



Pauli- und Dirac-Gleichung

(Theorie des spin)

1. Vorbemerkungen

Die (stationäre) Schrödingergleichung

$$(\vec{p}^2 - 2mE) |\Psi\rangle = 0$$

für ein freies Teilchen der Masse m ist nicht relativistisch kovariant, da die relativistische Kinematik die Energie-Impuls-Beziehung

$$\left(\frac{E}{c}\right)^2 - \vec{p}^2 = (mc)^2 \quad \left[\text{und nicht } E = \frac{\vec{p}^2}{2m} \right]$$

erfordert. In kovarianter Schreibweise (d.h. mit "Vierervektoren")¹

¹Diese Schreibweise wird besonders übersichtlich, wenn man für das Schreiben von Skalarprodukten ab im Minkowski-Raum die "Einsteinsche Summationskonvention" benutzt. Sie lautet: *über jedes Paar von kontravarianten ('oberen') und kovarianten ('unteren') Indizes, das in einem multiplikativen Ausdruck vorkommt, ist zu summieren* — in Formeln:

$$a_\mu b^\mu \text{ ist eine Abkürzung für } \sum_{\mu=0}^3 a_\mu b^\mu \quad (= ab) \text{ .}$$

Kommt dagegen ein bestimmter Index auf den beiden Seiten einer Gleichung (oder auch in jedem Term einer Addition) vor, so ist die Gleichung als *Vektorgleichung* zu verstehen (für jeden möglichen Wert des Index eine Gleichung) — in Formeln:

$$a_\mu + b_\mu = c_\mu \text{ ist eine Abkürzung für die 4 Gleichungen} \\ a_0 + b_0 = c_0 \text{ , } a_1 + b_1 = c_1 \text{ , } a_2 + b_2 = c_2 \text{ , } a_3 + b_3 = c_3 \text{ .}$$

Für ein gewöhnliches (dreidimensionales Euklidisches) Skalarprodukt $a \cdot b$ kann man zweckmäßig eine analoge Schreibweise benutzen:

$$a_k b_k \text{ ist eine Abkürzung für } \sum_{k=1}^3 a_k b_k \quad (= \vec{a} \cdot \vec{b}) \text{ .}$$

Wir werden beide Konventionen in diesem Skript durchgängig benutzen, wobei noch die zusätzliche Vereinbarung gilt:

Dreidimensionale (Euklidische) Vektoren werden mit lateinischen Buchstaben notiert, vierdimensionale (Minkowski-) Vektoren mit griechischen Buchstaben.

[Für eine ausführlichere Behandlung s. z.B. **J.D. Jackson**: *Classical Electrodynamics* (2nd Ed.), J. Wiley & Sons, New York 1975, Sec. 11.6; noch gründlicher in **C.C. Noack**: *Tensoranalysis: eine Einführung*, Universität Bremen 1994; und schließlich im differentialgeometrischen Glanz und Gloria in **C.W. Misner, K.S. Thorne, J.A. Wheeler**: *Gravitation*, W.H. Freeman and Comp. San Francisco 1973.]

$$x^\mu := \{x^0, x^1, x^2, x^3\} = (ct, \vec{r}) \quad (1)$$

$$p^\mu := \{p^0, p^1, p^2, p^3\} = (E/c, \vec{p}) \quad (2)$$

bzw.

$$x_\mu := g_{\mu\nu} x^\nu = (ct, -\vec{r})$$

$$p_\mu := g_{\mu\nu} p^\nu = (E/c, -\vec{p})$$

läßt sich die relativistische Energie-Impuls-Beziehung (“Dispersionsrelation”) also besonders übersichtlich schreiben als

$$(\mathbf{p}^2 - m^2 c^2) = (\mathbf{p}_\mu \mathbf{p}^\mu - m^2 c^2) = 0 \quad .$$

Eine ‘relativistische Schrödinger-Gleichung’ könnte also lauten

$$(\mathbf{p}^2 - m^2 c^2) |\Psi\rangle = 0$$

oder, mit der Schrödingerschen Quantisierungsvorschrift in Ortsdarstellung

$$E \implies p_0 = (i\hbar) \frac{\partial}{\partial t} \quad (3)$$

$$\vec{p} \implies p_k = (-i\hbar) \frac{\partial}{\partial x^k} \quad , \quad (4)$$

(die sich kovariant einfach als $\mathbf{p}_\mu \implies (i\hbar) \frac{\partial}{\partial x^\mu}$ schreibt)

$$\left\langle x \left| \left(\hbar^2 \frac{\partial}{\partial x_\mu} \frac{\partial}{\partial x^\mu} + m^2 c^2 \right) \right| \Psi \right\rangle = 0 \quad .$$

Diese Gleichung (die “Klein-Gordon-Gleichung”) wurde bereits von Schrödinger selbst so aufgestellt, und zwar noch *vor ihrer nichtrelativistischen Form* (der heutigen “Schrödinger-Gleichung”)!

Vom Standpunkt der Relativitätstheorie aus erscheint die Klein-Gordon-Gleichung zwingend als die einzig mögliche Form der quantenmechanischen Bewegungsgleichung (für ein freies Teilchen der Masse m). Es gibt dabei jedoch ein ernstes **Problem**: Die Zeitentwicklung ist *nicht linear*; das aber steht im Gegensatz zur Gruppenstruktur der Zeittranslationen zusammen mit der Wahrscheinlichkeitsinterpretation² der Hilbertraum-Vektoren.

Eine vollständige Lösung dieser Probleme gelingt erst im Rahmen der **Quantenfeldtheorie**, in der man die Idee von einzelnen, isoliert voneinander existierenden Teilchen aufgibt (oder zumindest stark modifiziert). Einen ersten Schritt dorthin machte P.A.M. Dirac [Dir 28] mit der Idee, die Klein-Gordon-Gleichung zu ‘linearisieren’, d.h.

eine im Viererimpuls *lineare* Gleichung zu finden, die die Klein-Gordon-Gleichung zur *Folge* hat.

Viele Jahre später zeigte Levy-Leblond [Lev 67], dass (und wie) man mit dem gleichen Recht eine solche ‘Linearisierung’ auch in der nichtrelativistischen (Schrödinger-)Theorie durchführen kann.

Da beides völlig analog funktioniert, werden wir in diesem Skript die beiden Fälle parallel behandeln.

²Das war der Grund für Schrödinger, die – von ihm selbst aufgestellte – Klein-Gordon-Gleichung schließlich wieder zu verwerfen und sich (resignierend!) auf den nichtrelativistischen Fall zu beschränken.

2. Linearisierung

Pauli

Dirac

Zur Linearisierung von

$$\vec{\mathbf{p}}^2 - 2mE$$

$$\mathbf{p}^2 - m^2c^2$$

setzen wir an $\left[\epsilon := \sqrt{2mE} \right]^3$

$$\begin{aligned} \mathbf{P} &:= \vec{\alpha} \cdot \vec{\mathbf{p}} - \beta \epsilon \\ \mathbf{P}' &:= \vec{\alpha}' \cdot \vec{\mathbf{p}} + \beta' \epsilon \end{aligned}$$

$$\begin{aligned} \mathbf{D} &:= \alpha \mathbf{p} - \beta mc \\ \mathbf{D}' &:= \alpha' \mathbf{p} + \beta' mc \end{aligned} \quad (5)$$

mit Unbekannten

$$\begin{aligned} \alpha_1, \alpha_2, \alpha_3, \beta \\ \alpha_1', \alpha_2', \alpha_3', \beta' \end{aligned}$$

$$\begin{aligned} \alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta \\ \alpha_0', \alpha_1', \alpha_2', \alpha_3', \beta' \end{aligned}$$

und verlangen nun

$$\mathbf{P}' \mathbf{P} \stackrel{!}{=} \vec{\mathbf{p}}^2 - \epsilon^2$$

$$\mathbf{D}' \mathbf{D} \stackrel{!}{=} \mathbf{p}^2 - m^2c^2 \quad .$$

Wenn wir dann nämlich als neue Bewegungsgleichung fordern

$$\mathbf{P}|\Psi\rangle = 0$$

$$\mathbf{D}|\Psi\rangle = 0 \quad (6)$$

so folgt daraus die Schrödinger- bzw. die Klein-Gordon-Gleichung

$$\mathbf{P}' \mathbf{P}|\Psi\rangle = (\vec{\mathbf{p}}^2 - \epsilon^2)|\Psi\rangle = 0$$

$$\mathbf{D}' \mathbf{D}|\Psi\rangle = (\mathbf{p}^2 - m^2c^2)|\Psi\rangle = 0 \quad . \quad (7)$$

Es soll also gelten⁴

³ E als der Eigenwert der Energie eines freien Teilchens ist immer nicht-negativ, ϵ also immer eine reelle Zahl.

⁴In weiser Voraussicht wird beim Ausmultiplizieren die Multiplikationsreihenfolge der Unbekannten strikt beibehalten!

$$\begin{array}{l|l}
\alpha_i' \alpha_k \cdot \mathbf{p}_i \mathbf{p}_k & \alpha_\mu' \alpha_\nu \cdot \mathbf{p}^\mu \mathbf{p}^\nu \\
+ (-\alpha_k' \beta \epsilon + \beta' \epsilon \alpha_k) \cdot \mathbf{p}_k & + (-\alpha_\mu' \beta m c + \beta' m c \alpha_\mu) \cdot \mathbf{p}^\mu \\
- \beta' \beta \epsilon^2 & - \beta' \beta m^2 c^2 \\
\stackrel{!}{=} \vec{\mathbf{p}}^2 - \epsilon^2 & \stackrel{!}{=} \mathbf{p}^2 - m^2 c^2 \quad .
\end{array}$$

Daraus folgt als Bestimmungsgleichungen für die Unbekannten⁵

$$\begin{array}{l|l}
\alpha_i' \alpha_k + \alpha_k' \alpha_i \stackrel{!}{=} 2\delta_{ik} & \alpha_\mu' \alpha_\nu + \alpha_\nu' \alpha_\mu \stackrel{!}{=} 2g_{\mu\nu} \quad (8) \\
\beta' \alpha_k - \alpha_k' \beta \stackrel{!}{=} 0 & \beta' \alpha_\mu - \alpha_\mu' \beta \stackrel{!}{=} 0 \quad (9) \\
\beta' \beta \stackrel{!}{=} 1 & \beta' \beta \stackrel{!}{=} 1 \quad (10) \\
(i, k = 1, 2, 3) & (\mu, \nu = 0, 1, 2, 3)
\end{array}$$

Wie man sieht, unterscheiden sich die beiden Fälle ausschließlich durch Dimension und Metrik:

$$\begin{array}{l|l}
3 \text{ Dimensionen} & 4 \text{ Dimensionen} \\
\text{Euklidische Metrik} & \text{Minkowski-Metrik}
\end{array}$$

Man sieht leicht, dass das Gleichungssystem mit

$$\begin{array}{l|l}
10 \text{ Gleichungen} & 15 \text{ Gleichungen} \\
8 \text{ Unbekannten} & 10 \text{ Unbekannten}
\end{array}$$

für reelle oder komplexe Werte *keine Lösungen* hat, z.B. so⁶:

$$\begin{array}{llll}
\text{aus } \beta' \beta = 1 & \text{folgt} & : & \beta' = \beta^{-1} \quad (11) \\
\text{aus } \alpha_i' \alpha_k = 1 \quad \forall i = k & \text{folgt} & : & \alpha_i' = \alpha_i^{-1} \quad \forall i \quad (12) \\
\text{aus } \alpha_i' \alpha_k = -\alpha_k' \alpha_i, \quad \forall i \neq k & \text{folgt [mit (12)]} & : & \alpha_i^2 = -\alpha_k^2 \quad \forall i \neq k \quad ; \quad (13)
\end{array}$$

andererseits folgt aus

$$\beta' \alpha_i = \alpha_i' \beta \quad [\text{mit (11,12)}]: \quad \beta^{-1} \alpha_i = \alpha_i^{-1} \beta \quad \forall i \quad ,$$

⁵Die Symmetrisierung auf der linken Seite der ersten dieser Gleichungen ist notwendig, weil die rechte Seite symmetrisch definiert ist — sonst würden wir zuviel verlangen!

⁶Der Beweis ist hier für den “Pauli-Fall” (Euklidische Metrik) hingeschrieben; im “Dirac-Fall” (Minkowski-Metrik) geht aber alles völlig analog.

also⁷

$$\alpha_i^2 = \beta^2 \quad \forall i \quad ;$$

das aber ist ein Widerspruch zu (13) .

Man muss die Lösung des Problems also mit komplexeren algebraischen Objekten versuchen — z.B. mit (nicht vertauschbaren) *Matrizen*.

Für die weitere Rechnung benutzt man zweckmäßigerweise andere Variable. Dazu bemerkt man zunächst, dass wegen Gl. (10) β' überhaupt nur dann existiert, wenn β nicht-singulär ist — wir können das im weiteren also voraussetzen. Dann folgt zunächst aus $\beta' \cdot \beta = 1$ [Gl. (10)] auch $\beta \cdot \beta' = 1$. Damit kann man neue Variable γ wie folgt einführen⁸:

$$\left. \begin{array}{l} \gamma_k := \beta' \alpha_k \quad (= \alpha_k' \beta) \\ (k = 1, 2, 3) \end{array} \right| \begin{array}{l} \gamma_\mu := \beta' \alpha_\mu \quad (= \alpha_\mu' \beta) \quad ; \quad [\gamma^\mu = g^{\mu\rho} \gamma_\rho] \\ (\mu = 0, 1, 2, 3) \end{array}$$

Mit diesen Variablen schreiben sich die Bestimmungsgleichungen (8) bis (10) in der einfachen Form⁹

$$\left. \begin{array}{l} \gamma_i \gamma_k + \gamma_k \gamma_i \stackrel{!}{=} 2\delta_{ik} \end{array} \right| \begin{array}{l} \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu \stackrel{!}{=} 2g_{\mu\nu} \end{array} \quad (14)$$

▷ **Übung:** Nachrechnen!

Wir suchen nun eine Lösung unseres Problems, bei denen die γ *endlich-dimensionale Matrizen* (der Dimension N) sind. Aufgrund der Bedingungen (14) haben diese Matrizen die folgenden Eigenschaften¹⁰:

① Aus

$$\gamma_\mu \gamma^\mu = \mathbb{1} \quad (\mu = 0, 1, 2, 3)$$

folgt

$$\boxed{\text{Tr}(\gamma_\mu \gamma^\mu) = N} \quad ,$$

⁷Man beachte, dass hier $\beta' \alpha_i = \alpha_i \beta'$, falls die Variablen α_μ, β , wie vorausgesetzt, *Zahlen* sind!

⁸In diesen neuen Variablen haben wir nur noch 6 Gleichungen mit 3 Unbekannten (Pauli-Fall) bzw. 10 Gleichungen mit 4 Unbekannten (Dirac-Fall) — die Variablen β und β' “sind herausgefallen”. Präzise gesagt bedeutet das: man erhält für *jedes beliebige* (nicht-singuläre) β eine Lösung (wenn es überhaupt eine gibt). Dass es im Weiteren auf β nicht ankommt, sieht man besonders deutlich an den Ausdrücken in Gl. (16), zusammen mit den Bewegungsgleichungen (6).

⁹Mathematiker nennen diese algebraische Struktur eine “Clifford-Algebra” (der Begriff kommt außer an dieser Stelle sonst in der Physik nicht vor).

¹⁰ACHTUNG: In diesem Abschnitt [Punkte ① bis ⑥] soll die Summationskonvention *nicht gelten* — soweit Summen vorkommen, werden sie explizit ausgeschrieben!

Wie man am Beweisgang sieht, sind diese Eigenschaften ganz unabhängig davon, ob wir den Pauli- oder den Dirac-Fall betrachten. Sie gelten daher, obwohl hier mit griechischen Indizes (Minkowski-Metrik) geschrieben, in gleicher Weise auch mit lateinischen Indizes (Euklidischer Fall).

- ② aus $\text{Tr}(\gamma_\mu \gamma_\nu) = \text{Tr}(-\gamma_\nu \gamma_\mu) = -\text{Tr}(\gamma_\nu \gamma_\mu) = -\text{Tr}(\gamma_\mu \gamma_\nu)$ für $\mu \neq \nu$ folgt

$$\boxed{\text{Tr}(\gamma_\mu \gamma_\nu) = 0 \quad \text{für} \quad \mu \neq \nu} \quad ,$$

- ③ aus $\gamma_\mu \gamma_\nu = -\gamma_\nu \gamma_\mu$ für $\mu \neq \nu$ (die Matrizen “antikommutieren”) folgt

$$\det(\gamma_\mu \gamma_\nu) = \det(\gamma_\mu) \cdot \det(\gamma_\nu) = \det(-\gamma_\nu \gamma_\mu) = (-1)^N \cdot \det(\gamma_\nu) \cdot \det(\gamma_\mu) \quad ,$$

d.h.

$$\boxed{N \text{ ist gerade}} \quad ,$$

- ④ für $\mu \neq \nu$ gilt $\gamma_\mu \gamma_\nu \gamma_\nu = -\gamma_\nu \gamma_\mu \gamma_\nu$, woraus folgt [beachte $\text{Tr}(ABC) \equiv \text{Tr}(CAB)$!]

$$\text{Tr}(\gamma_\mu \gamma_\nu \gamma_\nu) = 0 \quad .$$

Wegen ① folgt hieraus

$$\boxed{\text{Tr}(\gamma_\mu) = 0} \quad ,$$

- ⑤ die Matrizen sind

$$\boxed{\text{linear unabhängig}} \quad .$$

Das sieht man so: sei $\sum_\mu c^\mu \gamma_\mu = 0$, mit irgendwelchen (reellen oder komplexen) Konstanten c^μ . Dann folgt trivial

$$0 = \gamma^\nu \sum_\mu c^\mu \gamma_\mu = \sum_\mu c^\mu \gamma^\nu \gamma_\mu \quad ,$$

also ist auch

$$\begin{aligned} 0 &= \text{Tr} \sum_\mu c^\mu \gamma^\nu \gamma_\mu = \text{Tr} \left(c^\nu \gamma^\nu \gamma_\nu + \sum_{\mu \neq \nu} c^\mu \gamma^\nu \gamma_\mu \right) \\ &= c^\nu \cdot \text{Tr}(\gamma^\nu \gamma_\nu) + \sum_{\mu \neq \nu} c^\mu \cdot \text{Tr}(\gamma^\nu \gamma_\mu) \\ &= c^\nu \cdot N + 0 = N c^\nu \quad . \end{aligned}$$

Es folgt, dass alle c^ν verschwinden müssen.

- ⑥ Für festes N gibt es gerade $(N^2 - 1)$ unabhängige spurlose $N \times N$ -Matrizen, d.h. Matrizen mit den Eigenschaften ④ und ⑤.

▷ **Übung:** Alles nachrechnen!

In der weiteren Analyse des Problems müssen wir jetzt die beiden Fälle unterscheiden.

2.1 Pauli-Matrizen

Gesucht werden 3 Matrizen mit den Eigenschaften ① bis ④. Es muss also gelten $N^2 - 1 \geq 3$, N gerade — für die Lösung mit minimalem N wäre also $N = 2$. Man findet (wohl am schnellsten durch geschicktes Probieren), dass die folgenden drei 2×2 -Matrizen tatsächlich eine Lösung der Clifford-Algebra-Gleichungen (14) darstellen:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Sie heißen “die **Pauli-Matrizen**”¹¹.

▷ **Übung:** Nachrechnen!

Die Lösung ist *nicht* eindeutig; z.B. ist ja offensichtlich eine zyklische Vertauschung der drei Matrizen trivialerweise wieder eine Lösung. Aber auch “Drehungen des Koordinatensystems” führen auf (äquivalente) Lösungen.

Die so aus dem Linearisierungsprogramm hergeleiteten Pauli-Matrizen haben eine bemerkenswerte Eigenschaft: ihre Vertauschungsrelationen untereinander sind (fast) die eines *Drehimpulses*:

$$[\sigma_j, \sigma_k] = 2\varepsilon_{jkm}\sigma_m \quad ; \quad (15)$$

▷ **Übung:** Nachrechnen!

die Matrizen $\vec{s} := \frac{1}{2}\vec{\sigma}$ haben also Drehimpuls-Vertauschungsrelationen. Überdies folgt aus den Beziehungen in Gl. (14)

$$\vec{s}^2 = \frac{3}{4} \cdot \mathbb{1} = \frac{1}{2} \left(\frac{1}{2} + 1 \right) \cdot \mathbb{1} \quad ,$$

der Eigenwert dieses ‘Drehimpulses’¹² ist also *immer* $s = 1/2$.

Man nennt diesen ‘inneren Drehimpuls’ den **spin** des Teilchens [vgl. hierzu Abschn. 2.3].

2.2 Dirac-Matrizen

Hier haben wir 4 Matrizen. Es muss also gelten $N^2 - 1 \geq 4$, N gerade; $N = 2$ scheidet deshalb aus. Die nächste Möglichkeit ist $N = 4$. Hier findet man (wieder durch richtiges Probieren) als Lösung die 4×4 -Matrizen

¹¹Diese 3 Matrizen sind also die gesuchten Lösungen der Gl. (14) für den Pauli-Fall. Entsprechend der universell üblichen Notation haben wir sie hier mit dem Buchstaben σ bezeichnet und reservieren den Buchstaben γ ab jetzt für die “Dirac-Matrizen” (s. Abschn. 2.2).

¹²in einem geeigneten System von Eigenzuständen; vgl. Abschn. 2.3.

$$\gamma_0 := \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{pmatrix} \quad \vec{\gamma} := \begin{pmatrix} \mathbf{0} & \vec{\sigma} \\ -\vec{\sigma} & \mathbf{0} \end{pmatrix}$$

Sie heißen “die **Dirac-Matrizen**” (jeder Eintrag steht für eine 2×2 -Matrix; die Vektorpfeile notieren wie üblich einen räumlichen Dreiervektor, und $\vec{\sigma}$ steht für die drei Pauli-Matrizen).

▷ **Übung:** Nachrechnen!

Auch sie sind nicht eindeutig¹³.

Es ergibt sich jetzt als Lösung unseres Ausgangsproblems (der ‘Linearisierung’ der Bewegungsgleichung) – Gl. (5) bzw. (6) –

$$\begin{array}{l|l} \beta' \mathbf{P} & := \vec{\sigma} \cdot \vec{\mathbf{p}} - \epsilon \\ \mathbf{P}' \beta & := \vec{\sigma} \cdot \vec{\mathbf{p}} + \epsilon \end{array} \quad \left| \quad \begin{array}{l} \beta' \mathbf{D} := \gamma_\mu \mathbf{p}^\mu - mc \\ \mathbf{D}' \beta := \gamma_\mu \mathbf{p}^\mu + mc \end{array} \right. \quad (16)$$

▷ **Übung:** Nachrechnen!

Diese Operatoren sind nicht einfach nur Operatoren in dem Hilbertraum, der strukturlose Teilchen beschreibt, sondern zusätzlich $N \times N$ -Matrizen in einem abstrakten N -dimensionalen Vektorraum. Damit das Sinn ergibt, müssen also die Zustände $|\Psi\rangle$, auf die diese Operatoren wirken, beschrieben werden durch *Vektoren* in diesem N -dimensionalen Vektorraum [$N = 2$ im Pauli-Fall, $N = 4$ im Dirac-Fall], deren Elemente ihrerseits gewöhnliche Zustandsvektoren (für strukturlose Teilchen) sind. Solche Objekte nennt man **Spinoren**. Mathematisch ist klar, dass ihr Auftreten die Einführung eines neuen, zusätzlichen *Freiheitsgrades* bedeutet; dessen physikalische Interpretation ist allerdings erst noch zu leisten¹⁴.

2.3 Pauli-Gleichung

Für die gesuchte in der Zeitableitung *lineare* Bewegungsgleichung ergibt sich also

$$(\vec{\sigma} \cdot \vec{\mathbf{p}} - \epsilon)|\Psi\rangle = 0 \quad , \quad (17)$$

die “**Pauli-Gleichung**”. Notieren wir einen ‘Pauli-spinor’ durch

$$|\Psi\rangle := \begin{pmatrix} |\psi_+\rangle \\ |\psi_-\rangle \end{pmatrix} \quad ,$$

so schreibt sich die Pauli-Gleichung in expliziter Form¹⁵:

$$\begin{pmatrix} \mathbf{p}_z - \epsilon & \mathbf{p}_x - i\mathbf{p}_y \\ \mathbf{p}_x + i\mathbf{p}_y & -(\mathbf{p}_z + \epsilon) \end{pmatrix} \begin{pmatrix} |\psi_+\rangle \\ |\psi_-\rangle \end{pmatrix} = 0 \quad , \quad (18)$$

¹³Die 4 Matrizen bilden zusammen einen Minkowski-Vierervektor; entsprechend führen hier z.B. *Lorentz-Transformationen* auf (äquivalente) Lösungen.

¹⁴Siehe hierzu Abschn. 3..

¹⁵Um der physikalischen Deutlichkeit willen schreiben wir hier (x, y, z) für die 3 räumlichen Indices.

▷ **Übung:** Nachrechnen!

woraus man durch Lösen dieses Gleichungssystems erhält

$$\begin{aligned}(\vec{\mathbf{p}}^2 - 2mE) |\psi_+\rangle &= 0 \\(\vec{\mathbf{p}}^2 - 2mE) |\psi_-\rangle &= 0 \quad ;\end{aligned}$$

▷ **Übung:** Nachrechnen!

d.h. also: jede der beiden Komponenten *erfüllt schon für sich die Schrödinger-Gleichung*. Physikalisch bedeutet das: das System ist in Bezug auf den zusätzlichen Freiheitsgrad *entartet*¹⁶.

Wie schon im Abschn. 2.1 festgestellt, ist *jeder* solche spinor ein Eigenzustand von \vec{s}^2 mit $s = 1/2$. Zusätzlich zu \vec{s}^2 kann man nun – wie immer bei einem Drehimpuls – auch *eine Komponente* diagonalisieren, z.B. (wie üblich) die z -Komponente. Wegen $\sigma_k^2 = \mathbb{1}$ [Gl. (14)] ist der Eigenwert $\pm \frac{1}{2}$. Man hat also die Eigenwertgleichung

$$\sigma_z |\Psi_\pm\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} |\Psi_\pm\rangle \stackrel{!}{=} \pm |\Psi_\pm\rangle$$

zu lösen. Die beiden Lösungen sind offenbar

$$\begin{aligned}|\Psi_+\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot |\psi\rangle \\|\Psi_-\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot |\psi\rangle \quad ,\end{aligned}$$

wobei $|\psi\rangle$ ein gewöhnlicher Zustandvektor für strukturlose Teilchen ist. Eine solche Notation hatten wir in Gl. (18) schon vorweggenommen.

Es stellt sich also heraus:

Der im Abschn. 2.2 schon erwähnte zusätzliche Freiheitsgrad entspricht im Pauli-Fall der räumlichen Orientierung einer drehimpulstigen Größe, die *spin* genannt wird.

Die *physikalische* Bedeutung des ‘spin’ als eines ‘inneren’ Drehimpulses wird im Abschn. 3. weiter klar werden.

2.4 Dirac-Gleichung

Noch interessanter ist der Dirac-Fall. Da wir es hier ständig mit Skalarprodukten von irgendwelchen Vierer-Vektoren mit den γ -Matrizen zu tun haben, führen wir zunächst eine Abkürzung ein, den “Feynman-dagger” (“Feynman-Dolch”):

$$\not{a} := \gamma_\mu a^\mu = \gamma^\mu a_\mu \quad .$$

Mit dieser Kurzschrift erhält man die gesuchte in der Zeitableitung *lineare* Bewegungsgleichung in einer höchst einprägsam Form¹⁷:

$$\boxed{(\not{p} - mc)|\Psi\rangle = 0} \quad , \quad (19)$$

¹⁶Das ist natürlich kein Zufall, sondern mit dieser Absicht war das ganze Verfahren ja konstruiert!

¹⁷In Worten: “der Impuls des Teilchens ist gleich mc ”.

Sie heißt die “**Dirac-Gleichung**”. Schreibt man – analog zu Gl. (18) – ‘obere’ und ‘untere’ Komponenten explizit aus

$$|\Psi\rangle =: \begin{pmatrix} |u\rangle \\ |v\rangle \end{pmatrix} ,$$

[$|u\rangle, |v\rangle$ sind hier Zweierspinoren], so erhält man

$$\begin{pmatrix} E - mc^2 & -c\vec{\sigma} \cdot \vec{\mathbf{p}} \\ c\vec{\sigma} \cdot \vec{\mathbf{p}} & -(E + mc^2) \end{pmatrix} \begin{pmatrix} |u\rangle \\ |v\rangle \end{pmatrix} = 0 . \quad (20)$$

▷ **Übung:** Nachrechnen!

Lösen dieses Gleichungssystems führt auf

$$\begin{aligned} \left(\frac{E^2}{c^2} - (\vec{\sigma} \cdot \vec{\mathbf{p}})^2 - m^2c^2 \right) |u\rangle &= 0 \\ \left(\frac{E^2}{c^2} - (\vec{\sigma} \cdot \vec{\mathbf{p}})^2 - m^2c^2 \right) |v\rangle &= 0 . \end{aligned}$$

▷ **Übung:** Nachrechnen!

Da $(\vec{\sigma} \cdot \vec{\mathbf{p}})^2 \equiv \vec{\mathbf{p}}^2$,

▷ **Übung:** Nachrechnen! [siehe dazu auch Abschn. 3..]

folgt

$$\begin{aligned} (\mathbf{p}^2 - m^2c^2) |u\rangle &= 0 \\ (\mathbf{p}^2 - m^2c^2) |v\rangle &= 0 . \end{aligned}$$

▷ **Übung:** Nachrechnen!

Alle 4 Komponenten von $|\Psi\rangle$ erfüllen also gleichermaßen die Klein-Gordon-Gleichung; das System ist jetzt aber *4-fach* entartet.

Neue Physik ergibt sich aber erst, wenn man *wechselwirkende Systeme* betrachtet, z.B. die Bewegung eines geladenen Teilchens in einem (äußeren, vorgegebenen) elektromagnetischen Feld. Das soll im nächsten Abschnitt geschehen.

3. Die Wechselwirkung mit dem elektromagnetischen Feld

3.1 Vorbemerkung

Die Wechselwirkung eines (mit der Ladung e) geladenen Teilchens mit einem vorgegebenen äußeren elektromagnetischen Feld erhält man aufgrund des Eichprinzips der Elektrodynamik durch die sogenannte “minimale Kopplung”. Man schreibt hierzu nicht nur Energie und Impuls des Teilchens in Form eines Vierervektors [Gl. (2)]

$$p^\mu := \{p^0, p^1, p^2, p^3\} = (E/c, \vec{\mathbf{p}}) ,$$

sondern auch die ‘elektromagnetischen Potentiale’:

$$a^\mu := \{a^0, a^1, a^2, a^3\} = \left(\Phi/c, \vec{A} \right) \quad ;$$

die Bewegungsgleichung für das *geladene* Teilchen erhält man dann aus der *freien* Gleichung, wenn man überall den ‘kinematischen’ Impuls p^μ durch den ‘kanonischen’ Impuls ersetzt:

$$p^\mu \longrightarrow \hat{p}^\mu := p^\mu - e a^\mu \quad , \quad (21)$$

oder, in nichtrelativistische Schreibweise,

$$\begin{aligned} E &\longrightarrow \hat{E} := E - e\Phi \\ \vec{p} &\longrightarrow \hat{p} := \vec{p} - e\vec{A} \quad . \end{aligned}$$

Die Schrödinger-Quantisierung (Übergang zur Ortsdarstellung) erhält man daraus wie vorher durch die Ersetzungsvorschrift in Gl. (4).

3.2 Pauli-Identität

Wollen wir diese “minimale Kopplung” auf die Pauli- oder Dirac-Gleichung anwenden, so ist Vorsicht geboten bei der Berechnung von Ausdrücken, die Pauli- oder Dirac-Matrizen enthalten, denn es handelt sich im Allgemeinen um *nicht-vertauschbare* Größen. So ist z.B. der Ausdruck $(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B})$ im Allgemeinen *nicht* gleich $(\vec{A} \cdot \vec{B})$, wie man das nach der gewöhnlichen Vektorrechnung vielleicht erwarten würde. Der Ausdruck ist vielmehr – für allgemeine Vektor-Operatoren \vec{A}, \vec{B} – unter strikter Beachtung der Reihenfolge auszuwerten.

Zunächst schreiben wir das sogenannte “Vektorprodukt” $(\vec{A} \times \vec{B})$ – das ja in Wahrheit ein *antisymmetrischer Tensor 2.Stufe* ist – in kovarianter Schreibweise als einen solchen Tensor¹⁸:

$$(\vec{A} \times \vec{B})_k := \varepsilon_{kmn} A_m B_n \quad .$$

Man kann daher die Summe zweier solcher Vektorprodukte mit vertauschter Reihenfolge schreiben in der Form¹⁹

$$(\vec{A} \times \vec{B})_k + (\vec{B} \times \vec{A})_k = \varepsilon_{kmn} [A_m, B_n] \quad .$$

▷ **Übung:** Nachrechnen!

Insbesondere gilt (!)

$$(\vec{A} \times \vec{A})_k = \frac{1}{2} \varepsilon_{kmn} [A_m, A_n] \quad .$$

Wir berechnen nun den Ausdruck

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = \sigma_k A_k \sigma_m B_m \quad .$$

Die Paulimatrizen erfüllen die algebraischen Beziehungen

$$\sigma_j \sigma_k = \begin{cases} \mathbb{1} & \text{für } j = k \\ i\varepsilon_{jkm} \sigma_m & \text{für } j \neq k \end{cases} \quad ;$$

¹⁸Beachte die Summationskonvention!

¹⁹Man sieht hier besonders deutlich, dass (und warum) die aus der elementaren Vektorrechnung bekannte Antisymmetrie des Vektorprodukts nur für *vertauschbare* Größen gilt!

— das ergibt sich aus ihren Vertauschungsrelationen Gl. (15), zusammen mit der Algebra in Gl. (14).

Sofern die Matrizen σ_k mit den Vektorkomponenten von \vec{A} und \vec{B} vertauschen, ergibt sich also

$$\begin{aligned}
 (\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) &= \sigma_k A_k \sigma_m B_m = \sigma_k \sigma_m A_k B_m = A_k B_k + i \varepsilon_{kmn} \sigma_n A_k B_m \\
 &= (\vec{A} \cdot \vec{B}) + i \sigma_n \varepsilon_{kmn} A_k B_m = (\vec{A} \cdot \vec{B}) + i \sigma_n (\vec{A} \times \vec{B})_n \\
 &= \boxed{(\vec{A} \cdot \vec{B}) + i \vec{\sigma} \cdot (\vec{A} \times \vec{B})} .
 \end{aligned} \tag{22}$$

Diese wichtige allgemeine Formel heißt die “**Pauli-Identität**”.

Für den uns interessierenden Spezialfall $\vec{A} := \vec{B} := \hat{\mathbf{p}}$ haben wir ²⁰

$$\begin{aligned}
 (\vec{\sigma} \cdot \hat{\mathbf{p}})^2 &= (\hat{\mathbf{p}})^2 + \frac{i}{2} \varepsilon_{kmn} \sigma_k [\mathbf{p}_m - eA_m, \mathbf{p}_n - eA_n] \\
 &= (\hat{\mathbf{p}})^2 + \frac{i}{2} \varepsilon_{kmn} \sigma_k \{-e [\mathbf{p}_m, A_n] + e [\mathbf{p}_n, A_m]\} \\
 &= (\hat{\mathbf{p}})^2 - ie \varepsilon_{kmn} \sigma_k [\mathbf{p}_m, A_n] .
 \end{aligned}$$

In der Ortsdarstellung ist nun (das ist der Witz der ganzen Rechnerei!) $\vec{\mathbf{p}} = -i\hbar\nabla$. Da aber das Vektorpotential \vec{A} in der Regel *ortsabhängig* ist, *vertauscht $\vec{\mathbf{p}}$ nicht mit \vec{A}* ; vielmehr gilt

$$\begin{aligned}
 \varepsilon_{kmn} [\mathbf{p}_m, A_n] &= \varepsilon_{kmn} (-i\hbar) \nabla_m A_n = (-i\hbar) (\text{rot } \vec{A})_k \\
 &= (-i\hbar) \vec{B}_k \quad !
 \end{aligned}$$

Insgesamt erhalten wir also

$$(\vec{\sigma} \cdot \hat{\mathbf{p}})^2 = (\hat{\mathbf{p}})^2 - e\hbar(\vec{\sigma} \cdot \vec{B}) .$$

3.3 Volle Pauli-Gleichung

Multipliziert man die Pauli-Gleichung

$$(\vec{\sigma} \cdot \hat{\mathbf{p}} - \sqrt{2m\hat{E}}) |\Psi\rangle = 0$$

jetzt mit $\mathbf{P}'\beta = \vec{\sigma} \cdot \hat{\mathbf{p}} + \sqrt{2m\hat{E}}$, so hat man

$$\left\{ \frac{(\vec{\sigma} \cdot \hat{\mathbf{p}})^2}{2m} - \hat{E} \right\} |\Psi\rangle = \left\{ \frac{(\hat{\mathbf{p}})^2}{2m} - \frac{e\hbar}{2m} (\vec{\sigma} \cdot \vec{B}) - \hat{E} \right\} |\Psi\rangle = 0$$

oder, voll ausgeschrieben,

$$\boxed{\left\{ \frac{(\mathbf{p} - e\vec{A})^2}{2m} + e\Phi - \frac{e\hbar}{2m} (\vec{\sigma} \cdot \vec{B}) \right\} |\Psi\rangle = E|\Psi\rangle} \tag{23}$$

²⁰Zur Erinnerung: $\hat{\mathbf{p}} := \mathbf{p} - e\mathbf{A}$.

Das ist die Pauli-Gleichung in ihrer vollen Schönheit. Sie zeigt die energetische Aufspaltung der beiden spin-Richtungen im Magnetfeld, wie man sie auch im Experiment beobachtet (z.B. beim anomalen Zeeman-Effekt am Wasserstoffatom) — der spin ist also ein physikalisch reales Phänomen!

Man kann das noch etwas weiter interpretieren. In der klassischen Elektrodynamik ist die Energie eines magnetischen Dipolmoments $\vec{\mu}$ in einem Magnetfeld \vec{B} gegeben durch $E = \vec{\mu} \cdot \vec{B}$; das magnetische Dipolmoment eines Teilchens mit Ladung e , Masse m und Drehimpuls \vec{J} ($= \hbar \frac{\vec{J}}{\hbar}$) ist $\vec{\mu} = \frac{e}{2m} \vec{J}$ ($= \frac{e\hbar}{2m} \frac{\vec{J}}{\hbar}$).

Die Größe $\mu_0 := \frac{e\hbar}{2m}$ ist also der Skalenfaktor, der klassisch den Drehimpuls des Teilchens (in Einheiten von \hbar) mit seinem magnetischen Moment verknüpft; er heißt *Bohrsches Magneton*.

Vergleicht man nun diesen klassischen Sachverhalt mit dem Term $-\frac{e\hbar}{2m}(\vec{\sigma} \cdot \vec{B})$ in der Pauli-Gleichung (23), so sieht man zweierlei:

1. Das Elektron, obwohl es offenbar auch quantenmechanisch ein Teilchen ‘ohne Ausdehnung’ zu sein scheint, besitzt tatsächlich einen ‘inneren Drehimpuls’ — daher der (etwas *zu* anschauliche) Name ‘spin’,
2. das durch diesen ‘spin’ erzeugte magnetische Moment ist offenbar *doppelt so groß*, als man aus der klassischen Vorstellung erwarten würde, denn der ‘innere Drehimpuls’ ist ja $\vec{s} = \frac{1}{2}\vec{\sigma}$ [vgl. Gl. (15)]. Dieses ‘doppelte’ magnetische Moment aber ist gerade, was man experimentell findet (anomaler Zeeman-Effekt)!

Da das magnetische Moment auch der anderen Elementarteilchen (Proton, Neutron usw.) sich vom klassisch erwarteten Wert μ_0 unterscheidet (für die verschiedenen Teilchen verschieden!), definiert man allgemein eine Art ‘Korrektur-Faktor’, den “Landé-schen g -Faktor”, durch

$$\vec{\mu} =: g \cdot \mu_0 \vec{J} \quad ;$$

der jeweilige Wert des g -Faktors ist dann als ‘Eigenschaft’ des betreffenden Teilchens anzusehen — wobei $g = 2$ für spin- $\frac{1}{2}$ -Teilchen eben die Voraussage der Pauli-Dirac-Theorie ist²¹.

3.4 Volle Dirac-Gleichung

Führt man in der Dirac-Gleichung (20) die ‘minimale Kopplung’ durch, so findet man

$$\begin{pmatrix} E - e\Phi - mc^2 & -c\vec{\sigma} \cdot (\vec{\mathbf{p}} - e\vec{A}) \\ c\vec{\sigma} \cdot (\vec{\mathbf{p}} - e\vec{A}) & -(E - e\Phi + mc^2) \end{pmatrix} \begin{pmatrix} |u\rangle \\ |v\rangle \end{pmatrix} = 0 \quad . \quad (24)$$

²¹Diese Vorhersage – historisch zuerst von Dirac anhand der Dirac-Gleichung gemacht (obwohl sie, wie wir hier gesehen haben, bereits aus der *nichtrelativistischen* Pauli-Gleichung folgt!) – ist einer der großen Erfolge der Quantentheorie und hat das Vertrauen in die Richtigkeit der Dirac-Gleichung für das Elektron von Anbeginn an ganz entscheidend gestärkt.

Ganz genau genommen zeigt das Experiment allerdings eine geringfügige Abweichung des g -Faktors vom Dirac-schen Wert 2; der heutige Bestwert(!) ist

$$\frac{g-2}{2} = (0.001\,159\,652\,193 \pm 0.000\,000\,000\,010) \quad .$$

Eine befriedigende Erklärung für diese Abweichung liefert erst die voll-relativistische (feldtheoretische) Behandlung des Problems.

Diese Gleichung hat eine bemerkenswerte Symmetrie (die im Pauli-Fall *nicht* auftritt!). Geht man nämlich mit der Substitution

$$\begin{aligned} E &\longrightarrow \tilde{E} := -E \\ e &\longrightarrow \tilde{e} := -e \\ \vec{p} &\longrightarrow \tilde{\vec{p}} := -\vec{p} \end{aligned} \tag{25}$$

in die Gleichung ein, so folgt mit Einsetzen und Umordnen

$$\begin{pmatrix} \tilde{E} - \tilde{e}\Phi - mc^2 & -c\vec{\sigma} \cdot (\tilde{\vec{p}} - \tilde{e}\vec{A}) \\ c\vec{\sigma} \cdot (\tilde{\vec{p}} - \tilde{e}\vec{A}) & -(\tilde{E} - \tilde{e}\Phi + mc^2) \end{pmatrix} \begin{pmatrix} |v\rangle \\ |u\rangle \end{pmatrix} = 0 \quad ;$$

das ist aber die *ursprüngliche Gleichung*, wenn man zusätzlich zu den obigen Substitutionen (25) auch noch *obere und untere* Komponenten $|u\rangle$ und $|v\rangle$ vertauscht:

$$\begin{aligned} |u\rangle &\longrightarrow |\tilde{u}\rangle := |v\rangle \\ |v\rangle &\longrightarrow |\tilde{v}\rangle := |u\rangle \end{aligned} \tag{26}$$

Mit dem Übergang $e \longrightarrow -e$ scheinen wir ein völlig anderes System zu beschreiben, nämlich ein Teilchen *mit der umgekehrten Ladung*. Beides ist aber *in derselben Gleichung* enthalten, wenn man nur etwas uminterpretiert:

Die Dirac-Gleichung beschreibt mit ihren 4 Komponenten Teilchen *entgegengesetzter Ladung* (aber *gleicher* Masse!) in *einer gemeinsamen* Gleichung. Die Vertauschung der oberen und unteren Komponenten bedeutet physikalisch den Übergang zum ‘Gegenteil’: *umgekehrte* Ladung, *umgekehrter* Impuls, *negative* Energie, alles in allem einfach das ‘Fehlen’ eines Teilchens \Rightarrow ein **Antiteilchen!**

Ein Antiteilchen ist *das Fehlen* eines Teilchens.

Geht man zu niedrigen Energien ($E \sim mc^2$, d.h. $|\vec{p}| \ll mc$), so sieht man, dass die *unteren* Komponenten, $|v\rangle$, klein gegen die oberen, $|u\rangle$, werden. Denn dann wird

$$|v\rangle = \frac{c}{E + mc^2} (\vec{\sigma} \cdot \vec{p}) |u\rangle \sim \frac{c}{2mc^2} (\vec{\sigma} \cdot \vec{p}) |u\rangle \ll |u\rangle \quad .$$

Das bedeutet: Antiteilchen spielen bei niedrigen Energien keine Rolle; sie mischen sich erst ein – im Wortsinn: nämlich in den Zustand! –, wenn untere und obere Komponenten vergleichbar werden, d.h. bei kinetischen Energien, die vergleichbar sind mit der Ruhemasse.

Das ist das “relativistische Phänomen”, das die Dirac-Gleichung neu liefert gegenüber der nichtrelativistischen Theorie! Der ‘spin’ (und das magnetische Moment!) ist es – entgegen mancher Behauptung – *nicht*.²²

Ich danke Herrn *Oliver Henke* für eine sorgfältige Durchsicht und die Entdeckung zahlreicher Satzfehler in einer früheren Version dieses Skripts.

²²Der erste Hinweis darauf, dass man die Pauli-Gleichung (samt magnetischem Moment des Elektrons!) auch ganz im Rahmen der nichtrelativistischen Theorie herleiten kann, stammt von Galindo [Gal 61].

Literatur

- [Dir 28] P.A.M. Dirac, Proc.Roy.Soc.(London) **A117**, 610 (1928); Proc.Roy.Soc.(London) **A118**, 351 (1928)
- [Lev 67] J.-M. Lévy-Leblond, Comm.math.Phys. **6**, 286 (1967)
- [Gal 61] A. Galindo, Sánchez del Río, Am.J.Phys. **29**, 582 (1961)