Entropie und Richtung der Zeit

Material zu Kurs 6.3 der SchülerAkademie 2005 in Rostock

Thomas Neusius

Inhaltsverzeichnis

1 Grundlagen der Mechanik

In der klassischen Mechanik waren die Größen Zeit t, Länge x oder im dreidimensionalen $\vec{x}$ und Masse m gegeben. Man ging davon aus, daß Zeit und Raum beliebig unterteilbar sind, daß sie also *kontinuierlich* sind. Deswegen kann man die Zeit mit den reelen Zahlen $\mathbb{R}$ beschreiben. Den Ort eines Teilchens, von dem man in den theoretischen Überlegungen annimmt, es habe keine räumliche Ausdehnung, kann man nun zu allen Zeitpunkten angeben und erhählt so eine Ort-Zeit-Funktion $\vec{x}(t)$ für das Teilchen. Da sich der Ort nicht sprunghaft ändert und die Zeit durch die reelen Zahlen beschrieben wird, kann man nun die Ableitung der Funktion $\vec{x}(t)$ nach der Zeit betrachten. Sie ist ein Maß für die Stärke der Änderung des Ortes in einem bestimmten Moment. Dies bezeichnet man in der Leibnizschen Schreibweise[1] als

$$\frac{\mathrm{d}x}{\mathrm{d}t}.$$

Die Ableitung der Ort-Zeit-Funktion nach der Zeit nennen wir die Geschwindigkeit

$$\vec{v}(t) := \frac{\mathrm{d}x}{\mathrm{d}t},$$

die Ableitung der Geschwindiglkeit nach der Zeit wird als Beschleunigung bezeichnet

$$\vec{a}(t) := \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}^2x}{\mathrm{d}t^2}.$$

Außerdem definiert man den Impuls als das Produkt aus Geschwindigkeit und Masse eines Teilchens.

$$\vec{p} = m\vec{v}.$$

Nach dem Trägheitssatz bewegen sich Körper, die eine Masse besitzen mit konstanter Geschwindigkeit auf einer geraden Bahn, solange, bis eine Kraft auf sie einwirkt und ihre Bahn oder Geschwindigkeit ändert. Es ist eines der wichtigsten Gesetze der Physik, das zweite Newtonsche Gesetz, das besagt: Eine Kraft ist das Produkt aus Masse und Beschleunigung

$$\vec{K} = m\vec{a}.$$

[1] In der Schule schreibt man meist $f'(x)$ für die Ableitung von $f(x)$ nach x, die wir als $\frac{\mathrm{d}f}{\mathrm{d}x}$ bezeichnen werden.

Dies bedeutet, die Masse ist die Eigenschaft, die sich einer Änderung des Bewegungszustandes (Änderung der Geschwindigkeit bzw. der Bewegungsrichtung) widersetzt. Die Einheit der Kraft ist das Newton [N=kg m/s^2]. Auf eine 100g-Tafel Schokolade wirkt am Erdboden eine Kraft von ca. 1N. Beispiele für Kräfte sind die Erdanziehung, die Anziehung oder Abstoßung elektrisch geladener Körper oder die Reibungskraft, die proportional zur Geschwindigkeit ist

$$\vec{K} = -m\zeta\vec{v}.$$

Dabei bezeichnet ζ den Reibungskoeffizienten, eine Eigenschaft der verwendeten Materialien, m die Masse und $\vec{v}$ die Geschwindigkeit.

Die Energie ist das Produkt aus Kraft und Weg, über den die Kraft wirkt. Genauer: Auf einen 10kg schweren Koffer wirkt am Erdboden eine Kraft von 981 Newton. Hebe ich diesen Koffer um einen Meter an, so habe ich eine Arbeit von 981 Joule [J=Nm] verrichtet. Wenn ich aber den Koffer einen Meter nach rechts verschiebe, habe ich keine Hubarbeit geleistet, da ich nicht gegen die Schwerkraft gearbeitet habe. Wenn die Kraft vom Ort abhängt, muß man zur Berechnung der Energie ein Integral verwenden. Dies bedeutet, man nimmt ganz kleine Wegstücke dx auf denen man sich die Kraft als konstant denkt, dann multipliziert man diese Wegstücke mit der dortigen Kraft und summiert die vielen kleinen Energien auf zur Gesamtenergie

$$E = \int \vec{K}(\vec{x})\mathrm{d}\vec{x}.$$

Es gibt viele verschiedene Energieformen: Eine Energieform ist die Lageenergie, meist als potentielle Energie bezeichnet. Ein Körper, der auf einem hohen Berg liegt hat diese Energie. Sie berechnet sich in der Nähe des Erdbodens als Produkt aus der Erdbeschleunigung $g = 9,81 m/s^2$, der Höhe h und der Masse des Körpers m zu

$$E_{pot} = mgh.$$

Wenn man ihn vom Berg wirft, dann verliert er die Lageenergie und seine Geschwindigkeit nimmt zu. Dies entspricht einer anderen Energieform, der kinetischen Energie

$$E_{kin} = \frac{mv^2}{2},$$

mit der Geschwindigkeit v.

2 Differentialgleichungen

Weil wir den Ort, die Geschwindigkeit und andere Eigenschaften von Teilchen als Funktionen der Zeit definiert haben, ist es nun Ziel, Zusammenhänge zwischen diesen Größen zu finden. Zum Beispiel können wir angeben, wie groß die Anziehungskraft zwischen zwei Massen in einem bestimmten Abstand oder die elektrische Kraft ist. Wenn wir die Kräfte kennen, möchten wir gerne berechnen, wie die Bewegungen der Teilchen oder Körper unter diesen Kräften aussehen. Ein einfaches Beispiel ist eine Feder.

Wir wissen, daß Feder eine Masse mit der Kraft $K = -kx$ in Richtung Ruhelage zieht, wenn sich die Masse in der Entfernung x von dieser Ruhelage befindet. Nach Newtons Gesetz für die Kraft heißt das

$$m\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} = -kx(t).$$

Das gibt uns einen Zusammenhang zwischen der Beschleunigung und dem Ort, beide als Funktionen der Zeit. Da aber die Beschleunigung auch die zweite Ableitung des Ortes nach der Zeit ist, können die beiden nicht völlig unabhängig sein. Eine Beziehung wie hier, die die Ableitungen einer Größe mit anderen Größen oder anderen, niedrigeren Ableitungen verknüpft, heißt eine *Differentialgleichung*. Viele wichtigen Beziehungen in der Physik bestehen aus solchen Gleichungen. Ziel ist es dann, eine Funktion (hier $x(t)$) zu finden, die so beschaffen ist, daß die Gleichung gilt. Dafür gibt es einige Tricks, aber es ist oft gar nicht möglich, eine Lösung zu finden. Manchmal braucht man auch einfach Glück, um die richtige Lösung zu erraten. Hat man aber die Lösung, so kann man leicht feststellen, ob diese richtig ist, denn Ableiten ist viel einfacher als das Lösen von Differentialgleichungen.

Die Lösung für die Gleichung hier ist einfach eine Sinus- oder Kosinusfunktion

$$x(t) = A\sin(\sqrt{\frac{k}{m}}t).$$

Die Amplitude A, das bedeutet die maximale Auslenkung der Masse, ist in dieser Lösung jedoch noch nicht enthalten. Wenn man nur die Differentialgleichung kennt, genügt dies also doch nicht für die ganze Lösung. Hier könnten wir jedes A wählen. Erst wenn wir noch zu einem Zeitpunkt den Ort und die Geschwindigkeit der Masse an der Feder kennen, können wir A ausrechnen und damit die Bewegung der Masse für alle Zukunft vorher sagen. Dies ist ein wichtiges Konzept der klassischen Physik: Bei Kenntnis eines Anfangszustandes und aller Kräfte sind wir grundsätzlich in der Lage, das System für alle Zeiten korrekt zu beschreiben. Mathematisch mag dies schwierig sein, aber die Bewegungen sind alle vorherbestimmt, sie sind *deterministisch*. Wenn die ganze Natur so beschaffen wäre, stünde die Zukunft heute schon fest. Der französische Mathematiker Pierre-Simon de Laplace hat sich ein Wesen ausgedacht, daß intelligent genug sein sollte, die ganze Welt in allen Einzelheiten zu erfassen und dann mit den Gesetzen der Mechanik die Zukunft für alle Zeit vorherzusagen. Nach ihm nennt man solch ein übermenschliches Wesen einen *Laplaceschen Dämon*.

Die meisten Gleichungen der Mechanik haben eine weitere wichtige Eigenschaft: Sie enthalten nur Zeitableitungen der Ordnung zwei, vier, sechs und so weiter. Das bedeutet, daß zu einer Lösung der Differentialgleichung jederzeit eine zweite angegeben werden kann, nämlich indem man für die Zeit t die Zeit $t' = -t$ einsetzt. In diesem Sinne sind durch solche Gleichungen Zukunft und Vergangenheit nicht unterscheidbar. Jeder Vorgang der durch solche Differentialgleichungen beschrieben wird kann in zwei Zeitrichtungen ablaufen. Diese Aussage gilt für alle grundlegenden mechanischen Vorgänge wie Stöße, Schwingungen oder sonstige Bewegungen ohne Reibung.[2]

3 Thermodynamik

Aus unserer Alltagswelt sind uns Begriffe wie Temperatur und Druck geläufig. Aber was genau ist Temperatur? Es wird sich zeigen, daß dies keine leicht zu beantwortende Frage ist. Wir werden uns deshalb ersteinmal auf folgenden Standpunkt stellen: Die Temperatur ist die Größe, die wir mit einem Thermometer messen. Dies geht jedoch nur, wenn wir nach dem berühren eines Gegenstandes einige Augenblicke warten, bis das Thermometer die Temperatur des zu messenden Gegenstandes angenommen hat. Es muß sich erst ein Gleichgewichtszustand einstellen. Erst dann können wir die Temperatur messen. Zum jetzigen Zeitpunkt können wir über die Temperatur außerhalb des Gleichgewichtes keine Aussage machen, da wir Temperatur dann nicht definieren können. Den Druck definieren wir als Kraft pro Fläche $P = K/A$.

Der Begriff Gleichgewicht soll hier nicht nur beiläufig erwähnt werden. Wenn von Gleichgewicht die Rede ist, so beinhaltet dies bereits die Annahme

[2] Diese Zeitsymmetrie ist für die elementaren Reaktionen von Teilchen ebenfalls erfüllt nach dem sogenannten CPT-Theorem. Es muß jedoch neben der Zeitumkehr auch eine Ladungsinversion und ein Wechsel der Parität miteingeschlossen werden.

einer Zeitrichtung. Denn in einen Gleichgewichtszustand läuft ein System hinein, es kann aber nicht wieder von selbst hinaus. Solche Gleichgewichte sind uns aus unserem Alltag wohlbekannt, und sie sind ein wichtiger Hinweis für Widersprüche zwischen einem auf mechanischen Gleichungen basierten Weltbild und unserer Erfahrung.

Bereits in der Mechanik stellt man fest, daß viele Prozesse so ablaufen, daß eine Energieform in eine andere umgewandelt wird. Eine Feder mit einem schwingenden Gewicht wandelt permanent kinetische Energie und Spannungsenergie der Feder ineinander um, ein fallender Stein wandelt seine potentielle Energie in kinetische Energie usw. Wenn aber Reibung mit hinzukommt, so scheint auch Energie zu verschwinden: Ein Ball rollt über eine Wiese und bleibt irgendwann liegen: Wo ist die kinetische Energie dieses Balles? Tatsächlich ist diese Energie nicht verloren, sondern hat sich in Wärme umgewandelt. Alle Reibungsprozesse verwandeln mechanische Energie in Wärme, wie man z. B. am Fahrrad sehen kann, wenn die Bremsen heiß werden. Es war eine wichtige Entdeckung, daß Wärme eine Energieform ist, die vom Heilbronner Arzt Julius Robert Mayer 1842 gemacht wurde. Wenn man die Wärme mit hinzu nimmt, so ist die Summe aller Energien immer konstant. Diese Tatsache bezeichnet man als den *ersten Hauptsatz der Thermodynamik*. Er besagt zum Beispiel, daß man keine Maschine bauen kann, die Energie erzeugt (ein sogenanntes *perpetuum mobile*). Es geht aber auch keine Energie verloren, sie ändert nur ihre Erscheinung. Bis heute wurde nie beobachtet, daß die Natur gegen den ersten Hauptsatz verstößt.

Man stellte sich aber nach Mayers Ergebnis die Frage: Kann man jede beliebige Energieform in jede andere überführen? Kann man eine Maschine bauen, die zwar keine Energie erzeugt, aber deren Energiegehalt konstant bleibt, so daß sie sich unendlich lange bewegt, ein sog. *perpetuum mobile 2. Art*.

Es waren Beobachtungen an der Dampfmaschine, die den französischen Ingenieur Sadi Carnot bereits 1824 auf den Gedanken brachten, daß nur ein Teil der Wärme in mechanische Energie umgewandelt werden kann. Er erkannte, daß es einen theoretisch optimalen Wirkungsgrad gibt. Der Wirkungsgrad ist der Teil der Wärmeenergie, der in mechanische Arbeit umgewandelt wird. Er berechnet sich für ideale Maschinen nach Carnot zu

$$\eta = \frac{T_2 - T_1}{T_2}.$$

Die Temperatur T_2 ist die eines heißen Bades (z. B. Dampfkessel), T_1 die Temperatur eines kalten Bades (z. B. Umgebungstemperatur oder Kühlwasser). Diese Temperaturen können jedoch nicht in einer beliebigen Skala gemessen werden, sondern müssen auf die absolute Temperaturskala bezogen werden, die ihren Nullpunkt bei -273° C hat. Tiefere Temperaturen existieren nicht, da sonst der Wirkungsgrad größer eins würde, was gegen den ersten Hauptsatz verstieße.

1850 veröffentlichte der deutsche Physiker Robert Clausius einen Artikel, in dem er eine Größe einführte, die er als *Entropie* bezeichnete. Diese definierte er als das Verhältnis von übertragener Wärme $\mathrm{d}Q$ zu absoluter Temperatur T

$$\mathrm{d}S = \frac{\mathrm{d}Q}{T}.$$

Er wies darauf hin, daß in allen beobachteten Situationen diese Größe nur gleichbleibt oder zunimmt. Niemals aber wurde beobachtet, daß die Entropie abnimmt. Diese Erfahrungstatsache wurde auch vom britischen Lord Kelvin formuliert und zeigte sich als sehr nützlich zur Beschreibung der Vorgänge in der Thermodynamik. Sie wird als *zweiter Hauptsatz der Thermodynamik* bezeichnet. Im Gegensatz zu den Gesetzen der Mechanik war hier erstmals eine physikalische Größe definiert, die eine Unterscheidung von Zukunft und Vergangenheit in physikalischer Sicht ermöglichte. Anschaulich gesprochen besagte der zweite Hauptsatz, daß ein perpetuum mobile 2. Art nicht existiert. Man kann z. B. kein Schiff bauen, was dadurch angetrieben wird, daß dem Ozean Wärme entzogen wird. Auch gibt es, wie oben erwähnt, einen Prozeß, der mechanische Energie in Wärme verwandelt. Einen umgekehrten Prozeß gibt es aber nur, wenn zwei Wärmebäder vorhanden sind, wie von Carnot festgestellt. Wie der zweite Hauptsatz und die Zunahme der Entropie begriffen werden kann und welche Schwierigkeiten diese verursacht hat, werden wir noch ausgibig besprechen.

Eine wichtige Beziehung der Thermodynamik soll nun noch erläutert werden: Ein Mol Gas[3], welches unter dem Druck P im Volumen V eingeschlossen ist, genügt der Gleichung

$$PV = RT,$$

wobei T die absolute Temperatur und R eine Konstante (sog. Gaskonstante) ist. Diese Beziehung verknüpft die Temperatur mit dem Druck und dem

[3]Genaugenommen handelt es sich um ein ideales Gas, also ein Gas, dessen Atome keine Ausdehnung und keine gegenseitige Wechselwirkung außer dem zentralen Stoß von Massenpunkten haben.

Volumen. Für ein Gas könnten wir nun allgemeiner die Temperatur durch diese Formel definieren, was uns ermöglichte, auch dann von Temperatur zu sprechen, wenn wir kein Thermometer im Gleichgewicht mit dem Gas haben.

4 Entropie aus mikroskopischer Sicht

Ende des neunzehnten Jahrhunderts versuchte der Wiener Physiker Ludwig Boltzmann aus den mechanischen Gesetzmäßigkeiten einzelner Teilchen auf komplexes Verhalten von Systemen vieler Teilchen zu schließen. Er war überzeugt, daß den Erscheinungen der Thermodynamik, welche man als Makrozustand bezeichnen kann, Mikrozustände entsprechen, die den Zustand einer großen Zahl von Atomen oder Molekülen beschreiben. Jeder der möglichen Mikrozustände ist, so nahm Boltzmann an, an sich mit gleicher Wahrscheinlichkeit anzutreffen. Er glaubte, daß die Moleküle sich nach den Gesetzen der Mechanik verhalten und so die Mechanik eine Grundlage des Verständnisses der Thermodynamik bildete. Auf diesem Weg wollte er Größen wie Druck und Temperatur erklären, die er als kollektives Verhalten vieler Teilchen deutete.

Hier wird ein äquivalenter Weg beschritten, der Ende der vierziger Jahre vom Amerikaner Claude Shannon entdeckt wurde.

Nehmen wir an, wir haben ein System, daß die Zustände $z_1, z_2, ..., z_n$ annehmen kann. Die Gesamtheit der möglichen Zustände sei die Menge M. Die Wahrscheinlichkeit, daß es sich im Zustand z_i befinde, sei p_i, mit

$$\sum_{i \in M} p_i = 1,$$

was bedeutet, daß die Summe über die Wahrscheinlichkeiten gleich eins ist: Das System muß sich also in irgendeinem Zustand z_i aus M befinden.

Angenommen, wir kennten die Wahrscheinlichkeitsverteilung, und wollten nun durch Ja-Nein-Fragen den Zustand des Systems herausfinden, wie sollten wir vorgehen?

Um unsere Fragen zu symbolisieren, teilen wir M sukzessive in Untermengen, zwischen denen jeweils durch eine Frage entschieden wird, z. B. M_0 und M_1 und erstere weiter in M_{00} und M_{01} usw. bis jede Menge nur noch ein Ereignis enthält. Jedes Ereignis ist nun durch eine Folge von Einsen und Nullen bestimmt, die jedoch keinsfalls gleichlang sein mssen.

Unsere Taktik hat das Ziel, im Mittel möglichst wenige Fragen stellen zu müssen. Dazu legen wir die Mengen möglichst so an, daß jede Frage in 50%

mit Ja bzw. in 50% mit Nein beantwortet werden wird.

Sei nach l Fragen die Menge M_b erreicht worden; dann gilt

$$\sum_{i \in M_{b0}} p_i = \sum_{i \in M_{b1}} p_i = 2^{-(l+1)}.$$

Wie gesagt, unterteilen wir M so lange bis in jeder Menge nur noch ein Ereignis z_i enthalten ist. Bis dahin sind l_i Fragen notwendig und wir erhalten schließlich

$$p_i = 2^{-l_i}$$

So berechnet sich umgekehrt l_i aus p_i

$$l_i = -\log_2 p_i.$$

Der so gewonnene Wert für l_i ist im allgemeinen nicht ganzzahlig, dies stört jedoch bei unserer Definition nicht. Als *fehlendende Information I* bezeichnen wir den Erwartungswert von l_i, also die Zahl an Fragen, die wir im Mittel stellen mssen, um den Zustand des Systems mit Ja-Nein-Fragen herauszubekommen. Diese berechnet sich nach den Regeln der Wahrscheinlichkeitsrechnung

$$I = -\sum_i p_i \log_2 p_i = -\frac{1}{\ln 2} \sum_i p_i \ln p_i.$$

Die fehlende Information gibt uns nun eine Verbindung von Mikrosicht und Makrosicht an die Hand: Durch Beobachtung des Makrozustandes kann die Zahl der möglichen Mikrozustände begrenzt werden. Wie weit unsere makroskopische Sicht von der Kenntnis der mikroskopischen Details entfernt ist, kommt in der fehlenden Information zum Ausdruck.

Gibt es sehr viele Mikrozustände, die zu einem Makrozustand gehören, wird es viele Fragen erfordern, den Mikrozustand zu bestimmen. Umgekehrt ist die fehlende Information klein, wenn es nur wenige Mikrozustände gibt, die mit dem Makrozustand vereinbar sind.

Die fehlende Information hat immer auch einen Bezug zum Beobachter bzw. zu seinem Wissen über das System. Damit ist bereits hier klar, daß die fehlende Information keiner Meßgröße entsprechen kann.

4.1 Relevanzkonzept

Würden wir jeden Teilchenort und jede Teilchengeschwindigkeit mit unendlicher Genauigkeit als Wert in den reelen Zahlen $\mathbb{R}$ oder rationalen Zahlen $\mathbb{Q}$ angeben wollen, so müssten wir bereits bei einem Teilchen unendlich viele Fragen stellen. Deswegen werden wir uns in der Genauigkeit beschränken müssen und unseren Mikrozustand nur ungenau beschreiben. Wir sagen, daß Information unterhalb einer gewissen Schwelle für unsere fehlende Information irrelevant geworden ist.

Tatsächlich werden wir die Relevanzschwelle nicht beliebig festsetzen können, da die Beobachtungsgenauigkeit eingeschränkt ist und wir bei jeder Messung mit Meßfehlern rechnen müssen. Diese Meßfehler und weitere äußere Störungen, die aus der mangelnden Abgeschlossenheit unseres Systems folgen, zwingen uns, die Information ab einer bestimmten Schwelle als ungenau zu betrachten. Natürlicherweise sollte unsere Relevanzschwelle also an genau dieser Grenze der Meßgenauigkeit ausgerichtet sein.

4.2 Zeitliche Zunahme der fehlenden Information

Wie bereits angemerkt wurde, haben Makrozustände mit vielen Mikrozuständen eine hohe fehlende Information und umgekehrt ist die fehlende Information für Makrozustände mit kleiner Anzahl von Mikrozuständen klein. Wenn wir von einem abgeschlossenen System vieler Teilchen ausgehen, die sich nach den Gesetzen der Mechanik verhalten, sollte die fehlende Information im Zeitverlauf symmetrisch verhalten, da nicht einzusehen ist, wieso die zeitsymmetrischen mechanischen Gesetze eine zeitassymetrische Größe hervorbringen sollten. Da die äußeren Störungen jedoch niemals gänzlich kontrollierbar sind und unser System durch die große Teilchenzahl sehr sensibel[4] auf solche Störungen reagiert. Die Störungen werden also dazu führen, daß sich unser System zufällig in einen anderen Mikrozustan begibt, der mit großer Wahrscheinlichkeit einem Makrozustand angehört, welcher aus sehr vielen Mikrozuständen besteht. Da der Gleichgewichtszustand die meisten Mikrozustände umfaßt, wird das System wahrscheinlich nach einiger Zeit in genau diesem Makrozustand sich befinden. Die fehlende Information nimmt

[4]Es ist ein System, welches deterministisches Chaos hervorbringt: Damit ist gemeint, daß bereits minimal unterschiedliche Anfangsbedingungen auch bei Gültigkeit deterministischer Gesetze (wie in der Mechanik), zu völlig unterschiedlichen Entwicklungen des Systems führen werden. In diesem Sinne ist das Verhalten des Systems nicht vorhersagbar.

also, bedingt durch die äußeren Störungen sukzessive zu.

Erst zusammen mit diesen Störungen und einem Relevanzkonzept bringen die deterministischen Gesetze der Mechanik für Vielteilchensysteme eine Größe wie die fehlende Information hervor, die im Zeitverlauf immer zunehmen wird.

4.3 Vorurteilsfreies System

Dieses Kapitel kam im Kurs nur als Randnotiz zur Sprache. Es ist mathematisch an einigen Stellen stark vereinfacht und kann andererseits bis auf den letzten Absatz problemlos übersprungen werden.

Nun wollen wir eine Beschreibung eines Systems finden, in der wir neben unserem Wissen über den Makrozustand keine weiteren Annahmen über das System machen. Diese *vorurteilsfreie Beschreibung* können wir mathematisch realisieren, in dem wir die fehlende Information maximieren. Dies bedeutet anschaulich, daß wir möglichst viele Mikrozustände zulassen wollen, sofern diese nur mit unserm makroskopischen Wissen über das System vereinbar sind.

Mathematisch betrachtet man dazu alle möglichen Wahrscheinlichkeitsverteilungen $\{p_i | i \in M : \sum_{i\in M} p_i = 1\}$ bzw. im kontinuierlichen[5] Fall $P(x)$ mit $\int P(x)\mathrm{d}x = 1$. Wir suchen nun die Verteilung, welche die fehlende Information maximiert.

Damit gleichzeitig unser Wissen über den Makrozustand berücksichtigt wird, führen wir die sogenannten *Lagrange-Parameter* λ_i ein, die nach unserer Rechnung jeweils so gewählt werden, daß die Nebenbedingungen (also unser makroskopisches Wissen) mit der Beschreibung des Systems kompatibel ist.

Wir variieren nun die Wahrscheinlichkeitsverteilung P ein wenig, was durch δP angedeutet werden soll. Wir nehmen an, daß die Mittelwerte der Größen $A_i(x)$ bekannt sei. Wir kennen also die Werte

$$\langle A_i \rangle = \int P(x) A_i(x) \mathrm{d}x.$$

Da die Bedingungen für die A_i unter dieser Variation erfüllt sein sollen, gilt

$$\frac{\partial}{\partial P}[\int A_i(x) P(x) \mathrm{d}x - \langle A_i \rangle] \delta P = 0.$$

[5]Es wird hier vereinfachend ein Integral über x geschrieben. Genaugenommen müsste es sich um $6N$ Integrale handeld, die sich über die drei Ortskoordinaten und die drei Geschwindigkeitskoordinaten jedes Teilchens erstrecken.

Ebenso ist

$$\int P(x)\mathrm{d}x = 1,$$

woraus sich die Nebenbedingung

$$\frac{\partial}{\partial P}[\int P(x)\mathrm{d}x - 1]\delta P = 0,$$

ableiten läßt. Weiterhin ist

$$\delta I = \frac{\partial I}{\partial P}\delta P = 0.$$

Zusammen kann man schreiben

$$\frac{\partial}{\partial P}\left(I - \alpha[\int P(x)\mathrm{d}x - 1] - \sum_i \lambda_i[\int A_i(x)P(x)\mathrm{d}x - \langle A_i\rangle]\right)\delta P = 0,$$

wobei die Lagrangeparameter α und λ_i eingesetzt wurden, die so gewählt werden, daß die Mittelwerte der Größen A_i die korrekten Werte annehmen. Tatsächlich variieren wir hier die Abhängigkeit der Verteilung $P(x)$ von den Koordinaten, die kurz als x geschrieben wurden. Bei N Teilchen wären dies $6N$ unabhängige Variablen. Jeder Mittelwert wird jedoch die Unabhängigkeit der Variablen beeinflussen. Um diese Beeinflussung handhaben zu können, verschieben wir sie quasi auf die λ_i, die folglich nicht frei gewählt werden können.

Es ergibt sich

$$\begin{aligned} 0 &= \left(-\int(\ln P(x) + 1)\mathrm{d}x - \alpha\int \mathrm{d}x - \sum_i \lambda_i[\int A_i(x)\mathrm{d}x]\right)\delta P \\ &= \left(\int[-\ln P(x) - 1 - \alpha - \sum_i \lambda_i A_i(x)]\mathrm{d}x\right). \end{aligned}$$

Das bedeutet in unserem Fall, daß der Ausdruck in der eckigen Klammer Null werden muß

$$0 = -\ln P(x) - 1 - \alpha - \sum_i \lambda_i A_i(x).$$

Zur Abkürzung definieren wir

$$Z(\{\lambda_i\}) = \mathrm{e}^{1+\alpha} = \int \exp(-\sum_i \lambda_i A_i(x))\mathrm{d}x,$$

die sogenannte *Zustandssumme*, die auch im englischen Sprachraum mit Z abgekürzt wird. Damit folgt für die Verteilung $P(x)$

$$P(x) = \frac{1}{Z} \exp(-\sum_i \lambda_i A_i(x)).$$

Die Kenntnisse des Zustandes sind nun in der Zustandssumme enthalten, denn es gilt

$$-\frac{\partial}{\partial \lambda_i} \ln Z(\{\lambda_i\}) = \int A_i(x) \exp(-\sum_i \lambda_i A_i(x)) \mathrm{d}x = \langle A_i(x) \rangle.$$

Die fehlende Information zu der so gefundenen Wahrscheinlichkeitsverteilung ist

$$\begin{aligned}
I \ln 2 &= -\int P(x) \ln P(x) \mathrm{d}x \\
&= -\frac{1}{Z} \int \exp(-\sum_i \lambda_i A_i(x))[-\sum_i \lambda_i A_i(x) - \ln Z] \mathrm{d}x \\
&= \frac{1}{Z} \left\{ \int \exp(-\sum_i \lambda_i A_i(x))[\sum_i \lambda_i A_i(x) \mathrm{d}x] + Z \ln Z \right\} \\
&= \int P(x)[\sum_i \lambda_i A_i(x) \mathrm{d}x] + \ln Z \\
&= \sum_i \lambda_i \langle A_i(x) \rangle + \ln Z
\end{aligned}$$

Man kann zeigen, daß dies tatsächlich die größte fehlende Information ist, die mit den bekannten Mittelwerten vereinbar ist.

Eine eingehende Betrachtung zeigt, daß der Lagrangeparameter für den Energiemittelwert mit der Temperatur T verbunden ist

$$\lambda_1 = \frac{1}{k_B T},$$

wobei k_B die Boltzmannkonstante ist. Der Lagrangeparameter für den Volumenmittelwert ist mit dem Druck P verbunden

$$\lambda_2 = \lambda_1 P = \frac{P}{k_B T}.$$

Auf diesem Wege lassen sich auch weitere Zustandsgleichungen ermitteln, die man mit dem Ersten und Zweiten Hauptsatz vergleichen kann, um schließlich für die Entropie folgenden Zusammenhang zu finden

$$S = -k_B I \ln 2.$$

Damit gelten auch die Überlegungen zu Relevanzkonzept und Zeitrichtung für die Entropie. Diese verbindet die beiden Sichtweisen des Mikro- und Makrozustandes, die je der Mechanik und Thermodynamik entsprechen. Dieses Konzept ist ein sehr allgemeines Vorgehen: Überall wo eine von Zufälligkeiten beeinflußte Mikrosicht ein kollektives Verhalten erzeugt, kann mit der Entropie bzw. fehlenden Information argumentiert werden, wie zum Beispiel bei der Modellierung von Transaktionen am Börsenmarkt, bei der Beschreibung des Verkehrsflusses, bei der Analyse von Netzwerkstrukturen etc.

5 Barometrische Höhenformel und Diffusion

Wir möchten hier noch zwei Beispiele für Differentialgleichungen herleiten und ihre Lösung besprechen.

In der Atmosphäre entsteht der Luftdruck durch das Gewicht der Luft im Schwerefeld der Erde. Der Druck in einer um $\mathrm{d}x$ höheren Schicht ist also gerade soviel niedriger, wie der Druck der Luft im Volumen $A\mathrm{d}x$ auf die Fläche A. Die Masse im Volumen ist $\varrho A\mathrm{d}x$, wobei ϱ die Dichte der Luft ist. Dies bedeutet

$$\mathrm{d}P = -\varrho g \mathrm{d}x.$$

Nach dem Gasgesetz wissen wir zudem, daß

$$\varrho = \frac{\tilde{m}}{V} = \frac{\tilde{m}P}{RT},$$

wobei $\tilde{m}$ die Masse eines Mols ist. Also folgt die Differentialgleichung

$$\frac{\mathrm{d}\varrho}{\mathrm{d}x} = -\frac{\tilde{m}g}{RT}\varrho.$$

Die Lösung lautet

$$\varrho(x) = \varrho_0 \mathrm{e}^{-\frac{\tilde{m}gx}{RT}}.$$

Wenn man einen Tropfen Milch ein eine Tasse Kaffee gibt, so kann man ein weiteres, sehr wichtiges Phänomen beobachten: Diffusion. Je nach Temperatur zerläuft der Tropfen schneller oder langsamer und vermischt sich mit dem Kaffee. Die zeitliche Änderung der Dichte der Milch in einem kleinen Volumen V^* ist gleich dem Strom von Milch über die Begrenzung des Volumens

$$\frac{\partial}{\partial t}\int_{V^*}\varrho \mathrm{d}^3x = \int_{\partial V^*}\vec{j}\mathrm{d}\sigma.$$

Mit ∂V^* bezeichnen wir die Oberfläche des Volumens, $\vec{j}$ ist der Strom, den wir als Produkt aus Dichte und Geschwindigkeit definieren $\vec{j} := \varrho\vec{v}$, $\mathrm{d}\sigma$ bedeutet, daß über die Oberfläche des Volumens intergriert wird. Es ist ein Gesetz aus der Mathematik (Satz von Stokes, der in diesem Zusammenhang auch Satz von Gauß genannt wird), daß man das zweite Integral umschreiben kann in ein Volumenintegral

$$\int_{\partial V^*}\vec{j}\mathrm{d}\sigma = \int_{V^*}\left(\frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}\right)\mathrm{d}^3x.$$

Wir nehmen an dieser Stelle an, daß $j_y = j_z = 0$ ist. Dies bedeutet, wir haben nur einen Strom in x-Richtung. Also ergibt sich

$$\frac{\partial}{\partial t}\int_{V^*}\varrho \mathrm{d}^3x = \int_{V^*}\frac{\partial j_x}{\partial x}\mathrm{d}^3x,$$

was für alle Volumen V^* gilt und damit folgt

$$\frac{\partial \varrho}{\partial t} = \frac{\partial j_x}{\partial x}.$$

Andererseits ist der Fluß da besonders groß, wo große Dichteunterschiede herrschen, was in der Gleichung

$$j_x = -D\frac{\partial \varrho}{\partial x}$$

zum Ausdruck kommt, wobei D als *Diffusionskonstante* durch diese Gleichung definiert ist. Da wir eben nur einen Fluß in x-Richtung betrachtet haben, sind Dichteunterschiede in y- und z-Richtung nicht zulässig. Es ergibt sich so die Differentialgleichung der Diffusion

$$\frac{\partial \varrho}{\partial t} = -D\frac{\partial^2 \varrho}{\partial t^2}.$$

Die Lösung lautet

$$\varrho(x,t) = \frac{\varrho_0}{\sqrt{4\pi Dt}}\mathrm{e}^{-\frac{x^2}{4Dt}}.$$

Diese Funktion ist gekoppelt an die Wahrscheinlichkeit, ein Teilchen, daß zum Zeitpunkt $t = 0$ am Ort $x = 0$ war, zum Zeitpunkt t am Ort x zu finden.

$$P(x,t) = \frac{\varrho(x,t)}{\varrho_0} = \frac{1}{\sqrt{4\pi Dt}}\mathrm{e}^{-\frac{x^2}{4Dt}}.$$

Das mittlere Auslenkungsquadrat eines solchen Teilchens nach der Zeit t kann man deswegen berechnen zu[6]

$$\langle x^2(t)\rangle = \int_{-\infty}^{\infty} x^2\varrho(x,t)\mathrm{d}x = \frac{1}{\sqrt{4\pi Dt}}\frac{1}{2}\sqrt{\pi(4Dt)^3} = 2Dt.$$

6 Brownsche Bewegung

Ähnlich wie ein ideales Gas übt auch ein in einer Flüssigkeit gelöster Stoff einen Druck aus, den sogenannten *osmotischen Druck*. Falls der gelöste Stoff aus Atomen oder Molekülen besteht, müßte auch eine Menge von größeren Teilchen einen solchen Druck ausüben, wie Einstein argumentierte. Dann kann man die Gleichung für des ideale Gas schreiben als

$$PV = Lk_BT,$$

mit der Loschmidtzahl[7] L ist dabei $k_B = R/L$ und die Loschmidtzahl ist die Anzahl in einem Mol enthaltener Teilchen.

Einstein betrachtet nun zwei Ströme, einen Diffusionsstrom und einen durch die Gravitation hervorgerufenen Strom fallender Teilchen, die im thermischen Gleichgewicht gerade so sein müssen, daß sie sich gegenseitig kompensieren. Der Diffusionstrom erfüllt die Gleichung

$$j_{diff} = -D\frac{\partial\varrho}{\partial x},$$

wobei $\varrho = L/V$ die Teilchendichte bezeichnet. Der Strom der fallenden Teilchen ist das Produkt aus Geschwindigkeit und Dichte der fallenden Teilchen $\vec{j}_{fall} = \vec{v}\varrho$. Die Geschwindigkeit der fallenden Teilchen ist gerade so groß, daß

[6] Es gilt $\int_{\infty}^{\infty} x^2\mathrm{e}^{-a^2x^2}\mathrm{d}x = \sqrt{\pi}/2a^3$.

[7] In der neueren Literatur wird $L = 6,0221\times10^{23}$ häufig wie im englischen als *Avogadro*-Konstante bezeichnet.

die Reibungskraft und die Schwerkraft gleichgroß werden. Mit der Schwerkraft $F_{grav} = mg$ und der Reibungskraft $F_{Reib} = m\zeta v$ kann man so die Geschwindigkeit der fallenden Teilchen zu

$$v = \frac{g}{\zeta}$$

berechnen. Eingesetzt für den Strom der fallenden Teilchen ergibt sich aus $j_{fall} = j_{diff}$ die Gleichung

$$-D\frac{\partial\varrho}{\partial x} = \frac{\varrho g}{\zeta}.$$

Diese Differentialgleichung für die Dichte ϱ hat folgende Lösung

$$\varrho(x) = \varrho_0 \mathrm{e}^{-\frac{gx}{D\zeta}}.$$

Andererseits ist die barometrische Höhenformel bekannt, nach der gilt

$$\varrho(x) = \varrho_0 \mathrm{e}^{-\frac{mgx}{k_B T}}$$

mit m der Masse eines Teilchens ($m/k_B = \tilde{m}/R$). Daraus schließen wir mit Einstein die wichtige Relation

$$D = \frac{k_B T}{m\zeta},$$

die sogenannte *Einstein-Beziehung.*

Wir wissen aber aus der Lösung der Diffusionsgleichung, daß die mittlere quadratische Auslenkung eines Testteilchens zur Zeit t gegeben ist, durch

$$\sqrt{\langle x^2 \rangle} = \sqrt{2Dt} = \sqrt{\frac{2k_B T}{m\zeta}t},$$

wobei wir die Einstein-Beziehung ausgenutzt haben. Die Reibungskonstante ζ kann auf anderem Wege bestimmt werden, die quadratische Auslenkung aber ist im Mikroskop zu beobachten und zu messen. Auf diesem Weg kann experimentell überprüft werden, ob die obige Herleitung richtig ist, die eine Konsequenz aus der Annahme ist, daß eine Flüssigkeit aus Atomen aufgebaut ist.