

2. Teil: Hilbertraum und Quantisierung

Es gibt einen heuristischen und einen axiomatischen Zugang zur Quantenphysik. Hier ist nicht der Platz, die Grundzüge der Quantenmechanik aus Experimenten und physikalischen Untersuchungen spezieller Systeme herzuleiten, zumal das auch nur in eingeschränktem Maße möglich ist. Tatsächlich hat man, motiviert durch experimentelle Erkenntnisse, eine axiomatische Quantentheorie aufgestellt, die sich bewährt hat. Viele Ergebnisse dieser eher mathematischen Theorie konnten nachträglich physikalisch bestätigt werden. Dazu muss noch gesagt werden, dass es auch unterschiedliche axiomatische Ansätze gibt (ich gehe hier z.B. nicht auf die Gelfand-Segal-Konstruktion der Observablen-Algebra als $*$ -Algebra ein), dass aber letzten Endes immer die gleiche Theorie herauskommt. Im Folgenden beschreibe ich den in den Lehrbüchern am häufigsten verwendeten Zugang.

Das klassische Konzept der *Bahn* eines Teilchens wird in der Quantenphysik durch den *zeitabhängigen Zustand* ersetzt, und dieser Quantenzustand wird durch eine Wellenfunktion $\psi(\mathbf{x}, t)$ beschrieben, wie man sie schon aus der Elektrodynamik kennt. Da es jetzt aber um Partikel geht (im Sinne des Teilchen-Welle-Dualismus), spricht man auch von *Materiewellen*. Einfachster Typ einer Welle ist die *ebene, monochromatische Welle*

$$\psi(\mathbf{x}, t) := e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}.$$

Das ist eine Schwingung mit der Frequenz ω , die sich mit konstanter Geschwindigkeit in Richtung des „Wellenvektors“ \mathbf{k} ausbreitet. Jede allgemeine Welle ist Überlagerung von solchen ebenen, monochromatischen Wellen. Jede Frequenz ω liefert über die Gleichung $E = \hbar\omega$ eine wohlbestimmte Teilchenenergie.

Um Teilchen durch Wellen von begrenzter Ausdehnung beschreiben zu können, muss man sogenannte *Wellenpakete* einführen:

$$\psi(\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \int f(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \omega(\mathbf{k})t)} d\mathbf{k}.$$

Das ist eine räumliche Fourier-Transformation, die man sich als Überlagerung eines ganzen Kontinuums von monochromatischen Wellen vorstellen muss. Die Funktion f bestimmt die Gestalt des Wellenpakets. Die Funktion ψ soll zumindest quadratintegrabel sein (was uns zum Hilbertraum-Konzept führt), in gewissen Fällen auch hinreichend oft differenzierbar.

Die Wellenfunktion $\psi(\mathbf{x}, t)$ bezeichnet man als *Wahrscheinlichkeitsamplitude* für den Aufenthalt des betrachteten Teilchens. Die Größe $|\psi(\mathbf{x}, t)|^2$ interpretiert man als *Wahrscheinlichkeitsdichte*.

Nun postuliert man, dass die Wellenfunktion eines Quantensystems seinen dynamischen Zustand bestimmt, dass sich also aus der Kenntnis der Funktion $\mathbf{x} \mapsto \psi(\mathbf{x}, t)$ (zu einem festen Zeitpunkt t) alle Vorhersagen über das künftige dynamische Verhalten ableiten lassen (sofern solche Vorhersagen überhaupt gemacht werden können).

Man muss also eine (Differential-)Gleichung für die Ausbreitung der Welle ψ kennen. Man kann hier Plausibilitätsüberlegungen anstellen, die Gleichung kann aber nicht wirklich deduktiv hergeleitet werden. Die Bewegungsgleichung muss postuliert werden, so wie man das aus anderen Bereichen der Physik kennt (z.B. im Falle der Newtonschen Bewegungsgleichung „Kraft = Masse · Beschleunigung = zeitliche Änderung des Impulses“). Es gibt allerdings ein paar Indizien: Da man für die Wellen ein *Superpositions-Prinzip* haben möchte, muss die Gleichung (homogen-) linear sein. Und damit die spätere Entwicklung eines Systems aus der Kenntnis eines Anfangszustandes hergeleitet werden kann, muss es sich um eine Differentialgleichung erster Ordnung in Bezug auf die Zeit handeln. Schließlich muss eine formale Analogie zu klassischen Bewegungsgleichungen bestehen, das besagt das *Korrespondenzprinzip*:

Da die klassische Physik „makroskopisch richtig“ ist, muss die Quantentheorie für den Grenzfall großer Quantenzahlen asymptotisch in die klassische Theorie übergehen.

Alles zusammen führte schließlich zu der bekannten *Schrödinger-Gleichung* für ein Teilchen mit Masse m in einem Potential V :

$$i \hbar \frac{\partial}{\partial t} \psi(\mathbf{x}, t) = -\frac{\hbar^2}{2m} \Delta \psi(\mathbf{x}, t) + V(\mathbf{x}, t) \psi(\mathbf{x}, t).$$

Dabei ist $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ der Laplace-Operator. Der Operator

$$H := -\frac{\hbar^2}{2m} \Delta + V(\mathbf{x}, t)$$

wird aus Gründen, die weiter unten zur Sprache kommen werden, als *Hamilton-Operator* bezeichnet. Er soll der klassischen Hamilton-Funktion $H = \frac{p^2}{2m} + V$ eines mechanischen Systems entsprechen. Das macht nur Sinn, wenn man den Operator

$$P := -i \hbar \nabla = \left(-i \hbar \frac{\partial}{\partial x}, -i \hbar \frac{\partial}{\partial y}, -i \hbar \frac{\partial}{\partial z} \right)$$

als *Impuls* auffasst.

Wieso werden jetzt klassische Beobachtungsgrößen zu Operatoren? Das hängt mit folgender Feststellung zusammen:

Jede Messung im mikroskopischen Bereich verändert den Zustand des Systems. Deshalb sind die Observablen (beobachtbare Größen) in der Quantenphysik lineare Operatoren auf dem Hilbertraum der Zustände.

Auf den genauen Mechanismus, wie den makroskopischen Größen quantenphysikalische Observablen zuzuordnen sind, gehe ich später noch ein. Der Hamilton-Operator misst die Energie des Systems. Da sich herausgestellt hat, dass nur die Eigenwerte von H als Messergebnisse von Energiezuständen auftreten, fordert man

ganz allgemein, dass die Messergebnisse von Observablen unter deren Eigenwerten zu suchen sind. Da die Messergebnisse reell sein müssen, sollten die Observablen hermitesche Operatoren sein. Kann eine Größe exakt gemessen werden, so müssen die Werte im diskreten Spektrum des entsprechenden Operators liegen. Kann für die Werte nur ein bestimmter Bereich angegeben werden, so liegt dieser im kontinuierlichen Spektrum (zur Spektraltheorie von hermiteschen Operatoren im Hilbertraum vgl. Anhang E).

Definition.

Ein *quantenmechanisches System* ist durch ein Paar (\mathcal{H}, H) gegeben, welches den folgenden vier Axiomen genügt:

1. \mathcal{H} ist ein separabler (komplexer) Vektorraum. Die Zustände des Systems werden durch die 1-dimensionalen Unterräume (sogenannte *Strahlen*) repräsentiert. $\mathbb{P}(\mathcal{H})$ bezeichne die Menge der Zustände.
2. Die *Observablen* des Systems werden durch die (dicht definierten) selbstadjungierten Operatoren auf \mathcal{H} repräsentiert. \mathcal{O} bezeichne die Menge der Observablen.
3. Der *Hamilton-Operator* H ist eine Observable und bestimmt die Dynamik des Systems in folgendem Sinne: Ist $\varphi_0 \in \mathbb{P}(\mathcal{H})$ ein Zustand des Systems, der durch $f_0 \in \mathcal{H}$ mit $\|f_0\| = 1$ repräsentiert wird, so wird die zeitliche Veränderung φ_t des Zustandes repräsentiert durch die eindeutig bestimmte Lösung $f(t)$ der Differentialgleichung

$$f'(t) = -i H(f(t)),$$

mit der Anfangsbedingung $f(0) = f_0$.

4. Sei $\varphi \in \mathbb{P}(\mathcal{H})$ ein Zustand des Systems, der durch $f \in \mathcal{H}$ mit $\|f\| = 1$ repräsentiert wird, sowie T eine Observable mit Spektralschar (E_λ) . Für ein Intervall $J = (a, b]$ ist dann

$$p(\varphi, T, J) := \int_a^b d(\|E_\lambda f\|^2)$$

die Wahrscheinlichkeit dafür, dass ein Messwert der Observablen T im Zustand φ im vorgegebenen Intervall J liegt.

Bemerkungen:

1) Dass man einen separablen Hilbertraum wählt, hat eher praktische Gründe. Die Menge $\mathbb{P}(\mathcal{H})$ der „Strahlen“ ist der dem Vektorraum \mathcal{H} zugeordnete projektive Raum. Da man jeden Zustand $\psi \in \mathcal{H}$ auf die Länge 1 normieren kann, versteht man unter den Strahlen manchmal auch nur Mengen der Gestalt $\{\psi = e^{i\tau} \cdot \psi_0 : \tau \in \mathbb{R}\}$. Völlig eindeutig kann man die Zustände allerdings nicht machen, denn der *Phasenfaktor* $e^{i\tau}$ kann nicht gemessen werden. Leider wird die Theorie unnötig kompliziert, wenn man mit Strahlen an Stelle von Vektoren arbeiten muss. Deshalb

arbeitet man in der Praxis doch mit den Vektoren in \mathcal{H} (und bezeichnet diese Vektoren als Zustände). Die dadurch entstehende Mehrdeutigkeit muss man im Kopf behalten.

2) Den Mathematiker interessieren keine physikalischen Konstanten. Deshalb taucht in der obigen Schrödinger-Gleichung kein \hbar mehr auf.

3) Das Quadrat des Skalarproduktes $(\varphi|\psi)^2$ zweier Elemente $\varphi, \psi \in \mathcal{H}$ interpretiert man als Übergangswahrscheinlichkeit von φ nach ψ . Das Skalarprodukt soll linear im zweiten Faktor und antilinear im ersten Faktor sein (anders als im Anhang E)!

Der Dirac-Kalkül

Nach Dirac versteht man unter einem *Ket* einen Vektor x des Hilbertraums \mathcal{H} , und man schreibt $x = |Name\rangle$. Durch den Namen wird das Ket näher bezeichnet, das ergibt Symbole wie

$$|\psi\rangle, \quad |\varphi\rangle, \quad |\varphi_{n,p}\rangle, \quad |k, j, m+1\rangle \text{ etc.}$$

Ein *Bra* $\langle Name|$ ist eine Linearform auf \mathcal{H} . Auch hier findet man Symbole wie

$$\langle\varphi|, \quad \langle\psi|, \quad \langle\varphi_n| \quad \langle 1| \text{ etc.}$$

Bekanntlich wird jedem Vektor $x \in \mathcal{H}$ eine stetige Linearform λ_x zugeordnet, mit $\lambda_x(y) := (x|y)$. Ist $x = |\psi\rangle$, so wird λ_x mit $\langle\psi|$ bezeichnet. Im Dirac-Kalkül bleiben leider alle Definitionen etwas vage. Es scheint zunächst so, als könne ein Bra eine **beliebige** Linearform sein. Aber nur stetige Linearformen können durch einen Vektor beschrieben werden. Jedes Ket definiert ein Bra, aber nicht umgekehrt (die Delta-Distribution $\psi \mapsto \psi(\mathbf{x}_0)$ ist z.B. ein Bra ohne Ket). Dirac selbst scheint nur Bras von der Gestalt λ_x betrachtet zu haben, er nennt $\langle\psi|$ den zu dem „Ket-Vektor“ $|\psi\rangle$ konjugierten „Bra-Vektor“. Dann ist der Hilbertraum \mathcal{H} der Raum der Ket-Vektoren und der Raum $L(\mathcal{H}, \mathbb{C})$ der stetigen Linearformen auf \mathcal{H} der Raum der Bra-Vektoren.

Ist $c \in \mathbb{C}$, so wird die konjugiert-komplexe Zahl \bar{c} im Dirac-Kalkül mit c^* bezeichnet. Man beachte, dass die Zuordnung $|\psi\rangle \mapsto \langle\psi|$ (also $x \mapsto \lambda_x$) antilinear ist. Das bedeutet: Ist $|\psi\rangle = c \cdot |\varphi\rangle$, so ist $\langle\psi| = c^* \cdot \langle\varphi|$ (denn es ist $\lambda_{cx}(y) = (cx|y) = \bar{c} \cdot (x|y) = \bar{c} \cdot \lambda_x(y)$).

Ist $x = |\varphi\rangle$ und $y = |\psi\rangle$, so wird die Zahl $\lambda_x(y) = (x|y)$ mit $\langle\varphi|\psi\rangle$ bezeichnet („Skalarprodukt“ oder „Bracket“). Klar: Ein Bracket besteht aus einem Bra und einem Ket. Sind $|\varphi\rangle, |\psi\rangle$ zwei Ket-Vektoren, so erklärt Dirac deren *Skalarprodukt* als das Skalarprodukt $\langle\varphi|\psi\rangle$. Das Skalarprodukt erfüllt die üblichen Regeln:

Ist $|\psi_0\rangle = c \cdot |\psi\rangle$, so ist

$$\begin{aligned} \langle\varphi|\psi_0\rangle &= c \cdot \langle\varphi|\psi\rangle \\ \text{und } \langle\psi_0|\varphi\rangle &= c^* \cdot \langle\psi|\varphi\rangle. \end{aligned}$$

Außerdem ist

$$\langle \varphi | \psi \rangle^* = \langle \psi | \varphi \rangle.$$

Sei nun A ein linearer Operator auf \mathcal{H} . Wir wissen (Anhang E, ab Seite 18), dass man sehr sorgfältig zwischen beschränkten und unbeschränkten Operatoren unterscheiden muss. Das geschieht beim Dirac-Kalkül normalerweise leider nicht, obwohl viele Observablen durch unbeschränkte Operatoren beschrieben werden müssen. Im folgenden wollen wir immer stillschweigend voraussetzen, dass die betrachteten Kets im Definitionsbereich der betrachteten Operatoren liegen, und dass außerdem alle vorkommenden Bras stetige Linearformen sind.

$A | \psi \rangle$ steht für das Bild $A(| \psi \rangle)$ des Kets $| \psi \rangle$ unter dem Operator A . Umgekehrt soll $\langle \varphi | A$ die Linearform $\langle \varphi | \circ A$ bezeichnen. Hier ergibt sich ein Problem: Ist A kein beschränkter Operator, so ist $\langle \varphi | A$ kein Bra-Vektor im engeren Sinne. Wir müssen also entweder bei beschränkten Operatoren bleiben (was physikalisch nicht sinnvoll ist) oder doch Kets zulassen, die keine Ket-Vektoren im Diracschen Sinne sind.

Ist $\langle \varphi | = \lambda_x$ und $| \psi \rangle = y$, so ist $\langle \varphi | A = \lambda_x \circ A$ und $A | \psi \rangle = A(y)$. Der Ausdruck

$$\langle \varphi | A | \psi \rangle := \lambda_x \circ A(y)$$

kann also auf zweierlei Weise interpretiert werden. Die Bedeutung ist jedoch in beiden Fällen die gleiche. Es ist $(\langle \varphi | A)(| \psi \rangle) = \langle \varphi | (A | \psi \rangle)$.

Ein spezieller Operator ist $A := | \psi \rangle \langle \varphi |$. Was soll das sein? Ist $x = | \psi \rangle$ und $y = | \varphi \rangle$, so ist $Az = (x \cdot \lambda_y)(z) = (y|z) \cdot x$. Ist $x = y$, so ist $Az = (x|z) \cdot x$ die orthogonale Projektion von z auf die durch x festgelegte Gerade. Ist $| \psi \rangle$ normiert, so schreibt man:

$$P_\psi := | \psi \rangle \langle \psi |.$$

Ist etwas allgemeiner ein ON-System $\mathcal{S} = \{ | \varphi_1 \rangle, \dots, | \varphi_q \rangle \}$ gegeben, so ist

$$P_{\mathcal{S}} := \sum_{i=1}^q | \varphi_i \rangle \langle \varphi_i |$$

die orthogonale Projektion auf den von \mathcal{S} aufgespannten Untervektorraum.

Der Vorteil der Dirac-Notation besteht darin, dass man in einem Ket-Symbol allerlei Bezeichnungen unterbringen kann, z.B. $| n, n+1, 2 \rangle$. Wahrscheinlich kann man auch – wenn man nicht darüber nachdenkt – ganz nett damit rechnen. Sobald man aber über die Bedeutungen genauer nachdenkt, erweist sich diese Notation als schwer durchschaubar, eher schwerfällig und altmodisch. Dank Bourbaki hat sich in der heutigen Mathematik ein ganz anderer Umgang mit Mengen, Vektoren und Abbildungen durchgesetzt.

Ich denke, dass Dirac zu seiner Zeit (1930) an die Matrizenrechnung anknüpfen wollte. Um das zu erläutern, betrachten wir den endlich-dimensionalen Raum \mathbb{C}^n .

Wie üblich sei $\overline{\mathbb{C}}^n$ der \mathbb{C} -Vektorraum, der als abelsche Gruppe mit \mathbb{C}^n übereinstimmt, bei dem aber die Multiplikation mit Skalaren durch $\mu(c, \mathbf{z}) := \bar{c}\mathbf{z}$ erklärt wird. Dann liefert $\varphi(\mathbf{z})(\mathbf{w}) := (\mathbf{z}|\mathbf{w})$ einen Isomorphismus $\varphi : \overline{\mathbb{C}}^n \rightarrow (\mathbb{C}^n)^*$, denn es gilt:

$$\varphi(\mu(c, \mathbf{z}))(\mathbf{w}) = (\bar{c}\mathbf{z}|\mathbf{w}) = c \cdot (\mathbf{z}|\mathbf{w}) = c \cdot \varphi(\mathbf{z})(\mathbf{w}).$$

Damit wird $\mathbf{z} \in \mathbb{C}^n$ mit der Linearform $\mathbf{w} \mapsto \bar{\mathbf{z}} \cdot \mathbf{w}^\top$ identifiziert. Das legt in diesem Falle folgende Interpretation von Ket- und Bra-Vektoren nahe:

$$\langle \mathbf{x} | = \bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n) \quad \text{und} \quad | \mathbf{x} \rangle = \mathbf{x}^\top = \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

sowie $\langle \mathbf{x} | \mathbf{y} \rangle = \bar{\mathbf{x}} \cdot \mathbf{y}^\top = \sum_\nu \bar{x}_\nu y_\nu$. Ein Operator kann im endlich-dimensionalen Fall durch eine (hermitesche) Matrix A beschrieben werden. Dann machen die Bezeichnungen $\langle \mathbf{x} | A$ und $A | \mathbf{x} \rangle$ auch etwas mehr Sinn, es ist

$$\langle \mathbf{x} | A = \bar{\mathbf{x}} \cdot A \quad \text{und} \quad A | \mathbf{x} \rangle = A \cdot \mathbf{x}^\top.$$

Weil A hermitesch ist, ist $\bar{\mathbf{x}} \cdot A = \bar{\mathbf{x}} \cdot \bar{A}^\top = \overline{\mathbf{x} \cdot A^\top}$. Also ergibt sich kein Widerspruch, und insbesondere ist

$$\langle \mathbf{x} | A | \mathbf{y} \rangle = \bar{\mathbf{x}} \cdot A \cdot \mathbf{y}^\top.$$

Doch nun zurück zum Hilbertraum der Quantenphysik!

Der *adjungierte Operator* wird in der Dirac-Notation mit A^\dagger bezeichnet und wie folgt definiert:

Wenn es zu dem (verallgemeinerten) Bra $\langle \varphi | A$ ein konjugiertes Ket gibt, so bezeichnet man dieses mit $A^\dagger | \varphi \rangle$. Üblicherweise wird die Problematik der unbeschränkten Operatoren verschwiegen. Genauer muss man – gemäß Anhang E – folgendermaßen vorgehen:

Sei $\mathcal{D} \subset \mathcal{H}$ der Definitionsbereich von A und $\mathcal{D}^* \subset \mathcal{H}$ die Menge der Ket-Vektoren $|\varphi\rangle$, für die $\langle \varphi | A$ stetig auf \mathcal{D} ist. Nur für solche $|\varphi\rangle$ kann $A^\dagger | \varphi \rangle$ definiert werden. Es gibt dann nämlich ein $|\varphi_A\rangle \in \mathcal{H}$, so dass das (stetige) Bra $\langle \varphi_A |$ eine Fortsetzung von $\langle \varphi | A$ auf ganz \mathcal{H} ist. Man setzt $A^\dagger | \varphi \rangle := |\varphi_A\rangle$.

Behauptung: Für $|\psi\rangle \in \mathcal{D}$ und $|\varphi\rangle \in \mathcal{D}^*$ ist $\langle \psi | A^\dagger | \varphi \rangle = \langle \varphi | A | \psi \rangle^*$.

BEWEIS: Es ist

$$\begin{aligned} \langle \psi | A^\dagger | \varphi \rangle &= \langle \psi | \varphi_A \rangle \\ &= \langle \varphi_A | \psi \rangle^* \\ &= \{ \langle \varphi_A | (|\psi\rangle) \}^* \\ &= \{ \langle \varphi | (A | \psi) \}^* \\ &= \langle \varphi | A | \psi \rangle^*. \end{aligned}$$

In moderner Schreibweise haben wir gezeigt:

$$\lambda_x(A^\dagger y) = (x|A^\dagger y) = (Ax|y) = (y|Ax)^* = \{\lambda_y(Ax)\}^*.$$

Jeder möge selbst entscheiden, was er besser findet. ■

Ein Operator A heißt *hermitesch* oder *selbstadjungiert*, falls $A = A^\dagger$ ist. Diese Definition stimmt mit der von Anhang E überein.

Der *Erwartungswert* $\langle A \rangle_\psi$ einer Observablen A im Zustand $|\psi\rangle$ ist der Mittelwert aller Ergebnisse, die man erhält, wenn man sehr viele Messungen von A an Systemen im Zustand $|\psi\rangle$ vornimmt. Da man bei gegebenem $|\psi\rangle$ die Wahrscheinlichkeiten für das Auftreten aller möglichen Resultate kennt, ist $\langle A \rangle_\psi$ vorhersagbar. Man kann durch Untersuchung des Spektrums von A zeigen, dass $\langle A \rangle_\psi = \langle \psi|A|\psi \rangle$ ist.

Die Größe $(\Delta A)^2 := \langle (A - \langle A \rangle_\psi)^2 \rangle_\psi$ nennt man die *Standardabweichung*. Sie misst die Streuung der Ergebnisse im Vergleich zum Mittelwert. Es ist

$$\Delta A = \sqrt{\langle \psi|(A - \langle A \rangle_\psi)^2|\psi \rangle}.$$

6.1 Satz (Heisenbergsche Unschärferelation). *Wenn P und Q Observablen mit Kommutator $[P, Q] = i$ sind, so ist $\Delta P \cdot \Delta Q \geq \frac{1}{2}$.*

BEWEIS: Sei $|\varphi\rangle$ ein fester normierter Zustand (auf den sich die Erwartungswerte beziehen), $c \in \mathbb{C}$ und $|\varphi\rangle := (Q + icP)|\psi\rangle$. Dann gilt:

$$\begin{aligned} 0 &\leq \langle \varphi|\varphi \rangle = \langle \psi|(Q - icP)(Q + icP)|\psi \rangle \\ &= \langle \psi|Q^2|\psi \rangle + \langle \psi|i c[Q, P]|\psi \rangle + \langle \psi|c^2 P^2|\psi \rangle \\ &= \langle Q^2 \rangle + ic \langle [Q, P] \rangle + c^2 \langle P^2 \rangle \\ &= \langle Q^2 \rangle + c + c^2 \langle P^2 \rangle. \end{aligned}$$

Die Diskriminante des so erhaltenen quadratischen Polynoms in c muss also ≤ 0 sein. Das bedeutet:

$$1 - 4 \langle P^2 \rangle \langle Q^2 \rangle \leq 0, \text{ also } \langle P^2 \rangle \langle Q^2 \rangle \geq \frac{1}{4}.$$

Sei nun $P' := P - \langle P \rangle$ und $Q' := Q - \langle Q \rangle$. Weil ein skalarer Operator $k = k \cdot \text{id}$ mit jedem anderen Operator vertauschbar ist, ist $[P', Q'] = [P, Q] = i$. Daraus folgt, dass auch $\langle (P')^2 \rangle \langle (Q')^2 \rangle \geq \frac{1}{4}$ ist, also $\Delta P \cdot \Delta Q \geq \frac{1}{2}$. ■

Zwei Operatoren heißen *kompatibel*, wenn sie gleichzeitig scharf messbar sind.

6.2 Folgerung. *Sind zwei Operatoren kompatibel, so muss ihr Kommutator verschwinden.*

Unter einem *vollständigen Satz kommutierender Observablen* versteht man ein System von Observablen A, B, C, \dots , so dass gilt:

1. Die Observablen vertauschen paarweise.
2. Es gibt eine (bis auf die Phase) eindeutig bestimmte Hilbertbasis aus gemeinsamen Eigenvektoren dieser Observablen.

Man nimmt an, dass es so etwas immer gibt.

Quantisierung

Einfache Quantensysteme erhält man, indem man entsprechende klassische Systeme nach einem festen Verfahren „quantisiert“. Um das zu verstehen, müssen wir einen kurzen Ausflug in die klassische (makroskopische) Mechanik machen.

Wir beginnen mit der Mechanik eines Massenpunktes. Wir betrachten ein punktförmiges Teilchen mit der Masse m in einem Kraftfeld. Es sei $t \mapsto \mathbf{r}(t)$ die Bahnkurve des Teilchens im \mathbb{R}^3 . Der Tangentialvektor $\mathbf{v}(t) = \mathbf{r}'(t)$ wird als *Geschwindigkeit* des Teilchens bezeichnet, das Produkt $\mathbf{p}(t) := m \cdot \mathbf{v}(t)$ als *Impuls*.

Das Kraftfeld wird durch ein Vektorfeld $\mathbf{F}(\mathbf{x}, t)$ beschrieben. Das *Newton'sche Gesetz der Bewegung* besagt:

$$\mathbf{F}(\mathbf{r}(t), t) = \mathbf{p}'(t) = m \cdot \mathbf{a}(t),$$

wenn $\mathbf{a}(t) = \mathbf{r}''(t)$ die *Beschleunigung* des Teilchens bezeichnet. Ist $\mathbf{F}(\mathbf{x}, t) \equiv \mathbf{0}$ (man spricht dann von einem „freien“ Teilchen), so ist $\mathbf{p}(t)$ konstant. Das ist der *Impulserhaltungssatz*.

Den Vektor $\mathbf{L}(t) := \mathbf{r}(t) \times \mathbf{p}(t)$ nennt man den *Drehimpuls*.

Ist etwa $\mathbf{r}(t) = (r \cos t, r \sin t, 0)$ eine um die z -Achse kreisende Bewegung, so ist

$$\mathbf{L}(t) = (r \cos t, r \sin t, 0) \times (-mr \sin t, mr \cos t, 0) = (0, 0, mr^2).$$

Der Drehimpuls zeigt in diesem Falle in die Richtung der Drehachse, und er wächst quadratisch mit dem Radius der Drehbewegung.

Ist dagegen $\mathbf{r}(t) = \mathbf{x}_0 + t \mathbf{w}$ eine geradlinige Bewegung, so verschwinden Impuls und Drehimpuls.

Den Vektor $\mathbf{N}(t) := \mathbf{r}(t) \times \mathbf{F}(\mathbf{r}(t), t) = \mathbf{r}(t) \times \mathbf{p}'(t)$ nennt man das *Drehmoment* des Teilchens um den Nullpunkt. Dann ist

$$\mathbf{L}'(t) = \mathbf{r}'(t) \times (m \cdot \mathbf{r}'(t)) + \mathbf{r}(t) \times (m \cdot \mathbf{r}''(t)) = \mathbf{N}(t).$$

Ist $\mathbf{N}(t) \equiv \mathbf{0}$, so ist $\mathbf{L}(t)$ konstant.

Im allgemeinen wird der Zustand eines Systems in der klassischen Mechanik durch einen Punkt im *Konfigurationsraum* beschrieben (i.a. ist das eine n -dimensionale differenzierbare Mannigfaltigkeit M). Lokale Koordinaten q_1, \dots, q_n werden als Ortskoordinaten aufgefasst. Im Tangentialbündel $T(M)$ (dem *Phasenraum*) hat

man dann $2n$ Koordinaten $q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n$. Die *Lagrange-Funktion* $L = T - V$ (Differenz aus kinetischer und potentieller Energie) wird als Funktion auf $T(M) \times \mathbb{R}$ aufgefasst. Das *Prinzip der kleinsten Wirkung* besagt, dass sich das System nur entlang von Wegen $\alpha : [t_1, t_2] \rightarrow M$ bewegt, für die das *Wirkungsfunktional* $S[\alpha] := \int_{t_1}^{t_2} L(\alpha(t), \alpha'(t), t) dt$ minimal wird. Daraus ergeben sich die *Euler-Lagrange-Gleichungen*

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i}(q(t), \dot{q}(t), t) \right) - \frac{\partial L}{\partial q_i}(q(t), \dot{q}(t), t) \equiv 0, \quad i = 1, \dots, n.$$

Unter bestimmten Voraussetzungen gibt es einen Diffeomorphismus $D : T(M) \rightarrow T^*(M)$ (die „Legendre-Transformation“), mit $(q, \dot{q}) \mapsto (q, p)$, unter dem sich die Euler-Lagrange-Gleichungen in die *Hamilton-Gleichungen* transformieren:

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \quad \text{und} \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad k = 1, \dots, n.$$

Dabei steht die *Hamilton-Funktion* H für die Energie des Systems.

Eine geeignete Voraussetzung für die Existenz der Legendre-Transformation ist die „Legendre-Bedingung“

$$\det \left(\frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i} \right) \neq 0.$$

In lokalen Koordinaten ist dann

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad \text{und} \quad H(q_1, \dots, q_n, p_1, \dots, p_n, t) = \sum_i p_i \dot{q}_i - L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t).$$

Beispiel.

Betrachtet man einen Massenpunkt in einem Potentialfeld U , so geht es um die Lagrangefunktion

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{m}{2} \|\dot{\mathbf{q}}\|^2 - U(\mathbf{q}).$$

Die Bewegung des Punktes wird durch die Lagrange-Gleichungen beschrieben:

$$m \cdot q_i''(t) = -\frac{\partial U}{\partial q_i}(\mathbf{q}(t)), \quad i = 1, 2, 3.$$

Das sind einfach die Newtonschen Bewegungsgleichungen.

Weiter sind $p_i = m\dot{q}_i$ die Impulskoordinaten, und es ist

$$H(\mathbf{p}, \mathbf{q}) = m \cdot \sum_i (\dot{q}_i)^2 - \left(\frac{m}{2} \|\dot{\mathbf{q}}\|^2 - U(\mathbf{q}) \right) = \frac{1}{2m} \|\mathbf{p}\|^2 + U(\mathbf{q}).$$

Die Mannigfaltigkeit $T^*(M)$ (auch *Impulsraum* genannt) trägt eine sogenannte *symplektische Struktur*, gegeben durch die *Liouville-Form* λ , eine 1-Form auf der Mannigfaltigkeit $T^*(M)$. Ist $\omega \in T_x^*(M)$ und $v \in T_\omega(T^*(M))$, so setzt man $\lambda(v) := \omega(\pi_*v)$, wenn $\pi : T^*(M) \rightarrow M$ die kanonische Projektion ist. In lokalen Koordinaten ist

$$\lambda = \sum_{i=1}^n p_i dq_i.$$

Die auf $T^*(M)$ gegebene 2-Form

$$\omega := -d\lambda = \sum_{i=1}^n dq_i \wedge dp_i$$

ist geschlossen (d.h. $d\omega = 0$) und nicht ausgeartet (d.h., die durch $v \mapsto (w \mapsto \omega_x(v, w))$ gegebene lineare Abbildung $T_x(T^*M) \rightarrow T_x^*(T^*M)$ ist für jedes $x \in T^*M$ ein Isomorphismus. Wir kennen diese Situation von symmetrischen Bilinearformen, hier liegt nun eine alternierende Bilinearform vor.

Ist f eine differenzierbare Funktion auf T^*M , so gibt es ein Vektorfeld $\nabla^a f$ auf T^*M , so dass $\omega(\nabla^a f, \xi) = df(\xi)$ für alle Vektorfelder ξ auf T^*M gilt. Wir können $\nabla^a f$ als „alternierenden Gradienten“ von f bezeichnen. Man kann nachrechnen, dass folgendes gilt:

$$\nabla^a f = \sum_{\nu=1}^n \frac{\partial f}{\partial p_\nu} \frac{\partial}{\partial q_\nu} - \sum_{\nu=1}^n \frac{\partial f}{\partial q_\nu} \frac{\partial}{\partial p_\nu}.$$

Definition.

Die Funktion $\{f, g\} := -(\nabla^a f)(g)$ nennt man die *Poissonklammer* von f und g .

Offensichtlich ist

$$\begin{aligned} \omega(\nabla^a f, \nabla^a g) &= \sum_{\nu=1}^n \frac{\partial g}{\partial p_\nu} \cdot \omega(\nabla^a f, \frac{\partial}{\partial q_\nu}) - \sum_{\nu=1}^n \frac{\partial g}{\partial q_\nu} \cdot \omega(\nabla^a f, \frac{\partial}{\partial p_\nu}) \\ &= \sum_{\nu=1}^n \frac{\partial f}{\partial q_\nu} \cdot \frac{\partial g}{\partial p_\nu} - \sum_{\nu=1}^n \frac{\partial f}{\partial p_\nu} \cdot \frac{\partial g}{\partial q_\nu} \\ &= \{f, g\}. \end{aligned}$$

Man kann zeigen: $\mathcal{C}^\infty(T^*M)$ wird mit der Poissonklammer zu einer Liealgebra. Und die Hamiltonschen Gleichungen bekommen die folgende Gestalt:

$$\dot{q}_i = \{q_i, H\} \quad \text{und} \quad \dot{p}_i = \{p_i, H\}.$$

Jetzt können wir sagen, was man unter der *kanonischen Quantisierung* eines klassischen Systems versteht.

Gegeben sei ein Hamiltonsches System (P, H) , mit Impulsraum P und Hamiltonfunktion H , sowie eine Menge $\mathcal{A} \subset \mathcal{C}^\infty(P; \mathbb{C})$ von klassischen Observablen. H gehöre zu den Observablen. Eine *Quantisierung* dieses Systems ist eine Abbildung $\varrho : \mathcal{A} \rightarrow \mathcal{O}$ in die Menge der Observablen auf einem Hilbertraum \mathcal{H} , so dass gilt:

1. $[\varrho(f), \varrho(g)] = i \varrho(\{f, g\})$ für $f, g \in \mathcal{A}$ mit $\{f, g\} \in \mathcal{A}$.
2. $\varrho(1) = \text{id}_{\mathcal{H}}$, falls $1 \in \mathcal{A}$.

Speziell gilt dann für die „kanonischen Koordinaten“ q_i und p_i und die zugeordneten Operatoren $Q_i := \varrho(q_i)$ und $P_i := \varrho(p_i)$:

$$[Q_i, Q_j] = [P_i, P_j] = 0 \quad \text{und} \quad [Q_i, P_j] = i \delta_{ij}.$$

Das sind die *kanonischen Vertauschungsregeln*, die schliesslich zur Heisenbergschen Unschärferelation führen.

Es gibt verschiedene Möglichkeiten, die Operatoren Q_i und P_i zu definieren. Üblich ist die folgende *kanonische Quantisierung*:

Als Hilbertraum nimmt man den Raum \mathcal{H} der quadratintegrierbaren komplexwertigen Funktionen in den Variablen q_1, \dots, q_n , mit dem Skalarprodukt

$$\langle \varphi | \psi \rangle := \int \bar{\varphi} \psi d\mathbf{q}.$$

Der Operator Q_i ist gegeben als Multiplikation mit der Variablen q_i , der Operator P_i ist definiert durch

$$P_i := -i \frac{\partial}{\partial q_i}.$$

Die Motivation dafür ergibt sich aus der Theorie der Fouriertransformationen, auf die Details kann ich hier nicht eingehen. Offensichtlich gelten bei dieser Festlegung die kanonischen Vertauschungsregeln.

Der Hamiltonoperator wird nun definiert als

$$\hat{H} := H(Q_1, \dots, Q_n, P_1, \dots, P_n).$$

Wie man sieht, kann \hat{H} nicht immer eindeutig definiert sein, weil man ja nicht-kommutierende Operatoren an die Stelle der kommutierenden Koordinaten q_i und p_i einsetzt. In einfachen Fällen, etwa bei der Quantisierung eines Massenpunktes im Potentialfeld, tritt diese Schwierigkeit nicht auf.

Wir können jetzt auch den Drehimpuls quantisieren. Wir schreiben $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{p} = (p_1, p_2, p_3)$ und $\mathbf{L} = (L_1, L_2, L_3)$. Die Gleichung $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ liefert dann

$$\begin{aligned} L_1 &= x_2 \cdot p_3 - x_3 \cdot p_2, \\ L_2 &= x_3 \cdot p_1 - x_1 \cdot p_3 \\ \text{und} \quad L_3 &= x_1 \cdot p_2 - x_2 \cdot p_1. \end{aligned}$$

Bei der kanonischen Quantisierung benutzen wir die Übergänge

$$x_1 \mapsto x_1 \cdot, \quad x_2 \mapsto x_2 \cdot \quad \text{und} \quad x_3 \mapsto x_3 \cdot,$$

sowie

$$p_1 \mapsto -i \frac{\partial}{\partial x_1}, \quad p_2 \mapsto -i \frac{\partial}{\partial x_2} \quad \text{und} \quad p_3 \mapsto -i \frac{\partial}{\partial x_3}.$$

Aus den Komponenten des Drehimpulses werden damit die Operatoren

$$\begin{aligned} \widehat{L}_1 &= -i \left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2} \right), \\ \widehat{L}_2 &= -i \left(x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3} \right) \\ \text{und } \widehat{L}_3 &= -i \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \right). \end{aligned}$$

Man kann nun die Poisson-Klammern der L_i ausrechnen. Zunächst ist

$$\{x_i p_j, x_k p_l\} = \delta_{il} p_j x_k - \delta_{jk} x_i p_l.$$

Ist nun (i, j, k) eine zyklische Vertauschung von $(1, 2, 3)$, so ist

$$\{L_i, L_j\} = \{x_j p_k - x_k p_j, x_k p_i - x_i p_k\} = x_i p_j - x_j p_i = L_k.$$

Daraus folgt:

$$[\widehat{L}_i, \widehat{L}_j] = i \widehat{L}_k, \quad \text{für jede zyklische Vertauschung } (i, j, k).$$

Taucht irgendwo ein System von drei Operatoren auf, das diesen Vertauschungsregeln genügt, so spricht man von *verallgemeinerten Drehimpuls-Operatoren* und bezeichnet die Operatoren mit J_1, J_2, J_3 . Der Operator $J^2 := J_1^2 + J_2^2 + J_3^2$ vertauscht mit allen drei J_i und spielt daher eine besondere Rolle.¹

Man führt auch noch die sogenannten *Leiteroperatoren* ein,

$$J_+ := J_1 + i J_2 \quad \text{und} \quad J_- := J_1 - i J_2.$$

Sie sind nicht hermitesch, sondern zueinander konjugiert. Bei Mehrteilchensystemen dienen sie als Erzeugungs- und Vernichtungsoperatoren. Es ist $[J_+, J_-] = 2J_3$, $[J_+, J_3] = -J_+$ und $[J_-, J_3] = J_-$.

Weil J^2 und J_3 vertauschen, gibt es eine ON-Basis aus gemeinsamen Eigenvektoren. Sei nun $|\psi\rangle$ ein gemeinsamer Eigenzustand von J^2 (zu einem Eigenwert λ) und von J_3 (zu einem – eventuell von λ verschiedenen – Eigenwert m). Man schreibt dann gerne auch $|\lambda, m\rangle$. Es ist

¹ J^2 ist ein sogenannter *Casimir-Operator*, der nicht zur Liealgebra der Observablen gehört, sondern zur einhüllenden Algebra (vgl. Aufgaben-Teil) der Liealgebra.

$$J^2(J_+ | \lambda, m \rangle) = J_+(J^2 | \lambda, m \rangle) = \lambda(J_+ | \lambda, m \rangle).$$

Weiter ist $J_3 J_+ = J_+ J_3 + J_+ = J_+ (J_3 + \text{id})$ und daher

$$J_3(J_+ | \lambda, m \rangle) = J_+ (J_3 + \text{id}) | \lambda, m \rangle = (m + 1) J_+ | \lambda, m \rangle.$$

Das bedeutet, dass $J_+ | \lambda, m \rangle$ immer noch ein gemeinsamer Eigenzustand von J^2 und J_3 ist, jetzt aber zu den Eigenwerten λ und $m + 1$.

Analog ist $J_- | \lambda, m \rangle$ gemeinsamer Eigenzustand von J^2 und J_3 zu den Eigenwerten λ und $m - 1$. Diese Eigenschaften führten zu der Bezeichnung „Leiteroperatoren“. Die Operatoren führen jeweils in einen höheren oder niedrigeren Quantenzustand über. Da es nur endlich viele Quantenzustände gibt, bewirken die Leiteroperatoren am Ende der „Leiter“ Erzeugung oder Vernichtung.

Die 2. Quantisierung

Sind H_1, H_2 zwei Hilberträume, so wird das Tensorprodukt $H_1 \otimes H_2$ zu einem Prähilbertraum,² mit dem Skalarprodukt

$$(x_1 \otimes x_2 | y_1 \otimes y_2) := (x_1 | y_1) \cdot (x_2 | y_2).$$

Die Vervollständigung von $H_1 \otimes H_2$ nennt man das hilbertsche Tensorprodukt von H_1 und H_2 . Es wird mit $H_1 \widehat{\otimes} H_2$ bezeichnet.

$S^n(H)$ und $A^n(H)$ seien die symmetrischen und antisymmetrischen Tensorpotenzen eines Hilbertraumes H . Jeder Tensor aus $T^n(H)$ kann in einen symmetrischen und einen antisymmetrischen Anteil zerlegt werden. Auch hier muss man wieder zur Vervollständigung übergehen, um Hilberträume zu erhalten.

Sei nun \mathcal{H} der Hilbertraum der Einteilchenzustände einer bestimmten Teilchensorte. Handelt es sich um Bosonen, so ist $S^n(\mathcal{H})$ der „Bosonen-Fock-Raum“ der n -Teilchensysteme. Handelt es sich um Fermionen, so ist $A^n(\mathcal{H})$ der „Fermionen-Fock-Raum“ der n -Teilchensysteme.

Um das starre Schema von n -Teilchensystemen verlassen zu können, führt man den „Vakuum-Zustand“ $|0\rangle$ und Erzeugungs- und Vernichtungsoperatoren ein, die zur Erhöhung bzw. Erniedrigung der Teilchenzahl führen. Im Falle von Bosonen kommutieren diese Operatoren, bei Fermionen antikommutieren sie. Wechselwirkungen werden durch Vernichtung und Erzeugung beschrieben, so wie z.B. der Farbaustausch von Quarks mit Hilfe von Gluonen.

Für eine adäquate Beschreibung dieser Theorie sollte man Quantenfeldtheorie benutzen. Weil dann die Quantisierung von Feldern eine Rolle spielt, hat man – etwas missverständlich – den Begriff von der 2. Quantisierung geprägt. Es ist leider unmöglich, hier auf die Einzelheiten einzugehen.

²Das Tensorprodukt unendlich-dimensionaler Vektorräume muss etwas anders als im endlich-dimensionalen Fall erklärt werden. Wir können hier nicht näher darauf eingehen.