

7. MATHEMATISCH-PHYSIKALISCHER ANHANG

Zu 2.1: Relativität der Gleichzeitigkeit

Betrachten wir das anschaulich Beispiel eines Eisenbahnzugs im Detail (Die Bezeichnungen sind jeweils für vom Zug aus gemessene Größen einfache Buchstaben (z.B. x, t), für vom Bahndamm aus gemessene Größen dieselben Buchstaben „gestrichen“ (z.B. x', t'):

Sei x' der Abstand¹ zwischen A und B, v ($= v'$) die Geschwindigkeit des Zugs, c die Lichtgeschwindigkeit und $\Delta t'$ der Zeitunterschied zwischen der Ankunft des Blitzes bei A und der Ankunft des Blitzes bei B – alles vom Bahndamm aus gemessen. Der Lichtweg x_B' bis zu B ist etwas kürzer als der Abstand $x'/2$, nämlich um die Strecke kürzer, die der Zug in der Zeit α_B' weiterfährt. Dabei ist α_B' die Zeit, die das Licht für den Weg x_B' braucht. Es gilt also

$$x_B' = c \Delta \tau_B'; \quad x_B' = x'/2 - v \Delta \tau_B';$$

daraus folgt

$$\Delta \tau_B' = x'/(2(c+v)).$$

Entsprechend kann man die Zeit berechnen, die das Licht zum Gerät A braucht:

$$\Delta \tau_A' = x'/(2(c-v));$$

Der Zeitunterschied am Bahndamm zwischen den beiden im Zug gleichzeitigen Ereignissen ist also

$$\Delta t' = \Delta \tau_A' - \Delta \tau_B' = x' v / (c^2 - v^2). \quad (1)$$

Dieser Zeitunterschied wäre im Fall eines wirklichen Zuges nicht zu bemerken: Sei z. B. $x = 100$ m und $v = 252$ km/h $= 70$ m/s; es ist $c = 3 \cdot 10^8$ m/sec, also $c^2 = 10^{17}$ m²/sec². Dann ist

$$\Delta t' = x' v / (c^2 - v^2) \approx 10^{-13} \text{ sec.}$$

(Das ist die Zeit, in der ein Lichtsignal eine Strecke von $30 \mu\text{m}$ zurücklegt!)

Zu 2.2: Lorentz-Kontraktion

Nach der Speziellen Relativitätstheorie erscheinen Längen im bewegten System kürzer als im ruhenden. Quantitativ läßt sich diese „Lorentz-Kontraktion“ berechnen, wie wir am eben behandelten Beispiel des Eisenbahnzugs zeigen wollen. Die Länge einer Strecke am Bahndamm werde vom Zug aus dadurch festgelegt, daß die beiden Apparate bei A und B *zugleich* Marken auf dem Gleis anbringen. Der Abstand dieser Marken ist – vom Zug aus gemessen – definitionsgemäß gleich dem Abstand der Geräte A und B:

Sei x' der Abstand der Geräte A und B, vom Bahndamm aus betrachtet; und sei m' der Abstand der beiden Marken auf dem Gleis. Die vordere Marke wird um die Zeit $\Delta t'$ später angebracht als die hintere, und dann ist der Zug schon die Strecke $v \Delta t'$ weitergefahren:

$$m' = x' + v \Delta t' = x' + (x' v^2) / (c^2 - v^2) \quad \text{nach (1)}$$

also, umgerechnet,

$$m' = x' / (1 - (v/c)^2). \quad (2)$$

Nun ist die Beziehung zwischen dem Abstand x im Zug und der entsprechenden Länge x' am Bahndamm dieselbe wie die zwischen dem Abstand m' am Bahndamm und der entsprechenden Länge x im Zug: Jeweils wird der Abstand zwischen zwei zueinander ruhenden Objekten gemessen von einem System aus, das sich gegen sie mit der Geschwindigkeit v parallel zur Verbindungslinie bewegt (alle Relativbewegungen sind ja nach Voraussetzung gleichberechtigt). Die im

¹ gemessen im Bahndamm-System, wie man jetzt dazusagen muß. Es wird unten gezeigt, daß die Länge im Zug-System eine andere ist..

bewegten System festgestellte (Lorentzkontrahierte) Länge sei eine Funktion L der „Ruhlänge“; dann gilt:

$$m' = L(x); x = L(x').$$

Also

$$m' = L(L(x')) = x' / (1 - (v/c)^2) \quad \text{nach (2).}$$

Denken wir uns nun x in zwei Hälften aufgeteilt: die Hälften werden gleichmäßig Lorentz-kontrahiert und müssen zusammen wieder das Lorentz-kontrahierte Ganze ergeben

$$m' = L(x) = 2 \cdot L(x/2);$$

oder allgemein, bei Teilung in n gleiche Teile:

$$L(x) = n \cdot L(x/n)$$

also:

$$L(x/n) = 1/n \cdot L(x).$$

Das heißt, L ist linear (ohne konstanten Summanden), also

$$m' = \alpha \cdot x$$

und daher

$$m' = \alpha^2 \cdot x^2$$

d. h. $\alpha = 1/\sqrt{1-(v/c)^2}$ ist der Faktor der Lorentz-Kontraktion. Im obigen Beispiel, $v = 252$ km/h, wäre die Lorentz-Kontraktion ein Teil entsprechend $\alpha - 1 = 2 \cdot 10^{-14}$ der Länge – also unmeßbar klein.

Zu 2.3: Trägheit der Energie

Um die Trägheit der Energie (d.h. die Formel $E = m \cdot c^2$) aus der Speziellen Relativitätstheorie abzuleiten, müssen wir berechnen, wie sich Geschwindigkeiten addieren. Wegen der Lorentz-Kontraktion und der Zeit-Dilatation ergibt das relativ komplizierte Formeln, deren Ableitung wir hier nicht vorführen können. Betrachten wir dann *inelastische* Stöße von zwei Körpern (z. B. Plastinkugeln) in verschiedenen Inertialsystemen, und bedenken wir, daß der Impuls bei solchen Stößen erhalten bleibt, dann folgt², daß der Impuls p so von der Geschwindigkeit v abhängt:

$p = m_0 \cdot v / \sqrt{1-(v/c)^2} = (m_0 + \Delta m) \cdot v = m \cdot v$, wobei m_0 die Masse des ruhenden Körpers („Ruhmasse“) ist. Δm ist die zusätzliche Masse, die nur dem bewegten Körper zukommt. Für kleine Geschwindigkeiten v kann man genähert schreiben:

$$m = m_0 + m_0 / 2 \cdot (v^2 / c^2)$$

so daß sich der Zuwachs der Masse, Δm , durch die kinetische Energie $T = m \cdot v^2 / 2$ ausdrücken läßt als:

$$\Delta m = T / c^2.$$

Wegen der Erhaltung von Energie und Impuls bleibt der Massenzuwachs Δm derselbe, wenn die kinetische Energie beim Stoß in Wärme des Körpers verwandelt wird; er hängt überhaupt nur vom Energie-Zuwachs ΔE des Körpers ab:

$$\Delta m = \Delta E / c^2.$$

Aus der vollständigen Verwandlung massiver Teilchen in Bewegungsenergie und Strahlung, bzw. umgekehrt, ist die Formel sogar für die *Gesamtmasse* gut bestätigt. Es gilt also ganz allgemein die Gleichheit (bis auf Maßfaktoren) von Masse und Energie, also die berühmte Einsteinsche Formel:

$$E = m \cdot c^2.$$

Zu 2.5: Poincaré-Transformationen

Die Gruppe der Symmetrie-Transformationen der Speziellen Relativitätstheorie heißt Poincaré-Gruppe. Sie besteht aus Translationen, Drehungen und Übergängen zu bewegten Bezugssystemen.

² M. Born (1964) rechnet das sehr verständlich vor; eine allgemeinere Ableitung referiert er nach Einstein auf S. 244ff.

men („boosts“). Ihre Transformationen wirken auf Funktionen der Raum- und Zeit-Koordinaten. Und zwar wirken die Transformationen zunächst auf die Koordinaten selbst: sie ordnen einem Ereignis neue Koordinaten zu. Dafür gibt es zwei mögliche physikalische Interpretationen: Die „passive“ Interpretation sieht die Transformation als Beschreibung der gleichen wirklichen Ereignisse in einem neuen Koordinatensystem, so wie wir es an den Beispielen vom „Zug“ und „Bahndamm“ gezeigt haben (vgl. Abb. 2.5); die „aktive“ Interpretation sieht in der Transformation eine Verschiebung der wirklichen Ereignisse an andere Orte und zu anderen Zeiten. Bei einer rein formalistischen Sicht macht das keinen Unterschied³. Beschreiben wir die Transformation (sie heie L) neutral: sie ordnet den Koordinaten (x_0, t_0) eines Ereignisses („Urbild“) neue Koordinaten (x_0', t_0') („Bild“) zu. Fur die neuen Koordinaten gilt also: $(x_0', t_0') = L(x_0, t_0)$; und entsprechend bezeichnet man die umgekehrte Transformation (von Bild zu Urbild; angenommen es gibt sie) mit L^{-1} :

$$(x_0, t_0) = L^{-1}(x_0', t_0')$$

Nehmen wir nun eine beliebige Funktion der Ereignis-Koordinaten, $F(x, t)$. Die Transformation der Koordinaten kann man zur Definition einer neuen Funktion benutzen, indem man die Funktionswerte mit den Koordinatenwerten „mitnimmt“: Die neue Funktion hat an einem Bildpunkt (x', t') den Wert, den die ursprungliche Funktion am Urbild-Punkt $L^{-1}(x, t)$ hatte. Die neue Funktion F' lat sich also schreiben (als Funktion der allgemeinen Koordinaten x, t – wir lassen ihre „Striche“ jetzt weg)

$$F'(x, t) = F(L^{-1}(x, t)).$$

Das ist die Poincaré-Transformierte der ursprunglichen Funktion $F(x, t)$.

Zu 2.7: Schwere Masse

Fur die schwere Masse gelten folgende Formeln:

Wenn sich zwei Korper mit den Massen M_1 und M_2 im Abstand r voneinander befinden, dann ist die Kraft K , mit der die beiden Korper zueinander gezogen werden (vgl. zu 2.9: 2. Newtonsche Gleichung):

$$K = G \cdot \frac{M_1 \cdot M_2}{r^2}$$

Dabei ist G eine Konstante (also fur alle Massen und Abstande dieselbe Zahl), deren Zahlwert von den Einheiten abhangt, in denen Masse, Abstand und Kraft gemessen wird.

Sei z.B. $M_1 = 1$ kg, $M_2 = 100$ g und $r = 5$ cm (Groenordnung des Cavendish-Experiments); G hat den Zahlwert $(6,67 \cdot 0,01) \cdot 10^{-11}$ N m²/kg². Also ist die Kraft im Beispiel $K = 2,7 \cdot 10^{-9}$ N. Die Anziehungskraft zwischen Erde (Masse ca. $6 \cdot 10^{21}$ t) und Mond (Masse ca. $5 \cdot 10^{19}$ t), Abstand ca. 400.000 km, ist nach dieser Formel $K = 5 \cdot 10^{25}$ N, also 10^{34} mal so gro – unvorstellbarer Bereich von Groenordnungen, und immer gilt in guter Naherung dieselbe Formel.

Die elektrostatische Anziehung zwischen zwei Ladungen, Q_1 und Q_2 , im Abstand r ist, analog,

$$K = -C \cdot \frac{Q_1 \cdot Q_2}{r^2},$$

wobei C wieder eine Konstante ist, deren Zahlwert von den verwendeten Maeinheiten abhangt. Es gibt allerdings positive und negative Ladungen; gleichartige stoen sich ab und verschiedenartige ziehen sich an. [Fur einzelne Elementarteilchen (Proton-Elektron) ist die elektrostatische Anziehung um einen Faktor von ca. 10^{40} starker als die Gravitation. Aber im Groen heben sich die Ladungen in ihrer Wirkung gegenseitig auf (zumal sich *ungleiche* anziehen) so, da bei groen

³ Vgl. Abb. 2.7. – Genau genommen haben wir Abb. 2.7 eingefuhrt als Ergebnis einer *passiven* Transformation, bei der die *Zeichnung* so verandert wird, da die neuen Koordinatenachsen die bliche Figur ergeben; das Ergebnis einer *aktiven* Transformation wurde man aber genauso zeichnen.

Ansammlungen von Elementarteilchen – etwa bei Planeten – die elektrostatischen Kräfte praktisch keine Rolle spielen.]

Die schwere Masse spielt also in der Gravitation eine ähnliche Rolle wie die *Ladung* in der Elektrostatik.

Zu 2.7: Äquivalenz von träger und schwerer Masse

In Formeln läßt sich der Unterschied zwischen träger und schwerer Masse im Rahmen der Newtonschen Theorie ausdrücken: Betrachten wir einen Körper der schweren Masse M und der trägen Masse m in einem Gravitationsfeld.

Die Schwerkraft auf den Körper ist (in Vektoren)

$\vec{K}_s = M \cdot \vec{g}$, abhängig von der *schweren* Massen, wobei \vec{g} das Gravitationsfeld kennzeichnet⁴, das für alle Körper gleich ist.

Die Beschleunigung ist nach der Newtonschen Formel

$$\vec{K} = m \cdot \vec{b}$$

mit der *trägen* Masse zu berechnen, also

$$M \cdot \vec{g} = m \cdot \vec{b}$$

oder

$$\vec{b} = \frac{M}{m} \cdot \vec{g}.$$

Die Beschleunigung ist proportional zum Schwerefeld und zum *Verhältnis* von schwerer zu träger Masse. Die Gleichheit der Beschleunigung aller Körper im Schwerefeld zeigt, daß dieses Verhältnis für alle Körper gleich ist.

Zu 2.9: 2. Newtonsche Gleichung

Sei $\vec{g}(\vec{x})$ die Feldstärke am Ort \vec{x} , dann ist die Kraft des Feldes auf den Körper Nr. r $\vec{K}_r = M_r \cdot \vec{g}(\vec{x}_r)$. Die Feldstärke aber rührt von allen anderen Körpern her, wie man aus dem Zweiten Gesetz durch Superposition der Kräfte leicht herleitet. Wenn M_i die Masse des Körpers Nr. i und \vec{x}_i sein Ort ist, gilt:

$$\vec{g}(\vec{x}_r) = G \cdot \sum_{i \neq r} \frac{M_i}{(\vec{x}_r - \vec{x}_i)^2} \cdot \frac{\vec{x}_r - \vec{x}_i}{|\vec{x}_r - \vec{x}_i|}.$$

Dabei ist über alle anderen Körper zu summieren (alle $i \neq r$). Der erste Faktor in der Summe gibt den *Betrag* der Anziehung des Körpers Nr. i an, gemäß der obigen Gleichung, der zweite Faktor die *Richtung* zu diesem Körper; die Kräfte müssen ja vektoriell, gemäß dem „Parallelogramm der Kräfte“, addiert werden).

Zu 3.5.: Voraussage relativer Häufigkeit

a) Additionsregel

Die Beziehungen zwischen Wahrscheinlichkeiten sind dieselben wie zwischen relativen Häufigkeiten. Seien E und F zwei Ereignisse, die nicht beide zusammen eintreten können. Das Ereignis $E \vee F$ ist das Ereignis, daß entweder E oder F eintritt. Die Anzahl der Ereignisse $E \vee F$ ist dann die Zahl der Ereignisse E plus Zahl der Ereignisse F . Entsprechend ist die Wahrscheinlichkeit des Ereignisses $E \vee F$ die Summe der Einzelwahrscheinlichkeiten:

$$w(E \vee F) = w(E) + w(F).$$

Das ist die Additionsregel der Wahrscheinlichkeitsrechnung.

b) Produktregel

Die Produktregel der Wahrscheinlichkeitsrechnung wird i.a. (z.B. in der Axiomatik von Kolmogoroff) als Definition der Unabhängigkeit von Ereignissen eingeführt:

⁴ Das Gravitationsfeld rührt i.a. von anderen Körpern her; vgl. *Zu 2.9* unten.

Man nennt zwei Versuche E und F gegenseitig unabhängig, wenn die Gleichung

$$w(E \wedge F) = w(E) \cdot w(F)$$

gilt.⁵ Man kann die Produktregel aber auch ableiten, wenn man die Unabhängigkeit von E und F definiert durch die Bedingung: „Die Wahrscheinlichkeit von E ist dieselbe, gleichgültig ob F oder $\neg F$ (nicht- F) gilt.“⁶

c) Wahrscheinlichkeit der relativen Häufigkeit

Nehmen wir an, ein Experiment, in dem das Ergebnis E die Wahrscheinlichkeit $w(E)=p$ hat, sei in einer Versuchsreihe N -mal gemacht worden. (Die Wahrscheinlichkeit, daß E nicht eintritt, ist $w(\neg E) = 1-p$.) Ein mögliches Ergebnis dieser Versuchsreihe wäre, daß bei den ersten k Versuchen das Ergebnis E eingetreten ist, bei allen weiteren $(N - k)$ Versuchen nicht mehr (daß also das Ergebnis $\neg E$ eingetreten ist). Die einzelnen Versuche sollen statistisch unabhängig sein. Dann gilt die Produktregel der Wahrscheinlichkeitsrechnung: Die Gesamtwahrscheinlichkeit ist das Produkt aller Einzelwahrscheinlichkeiten. Die Wahrscheinlichkeit dieser speziellen Abfolge von Ergebnissen ist also $p^k \cdot (1 - p)^{N-k}$. – Bei dieser Versuchsreihe ist die relative Häufigkeit h des Ergebnisses E : $h(E)=k/N$. Aber auch bei anderen Versuchsreihen kann $h(E)=k/N$ die relative Häufigkeit von E sein, wenn nämlich die Ergebnisse E und $\neg E$ in derselben Anzahl, aber in anderer Reihenfolge auftreten. Die Anzahl verschiedener Möglichkeiten, k Ergebnisse E auf die N

Versuche einer Reihe zu verteilen (die übrigen Ergebnisse sind dann $\neg E$), nennt man $\binom{N}{k}$, „ k unter N “⁷. Das ist also die Anzahl der möglichen Reihenfolgen von k Ergebnissen E und $(n-k)$ Ergebnissen $\neg E$. Nach der Summenregel ist die Wahrscheinlichkeit, daß *irgendeines* dieser Ergebnisse eintritt, die Summe der einzelnen Wahrscheinlichkeiten – und die sind alle gleich, nämlich wie oben gezeigt $p^k \cdot (1 - p)^{N-k}$. Bezeichnen wir mit E_k^N das Ereignis, daß unter den N Versuchen k das Ergebnis E haben, unabhängig von der Reihenfolge, dann gilt also:

$$w(E_k^N) = \binom{N}{k} \cdot p^k \cdot (1 - p)^{N-k}.$$

Bei einer Reihe von 20 Münzwürfen [$w(\text{Kopf}) = w(\text{Zahl}) = 0,5$], zum Beispiel, ergeben sich aus dieser Formel die folgenden Wahrscheinlichkeiten: Genau 10-mal (also in der Hälfte der Fälle) Kopf: 17,6%; 9-mal Kopf: 16% (also kaum weniger); keinmal Kopf: Ein Millionstel!

d) Erwartungswert der relativen Häufigkeit

Mit diesen Wahrscheinlichkeiten können wir den *Erwartungswert* und die *Streuung* der relativen Häufigkeit berechnen:

Der Erwartungswert $E(X)$ einer Größe X ist der vorausgesagte Mittelwert, entsprechend der vorausgesagten Häufigkeit (= Wahrscheinlichkeit) der mögliche Einzelwerte x_i der Größe X :

$$E(X) = \sum_k x_k \cdot w(x_k).$$

In unserem Fall ist die Größe X die relative Häufigkeit, die Werte x_k sind also hier die Werte (k/N) , bezogen auf eine Versuchsreihe mit der festen Anzahl N von Versuchen. Das in die obige Formel eingesetzt ergibt:

$$E\left(\frac{k}{N}\right) = \sum_{k=0}^N \left[\frac{k}{N} \cdot w(E_k^N) \right]$$

⁵ Vgl. Kolmogoroff (1933), §5, für unsere Zwecke angepaßt.

⁶ Vgl. für die Einzelheiten Drieschner (1979).

⁷ Man kann sich überlegen, daß gilt: $\binom{N}{k} = \frac{N!}{k!(N-k)!}$ wobei $i!$ („ i -Fakultät“) das Produkt aller ganzen Zahlen von 1 bis i ist.

Man kann ausrechnen, wenn man den obigen Wert von $w(E_k^N)$ einsetzt, daß der Erwartungswert der relativen Häufigkeit die *Wahrscheinlichkeit* des Ereignisses E ist:

$$E\left(\frac{k}{N}\right) = \sum_{k=0}^N \left[\frac{k}{N} \cdot w(E_k^N) \right] = w(E)$$

e) *Streuung der relativen Häufigkeit*

Entsprechend kann man auch die Streuung σ der relativen Häufigkeit ausrechnen. Dazu errechnet man σ^2 , den Erwartungswert für das Quadrat der Abweichung der einzelnen relativen Häufigkeit von der Wahrscheinlichkeit:

$$\sigma^2 = E\left[\left(\frac{k}{N} - p\right)^2\right]$$

Es ergibt sich (nach etwas mühsamerer Rechnung als oben) die bekannte Formel

$$\sigma = (1/\sqrt{N}) \cdot \sqrt{p \cdot (1-p)}.$$

Daran kann man ablesen: Die Streuung wird (mit $1/\sqrt{N}$) geringer, wenn die Versuchsreihe länger wird, und sie ist um so geringer, je näher p an 1 oder 0 ist.

Für das obige Beispiel von 20 Münzwürfen ergibt sich folgendes: Betrachtet man 10 Versuchsreihen von je 20 Münzwürfen, dann wird die Streuung des Ergebnisses »10-mal Kopf« ca. 12% ausmachen, bei 100 solchen Versuchsreihen ist die Streuung nur noch 3,8%.

f) *Stufenstruktur der Wahrscheinlichkeit*

Man kann aus der Wahrscheinlichkeit eines Ereignisses E die Wahrscheinlichkeit dafür ausrechnen, daß das Ereignis k -mal bei N Versuchen auftritt, für alle k zwischen 0 und N . (s.o.) Diese Wahrscheinlichkeit ist wieder eine vorausgesagte relative Häufigkeit, diesmal für zusammengesetzte Ereignisse. Sie ist eine Prognose für viele *Versuchsreihen* von je N Versuchen, nämlich dafür, welcher Anteil dieser Versuchsreihen gerade je k -mal das Ereignis E enthalten wird. Diese allgemeine Prognose kann wiederum nur *ungefähr* gelten. Wer es genauer wissen will, kann wie oben die Wahrscheinlichkeit berechnen, daß in einem Kollektiv von M Versuchsreihen (mit je N Versuchen) gerade l sind, die je k -mal das Ereignis E enthalten. Diese Wahrscheinlichkeit ist wiederum eine vorausgesagte relative Häufigkeit, die sich wie oben präzisieren läßt in Kollektiven der nächsten Stufe, etc.⁸; man kann beliebig so fortfahren. Für den praktischen Gebrauch genügt aber der Präzisierungsschritt auf die zweite, allenfalls die dritte Stufe, denn dann kann man die möglichen Ergebnisse der nächsten Stufe so zusammenfassen, daß nur Wahrscheinlichkeiten sehr nahe an Null oder Eins auftreten, die man als „praktisch sicher“ bzw. „praktisch unmöglich“ behandeln kann.

Zu 3.6c: Bayessche Regel

Sei E das prognostizierte Ereignis, und seien H_1, H_2, \dots, H_n die Hypothesen, die zur Auswahl stehen; sei außerdem $w(E|H_i)$ die Wahrscheinlichkeit des Ereignisses E gemäß Hypothese H_i (dabei bedeutet $w(a|b)$: »Wahrscheinlichkeit von a unter der Bedingung b «). Das Standardbeispiel ist eine Urne mit 10 Kugeln, die rot oder schwarz sein können. Man soll blind eine Kugel herausgreifen, ihre Farbe feststellen, und dann Wahrscheinlichkeiten für das Mischungsverhältnis von schwarz und rot angeben. Es gibt 11 verschiedene Hypothesen, H_0, H_1, \dots, H_{10} für das Mischungsverhältnis, nämlich »0,1,2,..., 10 der 10 Kugeln sind schwarz«; und es gibt zwei mögliche Ergebnisse, nämlich daß die gezogene Kugel rot oder schwarz ist, E_r und E_s . Die Wahrscheinlichkeit, eine schwarze Kugel zu ziehen, ist $x/10$, wenn x der zehnten Kugeln schwarz sind, also $w(E_s|H_x) = x/10$.

Das Beispiel sollte zur Veranschaulichung der Größen dienen. Wir geben die Bayessche Regel gleich für den allgemeinen Fall:

Sei $w(E|H_i)$ die Wahrscheinlichkeit des Ereignisses E gemäß der Hypothese H_i . Wenn man das Ergebnis E festgestellt hat, dann ist die Wahrscheinlichkeit, daß Hypothese H_i die richtige ist:

$$w(H_i | E) = \frac{w(H_i) \cdot w(E | H_i)}{\sum_{i=1}^N w(H_i) \cdot w(E | H_i)}.$$

Das ist leicht aus den Grundgesetzen der Wahrscheinlichkeitsrechnung abzuleiten. Es ist eine Art Umkehrung der Abhängigkeiten: Während man gewöhnlich eine Hypothese benutzt, um die bedingte Wahrscheinlichkeit eines Ereignisses unter der Bedingung der Hypothese abzuleiten, gibt die Bayessche Regel die bedingte Wahrscheinlichkeit unter der Bedingung des Ereignisses. Wie man sieht, hängt diese Wahrscheinlichkeit noch von Ausdrücken der Form $w(H_i)$ ab, der „a-priori-Wahrscheinlichkeit“ der Hypothese H_i . Man benutzt die Bayessche Regel zur Verbesserung von schon vor dem Versuch geschätzten Wahrscheinlichkeiten. Es läßt sich zeigen, daß *auf die Dauer* die Wahrscheinlichkeit einer Hypothese, wenn sie auf Grund von empirischen Befunden immer wieder gemäß der Bayesschen Regel korrigiert wird, von den ursprünglichen „a-priori-Wahrscheinlichkeiten“ ganz unabhängig wird. Dabei erhält man eine Erweiterung der Möglichkeiten, wenn man als Ereignis E auch ein Kollektiv mit bestimmten relativen Häufigkeiten zuläßt:

Dann ist z.B. „Zweimal rot und einmal schwarz“ ein einziges Ereignis⁹.

Zu 3.8b: Ehrenfest'sches Kugelspiel

Das Ehrenfest'sche Kugelspiel¹⁰ ist ein Modell, mit dem man Überlegungen der statistischen Thermodynamik vereinfacht darstellen kann. Es verwendet diskrete Mikrozustände und diskrete Zeitschritte. Allerdings ist die reversible „Mikro“-Theorie des Kugelspiels (die hier an die Stelle der Mechanik tritt) nicht deterministisch, aber das spielt für unsere Erwägungen keine Rolle.

Dieses Kugelspiel geht so:

N Kugeln, numeriert mit $1, 2, \dots, N$, sind auf 2 Urnen verteilt. Außerdem gibt es eine Nummernlotterie, welche bei jeder „Ziehung“ eine der Zahlen $1, 2, \dots, N$ produziert, jede mit der Wahrscheinlichkeit $1/N$. Nach jeder solchen Ziehung wechselt die Kugel, deren Nummer gezogen ist, die Urne.

Wie man leicht ausrechnet, ist die Wahrscheinlichkeit, daß eine Urne mit k Kugeln eine weitere dazubekommt, also dann $k+1$ Kugeln enthält:

$$w(k+1 | k) = (N - k)/N;$$

diese Wahrscheinlichkeit ist proportional der Zahl $(N - k)$ der Kugeln in der anderen Urne; es ist die Wahrscheinlichkeit, daß die Nummer einer der Kugeln in der anderen Urne gezogen wird, so daß diese Kugel in „unsere“ Urne wechselt. Sie ist umso größer, je *kleiner* k ist.

Die Wahrscheinlichkeit, daß die Urne eine Kugel verliert, ist die Wahrscheinlichkeit, daß die Nummer einer der Kugeln in dieser Urne gezogen wird:

$$w(k-1 | k) = k/N;$$

diese Wahrscheinlichkeit ist umso größer, je *größer* k ist.

Die Zahl der Kugeln tendiert also zum mittleren Wert, dem Gleichgewichtswert $k_G = N/2$. Dann ist $w(k_G-1 | k_G) = w(k_G+1 | k_G) = 1/2$; die Wahrscheinlichkeit der Zu- oder Abnahme der Kugelzahl ist symmetrisch.

Nennen wir den *Makrozustand* die Anzahl k von Kugeln in der, sagen wir, linken Urne. Der *Mikrozustand* ist die Angabe zu jeder Kugel, in welcher Urne sie ist. Die Mikrozustände sind gleichwahrscheinlich, wenn die Kugeln „statistisch“ auf die Urnen, also jedenfalls nach sehr langer Spieldauer.

⁹ Eine sehr anregende Diskussion dieser Probleme gibt S. Watanabe (1969).

¹⁰ Ehrenfest (1906). Vgl. z.B. Eigen / Winkler (1975).

Die Anzahl der Mikrozustände in einem Makrozustand ist die Anzahl der Möglichkeiten, k Kugeln unter N herauszusuchen, geschrieben $\binom{N}{k}$, und diese Zahl ist¹¹: $\binom{N}{k} = \frac{N!}{k!(N-k)!}$.

Die Gesamtzahl aller Möglichkeiten, N Kugeln auf zwei Urnen zu verteilen, ist 2^N . Damit müssen wir die Zahl der Mikrozustände normieren, um die „thermodynamische Wahrscheinlichkeit“ des Zustands k zu erhalten:

$$w_{th}(k) = 2^{-N} \cdot \binom{N}{k}.$$

Seine Entropie ist $H(k) = \alpha \cdot \log w_{th}(k)$, mit einem bestimmten Faktor α . Das Maximum der Entropie liegt beim Gleichgewichtswert von $k = k_G = N/2$.

An dem Kugelspiel-Modell sieht man, was wir oben allgemein behauptet haben: Die Wahrscheinlichkeit, daß die Entropie zunimmt, ist größer, als daß sie abnimmt – in Zukunft. Für die Vergangenheit gilt das nur, wenn ohnehin der Gleichgewichtszustand erreicht war und jeder Wert $k \approx N/2$ am wahrscheinlichsten das Entropie-Minimum in einer Schwankung ist. Auch das kann man leicht ausrechnen, ebenso wie die Wahrscheinlichkeit vergangener Zustände unter anderen Voraussetzungen, etwa wenn das Kugelspiel eben erst angefangen hat.¹²

Zu 3.13a: Verband der klassischen Aussagenlogik

Wir betrachten als grundlegende Relation die *Implikation* $A \rightarrow B$: „Wenn A wahr ist, dann ist auch B wahr“. In der klassischen Aussagenlogik wird die gesamte Aussage „ $A \rightarrow B$ “ wieder als Aussage vom selben Typ wie A oder B behandelt; sie kann wahr oder falsch sein, und welches von beiden sie ist, hängt von den Wahrheitswerten von A oder B ab: Die Implikation ist eine *Wahrheitswertfunktion* („ $A \rightarrow B$ “ ist falsch, wenn „ A “ wahr ist und „ B “ falsch, sonst wahr; „ $A \rightarrow B$ “ ist also als Wahrheitswertfunktion äquivalent der Funktion „ $\neg A \vee B$ “).

Wir betrachten in diesem Kapitel, anders als in der klassische Aussagenlogik, *Voraussagen* A und B ; die Implikation muß daher hier lauten: „Wenn A notwendig ist, dann ist auch B notwendig“. Diese Implikation ist selbst *nicht* eine Voraussage, sondern eine (immer-wahre oder immer-falsche) Aussage *über* mögliche Voraussagen („Metaaussage“).

Die Implikation gibt der Menge der möglichen Voraussagen eine (Halb-) Ordnung – im mathematischen Sinn –, denn:

1. Es gilt für beliebige A : $A \rightarrow A$ („Identität“);
2. Wenn $A \rightarrow B$ und $B \rightarrow C$ dann gilt auch $A \rightarrow C$ („Transitivität“); und
3. Es kann für verschiedene A und B nur entweder $A \rightarrow B$ oder $B \rightarrow A$ gelten, nicht beides („Antisymmetrie“; $A \rightarrow B$ und $B \rightarrow A$ bedeutet: A und B sind logisch äquivalent, also *dieselbe* Aussage)

Wenn es in dieser Ordnung zu je zwei Elementen A und B auch das *Infimum* ($A \cap B$) und das *Supremum* ($A \cup B$) gibt, dann nennt man die Struktur einen *Verband*. Das Infimum $A \cap B$ ist dabei definiert durch die 3 Regeln:

- i) $A \cap B \rightarrow A$,
- ii) $A \cap B \rightarrow B$.

iii) Wenn für irgendein X gilt: $X \rightarrow A$ und $X \rightarrow B$, dann gilt auch $X \rightarrow A \cap B$.

Das sind die Regeln, die in der Aussagenlogik die Verknüpfung mit „und“ definieren (Konjunktion). Ganz analog wird das Supremum definiert: man ersetze nur „ \cap “ durch „ \cup “ und „ \rightarrow “ durch „ \leftarrow “; in der Aussagenlogik entspricht dem die Verknüpfung durch „oder“ (Disjunktion). Die klassische Aussagenlogik bildet einen Verband. Ihre Struktur entspricht speziell der eines *Booleschen Verbandes* oder einer Booleschen Algebra.

¹¹ Vgl. oben, zu 3.6a.

¹² Vgl. dazu Drieschner (1979), Kap. III, und Weizsäcker (1972).

Zu 3.13c: Verband der Quantenlogik

Der quantenmechanische Aussagenverband hat, insbesondere wegen des Indeterminismus, die Struktur einer *Komplexen Projektiven Geometrie*¹³; bei vollständigem Indeterminismus (siehe unten) ist diese projektive Geometrie *irreduzibel* und damit isomorph (strukturgleich) dem Verband der Unterräume eines *komplexen Vektorraums*, in diesem Fall (wegen der Orthokomplementarität, die die Negation abbildet) genauer eines *Hilbert-Raums*.

Die Wahrscheinlichkeitsbeziehungen sind – interessanterweise – durch die Verbandsstruktur schon vollständig festgelegt: In dem Verband der Unterräume eines Hilbertraums kann eine Wahrscheinlichkeitsverteilung nur die spezielle Form haben, die in der Quantenmechanik wirklich benutzt wird, nämlich

$$w(A) = \text{Sp}(W \cdot P_A),$$

was folgendes bedeutet: A sei ein Element des Verbands (entsprechend einer Eigenschaft des Objekts), das durch einen Unterraum des Hilbertraums dargestellt wird. Die Wahrscheinlichkeit $w(A)$, daß bei einer Nachprüfung der Eigenschaft A diese gefunden wird, hängt ab vom augenblicklichen Zustand des Objekts, der durch den „statistischen Operator“ W (die „Dichtematrix“) beschrieben wird; und zwar ist die Wahrscheinlichkeit die *Spur* des Produkts der Dichtematrix mit dem Projektor P_A auf den oben genannten Unterraum zu A .

Wird der Zustand des Objekts durch ein Atom des Verbandes (einen eindimensionalen Unterraum) ψ richtig gekennzeichnet, dann ist $W = P_\psi$ selbst ein Projektor, der Zustand heißt ein „reiner Fall“. Ist dann auch noch A eine atomare Eigenschaft φ , dann vereinfacht sich die obige Formel zu dem bekannten

$$w(A) = |(\varphi, \psi)|^2,$$

wobei (φ, ψ) das innere Produkt im Hilbertraum ist.

Im allgemeinen Fall kann W als Linearkombination von Projektoren dargestellt werden,

$$W = \sum_i \alpha_i \cdot P_{\psi_i}$$

wobei $\alpha_i \geq 0$ und

$$\sum_i \alpha_i = 1.$$

W heißt dann ein „Gemenge“ oder „Gemisch“. Ich möchte diesen Fall interpretieren wie in der klassischen Physik: Wenn man genügend über das Objekt wüßte, könnte man den zutreffenden reinen Fall (ein Atom des Verbandes) angeben; man weiß aber nur, daß verschiedene reine Fälle möglich sind, mit zugehörigen Wahrscheinlichkeiten; dieses unvollkommene Wissen wird durch das Gemenge ausgedrückt¹⁴. –

Es ist in den Überlegungen zur Quantenmechanik immer wieder gefragt worden, was denn die einzelnen Vektoren des Hilbertraums physikalisch bedeuten, was z. B. die Addition von zwei Vektoren. Nach dem eben dargestellten Bild, das wir uns heute von der Quantenmechanik machen können, müssen wir antworten: „Nichts“. Die Struktur der Quantenmechanik wird durch den Verband ausgedrückt (die komplexe Projektive Geometrie mit der Wahrscheinlichkeitsfunktion). Dieser Verband wird, zum leichteren Rechnen, als Verband der Unterräume eines Hilbertraums dargestellt. Die weiteren Eigenschaften dieses Hilbertraums sind aber für die Quantenmechanik irrelevant. Man kann sich das etwa analog wie bei den Eigenschaften von Koordinaten vorstellen: Daß der Breitengrad, auf dem ich lebe, die Summe der Breitengrade von Kopenhagen und Daressalam ist, hat *physikalisch* keine Bedeutung.

¹³ Die Punkte dieser Geometrie entsprechen den Atomen des Verbandes. – Für eine detaillierte Diskussion vgl. Drieschner (1979).

¹⁴ Bei „entarteten Eigenwerten“ ist das nicht ganz so einfach!

Zu 3.13d: Zusammensetzung von quantenmechanischen Objekten

Betrachten wir ein Objekt X , das durch eine der Eigenschaften aus der (atomaren) Alternative $\{x_1, x_2, \dots, x_n\}$ vollständig beschrieben wird, und ein Objekt Y , zu dem entsprechend die Alternative $\{y_1, y_2, \dots, y_k\}$ gehört. Wir können die beiden Objekte rein gedanklich zu einem neuen Objekt zusammenfassen, nennen wir es $X&Y$, das durch eine Eigenschaft $x_i \wedge y_j$ („ \wedge “ bedeutet „und“) vollständig beschrieben werden kann. Ist die atomare Alternative jeweils die einzige zu ihrem Objekt, wie in der klassischen Physik, dann ist damit der Fall erledigt: x_i ist die vollständige Angabe aller Eigenschaften von X , y_j die vollständige Angabe aller Eigenschaften von Y – also gibt $x_i \wedge y_j$ vollständig alle Eigenschaften von $X&Y$ an. Die (einzige) atomare Alternative von $X&Y$ ist: $\{x_1 \wedge y_1, x_1 \wedge y_2, \dots, x_1 \wedge y_k, x_2 \wedge y_1, x_2 \wedge y_2, \dots, x_2 \wedge y_k, \dots, x_n \wedge y_1, x_n \wedge y_2, \dots, x_n \wedge y_k\}$, das „direkte Produkt“ der beiden Alternativen. Sind z.B. X und Y Massenpunkte der Klassischen Mechanik, mit je einem 6-dimensionalen Phasenraum als atomarer Alternative, dann ist die atomare Alternative des Systems $X&Y$ (zwei Massenpunkte) der 12-dimensionale Phasenraum, direktes Produkt der beiden 6-dimensionalen.

Sind X und Y quantenmechanische Objekte, dann sieht das anders aus: Jedes quantenmechanische Objekt hat *mehrere* („inkompatible“) Alternativen von vollständigen Beschreibungen – i.a. unendlich viele. Das direkte Produkt *jeder* atomaren Alternative von X mit *jeder* atomaren Alternative von Y ergibt eine atomare Alternative von $X&Y$. Sind die beiden Objekte statistisch unabhängig, dann muß außerdem die Wahrscheinlichkeit für das zusammengefaßte Objekt das Produkt aus den Wahrscheinlichkeiten für die Einzelobjekte sein. Da zudem der Verband des Gesamtobjekts wieder ein Hilbertraum-Verband sein muß, liegt seine Struktur fest: Der Hilbertraum des Gesamtobjekts $X&Y$ ist das *Tensorprodukt* der beiden Hilberträume \mathfrak{H}_X und \mathfrak{H}_Y der Einzelobjekte X und Y , in Zeichen:

$$\mathfrak{H}_{X&Y} = \mathfrak{H}_X \otimes \mathfrak{H}_Y.$$

Auf die Einzelheiten des Tensorprodukts kann ich hier nicht eingehen¹⁵. Der wichtigste Punkt ist der Unterschied zwischen dem direkten Produkt und dem Tensorprodukt: Das direkte Produkt ist im Tensorprodukt enthalten, letzteres umfaßt aber viel mehr. Das bedeutet: Jede Angabe eines Zustands φ des Objekts X und eines Zustands ψ des Objekts Y beschreibt zusammen einen Zustand des Gesamtobjekts $X&Y$; darüber hinaus gibt es aber noch viele weitere Zustände des Gesamtobjekts $X&Y$, die sich nicht als Kombination aus je einem Zustand der Teilobjekte beschreiben lassen, mathematisch gesagt sind das sogar *fast alle*.

Zu 3.14: Begründung der Quantenmechanik a priori

Physik ist eine Theorie für Gesetze über Voraussagen, und das heißt notwendigerweise, eine Theorie, welche die Voraussagen über Objekte mit *Wahrscheinlichkeiten* belegt (vgl. 3.6b). Das Objekt ist selbst definiert durch die Gesamtheit seiner möglichen kontingenten Eigenschaften, und es ist vollständig beschrieben, wenn zu jeder dieser Eigenschaften eine Wahrscheinlichkeit angegeben wird. Damit läßt sich aus unserem Objektbegriff die Existenz von Verbands-Atomen und die Behauptung, daß zu jeder Zeit ein Atom des Verbandes das Objekt richtig beschreibt, begründen. Dies ist ein grundlegender Zug der Quantenmechanik; uns erscheint er hier als eine Eigenschaft, die a priori jeder Physik zukommt.

Die Struktur der *Negation* ist für jede empirische Theorie vorausgesetzt: Ein Experiment entscheidet von jedem Ergebnis, ob es herauskommt oder *nicht* herauskommt. Formal findet sich das auch in der Wahrscheinlichkeitsstruktur wieder: Wenn zwei mögliche Ergebnisse für einander Wahrscheinlichkeit Null implizieren, dann ist jeweils wenn eines wahr ist, das andere falsch: es gehört zur Negation des ersteren. – Aus der Struktur der Negation folgt der *orthomodulare* Charakter der Eigenschaftsmenge in der Theorie, also ihre Zusammensetzung aus Booleschen Verbänden. Wir lassen hier die („klassische“) Möglichkeit zu, daß die Eigenschaften insgesamt nur *einen*

¹⁵ Für eine mathematisch ordentliche Konstruktion vgl. z. B. Jauch (1968).

Booleschen Verband bilden, aber wir wollen uns wegen des Indeterminismus nicht auf diesen Fall beschränken.

Ein nächster großer Schritt wäre die Begründung a priori des sogenannten *Projektionspostulats*. Es besteht eine gewisse Hoffnung, daß man aus Forderungen an die Zusammensetzung von Objekten (vgl. 3.13d) das Projektionspostulat ableiten kann. Denn das Tensorprodukt zweier Vektorräume vereinigt ganz wunderbar die notwendigen Eigenschaften, das heißt, es enthält das direkte Produkt in der richtigen Weise und ergibt dazu noch die richtigen Wahrscheinlichkeiten für unabhängige Objekte. Es ist bisher nicht gelungen, etwas Analoges für allgemeinere orthomodulare Verbände zu konstruieren, und es ist denkbar, daß es diese Struktur *nur* dann gibt, wenn das Projektionspostulat gilt. Damit wäre die Notwendigkeit des Projektionspostulats begründet, denn Objekte zusammensetzen, das muß man doch können!

Das bisher Genannte genügt, um die Verbandsstruktur der Projektiven Geometrie abzuleiten, und erst jetzt trennen sich in diesem Aufbau die Wege von klassischer Physik und Quantenmechanik: Der klassischen Physik entspricht, in dieser Sprache, eine vollständig *reduzible* Projektive Geometrie. Das ist sozusagen ein Sack voll Punkte, ohne jede weitere Struktur, bei dem je zwei Punkte eine Gerade heißen, beliebige drei Punkte eine Ebene, etc.; wir finden wieder *jede* Teilmenge der Punktmenge als Element des Verbands, der Verband hat die Struktur *eines* Booleschen Verbands. – Dem steht die *irreduzible* Projektive Geometrie gegenüber, dadurch definiert, daß jede Gerade wenigstens drei Punkte enthält. Dieses geometrische Kriterium entspricht dem *Indeterminismus*: Der Eigenschaftsverband mit der Struktur einer Projektiven Geometrie ist genau dann irreduzibel, wenn es zu jedem seiner Elemente wenigstens einen „reinen Fall“ gibt, in dem die Wahrscheinlichkeit dieses Elements weder 0 noch 1 ist – d. h., in dem der Ausgang einer Messung der entsprechenden Eigenschaft „indeterminiert“ ist.

Im Rahmen unserer transzendenten Argumentation müssen wir jedenfalls die Möglichkeit des Indeterminismus offenhalten. In unserer Diskussion der Voraussetzungen (3.4) fanden wir, daß die Offenheit der Zukunft ein Prinzip aller Erfahrung ist, so daß wir uns nicht auf den deterministischen, also den vollständig reduzierbaren Fall beschränken dürfen. – In der Quantenmechanik, wie wir sie heute kennen, kommen tatsächlich alle Arten von Übergangsformen vor, teilweise reduzierbare Geometrien oder, anders ausgedrückt, *Familien* von mehreren irreduzierbaren Geometrien (in der physikalischen Terminologie meist als „Superauswahlregeln“ bezeichnet).

Eine irreduzible Projektive Geometrie ist isomorph dem Verband aller Unterräume eines Vektorraums über einem Zahlkörper K ; das ist ein bekannter „Hauptsatz der Projektiven Geometrie“. Bleibt zur vollständigen Charakterisierung unseres Eigenschaftsverbands also nur noch der Zahlkörper zu wählen. Ein p -adischer Körper kommt nicht in Frage, da er kein Orthokomplement zulassen würde, das die Negation darstellt. Sehr allgemeine topologische Argumente lassen nur drei Möglichkeiten übrig, nämlich die reellen Zahlen, die komplexen Zahlen und die Quaternionen. Letztere lassen kein Tensorprodukt mit den richtigen Wahrscheinlichkeiten zu, so daß erstlich nur die reellen und komplexen Zahlen in Frage kommen. In einer „reellen“ Quantenmechanik scheint immer die Möglichkeit zu bestehen, je zwei Koordinaten zu einer komplexen zusammenzufassen. Also bleiben praktisch nur die komplexen Zahlen übrig.

Mit dieser letzten Bedingung, daß *komplexe* Zahlen verwendet werden, ist die Struktur des quantenmechanischen Eigenschaftsverbandes festgelegt, einschließlich der Wahrscheinlichkeitsbeziehungen, welche aus der Negationsrelation folgen – es ist die Struktur des Verbands der (abgeschlossenen) Unterräume eines *Hilbertraums*. Die allgemeine Form der zeitlichen Änderung ergibt sich daraus zwangsläufig so, wie sie in der *Schrödingergleichung* beschrieben wird.

Technisch wird die zeitliche Änderung dargestellt durch eine einparametrische Gruppe von Automorphismen, mit dem reellen Parameter t , die Zeit darstellend. D. h., wenn $\psi(0)$ der Zustand zur Zeit 0 ist, dann gilt für den Zustand $\psi(t)$ zur Zeit t :

$$\psi(t) = e^{iHt} \psi(0).$$

Das ist die („zeitabhängige“) Schrödingergleichung mit dem selbstadjungierten „Hamiltonoperator“ H , der den speziellen Charakter der zeitlichen Änderung beschreibt. Dabei ist i die imaginäre Einheit und t der Zeitparameter.

Soweit die Andeutungen zur Axiomatik. Man sieht, es ist bisher keine transzendente Begründung der abstrakten Struktur der Quantenmechanik wirklich vorgelegt, aber nach den vorgeführten Argumenten sind wir immerhin einer solchen Begründung nahe gekommen.

Zu 3.15c Verschwinden der Interferenzterme im Meßprozeß

Zwei Postulate, welche beide für die Beschreibung der Messung unentbehrlich scheinen, widersprechen sich unvermeidlich innerhalb der Quantenmechanik selbst. Das läßt sich nur mit einigem Formalismus näher erklären:

Nehmen wir an, wir messen eine Größe A mit (der Einfachheit halber) zwei möglichen Werten a_1 und a_2 . Die Zustände des Objekts, die zu diesen Werten gehören, seien ψ_1 bzw. ψ_2 ; die Zustände des Meßgeräts, die das Vorliegen der Werte anzeigen, seien χ_1 bzw. χ_2 . Die Postulate der Quantenmechanik besagen folgendes: Ist das gemessene Objekt vor der Messung im Zustand ψ_1 , dann muß auf Grund der Meßwechselwirkung schließlich der Zustand des Gesamtobjekts (gemessenes Objekt & Meßgerät) $\psi_1 \wedge \chi_1$ sein, denn das Meßergebnis a_1 war ja *notwendig* (das besagt der Zustand ψ_1 , nach Voraussetzung), und χ_1 zeigt dieses Ergebnis a_1 an; analog muß der Endzustand $\psi_2 \wedge \chi_2$ werden, wenn der Zustand des gemessenen Objekts vor der Messung ψ_2 war. Nach der Quantenmechanik erzeugt dann die Meßwechselwirkung, wenn das gemessene Objekt ursprünglich in dem beliebigen Zustand

$$\psi_0 = \beta_1 \cdot \psi_1 + \beta_2 \cdot \psi_2 \quad (*)$$

war (mit beliebigen komplexen Zahlen β_1, β_2 so, daß $|\beta_1|^2 + |\beta_2|^2 = 1$), den Endzustand des Gesamtobjekts

$$\Phi = \beta_1 \cdot \psi_1 \wedge \chi_1 + \beta_2 \cdot \psi_2 \wedge \chi_2.$$

Dies ist eine für die Quantenmechanik grundlegende Argumentation.

Wir können aber auch etwas anders argumentieren, ebenso grundlegend für die Quantenmechanik, und erhalten etwas anderes:

Der Zustand ψ_0 des gemessenen Objekts (wie in Gl.(*) beschrieben) bedeutet, daß man mit der Wahrscheinlichkeit $|\beta_1|^2$ das Ergebnis a_1 und mit der Wahrscheinlichkeit $|\beta_2|^2$ das Ergebnis a_2 finden wird. D. h. nach der Messung wird das Gesamtobjekt mit der Wahrscheinlichkeit $|\beta_1|^2$ in dem Zustand sein, welcher dem Meßergebnis a_1 entspricht, d. h. $\psi_1 \wedge \chi_1$, und entsprechend für a_2 . Das Gesamtobjekt wird also nach der Messung – solange das Meßergebnis nicht berücksichtigt ist – durch den statistischen Operator W beschrieben:

$$W = |\beta_1|^2 \cdot P_{\psi_1 \wedge \chi_1} + |\beta_2|^2 \cdot P_{\psi_2 \wedge \chi_2},$$

wobei P_x der Projektor auf den Zustand x ist. W ist die gewichtete Summe von zwei Projektoren und stellt i.a. *nicht* den Zustand Φ dar; dem Zustand Φ entspricht vielmehr der statistische Operator

$$W' = P_{\Phi},$$

der Projektor auf Φ . $W = W'$ gilt nur dann, wenn eins der β_i verschwindet.

Für die Beschreibung sowohl des gemessenen Objekts wie auch des Meßgeräts, jeweils für sich betrachtet, gibt es zwischen den beiden statistischen Operatoren W und W' allerdings keinen Unterschied. Der Unterschied liegt in der Korrelation *zwischen* beiden Objekten. Diese Korrelation kann bei einer guten Messung sehr klein sein, ist aber niemals Null!

Es wird deshalb argumentiert, zur Messung gehöre der gedankliche *Schnitt*¹⁶ zwischen gemessenem Objekt und dem Meßgerät, oder auch zwischen Objekt und beobachtendem Subjekt. Das mag richtig sein, aber es ändert nichts an der Tatsache, daß die Postulate, die zu Recht an die

¹⁶ Vgl. Süßmann (1958), Mittelstaedt (1963).

Quantenmechanik gestellt werden, formalistisch konsequent durchgezogen zu einem Widerspruch führen.

Zu 3.15f: a) Stern-Gerlach-Versuch

Betrachten wir den Eigendrehimpuls (Spin) eines Elektrons! O. Stern und W. Gerlach haben in ihrem berühmten Experiment gezeigt, daß sich die Komponente dieses Spins in bezug auf ein vorgegebenes Magnetfeld nur als $+1/2$ oder als $-1/2$ herausstellen kann (vgl. 3.2). Messungen mit Magnetfeldern *verschiedener* Richtung sind quantenmechanisch unverträglich. Ein Spinzustand ist z. B. vollständig beschrieben durch die Angabe, daß in einer bestimmten Orientierung des Magnetfeldes mit Sicherheit die Komponente $+1/2$ gefunden würde; für alle anderen Orientierungen sind dann beide Ergebnisse ($+1/2$ und $-1/2$) möglich, mit Wahrscheinlichkeiten *zwischen* Null und Eins; eine vollständigere Beschreibung gibt es nicht. Nehmen wir etwa an, der Spin eines Elektrons sei nach oben orientiert, das heißt, bei einem senkrecht orientierten Magnetfeld könnte man mit Sicherheit voraussagen, daß die Spinkomponente sich als „nach oben“ herausstellen wird und nicht „nach unten“. Wird nun das Magnetfeld des Stern-Gerlach-Versuchs waagrecht statt senkrecht orientiert, dann ist die Wahrscheinlichkeit, die Spin-Komponente „nach rechts“ oder „nach links“ zu finden, je $1/2$; bei einer irgendwie schrägen Orientierung des Magnetfeldes ist die Wahrscheinlichkeit, die Komponente in der Richtung x zu finden:

$$w(x) = \cos^2 \vartheta/2,$$

wenn ϑ der Winkel zwischen der Richtung x und der vorliegenden Richtung (im Beispiel „nach oben“) ist.

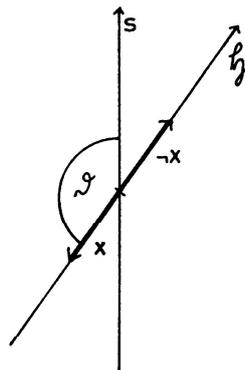


Abb. 20: Orientierung des Stern-Gerlach-Versuchs (Magnetfeld ξ) zum vorliegenden Spin s . Wenn $\vartheta = 120^\circ$, dann ist $w(x) = \cos^2 60^\circ = 1/4$; $w(-x) = \cos^2 30^\circ = 3/4$.

Für das Verständnis des EPR-Arguments ist es entscheidend, den speziell quantenmechanischen Charakter dieser Größen zu sehen. In der klassischen Mechanik ist die Observable „Größe und Richtung des Drehimpulses“ eine Vektorgröße; man kann etwa einem Kreisel die Eigenschaft „Drehimpuls 2 erg · sec links um eine senkrechte Achse“ zuordnen; diese Eigenschaft hat er „an sich“, alle anderen Drehimpulse hat er nicht. – In der Quantenmechanik ist die *Größe* des Drehimpulses charakteristisch für das Objekt und unveränderlich – sie ist keine „kontingente Größe“. Bleibt die Richtung: Tatsächlich stellt jede mögliche Richtung des dreidimensionalen Raumes einen „Zustand“ des Elektronenspins dar¹⁷. Der entscheidende Unterschied zur klassischen Mechanik ist aber, daß diese Richtungen nicht in *einer* Messung unterscheidbar sind; sie gehören, technisch gesprochen, nicht zu einem Booleschen Verband (3.13). Vielmehr kann ein Experimentator willkürlich die Orientierung des Magnetfeldes in seinem Stern-Gerlach-Versuch vorgeben; dann kann er nur entweder finden, daß der Spin parallel oder daß er antiparallel zum Magnetfeld orientiert ist – nur zwei Möglichkeiten. Die *Wahrscheinlichkeiten* der Ergebnisse aller möglichen Versuche sind im „Zustand“ zusammengefaßt, und diesen Zustand kann man (im „reinen“ Fall, s.o.) als Richtung im Raum angeben. Die verschiedenen Observablen, unter denen der Ex-

¹⁷ Für Spins anderer Objekte, die größer als $1/2 \hbar$ sind, ist die Situation komplizierter. Dann gibt es für jede Messung mehr als 2 Möglichkeiten. Wir können darauf hier nicht eingehen und verweisen auf Lehrbücher der Quantenmechanik

perimentator durch die Orientierung des Magnetfeldes wählen kann, sind quantenmechanisch nicht verträglich, es kann immer nur für *eine* Richtung entschieden sein, ob parallel oder antiparallel; und wenn es entschieden ist, dann ist damit der „Zustand“ festgelegt.

Zu 3.15f: b) EPR-Experiment nach Bohm

Ich schildere das Argument an einem leicht veränderten Beispiel, das zuerst David Bohm angegeben hat¹⁸. Das Entscheidende ist beiden Beispielen gemeinsam: Man betrachtet meßbare Größen, die quantenmechanisch unverträglich sind; jede dieser Größe kann an verschiedenen Teil-Objekten korreliert sein.

Nun zu dem Gedankenexperiment von EPR bzw. Bohm: Man betrachtet ein Objekt mit dem Spin Null, das in zwei Objekte mit je Spin $\frac{1}{2}$ zerfällt. Die Zerfallsprodukte werden in Stern-Gerlach-Apparaten (A und B) analysiert, die unabhängig voneinander um die Einfallrichtung drehbar sind. Es gilt der Satz von der *Erhaltung des Drehimpulses*, aus dem folgt: Wenn der Apparat A senkrecht orientiert ist und man dort den Spin „nach oben“ festgestellt hat, dann ist der *Spin-zustand* bei B „nach unten“; das heißt, *wenn* Apparat B auch senkrecht orientiert ist, dann kann man mit Sicherheit voraussagen, daß Spin „nach unten“ gefunden wird, für schräge Orientierung hat man die oben angeführten Wahrscheinlichkeitsvoraussagen.

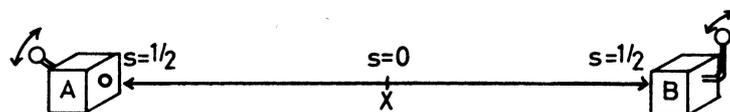


Abb. 21: Stern-Gerlach-Apparate (A und B) zur Messung der Spin-Komponente der Zerfallsprodukte.

Nun ist es vernünftig, diesen „Zustand“ als eine Eigenschaft anzusehen, die das Objekt an sich hat, denn man kann sagen: „Wenn ich die Eigenschaft (hier: ob der Spin nach unten orientiert ist) nachprüfe, werde ich sie mit Sicherheit finden.“ Einstein, Podolsky und Rosen nennen das „ein Element der Wirklichkeit“ („an element of reality“): „Wenn wir an einem System, ohne es irgendwie zu stören, den Wert einer physikalischen Größe mit Sicherheiten (d.h. mit Wahrscheinlichkeit Eins) voraussagen können, dann gibt es ein Element der physischen Wirklichkeit, das dieser physikalischen Größe entspricht.“ – Nach dieser Definition ist die Spin-Richtung ein Element der physischen Wirklichkeit; denn wenn die beiden Apparate A und B in derselben Richtung orientiert sind, dann kann ich das Ergebnis bei B mit Sicherheit voraussagen, sobald ich bei A nachgeschaut habe.

So weit ist die Sache einleuchtend. Es entsteht aber ein paradoxer Eindruck, weil der Spin-Zustand des Teilchens bei B, also das „Element der Wirklichkeit“, von der Willkür des Experimentators bei A abhängt: Er kann den Apparat A irgendwie drehen, wenn das andere Teilchen (B) schon so weit weg ist, daß sein Zustand (gemäß der Relativitätstheorie) nicht mehr beeinflusst werden kann; trotzdem hängt dieser Zustand (B) von der Orientierung des Apparates A ab.

Um das zu verstehen, muß man sich wieder erinnern, daß der „Zustand“ zu einer abgekürzten Sprechweise gehört: Es handelt sich eigentlich um eine Zusammenfassung vieler Wahrscheinlichkeits-Voraussagen. Man kann bei genauerer Betrachtung sehen, daß die Wahrscheinlichkeiten kein besonderes Geheimnis bergen; die Ergebnisse sind in jedem Fall konsistent, wie man sich leicht überzeugen kann, denn das *Ergebnis* »parallel oder antiparallel zum Magnetfeld« hängt nicht von der Willkür des Experimentators ab. Das Ungewohnte ist, daß man dem anderen Teilchen (B) erst *nach* der Messung bei A einen „wirklichen“ Zustand zuschreiben kann, nicht vorher, solange die Orientierung des Apparates A noch nicht endgültig und die Messung noch nicht entschieden ist.

Die Voraussagen für Messungen bei B sind von den Manipulationen bei A ganz unabhängig; man kann also den Effekt nicht zur Übermittlung von Nachrichten benutzen. Die Korrelation kann man nur *nachträglich* beim Vergleich zusammengehöriger Paare feststellen. Man wird also nicht so reden, daß „etwas Reales“ von A nach B übermittelt wird.

¹⁸ Bohm (1952).

Einstein, Podolsky und Rosen ziehen aus diesen Erwägungen den Schluß, daß die Quantenmechanik nicht vollständig (“complete”) sei, wenn sie ein offensichtlich vorhandenes Element der Wirklichkeit nicht zu beschreiben gestatte. In Wahrheit ist es viel schlimmer: Allein die *Annahme*, das andere Teilchen (B) habe vor der Messung bei A irgendeinen Spinzustand, auch wenn man sonst alles offen läßt, ist mit den quantenmechanischen Voraussagen nicht vereinbar. J. S. Bell rechnet in einem berühmten Artikel¹⁹ vor, daß allein die Annahme, jedem der Teilchen komme gleich nach dem Zerfall irgendein Spin-Zustand zu (sogar mit ganz beliebiger Wahrscheinlichkeitsverteilung über die möglichen Spinzustände), zu einer Abweichung von der quantenmechanischen Voraussage führt, die experimentell nachweisbare Größenordnung hat. Dabei setzt er nur voraus, daß ein Drehen am Apparat A nicht den Zustand des Teilchens B auf die Ferne beeinflusst. Bell zeigt also, daß keine „lokale“ Theorie mit verborgenen Parametern in diesem Punkt die Voraussagen der Quantenmechanik reproduzieren kann.

Wegen der prinzipiellen Bedeutung dieser Frage ist die Bellsche Korrelation in umfangreichen Experimenten nachgeprüft worden. Und zwar hat man als „Teilchen“ Photonen verwendet, die paarweise von Atomen emittiert werden, so daß ihr Gesamtspin 0 ist („Kaskade“); anstelle der Stern-Gerlach-Apparate sind bei A und B Polarisatoren aufgestellt. Die Struktur des Arguments ist aber genau dieselbe wie die oben vorgeführte. Es bestätigte sich in diesen Experimenten genau die quantenmechanische Voraussage, während jede Möglichkeit für lokale verborgene Parameter außerhalb der Fehlergrenzen lag²⁰.

Wenn man EPR *realistisch* interpretieren will, d. h. so, daß der „Zustand“ des Teilchens B etwas Wirkliches ist, dann muß man eine Fernwirkung postulieren, welche (mit Überlichtgeschwindigkeit) dafür sorgt, daß das Teilchen B den richtigen Zustand annimmt. Das Argument ist dann, daß eine solche Fernwirkung nicht der Relativitätstheorie widerspreche, weil sie keine Energie und keine Signale übermitteln kann. – Das ist freilich eine möglich Sprechweise, aber man spricht damit von *nichts*: Die so eingeführte Fernwirkung ist eigentlich gar keine Wirkung; sie ist nur ein Wort, das seine Daseinsberechtigung allein daher hat, daß es eine „realistische“ Sprachregelung ermöglicht.

¹⁹ Bell (1964); Espagnat (1980) gibt eine sehr übersichtliche Darstellung.

²⁰ Clauser (1976) und die dort zitierte Literatur. Sehr schön ist die Argumentation wiederum bei Espagnat (1980) dargestellt.