbb{R} \rightarrow \mathbb{R}^n \\ t \mapsto \mathbf{u}(t) \end{cases} \quad (6.72)$$

Falls wir einen Eigenvektor \mathbf{x} von \mathbf{K} zum Eigenwert λ kennen,

$$\mathbf{K} \mathbf{x} = \lambda \mathbf{x}, \quad (6.73)$$

dann definiert

$$\mathbf{u}(t) = (a \cos \omega t + b \sin \omega t) \mathbf{x} \quad \text{mit } a, b \in \mathbb{R} \text{ und } \omega^2 = \lambda \quad (6.74)$$

eine Lösung von (6.71) (nachrechnen!). Eine solche Lösung heißt eine *Eigenschwingung* des Systems.

Für eine positiv definite $n \times n$ -Matrix gibt es n linear unabhängige Eigenvektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$, die zu positiven Eigenwerten $\omega_1^2, \dots, \omega_n^2$ von \mathbf{K} gehören. Die positiven Zahlen $\omega_1, \dots, \omega_n$ heißen die *Eigenfrequenzen* des Systems. Da die Gleichung linear ist, ist eine beliebige Linearkombination von Lösungen wieder eine Lösung. Man nennt

$$\mathbf{u}(t) = \sum_{j=1}^n (a_j \cos \omega_j t + b_j \sin \omega_j t) \mathbf{x}_j \quad (6.75)$$

die *allgemeine Lösung* des Gleichungssystems (6.71). Die Koeffizienten a_j und b_j bestimmt man aus den Anfangsbedingungen:

$$\mathbf{u}(0) = \mathbf{u}_0 \in \mathbb{R}^n, \quad \left. \frac{d}{dt} \mathbf{u}(t) \right|_{t=0} = \mathbf{v}_0 \in \mathbb{R}^n. \quad (6.76)$$

Das ergibt die $2n$ linearen Gleichungen

$$\mathbf{u}_0 = a_1 \mathbf{x}_1 + \dots + a_n \mathbf{x}_n \quad (6.77)$$

$$\mathbf{v}_0 = b_1 \omega_1 \mathbf{x}_1 + \dots + b_n \omega_n \mathbf{x}_n \quad (6.78)$$

für die $2n$ unbekanntenen Koeffizienten.

Sei \mathbf{S} die Matrix aus den Eigenvektoren, die die Matrix \mathbf{K} diagonalisiert, $\mathbf{S}^{-1} \mathbf{K} \mathbf{S} = \mathbf{\Lambda}$. Sei weiters $\mathbf{\Omega} = \mathbf{\Lambda}^{1/2}$, die Diagonalmatrix mit ω_i als i -tem Element in der Diagonale. Wir fassen die Koeffizienten zu Vektoren \mathbf{a} und \mathbf{b} zusammen. Dann lauten die Anfangsbedingungen:

$$\mathbf{u}_0 = \mathbf{S} \mathbf{a}, \quad \mathbf{v}_0 = \mathbf{S} \mathbf{\Omega} \mathbf{b} \quad (6.79)$$

und daraus ergeben sich die Koeffizientenvektoren zu

$$\mathbf{a} = \mathbf{S}^{-1} \mathbf{u}_0, \quad \mathbf{b} = \Omega^{-1} \mathbf{S}^{-1} \mathbf{v}_0. \quad (6.80)$$

6.6.3. Beispiel: Federkette

Wir betrachten hier das in Abbildung 6.1 dargestellte Federsystem als Beispiel. Die Geometrie wird durch die Inzidenzmatrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} \quad (6.81)$$

beschrieben. (Zeilenzahl = Zahl der Kanten/Federn, Spaltenzahl = Zahl der Knoten/Massen). Wir betrachten das System in Abwesenheit äußerer Kräfte. Wir bezeichnen die Auslenkungen der Massen zur Zeit t mit $u_i(t)$, $i = 1, 2$. Aus dem Vektor $\mathbf{u}(t) = (u_1(t), u_2(t))^T$ erhalten wir die Federelongationen als $\mathbf{A} \mathbf{u}(t)$ (unter der Annahme, daß die Federn entspannt sind, falls $\mathbf{u} = 0$ ist).

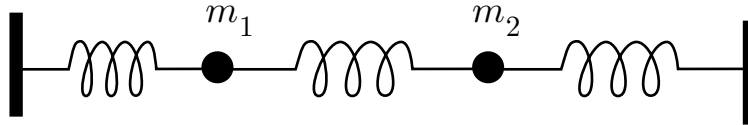


ABBILDUNG 6.1. Ein schwingfähiges System aus zwei Massen und drei Federn.

Wir nehmen an: Alle Massen und Federkonstanten haben den Zahlenwert 1. Die Matrix \mathbf{C} der Federkonstanten ist daher die Einheitsmatrix. Die Steifigkeits-Matrix $\mathbf{K} = \mathbf{A}^T \mathbf{C} \mathbf{A}$ ergibt sich somit zu

$$\mathbf{K} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \quad (6.82)$$

Die Steifigkeits-Matrix beschreibt auch die potentielle Energie des Systems als Funktion der Lage \mathbf{u} der Massen, siehe (4.18):

$$P(\mathbf{u}) = \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u}. \quad (6.83)$$

Die kinetische Energie ist übrigens einfach die Summe der kinetischen Energien der beiden Massen. Wir können die kinetische Energie als quadratische Form der Geschwindigkeiten schreiben. Mit

$$\dot{\mathbf{u}} = \frac{d}{dt} \mathbf{u}(t) \quad (6.84)$$

ist die kinetische Energie

$$T(\dot{\mathbf{u}}) = \frac{1}{2} \dot{\mathbf{u}}^\top \mathbf{T} \dot{\mathbf{u}} = \frac{m_1 \dot{u}_1^2}{2} + \frac{m_2 \dot{u}_2^2}{2}, \quad \mathbf{T} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}. \quad (6.85)$$

In unserem Beispiel ist \mathbf{T} wieder die Einheitsmatrix.

Die Bewegungsgleichung ergibt sich aus der Newtonschen Bewegungsgleichung: Masse mal Beschleunigung = Kraft = negativer Gradient der potentiellen Energie. Nun ist die Kraft

$$-\nabla P(\mathbf{u}) = - \begin{pmatrix} \frac{\partial}{\partial u_1} P(\mathbf{u}) \\ \frac{\partial}{\partial u_2} P(\mathbf{u}) \end{pmatrix} = -\mathbf{K} \mathbf{u}. \quad (6.86)$$

Daher ist die Bewegungsgleichung

$$\frac{d^2}{dt^2} \mathbf{u}(t) + \mathbf{K} \mathbf{u}(t) = 0. \quad (6.87)$$

Die Matrix \mathbf{K} hat die Eigenwerte 1 und 3, also sind die Frequenzen der Eigenschwingungen $\omega_1 = 1$ und $\omega_2 = \sqrt{3}$. Die Eigenvektoren wurden schon in (6.36) und (6.37) angegeben. Die allgemeine Lösung der Bewegungsgleichung lautet daher

$$\mathbf{u}(t) = (a_1 \cos t + b_1 \sin t) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (a_2 \cos \sqrt{3}t + b_2 \sin \sqrt{3}t) \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (6.88)$$

Dabei beschreibt die erste (bzw. zweite) Komponente von \mathbf{u} die Bewegung der ersten (bzw. zweiten) Masse. Im allgemeinen ist die Bewegung eine Überlagerung zweier Oszillationen mit einem irrationalen Frequenzverhältnis, ist also im allgemeinen nicht periodisch (es sei denn, die Anfangsbedingung verlangt zum Beispiel $a_2 = b_2 = 0$). Wenn Eigenschwingungen mit irrationaler Frequenzverhältnis zur Überlagerung kommen, gerät das System nie mehr in den Anfangszustand zurück. (Allerdings kommt das System zu späteren Zeiten immer wieder dem Anfangszustand sehr nahe – man bezeichnet eine solche Bewegung als quasiperiodisch). Bei der zur Frequenz $\omega_1 = 1$ gehörigen Eigenschwingung bewegen sich die beiden Massen “im Gleichtakt”. Bei der Eigenschwingung mit $\omega_2 = \sqrt{3}$ bewegen sich die Massen entgegengesetzt.

Die Anfangsbedingung

$$\mathbf{u}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}'_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (6.89)$$

führt zum Beispiel auf die quasiperiodische Bewegung

$$\mathbf{u}(t) = \frac{1}{2} \begin{pmatrix} \cos t + \cos \sqrt{3}t \\ \cos t - \cos \sqrt{3}t \end{pmatrix}. \quad (6.90)$$

Diese Bewegung kehrt niemals mehr in den Anfangszustand zurück. Die Bedingung daß zur Zeit t der Anfangszustand erreicht wird wäre nämlich, daß es ganze Zahlen n_1 und n_2 gibt mit

$$t = 2\pi n_1, \quad \sqrt{3}t = 2\pi n_2, \quad \text{oder} \quad \sqrt{3} = \frac{n_2}{n_1} \quad (\text{Widerspruch!}) \quad (6.91)$$

Allerdings kann das System zu späteren Zeiten t sehr nahe an die Anfangslage herankommen, wenn n_2/n_1 eine gute rationale Approximation an $\sqrt{3}$ ist.

6.7. Gedämpfte Schwingungen

Gleichungen höherer Ordnung führen auf natürliche Weise auf Gleichungssysteme erster Ordnung. Betrachte z.B. die Gleichung

$$\frac{d^2u(t)}{dt^2} + \mu \frac{du(t)}{dt} + \lambda u(t) = 0. \quad (6.92)$$

Wir führen eine weitere unbekannt Funktion die Ableitung

$$v(t) = \frac{du(t)}{dt} \quad (6.93)$$

ein. Dann bekommen wir das Gleichungssystem

$$\frac{du(t)}{dt} = v(t) \quad (6.94)$$

$$\frac{dv(t)}{dt} = -\lambda u(t) - \mu v(t). \quad (6.95)$$

Die Matrix dieses Gleichungssystems ist

$$A = \begin{pmatrix} 0 & 1 \\ -\lambda & -\mu \end{pmatrix}. \quad (6.96)$$

Diese Matrix hat die Eigenwerte und Eigenvektoren:

$$\lambda_1 = \frac{1}{2}(-\mu - \sqrt{\mu^2 - 4\lambda}), \quad \mathbf{x}_1 = \begin{pmatrix} -\mu + \sqrt{\mu^2 - 4\lambda} \\ 2\lambda \end{pmatrix}, \quad (6.97)$$

$$\lambda_2 = \frac{1}{2}(-\mu + \sqrt{\mu^2 - 4\lambda}), \quad \mathbf{x}_2 = \begin{pmatrix} -\mu - \sqrt{\mu^2 - 4\lambda} \\ 2\lambda \end{pmatrix}. \quad (6.98)$$