Mathematisches Institut der Universität Bonn

Sommersemester 2005

Lineare Algebra II

Ingo Lieb

Letzte Änderung: 2. September 2005

Inhaltsverzeichnis

Einleitung

Behandelt werden die Theorie euklidischer und unitärer Vektorräume (teilweise in Überschneidung mit den letzten Kapiteln der Vorlesung des Wintersemesters), der Spektralsatz mit seinen wichtigsten Anwendungen, die Klassifikation von Endomorphismen und eine Einführung in die Darstellungstheorie. Die Vorlesung ist Grundlage für ein weiterführendes Studium der Algebra, aber auch Grundlage für Teile der theoretischen Physik und der Analysis.

Natürlich beansprucht die Vorlesung keine Originalität. Ich habe mich in Begriffsbildungen und Bezeichnungen möglichst genau an allgemein übliche Konventionen gehalten, wie sie etwa im Buch von Bosch verwandt werden. Die Theorie der Normalformen von Endomorphismen entwickle ich als Spezialfall der Strukturtheorie von endlich erzeugten Moduln über Hauptidealringen. Soweit hier einige Beweise nicht in der Literatur enthalten sind, gehen sie auf eine frühere Vorlesung von S. Bosch und mir zurück. Im Kapitel über Darstellungstheorie folge ich dem entsprechenden Kapitel in Artins ”Algebra”, wobei sich aber eine Reihe von Änderungen und Ergänzungen in den Abschnitten über kompakte Gruppen als zweckmäßig erwiesen haben. Bei den physikalischen Interpretationen der Darstellungstheorie stütze ich mich auf van der Waerden, Wigner und Shafarevich, ferner auf geduldige Erläuterungen meines Physik-Kollegen J. Kroha.

Obwohl physikalische Anwendungen Motivation für einen großen Teil des Stoffes bilden, kommen sie explizit nur an zwei Stellen vor: im ersten Kapitel (Theorie kleiner Schwingungen) und im letzten Abschnitt des dritten Kapitels (Quantenmechanik und Elementarteilchen). Eine ausführliche Diskussion der Rolle unitärer Räume in der Quantenmechanik hat bereits in der – nicht ausgearbeiteten – Vorlesung des Wintersemesters (Lineare Algebra I) stattgefunden. Dieser erste Teil enthielt auch elementare für die Physik wichtige Begriffsbildungen wie Volumina, Winkel, Vektorprodukte und dergleichen. Ebenso sind die Begriffe, an die ich in der vorliegenden Ausarbeitung ”erinnere”, im ersten Semester eingeführt worden: Gruppen, Faktorgruppen, Ideale, endliche Körper, Eigenwerte u.a.

Die Ausarbeitung ist für die Hörerinnen und Hörer meiner Vorlesung bestimmt und soll ihnen bei ihrer weiteren Arbeit helfen. Insgesamt war der in den beiden Vorlesungen dargestellte Stoff etwas umfangreicher und schwieriger als sonst üblich; die Übungen haben gezeigt, daß die Anforderungen nicht zu hoch waren. Für das stets gezeigte Interesse und die gute Mitarbeit bedanke ich mich sehr herzlich bei allen Teilnehmern.

Ebenso herzlich danke ich Frau Lütz für die Erstellung des ”Texoskript” und Herrn Doberenz für die sorgfältige Korrektur.

Literatur

M. Artin. *Algebra.* Basel etc. 1993

S. Bosch. *Lineare Algebra.* Berlin etc. 2001

E. Brieskorn. *Lineare Algebra und analytische Geometrie.* 2 Bde Braunschweig 1983 u. 1985

G. Fischer. *Analytische Geometrie.* Hamburg 1978

J.P. Serre. Représentations linéaires des groupes finis. Paris 1978 und 1998

B.L. van der Waerden. *Group theory and quantum mechanics.* New York etc. 1974

I.R. Shafarevich. *Algebra 1.* In: Encyclopedia of mathematical sciences, vol 11. Berlin etc. 1990

E. Wigner *Group theory and its applications to the quantum mechanics of atomic spectra.* New York etc. 1959

1 Der Spektralsatz

§1 Skalarprodukte und Sesquilinearformen

1.) Sei **K** immer einer der beiden Körper **R** oder **C** (reelle bzw. komplexe Zahlen); für $z \in \mathbf{C}$ ist $\bar{z}$ die konjugiert komplexe Zahl zu z und $|z|$ der Absolutbetrag. Wir notieren die

Regeln: (1) $z = x + iy$, dann $\bar{z} = x - iy$ für $x, y \in \mathbf{R}$.
(2) $|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} = |\bar{z}|$.
(3) $|z| = 0$ genau für $z = 0$.
(4) $|zw| = |z||w|$.
(5) $|z + w| \leq |z| + |w|$.

Eine Zahl z ist reell genau dann, wenn $z = \bar{z}$ ist. Alle auftretenden Vektorräume sind (nicht notwendig endlichdimensionale) **K**-Vektorräume. Die folgenden Definitionen sind aus dem vergangenem Semester bekannt: ich wiederhole sie hier.

Definition 1.1. *Eine Bilinearform ist eine Abbildung*

$$s : V \times V \to \mathbf{K},$$

*(V ein **K**-Vektorraum), für die gilt:*
(1) Für $y \in V$ ist $x \longmapsto s(x, y)$ linear in x.
(2) Für $x \in V$ ist $y \longmapsto s(x, y)$ linear in y.

(Diese Definition ist natürlich für beliebige Grundkörper sinnvoll).

Definition 1.2. *Eine Abbildung $\varphi : V \to W$ zweier **K**-Vektorräume heißt semilinear, wenn gilt:*

(1) $\varphi(x + y) = \varphi(x) + \varphi(y)$ für $x, y \in V$,
(2) $\varphi(\lambda x) = \bar{\lambda}\varphi(x)$ für $\lambda \in \mathbf{K}$, $x \in V$.

Für $\mathbf{K} = \mathbf{R}$ ist also „linear“ und „semilinear“ dasselbe, nicht aber für $\mathbf{K} = \mathbf{C}$.

Definition 1.3. *Eine Sesquilinearform auf V ist eine Abbildung*

$$s : V \times V \to \mathbf{K}$$

mit:

(1) Für $y \in V$ ist $x \longmapsto s(x, y)$ semilinear in x.
(2) Für $x \in V$ ist $y \longmapsto s(x, y)$ linear in y.

Definition 1.4. *Die Sesquilinearform s heißt hermitisch, wenn stets*

$$s(x, y) = \overline{s(y, x)}$$

ist. Kurz: s ist eine hermitische Form.

Für $\mathbf{K} = \mathbf{R}$ ist „sesquilinear" dasselbe wie „bilinear" und „hermitisch" gleich „symmetrisch":

$$s(x, y) = s(y, x).$$

Schließlich als vorläufig letzte

Definition 1.5. *Eine Sesquilinearform heißt nicht entartet, wenn gilt:*
(1) Aus $s(x, y) = 0$ für alle y folgt: $x = 0$.
(2) Aus $s(x, y) = 0$ für alle x folgt: $y = 0$.

Ein Skalarprodukt auf V ist eine nicht entartete hermitische Sesquilinearform.

2.) Die folgenden Beispiele sind fundamental:
$V = \mathbf{K}^n$. Für $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ sei

$$s(x, y) = \langle x, y \rangle = \sum_{j=1}^{n} \overline{x}_j y_j.$$

Faßt man x, y als Spalten auf, dann ist also

$$\langle x, y \rangle = {}^t\overline{x} \cdot y \quad \text{(Matrizenprodukt, } {}^t \cdot \text{ steht für Transposition).}$$

Satz 1.1 $\langle\ ,\ \rangle$ *ist ein positiv definites Skalarprodukt auf* $\mathbf{K}^n$.

Dabei benutzen wir

Definition 1.6. *Eine hermitische Form s heißt positiv semidefinit, wenn* $s(x,x) \geq 0$ *für alle* x *ist; s heißt positiv definit, wenn* $s(x,x) > 0$ *für alle* $x \neq 0$ *ist.*

Beachte, daß natürlich $s(x,y) = 0$ für x oder $y = 0$ gilt, ferner daß bei hermitischen Formen

$$s(x,x) = \overline{s(x,x)} \in \mathbf{R}$$

ist. Positiv definite Formen sind klarerweise nicht entartet.

Beweis von Satz 1.1.

$$\langle x,x\rangle = \sum_{j=1}^{n} \overline{x}_j x_j = \sum_{j=1}^{n} |x_j|^2 > 0 \text{ für } x \neq 0.$$

Definition 1.7. *Sei s eine positiv definite hermitische Form auf V. Dann heißt*

$$V, s \text{ ein } \left\{ \begin{array}{l} \textit{euklidischer } \mathbf{R} \\ \textit{unitärer } \mathbf{C} \end{array} \right\} \quad \textit{Vektorraum}$$

und s das zugehörige Skalarprodukt. –

Meist schreibt man $s(x,y) = \langle x,y\rangle$ – "Unitärer **K**-Vektorraum" verwenden wir für unitäre und euklidische Räume.
Ich erinnere an weitere Beispiele:

(1) $V = C^0_{\mathbf{C}}[a,b]$ der Vektorraum der stetigen komplexwertigen Funktionen auf dem Intervall $[a,b]$ und

$$\langle f,g\rangle = \int_a^b \overline{f(x)}g(x)dx$$

als positiv definites Skalarprodukt.

(2) $V = \mathbf{R}^n$,

$$\langle x,y\rangle = \sum \varepsilon_j x_j y_j, \quad \varepsilon_j = \pm 1.$$

Dies ist ein Skalarprodukt, das positiv definit nur für $\epsilon_j = +1,\ j = 1 \ldots n$, ist.

3.) Im endlich-dimensionalen Fall stellen wir nun einen Zusammenhang zwischen Matrizen und Formen her.
Es sei $\mathcal{A} = \{a_1, \ldots, a_n\}$ eine Basis des n-dimensionalen **K**-Vektorraums V. Ferner sei s eine Sesquilinearform auf V. Wir definieren

$$A_{s;\mathcal{A}} = (s(a_i,a_j))_{i=1\ldots n, j=1\ldots n} \in \mathbf{K}^{n\times n}$$

als Matrix von s bezüglich der Basis $\mathcal{A}$ und notieren die folgenden Aussagen:

Aussage 1. Ist $x = \sum_{j=1}^{n} \alpha_j a_j$, $y = \sum_{j=1}^{n} \beta_j a_j$, bezeichnet

$$\mathfrak{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \mathfrak{y} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

die Koordinatenvektoren von x und y im $\mathbf{K}^n$ und A die Matrix von s bez. $\mathcal{A}$, so ist

$$s(x, y) = {}^t\overline{\mathfrak{x}} A \mathfrak{y}.$$

Beweis.

$$\begin{aligned} s(x, y) &= s\left(\sum_i \alpha_i a_i, \sum_j \beta_j a_j\right) \\ &= \sum_{i,j} \overline{\alpha}_i \beta_j s(a_i, a_j) = {}^t\overline{\mathfrak{x}} A \mathfrak{y}. \end{aligned}$$

Aussage 2. Seien $x, y, \mathfrak{x}$ und $\mathfrak{y}$ wie in Aussage 1. Es sei $A \in \mathbf{K}^{n \times n}$ beliebig. Definitere

$$s(x, y) = {}^t\overline{\mathfrak{x}} A \mathfrak{y}.$$

Dann ist s sesquilinear, mit $A_{s;\mathcal{A}} = A$.

Aussage 3. Es gilt $A_{s;\mathcal{A}} = 0$ genau für $s \equiv 0$.

Aussage 4. Es ist s hermitisch genau dann, wenn $A_{s,\mathcal{A}} = A$ der Bedingung

$${}^t\overline{A} = A$$

genügt.

Beweis. Die Beziehung $s(a_i, a_j) = \overline{s(a_j, a_i)}$ ist notwendig und hinreichend für Hermitizität und sagt gerade $A = {}^t\overline{A}$.

Aussage 5. Es ist s genau dann nicht-entartet, wenn $A_{s;\mathcal{A}}$ regulär ist.

Beweis. (1) Es sei s entartet. D.h. es gibt ein $x \neq 0$ mit $s(x, y) \equiv 0$ (mit „$\equiv$" meinen wir Gleichheit für alle y). Es sei $\mathfrak{x}$ der Koordinatenvektor von x, also $\mathfrak{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ mit $\sum \alpha_j a_j = x$, entsprechend $\mathfrak{y} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ der von y. Dann gilt mit

$$A = (a_{ij}) = A_{s;\mathcal{A}} :$$

$$\sum_{i,j} \overline{\alpha_i} a_{ij} \beta_j \equiv 0 \quad \text{für alle} \quad \mathfrak{y} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

$$\sum_j \left(\sum_i \overline{\alpha_i} a_{ij} \right) \beta_j = 0.$$

Wir setzen

$$\mathfrak{z} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \quad \text{mit} \quad \gamma_j = \sum \overline{\alpha_i} a_{ij}$$

und erhalten

$$\sum \overline{\gamma_j} \beta_j = 0 \quad \text{für alle} \quad (\beta_1, \ldots, \beta_n),$$

also sind alle $\overline{\gamma_j} = 0$. Damit ist $\sum_{i=1}^n \overline{\alpha_i} a_{ij} = 0$, nicht alle $\overline{\alpha_i} = 0$, und somit ist A singulär.

(2) Der Fall $s(x, y) \equiv 0$ für ein $y \neq 0$ wird genauso behandelt.

(3) $A = A_{s;\mathcal{A}}$ sei singulär. Dann existiert ein $\mathfrak{x} \neq 0$ mit ${}^t\overline{\mathfrak{x}} A = 0$. Ist $x \in V$ mit $\mathfrak{x}$ als Koordinatenvektor, so ist für alle $y \in V$ also (y hat den Koordinatenvektor $\mathfrak{y}$):

$$s(x, y) = {}^t\overline{\mathfrak{x}} A \mathfrak{y} = 0,$$

d.h. s ist entartet. – Analog findet man $y \neq 0$ mit $s(x, y) \equiv 0 \quad \text{für alle} \quad x$.

Bemerkung. Im endlichdimensionalen Fall sind die Bedingungen (1) und (2) aus Definition 5 also äquivalent.

Wir beachten nun, daß die Sesquilinearformen auf V einen $\mathbf{K}$-Vektorraum bilden, unter den Verknüpfungen

$$(s_1 + s_2)(x, y) = s_1(x, y) + s_2(x, y)$$
$$(\lambda s)(x, y) = \lambda \cdot s(x, y);$$

wir nennen ihn $\mathrm{Sq}(V)$; die hermitischen Sesquilinearformen bilden einen $\mathbf{R}$-Untervektorraum hiervon (nicht einen $\mathbf{C}$-Untervektorraum!) Damit können wir die bisherigen Aussagen (mit einer kleinen Erweiterung) zusammenfassen zu

Satz 1.2 *Es sei $\mathcal{A}$ eine Basis von V. Die Zuordnung*

$$s \longmapsto A_{s;\mathcal{A}}$$

definiert einen Isomorphismus von $\mathbf{K}$-Vektorräumen zwischen

$$\mathrm{Sq}(V) \to \mathbf{K}^{n\times n}.$$

Hierbei gehen die hermitischen Formen in den $\mathbf{R}$-Unterraum der hermitischen Matrizen über, die nicht-entarteten Formen werden auf die regulären Matrizen abgebildet.

Wir holen eine Definition nach:

Definition 1.8. *$A \in \mathbf{K}^{n\times n}$ heißt hermitisch, wenn*

$${}^t\overline{A} = A$$

ist.

Für $\mathbf{K} = R$ ist A also eine symmetrische Matrix.

Definition 1.9. *Eine hermitische Matrix A heißt positiv definit (-semidefinit), wenn die Form*

$$s(\mathfrak{x}, \mathfrak{y}) = {}^t\overline{\mathfrak{x}} A \mathfrak{y}$$

positiv definit (-semidefinit) ist.

4.) Wie im Falle von Endomorphismen müssen wir nun die Transformationsformeln bei Basiswechsel aufstellen.
$\mathcal{A} = \{a_1, \ldots, a_n\}$ und $\mathcal{B} = \{b_1, \ldots, b_n\}$ seien Basen,

$$S = A_{id;\mathcal{B},\mathcal{A},}$$

d.h. $b_i = \sum_j s_{ji} a_j$. Für eine Sesquilinearform s ist dann

$$\begin{aligned} s(b_i, b_j) &= s\left(\sum_k s_{ki} a_k, \sum_l s_{lj} a_l\right) \\ &= \sum_{kl} \overline{s}_{ki} s_{lj} s(a_k, a_l), \end{aligned}$$

d.h. wir haben

Satz 1.3

$$A_{s;\mathcal{B}} = {}^t\overline{S} A_{s;\mathcal{A}} S$$

Sowohl Endomorphismen als auch Sesquilinearformen werden durch $n \times n$- Matrizen beschrieben. Daß es sich um unterschiedliche Objekte handelt, kommt in dieser Beschreibung durch die unterschiedlichen Tranformatiosgesetze bei Basiswechsel heraus. Genauer:

Definition 1.10. *$A, B \in \mathbf{K}^{n \times n}$ heißen kongruent, wenn es eine reguläre Matrix $S \in \mathbf{K}^{n \times n}$, mit*

$${}^t\overline{S} A S = B$$

gibt.

Offensichtlich haben wir eine Äquivalenzrelation definiert. Die Gruppe $Gl(V)$ operiert auf $\mathrm{Sq}(V)$ durch

$$s_\varphi(x, y) = s(\varphi(x), \varphi(y))$$

(für $\varphi \in Gl(V)$ und $s \in \mathrm{Sq}(V)$.) Nennt man Sesquilinearformen kongruent, $s_1 \equiv s_2$, wenn es ein φ gibt mit $s_2 = s_{1,\varphi}$, so sagt Satz 1.3 (etwas erweitert) aus, daß die Zuordnung

$$s \longmapsto A_{s;\mathcal{A}}$$

eine umkehrbare eindeutige (basisunabhängige) Abbildung zwischen Kongruenzklassen von Sesquilinearformen und Kongruenzklassen von Matrizen definiert.

Vergleichen wir mit der Situation bei Endomorphismen eines Vektorraumes: $\dim V = n$, der Grundkörper darf beliebig sein. Auf dem Endomorphismenring $\operatorname{End} V$ operiert dann $\mathrm{Gl}(V)$ durch Konjugation: $f, g \in \operatorname{End} V$ heißen ähnlich, wenn es ein $\varphi \in \mathrm{Gl}(V)$ gibt mit

$$g = \varphi^{-1} \circ f \circ \varphi;$$

Matrizen $A, B \in K^{n\times n}$ hatten wir schon früher ähnlich genannt, wenn es $S \in \mathrm{Gl}(n, K)$ gibt mit

$$B = S^{-1}AS.$$

Dann entnehmen wir den früheren Überlegungen: der Isomorphismus

$$\begin{aligned} \operatorname{End} V &\longrightarrow K^{n\times n} \\ f &\longmapsto A_{f;\mathcal{A},\mathcal{A}}, \end{aligned}$$

der mittels einer Basis $\mathcal{A}$ definiert wird, liefert eine basisunabhängige umkehrbar eindeutige Abbildung zwischen den Mengen der Ähnlichkeitsklassen von Endomorphismen und den Mengen der Ähnlichkeitsklassen von Matrizen.
Ähnlichkeit und Kongruenz sind durchaus verschiedene Äquivalenzrelationen auf $\mathbf{K}^{n\times n}$.

5.) Wir schauen nochmal den $\mathbf{K}^n$ an. Jedes Skalarprodukt wird durch eine reguläre hermitische Matrix A nach der Formel

$$s(\mathfrak{x}, \mathfrak{y}) = {}^t\overline{\mathfrak{x}}A\mathfrak{y}$$

gegeben. Ist $\mathcal{E}$ die übliche Basis $e_j = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftarrow j$, so ist $A_{s;\mathcal{E}} = A$.
Für $A = E =$ Einheitsmatrix hat man das übliche Skalarprodukt

$$\langle \mathfrak{x}, \mathfrak{y} \rangle = {}^t\overline{\mathfrak{x}} \cdot \mathfrak{y} = \sum \overline{x}_i y_i.$$

Schließlich noch eine terminologische Bemerkung: in der Literatur werden oft nur positiv definite hermitische Formen als Skalarprodukte bezeichnet. Ich werde mich in den folgenden Paragraphen dieser Konvention anschließen.

§2 Normen

1.) Wir beginnen mit der folgenden

Definition 2.1. *Eine Norm* $\|\cdot\|$ *auf dem* $\mathbf{K}$*-Vektorraum* V *ist eine Abbildung*

$$V \to \mathbf{R}$$

$$x \longmapsto \|x\|$$

mit:

(1) $\|x\| \geq 0$, *und zwar* $= 0$ *genau dann, wenn* $x = 0$,

(2) $\|\lambda x\| = |\lambda| \|x\|$ *für* $\lambda \in \mathbf{K}$, $x \in V$,

(3) $\|x+y\| \leq \|x\| + \|y\|$.

V, $\|\cdot\|$ *heißt normierter Vektorraum.*

Wir geben einige Beispiele.

Beispiel 1 $V = \mathbf{K}$, $\|\cdot\| = |\cdot|$ Absolutbetrag.

Beispiel 2 $V = \mathbf{K}^n$; $\|x\| = \max |x_i|$ (mit $x = (x_1, \ldots, x_n)$).

Beispiel 3 $V = \mathbf{K}^n$; $\|x\| = \sum |x_i|$.

Beispiel 4 $V = C^0_{\mathbf{C}}[a,b]$, $\|f\| = \max_{[a,b]} |f(x)|$.

Beispiel 5 Wie Beispiel 4, $\|f\| = \int_a^b |f(x)| dx$.

Wenn wir mit mehreren dieser Normen gleichzeitig zu tun haben, müssen wir sie natürlich unterschiedlich bezeichnen: $\|\cdot\|$, $|||\cdot|||$, $\| \ \|_1$, $\|\cdot\|_\infty$ zum Beispiel.

2.) Es folgt das wichtigste Beispiel.

Satz 2.1 *Es sei* V, $\langle\,,\rangle$ *ein unitärer Vektorraum. Durch*

$$\|x\| = \langle x, x\rangle^{1/2}$$

wird eine Norm definiert.

Satz 2.2 *(Schwarzsche Ungleichung). Voraussetzungen wie in Satz 2.1. Für $x, y \in V$ ist dann*

$$|\langle x, y\rangle|^2 \leq \|x\|^2 \|y\|^2$$
$$(= \langle x, x\rangle\langle y, y\rangle);$$

Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Zusatz: Die Ungleichung von Satz 2.1 gilt auch für positiv semidefinite hermitische Formen (nicht aber die Aussage über Bestehen der Gleichheit).
Beide Sätze sind im letzten Semester bewiesen worden – ich wiederhole hier den Beweis.

Beweis von Satz 2.2. (1) x oder $y = 0$: dann ist die Aussage klar.
(2) Es sei $\langle x, x\rangle \neq 0$, $\lambda \in \mathbf{K}$ beliebig. Dann ist

$$\begin{aligned} 0 \leq \langle \lambda x + y, \lambda x + y\rangle &= |\lambda|^2 \langle x, x\rangle + \overline{\lambda}\langle x, y\rangle \\ &\quad + \lambda\langle y, x\rangle + \langle y, y\rangle \end{aligned}$$

Wir setzen $\lambda = -\langle x, y\rangle / \langle x, x\rangle$ und erhalten

$$0 \leq \frac{|\langle x, y\rangle|^2}{\langle x, x\rangle} - \frac{|\langle x, y\rangle|^2}{\langle x, x\rangle} - \frac{|\langle x, y\rangle|^2}{\langle x, x\rangle} + \langle y, y\rangle,$$

also

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

(3) Gleichheit heißt: für obiges λ ist $\langle \lambda x + y, \lambda x + y\rangle = 0$, also $\lambda x = -y$.

Beweis von Satz 2.1. Nur die Dreiecksungleichung (3) ist nicht offensichtlich. Nun ist

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\,\mathrm{Re}\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\| + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \quad \text{(Satz 2.2)} \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

3.) Es sei $V, \|\cdot\|$ ein normierter Vektorraum. Durch

$$d(x,y) = \|x-y\|$$

wird dann eine Metrik (Abstand zwischen zwei Vektoren) eingeführt, und die Eigenschaften einer Norm übersetzen sich in

(1) $d(x,y) \geq 0, = 0$ genau für $x = y$,

(2) $d(x,y) = d(y,x)$,

(3) $d(x,y) \leq d(x,z) + d(z,y)$.

Mittels d werden nun die Begriffe „Konvergenz“, „Stetigkeit“, „offen“, „abgeschlossen“, „gleichmäßige Konvergenz (Stetigkeit)“ in V definiert: die Formulierungen für den $\mathbf{R}^n$ bleiben wörtlich dieselben. Insbesondere steht der Kompaktheitsbegriff zur Verfügung, in folgender Formulierung:

Definition 2.2. *$M \subset V$ heißt kompakt, wenn jede offene Überdeckung von M eine endliche Teilüberdeckung enthält.*

Definition 2.3. *Eine Teilmenge $M \subset V$ heißt beschränkt, wenn es ein $C > 0$ gibt, so daß $\|x\| \leq C$ für alle $x \in M$ gilt.*

Wie im $\mathbf{R}^n$ sieht man, daß kompakte Mengen abgeschlossen und beschränkt sind. Vorsicht: die Umkehrung gilt nur in endlichdimensionalen Vektorräumen – siehe später. Wir untersuchen nun lineare Abbildungen im Rahmen dieser Begriffe.

Hilfssatz 1. *Eine lineare Abbildung $f : V \to W$ zwischen normierten Vektorräumen ist genau dann überall stetig, wenn sie im Nullpunkt stetig ist. Das gilt genau dann, wenn es ein $C \geq 0$ gibt mit*

$$\|f(x)\| \leq C\|x\|$$

für alle $x \in V$. (f heißt dann beschränkt.)
(Links steht die Norm auf W, rechts die von V).

Beweis. f sei in 0 stetig. Dann existiert zu $\epsilon = 1$ ein $\delta > 0$ mit

$$\|f(x)\| < 1 \text{ für alle } \|x\| < \delta.$$

Ist nun $y \neq 0$ beliebig, so ist

$$\begin{aligned} \| f(y) \| &= \| f\left(\frac{y}{\|y\|}\right) \| \, \|y\| \\ &= \| f\left(\frac{\delta y}{\|y\|}\right) \| \frac{1}{\delta} \|y\| \\ &\leq \frac{1}{\delta} \|y\| . \end{aligned}$$

Somit ist f beschränkt mit $\frac{1}{\delta} = C$. Daraus folgt, wenn $x_0 \in V$ und $\epsilon > 0$ beliebig sind:

$$\begin{aligned} \| f(x) - f(x_0) \| &= \| f(x - x_0) \| \\ &\leq C \| x - x_0 \| \quad < \epsilon, \end{aligned}$$

falls $\| x - x_0 \| < \frac{\epsilon}{C}$, d.h. die Stetigkeit von f in x_0.

Vorsicht: ist V ∞-dimensional, so brauchen lineare Abbildungen nicht stetig zu sein! Ist etwa

$$\begin{aligned} V &= C^1_{\mathbf{R}}[a, b], \quad \| f \| = \max |f(x)|, \\ W &= C^0_{\mathbf{R}}[a, b], \quad \| f \| \text{ wie oben}, \end{aligned}$$

und $D : V \to W$ die lineare Abbildung

$$(Df)(x) = f'(x),$$

so gilt z.B.

$$\lim_{n\to\infty} \frac{1}{n} \sin nx = 0 \quad \text{in } V,$$

aber $\lim_{n\to\infty} \left(\frac{1}{n} \sin nx\right)' = \lim_{n\to\infty} \cos nx$ ist nicht existent. – Das Beispiel ist typisch für die Verhältnisse bei linearen Differentialoperatoren.

Wir wenden uns nun endlichdimensionalen normierten Vektorräumen zu.

Hilfssatz 2. *Es sei $a_1, \ldots, a_n$ eine Basis des $\mathbf{K}$-Vektorraumes V. Setze für*

$$\begin{aligned} x &= \sum_{i=1\ldots n} \lambda_i a_i \quad \in V : \\ |x| &= \sum |\lambda_i|. \end{aligned}$$

Dann ist $|\cdot|$ eine Norm auf V, und die Abbildung

$$x \longmapsto (\lambda_1, \ldots, \lambda_n) \in \mathbf{K}^n$$

ist ein Homeomorphismus.

Hieraus folgt
Hilfssatz 3. *Die Mengen $H = \{x : |x| = 1\}$ und $M_c = \{x : |x| \leq c\}$ sind – bezüglich der „Summennorm“$|\cdot|$ – kompakt.*

Hilfssatz 4. *$V, \|\cdot\|$ sei ein n-dimensionaler normierter $\mathbf{K}$ -Vektorraum und $|\cdot|$ die in Hilfssatz 2 eingeführte Summennormen. Dann sind die Abbildungen*

$$id : \ V, |\cdot| \to V, \|\cdot\|$$

und

$$\|\cdot\| \ : \ V, |\cdot| \to \mathbf{R}$$

stetig.

Beweis (1) $\|x\| = \|\sum_j \lambda_j a_j\| \leq \sum |\lambda_j| \|a_j\| \leq C|x|$, mit

$$C = \max \|x_j\|.$$

(2) $\|\cdot\| : \quad V, \|\cdot\| \to \mathbf{R}$ ist stetig. Denn: Es sei etwa $\|x_0\| \leq \|x\|$. Dann ist $\|x\| = \|x - x_0 + x_0\| \leq \|x - x_0\| + \|x_0\|$, also $\|x\| - \|x_0\| \leq \|x - x_0\|$.
Für $\|x_0\| \geq \|x\|$ ist entsprechend

$$\|x_0\| - \|x\| \leq \|x - x_0\|,$$

insgesamt

$$|\|x\| - \|x_0\|| \leq \|x - x_0\|.$$

Daraus folgt aber die Stetigkeit von $\|\cdot\|$.

(3) $\|\cdot\| : \ V, |\cdot| \to \mathbf{R}$ ist das Kompositum

$$V, |\cdot| \xrightarrow{id} V, \|\cdot\| \xrightarrow{\|\cdot\|} \mathbf{R},$$

also stetig.

Hilfssatz 5. *$V, \|\cdot\|$ sei ein n-dimensionaler normierter Vektorraum mit einer Basis $a_1, \dots, a_n$. Für $x = \sum_j \lambda_j a_j \in V$ sei*

$$f_j(x) = \lambda_j.$$

Dann sind die f_j linear und stetig von $V, \|\cdot\|$ nach $\mathbf{K}$, und es gibt ein $C > 0$ mit

$$|x| \leq C \| x \| .$$

(mit $|\cdot|$ als mittels der a_j definierten Summennorm).

Beweis. Nach Hilfssatz 4 ist $\|\cdot\|$ stetig auf $V, |\cdot|$. Nach Hilfssatz 3 ist $M = \{x : |x| = 1\}$ kompakt. Wegen $\| x \| > 0$ auf M gibt es ein $c > 0$ mit $\| x \| \geq c$ auf M. Nun sei $x \neq 0$ beliebig.Dann ist

$$\| x \| = \| \frac{x}{|x|} |x| \| = |x| \| \frac{x}{|x|} \| \geq c|x|.$$

Es folgt weiter

$$| f_j(x)| = |\lambda_j| \leq |x| \leq \frac{1}{c} \| x \| .$$

Hilfssatz 6 *$f : V \to W$ sei eine lineare Abbildung normierter Vektorräume;* $\dim V = n < \infty$. *Dann ist f stetig.*

Beweis. $a_1, \dots, a_n$ sei eine Basis und $|\cdot|$ die zugehörige Summennorm. Es sei $x = \sum_j \lambda_j a_j \in V$. Dann

$$\begin{aligned}
\| f(x) \| &= \| f(\sum_j \lambda_j a_j) \| \\
&= \| \sum_j \lambda_j f(a_j) \| \\
&\leq \sum |\lambda_j| \| f(a_j) \| \\
&\leq C|x| \quad \text{mit } C = \max \| f(a_j) \| \\
&\leq C' \| x \| \quad \text{nach Hilfssatz 5.}
\end{aligned}$$

Wir fassen die Ergebnisse zusammen in
Satz 2.3 *Es sei V ein endlichdimensionaler Vektorraum mit Norm $\|\cdot\|$. Dann gilt:*

(1) Jede lineare Abbildung $f: V \to W$ in irgendeinen normierten Vektorraum ist stetig.
(2) Ist $|||\cdot|||$ eine weitere Norm auf V, so gibt es positive Zahlen c, C mit

$$c\,\|x\| \leq |||x||| \leq C\,\|x\|.$$

(3) Die Mengen $S_r = \{x : \ \|x\| = r\}$ und $B_r = \{x : \ \|x\| \leq r\}$ sind kompakt.
(4) Stetigkeit und Konvergenz bez. einer Norm implizieren Stetigkeit und Konvergenz bez. aller Normen.

Beweis. (1) ist gerade Hilfssatz 6. und (2) folgt aus (1): setze $f = id$. (4) folgt aus (2). Schließlich ergibt sich (3) aus Hilfssatz 3 und aus (4).
Aus diesem Satz folgern wir

Satz 2.4 *s sei eine hermitische Sesquilinearform auf dem normierten Vektorraum V, $\|\cdot\|$. Dann sind folgende Aussagen äquivalent:*
(1) s ist positiv definit.
(2) Es gibt ein $c > 0$ mit

$$s(x, x) \geq c\,\|x\|^2.$$

Beweis. Offensichtlich folgt aus (2) die positive Definitheit. Umgekehrt sei $|||\cdot|||$ die durch s definite Norm, also $|||x|||^2 = s(x, x)$. Nach Satz 2.3 gilt dann die Ungleichung (2).

Folgerung. Für eine hermitische Matrix A sind äquivalent:
(1) A ist positiv definit.
(2) Es gibt ein positives c mit

$${}^t\overline{x}Ax \geq c \sum |x_j|^2$$

für alle $x \in \mathbf{K}^n$.

Einen Teil von Satz 2.3 formulieren wir mit einem neuen Begriff.

Definition 2.4. *Zwei Normen $\|\cdot\|, |\cdot|$ auf V heißen äquivalent, wenn es $c, C > 0$ mit*

$$c|x| \leq \|x\| \leq C|x|$$

für alle $x \in V$ gibt.

(Es handelt sich in der Tat um eine Äquivalenzrelation!). Damit lautet 2.3 (3): Je 2 Normen auf einem endlichdimensionalen Vektorraum sind äquivalent.

Man überzeugt sich leicht, daß die Normen

$$\| f \|_1 = \int_a^b |f(x)| dx, \quad \| f \|_2 = \left[\int_a^b |f(x)|^2 dx \right]^{1/2}$$

und $\| f \|_\infty = \max_{x \in [a,b]} |f(x)|$ paarweise inäquivalent sind.

§3 Orthogonalität

1.) $V, \langle\,,\rangle$ sei immer ein unitärer bzw. euklidischer Vektorraum. $\langle\,,\rangle$ ist also das positiv-definite Skalarprodukt, $\|\cdot\|$ bezeichnet die zugehörige Norm.

Definition 3.1. *(1) Wir schreiben $x \perp y$ für $\langle x, y\rangle = 0$ und nennen x und $y \in V$ dann orthogonal (senkrecht) zueinander.*

(2) $M, N \subset V$. $M \perp N$ („senkrecht“) heißt: für alle $x \in M$ und $y \in N$ ist $x \perp y$.

(3) $M \subset V$. $M^\perp = \{x : \; x \perp M\}$ heißt orthogonales Komplement von M.

Die folgenden Aussagen sind unmittelbare Konsequenzen der Definitionen:

(1) $M^\perp$ ist ein Unterraum von V.

(2) $M \perp N$ impliziert für die linearen Hüllen $L(M)$ und $L(N)$ ebenfalls $L(M) \perp L(N)$.

(3) $M \cap M^\perp \subset \{0\}$.

(4) Wenn $x \perp y$, dann gilt der „Satz des Pythagoras“

$$\| x + y \|^2 = \| x \|^2 + \| y \|^2$$

(für $\mathbf{K} = \mathbf{R}$ gilt auch die Umkehrung).

Definition 3.2. *Ein Orthonormalsystem (ONS) ist eine endliche oder unendliche Folge von Vektoren* $e_1, e_2, e_3, \ldots$ *mit*

$$\langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j. \end{cases}$$

D.h. also: die e_i stehen paarweise senkrecht aufeinander und haben jeweils die Norm 1. Offenbar ist ein ONS linear unabhängig.

Satz 3.1 *Es sei* $a_1, a_2, a_3 \ldots \in V$ *ein linear unabhängiges (endliches oder unendliches) System von Vektoren. Dann existiert genau ein ONS* $e_1, e_2, e_3, \ldots$ *mit folgenden Eigenschaften:*

(1) $L(a_1, \ldots, a_k) = L(e_1, \ldots, e_k), \quad k = 1, 2, 3, \ldots$
(2) $e_k = \sum_{j=1 \ldots k} \lambda_{jk} a_j$ *mit* $\lambda_{kk} > 0$.

(Mit $L(a_1, \ldots, a_k)$ *haben wir den von* $a_1, \ldots, a_k$ *erzeugten Unterraum bezeichnet.)*

Der folgende Beweis konstruiert die e aus den a effektiv; das Verfahren heißt das Gram-Schmidtsche Orthonormalisierungsverfahren. „Die e entstehen aus den a durch Orthonormalisierung“.

Beweis von Satz 3.1. (Induktive Konstruktion der e_k).

Für $k = 1$ setze man $e_1 = a_1 / \| a_1 \|$.

Nun seien $e_1, \ldots, e_k$ bereits so konstruiert, daß die Aussagen (1) und (2) für alle $j \leq k$ gelten. Wir setzen mit noch zu wählenden $\lambda'_j \in \mathbf{K}$, $j = 1, \ldots, k$,

$$e'_{k+1} = \lambda'_1 e_1 + \cdots + \lambda'_k e_k + a_{k+1}.$$

Unabhängig von den λ'_j ist dann $e'_{k+1} \notin L(e_1, \ldots, e_k) = L(a_1, \ldots, a_k)$, da a_{k+1} linear unabhängig von $a_1, \ldots, a_k$ ist; insbesondere ist $e'_{k+1} \neq 0$.

Als nächstes erreichen wir $e'_{k+1} \perp e_j$ für $j = 1, \ldots, k$ durch Wahl von λ'_j:

$$0 = \langle e'_{k+1}, e_j \rangle = \lambda'_j + \langle a_{k+1}, e_j \rangle.$$

Damit sind die λ'_j festgelegt; wir setzen nun

$$e_{k+1} = e'_{k+1} / \| e'_{k+1} \|$$

und können dann leicht die gewünschten Beziehungen für $e_1, \ldots, e_{k+1}$ und $a_1, \ldots, a_{k+1}$ verifizieren.

Analyse der Konstruktion zeigt auch die eindeutige Bestimmtheit der e_j.

Definition 3.3. *Es sei* $\dim V = n$. *Eine Orthonormalbasis ONB von V ist eine Basis* $e_1, \ldots, e_n$, *die ein Orthonormalsystem bildet.*

Folgerung 1. Jeder endlichdimensionale unitäre Vektorraum hat ONBen.

Folgerung 2. Es sei $A \in \mathbf{K}^{n \times n}$ eine positiv definite hermitische Matrix. Dann existiert eine reguläre Matrix S, und zwar eine Dreiecksmatrix mit positiven Diagonalelementen, so daß

$$ {}^t\overline{S} A S = E \quad \text{(Einheitsmatrix)}$$

ist. – D.h. positiv definite hermitische Matrizen sind kongruent zur Einheitsmatrix.

Beweis. Definiere im $\mathbf{K}^n$ als Skalarprodukt

$$s(x, y) = {}^t\overline{x} A y.$$

Es seien a_i die Vektoren $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$, und $\mathcal{A} = a_1, \ldots, a_n$ die daraus gebildete Basis, dann ist

$$A_{s;\mathcal{A}} = A.$$

Orthonormalisiert man diese Basis zu $\mathcal{E} = \{e_1, \cdots, e_n\}$, so ist

$$A_{s;\mathcal{E}} = E,$$

also, mit Basistransformationsmatrix $S = A_{id;\mathcal{E},\mathcal{A}}$:

$$E = {}^t\overline{S}AS.$$

Beachte, daß der Gram-Schmidtsche Basisübergang durch eine Dreiecksmatrix S beschrieben wird.

Zu einer Matrix A wollen wir die linke obere $k \times k$-Untermatrix wieder mit A_k bezeichnen.

Satz 3.2 *Für eine hermitische Matrix $A = (a_{ij})$ sind folgende Aussagen äquivalent:*

(1) A ist positiv definit.
(2) Für jedes k mit $1 \le k \le n$ ist

$$\det A_k = \det(a_{ij})_{i,j=1\ldots k} > 0$$

(Insbesondere sind alle diese Determinanten reell).

Beweis. Es sei A positiv definit. Nach Folgerung 2 gibt es dann $S \in \mathrm{Gl}(n, \mathbf{K})$ mit

$${}^t\overline{S}AS = E,$$

also

$$|\det S|^2 \det A = 1.$$

Damit ist $\det A > 0$. Betrachte nun das durch A definierte Skalarprodukt auf $L(e_1, \ldots, e_n)$; dabei sei $e_i = (0, \ldots, 1, \ldots, 0)$. A_k ist die Matrix des Skalarproduktes bez. der Basis $e_1, \ldots, e_k$, also hat A_k positive Determinante.

Umgekehrt sei (2) erfüllt. Wir führen Induktion nach n. Für $n = 1$ ist die Behauptung trivial, sei sie also für $n-1 \ge 1$ schon bewiesen. Es sei s die durch A definierte Sesquilinearform auf dem $\mathbf{K}^n$, ihre Einschränkung auf $\mathbf{K}^{n-1} = L(e_1, \ldots, e_{n-1})$ bezeichnen wir auch mit s. Die Matrix dieser Einschränkung bezüglich der Basis $e_1, \ldots, e_{n-1}$ ist gerade A_{n-1}, also ist s dort nach Induktionsannahme positiv definit. Es sei $a_1, \ldots, a_{n-1} \in L(e_1, \ldots, e_{n-1})$ eine ONB von $L(e_1, \ldots, e_{n-1})$ für s, die nach Gram-Schmidt konstruiert wird.

Ferner sei

$$a_n = e_n - \sum_{i=1}^{n-1} s(e_n, a_i)a_i.$$

Dann ist $a_1, \ldots, a_{n-1}, a_n$ eine Basis des $\mathbf{K}^n$, und die Matrix von s bez. dieser Basis ist $\mathrm{diag}(1, \ldots, 1; \beta)$, mit einem $\beta \in \mathbf{K}$. Wir haben nun

$$\begin{aligned} {}^t\overline{S}AS &= \mathrm{diag}(1, \ldots, 1; \beta) \\ |\det S|^2 \det A &= \det \mathrm{diag}(1, \ldots, \beta) = \beta, \end{aligned}$$

also $\beta > 0$. Damit ist s positiv definit, also auch A.

2.) V sei weiterhin ein unitärer resp. euklidischer Vektorraum.

Definition 3.4. *V ist die orthogonale Summe der Unterräume $V_1, \ldots, V_r$, wenn*

$$V = V_1 + \ldots + V_r$$

und $V_i \perp V_j$ für $i \neq j$ gilt.

Wir schreiben dann

$$V = V_1 \overline{\oplus} \ldots \overline{\oplus} V_r = V_1 \perp \ldots \perp V_r.$$

Offenbar ist eine orthogonale Summe auch direkt, und in der eindeutigen Zerlegung

$$x = x_1 + x_2 + \ldots + x_r \quad , \; x_i \in V_i$$

gilt

$$\| x \|^2 = \| x_1 \|^2 + \ldots + \| x_r \|^2,$$

(Satz des Pythagoras.)

Satz 3.3 *Es sei W ein endlichdimensionaler Teilraum eines unitären Vektorraums V. Dann ist*

$$V = W \overline{\oplus} W^{\perp}$$

und

$$W^{\perp\perp} = W.$$

Beweis. (1) Offenbar ist $W \cap W^{\perp} = \{0\}$, wir zeigen $W + W^{\perp} = V$.

Dazu sei $e_1, \ldots, e_n$ eine ONB von W. Ist $x \in V$, so sei

$$\begin{aligned} x_1 &= \sum_{i=1}^{n} \langle x, e_i \rangle e_i \qquad \in W, \\ x_2 &= x - x_1. \end{aligned}$$

Dann ist

$$\begin{aligned} \langle x_2, e_j \rangle &= \langle x, e_j \rangle - \sum_{i=1}^{n} \langle x, e_i \rangle \langle e_i, e_j \rangle \\ &= 0. \end{aligned}$$

Somit ist $x_2 \in W^{\perp}$, und $x = x_1 + x_2$.

(2) Sicher ist $W^{\perp\perp} \supset W$. Es sei nun $x \in W^{\perp\perp}$. Nach Teil (1) des Beweises ist

$$x = x_1 + x_2, \qquad x_1 \in W,\ x_2 \in W^{\perp},$$

also

$$0 = \langle x, x_2 \rangle = \langle x_1, x_2 \rangle + \langle x_2, x_2 \rangle = \langle x_2, x_2 \rangle,$$

damit folgt $x = x_1 \in W$.

Bemerkung Ist in der Situation von Satz 3

$$x = x_1 + x_2, \quad x_1 \in W,\ x_2 \in W^\perp,$$

so prüft man sofort die Beziehung

$$\| x_2 \| = \|x - x_1\| = \min\{ \| x - u \| : u \in W\}$$

nach. Es ist ja für $u \in W$

$$\begin{aligned} \| x - u \|^2 &= \| x_1 - u + x_2 \|^2 = \| x_1 - u \|^2 + \| x_2 \|^2 \\ &\geq \| x_2 \|^2, \end{aligned}$$

mit Gleichheit genau für $u = x_1$. Die Zuordnung $x \mapsto x_1 = p(x)$ ist linear, mit Kern $W^\perp$. Wir nennen sie die orthogonale Projektion auf den (endlichdimensionalen) Unterraum W. Die Norm $\| x_2 \|$ kann als Abstand des Punktes x vom Unterraum W angesehen werden.

Ist V selbst endlichdimensional, so natürlich auch $W^\perp$, und die Rollen von W und $W^\perp$ können im Satz vertauscht werden. Im allgemeinen Fall notieren wir die topologischen Aussagen

Satz 3.4 *(1) Ein endlichdimensionaler Teilraum eines normierten Vektorraumes ist abgeschlossen.*

(2) $M^\perp$ ist ein abgeschlossener Unterraum des unitären Vektorraumes V.

Beides ist leicht zu zeigen, wird aber im folgenden keine Rolle spielen.

§4 Der Spektralsatz für normale Endomorphismen

1.) Wir studieren lineare Abbildungen zwischen unitären (euklidischen) Vektorräumen. Grundlegend ist

Satz 4.1 *(Riesz/Fischer). Es sei V ein endlich-dimensionaler unitärer* **K***-Vektorraum. Für jedes $x \in V$ sei l_x die Linearform*

$$l_x(y) = \langle x, y \rangle.$$

Dann ist durch

$$x \longmapsto l_x$$

ein Semi-Isomorphismus von V auf den Dualraum V^* erklärt.

Das heißt also insbesondere: ist l irgendeine Linearform auf V, so existiert genau ein $x \in V$ mit

$$l(y) = \langle x, y \rangle$$

für alle $y \in V$: „l kann durch Skalarprodukt mit einem Vektor $x \in V$ gegeben werden“.

Beweis. Offenbar ist l_x linear und die Zuordnung $x \longmapsto l_x$ semilinear. Es reicht nun, ihre Injektivität zu zeigen. Aus $l_x = 0$ folgt

$$\langle x, y \rangle = 0$$

für alle $y \in V$ und damit $x = 0$.

(Wir sehen, daß die Aussage also für nichtentartete Skalarprodukte gilt. Wir bleiben aber im positiv definiten Fall).

Es sei nun

$$f : V \to W$$

eine lineare Abbildung endlichdimensionaler unitärer **K**-Vektorräume. Für $y \in W$ ist

$$x \longmapsto l_y(x) = \langle y, f(x) \rangle$$

eine Linearform l_y auf V. Nach Satz 4.1 gibt es also ein wohlbestimmtes $z \in V$ mit

$$l_y(x) = \langle z, x \rangle$$

für alle x. Wir nennen dieses z, das ja von y abhängt, $f^*(y)$ und haben damit eine (offensichtlich lineare) Abbildung von W nach V definiert. Insgesamt:

Satz 4.2 *Zu einer linearen Abbildung $f : V \to W$ unitärer* **K***-Vektorräume endlicher Dimension existiert genau eine lineare Abbildung $f^* : W \to V$ mit*

$$\langle y, f(x) \rangle = \langle f^*(y), x \rangle$$

für alle $y \in W$ und $x \in V$.

Definition 4.1. *f^* heißt die zu f adjungierte lineare Abbildung.*

f^* hängt folgendermaßen mit der früher betrachteten *dualen* Abbildung, die wir momentan mit $f^\#$ bezeichnen wollen, zusammmen: Der Satz von Riesz/Fischer definiert Semiisomorphismen $\iota_V : V \to V^*$ bzw. $\iota_W : W \to W^*$. Dann ist

$$f^* = \iota_V^{-1} \circ f^\# \circ \iota_W.$$

(Da 2 Semiisomorphismen in der Formel auftreten, ist das Kompositum linear.)

Auch im folgenden seien alle auftretenden Vektorräume endlichdimensional. Die Bildung des Adjungierten hat folgende leicht zu verifizierende Eigenschaften:

(1) $0^* = 0$, $id^* = id$, $f^{**} = f$, $(g \circ f)^* = f^* \circ g^*$.

(2) $(\text{Kern } f)^\perp = \text{Bild } f^*$, $(\text{Bild } f)^\perp = \text{Kern } f^*$.
(3) $\text{rg } f = \text{rg } f^*$.

Beweis von (3): $\text{rg } f = \dim \text{Bild } f = \dim(\text{Kern } f^*)^\perp = \text{codim}\,\text{Kern } f^* = \text{rg } f^*$. Alternativ kann man natürlich auch $\text{rg } f^* = \text{rg } f^\# = \text{rg } f$ benutzen.

2.) Wir müssen die obigen Ergebnisse jetzt in der Sprache der Matrizen interpretieren.

Es seien $\mathcal{A} = \{a_1, \ldots, a_n\}$, $\mathcal{B} = \{b_1, \ldots, b_m\}$ ONBen von V bzw. W,

$$\begin{aligned} A &= A_{f;\mathcal{A},\mathcal{B}} = (\alpha_{ij}) \\ B &= A_{f^*;\mathcal{B},\mathcal{A}} = (\beta_{ij}). \end{aligned}$$

Also hat man

$$\begin{aligned} f(a_i) &= \sum \alpha_{ji} b_j \\ f^*(b_k) &= \sum \beta_{lk} a_l \\ \langle f(a_i), b_k \rangle &= \sum_j \overline{\alpha}_{ji} \langle b_j, b_k \rangle = \overline{\alpha}_{ki}, \end{aligned}$$

aber auch

$$\langle a_i, f^*(b_k) \rangle = \sum_l \beta_{lk} \langle a_i, a_l \rangle = \beta_{ik}.$$

Somit ist $\beta_{ik} = \overline{\alpha}_{ki}$, d.h. wir haben

Satz 4.3 *Bez. ONBen $\mathcal{A}$ und $\mathcal{B}$ gilt für die Matrizen von f und f^* die Beziehung*

$$A_{f^*;\mathcal{B},\mathcal{A}} = {}^t\overline{A}_{f;\mathcal{A},\mathcal{B}}.$$

3.) Wir kehren zu Endomorphismen zurück.

Definition 4.2. *Ein Endomorphismus f eines unitären* **K**-*Vektorraums heißt normal, wenn er mit seinem Adjungierten kommutiert:*

$$f \circ f^* = f^* \circ f.$$

Z.B. sind 0 und *id* normale Endomorphismen. Wir können die Bedingung auch anders formulieren:

f ist genau dannn normal, wenn stets

$$\langle f(x), f(y) \rangle = \langle f^*(x), f^*(y) \rangle$$

ist.

Ist nämlich f normal, so gilt

$$\begin{aligned} \langle f(x), f(y) \rangle &= \langle x, f^* f(y) \rangle \\ &= \langle x, f f^*(y) \rangle \\ &= \langle f^*(x), f^*(y) \rangle. \end{aligned}$$

Umgekehrt berechnen wir

$$\begin{aligned} \langle f^* f(x), y \rangle &= \langle f(x), f(y) \rangle \\ &= \langle f^*(x), f^*(y) \rangle \\ &= \langle f f^*(x), y \rangle \end{aligned}$$

und schließen $f \circ f^* = f^* \circ f$.

Folgerung Ist f normal, so ist Kern f = Kern f^* und damit Kern $f \perp$ Bild f.

4.) Ich erinnere jetzt an den Begriff des Eigenwertes und die damit zusammenhängenden Begriffe und Ergebnisse. Der Grundkörper K sei beliebig. V sei ein K-Vektorraum, $f : V \to V$ ein Endomorphismus. Ein Eigenwert (EW) $\lambda \in K$ von f ist ein Körperelement, für das

$$V(\lambda) = V_f(\lambda) = \mathrm{Kern}(\lambda\, id - f) \neq 0$$

ist. Die Elemente $x \neq 0$ in $V(\lambda)$ heißen Eigenvektoren (EV) von f zum Eigenwert $\lambda, V(\lambda)$ der Eigenraum zu λ. Ist $A \in K^{n\times n}$ eine Matrix, so sind die Eigenwerte von A per definitionem die Eigenwerte des Endomorphismus $x \longmapsto Ax$ von K^n; entsprechend werden die Eigenvektoren und Eigenräume von A erklärt.

Ist A eine beschreibende Matrix für f bez. einer Basis, so sind die EWe von f genau die von A, und die Eigenräume in V bzw. K^n entsprechen sich unter dem Koordinatenisomorphismus $V \xrightarrow{\sim} K^n$. Die Eigenwerte von f hängen nur von der Ähnlichkeitsklasse von f ab, dito für A.

Die Eigenwerte von A sind die Nullstellen des charakterisierten Polynoms

$$\begin{aligned} \chi_A(x) &= \det(EX - A) \quad \in K[X] \\ &= X^n + \ldots, \end{aligned}$$

somit gibt es höchstens n Eigenwerte in K für $A \in K^{n\times n}$. Ist $L \supset K$ eine Körpererweiterung, in der $\chi_A(X)$ in Linearfaktoren zerfällt – so eine Erweiterung gibt es immer –,

$$\chi_A(X) = \prod_{j=1\ldots k} (X - \lambda_j)^{m_j}$$

$\lambda_j \in L$, $\lambda_j \neq \lambda_i$ für $i \neq j$, $\sum m_j = n$, so sind die λ_j die EWe von A, aufgefaßt als Matrix in $L^{n\times n}$; m_j nennt man die algebraische Vielfachheit von λ_j.

Notieren wir noch:

$$\begin{aligned} \det A &= \lambda_1^{m_1} \ldots \lambda_k^{m_k} \quad \in K \\ \mathrm{Sp}\, A &= m_1\lambda_1 + \ldots m_k\lambda_k \in K. \end{aligned}$$

Für einen Endomorphismus f ist das charakteristische Polynom

$$\chi_f(X) = \chi_A(X),$$

wobei A eine beschreibende Matrix für f ist; es hängt nur von der Ähnlichkeitsklasse von A und von f ab. Die Gleichung

$$\chi_f(X) = 0 \quad \text{bzw.} \quad \chi_A(X) = 0$$

heißt charakteristische Gleichung oder Säkulargleichung.

Schließlich erinnere ich an den Begriff „diagonalisierbar“: f (bzw. A) heißt so, wenn es eine Basis aus EVen von f (bzw. A) gibt.

5.) Ich kehre jetzt zu unitären $\mathbf{K}$-Vektorräumen V der Dimension n zurück.

Satz 4.4 *Es sei f ein Endomorphismus von V, f^* sein adjungierter Endomorphismus. Dann gilt für $\lambda \in \mathbf{K}$*

$$(\lambda id + f)^* = \overline{\lambda} id + f^*.$$

Ist ferner f normal, λ ein EW von f und a ein EV zu λ, so ist a auch ein EV von f^*, und zwar zum EW $\overline{\lambda}$. Die EWe von f^* sind konjugiert komplex zu denen von f, die Eigenräume sind identisch.

Zum Beweis berechnen wir

$$\begin{aligned} \langle(\lambda\, id + f)(x), y\rangle &= \langle \lambda x, y\rangle + \langle f(x), y\rangle \\ &= \langle x, \overline{\lambda} y\rangle + \langle x, f^*(y)\rangle \\ &= \langle x, (\overline{\lambda}\, id + f^*)(y)\rangle. \end{aligned}$$

Der Rest ergibt sich aus dieser Formel und der Folgerung nach Definition 4.2.

Hauptresultat dieses Kapitels ist der folgende

Satz 4.5 *(Spektralsatz für normale Endomorphismen) Folgende Aussagen sind für einen Endomorphismus $f : V \to V$ eines unitären $\mathbf{K}$-Vektorraumes äquivalent.*

(1) f ist normal, und alle EWe von f gehören zu **K**.

(2) Es gibt eine ONB von EVen für f.

(Im Falle $\mathbf{K} = \mathbf{C}$ ist die Voraussetzung über die EWe von f immer erfüllt; für $\mathbf{K} = \mathbf{R}$ lautet sie: „alle EWe von f seien reell“).

Beweis. durch Induktion nach $n = \dim V$. Für $n = 1$ ist die Aussage trivial; wir nehmen den Satz für $n - 1$ nun an.
Es sei $\lambda_n \in \mathbf{K}$ ein EW von f und e_n ein EV dazu mit $\| e_n \| = 1$. Dann ist

$$U = e_n^{\perp}$$

ein $n - 1$-dimensionaler unitärer VR und $f|U : U \to U$ noch immer ein normaler Endomorphismus. Denn für $x \in U$ ist

$$\langle f(x), e_n \rangle = \langle x, f^*(e_n) \rangle = \overline{\lambda}_n \langle x, e_n \rangle = 0,$$

also ist $f(x) \in U$, folglich U invariant unter f, und $(f|U)^* = f^* \mid U$.

Nach Induktionsannahme existiert eine ONB $e_1, \ldots, e_{n-1}$ von EVen von $f \mid U$ in U, insgesamt also ist $e_1, \ldots, e_{n-1}, e_n$ eine ONB von EVen für f.

Ist umgekehrt $\mathcal{A} = \{e_1, \ldots, e_n\}$ eine ONB von EVen für f, so gilt

$$A_{f;\mathcal{A},\mathcal{A}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

wobei $\lambda_j \in \mathbf{K}$ die EWe von f (nicht notwendig alle verschieden) sind. Dann ist

$$A_{f^*;\mathcal{A},\mathcal{A}} = \operatorname{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_n)$$

offenbar mit $A_{f;\mathcal{A},\mathcal{A}}$ vertauschbar.

Zusatz. *f sei normal. Sind dann $V(\lambda)$ und $V(\mu)$ die Eigenräume zu verschiedenen EWen $\lambda \neq \mu$, so ist $V(\lambda) \perp V(\mu)$.*

Beweis. Es sei $x \in V(\lambda)$, $y \in V(\mu)$. Dann ist

$$\begin{aligned}\overline{\lambda}\langle x,y\rangle &= \langle \lambda x,y\rangle = \langle f(x),y\rangle = \langle x, f^*(y)\rangle \\ &= \langle x,\overline{\mu}y\rangle = \overline{\mu}\langle x,y\rangle,\end{aligned}$$

also $\langle x,y\rangle = 0$.

§5 Selbstadjungierte Abbildungen und Isometrien

1.) Nach wie vor sei V ein endlich-dimensionaler unitärer $\mathbf{K}$-Vektorraum, also euklidisch für $\mathbf{K} = \mathbf{R}$. $\langle\ ,\ \rangle$ ist das Skalarprodukt.

Definition 5.1. *$f : V \to V$ heißt selbstadjungiert, wenn $f = f^*$ ist.*

Natürlich ist f dann normal. Wir haben mehr als das:

Satz 5.1 *Ein selbstadjungierter Endomorphismus hat nur reelle Eigenwerte.*

Als Folgerung aus dem Spektralsatz ergibt sich nun

Satz 5.2 *(Diagonalisierung selbstadjungierter Endomorphismen)*
Zu jedem selbstadjungierten Endomorphismus gibt es eine ONB von EVen.

Gibt es umgekehrt zu f eine ONB von EVen und sind alle EWe reell, so ist f natürlich selbstadjungiert.

Beweis von Satz 5.1. Die EWe von f^* sind konjugiert zu denen von f. Aber $f = f^*$, d.h. alle EWe sind reell. Genauer: ist λ EW von f, so ist $\overline{\lambda}$ EW von f^* mit demselben Eigenraum $V(\lambda)$. Wegen $f = f^*$ gilt für einen EV x in $V(\lambda)$: $\lambda x = f(x) = \overline{\lambda}x$, also $\lambda = \overline{\lambda}$.

2.) Der folgende Begriff war schon im vorigen Semester eingeführt worden:

Definition 5.2. *$f : V \to V$ heißt eine Isometrie, wenn für alle $x, y \in V$*

$$\langle f(x), f(y)\rangle = \langle x,y\rangle$$

ist. Für $\mathbf{K} = \mathbf{R}$ heißen Isometrien orthogonale, für $\mathbf{K} = \mathbf{C}$ unitäre Abbildungen.

Satz 5.3 *Für $f : V \to V$ sind folgende Aussagen äquivalent:*

(1) f ist eine Isometrie.
(2) Für alle $x \in V$ ist $\| f(x) \| = \| x \|$.
(3) Für alle x mit $\| x \| = 1$ ist $\| f(x) \| = 1$.
(4) Für jede ONB von V gilt: das Bild unter f ist wieder eine ONB.
(5) Es gibt eine ONB von V, die unter f auf eine ONB abgebildet wird.

Den einfachen Beweis aus dem vergangenem Semester reproduziere ich nicht.

Klarerweise sind Isometrien bijektiv und normal. Genauer gilt

Satz 5.4 *Äquivalent für f sind die folgenden Aussagen:*

(1) f ist eine Isometrie.
(2) f ist normal mit EWen vom Betrag 1.
(3) $f^{-1} = f^*$.

Beweis. (1) $\Leftrightarrow$ (3): f^{-1} existiert, und

$$\langle f(x), y \rangle = \langle f(x), f f^{-1}(y) \rangle = \langle x, f^{-1}(y) \rangle,$$

also $f^{-1} = f^*$. f vertauscht aber mit f^{-1}.

Ist umgekehrt $f^{-1} = f^*$, so ist f natürlich normal, und wir haben

$$\begin{aligned} \langle f(x), f(y) \rangle &= \langle x, f^* f(y) \rangle \\ &= \langle x, f^{-1} f(y) \rangle = \langle x, y \rangle. \end{aligned}$$

(1) $\Leftrightarrow$ (2): Diese Äquivalenz beweisen wir zunächst nur im Falle $\mathbf{K} = \mathbf{C}$.

Es sei λ ein EW von f, $x \neq 0$ ein EV zu λ. Dann haben wir

$$\| x \|^2 = \| f(x) \|^2 = |\lambda|^2 \| x \|^2,$$

also $|\lambda| = 1$. Ist umgekehrt (2) erfüllt, so ist nach dem Spektralsatz eine ONB von EVen vorhanden, wobei alle EWe den Betrag 1 haben. Damit führt f diese ONB wieder in eine ONB über und ist somit eine Isometrie.

Den Fall $\mathbf{K} = \mathbf{R}$ behandeln wir etwas später.

Aus dem Spektralsatz folgt nun

Satz 5.5 *Es sei $f : V \to V$ eine Isometrie, deren EWe alle zu* $\mathbf{K}$ *gehören. Dann existiert eine ONB von EVen von f.*

Als weitere Folgerung notieren wir: die Determinante einer orthogonalen Abbildung ist ± 1.

3.) Wenden wir uns nun der Interpretation der obigen Aussagen im Matrizenkalkül zu; dabei wird auch der Beweis von Satz 5.4 vervollständigt werden.

Es sei $\mathcal{A} = \{a_1, \ldots, a_n\}$ eine ONB von V. $f : V \to V$ sei selbstadjungiert und $A = A_{f;\mathcal{A},\mathcal{A}}$ die Matrix von f bez. $\mathcal{A}$. Dann ist ${}^t\overline{A} = A_{f^*;\mathcal{A},\mathcal{A}}$; und wegen $f = f^*$:

Satz 5.6 *f ist genau dann selbstadjungiert, wenn bez. einer ONB die Matrix f hermitisch ist. Dann ist bez. jeder ONB die Matrix von f hermitisch.*

Für $\mathbf{K} = \mathbf{R}$ sind die hermitischen Matrizen gerade die symmetrischen. Wir haben mitbewiesen:

Folgerung. Die EWe einer hermitischen (bzw. reellen symmetrischen) Matrix sind alle reell.

Nun möge $f : V \to V$ eine Isometrie sein. Dann folgt aus $A_{f^*;\mathcal{A},\mathcal{A}} = {}^t\overline{A}$ und $f^* = f^{-1}$:

Satz 5.7 *f ist genau dann eine Isometrie, wenn bez. einer (und damit aller) ONB die Matrix $A = A_{f;\mathcal{A},\mathcal{A}}$ die Beziehung*

$${}^t\overline{A} = A^{-1}$$

erfüllt.

Definition 5.3. *(1) $A \in \mathbf{C}^{n\times n}$ heißt unitär, wenn*

$$^t\overline{A} = A^{-1}$$

gilt.
(2) $A \in \mathbf{R}^{n\times n}$ heißt orthogonal, wenn die Beziehung

$$^tA = A^{-1}$$

besteht (d.h. A reell-unitär ist).
(3) $A \in \mathbf{K}^{n\times n}$ heißt normal, wenn $A \cdot {}^t\overline{A} = {}^t\overline{A}A$ ist.

Wir bemerken noch, daß die Isometrien eines $\mathbf{K}$-Vektorraumes V bei der Komposition offenbar eine Untergruppe

$$\mathrm{Iso}\, V \subset \mathrm{Gl}(V)$$

bilden. Demnach bilden die unitären bzw. orthogonalen Matrizen Untergruppen von $\mathrm{Gl}(n, \mathbf{K})$; Bezeichnung:

$$\left.\begin{array}{ll} O(n) \subset \mathrm{Gl}(n, \mathbf{R}) & \text{orthogonale} \\ U(n) \subset \mathrm{Gl}(n, \mathbf{C}) & \text{unitäre} \end{array}\right\} \text{Gruppe.}$$

Nun führen wir den Beweis von Satz 5.4 zu Ende. Zunächst folgt aus Satz 5.7 und dem schon bewiesenen Teil von Satz 5.4

Satz 5.8 *Die Eigenwerte einer unitären (und daher erst recht einer orthogonalen) Matrix haben den Betrag 1.*

Damit hat jede Isometrie $f : V \to V$ eines euklidischen Vektorraums EWe nur vom Bertrag 1 (in $\mathbf{C}$). Ist umgekehrt f normal mit komplexen EWen vom Betrag 1, so folgt: f wird bezüglich einer ONB von V durch eine Martrix A mit

$$A \cdot {}^tA = {}^tAA$$

gegeben. A definiert einen normalen Endomorphismus des $\mathbf{C}^n$ (mit üblichem Skalarprodukt) mit EWen vom Betrag 1 und ist daher eine unitäre Matrix, also, da reell, eine orthogonale Matrix. Damit ist f eine Isometrie.

4.) Jetzt können wir den Spektralsatz zur Ähnlichkeitsklassifikation normaler, insbesondere hermitischer Matrizen, heranziehen. A und $B \in \mathbf{K}^{n\times n}$ heißen ähnlich, $A \sim B$, wenn es ein $S \in \mathrm{Gl}(n, \mathbf{K})$ mit

$$B = S^{-1}AS$$

gibt. Dann stimmen natürlich die charakteristischen Polynome χ_A und χ_B überein. Kann man S sogar als unitäre Matrix wählen, so sollen A und B unitär ähnlich heißen – im reellen Fall orthogonal ähnlich. Notieren wir:

(1) Jedes B, welches unitär ähnlich zu einer normalen Matrix A ist, ist selbst normal (für $\mathbf{K} = \mathbf{R}$: „orthogonal ähnlich“).
(2) Ist A normal mit EWen $\lambda_1, \dots, \lambda_n \in \mathbf{K}$, so ist A unitär ähnlich zu $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$, resp. orthogonal ähnlich, wenn A reell und $\lambda_1, \dots, \lambda_n$ reell sind.
(3) $A, B \in \mathbf{C}^{n\times n}$ seien ähnlich. Wenn A, B beide normal sind, so sind sie auch unitär ähnlich.

Der wichtigste Fall ist

Satz 5.9 *Jede hermitische Matrix ist unitär ähnlich zu einer reellen Diagonalmatrix; jede reelle symmetrische Matrix ist orthogonal ähnlich zu einer Diagonalmatrix.*

Wie stellt man nun bei einer normalen Matrix A die Diagonalform her? Die Frage haben wir schon im letzten Semester allgemein beantwortet – hier nochmals:

Es sei

$$S^{-1}AS = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

mit $S \in U(n)$. Dann sind die λ_i die EWe von A. Wir haben weiter

$$\begin{aligned} AS &= S \cdot D \\ &= (\lambda_1 u_1, \dots, \lambda_n u_n), \\ \text{also} \quad Au_i &= \lambda_i u_i, \\ \text{wenn} \quad S &= (u_1 u_2 \dots u_n) \text{ ist} \end{aligned}$$

(u_i Spaltenvektoren). Da S unitär ist, müssen – siehe Schluß des Paragraphen – die u_i eine ONB von EVen von A bilden, für den $\mathbf{C}^n$ mit dem üblichen hermitischen Skalarprodukt.

Somit hat man folgendes zu tun:

1. Lösung der charakteristischen Gleichung

$$\chi_A(X) = 0$$

zur Bestimmung der λ_i. Das ist das Hauptproblem.

2. Bestimmung der Eigenräume zu den verschiedenen λ_i; sie sind automatisch orthogonal und spannen den $\mathbf{C}^n$ auf.

3. Bestimmung einer ONB zu jedem Eigenraum.

5.) Abschließend eine Bemerkung, die wir schon früher hätten machen können:

(1) Die Transformationsmatrix zwischen zwei ONBen ist unitär resp. orthogonal.

(2) Die Spalten und Zeilen einer unitären Matrix (oder orthogonalen Matrix) bilden eine ONB des $\mathbf{K}^n$ (für das übliche Skalarprodukt). Hierdurch sind unitäre resp. orthogonale Matrizen charakterisiert.

Bemerkung (2) ist nichts weiter als die Ausführung der Bedingung

$${}^t\overline{S}S = E,$$

die unitäre Matrizen charakterisiert.

§6 Klassifikation hermitischer Formen

1.) Hauptergebnis ist die folgende Interpretation des Spektralsatzes:

Satz 6.1 *Es sei $V, \langle\, ,\rangle$ ein unitärer n-dimensionaler* $\mathbf{K}$*-Vektorraum und s eine hermitische Sesquilinearform auf V. Dann existiert eine ONB von V, bezüglich welcher s durch eine Diagonalmatrix D mit reellen Diagonalelementen beschrieben wird. Bis auf die Reihenfolge der Diagonalelemente ist D eindeutig bestimmt.*

Beweis. (1) Es sei $\mathcal{B} = \{b_1, \dots, b_n\}$ eine ONB und $B = A_{s;\mathcal{B}}$ die Matrix von S bez. $\mathcal{B}$. Nach dem Klassifikationssatz 5.9 gibt es eine unitäre Matrix S mit

$$S^{-1}BS = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

Aber $S^{-1} = {}^t\overline{S}$, also

$${}^t\overline{S}BS = D.$$

Nun sei $\mathcal{A} = \{a_1, \dots, a_n\}$ die durch

$$A_{id;\mathcal{A},\mathcal{B}} = S$$

definierte Basis, d.h.

$$a_i = \sum s_{ji} b_j, \quad S = (s_{ij}).$$

Da S unitär ist, ist $\mathcal{A}$ eine ONB, und

$$A_{s;\mathcal{A}} = {}^t\overline{S}BS = D.$$

(2) D ist durch das charakteristische Polynom χ_B festgelegt (bis auf die Reihenfolge der Diagonalelemente). Ist C irgendeine ONB und $C = A_{s;C}$, so ist

$$C = {}^t\overline{T}BT = T^{-1}BT$$

mit einer unitären Matrix T, also $\chi_C = \chi_B$; somit hängt χ_B nur von s ab.

Satz 6.2 *Es sei s eine hermitische Sesquilinearform auf einem n-dimensionalen* **K**-*Vektorraum V. Dann existiert eine Bais $\mathcal{A}$ von V, bezüglich welcher die Matrix $A_{s;\mathcal{A}}$ die Gestalt*

$$\mathrm{diag}(\underbrace{1,\ldots,1}_{k_+},\underbrace{-1,\ldots,-1}_{k_-},\underbrace{0,\ldots,0}_{k_0})$$

hat. Der Tripel (k_+, k_-, k_0) *mit* $0 \leq k_{+,-,0}$, $k_+ + k_- + k_0 = n$ *ist durch* s *eindeutig bestimmt.*

Definition 6.1. (k_+, k_-, k_0) *heißt Signatur von* s, $k_+ + k_-$ *Rang von* s, $k_- - k_+$ *Index von* s.

Satz 6.2 wird auch Sylvesterscher Trägheitssatz genannt. Die in der Definition eingeführten Begriffe werden in der Literatur unterschiedlich gebraucht.

Beweis von Satz 6.2 Wir führen mittels einer willkürlich gewählten Basis, die per definitionem orthogonal sei, ein Skalarprodukt auf V ein. Nach Satz 6.1 gibt es eine Basis $\mathcal{A}$, bezüglich welcher s durch die Diagonalmatrix.

$$D = \mathrm{diag}(\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_l, 0, \ldots, 0)$$

gegeben wird, mit $\lambda_\kappa > 0$, $\mu_\lambda < 0$.

Sei nun

$$S = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_k}}, \frac{1}{\sqrt{|\mu_1|}}, \ldots, \frac{1}{\sqrt{|\mu_l|}}, 1, \ldots, 1\right).$$

und $D_0 = {}^t\overline{S}DS$. Dann beschreibt D_0 die Form s bez. der mittels S transformierten Basis, und es ist

$$D_0 = \mathrm{diag}\left(\underbrace{1,\ldots,1}_{k},\underbrace{-1,\ldots,-1}_{l},\underbrace{0,\ldots,0}\right).$$

2) Es bleibt die Eindeutigkeit von D_0 zu zeigen. Es seien also

$$\mathcal{A} = \{a_1, \ldots, a_k; a'_1, \ldots, a'_l; a''_1, \ldots, a''_m\}$$
$$\mathcal{B} = \{b_1, \ldots, b_r; b'_1, \ldots, b'_s; b''_1, \ldots, b''_t\}$$

zwei Basen mit

$$A_{s;\mathcal{A}} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{k}; \underbrace{-1, \ldots, -1}_{l}; \underbrace{0, \ldots, 0}_{m})$$
$$A_{s;\mathcal{B}} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r}; \underbrace{-1, \ldots, -1}_{s}; \underbrace{0, \ldots, 0}_{t})$$

Also ist $k + l + m = n = r + s + t$.

Es sei

$$\begin{aligned} V_+ &= L(a_1, \ldots, a_k) \\ V_- &= L(a'_1, \ldots, a'_l) \\ V_0 &= L(a''_1, \ldots, a'_m), \end{aligned}$$

W_0, W_+, W_- seien entsprechend mittels $\mathcal{B}$ definiert. Es sei $x \in V_0$. D.h. für beliebiges $y \in V$ gilt

$$\begin{aligned} s(x, y) &= s\left(\sum \alpha''_i a''_i, \sum \beta_j a_j + \sum \beta'_j a'_j + \sum \beta''_j a''_j\right) \\ &= 0. \end{aligned}$$

Umgekehrt: aus $s(x, y) \equiv 0$ für alle y folgt $x \in V_0$. Damit ist

$$V_0 = \{x : s(x, y) \equiv 0\} = W_0 \quad \text{und} \quad m = t.$$

Nun sei $x \in W_- \cap (V_0 \oplus V_+)$. Dann ist

$$
\begin{aligned}
s(x,x) &\leq 0 \quad , \text{da} \quad x \in W_- \\
s(x,x) &\geq 0 \quad , \text{da} \quad x \in V_+ \oplus V_0,
\end{aligned}
$$

also $s(x,x) = 0$. Aber auf W_- ist $-s$ positiv definit, also $x = 0$. Demnach

$$
\dim W_- = s \leq n - (k+m) = l.
$$

Analog ist $l = \dim V_- \leq n - (r+t) = s$, also $l = s$ und schließlich $k = r$.

Eine Umformulierung des Sylvesterschen Trägheitssatzes ist

Satz 6.2' *Es sei s eine hermitische Sesquilinearform auf dem* $\mathbf{K}$*-Vektorraum V. Dann existiert eine direkte Zerlegung*

$$
V = V_0 \oplus V_+ \oplus V_-,
$$

so daß gilt:

1) $s(x,y) \equiv 0$ für $x \in V_0$ und $y \in V$.
2) s ist positiv definit auf V_+.
3) $-s$ ist positiv definit auf V_-.
4) $s(x,y) = 0$ für $x \in V_+, y \in V_-$.

V_0 ist durch s eindeutig bestimmt, ebenso $\dim V_+$ *und* $\dim V_-$.

Noch eine Bemerkung zur Interpretation dieser Sätze. Der Sylvestersche Trägheitssatz garantiert, daß jede hermitische Sesquilinearform durch eine (sehr spezielle) Diagonalmatrix beschrieben werden kann, bezüglich einer geeigneten Basis; ist die Form positiv definit, wird sie (bez. einer ONB) durch die Einheitsmatrix beschrieben.

Der Spektralsatz (Satz 6.1) bezieht sich auf Paare von hermitischen Sesquilinearformen; ist eine dieser Formen positiv definit, so gibt es eine Basis, in der beide Formen gleichzeitig durch Diagonalmatrizen beschrieben werden, und zwar eine positiv definite durch die Einheitsmatrix; die zweite Form wird in einer geeigneten

ONB für die erste Form durch eine reelle Diagonalmatrix gegeben. In der Diagonale stehen die EWe der Matrix, welche die zweite Form in einer willkürlichen ONB bez. der ersten Form beschreibt.

2.) Wir übersetzen die obigen Sätze nun in die Sprache der Matrizen.

Definition 6.2. *$A, B \in \mathbf{K}^{n\times n}$ heißen kongruent, $A \equiv B$, wenn es ein reguläres S mit*

$$B = {}^t\overline{S}AS$$

gibt. $A \equiv_u B$ („unitär kongruent“), wenn $S \in U(n)$ existiert mit

$$B = {}^t\overline{S}AS$$

(vgl. Def. 1.10). Wir notieren:

Ist A hermitisch und $B \equiv A$, so ist B hermitisch. Ist $B \equiv_u A$, so ist $\chi_A = \chi_B$ (charakteristische Polynome).

Jetzt gilt

Satz 6.3 *(1) Jede hermitische Matrix A ist kongruent zu einer Diagonalmatrix*

$$\mathrm{diag}(\underbrace{1,\dots,1}_{k}, \underbrace{-1,\dots,-1}_{l}, \underbrace{0,\dots,0}_{m}).$$

(k, l, m) mit $k + l + m = n$ sind durch A eindeutig bestimmmt als Signatur von A.

(2) Jede hermitische Matrix A ist unitär kongruent zu einer Diagonalmatrix

$$\mathrm{diag}(\lambda_1,\dots,\lambda_k, \mu_1,\dots,\mu_l, \underbrace{0,\dots,0}_{m}),$$

mit $\lambda_\kappa > 0$, $\mu_\lambda < 0$. (k, l, m) ist die Signatur von A, die Einträge sind gerade die EWe von A.

(3) $A \equiv B$ genau, wenn Signatur A =Signatur B.

(4) $A \equiv_u B$ genau, wenn $\chi_A = \chi_B$.

3.) Wie berechnet man die Invarianten?

(1) Die Invarianten für unitäre Kongruenz: siehe früher.

(2) Die Berechnung der Signatur. Hier wäre es dumm, zunächst die EWe von A zu berechnen und dann die Signatur als Anzahl der positiven , der negativen und der Null-EWe festzustellen. Stattdessen bringen wir A durch Transformation mit regulären Matrizen in der folgenden Weise auf Diagonalgestalt:

Es sei $A = (\alpha_{ij})$.

Fall 1. $\alpha_{11} \neq 0$.

Setze

$$S = \begin{pmatrix} 1 & \beta_2 & \cdots & \beta_n \\ 0 & & & \\ \vdots & & E_{n-1} & \\ 0 & & & \end{pmatrix}$$

mit E_{n-1} als Einheitsmatrix in $\mathbf{K}^{(n-1)\times(n-1)}$ und geeignet zu wählenden $\beta_2, \ldots, \beta_n \in \mathbf{K}$. Bilde

$$A' = {}^t\overline{S}AS.$$

Ausmultiplizieren zeigt: wählt man

$$\beta_i = -\alpha_{1i}/\alpha_{11},$$

so hat A_1 die Gestalt

$$A_1 = \begin{pmatrix} a_{11} & 0 \\ 0 & A_2 \end{pmatrix}$$

mit einer hermitischen Matrix $A_2 \in \mathbf{K}^{(n-1)\times(n-1)}$.

Fall 2. $\alpha_{11} = 0$, aber nicht alle $\alpha_{1j} = 0$. Sei etwa $\alpha_{ij} \neq 0$ für $j > 1$. Setze dann

$$S = \begin{pmatrix} 1 & 0 \cdots 0 & 1 & 0 \cdots 0 \\ 0 & & & \\ \vdots & & E_{n-1} & \\ 0 & & & \end{pmatrix};$$

dabei steht die 1 an der Stelle j; dann ist

$$A_1 = {}^t\overline{S}AS = (\alpha'_{ij})$$

hermitisch mit $\alpha'_{11} \neq 0$.

Fall 3. $\alpha_{1j} = 0$ für alle j. Dann hat A die Gestalt

$$A = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & A_1 & \\ 0 & & \end{pmatrix}$$

mit einer hermitischen Matrix $A_1 \in \mathbf{K}^{(n-1)\times(n-1)}$.

In jedem Falle reduziert man das Problem um eine Dimension und ermittelt so sukzessive eine reelle Diagonalmatrix, die kongruent zu A ist und an der man die Signatur von A abliest.

In anderer Sprechweise kann man die Ergebnisse so formulieren: jede hermitische Bilinearform kann durch Einführung neuer Koordinaten in die Gestalt

$$\sum_{j=1}^{n} \epsilon_j \bar{z}_j w_j$$

mit $\epsilon_j = 0, \pm 1$ überführt werden, jede relle symmetrische Bilinearform in die Gestalt

$$\sum_{j=1}^{n} \epsilon_j x_j y_j$$

mit $\epsilon_j = 0, \pm 1$. Die ϵ_j sind jeweils durch die Form bis auf die Reihenfolge bestimmt. Das obige Verfahren liefert die gewünschte Umformung.

§7 Anwendungen

1.) Wir beginnen mit einer geometrischen Anwendung.

Definition 7.1. *Eine Quadrik im* $\mathbf{R}^n$ *ist die Nullstellenmenge eines quadratischen Polynoms.*

Offensichtlich kann ein solches Polynom folgendermaßen gegeben werden:

$$q(x) = q(x_1, \ldots, x_n) = {}^t x A x + {}^t a x + b$$

mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $A \in \mathbf{R}^{n \times n}$ eine symmetrische Matrix, $a \in \mathbf{R}^n$ und $b \in \mathbf{R}$. Um eine Vorstellung von der Gestalt der Quadrik

$$Q = \{x : q(x) = 0\}$$

zu bekommen, führen wir geeignete neue Koordinaten mittels der Ergebnisse des vorigen Paragraphen ein. Zunächst existiert nach dem Sylvesterschen Trägheitssatz ein $S \in \mathrm{Gl}(n, \mathbf{R})$ mit

$$ {}^t S A S = \mathrm{diag}\left(\underbrace{1, \ldots, 1}_{k}, \underbrace{-1, \ldots, -1}_{l}, \underbrace{0, \ldots, 0}_{m} \right).$$

Ersetzt man also x durch Sx und schreibt für ${}^taS$ wiederum ${}^ta$, so wird die Gleichung in den neuen, wieder mit x bezeichneten Koordinaten,

$$x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2 + a_1x_1 + \ldots + a_nx_n + b = 0.$$

Wir können noch vereinfachen, indem wir

$$\begin{aligned} x_1^2 + a_1x_1 &= \left(x_1 + \frac{a_1}{2}\right)^2 - \frac{a_1^2}{4}, \\ -x_{k+1}^2 + a_{k+1}x_{k+1} &= -\left(x_{k+1} - \frac{a_{k+1}}{2}\right)^2 + \frac{a_{k+1}^2}{4} \end{aligned}$$

schreiben, entsprechend an den anderen Stellen, die neuen Koordinaten $x_1 + a_1/2$ usw. einführen und wieder mit x_1 usw. bezeichnen; dann wird die Gleichung

$$x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2 + a_{k+l+1}x_{k+l+1} + \ldots + a_nx_n + b = 0,$$

wobei b eine neue Bedeutung hat. Falls nicht alle a_j für $j > k + l$ Null sind, führen wir als neue Koordinate

$$\widetilde{x}_{k+l+1} = a_{k+l+1}x_{k+l+1} + \ldots + a_nx_n$$

ein. Schließlich normieren wir die Koordinaten noch so, daß $b = 0$ oder ± 1 wird. Damit wird die Gleichung

$$x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2 + \epsilon x_{k+l+1} + b = 0$$

mit $\epsilon = 0$ oder 1, $b = 0$ oder ± 1, $0 \leq k \leq n$, $0 \leq l \leq n$, $k + l \leq n$.

Die Gestalt von Q kann hieraus abgelesen werden – das ist für $n = 2$ oder 3 Gegenstand elementarer Geometrie. Z.B. wird durch

$$x_1^2 + x_2^2 + x_3^2 - 1$$

eine Sphäre gegeben, durch $x_1^2 + x_2^2 - x_3^2 = 1$ ein einschaliges Hyperboloid, usw. Das beschriebene Verfahren wird in geometrischer Sprechweise als Hauptachsentransformation quadratischer Hyperflächen bezeichnet (genauer: affine Hauptachsentransformation). Arbeitet man nur mit unitären Abbildungen, so spricht man von euklidischer Hauptachsentransformation. Man vergleiche G. Fischer: Analytische Geometrie (Hamburg 1978).

2.) Als zweites Beispiel betrachten wir kleine Schwingungen eines mechanischen Systems. Es werde durch die Koordinaten $(x_1, \ldots, x_n) = x$ beschrieben und befinde sich bei $x = 0$ in der Ruhelage. Die kinetische Energie sei eine positiv definite quadratische Form in den Geschwindigkeiten $\dot{x}_j$, die potentielle Energie eine in den Auslenkungen x_j :

$$T = \frac{1}{2}{}^t\dot{x}T\dot{x}, \quad V = \frac{1}{2}{}^t xVx,$$

T und V positiv definite $n \times n$-Matrizen. Die Bewegungsgleichungen sind dann

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}_j} - \frac{\partial T}{\partial x_j} = -\frac{\partial V}{\partial x_j} \quad , \quad j = 1 \ldots n,$$

also nach leichter Rechnung

$$T\ddot{x} + Vx = 0.$$

Nach dem Spektralsatz existiert ein $S \in \mathrm{Gl}(n, \mathbf{R})$ mit

$${}^tSTS = E \quad \text{und} \quad {}^tSVS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \quad \lambda_i > 0.$$

Einführen der neuen Variablen

$$x = Sq$$

liefert nach Multiplikation mit ${}^tS$:

$$\ddot{q} + Dq = 0, \quad D = \text{diag}\,(\lambda_1, \ldots, \lambda_n).$$

Hieraus liest man den Bewegungsablauf des Systems ab:

$$\begin{aligned} q_j &= A_j e^{i\omega_j t} + B_j e^{-i\omega_j t}, \\ \omega_j^2 &= \lambda_j. \end{aligned}$$

§8 Topologie von klassischen Gruppen

1.) $\mathbf{K} = \mathbf{R}, \mathbf{C}$. Es sei $G = \mathrm{Gl}(n, \mathbf{K})$.

Satz 8.1 *Die folgenden Abbildungen sind reell-analytisch, im Falle* $\mathbf{K} = \mathbf{C}$ *sogar holomorph.*

(1)

$$\begin{aligned} \mu : \mathbf{K}^{n\times n} \times \mathbf{K}^{n\times n} &\to \mathbf{K}^{n\times n} \\ A, B &\longmapsto AB. \end{aligned}$$

(2)

$$\det : K^{n\times n} \to \mathbf{K}.$$

(3)

$$\begin{aligned} \iota : G &\to G \\ A &\longmapsto A^{-1}. \end{aligned}$$

Insbesondere sind die Gruppenoperationen auf G reell-analytisch bzw. holomorph, und G ist offen in $\mathbf{K}^{n\times n}$.

2.) Wir verwenden jetzt einige topologische Begriffe, die aus der Analysis bekannt sein sollten.

Satz 8.2 $\mathrm{Gl}(n, \mathbf{R})$ *hat* 2 *Wegkomponenten,* $\mathrm{Gl}(n, \mathbf{C})$ *ist zusammenhängend.*

Beweis. 1) $\det : \mathrm{Gl}(n, \mathbf{R}) \to \mathbf{R}^*$ ist stetig surjektiv, also hat $\mathrm{Gl}(n, \mathbf{R})$ mindestens 2 Komponenten.

Hilfssatz. *Es sei* $E^n_\pm = \mathrm{diag}(1, \ldots, 1, \pm 1) \in G_n = \mathrm{Gl}(n, \mathbf{K})$*. Dann ist jedes* $A \in G_n$ *mit* E^n_+ *oder* E^n_- *verbindbar (i.e. es gibt einen Weg in* G_n *von* A *nach* E^n_+ *bzw.* E^n_-*).*

Beweis. Es sei $A = (\ldots \mathfrak{a} \ldots \mathfrak{b} \ldots) \in G_n$, wobei $\mathfrak{a}, \mathfrak{b}$ Spaltenvektoren sind. Dann ist A in G_n verbindbar

$$\begin{aligned} &(a) \text{ mit} \quad A' = (\ldots \mathfrak{a} \ldots \mathfrak{b} + s\mathfrak{a} \ldots), \quad s \in \mathbf{R} \\ &(b) \text{ mit} \quad A' = (\ldots -\mathfrak{a} \ldots -\mathfrak{b} \ldots). \end{aligned}$$

Um (a) zu zeigen, wähle man $t \in [0, 1]$ und bilde

$$A(t) = (\ldots \mathfrak{a} \ldots \mathfrak{b} + ts\mathfrak{a} \ldots).$$

Dann ist $\det A(t) = \det A$ und $A(0) = A$, $A(1) = A'$.

(b) folgt aus (a), indem wir der Reihe nach die folgenden Matrizen gemäß (a) miteinander verbinden:

$$\begin{aligned} A, &(\ldots \mathfrak{a} \ldots \mathfrak{b} + 2\mathfrak{a} \ldots), \\ &(\ldots -\mathfrak{a} - \mathfrak{b} \ldots \mathfrak{b} + 2\mathfrak{a} \ldots), \\ &(\ldots -\mathfrak{a} - \mathfrak{b} \ldots \mathfrak{a} \ldots), \\ &(\ldots -\mathfrak{a} - \mathfrak{b} \ldots -\mathfrak{b} \ldots), \\ &(\ldots -\mathfrak{a} \ldots -\mathfrak{b} \ldots) = A'. \end{aligned}$$

Den Hilfssatz zeigen wir nun durch Induktion nach n. Der Fall $n = 1$ ist trivial.

Es sei $A = (\mathfrak{a}_1 \ldots \mathfrak{a}_n) \in G_n$, mit

$$\mathfrak{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}.$$

Ist $a_{nn} = 0$, so läßt sich A mit $A' = (a'_{ij}) \in G_n$ gemäß a) verbinden, wobei $a'_{nn} \neq 0$ ist. Nehmen wir also gleich $a_{nn} \neq 0$ an. Gemäß (a) verbinden wir nun A mit

$$A_1 = \begin{pmatrix} & & & * \\ & A' & & \vdots \\ & & & \vdots \\ & & & * \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix};$$

wegen $\det A_1 = a_{nn} \cdot \det A'$ ist $A' \in G_{n-1}$. Nach Induktionsannahme ist A' in G_{n-1} mit $E_\pm^{n-1}$ verbindbar, etwa durch einen Weg $A'(t)$, $0 \leq t \leq 1$. Dann ist

$$A(t) = \begin{pmatrix} & & & * \\ & A'(t) & & \vdots \\ & & & \vdots \\ & & & * \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

eine Verbindung von A_1 mit

$$A_2 = \begin{pmatrix} & & & * \\ & E_\pm^{n-1} & & \vdots \\ & & & \vdots \\ & & & * \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Der Weg $A_2(t)$:

$$A_2(t) = \begin{pmatrix} & & & t* \\ & E_\pm^{n-1} & & \vdots \\ & & & \vdots \\ & & & t* \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix},$$

verbindet dann A_2 mit

$$A_3 = \mathrm{diag}(1, \ldots, 1, \pm 1, a_{nn}).$$

Ist a_{nn} keine reelle negative Zahl, so bilden wir

$$A_3(t) = \mathrm{diag}(1, \ldots, 1, \pm 1, 1 + t(a_{nn} - 1))$$

als Verbindung in G_n zwischen A_3 und

$$A_4 = \mathrm{diag}(1, \ldots, \pm 1, 1).$$

Ist $a_{nn} < 0$, so sei

$$A_3(t) = \mathrm{diag}(1, \ldots, \pm 1, -1 + t(a_{nn} + 1))$$

eine Verbindung in G_n zwischen A_3 und

$$A_4 = \mathrm{diag}(1, \ldots, \pm 1, -1).$$

Ändert man nun A_4 gemäß (b) ab, so erhält man die Behauptung des Hilfsatzes.

Beweis von Satz 8.2. Ende. (a) $\mathbf{K} = \mathbf{R}$. A ist mit E_+^n bzw. E_-^n verbindbar, je nachdem, ob $\det A > 0$ oder < 0 ist. E_+^n und E_-^n sind nicht verbindbar, da die Determinanten unterschiedliches Vorzeichen haben.

(b) $\mathbf{K} = \mathbf{C}$. Jetzt sind E_+^n und E_-^n in G_n verbindbar:

$$E(t) = \mathrm{diag}(1, \ldots, 1, e^{it}), \quad 0 \leq t \leq \pi,$$

stellt eine Verbindung her.

3.) **Satz 8.3** *$U(n)$ und $O(n)$ sind kompakt.*

Beweis. (1) Ist $A \in U(n)$ oder $O(n)$, so sind die Matrixeinträge a_{ij} vom Betrag ≤ 1, d.h. $O(n)$ und $U(n)$ sind beschränkt.

(2) $A = (\mathfrak{a}_1 \ldots \mathfrak{a}_n) \in U(n)$ bedeutet: $\langle \mathfrak{a}_j, \mathfrak{a}_k \rangle = \delta_{jk}$. Das sind Gleichungen zwischen stetigen Funktionen; die Lösungsmenge ist also abgeschlossen.

Ohne Beweis notieren wir noch: $O(n)$ hat genau 2 Wegkomponenten, $U(n)$ ist zusammenhängend.

2 Normalformen von Endomorphismen

§0 Algebraische Grundlagen

1.) Ich erinnere zunächst an grundlegende Begriffe und Konstruktionen aus der Theorie kommutativer Ringe. R sei immer ein solcher Ring mit 1-Element.

Definition 0.1. *Ein Ideal* $\mathfrak{a} \subset R$ *ist eine nichtleere Teilmenge, für die gilt:*
(1) Mit $x, y \in \mathfrak{a}$ *ist auch* $x + y \in \mathfrak{a}$.
(2) Mit $x \in \mathfrak{a}$ *und* $r \in R$ *ist* $rx \in \mathfrak{a}$.

Definition 0.2. *(1) Ein Ideal* $\mathfrak{p}$ *heißt prim, wenn aus* $xy \in \mathfrak{p}$ *stets folgt:* x *oder* $y \in \mathfrak{p}$ *(und* $\mathfrak{p} \neq R$*).*
(2) Ein Ideal $\mathfrak{m}$ *heißt maximal, wenn* $\mathfrak{m} \neq R$ *ist, aber aus der Inklusion* $\mathfrak{m} \subset \mathfrak{a} \subset R$, $\mathfrak{a}$ *ein Ideal, folgt, daß* $\mathfrak{a} = \mathfrak{m}$ *oder* R *ist.*

Definition 0.3. $M \subset \mathfrak{a}$ *heißt Erzeugendensystem von* $\mathfrak{a}$, *wenn* $\mathfrak{a}$ *das kleinste Ideal mit* $M \subset \mathfrak{a}$ *ist. Wir schreiben* $\mathfrak{a} = (M)$. *Ist* M *einelementig,* $M = \{x\}$, *so heißt* $\mathfrak{a} = (M) = (x)$ *Hauptideal.*

Z.B. ist $R = (1)$; (0) das Nullideal. Alle Ideale in $\mathbf{Z}$, dem Ring der ganzen Zahlen, sind Hauptideale, ebenso in $K[X]$, dem Polynomring in einer Unbestimmten über dem Körper K.

Ideale treten als Kerne von Ringhomomorphismen auf. Zu jedem Ideal $\mathfrak{a}$ kann man den Faktorring $R/\mathfrak{a}$ und den Quotientenhomomorphismus

$$\pi : R \to R/\mathfrak{a}$$

bilden: $R/\mathfrak{a}$ besteht aus den Restklassen oder Kongruenzklassen

$$x + \mathfrak{a} = \{x + u : u \in \mathfrak{a}\},$$

und π ordnet jedem x seine Restklasse zu. Für $\mathfrak{a} = (0)$ ist $R/\mathfrak{a} \cong R$ (isomorph), für $\mathfrak{a} = R$ ist $R/\mathfrak{a}$ der Nullring. Ein Ideal $\mathfrak{p}$ ist genau dann prim, wenn $R/\mathfrak{p}$ nullteilerfrei ist, $\mathfrak{m}$ ist genau dann maximal, wenn $R/\mathfrak{m}$ ein Körper ist.

Definition 0.4. *Eine Einheit in R ist ein $u \in R$, zu dem es ein (multiplikativ) Inverses u^{-1} (mit $uu^{-1} = 1$) gibt.*

Die Einheiten bilden eine multiplikative Gruppe R^*. Wir nennen Elemente $x, y \in R - \{0\}$ assoziiert, wenn es eine Einheit u mit $y = ux$ gibt. Jede Einheit erzeugt das Ideal $R = (1)$, je 2 assoziierte Elemente erzeugen dasselbe Hauptideal.

2.) R sei nun nullteilerfrei.

Definition 0.5. *(1) $a|b$ „teilt", wenn $b = ra$ mit $r \in R$ ist.*

(2) p heißt irreduzibel, wenn p keine Einheit ist und wenn aus der Zelegung $p = ab$ folgt, daß a oder b Einheiten sind.

(3) Eine Nichteinheit p heißt prim, wenn aus der Teilbarkeit $p|ab$ folgt, $p|a$ oder $p|b$.

Primelemente sind irreduzibel. Ist ferner p' assoziiert zu p, so ist p' genau dann prim bzw. irreduzibel, wenn p es ist. Das von einem Primelement erzeugte Hauptideal ist ein Primideal.

Definition 0.6. *Ein Ring R heißt faktoriell, wenn er nullteilerfrei ist und jedes Element $\neq 0$ Produkt von Primelementen und Einheiten ist.*

In faktoriellen Ringen fallen die Begriffe „prim“ und „irreduzibel“ zusammen; außerdem ist jede Primzerlegung bis auf Einheiten eindeutig.

Um Beispiele zu erhalten, beweisen wir

Satz 0.1. *Jeder Hauptidealring ist faktoriell.*

Dabei benutzen wir natürlich

Definition 0.7. *Ein Hauptidealring ist ein nullteilerfreier Ring, in dem jedes Ideal Hauptideal ist.*

Beweis von Satz 0.1 (1) Zunächst zeigen wir, daß jedes irreduzible Element in R prim ist.

Es sei also q irreduzibel und $q|ab$. Nehmen wir an, daß q nicht a teilt. Das von q und a erzeugte Ideal (q,a) ist dann ein Hauptideal, etwa $= (r)$. Also ist $q = \rho \cdot r$. Wäre ρ eine Einheit, so hätte man wegen $r|a$ auch $q|a$. Da q irreduzibel ist, ist somit r eine Einheit, d.h. $(q,a) = R$, und die Gleichung

$$1 = \pi q + \alpha a$$

ist in R lösbar. Es folgt

$$b = b \cdot 1 = b(\pi q + \alpha a) = b\pi q + \alpha ab,$$

und beide Summanden rechts sind durch q teilbar. Es folgt: $q|b$ und damit ist q prim.

2) Als zweites zeigen wir, daß jedes $a \in R$ Produkt irreduzibler Elemente ist (wobei wir $a \neq 0$, a keine Einheit, annehmen). Nennen wir a schlecht, wenn es kein solches Produkt ist, und weisen wir nach, daß es keine schlechten Elemente gibt!

a sei schlecht. Insbesondere ist a nicht irreduzibel, und wir haben eine echte Zerlegung

$$a = a_0 = a_1 \cdot a_1'. \quad (a_1, a_1' \text{ keine Einheiten })$$

Hier muß mindestens ein Faktor wieder schlecht sein, etwa a_1, und es folgt

$$a_1 = a_2 \cdot a_2', \quad (a_2, a_2' \text{ keine Einheiten })$$

mit einem schlechten Faktor, etwa a_2. Sukzessive erhalten wir eine Folge

$$a = a_0,\ a_1,\ a_2,\ a_3,\ \ldots$$

von schlechten Elementen, von denen jedes ein echter Teiler des vorhergehenden ist:

$$a_j = a_{j+1}a'_{j+1}, \quad a'_{j+1} \text{ keine Einheit.}$$

Betrachten wir die hiervon erzeugten Ideale: man hat also

$$(a_0) \subsetneqq (a_1) \subsetneqq (a_2) \subsetneqq (a_3) \subsetneqq \ldots$$

Setzen wir

$$\mathfrak{a} = \bigcup_{j\geq 0}(a_j),$$

so ist, da die Idealfolge aufsteigt, $\mathfrak{a}$ wieder ein Ideal, also n.V. ein Hauptideal

$$\mathfrak{a} = (A).$$

Das Element A liegt nun in einem der (a_j), etwa in (a_k). D.h. aber: für $j \geq k$ ist einerseits

$$a_j = \alpha_j A,$$

andererseits ist

$$A = \alpha a_k,$$

$$\text{d.h.} \qquad a_j = \alpha\alpha_j a_k \quad (\text{für } j \geq k).$$

Damit wäre

$$(a_k) = (a_{k+1}) = (a_{k+2}) = \ldots,$$

im Widerspruch zur Konstruktion der a_j.

3) Wir wissen aus dem vorigen Semester, daß aus 1) und 2) die Faktorialität von R folgt.

Als Konsequenz notieren wir: die Ringe $\mathbf{Z}$, sowie $K[X]$, wobei K ein Körper ist, sind faktoriell. Weiter sei noch erwähnt, daß in einem Hauptidealring jedes von 0 verschiedene Primideal maximal ist. Ist also p eine Primzahl, so ist $\mathbf{Z}/(p)$ ein Körper: wir kennen ihn längst als Primkörper $\mathbf{F}_p$; ist $f(x)$ ein irreduzibles Polynom $\neq 0$, so ist

$$K[X]/(f(x))$$

ein Körper. Auf diese Weise konstruiert man in der Algebra zahllose neue Körper; wir kennen die Beispiele

$$\begin{aligned} \mathbf{Q}(\sqrt{2}) &\cong \mathbf{Q}[X]/(X^2-2), \\ \mathbf{C} &\cong \mathbf{R}[X]/(X^2+1), \end{aligned}$$

und viele mehr.

§1 Problemstellung und elementare Ergebnisse

1.) Es sei $f : V \to V$ ein Endomorphismus eines n-dimensionalen K-Vektorraums. Wir suchen eine Basis von V, etwa $\mathcal{A}$, so daß f durch eine „möglichst einfache" Matrix in dieser Basis beschrieben wird, im Idealfall durch eine Diagonalmatrix. Wir wissen schon, daß der Idealfall nicht immer eintritt. In der Sprechweise der Matrizen suchen wir zu $A \in K^{n\times n}$ eine ähnliche Matrix $S^{-1}AS$ von „möglichst einfacher" Gestalt. Analog können wir fragen, unter welchen Umständen zwei gegebene Matrizen oder Endomorphismen ähnlich sind.

Mit f bzw. A verbunden sind folgende Objekte.

(1)

$$\chi_f(X) = \det(X \cdot id - f),$$

$$\chi_A(X) = \det(X \cdot E - A),$$

das charakteristische Polynom von f bzw. A. Es ist ein Polynom n-ten Grades,

$$\chi_f(X) = X^n + a_1 X^{n-1} + \ldots + a_n,$$

mit $a_j \in K$, und zwar ist $a_1 = -\operatorname{Sp} f$, $a_n = (-1)^n \det f$; analog für A.

(2) Die Nullstellen λ_j von $\chi_f(X)$ bzw. $\chi_A(X)$ (in einem Erweiterungskörper von K): das sind die Eigenwerte von f bzw. A. In einer geeigneten Körpererweiterung zerfällt das charakteristische Polynom

$$\chi_f(X) = \prod_{j=i\ldots k} (X - \lambda_j)^{m_j}, \quad \lambda_j \neq \lambda_i;$$

m_j heißt algebraische Vielfachheit des EW λ_j.
(3) Zu jedem $\lambda_j \in K$ der Eigenraum

$$V(\lambda_j) = \text{Kern } (\lambda_j id - f);$$

notwendig ist

$$\dim V(\lambda_j) \leq m_j$$

(denn $f|V(\lambda_j)$ ist die „Homothetie“ $x \longmapsto \lambda_j x$; demgemäß spaltet $\chi_f(X)$ den Faktor $(X - \lambda_j)^r$ mit $r = \dim V(\lambda_j)$ ab).
f ist genau dann diagonalisierbar, wenn alle $\lambda_j \in K$ und immer

$$m_j = \dim V(\lambda_j)$$

gilt.

2.) Es sei nun $p(X) \in K(X)$ ein beliebiges Polynom,

$$p(X) = a_0 + a_1 X + \ldots + a_l X^l, \quad a_\lambda \in K.$$

Wir definieren

$$p(f) = a_0 \cdot id + a_1 f + a_2 \cdot f^2 + \ldots + a_l f^l,$$

mit

$$f^\lambda = \underbrace{f \circ f \circ \cdots \circ f}_{\lambda}, \quad f^0 = id.$$

Dann ist $p(f) : V \to V$ wieder ein Endomorphismus:

$$p(f)(x) = a_0 x + a_1 f(x) + \ldots + a_l f^l(x).$$

Entsprechend können wir eine Matrix $A \in K^{n\times n}$ in $p(X)$ einsetzen und erhalten

$$p(A) \in K^{n\times n}.$$

Wichtig ist

Satz 1.1 *Es gibt ein Polynom* $p(x) \neq 0$, *so daß* $p(f) = 0 \in \operatorname{End} V$. *Analoges gilt für Matrizen.*

Beweis. Die Endomorphismen

$$id = f^0, f, f^2, \ldots, f^l, \ldots$$

sind linear abhängig, da $\dim \operatorname{End} V = n^2$ ist. Somit existiert ein $l \leq n^2$ und $a_0, \ldots, a_l$, nicht alle 0, so daß

$$a_0 id + a_1 f + \ldots + a_e f^l = 0$$

ist. D.h. aber, für

$$p(X) = a_0 + a_1 X + \ldots + a_l X^2$$

gilt $p(f) = 0$.

In Wahrheit kann l viel kleiner als n^2 gewählt werden:

Satz 1.2 *(Satz von Cayley-Hamilton) Jeder Endomorphismus ist Nullstelle seines charakteristischen Polynoms, jede Matrix entsprechend.*

Bemerkung. E sei $A \in K^{n\times n}$. Natürlich ist

$$\det(AE - A) = 0 \in K;$$

es ist aber etwas anderes zu zeigen:

$$\chi_A(A) = 0 \in K^{n\times n}.$$

Beweis des Satzes. (1) Wir bezeichnen mit $\widetilde{\ }$ die adjungierte Matrix: ist $A = (\alpha_{ij})$, so ist $\widetilde{A} = (\widetilde{\alpha}_{ij})$ mit

$$\widetilde{\alpha}_{ij} = (-1)^{i+j} \det A^{ji};$$

A^{ji} entsteht durch Streichen der Zeile j und Spalte i aus A.

(2) $XE - A$ können wir als Matrix in $K[X]^{n\times n}$ auffassen, aber auch als Polynom von Grad 1 im Polynomring $K^{n\times n}[X]$. Beides werden wir tun, wobei wir darauf achten müssen, daß $K^{n\times n}$ nicht kommutativ ist.

$\widetilde{XE-A}$ sei nun die adjungierte Matrix zu $XE - A$ in $K[X]^{n\times n}$. Weiter ist

$$\det(XE - A) = \chi_A(X) \in K[X],$$

und wir haben die in $K[X]^{n\times n}$ gültige Relation

$$(\widetilde{XE-A})(XE - A) = X_A(X) \cdot E$$

(Cramersche Regel). Diese Relation interpretieren wir als eine Relation in $K^{n\times n}[X]$:

$$\begin{aligned} \chi_A(X) \cdot E &= c_0E + c_1EX + \ldots + E \cdot X^n \in K^{n\times n}[X] \\ XE - A &\qquad\qquad\qquad\qquad\quad\; \in K^{n\times n}[X] \\ \widetilde{XE-A} &= A_0 + A_1X + \ldots + A_mX^m \in K^{n\times n}[X], \end{aligned}$$

also lautet die obige Identiät

$$\left(\sum_{j\geq 0} A_jX^j\right) \cdot (XE - A) = \sum_{j\geq 0}(c_jE)X^j,$$

oder – da X mit A vertauscht! –

$$-A_0A + \sum_{j=0}^{n-2}(A_j - A_{j+1}A)X^{j+1} + A_{n-1}X^n = \sum_{j=0}^{n}(c_jE)X^j.$$

Damit haben wir die Koeffizienten c_jE des charakteristischen Polynoms neu berechnet: es ist

$$\chi_A(X) \cdot E = -A_0A + (A_0 - A_1A)X + \ldots + (A_{n-2} - A_{n-1}A)X^{n-1} + A_{n-1}X^n.$$

Setzt man $X = A$, so folgt unmittelbar $\chi_A(A) = 0$.

Bemerkung: Aus der Identität

$$\left(\sum_{j=0}^{n-1} A_jX^j\right)(XE - A) = \chi_A(X) \cdot E$$

folgt *nicht*, daß für beliebige Matrizen $B \in K^{n\times n}$

$$\left(\sum_{j=0}^{n-1} A_j B^j\right)(BE - A) = \chi_A(B) \cdot E$$

ist – richtig ist das nur für mit A vertauschbare Matrizen.

Die Menge $\mathrm{Ann}(f)$ der Polynome p mit $p(f) = 0$ ist ein Ideal, welches nach Satz 1 nicht das Nullideal ist und nach Satz 2 das charakteristische Polynom enthält. Da es ein Hauptideal ist, gibt es ein – bis auf Assoziiertheit – eindeutig bestimmtes Erzeugendes $P_f(X)$ in $\mathrm{Ann}\, f$,

$$\mathrm{Ann}\, f = (P_f(X));$$

es ist ein Polynom minimalen Grades in diesem Ideal.

Definition 1.1. *Das Minimalpolynom P_f von f ist das normierte Polynom minimalen Grades mit $P_f(f) = 0$.*

„normiert“ soll heißen: der Leitkoeffizient ist 1.

Nach Satz 1.2 gilt dann $P_f \mid \chi_f$, insbesondere ist $\deg P_f \leq n$.

Alle diese Begriffe sind natürlich genauso auf Matrizen $A \in K^{n\times n}$ anwendbar. Als einfache Beispiele betrachten wir

(1)
$$A = \mathrm{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{l_1}, \underbrace{\lambda_2, \ldots, \lambda_2}_{l_2}, \ldots, \underbrace{\lambda_k, \ldots, \lambda_k}_{l_k}).$$

Man prüft sofort

$$\chi_A(X) = \prod_{i=1\ldots k} (X - \lambda_i)^{l_i}$$
$$P_A(X) = \prod_{j=1\ldots k} (X - \lambda_i)$$

nach.

(2) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Hier ist

$$\chi_A(X) = P_A(X)) = (X-1)^2$$

Wir können nun eine weitere Ähnlichkeitsinvariante angeben:

Satz 1.3 *Ähnliche Matrizen (oder Endomorphismen) haben dasselbe Minimalpolynom.*

Beweis. Es sei $B = S^{-1}AS$, es sei $p(X) = \sum a_j X^j$ ein Polynom mit $p(A) = 0$. Dann ist

$$\begin{aligned} p(B) &= \sum_j a_j \left(S^{-1}AS\right)^j \\ &= \sum_j a_j S^{-1}A^j S \\ &= S^{-1}\left(\sum a_j A^j S\right) \\ &= S^{-1}\left(\sum a_j A^j\right) S \\ &= S^{-1}p(A)S = 0. \end{aligned}$$

3.) Ich beschreibe nun die Hauptidee der folgenden Überlegungen.

$$f : V \to V$$

sei ein Endomorphismus des K-Vektorraumes V. Ist $p = p(X) \in K[X]$ ein Polynom, so definieren wir für $x \in V$

$$p \cdot x \overset{\text{def}}{=} p(f)(x) = a_0 x + a_1 f(x) + \ldots + a_n f^n(x),$$

wobei $p = p(X) = a_0 + a_1 X + \ldots + a_n X^n$ ist. Das ist eine Verknüpfung zwischen Polynomen und Vektoren, d.h. eine Abbildung

$$\begin{gathered} K[X] \times V \to V \\ (p(X), x) \longmapsto p(X) \cdot x. \end{gathered}$$

Es ist leicht zu sehen, daß diese Verknüpfung dieselben formalen Eigenschaften haben wie die Multiplikation von Körperelementen mit Vektoren,

$$K \times V \to V$$
$$(\lambda, x) \longmapsto \lambda x.$$

Die Verknüpfung ist von f abhängig und bringt daher, wie wir sehen werden, wesentliche Eigenschaften von f zum Vorschein. Neu ist nun, daß der „Operatorbereich“, den wir für V einführen, nicht mehr ein Körper ist, sondern lediglich ein Ring, nämlich der Polynomring $K[X]$. Das ist der Grund für unser vorheriges Studium von Ringen und Teilbarkeit. Im nächsten Paragraphen systematisieren wir diese Ideen.

§2 Moduln

1.) R sei ein kommutativer Ring mit 1-Element.

Definition 2.1. *Ein R-Modul M (“Modul über R“) ist eine nichtleere Menge M zusammen mit zwei Verknüpfungen*

$$\begin{aligned} M \times M &\xrightarrow{+} M \\ a, b &\longmapsto a + b, \\ R \times M &\xrightarrow{\cdot} M, \\ \lambda, a &\to \lambda a, \end{aligned}$$

für die gilt:

(1) $M, +$ ist eine abelsche Gruppe

(2)

$$\begin{aligned} \lambda(x + y) &= \lambda x + \lambda y \\ (\lambda + \mu)x &= \lambda x + \mu x \\ (\lambda\mu)x &= \lambda(\mu x) \\ 1 \cdot x &= x \end{aligned}$$

für $\lambda, \mu, 1 \in R$, $x, y \in M$.

Beispiele
1. Vektorräume sind Moduln über einem Körper.
2. Jede abelsche Gruppe "ist" ein **Z**-Modul mit der Verknüpfung

$$nx = \underbrace{x + \cdots + x}_{n-\text{mal}} \quad , \quad n \geq 1,$$

$$(-n)x = n \cdot (-x)$$
$$0 \cdot x = 0$$

3. Die Ideale von R sind die R-Untermoduln. Dabei verwenden wir natürlich

Definition 2.2. *$N \subset M$ heißt Untermodul des R-Moduls M, wenn N mit den induzierten Verknüpfungen $+, \cdot$ ein R-Modul ist.*

4. Es sei V ein K-Vektorraum, $f : V \to V$ ein Endomorphismus. Für $p(X) \in K[X]$ definieren wir wie früher

$$p(X) \cdot x = p(f)(x), \quad x \in V.$$

Dann ist V ein $K[X]$-Moldul, wie man sofort verifiziert.

Wir entwickeln nun die Theorie weiter.

Definition 2.3. *(1) $\varphi : M \to N$ ist ein Homomorphismus von R-Moduln, wenn gilt:*

$$\varphi(x+y) = \varphi(x) + \varphi(y) \quad , \quad \varphi(\lambda x) = \lambda\varphi(x),$$

für $\lambda \in R$, $x, y \in M$.
(2) Ein Isomorphismus ist ein bijektiver Homomorphismus.

$\operatorname{Kern}\varphi = \{x \in M : \varphi(x) = 0\}$ sowie $\operatorname{Bild}\varphi = \{\varphi(x) : x \in M\}$ sind dann Untermoduln von M bzw. N.

Satz 2.1 *Es sei $N \subset M$ ein Untermodul. Auf der Faktorgruppe M/N wird durch*

$$\lambda \cdot [x] = [\lambda x],$$

wobei $[x]$ die Kongruenzklasse von $x \mod N$ bedeutet, eine R-Modul- Struktur erklärt.

$$\begin{aligned} \pi : M &\longrightarrow M/N \\ x &\longmapsto [x] \end{aligned}$$

ist ein Modulhomomorphismus für diese Struktur. Ist

$$\varphi : M \to P$$

irgendein Modulhomomorphimus und $N = \operatorname{Kern} \varphi$, so gibt es genau einen Homomorphismus $\overline{\varphi}$, der das Diagramm

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[d] & P \\ M/N \arrow[ur, "\overline{\varphi}"'] & \end{tikzcd}$$

kommutativ macht; $\overline{\varphi}$ ist injektiv und bildet M/N isomorph auf $\operatorname{Bild} \varphi$ ab.

Der Beweis besteht in den üblichen Verifikationen.

2.)

Definition 2.4. *N sei eine Teilmenge eines R-Moduls M. Der von N erzeugte Untermodul von M ist*

$$L(N) = \bigcap \{M' : M' \supset N, M' \text{ Untermodul}\}.$$

Wie im Falle von Vektorräumen zeigt man

$$L(N) = \left\{ \sum_j \lambda_j x_j \;:\; x_j \in N, \lambda_j \in R \right\},$$

d.h. $L(N)$ ist die Menge aller endlichen Linearkombinationen von Elementen von N (mit Koeffizienten in R). Ist N endlich, so schreibt man auch, für $N = \{x_1, \ldots, x_k\}$:

$$L(N) = Rx_1 + \ldots + Rx_k = \sum Rx_i.$$

Definition 2.5. *Ein Erzeugendensystem für M ist eine Teilmenge N von M mit*

$$M = L(N).$$

Besitzt M ein endliches Erzeugendensystem, so heißt M endlich erzeugt.

Z.B. ist $R = R \cdot 1$ endlich erzeugt. In einem Hauptidealring ist jedes Ideal endlich erzeugt (nämlich von einem Element).

Definition 2.6. *Ein Untermodul, der von einem Element erzeugt wird, heißt zyklisch:*

$$N = Rx \subset M, \quad x \in M.$$

Definition 2.7. *$M_1, \ldots, M_r$ seien Untermoduln von M. Dann ist ihre Summe*

$$\sum_{j=1\ldots r} M_j = M_1 + \cdots + M_r = L(M_1 \cup \ldots \cup M_r).$$

Man sieht wie bei Vektorräumen, daß die Summe der M_j aus allen x der Gestalt

$$x = x_1 + \ldots + x_r, \quad x_j \in M_j,$$

besteht.

Ebenso leicht zeigt man

Satz 2.2 *Es sei $M' = \sum_{j=1}^{r} M_j \subset M$. Folgende Aussagen sind äquivalent:*

(1) Jedes $x \in M'$ läßt sich auf genau eine Weise als

$$x = \sum_{j=1}^{r} x_j, \quad x_j \in M_j,$$

darstellen.
(2) $M_i \cap \sum_{j \neq i} M_j = 0$ für alle i.

Definition 2.8. *Gilt für $M' = \sum M_j$ die Eigenschaft (1) oder (2) aus Satz 2.2, so heißt M' die direkte Summe der Untermoduln M_j,*

$$M' = \bigoplus_{j=1\ldots r} M_j = M_1 \oplus \ldots \oplus M_r;$$

die Darstellung $M' = M_1 \oplus \ldots \oplus M_r$ ist eine direkte Zerlegung von M'.

Für endlich viele Moduln läßt sich auch die äußere oder konstruierte direkte Summe einführen: $M_1, \ldots, M_r$ seien R-Moduln. Man erklärt

$$M = \{(x_1, \ldots, x_r) : x_j \in M_j\}$$

mit komponentenweiser Addition und Multiplikation mit Ringelementen. Dann heißt M die direkte Summe der M_j :

$$M = M_1 \oplus \cdots \oplus M_r.$$

Jedes M_j ist isomorph zum Untermodul

$$M'_j = \{0, \cdots, 0, x, 0, \ldots, 0) : x \in M_j\},$$

und $M = M'_1 \oplus \ldots \oplus M'_r$ gemäß Def. 2.8. Insbesondere ist

$$R^k = \underbrace{R \oplus \ldots \oplus R}_{k\,\text{mal}}$$

ein R-Modul.

Definition 2.9. *M heißt frei vom Rang k, wenn M isomorph zu R^k ist.*

3.)
Satz 2.3 *M sei ein R-Modul. Die Menge*

$$\operatorname{Ann} M = \{\lambda \in R : \lambda x = 0 \textit{ für alle } x \in M\}$$

ist ein Ideal in R.

Das verifiziert man sofort.

Definition 2.10. *(1)* $\operatorname{Ann} M$ *heißt Annullatorideal von* M*.*

(2) $\operatorname{Ann} x = \operatorname{Ann} Rx$ *heißt Annullatorideal von* x*.*

(3) $x \in M$ *ist ein Torsionselement, wenn* $\operatorname{Ann} x \neq (0)$*, wenn es also ein* $\lambda \neq 0$ *in* R *mit* $\lambda x = 0$ *gibt.*

(4) M *heißt Torsionsmodul, wenn* M *nur aus Torsionselementen besteht.*

(5) M *heißt torsionsfrei, wenn* $0 \in M$ *das einzige Torsionselement ist.*

Schauen wir uns kurz einige einfache Beziehungen zwischen diesen Begriffen an:

(1) Endlichdimensionale Vektorräume über K sind freie K-Moduln vom Rang gleich der Dimension des Raumes; sie sind torsionsfrei.

(2) Ist R nullteilerfrei, so ist jeder freie Modul auch torsionsfrei.

(3) Eine endliche abelsche Gruppe ist ein Torsions-$\mathbf{Z}$-Modul.

(4) Die Torsionselemente in R sind gerade die Nullteiler.

(5) $\mathbf{Q}, +$ ist ein torsionsfreier nicht freier $\mathbf{Z}$-Modul (d.h. kein freier Modul vom Rang $h < \infty$. Definiert man auch freie Moduln unendlichen Ranges in naheliegender Weise, so erkennt man auch, daß $\mathbf{Q}$ überhaupt kein freier $\mathbf{Z}$-Modul ist).

4.) Nun sei $f : V \to V$ ein Endomorphismus eines n-dimensionalen K-Vektorraums. Wir fassen V als $K[X]$-Modul auf - siehe §1:

$$p(X) \cdot x = p(f)(x).$$

Außerdem ist V natürlich ein freier K-Modul vom Rang n: wir vergleichen die beiden Strukturen.

(1) Ist $a_1, \ldots, a_n$ eine Basis von V, so ist

$$V = K[X]a_1 + \ldots + K[X]a_n,$$

also ein endlich erzeugter $K[X]$-Modul.

(2) Ist $P_f(X)$ das Minimalpolynom von f - es ist $\neq 0$ -, so ist

$$\operatorname{Ann} V = K[X]P_f(X) = (P_f(X));$$

insbesondere ist V ein Torsions-$K[X]$-Modul.

Satz 2.4 *Für* $W \subset V$ *sind äquivalent:*

i) W ist ein Unter-$K[X]$-Modul.
ii) W ist ein f-invarianter Untervektorraum - i.e. $f(W) \subset W$.

Satz 2.5 *Für V ist äquivalent*

i) $V = U \oplus W$ als direkte Summe von $K[X]$-Moduln.
ii) $V = U \oplus W$ als direkte Summe f-invarianter K-Unterräume.

Beide Sätze verifiziert man unmittelbar. Wir wollen nun die Situation von Satz 2.5 genauer ansehen.

Es sei also $V = U \oplus W$ die direkte Summe f-invarianter Unterräume. Wir wählen Basen $a_1, \ldots, a_k$ in U, $a_{k+1}, \ldots, a_n$ in W, dann ist also $a_1, \ldots, a_n$ eine Basis von V. In dieser Basis gilt

$$f(a_i) = \sum_{j=1\ldots n} \alpha_{ji} a_j,$$

und wegen der f-Invarianz von U und V ist

$$\begin{aligned} \alpha_{ji} = 0 \text{ für } i = 1, \ldots, k \text{ und } j > k, \\ \text{für } i = k+1, \ldots, n \text{ und } j \leq k, \end{aligned}$$

d.h. die Matrix A von f hat bez. dieser Basis die Gestalt

$$A = \begin{pmatrix} * & * & & 0 \\ * & * & & \\ & & * & * \\ 0 & & * & * \end{pmatrix} = \operatorname{diag}(B, C),$$

mit $B \in K^{k \times k}, C \in K^{(n-k)\times(n-k)}$. — Hat umgekehrt A bez. einer Basis $a_1, \ldots, a_k, a_{k+1}, \ldots, a_n$ diese Gestalt, so ist, mit

$$U = L(a_1, \ldots, a_k); \quad W = L(a_{k+1}, \ldots, a_n),$$

$$V = U \oplus W$$

als direkte Summe von $K[X]$-Moduln.

§3 Torsionsmoduln über Hauptidealringen

Hauptergebnis dieses Kapitels ist der in diesem Paragraphen bewiesene “Struktursatz“ für endlich erzeugte Torsionsmoduln über Hauptidealringen. Wir benötigen alle bisherigen Informationen aus der Teilbarkeitslehre.

1.) Es sei $\mathfrak{a}$ ein Ideal im Ring R. Dann ist der Faktorring $R/\mathfrak{a}$ ein R-Modul, der von genau einem Element, nämlich der Restklasse $1 + \mathfrak{a}$, erzeugt wird. Ist umgekehrt $M = Ra$, d.h. ein von einem Element erzeugter R-Modul, so ist

$$M \cong R/\mathfrak{a},$$

mit $\mathfrak{a} = \operatorname{Ann} a$.

Definition 3.1. *Der Modul M heißt zyklisch, wenn es ein $a \in M$ mit $M \cong Ra$ bzw. ein Ideal $\mathfrak{a} \subset R$ mit $M \cong R/\mathfrak{a}$ gibt.*

Z.B. sind die $\mathbf{Z}_m = \mathbf{Z}/(m)$ zyklische $\mathbf{Z}$-Moduln, die uns ja schon gut bekannten endlichen zyklischen Gruppen.

Falls $\mathfrak{a} \neq (0)$ gilt, ist $M = R/\mathfrak{a}$ natürlich immer ein Torionsmodul, ist $\mathfrak{a} = (0)$, so ist $M = R/(0) \cong R$ frei und kein Torsionsmodul ($1 \in R$ ist kein Torsionselement); für nullteilerfreies R ist $R \cong M$ sogar torsionsfrei.

Die zyklischen Moduln $R/\mathfrak{a}$ sind die einfachsten Torsionsmoduln. Unser Ziel ist, beliebige (endlich erzeugte) Torsionsmoduln aus ihnen aufzubauen.

2.) Ab sofort sei R ein Hauptidealring und M ein endlich erzeugter Torsionsmodul über R, also

$$M = Rx_1 + \cdots + Rx_k.$$

Hilfsatz 1. *Es gibt $a \neq 0$ in R mit $aM = 0$.*

Beweis. Zu x_i existiert $a_i \neq 0$ in R mit $a_i x_i = 0$. Dann ist für $a = a_1 \cdot a_2 \ldots a_k$ sicher $aM = 0$, und es ist $a \neq 0$ wegen der Nullteilerfreiheit von R.

Wir definieren nun für $a \in R$:

$$M(a) = \{x \in M : ax = 0\}$$

und bemerken sofort, daß $M(a)$ ein Untermodul von M ist. Ist ferner e eine Einheit, so ist $M(a) = M(ea)$.

Definition 3.2. *Es sei $p \in R$ ein Primelement. Der Modul*

$$M_p = \bigcup_{n \in \mathbf{N}} M(p^n) = \{x \in M : \exists\, n \in \mathbf{N} \text{ mit } p^n x = 0\}$$

heißt p-primäre Komponente von M. Ist $M = M_p$, so heißt M p-primär.

Fundamental ist nun

Hilfssatz 2. *a und b seien teilerfremde Ringelemente. Dann ist*

$$M(ab) = M(a) \oplus M(b).$$

Beweis. Nach Voraussetzung ist das Ideal (a, b) der ganze Ring, d.h. es gibt $\alpha, \beta \in R$ mit

$$\alpha a + \beta b = 1.$$

Nun sei $x \in M(ab)$, dann gilt:

$$x = 1 \cdot x = (\alpha a + \beta b)x = \alpha a x + \beta b x = x_1 + x_2,$$

und es ist

$$\begin{aligned} bx_1 &= \alpha abx = 0, \\ ax_2 &= \beta bax = 0, \end{aligned}$$

somit ist $M(a, b) = M(a) + M(b)$.
Um zu zeigen, daß die Summe direkt ist, betrachten wir $x \in M(a) \cap M(b)$):

$$x = 1 \cdot x = \alpha a x + \beta b x = 0 + 0 = 0.$$

Satz 3.1 (1. Zerlegungssatz). *Es sei M ein endlich erzeugter Torsionsmodul über dem Hauptidealring R. Dann gilt*

(1) Es gibt nur endlich viele Primelemente $p \in R$ mit $M_p \neq 0$ (bis auf Assozierte genau).

(2) Ist $a \in \operatorname{Ann} M$ und

$$a = ep_1^{r_1} \dots p_k^{r_k}$$

die Primzerlegung von a, so ist $M_p \neq 0$ höchstens für $p = p_i$, und man hat

$$M = \bigoplus_{i=1\dots k} M_{p_i}.$$

(3) In der Situation von (2) ist

$$M_{p_i} = M(p_i^{r_i}), \quad i = 1\dots k$$

(4) Ist insbesondere a ein erzeugendes Element des Annullatorideals von M,

$$a = ep_1^{r_1} \dots p_m^{r_m}$$

die Primzerlegung, so sind alle $M_{p_i} \neq 0$, und

$$M = M_{p_1} \oplus \dots \oplus M_{p_m}.$$

(5) Ist

$$M = M_{q_1} \oplus \dots \oplus M_{q_l}$$

eine direkte Zerlegung von M in q_i-primäre Untermoduln $\neq 0$ und sind die q_i paarweise nicht-assoziiert, so ist $l = m$ und (nach Umnumerierung) $p_i = q_i$. - Dabei sind die p_i die Primfaktoren aus (4).

Beweis. (a) Nach Hilfssatz 1 ist $\operatorname{Ann} M \neq 0$, also folgt (1) aus (2).

(b) Wir zeigen nun die Aussagen (2) und (3). Es sei also

$$a = ep_1^{r_1} \dots p_k^{r_k} \in \operatorname{Ann} M,$$

e Einheit, p_i paarweise nichtassoziierte Primelemente in R. Nach Voraussetzung ist $M = M(a)$, nach Hilfssatz 2 also

$$\begin{aligned} M = M(a) &= M(p_1^{r_1}) \oplus M(p_2^{r_2} \dots p_k^{r_k}) \\ &= \dots \\ &= \bigoplus_{i=1\dots k} M(p_i^{r_i}). \end{aligned}$$

Nun gilt natürlich $M_{p_i} \supset M(p_i^{r_i})$. Ist andererseits $x \in M_{p_i}$, so besitzt x die Summenzerlegung

$$x = x_1 + \ldots + x_k$$

mit $x_j \in M(p_j^{r_j})$. Dann ist weiter für l groß genug

$$0 = p_i^l x = p_i^l x_1 + \ldots + p_i^l x_k.$$

Damit folgt

$$p_i^l x_j = 0 \quad \text{für alle } j,$$

d.h.

$$x_j \in M(p_i^l) \cap M(p_j^{r_j}).$$

Für $j \neq i$ ist damit $x_j = 0$, und wir sehen $x = x_i \in M(p_i^{r_i})$.

Damit ist (3) bewiesen. Ist schließlich p irgendein Primelement und $x \in M_p$, so gilt $p^l x = 0$ für passendes l. Wir zerlegen wieder

$$x = x_1 + \ldots + x_k$$

mit $x_i \in M(p_i^{r_i})$ und erhalten wie oben

$$x_i \in M_p \cap M(p_i^{r_i})$$

und damit (nach Hilfssatz 2) $x_i = 0$ für $p \neq p_i$, also entweder $M_p = 0$, falls nämlich $p \neq p_i$, $i = 1 \ldots k$, oder $p = p_i$ für genau ein i.

c) Es sei nun $(a) = \operatorname{Ann} M$ und

$$a = ep_1^{r_1} \ldots p_m^{r_m}$$

die Primzerlegung von a. Nehmen wir $M_{p_m} = 0$ an und setzen

$$b = ep_1^{r_1} \ldots p_{m-1}^{r_{m-1}}.$$

Dann haben wir $b \mid a$ und, wegen $M_{p_m} = 0$, auch $b \in \operatorname{Ann} M$. Aber $a \nmid b$, im Widerspruch zur Wahl von a. Damit ist (4) bewiesen.

d) Wir zeigen: es seien

$$M = M'_{p_1} \oplus \ldots \oplus M'_{p_n} = M''_{q_1} \oplus \ldots \oplus M''_{q_t}$$

zwei Zerlegungen von M in p_i bzw. q_j-primäre Untermoduln $\neq 0$; dann ist $n \leq t$ und $p_1 = q_1, \ldots, p_n = q_n$, (nach geeigneter Numerierung) und $M'_{p_i} \subset M''_{q_i}$, $i = 1 \ldots n$. Daraus folgt die letzte Behauptung des Satzes.

Wir führen Induktion nach n durch.
Es sei $n = 1$. Dann ist $M = M'_{p_1} \neq 0$, also gibt es ein $M''_{q_j} \neq 0$, ohne Einschränkung etwa M''_{q_1}, d.h. $1 \leq t$. Nun sei $x \in M = M'_{p_1}$. Gemäß der zweiten Summenzerlegung gilt

$$x = x_1 + x_2 + \cdots + x_t, \quad x_j \in M''_{q_j}.$$

Es ist für passendes l sicher $p_1^l x = 0$. Daraus folgt wie früher

$$p_1^l x_j = 0,$$

also

$$x_j \in M(p_1^l) \cap M''_{q_i},$$

und damit

$$x_j = 0 \quad \text{oder} \quad p_1 = q_i.$$

Wir dürfen dann $p_1 = q_1$ annehmen und haben

$$M'_{p_1} \subset M''_{q_1}.$$

(sogar Gleichheit).
Nun sei die Aussage schon für $n-1$ bewiesen, und $x \in M'_{p_n}$. Wir zerlegen wieder

$$x = x_1 + \ldots + x_{n-1} + x_n + \cdots + x_t$$

mit $x_j \in M''_{q_j}$, wobei wir $p_j = q_j$ für $j < n$ und $M'_{p_j} \subset M^n_{q_j}$, $j < n$, annehmen. Multiplikation mit p_n^l für großes l liefert

$$p_n^l x = 0,$$

also $p_n^l x_j = 0$ für alle j, also $x_j = 0$ für $j \neq n$ (bei geeigneter Numerierung!) und $p_n = q_n$, ferner $M'_{p_n} \subset M''_{q_n}$.

3.) Wir beachten nun den unmittelbar klaren

Hilfssatz 3. *Jeder direkte Summand eines endlich erzeugten Moduls ist wieder endlich erzeugt.*

Damit bleibt die Struktur endlich erzeugter p-primärer Moduln zu untersuchen. Hier gilt

Satz 3.2 *(2. Zerlegungssatz). Es sei M endlich erzeugt p- primär über dem Hauptidealring R. Dann existiert eine Folge natürlicher Zahlen*

$$n_1 \leq n_2 \leq \cdots \leq n_r$$

mit

$$M \cong \bigoplus M_i, \quad M_i \cong R/(p^{n_i}).$$

Die n_i sind durch M eindeutig bestimmt, insbesondere ist jedes M_i direkt unzerlegbar, und die M_i sind bis auf Isomorphie eindeutig durch M bestimmt.

Beweis. Es ist $M = \sum_{i=1\dots l} Rx_i$. Für die Existenzaussage führen wir Induktion nach l durch.

(i) $l = 1$. Jetzt ist $M = Rx \cong R/\operatorname{Ann} x = R/(p^n)$.

(ii) Induktionsschluß von $l-1$ auf l.

Wir wissen, daß

$$\operatorname{Ann} x_i = (p^{r_i})$$

gilt. O.E. sei $r_l = n$ das Maximum der r_i, damit

$$M_l = Rx_l \cong R/(p^n).$$

Gehen wir zum Faktormodul $\overline{M} = M/M_l$ über. Bezeichnet $\overline{x}$ immer die Restklasse von x bezüglich M_l, so hat man

$$\overline{M} = R\overline{x}_1 + \cdots + R\overline{x}_{l-1};$$

da $\overline{M}$ wieder p-primär ist, liefert die Induktionsannahme Elemente $y_1, \dots, y_k \in M$ mit

$$\overline{M} = R\overline{y}_1 \oplus \cdots \oplus R\overline{y}_k,$$

und

$$R\bar{y}_i \cong R/(p^{n_i});$$

offensichtlich ist $n_i \le n$, da M und damit $\overline{M}$ durch p^n annulliert wird.

Die y_i sind durch $R\bar{y}_i$ nicht eindeutig bestimmt: wir können sie immer noch durch Vielfache von x_l abändern. Wir wollen das so tun, daß für die abgeänderten y_i, wir nennen sie dann z_i, die direkte Zerlegung

$$M = \bigoplus_{i=1\ldots k} Rz_i \oplus M_l$$

besteht. Dazu beachten wir:

$$p^{n_i}\bar{y}_i = 0,$$

d.h.

$$p^{n_i}y_i = a_i x_l = h_i p^{m_i} x_l,$$

wobei p^{m_i} die höchste p-Potenz in a_i ist, also

$$(h_i, p) = 1.$$

Damit ist

$$0 = p^n y_i = p^{n-n_i} p^{n_i} y_i = p^{n-n_i+m_i} h_i x_l.$$

Das impliziert

$$p^{n-n_i+m_i} x_l \in M(h_i),$$

aber $M(h_i) = 0$ nach Hilfssatz 2, und damit muß

$$n - n_i + m_i \ge n$$

sein, also $n_i \le m_i$. Jetzt setzen wir

$$z_i = y_i - h_i p^{m_i - n_i} x_l$$

und zeigen:

(iii) $M = \bigoplus Rz_i \oplus M_l$.

(a) Zunächst gilt $\overline{y}_i = \overline{z}_i$, ferner

$$p^{n_i} z_i = p^{n_i}(y_i - h_i p^{m_i - n_i} x_l) = p^{n_i} y_i - h_i p^{m_i} x_l = 0$$

also $Rz_i \cong R\overline{z}_i \cong R/(p^{n_i})$.

(b) Ist $x \in M$, so ist

$$\overline{x} = a_1 \overline{z}_1 + \cdots + a_k \overline{z}_k,$$

also

$$x = a_1 z_1 + \ldots + a_k z_k + v, \quad v \in M_l.$$

Damit ist $M = Rz_1 + \ldots + Rz_k + M_l$.

(c) Die Direktheit der Zerlegung ergibt sich so: es sei

$$u_1 + \ldots + u_k + v = 0, \quad u_i \in Rz_i, v \in M_l.$$

Dann ist erst recht

$$\overline{u}_1 + \cdots + \overline{u}_k = \overline{0},$$

also $\overline{u}_i = 0$. Ist also $n_i = a_i z_i$, so hat man $\overline{u}_i = a_i \overline{z}_i = 0$, und das bedeutet

$$p^{n_i} \mid a_i.$$

Damit ist aber $a_i z_i = 0$ (wegen Teil a), also alle $u_i = 0$, also $v = 0$.

(iv) Es bleibt die Eindeutigkeit der Zerlegung zu beweisen.

$$M = M_1 \oplus \ldots \oplus M_r = N_1 \oplus \ldots \oplus N_s$$

seien zwei Zerlegungen der obigen Art, mit

$$M_\rho \cong R/(p^{l_\rho}), \quad N_\sigma \cong R/(p^{m_\sigma}).$$

Wir dürfen $l_1 \leq l_2 \leq \ldots \leq l_r$ und $m_1 \leq \ldots \leq m_s$ annehmen. Zu zeigen ist:

$$r = s \quad \text{und} \quad l_\rho = m_\rho \quad \text{für} \quad \rho = 1, \ldots, r.$$

Für $r = 0$ ist $M = 0$; also auch $s = 0$. Die Aussage sei nun für $r - 1 \geq 0$ schon bewiesen, wir folgern sie für r.

Offenbar ist Ann $M = (p^{l_r})$, ebenso $= (p^{m_s})$, damit folgt $m_s = l_r$. Es sei nun

$$N_s = R_\rho, \quad M_\rho = Rx_\rho.$$

Wir wählen t so, daß für $t \leq \rho \leq r$ immer $l_\rho = l_r$ gilt, für $\rho < t$ aber $l_\rho < l_r$ ist, und betrachten die Darstellung von y als Summe von Elementen in M_ρ:

$$y = a_1x_1 + \ldots + a_tx_t + \ldots + a_rx_r$$

mit $a_\rho \in R$. Wäre p ein Teiler der a_ρ für alle ρ mit $t \leq \rho \leq r$, so hätte man

$$p^{l_r-1}y = 0,$$

im Widerspruch zur Annahme

$$\operatorname{Ann} y = (p^{l_r}).$$

O.E. sei nun $p \nmid a_r$. Damit ist auch

$$(p^{l_r}, a_r) = 1,$$

und es gibt $\alpha, \beta \in R$ mit

$$\alpha a_r + \beta p^{l_r} = 1.$$

Das impliziert

$$\alpha a_r x_r = (1 - \beta p^{l_r})x_r = x_r.$$

Ich zeige nun, daß

$$M = M_1 \oplus \cdots \oplus M_{r-1} \oplus N_s$$

gilt.

(a) Es ist

$$a_r x_r = y - a_1x_1 - \cdots - a_{r-1}x_{r-1};$$

Mulitplikation mit α liefert

$$\alpha a_r x_r = x_r = \alpha y - \alpha a_1 x_1 - \cdots - \alpha a_{r-1}x_{r-1}.$$

Also liegt x_r in $M_1 + \cdots + M_{r-1} + N_s$, und wir sehen

$$M = M_1 + \cdots + M_{r-1} + N_s.$$

(b) Um die Direktheit der Zerlegung zu zeigen, nehmen wir eine Relation

$$b_1x_1 + \cdots + b_{r-1}x_{r-1} + by = 0$$

mit $b_\rho, b \in R$ an. Einsetzen der früheren Darstellung von y durch die x_ρ ergibt

$$b_1x_1 + \cdots + b_{r-1}x_{r-1} + b(a_1x_1 + \cdots + a_rx_r) = 0$$

$$(b_1 + ba_1)x_1 + \cdots + (b_{r-1} + ba_{r-1})x_{r-1} + ba_rx_r = 0.$$

Hieraus folgt zunächst

$$ba_rx_r = 0, \quad \alpha ba_rx_r = 0,$$

also wegen $\alpha a_rx_r = x_r$ auch $bx_r = 0$ und damit die Teilbarkeit

$$p^{l_r} \mid b.$$

Damit ist aber

$$ba_\rho x = 0 \quad \text{für alle } \rho,$$

$$b_1x_1 + \cdots + b_{r-1}x_{r-1} = 0,$$

somit $b_\rho x_\rho = 0$: die Zerlegung ist direkt!

Jetzt können wir den Induktionsbeweis zu Ende führen. Wir setzen

$$M = M_1 \oplus \cdots \oplus M_{r-1} \oplus N_s = N_1 \oplus \cdots \oplus N_{s-1} \oplus N_s.$$

Daraus folgt

$$M/N \cong M_1 \oplus \cdots \oplus M_{r-1} \cong N_1 \oplus \cdots \oplus N_{s-1}.$$

Es gibt also Untermoduln $N'_\sigma \cong N_\sigma$ von $M_1 \oplus \cdots \oplus M_{r-1}$, so, daß

$$M_1 \oplus \cdots \oplus M_{r-1} = N'_1 \oplus \cdots \oplus N'_{s-1}$$

ist. Nach Induktionsannahme ist dann $r - 1 = s - 1$ und $l_\rho = m_\rho$, $\rho = 1, \ldots, r - 1$.

4.) Wir fassen die beiden Ergebnisse zusammen in

Satz 3.3 *(Struktursatz für endlich erzeugte Torsionsmoduln über Hauptidealringen) Es sei M ein endlich erzeugter Torsionsmodul über dem Hauptidealring R. Dann gibt es eindeutig bestimmte Primelemente $p_1, \ldots, p_k$ in R und p_i-primäre Untermoduln M_i von M mit*

$$M = \bigoplus_{i=1\ldots k} M_i.$$

Ferner gibt es eindeutig bestimmte natürliche Zahlen n_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, s_i$ mit

$$1 \leq n_{i1} \leq \cdots \leq n_{is_i} \stackrel{\text{def}}{=} r_i,$$

so daß

$$M_i \cong \bigoplus_{j=1\ldots s_i} R/(p_i^{n_{ij}}).$$

Es ist $\operatorname{Ann} M = (\prod p_i^{r_i})$.

Insgesamt ist also

$$M \cong \bigoplus_{i,j} R/(p_i^{n_{ij}}),$$

d.h. jeder endlich erzeugte Torsionsmodul über einem Hauptidealring ist eine direkte Summe p-primärer zyklischer Untermoduln; die Summanden sind bis auf Isomorphie eindeutig bestimmt. Insbesondere ist ein Modul $R/(p^m)$ direkt unzerlegbar.

5.) Wir notieren zwei wichtige Spezialfälle; der für die lineare Algebra wichtigste Fall wird im nächsten Paragraphen besprochen.

Für $R = \mathbf{Z}$, Ring der ganzen Zahlen, ist ein endlich erzeugter Torsionsmodul eine endliche abelsche Gruppe und die zyklischen Moduln

$$R/(p^m)$$

sind zyklische Gruppen der Primpotenzordnung p^m. Wir haben damit

Satz 3.4 *(Satz von Gauß) Jede endliche abelsche Gruppe ist direkte Summe zyklischer Gruppen von Primpotenzordnung. Die Zerlegung ist bis auf Isomorphie eindeutig bestimmt.*

Die nächste Anwendung bezieht sich auf den einfachen Fall zyklischer Torsionsmoduln und ist als chinesischer Restsatz bekannt:

Satz 3.5 *Es sei $a \neq 0$ ein Element eines Hauptidealringes R.*

$$a = ep_1^{r_1} \dots p_k^{r_k}$$

sei die Primzerlegung von a. Dann ist

$$R/(a) \cong \bigoplus_{i=1\dots k} R/(p_i^{r_i})$$

Beweis. Nach dem ersten Zerlegungssatz ist

$$R/(a) = M \cong \bigoplus M(p_i^{r_i})$$

mit Ann $M(p_i^{r_i}) = (p_i^{r_i})$. Da M zyklisch ist, gilt dasselbe für $M(p_i^{r_i})$, also ist

$$M(p_i^{r_i}) \cong R/(p_i^{r_i}).$$

§4 Normalformen von Endomorphismen

1.) Wir kehren zum Normalformproblem für einen Endomorphismus

$$f : V \to V$$

eines endlichdimensionalen K-Vektorraumes zurück und wenden die Ergebnisse des vorigen Paragraphen an. Durch

$$Qx = Q(f)(x) \quad , \; Q \in K[X],$$

wird V ein $K[X]$-Modul, wie wir wissen, ein Torsionsmodul über $K[X]$.

Definition 4.1. *Es sei $W \subset V$ ein Unterraum.*

(i) W heißt f-invariant, wenn $f(W) \subset W$.

(ii) W heißt f-zyklisch, wenn es ein x in W gibt mit $W = L(x, f(x), f^2(x), \cdots)$, wenn also die Bilder von x unter den Iterierten von f den Unterraum W erzeugen.

(iii) W heißt f-irreduzibel, wenn W f-invariant, aber nicht direkte Summe echter f-invarianter Teilräume ist.

Jetzt können wir ein Wörterbuch anlegen, welches die Begriffe der Modultheorie – V als $K[X]$-Torsionsmodul – in die der Vektorraumtheorie (Unterräume usf.) übersetzt.

Satz 4.1 *Für einen Unterraum $W \subset V$ gelten folgende Übersetzungsregeln*

	$K[X]$-*Modul*	K-*Vektorraum*
(i)	*W Untermodul*	*W f-invarianter Unterraum*
(ii)	*W zyklischer Untermodul*	*W f-zyklischer Unterraum*
(iii)	*W direkt unzerlegbar*	*W f-irreduzibel*
(iv)	$\operatorname{Ann} W = (Q)$	*W f-invariant, und $f\|W$ hat Q als Minimalpolynom*

Beweis. Nur (ii) und (iv) sind noch zu zeigen. Ist W zyklischer Untermodul, so ist also $W = K[X]a$, $a \in W$. Zu jedem $x \in W$ existiert somit ein Polynom $Q(X) = a_0 + a_1 X + \cdots + a_l X^l$ mit $Q(X)a = x$, d.h.

$$x = a_0 \cdot a + a_1 f(a) + \cdots + a_l f^l(a).$$

Damit erzeugen die $f^l(a)$ den Unterraum W. Die Umkehrung folgt entsprechend.

Zu (iv). Es sei $R(X)$ das Minimalpolynom von $f|W$, und es sei $(Q) = \operatorname{Ann} W$. Dann ist $Q(f) = 0$ auf W, also $R|Q$. Da $R(f) = 0$ auf W ist, ist $R \in \operatorname{Ann} W$, d.h. $Q|R$. Somit ist $Q = R$ bis auf Einheiten genau.

2.) Wir wenden nun Satz 3.3 auf den $K[X]$-Modul V an:

Es gibt eindeutig bestimmte Primpolynome $P_1, \ldots, P_k$ und P_i-primäre Untermoduln V_i von V mit

$$V = \bigoplus_{i=1\ldots k} V_i$$

Ferner gibt es eindeutig bestimmte ganze Zahlen n_{ij}, $i = 1, \dots, k$, $j = 1, \dots, s_i$ mit

$$1 \le n_{i1} \le \dots \quad \le n_{is_i} \overset{\text{def}}{=} r_i,$$

so daß

$$V_i = \bigoplus_{j=1\dots s_i} V_{ij}$$

in zyklische Untermoduln

$$V_{ij} \cong K[X]/(P_i^{n_{ij}})$$

zerfällt. Die V_i sind die P_i-primären Komponenten von V,

$$V_i = V_{P_i} = V(P_i^{r_i}),$$

und

$$P = P_1^{r_1} \dots P_k^{r_k}$$

erzeugt das Annullatorideal von V.

Übersetzen wir diese Aussagen nun in die Sprache der Vektorräume.
(1) P ist das Minimalpolynom von f.
(2) $V_i =$ Kern $P_i^{r_i}(f)$
(3) V_{ij} ist f-zyklisch und f-irreduzibel.
(4) $P_i^{n_{ij}}$ ist das Minimalpolynom von $f|V_{ij}$.

Es bleiben noch die n_{ij} zu interpretieren. Dazu konstruieren wir eine spezielle Basis von V_{ij}.

Nach Voraussetzung ist

$$V_{ij} = K[X]x,$$

mit einem passenden $x \in V_{ij}$. Nach dem Beweis von Satz 4.1 erzeugen dann

$$x, f(x), f^2(x), ...$$

den Vektorraum V_{ij} über K. Es sei nun

$$\deg\ P_i^{n_{ij}} = l.$$

Dann ist

$$0 = P_i^{n_{ij}} x = f^l(x) + a_{l-1} f^{l-1}(x) + \cdots + a_0 x,$$

also ist $f^l(x)$ eine Linearkombination der $f^\lambda(x)$, mit $\lambda \leq l - 1$. Induktiv folgt, daß

$$x, f(x), \ldots, f^{l-1}(x)$$

den Vektorraum V_{ij} erzeugen; wir erkennen auch leicht die lineare Unabhängigkeit des Systems: ist

$$\sum_{\lambda=0}^{l-1} b_\lambda f^\lambda(x) = 0,$$

so gilt für das Polynom $\sum_{\lambda=0}^{l-1} b_\lambda X^\lambda = Q$ die Beziehung

$$Qx = 0,$$

damit $Q \cdot V_{ij} = 0$ und $Q \in \operatorname{Ann} V_{ij}$. Wegen $\deg\ Q < l = \deg\ P_i^{n_{ij}}$ ist dann $Q = 0$. Somit ist $x, \ldots, f^{l-1}(x)$ eine Basis von V_{ij}, und wir sehen

(5) $\dim V_{ij} = n_{ij}\ \deg\ P_i$.

Zusammenfassend gilt

Satz 4.2 *(Zerlegungssatz für Endomorphismen)*

Es sei $f : V \to V$ ein Endomorphismus eines n-dimensionalen K-Vektorraums. Dann existieren

(1) Primpolynome $P_1(X), \ldots, P_k(X) \in K[X]$,

(2) $n_{ij} \in \mathbf{N}$, $1 \leq n_{i1} \leq \cdots \leq n_{is_i} \stackrel{\text{def}}{=} r_i$, $i = 1, \ldots, k$,

(3) f-zyklische f-irreduzible Unterräume $V_{ij} \subset V$ mit $1 \leq i \leq k$, $1 \leq j \leq s_i$, so daß gilt:

(a)

$$V = \bigoplus_{i=1\ldots k} \bigoplus_{j=1\ldots s_i} V_{ij}$$

(b) $P_i^{n_{ij}}(X)$ ist das Minimalpolynom von $f|V_{ij}$.

Durch f bestimmt sind
(1) die $P_i(X)$ (bis auf Assoziierte),
(2) die n_{ij},
(3) die Räume

$$V_i = \bigoplus_{j=1\ldots s_i} V_{ij},$$

(4) $\dim V_{ij}$,
und zwar ist

$$P(X) = P_1(X)^{r_1} \ldots P_k(X)^{r_k}$$

die Primzerlegung des Minimalpoynoms $P(X)$ von f,

$$V_i = \operatorname{Kern} P_i^{r_i}(f),$$

$$\dim V_{ij} = \deg P_i(X) \cdot n_{ij}.$$

3.) Wie sieht die Matrix von f aus, wenn wir mittels des vorigen Satzes eine geschickte Basis wählen? Es sei zunächst $\mathcal{X}_{ij}$ eine beliebige Basis von V_{ij} und

$$A_{ij} = A_{f|V_{ij}}; \mathcal{X}_{ij}, \mathcal{X}_{ij}.$$

Dann ist

$$\mathcal{X} = \bigcup_{i,j} \mathcal{X}_{ij}$$

eine Basis von V, und bei entsprechender Anordnung der A_{ij} gilt:

$$A_{f;\mathcal{X},\mathcal{X}} = \operatorname{diag}(A_{11}, \ldots, A_{1s_1}, \ldots, A_{k1}, \ldots, A_{ks_k}), \tag{2.1}$$

also eine “Diagonalmatrix von Matrizen“ - cf. den Schluß von §2. Für $\mathcal{X}_{ij}$ wählen wir nun die bei Beweis von (5) eingeführte Basis von V_{ij}. Es sei also

$$\deg P_i \cdot n_{ij} = \dim V_{ij} = l_{ij} \overset{\text{def}}{=} l$$

und

$$P_i(X)^{n_{ij}} = a_0 + a_1 X + \cdots + a_{l-1} X^{l-1} + X^l$$

(wir können natürlich 1 als Leitkoeffizient wählen). Dann gibt es einen Vektor x, so daß

$$x, f(x), f^2(x), \ldots, f^{l-1}(x)$$

eine Basis von V_{ij} ist, wir numerieren sie neu

$$x_1, x_2, x_3, \ldots, x_l,$$

also $x_\lambda = f^{\lambda-1}(x)$. Dann gilt:

$$f(x_\lambda) = x_{\lambda+1}, \quad \lambda = 1, \ldots, l-1$$

$$\begin{aligned} f(x_l) = f^l(x) &= (-a_0 - a_1 X - \ldots - a_{l-1}X^{l-1})x \\ &= -a_0 x_1 - a_1 x_2 - \ldots - a_{l-1} x_l. \end{aligned}$$

Somit wird

$$A_{ij} = \begin{pmatrix} 0 & 0 & 0 & \cdots & \cdots & \cdots & -a_0 \\ 1 & 0 & 0 & \cdots & \cdots & \cdots & -a_1 \\ 0 & 1 & 0 & & & & \vdots \\ 0 & 0 & 1 & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & 0 & a_{l-2} \\ 0 & 0 & 0 & \cdots & \cdots & 1 & a_{l-1} \end{pmatrix}; \qquad (2.2)$$

dabei ist $l = l_{ij}$ und auch die a_λ hängen von i und j ab. Also

Satz 4.3 *Bei geeigneter Basiswahl wird f durch eine Matrix der Gestalt (2.1) (2.2) beschrieben. Dabei sind die a_λ so, daß*

$$X^l + \sum_{\lambda=0}^{l-1} a_\lambda X^\lambda = P_i(X)^{n_{ij}}$$

Potenz eines Primpolynoms ist. Die A_{ij} sind bis auf die Reihenfolge eindeutig durch f bestimmt.

Definition 4.2. *Die Matrizen A_{ij} der Form (2.2) heißen Begleitmatrizen zur Primpolynompotenz $P_i(X)^{n_{ij}}$; die Matrix A der Form (2.1) (2.2) heißt allgemeine Normalform für den Endomorphismus f.*

Bemerkung. Die Existenzaussage des Satzes haben wir gerade bewiesen. Ist nun A eine Matrix für f in allgemeiner Normalform (2.1) (2.2), so zerfällt V entsprechend den A_{ij} in eine direkte Summe f-invarianter Unterräume V_{ij}. Wir wählen die Numerierung so, daß A_{ij} jeweils Begleitmatrix zu einer Primpolynompotenz $P_i(X)^{n_{ij}}$ mit demselben P_i ist. Dann sind die V_{ij} f-zyklisch und f-irreduzibel, und die Zerlegung muß mit einer Zerlegung wie in Satz 4.2 beschrieben übereinstimmen. Das legt die A_{ij} sämtlich fest.

Naürlich sagt Satz 4.3 matrizentheoretisch folgendes aus:

Satz 4.3' *Jede Matrix $A \in K^{n \times n}$ ist ähnlich zu einer (bis auf die Reihenfolge der "Diagonalkästchen" eindeutig bestimmten) Matrix in allgemeiner Normalform.*

4.) Der Fall linearer Primpolynome P_i vereinfacht die Situation. Er tritt bei algebraisch abgeschlossenem Grundkörper immer ein.

Es sei also

$$P(X) = (X - \lambda_1)^{r_1} \dots (X - \lambda_k)^{r_k}$$

das Minimalpolynom von f. Wieder zerlegt sich V in nur von f abhängiger Weise zu

$$V = \bigoplus_{i=1\dots k} V_i$$

mit $V_i = \text{Kern } P_i^{r_i}(f) = \text{Kern}\,(f - \lambda_i id)^{r_i}$, und es ist

$$V_i = \bigoplus_{j=1\dots s_i} V_{ij},$$

wobei die V_{ij} f-zyklisch und f-irreduzibel sind,

$$\dim V_{ij} = n_{ij};$$

und es ist

$$(X-\lambda_i)^{n_{ij}}$$

das Minimalpolynom von $f|V_{ij}$.

Wie früher wählen wir eine Basis von V_{ij} der Gestalt

$$x, f(x), f^2(x)\ldots, f^{n_{ij}-1}(x),$$

die wir durchnumerieren:

$$y_1 = x, \quad y_\lambda = f^{\lambda-1}(x), \quad \lambda = 1,\ldots,n_{ij}.$$

Nun gehen wir aber zu einer besseren Basis über.

Hilfssatz. *Durch*

$$x_1 = x \quad , \quad x_{j+1} = (f-\lambda_i id)x_j,$$

$j = 1,\ldots,n_{ij}-1$, *wird eine Basis von* V_{ij} *gegeben.*

Beweis. Wir zeigen induktiv: $y_j \in L(x_1\ldots,x_j)$, der von $x_1,\ldots,x_j$ erzeugte Unterraum. Für $j=1$ ist das klar: $y_1 = x_1$. Ist die Aussage für $j-1$ schon bewiesen, so ist also

$$y_{j-1} = \sum_{m=1}^{j-1} \alpha_m x_m, \quad \alpha_m \in K.$$

$$\begin{aligned} y_j = f(y_{j-1}) &= \sum \alpha_m f(x_m) \\ &= \sum_{m=1\ldots j-1} \alpha_m(x_{m+1}+\lambda_i x_m) \\ &= \sum_{m=1\ldots j} \beta_m x_m. \end{aligned}$$

Jetzt ergibt sich

$$f(x_j) = \lambda_j x_j + x_{j+1}\ , \quad 1 \le j \le n_{ij}-1,$$

$$(f-\lambda_i)x_{n_{ij}} = (f-\lambda_i)^{n_{ij}} x = 0$$

d.h. $f(x_{n_{ij}}) = \lambda_i x_{n_{ij}}$.

Damit wird die Matrix von $f|V_{ij}$ bez. dieser Basis

$$A_{ij} = \begin{pmatrix} \lambda_i & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i \end{pmatrix} \in K^{n_{ij} \times n_{ij}} \tag{2.3}$$

Definition 4.3. *Für $\lambda \in K$ und $m \in \mathbf{N}$ heißt*

$$J_m(\lambda) = \begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix} \in K^{m \times m}$$

Jordankästchen der Größe m zum Körperelement λ.

Es ist also $J_1(\lambda) = \lambda \in K = K^{1\times 1}$.

Wir sehen:

Satz 4.4. *Es sei $f : V \to V$ ein Endomorphismus mit einem über K zerfallenden Minimalpolynom.*

$$P_f(X) = (X - \lambda_1)^{r_1} \dots (X - \lambda_k)^{r_k}.$$

Dann exitieren eindeutig bestimmte ganze Zahlen

$$1 \le n_{i1} \le \dots \le n_{is_i} = r_i,$$

für $i = 1, \dots, k$, so daß bez. einer geeigneten Basis f durch eine Matrix

$$J = \operatorname{diag}(J_{11}, \dots, J_{1s_1}, \dots, J_{k1}, \dots, J_{ks_k})$$

mit

$$J_{ij} = J_{n_{ij}}(\lambda_i)$$

beschrieben wird.

Definition 4.4. *Eine Matrix*

$$J = \operatorname{diag}(J_1, \ldots, J_t),$$

wobei jedes J_τ ein Jordankästchen ist, heißt Jordanmatrix.

Wir können also abkürzend sagen: unter den Voraussetzungen von Satz 4.4 kann f durch eine Jordanmatrix beschrieben werden. Berechnet man das Minimalpolynom einer Jordan-Matrix, so folgt der Eindeutigkeitssatz

Satz 4.5 *Je zwei Jordanmatrizen zu f unterscheiden sich nur in der Reihenfolge der Jordankästchen.*

Man sagt auch, f sei auf Jordansche Normalform gebracht.

Alle diese Aussagen besitzen eine äquivalente matrizentheoretische Formulierung.

Satz 4.4' und 5'. *Zu $A \in K^{n\times n}$ mit zerfallendem Minimalpolynom existiert $S \in \mathrm{Gl}(n, K)$, so daß $S^{-1}AS$ eine Jordanmatrix ist. Bis auf die Reihenfolge der Jordankästchen ist diese Jordanmatrix eindeutig bestimmt.*
Kurz: "A läßt sich auf Jordansche Normalform bringen."

§5 Anwendungen und Ergänzungen

1.) Es sei $A = \operatorname{diag}(A_{ij})$ eine Matrix in allgemeiner Normalform. Dann gilt für das charakteristische Polynom

$$\chi_A(X) = \det(XE - A) = \prod_{i,j} \det(XE - A_{ij}),$$

wobei E für die Einheitsmatrix der betreffenden Dimension steht. Nun ist

$$XE - A_{ij} = \begin{pmatrix} X & & & & a_0 \\ -1 & X & & & \vdots \\ & -1 & X & & \vdots \\ & & \ddots & & a_{l-2} \\ & & & -1 & X + a_{l-1} \end{pmatrix}$$

Dabei ist

$$a_0 + a_1 X + \cdots + a_{l-1} X^{l-1} + X^l = P_i(X)^{n_{ij}},$$

also $l = n_{ij}$, und $n_{ij} \leq r_i$ mit

$$\prod P_i(X)^{r_i} = P(X),$$

das Minimalpolynom von A. Damit gilt:

$$\begin{aligned} \det(XE - A_{ij}) &= \pm(a_0 + \cdots + a_{l-1} X^{l-1} + X^l) \\ &= \pm P_i(X)^{n_{ij}} \end{aligned}$$

(Entwicklung nach letzter Spalte). Da r_i das Maximum aller n_{ij} ist, taucht insbesondere der Faktor $P_i(X)^{r_i}$ in $\chi_A(X)$ auf, und wir haben als Präzisierung des Satzes von Cayley/Hamilton

Satz 5.1 *Das charakteristische Polynom hat dieselben Primfaktoren wie das Minimalpolynom, evtl. in höherer Potenz.*

Betrachten wir nun eine Jordanmatrix

$$J = \mathrm{diag}(J_{n_{ij}}(\lambda_i)).$$

Es sei wieder

$$P(X) = \prod_{i=1\dots k} (X - \lambda_i)^{r_i}$$

das Minimalpolynom von J.

Satz 5.2 *(i) Die λ_i sind die EWe von J.*

(ii) Genau dann, wenn alle $r_i = 1$ sind, ist J eine Diagonalmatrix.

(iii) Eine Matrix A ist genau dann diagonalisierbar, wenn ihr Minimalpolynom nur einfache Nullstellen hat.

Beweis. (iii) folgt aus (ii), (i) folgt aus Satz 5.1
Zu (ii). Ist $r_i = 1$, so ist wegen $n_{ij} \leq r_i$ jedes $J_{n_{ij}}(\lambda) = (\lambda_i)$, eine 1×1-Matrix. Umgekehrt: ist J eine Diagonalmatrix, so darf niemals ein $n_{ij} > 1$ sein, da sonst $J_{n_{ij}}(\lambda_i)$ nicht diagonal ist.

Aus dem vorigen Paragraphen folgt schließlich noch:

Satz 5.3. *A und B sind genau dann ähnlich, wenn ihre allgemeinen Normalformen übereinstimmen.*

2.)

Definition 5.1. *(i) $f : V \to V$ heißt nilpotent, wenn es ein $N \in \mathbf{N}$ mit $f^N = 0$ gibt.*
(ii) $f : V \to V$ heißt unipotent, wenn $f - id$ nilpotent ist.

Aus der Existenz der Jordanschen Normalform ergibt sich nun

Satz 5.4. *Es sei K algebraisch abgeschlossen. Dann ist $f : V \to V$ eine Summe*

$$f = f_d + f_n$$

eines diagonalisierbaren Endomorphismus f_d und eines nilpotenten Endomorphismus f_n. f_n und f_d vertauschen. Die EWe von f sind die von f_d.

Ohne Beweis notieren wir den

Zusatz. *f_d und f_n sind durch die obigen Bedingungen eindeutig bestimmt.*

Ist insbesondere f invertierbar, so auch f_d, und wir haben

$$f = f_d + f_n = f_d(id + f_d^{-1} f_n).$$

Da $f_d^{-1} f_n$ nilpotent ist, ist $id + f_d^{-1} f_n$ unipotent, und es ergibt sich

Satz 5.5. *K sei wie oben, $f : V \to V$ ein Automorphismus. Dann ist*

$$f = f_d \circ f_u,$$

wobei f_d diagonalisierbar und f_u unipotent ist, mit $f_d \circ f_u = f_u \circ f_d$. Diese Zerlegung ist eindeutig bestimmt.

(Den Eindeutigkeitsbeweis führen wir nicht).

3.) Berechnung der Jordanschen Normalform.

Wir lösen diese Aufgabe an einem Beispiel, sagen aber nichts zur Berechnung der Transformationsmatrizen in der Darstellung

$$J = S^{-1}AS.$$

Es sei also

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathbf{C}^{4\times 4}$$

Gesucht ist die Jordansche Normalform J von A.

Zunächst berechnen wir die Nullstellen des charakteristischen Polynoms χ_A, die Eigenwerte. Man berechnet

$$X_A = \det(XE - A) = (X-2)^4.$$

Damit läßt sich A über $\mathbf{R}$ in Jordansche Normalform bringen; es ist

$$J = \begin{pmatrix} 2 & & & \\ * & 2 & 0 & \\ 0 & * & 2 & \\ & & * & 2 \end{pmatrix} \quad \text{mit} \quad * = 0 \text{ oder } 1.$$

Um über $*$ zu entscheiden, beachten wir:

$$J = S^{-1}AS$$

$$\begin{aligned} rg(J-\lambda E)^k &= rg(S^{-1}AS - \lambda E)^k \\ &= rg S^{-1}(A-\lambda E)^k S \\ &= rg(A-\lambda E)^k, \end{aligned}$$

für jedes $\lambda \in \mathbf{C}$. Insbesondere für $\lambda = 2$:

$$rg(J-2E) = rg(A-2E).$$

Aber offenbar ist der Rang von $J-2E$ gerade die Anzahl der Einsen in der Matrix J unter der Diagonale. Die Rechnung liefert

$$rg(A-2E) = 1,$$

womit J bestimmt ist:

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

Wäre der Rang 2 gewesen, so hätte man $(J-2E)^2$ betrachten müssen, um die Fälle

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

und

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

voneinander zu trennen. Für großes n sind ggf. höhere Potenzen von $A-\lambda E$ zu betrachten - man sieht, daß das Verfahren immer funktioniert.

Beispiel. $A \in \mathbf{R}^{20\times 20}$, $\chi_A(X) = (X-2)^{18}(X-3)^2$, also

$$J = \operatorname{diag}(J_1 J_2), \quad J_1 \in \mathbf{R}^{18\times 18}, J_2 \in \mathbf{R}^{2\times 2}.$$

Sei

$$r_t = rg(A-2E)^t$$
$$s_t = rg(A-3E)^t$$

$s_1 \geq 18$.

$$s_1 = \begin{cases} 18 \quad \text{bedeutet } J_2 = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ 19 \quad \text{bedeutet } J_2 = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}. \end{cases}$$

$r_t \geq 2$. Wir nehmen an, daß wir folgende Informationen erhalten haben:

t	r_t	$r_{t-1} - r_t$	Interpretation
1	14	-	J_1 enthält 12 1-en
2	9	5	5 Kästchen der Größe ≥ 2
3	5	4	4 Kästchen der Größe ≥ 3
4	2	3	3 Kästchen der Größe ≥ 4

§6 Endlich erzeugte Moduln über Hauptidealringen

1.) Wir behandeln beliebige endlich erzeugte Moduln über einem Hauptidealring R und erinnern an den Begriff des freien Moduls.

M heißt frei, wenn es Elemente $x_1, \ldots, x_k \in M$ mit

$$M = \bigoplus_{i=1}^{k} Rx_i$$

gibt, und $Rx_i \cong R$ gilt. Die x_i heißen dann eine Basis von M, und $M \cong R^k$. Grundlegend ist

Satz 6.1 *Je zwei Basen eines freien Moduls haben gleiche Elementzahl (Länge).*

Beweis. Es sei $M = \bigoplus_{i=1\cdots k} Rx_i$ und $p \in R$ ein Primelement. Wir setzen

$$K = R/(p), \quad pM = \{px : x \in M\}, \ \overline{M} = M/pM.$$

Dann ist pM ein Untermodul von M, und der Faktormodul $\overline{M}$ wird durch die Festsetzung

$$\overline{r}\ \overline{x} = \overline{rx},$$

wobei Querstriche Restklassen bezüglich (p) oder pM bedeuten, ein K-Vektorraum. Jetzt ist $\overline{x}_1, \ldots, \overline{x}_k \in \overline{M}$ eine Basis von $\overline{M}$ über K. Klarerweise handelt es sich um ein Erzeugendensystem. Falls nun eine Relation

$$\overline{r}_1\overline{x}_1 + \cdots + \overline{r}_k\overline{x}_k = \overline{0} \in \overline{M}$$

besteht, so heißt das

$$r_1x_1 + \cdots + r_kx_k = px \in M,$$

also, mit $x = s_1x_1 + \cdots + s_kx_k$,

$$(r_1 - ps_1)x_1 + \cdots + (r_k - ps_k)x_k = 0,$$

damit $r_i = ps_i$, d.h. $\overline{r}_i = \overline{0}$. Es folgt

$$k = \dim_K \overline{M},$$

und damit liegt k fest.

Definition 6.1. *Die obige Zahl heißt Rang des freien Moduls M,* $\operatorname{rg} M$*.*

Bei Vektorräumen ist der Rang gerade die Dimension.

Satz 6.2 *Es sei $N \subset M$ ein Untermodul des freien Moduls M. Dann ist auch N frei, und* $\operatorname{rg} N \leq \operatorname{rg} M$*.*

Beweis. Es sei $M = \bigoplus_{j=1\ldots k} Rx_i$, und $p_i : M \to R$ werde durch

$$p_i\left(\sum \lambda_j x_j\right) = \lambda_i$$

definiert. Ferner setzen wir

$$M_i = \bigoplus_{j=1}^{i} Rx_j, \quad N_i = N \cap M_i.$$

Also ist $N_k = N$. Dann ist

$$p_i(N_i) = \mathfrak{a}_i$$

ein Ideal in R, also ein Hauptideal (α_i). Wir wählen $y_i \in N_i$ mit $p_i(y_i) = \alpha_i$; für $\alpha_i = 0$ sei auch $y_i = 0$. Nun folgt:

$$N = Ry_1 \oplus \ldots \oplus Ry_k.$$

Zum Beweis zeigen wir zunächst induktiv

$$N_i = Ry_1 + \ldots + Ry_i.$$

<u>$i = 1$.</u> Es sei $y \in N_1$, dann ist

$$p_1(y) = \alpha\alpha_1.$$

Somit wird

$$p_1(y - \alpha y_1) = 0,$$

damit aber $y = \alpha y_1$, denn $p_1|Rx_1$ ist injektiv.

<u>Schluß von $i-1$ auf i.</u> Es sei $y \in N_i$. Dann ist $p_i(y) = \alpha\alpha_i$ und daher

$$p_i(y - \alpha y_i) = 0.$$

Das heißt aber: $y - \alpha y_i \in M_{i-1} \cap N = N_{i-1}$, und nach Induktionsannahme ist

$$y = \beta_1 y_1 + \cdots + \beta_{i-1} y_{i-1} + \alpha y_i.$$

Um die Direktheit der Darstellung zu zeigen, nehmen wir eine Relation

$$0 = \beta_1 y_1 + \cdots + \beta_i y_i$$

an. Es sei $i = 1$. Aus $\beta_1 y_1 = 0$ folgt β_1 oder $y_1 = 0$, da M torsionsfrei ist. Für $i > 1$ erhalten wir

$$0 = p_i(\beta_1 y_1 + \cdots + \beta_i y_i) = \beta_i p_i(y_i) = \beta_i \alpha_i,$$

also $\beta_i = 0$ oder $\alpha_i = 0$, damit dann auch $y_i = 0$. In beiden Fällen ist $\beta_i y_i = 0$, somit

$$\beta_1 y_1 + \cdots + \beta_{i-1} y_{i-1} = 0;$$

durch Induktion folgt die Behauptung.

2.) **Satz 6.3** *N ein Untermodul eines endlich erzeugten Moduls M. Dann ist auch N endlich erzeugt.*

Beweis. Es sei $M = \sum_{i=1\ldots k} Rx_i$. Wir haben dann einen surjektiven Homomorphismus

$$\begin{aligned} p : R^k &\longrightarrow M \\ (\alpha_1, \ldots, \alpha_k) &\longmapsto \sum \alpha_i x_i. \end{aligned}$$

$p^{-1}(N) \subset R^k$ ist (frei und) endlich erzeugt, also auch $N = p(p^{-1}(N))$.

Definition 6.2. $T = T(M)$, *die Menge der Torsionselemente von* M, *heißt Torsionsmodul von* M.

Satz 6.4 *(i)* M/T *ist frei. (ii)* T *ist endlich erzeugt, und es gibt einen endlich erzeugten freien Untermodul* $F \subset M$ *mit*

$$M = T \oplus F$$

Der Rang von F *ist durch* M *eindeutig bestimmt.*

Definition 6.3. $\operatorname{rg} M = \operatorname{rg} M/T = \operatorname{rg} F$ *heißt Rang des endlich erzeugten Moduls* M.

Beweis von Satz 6.4 Wir zeigen zunächst

Satz 6.5 *Ein endlich erzeugter torsionsfreier Modul* M *ist frei.*

Beweis. Es sei $M = \sum_{i=1\ldots k} Rx_i$. Für $k = 0$ ist nichts zu zeigen. Die Aussage sei nun für alle Moduln, die von weniger als k Elementen erzeugt werden können, bewiesen. Dann schreiben wir

$$M = \sum_{i=1\ldots k-1} Rx_i + Rx_k.$$

Der erste Summand ist nach Induktionsannahme frei, d.h.

$$M = \bigoplus_{j=1\ldots l} Ry_j + Rx_k = M' + Rx_k.$$

Falls $M' \cap Rx_k = \{0\}$, sind wir fertig. Andernfalls gibt es ein $r \neq 0$ in R mit

$$r x_k \in M'.$$

Wir setzen dann $N = rM$. Das ist ein zu M isomorpher Untermodul von M, da M torsionsfrei ist. Weiter ist $N \subset M'$ und M' frei, also ist N frei, also M auch.

Um Teil (ii) von Satz 6.4 zu zeigen, wählen wir $x_1, \ldots, x_k \in M$ so, daß ihre Restklassen $\overline{x}_i$ bezüglich T eine Basis des freien Moduls M/T bilden und setzen

$$F = Rx_1 + \ldots + Rx_k.$$

Offenbar ist $F + T = M$. Ist ferner

$$r_1 x_1 + \cdots + r_k x_k = 0,$$

so ist erst recht

$$r\overline{x}_1 + \ldots + r_k \overline{x}_k = 0,$$

also alle $r_i = 0$. Damit ist F frei. Schließlich folgt aus $x \in F \cap T$:

$$\begin{aligned} x &= r_1 x_1 + \ldots + r_k x_k \\ \overline{0} = \overline{x} &= r_1 \overline{x}_1 + \ldots + r_k \overline{x}_k, \end{aligned}$$

also alle $r_i = 0$ und damit $x = 0$. Wir sehen:

$$M = F \oplus T.$$

3 Darstellungen von Gruppen

§0 Gruppen

1.) Zunächst erinnere ich an Grundbegriffe aus der Gruppentheorie.

$\varphi : G \to H$ sei ein Gruppenhomomorphismus. Dann ist $\operatorname{Kern} \varphi = \{x \in G : \varphi(x) = e\} = N$ ein Normalteiler von G. Das bedeutet, daß Links- und Rechtsnebenklassen übereinstimmen:

$$aN = Na,$$

ferner, daß man auf G/N, der Menge der Nebenklassen, repräsentantenweise eine Gruppenstruktur erklären kann, so daß die Quotientenprojektion

$$\begin{aligned} \pi : G &\to G/N \\ a &\longmapsto aN \end{aligned}$$

ein surjektiver Gruppenhomomorphismus ist. Damit ist jeder Normalteiler Kern eines Homomorphismus, und man beweist leicht den "Faktorisierungssatz":

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ \pi \downarrow & & \nearrow \overline{\varphi} \\ G/\operatorname{Kern}\varphi & & \end{array}$$

d.h. die Existenz des obigen kommutativen Diagramms mit injektivem Gruppenhomomorphismus $\overline{\varphi}$. - Im abelschen Falle ist jede Untergruppe Normalteiler.

2.) Die Ordnung einer Gruppe ist ihre Elementanzahl $\#G = \operatorname{ord} G$ ($\in \mathbf{N}$ oder ∞). Durch Zerlegung von G in die Nebenklassen (rechts oder links) bez. einer Untergruppe G beweist man:

$$\operatorname{ord} G = \operatorname{ord} U \cdot [G : U],$$

wobei der Index von U in G, $[G : U]$, die Anzahl der Nebenklassen ist. Die Formel ist nur im Falle $\operatorname{ord} G < \infty$ von Interesse.

3.) Eine Teilmenge M einer Gruppe G erzeugt die Untergruppe $\langle M \rangle$ mit

$$\langle M \rangle = \bigcap \{U : U \supset M,\ U \text{ Untergruppe}\};$$

ist $\langle M \rangle = G$, so heißt M Erzeugendensystem von G. G ist zyklisch, wenn $G = \langle a \rangle$ für ein geeignetes $a \in G$ ist.

Dann ist $G = \{a^n : n \in \mathbf{Z}\}$, die Menge aller Potenzen von a. Eine zyklische Gruppe kann die Ordnungen $1, 2, \dots$ oder ∞ haben; sie ist dann isomorph zur trivialen Gruppe, zu $\mathbf{Z}_n$ oder $\mathbf{Z}$ (Menge der Kongruenzklassen mod n bzw. die ganzen Zahlen mit Addition als Verknüpfung). Dabei heißen 2 Gruppen isomorph, wenn es einen bijekiven Homomorphismus zwischen ihnen gibt.

Jedes Element $a \in G$ erzeugt die zyklische Untergruppe $\langle a \rangle$, die endlich oder unendlich sein kann. Die Ordnung von $\langle a \rangle$ heißt Ordnung des Elementes a.

4.) Grundlegend für die Gruppentheorie und Ursache für die flexible Anwendbarkeit des Gruppenbegriffes ist folgende Tatsache:

Isomorphe Gruppen können in vielen verschiedenen "Verkleidungen" auftreten.

Z.B. taucht die zyklische Gruppe C_n als $\mathbf{Z}_n, +$ auf, aber auch als Gruppe der Drehungen eines regelmäßigen n-Ecks, als Gruppe der n-ten Einheitswurzeln im Körper $\mathbf{C}$; für $n = p^m - 1$ "ist" C_n auch die multiplikative Gruppe des endlichen Körpers $\mathbb{F}_{p^m}$.

§1 Gruppenoperationen

1.)

Definition 1.1. *Es sei G eine Gruppe, X eine Menge. Eine Operation von G auf X ist eine Abbildung*

$$\begin{aligned} G \times X &\longrightarrow X \\ a, x &\longmapsto ax \end{aligned}$$

mit:
(1) $e \cdot x = x$ für alle $x \in X$
(2) $a(bx) = (ab)x$, $x \in X$, $a, b \in G$.

Die Operation heißt treu, wenn aus $ax = x$ für alle $x \in X$ folgt: $a = e$. — X heißt eine G-Menge, wenn eine Operation von G auf X gegeben ist.

Bemerkungen

(1) Für $a \in G$ ist die Abbildung $x \longmapsto ax$ bijektiv. Denn aus (1) und (2) oben folgt:

$$a^{-1}(ax) = (a^{-1}a)x = e \cdot x = x.$$

Damit ist jedem $a \in G$ also eine Permutation $\sigma \in \mathfrak{S}(X)$ zugeordnet. Klarerweise ist so ein Homomorphismus von G in $\mathfrak{S}(X)$ definiert, der genau für treue Operationen injektiv ist.

(2) $\mathfrak{S}(X)$ operiert in natürlicher Weise auf X.

Definition 1.2. *Es sei X eine G-Menge. Für $x_0 \in X$ ist*

$$G_{x_0} = \{a \in G : ax_0 = x_0\}$$

die Isotropiegruppe von x_0 und

$$B_G(x_0) = \{ax_0 : a \in G\}$$

die Bahn von x_0 unter G.

Satz 1.1 *X zerfällt unter G in disjunkte Bahnen.*

Das ist klar - aber wichtig. Ist nämlich insbesondere X endlich, so gibt es endlich viele Bahnen der Länge (Elementzahl) l_i, und man hat die "Klassengleichung"

$$\#X = l_1 + l_2 + \cdots + l_r.$$

Besteht X nur aus einer einzigen Bahn, so heißt G transitiv.

Satz 1.2 *G operiere auf X, x und $y \in X$ seien Elemente derselben Bahn, G_x, G_y die Isotropiegruppen. <*

(i) Die Menge $\{a \in G : ax = y\}$ ist eine Linksnebenklasse von G_x.

(ii) $G/G_x \longleftrightarrow B_G(x)$ umkehrbar eindeutig.

(iii) G_x und G_y sind konjugierte Untergruppen:

$$G_y = a^{-1}G_x a.$$

Beweis. (i) (ii) Wähle $a \in G$ fest mit $ax = y$. Annahme, $bx = y$. Dann ist $b^{-1}ax = x$, d.h. $b^{-1}a \in G_x$; damit $a \in bG_x$ und $b \in aG_x$. Umgekehrt, wenn $b = au$, $u \in G_x$, so ist $bx = ax = y$.

(iii) Wähle a mit $ax = y$. Dann ist

$$G_y = aG_x a^{-1}.$$

Folgerung (Bahnformel):

$$\operatorname{ord} G = \operatorname{ord} G_x \cdot l_x, \quad l_x = \text{Länge der Bahn von } x$$

2.) Ich gebe einige einfache Anwendungen dieser Überlegungen.

Jede Gruppe operiert auf sich selbst durch Multiplikation:

$$a, x \longmapsto ax.$$

Die Operation ist treu, damit

Satz 1.3 *(Cayley)* *$G \cong$ Untergruppe von $\mathfrak{S}(G)$.*

G operiert auch durch Konjugation auf sich:

$$a, x \longmapsto a * x \stackrel{\text{def}}{=} axa^{-1}.$$

Definition 1.3. *Der Zentralisator von x ist*

$$Z(x) = \{a : a * x = x\},$$

die Menge aller mit x vertauschbaren Elemente.

Definition 1.4. *Die Konjugiertenklasse von x ist die Menge $K(x) = \{a * x : a \in G\}$.*

Damit haben wir also Bahn und Isotropiegruppe unter der $*$-Operation bezeichnet; Klassengleichung und Bahnformel lauten bei einer endlichen Gruppe:

$$\begin{aligned} \operatorname{ord} G = & n = k_1 + \ldots + k_m, \\ & n = \operatorname{ord} Z(x) \cdot k(x); \end{aligned}$$

dabei ist k_μ die Länge der Bahn (Konjugiertenklasse) K_μ, entsprechend $k(x)$ die Länge der Bahn $K(x)$.

Wir notieren

(1) Die k_μ sind Teiler von n.
(2) Ist $N \subset G$ ein Normalteiler, so ist N eine Vereinigung von Konjugiertenklassen. Dadurch sind Normalteiler unter den Untergruppen charakterisiert.

Schließlich ist $K(e) = \{e\}$, d.h. in der Klassengleichung taucht mindestens einmal die 1 auf. Genauer:

Definition 1.5. *Das Zentrum von G, $Z(G)$, ist die Menge aller mit allen $x \in G$ vertauschbaren Elemente.*

Offenbar ist $Z(G)$ ein Normalteiler, er besteht aus allen Konjugationsklassen der Länge 1. Ist $z = \operatorname{ord} Z(G)$, so kann die Klassengleichung auch als

$$n = z + k_1 + \cdots k_r$$

geschrieben werden, mit $k_\delta > 1$, $z \geq 1$, und z, k_δ Teiler von n.

Diese einfachen Abzählspiele haben überraschende Anwendungen
- siehe die folgenden Paragraphen.

§2 Beispiele und Anwendungen

1.) Die einfachsten Gruppen sind die zyklischen Gruppen C_n der Ordnung n und C_∞, die unendliche zyklische Gruppe; sie bestehen jeweils aus den Potenzen eines Elementes, und man hat die Isomorphien

$$C_n \cong \mathbf{Z}_n, + \quad , \quad C_\infty \cong \mathbf{Z}, +$$

2.) Mit D_n bezeichne ich die Gruppe der Bewegungen, die ein regelmäßiges n-Eck in sich überführen. D_n ist nicht abelsch (für $n > 2$) und hat $2n$ Elemente; $D_n \supset C_n$ als Normalteiler der Drehungen um den Winkel $2\pi/n$. Die D_n heißen Diedergruppen. Die Gruppe wird erzeugt von einer Drehung um $2\pi/n$, nennen wir sie x, und einer Spiegelung y an einer der Symmetrieachsen des n-Ecks; es bestehen die Relationen $x^n = e$, $y^2 = e$ und $xy = yx^{-1}$. D_3 etwa ist isomorph zur symmetrischen Gruppe $\mathfrak{S}_3$; die Konjugationsklassen sind $\{e\}$, $\{x, x^2\}$ sowie $\{y, xy, x^2y\}$. Die Klassengleichung ist

$$1 + 2 + 3 = 6.$$

3.) Mit T, W, J bezeichnen wir die Gruppen der Bewegungen, die ein regelmäßiges Tetraeder, Hexaeder (Würfel) bzw. Ikosaeder in sich überführen. Wir fassen sie als Untergruppen von $SO(3)$ auf. Sie haben $12, 24$ bzw. 60 Elemente. Für ihre genaue Beschreibung verweise ich auch auf Lehrbücher der Gruppentheorie bzw. der Geometrie, notiere hier zunächst die Isomorphien

$$T \cong \mathfrak{A}_4$$

$$J \cong \mathfrak{A}_5,$$

wobei $\mathfrak{A}_4, \mathfrak{A}_5$ die alternierenden Gruppen von 4 bzw. 5 Objekten sind. Außerdem ist klar: die Symmetriegruppe eines Oktaeders ist die des einbeschriebenen Würfels, die eines Dodekaeders die des einbeschriebenen Ikosaeders. Zusammen mit den zyklischen Gruppen C_n und den Diedergruppen handelt es sich bei diesen Gruppen um (bis auf Konjugation) alle endlichen Untergruppen von $SO(3)$, d.h. jede endliche Untergruppe von $SO(3)$ ist isomorph zu einer der obigen Gruppen, und sind zwei endliche Untergruppen von $SO(3)$ isomorph, etwa $G_1 \cong G_2$, so existiert ein $A \in SO(3)$ mit $G_2 = A^{-1}G_1A$. Vgl. hierzu Brieskorn: Lineare Algebra bzw. Speiser: Gruppentheorie.

4.) Wir beschreiben T und J nun etwas genauer. T enhält zunächst eine Drehung um 120° um einen Eckpunkt (genauer: die Drehachse geht durch einen Eckpunkt und den Schwerpunkt der gegenüberliegenden Seite). Verbindet man ferner die Mittelpunkte zweier sich gegenüberliegenden Kanten (es gibt 3 Paare solcher Kanten), so enthält T die 180° Drehungen um diese Verbindungsstrecken. Wir bezeichnen diese 4 Elemente mit x, y_1, y_2, y_3. Bzeichnen wir die Eckpunkte mit $1, 2, 3, 4$, so permutieren die Elemente von T diese Eckpunkte, und zwar, wenn 1 unter x fest bleibt, in der folgenden Weise:

$$\begin{aligned} x: &\quad (234) \\ y_1: &\quad (12)(34) \\ y_2: &\quad (13)(24) \\ y_3: &\quad (14)(23) \end{aligned}$$

(Wir haben Zykeldarstellung für Permutationen verwandt). Die x, y_i erzeugen T, und ihre Wirkung auf die Eckpunkte liefert einen Isomprhismus von T mit $\mathfrak{A}_4$. Bei geeigneter Koordinatenwahl können wir die x, y_i, durch die folgenden Drehmatrizen beschreiben:

$$\begin{aligned} y_1 &= \operatorname{diag}(\ 1 \ -1 \ -1) \\ y_2 &= \operatorname{diag}(-1 \ \ 1 \ -1) \\ y_3 &= \operatorname{diag}(-1 \ -1 \ \ 1) \end{aligned}$$

$x = (e_2 e_3 e_1)$ mit e_i als i-tem Einheitsvektor.

J hat 60 Elemente. Es sind (a) Drehungen um $2\pi/5$ um die Zentren der Dodekaederseiten, (b) Drehungen um $2\pi/3$ um die Dodekaeder-Eckpunkte, (c) Drehungen um π um die Kantenmittelpunkte - und Iterationen dieser Prozesse. (Wir fassen J als Symmetriegruppe des Dodekaeders auf!). Jeder Eckpunkt wird von einer Untergruppe der Ordnung 3 festgelassen, und die läßt auch den gegenüberliegenden Eckpunkt fest. Das liefert 10 Untergruppen der Ordnung 3 und 20 Elemente der Ordnung 3. Ebenso gibt es 6 Untergruppen der Ordnung 5, die jeweils ein Paar gegenüberliegender Seiten in sich überführen und damit 24 Elemente der Ordnung 5. Schließlich hat man 15 Untergruppen der Ordnung 2, die Paare gegenüberliegender Kanten in sich überführen, und es gibt noch die Identität. Damit haben wir alle Elemente mit ihren Ordnungen (in Klammern darunter):

$$\begin{array}{ccccccccc} 60 & = & 1 & + & 15 & + & 20 & + & 24 \\ & & (1) & & (2) & & (3) & & (5) \end{array}$$

Wir wollen die Klassengleichung aufstellen.

Zunächst operiert J auf den 20 Eckpunkten transitiv; daher sind die Isotropiegruppen der Eckpunkte alle zueinander konjugiert - siehe §1. Aus dem ensprechenden Grunde sind die Untergruppen der Ordnung 5 und die der Ordnung 2 jeweils untereinander konjugiert. Damit bilden zunächst einmal die 15 Elemente der Ordnung 2 eine Konjugiertenklasse. Bei den Elementen der Ordnung 3 haben wir Drehungen um $2\pi/3$ und um $4\pi/3$ um die Achse durch zwei gegenüberliegende Eckpunkte. Unter der Konjugation der Isotropiegruppen der Eckpunkte gehen natürlich 120°-Drehungen in ebensolche über, entsprechend Drehungen um 240°. Ist nun P' der Eckpunkt, der P genau gegenüberliegt, so ist die 120° Drehung um P' dasselbe wie die 240° Drehung um P; unter der Konjugation der beiden Isotropiegruppen geht also die 120°-Drehung um P in die 240°-Drehung um P über, damit sind auch die Elemente der Ordnung 3 eine volle Konjugiertenklasse. Genauso erkennt man, daß die $2\pi/5$-Drehungen zu den $8\pi/5$-Drehungen konjugiert sind und die $4\pi/5$-Drehungen zu den $6\pi/5$-Drehungen. Damit lautet die Klassengleichung von J :

$$\begin{array}{ccccccccccc} 60 & = & 1 & + & 15 & + & 20 & + & 12 & + & 12 \\ & & (1) & & (2) & & (3) & & (5) & & (5) \end{array}$$

Es können **nicht** alle Elemente der Ordnung 5 untereinander konjugiert sein, da 24 kein Teiler von 60 ist!

Eine Konsequenz ist

Satz 2.1 *Die Gruppe J ist einfach.*

Definition 2.1. *Eine Gruppe heißt einfach, wenn sie keine echten Normalteiler besitzt.*

Beweis von Satz 2.1. Ein echter Normalteiler muß eine Ordnung haben, die ein echter Teiler von 60 ist und die sich additiv aus den Zahlen $1, 15, 20, 12, 12$ zusammensetzt, wobei 1 vorkommen muß. So eine Zahl gibt es nicht.

Wegen $J \cong \mathfrak{A}_5$ ist also $\mathfrak{A}_5$ einfach - dieser von E. Galois bewiesene Satz steht am Anfang der modernen Algebra.

5.) Wir wenden uns nun einigen unendlichen Gruppen zu.

$\mathbf{R}, +$ ist die additive Gruppe der reellen Zahlen. Durch

$$z = \varphi(t) = e^{2\pi it}$$

wird sie homomorph auf $S^1 = \{z \in \mathbf{C} : |z| = 1\}$ abgebildet; der Kern von φ ist $\mathbf{Z}$, d.h.

$$S^1 \cong \mathbf{R}/\mathbf{Z}.$$

(Dabei ist auf S^1 die Multiplikation komplexer Zahlen als Verknüpfung zu wählen). Natürlich können wir S^1 auch anders interpretieren:

$$S^1 \cong U(1, \mathbf{C}) \cong SO(2)$$

als unitäre Gruppe der Dimension 1 oder als Drehgruppe im $\mathbf{R}^2$. Für die Physik das wichtigste Beispiel ist die Gruppe $SU(2, \mathbf{C})$: wir untersuchen sie nun.

$A \in SU(2)$, mit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

bedeutet

$${}^t\overline{A} = A^{-1} \quad \text{und} \quad \det A = 1.$$

Das heißt

$$\begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

also $\overline{a} = d$ und $\overline{c} = -b$. Damit:

Satz 2.2. *Die Matrizen in* $SU(2)$ *haben die Gestalt*

$$A = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}, \ |a|^2 + |b|^2 = 1.$$

Durch $\varphi(A) = \binom{a}{b}$ wird eine umkehrbare in beiden Richtungen reell-analytische Abbildung

$$\varphi : SU(2, \mathbf{C}) \to S^3 \subset \mathbf{C}^2$$

(mit $S^3 = \{(a, b) : |a|^2 + |b|^2 = 1\}$) definiert.

Die Menge S^3 ist die 3-dimensionale Einheitssphäre - sie trägt also eine analytische Gruppenstruktur. Das hat sie mit der Sphäre S^1 gemein. Man weiß, daß alle anderen Sphären S^n keine solche Struktur zulassen.

Wir fassen jetzt $SU(2, \mathbf{C})$ als S^3 auf und beschreiben geometrisch die Konjugationsklassen. Das wird sich als hübsche Anwendung des Spektralsatzes herausstellen.

Für $-1 \leq \gamma \leq 1$ sei

$$S^2(\gamma) = \{(a, b) : \operatorname{Re} a = \gamma\} \subset S^3;$$

für $\gamma = \pm 1$ ist also $S^2(\gamma)$ Nord- bzw. Südpol der Sphäre S^3, für $\gamma \neq \pm 1$ ist $S^2(\gamma)$ eine 2-dimensionale Sphäre vom Radius $\sqrt{1 - \gamma^2}$, und zwar ein “Breitenkreis“ der S^3. S^3 ist die disjunkte Vereinigung der $S^2(\gamma)$.

Satz 2.3. *Unter* φ *gehen die Konjugationsklassen von* $SU(2, \mathbf{C})$ *in die Breitenkreise* $S^2(\gamma)$ *über.*

Kürzer: die “Breitenkreise“ sind die Konjugationsklassen.

Beweis. Ist $A \in S^2(\gamma)$, so ist

$$\begin{aligned}\chi_A(X) &= X^2 - (2\operatorname{Re} a)X + 1\\ &= X^2 - 2\gamma X + 1;\end{aligned}$$

damit ist $2\gamma = \operatorname{Sp} A$. Konjugierte Elemente haben dieselbe Spur, somit sind die $S^2(\gamma)$ jeweils Vereinigung voller Konjugationsklassen. Es seien nun λ und $\overline{\lambda}$ die Eigenwerte von A (beachte, daß die Nullstellen von χ_A konjugiert komplex zueinander sind!). Nach dem Spektralsatz existiert eine unitäre Matrix S mit

$$S^{-1}AS = {}^t\,\overline{S}AS = \operatorname{diag}(\lambda, \overline{\lambda}).$$

Indem man mit $(\det S)^{-1}$ multipliziert, kann man $S \in SU(2, \mathbf{C})$ erreichen. Also ist A konjugiert zu $\operatorname{diag}(\lambda, \overline{\lambda})$; die Eigenwerte sind aber durch die Spur bestimmt, wegen

$$\begin{aligned}\lambda + \overline{\lambda} &= \operatorname{Sp} A = 2\gamma\\ \lambda\overline{\lambda} &= 1\end{aligned}$$

Damit ist jedes $A \in S^2(\gamma)$ zu $\operatorname{diag}(\lambda, \overline{\lambda})$ konjugiert.

§3 Darstellungen

1.) Es sei G eine Gruppe, K ein Körper.

Definition 3.1. *Eine Darstellung einer Gruppe G ist ein Homomorphismus.*

$$\begin{aligned}R : G &\to \operatorname{Gl}(n, K)\\ g &\longmapsto R(g) = R_g\end{aligned}$$

Genauer: "n-dimensionale Matrixdarstellung über dem Körper K".

Definition 3.2. *Eine Darstellung von G auf dem Vektorraum V ist ein Homomorphismus*

$$\rho : G \to \operatorname{Gl}(V).$$

Die Dimension von V heißt Dimension der Darstellung.

Jede n-dimensionale Matrixdarstellung ist natürlich eine Darstellung auf dem Vektorraum K^n, und umgekehrt kann durch Wahl einer Basis in V eine n-dimensionale Darstellung auf eine n-dimensionale Matrixdarstellung zurückgeführt werden.

Definition 3.3. *Zwei n-dimensionale Matrixdarstellungen R, R' heißen konjugiert, wenn es ein $S \in \mathrm{Gl}(n, K)$ mit*

$$R'_g = S^{-1} R_g S$$

gibt.

Ist $\rho : G \to \mathrm{Gl}(V)$ eine n-dimensionale Darstellung, R und R' zwei Matrixdarstellungen von ρ bezüglich verschiedener Basen, so sind R' und R offenbar konjugiert.

Definition 3.4. *Eine lineare G-Operation auf dem Vektorraum V ist eine Abbildung*

$$G \times V \to V,$$

die eine Gruppenoperation (siehe Def. 1.1) ist und zusätzlich den Bedingungen

$$g(x + y) = gx + gy$$
$$g(\lambda x) = \lambda gx$$

für $g \in G$, $\lambda \in K$, $x, y \in V$, genügt.

Offensichtlich sind lineare G-Operationen auf V dasselbe wie Darstellungen von G über V. Wir können die Begriffe aus §1 damit anwenden und von "Isotropiegruppe", "Bahn" usw. sprechen.

Definition 3.5. *Eine injektive Darstellung heißt treu.*

Wir werden bei Darstellungen sehr oft statt $\rho(g)$, $R(g)$, ρ_g oder R_g einfach g schreiben: also

$$\rho_g x = gx \quad \text{usw.}$$

Das ist kürzer, allerdings gelegentlich mehrdeutig.

2.) Beispiele von Darstellungen kennen wir natürlich schon:

$$id : \mathrm{Gl}(V) \to \mathrm{Gl}(V)$$

ist die "tautologische" Darstellung von $\mathrm{Gl}(V) = G$. Wählt man eine Basis in V (falls $\dim V < \infty$), so erhält man eine Matrixdarstellung von G.

Die in §2 angegebenen Matrizen in $\mathbf{R}^{3\times 3}$ liefern eine 3- dimensionale Matrixdarstellung der Tetraedergruppe T.

Alle diese Dartstellungen sind treu. Jede Gruppe besitzt natürlich die triviale Darstellung

$$\begin{aligned} \rho : G &\to \mathrm{Gl}(V) \\ g &\longmapsto id \end{aligned}$$

bzw.

$$\begin{aligned} R : G &\to \mathrm{Gl}(n, K) \\ g &\to E \text{ Einheitsmatrix} \end{aligned}$$

Wir werden mehr Beispiele kennenlernen.

3.)

Definition 3.6. *Es seien ρ, ρ' Darstellungen von G auf Vektorräumen V, V'. Eine lineare Abbildung $\varphi : V \to V'$ heißt G-äquivariant, wenn für alle $g \in G$ und $x \in V$ gilt:*

$$\varphi\rho_g x = \rho'_g\varphi x.$$

M.a.W.:"φ vertauscht mit der G-Operation". Bezeichnen wir ρ, ρ' nicht extra, sondern schreiben $\rho_g = g$, so lautet die Bedingung kürzer – und etwas unklarer –

$$\varphi \circ g = g \circ \varphi.$$

Definition 3.7. *ρ und ρ' heißen isomorph, wenn es einen G -äquivarianten Isomorphismus $\varphi : V \to V'$ gibt.*

In diesem Fall kann man durch Wahl geeigneter Basen in V und V' - welcher nämlich? - erreichen, daß ρ und ρ' durch dieselbe Matrixdarstellung beschrieben werden. Es folgt leicht:

Satz 3.1 *Die Isomorphieklassen von Darstellungen auf n-dimensionalen Vektorräumen entsprechen umkehrbar eindeutig den Konjugationsklassen von n-dimensionalen Matrixdarstellungen.*

§4 Unitäre Darstellungen

1.) Ab sofort werden nur noch komplexe Vektorräume endlicher Dimension betrachtet: $K = \mathbf{C}$. Bis auf weiteres sei G eine **endliche** Gruppe der Ordnung N.

2.)

$$\rho : G \to \mathrm{Gl}(V)$$

sei eine Darstellung. Wir können auf V ein Skalarprodukt $\langle\, ,\rangle$ wählen. Mittels ρ definieren wir dann eine neue hermitische Form durch “Mittelung über die Gruppe“:

$$(x, y) = \frac{1}{N} \sum_{g \in G} \langle gx, gy\rangle, \quad N = \mathrm{ord}\, G$$

für x und $y \in V$, und beweisen den

Hilfssatz 1.*(1)* $(\, ,)$ *ist eine positiv definite hermitische Form.*
(2) Für alle $g \in G$ und $x, y \in V$ ist

$$(gx, gy) = (x, y).$$

Kürzer: $(\, ,)$ *ist ein G-invariantes Skalarprodukt auf V.*

Beweis. Offenbar ist $(\, ,)$ eine hermitische Form. Wegen

$$(x, x) = \frac{1}{N} \sum_{g} \langle gx, gx\rangle \geq \frac{1}{N} \langle x, x\rangle$$

ist sie positiv definit. Ist schließlich $h \in G$, so hat man

$$\begin{aligned}(hx, hy) = \frac{1}{N}\sum_g \langle ghx, ghy\rangle &= \frac{1}{N}\sum_g \langle gx, gy\rangle \\ &= (x, y),\end{aligned}$$

denn mit g durchläuft auch gh alle Gruppenelemente, d.h. der 2-te und 3-te Term weichen nur in der Reihenfolge der Summanden voneinander ab.

Es sei nun R die der Darstellung ρ nach Wahl eine Basis von V zugeordnete Matrixdarstellung, S sei die Matrixdarstellung von ρ bezüglich einer Orthonormalbasis für $(\,,\,)$. Da ρ_g eine Isometrie ist, sind alle Matrizen S_g unitär, und S ist andererseits konjugiert zu R. Wir haben damit den fundamentalen

Satz 4.1 *Jede Matrixdarstellung einer endlichen Gruppe G ist konjugiert zu einer unitären Darstellung (i.e. $S : G \to U(n)$).*

Äquivalent hierzu ist die

Folgerung 1. *Jede endliche Untergruppe von* $\mathrm{Gl}(n, \mathbf{C})$ *ist konjugiert zu einer endlichen Untergruppe von $U(n)$.*

Dazu wende man den Satz auf die tautologische Darstellung

$$id : G \to \mathrm{Gl}(n, \mathbf{C})$$

an. Ebenso ergibt sich

Folgerung 2. *Ist $\rho : G \to \mathrm{Gl}(V)$ eine Darstellung der endlichen Gruppe G, so sind alle Automorphismen σ_g diagonalisierbar. Insbesondere ist jeder Automorphismus f endlicher Ordnung (i.e. $f^k = id$) diagonalisierbar.*

Das folgt, da unitäre Matrizen diagonalisierbar sind.

3.) Wenn auch jede einzelne Matrix R_g einer Darstellung $R : G \to \mathrm{Gl}(n, \mathbf{C})$ diagonalisierbar ist, kann man i.a. keine Basis finden, in der alle R_g gleichzeitig Diagonalgestalt haben. Dazu wäre notwendig, daß G abelsch ist - ob das auch hinreicht, lasse ich zunächst offen.

§5 Irreduzible Darstellungen

1.) Die Gruppe G und der Grundkörper K seien zunächst beliebig, V ein endlichdimensionaler K-Vektorraum. Es sei

$$\rho : G \to \mathrm{Gl}(V)$$

eine Darstellung, die wir als G-Operation auf V schreiben.

Definition 5.1. *Ein Unterraum $W \subset V$ heißt ρ-invariant oder G-invariant, wenn $gW \subset W$ für alle $g \in G$ gilt.*

In diesem Fall erhält man durch Einschränkung von ρ_g eine Darstellung ρ' von G über W. Ist $a_1, \ldots, a_k$ eine Basis von W und ergänzt man diese durch $a_{k+1}, \ldots, a_n$ zu einer Basis von V, so hat die zugehörige Matrixdarstellung R von ρ die Gestalt

$$R_g = \begin{pmatrix} R'_g & * \\ 0 & * \end{pmatrix},$$

wobei R' eine k-dimensionale Martrixdarstellung von G ist, welche zu ρ', der "Einschränkung" von ρ auf W, gehört. Umgekehrt: findet man eine Basis, in der alle R_g die obige Gestalt haben, so ist die lineare Hülle der ersten k Basisvektoren ein ρ- invarianter Unterraum.

V und 0 sind immer invariant.

Definition 5.2. *Eine Darstellung heißt irreduzibel, wenn 0 und V die einzigen invarianten Unterräume sind, sonst reduzibel. V heißt dann auch irreduzibel (bez. ρ).*

2.) **Beispiel 1.** D_n Diedergruppe, erzeugt durch die Drehung um $2\pi/n$ und Spiegelung an einer Symmetrieachse des regelmäßigen n-Ecks.

$$\rho : \quad D_n \to O(2) \subset SU(2)$$

wird so definiert: das n-Eck ist in den Einheitskreis einbeschrieben, eine Ecke ist der Vektor $(1, 0)$. Sind dann g und h die obigen Erzeugenden, so ist

$$\rho(g) = \begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix},$$

$$\rho(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Die Darstellung ist offensichtlich treu, und, da es keinen D_n-invarianten Unterraum des $\mathbf{R}^2$ gibt ($n \geq 3$), ist sie auch irreduzibel.

Beispiel 2 Wir sehen jetzt das n-Eck als in eine Ebene des $\mathbf{R}^3$ eingebettet an; als Basis des $\mathbf{R}^3$ wählen wir eine ONB e_1, e_2, e_3, wobei e_1 im Mittelpunkt des n-Eckes senkrecht auf Ebene E steht, in der das n-Eck liegt; e_2 und e_3 sind eine ONB in der n-Eck-Ebene, e_2 sei wieder eine Ecke des n-Ecks, das in den Einheitskreis einbeschrieben ist. Diese Konstruktion liefert eine Darstellung.

$$\rho : D_n \to SO(3)$$

mit

$$\rho g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\pi/n & -\sin 2\pi/n \\ 0 & \sin 2\pi/n & \cos 2\pi/n \end{pmatrix}$$

$$\rho h = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Diese 3-dimensionale Darstellung ist reduzibel: E und $E^\perp = \mathbf{R}e_1$ sind invariante Unterräume. Die Einschränkung von ρ auf $E^\perp$ ist eine 1-dimensionale Darstellung von D_n:

$$\rho_1 : D_n \to O(1),$$

mit

$$\rho_1 g = 1, \quad \rho h = -1;$$

die Einschränkung von ρ auf E ist die Darstellung $\rho_2 = \rho$ aus Beispiel 1.

Definition 5.3. *$\rho : G \to \mathrm{Gl}(V)$ sei eine Darstellung. Ist $V = V_1 \oplus V_2$ mit G-invarianten Unterräumen, sind ρ_1 und ρ_2 die zugehörigen Einschränkungen von ρ, so heißt ρ direkte Summe von ρ_1 und ρ_2:*

$$\rho = \rho_1 \oplus \rho_2$$

Wählt man in diesem Fall Basen von V_1 und V_2, so gilt für die zugehörigen Matrixdarstellungen (in diesen Basen!)

$$R_g = R_{1,g} \oplus R_{2,g}$$

in demSinn

$$Rg = \begin{pmatrix} R_{1,g} & 0 \\ 0 & R_{2,g} \end{pmatrix}$$

In Beispiel 2 ist

$$\rho = \rho_1 \oplus \rho_2,$$

wobei ρ_1 und ρ_2 irreduzibel sind.

3.) Wir bleiben nun bei unitären Darstellungen, damit ist wieder $K = \mathbf{C}$. Grundlegend ist der einfache

Satz 5.1 *Es sei ρ eine unitäre Darstellung von G auf dem unitären (endlichdimensionalen) Vektorraum V und W ein G-invarianter Unterraum. Dann ist auch* $W^\perp$ *G- invariant.*

Beweis. Wir schreiben wieder g für ρg, $g \in G$. Jedes g ist unitär. Ist $x \perp W$, so gilt für $y \in W$ auch $< gx, y >=< x, g^{-1}y >= 0$, da auch $g^{-1}y \in W$.

Induktiv können wir V nun in eine endliche direkte Summe G- invarianter Unterräume zerlegen, von denen keiner weiter invariant zerlegbar ist. Also:

Satz 5.2 *Jede unitäre Darstellung ist direkte Summe irreduzibler unitärer Darstellungen.*

Zusammen mit Satz 4.1 liefert das den

Satz 5.3 *(Satz von Maschke) Jede Darstellung einer endlichen Gruppe (auf einem endlichdimensionalen* **C***-Vektorraum) ist direkte Summe irreduzibler Darstellungen.*

4.) Schlußbemerkung. Es sei $R : G \to \mathrm{Gl}(n, k)$ eine Matrixdarstellung, also eine Darstellung über dem Vektorraum k^n. $K \supset k$ sei ein Erweiterungskörper von k. Wegen $\mathrm{Gl}(n, k) \subset \mathrm{Gl}(n, K)$ kann R auch als Darstellung von G über K^n angesehen werden. Dann ist möglicherweise R über k^n irreduzibel, über K^n aber reduzibel! Der Fall $k = \mathbf{R}$, $K = \mathbf{C}$ zeigt das - siehe Übungsaufgaben!

§6 Gruppencharaktere

1.) Es sei $\rho : G \to \mathrm{Gl}(V)$ eine Darstellung von G auf dem n-dimensionalen Vektorraum V (über $\mathbf{C}$). G braucht zunächst nicht endlich zu sein.

Definition 6.1. *Der Charakter von ρ wird durch*

$$\chi_\rho(g) = \mathrm{Spur}\,\rho(g) \quad (\in \mathbf{C})$$

definiert. Die Dimension von χ ist die von ρ (d.h. die von V). χ heißt irreduzibel, wenn ρ es ist.

Ein Charakter ist also eine Abbildung von G nach $\mathbf{C}$. Wir notieren

Satz 6.1 *Es sei χ der Charakter von ρ. Dann gilt*

(i) $\chi(e) = \dim\rho = \dim\chi$ *(mit e als 1-Element)*
(ii) $\chi(a^{-1}ga) = \chi(g)$ *für alle* $a, g \in G$
(iii) $\chi(g^{-1}) = \overline{\chi(g)}$
(iv) Ist $\rho = \rho' \oplus \rho''$, *so ist* $\chi_\rho = \chi_{\rho'} + \chi_{\rho''}$

Beweis. (i) (ii) und (iv) sind evident. Zu (iii) Es seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von $\rho(g)$, dann sind $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ die von $\rho(g^{-1})$. Da $\rho(g)^m = id$ für passendes m ist, sind die λ_j Einheitswurzeln, und $|\lambda_j| = 1$. Dann ist aber $\lambda_j^{-1} = \overline{\lambda_j}$, und die Behauptung folgt.

Ferner notieren wir

Satz 6.2 *Isomorphe Darstellungen haben denselben Charakter.*

2.) Zum weiteren Studium von Charakteren benötigen wir

Satz 6.3 *(Schursches Lemma). Es seien ρ und ρ′ irreduzible Darstellungen einer (beliebigen) Gruppe G auf endlichdimensionalen* **C***-Vektorräumen V und V′*;

$$\varphi : V \to V'$$

sei eine G-äquivariante lineare Abbildung. Dann gilt *(i) φ ist ein Isomorphismus oder die Nullabbildung.*

(ii) Für $V = V'$ und $\rho = \rho'$ ist φ eine Homothetie, d.h. $\varphi(x) = c \cdot x$ für $c \in \mathbf{C}$ geeignet.

Beweis. (1) Es sei $K = \operatorname{Kern} \varphi$. Für $x \in K$ und $g \in G$ ist dann

$$\varphi g x = g \varphi x = g 0 = 0,$$

d.h. K ist G-invariant. Ebenso ist $B = \operatorname{Bild} \varphi$ G-invariant.

(2) Da V keine G-invarianten Unterräume außer 0 und V enthält, ist entweder $K = V$, also $\varphi = 0$, oder $K = 0$. Dann ist φ injektiv, $B = \operatorname{Bild} \varphi \neq 0$, daher $B = V'$, und φ ist auch surjektiv.

(3) Im Fall (ii) sei c ein EW von φ. Dann ist auch $\varphi - c \cdot id$ G-äquivariant, mit Kern $\neq 0$, also muß $\varphi - c\, id = 0$ sein.

Folgerung 1. *Es sei ρ eine Darstellung von G über V; $\rho = \rho_1 \oplus \ldots \oplus \rho_m$ eine Zerlegung von ρ in irreduzible Darstellungen ρ_μ über $V_\mu \subset V$; also*

$$V = V_1 \oplus \ldots \oplus V_m$$

mit G-invarianten irreduziblen Summanden. Die ρ_μ seien paarweise nichtisomorph. Es sei $W \subset V$ ein beliebiger G-invarianter Teilraum. Dann ist W direkte Summe einiger der V_μ.

Beweis. Die Einschränkung ρ_W von ρ nach W zerfällt in irreduzible Darstellungen $\rho_{W,\lambda}$ über Unterräumen W_λ von W mit

$$W = W_1 \oplus \ldots W_l.$$

O.E. sei $l = 1$, $W = W_1$, $\rho_W = \rho_{W,1}$, d.h. ρ_W irreduzibel. Es bezeichne p_μ die Projektion von W nach V_μ. Es können nicht alle ρ_μ Null sein (im Falle $W = 0$ haben wir nichts zu zeigen); da die p_μ G-äquivariant sind, sind sie isomorph oder 0; wegen der Nichtisomorphie der p_μ ist dann genau ein p_μ ein Isomorphismus, etwa p_1. Dann ist für $x \in W$

$$x = \sum_{i=1}^{m} p_\mu x = p_1 x,$$

d.h. $W = V_1$.

Folgerung 2. *Es sei $\varphi : V \to V'$ eine beliebige lineare Abbildung. G sei endlich, $\operatorname{ord} G = N$. Sei*

$$\widetilde{\varphi} = \frac{1}{N}\sum_{g} \rho'(g^{-1})\,\varphi\,\rho(g)$$

Dann: (1) Wenn ρ und ρ' nicht isomorph sind, so ist $\widetilde{\varphi} = 0$.

(2) Für $\rho = \rho'$ ist $\widetilde{\varphi}$ eine Homothetie mit Faktor

$$\frac{1}{n}\operatorname{Sp}\varphi, \quad n = \dim\varphi.$$

Beweis. (1) Wir zeigen, daß $\widetilde{\varphi}$ G-äquivariant ist:

$$\begin{aligned}\rho'(h^{-1})\,\widetilde{\varphi}\rho(h) &= \frac{1}{N}\sum_{g} \rho'(h^{-1})\rho'(g^{-1})\,\varphi\rho(g)\rho(h)\\ &= \frac{1}{N}\sum_{g} \rho'(gh)^{-1}\,\varphi\rho(gh)\\ &= \widetilde{\varphi}\end{aligned}$$

(2) Jetzt folgt (1) aus 6.3.(i), Aussage (2) aus 6.3.(ii), wobei wir den Skalar noch bestimmen müssen:

$$\operatorname{Sp}\widetilde{\varphi} = \frac{1}{N}\sum \operatorname{Sp}(\rho(g^{-1})\varphi\rho(g)) = \operatorname{Sp}\varphi;$$

da bei einer Homothetie mit Faktor c die Spur gerade $c \cdot \dim V = nc$, ist, folgt

$$c = \frac{1}{n}\operatorname{Sp}\varphi.$$

Für die nächsten Folgerungen gehen wir -nach Wahl von Basen- zu den zu ρ' und ρ assoziierten Matrixdarstellungen über, die wir mit S und R bzeichnen. Das Matrixelement $S(g)_{ij}$ wird auch mit $S_{ij}(g)$ bezeichnet, entsprechend für R. Die lineare Abbildung φ werde durch die Matrix $X = (X_{ij})$ gegeben, $\widetilde{\varphi}$ durch $\widetilde{X}$. Dann gilt:

$$\widetilde{X} = \frac{1}{N} \sum_{g} S(g)^{-1} X R(g)$$

$$\begin{aligned}
\widetilde{X}_{im} &= \frac{1}{N} \sum_{g,k,l} S_{i,k}(g^{-1}) X_{k,l} R_{l,m}(g) \\
&= \sum_{k,l} \left(\frac{1}{N} \sum_{g} S_{i,k}(g^{-1}) R_{l,m}(g) \right) X_{k,l}.
\end{aligned}$$

Sind ρ und ρ' nicht isomorph, so ist $\widetilde{X}_{im} = 0$ für alle Wahlen von X_{kl}. Das bedeutet, daß die Koeffizienten 0 sind, also

Folgerung 3. *ρ und ρ' seien nichtisomorphe irreduzible Darstellungen der endlichen Gruppe G und R, S zugehörige Matrixdarstellungen. Dann ist für alle i, k, l, m*

$$\frac{1}{N} \sum_{g} S_{ik}(g^{-1}) R_{lm}(g) = 0.$$

Nun sei $\rho = \rho'$, damit ist $\widetilde{\varphi}$ eine Homothetie mit Faktor $c = \mathrm{Sp}\, \varphi / n$. Dann ist

$$\begin{aligned}
\widetilde{X}_{im} &= \frac{1}{n} (\mathrm{Sp}\, X) \delta_{im} \\
&= \frac{1}{n} \sum_{k,l} \delta_{kl} X_{kl} \delta_{im},
\end{aligned}$$

und es folgt

$$0 = \sum_{k,l} X_{kl} \left(\frac{1}{N} \sum_{g} R_{ik}(g^{-1}) R_{lm}(g) - \frac{1}{n} \delta_{kl} \delta_{im} \right)$$

also, da X_{kl} beliebig sein kann:

Folgerung 4. *Für jede irreduzible Darstellung ρ auf einem n- dimensionalen Vektorraum V der endlichen Gruppe G der Ordnung N gilt:*

$$\begin{aligned}
\frac{1}{N} \sum_{g} R_{ik}(g^{-1}) R_{lm}(g) &= \frac{1}{n} \delta_{kl} \delta_{im} \\
&= \begin{cases} \frac{1}{n} & \text{für } k = l, i = m \\ 0 & \text{sonst} \end{cases}
\end{aligned}$$

3.) Wir interpretieren nun die Folgerungen.

Definition 6.2. *(1) Es seien $\varphi, \psi : G \to \mathbf{C}$ komplexwertige Funktionen. Das Skalarprodukt von φ und ψ ist*

$$\langle \varphi, \psi \rangle \overset{\text{def}}{=} \frac{1}{N} \sum_{g \in G} \overline{\varphi(g)} \psi(g)$$

(2) φ heißt eine Klassenfunktion, wenn φ auf allen Konjugiertenklassen konstant ist:

$$\varphi(h^{-1}gh) = \varphi(g) \quad \textit{für} \quad \textit{alle}\, g, h \in G.$$

Nach Satz 6.1(ii) sind Charaktere also spezielle Klassenfunktionen.

Satz 6.4 *Es seien χ und χ' irreduzible Charaktere. Dann ist für $\chi \neq \chi'$ das Skalarprodukt*

$$\langle \chi, \chi' \rangle = 0;$$

für $\chi = \chi'$ ist

$$\langle \chi, \chi \rangle = 1.$$

Beweis. Es seien χ, χ' Charaktere der irreduziblen Matrixdarstellungen R und S. Dann ist für $g \in G$

$$\chi(g) = \sum_{i=1}^{n} R_{ii}(g),$$

also

$$\langle \chi', \chi \rangle = \frac{1}{N} \sum_{i,j,g} S_{ii}(g^{-1}) R_{jj}(g) = 0$$

nach Folgerung 3. Dabei haben wir $\overline{\chi'(g)} = \chi'(g^{-1})$ ausgenutzt. Entsprechend liefert Folgerung 4 die Beziehung

$$
\begin{aligned}
<\chi,\chi> &= \frac{1}{N}\sum_{i,j,g} R_{ii}(g^{-1})R_{jj}(g) \\
&= \sum_i \frac{1}{n} = 1.
\end{aligned}
$$

Wir sehen also, daß die irreduziblen Charaktere ein Orthonormalsystem im Raum $\mathcal{C}$ der Klassenfunktionen bilden; insbesondere gibt es nur endlich viele solcher Charaktere.

4.) Die Orthogonalitätsrelationen haben in Verbindung mit dem Satz von Maschke wichtige Konsequenzen.

Satz 6.5 *1) Es gibt nur endlich viele Isomorphieklassen irreduzibler Darstellungen von G.*

2) Darstellungen mit demselben Charakter sind isomorph.

Beweis. 1) ist klar: für nichtisomorphe irreduzible Darstellungen sind die Charaktere orthogonal, insbesondere also verschieden. Um 2) zu zeigen, stellen wir nach Satz 5.3 eine Darstellung ρ als direkte Summe

$$\rho = \rho_1 \oplus \cdots \oplus \rho_l$$

irreduzibler Darstellungen dar. Indem wir isomorphe Darstellungen in der Summe zusammenfassen und zu einer isomorphen Darstellung übergehen, können wir

$$\rho = n_1\rho_1 \oplus n_2\rho_2 \oplus \cdots \oplus n_r\rho_r$$

schreiben, wobei die ρ_j die verschiedenen Isomorphieklassen irreduzibler Darstellungen repräsentieren und $n_j \geq 0$ ganz ist, mit

$$n_j\rho_j = \underbrace{\rho_j \oplus \cdots \oplus \rho_j}_{n_j}.$$

Für die Charaktere von ρ und ρ_j folgt

$$\chi = n_1\chi_1 + \cdots + n_r\chi_r.$$

Da die χ_j ein ONS bilden, ist

$$n_j = \langle \chi, \chi_j \rangle$$

eindeutig durch χ bestimmt: hieraus folgt Aussage 2).

Folgerung. *Es sei ρ eine Darstellung mit Charakter χ. Es sei*

$$\rho = n_1\rho_1 \oplus \cdots \oplus n_r\rho_r$$

ihre Zerlegung in irreduzible Darstellungen. Dann ist

$$n_j = \langle \chi, \chi_j \rangle,$$
$$\langle \chi, \chi \rangle = n_1^2 + \cdots + n_r^2.$$

Insbesondere ist ρ genau dann irreduzibel, wenn $< \chi, \chi >= 1$ ist.

(Die letzte Bemerkung ist ein nützliches Irreduzibilitätskriterium!).

5.) Wir benötigen noch Informationen über die Anzahl und die Dimensionen der irreduziblen Darstellungen. Dazu führen wir die sogenannte reguläre Darstellung ein:

Es sei N die Gruppenordnung und V_G der von der Gruppe G aufgespannte N-dimensionale **C**-Vektorraum - G ist also eine Basis von V_G. G operiert auf V_G durch Linksmultiplikation:

$$h(c_1g_1 + \cdots + c_Ng_N) = c_1hg_1 + \cdots + c_Nhg_N.$$

Diese Darstellung heißt reguläre Darstellung von G, wir nennen sie ρ^{reg}; ihren Charakter bezeichnen wir entsprechend mit χ^{reg}. Er ist leicht zu berechnen: ist $g = e$, also das 1-Element der Gruppe, so ist $\rho^{\text{reg}}(e) = id$ und damit

$$\chi^{\text{reg}}(e) = \operatorname{Spur} \rho^{\text{reg}}(e) = N.$$

Ist andererseits $h \neq e$ in G, so ist immer $hg_i \neq g_i$, d.h. die Diagonalelemente der Matrix von $\rho^{\text{reg}}(h)$ bez. der Basis G sind alle 0, und wir haben

$$\chi^{\text{reg}}(h) = 0 \quad \text{für}\, h \neq e.$$

Es folgt

Satz 6.6. *Es seien $\rho_1, \ldots, \rho_r$ die verschiedenen irreduziblen Darstellungen von G, χ_j ihre Charaktere, d_j ihre Dimensionen. Dann gilt:*

$$\begin{aligned} \langle \chi^{\text{reg}}, \chi_j \rangle &= d_j, \\ \rho^{\text{reg}} &= d_1\rho_1 \oplus \cdots \oplus d_r\rho_r \end{aligned}$$

In der Tat ist

$$\begin{aligned} < \chi^{reg}, \chi_j > &= \frac{1}{N} \sum \chi^{reg}(g^{-1})\chi_j(g) \\ &= \frac{1}{N} \cdot N \cdot d_j. \end{aligned}$$

Folgerung. *Mit den Bezeichnungen von eben gilt*

$$d_1^2 + \cdots + d_r^2 = N.$$

Dazu müssen wir in der Formel für ρ^{reg} nur die Dimensionen zählen.

Es bleibt r zu bestimmen, die Anzahl der irreduziblen Darstellungen. Dazu zeigen wir den

Hilfssatz. *Es sei F eine Klassenfunktion, die orthogonal zu allen irreduziblen Charakteren ist. Dann ist $F = 0$.*

Beweis. (1) Wir bilden für eine beliebige Darstellung ρ über V mit Charakter χ den Endomorphismus

$$\varphi = \varphi_\rho = \frac{1}{N} \sum_g \overline{F(g)}\rho(g);$$

φ ist ein Endomorphismus von V.

(2) φ ist G-äquivariant. Denn für $h \in G$ gilt

$$\rho(h^{-1})\,\varphi\,\rho(h) = \frac{1}{N}\sum_g \overline{F(g)}\,\rho(h^{-1}gh)$$

$$= \frac{1}{N}\sum_g \overline{F(h^{-1}gh)}\,\rho(h^{-1}gh)$$

(da F eine Klassenfunktion ist)

$$= \frac{1}{N}\sum_g \overline{F(g)}\,\rho(g).$$

(3) Weiter ist

$$\operatorname{Spur}\varphi = \frac{1}{N}\sum \overline{F(g)}\ \operatorname{Spur}\rho(g) = \langle F, \chi\rangle.$$

Falls ρ irreduzibel ist, ist Spur $\varphi = 0$.

(4) Damit ist, falls ρ eine irreduzible Darstellung ist, $\varphi = c \cdot id$ nach dem Schurschen Lemma, $c = 0$ wegen $\operatorname{Sp}\varphi = 0$. Im Allgemeinfall ist $\rho = n_1\rho_1 \oplus \cdots \oplus n_r\rho_r$ direkte Summe irreduzibler Darstellungen, und man hat

$$\varphi_\rho = n_1\varphi_{\rho_1} + \cdots + n_r\varphi_{\rho_r},$$

also wegen $\varphi_{\rho_j} = 0$ auch jetzt $\varphi_\rho = 0$.

(5) Wir wenden (1) auf die reguläre Darstellung an:

$$0 = \frac{1}{N}\sum_g \overline{F(g)}\,\rho^{reg}(g).$$

Das 1-Element e ist ein Basisvektor von V_G, und ferner ist für jedes $g \in G$

$$\rho^{reg}(g)(e) = g.$$

Somit ist

$$0 = \frac{1}{N}\sum \overline{F(g)}\,\rho^{reg}(g)(e)$$
$$= \frac{1}{N}\sum \overline{F(g)}\cdot g \qquad \in V_G.$$

Da die $g \in G$ eine Basis von V_G bilden, folgt

$$\overline{F(g)} = 0$$

für alle g, wie behauptet.

(6) Wir fassen die wesentlichen Resultate zusammen:

Satz 6.7 *G sei eine Gruppe der Ordnung N, $\rho_1, \ldots$ seien Repräsentanten der Isomorphieklassen irreduzibler Darstellungen von G, d_i die Dimensionen, χ_i die zugehörigen Charaktere. Dann gilt:*

(1) Es gibt genau r Isomorphieklassen irreduzibler Darstellungen, wobei r die Anzahl der Konjugationsklassen von G ist.

(2) Die irreduziblen Charaktere bilden eine ONB des Raumes der Klassenfunktionen.

(3) Es ist $N = d_1^2 + \cdots + d_r^2$.

(4) Jedes d_i ist ein Teiler von N.

Wir haben (2) bewiesen, daraus folgt (1), da der Raum der Klassenfunktionen natürlich die Dimension r hat. (3) ist die Folgerung aus Satz 6.6. Die Teilbarkeitsrelation (4) beweisen wir nicht - sie erfordert mehr Theorie.

§7 Beispiele

1.) Eindimensionale Darstellungen sind natürlich irreduzibel und mit ihrem Charakter identisch: es sind einfach Homomorphismen von G in die multiplikative Gruppe $\mathbf{C}^*$. Da jedes $g \in G$ endliche Ordnung hat ($\leq N = \text{ord } G$), liegt das Bild in der Gruppe der N-ten Einheitswurzeln. Ist G abelsch der Orndung N, so hat G genau N Konjugationsklassen, also genau N irreduzible Darstellungen, die nach Satz 6.7 (3) alle 1-dimensional sein müssen. Umgekehrt: sind alle Darstellungen 1-dimensional, so müssen sie in der Anzahl N vorliegen, damit hat G genau N Konjugationsklassen und ist abelsch.

Satz 7.1 *G ist genau dann abelsch, wenn jede irreduzible Darstellung 1-dimensional ist.*

Ein direkter Beweis eines Teiles von Satz 7.1:

Es sei ρ eine irreduzible Darstellung der (evtl. unendlichen) abelschen Gruppe G über V. Ist $g \in G$ und $V(\alpha) \subset V$ ein Eigenraum von $\rho(g)$ zum Eigenwert α, so ist $V(\alpha)$ G-invariant, da die $\rho(h)$, $h \in G$, mit $\rho(g)$ vertauschen. Somit ist $V(\alpha) = V$ und $\rho(g)$ eine Homothetie. Das gilt für jedes $g \in G$, d.h. alle $\rho(g)$ sind Homothetien. Ein Darstellung durch Homothetien ist aber nur im 1-dimensionalen Fall irreduzibel.

2.) Die irreduziblen Darstellungen (Charaktere) der zyklischen Gruppen sind leicht berechenbar. Sie sind 1-dimensional, und zwar Homomorphismen $\chi : C_n \to S^1 = \{z : |z| = 1\}$. Wegen $x^n = e$ für jedes $x \in C_n$ muß $\chi(z)$ eine n-te Einheitswurzel sein. Setzt man also

$$\zeta = e^{2\pi i/n}$$

und ist x ein Erzeugendes von C_n, so definiert

$$\chi_h(x) = \zeta^h, \quad h = 0, 1, \ldots, n-1$$

sämtliche irreduziblen Charaktere von C_n.

Für eine abelsche Gruppe G definieren wir als Multiplikation zweier irreduzibler Charaktere

$$\chi_1\chi_2(x) \overset{\text{def}}{=} \chi_1(x)\chi_2(x).$$

Satz 7.2 *Durch die obige Formel wird auf der Menge $\widehat{G}$ der irreduziblen Charaktere eine abelsche Gruppenstruktur definiert.*

Definition 7.1. *$\widehat{G}$ heißt die Charaktergruppe von G.*

Jetzt ergibt sich aus der obigen Beschreibung der Charaktere von C_n

Satz 7.3. *$\widehat{C_n}$ ist zyklisch von der Ordnung n, also isomorph zu C_n.*

Diese Aussage gilt allgemeiner:

Satz 7.4. *Die Charaktergruppe jeder endlichen abelschen Gruppe ist zur Gruppe isomorph.*

Zum Beweis bemerken wir, daß die Aussage für zyklische Gruppen richtig ist; da jede endliche abelsche Gruppe direktes Produkt zyklischer Gruppen ist, folgt sie aus dem

Hilfssatz. *Es sei $G = G_1 \times G_2$ ein direktes Produkt endlicher abelscher Gruppen. Dann gilt für die Charaktergruppen*

$$\widehat{G} \cong \widehat{G_1} \times \widehat{G_2}$$

Beweis. Es sei $\chi_i \in \widehat{G_i}$. Definiere $\chi \in \widehat{G}$ durch

$$\chi(a_1, a_2) = \chi_1(a_1)\chi_2(a_2)$$

(mit $a_i \in G_i$). χ ist ein Charakter, die Abbildung

$$\chi_1, \chi_2 \to \chi$$

ist ein Homomorphismus, der offensichtlich injektiv ist, also, da die Gruppen dieselbe Ordnung haben, ein Isomorphismus.

Wir definieren - ähnlich wie bei Vektorräumen - eine “Paarung“ zwischen G und $\widehat{G}$ durch

$$\begin{aligned} G \times \widehat{G} &\xrightarrow{\phi} S^1 \\ g, \chi &\longmapsto \chi(g). \end{aligned}$$

Dann läßt sich ein Homomorphismus von G auf die Bicharaktergruppe $\widehat{\widehat{G}}$ durch

$$\varphi(g)(\chi) = \phi(g, \chi) = \chi(g)$$

definieren. Der Kern von φ ist die Untergruppe

$$G_0 = \{g \in G : \chi(g) = 1 \quad \text{für} \quad \text{alle}\, \chi \in \widehat{G}\}.$$

Jeder Charakter von G faktorisiert sich über einen Charakter von G/G_0 (nach dem üblichen Homomorphiesatz). Daher muß

$$\text{ord } G = \text{ord } \widehat{G} = \text{ord } G/G_0 = \text{ord}\, \widehat{G/G_0}$$

gelten, d.h. $G_0 = \{e\}$, und φ ist damit injektiv, also (wegen Gleichheit der Ordnungen) auch surjektiv. Wir sehen

Satz 7.4' *φ ist ein Isomorphismus von G auf $\widehat{\widehat{G}}$, d.h. ϕ "definiert" G als Charaktergruppe von $\widehat{G}$.*

3.) Wir bestimmen nun die irreduziblen Charaktere von D_3 und D_4.

$$D_3 = \{e, x, x^2, y, yx, yx^2\}$$

- siehe §2 - ; die Konjugationsklassen sind

$$K_1 = \{e\}, \quad K_2 = \{x, x^2\}, \quad K_3 = \{y, yx, yx^2\}.$$

Die einzige Lösung der Gleichung 6.7.3 ist

$$6 = 1^2 + 1^2 + 2^2,$$

demnach gibt es 2 1-dimensionale Charaktere und einen irreduziblen 2- dimensionalen. Die zugehörigen Darstellungen seien ρ_i, $i = 1, 2, 3$. ρ_1 sei die triviale Darstellung mit $\rho(g) = 1$ für alle g; ρ_2 definieren wir durch

$$\rho_2(x) = 1, \quad \rho_2(y) = -1;$$

das ist gerade die Signum-Funktion, wenn wir die $g \in D_3$ als Permutationen der Eckpunkte eines Dreiecks auffassen. Schließlich definieren wir als ρ_3 die 2-dimensionale Darstellung aus §5, Beispiel 1. Das sind alle irreduziblen Darstellungen - bis auf Isomorphie genau; die Werte ihrer Charaktere χ_i notieren wir in der folgenden Charaktertafel:

	(1) e	(2) x	(3) y
χ_1	1	1	1
χ_2	1	1	-1
χ_3	2	-1	0

Wir haben aus jeder Konjugationsklasse einen Repräsentanten angegeben, in Klammern die Elementzahl der Konjugationsklasse darüber notiert, und mit χ_1 den trivialen Charakter bezeichnet. Dementsprechend stehen in der ersten Zeile nur Einsen, in der ersten Spalte mit $\chi_i(e) = d_i$ die Dimensionen der Darstellungen. Aus der Tafel liest man natürlich sofort die Orthogonalitätsrelationen der Charaktere ab. Die entsprechende Tafel für D_4 gebe ich ohne Kommentar an:

	(1) e	(1) x^2	(2) x	(2) y	(2) yx
χ_1	1	1	1	1	1
χ_2	1	1	1	-1	-1
χ_3	1	1	-1	1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

Der Leser sollte sie aus den Informationen über D_n leicht begründen können.

Als letztes Beispiel betrachte ich die Tetraedergruppe T. Sie wird von den Tranformationen x, y_1, y_2, y_3 aus §2 erzeugt; ich gebe nochmal die Matrizen einer Darstellung an und die Wirkung auf die 4 Eckpunkte:

$$x = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

$$y_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

$$y_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

$$y_3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

Die Gleichheitszeichen bedeuten natürlich, daß diese Elemente unter dem fraglichen Isomorphismus sich entsprechen; für Permutationen habe ich die ausführliche Darstellung

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma 1 & \sigma 2 & \dots & \sigma n \end{pmatrix}$$

benutzt. Es gibt genau 4 Konjugationsklassen, repräsentiert durch $e, x, x^2, y = y_1$, von der Länge 1,4,4,3. Dementsprechend gibt es 4 irreduzible Chraktere; die einzige Lösung von

$$d_1^2 + d_2^2 + d_3^2 + d_4^2 = 12 \quad , d_1 = 1,$$

ist

$$1^2 + 1^2 + 1^2 + 3^2 = 12.$$

Die Elemente e, y_1, y_2, y_3 bilden einen Normalteiler V Ordnung 4, isomorph zu $C_2 \times C_2$, und $T/V \cong C_3$. Jeder 1-dimensionale Charakter von C_3, etwa $\overline{\chi}$, liefert zusammen mit $\pi : T \to T/V$ einen 1-dimensionalen Charakter χ von T:

$$T \xrightarrow{\pi} T/V \xrightarrow{\overline{\chi}} \mathbf{C}^*,$$

$\chi = \overline{\chi} \circ \pi$. Damit hat man die 3 1-dimensionalen Darstellungen von T gefunden; die 3-dimensionale Darstellung ist dann die durch die obigen Matrizen beschriebene, deren Spur dann den noch fehlenden Charakter liefert. Insgesamt ergibt sich mit den Ergebnissen über C_3 aus dem Anfang der Paragraphen die folgende Charaktertafel:

	(1) e	(4) x	(4) x^2	(3) y
χ_1	1	1	1	1
χ_2	1	ζ	ζ^2	1
χ_3	1	ζ^2	ζ	1
χ_4	3	0	0	-1

Es ist eine hübsche Übung, aufgrund dieser Charaktertafel die Gruppe zu identifizieren.

Abschließend in diesem Paragraphen machen wir noch eine Bemerkung zu den Darstellungen der Ikosaedergruppe. Wir kennen die triviale Darstellung und die “tautologische“ Darstellung als Untergruppe der $SO(3)$; die Dimensionen sind 1 und 3. Nach Satz 6.7. (3) und (4) hat man für die übrigen Darstellungen

$$60 = 1 + 3^2 + d_3^2 + d_4^2 + d_5^2;$$

die einzigen Lösungen sind 3,4 und 5. Es gibt also noch 3 weitere irreduzible Darstellungen in diesen Dimensionen. Ich gebe sie hier nicht mehr an - siehe [Artin].

§8 Kompakte Gruppen

1.) Wir übertragen eine Reihe wesentlicher Ergebnisse der vorigen Abschnitte auf gewisse unendliche Gruppen, und zwar kompakte Untergruppen von $\mathrm{Gl}(n, \mathbf{C})$ oder $\mathrm{Gl}(n, \mathbf{R})$. Der Fall endlicher Gruppen ist natürlich hierin enthalten: er dient uns als Modell.

Hauptwerkzeug war bisher die Mittelbildung über die endliche Gruppe G, die wir zur Konstruktion G-äquivarianter Objekte verwandt haben. Wir brauchen einen Ersatz hierfür. Ohne Beweis benutze ich

Satz 8.1 *Es sei $C^0(G)$ der Vektorraum der stetigen komplexwertigen Funktionen auf der kompakten Gruppe G. Es existiert genau eine stetige Linearform*

$$I : C^0(G) \to \mathbf{C}$$

mit folgenden Eigenschaften:
(1) $I(1) = 1$.
(2) Bezeichnet f_a, für $f \in C^0(G)$ und $a \in G$, die Funktion

$$f_a(x) \stackrel{def}{=} f(ax),$$

so ist immer

$$I(f_a) = I(f).$$

Darüberhinaus gilt mit $f^*(x) = f(x^{-1})$ und $f^a(x) = f(xa)$:
3) $I(f^a) = I(f) = I(f^*)$.

Erläuterungen (1) Die Stetigkeit von I bezieht sich auf gleichmäßige Konvergenz: ist $f_j \to f$ eine gleichmäßig konvergente Folge, so gilt $I(f_j) \to I(f)$.

(2) $I(1)$ definiert das Volumen von G: er wird also auf 1 normiert.

(3) I heißt invariante Integration auf G; i.a. schreibt man

$$I(f) = \int_G f(x)\,\mathrm{d}\mu(x)$$

und nennt $\mathrm{d}\mu$ ein invariantes oder Haarsches Maß.

(4) Der Satz ist (sinngemäß) auch für sogenannte lokal-kompakte Gruppen richtig: hier trifft man sogar den einfachsten und wichtigsten Spezialfall: $G = \mathbf{R}, +, \ \mathrm{d}\mu = \mathrm{d}x$, also

$$I(f) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x;$$

$C^0(G)$ muß dabei als Vektorraum der stetigen Funktionen mit kompaktem Träger gewählt werden. Eigenschaft (3) des Maßes ist dann nicht mehr garantiert, und die Stetigkeit muß etwas anders definiert werden, ebenso die Normierungsbedingung (1).

Beispiel 1. Eine endliche Gruppe (der Ordnung N) ist (bezüglich der diskreten Topologie) kompakt. Als Maß verwenden wir das “Zählmaß“, dividiert durch N. Jede Funktion f von G nach $\mathbf{C}$ ist stetig, und

$$I(f) = \int_G f(x)\,\mathrm{d}x = \frac{1}{N}\sum_{x\in G} f(x),$$

d.h. $I(f)$ ist der Mittelwert von f über G. Diese Bildung hatten wir in den vorigen Paragraphen benutzt.

Beispiel 2. Es sei $G = S^1 = U(1)$, also

$$S^1 = \{z : z = e^{i\vartheta} : 0 \le \vartheta \le 2\pi\}.$$

Für eine stetige Funktion f auf S^1 definieren wir

$$\widetilde{f}(\vartheta) = f(e^{i\vartheta}).$$

Dann ist $\widetilde{f}(\vartheta)$ eine stetige Funktion auf dem Intervall $[0, 2\pi]$, die wir periodisch zu einer stetigen 2π-periodischen Funktion $\widetilde{f}$ auf $\mathbf{R}$ fortsetzen können. Wir definieren

$$I(f) = \int_{S^1} f(x)\,\mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi} \widetilde{f}(\vartheta)\,\mathrm{d}\vartheta.$$

Der Koeffizient $1/2\pi$ sorgt für die Normierung des Maßes (Bedingung 1). I ist stetig und linear, wie in der Integralrechnung gezeigt wird. Schließlich gilt auch Regel 2 des Satzes: es sei $y = e^{i\alpha}$ fest, dann ist

$$\begin{aligned} I(f_y) = \int_{S^1} f(yx)\,\mathrm{d}x = \int_{S^1} f(e^{i(\alpha+\vartheta)})\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_0^{2\pi} \widetilde{f}(\alpha+\vartheta)\,\mathrm{d}\vartheta \\ &= \frac{1}{2\pi}\int_\alpha^{\alpha+2\pi} \widetilde{f}(\widetilde{\vartheta})\,\mathrm{d}\widetilde{\vartheta} \quad \text{(Substitution!)} \\ &= \frac{1}{2\pi}\int_0^{2\pi} \widetilde{f}(\vartheta)\,\mathrm{d}\vartheta \quad \text{(Periodizität)} \\ &= I(f). \end{aligned}$$

Beispiel 3 $SU(2)$ ist kompakt (homeomorph zu $S^3 \subset \mathbf{R}^4$) und besitzt also ein invariantes Maß. Wir brauchen keine explizite Formel, sondern nur die Existenz!

2.)

Definition 8.1. *Eine Darstellung einer kompakten Gruppe G auf dem endlich-dimensionalen Vektorraum V ist ein* **stetiger** *Homomorphismus*

$$\rho : G \to \mathrm{Gl}(V).$$

ρ heißt irreduzibel, wenn 0 und V die einzigen invarianten Unterräume sind; sonst heißt ρ reduzibel.

Nun können wir den größten Teil der bisherigen Resultate auf stetige Darstellungen kompakter Gruppen übertragen. Dazu zeigen wir zunächst

Hilfssatz 1. *Durch*

$$(x,y) = \int_G \langle gx, gy\rangle\,\mathrm{d}g, \qquad x,y \in V$$

wird ein G-invariantes Skalarprodukt auf V erklärt.

Dabei war $\rho : G \to \mathrm{Gl}(V)$ eine (stetige!) Darstellung, und $\mathrm{d}g$ bezeichnet das Haarsche Maß auf G. Statt $\rho(g)(x)$ haben wir gx geschrieben.

Der Beweis ist mit dem Beweis von §4, Hs 1, identisch. Wir können damit auch die Folgerungen aus diesem Hilfssatz in unsere Situation übernehmen.

Satz 8.2 *Es sei G eine kompakte Gruppe.*
(1) Jede (stetige) Matrixdarstellung von G ist konjugiert zu einer unitären Darstellung.
(2) Jede kompakte Untergruppe $G \subset \mathrm{Gl}(n, \mathbf{C})$ ist zu einer (kompakten) Untergruppe G_0 von $U(n)$ konjugiert.
(3) Alle Elemente einer kompakten Untergruppe sind diagonalisierbar.

Als Folgerung haben wir den Zerlegungssatz

Satz 8.3 *Jede (stetige endlichdimensionale) Darstellung einer kompakten Gruppe G ist direkte Summe irreduzibler Darstellungen.*

Charaktere werden wie früher definiert als

$$\chi_\rho = \mathrm{Spur}\rho, \quad \rho \text{ eine Darstellung},$$

es sind stetige Klassenfunktionen, die der Gleichung

$$\chi_{\rho\oplus\rho'} = \chi_\rho + \chi_{\rho'}$$

genügen. Ist ρ irreduzibel, so heißt χ_ρ irreduzibel.

Das Schursche Lemma (Satz 6.3) hatten wir gleich für beliebige Gruppen bewiesen. Wir führen nun noch das richtige Skalarprodukt ein:

Definition 8.2. *Für stetige Funktionen φ und ψ auf G sei*

$$\langle\varphi, \psi\rangle = \int_G \overline{\varphi(x)}\psi(x)\,\mathrm{d}x.$$

(Dabei ist $\mathrm{d}x$ *das Haarsche Maß von G).*

Jetzt folgt wie in §6 der fundamentale

Satz 8.4 *(1) Die irreduziblen Charaktere sind ein Orthonormalsystem im Raum der Klassenfunktionen.*

(2) Ein Charakter χ ist genau dann irreduzibel, wenn $<\chi,\chi>= 1$ ist.

(3) Zwei Darstellungen ρ und ρ' sind genau dann isomorph, wenn ihre Charaktere gleich sind.

(4) Ist G abelsch, so ist jede irreduzible Darstellung 1-dimensional.

Wir hatten für (4) einen Beweis gegeben - siehe Satz 7.1.

Nicht übertragen können wir die Teile von Satz 6.7, die Anzahlaussagen machen; in der Tat gibt es i.a. unendlich viele Konjugationsklassen und irreduzible Charaktere. Um hier Genaueres zu erfahren, müßten wir wieder die reguläre Darstellung von G definieren, die aber notwendig unendliche Dimension hat. Dazu verweise ich auf die Literatur.

§9 Darstellungen von Drehgruppen

1.) Es sei $U(1) = S^1 = \{z \in \mathbf{C} : |z| = 1\}$. Alle irreduziblen Charaktere sind 1-dimensional; da irreduzible Darstellungen konjugiert zu unitären sind, müssen sie wegen der Kommutativität von $\mathbf{C}^* = \mathrm{Gl}(1,\mathbf{C})$ selbst unitär sein, also sind die irreduziblen Charaktere einfach die (stetigen!) Homomorphismen

$$\chi : U(1) \to U(1).$$

Einige solche kennen wir: $\chi_k(z) = z^k$, für $k \in \mathbf{Z}$.

Satz 9.1 *Die irreduziblen Charktere von $U(1) = S^1$ sind die obigen Homomorphismen*

$$\chi_k(z) = z^k, \quad k = 0, \pm 1, \dots$$

Folgerung. *Die Charaktergruppe von S^1 ist isomorph zu $\mathbf{Z}$.*

Denn, bei der üblichen Multiplikation von Charakteren:

$$(\chi_k \cdot \chi_l)(z) = \chi_k(z)\chi_l(z)$$

ist natürlich

$$\chi_k \cdot \chi_l = \chi_{k+l},$$

und durch $\chi_k \mapsto k$ ist ein Isomorphismus definiert. Anders als bei endlichen abelschen Gruppen ist also die Charaktergruppe nicht isomorph zur Ausgangsgruppe. Die weitere Untersuchung dieser Zusammenhänge führt auf die Pontrjaginsche Dualitätstheorie

Beweis von Satz 9.1. Wir zeigen eine Reihe von Hilfssätzen:

Hilfssatz 1. *Jeder stetige Homomorphismus von* $\mathbf{R}, +$ *in sich ist von der Gestalt*

$$f(t) = at \quad , \ a \in \mathbf{R} \text{ fest}.$$

Beweis. Es sei $a = f(1)$. Ist $n \in \mathbf{N}$, so ist

$$f(n) = f(1 + \cdots + 1) = n \cdot f(1) = na;$$

für $-n \in \mathbf{N}$ gilt ebenfalls

$$f(n) = -f(-n) = -(-n)a = na,$$

und natürlich ist

$$f(0) = 0.$$

Damit gilt $f(t) = a \cdot t$ für alle $t \in \mathbf{Z}$. Ist $s = r/t$ rational, so folgt

$$tf(s) = f(ts) = f(r) = ra,$$

also

$$f(s) = \frac{r}{t} \cdot a = sa.$$

Ist schließlich $t \in \mathbf{R}$ beliebig, so existiert eine Folge $t_j \in \mathbf{R}$ mit $t_j \to t$, und man hat wegen der Stetigkeit

$$f(t) = \lim_{j\to\infty} f(t_j) = a \lim_{j\to\infty} t_j = at.$$

Hilfssatz 2. *Jeder stetige Homomorphismus*

$$\varphi : \mathbf{R}, + \to S^1 = \{z : |z| = 1\} \subset \mathbf{C}$$

hat die Gestalt

$$\varphi(t) = e^{i\alpha t} \quad , \ \alpha \in \mathbf{R}.$$

Beweis. Es sei $\epsilon : \mathbf{R} \to S^1$ durch

$$\epsilon(t) = e^{it}$$

gegeben. Wir haben dann das Diagramm

$$\begin{tikzcd} \mathbf{R} \arrow[dr, "\epsilon"'] & & \mathbf{R} \arrow[dl, "\varphi"] \\ & S^1 & \end{tikzcd}$$

zu untersuchen. Nun kann man genau eine stetige Abbildung $\overline{\varphi} : \mathbf{R} \to \mathbf{R}$ mit $\epsilon \circ \overline{\varphi} = \varphi$ und $\overline{\varphi}(0) = 0$ finden, die also zu einem kommutativen Diagramm

$$\begin{tikzcd} \mathbf{R} \arrow[dr, "\epsilon"'] & & \mathbf{R} \arrow[ll, "\overline{\varphi}"'] \arrow[dl, "\varphi"] \\ & S^1 & \end{tikzcd}$$

führt. Konstruktion von $\overline{\varphi}$: Wir setzen $\overline{\varphi}(0) = 0$ und definieren

$$I = \{t \geq 0 : \overline{\varphi} : [0, t] \to \mathbf{R} \text{ ist stetig }, \overline{\varphi}(0) = 0, \epsilon\overline{\varphi} = \varphi\},$$

als die Menge der positiven t, für die $\overline{\varphi}$ mit den gewünschten Eigenschaften schon definiert ist. Es ist also $0 \in I$; wir zeigen

$$s = \sup I = +\infty.$$

Nehmen wir an, es sei $s \in \mathbf{R}$. Dann ist $z = \varphi(s) \in S^1$. Um $z = e^{i\alpha}$ wählen wir einen Kreisbogen $V = \{e^{i\tau} : \tau - \delta < \alpha < \tau + \delta\}$, der so klein ist, daß

$$\epsilon^{-1}(V) = \bigcup_{k\in\mathbf{Z}} U_\delta(\alpha + k\cdot 2\pi)$$

gilt, wobei die Intervalle $U_k = U_\delta(\alpha + k \cdot 2\pi)$ mit Mittelpunkt $\alpha + k \cdot 2\pi$ paarweise disjunkt sind und durch ϵ umkehrbar stetig auf den Kreisbogen V abgebildet werden. W sei ein so kleines offenes Intervall um s, daß $\varphi(W) \subset V$ gilt. In $W \cap I$ gibt es einen Punkt t, mit $0 \le t \le s$.

Es sei etwa $\overline{\varphi}(t) = U_k$, $\epsilon_k = \epsilon|U_k$ und ϵ_k^{-1} die Umkehrung von ϵ_k. Dann ist für $t' \le t$ mit $t' \in W$ sicher $\overline{\varphi}(t') = \epsilon_k^{-1} \circ \varphi(t')$. Die rechte Seite ist aber auf ganz W definiert und stellt eine stetige Fortsetzung von $\overline{\varphi}$ über s hinaus dar: das widerspricht der Definition von s. – Entsprechend behandelt man die $t \le 0$ und erkennt, daß $\overline{\varphi}$ auf ganz $\mathbf{R}$ mit den gewünschten Eigenschaften existiert.

Die Konstruktion funktioniert für jedes stetige φ. Ist nun φ sogar homomorph, so hat man für $t, s \in \mathbf{R}$

$$\begin{aligned}
\epsilon\left(\overline{\varphi}(s+t) - \overline{\varphi}(s) - \overline{\varphi}(t)\right) &= \frac{\epsilon(\overline{\varphi}(s+t))}{\epsilon(\overline{\varphi}(s))\epsilon(\overline{\varphi}(t))},\\
&= \frac{\varphi(s+t)}{\varphi(s)\varphi(t)}\\
&= 1,
\end{aligned}$$

also gibt es eine ganze, von s, t abhängige Zahl $k(s, t)$ mit

$$\overline{\varphi}(s+t) = \overline{\varphi}(s) + \overline{\varphi}(t) + k(s,t) \cdot 2\pi.$$

Die Funktion k ist auf $\mathbf{R}$ stetig, ganzzahlig, also konstant. Für $s = t = 0$ ist sie 0, also überhaupt.

Ende des Beweises. Es sei $\chi : S^1 \to S^1$ ein Charakter. Mit $\epsilon(t) = \exp 2\pi i t$ ist dann $\chi \circ \epsilon$ ein stetiger Homomorphismus von $\mathbf{R}$ nach S^1, also

$$\chi \circ \epsilon(t) = e^{2\pi i \alpha t} \quad , \text{ für } \alpha \in \mathbf{R}.$$

Damit gilt

$$\chi(e^{2\pi i t}) = e^{2\pi i \alpha t};$$

für $t = 1$ insbesondere

$$1 = \chi(1) = e^{2\pi i\alpha}.$$

Daraus folgt $\alpha = k \in \mathbf{Z}$, somit

$$\chi(z) = \chi(e^{2\pi it}) = e^{2\pi it\cdot k} = z^k.$$

2.) Es sei nun $G = SU(2)$. Wir suchen die irreduziblen Darstellungen. Dazu führen wir die folgenden Vektorräume ein:

$$V_n = \{f \in \mathbf{C}[u, v] : f \text{ homogen vom Grad } n\}.$$

Ein Element $f \in V_n$ sieht also so aus:

$$f = f(u, v) = x_0u^n + x_1u^{n-1}v + \ldots + x_{n-1}uv^{n-1} + x_nv^n,$$

wobei die x_i komplexe Zahlen sind. V_n ist (mit der Addition von Polynomen und Multiplikation mit komplexen Zahlen) ein $n+1$-dimensionaler $\mathbf{C}$- Vektorraum, in dem die Polynome

$$u^n, u^{n-1}v, \ldots, uv^{n-1}, v^n$$

offenbar eine Basis bilden: $V_0 = \mathbf{C}$. Wir definieren eine Darstellung ρ_n von $\mathrm{Gl}(n, \mathbf{C})$ auf V_n durch die Formeln:

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Gl}(n, \mathbf{C})$$

$$\begin{aligned} (u'v') &= (u\ v) \cdot P \\ f = f(u, v) &= x_0u^n + \cdots + x_nv^n \\ (\rho_nP)f &= x_0u'^n + \cdots + x_nv'^n, \end{aligned}$$

also

$$(\rho_nP)u^iv^j = u'^iv'^j, \qquad i + j = n.$$

Für $n = 0$ ist ρ_0 die triviale Darstellung von $\mathrm{Gl}(n, \mathbf{C})$ auf $V_0 = \mathbf{C}$. Für $n = 1$ ist ρ_1 die Standarddarstellung auf dem $\mathbf{C}^2$:

Wir wählen $\mathcal{A} = \{u, v\}$ als Basis von V_1 und sehen, daß die zu ρ_1 gehörige Matrixdarstellung $R_{1,P}$ durch

$$R_{1,P} = A_{\rho_1(P);\mathcal{A},\mathcal{A}} = P$$

gegeben wird. Für $n = 2$ erhält man entsprechend bezüglich der Basis $\mathcal{A} = \{u^2, uv, v^2\}$:

$$R_{2,P} = \begin{bmatrix} a^2 & ab & b^2 \\ 2ac & ad + bc & 2bd \\ c^2 & cd & d^2 \end{bmatrix};$$

in den Spalten stehen die Koordinatenvektoren von $u'^2, u'v', v'^2$ bezüglich $\mathcal{A}$.

Nun ist $SU(2)$ eine Untergruppe von $\mathrm{Gl}(n, \mathbf{C})$; die Einschränkungen von ρ_n auf $SU(2)$ mögen auch mit ρ_n bezeichnet werden.

Satz 9.2 *Die ρ_n, $n = 0, 1, 2, \ldots$ sind die irreduziblen Darstellungen von $G = SU(2)$.*

a) Zum Beweis brauchen wir zunächst

Hilfssatz 3. *Die ρ_n sind irreduzibel.*

Beweis. Es sei

$$T \cong U(1)$$

die Untergruppe der Diagonalmatrizen in $SU(2)$, also

$$T = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix} : |\lambda| = 1 \right\}.$$

Die Einschränkung von ρ_n auf T ist eine Darstellung von T, die vollständig in 1-dimensionale irreduzible Darstellungen zerfällt. Wir werden in Hilfssatz 5 zeigen, daß diese irreduziblen Komponenten paarweise nicht-isomorph sind; es ist

$$\begin{aligned} V_n &= W_0 \oplus \cdots \oplus W_n, \\ W_i &= \mathbf{C} \cdot u^{n-i} v^i \quad , \quad i = 0, \ldots, w, \end{aligned}$$

die irreduzible Zerlegung von ρ_n (eingeschränkt auf T). Nach Folgerung 1 zum Schurschen Lemma ist dann jeder T-invariante Unterraum von V_n eine direkte

Summe einiger oder aller W_i. Das gilt erst recht für jeden $SU(2)$−invarianten Unterraum. Es sei also W ein solcher Unterraum $\neq 0$. W enthält dann ein Monom $u^{n-i}v^i$. Ist dann

$$P = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}$$

eine unitäre Matrix mit $ab \neq 0$, so ist

$$\begin{aligned} P(u^{n-i}v^i) &= (au - \overline{b}v)^{n-i}(bu + \overline{a}v)^i \\ &= a^{n-i}b^i u^n + \dots; \end{aligned}$$

damit muß W also auch u^n enthalten. Weiter ist

$$\begin{aligned} Pu^n &= (au - \overline{b}v)^n \\ &= \sum_{k=0}^{n} \binom{n}{k} a^k (\overline{b})^{n-k} u^k v^{n-k}; \end{aligned}$$

da sämtliche auftretende Koeffizienten $\neq 0$ sind, muß W alle Monome $u^k v^{n-k}$ enthalten, d.h. $W = V_n$.

b) Wir betrachten weiter die Gruppe T. Jede Konjugiertenklasse von $SU(2)$ schneidet T in den beiden Matrizen

$$\begin{pmatrix} \lambda & \\ & \overline{\lambda} \end{pmatrix}, \begin{pmatrix} \overline{\lambda} & \\ & \lambda \end{pmatrix},$$

die nur für $\lambda = \pm 1$ übereinstimmen.

Damit haben wir

Hilfssatz 4. *1) Jede Klassenfunktion ϕ auf $SU(2)$ ist durch ihre Einschränkung auf T eindeutig festgelegt.*

2) Faßt man φ als Funktion von λ auf, so muß die Funktionsgleichung

$$\varphi(\lambda) = \varphi(\overline{\lambda})$$

bzw., mit $\lambda = e^{i\vartheta}$ und φ als Funktion von ϑ,

$$\varphi(\vartheta) = \varphi(-\vartheta)$$

erfüllt sein.

c) Jetzt berechnen wir die Charaktere χ_n der ρ_n. Da es Klassenfunktionen sind, brauchen wir nur ihre Einschränkungen auf T zu kennen, und das ist jeweils der Charakter der Einschränkung von ρ_n auf T. Ist nun $P = \mathrm{diag}(\lambda, \overline{\lambda}) \in T$, so ist

$$Pu^k v^l = \lambda^{k-l} u^k v^l;$$

dabei haben wi $\overline{\lambda} = \lambda^{-1}$ ausgenutzt. P wirkt also als Diagonalmatrix auf V_n mit den Eigenwerten

$$\lambda^n, \lambda^{n-2}, \ldots, \lambda^{-n}$$

und der Spur

$$\chi_n(\lambda) = \sum_{k=-n/2}^{n/2} \lambda^{2k} = \sum_{k=-n/2}^{n/2} e^{2ik\vartheta}.$$

Also

$$\begin{aligned}
\chi_0(\vartheta) &= 1 \\
\chi_1(\vartheta) &= e^{i\vartheta} + e^{-i\vartheta} = 2\cos\vartheta \\
\chi_2(\vartheta) &= e^{2i\vartheta} + 1 + e^{-2i\vartheta} = 1 + 2\cos 2\vartheta,
\end{aligned}$$

usw.

Wir notieren insbesondere

Hilfssatz 5. *Die irreduziblen Komponenten von $\rho_n|T$ aus Hilfssatz 3 sind paarweise nichtisomorph.*

Damit ist Hilfssatz 3 vollständig bewiesen.

d) Jetzt sei χ ein beliebiger irreduzibler Charakter von $SU(2)$. Wir weisen nach, daß χ mit einem χ_n übereinstimmt. Wie im Beweis von Hilfssatz 2 schränken wir die Darstellung ρ, zu der χ gehört, auf T ein; dort zerfällt sie in eine Summe von

Exponentialfunktionen $e^{in\vartheta}$, weil das ja die irreduziblen Charaktere von $U(1)$ sind; damit ist $\chi|T$ eine Summe von Exponentialfunktionen, die zusätzlich gerade sein muß: $\chi(\vartheta) = \chi(-\vartheta)$. Somit kommen in dieser Summe die Terme $e^{in\vartheta}$ und $e^{-in\vartheta}$ mit denselben Koeffizienten vor, und damit kann sie als Linearkombination der χ_n mit rationalen Koeffizienten dargestellt werden:

$$\chi = \sum_{j=0}^{t} r_j \chi_j.$$

Da diese Gleichung auf T gilt, gilt sie nach Hilfssatz 4 auf ganz $SU(2)$. Die Funktionen $\chi, \chi_0, \chi_1, \ldots, \chi_t$ sind also linear abhängig, bilden andererseits ein Orthonormalsystem. Das ist nur möglich, wenn χ eins der χ_j, $0 \le j \le t$, ist.

Damit ist der Satz bewiesen.

3.) Wir untersuchen als nächstes die Gruppe $SO(3)$. Dazu betrachten wir die Konjugationsklassen von $SU(2)$, die bei der Identifikation

$$SU(2) = S^3$$

gerade die Breitenkreise der S^3 sind, also 2-Spären (oder 2 einzelne Punkte). Insbesondere ist der Äquator der S^3 eine solche Konjugationsklasse; er besteht (als Teilmenge von $SU(2)$) aus den Matrizen der Gestalt

$$M = \begin{pmatrix} x & y \\ -\overline{y} & \overline{x} \end{pmatrix}, |x|^2 + |y|^2 = 1,\ x + \overline{x} = 0.$$

Führen wir Real- und Imaginärteile der Koeffizienten ein, so können wir

$$M = \begin{pmatrix} ix_2 & y_1 + iy_2 \\ -y_1 + iy_2 & -ix_2 \end{pmatrix},$$

mit

$$x_2^2 + y_1^2 + y_2^2 = 1,$$

schreiben. Auf dem Äquator $S^2 = \{M\}$ operiert $SU(2)$ also durch Konjugation:

$$M \longmapsto P^{-1}MP$$

mit $M \in S^2$, $P \in SU(2)$. Wir betrachten allgemeiner den reellen Vektorraum H der “schiefhermitischen“ Matrizen mit Spur 0.

$$M = \begin{pmatrix} ix_2 & y_1 + iy_2 \\ -y_1 + iy_2 & -ix_2 \end{pmatrix}, \quad x_2, y_1, y_2 \in \mathbf{R};$$

er hat die Dimension 3. Auf H operiert $SU(2)$ ebenfalls durch Konjugation. In der Tat: M ist genau dann schiefhermitisch, wenn ${}^t\overline{M} = -M$ ist. Dann ist für unitäres P

$$\begin{aligned} {}^t\overline{(P^{-1}MP)} &= {}^t(\overline{P}^{-1}\overline{MP}) \\ &= {}^t\overline{P}{}^t\overline{M}{}^t\overline{P}^{-1} \\ &= P^{-1}(-M)P \\ &= -P^{-1}MP. \end{aligned}$$

und die Spur hat sich nicht geändert.

Darüberhinaus ist die Operation linear:

$$\begin{aligned} P^{-1}(M_1 + M_2)P &= P^{-1}M_1P + P^{-1}M_2P \\ P^{-1}\lambda MP &= \lambda P^{-1}MP, \quad \lambda \in \mathbf{R}. \end{aligned}$$

Also haben wir den

Hilfssatz 6. *Durch $\sigma(P)(M) = P^{-1}MP$ wird eine 3-dimensionale reelle Darstellung von $SU(2)$ definiert.*

Da ferner für jedes P der Isomorphismus $\sigma(P)$ die S^2 in sich überführt, ist $\sigma(P)$ eine Isometrie des Raumes H (bezüglich der euklidischen Norm $\|\cdot\|^2 = x_2^2 + y_1^2 + y_2^2$). In einer zugehörigen ONB, die ich gleich explizit angebe, ist damit $\sigma(P)$ durch eine orthogonale (reelle!) 3×3-Matrix gegeben. Wir wählen als Basis die den Basisvekoren $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ entsprechenden Matrizen

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

Also:

Hilfssatz 7. *σ “ist“ ein Homomorphismus von $SU(2)$ nach $O(3)$.*

“ist“: bedeutet: “definiert in der obigen Basis.“

Es sei wieder $T \subset SU(2)$ die Untergruppe der Diagonalmatrizen. Wir berechnen für $P = \operatorname{diag}(\lambda \ \overline{\lambda}) \in T$:

$$P^{-1}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} P = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$
$$P^{-1}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} P = \begin{pmatrix} 0 & \overline{\lambda}^2 \\ -\lambda^2 & 0 \end{pmatrix}$$
$$P^{-1}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} P = \begin{pmatrix} 0 & i\overline{\lambda}^2 \\ i\lambda^2 & 0 \end{pmatrix}.$$

Das bedeutet, mit $\lambda = e^{i\vartheta}$, für $\sigma(P)$:

$$\sigma(P)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \sigma(P)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \cos 2\vartheta \\ -\sin 2\vartheta \end{pmatrix}$$
$$\sigma(P)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \sin 2\vartheta \\ \cos 2\vartheta \end{pmatrix}.$$

Also: das Bild von $SU(2)$ unter σ enthält die Isotropiegruppe von $(1, 0, 0)$ in S^2, die ja aus den Matrizen der Gestalt

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\vartheta & \sin\vartheta \\ 0 & -\sin\vartheta & \cos\vartheta \end{pmatrix}$$

besteht. Da ferner $SU(2)$ transitiv auf S^2 operiert - S^2 ist ja eine Konjugationsklasse von $SU(2)$ - , folgt

Hilfssatz 8 $\sigma : SU(2) \to SO(3)$ *ist surjektiv.*

Daß wirklich $SO(3)$ als Bild sich ergibt, nicht etwa $O(3)$, ist klar: die Funktion $\det \circ \sigma$ ist auf $SU(2)$ stetig, ganzzahlig und im Punkte E, der Einheitsmatrix, gleich 1.

Hilfssatz 9 Kern $\sigma = \{\pm E\}$.

Zum Beweis werte man einfach die Bedingungen $\sigma(P) = id$ aus: es muß für ein solches P gelten:

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} ;$$

daraus folgt sofort $b = 0$, also $a = \lambda = e^{i\vartheta}$. Ferner muß

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

sein, woraus $\lambda = \overline{\lambda}$ folgt. Wegen $|\lambda| = 1$ ist dann $\lambda = \pm 1$, also $P = \pm E$.

Insgesamt

Satz 9.3 $\sigma : SU(2) \to SO(3)$ *ist ein surjektiver Homomorphismus mit Kern* $Z = \{\pm E\}$*; darüberhinaus ist der Kern Z abgeschlossen, und* σ *ist stetig.*

$SU(2)$ ist also (in der Sprache der Topologie) eine 2-blättrige unverzweigte Überlagerung von $SO(3)$.

Nach dem Homomorphiesatz der Gruppentheorie folgt jetzt leicht

Satz 9.4 *1) Die Darstellungen von* $SU(2)$*, die auf Z trivial sind, sind von der Gestalt* $\rho = \overline{\rho} \circ \sigma$*, wobei* $\overline{\rho}$ *eine beliebige Darstellung von* $SO(3)$ *ist.* $\overline{\rho}$ *ist durch* ρ *eindeutig bestimmt.*
2) ρ *ist genau dann irreduzibel, wenn* $\overline{\rho}$ *es ist.*

Zum Beweis müssen wir uns nur klarmachen, daß ρ genau dann stetig ist, wenn $\overline{\rho}$ es ist. Das folgt aus der Stetigkeit und Abgeschlossenheit des Homomorphismus σ, die sich wiederum aus der Konstruktion von σ ergeben.

Damit kennen wir die irreduziblen Darstellungen von $SO(3)$: sie sind gegeben durch die ρ_j aus Satz 9.2 mit *geradem* j , denn genau für diese ist $\rho_j(-E) = id$. Wir bezeichnen sie mit $\overline{\rho}_j$.

4.) Als letztes in diesem Paragraphen erweitern wir noch die Theorie in einem Punkt: wir wollen Produkte von Darstellungen einführen.

Definition 9.1. *Es seien V und W n- bzw. m- dimensionale K-Vektorräume,* $\mathcal{A} = \{a_1, \ldots, a_n\}$, $\mathcal{B} = \{b_1, \ldots, b_m\}$ *Basen von V bzw. W. Der von der Produktmenge* $\mathcal{A} \times \mathcal{B}$ *erzeugte* $n \cdot m$*-dimensionale Vektorraum heißt Tensorprodukt von V und W und wird mit* $V \otimes W$ *bezeichnet.*

Der Grundkörper K darf beliebig sein. Durch

$$\tau(a_i, b_j) \stackrel{\text{def}}{=} (a_i, b_j) \in V \otimes W$$

wird eine bilineare Abbildung von $V \times W$ nach $V \otimes W$ definiert; Schreibweise: $\tau(x, y) = x \otimes y$. Es gilt also

$$\left(\sum \alpha_i a_i\right) \otimes \left(\sum \beta_j b_j\right) = \sum_{i,j} \alpha_i \beta_j a_i \otimes b_j.$$

Jede bilineare Abbildung

$$\varphi : V \times W \to Z$$

in irgendeinen K-Vektorraum Z faktorisiert sich eindeutig über das kommutative Diagramm

$$\begin{array}{ccc} V \times W & \stackrel{\varphi}{\longrightarrow} & Z \\ \tau \downarrow & \nearrow \overline{\varphi} & \\ V \otimes W & & \end{array}$$

mit einer **linearen** Abbildung $\overline{\varphi}$. Insbesondere hängt $V \otimes W$ “bis auf kanonische Isomorphie genau” nicht von den gewählten Basen ab.

Sind insbesondere $f : V \to V$ und $g : W \to W$ Endomorphismen, so definiert

$$(x, y) \longmapsto f(x) \otimes g(y)$$

eine bilineare Abbildung von $V \times W$ in $V \otimes W$, die sich über eine lineare Abbildung, genannt $f \otimes g$, von $V \otimes W \to V \otimes W$ faktorisiert.

Definition 9.2. *$f \otimes g$ heißt Tensorprodukt der Endomorphismen f und g.*

Es ist leicht, aus den Matrizen für f und g bezüglich gewählter Basen die Matrix des Tensorproduktes auszurechnen: Es sei

$$\begin{aligned} A &= A_{f;\mathcal{A},\mathcal{A}} = (\alpha_{ij}) \\ B &= A_{g;\mathcal{B},\mathcal{B}} = (\beta_{kl}) \\ C &= A_{f \otimes g;\mathcal{C},\mathcal{C}} = (\gamma_{kl;ij}), \\ \text{mit } \mathcal{C} &= \mathcal{A} \times \mathcal{B}; \end{aligned}$$

dann ist

$$\begin{aligned}(f \otimes g)(a_i \otimes b_j) &= f(a_i) \otimes g(b_j)\\ &= \sum \alpha_{ki} \cdot a_k \otimes \sum \beta_{lj} b_l\\ &= \sum_{k,l} \alpha_{ki}\beta_{lj} a_k \otimes b_l,\\ \text{also } \gamma_{kl;ij} &= \alpha_{ki}\beta_{lj}.\end{aligned}$$

Übersichtlich läßt sich C als Matrix von Matrizen schreiben:

$$C = \begin{pmatrix} \alpha_{11}B & \alpha_{12}B & \dots & \alpha_{1n}B \\ \vdots & & & \\ \alpha_{n1}B & \alpha_{n2}B & \dots & \alpha_{nn}B \end{pmatrix} \in K^{nm \times nm}.$$

Definition 9.3. *Die oben eingeführte Matrix heißt Kronecker-Produkt von A und B, in Zeichen:*

$$A \otimes B.$$

Damit hat man

$$A_{f\otimes g;\mathcal{C},\mathcal{C}} = A_{f;\mathcal{A},\mathcal{A}} \otimes \mathcal{A}_{g;\mathcal{B},\mathcal{B}}.$$

Schließlich notieren wir den einfachen

Hilfssatz 10. *Die Spur von $f \otimes g$ ist das Produkt der Spuren von f und von g.*

5.) Kehren wir nun zu Darstellungen zurück!

Definition 9.4. *Es seien ρ_1 und ρ_2 Darstellungen von G auf Vektorräumen V_1 bzw. V_2. Durch*

$$\rho_1 \otimes \rho_2(g) = \rho_1(g) \otimes \rho_2(g)$$

wird eine Darstellung auf $V_1 \otimes V_2$ gegeben: das Tensorprodukt von ρ_1 und ρ_2.

Aus Hilfssatz 10 folgt nun

Satz 9.5 *Der Charakter* $\chi_{\rho_1\otimes\rho_2}$ *ist das Produkt der Charaktere von* ρ_1 *bez.* ρ_2*:*

$$\chi_{\rho_1\otimes\rho_2} = \chi_{\rho_1}\cdot\chi_{\rho_2}.$$

(Alle diese Sätze und Definitionen gelten für stetige Darstellungen komplexer Gruppen, insbesondere also im Falle endlicher Gruppen).

Das Produkt irreduzibler Charaktere ist im allgemeinen nicht irreduzibel. Wir studieren die Situation am Beipsiel der Gruppe $SU(2)$. Die irreduziblen Darstellungen waren in Satz 9.2 aufgezählt worden: für $n = 0, 1, 2, \ldots$ gibt es bis auf Isomorphie genau eine solche Darstellung ρ_n der Dimension $n + 1$; für ihren Charakter χ_n gilt auf der Untergruppe T der Diagonalmatrizen

$$T = \{\mathrm{diag}(\lambda, \overline{\lambda}) : |\lambda| = 1\}$$

die Formel: $\chi_n(\lambda)$ ist eine Summe von Potenzen von λ, deren Summanden die Zeilen des folgenden Dreieckschemas sind:

$$\begin{array}{cccccccccc}
0 & & & & & 1 & & & & \\
1 & & & & \lambda^1 & & \lambda^{-1} & & & \\
2 & & & \lambda^2 & & 1 & & \lambda^{-2} & & \\
3 & & \lambda^3 & & \lambda^1 & & \lambda^{-1} & & \lambda^{-3} & \\
4 & \lambda^4 & & \lambda^2 & & 1 & & \lambda^{-2} & & \lambda^{-4}
\end{array}$$

usw. Also ist z.B. $\chi_2(\lambda) = \lambda^2+1+\lambda^{-2}$. Sind nun χ_n und χ_m irreduzible Charaktere, so gilt auf T

$$\begin{aligned}
\chi_n(\lambda)\chi_m(\lambda) &= (\lambda^{+n} + \ldots + \lambda^{-n})(\lambda^m + \ldots + \lambda^{-m}) \\
&= \chi_{n+m}(\lambda) + \chi_{n+m-2}(\lambda) + \cdots + \chi_{|n-m|}(\lambda),
\end{aligned}$$

wie man unmittelbar nachrechnet. Da wir es mit Klassenfunktionen zu tun haben, gilt dieselbe Beziehung

$$\chi_n\chi_m = \chi_{n+m} + \chi_{n+m-2} + \ldots + \chi_{|n-m|}$$

auf ganz $SU(2)$, und wir haben

Satz 9.6 *(Clebsch/Gordan-Formel). Für das Tensorprodukt der irreduziblen Darstellungen* ρ_n *von* $SU(2)$ *gilt die Zerlegung in irreduzible Summanden*

$$\rho_n \otimes \rho_m = \rho_{n+m} \oplus \rho_{n+m-2} \oplus \ldots \oplus \rho_{|n-m|}.$$

§10 Quantenmechanik und Gruppendarstellungen

1.) Ich beginne mit einer mathematisch besonders einfachen, physikalisch aber besonders wichtigen Situation.

Es sei ein System aus n identischen (d.h. ununterscheidbaren) Teilchen gegeben. Es kann sich um Photonen, Elektronen, Mesonen, α-Teilchen und dergleichen handeln. Der Zustand des Systems wird durch eine **C**-wertige Funktion der Koordinaten $x_1, \ldots, x_n$ beschrieben:

$$\psi(x_1, \ldots, x_n);$$

es steht x_i für die Koordinaten des i-ten Teilchens. Dabei definieren zwei Funktionen denselben Zustand, wenn sie sich nur um einen Skalar λ vom Betrag 1 unterscheiden. - Bei einer Permutation σ der Teilchen ändert sich der Zustand also nicht, und ψ multipliziert sich nur mit einem Faktor $\lambda(\sigma)$:

$$\psi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \lambda(\sigma)\psi(x_1, \ldots, x_n).$$

Offenbar muß $\lambda(\sigma_1\sigma_2) = \lambda(\sigma_1)\lambda(\sigma_2)$ sein, d.h. $\lambda : \mathfrak{S}_n \to S^1$ ist eine 1-dimensionale Darstellung der symmetrischen Gruppe. Wir wissen aus den Übungen, daß es nur zwei solche Darstellungen gibt:

$$\lambda(\sigma) = 1 \text{ für alle } \sigma \text{ oder } \lambda(\sigma) = \operatorname{sign}\sigma.$$

Beide Möglichkeiten werden in der Natur realisiert. Teilchen mit $\lambda(\sigma) = 1$ für alle σ heißen Bosonen, mit $\lambda(\sigma) = \operatorname{sign}\sigma$ Fermionen; der erste Fall ist durch Photonen, α-Teilchen u.a. realisiert, der zweite Fall durch Elektronen, Protonen u.a.

Die Quantenmechanik zeigt, daß die Unterscheidung grundlegend für unser Verständnis des Aufbaus der Materie ist. Die Untersuchung hängt mit Spin-Eigenschaften der Teilchen zusammen.

2.) Als nächstes betrachte ich rotationssymmetrische Probleme der Quantenmechanik. Ein solches liegt z.B. mit der Beschreibung des Wasserstoffatoms vor. Die Zustandsfunktion ψ eines Wasserstoffatoms ist eine von 0 verschiedene Lösung der zeitunabhängigen Schrödingergleichung.

$$(S) \quad H\psi = E\psi.$$

Dabei ist H der Hamilton-Operator.

$$\begin{aligned} H &= -a\Delta - \frac{b}{r}, \\ \Delta &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\ r &= \sqrt{x^2 + y^2 + z^2}, \end{aligned}$$

$a, b > 0$ physikalische Konstanten. Die Gleichung kann explizit gelöst werden, und man erhält folgendes Ergebnis:

Verlangt man, daß ψ quadratintegrierbar ist, so gibt es eine Folge E_n negativer Zahlen

$$\begin{aligned} E_1 < E_2 < E_3 < \cdots \qquad &< 0, \\ \lim E_n &= 0, \end{aligned}$$

für die (S) von 0 verschiedene Lösungen hat. Für festes n ist der Lösungsraum (Eigenraum) V_n genau n^2-dimensional (über $\mathbf{C}$). Die E_n sind die Energien der Zustände; $n = 1, 2, \ldots$ heißt Hauptquantenzahl des Zustandes $\psi \in V_n$, die $\psi \in V_n$ werden stationäre Zustände genannt.

Da das Problem streng rotationssymmetrisch ist, d.h. H ist invariant unter Operationen der Drehgruppe $SO(3)$ - sogar unter $O(3)$ -, operiert $SO(3)$ auf allen Eigenräumen V_n durch die Formel

$$(A\psi)(x,y,z) = \psi(A^{-1}(x,y,z)),$$

$A \in SO(3)$. Man erhält also n^2-dimensionale Darstellungen ρ^n des $SO(3)$ auf den Vektorräumen V_n; diese Darstellungen müssen in irreduzible Darstellungen zerfallen, und zwar wird (mit den Bezeichnungen des vorigen Paragraphen)

$$\rho^n = \overline{\rho}_0 \oplus \overline{\rho}_2 \oplus \cdots \oplus \overline{\rho}_{2l} \oplus \cdots \oplus \overline{\rho}_{2(n-1)}.$$

Die Darstellungsräume, in die V_n zerfällt, sind $2l{+}1$-dimensional:

$$V_n = \bigoplus_{l=0}^{n-1} V_{n,l}.$$

Die Zahl l, welche die irreduzible Darstellung charakterisiert, heißt Bahnquantenzahl (orbitale, azimuthale Quantenzahl).

In der ursprünglichen Fassung der Quantenmechanik ordne man dem Zustand ψ eine Elektronenbahn zu; dann ist $\sqrt{l(l+1)}$ proportional zum Drehimpuls der Bahn.

3.) Bei einem Mehrelektronen-Atom ist in “0-ter Näherung“ die Situation nicht anders als beim Wasserstoffatom: man behandelt es wie mehrere unabhängige 1-Elektronen-Atome. Die Energien der stationären Zustände sind Summen von Energien der einzelnen Komponenten, die Zustände selbst Produkte der Einzelzustände. Wieder hat man Hauptquantenzahlen N und Bahnquantenzahlen L, die die irreduziblen Darstellungen der $SO(3)$ durchnumerieren. Die Energien hängen nur von N, nicht von L ab.

Natürlich ist das eine kümmerliche Näherung an die Realität: im Hamilton-Operator müssen wir noch die Wechselwirkung zwischen den Elektronen berücksichtigen! Das bedeutet Hinzunahme von Termen der Gestalt

$$\text{const} \cdot \frac{1}{d(\mathfrak{x}, \mathfrak{x}')}$$

mit

$$d(\mathfrak{x}, \mathfrak{x}') = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2};$$

das Problem (S) ist dann nicht mehr explizit lösbar, aber die Operation von $SO(3)$ bleibt nützlich: die Eigenräume V_N des Näherungsproblems zerfallen in $SO(3)$-irreduzible Eigenräume $V_{N,L}$; bei Berücksichtigung der Störung "spalten die Energieterme auf", d.h. die Energien der Zustände in einem festen Raum $V_{N,L}$ sind zwar alle gleich, aber sie hängen nun auch von L ab. Die Dimensionen der $V_{N,L}$ werden durch die Darstellungstheorie der $SO(3)$ vorausgesagt!

In einem Mehrelektronensystem operieren neben $SO(3)$ noch weitere Gruppen (z.B. $\mathfrak{S}_n$), deren Darstellungen wiederum Informationen über die Lösungen der Schrödinger-Gleichung liefern - cf. Lehrbücher der Quantenmechanik.

4.) Übergänge zwischen verschiedenen Zuständen eines Ein- oder Mehrelektronensystems führen zur Emission von Lichtquanten; die Bestimmung der Frequenzen und Intensitäten durch Theorie und Messung ist Aufgabe der Spektroskopie - und eine der Wurzeln der Quantenmechanik. Aus physikalischen Gründen sind nur solche Übergänge möglich, etwa von $V_{N,L}$ nach $V_{N',L'}$, bei denen $\overline{\rho}_{2L'}$ ein irreduzibler Summand von $\overline{\rho}_2 \otimes \overline{\rho}_{2L}$ ist. Nach der Clebsch/Gordan-Formel ist aber

$$\overline{\rho}_2 \otimes \overline{\rho}_{2L} = \overline{\rho}_{2L+2} \oplus \overline{\rho}_{2L} \oplus \overline{\rho}_{2L-2}.$$

Damit hat man die spektroskopische Auswahlregel (für die sogenannte Dipolstrahlung):

$$|L' - L| \leq 1.$$

Die Regel ist experimentell gut bestätigt, umso besser, je weniger Elektronen da sind.

5.) Bisher haben wir nur Darstellungen der $SO(3)$ benötigt; die Darstellungstheorie von $SU(2)$ war nur Hilfsmittel für das Studium der othogonalen Gruppe. In der Kenrphysik tritt aber die Gruppe $SU(2)$ direkt auf:

In einem ersten Schritt betrachten die Physiker Proton und Neutron als 2 Zustände eines Elementarteilchens N, des Nukleon. Der Zustandsraum V von N ist dann ein 2-dimensionaler unitärer Vektorraum, auf dem die unitäre Gruppe $SU(2)$ durch die Standarddarstellung ρ_1 operiert. Ein System von n Nukleonen hat dann einen Zustandsraum

$$\underbrace{V \otimes V \otimes \ldots \otimes V}_{n-\text{mal}};$$

$SU(2)$ operiert durch $\rho_1 \otimes \ldots \otimes \rho_1$ hierauf. Dieses Tensorprodukt läßt sich in irreduzible Darstellungen zerlegen, nach der Clebsch-Gordan-Formel ist etwa im einfachsten Fall

$$\rho_1 \otimes \rho_1 = \rho_2 \oplus \rho_0;$$

für mehr Faktoren iteriert man die Formel. Hierin liegt bereits physikalische Information: die Darstellungsräume der irreduziblen Komponenten sind jeweils orthogonal zueinander, und das bedeutet, daß Zustände aus verschiedenen Darstellungsräumen nicht ineinander übergehen (“Gesetz der verbotenen Übergänge“).

Dieselbe Idee wird in einem zweiten Schritt auf 6 weitere Elementarteilchen angewandt: das Λ-Hyperon, das Paar Ξ^+ und Ξ^- sowie das Tripel Σ^+, Σ^0, Σ^-. Ähnlich bezeichnete Teilchen werden als Quantenzustände desselben Teilchens angesehen: die Zustandsräume sind also 1,2 und 3-dimensional, mit Darstellungen ρ_0, ρ_1 und ρ_2 der $SU(2)$. Bei mehreren Teilchen sind die Zustandsräume wieder die Tensorprodukte aus den Räumen der einzelnen Partner, die Darstellung von $SU(2)$ das Tensorprodukt der Komoponenten, das wiederum nach der Clebsch/Gordan-Formel irreduzibel zerlegt wird. Die auftretenden Indizes der Darstellungen heißen Isospin des betreffenden Zustandes: ein Zustand hat den Isospin $j/2$, wenn er im Darstellungsraum von ρ_j liegt. Wieder lassen sich Übergangsverbote aus dieser Theorie ableiten.

In einem letzten Schritt werden alle 8 oben genannten Teilchen einheitlich betrachtet. Das läuft mathematisch gesehen auf folgendes (lösbares) Problem hinaus: Suche eine irreduzible 8-dimensionale Darstellung ω einer kompakten Gruppe G, die $SU(2)$ als Untergruppe enthält, so daß die Beschränkung von ω auf $SU(2)$ zerfällt in

$$\omega/SU(2) = \rho_0 \oplus \rho_1 \oplus \rho_1 \oplus \rho_2.$$

Für G kann man $SU(3)$ wählen, und auch ω ist nicht schwer zu finden. Auf die physikalische Bedeutung der Konstruktionen gehe ich nicht mehr ein, da spätestens hier meine Physikkenntnisse enden.