

Inhaltsverzeichnis

1	Einleitende Worte	25
2	Physikalische und mathematische Voraussetzungen	26
2.1	Das 1. Keplersche Gesetz	26
2.2	Das 2. Keplersche Gesetz	26
2.3	Das 3. Keplersche Gesetz	27
2.4	Das Gravitationsgesetz	28
2.5	Bestimmung der kosmischen Geschwindigkeiten	28
3	Die Sonne als „heiße Gaskugel“	30
3.1	Einleitung	30
3.2	Die Masse eines Sterns	30
3.3	Das hydrostatische Gleichgewicht	31
3.4	Nukleare Energieerzeugung und die Leuchtkraft eines Sterns	32
3.4.1	Hintergründe zur Energieerzeugung	32
3.4.2	Mathematische Hintergründe zur Energieproduktion und Leuchtkraft	33
3.4.3	Das Ende eines Sterns	33
3.5	Energietransport	33
3.5.1	Hintergründe zum Energietransport	33
3.5.2	Mathematische Betrachtung des Energietransportes durch Strahlung	34
3.5.3	Mathematische Betrachtung des Energietransportes durch Konvektion	35
4	Experimente und Ergänzendes	36
4.1	Experimentelle Bestimmung der Rotationszeit der Sonne	36

Kapitel 1

Einleitende Worte

Die Gruppe beschäftigte sich mit Gleichgewichtszuständen kosmischer Systeme. Es wurde deutlich, dass in längerfristig stabilen Systemen immer ein dynamisches Gleichgewicht zwischen der Gravitation und Bewegungsgrößen besteht. Während die Gravitation als immer anziehende Kraft stets eine Annäherung der Materie bewirkt, sucht die mit Impuls und kinetischer Energie behaftete Materie stets das Weite. Dies gilt in besonderem Maße auch für die elektromagnetische Strahlung, die nur in Bewegung - mit Lichtgeschwindigkeit - existiert. Als Beispiel für ein dynamisches System wurde ein Stern (wie die Sonne) betrachtet. Zur Klärung der Begriffe Potentielle und Kinetische Energie, Entweichgeschwindigkeit, Kräfte- und Energiegleichgewicht, Drehimpulserhaltung wurde die Bewegung eines Planeten auf einer elliptischen bzw. Kreisbahn um die Sonne untersucht (Keplersche Gesetze).

Im Stern, einer heißen Gaskugel, lassen sich verschiedene Gleichgewichte definieren:

1. Das hydrostatische Gleichgewicht bedeutet das Gleichgewicht zwischen der durch die Masse des Sterns bewirkten Gravitationskraft (Schwerkraft) und den Druckkräften im Sterninneren, in welchen die „Bewegungsgröße“ als kinetische Energie der Atome auftritt, charakterisiert durch die Temperatur.
2. Ein zweites Gleichgewicht, das man auch als Erhaltungssatz beschreiben kann, tritt zutage, wenn man die auffälligste Eigenschaft eines Sterns, nämlich dass er leuchtet, d.h. ständig Strahlungsenergie abgibt, näher betrachtet. Die Strahlung wird im Innern des Sterns produziert und muss nach außen gelangen und den Stern verlassen, damit dieser nicht aus dem (hydrostatischen) Gleichgewicht gerät. Die Energieerzeugungsrate muss also zu jeder Zeit gleich der abgestrahlten Energie sein.

Aus diesen Bedingungen ergeben sich Differentialgleichungen für Temperatur-, Druck- und Leuchtkraftverlauf vom Zentrum des Sterns bis zum Rand. Die Lösung dieses Systems von Differentialgleichungen erfordert noch weitere Kenntnis der Energieerzeugung und das Absorptions- und Emissionsverhaltens der Sternmaterie, die ebenfalls diskutiert wurden.

Kapitel 2

Physikalische und mathematische Voraussetzungen



Abbildung 2.1: Portrait von Kepler

Johannes Kepler (*1571 - †1630) formulierte aufgrund seiner Beobachtungen folgende Gesetzmäßigkeiten:

2.1 Das 1. Keplersche Gesetz

Die Planeten laufen auf Ellipsen um die Sonne und diese steht in einem Brennpunkt dieser Ellipse.

2.2 Das 2. Keplersche Gesetz

Der Leitstrahl vom Planeten zur Sonne überstreicht in gleichen Zeiten gleiche Flächen. D.h. der Quotient aus der Fläche A und der benötigten Zeit ist konstant:

$$\frac{A}{t} = \textit{konstant}$$

Der Leitstrahl ist diejenige gedachte Verbindungslinie zwischen Brennpunkt der Ellipse und einem Planeten, der sich auf der Ellipsenbahn befindet.

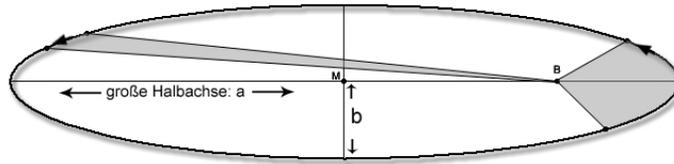


Abbildung 2.2: Illustration des 2. Keplerschen Gesetzes

2.3 Das 3. Keplersche Gesetz

Die Quadrate der Umlaufzeiten T verhalten sich wie die Kuben der großen Halbachsen. Es gilt:

$$T^2 \sim a^3$$

Wir bezeichnen hier a als große Halbachse, womit der größte Abstand vom Mittelpunkt zur Ellipsenbahn gemeint ist. Demgegenüber sei b , die kleine Halbachse, gerade der kleinste Abstand vom Mittelpunkt zur Ellipsenbahn.

Wir folgern:

$$\frac{(T_1)^2}{(T_2)^2} = \frac{(a_1)^3}{(a_2)^3}$$

Dem zuletzt am 08.06.2004 beobachteten Venustransit wurde seit dem 17. Jh. eine große Bedeutung zugemessen, da man durch Beobachtung des Transits von zwei entfernten Punkten auf der Erde aus den Abstand der Erde zur Sonne ermitteln kann. Am besten gelang dies im Jahre 1882. Mit dieser einen Größe war es nun möglich (die Umlaufzeiten waren ja bekannt) alle anderen Entfernungen zu berechnen:

$$(a_{Planet\ 1})^3 = \frac{(T_{Planet\ 1})^2 \cdot (a_{Planet\ 2})^3}{(T_{Planet\ 2})^2}$$

Dass die Keplerschen Gesetze zutreffen, kann man mithilfe des Satzes über die Erhaltung des Drehimpulses zeigen. Daher wollen wir zuerst einmal diesen Satz erwähnen um ihn anschließend zu benutzen um das 2. Keplersche Gesetz zu begründen.

Der Satz von der Erhaltung des Drehimpulses (die Rotationsbewegung in einem physikalischen System) L besagt, dass dieser stets konstant ist:

$$L = m \cdot v \cdot r = m \cdot \frac{ds}{dt} \cdot r$$

Man kann sich dies leicht durch folgendes Beispiel verdeutlichen: Der Drehimpuls beim Wirbeln einer Schnur mit einem Stein daran ergibt sich aus dem Produkt von Schnurlänge (r), Masse des Steins (m) und der Geschwindigkeit des Steins (v).

Das zweite Keplersche Gesetz besagt nun, dass die Flächengeschwindigkeit $\frac{dF}{dt}$ konstant bleibt (eben die überstrichenen Flächen des Leitstrahls in einer Zeit t).

Aus

$$\frac{dF}{dt} = \frac{ds}{dt} \cdot \frac{r}{2}$$

folgt nun, dass $\frac{dF}{dt}$ konstant sein muss, weil wir aus obiger Überlegung wissen, dass $\frac{ds}{dt} \cdot r$ konstant bleibt.

2.4 Das Gravitationsgesetz

Wir bezeichnen zwei unterschiedliche Massen mit M und m sowie den Abstand der Kerne dieser beiden Massen mit r . Des Weiteren sei K die Anziehungskraft, bzw. die *Gravitationskraft*:

$$K \sim \frac{M \cdot m}{r^2} \Rightarrow K = G \cdot \frac{M \cdot m}{r^2} \quad (2.1)$$

G heißt Gravitationskonstante.

$\left[\frac{G \cdot M}{r^2}\right]$ heißt Schwerebeschleunigung von M , sowie $\left[\frac{G \cdot m}{r^2}\right]$ Schwerebeschleunigung von m .

Beispiel 1. Die Schwerebeschleunigung der Erde ist für unsere Breitengrade bekannt ($g \approx 9,81 \frac{m}{s^2}$), sodass sich o.g. Formel leicht am Beispiel der Erde verdeutlichen lässt.

Wir kennen die Gravitationskonstante $G \approx 6,668 \cdot 10^{-11} \frac{Nm^2}{kg^2}$, die Masse der Erde $M \approx 5,977 \cdot 10^{24} kg$ und den Radius $r \approx 6,3782 \cdot 10^6 m$. Dadurch erhalten wir einen genäherten Wert für die Schwerebeschleunigung der Erde (Erdbeschleunigung) g :

$$g \approx \frac{6,668 \cdot 10^{-11} \frac{Nm^2}{kg^2} \cdot 5,977 \cdot 10^{24} kg}{(6,3782 \cdot 10^6 m)^2} \approx 9,796 \frac{m}{s^2}$$

Dieser Wert gilt für die Beschleunigung in Äquatornähe. Der uns bekannte Wert gilt dagegen für unsere Breitengrade.

Beispiel 2. Als zweites Beispiel wollen wir die Kräfte von der Sonne sowie der Erde auf den Mond vergleichen und betrachten, inwiefern sich die Kräfte unserer Erde und die des für uns essentiellen Sterns - der Sonne - auf den Mond unterscheiden. Um etwas Platz zu sparen, werden wir die Werte gleich einsetzen, die Reihenfolge allerdings beibehalten, sodass Sie die Werte aus der Formel ablesen können:

$$K_{[Erde \leftrightarrow Mond]} \approx \frac{6,668 \cdot 10^{-11} \frac{Nm^2}{kg^2} \cdot 5,9756 \cdot 10^{24} kg \cdot 7,35 \cdot 10^{22} kg}{(400\,000 \cdot 10^3 m)^2} \approx 1,83 \cdot 10^{20} N$$

$$K_{[Sonne \leftrightarrow Mond]} \approx \frac{6,668 \cdot 10^{-11} \frac{Nm^2}{kg^2} \cdot 2 \cdot 10^{30} kg \cdot 7,35 \cdot 10^{22} kg}{(150 \cdot 10^9 m)^2} \approx 4,36 \cdot 10^{20} N$$

Wir erkennen, dass der Mond stärker von der Sonne als von der Erde angezogen wird, d.h., der Mond wird von der Erde in seinem Umlauf um die Sonne etwas gestört.

2.5 Bestimmung der kosmischen Geschwindigkeiten

Die 1. kosmische Geschwindigkeit ist die minimale Kreisbahngeschwindigkeit, d.h., dass sich ein Körper mit dieser Geschwindigkeit in einer Kreisbahn um den anderen Körper bewegt. Wenn man diese minimale Kreisbahngeschwindigkeit berechnet, muss man die Gravitationskraft mit der Radialkraft (Zentrifugalkraft) gleichsetzen. Das Gleichsetzen dieser beiden Kräfte bedeutet, dass ein Planet in einer Kreisbahn um den anderen Planeten rotiert.

Gravitationskraft = Radialkraft (Zentrifugalkraft)

$$G \cdot \frac{Mm}{r^2} = \frac{m \cdot r^2}{r} \Rightarrow v_{\text{Kreis}} = \sqrt{\frac{G \cdot M}{r}} \quad (1. \text{ kosmisches Gesetz})$$

Wir halten folglich fest: Die Kreisbahngeschwindigkeit beträgt $\sqrt{\frac{G \cdot M}{r}}$.

Die 2. Kosmische Geschwindigkeit ist die Entweichgeschwindigkeit, d.h. wenn ein Körper diese Geschwindigkeit erreicht, wird z.B. ein Planet aus seiner ellipsenförmigen Kreisbahn in eine parabelförmige Umlaufbahn übergehen.

Dazu sind wieder ein paar Vorgesankten nötig. So stellt sich die Frage, wie es passieren kann, dass ein Planet aus seiner Umlaufbahn gerissen wird. Dieser Fall kann nur eintreten, wenn die durch die Geschwindigkeit des Planeten bestimmte kinetische Energie (E_{kin}) größer ist als das Potential (E_{pot}) des anderen Planeten ihn wieder anzuziehen. Deswegen werden wir die potentielle Energie mit der kinetischen Energie vergleichen. Die potentielle Energie überwiegt in einem Planetensystem, weil das Potential vorhanden sein muss den anderen Planeten wieder anzuziehen. Wenn nun die Energien gleich sind, wird der Verlauf des Planeten nicht großartig unterschiedlich sein, aber zum Zeitpunkt der Rückkehr ist das Potential nicht mehr vorhanden, den Planeten zurück zu holen. Deswegen reißt sich der Planet aus seiner Umlaufbahn und fliegt ins Universum. Die entstandene Umlaufbahn ist eine parabelförmige Umlaufbahn.

$$E_{kin} = E_{pot}$$

$$\frac{m \cdot v^2}{2} = G \cdot \frac{Mm}{r} \Rightarrow v_{\text{Entweichung}} = \sqrt{\frac{2GM}{r}} \quad (2. \text{ kosmisches Gesetz})$$

Wie man leicht erkennen kann, unterscheiden sich die beiden Gleichungen nur um den Faktor $\sqrt{2}$. Als Veranschaulichung soll ein kleines Beispiel dienen.

Beispiel 1. Die Erde bewegt sich heutzutage mit $30km/s$ auf ihrer fast kreisförmigen elliptischen Bahn. Die Entweichgeschwindigkeit der Erde beträgt folglich $42,43km/s$.

Kapitel 3

Die Sonne als „heiße Gaskugel“

3.1 Einleitung

In diesem Kapitel werden wir uns mit dem Leitthema „Die Sonne ist eine heiße Gaskugel“ beschäftigen und erklären, welche Vorgänge in der Sonne vorstatten gehen. Wir wollen uns daher z.B. mit dem Temperaturverlauf, Druckverlauf, Dichteverlauf und der Energiequelle beschäftigen. Das Modell eines Sterns wird grundlegend von drei Aspekten gestützt:

- Theorie der Sternatmosphären
- Theorie des Sternaufbaus
Die Temperatur, der Druck und die Dichte in einem Stern nehmen von innen nach außen ab.
- Langsame Verwandlung des Kerns durch Kernfusionsprozesse

Die letzten beiden Punkte werden wir hier näher erläutern, wohingegen der erste Punkt aufgrund seiner hohen Komplexität ausgelassen wird.

Eine Rolle wird auch die Frage spielen, **warum ein Stern stabil existieren kann.**

3.2 Die Masse eines Sterns

In Sternen verteilt sich die Masse in Kernnähe. Bei der Sonne befinden sich 95% der Masse in Kugel mit dem halben Sonnenradius:

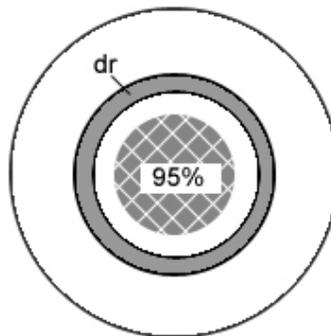


Abbildung 3.1: Masseverteilung der Sonne

Unser Ziel ist es nun eine mathematische Formulierung dieser Masseverteilung zu finden. Hierzu stellen wir folgende Überlegung an: Die Oberfläche einer Kugel berechnen wir bekanntlich mit $O = 4\pi r^2$. Das Volumen dieses Kugelausschnittes berechnen wir folglich durch $V = O \cdot \Delta r$, wobei Δr die Dicke des betrachteten Ausschnittes darstellt. Durch eine Zusatzüberlegung erhalten wir schließlich unsere gesuchte Formel zur Berechnung der Masse in einem Stern. Um die erforderliche Dichte und die Masse mit einzubringen, nehmen wir uns die Definition der Dichte

$$\rho = \frac{\text{Masse}}{\text{Volumen}} = \frac{M}{V}$$

zu Hilfe. Durch Ersetzen des Volumens

$$V = O \cdot \Delta r$$

erhalten wir

$$\frac{M}{\rho} = 4\pi r^2 \cdot \Delta r$$

und nach Umstellen die Differentialgleichung zur Bestimmung der Masse:

$$\frac{dM(r)}{dr} = \rho 4\pi r^2$$

Nun noch einige Erläuterungen zur Masse:

- Verkleinert man den Radius eines Sterns weit genug und lässt die Masse konstant, so überschreitet er den Schwarzschildradius. Dieser Radius gibt an, ab welchem Verhältnis von Masse zu Radius ein Objekt zum Schwarzen Loch werden kann. Man kann diesen Radius berechnen durch: $r_s = \frac{2GM}{c^2}$. Diese Formel soll hier allerdings keine weitere Erläuterung erfahren.
- Die Sonnenmasse beträgt $2 \cdot 10^{30} \text{ kg}$ und wird mit M_\odot bezeichnet.

3.3 Das hydrostatische Gleichgewicht

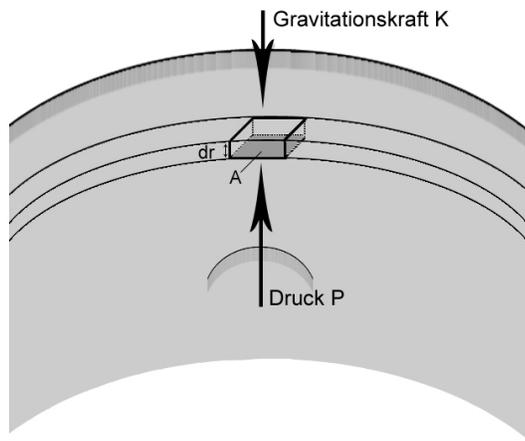


Abbildung 3.2: Ein Volumenelement in einem Stern - Kräftekompensation von K und P

Das hydrostatische Gleichgewicht ist essentiell für das Überleben eines Sternes, denn würden sich die in einem Stern wirksamen Kräfte nicht ausgleichen, würde der Stern nicht mehr seine

Beständigkeit erhalten können und so entweder auseinanderdriften oder zusammenfallen. Daher wollen wir die Gleichung für das hydrostatische Gleichgewicht herleiten. Auf das in Abbildung 3.2 eingezeichnete Volumenelement wirkt die Gravitationskraft K (siehe Kapitel 2.4):

$$K = \frac{G \cdot M \cdot m}{r^2} \quad (3.1)$$

des Weiteren wirkt auf diesen Abschnitt der Druck ΔP :

$$\Delta P = \frac{Kraft}{Fläche} = \frac{K}{A} \quad (3.2)$$

Durch die Dichte ρ des Volumenelementes mit

$$\rho = \frac{Masse}{Volumen} = \frac{m}{V}$$

und das Volumen desselbigen mit

$$V = A \cdot \Delta r$$

ergibt sich für die Dichte:

$$\rho = \frac{Masse}{Volumen} = \frac{m}{A \cdot \Delta r} \quad (3.3)$$

Durch Einsetzen von (3.2) und (3.3) in (3.1) ergibt sich dann:

$$dP \cdot A = G \frac{M \cdot \rho \cdot A \cdot dr}{r^2}$$

Umgestellt erhalten wir die Formel für das hydrostatische Gleichgewicht:

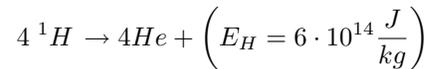
$$\frac{dP}{dr} = -\rho \cdot \frac{G \cdot M(r)}{r^2}$$

Das Minuszeichen ist lediglich formeller Natur, weil wir den Radius von innen nach außen verändern und der Druck im Stern nach außen hin abnimmt.

3.4 Nukleare Energieerzeugung und die Leuchtkraft eines Sterns

3.4.1 Hintergründe zur Energieerzeugung

Die Energie in der Sonne wird durch die Kernfusion von Wasserstoffatomen zu Heliumatomen erzeugt. Die summarische Vereinfachung der Reaktion ist folgende:



Die dabei frei werdende Energie E_H entspricht

$$E_H = mc^2$$

Durch die hohen Temperaturen im Kern der Sonne (ca. $15 \cdot 10^6 \text{K}$), dem Ort der Energieerzeugung, liegt der Wasserstoff in ionisierter Form vor und es kommt dadurch zur Protonenreaktion.



In (3.4) verschmelzen zwei Protonen zum schweren Wasserstoff, dem Deuterium. Dadurch wird ein Positron frei; die weiteren Verschmelzungsschritte sehen wie folgt aus:



Die bei der Kernfusion von von 1 g Wasserstoff frei werdende Energie E_H könnte einen Einpersonenhaushalt 150 Tage lang mit Energie versorgen.

3.4.2 Mathematische Hintergründe zur Energieproduktion und Leuchtkraft

Die Energieerzeugung in einem Stern bezeichnen wir als ϵ und sie ist definiert als $\epsilon = \frac{\text{Joule}}{\text{Sekunde} \cdot \text{Gramm}} = \frac{J}{s \cdot g}$. Die Leuchtkraft definieren wir uns als $L = \frac{\text{Joule}}{\text{Sekunde}}$. Zum Aufstellen der Differentialgleichung betrachten wir die Änderung dieser Leuchtkraft ΔL in einer Kugelschale der Dicke Δr . Die Masse dieser Kugelschale ergibt sich dann aus $4\pi r^2 \cdot \Delta r$ multipliziert mit der Dichte ρ . Wiederum multipliziert mit der Energieerzeugungsrate ϵ erhalten wir die gesuchte Differentialgleichung der Kugelschale:

$$\frac{dL(r)}{dr} = 4\pi r^2 \cdot \rho \cdot \epsilon$$

3.4.3 Das Ende eines Sterns

Die Kernfusion von Helium zu Wasserstoff läuft in dem Kern der Sonne ab. Hierbei sammelt sich über einen sehr langen Zeitraum das entstandene Helium im Kern. Um das im Kern entstandene Helium wird die Kernfusion in einer Schale um den Kern fortgesetzt. Durch die Gravitation verdichtet sich das Helium und es entstehen höherwertige Atome, wie Kohlenstoff und Sauerstoff. Die Kernfusion läuft stetig weiter, wodurch sich die o.g. Schale weiter in die äußeren Teile der Atmosphäre hinein vergrößert. Diese enthält 90% Wasserstoff und 10% Helium und einige Promille schwere Atome. Dabei löst sich diese Atmosphäre von dem verdichteten Kern, welcher als weißer Zwerg zurück bleibt.

Bei unserer Sonne, welche das Schicksal auch erleiden wird, hätte der weiße Zwerg einen Durchmesser von 0,01 Sonnenradien.

Das oben beschriebene Phänomen tritt bei Sternen auf, die eine Masse von weniger als acht Sonnenmassen besitzen. Sterne über diesem Wert werden bei dem o.g. Prozess der Verdichtung, aufgrund der Bildung von noch höher wertigen Atomen, wie Eisen und der erhöhten Gravitation des Kerns zu Neutronensternen oder schwarzen Löchern.

Bei den genannten Absprengungsprozessen werden Planeten von über 200 Sonnenradien Entfernung von dem abgesprengten Material umfasst und stark erhitzt.

3.5 Energietransport

3.5.1 Hintergründe zum Energietransport

Die pro Zeiteinheit produzierte Energie muss schließlich auch in dieser nach außen transportiert werden. Angetrieben vom Temperaturgefälle gelangt die im Zentrum der Sterne frei werdende Energie an die Oberfläche, von wo sie ins All abgestrahlt wird. Der Energietransport von innen nach außen kann sehr lange dauern. So dauert es zum Beispiel 170 000 - 1 000 000 Jahre bis die beim Wasserstoffbrennen im Innern frei gesetzte Energie bis zur von uns sichtbaren Oberfläche gelangt ist.

Prinzipiell gibt es für den Transport von Energie drei verschiedene Mechanismen: Wärmeleitung, Strahlung und Konvektion. Wärmeleitung tritt bei Sternen nur in extremen Materiezuständen auf, wie sie z.B. bei Weißen Zwergen und Neutronensternen zu finden sind. In „normalen“ Sternen dagegen sind für den Energietransport im Wesentlichen nur Strahlung und Konvektion

von Bedeutung. Beim konvektiven Wärmetransport wird die Energie in Form von aufsteigender heißer Materie, die abgekühlt wieder herunter sinkt, durch die Materie hindurch diffundiert. Während des Hauptreihenstadiums hängt die Art des Energietransports in den verschiedenen Tiefen eines Sterns vor allem von dessen Masse ab. Sterne mit kleiner Masse sind vollständig konvektiv, Sterne von etwa Sonnengröße haben einen radiativen Kern und eine konvektive Hülle, und Sterne mit großer Masse haben einen konvektiven Kern und eine radiative Hülle. In folgender Abbildung sind diese Sterne dargestellt:

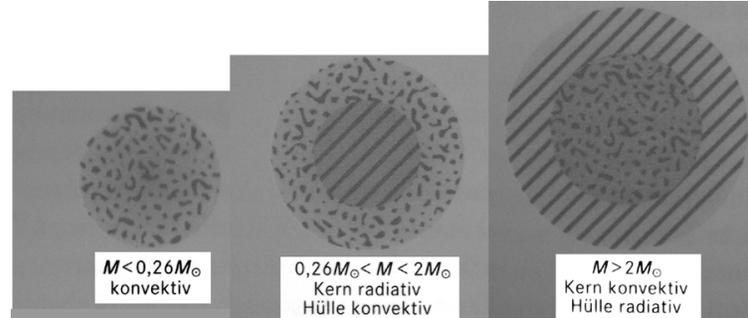


Abbildung 3.3: Sternkonfigurationen - Energietransport

3.5.2 Mathematische Betrachtung des Energietransportes durch Strahlung

In der Natur kommt es zu Konzentrationsausgleichen durch Diffusion. Diffusionsvorgänge sind so gekennzeichnet, dass der Diffusionsvorgang vom Ort der höheren Dichte zum Ort der niederen Dichte gerichtet ist. Bei Sternen nennen wir die zu betrachtende Dichte Strahlungsenergiedichte $u = a \cdot T^4$, wobei T die Temperatur bezeichnet, sowie a ein Proportionalitätsfaktor. Das Konzentrationsgefälle der Energiedichte schreiben wir als Differentialgleichung:

$$\frac{du}{dr} = \frac{du}{dT} \cdot \frac{dT}{dr} = 4aT^3 \cdot \frac{dT}{dr} \quad (3.7)$$

Den schon angesprochenen Diffusionsstrom bezeichnen wir mit j . Wir schreiben:

$$j = \frac{du}{dr} \cdot \frac{1}{3} \cdot v \cdot l \quad (3.8)$$

Hierzu sind noch einige kurze Erläuterungen vonnöten. Das Differential $\frac{du}{dr}$ heißt das Gefälle, v ist die Geschwindigkeit der diffundierenden Teilchen und l die sogenannte freie Weglänge (der Weg, den ein Teilchen zurück legt, bis es mit einem anderen Teilchen kollidiert). Des Weiteren ist es wichtig fest zu stellen, dass der Diffusionsstrom in einem Stern als

$$j = \frac{\text{Leuchtkraft}}{\text{Oberfläche}} = \frac{L}{4\pi r^2}$$

geschrieben wird. Man nennt ihn den Strahlungsstrom. Da wir den Transport von Energie betrachten, setzen wir die Geschwindigkeit $v = c$ (Lichtgeschwindigkeit). Weiterhin ist die freie Weglänge als

$$l = \frac{1}{\rho \cdot \kappa}$$

definiert. Hierbei ist ρ die Dichte und κ das Vermögen der Materie, Strahlung zu absorbieren. Durch äquivalente Umformung der Gleichung (3.8) nach $\frac{du}{dr}$ erhalten wir durch Gleichsetzen mit (3.7) die Gleichung:

$$4aT^3 \cdot \frac{dT}{dr} = \frac{3L}{4\pi r^2 \cdot c \cdot \frac{1}{\rho \cdot \kappa}}$$

Nach Umstellung nach $\frac{dT}{dr}$ erhalten wir

$$\frac{dT}{dr} = \frac{3\rho\kappa L(r)}{16\pi r^2 a c T^3}, \quad (3.9)$$

was unserer gesuchten Differentialgleichung zur Beschreibung des Energietransportes durch Strahlung entspricht.

3.5.3 Mathematische Betrachtung des Energietransportes durch Konvektion

Der Energietransport durch Konvektion in der Sonne erfolgt vorwiegend durch Konvektion an den Stellen des Sterns, an denen der adiabatische (ohne Wärmeverlust) Temperaturgradient $\left(\frac{dT}{dr}\right)_{ad}$ flacher verläuft als der Gradient, der sich im Strahlungsgleichgewicht einstellt:

$$\left(\frac{dT}{dr}\right)_{ad} < \frac{dT}{dr} \quad (3.10)$$

Maßgebend ist hierbei die Größe γ (das Verhältnis der spezifischen Wärmen bei konstantem Druck und konstantem Volumen), welche sich z.B. durch Ionisation von Wasserstoff stark ändern kann.

Bei Überwiegen der Konvektion stellt sich der Temperaturgradient wie folgt ein:

$$\frac{dT}{dr} = \left(1 - \frac{1}{\gamma}\right) \cdot \frac{T}{P} \cdot \frac{dP}{dr} \quad (3.11)$$

Kapitel 4

Experimente und Ergänzendes

4.1 Experimentelle Bestimmung der Rotationszeit der Sonne

Wir haben experimentell mit einem Sonnenprojektor ermittelt wie lange es dauert, bis sich die Sonne einmal um die eigene Achse gedreht hat.



Abbildung 4.1: Der Sonnenprojektor

Um diese Rotationszeit zu berechnen, benötigt man den Winkel zum Mittelpunkt der Sonne den der Sonnenfleck innerhalb einer bestimmten Zeit T zurück gelegt hat.

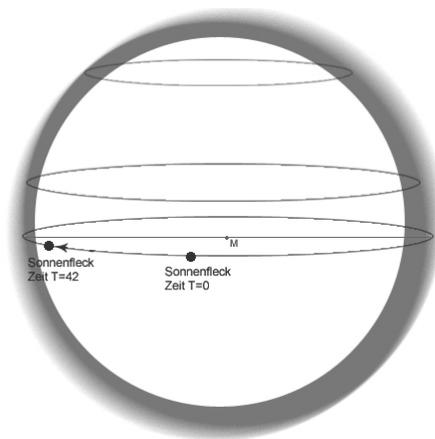


Abbildung 4.2: Sonnenflecken

Um den Winkel zu messen kippt man die Fläche, auf dem sich der Sonnenfleck befindet, geometrisch nach oben.

Nun befinden sich Anfangs- und Endpunkt am Rand des umgekippten Sonnenausschnitts. Anschließend verbindet man die Punkte mit dem Mittelpunkt, wodurch wir einen Winkel β messen können.

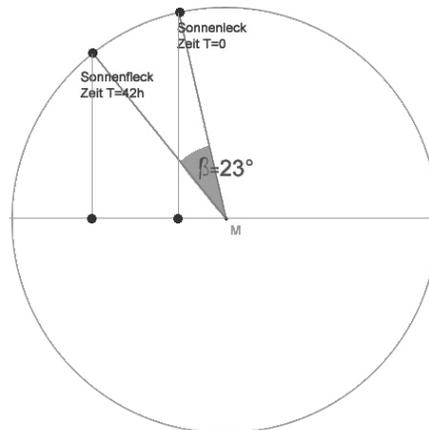


Abbildung 4.3: Umgekippte Ausschnittsfläche

Wir wissen nun, dass sich die Sonne nach einer bestimmten Zeit T um den Winkel β bewegt hat. Durch den Dreisatz können wir nun die gesuchte Rotationszeit T_R berechnen:

$$\frac{\beta}{T} = \frac{360^\circ}{T_R} \Leftrightarrow T_R = \frac{360^\circ \cdot T}{\beta}$$

Aus unseren experimentellen Aufzeichnungen haben wir eine Rotationszeit von 27,4 Tagen berechnet, was dem eigentlichen Wert für diese solare Breite von ca. 28 Tagen sehr nahe kommt. Unser Experiment mit einer einfachen Pappkonstruktion war letztendlich sehr erfolgreich - die Abweichung betrug lediglich wenige Prozent.

DIOPHANTISCHE APPROXIMATION

Teilnehmer:

Franz Arnold	Andreas-Oberschule
Mikolaj Czuchaj	Herder-Oberschule
Alexander Fauck	Heinrich-Hertz-Oberschule
Gabriel Flemming	OSZ KIM
Wiktor Pronobis	Herder-Oberschule
Christian Rekittke	Andreas-Oberschule
Robert Waniek	Heinrich-Hertz-Oberschule

Gruppenleiter:

Jürg Kramer	Humboldt-Universität, Mitglied im DFG-Forschungszentrum „Mathematik für Schlüsseltechnologien“
-------------	--

Die Gruppe beschäftigte sich mit verschiedenen Darstellungsformen reeller Zahlen und der Algebraizität bzw. Transzendenz reeller Zahlen.

Eine Untergruppe untersuchte dazu die Darstellung reeller Zahlen durch Dezimalbrüche, deren Güte zur Approximation reeller Zahlen als auch die Charakterisierung rationaler Zahlen durch abbrechende bzw. periodische Dezimalbrüche.

Eine zweite Untergruppe untersuchte die Darstellung reeller Zahlen durch Kettenbrüche, deren Güte zur Approximation reeller Zahlen als auch die Charakterisierung rationaler bzw. quadratisch-irrationaler Zahlen durch abbrechende bzw. periodische Kettenbrüche.

Eine dritte Untergruppe beschäftigte sich mit den (reellen) algebraischen und transzendenten Zahlen. Es wurde festgestellt, dass es abzählbar unendlich viele algebraische, dagegen aber überabzählbar viele transzendente Zahlen gibt. Mit Hilfe des Satzes von Liouville wurden transzendente Zahlen konstruiert. Als Beispiel eines typischen Transzendenzbeweises wurde die Transzendenz der Eulerschen Zahl $e = 2,71828\dots$ studiert.

Dokumentation der Arbeitsgruppe “Diophantische Approximation”

1 Dezimalbruchentwicklung

Definition. Die Menge \mathbb{R} der reellen Zahlen ist die Menge aller Punkte auf der Zahlengeraden. Jede reelle Zahl ist durch das Intervallschachtelungsprinzip eindeutig bestimmt.

Satz. Jede reelle Zahl α lässt sich eindeutig in der Form

$$\alpha = [\alpha] + \sum_{i=1}^{\infty} c_i \cdot 10^{-i}$$

darstellen, wobei $c_i \in \{0, 1, \dots, 9\}$ gilt und ab keinem Index i_0 alle c_i für $i \geq i_0$ gleich 9 sind. Diese Darstellung heißt die *Dezimalbruchentwicklung von α* .

Beweis. Im folgenden können wir ohne Beschränkung der Allgemeinheit annehmen, dass der ganzzahlige Anteil $[\alpha]$ von α verschwindet und somit α mit dem gebrochenen Anteil $\{\alpha\}$ übereinstimmt.

Die c_i ergeben sich nun rekursiv aus

$$\begin{array}{ll} \alpha_1 = \{\alpha\}, & c_1 = [10 \cdot \alpha_1], \\ \alpha_2 = \{10 \cdot \alpha_1\}, & c_2 = [10 \cdot \alpha_2], \\ \vdots & \vdots \\ \alpha_i = \{10 \cdot \alpha_{i-1}\}, & c_i = [10 \cdot \alpha_i], \\ \vdots & \vdots \end{array}$$

Da für alle i die Ungleichung $0 \leq \alpha_i < 1$, also $0 \leq 10 \cdot \alpha_i < 10$ gilt, haben wir $c_i \in \{0, 1, \dots, 9\}$. Indem wir noch beachten, dass eine 9-er Periode nicht auftreten kann, erkennen wir, dass die behauptete Darstellung existiert.

Darüber hinaus existiert stets höchstens eine (mithin genau eine) solche Darstellung, denn wäre α auf zwei Arten als Dezimalbruch entwickelbar, so gäbe es zwei Darstellungen

$$\alpha = \sum_{i=1}^{\infty} c_i \cdot 10^{-i} = \sum_{i=1}^{\infty} c'_i \cdot 10^{-i},$$

wobei mindestens für ein i $c_i \neq c'_i$ gilt. Es sei nun $j \geq 1$ der kleinste Index, für den die Ungleichung $c_j \neq c'_j$ erfüllt ist. Damit erhalten wir durch Vergleich

$$\frac{c_j}{10^j} - \frac{c'_j}{10^j} = \frac{c'_{j+1}}{10^{j+1}} + \dots - \frac{c_{j+1}}{10^{j+1}} - \dots,$$

d.h. nach Multiplikation mit 10^j

$$c_j - c'_j = 10^j \left(\frac{c'_{j+1} - c_{j+1}}{10^{j+1}} + \frac{c'_{j+2} - c_{j+2}}{10^{j+2}} + \dots \right) = \sum_{i=1}^{\infty} (c'_i - c_i) \cdot 10^{-i}.$$

Da nun die Differenz zweier verschiedener Elemente c_j, c'_j von $\{0, 1, \dots, 9\}$ betragsmäßig mindestens 1 beträgt, ergibt sich die Abschätzung

$$\begin{aligned} 1 &\leq |c_j - c'_j| = \left| \sum_{i=1}^{\infty} (c'_i - c_i) \cdot 10^{-i} \right| \\ &\leq \sum_{i=1}^{\infty} |c'_i - c_i| \cdot 10^{-i} \leq 9 \cdot \sum_{i=1}^{\infty} 10^{-i} = 1. \end{aligned}$$

Damit muss in obiger Abschätzung überall das Gleichheitszeichen gelten. Insbesondere erkennt man dabei, dass die Differenzen $(c'_i - c_i)$ für alle i entweder positiv oder negativ sind; nehmen wir an, sie seien positiv. Damit dann die letztere Ungleichung eine Gleichung ist, muss $c'_i - c_i = 9$ gelten, was nur möglich ist, wenn $c'_i = 9$ und $c_i = 0$ ist. Dies ist aber ein Widerspruch. \square

Satz. Es sei a/b eine rationale Zahl, wobei wir annehmen, dass die ganzen Zahlen a, b teilerfremd sind. Der Dezimalbruch von a/b bricht genau dann ab, wenn der Nenner b nur die Primteiler 2 und 5 besitzt.

Beweis. Falls b nur die Primteiler 2 und 5 besitzt, so erkennt man durch geeignetes Erweitern des Bruchs a/b sofort, dass der zugehörige Dezimalbruch abbricht.

Falls andererseits der Dezimalbruch zu a/b abbricht, so haben wir

$$\frac{a}{b} = \frac{c}{10^k} \iff 10^k \cdot a = b \cdot c$$

mit einer geeigneten ganzen Zahl c und einer geeigneten natürlichen Zahl k . Ist nun r ein Primteiler von b , so teilt r das Produkt $b \cdot c$, also auch das Produkt $10^k \cdot a$. Da nun aber a, b teilerfremd sind, muss r den Faktor 10^k teilen, so dass $r = 2$ oder $r = 5$ gilt. \square

Falls der Nenner b der rationalen Zahl a/b Primfaktoren besitzt, die verschieden von 2 und 5 sind, so ist die Dezimalbruchentwicklung von a/b periodisch. Zur Vorperiodenlänge bzw. Periodenlänge stellen wir zum Abschluss dieses ersten Abschnitts ohne Beweis folgendes fest.

Satz. (i) Die Vorperiodenlänge des Dezimalbruchs der rationalen Zahl a/b ist genau dann gleich 0, wenn für den größten gemeinsamen Teiler $\text{ggT}(10, b)$ von 10 und b gilt

$$\text{ggT}(10, b) = 1.$$

(ii) Gilt $\text{ggT}(10, b) = 1$, so ist die Periodenlänge gegeben durch die kleinste natürliche Zahl l mit der Eigenschaft

$$10^l \equiv 1 \pmod{b}.$$

Beispiel. Dies sei kurz am Beispiel $a/b = 5/7$ erläutert:

l	Restklasse
1	$10^1 \equiv 3 \pmod{7}$
2	$10^2 \equiv 2 \pmod{7}$
3	$10^3 \equiv 6 \pmod{7}$
4	$10^4 \equiv 4 \pmod{7}$
5	$10^5 \equiv 5 \pmod{7}$
6	$10^6 \equiv 1 \pmod{7}$

Daraus folgt, dass die Periodenlänge des Bruches $5/7$ gleich 6 ist. In der Tat rechnet man leicht nach

$$\frac{5}{7} = 0,\overline{714285}.$$

2 Kettenbruchentwicklung

2.1 Einführung in die Theorie der Kettenbrüche

Gegeben sei $\alpha \in \mathbb{Q}$. Dann heißt $[a_0; a_1, a_2, \dots, a_n]$ Kettenbruchentwicklung von α genau dann, wenn gilt:

$$\begin{aligned} \alpha_0 &= \alpha, & a_0 &= [\alpha_0]; \\ \alpha_1 &= \{\alpha_0\}^{-1}, & a_1 &= [\alpha_1]; \\ & \vdots & & \vdots \\ \alpha_n &= \{\alpha_{n-1}\}^{-1}, & a_n &= [\alpha_n]; \end{aligned}$$

hierbei ist $[\alpha_i]$ der ganzzahlige Anteil von α_i und $\{\alpha_i\} = \alpha_i - [\alpha_i]$ ($i = 0, \dots, n$). Es ergibt sich damit

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}}$$

Beispiel.

$$\frac{17}{5} = 3 + \frac{2}{5} = 3 + \frac{1}{\frac{5}{2}} = 3 + \frac{1}{2 + \frac{1}{2}},$$

bzw.

$$\frac{17}{5} = [3; 2, 2].$$

Analog findet man

$$\frac{159}{46} = [3; 2, 5, 4].$$

Verallgemeinert kann man nun auch Kettenbrüche für irrationale Zahlen definieren:

Statt $\alpha \in \mathbb{Q}$ ist nun $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, wodurch alle α_i größer als 1 und irrational sind ($i = 1, 2, \dots$). Wir erhalten damit für α den unendlichen Kettenbruch

$$[a_0; a_1, a_2, \dots, a_i, \dots].$$

Satz. Der Kettenbruch einer reellen Zahl α bricht genau dann ab, wenn α rational ist.

Beweis. Die Behauptung ergibt sich aus der Beobachtung, dass die Folge der Teilnenner, deren Glieder natürlich sind, streng monoton fällt. \square

Aufgrund des vorhergehenden Satzes sind die irrationalen Zahlen dadurch charakterisiert, dass ihre Kettenbrüche nicht abbrechen.

2.2 Konvergenz der unendlichen Kettenbrüche

Sei $[a_0; a_1, \dots, a_n, (\dots)]$ ein Kettenbruch (mit $a_0 \in \mathbb{Z}$ und $a_i \in \mathbb{N}$, $i = 1, 2, \dots$), dessen letztes Glied, sofern er abbricht, a_n lautet.

Dazu seien die beiden ganzzahligen Folgen $(p_i), (q_i)$ rekursiv definiert durch

$$\begin{aligned} p_{-2} &= 0, & p_{-1} &= 1, & p_i &= a_i \cdot p_{i-1} + p_{i-2}, \\ q_{-2} &= 1, & q_{-1} &= 0, & q_i &= a_i \cdot q_{i-1} + q_{i-2}. \end{aligned}$$

Bemerkung. Dabei ist $q_0 = 1$ und $q_i \geq q_{i-1} + q_{i-2}$ und somit ist $q_i \geq i$ bzw. q_i unbeschränkt für wachsende i . Der Sinn lässt sich anhand des folgenden Beispiels erahnen:

$$\begin{array}{lll} q_0 = 1, & p_0 = a_0, & \frac{p_0}{q_0} = \frac{a_0}{1} = [a_0], \\ q_1 = a_1, & p_1 = a_1 \cdot a_0 + 1, & \frac{p_1}{q_1} = \frac{a_1 \cdot a_0 + 1}{a_1} = a_0 + \frac{1}{a_1} = [a_0; a_1], \\ \vdots & \vdots & \vdots \end{array}$$

Allgemein lässt sich beweisen

Satz. Besitzt α den Kettenbruch $[a_0; a_1, \dots, a_i, \dots]$, so gilt für den endlichen Kettenbruch $A_i = [a_0; a_1, \dots, a_i]$ die Beziehung

$$A_i = \frac{p_i}{q_i}.$$

Aus diesem Grund nennen wir den Bruch p_i/q_i den i -ten Näherungsbruch der Zahl α . \square

Konvergenzsatz. Besitzt α den unendlichen Kettenbruch $[a_0; a_1, \dots, a_i, \dots]$, so gilt für die Näherungsbrüche $A_i = p_i/q_i$

$$A_0 < A_2 < \dots < \alpha < \dots < A_3 < A_1,$$

und α ist Grenzwert der Folge der Naherungsbruche.

Beweis. Man zeigt zunachst $p_{i-1}q_i - p_iq_{i-1} = (-1)^i$, woraus

$$A_{i-1} - A_i = \frac{(-1)^i}{q_{i-1}q_i} \quad (1)$$

fur $i \geq 1$ folgt. Damit lasst sich weiter zeigen, dass $p_{i-2}q_i - p_iq_{i-2} = (-1)^{i-1}a_i$ bzw. $A_i - A_{i-2} = (-1)^i a_i / (q_{i-2}q_i)$ fur $i \geq 2$ folgt. Daraus folgt fur gerades $i \geq 2$ die Ungleichung $A_i > A_{i-2}$ bzw. fur ungerades $i \geq 3$ die Ungleichung $A_i < A_{i-2}$. Sind jetzt s und t in \mathbb{N} mit $s \leq t$, bzw. $s > t$, so gilt deshalb nach Formel (1)

$$A_{2s} \leq A_{2t} < A_{2t+1} \quad \text{bzw.} \quad A_{2s} < A_{2s+1} < A_{2t+1}.$$

Damit ist die Folge A_0, A_2, \dots streng monoton wachsend und beschrankt (z.B. durch A_1), bzw. die Folge A_1, A_3, \dots streng monoton fallend und beschrankt (z.B. durch A_0). Beide Folgen konvergieren somit; seien α' , bzw. α'' die entsprechenden Grenzwerte. Damit ergibt sich

$$0 \leq \alpha'' - \alpha' < A_{2s-1} - A_{2s} \stackrel{(1)}{=} \frac{1}{q_{2s-1}q_{2s}} \leq \frac{1}{(2s-1)2s},$$

da α' das Supremum der Folge A_0, A_2, \dots , bzw. α'' das Infimum der Folge A_1, A_3, \dots ist, und die Ungleichung $q_i \geq i$ gilt. Der Grenzwert des letzten Ausdrucks ist Null, wenn s gegen unendlich strebt, womit nach dem Grenzwertsatz fur die Addition von Folgen $\alpha'' = \alpha' = \alpha$ folgt. \square

Corollar. Sei $\alpha \in \mathbb{R}$ und (p_i/q_i) die Folge der Naherungsbruche. Dann gilt

$$\left| \alpha - \frac{p_i}{q_i} \right| < \frac{1}{q_i^2},$$

d.h. die Approximation durch Kettenbruche hat im Gegensatz zur dezimalen Approximation eine doppelt so hohe Konvergenzgeschwindigkeit.

Beweis. Mit den vorhergehenden Abschatzungen findet man

$$\left| \alpha - \frac{p_i}{q_i} \right| < |A_{i+1} - A_i| = \frac{1}{q_i q_{i+1}} < \frac{1}{q_i^2}.$$

Dies beweist die Behauptung \square

Satz. Unter der Voraussetzung, dass der Kettenbruch nicht auf 1 endet, (was stets möglich ist, da man diese 1 mit dem vorletzten Glied des Kettenbruchs zusammenfassen kann,) ist die Kettenbruchentwicklung für jede reelle Zahl α eindeutig.

Beweis. Der Beweis für diese Aussage erfolgt induktiv über die Annahme, dass es zwei verschiedene Darstellungen gibt. Dabei wird nachgewiesen, dass $a_i = a'_i$ gilt, wenn bereits a_0, \dots, a_{i-1} und a'_0, \dots, a'_{i-1} gleich sind. Außerdem wird, unter Ausnutzung der Voraussetzung (das letzte Glied ist nicht 1), gezeigt, dass die beiden Kettenbrüche gleichlang sind. \square

2.3 Periodische Kettenbrüche (Sätze von Euler und Lagrange)

Im folgenden beschäftigen wir uns mit Kettenbrüchen, die in ihrer Entwicklung eine Periode aufweisen, also der Form $[a_0; a_1, \dots, \overline{a_{l+1}, \dots, a_{l+h}}]$ sind. Z.B. ist die Kettenbruchentwicklung von $\sqrt{2}$ wie folgt zu gewinnen:

$$\begin{aligned} (\sqrt{2} - 1)(\sqrt{2} + 1) = 1 &\Rightarrow \sqrt{2} - 1 = \frac{1}{\sqrt{2} + 1} \\ &\Rightarrow \sqrt{2} = 1 + \frac{1}{2 + \frac{1}{1 + \sqrt{2}}} \end{aligned}$$

Durch sukzessive Weiterführung obiger Rechnung ergibt sich schließlich

$$\sqrt{2} = [1; 2, 2, 2, \dots].$$

Das lässt sich verallgemeinern zum

Satz von Euler. Jeder unendliche, periodische Kettenbruch definiert eine reell-quadratische Irrationalzahl, d.h. eine algebraische Zahl vom Grad zwei (s. Abschnitt 3). \square

Andererseits lässt sich auch die Umkehrung dieses Satzes beweisen, nämlich

Satz von Lagrange. Jede reell-quadratische Irrationalzahl hat einen unendlichen, periodischen Kettenbruch.

Beweisskizze. Der Beweis hierfür nutzt aus, dass eine bestimmte, unendliche Folge von ganzzahligen Koeffizienten eines quadratischen Polynoms beschränkt ist und sich somit ein Teil der Glieder der Folge unendlich oft wiederholt, also periodisch ist. Damit ist auch der Kettenbruch periodisch. \square

3 Algebraizität und Transzendenz

3.1 Abzählbarkeit

Definition. Eine Menge M ist *abzählbar unendlich* genau dann, wenn eine bijektive Abbildung $f : M \longrightarrow \mathbb{N}$ existiert.

Satz. Eine abzählbar unendliche Vereinigung von abzählbar unendlichen Mengen ist abzählbar unendlich.

Beweis. Wir ordnen die abzählbar unendlich vielen abzählbar unendlichen Mengen A_1, A_2, A_3, \dots folgendermaßen an:

$$\begin{aligned} A_1 &= \{a_{11}, a_{12}, a_{13}, \dots\}, \\ A_2 &= \{a_{21}, a_{22}, a_{23}, \dots\}, \\ A_3 &= \{a_{31}, a_{32}, a_{33}, \dots\}, \\ &\vdots \end{aligned}$$

Nun kann man mittels Diagonalverfahren jedem Element in obiger Darstellung folgendermaßen eine Nummer zuordnen:

$$\begin{aligned} a_{11} &\mapsto n_1, \\ a_{12} &\mapsto n_2, \\ a_{21} &\mapsto n_3, \\ a_{31} &\mapsto n_4, \\ a_{22} &\mapsto n_5, \\ a_{13} &\mapsto n_6, \\ &\vdots \end{aligned}$$

Damit erkennt man die Vereinigung der abzählbar unendlichen Mengen A_1, A_2, A_3, \dots als gleichmächtig zur abzählbar unendlichen Menge

$$V = \{n_1, n_2, n_3, n_4, n_5, n_6, \dots\}.$$

□

3.2 Überabzählbarkeit

Definition. Eine Menge M ist *überabzählbar unendlich* genau dann, wenn M nicht abzählbar unendlich ist.

Satz. Die Menge \mathbb{R} der reellen Zahlen ist überabzählbar unendlich.

Beweis. Beweis durch Widerspruch. Wir treffen die Annahme: Das offene Intervall $I =]0, 1[$ ist abzählbar unendlich. Dann kann man alle Elemente $x \in I$ folgendermaßen untereinander anordnen:

$$\begin{aligned} x_1 &= 0, a_1 b_1 c_1 \dots, \\ x_2 &= 0, a_2 b_2 c_2 \dots, \\ x_3 &= 0, a_3 b_3 c_3 \dots, \\ &\vdots \\ x_n &= 0, a_n b_n c_n \dots, \\ &\vdots \end{aligned}$$

Man konstruiere nun eine reelle Zahl $\alpha \in I$ mit $\alpha = 0, a'_1 b'_2 c'_3 \dots$, wobei

$$a'_1 \neq a_1, b'_2 \neq b_2, c'_3 \neq c_3, \dots, z'_n \neq z_n, \dots$$

Da $\alpha \in I$ ist, müsste α in der obigen Aufzählung vorkommen, d.h. es müsste $\alpha = x_n$ für einen geeigneten Index n gelten. Dann müsste aber an der n -ten Stelle von α die Ziffer z_n stehen; im Gegensatz dazu steht aber an der n -ten Stelle z'_n . Damit kann also α in der obigen Aufzählung nicht vorkommen. Dies ist ein Widerspruch zu unserer Annahme. □

3.3 Algebraizität

Definition. Eine reelle Zahl $\alpha \in \mathbb{R}$ heißt (*reell*) *algebraisch vom Grad n* , wenn sie Nullstelle eines Polynoms

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

vom Grad n mit ganzzahligen Koeffizienten ist, aber es kein solches Polynom kleineren Grades gibt, das α zur Nullstelle hat.

Wir bezeichnen mit \mathbb{A} die Menge der reell algebraischen Zahlen.

Satz. Die Menge der algebraischen Zahlen \mathbb{A} ist abzählbar unendlich.

Beweis. Zunächst ordnen wir jedem Polynom

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in \mathbb{Z}[x]$$

die sogenannte Höhe $h(f)$ zu, welche durch

$$h(f) = |a_0| + |a_1| + \dots + |a_{n-1}| + |a_n| + n \in \mathbb{N}$$

definiert ist. Da es nur endlich viele Möglichkeiten gibt, eine natürliche Zahl als Summe natürlicher Zahlen darzustellen, gibt es zu einer gegebenen Höhe nur eine endliche Anzahl von Polynomen in $\mathbb{Z}[x]$ und somit nur eine endliche Anzahl algebraischer Zahlen, deren Polynome die gleiche Höhe besitzen.

Indem wir nun die Höhe alle natürlichen Zahlen durchlaufen lassen, erkennen wir die Menge \mathbb{A} der algebraischen Zahlen als abzählbar unendliche Vereinigung endlicher Mengen, welche nach dem Abschnitt 3.1 abzählbar unendlich sein muss. \square

3.4 Transzendenz

Definition. Eine reelle Zahl $\alpha \in \mathbb{R}$ heißt (*reell*) *transzendent*, wenn sie nicht algebraisch ist.

Wir bezeichnen mit \mathbb{T} die Menge der reell transzendenten Zahlen.

Satz. Es gibt überabzählbar unendliche viele transzendente Zahlen, d.h. die Menge \mathbb{T} ist insbesondere nicht leer.

Beweis. Da \mathbb{R} nach 3.2 überabzählbar unendlich ist und die Menge \mathbb{A} der algebraischen Zahlen Teilmenge von \mathbb{R} und nach Abschnitt 3.4 abzählbar ist, muss das Komplement $\mathbb{R} \setminus \mathbb{A}$ dieser beiden Mengen eine überabzählbar unendliche Menge sein. Diese Menge ist aber nichts anderes als die Menge \mathbb{T} der transzendenten Zahlen. \square

Satz von Liouville. Sei $z \in \mathbb{A}$ vom Grad $n > 1$. Dann gilt für alle $p \in \mathbb{Z}$

und hinreichend großen $q \in \mathbb{N}$:

$$\left| z - \frac{p}{q} \right| > \frac{1}{q^{n+1}}.$$

Wir verzichten hier auf den Beweis. Vielmehr weisen wir darauf hin, dass, wenn eine reelle Zahl $z \in \mathbb{R}$ das Liouvillesche Kriterium nicht erfüllt, diese Zahl transzendent sein muss.

Auf diese Art lässt sich beispielsweise zeigen, dass die sogenannte Liouvillesche Zahl

$$z = \sum_{i=1}^{\infty} 10^{-i!} = 0,1100010000000000000000000100\dots$$

transzendent ist.

3.5 Transzendenz von e

Satz. Die Eulersche Zahl $e = 2,71828\dots$ ist transzendent.

Strategie zum Beweis. Indirekter Beweis.

An dem Beweis der Transzendenz der Eulerschen Zahl e wird im Wesentlichen die Strategie für Transzendenzbeweise dargelegt. Derartige Beweise werden immer indirekt geführt.

Beweisidee: Wir nehmen an, dass e eine algebraische Zahl ist; e ist dann Nullstelle eines Polynoms, d.h. es gilt

$$\sum_{k=0}^m a_k \cdot e^k = 0$$

mit $a_k \in \mathbb{Z}$, $a_0, a_m \neq 0$ für $k \in \{0, 1, \dots, m\}$. Im Laufe des Beweises wird diese Annahme zu einem Widerspruch geführt. Zu diesem Zwecke konstruieren wir uns ein Hilfspolynom $H(x) \in \mathbb{Z}[x]$, das e^k approximiert und die folgenden weiteren Eigenschaften erfüllt:

I:

$$H(0) \neq 0.$$

Diese Eigenschaft wird benötigt, um Quotienten mit $H(0)$ im Nenner bilden zu können.

II:

$$\sum_{k=0}^m a_k \cdot \frac{H(k)}{H(0)} \neq 0.$$

Hiermit wird eine qualitative Aussage der Güte der Approximation getroffen, d.h. unser Hilfspolynom approximiert e^k nicht so gut, dass $H(k)/H(0)$ Nullstelle des Polynoms ist.

III:

$$\left| \sum_{k=0}^m a_k \cdot \left(e^k - \frac{H(k)}{H(0)} \right) \right| < \left| \frac{1}{H(0)} \right|.$$

Mit der dritten Voraussetzung wird eine quantitative Aussage über die Approximation getroffen, d.h. das Polynom nähert sich dem Wert e^k so gut an, dass die Summe der Produkte aus a_k und $e^k - H(k)/H(0)$ im Betrag kleiner ist als der Betrag von $1/H(0)$.

Eine Funktion mit diesen Eigenschaften existiert tatsächlich und wird im nächsten Abschnitt nachgewiesen.

Nachdem wir unser Hilfspolynom konstruiert haben, gilt es nun, den Widerspruch zu finden, um die Transzendenz von e zu beweisen. Dazu betrachten wir

$$\sum_{k=0}^m a_k \cdot e^k = \sum_{k=0}^m a_k \cdot \frac{H(k)}{H(0)} + \sum_{k=0}^m a_k \cdot \left(e^k - \frac{H(k)}{H(0)} \right).$$

Wegen der Annahme $\sum_{k=0}^m a_k \cdot e^k = 0$ und nach äquivalenter Umformung (Multiplikation mit $H(0)$) erhalten wir weiter

$$0 = \sum_{k=0}^m a_k \cdot H(k) + \sum_{k=0}^m a_k \cdot \left(e^k \cdot H(0) - H(k) \right).$$

Nach Eigenschaft II von H und wegen der Ganzzahligkeit der Koeffizienten von H gilt einerseits

$$\sum_{k=0}^m a_k \cdot H(k) \in \mathbb{Z}_{\neq 0};$$

andererseits muss aber nach Eigenschaft III

$$\sum_{k=0}^m a_k \cdot \left(e^k \cdot H(0) - H(k) \right) \in] -1, 1[$$

gelten. Eine ganze Zahl ungleich 0 und eine reelle Zahl im offenen Intervall $] - 1, 1[$ können sich nun aber nicht zu 0 addieren. Dies ist der gesuchte Widerspruch!

Für Transzendenzbeweise wird also immer angenommen, die fragliche Zahl sei algebraisch, und es wird versucht, eine approximierende Funktion zu konstruieren, um letzten Endes obigen Widerspruch zu finden.

Konstruktion des Hilfspolynoms H .

Ziel: Approximation von e^k ($k = 0, \dots, m$) durch $H(x)$

Vorüberlegung:

$$g(x) = a \cdot e^x \Leftrightarrow g'(x) = g(x) \quad (a \in \mathbb{R}),$$

$$g(x) = e^x \Leftrightarrow g'(x) = g(x) \wedge g(0) = 1.$$

Definition des Hilfspolynoms f :

$$f(x) := x^{p-1} \cdot (x-1)^p \cdot (x-2)^p \cdot \dots \cdot (x-m)^p,$$

dabei hat das Polynom f den Grad $N := mp + p - 1$; p ist eine große Primzahl.

Definition des Hilfspolynoms F :

$$F(x) := f(x) + f'(x) + \dots + f^{(N)}(x).$$

Wegen $f^{(N+1)}(x) = 0$ erhalten wir $F'(x) = F(x) - f(x)$. $F(x)$ ist deshalb zur Approximation von e^x geeignet, wenn $f(x)$ auf dem Intervall $[0, m]$ betragsmäßig "klein" ist, d.h.:

$$F'(x) \approx F(x).$$

Abschätzung von $f(x)$ auf dem Intervall $[0, m]$: Für alle $x \in [0, m]$ gilt:

$$|x \cdot (x-1) \cdot (x-2) \cdot \dots \cdot (x-m)|^p \leq (m^{m+1})^p,$$

also

$$\max_{0 \leq x \leq m} |f(x)| \leq (m^{m+1})^p.$$

Somit erhalten wir die Ungleichungen

$$0 \leq \frac{\max_{0 \leq x \leq m} |f(x)|}{(p-1)!} \leq \frac{(m^{m+1})^p}{(p-1)!},$$

also

$$0 \leq \lim_{p \rightarrow \infty} \frac{(m^{m+1})^p}{(p-1)!} = 0.$$

Da beide Grenzwerte existieren und gleich sind, muss auch der Grenzwert

$$\lim_{p \rightarrow \infty} \frac{\max_{0 \leq x \leq m} |f(x)|}{(p-1)!}$$

existieren und gleich Null sein. Folglich gilt für $x \in [0, m]$:

$$\lim_{p \rightarrow \infty} \frac{f(x)}{(p-1)!} = 0.$$

Wenn man die Gleichung $F'(x) = F(x) - f(x)$ durch $(p-1)!$ dividiert, so erhält man folgende Approximation:

$$\frac{F'(x)}{(p-1)!} \approx \frac{F(x)}{(p-1)!}.$$

Man bildet die gesuchte Hilfsfunktion in der Form

$$H(x) := \frac{F(x)}{(p-1)!},$$

und es lässt sich daher folgern

$$\frac{H(x)}{H(0)} \approx e^x.$$

Literatur

- [1] *P. Bundschuh*: Einführung in die Zahlentheorie. Springer-Verlag, 1996.
- [2] *F. Klein*: Elementarmathematik vom höheren Standpunkt aus I. Springer-Verlag, 1933.
- [3] *J. Kramer*: Vorlesungsnotizen, 2000.
- [4] *H. Schied*: Zahlentheorie. Spektrum Akademischer Verlag, 2003.

FOURIERREIHEN

Teilnehmer:

Paul Beyer	Herder-Oberschule
Paul Bismarck	Herder-Oberschule
Johannes Bleidorn	OSZ KIM
Martin Büttner	Heinrich-Hertz-Oberschule
Alexander Fromm	Heinrich-Hertz-Oberschule
Asar Hage-Ali	Heinrich-Hertz-Oberschule
Georg Wieselmann	OSZ KIM

Gruppenleiter:

Konrad Gröger	Humboldt-Universität
---------------	----------------------

Die Gruppe beschäftigte sich mit der Darstellung stetiger reellwertiger Funktionen durch Fourierreihen.

Zuerst lernte die Gruppe die Grundbegriffe „Skalarprodukt“ und „Norm“ für stetige Funktionen kennen. Zur Motivation wurde auf entsprechende Begriffe für gewöhnliche Vektoren eingegangen. Danach wurden Orthonormalsysteme behandelt, speziell ein mit Hilfe trigonometrischer Funktionen gebildetes System. Es wurde gezeigt, dass man zu jeder stetigen Funktion u eine Reihe angeben kann, deren Glieder skalare Vielfache der Elemente eines Orthonormalsystems sind und die in einem bestimmten Sinne gegen u konvergiert. Diese Reihe nennt man die Fourierreihe von u . Am Ende wurde ein Ausblick auf Möglichkeiten der theoretischen und praktischen Nutzung von Fourierreihen gegeben.

Fourierreihen

1. Grundbegriffe

Im Folgenden befassen wir uns immer mit stetigen reellwertigen Funktionen, die auf dem Intervall $[-1, 1]$ definiert sind. Zur Abkürzung bezeichnen wir die Menge aller dieser Funktionen mit V . Die Menge V hat folgende Eigenschaft: Sind u und v Elemente von V und ist t eine reelle Zahl, so liegen auch die Funktionen $u + v$ und tu in V ; dabei sind $u + v$ und tu durch die Vorschrift

$$(u + v)(x) := u(x) + v(x), \quad (tu)(x) := tu(x) \text{ für alle } x \in [-1, 1]$$

definiert.

Definition 1:

Für beliebige $u, v \in V$ nennt man die Zahl $(u|v) := \int_{-1}^1 u(x)v(x)dx$ **Skalarprodukt der Funktionen** u und v und die Zahl $\|u\| := \sqrt{(u|u)} = \left(\int_{-1}^1 (u(x))^2 dx \right)^{\frac{1}{2}}$ **Norm von** u .

Wer mit dem Vektorraum \mathbb{R}^3 oder allgemeiner mit dem Raum \mathbb{R}^n vertraut ist, erkennt, dass die eingeführten Begriffe zum Skalarprodukt $a \cdot b = \sum_{i=1}^n a_i b_i$ zweier

Vektoren a und b und zur Länge $|a| = \left(\sum_{i=1}^n a_i^2 \right)^{\frac{1}{2}}$ eines Vektors a analog sind. An die Stelle der Summe ist das Integral getreten, an die Stelle des Index i der Komponenten das Argument x der Funktionen.

Wir stellen die Eigenschaften des Skalarprodukts und der Norm, die im Folgenden benötigt werden, in einem Satz zusammen.

Satz 1. Für beliebige $u, v, w \in V$ und $t \in \mathbb{R}$ gilt

$$\|u\| = 0 \iff u = 0, \quad (u|v) = (v|u), \quad (1)$$

$$(u + v|w) = (u|w) + (v|w), \quad (tu|v) = t(u|v), \quad (2)$$

$$|(u|v)| \leq \|u\| \|v\| \text{ (Cauchy-Schwarzsche Ungleichung)}, \quad (3)$$

$$\|u + v\| \leq \|u\| + \|v\| \text{ (Dreiecksungleichung)}. \quad (4)$$

Beweis:

Schritt 1. Die Behauptungen (1) und (2) sind offensichtlich richtig.

Schritt 2. Für $u, v \in V$ und $t \in \mathbb{R}$ ergibt sich mit Hilfe von (2)

$$0 \leq \|u + tv\|^2 = (u + tv|u + tv) = (u|u) + 2t(u|v) + t^2(v|v) = \|u\|^2 + 2t(u|v) + t^2\|v\|^2. \quad (5)$$

Für $v = 0$ ist (3) sicher richtig. Für $v \neq 0$ und $t = -\frac{(u|v)}{\|v\|^2}$ folgt aus (5)

$$0 \leq \|u\|^2 - 2\frac{(u|v)^2}{\|v\|^2} + \frac{(u|v)^2}{\|v\|^2} = \|u\|^2 - \frac{(u|v)^2}{\|v\|^2}.$$

Also ist $(u|v)^2 \leq \|u\|^2\|v\|^2$, d.h., (3) ist richtig.

Schritt 3. Mit Hilfe der Ungleichung (3) ergibt sich

$$\|u + v\|^2 = (u + v|u + v) = \|u\|^2 + 2(u|v) + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2.$$

Daraus folgt die Dreiecksungleichung (4) durch Wurzelziehen. \square

Definition 2:

Gilt für $u, v \in V$ die Beziehung $(u|v) = 0$, so sagt man, u und v seien **zueinander senkrecht (orthogonal)**. Ist $\|u\| = 1$, so sagt man, u sei **normiert**. Ein System (e_0, e_1, e_2, \dots) von Funktionen aus V heißt **Orthonormalsystem**, wenn für $n, m = 0, 1, 2, \dots$ folgende

$$\text{Beziehung gilt: } (e_n|e_m) = \begin{cases} 0, & \text{falls } n \neq m, \\ 1, & \text{falls } n = m. \end{cases}$$

Orthonormalsysteme spielen in V die gleiche Rolle wie Systeme von Basisvektoren in gewöhnlichen Vektorräumen. Es wird gezeigt werden, dass man für die Elemente von V mit Hilfe eines Orthonormalsystems eine Art von „Koordinatendarstellung“ angeben kann.

2. Beispiel für ein Orthonormalsystem

Behauptung: Durch die Vorschriften $e_0(x) := \frac{1}{\sqrt{2}}$ für $x \in [-1, 1]$ und

$$e_{2k-1}(x) := \sin(k\pi x), \quad e_{2k}(x) := \cos(k\pi x) \quad \text{für } x \in [-1, 1] \text{ und } k = 1, 2, \dots \quad (6)$$

wird ein Orthonormalsystem (e_0, e_1, e_2, \dots) in V definiert.

Zum Beweis dieser Behauptung muss man die Skalarprodukte $(e_n|e_m)$ berechnen. Für $k, l = 1, 2, \dots$ ergibt sich durch partielle Integration

$$\begin{aligned} (e_{2k}|e_{2l}) &= \int_{-1}^1 \cos(k\pi x) \cos(l\pi x) dx = \frac{k}{l} \int_{-1}^1 \sin(k\pi x) \sin(l\pi x) dx \\ &= \frac{l}{k} \int_{-1}^1 \sin(k\pi x) \sin(l\pi x) dx. \end{aligned}$$

Für $k \neq l$ ist $\frac{k}{l} \neq \frac{l}{k}$. Deshalb muss in diesem Fall $(e_{2k}|e_{2l}) = (e_{2k-1}|e_{2l-1}) = 0$ sein. Für $k = l$ ergibt sich

$$\int_{-1}^1 \cos^2(k\pi x) dx = \int_{-1}^1 \sin^2(k\pi x) dx = \int_{-1}^1 (1 - \cos^2(k\pi x)) dx.$$

Daher ist

$$2 \int_{-1}^1 \cos^2(k\pi x) dx = \int_{-1}^1 dx = 2$$

und deshalb $\|e_{2k}\| = \|e_{2k-1}\| = 1$. Weiter ist

$$(e_{2k}|e_{2l-1}) = \int_{-1}^1 \cos(k\pi x) \sin(l\pi x) dx = 0,$$

weil der Integrand eine ungerade Funktion und das Integrationsintervall symmetrisch zum Nullpunkt ist. Die Aussagen $(e_0|e_0) = 1$ und $(e_0|e_{2k}) = (e_0|e_{2k-1}) = 0$ sind ebenfalls einfach zu beweisen. Darauf wird hier nicht eingegangen.

Das hier angegebene Orthonormalsystem nennen wir das **trigonometrische Orthonormalsystem**. Mit diesem System werden wir hauptsächlich arbeiten. Es gibt aber noch viele andere Orthonormalsysteme in V .

3. Approximation stetiger Funktionen

In diesem Abschnitt sei (e_0, e_1, e_2, \dots) ein beliebig gewähltes Orthonormalsystem in V . Für $n \in \mathbb{N}$ bezeichnen wir mit V_n die Menge der Elemente von V , die sich mit reellen Zahlen c_0, c_1, \dots, c_n in der Form $\sum_{i=0}^n c_i e_i$ schreiben lassen. Gehören u und v zu V_n , so auch $u + v$ und tu für reelles t .

Satz 2. *Zu jeder Funktion $u \in V$ gibt es genau eine Funktion $v_n \in V_n$, für die $\|u - v_n\|$ minimal wird. Es ist $v_n = \sum_{j=0}^n (u|e_j) e_j$, und es gilt*

$$\|u - v_n\|^2 = \|u\|^2 - \sum_{j=0}^n (u|e_j)^2. \quad (7)$$

Die Funktion $u - v_n$ ist senkrecht zu allen Funktionen aus V_n .

Beweis: Für $v = \sum_{j=0}^n c_j e_j$ gilt

$$\begin{aligned} \|u - v\|^2 &= \|u\|^2 - 2(u|v) + \|v\|^2 \\ &= \|u\|^2 - 2\left(u \left| \sum_{j=0}^n c_j e_j \right.\right) + \left(\sum_{j=0}^n c_j e_j \left| \sum_{i=0}^n c_i e_i \right.\right) \\ &= \|u\|^2 - 2 \sum_{j=0}^n c_j (u|e_j) + \sum_{j=0}^n \sum_{i=0}^n c_j c_i (e_j|e_i) \\ &= \|u\|^2 - 2 \sum_{j=0}^n c_j (u|e_j) + \sum_{j=0}^n c_j^2. \end{aligned}$$

Der Anteil von c_j an diesem Ausdruck ist eine quadratische Funktion in c_j , die genau für $c_j = (u|e_j)$ ihr Minimum annimmt. Setzt man diesen Wert für c_j in die obige Formel ein, so ergibt sich

$$\|u - v\|^2 = \|u\|^2 - \sum_{j=0}^n (u|e_j)^2.$$

Zugleich ist für $i = 0, \dots, n$

$$(u - v|e_i) = (u|e_i) - \left(\sum_{j=1}^n (u|e_j)e_j | e_i \right) = (u|e_i) - (u|e_i) = 0,$$

d.h., $u - v$ steht auf allen Elementen von V_n senkrecht. □

Definition 3:

Für eine Funktion $u \in V$ bezeichnet man die Zahlen $c_j := (u|e_j)$ als **Fourierkoeffizienten** der Funktion u .

Speziell für das trigonometrische Orthonormalsystem lassen sich die Fourierkoeffizienten durch folgende Formeln beschreiben:

$$(u|e_0) = \int_{-1}^1 u(x) \frac{1}{\sqrt{2}} dx,$$

$$(u|e_{2k-1}) = \int_{-1}^1 u(x) \sin(k\pi x) dx, \quad (u|e_{2k}) = \int_{-1}^1 u(x) \cos(k\pi x) dx, \quad k = 1, 2, \dots$$

Beispiele für Fourierkoeffizienten

Beispiel 1. Es sei $u(x) = x^2$ für $x \in [-1, 1]$. Dann gilt

$$(u|e_0) = \int_{-1}^1 x^2 \frac{1}{\sqrt{2}} dx = \frac{\sqrt{2}}{3},$$

$$(u|e_{2k-1}) = \int_{-1}^1 x^2 \sin(k\pi x) dx = 0, \quad \text{weil der Integrand eine ungerade Funktion ist,}$$

$$\begin{aligned} (u|e_{2k}) &= 2 \int_0^1 x^2 \cos(k\pi x) dx = -\frac{4}{k\pi} \int_0^1 x \sin(k\pi x) dx \\ &= -\frac{4}{k^2\pi^2} \int_0^1 \cos(k\pi x) dx + \left[\frac{4}{k^2\pi^2} x \cos(k\pi x) \right]_0^1 = \frac{4}{k^2\pi^2} \cos(k\pi) = \frac{4(-1)^k}{k^2\pi^2}. \end{aligned}$$

Beispiel 2. Es sei $u(x) = x$ für $x \in [-1, 1]$. Dann gilt

$$(u|e_0) = \int_{-1}^1 x \frac{1}{\sqrt{2}} dx = 0,$$

$$(u|e_{2k-1}) = 2 \int_0^1 x \sin(k\pi x) dx = \frac{2}{k\pi} (-1)^{k-1} \quad (\text{vgl. Beispiel 1}) \text{ und}$$

$$(u|e_{2k}) = \int_{-1}^1 x \cos(k\pi x) dx = 0, \quad \text{weil der Integrand eine ungerade Funktion ist.}$$

Die Darstellung von Funktionen und Näherungen sowie Satz 2 provozieren die Frage, ob eine Funktion $u \in V$ als *Grenzwert* der Näherungen v_n verstanden werden kann. Zur

Beantwortung dieser Frage ist es notwendig, zunächst erst einmal festzulegen, was man unter Konvergenz einer Folge von Funktionen verstehen will.

Definition 4:

Man sagt, eine Folge (u_n) aus V **konvergiere im (quadratischen) Mittel** gegen $u^* \in V$, wenn $\lim_{n \rightarrow \infty} \|u_n - u^*\| = 0$ ist. Man sagt, eine Folge (u_n) aus V **konvergiere gleichmäßig** gegen $u^* \in V$, wenn $\lim_{n \rightarrow \infty} \|u_n - u^*\|_\infty = 0$ ist; dabei ist $\|v\|_\infty := \sup_{x \in [-1,1]} |v(x)|$.

Dementsprechend nennt man eine *Reihe* $\sum_{j=0}^{\infty} v_j$ gleichmäßig (bzw. im quadratischen Mittel)

konvergent, wenn die Folge der Partialsummen $s_n := \sum_{j=0}^n v_j$ gleichmäßig (bzw. im quadratischen Mittel) konvergiert.

Bemerkung: Für beliebiges $v \in V$ ist $\|v\|^2 = \int_{-1}^1 (v(x))^2 dx \leq 2\|v\|_\infty^2$. Deshalb folgt aus der gleichmäßigen Konvergenz einer Folge (u_n) gegen u^* , dass (u_n) auch im Mittel gegen u^* konvergiert. Das Umgekehrte ist im Allgemeinen nicht der Fall. Aus $\|u_n - u^*\|_\infty \rightarrow 0$ folgt $u_n(x) \rightarrow u^*(x)$ für alle $x \in [-1, 1]$. Zudem gilt für beliebige $u, v \in V$ die Dreiecksungleichung $\|u + v\|_\infty \leq \|u\|_\infty + \|v\|_\infty$. Das zeigt man wie folgt: Für jedes $x \in [-1, 1]$ gilt

$$|(u + v)(x)| \leq |u(x)| + |v(x)| \leq \|u\|_\infty + \|v\|_\infty.$$

Daher ist auch

$$\|u + v\|_\infty = \sup_{x \in [-1,1]} |(u + v)(x)| \leq \|u\|_\infty + \|v\|_\infty.$$

Ist von der Konvergenz einer Folge aus V die Rede, so muss immer angegeben werden, welche Konvergenz gemeint ist.

Satz 3. *Gilt für eine Folge (u_n) aus V die Beziehung $\lim_{m,n \rightarrow \infty} \|u_n - u_m\|_\infty = 0$, so konvergiert die Folge (u_n) gleichmäßig gegen eine Funktion $u^* \in V$.*

Beweis: Aus der Voraussetzung folgt $u_n(x) - u_m(x) \rightarrow 0$ für $m, n \rightarrow \infty$ und jedes $x \in [-1, 1]$, d.h., für jedes $x \in [-1, 1]$ ist $(u_n(x))$ eine Cauchyfolge von reellen Zahlen. Deshalb existiert $u^*(x) := \lim_{n \rightarrow \infty} u_n(x)$ für $x \in [-1, 1]$. Wir zeigen, dass u^* stetig ist und dass die Folge (u_n) gleichmäßig gegen u^* konvergiert.

Es sei $\varepsilon > 0$ vorgegeben. Für hinreichend große n, m ist nach Voraussetzung $\|u_n - u_m\|_\infty < \varepsilon$. Mit Hilfe dieser Ungleichung und der Dreiecksungleichung für reelle Zahlen erhält man

$$|u_n(x) - u^*(x)| \leq |u_n(x) - u_m(x)| + |u_m(x) - u^*(x)| \leq \varepsilon + |u_m(x) - u^*(x)|.$$

Für $m \rightarrow \infty$ ergibt sich $|u_n(x) - u^*(x)| \leq \varepsilon$ für hinreichend große n . Da ε nicht von x abhängt, ist auch

$$\sup_{x \in [-1,1]} |u_n(x) - u^*(x)| \leq \varepsilon \quad \text{für hinreichend große } n. \tag{8}$$

Für beliebige $x, x_0 \in [-1, 1]$ ergibt sich (erneut mit Hilfe der Dreiecksungleichung für reelle Zahlen)

$$\begin{aligned} |u^*(x) - u^*(x_0)| &\leq |u^*(x) - u_n(x)| + |u_n(x) - u_n(x_0)| + |u_n(x_0) - u^*(x_0)| \\ &\leq 2 \sup_{y \in [-1, 1]} |u^*(y) - u_n(y)| + |u_n(x) - u_n(x_0)| \leq 2\varepsilon + |u_n(x) - u_n(x_0)|, \end{aligned}$$

falls n hinreichend groß gewählt wird. Ein solches hinreichend großes n fixieren wir. Weil u_n stetig ist, existiert ein $\delta > 0$ derart, dass $|u_n(x) - u_n(x_0)| < \varepsilon$ wird, falls $|x - x_0| < \delta$ ist. Folglich ist für $|x - x_0| < \delta$

$$|u^*(x) - u^*(x_0)| < 3\varepsilon,$$

d.h., die Funktion u^* ist stetig. Die Beziehung (8) besagt, dass $\lim_{n \rightarrow \infty} \|u_n - u^*\|_\infty = 0$ ist; die Folge (u_n) konvergiert also gleichmäßig gegen u^* . \square

Satz 4. Konvergiert eine Reihe $\sum_{j=0}^{\infty} c_j e_j$ im Mittel gegen eine Funktion $u \in V$, so sind die Zahlen c_0, c_1, \dots die Fourierkoeffizienten der Funktion u bezüglich (e_0, e_1, \dots) .

Beweis: Für $j = 0, 1, \dots$ und $n \geq j$ gilt aufgrund der Cauchy-Schwarzschen Ungleichung

$$|(u|e_j) - c_j| = \left| \left(u - \sum_{i=0}^n c_i e_i \middle| e_j \right) \right| \leq \left\| u - \sum_{i=0}^n c_i e_i \right\| \|e_j\| = \left\| u - \sum_{i=0}^n c_i e_i \right\|.$$

Nach Voraussetzung konvergiert die obere Schranke für $n \rightarrow \infty$ gegen 0, und somit ist $(u|e_j) = c_j$. Folglich sind die Zahlen c_j , wie behauptet, die Fourierkoeffizienten der Funktion u . \square

4. Spezielle Aussagen für das trigonometrische Orthonormalsystem

Von nun an sei (e_0, e_1, \dots) das trigonometrische Orthonormalsystem.

Satz 5. Haben zwei Funktionen u und v aus V die gleichen Fourierkoeffizienten bezüglich (e_0, e_1, e_2, \dots) , so sind sie gleich.

Beweis: Es sei $w := u - v$. Nach Voraussetzung ist $(w|e_j) = (u|e_j) - (v|e_j) = 0$ für $j = 0, 1, \dots$. Die Funktion w ist deshalb zu allen Elementen aller Räume V_n orthogonal. (Zur Definition von V_n siehe den Anfang von Abschnitt 3.) Zu zeigen ist $w = 0$. Wir nehmen an, es sei $w(\xi) \neq 0$ für ein $\xi \in]-1, 1[$. Wir dürfen zusätzlich annehmen, dass $w(\xi) = 1$ ist (erforderlichenfalls multiplizieren wir w mit $\frac{1}{w(\xi)}$). Wir wählen $\delta > 0$ so klein, dass das Intervall $I_\delta := [\xi - 2\delta, \xi + 2\delta]$ in $[-1, 1]$ liegt und dass $w(x) \geq \frac{1}{2}$ ist für $x \in I_\delta$. Es sei

$$p_n(x) := (1 + \cos(\pi(x - \xi)) - \cos(\pi\delta))^n \quad \text{für } x \in [-1, 1] \quad \text{und } n \in \mathbb{N}.$$

Mit Hilfe der Additionstheoreme für Winkelfunktionen zeigt man induktiv leicht, dass $p_n \in V_{2n}$ ist. Folglich ist $(w|p_n) = 0$ für alle $n \in \mathbb{N}$. Die Wahl von p_n ist so getroffen, dass

$p_n(x) \geq 1$ ist für $x \in [\xi - \delta, \xi + \delta]$, $p_n(x) \geq 0$ für $x \in I_\delta$ und $q := \sup_{x \in [-1, 1] \setminus I_\delta} |p_1(x)| < 1$. Es gilt

$$\begin{aligned} 0 &= (w|p_n) = \int_{-1}^{\xi-2\delta} w(x)p_n(x)dx + \int_{\xi-2\delta}^{\xi+2\delta} w(x)p_n(x)dx + \int_{\xi+2\delta}^1 w(x)p_n(x)dx \\ &\geq - \int_{-1}^{\xi-2\delta} |w(x)|q^n dx + \delta - \int_{\xi+2\delta}^1 |w(x)|q^n dx. \end{aligned}$$

Daraus folgt für $n \rightarrow \infty$ der Widerspruch $0 \geq \delta > 0$. Dieser Widerspruch widerlegt die Annahme, dass $w(\xi) \neq 0$ ist für ein $\xi \in]-1, 1[$. Deshalb muss $w = 0$ sein. \square

Satz 6. Konvergiert für gegebene Zahlen c_0, c_1, \dots die Reihe $\sum_{j=0}^{\infty} |c_j|$, so konvergiert die Reihe $\sum_{j=0}^{\infty} c_j e_j$ gleichmäßig gegen eine Funktion $u \in V$, und die Zahlen c_0, c_1, \dots sind die Fourierkoeffizienten dieser Funktion u .

Beweis:

Schritt 1. Nach Satz 3 genügt es zum Beweis der gleichmäßigen Konvergenz der Reihe $\sum_{j=0}^{\infty} c_j e_j$ zu zeigen, dass die Folge der Partialsummen $s_n := \sum_{j=0}^n c_j e_j$ eine Cauchyfolge im Sinne der gleichmäßigen Konvergenz ist. Dazu wird im Folgenden der Wert $\|s_n - s_m\|_\infty$ für $n, m \in \mathbb{N}$, $m < n$, betrachtet. Es ist

$$\begin{aligned} \|s_n - s_m\|_\infty &= \left\| \sum_{j=m+1}^n c_j e_j \right\|_\infty \leq \sum_{j=m+1}^n \|c_j e_j\|_\infty \quad (\text{Dreiecksungl. für die Supremumsnorm}) \\ &= \sum_{j=m+1}^n |c_j| \|e_j\|_\infty = \sum_{j=m+1}^n |c_j|. \end{aligned}$$

Da $\sum_{j=0}^{\infty} |c_j|$ konvergiert, ist die Folge der Partialsummen $a_n := \sum_{j=0}^n |c_j|$ nach dem Cauchyschen Konvergenzkriterium eine Cauchyfolge reeller Zahlen. Deshalb ist bei gegebenem $\varepsilon > 0$

$$\|s_n - s_m\|_\infty \leq \sum_{j=m+1}^n |c_j| = a_n - a_m < \varepsilon \quad \text{für hinreichend große } m, n.$$

Schritt 2. Da die gleichmäßige Konvergenz die Konvergenz im quadratischen Mittel nach sich zieht, gilt nach Satz 4 die Beziehung $c_j = (u|e_j)$ für $j = 0, 1, \dots$, d.h., die Zahlen c_j sind die Fourierkoeffizienten der Funktion u . \square

Satz 7. Es sei $u \in V$ eine zweimal stetig differenzierbare Funktion. Zusätzlich gelte die Beziehung $u(-1) = u(1)$. Dann konvergiert die Fourierreihe von u gleichmäßig gegen u .

Beweis: Für $k = 1, 2, \dots$ ergibt sich unter Anwendung der Methode der partiellen Integration

$$\begin{aligned} c_{2k-1} &= \int_{-1}^1 u(x) \sin(k\pi x) dx \\ &= \frac{1}{k\pi} \int_{-1}^1 u'(x) \cos(k\pi x) dx + \left[u(x) \frac{1}{k\pi} \cos(k\pi x) \right]_{-1}^1 \\ &= -\frac{1}{k^2\pi^2} \int_{-1}^1 u''(x) \sin(k\pi x) dx \end{aligned}$$

und

$$\begin{aligned} c_{2k} &= \int_{-1}^1 u(x) \cos(k\pi x) dx \\ &= -\frac{1}{k\pi} \int_{-1}^1 u'(x) \sin(k\pi x) dx + \left[u(x) \frac{1}{k\pi} \sin(k\pi x) \right]_{-1}^1 \\ &= -\frac{1}{k^2\pi^2} \int_{-1}^1 u''(x) \cos(k\pi x) dx + \left[\frac{1}{k^2\pi^2} u'(x) \cos(k\pi x) \right]_{-1}^1. \end{aligned}$$

Folglich ist $|c_{2k-1}| \leq \frac{2\|u''\|_\infty}{k^2\pi^2}$ und $|c_{2k}| \leq \frac{2(\|u''\|_\infty + \|u'\|_\infty)}{k^2\pi^2}$ und damit $|c_j| \leq \frac{M}{j^2}$ für $M := \frac{8(\|u''\|_\infty + \|u'\|_\infty)}{\pi^2}$ und $j > 0$. Es ist

$$\sum_{j=1}^{\infty} \frac{1}{j^2} \leq 1 + \sum_{j=2}^{\infty} \frac{1}{j(j-1)} = 1 + \sum_{j=2}^{\infty} \left(\frac{1}{j-1} - \frac{1}{j} \right) < \infty,$$

denn die Partialsummen der letzten Reihe sind $\sum_{j=2}^n \left(\frac{1}{j-1} - \frac{1}{j} \right) = 1 - \frac{1}{n}$. Daher konvergiert die Reihe $\sum_{j=0}^{\infty} |c_j|$, und die Behauptung des Satzes ist eine Konsequenz der Sätze 6 und 5.

Satz 6 garantiert, dass die Reihe $\sum_{j=0}^{\infty} c_j e_j$ gleichmäßig gegen eine Funktion v konvergiert, deren Fourierkoeffizienten die Zahlen c_j sind. Da die Zahlen c_j die Fourierkoeffizienten von u sind, ist $v = u$ nach Satz 5. \square

Beispiel. Es sei $u(x) = x^2$ für $x \in [-1, 1]$. Die so definierte Funktion u genügt den Voraussetzungen von Satz 7. Die Fourierkoeffizienten von u sind schon in Abschnitt 3 berechnet worden. Wegen der gleichmäßigen Konvergenz der Fourierreihe von u gegen u gilt für alle $x \in [-1, 1]$

$$\begin{aligned} x^2 &= c_0 \frac{1}{\sqrt{2}} + \sum_{k=1}^{\infty} c_{2k} \cos(k\pi x) \\ &= \frac{1}{3} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \cos(k\pi x). \end{aligned}$$

Aus der Gleichung für $x = 1$ folgt nach einfachen Umformungen

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Aus der Gleichung für $x = 0$ folgt ebenso

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - + \dots = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} = \frac{\pi^2}{12}.$$

Satz 8. *Es seien $u \in V$ und $\varepsilon > 0$ gegeben. Dann existiert eine Funktion $v \in V$ mit folgenden Eigenschaften:*

1. *Die Funktion v ist auf $[-1, 1]$ zweimal stetig differenzierbar und genügt der Bedingung $v(-1) = v(1)$.*
2. *Es ist $\|v - u\| < \varepsilon$.*

Beweis: Wir setzen zunächst

$$w(x) := \frac{1}{\eta^2} \int_x^{x+\eta} \int_y^{y+\eta} u(z) dz dy \quad \text{für } x \in [-1, 1];$$

dabei sei $u(z) := u(1)$ für $z > 1$, und $\eta > 0$ sei eine Zahl, über die später noch geeignet verfügt wird. Die Funktion w ist zweimal stetig differenzierbar: Es ist

$$w'(x) = \frac{1}{\eta^2} \left(\int_{x+\eta}^{x+2\eta} u(z) dz - \int_x^{x+\eta} u(z) dz \right)$$

und

$$w''(x) = \frac{1}{\eta^2} (u(x+2\eta) - 2u(x+\eta) + u(x)).$$

Weiter gilt

$$w(x) - u(x) = \frac{1}{\eta} \int_x^{x+\eta} \left(\frac{1}{\eta} \int_y^{y+\eta} u(z) dz - u(x) \right) dy = \frac{1}{\eta} \int_x^{x+\eta} \frac{1}{\eta} \int_y^{y+\eta} (u(z) - u(x)) dz dy.$$

Die Variable z in dem inneren Integral bleibt immer im Intervall $[x, x+2\eta]$. Deshalb wird

$$|w(x) - u(x)| \leq \frac{1}{\eta} \int_x^{x+\eta} \frac{1}{\eta} \int_y^{y+\eta} |u(z) - u(x)| dz dy \leq \sup_{|z-x| \leq 2\eta} |u(z) - u(x)|.$$

Wir wählen nun $\eta > 0$ (unabhängig von x) so klein, dass $|u(z) - u(x)| < \frac{\varepsilon}{4}$ wird, falls $|z - x| < 2\eta$ ist. Dann wird $|w(x) - u(x)| < \frac{\varepsilon}{4}$ für alle $x \in [-1, 1]$ und daher

$$\|w - u\|^2 = \int_{-1}^1 (w(x) - u(x))^2 dx \leq \frac{\varepsilon^2}{8}, \quad \text{also erst recht } \|w - u\| < \frac{\varepsilon}{2}.$$

Als nächstes wählen wir eine zweimal stetig differenzierbare Funktion $\psi \in V$ mit den Eigenschaften $0 \leq \psi(x) \leq 1$ für alle $x \in [-1, 1]$, $\psi(x) = 1$ für $x \in [-1 + \delta, 1 - \delta]$ sowie

$\psi(-1) = \psi(1) = 0$; dabei wird über $\delta > 0$ noch passend verfügt. Wir setzen schließlich $v := \psi w$. Dann ist v zweimal stetig differenzierbar und $v(-1) = v(1) = 0$. Außerdem gilt

$$\|v - w\|^2 = \int_{-1}^1 (\psi(x)w(x) - w(x))^2 dx \leq \int_{-1}^{-1+\delta} (w(x))^2 dx + \int_{1-\delta}^1 (w(x))^2 dx \leq 2\delta \|w\|_\infty^2.$$

Wir wählen nun $\delta > 0$ so, dass $2\delta \|w\|_\infty^2 \leq \frac{\varepsilon^2}{4}$ ist. Dann wird $\|v - w\| \leq \frac{\varepsilon}{2}$. Zusammen mit der Abschätzung für $\|w - u\|$ ergibt sich $\|v - u\| \leq \|v - w\| + \|w - u\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Die Funktion v hat also die gewünschten Eigenschaften. \square

Satz 9. Die Fourierreihe jeder Funktion $u \in V$ konvergiert im Mittel gegen u .

Beweis: Für den Beweis muss gezeigt werden, dass der Wert $\left\| u - \sum_{j=0}^n (u|e_j)e_j \right\|$ für hinreichend große n kleiner wird als jedes $\varepsilon > 0$. Dazu wird zunächst Satz 8 benutzt. Nach diesem existiert für jedes $\varepsilon > 0$ eine zweimal stetig differenzierbare Funktion u_ε mit den Eigenschaften $u_\varepsilon(-1) = u_\varepsilon(1)$ und $\|u_\varepsilon - u\| < \varepsilon$. Wir bestimmen $N \in \mathbb{N}$ so, dass

$$\left\| u_\varepsilon - \sum_{j=0}^n (u_\varepsilon|e_j)e_j \right\| < \varepsilon \quad \text{für } n \geq N.$$

Das ist möglich, weil die Fourierreihe von u_ε nach Satz 7 gleichmäßig (und damit erst recht im quadratischen Mittel) gegen u_ε konvergiert. Mit Hilfe der Dreiecksungleichung ergibt sich daraus und aus der Ungleichung $\|u - u_\varepsilon\| < \varepsilon$ die Aussage

$$\left\| u - \sum_{j=0}^n (u|e_j)e_j \right\| \leq \|u - u_\varepsilon\| + \left\| u_\varepsilon - \sum_{j=0}^n (u_\varepsilon|e_j)e_j \right\| < 2\varepsilon \quad \text{für } n \geq N.$$

Es ist $\sum_{j=0}^n (u|e_j)e_j \in V_n$. Nach Satz 2 ist die bestmögliche Approximation von u durch ein

Element aus V_n die Partialsumme $\sum_{j=0}^n (u|e_j)e_j$ der Fourierreihe von u . Folglich gilt auch

$$\left\| u - \sum_{j=0}^n (u|e_j)e_j \right\| < 2\varepsilon \quad \text{für } n \geq N.$$

Wegen der Willkür von $\varepsilon > 0$ beweist das, dass die Fourierreihe von u im Mittel gegen u konvergiert. \square

Ergänzende Bemerkung

Es gibt im Raum V der stetigen Funktionen auf dem Intervall $[-1, 1]$ Folgen (u_n) , für die $\lim_{m, n \rightarrow \infty} \|u_n - u_m\| = 0$ ist, die in V keinen Grenzwert im Sinne der Konvergenz im Mittel besitzen. Dieser Mangel lässt sich beseitigen, indem man von dem Raum V in geeigneter Weise zu einem größeren Raum U übergeht. Erst in diesem Raum U wird die

Theorie der Fourierreihen völlig befriedigend. Es lässt sich zeigen, dass eine Fourierreihe $\sum_{j=0}^{\infty} c_j e_j$ genau dann im Mittel gegen ein Element $u \in U$ konvergiert, wenn die Reihe $\sum_{j=0}^{\infty} c_j^2$ konvergiert. Es besteht dann eine bijektive Beziehung zwischen den Elementen u des Raums U und den Folgen (c_0, c_1, \dots) , für die $\sum_{j=0}^{\infty} c_j^2$ konvergiert. Die Zahlen c_0, c_1, \dots lassen sich als Koordinaten des Elements u bezüglich der Basis e_0, e_1, \dots verstehen. Rechenoperationen mit den Elementen von U lassen sich in Operationen für die Koordinaten „übersetzen“. Eine befriedigende Beschreibung des hier erwähnten Raums U erfordert allerdings die Kenntnis des sogenannten Lebesgue-Integrals.