

**Skriptum zur Vorlesung**  
**Angewandte Mathematik**

**Christian Schmeiser**  
Institut für Angewandte und  
Numerische Mathematik, TU Wien

# 1.KAPITEL Einige grundlegende Techniken (angewandt auf gewöhnliche Differentialgleichungen)

## 1.1. Dimensionslose Variable und Parameter

**Beispiel 1.1:** Wir betrachten die Bewegung eines Körpers, der von der Erdoberfläche senkrecht in die Höhe geworfen wird und interessieren uns für den Zeitpunkt  $T^*$ , an dem er wieder zur Erdoberfläche zurückkehrt. Der Körper bewegt sich nur auf einem vom Erdmittelpunkt ausgehenden Strahl. Den Abstand von der Erdoberfläche zum Zeitpunkt  $t^*$  bezeichnen wir mit  $x^*(t^*)$ . Dann ist  $x^*$  Lösung des Anfangswertproblems

$$\frac{d^2 x^*}{dt^{*2}} = -\frac{gR^2}{(x^* + R)^2}, \quad x^*(0) = 0, \quad \frac{dx^*}{dt^*}(0) = V, \quad (1.1)$$

wobei  $g$  die Gravitationsbeschleunigung an der Erdoberfläche,  $R$  der Erdradius und  $V$  die Anfangsgeschwindigkeit ist. Luftwiderstand wurde im obigen Problem vernachlässigt. Der Aufschlagzeitpunkt  $T^*$  ist Lösung der Gleichung  $x^*(T^*) = 0$ .

Es ist unser Ziel, mathematische Modelle für physikalische Probleme in *dimensionslose Form* zu transformieren. Dieser Vorgang läßt sich in Teilschritte zerlegen:

**1. Schritt:** Man fertige eine Liste aller im Problem auftretenden Variablen und Parameter zusammen mit ihren Dimensionen an.

**Beispiel 1.1:** Bei Verwendung des CGS-Einheitensystems ergibt sich:

	Dimension
Variable: $x^*$	cm
$t^*$	s
Parameter: $g$	cm s <sup>-2</sup>
$R$	cm
$V$	cm s <sup>-1</sup>

**2. Schritt:** Für jede Variable  $v$  bestimme man eine *intrinsische Referenzgröße*  $p$  als Kombination von Parametern und führe  $v/p$  als neue dimensionslose Variable ein.

Als Konsequenz ergibt sich, daß das transformierte Problem nur mehr dimensionslose Parameter enthält, die aus den ursprünglichen Parametern berechnet werden können.

**Beispiel 1.1:** Wir wählen  $R$  als *charakteristische Länge*, d.h. als intrinsische Referenzgröße für  $x^*$  und  $R/V$  als charakteristische Zeit. Wir führen also die dimensionslosen Variablen

$$y = \frac{x^*}{R}, \quad \tau = \frac{t^*}{R/V}$$

ein. Das dimensionslose Problem hat dann die Form

$$\varepsilon \frac{d^2 y}{d\tau^2} = -\frac{1}{(y+1)^2}, \quad y(0) = 0, \quad \frac{dy}{d\tau}(0) = 1, \quad (1.2)$$

mit dem dimensionslosen Parameter

$$\varepsilon = \frac{V^2}{gR}.$$

Eine allgemeine Aussage über dimensionslose Variable bzw. Parameter beinhaltet der folgende

**“Satz” 1.1.** *Mathematische Modelle physikalischer Probleme können auf dimensionslose Form gebracht werden. Die dimensionslosen Parameter und intrinsischen Referenzgrößen können als Produkte von Potenzen der ursprünglichen Parameter gewählt werden.*

Das Wort “Satz” steht unter Anführungszeichen, weil die Voraussetzungen viel zu ungenau formuliert sind, als daß diese Bezeichnung im mathematischen Sinn gerechtfertigt wäre. Wir wollen uns hier darauf beschränken, festzuhalten, daß diese Formulierung präzisiert und unter gewissen Annahmen sogar bewiesen werden kann.

“Satz” 1.1 hat starke Auswirkungen. Er gestattet uns, die für ein Problem in Frage kommenden Referenzgrößen und dimensionslosen Parameter ohne Kenntnis des mathematischen Modells nur aus der Liste der Variablen und Parameter zu bestimmen.

Dabei geht man vor wie folgt. Zunächst hat man *Grundeinheiten* festzulegen. Dabei hat man im allgemeinen gewisse Freiheit. Im *CGS*-System sind cm, g und s die Grundeinheiten. Die dimensionsbehafteten Parameter  $\alpha_k$ ,  $k = 1, \dots, N$  haben dann die Einheiten

$$\text{cm}^{l_k} \text{g}^{m_k} \text{s}^{t_k}, \quad k = 1, \dots, N.$$

Die Einheit einer Größe  $\alpha$  wird also bestimmt durch einen 3-Vektor  $(l, m, t)$ . So wird z.B. eine dimensionslose Größe durch den Nullvektor charakterisiert. Einer Größe mit der Dimension einer Masse entspricht der Vektor  $(0, 1, 0)$ . Ist  $\alpha$  eine intrinsische Referenzgröße oder ein dimensionsloser Parameter, so gibt es laut “Satz” 1.1  $a_k$ ,  $k = 1, \dots, N$ , sodaß

$$\alpha = \prod_{k=1}^N \alpha_k^{a_k}$$

gilt. Betrachtet man die entsprechende Gleichung für die Dimensionen, so erhält man

$$\text{cm}^l \text{g}^m \text{s}^t = \prod_{k=1}^N \text{cm}^{a_k l_k} \text{g}^{a_k m_k} \text{s}^{a_k t_k} = \text{cm}^{\sum a_k l_k} \text{g}^{\sum a_k m_k} \text{s}^{\sum a_k t_k}.$$

Exponentenvergleich liefert das lineare Gleichungssystem

$$\begin{aligned} \sum_{k=1}^N a_k l_k &= l, \\ \sum_{k=1}^N a_k m_k &= m, \\ \sum_{k=1}^N a_k t_k &= t \end{aligned}$$

für die Exponenten  $a_k$ . “Satz” 1.1 garantiert die Existenz intrinsischer Referenzgrößen, und damit die Existenz einer Lösung des Gleichungssystems, wenn  $\alpha$  die Dimension einer Variablen hat. Sucht man nach dimensionslosen Parametern, dann hat man das homogene System zu betrachten. Es gilt: *Die Anzahl  $N^*$  der relevanten dimensionslosen Parameter ist gleich der Dimension des Kernes der Matrix*

$$\begin{pmatrix} l_1 & \cdots & l_N \\ m_1 & \cdots & m_N \\ t_1 & \cdots & t_N \end{pmatrix}, \quad (1.3)$$

und daher  $N - 3 \leq N^* \leq N$ . Zwei linear abhängige Elemente des Kernes würden auf zwei Parameter  $\alpha$  und  $\beta$  führen, für die eine Beziehung der Form  $\beta = \alpha^\lambda$  gilt.  $\beta$  könnte also durch  $\alpha$  ersetzt werden. Der Übergang auf die dimensionslose Form führt immer auf eine Reduktion der Anzahl der Parameter ( $N^* < N$ ), solange die Matrix (1.3) nicht die Nullmatrix ist, d.h. solange die ursprünglich gegebenen Parameter nicht schon alle dimensionslos waren.

**Beispiel 1.1:** Die Matrix (1.3) ist gegeben durch

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -2 & 0 & -1 \end{pmatrix}$$

und hat den Rang 2. Daher wäre auch ohne Kenntnis der Modellgleichungen die Aussage möglich gewesen, daß das dimensionslose Problem nur einen Parameter enthält. Der Parameter  $\varepsilon$  entspricht dem Element  $(-1, -1, 2)$  des Kernes.

**Beispiel 1.2:** Wir betrachten ein Pendel in Form eines starren gewichtslosen Stabes der Länge  $L$  mit einer Punktmasse  $m$  an einem Ende. Am anderen Ende sei er frei drehbar um eine horizontale Achse aufgehängt. Die Bewegung des Pendels wird durch die Gravitation verursacht. Reibungseffekte werden vernachlässigt. Wir bringen das Pendel in horizontale Lage, lassen es los und interessieren uns für die Dauer  $T^*$  der ersten Schwingung (siehe Abb. 1.1).

Abb. 1.1: Pendel

Die Bewegung des Pendels kann durch die Funktion  $\theta^*(t^*)$  beschrieben werden, wobei  $\theta^*$  der Winkel zwischen der Vertikalen und dem Pendel ist. An Parametern haben wir neben der Länge des Pendels  $L$  und der Masse  $m$  die Gravitationskonstante  $g$  zu berücksichtigen. Die Liste der Variablen und Parameter ist also

	Dimension
Variable: $\theta^*$	1
$t^*$	s
Parameter: $L$	cm
$m$	g
$g$	$\text{cm s}^{-2}$

Die abhängige Variable  $\theta^*$  ist bereits dimensionslos. Für die unabhängige Variable  $t^*$  und das Resultat  $T^*$  wählen wir als Einheit die Referenzzeit  $\sqrt{L/g}$ . Das dimensionslose Resultat ist

$$T = \frac{T^*}{\sqrt{L/g}}.$$

Die Matrix (1.3) ist gegeben durch

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Die Regularität dieser Matrix impliziert, daß es keine dimensionslosen Parameter gibt. Das dimensionslose Problem und damit auch dessen Lösung  $T$  sind also von den Parametern unabhängig. Das bedeutet, daß wir ohne Kenntnis der zugrunde liegenden Gleichungen unser Problem bis auf

einen konstanten Faktor gelöst haben: Die Durchführung eines einzigen Experimentes, bei dem  $T^*$  gemessen und damit der Wert von  $T$  bestimmt wird, löst das Problem für beliebige Pendel vollständig. Als Nebenresultat erhalten wir, daß die Lösung unabhängig von der Masse  $m$  ist.

## 1.2. Skalierung

**Beispiel 1.1:** (siehe Abschnitt 1.1) Eine zweite Möglichkeit für die Einführung von dimensionslosen Variablen ist

$$z = \frac{x^*}{R}, \quad \tau_1 = \frac{t^*}{\sqrt{R/g}}.$$

Nun erhält das Problem die Form

$$\frac{d^2 z}{d\tau_1^2} = -\frac{1}{(z+1)^2}, \quad z(0) = 0, \quad \frac{dz}{d\tau_1}(0) = \sqrt{\varepsilon}. \quad (1.4)$$

Wir interessieren uns für den Fall, daß  $\varepsilon$  *klein ist im Vergleich zu 1*. In Symbolen drücken wir diesen Sachverhalt aus durch

$$\varepsilon \ll 1.$$

Man könnte annehmen, daß man die Lösung in diesem Fall approximieren kann, wenn man in (1.2) und (1.4)  $\varepsilon = 0$  setzt und die so entstehenden *reduzierten Probleme* löst. Es zeigt sich allerdings, daß mit  $\varepsilon = 0$  aus (1.2) ein unlösbares Problem entsteht, und daß die Lösung des reduzierten Problems zu (1.4) für  $\tau_1 > 0$  negativ ist und daher keine sinnvolle Approximation für die Lösung des ursprünglichen Problems sein kann.

Der Grund für das Versagen der reduzierten Probleme liegt in einer ungünstigen Wahl der Referenzgrößen. Beim Grenzübergang  $\varepsilon \rightarrow 0$  wurde implizit die Annahme verwendet, daß in den dimensionslosen Gleichungen vorkommende Variable und deren Ableitungen von der Größenordnung 1 sind, d.h. *daß die Größenordnungen der einzelnen Terme durch als Faktoren auftretende dimensionslose Parameter bestimmt sind*. Eine Wahl von Referenzgrößen, für die diese Annahme gerechtfertigt ist, nennt man *Skalierung*.

EINSCHUB. Landau'sche Ordnungssymbole—Präzisierung des Begriffes "Größenordnung": Seien  $f(\varepsilon)$ ,  $g(\varepsilon) \in \mathbb{R}$  für  $\varepsilon \in (0, \varepsilon_0]$ .

**Definition 1.2.** a)  $f(\varepsilon) = O(g(\varepsilon))$  für  $\varepsilon \rightarrow 0$  gilt genau dann, wenn es eine (von  $\varepsilon$  unabhängige) Konstante  $K > 0$  gibt, sodaß  $|f(\varepsilon)| \leq K|g(\varepsilon)|$  für  $\varepsilon \in (0, \varepsilon_0]$  gilt.

b)  $f(\varepsilon) = o(g(\varepsilon))$  für  $\varepsilon \rightarrow 0$  gilt genau dann, wenn  $\lim_{\varepsilon \rightarrow 0} |f(\varepsilon)/g(\varepsilon)| = 0$  gilt.

c)  $f(\varepsilon) = O_s(g(\varepsilon))$  für  $\varepsilon \rightarrow 0$  (in Worten:  $f(\varepsilon)$  und  $g(\varepsilon)$  haben dieselbe Größenordnung für  $\varepsilon \rightarrow 0$ ) gilt genau dann, wenn sowohl  $f(\varepsilon) = O(g(\varepsilon))$  als auch  $g(\varepsilon) = O(f(\varepsilon))$  für  $\varepsilon \rightarrow 0$  gelten.

Da im allgemeinen Probleme skaliert werden sollen, ohne sie vorher zu lösen, ist es oft eine schwierige Aufgabe, eine geeignete Skalierung zu finden.

**Beispiel 1.1:** (siehe Abschnitt 1.1) Eine einfache Überlegung führt auf eine Skalierung des Problems (1.1) für  $\varepsilon \ll 1$ . In diesem Fall ist zu erwarten, daß der maximale Abstand des Körpers von der Erdoberfläche klein ist im Vergleich zum Erdradius. Vernachlässigt man  $x^*$  im Vergleich zu  $R$  auf der rechten Seite der Differentialgleichung in (1.1), so tritt  $R$  nicht mehr als Parameter auf. Es ist daher naheliegend,  $R$  für die Konstruktion der Referenzgrößen nicht zu verwenden. Damit ergeben sich die Referenzlänge  $V^2/g$  und die Referenzzeit  $V/g$ . Bei Verwendung der dimensionslosen Variablen

$$x = \frac{x^*}{V^2/g}, \quad t = \frac{t^*}{V/g}$$

erhält das Problem die Form

$$\frac{d^2 x}{dt^2} = -\frac{1}{(\varepsilon x + 1)^2}, \quad x(0) = 0, \quad \frac{dx}{dt}(0) = 1. \quad (1.5)$$

Das reduzierte Problem

$$\frac{d^2 x_0}{dt^2} = -1, \quad x_0(0) = 0, \quad \frac{dx_0}{dt}(0) = 1,$$

hat die Lösung  $x_0 = t - t^2/2$ , die den Aufschlagzeitpunkt  $T^* = 2V/g$  vorhersagt. Die verwendete Näherung entspricht der vereinfachenden Annahme, daß die durch die Gravitation verursachte Beschleunigung unabhängig vom Abstand von der Erdoberfläche ist. Die Tatsache, daß das reduzierte Problem eine sinnvoll erscheinende Lösung liefert, läßt die Hoffnung zu, daß diese auch wirklich eine Approximation für die Lösung des vollen Problems (1.5) darstellt.

### 1.3. Asymptotische Entwicklungen

Wir betrachten für  $0 < \varepsilon \leq \varepsilon_0$  die Elemente  $u_\varepsilon$  und  $v_\varepsilon$  eines normierten Raumes  $(B, \|\cdot\|)$ , die von  $\varepsilon$  abhängen. Dann bezeichnen wir  $u_\varepsilon$  und  $v_\varepsilon$  als *asymptotisch äquivalent* für  $\varepsilon \rightarrow 0$ , wenn

$$\|u_\varepsilon - v_\varepsilon\| = o(\|u_\varepsilon\|) \quad \text{für } \varepsilon \rightarrow 0$$

gilt, d.h. wenn der relative Fehler bei der Approximation von  $u_\varepsilon$  durch  $v_\varepsilon$  (und umgekehrt) für  $\varepsilon \rightarrow 0$  gegen Null geht. Man beachte, daß der Ausdruck "äquivalent" gerechtfertigt ist, da man leicht zeigen kann, daß die obige Definition symmetrisch bezüglich  $u_\varepsilon$  und  $v_\varepsilon$  ist. Im folgenden wird zumeist der Zusatz "für  $\varepsilon \rightarrow 0$ " (auch bei Verwendung der Landau'schen Symbole) weggelassen. Betrachtet man  $v_\varepsilon$  als Approximation für  $u_\varepsilon$ , so wird auch die Sprechweise: " $v_\varepsilon$  ist eine *asymptotische Näherung* für  $u_\varepsilon$ " verwendet.

**Beispiel 1.3:** In  $B = \mathbb{R}$  sind

$$\begin{aligned} 3 + \varepsilon & \quad \text{und} \quad 3 - \varepsilon^2, \\ \frac{1}{\varepsilon} & \quad \text{und} \quad \frac{1}{\varepsilon} + 1, \\ (1 + \sqrt{\varepsilon}) \sin \frac{1}{\varepsilon} & \quad \text{und} \quad \sin \frac{1}{\varepsilon} \end{aligned}$$

asymptotisch äquivalent, aber nicht

$$\sin \frac{1 + \varepsilon \pi}{\varepsilon} \quad \text{und} \quad \sin \frac{1}{\varepsilon}.$$

**Beispiel 1.4:** Die reelle Funktion  $u_\varepsilon : [0, 1] \rightarrow \mathbb{R}$  ist gegeben durch

$$u_\varepsilon(x) = e^{-x/\varepsilon} + x.$$

Für positive  $x$  konvergiert  $u_\varepsilon(x)$  für  $\varepsilon \rightarrow 0$  gegen  $u_0(x) = x$ . Betrachtet man  $u_\varepsilon$  als Element des Banachraumes  $C([0, 1])$  mit der Norm

$$\|u\|_\infty = \max_{x \in [0, 1]} |u(x)|,$$

dann ist  $u_0$  nicht asymptotisch äquivalent mit  $u_\varepsilon$ . Man sieht leicht, daß es keine von  $\varepsilon$  unabhängige asymptotische Näherung für  $u_\varepsilon$  gibt. Wählt man dagegen als zugrunde liegenden Raum den Banachraum  $L^p((0, 1))$  mit der Norm

$$\|u\|_p = \left( \int_0^1 |u(x)|^p dx \right)^{1/p},$$

dann ist  $u_0$  eine asymptotische Näherung, weil

$$\|u_\varepsilon - u_0\|_p = \left( \int_0^1 e^{-xp/\varepsilon} dx \right)^{1/p} = \left( \frac{\varepsilon}{p} (1 - e^{-p/\varepsilon}) \right)^{1/p} = o(1)$$

gilt.

Seien  $u_k \in B$  und es gelte  $\|u_k\| = O(1)$  für  $k = 0, 1, \dots$ . Dann heißt die formale Reihe

$$\sum_{k=0}^{\infty} u_k \varepsilon^k$$

*asymptotische Entwicklung* nach Potenzen von  $\varepsilon$  für  $u_\varepsilon$ , wenn

$$\left\| u_\varepsilon - \sum_{k=0}^n u_k \varepsilon^k \right\| = o(\varepsilon^n), \quad n = 0, 1, \dots \quad (1.6)$$

gilt. In diesem Fall verwenden wir die Schreibweise

$$u_\varepsilon \sim \sum_{k=0}^{\infty} u_k \varepsilon^k.$$

Gilt (1.6) nur für  $n = 0, \dots, N$ , so spricht man von einer asymptotischen Entwicklung der Ordnung  $N$ .

**Bemerkung:** Statt  $\|u_\varepsilon - v_\varepsilon\| = O(f(\varepsilon))$  werden wir in Zukunft oft

$$u_\varepsilon = v_\varepsilon + O(f(\varepsilon))$$

schreiben, wenn kein Zweifel über die verwendete Norm besteht (analog für  $o$  und  $O_s$ ).

Aus dem Taylorschen Lehrsatz folgt, daß Größen, deren Ableitungen nach  $\varepsilon$  bis zur Ordnung  $N + 1$  in einer Umgebung von  $\varepsilon = 0$  beschränkt sind, asymptotische Entwicklungen der Ordnung  $N$  besitzen, deren Koeffizienten unabhängig von  $\varepsilon$  sind. Die Koeffizienten haben dann die Form

$$u_k = \frac{1}{k!} \left. \frac{d^k u_\varepsilon}{d\varepsilon^k} \right|_{\varepsilon=0}.$$

Die rechte Seite von (1.6) kann in diesem Fall durch  $O(\varepsilon^{n+1})$  ersetzt werden.

Der wesentliche Unterschied zwischen asymptotischen Entwicklungen und Taylorentwicklungen ist, daß man bei asymptotischen Entwicklungen am Grenzübergang  $\varepsilon \rightarrow 0$  bei fester Ordnung interessiert ist, während sich bei Taylorentwicklungen die Frage nach der Konvergenz der unendlichen Reihe für festes  $\varepsilon$  stellt. Man kann Funktionen angeben, für die asymptotische Entwicklungen beliebiger Ordnung existieren, wobei die entsprechenden unendlichen Reihen für jedes positive  $\varepsilon$  divergieren.

#### 1.4. Regulär gestörte Probleme

Sei  $F : B_1 \times [0, \varepsilon_0] \rightarrow B_2$  eine Abbildung, wobei  $(B_1, \|\cdot\|_1)$  und  $(B_2, \|\cdot\|_2)$  Banachräume sind. Wir betrachten für  $\varepsilon \ll 1$  das Problem

$$F(y_\varepsilon, \varepsilon) = 0. \quad (1.7)$$

Es ist naheliegend, das *reduzierte Problem*

$$F(y_0, 0) = 0 \quad (1.8)$$

zur Bestimmung von Näherungen  $y_0$  für Lösungen  $y_\varepsilon \in B_1$  des vollen Problems (1.7) zu verwenden. Hat das Problem (1.8) eine Lösung, dann stellt sich die Frage nach der *Konvergenz* der Approximation, d.h. ob

$$\lim_{\varepsilon \rightarrow 0} y_\varepsilon = y_0 \quad \text{in } B_1$$

gilt. Hängt  $F$  stetig von  $\varepsilon$  ab, dann ist die Näherung  $y_0$  zumindest *konsistent*. Konsistenz bedeutet, daß das *Residuum*  $r_\varepsilon = F(y_0, \varepsilon)$  gegen Null konvergiert:

$$\lim_{\varepsilon \rightarrow 0} r_\varepsilon = 0 \quad \text{in } B_2$$

In diesem Fall bezeichnen wir  $y_0$  auch als *formale asymptotische Näherung* für eine Lösung von (1.7). Im Abschnitt 1.6 beschäftigen wir uns mit der Frage, welche zusätzlichen Annahmen notwendig sind, damit aus Konsistenz Konvergenz folgt.

Die Approximation  $y_0$  könnte verbessert werden, indem man versucht, Terme in einer asymptotischen Entwicklung

$$y_\varepsilon \sim \sum_{k=0}^{\infty} y_k \varepsilon^k \quad (1.9)$$

zu berechnen. Dabei geht man folgendermaßen vor: Der Ansatz (1.9) wird in das Problem (1.7) eingesetzt, und die linke Seite nach Potenzen von  $\varepsilon$  entwickelt. Die Koeffizienten in dieser Entwicklung müssen verschwinden. Diese Bedingung liefert Gleichungen zur Bestimmung der  $y_k$  in (1.9).

Die Durchführbarkeit dieses Programmes wird garantiert durch die Annahme, daß  $F(y, \varepsilon)$  für festes  $y \in B_1$  eine asymptotische Entwicklung nach Potenzen von  $\varepsilon$  besitzt mit Koeffizienten, die glatte Funktionen von  $y$  sind:

**Lemma 1.3.** *Es gelte*

$$F(y, \varepsilon) = \sum_{k=0}^N F_k(y) \varepsilon^k + O(\varepsilon^{N+1})$$

mit Koeffizienten  $F_k(y)$ , deren Frechet-Ableitungen bis zur Ordnung  $N+1$  in einer Umgebung von  $y_0$  existieren. Weiters gelte

$$y_\varepsilon = \sum_{k=0}^N y_k \varepsilon^k + O(\varepsilon^{N+1}).$$

Dann besitzt  $F(y_\varepsilon, \varepsilon)$  eine asymptotische Entwicklung der Ordnung  $N$ , in der der Koeffizient von  $\varepsilon^0$  gegeben ist durch  $F_0(y_0) = F(y_0, 0)$  und die Koeffizienten von  $\varepsilon^k$ ,  $k \geq 1$  die Form

$$DF_0(y_0)y_k + G_k(y_0, \dots, y_{k-1})$$

besitzen. Hier bezeichnet  $DF_0(y_0)$  die Frechet-Ableitung von  $F_0$  an der Stelle  $y_0$ .

Der Beweis beruht auf einer Verallgemeinerung der Taylorschen Formel auf Funktionen in Banachräumen.

Als Konsequenz des Lemmas muß klarerweise der führende Term  $y_0$  in der asymptotischen Entwicklung das reduzierte Problem (1.8) lösen. Terme höherer Ordnung sind als Lösungen von linearisierten Versionen des reduzierten Problems zu bestimmen. Wegen der Gestalt der Inhomogenitäten können sie rekursiv berechnet werden, wenn der lineare Operator  $DF_0(y_0)$  invertierbar ist.



Verwenden wir Partialsummen

$$Y_n = \sum_{k=0}^n y_k \varepsilon^k$$

der konstruierten Entwicklung als Approximationen für die Lösung des vollen Problems (1.7), so ergibt sich für das Residuum

$$F(Y_n, \varepsilon) = O(\varepsilon^{n+1}), \quad n = 0, \dots, N.$$

Wir bezeichnen daher die konstruierte Reihe als *formale asymptotische Entwicklung* der Ordnung  $N$  für eine Lösung von (1.7). Ein Problem der Form (1.7), für das eine formale asymptotische Entwicklung konstruiert werden kann, nennen wir *regulär gestört*. Andernfalls bezeichnen wir das Problem als *singulär gestört*.

**Beispiel 1.1:** (siehe Abschnitt 1.1) Wir betrachten die skalierte Version

$$x'' = -\frac{1}{(\varepsilon x + 1)^2}, \quad x(0) = 0, \quad x'(0) = 1, \quad (1.10)$$

und versuchen, eine formale asymptotische Entwicklung für die Lösung zu konstruieren. Substitution des Ansatzes

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + O(\varepsilon^3)$$

in das Problem liefert

$$\begin{aligned} x_0'' + \varepsilon x_1'' + \varepsilon^2 x_2'' + O(\varepsilon^3) &= -(1 + \varepsilon x_0 + \varepsilon^2 x_1 + \varepsilon^3 x_2 + O(\varepsilon^4))^{-2}, \\ x_0(0) + \varepsilon x_1(0) + \varepsilon^2 x_2(0) + O(\varepsilon^3) &= 0, \\ x_0'(0) + \varepsilon x_1'(0) + \varepsilon^2 x_2'(0) + O(\varepsilon^3) &= 1. \end{aligned}$$

Nun entwickeln wir die rechte Seite der Differentialgleichung um  $\varepsilon = 0$ :

$$-(1 + \varepsilon x_0 + \varepsilon^2 x_1 + \varepsilon^3 x_2 + O(\varepsilon^4))^{-2} = -1 + \varepsilon 2x_0 + \varepsilon^2(2x_1 - 3x_0^2) + O(\varepsilon^3)$$

Koeffizientenvergleich liefert eine Folge von Problemen für die Bestimmung der  $x_k$ .

Koeffizienten von  $\varepsilon^0$ :

$$x_0'' = -1, \quad x_0(0) = 0, \quad x_0'(0) = 1$$

Koeffizienten von  $\varepsilon^1$ :

$$x_1'' = 2x_0, \quad x_1(0) = x_1'(0) = 0$$

Koeffizienten von  $\varepsilon^2$ :

$$x_2'' = 2x_1 - 3x_0^2, \quad x_2(0) = x_2'(0) = 0$$

Das Problem für  $x_0$  ist das reduzierte Problem, das wir bereits gelöst haben:

$$x_0(t) = t - \frac{t^2}{2}$$

Für die weiteren Koeffizienten ergibt sich

$$\begin{aligned} x_1(t) &= \frac{t^3}{3} \left(1 - \frac{1}{4}t\right), \\ x_2(t) &= -\frac{t^4}{4} \left(1 - \frac{11}{15}t + \frac{11}{90}t^2\right). \end{aligned}$$

Wir erhalten also die formale Entwicklung

$$x(t) = t - \frac{t^2}{2} + \varepsilon \frac{t^3}{3} \left(1 - \frac{1}{4}t\right) - \varepsilon^2 \frac{t^4}{4} \left(1 - \frac{11}{15}t + \frac{11}{90}t^2\right) + O(\varepsilon^3).$$

Das ist ein für Anwendungen zufriedenstellendes Resultat. Genaugenommen wissen wir allerdings noch nicht, ob die obige Beziehung für die Lösung von (1.10) wirklich gültig ist. Ja wir haben uns noch nicht einmal überlegt, ob für (1.10) eine Lösung existiert.

### 1.5. Singulär gestörte Probleme I: Die van der Pol-Gleichung

Die Schwingungen gewisser selbsterregter Systeme werden durch die van der Pol-Gleichung

$$\ddot{u} + u = \varepsilon(1 - u^2)\dot{u} \quad (1.11)$$

beschrieben. Diese Gleichung ist in dimensionsloser Form. Der Punkt bedeutet Ableitung nach der skalierten Zeit  $t$ , und  $\varepsilon$  ist ein kleiner positiver Parameter. Wir betrachten das Anfangswertproblem für (1.11) mit den Anfangsbedingungen

$$u(0) = \bar{u}, \quad \dot{u}(0) = 0, \quad (1.12)$$

und interessieren uns für das Verhalten der Lösung für  $t \rightarrow \infty$ .

Das reduzierte Problem

$$\ddot{u}_0 + u_0 = 0, \quad u_0(0) = \bar{u}, \quad \dot{u}_0(0) = 0,$$

hat die Lösung

$$u_0(t) = \bar{u} \cos t.$$

Das Residuum

$$\varepsilon(1 - u_0^2)\dot{u}_0 = -\varepsilon(1 - \bar{u}^2 \cos^2 t)\bar{u} \sin t$$

ist  $O(\varepsilon)$ , und damit ist  $u_0$  eine formale Näherung für die Lösung von (1.11), (1.12). Mit dem Ansatz  $u = u_0 + \varepsilon u_1 + O(\varepsilon^2)$  ist der Korrekturterm  $u_1$  Lösung des Problems

$$\begin{aligned} \ddot{u}_1 + u_1 &= \bar{u} \left( \frac{\bar{u}^2}{4} - 1 \right) \sin t + \frac{\bar{u}^3}{4} \sin 3t, \\ u_1(0) &= \dot{u}_1(0) = 0. \end{aligned}$$

Offensichtlich erzeugt der erste Summand in der Inhomogenität Resonanz. Die Lösung ist

$$u_1(t) = \frac{\bar{u}}{2} \left( 1 - \frac{\bar{u}^2}{4} \right) (t \cos t - \sin t) + \frac{\bar{u}^3}{32} (3 \sin t - \sin 3t).$$

Da  $u_1(t)$  unbeschränkt ist, ist die "verbesserte" Näherung  $u_0 + \varepsilon u_1$  inkonsistent in dem Sinn, daß das Residuum, als Funktion auf dem Intervall  $[0, \infty)$  betrachtet, für  $\varepsilon \rightarrow 0$  zwar punktweise aber nicht gleichmäßig gegen Null konvergiert.

Da die direkte Methode zur Approximation der Lösung hier versagt, ist das Problem (1.11), (1.12) singular gestört. Man beachte, daß das Problem regulär gestört ist, wenn wir uns auf ein endliches Zeitintervall einschränken. Um das Versagen der direkten Methode zu verstehen, betrachten wir zwei einfachere Beispiele.

**Beispiel 1.5:** Das Anfangswertproblem

$$\ddot{u} + u = \varepsilon u, \quad u(0) = \bar{u}, \quad \dot{u}(0) = 0$$

hat die Lösung

$$u(t) = \bar{u} \cos(\sqrt{1 - \varepsilon} t).$$

Offensichtlich hat die formale Näherung  $\bar{u} \cos t$  für  $t$  groß genug mit der Lösung nichts mehr zu tun. Auch der kleine Unterschied in der Frequenz bewirkt nach langer Zeit deutliche Unterschiede in der Lösung. Eine Verbesserung der reduzierten Lösung erreicht man durch asymptotische Entwicklung der Frequenz:

$$\sqrt{1 - \varepsilon} = 1 - \frac{\varepsilon}{2} + O(\varepsilon^2)$$

Die Näherung

$$U_0(t) = \cos\left(t - \frac{\varepsilon t}{2}\right)$$

ist zwar auch nicht gleichmäßig auf  $[0, \infty)$  gültig, aber ihr Gültigkeitsbereich hat sich wesentlich vergrößert. Während  $u_0(t)$  die exakte Lösung approximiert, solange  $t = o(\varepsilon^{-1})$  gilt, ist  $U_0(t)$  für  $t = o(\varepsilon^{-2})$  eine gültige Approximation. Durch Berücksichtigung von weiteren Termen in der Entwicklung der Frequenz kann der Gültigkeitsbereich weiter vergrößert werden. Im Gegensatz dazu führt die direkte Methode des vorigen Abschnittes zwar zu erhöhter Genauigkeit auf beschränkten Intervallen aber zu keiner Vergrößerung des Gültigkeitsintervalles.

**Beispiel 1.6:** Die Lösung des Problems

$$\ddot{u} + u = -2\varepsilon\dot{u}, \quad u(0) = \bar{u}, \quad \dot{u}(0) = 0$$

ist gegeben durch

$$u(t) = \bar{u}e^{-\varepsilon t} \cos\left(\sqrt{1 - \varepsilon^2} t\right).$$

Hier wird nicht nur die Frequenz, sondern auch die Amplitude der Schwingung durch die Störung verändert. Das Langzeitverhalten der exakten Lösung und das der Lösung des reduzierten Problems unterscheiden sich wesentlich voneinander. Während die Lösung des reduzierten Problems periodisch ist, konvergiert die Lösung des vollen Problems für  $t \rightarrow \infty$  gegen Null. Näherungen, die auf immer längeren Intervallen gültig sind, können auch hier durch Entwicklung der Frequenz bestimmt werden.

Das Auftreten der Kombination  $\varepsilon t$  in den Näherungslösungen bei den beiden Beispielen könnte auf die Vermutung führen, daß ein Grund für unsere Schwierigkeiten in einer ungünstigen Skalierung der Zeit liegt. Es zeigt sich allerdings, daß es auch nicht weiterhilft, statt  $t$  die umskalierte Zeitvariable  $T_1 = \varepsilon t$  zu verwenden. Aus Beispiel 1.6 ist ersichtlich, daß hier Vorgänge mit verschiedenen charakteristischen Zeiten gleichzeitig ablaufen.

Diese Erkenntnis ist die Grundlage für die *Methode der mehrfachen Skalierungen*. Sie besteht darin, für die Koeffizienten in der Entwicklung  $u = u_0 + \varepsilon u_1 + \dots$  einen Ansatz der Form

$$u_k = u_k(T_0, T_1, T_2, \dots)$$

zu machen, wobei die unabhängigen Variablen

$$T_i = \varepsilon^i t$$

verwendet werden. Mit Hilfe der Kettenregel erhalten wir

$$\begin{aligned} \dot{u}_k &= \frac{\partial u_k}{\partial T_0} + \varepsilon \frac{\partial u_k}{\partial T_1} + O(\varepsilon^2), \\ \ddot{u}_k &= \frac{\partial^2 u_k}{\partial T_0^2} + 2\varepsilon \frac{\partial^2 u_k}{\partial T_0 \partial T_1} + O(\varepsilon^2). \end{aligned}$$

Setzen wir einen derartigen Ansatz in die van der Pol-Gleichung ein, so ergibt sich

$$\frac{\partial^2 u_0}{\partial T_0^2} + 2\varepsilon \frac{\partial^2 u_0}{\partial T_0 \partial T_1} + \varepsilon \frac{\partial^2 u_1}{\partial T_0^2} + u_0 + \varepsilon u_1 = \varepsilon(1 - u_0^2) \frac{\partial u_0}{\partial T_0} + O(\varepsilon^2).$$

Koeffizientenvergleich liefert

$$\begin{aligned} \frac{\partial^2 u_0}{\partial T_0^2} + u_0 &= 0, \\ \frac{\partial^2 u_1}{\partial T_0^2} + u_1 &= -2 \frac{\partial^2 u_0}{\partial T_0 \partial T_1} + (1 - u_0^2) \frac{\partial u_0}{\partial T_0}. \end{aligned}$$

Zunächst scheint diese Methode das Problem wesentlich komplizierter zu machen, da sie auf Systeme von partiellen Differentialgleichungen führt. Man sieht jedoch leicht, daß die Koeffizienten in der Entwicklung

rekursiv durch Lösen gewöhnlicher Differentialgleichungen bestimmt werden können, die inhomogene Versionen der reduzierten Gleichung sind. Die Gleichung für  $u_0$  hat die allgemeine Lösung

$$u_0 = a(T_1, T_2, \dots) \cos(T_0 + b(T_1, T_2, \dots))$$

mit beliebigen  $a, b$ , die von  $T_0$  unabhängig sind. Aus den Anfangsbedingungen folgt

$$a(0, 0, \dots) = \bar{u}, \quad b(0, 0, \dots) = 0.$$

Die rechte Seite in der Gleichung für  $u_1$  ist damit

$$\begin{aligned} & 2 \frac{\partial a}{\partial T_1} \sin \phi + 2a \frac{\partial b}{\partial T_1} \cos \phi - a \sin \phi (1 - a^2 \cos^2 \phi) \\ &= \left( 2 \frac{\partial a}{\partial T_1} - a + \frac{a^3}{4} \right) \sin \phi + 2a \frac{\partial b}{\partial T_1} \cos \phi + \frac{a^3}{4} \sin 3\phi, \end{aligned}$$

wobei wir  $\phi = T_0 + b$  gesetzt haben. Ähnlich wie bei der direkten Methode würden die Terme mit  $\sin \phi$  und  $\cos \phi$  Resonanz erzeugen. Nun können wir aber die Freiheit bei der Wahl von  $a$  und  $b$  dazu nützen, diese Terme auszuschalten. Damit ergeben sich für  $a$  und  $b$  die gewöhnlichen Differentialgleichungen

$$\begin{aligned} \frac{\partial a}{\partial T_1} &= \frac{a}{8}(4 - a^2), \\ \frac{\partial b}{\partial T_1} &= 0. \end{aligned}$$

Suchen wir eine formale Näherung, die ein  $O(\varepsilon^2)$ -Residuum erzeugt, dann können wir in  $a$  und  $b$  die Abhängigkeit von den Variablen  $T_2, T_3, \dots$ , und in  $u_1$  die Abhängigkeit von  $T_1, T_2, \dots$  vernachlässigen. Die Gleichung für  $a$  hat die stabilen stationären Punkte  $a = \pm 2$  und den instabilen stationären Punkt  $a = 0$ . Mit den Anfangsbedingungen für  $a$  und  $b$  erhalten wir

$$a = \frac{2\bar{u}}{\sqrt{\bar{u}^2 - (\bar{u}^2 - 4)e^{-\varepsilon t}}}, \quad b = 0.$$

Die Approximation

$$u_0(t) = \frac{2\bar{u} \cos t}{\sqrt{\bar{u}^2 - (\bar{u}^2 - 4)e^{-\varepsilon t}}}$$

nähert sich für  $t \rightarrow \infty$  einer periodischen Schwingung mit Amplitude 2 und Periode  $2\pi$ , d.h. es tritt ein sogenannter *Grenzykel* auf. Die Existenz eines solchen Grenzykels kann auch für das volle Problem bewiesen werden.

## 1.6. Singulär gestörte Probleme II: Grenzschichten

**Beispiel 1.4:** (siehe Abschnitt 1.3) Die Funktion

$$u_\varepsilon(x) = e^{-x/\varepsilon} + x$$

ist die Lösung des Anfangswertproblems

$$\varepsilon u' = -u + x + \varepsilon, \quad u(0) = 1.$$

Das reduzierte Problem

$$0 = -u_0 + x, \quad u_0(0) = 1$$

hat keine Lösung. Das Problem ist daher singulär gestört. Die Lösung  $\bar{u}(x) = x$  der reduzierten Differentialgleichung approximiert die exakte Lösung überall außer in der Nähe von  $x = 0$ .

Dieses Beispiel motiviert die folgende Definition.

**Definition 1.4.** Sei  $D \subset \mathbb{R}^n$  ein Gebiet und für  $0 < \varepsilon \leq \varepsilon_0$  sei  $u_\varepsilon \in C(\bar{D})$  eine reelle Funktion.

a) Die Funktion  $u_\varepsilon$  heißt *regulär* in  $D$ , wenn der Grenzwert  $\lim_{\varepsilon \rightarrow 0} u_\varepsilon$  bezüglich der Norm

$$\|u\|_D := \sup_{\mathbf{x} \in D} |u(\mathbf{x})|$$

existiert.

b) Sei  $S \subset \bar{D}$  eine  $C^1$ -Mannigfaltigkeit mit Dimension kleiner als  $n$ . Die Funktion  $u_\varepsilon$  hat *Grenzschichtverhalten* an  $S$ , wenn  $u_\varepsilon$  nicht regulär in  $D$  ist, allerdings regulär in jedem  $D_1$  mit  $\bar{D}_1 \subset D \setminus S$  ist.

**Satz 1.5.** Hat  $u_\varepsilon$  Grenzschichtverhalten an  $S$ , dann gibt es eine von  $\varepsilon$  unabhängige Funktion  $\bar{u} \in C(\bar{D} \setminus S)$  mit

$$\lim_{\varepsilon \rightarrow 0} \|u_\varepsilon - \bar{u}\|_{D_1} = 0 \quad \text{für } \bar{D}_1 \subset \bar{D} \setminus S. \quad (1.13)$$

*Beweis.* Für jedes  $D_1$  mit  $\bar{D}_1 \subset \bar{D} \setminus S$  gibt es eine Funktion  $\bar{u}_{D_1} \in C(\bar{D}_1)$  mit

$$\lim_{\varepsilon \rightarrow 0} \|u_\varepsilon - \bar{u}_{D_1}\|_{D_1} = 0.$$

Weiters gilt für  $\bar{D}_1, \bar{D}_2 \subset \bar{D} \setminus S$  und  $\mathbf{x} \in D_1 \cap D_2$  offensichtlich

$$\bar{u}_{D_1 \cap D_2}(\mathbf{x}) = \bar{u}_{D_1}(\mathbf{x}) = \bar{u}_{D_2}(\mathbf{x}).$$

Daher können wir eine Funktion  $\bar{u} : \bar{D} \setminus S \rightarrow \mathbb{R}$  definieren durch

$$\bar{u}(\mathbf{x}) = \bar{u}_{U(\mathbf{x})}(\mathbf{x}),$$

wobei  $U(\mathbf{x})$  eine beliebige Umgebung von  $\mathbf{x}$  in  $\bar{D} \setminus S$  ist. Klarerweise gilt (1.13). Die Stetigkeit von  $\bar{u}$  folgt aus der gleichmäßigen Konvergenz. ■

**Beispiel 1.7:** a)

$$u_\varepsilon(x) = e^{-x/\varepsilon} + x, \quad D = (0, 1), \quad S = \{0\}, \quad \bar{u}(x) = x.$$

b)

$$u_\varepsilon(x, y) = \tanh \frac{x - y^2}{\varepsilon}, \quad D = \mathbb{R}^2, \quad S = \{(x, y) \mid x - y^2 = 0\},$$

$$\bar{u}(x, y) = \begin{cases} -1 & \text{für } x - y^2 < 0, \\ 1 & \text{für } x - y^2 > 0. \end{cases}$$

Die Funktion  $\bar{u}$  beschreibt das Verhalten von  $u_\varepsilon$  weg von Grenzschichten. Um die Funktion in der Grenzschicht (d.h. in der Nähe von  $S$ ) genauer betrachten zu können, verwenden wir eine Art mathematische Lupe in Form einer Transformation der unabhängigen Variablen.

Sei  $S$  eine  $C^1$ -Mannigfaltigkeit mit Dimension  $n - k$ . Dann führen wir lokale Koordinaten  $\mathbf{z} = \mathbf{z}(\mathbf{x})$  so ein, daß  $S = \{\mathbf{x} \mid z_1(\mathbf{x}) = \dots = z_k(\mathbf{x}) = 0\}$  gilt. Die Abbildung  $\mathbf{z}(\mathbf{x})$  ist in einer Umgebung  $U(S)$  von  $S$  definiert. *Lokale Koordinaten* (auch *lokale Variable*)  $\xi$  in  $U(S)$  sind dann gegeben durch

$$\begin{aligned} \xi_i &= z_i \varepsilon^{-\alpha_i} \quad \text{mit } \alpha_i > 0, \quad \text{für } i = 1, \dots, k; \\ \xi_i &= z_i, \quad \text{für } i = k + 1, \dots, n. \end{aligned}$$

Lokale Koordinaten erfüllen ihren Zweck dann, wenn sie *regularisierend* wirken, d.h. wenn es einen Bereich gibt, in dem die Funktion

$$U_\varepsilon(\xi) := u_\varepsilon(\mathbf{x}(\xi, \varepsilon))$$

regulär ist.

**Beispiel 1.4:** Für die Funktion  $u_\varepsilon(x) = e^{-x/\varepsilon} + x$  ist durch  $\xi = x\varepsilon^{-\alpha}$  eine lokale Variable in der Nähe von  $x = 0$  gegeben. Die transformierte Funktion

$$U_\varepsilon(\xi) = \exp(-\xi\varepsilon^{\alpha-1}) + \varepsilon^\alpha \xi$$

ist für  $\alpha \geq 1$  regulär auf  $\xi$ -Intervallen der Form  $(0, X)$  mit

$$\lim_{\varepsilon \rightarrow 0} U_\varepsilon(\xi) = \begin{cases} e^{-\xi} & \text{für } \alpha = 1, \\ 1 & \text{für } \alpha > 1, \end{cases}$$

und für  $0 < \alpha < 1$  regulär auf Intervallen der Form  $(X_1, X_2)$  mit  $X_1 > 0$  und  $\lim_{\varepsilon \rightarrow 0} U_\varepsilon(\xi) = 0$ .

**Beispiel 1.8:** Es muß nicht immer möglich sein, durch Einführung einer lokalen Variablen eine sinnvolle Regularisierung zu erreichen. Um das Verhalten der Funktion

$$u_\varepsilon(t) = e^{-\varepsilon t} \sin t$$

für  $t \rightarrow \infty$  zu studieren, setzen wir  $t = s^{-1}$  und betrachten

$$v_\varepsilon(s) = e^{-\varepsilon/s} \sin \frac{1}{s}$$

in der Nähe von  $s = 0$ . Offensichtlich hat  $v_\varepsilon$  Grenzschichtverhalten, und es gilt  $\bar{v}(s) = \sin(1/s)$ . Die lokale Variable  $\sigma = s\varepsilon^{-\alpha}$  führt auf

$$V_\varepsilon(\sigma) = \exp(-\varepsilon^{1-\alpha}/\sigma) \sin \frac{1}{\varepsilon^\alpha \sigma}.$$

Für  $0 < \alpha \leq 1$  gibt es kein  $\sigma$ -Intervall, in dem  $V_\varepsilon$  regulär ist. Für  $\alpha > 1$  erhalten wir bei  $\varepsilon \rightarrow 0$  die Näherung  $0 = v_\varepsilon(0)$ , die nicht sehr aussagekräftig ist.

Bei dem Beispiel  $u_\varepsilon(x) = e^{-x/\varepsilon} + x$  hat man das Gefühl, daß die lokale Variable  $\xi = x/\varepsilon$  vor anderen lokalen Variablen ausgezeichnet ist dadurch, daß die entsprechende Näherung  $e^{-\xi}$  die meiste Information enthält. Die anderen Näherungen 0 und 1 können mit  $\xi = 0$  und  $\xi \rightarrow \infty$  daraus berechnet werden. Diese Ideen wollen wir formalisieren.

**Definition 1.6.** a) Sei  $\xi$  eine lokale Variable. Dann definieren wir den lokalen Grenzwert von  $u_\varepsilon$  bezüglich  $\xi$  durch

$$\left( \lim_{\xi} u_\varepsilon \right) (\xi) := \lim_{\varepsilon \rightarrow 0} u_\varepsilon(\mathbf{x}(\xi, \varepsilon)).$$

Der Definitionsbereich von  $\lim_{\xi} u_\varepsilon$  ist die Vereinigung aller kompakten Mengen, in denen obiger Grenzübergang gleichmäßig ist.

b) Seien  $\xi_1$  und  $\xi_2$  lokale Variable und  $D_2$  und  $D_{12}$  die Definitionsbereiche der lokalen Grenzwerte bezüglich  $\xi_2$  von  $u_\varepsilon$  und  $\lim_{\xi_1} u_\varepsilon$ . Der lokale Grenzwert von  $u_\varepsilon$  bezüglich  $\xi_2$  ist enthalten in dem bezüglich  $\xi_1$ , wenn  $D_2 \subset D_{12}$  und

$$\lim_{\xi_2} \left( \lim_{\xi_1} u_\varepsilon \right) = \lim_{\xi_2} u_\varepsilon$$

gleichmäßig in kompakten Teilmengen von  $D_2$  gilt.

c) Ein lokaler Grenzwert heißt signifikant, wenn er in keinem anderen lokalen Grenzwert enthalten ist, d.h. wenn er bezüglich der in b) definierten Ordnungsrelation ein maximales Element ist. Die signifikanten Grenzwerten entsprechenden lokalen Variablen heißen Grenzschnittvariablen.

**Beispiel 1.9:** Es kann auch mehrere signifikante Grenzwerte einer Funktion geben. Sei für  $x \in [0, 1]$

$$u_\varepsilon(x) = x \left( \frac{e^{-x/\varepsilon}}{x + \varepsilon^2} + 1 \right).$$

Offensichtlich gilt  $\bar{u}(x) = x$ . Außerdem hat  $u_\varepsilon$  Grenzschnittverhalten, weil  $u_\varepsilon(\varepsilon^2) = 1/2$  gilt. Man zeigt leicht, daß  $\xi_1 = x/\varepsilon$  und  $\xi_2 = x/\varepsilon^2$  Grenzschnittvariable sind. Die entsprechenden Grenzwerte und Definitionsbereiche sind  $e^{-\xi_1}$  und  $(0, \infty)$  bzw.  $\xi_2/(1 + \xi_2)$  und  $[0, \infty)$ .

Zur Approximation von Funktionen, für die es genau eine Grenzschnittvariable  $\xi$  gibt, ist folgende Vorgangsweise naheliegend: Die Funktion

$$\tilde{u} = u_\varepsilon - \bar{u}$$

liefert nur in der Nähe von  $S$  einen signifikanten Beitrag. Wir verwenden das *heuristische Prinzip*, daß  $\tilde{u}$  durch seinen lokalen Grenzwert bezüglich  $\xi$  gleichmäßig in einer Umgebung von  $S$  approximiert werden kann:

$$\|\tilde{u} - \lim_{\xi} \tilde{u}\|_{U(S)} = o(1)$$

Gültigkeit des heuristischen Prinzips bedeutet also, daß

$$\bar{u} + \lim_{\xi} (u_\varepsilon - \bar{u})$$

eine asymptotische Näherung für  $u_\varepsilon$  in dem Banachraum  $C(\bar{U}(S))$  mit der Supremum-Norm ist.

Diese Idee ist auch im Fall der Existenz von mehreren Grenzschnittvariablen anwendbar:

**Beispiel 1.9:** Wir approximieren  $u_\varepsilon$  durch

$$\bar{u} + \lim_{\xi_1} (u_\varepsilon - \bar{u}) + \lim_{\xi_2} \left( u_\varepsilon - \bar{u} - \lim_{\xi_1} (u_\varepsilon - \bar{u}) \right).$$

Die Näherung besteht aus 3 Summanden, die von den unabhängigen Variablen  $x$ ,  $\xi_1$  und  $\xi_2$  abhängen. Eine einfache Rechnung zeigt

$$\left\| u_\varepsilon - x - e^{-x/\varepsilon} + \frac{1}{1 + x/\varepsilon^2} \right\|_{(0,1)} = O(\varepsilon).$$

Bessere Näherungen können gefunden werden, indem man den Fehler durch  $\varepsilon$  dividiert und dieselbe Vorgangsweise noch einmal anwendet. Fortsetzung dieser Idee liefert eine asymptotische Entwicklung für  $u_\varepsilon$  in der Form

$$u_\varepsilon(x) = \sum_{k=0}^N (\bar{u}_k(x) + \tilde{u}_k(\xi_1) + \hat{u}_k(\xi_2)) \varepsilon^k + O(\varepsilon^{N+1}).$$

Den Koeffizienten von  $\varepsilon^0$  haben wir schon berechnet:

$$\bar{u}_0(x) = \bar{u}(x) = x, \quad \tilde{u}_0(\xi_1) = e^{-\xi_1}, \quad \hat{u}_0(\xi_2) = -\frac{1}{1 + \xi_2}$$

Für den Koeffizienten von  $\varepsilon^1$  gilt:

$$\bar{u}_1(x) = 0, \quad \tilde{u}_1(\xi_1) = \frac{1 - e^{-\xi_1}}{\xi_1}, \quad \hat{u}_1(\xi_2) = -\frac{1}{1 + \xi_2}$$

Die Tatsache, daß die Grenzschichtterme  $\tilde{u}_k$  und  $\hat{u}_k$  nur in der Grenzschicht einen signifikanten Beitrag liefern, läßt sich in der Form

$$\lim_{\xi_1 \rightarrow \infty} \tilde{u}_k(\xi_1) = \lim_{\xi_2 \rightarrow \infty} \hat{u}_k(\xi_2) = 0$$

schreiben.

Wir wenden uns nun Problemen zu, deren Lösungen Grenzschichtverhalten zeigen.

**Beispiel 1.10:** Wir betrachten das Randwertproblem

$$\begin{aligned} -\varepsilon u'' + u' + u &= 0 & \text{für } x \in (0, 1), \\ u(0) &= 1, & u(1) = 0. \end{aligned} \tag{1.14}$$

Das Problem ist singular gestört, weil Lösungen der Differentialgleichung für  $\varepsilon = 0$

$$u' + u = 0$$

nicht beide Randbedingungen erfüllen können.

**Definition 1.7.** a) Ausgehend von einer Differentialgleichung, die einen kleinen Parameter  $\varepsilon$  enthält, nennt man die Gleichung, die man für  $\varepsilon = 0$  erhält, die *reduzierte Gleichung*.

b) Eine Differentialgleichung mit einem kleinen Parameter heißt *singular gestört*, wenn die entsprechende reduzierte Gleichung von einem anderen Typ ist.

Der "Typ" einer Differentialgleichung ist die Klasse aller Zusatzbedingungen, die in Verbindung mit der Differentialgleichung zu einem sachgemäß gestellten Problem führen. Der Typ von gewöhnlichen Differentialgleichungen ist bestimmt durch ihre Ordnung. Eine gewöhnliche Differentialgleichung ist daher singular gestört, wenn die reduzierte Gleichung geringere Ordnung hat.

Ein Problem, das aus einer singular gestörten Differentialgleichung und aus Anfangs- und/oder Randbedingungen besteht, ist im allgemeinen ein singular gestörtes Problem. Lösungen der reduzierten Gleichung werden

- 1) nicht in der Lage sein, alle Zusatzbedingungen zu erfüllen und/oder
- 2) nicht die Differenzierbarkeitseigenschaften haben, die von der exakten Lösung erwartet werden.



**Heuristische Annahme 1.8.** *Lösungen von Anfangs- und/oder Randwertproblemen für singular gestörte Differentialgleichungen haben Grenzschichten an Mannigfaltigkeiten, auf denen Zusatzbedingungen gegeben sind oder an denen ein Glattheitsverlust auftritt.*

Eine Konsequenz aus dieser Annahme ist, daß die oben definierte Funktion  $\bar{u}$ , die die Lösung weg von Grenzschichten approximiert, als Lösung der reduzierten Gleichung gewählt werden kann.

**Beispiel 1.10:** Die allgemeine Lösung der reduzierten Gleichung ist  $\bar{u} = Ae^{-x}$ . Im Moment ist noch unklar, wie die Integrationskonstante  $A$  zu wählen ist. Grenzschichten können der heuristischen Annahme zufolge bei  $x = 0$  und/oder  $x = 1$  auftreten.

Um Näherungen für die Lösung in einer Grenzschicht bei  $x = 0$  zu berechnen, transformieren wir auf die lokale Variable  $\xi = x\varepsilon^{-\alpha}$ . Die Differentialgleichung hat dann die Form

$$-\varepsilon^{1-2\alpha}\ddot{u} + \varepsilon^{-\alpha}\dot{u} + u = 0, \quad \left(\dot{u} = \frac{du}{d\xi}\right). \quad (1.15)$$

Um zu garantieren, daß die linke Seite für  $\varepsilon \rightarrow 0$  einen Grenzwert besitzt, müssen wir in Abhängigkeit von  $\alpha$  mit Potenzen von  $\varepsilon$  multiplizieren. Wir erhalten für  $\varepsilon \rightarrow 0$  die Differentialgleichungen

$$\begin{aligned} \dot{u} &= 0 && \text{für } 0 < \alpha < 1, \\ -\ddot{u} + \dot{u} &= 0 && \text{für } \alpha = 1, \\ -\ddot{u} &= 0 && \text{für } \alpha > 1. \end{aligned}$$

**Definition 1.9.** a) Sei  $L_\varepsilon$  ein linearer Differentialoperator, der bei Transformation auf die lokale Variable  $\xi$  in den Operator  $\mathcal{L}_\varepsilon$  übergeht. Sei  $\gamma \in \mathbb{R}$  so gewählt, daß

$$\varepsilon^\gamma \mathcal{L}_\varepsilon(u) = O_s(1).$$

Dann ist die lokale Degeneration von  $L_\varepsilon$  bezüglich  $\xi$  definiert durch

$$\lim_{\xi} L_\varepsilon(u) := \lim_{\varepsilon \rightarrow 0} \varepsilon^\gamma \mathcal{L}_\varepsilon(u).$$

b) Die lokale Degeneration von  $L_\varepsilon$  bezüglich  $\xi_2$  ist enthalten in der bezüglich  $\xi_1$ , wenn

$$\lim_{\xi_2} \left( \lim_{\xi_1} L_\varepsilon(u) \right) = \lim_{\xi_2} L_\varepsilon(u)$$

gilt. Eine signifikante Degeneration ist eine lokale Degeneration, die in keiner anderen enthalten ist.

**Korrespondenzprinzip 1.10.** *Ist  $u_\varepsilon$  Lösung einer Differentialgleichung, dann nehmen wir an, daß jeder signifikante Grenzwert von  $u_\varepsilon$  Lösung der entsprechenden signifikanten Degeneration der Differentialgleichung ist.*

Beim Auffinden von signifikanten Degenerationen ist die folgende **heuristische Regel** brauchbar: *Signifikante Degenerationen enthalten möglichst viele, mindestens jedoch zwei Terme.*

Daraus ergibt sich folgende Vorgangsweise: Kandidaten für signifikante Degenerationen erhält man durch Gleichsetzen der Exponenten von  $\varepsilon$  in verschiedenen Termen der auf lokale Variable transformierten Differentialgleichung unter der Nebenbedingung, daß die übrigen Exponenten größer oder gleich den verglichenen sind.

**Beispiel 1.10:** Obige Methode angewandt auf die transformierte Gleichung (1.15) ergibt die Möglichkeiten

$$\begin{aligned} 1 - 2\alpha &= -\alpha \leq 0, \\ 1 - 2\alpha &= 0 \leq -\alpha \quad \text{oder} \\ -\alpha &= 0 \leq 1 - 2\alpha. \end{aligned}$$

Nur die erste Zeile hat eine positive Lösung, nämlich  $\alpha = 1$ . Damit folgt aus dem Korrespondenzprinzip, daß eine Grenzschichtvariable in der Nähe von  $x = 0$  durch  $\xi = x/\varepsilon$  gegeben ist. Ähnlich erhält man für eine Grenzschicht bei  $x = 1$  die Grenzschichtvariable  $\eta = (1 - x)/\varepsilon$ . Aus den bisher besprochenen heuristischen Prinzipien ergibt sich ein Ansatz der Form

$$u(x, \varepsilon) = \bar{u}(x) + \tilde{u}(\xi) + \hat{u}(\eta) + O(\varepsilon)$$

mit  $\lim_{\xi \rightarrow \infty} \tilde{u}(\xi) = \lim_{\eta \rightarrow \infty} \hat{u}(\eta) = 0$ . Die Gleichungen für  $\bar{u}$ ,  $\tilde{u}$  und  $\hat{u}$  ermittelt man, indem man den Ansatz in die ursprüngliche Differentialgleichung einsetzt, die entstehende Gleichung in Abhängigkeit der unabhängigen Variablen  $x$ ,  $\xi$  bzw.  $\eta$  schreibt und in jeder dieser Gleichungen  $\varepsilon$  gegen Null gehen läßt. Das Resultat sind die Gleichungen

$$\begin{aligned} \bar{u}_x + \bar{u} &= 0, \\ -\tilde{u}_{\xi\xi} + \tilde{u}_\xi &= 0, \\ -\hat{u}_{\eta\eta} - \hat{u}_\eta &= 0. \end{aligned}$$

Die *Grenzschichtgleichungen* für  $\tilde{u}$  und  $\hat{u}$  unterscheiden sich durch ein Vorzeichen ganz wesentlich. Berücksichtigt man die Abklingbedingung, dann kommt für  $\tilde{u}$  nur die triviale Lösung  $\tilde{u} = 0$  in Frage, während im Fall von  $\hat{u}$  die eindimensionale Lösungsschar  $\hat{u}(\eta) = Be^{-\eta}$  verwendet werden kann. Unsere formale Vorgangsweise impliziert, daß  $u(x, \varepsilon)$  eventuell eine Grenzschicht an  $x = 1$ , aber keine an  $x = 0$  hat. Daraus folgt, daß  $\bar{u}$  die Randbedingung an  $x = 0$

$$\bar{u}(0) = 1$$

erfüllen muß. Das gibt  $\bar{u}(x) = e^{-x}$ . Die Randbedingung an  $x = 1$  liefert

$$\bar{u}(1) + \hat{u}(0) = 0,$$

woraus  $B = -e^{-1}$  folgt. Damit haben wir die formale Näherung

$$u_{as}(x, \varepsilon) = e^{-x} - e^{-1+(x-1)/\varepsilon}$$

für die Lösung von (1.14) berechnet. Da die exakte Lösung des Problems explizit angegeben werden kann, ist die Gültigkeit der Näherung leicht zu überprüfen.

## 1.7. Gültigkeit formaler asymptotischer Näherungen

Bisher haben wir uns mit der Herleitung formaler Näherungen für die Lösungen von Problemen mit kleinen Parametern beschäftigt. Formale Näherungen sind durch Kleinheit des Residuums ausgezeichnet, das sich ergibt, wenn man die Näherung in die das Problem bestimmenden Gleichungen einsetzt. Daß diese Bedingung im allgemeinen nicht hinreichend ist, zeigt das folgende Beispiel.

**Beispiel 1.11:** In der ersten Gleichung des Systems

$$0,01x + y = 0,1, \quad x + 101y = 11 \quad (1.16)$$

ist der Koeffizient von  $x$  klein im Vergleich zum Koeffizienten von  $y$ . Vernachlässigung des Termes  $0,01x$  führt auf die Näherungslösung

$$x_0 = 0,9, \quad y_0 = 0,1. \quad (1.17)$$

Das Residuum ist der Vektor  $\mathbf{r} = (0,009, 0)^{tr}$ . Trotz der Kleinheit des Residuums ist die exakte Lösung

$$x = -90, \quad y = 1$$

weit von der "Näherung" (1.17) entfernt. Was sind die Gründe für diesen Fehlschlag? Bezeichnen wir die Koeffizientenmatrix in (1.16) mit  $A$  und führen wir die Abkürzungen  $\mathbf{x} = (x, y)^{tr}$  und  $\mathbf{x}_0 = (x_0, y_0)^{tr}$  ein, dann gilt für den Approximationsfehler

$$\mathbf{x}_0 - \mathbf{x} = A^{-1}\mathbf{r}.$$

Die Inverse von  $A$  ist

$$A^{-1} = \begin{pmatrix} 10100 & -100 \\ -100 & 1 \end{pmatrix}$$

was den großen Approximationsfehler erklärt. Die Matrix  $A$  ist nicht weit von einer singulären Matrix entfernt.

Zu untersuchen ist also, ob das Problem *gut konditioniert* ist, d.h. ob es nicht zu sensitiv auf Änderungen in den Daten reagiert. Ähnliches gilt auch für nichtlineare Probleme. Die Sensitivität wird dann anhand des Verhaltens der Inversen von linearisierten Problemen gemessen.

Im folgenden bezeichnet  $(B_1, \|\cdot\|_1)$  einen Banachraum und  $(B_2, \|\cdot\|_2)$  einen normierten Raum. Das folgende Resultat gibt hinreichende Bedingungen dafür an, daß aus Konsistenz Konvergenz folgt.

**Satz 1.11.** Die Abbildung  $F : B_1 \rightarrow B_2$  sei an der Stelle  $u_{as} \in B_1$  Frechet-differenzierbar. Die Frechet-Ableitung  $DF(u_{as})$  sei invertierbar, und es gelte

$$\|DF(u_{as})^{-1}f\|_1 \leq K\|f\|_2 \quad \forall f \in B_2. \quad (1.18)$$

Für  $P(v) := F(u_{as} + v) - F(u_{as}) - DF(u_{as})v$  gelte

$$\|P(v_1) - P(v_2)\|_2 \leq L\delta\|v_1 - v_2\|_1 \quad \text{für } \|v_1\|_1, \|v_2\|_1 \leq \delta. \quad (1.19)$$

Gilt außerdem für das Residuum  $\rho := F(u_{as})$

$$\|\rho\|_2 \leq \frac{1}{4K^2L},$$

dann gibt es eine Lösung der Gleichung  $F(u) = 0$ , die eindeutig ist in der Kugel  $K_{\bar{\delta}}(u_{as}) \subset B_1$  mit Mittelpunkt  $u_{as}$  und Radius  $\bar{\delta} = \frac{1}{2KL}$ , und es gilt

$$\|u - u_{as}\|_1 \leq 2K\|\rho\|_2. \quad (1.20)$$

*Beweis.* Der Beweis ist eine Anwendung des Banachschen Fixpunktsatzes. Wir setzen  $R = u - u_{as}$ . Dann ist das Problem  $F(u) = 0$  äquivalent zu

$$F(u_{as} + R) - F(u_{as}) = -\rho.$$

Das wiederum gilt genau dann, wenn

$$DF(u_{as})R = -\rho - P(R)$$

gilt. Wir haben also die Lösbarkeit des Fixpunktproblems

$$R = G(R) := -DF(u_{as})^{-1}(\rho + P(R))$$

zu zeigen. Dazu zeigen wir, daß  $G(R)$  in  $K_{\bar{\delta}}(0)$  eine Kontraktion ist. Die Abschätzung

$$\|G(R)\|_1 \leq K\|\rho + P(R)\|_2 \leq K\left(\frac{1}{4K^2L} + L\bar{\delta}^2\right) = \bar{\delta}$$

beweist  $G : K_{\bar{\delta}}(0) \rightarrow K_{\bar{\delta}}(0)$ . Die Kontraktivität folgt aus

$$\begin{aligned} \|G(R_1) - G(R_2)\|_1 &\leq K\|P(R_1) - P(R_2)\|_2 \\ &\leq KL\bar{\delta}\|R_1 - R_2\|_1 = \frac{1}{2}\|R_1 - R_2\|_1. \end{aligned}$$

Damit ist die Existenz einer in  $K_{\bar{\delta}}(0)$  eindeutigen Lösung  $R$  bewiesen. Für diese gilt

$$\begin{aligned} \|R\|_1 &= \|G(R)\|_1 \leq K(\|\rho\|_2 + \|P(R)\|_2) \\ &\leq K(\|\rho\|_2 + L\bar{\delta}\|R\|_1) = K\|\rho\|_2 + \frac{1}{2}\|R\|_1, \end{aligned}$$

woraus die Abschätzung (1.20) folgt. ■

Zu den Voraussetzungen im Satz wäre zu sagen, daß (1.19) eine Aussage über die Glattheit von  $F$  ist, die gleichbedeutend mit Lipschitzstetigkeit der Frechet-Ableitung ist. Die zentrale Voraussetzung ist (1.18). Sie bedeutet, daß das an der formalen Näherung linearisierte Problem für beliebige Inhomogenitäten eindeutig lösbar sein muß und die Lösung stetig von den Daten abhängen muß. Diese Eigenschaft wird als Stabilität bezeichnet. Analog zur Analyse von gewissen numerischen Verfahren kann auch hier die Merkregel

$$\textit{Konsistenz} + \textit{Stabilität} = \textit{Konvergenz}$$

verwendet werden.

Bei Anwendbarkeit von Satz 1.11 kann das entsprechende Problem als *sachgemäß gestellt* bezeichnet werden: Es gibt eine lokal (in  $K_{\bar{\delta}}(u_{as})$ ) eindeutige Lösung, die stetig von Störungen ( $\rho$ ) abhängt.

Um den Satz auf regulär gestörte Probleme anwenden zu können, benötigen wir das folgende Resultat.

**Lemma 1.12.** *Seien  $A$  und  $B$  lineare Abbildungen von  $B_1$  nach  $B_2$ . Weiters besitze  $A$  eine beschränkte Inverse mit*

$$\|A^{-1}f\|_1 \leq K\|f\|_2 \quad \forall f \in B_2,$$

und für  $A^{-1}B$  gelte

$$\|A^{-1}Bu\|_1 \leq \delta\|u\|_1 \quad \forall u \in B_1, \quad \delta < 1.$$

Dann besitzt die Abbildung  $A + B$  eine beschränkte Inverse mit

$$\|(A + B)^{-1}f\|_1 \leq \frac{K}{1 - \delta}\|f\|_2 \quad \forall f \in B_2.$$

*Beweis.* Wir führen die Abbildungsnorm

$$\|A\|_{1 \rightarrow 2} := \sup_{\|u\|_1=1} \|Au\|_2$$

und analog  $\|\cdot\|_{2 \rightarrow 1}$  und  $\|\cdot\|_{1 \rightarrow 1}$  ein. Damit können die Voraussetzungen des Lemmas geschrieben werden als  $\|A^{-1}\|_{2 \rightarrow 1} \leq K$  und  $\|A^{-1}B\|_{1 \rightarrow 1} \leq \delta$ . Die Gleichung  $(A+B)u = f$  ist äquivalent zu

$$u = A^{-1}f - A^{-1}Bu.$$

Da die rechte Seite wegen der Voraussetzungen eine kontrahierende Abbildung ist, folgt die Existenz einer eindeutigen Lösung und damit die Invertierbarkeit von  $A+B$  aus dem Banachschen Fixpunktsatz.

Für die Abschätzung der Norm der Inversen verwenden wir

$$\|u\|_1 \leq K\|f\|_2 + \delta\|u\|_1 \implies \|(A+B)^{-1}f\|_1 = \|u\|_1 \leq \frac{K}{1-\delta}\|f\|_2,$$

womit das Lemma bewiesen ist. ■

Der folgende Satz ist ein Resultat über die Gültigkeit von formalen asymptotischen Entwicklungen für regulär gestörte Probleme der Form

$$F(u, \varepsilon) = 0. \tag{1.22}$$

**Satz 1.13.** Sei  $F : B_1 \times [0, \varepsilon_0] \rightarrow B_2$  eine Abbildung. Es existiere eine Lösung  $u_0 \in B_1$  des reduzierten Problems  $F(u_0, 0) = 0$ , sodaß  $F$  in einer Umgebung von  $(u_0, 0)$   $(N+1)$ -mal ( $N \geq 1$ ) Frechet-differenzierbar (nach beiden Argumenten) ist und die Frechet-Ableitung  $DF(u_0, 0)$  von  $F$  nach  $u$  an der Stelle  $(u_0, 0)$  eine beschränkte Inverse besitzt. Dann ist eine formale asymptotische Entwicklung der Form

$$U_n = \sum_{k=0}^n u_k \varepsilon^k \quad \text{für } n \leq N$$

für eine Lösung von (1.22) eindeutig bestimmt. Weiters besitzt das Problem (1.22) für  $\varepsilon$  klein genug eine in einer von  $\varepsilon$  unabhängigen Umgebung von  $u_0$  eindeutige Lösung  $u$ , für die

$$\|u - U_n\|_1 = O(\varepsilon^{n+1}) \quad \text{für } n \leq N$$

gilt.

*Beweis.* Die Berechenbarkeit der formalen asymptotischen Entwicklung wurde schon in Abschnitt 1.4 gezeigt. Es gilt also

$$\|F(U_n, \varepsilon)\|_2 = O(\varepsilon^{n+1}).$$

Wegen der Glattheitsannahmen an  $F$  ist die Voraussetzung (1.19) von Satz 1.11 mit einer von  $\varepsilon$  unabhängigen Konstanten  $L$  erfüllt. Außerdem benötigen wir die Invertierbarkeit der linearen Abbildung  $DF(U_n, \varepsilon)$ . Wegen der Differenzierbarkeitsannahmen für  $F$  gilt

$$DF(U_n, \varepsilon) = DF(u_0, 0) + \varepsilon C$$

mit einer linearen Abbildung  $C$ , die gleichmäßig in  $\varepsilon$  beschränkt ist. Für  $\varepsilon$  klein genug können wir daher Lemma 1.12 mit  $A = DF(u_0, 0)$  und  $B = \varepsilon C$  anwenden. Es folgt, daß die Inverse von  $DF(U_n, \varepsilon)$  existiert und gleichmäßig in  $\varepsilon$  beschränkt ist. Es ist also auch Voraussetzung (1.18) erfüllt. Anwendung von Satz 1.11 vervollständigt den Beweis. ■

**Beispiel 1.1:** Wir wollen Satz 1.13 auf das Problem

$$F(x, \varepsilon) := \begin{pmatrix} x'' + \frac{1}{(\varepsilon x + 1)^2} \\ x(0) \\ x'(0) - 1 \end{pmatrix} = 0$$

anwenden. Wir wählen  $B_1 = C^2([0, 2])$  und  $B_2 = C([0, 2]) \times \mathbb{R}^2$  mit den Normen

$$\|x\|_1 = \max_{0 \leq t \leq 2} (|x(t)| + |x'(t)| + |x''(t)|)$$

und

$$\|(f, a, b)\|_2 = \max_{0 \leq t \leq 2} |f(t)| + |a| + |b|.$$

Dabei wurde das Zeitintervall  $[0, 2]$  deswegen gewählt, weil  $t = 2$  der von der Näherung  $x_0(t) = t - t^2/2$  prognostizierte Aufschlagszeitpunkt ist. Die Differenzierbarkeitsforderungen des Satzes für beliebige  $N \in \mathbb{N}$  sind leicht zu überprüfen. Das linearisierte Problem

$$DF(x_0, 0)y = \begin{pmatrix} y'' \\ y(0) \\ y'(0) \end{pmatrix} = \begin{pmatrix} f \\ a \\ b \end{pmatrix}$$

hat die eindeutige Lösung

$$y(t) = a + bt + \int_0^t (t - \tau)f(\tau) d\tau.$$

Es ist leicht, die Abschätzung

$$\|y\|_1 \leq 7\|(f, a, b)\|_2$$

zu zeigen. Satz 1.13 garantiert die Existenz einer Lösung des vollen Problems, die durch die in Abschnitt 1.4 konstruierte asymptotische Entwicklung approximiert wird.

## 2.KAPITEL Die Grundgleichungen der Kontinuumsmechanik

### 2.1. Einige Resultate aus der Teilchenmechanik

Wir beschäftigen uns mit einem System von  $N$  Massenpunkten, deren Wechselwirkungen durch Zentralkräfte bestimmt sind, und auf die zusätzlich äußere Kräfte wirken. Wir bezeichnen den Ort zum Zeitpunkt  $t$  bzw. die Masse des  $i$ -ten Teilchens mit  $\mathbf{r}_i(t) \in \mathbb{R}^3$  bzw.  $m_i$ ,  $i = 1, \dots, N$ . Aus dem zweiten Newtonschen Gesetz folgt für die Impulse  $\mathbf{p}_i = m_i \dot{\mathbf{r}}_i$  ( $\dot{\mathbf{r}} = d\mathbf{r}/dt$ )

$$\dot{\mathbf{p}}_i = \sum_{j \neq i} \mathbf{f}_{ij} + \mathbf{f}_i, \quad i = 1, \dots, N, \quad (2.1)$$

wobei  $\mathbf{f}_{ij}$  die vom  $j$ -ten auf das  $i$ -te Teilchen ausgeübte Kraft und  $\mathbf{f}_i$  die auf das  $i$ -te Teilchen wirkende äußere Kraft bezeichnen. Aus dem dritten Newtonschen Gesetz folgt die Forderung  $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$ , die von den hier angenommenen Zentralkräften der Form

$$\mathbf{f}_{ij}(\mathbf{r}_i, \mathbf{r}_j) = \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|} g_{ij}(|\mathbf{r}_i - \mathbf{r}_j|)$$

mit  $g_{ij}(r) = g_{ji}(r)$  erfüllt ist. Für  $g_{ij} > 0$  ist die Wechselwirkung abstoßend, für  $g_{ij} < 0$  anziehend. Die äußeren Kräfte  $\mathbf{f}_i(t)$  werden als gegeben angenommen. Zwei Beispiele für  $g_{ij}$  sind

1. Gravitation:

$$g_{ij}(r) = -G \frac{m_i m_j}{r^2}$$

mit der Gravitationskonstanten  $G$ ,

2. Elektrische Ladungen:

$$g_{ij}(r) = k \frac{Q_i Q_j}{r^2}$$

mit der Ladung  $Q_i$  des  $i$ -ten Teilchens und der Proportionalitätskonstanten  $k$ . Da es sowohl positive als auch negative elektrische Ladungen gibt, kann diese Kraft sowohl abstoßend als auch anziehend sein.

Führt man den Gesamtimpuls  $\mathbf{p} = \sum_i \mathbf{p}_i$  und die Summe der äußeren Kräfte  $\mathbf{f} = \sum_i \mathbf{f}_i$  ein, so erhält man

$$\dot{\mathbf{p}} = \sum_i \sum_{j \neq i} \mathbf{f}_{ij} + \mathbf{f} = \sum_i \left( \sum_{j < i} \mathbf{f}_{ij} + \sum_{j < i} \mathbf{f}_{ji} \right) + \mathbf{f} = \mathbf{f} \quad (2.2)$$

wegen des dritten Newtonschen Gesetzes. Diese Eigenschaft nennen wir *Impulserhaltung*. Sie bedeutet, daß die Wechselwirkungskräfte zur Änderung der Gesamtimpulses nichts beitragen. Wir können also das gesamte Teilchenensemble selbst wieder als Massenpunkt mit Masse  $m = \sum_i m_i$ , Ort  $\mathbf{r} = (1/m) \sum_i m_i \mathbf{r}_i$  und Impuls  $\mathbf{p}$  betrachten, dessen Bewegung durch die Summe  $\mathbf{f}$  der äußeren Kräfte bestimmt wird. Treten keine äußeren Kräfte auf, dann vollführt der Schwerpunkt  $\mathbf{r}$  eine Trägheitsbewegung mit konstanter Geschwindigkeit auf einer geradlinigen Bahn.

Ein weiterer Erhaltungssatz gilt für den Drehimpuls. Wir führen die Drehimpulse  $\mathbf{L}_i = (\mathbf{r}_i - \mathbf{r}_0) \times \mathbf{p}_i$  und die Drehmomente  $\mathbf{M}_i = (\mathbf{r}_i - \mathbf{r}_0) \times \mathbf{f}_i$ ,  $i = 1, \dots, N$  um den ruhenden Punkt  $\mathbf{r}_0$  ein, wobei mit  $\mathbf{x} \times \mathbf{y}$  das Vektorprodukt bezeichnet wird. Außerdem definieren wir den Gesamtdrehimpuls  $\mathbf{L} = \sum_i \mathbf{L}_i$  und die Summe der Drehmomente  $\mathbf{M} = \sum_i \mathbf{M}_i$ . Eine leichte Rechnung zeigt  $\dot{\mathbf{L}}_i = (\mathbf{r}_i - \mathbf{r}_0) \times \dot{\mathbf{p}}_i$ . Daher gilt

$$\begin{aligned} \dot{\mathbf{L}} &= \sum_i (\mathbf{r}_i - \mathbf{r}_0) \times \dot{\mathbf{p}}_i = \sum_i \sum_{j \neq i} \mathbf{r}_i \times \mathbf{f}_{ij} - \mathbf{r}_0 \times \sum_i \sum_{j \neq i} \mathbf{f}_{ij} + \mathbf{M} \\ &= - \sum_i \sum_{j \neq i} \frac{\mathbf{r}_i \times \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|} g_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) + \mathbf{M} = \mathbf{M} \end{aligned} \quad (2.3)$$

wegen der Antisymmetrie des Vektorproduktes ( $\mathbf{r}_i \times \mathbf{r}_j = -\mathbf{r}_j \times \mathbf{r}_i$ ). Die Eigenschaft (2.3) nennen wir *Drehimpulserhaltung*. Ähnlich zur Impulserhaltung bedeutet sie, daß die Wechselwirkungskräfte nichts zur

Änderung des Gesamtdrehimpulses um einen beliebigen ruhenden Punkt  $\mathbf{r}_0$  beitragen. Auf den Gesamtdrehimpuls wirkt sich nur die Summe der durch die äußeren Kräfte verursachten Drehmomente aus.

Schließlich beschäftigen wir uns noch mit der *Energieerhaltung*. Wir definieren die *kinetische Energie*

$$T = \sum_i m_i \frac{|\dot{\mathbf{r}}_i|^2}{2}$$

und die von den äußeren Kräften pro Zeiteinheit verrichtete Arbeit

$$W = \sum_i \dot{\mathbf{r}}_i \cdot \mathbf{f}_i,$$

wobei  $\mathbf{x} \cdot \mathbf{y}$  das Skalarprodukt von zwei Vektoren im  $\mathbb{R}^3$  ist. Dann gilt

$$\begin{aligned} \dot{T} &= \sum_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{p}}_i = \sum_i \sum_{j \neq i} \dot{\mathbf{r}}_i \cdot \mathbf{f}_{ij} + W \\ &= \sum_i \sum_{j < i} \frac{(\dot{\mathbf{r}}_i - \dot{\mathbf{r}}_j) \cdot (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} g_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) + W \\ &= \sum_i \sum_{j < i} g_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) \frac{d}{dt} |\mathbf{r}_i - \mathbf{r}_j| + W. \end{aligned}$$

Sei für  $j < i$   $G_{ij}(r)$  eine Stammfunktion von  $g_{ij}(r)$ . Dann definieren wir die *potentielle Energie* durch

$$V = - \sum_i \sum_{j < i} G_{ij}(|\mathbf{r}_i - \mathbf{r}_j|).$$

Mit der Gesamtenergie  $E = T + V$  wird obige Gleichung zu

$$\dot{E} = W. \quad (2.4)$$

Ähnlich zu den bisher besprochenen Erhaltungssätzen gilt auch hier, daß die von den Wechselwirkungskräften verrichtete Arbeit nichts zur Veränderung der Gesamtenergie beiträgt.

Die Gleichungen (2.1) beschreiben ein System von Massenpunkten, die in einer dauernden Wechselwirkung miteinander stehen. Ein Extremfall dieser Situation ist ein System von harten Kugeln, deren Wechselwirkung sich auf diskrete Kollisionen beschränkt. Dieser Fall dient als Grundlage der klassischen Modelle für die Dynamik von verdünnten Gasen. Der Einfachheit halber beschränken wir uns auf den Fall lauter gleicher Kugeln mit Masse  $m$  und Radius  $R$ , die sich im  $\mathbb{R}^3$  bewegen. Sei  $\mathbf{r}_i(t)$  der Ort des Mittelpunktes der  $i$ -ten Kugel,  $i = 1, \dots, N$ . Solange keine Kollisionen stattfinden, wird die Bewegung der Kugeln nur durch die Wirkung äußerer Kräfte beeinflußt. Eine Kollision zwischen zwei Kugeln zum Zeitpunkt  $t_0$  findet statt, wenn für ein Paar  $(i, j)$

$$|\mathbf{r}_i(t_0) - \mathbf{r}_j(t_0)| = 2R$$

gilt. Die Kollision bewirkt eine augenblickliche Geschwindigkeitsänderung der beiden Kugelmittelpunkte. Die Annahme, daß Zentralkräfte zwischen den Kugeln wirken, bedeutet, daß die Geschwindigkeitsänderung in Richtung des Vektors

$$\mathbf{n} = \frac{\mathbf{r}_i(t_0) - \mathbf{r}_j(t_0)}{2R}$$

zwischen den Kugelmittelpunkten erfolgt. Bezeichnen wir mit  $\mathbf{v}'_i = \dot{\mathbf{r}}_i(t_0^-)$  bzw.  $\mathbf{v}_i = \dot{\mathbf{r}}_i(t_0^+)$  die Geschwindigkeiten vor bzw. nach der Kollision, dann gilt also

$$\begin{aligned} \mathbf{v}_i &= \mathbf{v}'_i + \lambda \mathbf{n}, \\ \mathbf{v}_j &= \mathbf{v}'_j + \mu \mathbf{n}. \end{aligned}$$



Die Konstanten  $\lambda$  und  $\mu$  werden durch die Forderungen der Impulserhaltung

$$\mathbf{v}_i + \mathbf{v}_j = \mathbf{v}'_i + \mathbf{v}'_j$$

und der Energieerhaltung

$$|\mathbf{v}_i|^2 + |\mathbf{v}_j|^2 = |\mathbf{v}'_i|^2 + |\mathbf{v}'_j|^2$$

bestimmt. Man erhält

$$\lambda = -\mu = \mathbf{n} \cdot (\mathbf{v}'_j - \mathbf{v}'_i).$$

Es ist leicht zu überprüfen, daß damit auch die Summe der Drehimpulse erhalten wird:

$$\mathbf{r}_i \times \mathbf{v}_i + \mathbf{r}_j \times \mathbf{v}_j = \mathbf{r}_i \times \mathbf{v}'_i + \mathbf{r}_j \times \mathbf{v}'_j$$

Auch für dieses "3-D Billiard" gelten also die Gesetze (2.2), (2.3) und (2.4).

## 2.2. Das kontinuierliche Medium

Für Punkte  $\mathbf{x}_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$  und  $h > 0$  bezeichnen wir den Würfel mit Mittelpunkt  $\mathbf{x}_0$  und Kantenlänge  $h$  mit

$$W_h(\mathbf{x}_0) = \{(x, y, z) \in \mathbb{R}^3 \mid \max(|x - x_0|, |y - y_0|, |z - z_0|) \leq h/2\}.$$

Betrachten wir ein großes ( $N \gg 1$ ) Ensemble von Massenpunkten (z.B. Moleküle) wie im vorigen Abschnitt, dann läßt sich eine Massendichte  $\rho_h$  definieren durch

$$\rho_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i. \quad (2.5)$$

Diese Dichte soll die Eigenschaft haben, daß das Integral

$$\int_R \rho_h(\mathbf{x}, t) d\mathbf{x} \quad (2.6)$$

als Näherung für die in  $R \subset \mathbb{R}^3$  enthaltene Masse verwendet werden kann. Dazu ist es klarerweise notwendig,  $h$  klein im Vergleich zu charakteristischen Längen in  $R$  zu wählen. Ist andererseits  $h$  so klein gewählt, daß sich in einem Würfel  $W_h(\mathbf{x}_0)$  nur wenige Moleküle befinden, dann ist die Dichtefunktion nicht sehr aussagekräftig, weil sie sehr stark von  $h$  abhängt. Die Grundannahme der Kontinuumsmechanik ist, daß es einen Bereich von Werten für  $h$  gibt, in dem  $\rho_h$  nur schwach von  $h$  abhängt. Im weiteren werden wir annehmen, daß  $h$  aus diesem Bereich gewählt wurde und daß das Integral (2.6) eine exakte Formel für die in  $R$  enthaltene Masse ist. Ähnlich zu (2.5) läßt sich auch eine Impulsdichte

$$\mathbf{p}_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \dot{\mathbf{r}}_i(t), \quad (2.7)$$

und damit eine mittlere Geschwindigkeit  $\mathbf{v}_h = \mathbf{p}_h / \rho_h$  definieren. Weiters ist die Energiedichte  $E_h = T_h + V_h$  als Summe aus der Dichte der kinetischen Energie

$$T_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \frac{|\dot{\mathbf{r}}_i(t)|^2}{2}$$

und der ähnlich definierten Dichte der potentiellen Energie gegeben. Definieren wir einen Energieanteil

$$\tilde{T}_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \frac{|\dot{\mathbf{r}}_i(t) - \mathbf{v}_h|^2}{2},$$

der durch die Abweichungen der Teilchengeschwindigkeiten von der mittleren Geschwindigkeit verursacht wird, dann gilt

$$T_h = \rho_h \frac{|\mathbf{v}_h|^2}{2} + \tilde{T}_h.$$

Schließlich führen wir die mikroskopische oder innere Energie pro Masseneinheit  $e_h$  durch die Gleichung  $\rho_h e_h = \tilde{T}_h + V_h$  ein. Dann kann die Energiedichte als Summe aus der Dichte der makroskopischen kinetischen Energie und der inneren Energiedichte geschrieben werden:

$$E_h = \rho_h \left( \frac{|\mathbf{v}_h|^2}{2} + e_h \right)$$

Es sei hier bemerkt, daß die Kontinuumsmechanik nicht unbedingt auf Mittelungen wie in (2.5) und (2.7) beruht. Es ist nicht weniger gerechtfertigt, die Materie als kontinuierlich im Raum verteilt anzunehmen und die Existenz einer Massendichte  $\rho(\mathbf{x}, t)$  und einer mittleren Geschwindigkeit  $\mathbf{v}(\mathbf{x}, t)$  zu postulieren, ohne Überlegungen über die mikroskopische Struktur des betrachteten Materials anzustellen.

Zunächst wollen wir uns mit der *Kinematik* von kontinuierlichen Medien beschäftigen. Ist die mittlere Geschwindigkeit  $\mathbf{v}(\mathbf{x}, t)$  vorgegeben, dann wird die Bewegung eines Materieteilchens beschrieben durch die Lösung des Anfangswertproblems

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, t), \quad \mathbf{x}(0) = \mathbf{A}, \quad (2.8)$$

wobei  $\mathbf{A}$  die Position des Teilchens zum Zeitpunkt  $t = 0$  ist. Wir berücksichtigen die Abhängigkeit der Teilchenbahn von der Anfangsposition durch die Schreibweise

$$\mathbf{x} = \mathbf{x}(\mathbf{A}, t). \quad (2.9)$$

Es gilt also:  $\mathbf{x}(\mathbf{A}, T)$  ist die Position zum Zeitpunkt  $t$  des Teilchens, das zum Zeitpunkt 0 die Position  $\mathbf{A}$  hatte. Wir postulieren die Undurchdringlichkeit des Materials. Das bedeutet, daß die Gleichung (2.9) eindeutig nach  $\mathbf{A}$  aufgelöst werden kann und äquivalent zu einer Gleichung der Form

$$\mathbf{A} = \mathbf{A}(\mathbf{x}, t)$$

ist. Daraus folgt, daß jede Funktion von  $(\mathbf{x}, t)$  auch als Funktion von  $(\mathbf{A}, t)$  geschrieben werden kann. Insbesondere führen wir für die Massendichte zum Zeitpunkt  $t$  des Teilchens, das zum Zeitpunkt 0 die Position  $\mathbf{A}$  hatte, die Bezeichnung

$$\delta(\mathbf{A}, t) = \rho(\mathbf{x}(\mathbf{A}, t), t)$$

ein. Je nach dem, ob man die *Euler-Koordinaten*  $\mathbf{x}$  oder die *Lagrange-Koordinaten*  $\mathbf{A}$  verwendet, spricht man von einer ortsbezogenen oder einer materialbezogenen Beschreibung.

Im folgenden werden wir eine Beziehung zwischen der partiellen Ableitung nach der Zeit bei festen Euler-Koordinaten und der bei festen Lagrange-Koordinaten benötigten. Es gilt

$$\frac{\partial \delta}{\partial t} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho,$$

wobei der Gradient  $\nabla \rho$  von  $\rho$  der Vektor der partiellen Ableitungen von  $\rho$  nach den Komponenten des Ortsvektors  $\mathbf{x}$  ist. Wir nennen die rechte Seite die *Materialableitung* von  $\rho$  und führen die Schreibweise

$$\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho \quad (2.10)$$

ein.

Als nächstes beschäftigen wir uns mit der Bewegung eines Teils des Materials, der zum Zeitpunkt  $t$  in dem Gebiet  $R(t) \subset \mathbb{R}^3$  liegt. In Abhängigkeit von  $R(0)$  ist  $R(t)$  gegeben durch

$$R(t) = \{\mathbf{x}(\mathbf{A}, t) \mid \mathbf{A} \in R(0)\},$$

wobei  $\mathbf{x}(\mathbf{A}, t)$  durch Lösen des Anfangswertproblems (2.8) ermittelt wird. Sei  $\mathcal{V}(t)$  das Volumen von  $R(t)$ . Dann gilt

$$\mathcal{V}(t) = \int_{R(t)} d\mathbf{x}, \quad \mathcal{V}(0) = \int_{R(0)} d\mathbf{A}.$$

Aus der Substitutionsregel folgt

$$\mathcal{V}(t) = \int_{R(0)} J(\mathbf{A}, t) d\mathbf{A}$$

mit der Funktionaldeterminanten

$$J(\mathbf{A}, t) = \det \left( \frac{\partial \mathbf{x}}{\partial \mathbf{A}} \right) = \sum_{\pi \in P_3} \text{sign}(\pi) \prod_{j=1}^3 \frac{\partial x_j}{\partial A_{\pi(j)}},$$

wobei  $P_3$  die Menge aller Permutationen von 3 Elementen ist und  $\mathbf{x} = (x_1, x_2, x_3)$  (analog für andere Vektoren) gilt. Im folgenden werden wir die Materialableitung der Funktionaldeterminanten brauchen. Offensichtlich gilt

$$\frac{\partial}{\partial t} \frac{\partial x_i}{\partial A_j} = \frac{\partial v_i}{\partial A_j}.$$

Aus der Produktregel und der Kettenregel folgt daher

$$\begin{aligned} \frac{\partial J}{\partial t} &= \sum_{\pi \in P_3} \text{sign}(\pi) \sum_{i=1}^3 \frac{\partial v_i}{\partial A_{\pi(i)}} \prod_{j \neq i} \frac{\partial x_j}{\partial A_{\pi(j)}} \\ &= \sum_{\pi \in P_3} \text{sign}(\pi) \sum_{i=1}^3 \sum_{k=1}^3 \frac{\partial v_i}{\partial x_k} \frac{\partial x_k}{\partial A_{\pi(i)}} \prod_{j \neq i} \frac{\partial x_j}{\partial A_{\pi(j)}} \\ &= \sum_{i=1}^3 \sum_{k=1}^3 \frac{\partial v_i}{\partial x_k} \sum_{\pi \in P_3} \text{sign}(\pi) \frac{\partial x_k}{\partial A_{\pi(i)}} \prod_{j \neq i} \frac{\partial x_j}{\partial A_{\pi(j)}}. \end{aligned}$$

Die letzte Umformung ist eine Vertauschung der Reihenfolge der Summationen. Die innerste Summe auf der rechten Seite ist die Determinante einer Matrix, die aus der Jacobi-Matrix  $\frac{\partial \mathbf{x}}{\partial \mathbf{A}}$  dadurch entsteht, daß die Zeile  $\frac{\partial x_i}{\partial \mathbf{A}}$  durch  $\frac{\partial x_k}{\partial \mathbf{A}}$  ersetzt wird. Außer im Fall  $k = i$  haben diese Matrizen zwei identische Zeilen, und ihre Determinante verschwindet daher. Es folgt die *Eulersche Entwicklungsformel*

$$\frac{\partial J}{\partial t} = \sum_{i=1}^3 \frac{\partial v_i}{\partial x_i} J = (\nabla \cdot \mathbf{v}) J. \quad (2.11)$$

Hier bezeichnet  $\nabla \cdot \mathbf{v} = \sum_{i=1}^3 \frac{\partial v_i}{\partial x_i}$  die Divergenz von  $\mathbf{v}$ . Als Anwendung berechnen wir die Ableitung des Volumens von  $R(t)$  nach der Zeit:

$$\mathcal{V}'(t) = \int_{R(0)} (\nabla \cdot \mathbf{v})(\mathbf{x}(\mathbf{A}, t), t) J(\mathbf{A}, t) d\mathbf{A} = \int_{R(t)} \nabla \cdot \mathbf{v} d\mathbf{x}$$

In den folgenden Abschnitten werden wir öfter die Zeitableitung von Integralen über den zeitabhängigen Bereich  $R(t)$  berechnen müssen. Sei  $F(\mathbf{x}, t)$  eine glatte Funktion. Dann gilt

$$\begin{aligned} \frac{d}{dt} \int_{R(t)} F(\mathbf{x}, t) d\mathbf{x} &= \frac{d}{dt} \int_{R(0)} F(\mathbf{x}(\mathbf{A}, t), t) J(\mathbf{A}, t) dt = \int_{R(0)} \left( \frac{DF}{Dt} J + F \frac{\partial J}{\partial t} \right) d\mathbf{A} \\ &= \int_{R(0)} \left( \frac{DF}{Dt} + F \nabla \cdot \mathbf{v} \right) J d\mathbf{A} = \int_{R(t)} \left( \frac{\partial F}{\partial t} + \nabla \cdot (F \mathbf{v}) \right) d\mathbf{x}. \end{aligned} \quad (2.12)$$

Wenden wir auf den zweiten Term auf der rechten Seite den Divergenzsatz an, so ergibt sich

$$\frac{d}{dt} \int_{R(t)} F(\mathbf{x}, t) d\mathbf{x} = \int_{R(t)} \frac{\partial F}{\partial t} d\mathbf{x} + \int_{\partial R(t)} F \mathbf{v} \cdot \mathbf{n} ds,$$

mit dem nach außen gerichteten Einheitsnormalvektor auf den Rand  $\mathbf{n}$ . Diese Version zeigt einerseits, daß (2.12) eine Verallgemeinerung der Formel für die Ableitung von eindimensionalen Integralen mit variablen Grenzen ist, andererseits hilft sie bei der physikalischen Interpretation von (2.12). Betrachtet man  $F$  als Dichte einer physikalischen Größe, dann ist durch  $F\mathbf{v}$  eine *Flußdichte* für diese Größe gegeben, und  $F\mathbf{v} \cdot \mathbf{n}$  ist die Komponente dieser Flußdichte normal auf den Rand von  $R(t)$ . Die Interpretation von (2.12) ist daher: Die Änderung der in  $R(t)$  enthaltenen Menge der Größe ist bestimmt durch Änderungen in der Dichte und durch den Fluß der Größe durch den Rand von  $R(t)$ .

### 2.3. Massenerhaltung

In der Teilchenmechanik bedeutet Massenerhaltung, daß kein Teilchen verschwindet und daß die Masse jedes Teilchens sich nicht mit der Zeit ändert. In der Kontinuumsmechanik wird die Massenerhaltung durch eine Differentialgleichung beschrieben. Wir werden zwei Methoden angeben, um diese Gleichung herzuleiten. Die erste Methode beruht auf der Aussage: Die Masse eines Teiles des Materials ändert sich nicht mit der Zeit. In eine Formel übersetzt, heißt das

$$\frac{d}{dt} \int_{R(t)} \rho(\mathbf{x}, t) d\mathbf{x} = 0$$

für beliebige  $R(t)$ . Verwendet man (2.12) mit  $F = \rho$ , so gilt

$$\int_{R(t)} \left( \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right) d\mathbf{x} = 0.$$

Das das Gebiet  $R(t)$  beliebig gewählt werden kann, folgt die *Divergenzform* der *Kontinuitätsgleichung*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (2.13)$$

Verwendet man die Definition (2.10) der Materialableitung, so erhält man die Version

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0. \quad (2.14)$$

Eine Formel für die zeitliche Änderung von Größen, deren Dichte proportional zu  $\rho$  ist, ergibt sich aus einer Kombination von (2.12) und (2.13):

$$\frac{d}{dt} \int_{R(t)} \rho F d\mathbf{x} = \int_{R(t)} \rho \frac{DF}{Dt} d\mathbf{x} \quad (2.15)$$

Die zweite Methode zur Herleitung der Kontinuitätsgleichung beruht auf der Aussage: Betrachtet man ein zeitlich konstantes Gebiet  $R \subset \mathbb{R}^3$ , dann ist die Änderung der in diesem Gebiet enthaltenen Masse durch den Massenfluß durch den Rand  $\partial R$  des Gebietes gegeben. Als Formel ausgedrückt, heißt das

$$\frac{d}{dt} \int_R \rho d\mathbf{x} = - \int_{\partial R} \rho \mathbf{v} \cdot \mathbf{n} ds.$$

Aus dem Divergenzsatz folgt

$$\int_R \left( \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right) d\mathbf{x} = 0,$$

woraus wegen der Beliebigkeit von  $R$  die Kontinuitätsgleichung (2.13) folgt.

## 2.4. Impulserhaltung

Wir wollen das zweite Newtonsche Gesetz für die Gesamtheit des in  $R(t)$  enthaltenen Materials formulieren. Daß das möglich ist, kann mit dem Impulserhaltungssatz (2.2) aus der Teilchenmechanik begründet werden, der es uns erlaubt, in einer Impulsbilanz die mikroskopischen Wechselwirkungskräfte zu vernachlässigen.

Wir definieren den Impuls des in  $R(t)$  enthaltenen Materials durch

$$\int_{R(t)} \rho \mathbf{v} \, d\mathbf{x}.$$

Der Impulserhaltungssatz besagt, daß die zeitliche Änderung dieser Größe durch die äußeren Kräfte gegeben ist, die auf das in  $R(t)$  enthaltene Material wirken. Es ist ein Postulat der Kontinuumsmechanik, daß zwei Arten von Kräften zu berücksichtigen sind. *Volumskräfte* werden beschrieben durch eine Kraftdichte  $\rho \mathbf{f}$ , die proportional zur Massendichte ist. Die auf  $R(t)$  wirkende Kraft ist also

$$\int_{R(t)} \rho \mathbf{f} \, d\mathbf{x}$$

mit der Kraft pro Masseneinheit  $\mathbf{f}$ . Weiters lassen wir *Oberflächenkräfte* der Form

$$\int_{\partial R(t)} \mathbf{t}(\mathbf{x}, t, \mathbf{n}) \, ds$$

zu. Der Vektor  $\mathbf{t}$  heißt *Spannungsvektor*. Er gibt die Kraft pro Flächeneinheit an, die am Punkt  $\mathbf{x} \in \partial R(t)$  auf das in  $R(t)$  enthaltene Material durch benachbartes Material ausgeübt wird. Umgekehrt übt das in  $R(t)$  enthaltene Material die Kraft  $\mathbf{t}(\mathbf{x}, t, -\mathbf{n})$  auf die Umgebung aus.

Sowohl Volumskräfte als auch Spannungskräfte lassen sich durch mikroskopische Wechselwirkungen zwischen dem in  $R(t)$  enthaltenen Material und dem Rest der Welt erklären. Volumskräfte werden verursacht von Wechselwirkungskräften mit einem großen Wirkungsbereich, während Wechselwirkungen mit kleinem Wirkungsbereich im Kontinuumsmodell Spannungskräfte bewirken. Beispiele für die erste Gruppe sind Gravitationskräfte und die Kräfte zwischen elektrischen Ladungen. In die zweite Gruppe fallen die Wechselwirkungen zwischen den in Abschnitt 2.1 besprochenen harten Kugeln.

Das zweite Newtonsche Gesetz hat also die Form

$$\frac{d}{dt} \int_{R(t)} \rho \mathbf{v} \, d\mathbf{x} = \int_{R(t)} \rho \mathbf{f} \, d\mathbf{x} + \int_{\partial R(t)} \mathbf{t}(\mathbf{x}, t, \mathbf{n}) \, ds.$$

Wendet man die Formel (2.15) auf die Komponenten der linken Seite an, so folgt

$$\int_{R(t)} \left( \rho \frac{D\mathbf{v}}{Dt} - \rho \mathbf{f} \right) d\mathbf{x} = \int_{\partial R(t)} \mathbf{t}(\mathbf{x}, t, \mathbf{n}) \, ds. \quad (2.16)$$

Näheres über den Spannungsvektor läßt sich herausfinden, wenn man (2.16) für eine Familie  $\{R_L \mid L > 0\}$  von ähnlichen Mengen (siehe Anhang) mit gemeinsamem Mittelpunkt betrachtet, wobei die das Gebiet beschreibende Funktion  $f_L$  (siehe Anhang) von der Form  $f_L(\phi, \theta) = L f_1(\phi, \theta)$  ist. Aus dem Lemma im Anhang folgt für das Volumen von  $R_L$  die Formel  $\mathcal{V}[R_L] = L^3 \mathcal{V}[R_1]$ .

Ersetzen wir in (2.16)  $R(t)$  durch  $R_L$ , dann folgt aus dem Mittelwertsatz der Integralrechnung, daß die linke Seite für  $L \rightarrow 0$  von der Ordnung  $O(L^3)$  ist. Dividieren wir die Gleichung durch  $L^2$  und lassen  $L \rightarrow 0$ , dann folgt

$$\lim_{L \rightarrow 0} \frac{1}{L^2} \int_{\partial R_L} \mathbf{t} \, ds = 0. \quad (2.17)$$

Diese Gleichung heißt *Prinzip des lokalen Spannungsgleichgewichtes*. Sie besagt, daß—lokal betrachtet—die Spannungskräfte einander aufheben. Durch Anwendung von (2.17) auf spezielle Familien  $\{R_L\}$  werden wir wichtige Aussagen über den Spannungsvektor herleiten.

### Abbildung 2.1

Sei  $R_L(\mathbf{x}, \mathbf{n})$  ein Quader mit quadratischer Grundfläche, die einen Eckpunkt in  $\mathbf{x}$  hat und normal auf  $\mathbf{n}$  steht. Die Seitenlänge sei  $L$ . Die Höhe des Quaders sei  $\varepsilon L$  mit  $\varepsilon \ll 1$ . Offensichtlich ist  $\{R_L(\mathbf{x}, \mathbf{n}) \mid L > 0\}$  eine Familie von ähnlichen Mengen mit Mittelpunkt  $\mathbf{x}$ . Aus dem Mittelwertsatz folgt, daß für das Integral in (2.17)

$$\int_{\partial R_L} \mathbf{t} \, ds = L^2 (\mathbf{t}(\mathbf{x}_D, t, \mathbf{n}) + \mathbf{t}(\mathbf{x}_G, t, -\mathbf{n}) + O(\varepsilon))$$

gilt, wobei  $\mathbf{x}_G$  und  $\mathbf{x}_D$  Punkte auf der Grund- und Deckfläche sind. Da diese Punkte für  $L \rightarrow 0$  gegen  $\mathbf{x}$  konvergieren, folgt mit (2.17) aus den sukzessiven Grenzübergängen  $L \rightarrow 0$  und  $\varepsilon \rightarrow 0$  die Gleichung

$$\mathbf{t}(\mathbf{x}, t, -\mathbf{n}) = -\mathbf{t}(\mathbf{x}, t, \mathbf{n}). \quad (2.18)$$

Das ist die Kontinuumsversion des dritten Newtonschen Gesetzes.

Nun wählen wir  $R_L(\mathbf{x}, \mathbf{n})$  als Tetraeder mit einem Eckpunkt in  $\mathbf{x}$ , sodaß  $\mathbf{n}$  der nach außen gerichtete Normalvektor auf die gegenüberliegende Seitenfläche mit Flächeninhalt  $S$  ist. Die übrigen Seitenflächen mit Flächeninhalt  $S_i$ ,  $i = 1, 2, 3$  seien normal auf die kanonischen Basisvektoren  $\mathbf{e}_i$ ,  $i = 1, 2, 3$ . Der Abstand zwischen  $\mathbf{x}$  und der geneigten Seitenfläche sei so gewählt, daß  $S = L^2$  gilt (siehe Abb. 2.1). Auch in diesem Fall ist offensichtlich, daß  $\{R_L(\mathbf{x}, \mathbf{n}) \mid L > 0\}$  eine Familie von ähnlichen Mengen mit Mittelpunkt  $\mathbf{x}$  ist.

Wegen des Mittelwertsatzes ist die Kraft, die auf die Seitenfläche mit Flächeninhalt  $S_i$  wirkt, gegeben durch

$$S_i \mathbf{t}(\mathbf{x}_i, t, -\mathbf{e}_i), \quad i = 1, 2, 3,$$

wobei  $\mathbf{x}_i$  ein Punkt auf dieser Fläche ist. Die auf die geneigte Seitenfläche wirkende Kraft ist

$$S \mathbf{t}(\mathbf{x}_0, t, \mathbf{n})$$

mit dem in der Fläche liegenden Punkt  $\mathbf{x}_0$ . Da  $S_i$  die Projektion von  $S$  auf die  $i$ -te Koordinatenebene ist, gilt  $S_i = (\mathbf{n} \cdot \mathbf{e}_i)S = n_i S$ . Aus dieser Beziehung und (2.17) folgt

$$\sum_{i=1}^3 n_i \mathbf{t}(\mathbf{x}, t, -\mathbf{e}_i) + \mathbf{t}(\mathbf{x}, t, \mathbf{n}) = 0,$$

oder, wegen (2.18),

$$\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = \sum_{i=1}^3 n_i \mathbf{t}(\mathbf{x}, t, \mathbf{e}_i). \quad (2.19)$$

Sei  $\mathbf{T}(\mathbf{x}, t)$  die Matrix mit den Zeilen  $\mathbf{t}(\mathbf{x}, t, \mathbf{e}_i)$ ,  $i = 1, 2, 3$ , d.h.

$$T_{ij}(\mathbf{x}, t) = t_j(\mathbf{x}, t, \mathbf{e}_i),$$

dann können wir (2.19) schreiben als

$$\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = \mathbf{n} \cdot \mathbf{T}(\mathbf{x}, t), \quad (2.20)$$

wobei  $\mathbf{n} \cdot \mathbf{T}$  der Vektor mit den Komponenten  $(\mathbf{n} \cdot \mathbf{T})_j = \sum_{i=1}^3 n_i T_{ij}$  ist. Die Matrix  $\mathbf{T}$  heißt *Spannungstensor*. Verwendet man (2.20) und den Divergenzsatz in (2.16), so erhält man

$$\int_{R(t)} \left( \rho \frac{D\mathbf{v}}{Dt} - \rho \mathbf{f} - \nabla \cdot \mathbf{T} \right) d\mathbf{x} = 0$$

und damit wegen der Beliebigkeit von  $R(t)$  die *Cauchyschen Differentialgleichungen*

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} + \nabla \cdot \mathbf{T}, \quad (2.21)$$

die die Kontinuumsversion des Impulserhaltungssatzes (2.2) darstellen. Die Anwendung der Divergenz auf eine Matrix ist komponentenweise definiert durch

$$(\nabla \cdot \mathbf{T})_j = \sum_{i=1}^3 \frac{\partial T_{ij}}{\partial x_i}.$$

Addiert man (2.21) und das Produkt der Kontinuitätsgleichung (2.13) mit  $\mathbf{v}$ , so erhält man die Divergenzform der Impulsgleichung

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{T}) = \rho \mathbf{f}. \quad (2.22)$$

Hier bezeichnet  $\mathbf{x} \otimes \mathbf{y} = \mathbf{xy}^{tr}$  das Tensorprodukt zweier Vektoren, d.h. die Matrix mit den Elementen  $x_i y_j$ ,  $i, j = 1, 2, 3$ . Der Ausdruck  $\rho \mathbf{v} \otimes \mathbf{v}$  kann als *Impulsflußdichte* interpretiert werden. So gesehen liefern auch die Spannungskräfte einen Beitrag zum Impulsfluß.

## 2.5. Drehimpulserhaltung

Wir definieren den Drehimpuls des in  $R(t)$  enthaltenen Materials um den ruhenden Punkt  $\mathbf{x}^{(0)}$  durch

$$\int_{R(t)} \rho(\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{v} d\mathbf{x},$$

und auf ähnliche Art die durch Volums- und Spannungskräfte verursachten Drehmomente. Das Postulat der Drehimpulserhaltung nimmt die Form

$$\frac{d}{dt} \int_{R(t)} \rho(\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{v} d\mathbf{x} = \int_{R(t)} \rho(\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{f} d\mathbf{x} + \int_{\partial R(t)} (\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{t} ds \quad (2.23)$$

an. Mit Hilfe der Formel (2.15) kann man zeigen, daß die linke Seite dieser Gleichung geschrieben werden kann als

$$\int_{R(t)} \rho(\mathbf{x} - \mathbf{x}^{(0)}) \times \frac{D\mathbf{v}}{Dt} d\mathbf{x}. \quad (2.24)$$

Wir wollen uns mit der ersten Komponente der Vektorgleichung (2.23) beschäftigen. Für die erste Komponente des Vektors  $(\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{t}$  gilt

$$\begin{aligned} ((\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{t})_1 &= (x_2 - x_2^{(0)}) \sum_{i=1}^3 n_i T_{i3} - (x_3 - x_3^{(0)}) \sum_{i=1}^3 n_i T_{i2} \\ &= \sum_{i=1}^3 n_i [(x_2 - x_2^{(0)}) T_{i3} - (x_3 - x_3^{(0)}) T_{i2}]. \end{aligned}$$

Der Divergenzsatz impliziert daher

$$\begin{aligned} \int_{\partial R(t)} ((\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{t})_1 ds &= \int_{R(t)} \sum_{i=1}^3 \frac{\partial}{\partial x_i} [(x_2 - x_2^{(0)}) T_{i3} - (x_3 - x_3^{(0)}) T_{i2}] d\mathbf{x} \\ &= \int_{R(t)} \left[ \sum_{i=1}^3 \left( (x_2 - x_2^{(0)}) \frac{\partial T_{i3}}{\partial x_i} - (x_3 - x_3^{(0)}) \frac{\partial T_{i2}}{\partial x_i} \right) + T_{23} - T_{32} \right] d\mathbf{x}. \end{aligned}$$

Nun betrachten wir die erste Komponente von (2.23) für eine Familie  $\{R_L\}$  von ähnlichen Mengen, die sich für  $L \rightarrow 0$  auf den Punkt  $\mathbf{x}^{(0)}$  zusammenziehen:

$$\int_{R_L} \rho \left( (\mathbf{x} - \mathbf{x}^{(0)}) \times \frac{D\mathbf{v}}{Dt} \right)_1 d\mathbf{x} = \int_{R_L} \rho ((\mathbf{x} - \mathbf{x}^{(0)}) \times \mathbf{f})_1 d\mathbf{x} \\ + \int_{R_L} \left[ \sum_{i=1}^3 \left( (x_2 - x_2^{(0)}) \frac{\partial T_{i3}}{\partial x_i} - (x_3 - x_3^{(0)}) \frac{\partial T_{i2}}{\partial x_i} \right) + T_{23} - T_{32} \right] d\mathbf{x} \quad (2.25)$$

Alle Integranden, die Komponenten von  $\mathbf{x} - \mathbf{x}^{(0)}$  als Faktoren enthalten, sind  $O(L)$  für  $L \rightarrow 0$ . Division von (2.25) durch  $L^3$  und  $L \rightarrow 0$  impliziert wegen des Mittelwertsatzes

$$T_{23}(\mathbf{x}^{(0)}, t) = T_{32}(\mathbf{x}^{(0)}, t).$$

Analog zeigt man  $T_{12} = T_{21}$  und  $T_{13} = T_{31}$ . Das bedeutet, daß die Forderung der Drehimpulserhaltung die Symmetrie des Spannungstensors impliziert:

$$\mathbf{T} = \mathbf{T}^{tr} \quad (2.26)$$

## 2.6. Energieerhaltung

Inspiziert durch das Resultat (2.4) aus der Teilchenmechanik postulieren wir die Energieerhaltungsgleichung für das in  $R(t)$  enthaltene Material:

$$\frac{d}{dt} \int_{R(t)} \rho \left( \frac{|\mathbf{v}|^2}{2} + e \right) d\mathbf{x} = \int_{R(t)} \rho \mathbf{f} \cdot \mathbf{v} d\mathbf{x} + \int_{\partial R(t)} (-h + \mathbf{t} \cdot \mathbf{v}) ds \quad (2.27)$$

Der Integrand auf der linken Seite ist die Energiedichte mit der inneren Energie pro Masseneinheit  $e(\mathbf{x}, t)$ . Weiters enthält die Gleichung die Arbeit, die von den Volums- und Spannungs Kräften verrichtet wird. Die Größe  $h$  ist der durch Wärmeleitung verursachte Energieverlust durch den Rand von  $R(t)$ .

Um aus (2.27) eine Differentialgleichung herzuleiten, ziehen wir mit Hilfe von (2.12) die Zeitableitung auf der linken Seite in das Integral. Die Randintegrale versuchen wir als Volumsintegrale zu schreiben. Es gilt

$$\int_{\partial R(t)} \mathbf{t} \cdot \mathbf{v} ds = \int_{\partial R(t)} \sum_{j=1}^3 \sum_{i=1}^3 n_i T_{ij} v_j ds = \int_{R(t)} \sum_{i=1}^3 \frac{\partial}{\partial x_i} \left( \sum_{j=1}^3 T_{ij} v_j \right) d\mathbf{x} \\ = \int_{R(t)} \nabla \cdot (\mathbf{T} \cdot \mathbf{v}) d\mathbf{x}.$$

Analog zum Beweis der Existenz des Spannungstensors zeigen wir

$$h = \mathbf{n} \cdot \mathbf{q}(\mathbf{x}, t)$$

durch Anwendung von (2.27) auf Familien von ähnlichen Mengen. Der Vektor  $\mathbf{q}$  ist die *Wärmeflußdichte*. Damit kann auch das Oberflächenintegral von  $h$  in ein Volumsintegral verwandelt werden. Da  $R(t)$  in (2.27) beliebig ist, folgt die Energieerhaltungsgleichung

$$\rho \frac{D}{Dt} \left( \frac{|\mathbf{v}|^2}{2} + e \right) = \rho \mathbf{f} \cdot \mathbf{v} - \nabla \cdot (\mathbf{q} - \mathbf{T} \cdot \mathbf{v}). \quad (2.28)$$

Unter Verwendung der Massenerhaltung läßt sich auch für diese Gleichung eine Divergenzform angeben:

$$\frac{\partial}{\partial t} \left[ \rho \left( \frac{|\mathbf{v}|^2}{2} + e \right) \right] + \nabla \cdot \left[ \rho \mathbf{v} \left( \frac{|\mathbf{v}|^2}{2} + e \right) + \mathbf{q} - \mathbf{T} \cdot \mathbf{v} \right] = \rho \mathbf{f} \cdot \mathbf{v} \quad (2.29)$$



## 2.7. Konstitutive Gleichungen

Fassen wir die in den Abschnitten 2.3–2.6 hergeleiteten Gleichungen noch einmal zusammen.

Massenerhaltung:

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \quad (2.30)$$

Impulserhaltung:

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} + \nabla \cdot \mathbf{T} \quad (2.31)$$

Drehimpulserhaltung:

$$\mathbf{T} = \mathbf{T}^{tr} \quad (2.32)$$

Energieerhaltung:

$$\rho \frac{D}{Dt} \left( \frac{|\mathbf{v}|^2}{2} + e \right) = \rho \mathbf{f} \cdot \mathbf{v} - \nabla \cdot (\mathbf{q} - \mathbf{T} \cdot \mathbf{v}) \quad (2.33)$$

Die Unbekannten sind die Dichte  $\rho$ , die Geschwindigkeit  $\mathbf{v}$ , der Spannungstensor  $\mathbf{T}$ , die spezifische innere Energie  $e$  und die Wärmeflußdichte  $\mathbf{q}$ . Das sind 17 unbekannte skalare Funktionen. Die Drehimpulserhaltung eliminiert 3 Unbekannte. Weiters haben wir 5 Differentialgleichungen zur Verfügung. Das zeigt, daß das System (2.30)–(2.33) stark unterbestimmt ist. Das ist auch nicht anders zu erwarten, wenn diese Gleichungen für eine große Klasse von Materialien gültig sein sollen. Wir haben noch keinerlei Informationen über die Art des Materials verwendet, also z.B. ob es fest, flüssig oder gasförmig ist.

Um das Gleichungssystem zu schließen, ist die Angabe von zusätzlichen Gleichungen, sogenannten *konstitutiven Beziehungen*, notwendig. Wir führen einige Beispiele dafür an:

In einem *starrten Körper* ist der Abstand zwischen zwei Teilchen konstant.

In einem *inkompressiblen Medium* ist die Dichte jedes Teilchens konstant. Das wird durch die Gleichung  $D\rho/Dt = 0$  ausgedrückt.

Für ein *reibungsfreies Fluid* (Flüssigkeit oder Gas) ist der Spannungsvektor gegeben durch  $\mathbf{t} = -p\mathbf{n}$ , wobei  $p$  der Druck ist.

Das *Newton-Fouriersche Abkühlungsgesetz* besagt, daß die Wärmeflußdichte proportional zum Gradienten der Temperatur ist:  $\mathbf{q} = -\kappa \nabla T$ , wobei  $T$  die Temperatur und  $\kappa$  die Wärmeleitfähigkeit ist.

Für ein *ideales Gas* gilt die Beziehung  $p = \rho RT$ , in der  $R$  eine von dem Gas abhängige Konstante ist. Weiters ist oft die Annahme begründet, daß Änderungen in der inneren Energie proportional zu Änderungen in der Temperatur sind:  $e = c_v T + \text{const}$ , mit der spezifischen Wärme bei konstantem Volumen  $c_v$ .

Konstitutive Gleichungen sind auf viele verschiedene Arten entstanden. Manche sind Vereinfachungen (Störungen) der wahren Situation. Manche können als durch Experimente gerechtfertigt gelten. In den wenigsten Fällen gibt es eine deduktive Herleitung aus dem mikroskopischen Verhalten.

## Anhang: Ähnliche Gebiete

**Definition.** Eine abgeschlossene Menge  $R \subset \mathbb{R}^3$  heißt *sternförmig*, wenn es einen Punkt  $\mathbf{x}_0 \in R$  (genannt *Mittelpunkt*) gibt, sodaß für jeden anderen Punkt  $\mathbf{x} \in R$  die Strecke zwischen  $\mathbf{x}_0$  und  $\mathbf{x}$  in  $R$  enthalten ist.

Seien  $(r, \phi, \theta)$  sphärische Koordinaten mit Zentrum in  $\mathbf{x}_0$ . Dann gibt es eine Funktion  $f(\phi, \theta)$ , sodaß  $R$  gegeben ist durch

$$0 \leq r \leq f(\phi, \theta) \quad \text{für } 0 \leq \phi \leq \pi, \quad 0 \leq \theta \leq 2\pi.$$

**Definition.** Zwei sternförmige Mengen  $R_1$  und  $R_2$  heißen *ähnlich*, wenn sie nach eventueller Rotation und Translation durch die Ungleichungen

$$0 \leq r \leq f_i(\phi, \theta), \quad i = 1, 2$$

beschrieben werden können, wobei  $f_2 = Lf_1$  für eine positive Konstante  $L$  gilt.

**Lemma.** Sei  $\mathcal{V}[R_i]$  das Volumen von  $R_i$ ,  $i = 1, 2$ . Dann gilt für zwei ähnliche Mengen  $R_1$  und  $R_2$

$$\mathcal{V}[R_2] = L^3 \mathcal{V}[R_1]$$

mit der Konstanten  $L$  aus obiger Definition.

*Beweis.* Es gilt

$$\begin{aligned} \mathcal{V}[R_2] &= \int_0^{2\pi} \int_0^\pi \int_0^{f_2(\phi, \theta)} r^2 \sin \phi \, dr \, d\phi \, d\theta = \frac{1}{3} \int_0^{2\pi} \int_0^\pi (f_2(\phi, \theta))^3 \sin \phi \, d\phi \, d\theta \\ &= \frac{L^3}{3} \int_0^{2\pi} \int_0^\pi (f_1(\phi, \theta))^3 \sin \phi \, d\phi \, d\theta = L^3 \mathcal{V}[R_1]. \end{aligned}$$

■

## 3.KAPITEL Strömungsmechanik

### 3.1. Die Navier-Stokes-Gleichungen

Die Strömungsmechanik beschäftigt sich mit der Bewegung von Fluiden (d.h. Flüssigkeiten und Gasen). Fluide sind ausgezeichnet durch die Art der in ihnen wirkenden Spannkraften. Zerlegt man den Spannungsvektor  $\mathbf{t}(\mathbf{x}, t, \mathbf{n})$  in eine normale Komponente (*Normalspannung*, parallel zu  $\mathbf{n}$ ) und in tangentialen Komponenten (*Scherspannungen*), dann gilt für ein bewegungsloses Fluid, daß die Scherspannungen verschwinden. Der Spannungsvektor ist also gegeben durch

$$\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = -p(\mathbf{x}, t, \mathbf{n})\mathbf{n},$$

wobei die skalare Größe  $p(\mathbf{x}, t, \mathbf{n})$  als *Druck* bezeichnet wird. Verwendet man die Darstellung des Spannungsvektors mit Hilfe des Spannungstensors  $\mathbf{T}$ , dann folgt

$$\mathbf{n} \cdot \mathbf{T}(\mathbf{x}, t) = -p(\mathbf{x}, t, \mathbf{n})\mathbf{n}. \quad (3.1)$$

Setzen wir in dieser Gleichung für  $\mathbf{n}$  den  $i$ -ten kanonischen Basisvektor  $\mathbf{e}_i$  ein und bilden wir das Skalarprodukt mit  $\mathbf{e}_j$ , so erhalten wir

$$T_{ij}(\mathbf{x}, t) = -p(\mathbf{x}, t, \mathbf{e}_i)\delta_{ij}, \quad i, j = 1, 2, 3,$$

mit dem Kronecker-Symbol  $\delta_{ij}$ . Damit wird die  $j$ -te Komponente von (3.1) zu

$$-n_j p(\mathbf{x}, t, \mathbf{e}_j) = -n_j p(\mathbf{x}, t, \mathbf{n}).$$

Aus dieser Gleichung folgt, daß der Druck unabhängig von  $\mathbf{n}$  ist, und daher für den Spannungstensor in einem ruhenden Fluid

$$\mathbf{T}(\mathbf{x}, t) = -p(\mathbf{x}, t)\mathbf{I} \quad (3.2)$$

gilt ( $\mathbf{I}$  ist die  $3 \times 3$ -Einheitsmatrix).

**Beispiel 3.1:** Um unser Vertrauen in (3.2) zu stärken, überprüfen wir, ob es zusammen mit der Impulsgleichung zu den bekannten Gesetzen der Hydrostatik führt. Wir setzen in der Impulsgleichung (2.31)  $\mathbf{v} = 0$  und berücksichtigen den Einfluß der Gravitation, indem wir für die Volumskraft  $\mathbf{f} = -g\mathbf{e}_3$  setzen, wobei  $g$  die Gravitationsbeschleunigung ist und wir annehmen, daß  $\mathbf{e}_3$  senkrecht nach oben zeigt. Die Impulsgleichung hat dann die Gestalt

$$\nabla p = -\rho g \mathbf{e}_3,$$

also, mit  $\mathbf{x} = (x, y, z)$ ,

$$p_x = p_y = 0, \quad p_z = -\rho g.$$

Der Druck ist nur von der Höhe abhängig, woraus folgt, daß dasselbe auch für die Dichte  $\rho$  gilt. Im Fall konstanter Dichte wächst der Druck linear mit der Tiefe:

$$p = p_a - \rho g z,$$

wobei  $z = 0$  die Oberfläche einer Flüssigkeit und  $p_a$  der Druck in der darüberliegenden Atmosphäre ist.

Eine weitere Annahme ist, daß in einem bewegten Fluid zusätzlich zum Druck Reibungskraften (*Viskosität*) zur Spannung beitragen:

$$\mathbf{T} = -p\mathbf{I} + \boldsymbol{\tau}$$

Die Matrix  $\boldsymbol{\tau}(\mathbf{x}, t)$  wird als *viskoser Spannungstensor* bezeichnet. Da bei einer Bewegung mit ortsunabhängiger Geschwindigkeit  $\mathbf{v} = \mathbf{v}(t)$  keine Reibungskraften zu erwarten sind, nimmt man an, daß  $\boldsymbol{\tau}$  von Ortsableitungen der Geschwindigkeit abhängig ist. Um das Modell möglichst einfach zu halten, wird weiter

angenommen, daß nur die Ableitungen erster Ordnung eingehen und daß die Abhängigkeit linear ist. Es gilt daher

$$\tau_{ij} = \sum_{k,l} C_{ijkl} \frac{\partial v_k}{\partial x_l}, \quad i, j = 1, 2, 3. \quad (3.3)$$

Aus der Symmetrie des Spannungstensors folgt die Bedingung

$$C_{ijkl} = C_{jikl}, \quad i, j, k, l = 1, 2, 3$$

für die Koeffizienten des Tensors vierter Stufe  $\mathbf{C}$ . Unter der Annahme der Homogenität des Materials sind diese Koeffizienten konstant. Die Annahme, daß nur bei einer Relativbewegung verschiedener Teile des Fluids Reibungskräfte auftreten, ist in (3.3) noch nicht vollständig berücksichtigt. Rotiert nämlich das Fluid wie ein starrer Körper, dann findet auch keine Relativbewegung statt, obwohl sich die Geschwindigkeit mit dem Ort ändert. Eine Rotation um die  $z$ -Achse mit konstanter Winkelgeschwindigkeit  $\omega$  ist gegeben durch

$$\mathbf{v} = \omega(-y, x, 0).$$

Die Forderung, daß der viskose Spannungstensor für dieses  $\mathbf{v}$  verschwindet, führt auf die Gleichung

$$C_{ij21} = C_{ij12}, \quad i, j = 1, 2, 3.$$

Analog zeigt man mit Rotationen um die  $x$ - bzw. die  $y$ -Achse

$$C_{ij32} = C_{ij23}, \quad C_{ij31} = C_{ij13}, \quad i, j = 1, 2, 3.$$

Damit vereinfacht sich (3.3) zu

$$\begin{aligned} \tau_{ij} = & C_{ij11}u_x + C_{ij22}v_y + C_{ij33}w_z + C_{ij12}(u_y + v_x) + C_{ij13}(u_z + w_x) \\ & + C_{ij23}(v_z + w_y), \quad i, j = 1, 2, 3. \end{aligned} \quad (3.4)$$

Bezeichnet man den symmetrischen Anteil

$$\mathbf{D} = \frac{1}{2}(\nabla \mathbf{v} + \nabla \mathbf{v}^{tr})$$

der Matrix  $\nabla \mathbf{v}$  als *Deformationstensor*, dann ergibt sich die Aussage, daß der viskose Spannungstensor nur vom Deformationstensor abhängig ist.

Die Darstellung (3.4) kann weiter vereinfacht werden, wenn man annimmt, daß das betrachtete Fluid *isotrop* ist. Das bedeutet, daß es keine ausgezeichneten Richtungen gibt. Für das mathematische Modell folgt daraus, daß es invariant unter Drehungen des Raumes sein muß. Drehungen werden durch Matrizen  $\mathbf{S}$  mit  $\mathbf{S}^{-1} = \mathbf{S}^{tr}$  beschrieben. Für die neuen Ortskoordinaten  $\mathbf{x}'$  nach der Drehung gilt

$$\mathbf{x}' = \mathbf{S}\mathbf{x}.$$

Genauso werden alle Vektoren, wie z.B. ein Geschwindigkeitsfeld  $\mathbf{v}$  oder die durch Reibung bedingte Kraft  $\mathbf{f} = \mathbf{n} \cdot \boldsymbol{\tau}$  auf ein Flächenstück mit Normalvektor  $\mathbf{n}$ , transformiert:

$$\mathbf{v}' = \mathbf{S}\mathbf{v}, \quad \mathbf{f}' = \mathbf{S}\mathbf{f} \quad (3.5)$$

Es gilt also

$$\begin{aligned} f'_i &= \sum_j S_{ij}(\mathbf{n} \cdot \boldsymbol{\tau})_j = \sum_{j,k} S_{ij}n_k \tau_{kj} = \sum_{j,k,l} S_{ij}S_{lk}n'_l \tau_{kj} \\ &= \sum_l n'_l \sum_{j,k} S_{lk} \tau_{kj} S_{ji}^{tr} = (\mathbf{n}' \cdot (\mathbf{S}\boldsymbol{\tau}\mathbf{S}^{tr}))_i, \quad i = 1, 2, 3, \end{aligned}$$

woraus wegen  $\mathbf{f}' = \mathbf{n}' \cdot \boldsymbol{\tau}'$  und der Beliebigkeit von  $\mathbf{n}'$

$$\boldsymbol{\tau}' = \mathbf{S}\boldsymbol{\tau}\mathbf{S}^{tr} \quad (3.6)$$

folgt. Komponentenweise geschrieben, gilt

$$\tau'_{ij} = \sum_{k,l} S_{ik} \tau_{kl} S_{jl} = \sum_{k,l,m,n} S_{ik} S_{jl} C_{klmn} \frac{\partial v_m}{\partial x_n}, \quad i, j = 1, 2, 3. \quad (3.7)$$

Für die ersten Ableitungen der Geschwindigkeitskomponenten gilt

$$\frac{\partial v'_i}{\partial x'_j} = \sum_k \frac{\partial v'_i}{\partial x_k} \frac{\partial x_k}{\partial x'_j} = \sum_k S_{jk} \frac{\partial v'_i}{\partial x_k} = \sum_{k,l} S_{jk} S_{il} \frac{\partial v_l}{\partial x_k} = \sum_{k,l} S_{il} \frac{\partial v_l}{\partial x_k} S_{kj}^{tr}, \quad i, j = 1, 2, 3,$$

und daher

$$\nabla' \mathbf{v}' = \mathbf{S} \nabla \mathbf{v} \mathbf{S}^{tr}, \quad (3.8)$$

wobei  $\nabla'$  der Gradient bezüglich  $\mathbf{x}'$  ist. So wie in (3.5) ein Gesetz zur Transformation von Vektoren verwendet wurde, sind (3.6) und (3.8) Anwendungen der Transformationsregel für Tensoren zweiter Stufe. Unter der Annahme, daß nach der Drehung der Zusammenhang zwischen dem viskosen Spannungstensor und den Ableitungen der Geschwindigkeitskomponenten unverändert ist, kann  $\boldsymbol{\tau}'$  mit Hilfe von (3.8) berechnet werden:

$$\tau'_{ij} = \sum_{k,l} C_{ijkl} \frac{\partial v'_k}{\partial x'_l} = \sum_{k,l,m,n} S_{km} S_{ln} C_{ijkl} \frac{\partial v_m}{\partial x_n}, \quad i, j = 1, 2, 3 \quad (3.9)$$

Da  $\nabla \mathbf{v}$  beliebig gewählt werden kann, folgt aus einem Vergleich von (3.7) und (3.9)

$$\sum_{k,l} S_{ik} S_{jl} C_{klmn} = \sum_{k,l} S_{km} S_{ln} C_{ijkl}, \quad i, j, m, n = 1, 2, 3. \quad (3.10)$$

Daß (3.10) für alle orthogonalen Matrizen  $S$  gelten soll, ist eine starke Forderung an den Tensor  $\mathbf{C}$ . Würde man in der rechten Seite von (3.10)  $C_{ijkl}$  durch  $C'_{ijkl}$  ersetzen, erhielte man die Transformationsregel für Tensoren vierter Stufe.

Gleichungen für die Komponenten von  $\mathbf{C}$  werden durch spezielle Wahl von  $\mathbf{S}$  in (3.10) hergeleitet. Ist z.B.  $\mathbf{S}$  eine Diagonalmatrix, also

$$\mathbf{S} = \begin{pmatrix} (-1)^{k_1} & 0 & 0 \\ 0 & (-1)^{k_2} & 0 \\ 0 & 0 & (-1)^{k_3} \end{pmatrix},$$

dann reduziert sich (3.10) auf

$$((-1)^{k_i+k_j+k_m+k_n} - 1) C_{ijmn} = 0, \quad i, j, m, n = 1, 2, 3.$$

Durch geeignete Wahl von  $k_1$ ,  $k_2$  und  $k_3$  zeigt man leicht, daß alle Komponenten von  $\mathbf{C}$  verschwinden müssen, bis auf die Komponenten der Form  $C_{iijj}$ ,  $i, j = 1, 2, 3$ , und

$$\begin{aligned} C_{1212} &= C_{1221} = C_{2121} = C_{2112}, \\ C_{1313} &= C_{1331} = C_{3131} = C_{3113}, \\ C_{2323} &= C_{2332} = C_{3232} = C_{3223}. \end{aligned}$$

Das bedeutet, daß  $\mathbf{C}$  maximal 12 verschiedene nicht verschwindende Komponenten hat.

Nun wählen wir

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

und eine Reihe spezieller Quadrupel  $(i, j, m, n)$  in (3.10). Die folgende Tabelle enthält diese Quadrupel und die sich aus (3.10) ergebenden Gleichungen:

$$\begin{array}{ll} (2, 2, 1, 1) & C_{1111} = C_{2222} \\ (1, 1, 3, 3) & C_{1111} = C_{3333} \\ (1, 1, 1, 1) & C_{1122} = C_{3311} \\ (2, 2, 2, 2) & C_{1122} = C_{2233} \\ (1, 1, 2, 2) & C_{1133} = C_{3322} \\ (2, 2, 3, 3) & C_{1133} = C_{2211} \\ (1, 2, 1, 3) & C_{1221} = C_{3113} \\ (1, 3, 2, 3) & C_{1331} = C_{3223} \end{array}$$

Schließlich folgt noch mit

$$\mathbf{S} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und  $(i, j, m, n) = (1, 1, 1, 1)$  die Gleichung

$$C_{1122} = C_{2211}.$$

Diese Gleichungen zeigen, daß die Komponenten von  $\mathbf{C}$  höchstens drei verschiedene nicht verschwindende Werte annehmen, und zwar

$$\kappa = C_{iii}, \quad i = 1, 2, 3, \quad \lambda = C_{iij}, \quad i \neq j, \quad \mu = C_{ijji} = C_{ijij}, \quad i \neq j.$$

Der viskose Spannungstensor ist daher gegeben durch

$$\boldsymbol{\tau} = \lambda(\nabla \cdot \mathbf{v})\mathbf{I} + 2\mu\mathbf{D} + (\kappa - \lambda - 2\mu)\text{diag}(\nabla\mathbf{v}),$$

wobei  $\text{diag}(\mathbf{A})$  die Diagonalmatrix mit den Diagonalelementen von  $\mathbf{A}$  ist. Die ersten beiden Summanden auf der rechten Seite erfüllen die Gleichung (3.6) wegen

$$\begin{aligned} \mathbf{I}' &= \mathbf{I} = \mathbf{S}\mathbf{I}\mathbf{S}^{tr}, \\ \mathbf{D}' &= \frac{1}{2}(\nabla'\mathbf{v}' + (\nabla'\mathbf{v}')^{tr}) = \frac{1}{2}(\mathbf{S}\nabla\mathbf{v}\mathbf{S}^{tr} + (\mathbf{S}\nabla\mathbf{v}\mathbf{S}^{tr})^{tr}) \\ &= \frac{1}{2}\mathbf{S}(\nabla\mathbf{v} + (\nabla\mathbf{v})^{tr})\mathbf{S}^{tr} = \mathbf{S}\mathbf{D}\mathbf{S}^{tr}. \end{aligned}$$

Da aber im allgemeinen

$$\text{diag}(\nabla'\mathbf{v}') = \text{diag}(\mathbf{S}\nabla\mathbf{v}\mathbf{S}^{tr}) \neq \mathbf{S}\text{diag}(\nabla\mathbf{v})\mathbf{S}^{tr}$$

gilt, folgt  $\kappa - \lambda - 2\mu = 0$ , und der viskose Spannungstensor ist gegeben durch

$$\boldsymbol{\tau} = \lambda(\nabla \cdot \mathbf{v})\mathbf{I} + 2\mu\mathbf{D}. \quad (3.11)$$

Zur Interpretation der Parameter  $\lambda$  und  $\mu$  betrachten wir zwei einfache Strömungen. Eine *isotrope Expansion* wird beschrieben durch das Geschwindigkeitsfeld  $\mathbf{v} = c\mathbf{x}$  mit einer positiven Konstanten  $c$ . Für die Gesamtspannung ergibt sich in diesem Fall  $\mathbf{T} = -(p - (3\lambda + 2\mu)c)\mathbf{I}$ . Es kann also von einem effektiven Druck

gesprochen werden, der durch Verminderung der thermodynamischen Druckes  $p$  um  $(3\lambda + 2\mu)c$  entsteht. Der Parameter

$$\mu_d = \frac{3\lambda + 2\mu}{3} = \lambda + \frac{2}{3}\mu$$

wird *Druckviskosität* genannt.

Außerdem betrachten wir eine einfache Scherströmung mit dem Geschwindigkeitsfeld  $\mathbf{v} = (\kappa x_2, 0, 0)$  mit der konstanten Schergeschwindigkeit  $\kappa$ . Der Spannungsvektor ist in diesem Fall gegeben durch  $\mathbf{t}(\mathbf{n}) = \mu\kappa(n_2, n_1, 0)$ . Das ist eine reine Scherkraft, und  $\mu$  wird als *Scherviskosität* bezeichnet. In Abhängigkeit von Druckviskosität und Scherviskosität schreiben wir den viskosen Spannungstensor auch als

$$\boldsymbol{\tau} = 2\mu \left( \mathbf{D} - \frac{1}{3}(\nabla \cdot \mathbf{v})\mathbf{I} \right) + \mu_d(\nabla \cdot \mathbf{v})\mathbf{I}.$$

Ein Beispiel für eine konstitutive Relation, die aus dem mikroskopischen Verhalten hergeleitet werden kann, ist das Verschwinden der Druckviskosität in idealen Gasen.

Die Impulserhaltungsgleichungen (2.31) mit dem soeben berechneten Spannungstensor sind die *Navier-Stokes-Gleichungen*

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} - \nabla p + \nabla(\lambda \nabla \cdot \mathbf{v}) + \nabla \cdot (2\mu \mathbf{D}). \quad (3.12)$$

Eine andere Schreibweise für konstante  $\lambda$  und  $\mu$  ist

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} - \nabla p + \nabla((\lambda + \mu)\nabla \cdot \mathbf{v}) + \mu \Delta \mathbf{v}, \quad (3.13)$$

in der die  $i$ -te Komponente des Vektors  $\Delta \mathbf{v}$  gegeben ist durch

$$(\Delta \mathbf{v})_i = \sum_j \frac{\partial^2 v_i}{\partial x_j^2}, \quad i = 1, 2, 3.$$

Unter der Annahme, daß die Volumskraft  $\mathbf{f}(\mathbf{x}, t)$  gegeben ist und für die Wärmeflußdichte das Newton-Fouriersche Abkühlungsgesetz  $\mathbf{q} = -\kappa \nabla T$  gilt, sind die Navier-Stokes-Gleichungen, kombiniert mit der Kontinuitätsgleichung (2.30) und der Energiegleichung (2.33), 5 skalare Differentialgleichungen für die Unbekannten  $\rho$ ,  $p$ ,  $T$ ,  $e$  und die 3 Komponenten von  $\mathbf{v}$ . Es fehlen also zwei Gleichungen, die das betrachtete Fluid näher bestimmen.

Eine für viele Fluide durch Experimente sehr gut abgesicherte Annahme ist, daß die spezifische innere Energie  $e$  nur von der Temperatur  $T$  abhängig ist. Insbesondere wird oft Konstanz der spezifischen Wärme bei konstantem Volumen  $c_v = de/dT$  angenommen, d.h.

$$e = c_v T + \text{const}. \quad (3.14)$$

Eine weitere Annahme, die für typische Flüssigkeiten wie z.B. Wasser verwendet wird, ist die der Inkompressibilität:

$$\frac{D\rho}{Dt} = 0$$

Sehr oft ist sogar die stärkere Annahme  $\rho = \text{const}$  gerechtfertigt. Man beachte, daß in diesem Fall die Kontinuitätsgleichung

$$\nabla \cdot \mathbf{v} = 0 \quad (3.15)$$

und die Navier-Stokes-Gleichungen

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} - \nabla p + \mu \Delta \mathbf{v} \quad (3.16)$$

ein geschlossenes System für die Bestimmung von Druck und Geschwindigkeit darstellen.

Für typische Gase wie z.B. Luft wird statt der Inkompressibilitätsannahme das ideale Gasgesetz

$$p = \rho R T \quad (3.17)$$

mit der Gaskonstanten  $R$  verwendet.

Schließlich stellen wir noch die Frage nach geeigneten Zusatzbedingungen. Befindet sich das Fluid in einem Behälter mit zeitabhängigem Rand, dann müssen die Differentialgleichungen für  $\mathbf{x} \in \Omega(t)$  betrachtet werden, wobei das Gebiet  $\Omega(t) \subset \mathbb{R}^3$  das Innere des Behälters zum Zeitpunkt  $t$  beschreibt. Nimmt man an, daß die Wände des Behälters undurchlässig sind, dann folgt die Randbedingung

$$\mathbf{v} \cdot \mathbf{n} = \mathbf{v}_w \cdot \mathbf{n} \quad \text{auf } \partial\Omega(t), \quad (3.18)$$

wobei  $\mathbf{n}$  der nach außen gerichtete Einheitsnormalvektor auf den Rand und  $\mathbf{v}_w$  die Geschwindigkeit der Wand ist.

Die Kontinuitätsgleichung kann als lineare Differentialgleichung erster Ordnung für die Dichte  $\rho$  betrachtet werden. Die Charakteristikenmethode zeigt, daß auf  $\partial\Omega(t)$  wegen (3.18) keine Randbedingungen für die Dichte vorgegeben werden müssen.

Zumindest im inkompressiblen Fall (3.16) ist es offensichtlich, daß die Navier-Stokes-Gleichungen ein System von parabolischen Gleichungen für die Komponenten des Geschwindigkeitsvektors sind. Die Randbedingungen (3.18) sind daher unzureichend, weil sie nur *eine* Geschwindigkeitskomponente auf dem Rand festlegen. Die Frage nach Randbedingungen für die Geschwindigkeit eines Fluids, das sich im Kontakt mit einer starren Wand befindet, ist äußerst schwierig. Trotzdem gibt es eine allgemein akzeptierte, einfache Antwort in Form der *Haftbedingung*

$$\mathbf{v} = \mathbf{v}_w \quad \text{auf } \partial\Omega(t). \quad (3.19)$$

Diese Bedingung bedeutet, daß zwischen der starren Wand und dem angrenzenden Fluid keine Relativbewegung stattfindet. (3.19) ist eine Dirichletbedingung für die Komponenten von  $\mathbf{v}$ .

Die Energieerhaltungsgleichung kann bei Verwendung von (3.14) und dem Newton-Fourierschen Abkühlungsgesetz als parabolische Gleichung für die Temperatur  $T$  aufgefaßt werden. Man muß also auch für die Temperatur Randbedingungen auf  $\partial\Omega(t)$  vorgeben. Alle für die Wärmeleitungsgleichung sinnvollen Arten von Randbedingungen sind hier denkbar.

Ein vollständig formuliertes mathematisches Problem erhält man bei zusätzlicher Angabe von Anfangsbedingungen, d.h.  $\rho$ ,  $\mathbf{v}$  und  $T$  müssen in  $\Omega(0)$  vorgegeben werden.

### 3.2. Einige exakte Lösungen für inkompressible viskose Strömungen

Wir betrachten ein Fluid mit konstanter vorgegebener Dichte, d.h. die Gleichungen (3.15), (3.16) mit  $\rho = \text{const.}$

Zunächst suchen wir eine stationäre Strömung zwischen zwei parallelen unendlichen Platten im Abstand  $d$ , von denen eine ruht und sich die andere mit konstanter Geschwindigkeit  $U$  bewegt.

Wir verwenden die Schreibweise  $\mathbf{x} = (x, y, z)$  und  $\mathbf{v} = (u, v, w)$  für die Komponenten des Orts- und des Geschwindigkeitsvektors. Das Koordinatensystem sei so gewählt, daß die  $y$ -Achse normal auf die Platten ist und die Bewegungsrichtung der bewegten Platte durch die positive  $x$ -Richtung gegeben ist. Wir vereinfachen das Problem, indem wir keine Volumskräfte berücksichtigen ( $\mathbf{f} = 0$ ) und nach einer zweidimensionalen Strömung suchen, d.h.  $w = 0$  und  $u, v$  und  $p$  sind unabhängig von  $z$ . Weiters scheint es plausibel, nach einer Lösung mit  $v = 0$  zu suchen. Mit all diesen Annahmen werden die Gleichungen (3.15), (3.16) zu

$$\begin{aligned} u_x &= 0, \\ 0 &= -p_x + \mu u_{yy}, \\ 0 &= -p_y. \end{aligned} \quad (3.20)$$

Aus der ersten Gleichung folgt, daß  $u$  nur von  $y$ , und aus der dritten, daß  $p$  nur von  $x$  abhängt. Damit impliziert die zweite Gleichung

$$p_x = \mu u_{yy} = \text{const.}$$

Wir sind an einer Strömung interessiert, die nur durch die Relativbewegung der Platten und nicht durch einen Druckabfall getrieben wird. Daher wählen wir  $p_x = 0$ . Aus den Haftbedingungen

$$u(0) = 0, \quad u(d) = U$$



folgt die Lösung

$$u = \frac{Uy}{d}, \quad v = w = 0, \quad p = \text{const.}$$

Das ist die schon oben beschriebene einfache Scherströmung. Sie wird auch als *ebene Couette-Strömung* bezeichnet. Eine andere exakte Lösung für eine ähnliche Situation ist die *Couette-Strömung* zwischen zwei rotierenden koaxialen Zylindern.

Wie oben berechnet, wirkt auf die ruhende Platte die Kraft  $(\mu U/d, 0, 0)$ . Zumindest im Prinzip könnte diese Formel verwendet werden, um die Viskosität  $\mu$  mit Hilfe eines Experimentes zu bestimmen. Typische Werte der Viskosität sind

$$\mu = 10^{-2} \text{gcm}^{-1}\text{s}^{-1} \quad \text{für Wasser}, \quad \mu = 2 \times 10^{-4} \text{gcm}^{-1}\text{s}^{-1} \quad \text{für Luft.}$$

Als zweites Beispiel suchen wir eine Strömung zwischen zwei ruhenden parallelen Platten, die von einem Druckabfall getrieben wird. Wir treffen dieselben Annahmen wie oben und erhalten wieder die Gleichungen (3.20). Allerdings setzen wir nun

$$p_x = -C < 0,$$

wobei das kleiner Werden des Druckes in positive  $x$ -Richtung eine Strömung in diese Richtung erwarten läßt. Aus der Gleichung  $C + \mu u_{yy} = 0$  und den Haftbedingungen  $u(0) = u(d) = 0$  folgt die Lösung

$$u = \frac{C}{2\mu}y(d-y), \quad v = w = 0, \quad p = -Cx + \text{const.}$$

Dieses parabolische Geschwindigkeitsprofil wird als *ebene Poiseuille-Strömung* bezeichnet. Der Name *Poiseuille-Strömung* wird für eine exakte Lösung verwendet, die die stationäre Strömung durch ein zylindrisches Rohr beschreibt, welche von einem axialen Druckgradienten getrieben wird.

Mit Hilfe dieser Lösungen kann die Viskosität aus Messungen der transportierten Masse des Fluids bestimmt werden. Für die ebene Poiseuille-Strömung ist die transportierte Masse pro Zeiteinheit und pro Längeneinheit in  $z$ -Richtung gegeben durch

$$\int_0^d \rho u \, dy = \frac{\rho C d^3}{12\mu}.$$

Als letztes Beispiel beschäftigen wir uns mit einer zeitabhängigen Strömung. Wir nehmen an, der Halbraum  $y > 0$  wäre gefüllt mit einem Fluid, das für  $t \leq 0$  ruht. Weiters sei angenommen, daß die das Fluid begrenzende unendliche Platte sich zum Zeitpunkt  $t = 0$  in ihrer Ebene mit der konstanten Geschwindigkeit  $U$  zu bewegen beginnt. Es ist anzunehmen, daß auch das Fluid durch Reibungseffekte in Bewegung gesetzt wird. Wie geschieht das?

Sei die Bewegungsrichtung der Platte die  $x$ -Richtung. Dann suchen wir eine Lösung mit  $v = w = 0$  und  $u = u(y, t)$ . Aus den Navier-Stokes-Gleichungen (3.16) folgt dann

$$\rho u_t = -p_x + \mu u_{yy}, \quad p_y = p_z = 0.$$

Daraus folgt wie oben, daß  $p_x$  konstant ist. Wir sind nicht an den Effekten eines Druckgefälles interessiert und setzen  $p_x = 0$ . Wir haben daher die Wärmeleitungsgleichung

$$u_t = \nu u_{yy} \tag{3.21}$$

mit der *kinematischen Viskosität*  $\nu = \mu/\rho$  zu betrachten. Als Anfangsbedingung ergibt sich

$$u(y, 0) = 0 \tag{3.22}$$

und als Randbedingung die Haftbedingung

$$u(0, t) = U. \tag{3.23}$$

Die dimensionslose Geschwindigkeit  $u^* = u/U$  ist Lösung des Anfangs-Randwertproblems

$$\begin{aligned}u_t^* &= \nu u_{yy}^*, \\u^*(y, 0) &= 0, \quad u^*(0, t) = 1.\end{aligned}$$

Da  $u^*$  dimensionslos ist, kann es nur von dimensionslosen Kombinationen von  $\nu$ ,  $y$  und  $t$  abhängen. Die Einheit der kinematischen Viskosität  $\nu$  ist  $\text{cm}^2\text{s}^{-1}$ , woraus folgt, daß  $u^*$  als Funktion der dimensionslosen Variablen

$$\eta = \frac{y}{2\sqrt{\nu t}}$$

geschrieben werden kann:  $u^* = F(\eta)$ . Für  $F$  folgt aus der Wärmeleitungsgleichung

$$F'' + 2\eta F' = 0.$$

Aus den Anfangs- und Randbedingungen werden

$$F(0) = 1, \quad F(\infty) = 0.$$

Die Lösung dieses Randwertproblems ist gegeben durch die komplementäre Fehlerfunktion

$$F(\eta) = \text{erfc}(\eta) = \frac{2}{\sqrt{\pi}} \int_{\eta}^{\infty} e^{-s^2} ds.$$

Die Geschwindigkeit des Fluids ist also

$$u = U \text{erfc} \left( \frac{y}{2\sqrt{\nu t}} \right).$$

Diese Lösung bedeutet, daß an einem festen Abstand von der Platte die Geschwindigkeit des Fluids für  $t \rightarrow \infty$  gegen die Geschwindigkeit der Platte konvergiert. Der Abstand von der Platte, an dem die Geschwindigkeit des Fluids einen bestimmten Prozentsatz der Geschwindigkeit der Platte erreicht, ist proportional zu  $\sqrt{\nu t}$ .

### 3.3. Reibungsfreie ideale Gase—Entropie—Rotation

Um die relative Größe verschiedener Terme in den Navier-Stokes-Gleichungen abschätzen zu können, skalieren wir die Impulsgleichungen (3.13) für ein ideales Gas, d.h. mit  $\mu_d = 0$ :

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p + \frac{\mu}{3} \nabla(\nabla \cdot \mathbf{v}) + \mu \Delta \mathbf{v} \quad (3.24)$$

Der Einfachheit halber wurden Volumskräfte vernachlässigt und Konstanz der Viskosität angenommen. Wir wählen die Referenzgeschwindigkeit  $U_0$ , -länge  $L_0$  und -dichte  $\rho_0$ . Die Zeit skalieren wir mit  $L_0/U_0$  und den Druck mit  $\rho_0 U_0^2$ . Verwenden wir in der skalierten Gleichung dieselben Variablennamen wie in (3.24), so ergibt sich

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p + \frac{1}{3\text{Re}} \nabla(\nabla \cdot \mathbf{v}) + \frac{1}{\text{Re}} \Delta \mathbf{v}, \quad (3.25)$$

mit der dimensionslosen *Reynoldszahl*

$$\text{Re} = \frac{\rho_0 U_0 L_0}{\mu}.$$

Eine typische Dichte für das fast ideale Gas Luft ist  $\rho_0 = 10^{-3} \text{g cm}^{-3}$ . Betrachtet man als Beispiel die Luftströmung um einen gehenden Menschen, dann sind  $U_0 = 100 \text{cm s}^{-1}$  und  $L_0 = 100 \text{cm}$  sinnvolle Referenzgrößen. Mit dem Wert  $\mu = 2 \times 10^{-4} \text{g cm}^{-1} \text{s}^{-1}$  für die Viskosität der Luft ergibt sich  $\text{Re} = 5 \times 10^4$ .

Nun betrachten wir die Energiegleichung (2.33) mit den Annahmen  $e = c_v T + \text{const}$  und  $\mathbf{q} = -\kappa \nabla T$ :

$$\rho \frac{D}{Dt} \left( \frac{|\mathbf{v}|^2}{2} + c_v T \right) = \nabla \cdot (\kappa \nabla T + \mathbf{T} \cdot \mathbf{v})$$

Wir wählen die Referenztemperatur  $U_0^2/c_v$  und skalieren sonst wie oben. Die dimensionslose Version der obigen Gleichung ist dann

$$\rho \frac{D}{Dt} \left( \frac{|\mathbf{v}|^2}{2} + T \right) = \nabla \cdot \left( \frac{1}{\text{RePr}} \nabla T + \mathbf{T} \cdot \mathbf{v} \right) \quad (3.26)$$

mit dem Spannungstensor

$$\mathbf{T} = -p\mathbf{I} + \frac{2}{\text{Re}} \mathbf{D} - \frac{2}{3\text{Re}} (\nabla \cdot \mathbf{v})\mathbf{I}$$

und der *Prandtlzahl*

$$\text{Pr} = \frac{\mu c_v}{\kappa}$$

als zusätzlichen dimensionslosen Parameter. Die Prandtlzahl gibt das Verhältnis zwischen der Stärke der Reibungseffekte und der Wärmeleitung an. Für Gase ist die Prandtlzahl gewöhnlich von der Größenordnung 1. Ein typischer Wert für Luft ist  $\text{Pr} = 0,72$ .

Unsere Überlegungen legen den Schluß nahe, daß in vielen Fällen die Effekte von Reibung und Wärmeleitung vernachlässigt werden können. Wir werden uns daher im Folgenden mit den *Eulergleichungen*

$$\begin{aligned} \frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} &= 0, \\ \rho \frac{D\mathbf{v}}{Dt} + \nabla p &= 0, \\ \rho \frac{D}{Dt} \left( \frac{|\mathbf{v}|^2}{2} + e \right) + \nabla \cdot (p\mathbf{v}) &= 0, \end{aligned} \quad (3.27)$$

mit den konstitutiven Relationen  $e = c_v T + \text{const}$  und  $p = \rho RT$  für ideale Gase beschäftigen.

Als nächsten Schritt wollen wir einige Überlegungen aus der Thermodynamik verwenden. Wir betrachten einen Behälter mit beweglichen Wänden, gefüllt mit einem räumlich homogenen idealen Gas. Der Zustand des Gases wird dann durch Angabe von zwei der drei *Zustandsvariablen* Druck  $P$ , Volumen  $V$  und Temperatur  $T$  beschrieben, die durch die Gleichung  $P = RT/V$  zusammenhängen. Die in dem Gas enthaltene Energie  $E$  ist bestimmt durch  $E = C_V T + \text{const}$ .

Der *erste Hauptsatz der Thermodynamik* sagt aus, daß die Differenz der Energien zweier Zustände gegeben ist durch die Differenz aus der bei der Zustandsänderung von außen zugeführten Wärmeenergie und der von dem Gas bei der Zustandsänderung geleisteten Arbeit. Eine differentielle Form dieser Aussage ist

$$dE = dQ - dW .$$

Es ist leicht einzusehen, daß für das Differential der Arbeit die Formel  $dW = P dV$  gilt. Die obige Gleichung kann daher auch geschrieben werden als

$$dQ = C_V dT + P dV . \quad (3.28)$$

Diese Form erklärt auch, warum die Konstante  $C_V$  als *spezifische Wärme bei konstantem Volumen* bezeichnet wird. Weiters gilt wegen des idealen Gasgesetzes  $P dV = R dT - V dP$  und daher

$$dQ = C_P dT - V dP$$

mit der *spezifischen Wärme bei konstantem Druck*  $C_P = C_V + R$ .

Division von (3.28) durch  $T$  ergibt

$$\frac{dQ}{T} = C_V \frac{dT}{T} + R \frac{dV}{V} .$$

Man beachte, daß der Ausdruck auf der rechten Seite ein totales Differential ist. Diese Darstellung motiviert die Definition der *Entropie*

$$S = C_V \ln T + R \ln V + \text{const} ,$$

die konstant ist für ein System, dem keine Wärme zugeführt wird.

Nach diesem Exkurs in die Thermodynamik definieren wir die Entropie für ein nicht homogenes Gas lokal durch

$$s = c_v \ln T - R \ln \rho + \text{const} .$$

Diese Relation kann auch in der Form

$$p = k \rho^\gamma e^{s/c_v} \quad (3.29)$$

mit einer Konstanten  $k$  und dem Verhältnis der spezifischen Wärmen  $\gamma = (c_v + R)/c_v$  geschrieben werden. Mit Hilfe der Eulergleichungen kann man zeigen, daß die Materialableitung von  $s$  verschwindet:

$$\frac{Ds}{Dt} = 0$$

Das bedeutet, daß die Entropie eines sich mit der Strömung bewegenden infinitesimalen Volumenelementes konstant ist. Diese Tatsache wird im Folgenden von Bedeutung sein.

Wir bezeichnen das Vektorfeld

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} := \left( \frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}, \frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}, \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right)$$

als *Rotation* des Geschwindigkeitsfeldes. Diese Bezeichnung wollen wir im Folgenden motivieren. Wir betrachten ein zum Zeitpunkt  $t = t_1$  in der  $x$ - $y$ -Ebene liegendes Rechteck  $ABCD$ , dessen Eckpunkte  $A$ ,  $B$  und  $C$  die Koordinaten  $(0, 0)$ ,  $(\Delta x, 0)$  bzw.  $(0, \Delta y)$  haben (siehe Abb. 3.1). Wir nehmen an, die Eckpunkte des Rechtecks bewegt sich bis zum Zeitpunkt  $t_2 = t_1 + \Delta t$  mit der Strömung, wobei  $\Delta x$ ,  $\Delta y$  und  $\Delta t$  klein sind. Die Projektionen der Punkte zum Zeitpunkt  $t_2$  in die  $x$ - $y$ -Ebene bezeichnen wir mit  $A'$ ,  $B'$  und  $C'$ .

Abb. 3.1: Drehung eines Rechtecks um die  $z$ -Achse

Sie haben näherungsweise die Koordinaten  $(\Delta t v_1(0, 0), \Delta t v_2(0, 0))$ ,  $(\Delta x + \Delta t v_1(\Delta x, 0), \Delta t v_2(\Delta x, 0))$  bzw.  $(\Delta t v_1(0, \Delta y), \Delta y + \Delta t v_2(0, \Delta y))$ .

Die Winkelgeschwindigkeit der Seite  $AB$  des Rechtecks um die  $z$ -Achse ist näherungsweise gegeben durch

$$\frac{\alpha}{\Delta t} \approx \frac{v_2(\Delta x, 0) - v_2(0, 0)}{\Delta x} \approx \frac{\partial v_2}{\partial x}.$$

Analog gilt für die Winkelgeschwindigkeit der Seite  $AC$  um die  $z$ -Achse

$$\frac{\beta}{\Delta t} \approx -\frac{\partial v_1}{\partial y}.$$

Die mittlere Winkelgeschwindigkeit um die  $z$ -Achse der beiden Seiten eines infinitesimalen Rechtecks ist also die Hälfte der  $z$ -Komponente der Rotation  $\boldsymbol{\omega}$ . Analog kann die  $x$ - bzw. die  $y$ -Komponente von  $\boldsymbol{\omega}$  als Maß für die Winkelgeschwindigkeit um die  $x$ - bzw.  $y$ -Achse interpretiert werden.

Für das Weitere benötigen wir eine Vektoridentität, die durch einfaches Nachrechnen verifiziert werden kann. Für den zweiten Term in der Materialableitung der Geschwindigkeit

$$\frac{D\mathbf{v}}{Dt} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v}$$

gilt die Formel

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = \nabla \left( \frac{|\mathbf{v}|^2}{2} \right) - \mathbf{v} \times \boldsymbol{\omega}. \quad (3.30)$$

Damit können wir die Impulsgleichung in (3.27) in der Form

$$\frac{\partial \mathbf{v}}{\partial t} + \nabla \left( \frac{|\mathbf{v}|^2}{2} \right) - \mathbf{v} \times \boldsymbol{\omega} + \frac{1}{\rho} \nabla p = 0 \quad (3.31)$$

schreiben. Aus der Formel (3.29) für die Darstellung des Druckes mit Hilfe der Entropie folgt

$$\frac{1}{\rho} \nabla p = \frac{\gamma}{\gamma - 1} \nabla \left( \frac{p}{\rho} \right) + \frac{p}{c_v} \left( 1 - \frac{\gamma}{(\gamma - 1)\rho} \right) \nabla s. \quad (3.32)$$

Aus diesen Formeln folgt eine wichtige Relation für stationäre Strömungen. Klarerweise gilt  $\mathbf{v} \cdot (\mathbf{v} \times \boldsymbol{\omega}) = 0$ . Außerdem gilt im stationären Fall  $\mathbf{v} \cdot \nabla s = 0$ . Verwenden wir daher (3.32) in der stationären Version von (3.31) und bilden das Skalarprodukt dieser Gleichung mit  $\mathbf{v}$ , so erhalten wir

$$\mathbf{v} \cdot \nabla \left( \frac{|\mathbf{v}|^2}{2} + \frac{\gamma p}{(\gamma - 1)\rho} \right) = 0.$$

Daraus folgt, daß entlang von Teilchenbahnen die *Bernoulli-Gleichung*

$$\frac{|\mathbf{v}|^2}{2} + \frac{\gamma p}{(\gamma - 1)\rho} = \text{const} \quad (3.33)$$

gilt.

### 3.4. Kompressionswellen in isentropen Gasen

In diesem Abschnitt beschäftigen wir uns mit reibungsfreien Gasen, in denen die Entropie einen konstanten Wert annimmt. Aus (3.29) folgt daher  $p = K\rho^\gamma$  bzw.—allgemeiner— $p = p(\rho)$ . Diese Relation ersetzt die Energiegleichung, und wir betrachten daher nur die ersten beiden Gleichungen in (3.27):

$$\begin{aligned}\frac{D\rho}{Dt} + \rho\nabla \cdot \mathbf{v} &= 0, \\ \rho\frac{D\mathbf{v}}{Dt} + p'(\rho)\nabla\rho &= 0\end{aligned}\tag{3.34}$$

Diese Gleichungen für ein reibungsfreies isentropes Gas werden oft als *p-System* bezeichnet. Die Ableitung des Druckes nach der Dichte hat die Dimension des Quadrates einer Geschwindigkeit. Wir führen daher die Bezeichnung

$$c^2 = p'(\rho)$$

ein.

Wir wollen uns zunächst mit Wellen mit kleiner Amplitude beschäftigen, d.h. wir betrachten kleine Störungen einer homogenen bewegungslosen Atmosphäre. Hat die ungestörte Atmosphäre die Dichte  $\rho_0$ , so machen wir den Ansatz

$$\frac{\rho}{\rho_0} = 1 + \varepsilon\sigma$$

mit  $0 < \varepsilon \ll 1$ . Weiters skalieren wir (3.34) durch Wahl der Referenzgrößen  $U_0$ ,  $L_0$  und  $t_0$  für die Geschwindigkeit  $\mathbf{v}$ , Länge  $\mathbf{x}$  und Zeit  $t$ . Die Geschwindigkeit  $c$  skalieren wir mit ihrem Wert  $c_0 = p'(\rho_0)$  in der ungestörten Atmosphäre. Balanciert man im skalierten Problem die führenden Terme, so ergeben sich die Forderungen

$$\frac{U_0 t_0}{\varepsilon L_0} = 1, \quad \frac{\varepsilon t_0}{U_0 L_0} c_0^2 = 1$$

für die Wahl von  $U_0$ ,  $L_0$  und  $t_0$ . Daraus folgt

$$\frac{L_0}{t_0} = c_0, \quad U_0 = \varepsilon c_0.\tag{3.35}$$

Mit diesen Annahmen hat das skalierte Problem die Form

$$\begin{aligned}\frac{\partial\sigma}{\partial t} + \varepsilon\mathbf{v} \cdot \nabla\sigma + (1 + \varepsilon\sigma)\nabla \cdot \mathbf{v} &= 0, \\ (1 + \varepsilon\sigma) \left( \frac{\partial\mathbf{v}}{\partial t} + \varepsilon(\mathbf{v} \cdot \nabla)\mathbf{v} \right) + c^2\nabla\sigma &= 0.\end{aligned}$$

Der Grenzübergang  $\varepsilon \rightarrow 0$  führt auf die Wellengleichung

$$\frac{\partial^2\sigma}{\partial t^2} = \Delta\sigma$$

für die Störung in der Dichte. Kehrt man zu dimensionsbehafteter Länge bzw. Zeit zurück, dann erhält sie die Form

$$\frac{\partial^2\sigma}{\partial t^2} = c_0^2\Delta\sigma.\tag{3.36}$$

In diesem Zusammenhang ist die Wellengleichung die Grundgleichung der Akustik. Daher erklärt sich auch die Bezeichnung *Schallgeschwindigkeit* für die Geschwindigkeit  $c = \sqrt{p'(\rho)}$ , mit der sich kleine Störungen in einer Atmosphäre mit der Dichte  $\rho$  ausbreiten.

Explizite Lösungen der dreidimensionalen Wellengleichung kann man mit Symmetrieanahmen berechnen. Für Lösungen mit einer Kugelsymmetrie gilt die Gleichung

$$\frac{\partial^2\sigma}{\partial t^2} = c_0^2 \left( \frac{\partial^2\sigma}{\partial r^2} + \frac{2}{r} \frac{\partial\sigma}{\partial r} \right),$$

die mit der Substitution  $\sigma = \beta/r$  auf die eindimensionale Wellengleichung

$$\frac{\partial^2 \beta}{\partial t^2} = c_0^2 \frac{\partial^2 \beta}{\partial r^2}$$

transformiert werden kann. Aus der allgemeinen Lösung dieser Gleichung erhält man alle kugelsymmetrischen Lösungen von (3.36):

$$\sigma = \frac{1}{r} (f(r - c_0 t) + g(r + c_0 t))$$

Sie bestehen aus wandernden Wellen, die sich mit Schallgeschwindigkeit fortbewegen und deren Amplitude proportional zum Kehrwert der Entfernung ist.

Der Störungsparameter  $\varepsilon$  ist wegen (3.35) der Quotient aus der charakteristischen Strömungsgeschwindigkeit und der Schallgeschwindigkeit. Dieser Quotient  $M = U/c_0$  wird als *Machzahl* bezeichnet. Die Gleichung (3.36) kann daher als Näherungsproblem für kleine Machzahl angesehen werden.

Die wesentlichste Eigenschaft der Näherungsgleichung (3.36) ist ihre Linearität. Im Folgenden sollen nichtlineare Effekte anhand spezieller Lösungen des  $p$ -Systems (3.34) demonstriert werden. Wir betrachten das eindimensionale  $p$ -System

$$\begin{aligned} \frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + \rho \frac{\partial u}{\partial x} &= 0, \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{c^2}{\rho} \frac{\partial \rho}{\partial x} &= 0, \end{aligned} \tag{3.37}$$

und suchen nach Lösungen in der Form *einfacher Wellen*

$$\rho = \rho(\alpha), \quad u = u(\alpha) \quad \text{mit } \alpha = \alpha(x, t).$$

Mit diesem Ansatz erhalten die Gleichungen (3.37) die Form

$$\begin{aligned} \rho'(\alpha) \left( \frac{\partial \alpha}{\partial t} + u \frac{\partial \alpha}{\partial x} \right) + \rho u'(\alpha) \frac{\partial \alpha}{\partial x} &= 0, \\ u'(\alpha) \left( \frac{\partial \alpha}{\partial t} + u \frac{\partial \alpha}{\partial x} \right) + \frac{c^2}{\rho} \rho'(\alpha) \frac{\partial \alpha}{\partial x} &= 0. \end{aligned}$$

Diese Gleichungen können als lineares Gleichungssystem zur Bestimmung von  $\rho'$  und  $u'$  bzw. von  $\partial\alpha/\partial x$  und  $\partial\alpha/\partial t$  aufgefaßt werden. Für die Existenz nichttrivialer Lösungen sind die Bedingungen

$$\frac{c^2}{\rho^2} (\rho')^2 = (u')^2, \quad \left( \frac{\partial \alpha}{\partial t} + u \frac{\partial \alpha}{\partial x} \right)^2 = c^2 \left( \frac{\partial \alpha}{\partial x} \right)^2$$

notwendig. Aus der ersten Bedingung folgt offensichtlich

$$u = \pm \int \frac{c(\rho)}{\rho} d\rho. \tag{3.38}$$

Die zweite Bedingung kann als quasilineare Differentialgleichung erster Ordnung für  $\alpha$  umgeschrieben werden:

$$\frac{\partial \alpha}{\partial t} + (u(\alpha) \pm c(\alpha)) \frac{\partial \alpha}{\partial x} = 0$$

Mit Hilfe des Charakteristikenverfahrens zeigt man, daß die allgemeine Lösung dieser Gleichung gegeben ist durch  $\alpha = F(x - (u \pm c)t)$  mit einer beliebigen Funktion  $F$ .

Mit der konstitutiven Relation  $p = K\rho^\gamma$  für ein ideales isentropes Gas folgt aus (3.38) die Beziehung

$$u = \pm \frac{2}{\gamma - 1} (c - c_0).$$

Abb. 3.2: Charakteristiken für a) monoton wachsendes  $u_0$ , b) monoton fallendes  $u_0$ .

Aus dieser Gleichung läßt sich  $c$  in Abhängigkeit von  $u$  berechnen. Wir erhalten schließlich

$$u = F \left[ x - \left( \frac{\gamma + 1}{2} u \pm c_0 \right) t \right].$$

Betrachten wir Anfangsbedingungen für  $u$  in der Form  $u(x, 0) = u_0(x)$ , dann gilt  $F = u_0$ . Durch jeden Punkt  $(x_0, 0)$  geht eine Charakteristik in Form der Geraden

$$x - x_0 = \left( \frac{\gamma + 1}{2} u_0(x_0) \pm c_0 \right) t,$$

auf der  $u$  den Konstanten Wert  $u_0(x_0)$  hat. Abhängig von  $u_0$  treten zwei typische Situationen auf: Ist  $u_0$  monoton wachsend, so laufen die Charakteristiken mit fortschreitender Zeit auseinander. Die Lösung des Anfangswertproblems ist für alle Zeiten eindeutig bestimmt. Für monoton fallendes  $u_0$  hingegen laufen die Charakteristiken zusammen (Abb. 3.2). Das bedeutet daß es einen Zeitpunkt  $t_0 > 0$  gibt, an dem zwei Charakteristiken, die verschiedene  $u$ -Werte tragen, einander schneiden.

Nach dem Zeitpunkt  $t_0$  kann daher keine stetige Lösung des Anfangswertproblems existieren. Es ist allerdings möglich, die Lösung über den Zeitpunkt  $t_0$  hinaus fortzusetzen, wenn man die Differentialgleichungen durch geeignete schwache Formulierungen ersetzt, die Lösungen mit Unstetigkeiten zulassen. Eine derartige Lösung besitzt eine Sprungunstetigkeit entlang einer Kurve in der  $x$ - $t$ -Ebene und wird *Stoßwelle* genannt.



Abb. 3.3: Zweidimensionale Strömung um eine Tragfläche

### 3.5. Potentialströmungen um Tragflächen

Als Modell für Strömungen um längliche Körper, deren Querschnitt schwach variiert, betrachten wir zweidimensionale Strömungen um Zylinder mit verschiedenen Querschnitten. Insbesondere ist unser Ziel die Behandlung von Querschnitten wie in Abb. 3.3 skizziert, die typisch für Tragflächen sind.

Wir verwenden ein Koordinatensystem, in dem sich die Tragfläche in Ruhe befindet, und nehmen an, daß die Strömung weit weg von der Tragfläche eine konstante Geschwindigkeit hat, die normal auf die Erzeugenden des Zylinders ist. Der Betrag dieser Geschwindigkeit sei  $U$ . Weiters habe die Atmosphäre weit weg von der Tragfläche konstante Dichte  $\rho_\infty$  und Druck  $p_\infty$  (siehe Abb. 3.3).

Das Koordinatensystem sei so gewählt, daß die  $x$ - $y$ -Ebene normal auf die Erzeugenden steht und die positive  $x$ -Richtung die Richtung der Strömung weg von der Tragfläche ist. Wir beschränken uns auf zweidimensionale Strömungen, d.h. die  $z$ -Komponente des Geschwindigkeitsfeldes verschwindet. Es gilt

$$\mathbf{v} = (u, v) \rightarrow (U, 0), \quad \rho \rightarrow \rho_\infty, \quad p \rightarrow p_\infty, \quad \text{für } |\mathbf{x}| = \sqrt{x^2 + y^2} \rightarrow \infty. \quad (3.39)$$

Das Innere der Querschnittsfläche bezeichnen wir mit  $\Omega$ . Offensichtlich muß die Randbedingung

$$\mathbf{v} \cdot \mathbf{n} = 0 \quad \text{auf } \partial\Omega \quad (3.40)$$

erfüllt sein. Hier bezeichnet  $\mathbf{n}$  den bezüglich  $\Omega$  nach innen gerichteten Einheitsnormalvektor auf den Rand.

Wir interessieren uns für stationäre Strömungen. Die in Abb. 3.3 eingezeichneten Kurven sind Teilchenbahnen oder *Stromlinien*, d.h. Lösungen der Differentialgleichung  $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$ .

Wegen der Bedingungen an  $\infty$  nimmt die Entropie dort einen konstanten Wert an. Da alle Stromlinien in  $\mathbb{R}^2 \setminus \Omega$  aus  $\infty$  kommen und die Entropie entlang von Stromlinien konstant bleibt, haben wir es mit einer isentropen Strömung zu tun. Es gilt die konstitutive Relation (siehe (3.29))

$$\frac{p}{p_\infty} = \left( \frac{\rho}{\rho_\infty} \right)^\gamma \quad (3.41)$$

und die Bernoulli-Gleichung (3.33) erhält die Form

$$\frac{|\mathbf{v}|^2}{2} + \frac{\gamma p}{(\gamma - 1)\rho} = \frac{U^2}{2} + \frac{\gamma p_\infty}{(\gamma - 1)\rho_\infty}. \quad (3.42)$$

Damit folgt aus der stationären Version der Impulsgleichung (3.31) und aus (3.32) die Beziehung

$$\mathbf{v} \times \boldsymbol{\omega} = 0. \quad (3.43)$$

Für die betrachteten zweidimensionalen Strömungen hat die Rotation nur eine nichtverschwindende Komponente. Die Gleichung (3.43) impliziert daher, daß die Strömung *wirbelfrei* ist, d.h. daß  $\boldsymbol{\omega} = \nabla \times \mathbf{v} = 0$  gilt. Jedes wirbelfreie Vektorfeld ist ein Gradientenfeld; es gibt daher ein *Geschwindigkeitspotential*  $\phi$  mit

$$\mathbf{v} = \nabla \phi. \quad (3.44)$$

Wirbelfreie Strömungen werden deshalb auch *Potentialströmungen* genannt.

Für einen stationären Zustand können wir das  $p$ -System (3.34) schreiben als

$$\begin{aligned}\rho \nabla \cdot \mathbf{v} + \mathbf{v} \cdot \nabla \rho &= 0, \\ \nabla \frac{|\mathbf{v}|^2}{2} + \frac{c^2}{\rho} \nabla \rho &= 0.\end{aligned}$$

Bilden wir das Skalarprodukt der zweiten Gleichung mit  $\mathbf{v}$  und verwenden wir die erste, so ergibt sich

$$c^2 \nabla \cdot \mathbf{v} = \mathbf{v} \cdot \nabla \frac{|\mathbf{v}|^2}{2}. \quad (3.45)$$

Zusammen mit der Bernoulli-Gleichung (3.42) bildet diese Gleichung ein System zur Bestimmung des Potentials  $\phi$  und des Druckes  $p$  (bzw. der Dichte  $\rho$ ).

Wir wollen das System (3.42), (3.45) skalieren, indem wir die Werte  $U$ ,  $p_\infty$  und  $\rho_\infty$  von Geschwindigkeit, Druck und Dichte an  $\infty$  als Referenzgrößen wählen. Die Referenzgröße für die Länge hat offensichtlich keinen Einfluß auf die skalierten Gleichungen. Wir wählen den Durchmesser des Querschnittes  $\Omega$ :

$$L = d(\Omega) := \sup_{(\mathbf{x}_1, \mathbf{x}_2) \in \Omega^2} |\mathbf{x}_1 - \mathbf{x}_2|$$

In der skalierten Version von (3.42), (3.45) verwenden wir für die dimensionslosen Variablen dieselben Namen wie für die dimensionsbehafteten und als dimensionslosen Parameter die Machzahl

$$M = \frac{U}{c_\infty} \quad \text{mit } c_\infty = \sqrt{p'(\rho_\infty)} = \frac{\gamma p_\infty}{\rho_\infty}.$$

Die skalierten Gleichungen sind

$$\begin{aligned}M^2 \frac{|\mathbf{v}|^2 - 1}{2} + \frac{1}{\gamma - 1} \left( p^{\frac{\gamma-1}{\gamma}} - 1 \right) &= 0, \\ p^{\frac{\gamma-1}{\gamma}} \nabla \cdot \mathbf{v} &= M^2 \mathbf{v} \cdot \nabla \frac{|\mathbf{v}|^2}{2}.\end{aligned} \quad (3.46)$$

Verwenden wir (3.44), so ist für gegebenen Druck die zweite Gleichung in (3.46) eine nichtlineare partielle Differentialgleichung zweiter Ordnung für das Potential. Um das Problem zu vereinfachen, beschränken wir uns auf *langsame Strömungen*, d.h. die Strömungsgeschwindigkeit sei klein im Vergleich zur Schallgeschwindigkeit. Die Machzahl ist also ein kleiner Parameter, und wir interessieren uns für Approximationen für die Lösung, die im Grenzfall  $M \rightarrow 0$  gültig sind.

Die skalierte Bernoulli-Gleichung zeigt, daß für  $M \rightarrow 0$  der skalierte Druck gegen den konstanten Wert 1 konvergiert. Wir machen daher den Ansatz

$$\mathbf{v} = \mathbf{v}_0 + O(M^2), \quad p = 1 + M^2 p_1 + O(M^4).$$

Einsetzen in (3.46) und Koeffizientenvergleich liefert

$$\frac{|\mathbf{v}_0|^2 - 1}{2} + \frac{p_1}{\gamma} = 0, \quad \nabla \cdot \mathbf{v}_0 = 0.$$

Schreibt man dieses vereinfachte System wieder in Abhängigkeit der dimensionsbehafteten Größen und verwendet man das Geschwindigkeitspotential, so erhält man

$$\rho_\infty \frac{|\nabla \phi|^2}{2} + p = \rho_\infty \frac{U^2}{2} + p_\infty, \quad \Delta \phi = 0. \quad (3.47)$$

Das Potential ist eine Lösung der Laplace-Gleichung. Als Zusatzbedingungen erhält man

$$\begin{aligned}\nabla\phi &\rightarrow (U, 0) \quad \text{für } |x| \rightarrow \infty, \\ \nabla\phi \cdot \mathbf{n} &= 0 \quad \text{auf } \partial\Omega.\end{aligned}$$

Die Annahme der langsamen Strömung hat zwei wesentliche Vereinfachungen bewirkt. Zunächst ist das Problem zur Bestimmung von  $\phi$  und  $p$  entkoppelt.  $\phi$  kann aus dem Randwertproblem für die Laplace-Gleichung berechnet und danach  $p$  aus der Bernoulli-Gleichung ermittelt werden. Außerdem ist nur mehr ein lineares Problem zu lösen. Die hier gemachten Annahmen entsprechen denen von Abschnitt 3.4 (Kompressionswellen mit kleiner Amplitude).

Als Beispiel betrachten wir die Strömung um einen Kreiszyylinder. Der Querschnitt  $\Omega$  ist ein Kreis mit Radius  $a$  und Mittelpunkt im Ursprung der  $x$ - $y$ -Ebene. Nach Transformation auf Polarkoordinaten  $(r, \theta)$  hat das Problem für das Geschwindigkeitspotential die Form

$$\phi_{rr} + \frac{1}{r}\phi_r + \frac{1}{r^2}\phi_{\theta\theta} = 0, \quad \text{für } r > a$$

mit den Randbedingungen

$$\begin{aligned}\phi_r(a, \theta) &= 0, \\ \phi_r &\rightarrow U \cos \theta, \quad \frac{1}{r}\phi_\theta \rightarrow -U \sin \theta, \quad \text{für } r \rightarrow \infty\end{aligned}$$

und den Stetigkeitsbedingungen für die Geschwindigkeit

$$\phi_r(r, \theta + 2\pi) = \phi_r(r, \theta), \quad \phi_\theta(r, \theta + 2\pi) = \phi_\theta(r, \theta).$$

Mit der Methode der Separation von Variablen läßt sich zeigen, daß

$$\phi(r, \theta) = -A\theta + U \left( r + \frac{a^2}{r} \right) \cos \theta$$

für jedes reelle  $A$  eine Lösung ist. Zusätzliche Lösungen können durch Addition von Konstanten erzeugt werden. Diese beschreiben aber klarerweise dieselbe Strömung. Ein echter Mangel an Eindeutigkeit ist allerdings durch die Möglichkeit der beliebigen Wahl von  $A$  gegeben. Im folgenden wollen wir uns auf Situationen beschränken, in denen die Parameter  $U$  und  $A$  nicht negativ sind.

Um die berechnete Strömung besser beschreiben zu können, bedienen wir uns der Tatsache, daß das berechnete Potential als harmonische Funktion als Realteil der komplexen analytischen Funktion

$$F(z) = Ai \ln \frac{z}{a} + U \left( z + \frac{a^2}{z} \right)$$

mit  $z = x + iy$  aufgefaßt werden kann. Umgekehrt kann für jede analytische Funktion

$$F(z) = \phi(x, y) + i\psi(x, y)$$

der Realteil  $\phi$  als Geschwindigkeitspotential interpretiert werden. In diesem Zusammenhang wird die Funktion  $F$  als *komplexes Geschwindigkeitspotential* bezeichnet. Aus den Cauchy-Riemannschen Differentialgleichungen

$$u = \phi_x = \psi_y, \quad v = \phi_y = -\psi_x$$

folgt, daß die Teilchenbahnen bzw. Stromlinien durch die Gleichung

$$\psi(x, y) = \text{const}$$

Abb. 3.4: Strömung um einen Kreiszyylinder für 1)  $A > 2aU$ , 2)  $A < 2aU$

gegeben sind. Der Imaginärteil  $\psi$  von  $F$  (die konjugierte harmonische Funktion zu  $\phi$ ) wird *Stromfunktion* genannt.

Bemerkenswert ist außerdem, daß der Geschwindigkeitsvektor aus der *komplexen Geschwindigkeit*

$$W = F' = u - iv$$

berechnet werden kann.

Für die Strömung um den Kreiszyylinder ist die Stromfunktion gegeben durch

$$\psi = A \ln \frac{r}{a} + U \left( r - \frac{a^2}{r} \right) \sin \theta .$$

Die Randbedingungen auf  $\partial\Omega$  bedeuten, daß  $\partial\Omega$  eine Stromlinie ist. Für den Kreiszyylinder ist  $\partial\Omega$  durch  $r = a$  gegeben, was der Stromlinie  $\psi = 0$  entspricht.

Um die Bedeutung des Parameters  $A$  zu verstehen, betrachten wir zunächst die Situation  $U = 0$ . In diesem Fall sind die Stromlinien Kreise. Es handelt sich um eine reine Drehbewegung, deren Stärke durch den Parameter  $A$  gegeben ist. Die Lösung des Strömungsproblems ist also bis auf die Stärke des drehenden Anteils eindeutig.

Für das qualitative Verhalten der Stromlinien im Fall  $U > 0$  sind die sogenannten *Staupunkte* von Bedeutung, an denen die Geschwindigkeit verschwindet. Die Forderung  $W(z_{st}) = 0$  liefert

$$z_{st} = -\frac{Ai}{2U} \pm \sqrt{a^2 - \frac{A^2}{4U^2}} .$$

Es ergeben sich zwei qualitativ verschiedene Fälle (siehe Abb. 3.4):

- 1)  $A > 2aU$ : In diesem Fall liegt nur einer der beiden Staupunkte außerhalb von  $\Omega$ . Er befindet sich auf der negativen imaginären Achse.
- 2)  $A < 2aU$ : Beide Staupunkte liegen auf  $\partial\Omega$ , im Sonderfall  $A = 0$  auf der reellen Achse.

Im weiteren wollen wir uns mit der Berechnung der Kraft

$$\mathbf{f} = \int_{\partial\Omega} p \mathbf{n} ds$$

beschäftigen, die auf den umströmten Körper wirkt. Es bietet sich an, die Bernoulli-Gleichung in (3.47) zu Berechnung des Druckes zu verwenden und das obige Integral auszuwerten. Hier soll jedoch ein anderer Zugang gewählt werden, der funktionentheoretische Überlegungen verwendet, die den Rechenaufwand erheblich verringern.

Bildet man den Gradienten der Bernoulli-Gleichung, so ergibt sich mit Hilfe der Identität (3.30) die Divergenzform der Impulsgleichung

$$\nabla \cdot (\rho_\infty \mathbf{v} \otimes \mathbf{v} + p \mathbf{I}) = 0 .$$

Abb. 3.5: Die Kontur  $C$

Nun betrachten wir eine im Strömungsgebiet verlaufende geschlossene Jordan-Kurve  $C$ , in deren Inneren sich der Querschnitt  $\Omega$  befindet (Abb. 3.5). Integrieren wir die obige Gleichung über das zwischen  $C$  und  $\partial\Omega$  liegende Gebiet und verwenden den Divergenzsatz, so erhalten wir

$$\int_{\partial\Omega} \mathbf{n} \cdot (\rho_\infty \mathbf{v} \otimes \mathbf{v} + p\mathbf{I}) ds + \int_C \mathbf{n} \cdot (\rho_\infty \mathbf{v} \otimes \mathbf{v} + p\mathbf{I}) ds = 0.$$

Wegen der Identität  $\mathbf{n} \cdot (\mathbf{v} \otimes \mathbf{v}) = (\mathbf{n} \cdot \mathbf{v})\mathbf{v}$  und der Randbedingung an  $\partial\Omega$  läßt sich die Kraft  $\mathbf{f}$  durch ein Kurvenintegral über  $C$  berechnen:

$$\mathbf{f} = - \int_C (\mathbf{n}p + \rho_\infty (\mathbf{n} \cdot \mathbf{v})\mathbf{v}) ds$$

Im Folgenden verwenden wir die Bernoulli-Gleichung für die Berechnung des Druckes und die Identitäten

$$\mathbf{n} ds = \begin{pmatrix} dy \\ -dx \end{pmatrix}, \quad \int_C \mathbf{n} ds = 0.$$

Die erste Komponente  $f_1$  der Kraft  $\mathbf{f}$ , d.h. die Komponente in Strömungsrichtung, genannt *Widerstand*, ist dann gegeben durch

$$f_1 = \frac{\rho_\infty}{2} \int_C [(v^2 - u^2)dy + 2uv dx].$$

Für die Komponente normal auf die Strömungsrichtung, den *Auftrieb*, gilt

$$f_2 = \frac{\rho_\infty}{2} \int_C [(v^2 - u^2)dx - 2uv dy].$$

Wegen der Relation

$$W^2 dz = (u - iv)^2(dx + i dy) = (u^2 - v^2)dx + 2uv dy + i[(u^2 - v^2)dy - 2uv dx]$$

folgt aus den beiden obigen Gleichungen die *Formel von Blasius*

$$i \frac{\rho_\infty}{2} \oint_C W^2 dz = f_1 - if_2.$$

Hat man eine Entwicklung von  $W$  um  $z = \infty$  der Form  $W = U + a_1/z + O(z^{-2})$  zur Verfügung, dann gilt  $W^2 = U^2 + 2Ua_1/z + O(z^{-2})$  und aus dem Residuensatz folgt

$$\oint_C W^2 dz = 4\pi i U a_1.$$

Außerdem gilt

$$\oint_C W dz = 2\pi i a_1.$$

Dieses Integral kann auch in der Form

$$\oint_C W dz = \int_C (u dx + v dy) + i \int_C (u dy - v dx) = \int_C \mathbf{v} \cdot d\mathbf{s} + i \int_C \mathbf{v} \cdot \mathbf{n} ds$$

geschrieben werden. Wegen der Divergenzfreiheit von  $\mathbf{v}$  verschwindet der Imaginärteil. Vergleich der letzten beiden Formeln liefert daher die Aussage, daß  $a_1$  rein imaginär sein muß. Damit sind die Komponenten von  $\mathbf{f}$  gegeben durch

$$f_1 = 0, \quad f_2 = -\rho_\infty 2\pi i a_1 U.$$

Die Tatsache, daß die Widerstandskraft verschwindet, ist das *d'Alembertsche Paradoxon*. Dieses Resultat zeigt eine Schwäche des verwendeten Modells auf. Für die im Experiment sehr wohl beobachteten Widerstandskräfte sind Reibungseffekte verantwortlich, die wir hier vernachlässigt haben. Für eine weitere Behandlung dieser Frage sei auf den folgenden Abschnitt verwiesen.

Für die Strömung um einen Kreiszyylinder ist die komplexe Geschwindigkeit gegeben durch

$$W = \frac{Ai}{z} + U \left( 1 - \frac{a^2}{z^2} \right).$$

Es gilt daher  $a_1 = Ai$ , und der Auftrieb ist gegeben durch

$$f_2 = 2\pi\rho_\infty AU. \quad (3.48)$$

Er ist also proportional zur Geschwindigkeit und zur Stärke des rotierenden Anteils.

Ein Kreiszyylinder ist keine sehr effektive Tragfläche. Andererseits können mit Hilfe von konformen Abbildungen beliebige Querschnitte auf den Einheitskreis abgebildet werden. Da außerdem die Laplace-Gleichung invariant unter konformen Abbildungen ist, haben wir zumindest theoretisch das Problem für beliebige Querschnitte gelöst. Eine spezielle konforme Abbildung, die in der Aerodynamik große Bedeutung hat, ist die *Joukowski-Transformation*

$$\zeta = z + \frac{c^2}{z}, \quad (3.49)$$

die das Äußere des Ursprungskreises mit Radius  $c$  in der  $z$ -Ebene auf die entlang der Strecke  $(-2c, 2c)$  geschlitzte  $\zeta$ -Ebene abbildet.

Mit Hilfe einer Joukowski-Transformation mit  $c = a$  können wir das Problem der Umströmung einer Platte lösen. Das komplexe Geschwindigkeitspotential ist gegeben durch

$$G(\zeta) = F(z(\zeta)) = Ai \ln \frac{z(\zeta)}{a} + U\zeta.$$

Verschwindet der drehende Anteil, dann wird die gleichmäßige Strömung durch die in Strömungsrichtung ausgerichtete Platte nicht gestört. Die komplexe Geschwindigkeit ist

$$V(\zeta) = \frac{W(z(\zeta))}{d\zeta/dz} = \frac{W(z(\zeta))}{1 - a^2/z(\zeta)^2}.$$

Die Punkte  $z = \pm a$  entsprechen den Kanten  $\zeta = \pm 2a$  der Platte. Die Geschwindigkeit wird dort unendlich groß.

Realistische Tragflächenquerschnitte sind vorne rund und haben eine Hinterkante (siehe Abb. 3.3). Die unbefriedigende Situation, daß die Hinterkante mit unendlicher Geschwindigkeit umströmt wird, kann durch geeignete Wahl der bisher beliebigen Rotation  $A$  behoben werden. Das bedeutet, daß der hintere Staupunkt mit der Hinterkante zusammenfällt. Diese von Kutta und Joukowski formulierte *Abflußbedingung* macht die

Lösung des Problems eindeutig. Eine rigorose theoretische Rechtfertigung dieser Bedingung existiert jedoch zur Zeit nicht.

Für die umströmte Platte ergibt sich die Forderung  $A = 0$ , also ein Verschwinden des drehenden Anteils. Ein Problem von größerer Relevanz für die Praxis ist die Strömung um eine Platte, die gegen die Strömungsrichtung geneigt ist. Das komplexe Geschwindigkeitspotential für einen unter dem Winkel  $\alpha$  angeströmten Kreiszyylinder ist

$$F(z) = Ai \ln \frac{z}{a} + U \left( ze^{-i\alpha} + \frac{a^2}{z} e^{i\alpha} \right).$$

Aus der Abflußbedingung folgt die Forderung  $W(a) = 0$  und damit

$$A = 2aU \sin \alpha.$$

Damit ist die komplexe Geschwindigkeit für eine unter dem Winkel  $\alpha$  angeströmte Platte gegeben durch

$$V(\zeta) = U \left( \cos \alpha - i \sin \alpha \frac{z(\zeta) - a}{z(\zeta) + a} \right).$$

Diese hat nur mehr an der Vorderkante eine Singularität.

Da die Joukowski-Transformation für  $z \rightarrow \infty$  bzw.  $\zeta \rightarrow \infty$  asymptotisch gleich der Identität ist, sind die Auftriebe in der  $z$ -Ebene und der  $\zeta$ -Ebene gleich. Aus Gleichung (3.48) ergibt sich also der Auftrieb

$$f_2 = 4\pi\rho_\infty aU^2 \sin \alpha.$$

Ein Kreis in der  $z$ -Ebene, der durch den Punkt  $c$  geht und dessen Mittelpunkt  $z_0$  im zweiten Quadranten liegt, wird durch die Joukowski-Transformation (3.49) auf ein sogenanntes *Joukowskisches Flügelprofil* abgebildet (siehe Abb. 3.6). Solche Profile kommen realistischen Tragflächenquerschnitten ziemlich nahe. Statt durch den Mittelpunkt  $z_0$  ist der Kreis auch durch seinen Radius  $a$  und den Anstiegswinkel  $\beta$  der Verbindungsstrecke zwischen  $z_0$  und  $c$  bestimmt (Abb. 3.6). Es gilt  $c = z_0 + ae^{-i\beta}$ .

Das komplexe Geschwindigkeitspotential für den unter dem Winkel  $\alpha$  angeströmten Kreiszyylinder ist

$$F(z) = Ai \ln \frac{z - z_0}{a} + U \left( (z - z_0)e^{-i\alpha} + \frac{a^2}{z - z_0} e^{i\alpha} \right).$$

Die Abflußbedingung  $F'(c) = 0$  ist für  $A = 2aU \sin(\alpha + \beta)$  erfüllt, woraus die Formel

$$f_2 = 4\pi\rho_\infty aU^2 \sin(\alpha + \beta)$$

für den Auftrieb einer Joukowski-Tragfläche folgt. Dieses Resultat ist erstaunlich gut durch Experimente belegt, wenn man berücksichtigt, daß sich das gelöste Problem aus einer starken Vereinfachung der ursprünglichen Gleichungen und der nur heuristisch motivierten Abflußbedingung zusammensetzt.

### 3.6. Grenzschichten

Das zu experimentellen Resultaten in krassem Widerspruch stehende d'Alembertsche Paradoxon motiviert eine eingehendere Betrachtung von Reibungseffekten. Wir wollen uns hier auf zweidimensionale stationäre Strömungen mit konstanter Dichte beschränken. Die skalierten Massen- und Impulserhaltungsgleichungen sind dann

$$\begin{aligned} u_x + v_y &= 0, \\ uu_x + vv_y + p_x &= \frac{1}{\text{Re}}(u_{xx} + u_{yy}), \\ uv_x + vv_y + p_y &= \frac{1}{\text{Re}}(v_{xx} + v_{yy}). \end{aligned} \quad (3.50)$$

Für große Reynoldszahlen sind diese Gleichungen singularär gestört. Es ist daher zu erwarten, daß die im vorigen Abschnitt berechneten Lösungen der reduzierten Gleichungen außerhalb von Grenzschichten brauchbare Approximationen darstellen. Hier nehmen wir an, wir hätten eine solche reduzierte Lösung  $\bar{u}$ ,  $\bar{v}$ ,  $\bar{p}$  gegeben. Es ist zu erwarten, daß Grenzschichten an Randstücken auftreten, an denen die reduzierte Lösung die Haftbedingung verletzt.

Sei als Beispiel die positive  $x$ -Achse ein solches Randstück. Dann gilt im allgemeinen  $\bar{v}(x, 0) = 0$ , aber  $\bar{u}(x, 0) \neq 0$ . Eine Grenzschichtvariable ist durch  $\eta = y\text{Re}^\alpha$ ,  $\alpha > 0$  gegeben. Außerdem ist anzunehmen, daß die Geschwindigkeitskomponente  $v$  überall in der Grenzschicht klein ist. Die Umskalierung  $w = v\text{Re}^\beta$ ,  $\beta > 0$  scheint daher angebracht. Führt man diese Transformationen in (3.50) durch, so zeigt sich, daß die Wahl  $\alpha = \beta = 1/2$  zu einer signifikanten Degeneration führt. Für die Grenzschichtlösung  $\tilde{u}$ ,  $\tilde{w}$ ,  $\tilde{p}$  gelten die Gleichungen

$$\begin{aligned} \tilde{u}_x + \tilde{w}_\eta &= 0, \\ \tilde{u}\tilde{u}_x + \tilde{w}\tilde{u}_\eta + \tilde{p}_x &= \tilde{u}_{\eta\eta}, \\ \tilde{p}_\eta &= 0. \end{aligned}$$

Die letzte Gleichung zeigt, daß der Druck über die Grenzschicht konstant ist. Beachtet man, daß für die reduzierte Lösung

$$\bar{u}(x, 0)\bar{u}_x(x, 0) + \bar{p}_x(x, 0) = 0$$

gilt, so ergibt sich das Grenzschichtproblem

$$\begin{aligned} \tilde{u}_x + \tilde{w}_\eta &= 0, \\ \tilde{u}\tilde{u}_x + \tilde{w}\tilde{u}_\eta - \bar{u}(x, 0)\bar{u}_x(x, 0) &= \tilde{u}_{\eta\eta}. \end{aligned} \quad (3.51)$$

Für positive  $\tilde{u}$  ist die zweite Gleichung eine parabolische Gleichung für  $\tilde{u}$ , während  $\tilde{w}$  durch Integration der ersten Gleichung in  $\eta$ -Richtung ermittelt werden kann. Die Zusatzbedingungen

$$\tilde{u}(x, 0) = \tilde{w}(x, 0) = 0, \quad \tilde{u}(x, \infty) = \bar{u}(x, 0), \quad \tilde{u}(0, \eta) = u_0(\eta) \quad (3.52)$$

scheinen daher auf ein sachgemäß gestelltes Problem zu führen.

Man kann zeigen, daß (3.51), (3.52) auch Grenzschichten entlang gekrümmter Ränder beschreibt, wobei  $x$  die Bogenlänge auf dem Rand und  $\eta$  der um  $\text{Re}^{1/2}$  gestreckte Abstand vom Rand ist. Hierbei muß man allerdings voraussetzen, daß der Krümmungsradius des Randes überall groß im Vergleich zu  $\text{Re}^{-1/2}$  ist.

Das Vertrauen des Lesers in das Problem (3.51), (3.52) soll durch die Betrachtung eines Beispiels gestärkt werden. Für eine in Anströmungsrichtung ausgerichtete Platte gilt  $\bar{u}(x, y) = U$ . Für die durch  $\tilde{u} = \psi_\eta$ ,  $\tilde{w} = -\psi_x$  definierte Stromfunktion  $\psi$  gilt

$$\begin{aligned} \psi_\eta\psi_{x\eta} - \psi_x\psi_{\eta\eta} &= \psi_{\eta\eta\eta}, \\ \psi_x(x, 0) = \psi_\eta(x, 0) &= 0, \quad \psi_\eta(x, \infty) = U, \quad \psi_\eta(0, \eta) = U. \end{aligned}$$

Von Blasius (1908) wurde entdeckt, daß dieses Problem eine Ähnlichkeitslösung der Form

$$\psi = \sqrt{2Ux}f(s) \quad \text{mit } s = \eta\sqrt{\frac{U}{2x}}$$



besitzt. Für die Funktion  $f$  ergibt sich das Problem

$$\begin{aligned} f''' + ff'' &= 0, \\ f(0) = f'(0) &= 0, \quad f'(\infty) = 1. \end{aligned}$$

Die Geschwindigkeitskomponenten lassen sich aus der Blasius-Lösung ermitteln:

$$\tilde{u} = Uf'(s), \quad \tilde{w} = \sqrt{\frac{U}{2x}}(sf'(s) - f(s))$$

Die mit Hilfe der Grenzschicht berechneten, auf einen umströmten Körper wirkenden viskosen Spannungskräfte sind  $O(\text{Re}^{-1/2})$ . Die in Experimenten gemessenen Widerstandskräfte sind oft viel größer. Unsere bisherigen Überlegungen sind daher noch nicht ausreichend.

Das d'Alembertsche Paradoxon erklärt sich dadurch, daß zwar vom vorderen Staupunkt weg der Druck auf den Körper abnimmt, sich diese Entwicklung jedoch an der Hinterseite des Körpers wieder umkehrt. Das bedeutet, daß der Term  $\bar{p}_x(x, 0) = -\bar{u}(x, 0)\bar{u}_x(x, 0)$  in (3.51) an der Hinterseite positiv ist. Aus der zweiten Gleichung in (3.51) wird es plausibel, daß diese Tatsache bewirken kann, daß die tangential Geschwindigkeitskomponente in der Grenzschicht negativ wird. Es tritt also eine Rückströmung auf, deren Effekt eine Abdrängung der Strömung von der Körperoberfläche ist. Dieser Effekt wird *Ablösung* der Grenzschicht genannt. Zwischen den an der Ober- und Unterseite abgelösten Grenzschichten bildet sich ein *Totwasser* mit viel geringerem Druck als durch die Potentialströmung vorausgesagt. Ob dieser Effekt auftritt, hängt von der Form des umströmten Körpers ab. Für an der Hinterseite spitz zulaufende (*stromlinienförmige*) Körper ist er weniger bedeutend. Die Resultate des vorigen Abschnittes haben daher zwar für Joukowski-Profile aber nicht für Kreiszyylinder praktische Bedeutung.

## 4. KAPITEL Halbleiterbauelemente

### 4.1. Die Halbleiterdiode

Eine Halbleiterdiode ist ein elektronisches Bauelement mit zwei Kontakten, das die Funktion eines Ventils hat, d.h. Stromfluß durch die Diode ist nur in eine Richtung möglich. Dieser Abschnitt enthält eine phänomenologische Erklärung ihrer Wirkungsweise.

Elektrischer Strom in einem Festkörper wird erzeugt durch Bewegung von *Ladungsträgern*. Das sind Elementarteilchen, für die Ladung eine Eigenschaft ist, die in zwei verschiedenen Arten auftritt, die mit positiv und negativ bezeichnet werden. Teilchen mit gleichartiger Ladung stoßen einander ab, solche mit ungleichartiger Ladung ziehen einander an.

Die in Atomkernen enthaltenen Protonen haben positive Ladung; die Elektronen in der Hülle sind negativ geladen. Die Atomkerne sowie die meisten zugehörigen Elektronen können innerhalb des Festkörpers als unbeweglich angenommen werden. Nur für die Valenzelektronen der äußeren Schale ist es möglich, daß sie ihre feste Bindung an einen Atomkern verlieren und sich dadurch im Festkörper frei bewegen können. Diese Elektronen werden *freie Elektronen* genannt. Sie sind für den Stromfluß verantwortlich. Die Anzahl der freien Elektronen in einem Festkörper hängt im wesentlichen von der Art der chemischen Bindung ab.

In der *Metallbindung* verlieren die Valenzelektronen ihre Bindung an den Atomkern. In Metallen ist also Stromfluß möglich, und sie werden als *Leiter* bezeichnet.

Im Gegensatz dazu wird die *Valenzbindung* durch Elektronenpaarbrücken hergestellt, in denen die Valenzelektronen gebunden, und damit unbeweglich, sind. Als Beispiel zeigt Abb. 4.1 eine vereinfachte (zweidimensionale) Darstellung eines Siliziumkristalles. Silizium hat 4 Valenzelektronen, und jedes Atom ist im Kristallgitter direkt mit 4 anderen verbunden.

Abb. 4.1: Valenzbindung im Siliziumkristall

In einem solchen Material kann es zu keinem Stromfluß kommen. Das gilt allerdings nur für den absoluten Temperaturnullpunkt. Bei positiver Temperatur kommt es zu Schwingungen des Kristallgitters, die verursachen, daß Elektronen aus Paarbindungen "herausbrechen" können (Abb. 4.2). Auf diese Art entstehen auch im Fall der Valenzbindung freie Elektronen, allerdings bei weitem nicht so viele wie in einem Leiter.

Abb. 4.2: Gitterschwingungen im Siliziumkristall

Je nachdem, ob dieses Herausbrechen aus Paarbindungen leichter oder weniger leicht möglich ist, spricht man von *Halbleitern* oder *Isolatoren*. Das betrachtete Element Silizium zählt zu den Halbleitern. Die Unterschiede in der Leitfähigkeit machen die folgenden Zahlen deutlich: Bei Zimmertemperatur gibt es in typischen Leitern ungefähr  $10^{22}$  freie Elektronen pro  $\text{cm}^3$ . Entsprechende Zahlen für typische Halbleiter bzw. Isolatoren sind  $10^{10}$  bzw.  $10^3$ .

Der Unterschied zwischen Halbleitern und Isolatoren kann mit Hilfe des *Bändermodells* erklärt werden. Die Quantenmechanik liefert die Aussage, daß Elektronen nur bestimmte Energiewerte annehmen können, sich also nur in bestimmten “erlaubten” Energiebändern aufhalten können. Das am absoluten Temperaturnullpunkt höchste besetzte Energieband heißt *Valenzband*. Das nächste darüberliegende Band heißt *Leitungsband*. Bei positiver Temperatur kommt es in Halbleitern und Isolatoren dazu, daß Elektronen aus dem Valenzband in das Leitungsband springen. Die Größe der sogenannten *Bandlücke* zwischen Valenzband und Leitungsband macht den Unterschied zwischen Halbleitern und Isolatoren aus (Abb. 4.3).

Abb. 4.3: Bänderstruktur bei a) Halbleiter und b) Isolator

Kehren wir zurück zum Siliziumkristall. An Stellen, an denen ein Elektron aus einer Paarbindung herausbricht, entsteht ein *Loch*. Außerdem entsteht an der entsprechenden Stelle ein positiver Ladungsüberschuß, weil ja ein negativ geladenes Elektron fehlt. Es ist eine sehr nützliche Betrachtungsweise, ein solches Loch als positiv geladenes Teilchen anzusehen (Abb. 4.4), das auch als *Defektelektron* bezeichnet wird.

Abb. 4.4: Loch als positiv geladenes Teilchen

Interessant werden Löcher durch die Tatsache, daß Elektronen aus benachbarten Paarbindungen diese verlassen und Löcher “stopfen” können. Dieser Vorgang kann auch als Bewegung des Loches in entgegengesetzte Richtung aufgefaßt werden (Abb. 4.5).

Zwei Dinge sind wesentlich. Erstens verläßt das Elektron während dieser Bewegung das Valenzband nicht. Zweitens können sich die Löcher auf diese Art im wesentlichen genauso frei bewegen wie die freien Elektronen. Damit kann die Bewegung der Löcher signifikant zum Stromfluß beitragen. Die *freien Ladungsträger* im Halbleiter sind also (freie) *Elektronen und Löcher*. Der Gesamtstrom setzt sich aus einem Elektronenstrom und einem Löcherstrom zusammen.

Wie wir gesehen haben, liegt die Leitfähigkeit eines reinen Halbleiters wesentlich unter der eines Leiters. Die Leitfähigkeit kann durch gezielte Verunreinigung (Einbau von Störstellen, *Dotierung*) stark erhöht

Abb. 4.5: Bewegung eines Loches

werden. Betrachtet man den Halbleiter Silizium, so eignen sich dafür besonders Elemente mit 3 oder 5 Valenzelektronen.

Betrachten wir als Beispiel eine Verunreinigung mit Phosphor. Phosphoratome übernehmen dann im Kristall die Stelle von Siliziumatomen. 4 der 5 Valenzelektronen des Phosphoratoms werden in Paarbindungen verwendet, das fünfte Valenzelektron findet keinen Platz im Valenzband und steht als freies Elektron zur Verfügung. Zusätzlich entsteht eine durch den Rumpf des Phosphoratoms verursachte ortsfeste positive Ladung. Da Elemente mit 5 Valenzelektronen freie Elektronen spenden, werden sie *Donatoren* genannt.

Verunreinigt man Silizium mit Aluminium, so werden die 3 Valenzelektronen jedes Aluminiumatoms in Paarbindungen verwendet. Für die vierte Paarbindung fehlt ein Elektron, es entsteht also ein Loch. Bewegt sich dieses vom Aluminiumatom weg, so bleibt eine ortsfeste negative Ladung zurück. Elemente mit 3 Valenzelektronen heißen *Akzeptoren*.

Abb. 4.6: Dotierung von Silizium mit a) Phosphor und b) Aluminium

In einem mit Donatoren dotierten (*n-dotierten*) Halbleiter wird der Stromfluß im wesentlichen von freien Elektronen bestritten. Ist der Halbleiter mit Akzeptoren dotiert (*p-dotiert*), dann werden die Löcher den größeren Anteil zum Stromfluß liefern.

Wir sind jetzt soweit, den Aufbau und die Funktion einer Halbleiterdiode beschreiben zu können. Sie besteht aus einem Stück Halbleiter, das auf einer Seite *p*-dotiert und auf der anderen *n*-dotiert ist (Abb. 4.7).

Da im *p*-dotierten Bereich nur wenige freie Elektronen und im *n*-dotierten Bereich nur wenige Löcher vorhanden sind, scheint zunächst weder ein signifikanter Fluß von Elektronen noch von Löchern möglich zu sein. Versucht man nun, einen Fluß von Elektronen von links nach rechts durch die Diode zu erzeugen, so bedeutet das, daß man rechts Elektronen und links Löcher "aus der Diode zieht". Die Anzahl der freien Ladungsträger in der Diode verringert sich also, und es kann zu keinem signifikanten Fluß kommen. Versucht man andererseits von rechts Elektronen und von links Löcher "in die Diode zu schieben", so werden einerseits Elektronen in den *p*-dotierten Bereich und Löcher in den *n*-dotierten Bereich vordringen, und in der ganzen Diode werden genügend freie Ladungsträger vorhanden sein, um große Ströme zuzulassen.

Während also Elektronen die Diode von rechts nach links mehr oder weniger ungehindert durchfließen können, kann es zu keinem nennenswerten Fluß von links nach rechts kommen.

Abb. 4.7: *pn*-Diode

Unter Ausnützung der speziellen Eigenschaften von dotierten Halbleitern werden verschiedenste elektronische Bauelemente (Transistoren, Thyristoren usw.) gebaut. Um deren Eigenschaften nicht nur qualitativ — wie oben — sondern auch quantitativ beschreiben zu können, benötigt man mathematische Modelle der Vorgänge in Halbleiterbauelementen.

## 4.2. Die Halbleitergrundgleichungen

Wir gehen aus vom *Coulombschen Gesetz*

$$\mathbf{f}_{ij} = \frac{Q_i Q_j (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3}$$

für die vom Teilchen mit der Position  $\mathbf{r}_j$  und der Ladung  $Q_j$  auf das Teilchen mit der Position  $\mathbf{r}_i$  und der Ladung  $Q_i$  ausgeübte Kraft. Die Wahl 1 für die Proportionalitätskonstante bedeutet, daß die obige Gleichung als Definitionsgleichung für die Ladungseinheit zu verstehen ist.

In einem Kontinuumsmodell betrachten wir eine *Ladungsdichte*  $\rho(\mathbf{x}, t)$  mit der Einheit Ladung/Volumen. Die auf ein Teilchen mit Ladung  $Q$  und Position  $\mathbf{x}$  wirkende Kraft ist

$$\mathbf{f} = Q \mathbf{E}(\mathbf{x}, t),$$

wobei das *elektrische Feld* gegeben ist durch

$$\mathbf{E}(\mathbf{x}, t) = \int \frac{\rho(\boldsymbol{\xi}, t)(\mathbf{x} - \boldsymbol{\xi})}{|\mathbf{x} - \boldsymbol{\xi}|^3} d\boldsymbol{\xi}.$$

Man sieht leicht, daß das elektrische Feld wirbelfrei ist ( $\nabla \times \mathbf{E} = 0$ ). Es ist daher ein Gradientenfeld, und das *elektrostatistische Potential*  $\psi(\mathbf{x}, t)$  ist durch die Gleichung

$$\mathbf{E} = -\nabla\psi$$

bis auf eine additive (eventuell zeitabhängige) Konstante eindeutig bestimmt. Eine leichte Rechnung zeigt

$$\psi(\mathbf{x}, t) = \int \frac{\rho(\boldsymbol{\xi}, t)}{|\mathbf{x} - \boldsymbol{\xi}|} d\boldsymbol{\xi}. \quad (4.1)$$

Eine physikalische Interpretation für das Potential ergibt sich aus der Gleichung

$$\psi(\mathbf{x}_2, t) - \psi(\mathbf{x}_1, t) = - \int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{E}(\mathbf{x}, t) \cdot d\mathbf{s}.$$

Die Wirbelfreiheit des elektrischen Feldes impliziert die Wegunabhängigkeit des Kurvenintegrals auf der rechten Seite, das die Arbeit angibt, die verrichtet werden muß, um ein Teilchen mit Ladung 1 von  $\mathbf{x}_1$  nach  $\mathbf{x}_2$  zu bewegen.

Da die Funktion

$$-\frac{1}{4\pi|\mathbf{x}|}$$

eine Fundamentallösung des Laplaceoperators ist, folgt aus (4.1) die *Poissongleichung*

$$\Delta\psi = -4\pi\rho.$$

In der Elektrotechnik ist diese Gleichung auch als differentielle Form des *Gaußschen Gesetzes der Elektrostatik* bekannt.

In einem dotierten Halbleiterkristall sei  $n(\mathbf{x}, t)$  die Anzahl der freien Elektronen pro Volumeneinheit. Ebenso sei  $p(\mathbf{x}, t)$  die Dichte der Löcher,  $N_D(\mathbf{x})$  die Dichte der Donatoratome und  $N_A(\mathbf{x})$  die Dichte der Akzeptoratome. Da die Donatoratome einen positiven und die Akzeptoratome einen negativen Ladungsüberschuß erzeugen, ergibt sich ein Beitrag

$$q(p - n + N_D - N_A)$$

zur Ladungsdichte  $\rho$ , wobei  $-q$  die Ladung eines Elektrons ist.  $q$  heißt *Elementarladung*, und es gilt  $q \approx 4,803 \times 10^{-10}$  esE. Hier bezeichnet esE =  $\sqrt{\text{gcm}^3/\text{s}}$  (elektrostatische Einheit) die Einheit der Ladung im CGS-System.

Nicht berücksichtigt wurden bis jetzt der Großteil der Atomkerne und die Elektronen, die fest an solche Kerne gebunden sind. Da die entsprechenden Ladungen einander aufheben und die jeweils einander aufhebenden Ladungen sehr nahe beisammen liegen, könnte man meinen, daß man aus unserer makroskopischen Sicht ihren Einfluß vernachlässigen kann. Das ist nicht ganz so.

Teilt man diese Ladungen in kleine Gruppen, die in Summe elektrisch neutral sind und deren Teile stark aneinander gebunden sind, so werden diese Gruppen durch ein vorhandenes elektrisches Feld ausgerichtet und verursachen dadurch ihrerseits einen Beitrag zum Feld, der im wesentlichen dem äußeren Feld entgegengerichtet ist. Da alle Gruppen in dieselbe Richtung ausgerichtet werden, ist die Summe ihrer Einflüsse nicht zu vernachlässigen. Diesem Effekt kann dadurch Rechnung getragen werden, daß das durch die schon berücksichtigten Ladungen erzeugte Feld mit einer Konstanten multipliziert wird, die vom Material abhängt. Ihren Kehrwert bezeichnet man mit  $\varepsilon$  und nennt ihn *Dielektrizitätskonstante*. In der folgenden Tabelle sind ihre Werte für verschiedene Materialien angegeben:

Luft	1,00059
Wasser	80
Kochsalz	6,12
Silizium	11,7

Wir können schließlich die Poissongleichung im dotierten Halbleiter schreiben als

$$\varepsilon\Delta\psi = 4\pi q(n - p - C), \quad (4.2)$$

wobei wir die als gegeben angenommene Funktion  $C(\mathbf{x}) = N_D(\mathbf{x}) - N_A(\mathbf{x})$  das *Dotierungsprofil* nennen. Wie schon erwähnt, unterscheiden wir zwischen  $n$ -dotierten Bereichen ( $C > 0$ ) und  $p$ -dotierten Bereichen ( $C < 0$ ). Flächen, die aus Nullstellen von  $C$  bestehen, heißen  $pn$ -Übergänge. Die Funktion

$$\rho(\mathbf{x}, t) = q(p(\mathbf{x}, t) - n(\mathbf{x}, t) + C(\mathbf{x}))$$

wird als *Raumladungsdichte* bezeichnet.

Als nächstes beschäftigen wir uns mit dem Transport von Ladungsträgern. Ladungserhaltung wird beschrieben durch die Kontinuitätsgleichung

$$\frac{\partial\rho}{\partial t} + \nabla \cdot \mathbf{J} = 0,$$

in der  $\mathbf{J}$  die *Stromdichte* bezeichnet, die sich aus der Elektronenstromdichte  $\mathbf{J}_n$  und der Löcherstromdichte  $\mathbf{J}_p$  zusammensetzt:  $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$ . Damit kann die Kontinuitätsgleichung geschrieben werden als

$$\nabla \cdot \mathbf{J}_n - q\frac{\partial n}{\partial t} = -\nabla \cdot \mathbf{J}_p - q\frac{\partial p}{\partial t}.$$

Setzen wir die beiden Seiten dieser Gleichung gleich einer Größe  $qR$ , so ergibt sich

$$\begin{aligned}\nabla \cdot \mathbf{J}_n - q \frac{\partial n}{\partial t} &= qR, \\ \nabla \cdot \mathbf{J}_p + q \frac{\partial p}{\partial t} &= -qR.\end{aligned}\tag{4.3}$$

Diese Aufteilung in zwei Gleichungen erscheint zunächst willkürlich und nicht sehr sinnvoll. Die Größe  $R$  hat jedoch eine einfache physikalische Bedeutung. Sie ist die Rate, mit der freie Elektronen und Löcher entstehen ( $R < 0$ ) bzw. verschwinden ( $R > 0$ ). Im ersten Fall springt ein Elektron aus dem Valenzband in das Leitungsband, wodurch ein freies Elektron und ein Loch entstehen. Dieser Vorgang heißt *Paarerzeugung* oder *Generation*. Umgekehrt kann ein freies Elektron in ein Loch fallen, wodurch beide verschwinden. Dieser Vorgang heißt *Rekombination*. Die Größe  $R$  wird als *Rekombinations-Generationsrate* bezeichnet. Die Gleichungen (4.3) heißen Elektronen- (bzw. Löcher-) Kontinuitätsgleichung.

Offen ist noch die Frage, wodurch die Ladungsträger nun wirklich bewegt werden, d.h. wie die Stromdichten  $\mathbf{J}_n$  und  $\mathbf{J}_p$  berechnet werden können. Wir wollen ein Modell verwenden, dessen rigorose Herleitung eine ausführliche Beschäftigung mit Festkörperphysik und Quantenmechanik notwendig machen würde. Hier wollen wir uns auf einfache heuristische Überlegungen beschränken, die das Modell auch plausibel machen.

In Metallen gibt es keine Löcher und die Dichte der freien Elektronen ist annähernd konstant. Der Grund dafür ist eine starke Tendenz, Ladungsneutralität zu erhalten. Eine experimentell gerechtfertigte Formel für die Elektronenstromdichte ist in diesem Fall das *Ohmsche Gesetz*

$$\mathbf{J}_n = \sigma \mathbf{E} = qn\mu_n \mathbf{E}$$

mit der Leitfähigkeit  $\sigma$  bzw. der Beweglichkeit  $\mu_n$ .

Durch die Gegenwart von Löchern bzw. durch ein mit dem Ort variierendes Dotierungsprofil ist es in Halbleitern möglich, Ladungsneutralität auch bei nichtkonstanten Elektronen- bzw. Löcherdichten herzustellen. Daher ist es naheliegend, daß auch Diffusionseffekte für den Ladungstransport von Bedeutung sein können. Diese Überlegungen führen auf die *Stromrelationen*

$$\begin{aligned}\mathbf{J}_n &= qn\mu_n \mathbf{E} + qD_n \nabla n, \\ \mathbf{J}_p &= qp\mu_p \mathbf{E} - qD_p \nabla p,\end{aligned}\tag{4.4}$$

wobei  $\mu_n$  ( $\mu_p$ ) bzw.  $D_n$  ( $D_p$ ) die Beweglichkeit bzw. die Diffusionskonstante der Elektronen (Löcher) ist. Man beachte die Vorzeichen! Aus einer tiefgehenden Herleitung dieser Gleichungen folgen die *Einsteinrelationen*

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{q},$$

mit der Temperatur  $T$  in Kelvin (K) und der *Boltzmannkonstanten*  $k \approx 1,38 \times 10^{-16}$  erg/K. Wir verwenden die Annahme einer konstanten Temperatur und bezeichnen  $U_T = kT/q$  als *Temperaturspannung*. Typische Werte für die Beweglichkeiten in Silizium bei Zimmertemperatur sind

$$\mu_n = 4,5 \times 10^5 \sqrt{\frac{\text{cm}^3}{g}}, \quad \mu_p = 1,8 \times 10^5 \sqrt{\frac{\text{cm}^3}{g}}.$$

Einsetzen der Stromrelationen in die Kontinuitätsgleichungen gibt

$$\begin{aligned}\nabla \cdot (\mu_n(n\mathbf{E} + U_T \nabla n)) - \frac{\partial n}{\partial t} &= R, \\ \nabla \cdot (\mu_p(p\mathbf{E} - U_T \nabla p)) + \frac{\partial p}{\partial t} &= -R.\end{aligned}\tag{4.5}$$

Für gegebenes elektrisches Feld sind das zwei parabolische Gleichungen für  $n$  und  $p$ . Zusammen mit der Poissongleichung (4.2) läge ein geschlossenes System vor, wenn wir ein Modell für die Rekombinations-Generationsrate  $R$  hätten.

Es gibt verschiedene physikalische Effekte, die für Rekombination und Generation verantwortlich sein können. Einer der wichtigsten soll hier beschrieben werden. Gewisse Verunreinigungen des Halbleiters haben den Effekt, daß im verbotenen Band zwischen Valenz- und Leitungsband zusätzliche erlaubte Energieniveaus auftreten. Diese können einem Elektron beim Überspringen des verbotenen Bandes als Zwischenstation dienen und damit Rekombination und Generation erleichtern.

Nehmen wir die Existenz eines Störstellenniveaus an, dann kann es zu 4 verschiedenen Vorgängen kommen:

- a) Ein Elektron fällt aus dem Leitungsband in das Störstellenniveau.
- b) Ein Elektron fällt aus dem Störstellenniveau in das Valenzband und stopft ein Loch.
- c) Ein Elektron springt aus dem Valenzband auf das Störstellenniveau, und ein Loch entsteht.
- d) Ein Elektron springt vom Störstellenniveau in das Leitungsband (Abb. 4.8).

Abb. 4.8: Rekombination-Generation über ein Störstellenniveau

Die Berechnung der Raten, mit denen diese Vorgänge stattfinden, beruht auf dem Massenwirkungsgesetz. Sei  $N_{St}$  die Dichte der verfügbaren Plätze und  $n_{St}$  die Dichte der besetzten Plätze auf dem Störstellenniveau. Dann ist die Anzahl der Vorgänge a), b), c) bzw. d) pro Volumeneinheit und Zeiteinheit gegeben durch

- a)  $C_a n(N_{St} - n_{St})$ ,
- b)  $C_b p n_{St}$ ,
- c)  $C_c (N_{St} - n_{St})$ ,
- d)  $C_d n_{St}$ ,

mit den Proportionalitätskonstanten  $C_a$ ,  $C_b$ ,  $C_c$  und  $C_d$ . Der Formel für die Rate in c) bzw. d) liegt die Annahme zugrunde, daß es im Überfluß Elektronen im Valenzband bzw. freie Plätze im Leitungsband gibt.

Da die Störstellen ortsfest sind, tragen die Elektronen an den Störstellenniveaus nichts zum Stromfluß bei. Wir haben daher die Differentialgleichungen

$$\begin{aligned} \frac{\partial n}{\partial t} - \frac{1}{q} \nabla \cdot \mathbf{J}_n &= C_d n_{St} - C_a n(N_{St} - n_{St}), \\ \frac{\partial p}{\partial t} + \frac{1}{q} \nabla \cdot \mathbf{J}_p &= C_c (N_{St} - n_{St}) - C_b p n_{St}, \\ \frac{\partial n_{St}}{\partial t} &= (C_a n + C_c)(N_{St} - n_{St}) - (C_b p + C_d) n_{St} \end{aligned}$$

zu betrachten. Durch die Einführung einer zusätzlichen Ladungsträgerspezies mit der Dichte  $n_{St}$  stimmen natürlich die Quellen- bzw. Senkenterme in den Gleichungen für freie Elektronen und Löcher nicht mehr notwendigerweise überein. Das Problem wird jedoch üblicherweise dadurch vereinfacht, daß man annimmt, daß die Dichte  $n_{St}$  viel schneller ihrem Gleichgewichtswert zustrebt als  $n$  und  $p$ . Daraus ergibt sich die Annahme, daß die rechte Seite der dritten Gleichung verschwindet. Aus der entstehenden algebraischen Gleichung kann  $n_{St}$  berechnet und damit aus dem System eliminiert werden.

Die Quotienten  $C_c/C_a$  und  $C_d/C_b$  haben die Dimension von Dichten ( $\text{cm}^{-3}$ ). Wir machen die Annahme, daß sie gleich und von der Anzahl der Störstellen unabhängig sind. Ihr gemeinsamer Wert heißt *intrinsische Dichte* und wird mit dem Symbol  $n_i$  bezeichnet.



Offensichtlich sind  $C_a N_{St}$  und  $C_b N_{St}$  typische Raten, mit denen Elektronen und Löcher verschwinden. Ihre Kehrwerte können daher als *Lebensdauern*

$$\tau_n = \frac{1}{C_a N_{St}}, \quad \tau_p = \frac{1}{C_b N_{St}}$$

interpretiert werden. Ihre Größe kann mit Art und Dichte der Störstellen sehr stark variieren. Typische Werte sind  $\tau_n = 10^{-6}$ s und  $\tau_p = 10^{-5}$ s.

Aus den obigen Annahmen folgt für die Rekombinations-Generationsrate die Formel

$$R = \frac{np - n_i^2}{\tau_p(n + n_i) + \tau_n(p + n_i)}. \quad (4.6)$$

Damit ist unsere Herleitung der *Halbleitergrundgleichungen* vollständig. Sie bestehen aus der Poissongleichung und den Kontinuitätsgleichungen für freie Elektronen und Löcher:

$$\begin{aligned} \varepsilon \Delta \psi &= 4\pi q(n - p - C), \\ \nabla \cdot (\mu_n(U_T \nabla n - n \nabla \psi)) - \frac{\partial n}{\partial t} &= R, \\ \nabla \cdot (\mu_p(U_T \nabla p + p \nabla \psi)) - \frac{\partial p}{\partial t} &= R, \end{aligned} \quad (4.7)$$

mit  $R$  gegeben durch (4.6).

Betrachten wir zunächst eine stationäre Situation ohne Stromfluß, die wir *thermisches Equilibrium* nennen. Das Potential und die Dichten in dieser Situation bezeichnen wir mit  $\psi_e$ ,  $n_e$  und  $p_e$ . Offensichtlich müssen die Gleichungen

$$\begin{aligned} U_T \nabla n_e - n_e \nabla \psi_e &= 0, \\ U_T \nabla p_e + p_e \nabla \psi_e &= 0, \\ n_e p_e - n_i^2 &= 0 \end{aligned}$$

gelten. Integration der ersten beiden Gleichungen liefert

$$n_e = A e^{\psi_e/U_T}, \quad p_e = B e^{-\psi_e/U_T}$$

mit den Integrationskonstanten  $A$  und  $B$ , für die wegen der dritten Gleichung  $AB = n_i^2$  gilt. Das Potential kann nur bis auf eine additive Konstante eindeutig bestimmt sein. Diese wird durch die Wahl

$$A = B = n_i$$

festgelegt. Damit ist das Equilibriumpotential Lösung der Poissongleichung

$$\varepsilon \Delta \psi_e = q \left( 2n_i \sinh \frac{\psi_e}{U_T} - C \right). \quad (4.8)$$

Um die Formulierung eines mathematischen Problems zu vervollständigen, sind noch Zusatzbedingungen notwendig. Die Gleichungen (4.7) werden für  $\mathbf{x} \in \Omega \subset \mathbb{R}^3$  betrachtet, wobei  $\Omega$  den Halbleiterteil eines Bauelementes repräsentiert. An der Oberfläche sind paarweise disjunkte Kontakte  $K_i \subset \partial\Omega$ ,  $i = 1, \dots, k$  angebracht, an denen der Halbleiterkristall mit leitendem Material in Berührung ist. Die Vereinigung der Kontakte bezeichnen wir mit

$$\partial\Omega_D = \bigcup_{i=1}^k K_i.$$

Für den Rest der Oberfläche

$$\partial\Omega_N = \partial\Omega \setminus \partial\Omega_D$$

nehmen wir an, daß er mit isolierendem Material in Berührung ist. Durch diese Randabschnitte fließt kein elektrischer Strom. Außerdem machen wir die idealisierende Annahme, daß die normale Komponente des elektrischen Feldes an isolierenden Randabschnitten verschwindet. Das könnte z.B. dadurch begründet sein, daß die Dielektrizitätskonstante im Isolator wesentlich kleiner ist als im Halbleiter. Es gelten also die Randbedingungen

$$\mathbf{E} \cdot \mathbf{n} = \mathbf{J}_n \cdot \mathbf{n} = \mathbf{J}_p \cdot \mathbf{n} = 0, \quad \text{auf } \partial\Omega_N.$$

Diese sind äquivalent zu den homogenen Neumannbedingungen

$$\nabla\psi \cdot \mathbf{n} = \nabla n \cdot \mathbf{n} = \nabla p \cdot \mathbf{n} = 0, \quad \text{auf } \partial\Omega_N. \quad (4.9)$$

Für die Kontakte nehmen wir an, daß es sich um sogenannte *Ohmsche Kontakte* handelt. Das bedeutet, daß der Kontakt zwischen Leiter und Halbleiter intensiv genug ist, daß die große Zahl von freien Elektronen im Leiter Abweichungen vom Gleichgewicht im Halbleiter immer sofort ausgleicht. Wir nehmen also an, daß an den Kontakten keine Raumladung auftritt,

$$n - p - C = 0, \quad \text{auf } \partial\Omega_D,$$

und daß thermisches Equilibrium,

$$np - n_i^2 = 0, \quad \text{auf } \partial\Omega_D,$$

gilt. Unter Berücksichtigung der Tatsache, daß  $n$  und  $p$  positiv sein müssen, erhält man aus diesen Gleichungen die Dirichletbedingungen

$$n = \frac{1}{2} \left( C + \sqrt{C^2 + 4n_i^2} \right), \quad p = \frac{1}{2} \left( -C + \sqrt{C^2 + 4n_i^2} \right), \quad \text{auf } \partial\Omega_D, \quad (4.10)$$

für die Dichten an den Kontakten.

Für das Equilibriumpotential folgt aus dem Verschwinden der Raumladung an den Kontakten die Randbedingung

$$\psi_e = \psi_{bi}, \quad \text{auf } \partial\Omega_D,$$

in der

$$\psi_{bi} = U_T \operatorname{arsinh} \frac{C}{2n_i}$$

das sogenannte *eingebaute Potential* (*built in potential*) ist. Damit ist das Problem für das Equilibriumpotential vollständig formuliert. Im allgemeinen Fall erfüllt das Potential die Randbedingungen

$$\psi = \psi_{bi} + U_i(t), \quad \text{auf } K_i, \quad i = 1, \dots, k, \quad (4.11)$$

wobei die Spannung zwischen den Kontakten  $K_i$  und  $K_j$  zum Zeitpunkt  $t$  durch  $U_i(t) - U_j(t)$  gegeben ist.

Durch die zusätzliche Angabe von Anfangsbedingungen für die Dichten entsteht aus (4.7), (4.9), (4.10), (4.11) ein Problem, von dem man annehmen kann, daß es sachgemäß gestellt ist.

### 4.3. Skalierung des stationären Problems

Das stationäre Halbleiterproblem besteht aus den Differentialgleichungen

$$\begin{aligned}\varepsilon \Delta \psi &= 4\pi q(n - p - C), \\ \nabla \cdot (\mu_n(U_T \nabla n - n \nabla \psi)) &= R, \\ \nabla \cdot (\mu_p(U_T \nabla p + p \nabla \psi)) &= R,\end{aligned}\tag{4.12}$$

mit der Rekombinations-Generationsrate gegeben durch (4.6), und den Randbedingungen an isolierenden Randabschnitten

$$\nabla \psi \cdot \mathbf{n} = \nabla n \cdot \mathbf{n} = \nabla p \cdot \mathbf{n} = 0, \quad \text{auf } \partial\Omega_N\tag{4.13}$$

und an Ohmschen Kontakten

$$\begin{aligned}n &= \frac{1}{2} \left( C + \sqrt{C^2 + 4n_i^2} \right), \quad p = \frac{1}{2} \left( -C + \sqrt{C^2 + 4n_i^2} \right), \quad \text{auf } \partial\Omega_D = \bigcup_{i=1}^k K_i, \\ \psi &= \psi_{bi} + U_i, \quad \text{auf } K_i, \quad i = 1, \dots, k.\end{aligned}\tag{4.14}$$

Die angelegten Spannungen (und damit die  $U_i$ ) müssen selbstverständlich als zeitunabhängig angenommen werden. Die freie Wahl einer additiven Konstanten im Potential erlaubt uns die Einschränkung  $U_1 = 0$ . Ein stationärer Zustand eines Bauelementes mit  $k$  Kontakten ist also durch die Angabe von  $k - 1$  Parametern eindeutig bestimmt.

Die folgende Tabelle enthält alle Variablen und Parameter des Problems zusammen mit ihren Dimensionen (siehe Abschnitt 1.1), sowie typische Werte der Parameter:

Variable :	Dimension	Typischer Wert
$n, p$	$\text{cm}^{-3}$	
$\psi$	statvolt	
$\mathbf{x}$	cm	
Parameter: $\varepsilon$	1	11,7
$q$	cm statvolt	$4,8 \times 10^{-10}$
$C(\mathbf{x})$	$\text{cm}^{-3}$	0 bis $10^{20}$
$\mu_n, \mu_p$	$\text{cm}^2 \text{statvolt}^{-1} \text{s}^{-1}$	$4,5 \times 10^5; 1,8 \times 10^5$
$U_T$	statvolt	$0,86 \times 10^{-4}$
$n_i$	$\text{cm}^{-3}$	$10^{10}$
$\tau_n, \tau_p$	s	$10^{-6}, 10^{-5}$
$U_1, \dots, U_k$	statvolt	
Geometrie	cm	

Hier bezeichnet statvolt =  $\sqrt{g \text{cm}}/\text{s} = eE/\text{cm} \approx 300\text{V}$  die Einheit der Spannung im CGS-System.

Was die typischen Werte betrifft, so liegen in der Größenordnung  $10^{20} \text{cm}^{-3}$  die höchsten Werte für die Dotierung, bei denen die Materialeigenschaften des Halbleiters durch die Verunreinigung noch nicht wesentlich beeinflusst werden. Der angegebene Wert für die Temperaturspannung entspricht Zimmertemperatur (300K). Typische Werte für angelegte Spannungen und Geometrieparameter sind für verschiedene Anwendungen so unterschiedlich, daß die Angabe von speziellen Werten nicht sehr aussagekräftig wäre. Bezüglich der Geometrie wäre zu sagen, daß die kleinsten bei modernen Technologien auftretenden typischen Längen zur Zeit (1991) in der Größenordnung von  $10^{-4} \text{cm}$  liegen.

Da die Parameter teilweise durch Funktionen gegeben sind (Dotierungsprofil, Geometrie), haben wir es mit unendlich vielen Parametern zu tun.

Für die Skalierung der Dichten verwenden wir Information aus den Randbedingungen an Ohmschen Kontakten. Der Sinn der Dotierung ist es, die Leitfähigkeit zu erhöhen. Daher wird der Absolutbetrag des Dotierungsprofils in  $n$ - und  $p$ -Gebieten im allgemeinen wesentlich größer als die intrinsische Dichte  $n_i$  sein. Daraus folgt, daß für Größenordnungsüberlegungen  $n_i$  in den Randbedingungen vernachlässigt werden kann. Damit bietet sich als Referenzgröße für die Dichten der maximale Wert des Absolutbetrages des Dotierungsprofils an:

$$\bar{C} = \max_{\mathbf{x} \in \Omega} |C(\mathbf{x})|$$

Die Größenordnung der Potentialvariationen wird natürlich wesentlich von den angelegten Spannungen bestimmt. Ein charakteristischer Spannungswert ist allerdings auch durch die Temperaturspannung gegeben. Die Stromrelationen zeigen, daß für Potentialunterschiede in der Größenordnung von  $U_T$  der Diffusions- und der Konvektionsstrom etwa gleichbedeutend sind. Wir wollen uns auf diese Situation konzentrieren und die Temperaturspannung  $U_T$  als Referenzgröße für Spannungen verwenden. Das bedeutet natürlich, daß wir implizit annehmen, daß die angelegten Spannungen auch in der Größenordnung der Temperaturspannung liegen.

Als Referenzlänge verwenden wir den *Durchmesser* des Halbleitergebietes

$$L = \text{diam}(\Omega) := \sup_{\mathbf{x}_1, \mathbf{x}_2 \in \Omega} |\mathbf{x}_1 - \mathbf{x}_2|.$$

Im skalierten Problem treten die dimensionslosen Parameter

$$\lambda^2 = \frac{\varepsilon U_T}{4\pi q L^2 \bar{C}}, \quad \delta^2 = \frac{n_i}{\bar{C}}, \quad (4.15)$$

sowie

$$\frac{\mu_p}{\mu_n}, \quad \frac{U_T \mu_n}{L^2} \tau_n, \quad \frac{U_T \mu_n}{L^2} \tau_p, \quad \frac{U_i}{U_T}, \quad i = 1, \dots, k \quad (4.16)$$

auf. Die Parameter in (4.16) fassen wir als skalierte Versionen der Löcherbeweglichkeit, der Lebensdauern bzw. der angelegten Spannungen auf und werden sie daher im skalierten Problem mit  $\mu_p$ ,  $\tau_n$ ,  $\tau_p$  bzw.  $U_i$ ,  $i = 1, \dots, k$ , bezeichnen. Ebenso verwenden wir für die skalierten Variablen sowie für das mit  $\bar{C}$  skalierte Dotierungsprofil und das skalierte Halbleitergebiet dieselben Namen wie für die entsprechenden dimensionsbehafteten Größen. Damit ist die skalierte Version des Problems (4.12), (4.13), (4.14) gegeben durch die Differentialgleichungen

$$\begin{aligned} \lambda^2 \Delta \psi &= n - p - C, \\ \nabla \cdot (\nabla n - n \nabla \psi) &= R(n, p), \\ \nabla \cdot (\mu_p (\nabla p + p \nabla \psi)) &= R(n, p), \end{aligned} \quad (4.17)$$

mit der Rekombinations-Generationsrate

$$R(n, p) = \frac{np - \delta^4}{\tau_p(n + \delta^2) + \tau_n(p + \delta^2)},$$

und den Randbedingungen

$$\begin{aligned} \nabla \psi \cdot \mathbf{n} &= \nabla n \cdot \mathbf{n} = \nabla p \cdot \mathbf{n} = 0, \quad \text{auf } \partial\Omega_N, \\ n &= \frac{1}{2} \left( C + \sqrt{C^2 + 4\delta^4} \right), \quad p = \frac{1}{2} \left( -C + \sqrt{C^2 + 4\delta^4} \right), \quad \text{auf } \partial\Omega_D = \bigcup_{i=1}^k K_i, \\ \psi &= \psi_{bi} + U_i, \quad \text{auf } K_i, \quad i = 1, \dots, k. \end{aligned} \quad (4.18)$$

Das skalierte eingebaute Potential ist gegeben durch

$$\psi_{bi} = \text{arsinh} \frac{C}{2\delta^2}.$$

Betrachten wir die Größenordnungen der dimensionslosen Parameter, so ist  $\mu_p$  von der Größenordnung 1. Dasselbe nehmen wir für die skalierten angelegten Spannungen an. Die skalierten Lebensdauern in der Rekombinations-Generationsrate hängen vom Durchmesser des Halbleitergebietes ab. Wir machen die Annahme, daß auch diese Parameter keine extremen Werte annehmen.

Für den Parameter  $\delta^2$  geht schon aus den bisherigen Überlegungen hervor, daß er für jedes funktionierende Halbleiterbauelement viel kleiner als 1 sein muß.

Der Parameter  $\lambda$  ist der Quotient aus der sogenannten *Debyelänge*

$$\sqrt{\frac{\varepsilon U_T}{4\pi q \bar{C}}}$$

und der Referenzlänge  $L$ . Wir werden im Folgenden die in vielen praktischen Situationen gerechtfertigte Annahme treffen, daß auch dieser Parameter klein im Vergleich mit 1 ist.

#### 4.4. Die Strom-Spannungs-Charakteristik der Halbleiterdiode

Abb. 4.9 zeigt einen zweidimensionalen Schnitt durch eine typische Halbleiterdiode. Das  $n$ -Gebiet ist mit  $\Omega_n$ , das  $p$ -Gebiet mit  $\Omega_p$  und der  $pn$ -Übergang mit  $\Gamma$  bezeichnet. Weiters sei  $K_1$  der am  $n$ -Gebiet und  $K_2$  der am  $p$ -Gebiet angebrachte Ohmsche Kontakt.

Abb. 4.9: Zweidimensionaler Schnitt durch eine  $pn$ -Diode

Ein typischer Verlauf des Dotierungsprofils entlang des durch die strichlierte Linie in Abb. 4.9 angedeuteten eindimensionalen Schnittes wird in Abb. 4.10 gezeigt. Auffallend und wichtig für das Funktionieren des Bauelementes ist die starke Variation im Bereich des  $pn$ -Überganges. Sie legt die Betrachtung eines idealisierten Problems nahe, in dem ein stückweise konstantes Dotierungsprofil betrachtet wird. Wir machen daher die Annahme

$$C(\mathbf{x}) = \begin{cases} -C_p < 0 & \text{für } \mathbf{x} \in \Omega_p, \\ C_n > 0 & \text{für } \mathbf{x} \in \Omega_n. \end{cases} \quad (4.19)$$

Da wir ein mit der maximalen Störstellenkonzentration skaliertes Profil betrachten, ist der größere der beiden Werte  $C_n$  und  $C_p$  gleich 1.

Abb. 4.10: Dotierungsprofil entlang eines eindimensionalen Schnittes

Wir betrachten das Problem (4.17), (4.18) mit  $k = 2$ ,  $U_1 = 0$  und  $U_2 = U$ , wobei  $U$  die angelegte Spannung ist. Das Dotierungsprofil ist durch (4.19) gegeben.

Die Kontinuitätsgleichungen sind Konvektions-Diffusionsgleichungen. Ist in einer solchen Gleichung das treibende Feld ein Gradientenfeld, dann gibt es eine Transformation, die die Gleichung auf formal selbstadjungierte Form bringt. Wir ersetzen die Elektronen- und Löcherdichte mit Hilfe der Transformation

$$n = \delta^2 e^\psi u, \quad p = \delta^2 e^{-\psi} v$$

durch die neuen Variablen  $u$  und  $v$ . Das transformierte Problem besteht aus den Differentialgleichungen

$$\begin{aligned} \lambda^2 \Delta \psi &= \delta^2 e^\psi u - \delta^2 e^{-\psi} v - C, \\ \nabla \cdot (\delta^2 e^\psi \nabla u) &= R(\delta^2 e^\psi u, \delta^2 e^{-\psi} v), \\ \nabla \cdot (\mu_p \delta^2 e^{-\psi} \nabla v) &= R(\delta^2 e^\psi u, \delta^2 e^{-\psi} v), \end{aligned} \quad (4.20)$$

und den Randbedingungen

$$\begin{aligned}\nabla\psi \cdot \mathbf{n} &= \nabla u \cdot \mathbf{n} = \nabla v \cdot \mathbf{n} = 0, & \text{auf } \partial\Omega_N, \\ u = v = 1, \quad \psi &= \psi_{bi}, & \text{auf } K_1, \\ u = e^{-U}, \quad v = e^U, \quad \psi &= \psi_{bi} + U, & \text{auf } K_2.\end{aligned}\tag{4.21}$$

Die Stromdichten sind im transformierten Problem gegeben durch

$$\mathbf{J}_n = \delta^{-2} e^\psi \nabla u, \quad \mathbf{J}_p = -\mu_p \delta^{-2} e^{-\psi} \nabla v.$$

Der Faktor  $\delta^{-2}$  resultiert aus der gewählten Skalierung, die sich im Folgenden als günstig erweisen wird. Der durch das Bauelement fließende Strom  $I$  kann durch ein Integral über einen der beiden Kontakte berechnet werden. Wegen der Ladungserhaltung ist das Resultat in beiden Fällen dasselbe. Eine andere Möglichkeit ist durch das Integral

$$I = \int_{\Gamma} (\mathbf{J}_n + \mathbf{J}_p) \cdot \mathbf{n} dF\tag{4.22}$$

über den  $pn$ -Übergang gegeben. Hier bezeichnet  $\mathbf{n}$  den Einheitsnormalvektor auf  $\Gamma$ , der in das  $n$ -Gebiet  $\Omega_n$  gerichtet ist. Das Ziel dieses Abschnittes ist, eine Näherung für die Strom-Spannungs-Charakteristik, d.h. die Abhängigkeit des Stromes  $I$  von der angelegten Spannung  $U$ , zu bestimmen.

Wir wollen im Problem (4.20), (4.21) die Grenzübergänge  $\lambda \rightarrow 0$  und  $\delta \rightarrow 0$  durchführen. In Bezug auf  $\lambda$  ist die Poissongleichung singular gestört, weil die entsprechende reduzierte Gleichung eine algebraische Gleichung ist. Um reduzierte Lösungen zu berechnen, eliminieren wir das Potential mit Hilfe dieser algebraischen Gleichung. Wenn wir die Komponenten der reduzierten Lösung mit denselben Symbolen bezeichnen wie die der vollen Lösung, dann lautet das reduzierte Problem

$$\begin{aligned}\nabla \cdot \left( \frac{C + \sqrt{C^2 + 4\delta^4 uv}}{2u} \nabla u \right) &= R(n(u, v), p(u, v)), \\ \nabla \cdot \left( \mu_p \frac{-C + \sqrt{C^2 + 4\delta^4 uv}}{2v} \nabla v \right) &= R(n(u, v), p(u, v)),\end{aligned}\tag{4.23}$$

mit den Randbedingungen

$$\begin{aligned}\nabla u \cdot \mathbf{n} = \nabla v \cdot \mathbf{n} &= 0, & \text{auf } \partial\Omega_N, \\ u = v = 1, & & \text{auf } K_1, \\ u = e^{-U}, \quad v = e^U, & & \text{auf } K_2.\end{aligned}\tag{4.24}$$

Das reduzierte Potential und die Dichten sind gegeben durch

$$\begin{aligned}\psi &= \ln \frac{C + \sqrt{C^2 + 4\delta^4 uv}}{2\delta^2 u}, \\ n(u, v) &= \frac{1}{2} \left( C + \sqrt{C^2 + 4\delta^4 uv} \right), \quad p(u, v) = \frac{1}{2} \left( -C + \sqrt{C^2 + 4\delta^4 uv} \right).\end{aligned}$$

Da das Dotierungsprofil am  $pn$ -Übergang unstetig ist, sind die in (4.23) auftretenden Ableitungen im klassischen Sinn nicht definiert. Diese Gleichungen sind daher im schwachen Sinn zu verstehen. Die schwache Formulierung liefert die Aussage, daß  $u$  und  $v$  sowie die Stromdichten in Richtung der Normalen am  $pn$ -Übergang stetig sind. Das bedeutet, daß das reduzierte Potential entlang von  $\Gamma$  Sprungunstetigkeiten hat. Ebenso sind die reduzierten Ladungsträgerdichten am  $pn$ -Übergang unstetig. In Anbetracht dessen, daß das Problem (4.20), (4.21) singular gestört ist, ist dieses Resultat nicht überraschend. Zusätzlich zu den Sprüngen an  $\Gamma$  wäre auch zu erwarten, daß die Randbedingungen an den Kontakten nicht erfüllt werden können. Hier hilft allerdings die spezielle Wahl der Randbedingungen, die mit dem reduzierten Problem kompatibel ist. Das reduzierte Problem entspricht nämlich der Forderung nach dem Verschwinden der Raumladung im ganzen Bauelement. Diese Forderung ist aber gerade eine der an Ohmschen Kontakten gestellten Bedingungen. Wenn  $u$  und  $v$  Lösungen des reduzierten Problems (4.23), (4.24) sind, dann erfüllt

auch das Potential (und damit die Dichten) die Kontaktrandbedingungen. Um eine auf ganz  $\Omega$  gleichmäßig gültige Näherung für das Potential zu bestimmen, müßten wir ein Grenzschichtproblem in der Nähe von  $\Gamma$  lösen. Die Kenntnis dieser Grenzschichtlösung ist allerdings für die näherungsweise Berechnung der Ströme nicht notwendig, und wir verzichten daher darauf.

Nun ist auch die Bedeutung der Transformation auf die Variablen  $u$  und  $v$  klar. Während im ursprünglichen Problem sowohl das Potential als auch die Dichten für  $\lambda \rightarrow 0$  gegen unstetige Funktionen konvergieren, bleiben  $u$  und  $v$  auch nach dem Grenzübergang stetig. Sprünge treten nur in ihren ersten Ableitungen auf. Da  $u$  und  $v$  im Vergleich zum Potential und den Dichten im Bereich der Grenzschicht langsam variieren, werden sie als *langsame Variable* bezeichnet (im Gegensatz zu den *schnellen Variablen*  $\psi$ ,  $n$  und  $p$ ). Hätten wir die Transformation nicht durchgeführt, dann müßten wir für die reduzierten Variablen  $\psi$ ,  $n$  und  $p$  Sprungbedingungen am  $pn$ -Übergang mit Hilfe von Grenzschichtlösungen herleiten. Die Kenntnis der langsamen Variablen erspart uns diese Arbeit. Die Sprünge über  $\Gamma$  müssen so beschaffen sein, daß  $e^{-\psi}n$  und  $e^{\psi}p$  stetig sind.

Wir wollen nun den Grenzübergang  $\delta \rightarrow 0$  durchführen. Dazu schreiben wir zunächst die reduzierten Differentialgleichungen (4.23) als ein System von Gleichungen erster Ordnung:

$$\begin{aligned}\delta^4 \mathbf{J}_n &= \frac{C + \sqrt{C^2 + 4\delta^4 uv}}{2u} \nabla u, \\ \delta^4 \mathbf{J}_p &= -\mu_p \frac{-C + \sqrt{C^2 + 4\delta^4 uv}}{2v} \nabla v, \\ \nabla \cdot \mathbf{J}_n &= -\nabla \cdot \mathbf{J}_p = \frac{uv - 1}{\tau_p(n(u, v) + \delta^2) + \tau_n(p(u, v) + \delta^2)}.\end{aligned}\quad (4.25)$$

Der Grenzübergang führt im  $n$ - und  $p$ -Gebiet jeweils auf verschiedene Resultate. Lassen wir die Namen der Variablen wieder unverändert, so ergibt sich für die Grenzwerte

$$\begin{aligned}\nabla u &= 0, \quad \mathbf{J}_p = -\frac{\mu_p u}{C_n} \nabla v, \quad -\nabla \cdot \mathbf{J}_p = \frac{uv - 1}{\tau_p C_n}, \quad \text{in } \Omega_n, \\ \nabla v &= 0, \quad \mathbf{J}_n = \frac{v}{C_p} \nabla u, \quad \nabla \cdot \mathbf{J}_n = \frac{uv - 1}{\tau_n C_p}, \quad \text{in } \Omega_p.\end{aligned}\quad (4.26)$$

Aus (4.26) und den Randbedingungen folgt unmittelbar

$$u = 1 \quad \text{in } \Omega_n, \quad v = e^U \quad \text{in } \Omega_p. \quad (4.27)$$

Damit können  $v$  in  $\Omega_n$  und  $u$  in  $\Omega_p$  als Lösungen der linearen Randwertprobleme

$$\begin{aligned}\Delta v &= \frac{v - 1}{\tau_p \mu_p} \quad \text{in } \Omega_n, \\ v &= 1 \quad \text{auf } K_1, \quad v = e^U \quad \text{auf } \Gamma, \\ \nabla v \cdot \mathbf{n} &= 0 \quad \text{auf } \partial\Omega_N \cap \partial\Omega_n,\end{aligned}\quad (4.28)$$

und

$$\begin{aligned}\Delta u &= \frac{u - e^{-U}}{\tau_n} \quad \text{in } \Omega_p, \\ u &= e^{-U} \quad \text{auf } K_2, \quad u = 1 \quad \text{auf } \Gamma, \\ \nabla u \cdot \mathbf{n} &= 0 \quad \text{auf } \partial\Omega_N \cap \partial\Omega_p,\end{aligned}\quad (4.29)$$

berechnet werden. Mit Hilfe des Superpositionsprinzips schreiben wir  $v$  und  $u$  als

$$v = 1 + (e^U - 1)\phi_v \quad \text{in } \Omega_n, \quad u = e^{-U} + (1 - e^{-U})\phi_u \quad \text{in } \Omega_p, \quad (4.30)$$

wobei für die Referenzfunktionen  $\phi_v$  und  $\phi_u$

$$\Delta \phi_v = \frac{\phi_v}{\tau_p \mu_p} \quad \text{in } \Omega_n, \quad \phi_v = 0 \quad \text{auf } K_1, \quad \phi_v = 1 \quad \text{auf } \Gamma,$$

und

$$\Delta\phi_u = \frac{\phi_u}{\tau_n} \quad \text{in } \Omega_p, \quad \phi_u = 0 \quad \text{auf } K_2, \quad \phi_u = 1 \quad \text{auf } \Gamma,$$

gilt. Das Interessante an der Darstellung (4.30) ist, daß die Referenzfunktionen unabhängig von der angelegten Spannung sind. Aus (4.30) ist also die Abhängigkeit von  $u$  und  $v$  von der Spannung direkt abzulesen. Die linearen Probleme für die Referenzfunktionen müssen für jede Diode nur einmal gelöst werden. Für die Stromdichten ergibt sich

$$\begin{aligned} \mathbf{J}_p &= -\frac{\mu_p}{C_n}(e^U - 1)\nabla\phi_v \quad \text{in } \Omega_n, \\ \mathbf{J}_n &= \frac{1}{C_p}(e^U - 1)\nabla\phi_u \quad \text{in } \Omega_p. \end{aligned}$$

Nun zeigt sich, warum wir zur Berechnung des Gesamtstromes in (4.22) ein Integral über  $\Gamma$  gewählt haben. Nach unserer bisherigen Rechnung ist  $\Gamma$  die einzige Fläche, entlang der wir sowohl die Elektronen- als auch die Löcherstromdichte kennen. Aus (4.22) erhalten wir die Strom-Spannungs-Charakteristik

$$I = I_s(e^U - 1) \tag{4.31}$$

mit

$$I_s = \int_{\Gamma} \left( \frac{1}{C_p}\nabla\phi_u - \frac{\mu_p}{C_n}\nabla\phi_v \right) \cdot \mathbf{n} dF.$$

Unser Resultat (4.31) ist die *Shockley-Gleichung*. Sie enthält die erstaunliche Aussage, daß sich die Charakteristiken für zwei verschiedene Dioden nur durch den Wert von  $I_s$  unterscheiden. Die Diode erfüllt ihre Funktion als Ventil. Während für positive angelegte Spannungen der Strom exponentiell mit der Spannung wächst, ist der Strom in Sperrichtung durch den *Sperrstrom*  $I_s$  begrenzt.

Eine weitergehende Analyse zeigt, daß die berechnete Charakteristik sowohl für stark positive als auch für stark negative Spannungen ihre Gültigkeit verliert. Es zeigt sich, daß für große positive Spannungen das exponentielle Wachstum des Stromes in ein algebraisches Wachstum übergeht. Im Bereich einer gewissen stark negativen Spannung beginnt der Strom plötzlich stark anzusteigen. Man sagt, die Diode verliert dann ihre Sperrfähigkeit.



## 5. KAPITEL Hindernisprobleme – Variationsungleichungen

### 5.1. Der endlichdimensionale Fall

Wir betrachten ein System von  $N$  Massenpunkten, die sich in der  $(u, x)$ -Ebene nur parallel zur  $u$ -Achse bewegen können (transversale Auslenkungen). Die Auslenkung des  $k$ -ten Massenpunktes zum Zeitpunkt  $t$  ist gegeben durch  $(u_k(t), x_k)$ ,  $k = 1, \dots, N$ . Dabei nehmen wir die  $x_k$  als äquidistant verteilt an:  $x_{k+1} - x_k = \Delta x$ . Die beiden äußersten Punkte halten wir fest:  $u_0 = u_{N+1} = 0$ .

Die Massenpunkte seien durch Federn verbunden sowie der Wirkung von zeitunabhängigen äußeren Kräften ausgesetzt. Damit ergibt sich für die auf den  $k$ -ten Massenpunkt wirkende Kraft:

$$F_k = \kappa(u_{k+1} - u_k + u_{k-1} - u_k) + f_k, \quad k = 1, \dots, N.$$

Haben die Punkte alle dieselbe Masse  $m$ , dann erfüllen die  $u_k(t)$  die Differentialgleichungen  $m\ddot{u}_k = F_k$ ,  $k = 1, \dots, N$ . Eine potentielle Energie  $V(u_1, \dots, u_N)$  mit der Eigenschaft  $\frac{\partial V}{\partial u_k} = -F_k$  ist gegeben durch

$$V = \sum_{k=0}^N \left( \kappa \frac{(u_{k+1} - u_k)^2}{2} - f_k u_k \right).$$

Für das zeitabhängige Problem gilt, daß die Gesamtenergie

$$\sum_{k=1}^N \frac{m\dot{u}_k^2}{2} + V$$

unabhängig von der Zeit ist. Stationäre Zustände kann man als stationäre Punkte der potentiellen Energie  $V$  berechnen.

$$\text{Notation: } L^2\text{-Norm } \|u\|_2^2 = \sum_{k=1}^N u_k^2. \quad \|\delta u\|_2^2 = \sum_{k=0}^N (u_{k+1} - u_k)^2$$

**Lemma 5.1.** (*Poincare-Ungleichung*) Sei  $u = (u_1, \dots, u_N) \in \mathbb{R}^N$  und  $u_1 = u_{N+1} = 0$ . Dann gilt

$$\|u\|_2 \leq N \|\delta u\|_2$$

*Beweis.*

$$\|u\|_2^2 = \sum_{k=1}^N \left( \sum_{l=0}^{k-1} (u_{l+1} - u_l) \right)^2 \leq N \left( \sum_{l=0}^N |u_{l+1} - u_l| \right)^2.$$

Letzter Schritt: Cauchy-Schwarz. ■

**Lemma 5.2.** (*Koerzivität*) Für  $0 \leq \gamma \leq 1$  und die oben definierte potentielle Energie  $V(u)$  gilt

$$\gamma V(u) + (1 - \gamma)V(v) - V(\gamma u + (1 - \gamma)v) \geq \frac{\kappa\gamma(1 - \gamma)}{2N^2} \|u - v\|_2^2$$

*Beweis.* Explizite Rechnung und Lemma 5.1. ■

**Korollar 5.3.**  $V$  ist streng konvex.

**Lemma 5.4.** Die oben definierte potentielle Energie  $V(u)$  ist nach unten beschränkt.

*Beweis.* Aus Lemma 5.1. und Cauchy-Schwarz folgt

$$V(u) \geq \frac{\kappa}{2N^2} \|u\|_2^2 - \|f\|_2 \|u\|_2 \geq -\frac{N^2 \|f\|_2^2}{2\kappa}.$$

■

**Lemma 5.5.** Sei  $V(u)$  konvex und stetig differenzierbar. Dann ist jeder stationäre Punkt ein globales Minimum.

*Beweis.* Sei  $\nabla V(v) = 0$  und  $u \in \mathbb{R}^N$ . Aus der Konvexität folgt

$$V(u) - V(v) \geq \frac{V(v + \gamma(u - v)) - V(v)}{\gamma},$$

und daraus mit  $\gamma \rightarrow 0$ :  $V(v) \leq V(u)$ . ■

**Satz 5.6.** Sei  $V : \mathbb{R}^N \rightarrow \mathbb{R}$  stetig, koerziv (Lemma 5.2) und nach unten beschränkt. Dann besitzt  $V$  ein eindeutiges globales Minimum in  $\mathbb{R}^N$ .

*Beweis.* Wegen der Beschränktheit gibt es eine Folge  $u^n$  mit  $\lim_{n \rightarrow \infty} V(u^n) = \inf_{v \in \mathbb{R}^N} V(v) = \underline{V}$ . Aus der Konvexität folgt, dass auch  $V(\frac{u^m + u^n}{2})$  für  $m, n$  groß genug beliebig nahe bei  $\underline{V}$  liegt. Die Koerzivität (mit  $u = u^m, v = u^n, \gamma = 1/2$ ) impliziert nun, dass  $u^n$  eine Cauchyfolge ist. Wegen der Stetigkeit von  $V$  gilt für den Grenzwert  $u$ :  $V(u) = \underline{V}$ . Die Eindeutigkeit ist eine direkte Konsequenz aus der strengen Konvexität. ■

Anwendung des Satzes für die oben definierte potentielle Energie zeigt, dass es einen eindeutigen stationären Punkt gibt, der durch Lösung eines Minimierungsproblems ermittelt werden kann. Dieses Prinzip soll nun auch zur Formulierung eines *Hindernisproblems* verwendet werden: Angenommen, die Bewegung der  $N$  Massenpunkte sei durch Hindernisse  $\psi_1, \dots, \psi_N$  eingeschränkt:  $u_k \geq \psi_k, k = 1, \dots, N$ . Dann definieren wir die Menge der erlaubten Zustände  $\mathbb{K} = \{u \in \mathbb{R}^N : u_k \geq \psi_k, k = 1, \dots, N\}$  und stationäre Zustände  $u$  durch Lösung des Problems

$$V(u) = \min_{v \in \mathbb{K}} V(v).$$

Es ist leicht zu sehen, dass  $\mathbb{K}$  konvex und abgeschlossen ist. Mit dem Beweis von Satz 5.6 zeigt man auch die folgende Erweiterung:

**Satz 5.7.** Sei  $B$  ein Banachraum und  $\mathbb{K} \subset B$  konvex und abgeschlossen. Sei  $V : \mathbb{K} \rightarrow \mathbb{R}$  stetig, koerziv (Lemma 5.2) und nach unten beschränkt. Dann besitzt  $V$  ein eindeutiges globales Minimum in  $\mathbb{K}$ .

Betrachten wir nun das Minimierungsproblem im eindimensionalen Fall mit  $\mathbb{K} = [a, b]$  und differenzierbarem  $V : [a, b] \rightarrow \mathbb{R}$ . Für die Minimalstelle  $u \in [a, b]$  gibt es dann 3 Möglichkeiten:

- a)  $a < u < b$  und  $V'(u) = 0$ ,
- b)  $u = a$  und  $V'(u) \geq 0$ ,
- c)  $u = b$  und  $V'(u) \leq 0$ .

In allen 3 Fällen ist die *Variationsungleichung*

$$V'(u)(v - u) \geq 0 \quad \forall v \in \mathbb{K}$$

erfüllt.

Im  $N$ -dimensionalen Fall ist für die Minimalstelle  $u$  und ein beliebiges anderes  $v \in \mathbb{K}$  die eindimensionale Funktion  $f(t) = V(u + t(v - u))$  auf dem Intervall  $[0, 1]$  definiert und nimmt an  $t = 0$  ihren minimalen Wert an. Die notwendige Bedingung  $f'(0) \geq 0$  führt auf die  $N$ -dimensionale Variationsungleichung

$$\nabla V(u) \cdot (v - u) \geq 0 \quad \forall v \in \mathbb{K}. \tag{5.1}$$

**Lemma 5.8.** Sei  $\mathbb{K} \subset \mathbb{R}^N$  konvex und  $V : \mathbb{K} \rightarrow \mathbb{R}$  stetig differenzierbar und streng konvex. Dann hat (5.1) höchstens eine Lösung.

*Beweis.* Aus der strengen Konvexität von  $V$  folgt die *strenge Monotonie* des Gradientenfeldes:

$$(\nabla V(v) - \nabla V(w)) \cdot (v - w) > 0, \quad \text{für } v \neq w.$$

Seien nun  $u_1 \neq u_2$  Lösungen von (5.1). Betrachten wir die Ungleichung (5.1) mit  $u = u_1$  und  $v = u_2$  sowie vice versa. Die Summe dieser Ungleichungen gibt einen Widerspruch zur strengen Monotonie von  $\nabla V$ . ■

Das letzte Resultat zeigt, dass für streng konvexe  $V$  das Minimierungsproblem und die Variationsungleichung äquivalent sind.

Für das Hindernisproblem ergibt sich die Variationsungleichung: Gesucht  $u \in \mathbb{K}$  mit

$$\sum_{k=1}^N (\kappa(2u_k - u_{k+1} - u_{k-1}) - f_k)(v_k - u_k) \geq 0,$$

für alle  $v \in \mathbb{K}$ . Setzen wir  $v_k = u_k + \varepsilon \delta_{kl}$ , so folgt

$$\varepsilon(\kappa(2u_l - u_{l+1} - u_{l-1}) - f_l) \geq 0. \quad (5.2)$$

Nun zerlegen wir die Indexmenge  $\{1, \dots, N\}$  in 2 Teile: Die *Koinzidenzmenge* (oder Kontaktmenge)  $K$ , bestehend aus allen  $k \in \{1, \dots, N\}$ , für die  $u_k = \psi_k$  gilt, und die *Nonkoinzidenzmenge*  $N = \{k : u_k > \psi_k\}$ . Ist  $l$  in der Nonkoinzidenzmenge, dann sind in (5.2) kleine  $\varepsilon$  mit beiden Vorzeichen zugelassen, was die Konsequenz

$$\kappa(2u_l - u_{l+1} - u_{l-1}) - f_l = 0, \quad l \in N,$$

hat, d.h. die Gleichung für das System ohne Hindernis. Für  $l \in K$  sind in (5.2) nur nichtnegative  $\varepsilon$  zugelassen mit der Konsequenz

$$\kappa(2\psi_l - u_{l+1} - u_{l-1}) - f_l \geq 0, \quad l \in K.$$

Ist z.B. die Höhe der benachbarten Hindernisse gleich, d.h.  $\psi_{l+1} = \psi_{l-1} = \psi_l$ , so folgt daraus die notwendige Bedingung  $f_l \leq 0$  für  $l \in K$ .

## 5.2. Das Hindernisproblem für Saite und Membran

Wir konstruieren einen kontinuierlichen Grenzfall für das endlichdimensionale Hindernisproblem, indem der maximale  $x$ -Wert  $L = (N + 1)\Delta x$  während des Grenzüberganges  $\Delta x \rightarrow 0$ ,  $N \rightarrow \infty$  festgehalten wird. Weiters ersetzen wir  $\kappa$  durch  $\kappa/\Delta x$  und setzen  $f_k = f(k\Delta x)\Delta x$ ,  $u_k = u(k\Delta x)$ ,  $\psi_k = \psi(k\Delta x)$ . Durch den Grenzübergang entsteht die neue potentielle Energie

$$V = \int_0^L \left( \kappa \frac{u'(x)^2}{2} - f(x)u(x) \right) dx.$$

Die Randbedingungen packen wir zusammen mit der Einschränkung durch das Hindernis  $\psi$  in die Definition der zulässigen Menge, die außerdem nur Funktionen enthalten darf, für die die potentielle Energie ausgewertet werden kann:

$$\mathbb{K} = \{u \in H_0^1((0, L)) : u \geq \psi\}.$$

Dabei machen wir für das Hindernis die Annahme  $\psi \in H^1((0, L))$  und für die Kraftdichte  $f \in L^2((0, L))$ . Eine Verallgemeinerung für mehrere Raumdimensionen ergibt sich, wenn wir das Intervall  $(0, L)$  durch ein beschränktes Gebiet  $\Omega \subset \mathbb{R}^d$  ersetzen:

$$V = \int_{\Omega} \left( \kappa \frac{|\nabla u(x)|^2}{2} - f(x)u(x) \right) dx,$$

$$\mathbb{K} = \{u \in H_0^1(\Omega) : u \geq \psi\},$$

mit den entsprechenden Annahmen an die Daten:  $\psi \in H^1(\Omega)$ ,  $f \in L^2(\Omega)$ . Um zu garantieren, dass  $\mathbb{K}$  nichtleer ist, ist auch die Annahme  $\psi \leq 0$  auf  $\partial\Omega$  notwendig. Mit Hilfe der kontinuierlichen Poincaré-Ungleichung

$$\|u\|_{L^2(\Omega)} \leq c \|\nabla u\|_{L^2(\Omega)}, \quad \forall u \in H_0^1(\Omega),$$

lassen sich die Voraussetzungen von Satz 5.7 zeigen, wobei  $B = H_0^1(\Omega)$  gewählt werden kann. Damit ist das Minimierungsproblem für  $V$  über  $\mathbb{K}$  eindeutig lösbar. Für die Lösung  $u \in \mathbb{K}$  gilt die Variationsungleichung

$$\int_{\Omega} (\kappa \nabla u \cdot \nabla(v - u) - f(v - u)) dx \geq 0, \quad \forall v \in \mathbb{K}. \quad (5.3)$$

Analog zum endlichdimensionalen Fall zeigt man die Eindeutigkeit der Lösung der Variationsungleichung.

Die nun folgenden Argumente sind formal in dem Sinn, dass ein gewisses (nicht bewiesenes) Maß an Glattheit der Lösung  $u$  des Problems angenommen wird. Der Beweis eines entsprechenden Glattheitsresultates würde den Rahmen dieser Vorlesung sprengen. Auch für das Hindernis  $\psi$  und die Kraftdichte  $f$  sind zusätzliche Annahmen notwendig.

Sind  $u$  und  $\psi$  stetig, dann gibt es eine (relativ zu  $\Omega$ ) abgeschlossene Koinzidenzmenge  $K = \{x \in \Omega : u(x) = \psi(x)\}$  und eine offene Nonkoinzidenzmenge  $N = \{x \in \Omega : u(x) > \psi(x)\}$ . Für  $\phi \in C_0^\infty(N)$  gilt dann  $v = u + \varepsilon\phi \in \mathbb{K}$  für  $|\varepsilon|$  klein genug. Die Variationsungleichung impliziert

$$\int_{\Omega} (\kappa \nabla u \cdot \nabla \phi - f\phi) dx \geq 0, \quad \forall \phi \in C_0^\infty(N),$$

d.h. die schwache Formulierung der Differentialgleichung

$$\kappa \Delta u + f = 0 \quad \text{in } N. \quad (5.4)$$

Wählen wir  $\phi \in C_0^\infty(K)$ ,  $\phi \geq 0$ , dann gilt  $v = u + \phi \in \mathbb{K}$  und daher

$$\int_{\Omega} (\kappa \nabla u \cdot \nabla \phi - f\phi) dx \geq 0, \quad \forall \phi \in C_0^\infty(K), \phi \geq 0,$$

d.h. die schwache Formulierung der Ungleichung

$$\kappa \Delta \psi + f \leq 0 \quad \text{in } K. \quad (5.5)$$

Für  $d = 1$  bzw.  $d = 2$  kann die Gleichung (5.4) als Gleichung für stationäre, transversale, kleine Auslenkungen einer gespannten Saite bzw. Membran interpretiert werden, und damit (5.3) als Formulierung des Hindernisproblems für die Saite bzw. Membran. Ähnlich zum endlichdimensionalen Fall erhält man für ein konstantes Hindernis  $\psi$  die Bedingung  $f \leq 0$  in  $K$ .

Teil der Lösung des Problems ist die Bestimmung von  $K$  und  $N$  bzw. des *freien Randes*  $\Gamma = K \cap \overline{N}$ . Ist dieser bekannt, dann kann  $u$  durch Lösung des Dirichletproblems (5.4) mit den Randbedingungen

$$u = \psi \quad \text{auf } \Gamma, \quad u = 0 \quad \text{auf } \partial N \cap \partial \Omega, \quad (5.6)$$

vollständig ermittelt werden. Nimmt man an, dass  $u$  und  $\psi$  stetig differenzierbar sind, dann muss die zusätzliche Randbedingung

$$\nabla u = \nabla \psi \quad \text{auf } \Gamma \quad (5.7)$$

gelten, die für glattes  $\Gamma$  durch eine Neumannbedingung ersetzt werden kann, nämlich dass die Normalableitungen von  $u$  und  $\psi$  auf  $\Gamma$  übereinstimmen. Die zusätzliche Bedingung (5.7) legt den freien Rand fest. Insgesamt könnte (5.4)–(5.7) als starke Formulierung der Variationsungleichung (5.3) bezeichnet werden.

### 5.3. Die Halbleiterdiode im thermischen Gleichgewicht

Das Equilibriumpotential in der Halbleiterdiode wird bestimmt durch das Problem

$$\begin{aligned}\lambda^2 \Delta \psi &= \delta^2 e^\psi - \delta^2 e^{-\psi} - C, \\ \psi &= \ln \frac{C + \sqrt{C^2 + 4\delta^4}}{2\delta^2} \quad \text{auf } K_1 \cup K_2, \quad \nabla \psi \cdot \mathbf{n} = 0 \quad \text{auf } \partial\Omega_N.\end{aligned}$$

Hier führen wir die neuen Parameter  $\gamma = (\ln \delta^{-2})^{-1}$  und  $\lambda_0^2 = \lambda^2/\gamma$  ein sowie die Umskalierung des Potentials durch  $\psi \rightarrow \psi/\gamma$ . Dadurch erhält das obige Problem die Form

$$\begin{aligned}\lambda_0^2 \Delta \psi &= \exp\left(\frac{\psi-1}{\gamma}\right) - \exp\left(\frac{-\psi-1}{\gamma}\right) - C, \\ \psi &= 1 + \gamma \ln \frac{C_n + \sqrt{C_n^2 + 4e^{-2/\gamma}}}{2} \quad \text{auf } K_1, \quad \psi = -1 - \gamma \ln \frac{C_p + \sqrt{C_p^2 + 4e^{-2/\gamma}}}{2} \quad \text{auf } K_2, \\ \nabla \psi \cdot \mathbf{n} &= 0 \quad \text{auf } \partial\Omega_N.\end{aligned}$$

Im Folgenden sind wir interessiert am Grenzübergang  $\gamma \rightarrow 0$  für festes  $\lambda_0^2 > 0$ . Mit Hilfe des Maximumprinzips zeigt man

$$-1 - \gamma \ln \frac{C_p + \sqrt{C_p^2 + 4e^{-2/\gamma}}}{2} \leq \psi \leq 1 + \gamma \ln \frac{C_n + \sqrt{C_n^2 + 4e^{-2/\gamma}}}{2} \quad \text{in } \Omega.$$

Die Dichten

$$n = \exp\left(\frac{\psi-1}{\gamma}\right), \quad p = \exp\left(\frac{-\psi-1}{\gamma}\right)$$

sind daher gleichmäßig in  $\gamma$  beschränkt, woraus die gleichmäßige Beschränktheit von  $\Delta \psi$  folgt. Mit Hilfe von Glattheitsresultaten für elliptische Gleichungen und dem Satz von Arzela und Ascoli kann man daraus folgern, daß die Grenzfunktion  $\psi_0 = \lim_{\gamma \rightarrow 0} \psi$  zumindest im Inneren von  $\Omega$  stetig differenzierbar ist.

Weiters folgt aus obiger Abschätzung, daß  $-1 \leq \psi_0 \leq 1$  gilt. Das Gebiet kann daher zerlegt werden in  $\Omega = K^+ \cup K^- \cup N$  mit

$$K^+ = \{x \in \Omega : \psi_0(x) = 1\}, \quad K^- = \{x \in \Omega : \psi_0(x) = -1\}, \quad N = \{x \in \Omega : -1 < \psi_0(x) < 1\}.$$

Aus  $\psi_0 = 1$  auf  $K_1$  und  $\psi_0 = -1$  auf  $K_2$  folgt, daß  $N$  nichtleer ist. In  $N$  verschwinden die Grenzdichten  $n_0$  und  $p_0$ , und daher gilt

$$\lambda_0^2 \Delta \Psi_0 = -C \quad \text{in } N. \quad (5.1)$$

Weiters zeigt man leicht

$$\begin{aligned}n_0 = C_n, \quad p_0 = 0 & \quad \text{in } K^+ \subset \Omega_n, \\ n_0 = 0, \quad p_0 = C_p & \quad \text{in } K^- \subset \Omega_p.\end{aligned}$$

Den Rand des Teilgebietes  $N$  zerlegen wir nun in  $\partial N = \Gamma^+ \cup \Gamma^- \cup (\partial N \cap \partial\Omega)$  mit  $\Gamma^+ = \overline{N} \cap \overline{K^+}$  und  $\Gamma^- = \overline{N} \cap \overline{K^-}$ . Dann gilt

$$\psi_0 = 1 \quad \text{auf } \Gamma^+, \quad \psi_0 = -1 \quad \text{auf } \Gamma^-,$$

sowie die ursprünglichen Randbedingungen auf  $\partial N \cap \partial\Omega$ . Wäre  $N$  bekannt, so könnte  $\psi_0$  aus diesen Randbedingungen und der Gleichung (5.1) berechnet werden.

Wir kennen zwar einerseits  $N$  noch nicht, haben aber andererseits noch Information zur Verfügung, nämlich die stetige Differenzierbarkeit von  $\psi_0$ . Sie impliziert

$$\nabla \psi_0 = 0 \quad \text{auf } \Gamma^+ \cup \Gamma^-.$$

Diese zusätzlichen Randbedingungen bestimmen die **freien Ränder**  $\Gamma^+$  und  $\Gamma^-$ .

Im eindimensionalen Fall mit  $\Omega_p \subset (-\infty, 0)$  und  $\Omega_n \subset (0, \infty)$  kann das Problem explizit gelöst werden:

$$\psi_0(x) = \begin{cases} -1 + \frac{C_p(x-\Gamma^-)^2}{2\lambda_0^2}, & \Gamma^- < x < 0, \\ 1 - \frac{C_n(x-\Gamma^+)^2}{2\lambda_0^2}, & 0 < x < \Gamma^+, \end{cases} \quad \Gamma^- = -2\lambda_0 \sqrt{\frac{C_n}{C_p(C_n + C_p)}}, \quad \Gamma^+ = 2\lambda_0 \sqrt{\frac{C_p}{C_n(C_n + C_p)}}.$$