LIE-GROEPEN IN DE FYSICA

Instituut voor Theoretische Fysica
Universiteit Utrecht
Faculteit Natuur- en Sterrenkunde

2003

M.J.G. Veltman
B.Q.P.J. de Wit
G. 't Hooft

INHOUD[1]

[1] Dit kollege is in zijn oorspronkelijke vorm opgezet door M. Veltman, en achtereenvolgens op enkele punten gewijzigd door G. 't Hooft en B. de Wit. Met ingang van de kursus 1987/88 is de kollegestof verder uitgebreid.

1. Inleiding

Veel systemen die we in de natuurkunde bestuderen vertonen een bepaalde vorm van symmetrie, en veel van de "natuurwetten" die we kennen zijn invariant onder bepaalde transformaties. De verzameling van alle mogelijke symmetrietransformaties vormt wat de wiskundigen een "groep" noemen. Zo vormen de spiegelingen ten opzichte van een vlak een groep die twee elementen bevat: de spiegelingstransformatie en de identiteit. Maar ook de verzameling van driedimensionale draaiingen, de verzameling van Lorentz-transformaties, of de verzameling van driedimensionale translaties zijn een groep, maar nu bestaande uit een oneindig aantal elementen. Om voor de hand liggende redenen noemt men groepen met een eindig aantal elementen *discrete* groepen; groepen van transformaties die op continue wijze van een aantal parameters afhangen, zoals de draaiingen, die gedefinieerd zijn in termen van hoekparameters, noemt men *continue* groepen.

De symmetrie van een systeem geeft aanleiding tot bepaalde relaties tussen observeerbare grootheden, waaraan met oneindig grote precisie is voldaan en die niet afhankelijk zijn van de eigenschappen van de in dat systeem werkzame krachten. In het waterstofatoom vindt men bijvoorbeeld dat de energieën van verschillende toestanden precies gelijk zijn als een gevolg van de rotatie-invariantie van dit systeem. Het komt echter ook vaak voor dat de symmetrie van een fysisch systeem slechts bij benadering is gerealiseerd. Een oneindig kristal is bijvoorbeeld invariant onder translaties met een veelvoud van de atomaire roosterafstand. In de werkelijkheid heeft een kristal echter eindige afmetingen, en dit verstoort de translatiesymmetrie. Niettemin, als het kristal een voldoende aantal atomen bevat, dan zullen deze verstoringen weinig invloed hebben op de eigenschappen van het kristal, tenzij men juist die eigenschappen van het kristal bestudeert die direkt samenhangen met het randoppervlak.

Een ander voorbeeld van een symmetrie die slechts bij benadering gerealiseerd is, vindt men in de elementaire-deeltjesfysika. Het z.g. Δ^+-deeltje, een van de aangeslagen toestanden van de nukleonen, blijkt te vervallen in een nukleon en een ander deeltje, π-meson, of ook wel pion genaamd. Er bestaan twee soorten nukleonen, neutronen en protonen, en drie soorten π-mesonen, namelijk de elektrisch geladen pionen π^+ en π^- en het neutrale pion π^0. Aangezien de elektrische lading in Δ^+-verval behouden moet zijn, kunnen we twee mogelijke vervalswijzen onderscheiden:

$$\Delta^+ \to n\,\pi^+ \quad \text{en} \quad \Delta^+ \to p\,\pi^0\,. \tag{1.1}$$

Het is opvallend dat het tweede verval twee maal zo vaak voorkomt als het eerste, een feit dat moeilijk verklaard kan worden uit het verschil in elektrische lading van de vervalsprodukten. Een natuurlijke verklaring voor deze faktor 2 zou kunnen volgen uit symmetriebeschouwingen. Dit is niet zo vreemd als op het eerste gezicht lijkt, aangezien protonen en neutronen bijna van gelijke massa zijn, evenals de drie soorten pionen en de vier soorten Δ-deeltjes die in de natuur voorkomen (zie tabel). Deze overeenkomst in massa, en, naar later zal blijken, ook de faktor 2 in de twee vervalswijzen (1.1), kan verklaard worden door aan te nemen dat de natuur invariant is onder z.g. *isospin*-transformaties. De term isobarische spin, of kortweg isospin, werd in 1932 geïntroduceerd door Heisenberg, die getroffen werd door het feit dat het proton en het neutron een bijna gelijke massa hebben,

nukleonen	pionen	Δ-deeltjes
$m_{\text{proton}} \approx 938$ MeV$/c^2$ $m_{\text{neutron}} \approx 939$ MeV$/c^2$	$m_{\pi^+} \approx 140$ MeV$/c^2$ $m_{\pi^0} \approx 135$ MeV$/c^2$ $m_{\pi^-} \approx 140$ MeV$/c^2$	$m_{\Delta^{++}} \approx 1231$ MeV$/c^2$ $m_{\Delta^+} \approx 1232$ MeV$/c^2$ $m_{\Delta^0} \approx 1233$ MeV$/c^2$ $m_{\Delta^-} \approx 1235$ MeV$/c^2$

Table 1: De massa's van nukleonen, pionen en Δ-deeltjes, uitgedrukt in MeV$/c^2$.

terwijl ze afgezien van hun elektrische lading, ook overeenkomen in andere eigenschappen. De nukleonen vormen dus een doublet, net zoals elektronen een doubletstruktuur vertonen als gevolg van het feit dat er twee mogelijke spinoriëntaties zijn voor elektrontoestanden (vandaar de naam isobarische *spin*). Later bleek dat elementaire deeltjes met bijna gelijke massa's altijd kunnen worden gegroepeerd in z.g. isospin-multipletten. Zo vormen de nukleonen een isospin-doublet, de pionen een isospin-triplet en de Δ-deeltjes een isospin-quadruplet. De deeltjes in een gegeven multiplet hebben allen ongeveer dezelfde massa, maar een verschillende elektrische lading. De ladingsverdeling is zoals aangegeven in de tabel: géén twee deeltjes in een multiplet hebben dezelfde lading en de deeltjes kunnen altijd zo geordend worden dat het verschil in lading tussen twee opeenvolgende deeltjes precies een elementaire ladingseenheid is.

Echter, het zal duidelijk zijn dat isospin invariantie op zijn best een benadering kan zijn, aangezien de massa's der nukleonen, pionen en Δ-deeltjes enigszins afhankelijk blijken te zijn van de elektrische lading van de deeltjes. De massaverschillen binnen een multiplet zijn slechts van de orde van een paar procent en dit is i.h.a. de graad van nauwkeurigheid die men kan verwachten voor theoretische voorspellingen gebaseerd op isospin-invariantie.

Het bovenstaande voorbeeld betreft een toepassing van de groepentheorie in de natuurkunde der elementaire deeltjes, maar invariantie-eigenschappen spelen in bijna iedere tak van de natuurkunde een grote rol. In de atoomfysika hebben we vaak te maken met de consekwenties van rotatie-invariantie, in de kernfysika met rotatie- en isospin-invariantie, in de vaste-stoffysika met de invariantie onder discrete translaties en draaiingen. Ook in de veldentheorie spelen symmetrietransformaties een grote rol. Een bijzonder soort transformatie komen we bijvoorbeeld tegen in de electrodynamica. Hier kunnen de electrische en magnetische velden worden uitgedrukt in de zogenaamde *vectorpotentiaal* $A_\mu(x)$, die we schrijven in een relativistische vier-vector notatie ($\mu = 1, 2, 3, 0$),

$$A_\mu = (\mathbf{A}(x), -c^{-1}\,\phi(x)), \qquad x^\mu = (\mathbf{x}, ct), \tag{1.2}$$

waarbij ϕ de potentiaal en $\mathbf{A}$ de driedimensionale vectorpotentiaal aangeeft; c is de lichtsnelheid. De electrische en magnetische velden zijn gedefinieerd door

$$\mathbf{E} = -\nabla\phi - c^{-1}\,\frac{\partial \mathbf{A}}{\partial t}, \tag{1.3}$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{1.4}$$

Men kan nu eenvoudig nagaan dat $\mathbf{E}$ en $\mathbf{B}$ niet veranderen onder zogenaamde *ijktrans-*

formaties

$$\mathbf{A}(x) \rightarrow \mathbf{A}(x) + \nabla\Lambda(x), \tag{1.5}$$

$$\phi(x) \rightarrow \phi(x) - c^{-1}\frac{\partial}{\partial t}\Lambda(x), \tag{1.6}$$

of, in vier-vector notatie,

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu\Lambda(x). \tag{1.7}$$

De karakteristieke eigenschap van ijktransformaties is, dat zij afhankelijk zijn van een willekeurige functie $\Lambda(x)$ van plaats en tijd. Het feit dat $\mathbf{E}$ en $\mathbf{B}$ niet veranderen onder ijktransformaties betekent dat elektromagnetische verschijnselen *ijkinvariant* zijn.

Laten we deze transformatie nu eens op een andere manier schrijven,

$$A_\mu \rightarrow i\,\mathrm{e}^{i\Lambda}(\partial_\mu - iA_\mu)\,\mathrm{e}^{-i\Lambda}, \tag{1.8}$$

zodat we nu te maken hebben met een ruimte-tijd-afhankelijke fasefactor $\exp(i\Lambda(x))$. Fasefactoren definiëren een groep, die $U(1)$ wordt genoemd: de groep van 1×1 unitaire matrices. In dit geval is dus sprake van een eenvoudige groep, maar het blijkt dat er ook zulke theorieën bestaan gebaseerd op andere (continuë) groepen, die veel minder eenvoudig zijn. Dergelijke theorieën worden aangeduid met de naam *ijktheorieën* en het veld A_μ wordt het *ijkveld* genoemd. Voor meer algemene groepen blijkt dat we verschillende van dergelijke ijkvelden nodig hebben. Zulke ijktheorieën blijken een belangrijke rol te spelen in de unificatie van de fundamentele wisselwerkingen tussen elementaire deeltjes. Electrodynamica, de ijktheorie van $U(1)$, blijkt de meest eenvoudige variant te zijn van deze klasse van theorieën.

Verrassend genoeg blijkt de theorie van de zwaartekracht, Einstein's algemene relativiteitstheorie, ook een ijktheorie te zijn, zij het van een wat ander type. Deze theorie kan worden beschouwd als de ijktheorie van de algemene coördinatentranformaties, algemene herparametrisaties van ruimte en tijd,

$$x^\mu \rightarrow x^\mu + \xi^\mu(x). \tag{1.9}$$

Het ijkveld is nu het gravitationele veld, dat de vorm aanneemt van een metriek, die gebruikt kan worden om afstanden en hoeken te definiëren in de vierdimensionale ruimte-tijd.

Het feit dat de ijktransformaties geassocieerd zijn met een abstracte groep, en tegelijkertijd ruimte-tijd-afhankelijk zijn, kan aanleiding geven tot interessante verschijnselen van topologische aard. Voorbeelden hiervan zijn de fluxquantisatie in supergeleiders, het Aharonov-Bohm effekt in de quantummechanika en magnetische monopolen. Om de relevantie van de topologie te illustreren, beschouwen we nogmaals de groep van $U(1)$ ijktransformaties, maar nu in een tweedimensionale ruimte (ofwel in een situatie die niet van de tijd en slechts van twee van de drie plaatscoördinaten afhangt). Laat $\psi(x,y)$ een complexe funktie zijn die transformeert onder deze ijktransformaties, m.a.w.,

$$\psi(x,y) \rightarrow \mathrm{e}^{i\Lambda(x,y)}\psi(x,y). \tag{1.10}$$

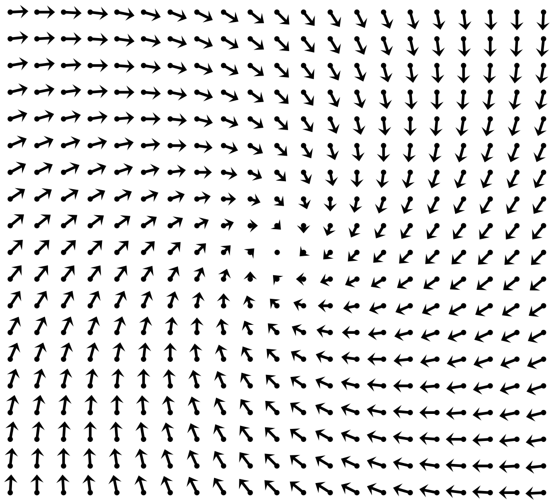

Figure 1: De fasehoek van $\psi(x,y)$ aangegeven door een pijltje (waarvan de lengte onbelangijk is, maar bijvoorbeeld gelijk genomen kan worden aan $|\psi(x,y)|$) op de verschillende plaatsen in het x-y vlak. De functie heeft een nulpunt in de oorsprong.

Een voorbeeld van een dergelijke funktie is de golffunktie in de quantummechanika. Uit het feit dat de fase van ψ op iedere plaats afzonderlijk kan worden veranderd door toepassing van een ijktransformatie, kunnen we konkluderen dat de fase van ψ in feite irrelevant is voor de beschrijving van het systeem. Toch is dit niet helemaal het geval. Bekijk bijvoorbeeld een funktie die in de oorsprong gelijk is aan nul. Neem nu een gesloten curve in het x-y vlak, en ga na hoe de fase van $\psi(x,y)$ verandert langs de curve. Na het volledig doorlopen van de curve zal de fase niet altijd dezelfde waarde aannemen als de beginwaarde; als we aannemen dat $\psi(x,y)$ enkelwaardig is op het vlak, dan moet het faseverschil echter gelijk zijn aan $2\pi n$, waarbij n een willekeurig *geheel* getal is. Dit getal wordt het *windingsgetal* genoemd. Een voorbeeld van een situatie met windingsgetal $n=1$ is afgebeeld in fig. 1; de fasehoek maakt een volledige draaiing over 2π wanneer we de functie $\psi(x,y)$ volgen langs een curve die een volledige draaiing over 2π beschrijft in het x-y vlak. Men kan zich gemakkelijk situaties voorstellen met andere windingsgetallen. Het geval $n=0$ treedt bijvoorbeeld op wanneer de fase van $\psi(x,y)$ constant is. Als we de functie $\psi(x,y)$ op continue wijze veranderen dan zal het windingsgetal gelijk blijven. Vandaar dat het windingsgetal een *topologische invariant* wordt genoemd. Dat betekent ook dat het windingsgetal niet zal veranderen onder de ijktransformaties (1.10), mits we ons beperken tot ijktransformaties die over het gehele vlak gedefinieerd zijn. Merk op dat de waarde van het windingsgetal niet afhangt van de keuze van de gesloten curve om de oorsprong, zolang we geen mogelijke andere nulpunten van de functie $\psi(x,y)$ gaan omsluiten. Uit dit alles volgt dat, hoewel we lokaal, d.w.z. in een punt en zijn direkte omgeving, de fase van ψ gelijk kunnen maken aan nul, dit alleen maar globaal, d.w.z. over het gehele vlak, bewerkstelligd kan worden als het windingsgetal gelijk is aan nul.

Een soortgelijke situatie doet zich voor met de vectorpotentiaal. Beschouw weer het tweedimensionale vlak, en neem aan dat we te maken hebben met een statisch magnetische veld dat overal gelijk is aan nul, met uitzondering van een beperkt gebied rond de oorsprong. In dit gebied zal $\mathbf{A}$ niet gelijk kunnen zijn aan aan nul, gezien de relatie (1.4). Echter, in het buitengebied waar $\mathbf{B}$ gelijk is aan nul, is er geen reden waarom $\mathbf{A}$ niet ook gelijk zou kunnen zijn aan nul. Inderdaad kunnen we laten zien dat we in ieder punt $\mathbf{A}$ gelijk kunnen maken aan nul middels een geschikt gekozen ijktransformatie. Dit resultaat geldt echter alleen lokaal, zoals we kunen zien door de volgende kringintegraal te beschouwen,

$$\Phi[C] = \oint_C A_i \, \mathrm{d}x^i, \tag{1.11}$$

waarbij C een gegeven (gesloten) curve is. Het is eenvoudig na te gaan dat $\Phi[C]$ niet verandert onder een ijktransformatie (1.5). Anderzijds weten we uit de theorie van het magnetisme dat $\Phi[C]$ evenredig moet zijn met de magnetische flux voor het door de curve C omsloten oppervlak. Passen we dit nu toe op de boven geschetste situatie. Neem de curve C om de oorsprong in het gebied waar $\mathbf{B}$ gelijk is aan nul. De totale magnetische flux door het omsloten gebied is nu niet noodzakelijk gelijk aan nul, waaruit volgt dat $\Phi[C]$ niet noodzakelijk gelijk is aan nul. Als dit laatste niet het geval is, kunnen we $\mathbf{A}$ dus niet wegtransformeren in het buitengebied, hoewel we $\mathbf{A}$ wel lokaal naar nul kunnen brengen[2]. Merk op dat de magnetische flux hier dezelfde rol speelt als het windingsgetal in het vorige voorbeeld. Inderdaad zijn er situaties waarbij deze twee grootheden lineair met elkaar samenhangen. We spreken dan van fluxquantisatie. Voorbeelden hiervan zijn de z.g. vortex-oplossingen in supergeleiders en de magnetische monopolen. Het zou ons echter te ver voeren om op de details daarvan nu in te gaan.

De rol die de groepentheorie in de fysica speelt, is hiermee enigszins aangegeven. In deze kolleges zullen we ons voornamelijk beperken tot de draaiingsgroep, en dit vooral in de context van de quantummechanika. Vele van de belangrijkste uitgangspunten kunnen aan dit geval duidelijk gemaakt worden, en de behandeling kan in het algemeen redelijk transparant, aanschouwelijk en mathematisch ongecompliceerd blijven. Het is geenszins de bedoeling om een komplete wiskundige behandeling te geven; het doel is om de relevantie van de groepentheorie zo duidelijk mogelijk aan te geven. We zullen dan ook verschillende fysische toepassingen uitgebreid behandelen. De draaiingsgroep is een voorbeeld van een z.g. compacte Lie-groep. In de meeste toepassingen zal het gaan om *representaties* van deze groep. De representatietheorie voor dergelijke groepen is geheel bekend in de wiskunde. Een redelijke voorkennis van de lineaire algebra (matrices, inprodukten, sporen, functies en afgeleiden van matrices, e.d.) is noodzakelijk. Volledigheidshalve zijn enkele van de belangrijkste eigenschappen van matrices samengevat in een aantal appendices.

[2]Dit veroorzaakt een interessant quantummechanisch effect bij elektronen die zich buiten een magnetisch veld bevinden, nl. het *Aharonov-Bohm effect*.

2. Quantummechanica en rotatie-invariantie

De quantummechanica leert dat elk fysisch systeem beschreven wordt door een (doorgaans complexe) golffunctie. Deze golffunctie is oplossing van een differentiaalvergelijking (bv. de Schrödingervergelijking indien de niet-relativistische benadering gemaakt kan worden) met door de fysische omstandigheden gesuggereerde randvoorwaarden. We zullen ons nu niet verdiepen in de problemen die zich voordoen bij het bepalen van de golffunctie in allerlei gevallen, maar we zijn geïnteresseerd in de eigenschappen van golffuncties welke voortvloeien uit het feit dat de natuur bepaalde symmetrieën vertoont.

Een van de duidelijkste symmetrieën welke zich in het dagelijks leven manifesteert, is de invariantie van de natuurwetten onder rotaties van de driedimensionale ruimte. Een waarnemer verwacht dat zijn meetresultaten onafhankelijk zullen zijn van de oriëntatie van zijn apparatuur in de ruimte, aangenomen althans dat de experimentele opstelling geen wisselwerking heeft met de omgeving. Zo verwacht men niet dat de tijd die een horloge aanwijst afhangt van de ruimtelijke oriëntatie van dat horloge, of dat de werking van een rekenmachine verandert na een draaiing. De rotatiesymmetrie is terug te vinden in de fundamentele vergelijkingen van de natuurkunde: de wetten van Newton en Maxwell en de Schrödingervergelijking zijn bijvoorbeeld oriëntatie-onafhankelijk. Preciezer geformuleerd: de natuurwetten zijn invariant onder rotaties in de driedimensionale ruimte.

Het is de bedoeling na te gaan wat de consequenties zijn van deze rotatie-invariantie voor de golffuncties. Uit de klassieke mechanica is het bekend dat rotatie-invariantie van een systeem dat geen wisselwerking met de omgeving heeft, aanleiding geeft tot de wet van behoud van impulsmoment: voor een dergelijk systeem is de impulsmoment-vector een constante van beweging. Een dergelijke behoudswet, die in feite niets te maken heeft met de details van het krachtenspel, maar volgt uit zeer algemene overwegingen, moet ook in de quantummechanica op een of andere wijze terug te vinden zijn. We verwachten daarom in de golffunctie iets aan te treffen dat correspondeert met impulsmoment, en dat constant is in de tijd. We zullen hier nu niet verder op in gaan, maar we merken op dat het gedrag van een golffunctie onder rotaties iets te maken heeft met impulsmoment.

Om nu wat concreter te kunnen zijn beschouwen we een golffunctie ψ die van allerlei variabelen kan afhangen en oplossing is van een of andere lineaire differentiaalvergelijking:

$$\mathcal{D}\psi = 0 \tag{2.1}$$

De precieze vorm van $\mathcal{D}$ is niet belangrijk, alleen het feit dat $\mathcal{D}$ invariant is onder rotaties is voor ons van belang. Een voorbeeld is de Schrödingervergelijking voor de zwaartepuntsaf- hankelijkheid van de golffunctie van een vrij deeltje,

$$\left[\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}\right)+i\hbar\frac{\partial}{\partial t}\right]\psi(\vec{x},t)=0. \tag{2.2}$$

Laten we nu aantonen dat deze differentiaalvergelijking rotatie-invariant is. Laat de plaatsvector $\vec{x}$ na een rotatie overgaan in een andere vector met coördinaten x_i', die geschreven kunnen worden als

$$x'_i = \sum_j R_{ij}\, x_j. \tag{2.3}$$

Een rotatie wordt gekarakteriseerd door een 3×3 matrix R, die orthogonaal is en determinant 1 heeft (orthogonale matrices met determinant -1 komen overeen met spiegelingen). De orthogonaliteitsconditie voor R betekent dat

$$\widetilde{R}\,R = R\,\widetilde{R} = \mathbf{1}\ , \qquad \text{ofwel} \quad \sum_i R_{ij}\ R_{ik} = \delta_{jk}\ ; \quad \sum_i R_{ij}\ R_{kj} = \delta_{ik}\ , \tag{2.4}$$

waarbij $\widetilde{R}$ de gespiegelde is van R.

Het is nu niet moeilijk om na te gaan dat de vergelijking (2.2) inderdaad rotatie-invariant is. Allereerst merken we op dat

$$\frac{\partial}{\partial x_i}\psi(\vec{x}\,',t) = \sum_j \frac{\partial x'_j}{\partial x_i}\frac{\partial}{\partial x'_j}\psi(\vec{x}\,',t) = \sum_j R_{ji}\frac{\partial}{\partial x'_j}\psi(\vec{x}\,',t), \tag{2.5}$$

waarbij we gebruik hebben gemaakt van (2.3). Vervolgens stellen we vast dat

$$\begin{aligned} \sum_i \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_i}\,\psi(\vec{x}\,',t) &= \sum_{i,j,k} R_{ji}\ R_{ki}\,\frac{\partial}{\partial x'_j}\frac{\partial}{\partial x'_k}\,\psi(\vec{x}\,',t) \\ &= \sum_i \frac{\partial}{\partial x'_i}\frac{\partial}{\partial x'_i}\,\psi(\vec{x}\,',t), \end{aligned} \tag{2.6}$$

waarbij we gebruik gemaakt hebben van (2.4). Uit het bovenstaande volgt dat de vergelijking (2.2) invariant is onder rotaties: indien $\psi(\vec{x},t)$ een oplossing is van (2.2), dan moet ook $\psi(\vec{x}\,',t)$ een oplossing zijn van deze vergelijking.

Laten we nu een belangrijke eigenschap van rotaties bespreken. Zoals hierboven al is gebruikt kunnen kunnen de rotaties worden voorgesteld door reële 3×3 matrices R met determinant 1 welke moeten voldoen aan de orthogonaliteitsconditie $R\,\widetilde{R} = \mathbf{1}$. Elke draaiing in 3 dimensies kan worden gekarakteriseerd door drie hoeken (we zullen dit in hoofdstuk 3 nader preciseren). Zijn R_1 en R_2 beide matrices behorende bij zekere rotaties, dan is het product $R_3 = R_1R_2$ eveneens een rotatie. Anders geformuleerd: twee draaiingen achtereenvolgend uitgevoerd geven weer een draaiing. Het bewijs van deze bewering is eenvoudig: veronderstel dat R_1 en R_2 orthogonale matrices zijn met determinant 1. Door gebruik te maken van

$$\widetilde{R}_1 = R_1^{-1}, \qquad \widetilde{R}_2 = R_2^{-1}, \tag{2.7}$$

kunnen we laten zien dat ook $R_3 = R_1R_2$ orthogonaal is:

$$\widetilde{R}_3 = \widetilde{R_1R_2} = \widetilde{R}_2\widetilde{R}_1 = R_2^{-1}R_1^{-1} = (R_1R_2)^{-1} = R_3^{-1}. \tag{2.8}$$

Aangezien

$$\det R_3 = \det(R_1R_2) = \det R_1\ \det R_2 = 1, \tag{2.9}$$

is R_3 dus ook een draaiing. Merk op dat ook het product $R_4 = R_2 R_1$ een draaiing is, maar dat R_3 en R_4 niet noodzakelijk identiek zijn. Met andere woorden, draaiingen zijn niet *commutatief*; uitgevoerd in een verschillende volgorde zullen ze i.h.a. niet leiden tot hetzelfde resultaat.

We kunnen nu vaststellen dat de rotaties een zogenaamde *groep* vormen. Een verzameling van elementen (hier de verzameling van reële 3×3 matrices R met determinant 1 en $\tilde{R}R = \mathbf{1}$) wordt een groep genoemd indien er een vermenigvuldiging gedefinieerd is (hier de gewone matrix vermenigvuldiging) zodanig dat aan de volgende eisen voldaan is:

1. Indien R_1 en R_2 elementen van de groep zijn dan is ook het product $R_1 R_2$ een element van de groep.

2. De vermenigvuldiging is associatief: $R_1(R_2 R_3) = (R_1 R_2)R_3$. We mogen dus hetzij eerst R_2 en R_3 met elkaar vermenigvuldigen en het resultaat daarvan dan met R_1 vermenigvuldigen, ofwel deze twee operaties in omgekeerde volgorde uitvoeren. Let erop dat de volgorde waarin de matrices in de uitdrukking staan wèl dezelfde moet blijven.

3. Er bestaat een eenheids-element $\mathbf{1}$, zodat $\mathbf{1}R = R$ voor alle elementen R van de groep. Dit eenheids-element is ook een element van de groep.

4. Bij elke R is er binnen de groep een inverse R^{-1} zodanig dat $R^{-1}R = \mathbf{1}$.

Aan al deze eigenschappen voldoet de verzameling van draaiingsmatrices en we konstateren daarom dat deze verzameling een groep vormt met oneindig veel elementen.

Iedere groep wordt volledig gekarakteriseerd door zijn vermenigvuldigingsstruktuur, de samenhang tussen de elementen via de vermenigvuldigingsregels. We zullen later trachten dit begrip "structuur" in formules vast te leggen. Merk op dat er in een groep niet zoiets bestaat als "optellen" en "aftrekken", maar alleen "vermenigvuldigen". Er hoort dan ook géén "nul-element" bij.

Overigens is het heel gebruikelijk dat de verzameling van alle mogelijke invarianties van een systeem een groep vormen. Aan twee invariantie-transformaties kunnen we namelijk onmiddellijk een derde toekennen, door de twee transformaties achtereenvolgens te laten werken op de grootheden in termen waarvan de theorie is gedefinieerd. Onder de resulterende transformatie moet de theorie vanzelfsprekend invariant zijn, en die moet dus ook een invariantie definiëren. Hiermee is voldaan aan de eerste van de bovenstaande groepseigenschappen; de overige eigenschappen zijn meestal vanzelfsprekend.

Laten we nu eerst wat preciezer aangeven hoe de functie ψ transformeert onder rotaties. Onder een draaiing gaan de coördinaten van $\vec{x}$ over in nieuwe coördinaten x'_i, die gegeven worden door

$$x_i \to x'_i = \sum_j R^{-1}_{ij}\, x_j. \tag{2.10}$$

De keuze van R^{-1} in plaats van R is slechts een kwestie van conventie en is van geen belang aangezien we R nog niet gespecificeerd hebben. De draaiing (2.10) kan het gevolg zijn van een draaiing van de fysische situatie. Bijvoorbeeld, als $\vec{x}$ de plaats van een deeltje aangeeft, dan wordt de plaats van het deeltje na een draaiing gegeven door $\vec{x}\,'$. De draaiing kan ook het gevolg zijn van een andere keuze van het coördinatenstelsel. Als we een nieuw coördinatenstelsel invoeren dat t.o.v. het oude stelsel is gedraaid (maar nu in tegengestelde richting) dan wordt een vector die in het oude stelsel werd beschreven door coördinaten x_i, in het nieuwe stelsel beschreven door de coördinaten x'_i. Vanzelfsprekend zijn beide uitgangspunten complementair, tenminste, zolang er geen voorkeursrichting is. Hier zullen we ervan uitgaan dat we een draaiing van het coördinaatstelsel uitvoeren zodat de fysische situatie niet verandert.[3] Een dergelijke draaiing van het coördinaatstelsel induceert een verandering van coördinaten, en dus ook van functies van die coördinaten. Een functie ψ gaat na draaiing over in een nieuwe functie ψ', zodanig dat de waarde van ψ' in de nieuwe coördinaten x'_i overeenkomt met de waarde van de oude functie ψ in de overeenkomstige oude coördinaten x_i. De relatie tussen beide functies wordt dus gegeven door

$$\psi'(\vec{x}\,') = \psi(\vec{x}). \tag{2.11}$$

Gebruikmakend van (2.10) konkluderen we dus dat de nieuwe functie wordt gegeven door

$$\psi'(\vec{x}\,') = \psi(R\,\vec{x}\,'). \tag{2.12}$$

Het is van belang om het effekt te bepalen van twee achtereenvolgende draaiingen. Stel, we draaien het koordinaatstelsel overeenkomend met de matrix R, en daarna voeren we nogmaals een draaiing uit overeenkomend met een matrix S. De coördinaten na de tweede draaiing worden dan gegeven door

$$x'' = S^{-1}_{ij}\, x'_j = S^{-1}_{ij}\, R^{-1}_{jk}\, x_k = (R\,S)^{-1}_{ij}\, x_j. \tag{2.13}$$

De veranderingen in functies zijn overeenkomstig, en volgen uit bovenstaande formules. Het gecombineerde effekt van de twee draaiingen geeft aanleiding tot een functie ψ'' gegeven door

$$\psi''(\vec{x}\,'') = \psi'(S\,\vec{x}\,'') = \psi(R\,S\,\vec{x}\,''). \tag{2.14}$$

waarbij we de aandacht vestigen op de volgorde waarin de matrices R en S voorkomen in de laatste term, die *tegengesteld* is aan de volgorde waarin de draaiingen zijn uitgevoerd. Omdat aan beide zijden van (2.14) $\vec{x}\,''$ voorkomt, schrijft men in die formule vaak x_i in plaats van x''_i. Twee achtereenvolgende draaiingen, eerst R en dan S, veranderen dan de functie $\psi(\vec{x})$ eerst in $\psi'(\vec{x}) = \psi(R\,\vec{x})$ en vervolgens in $\psi''(\vec{x}) = \psi(R\,S\,\vec{x})$. Met andere woorden, bij iedere draaiing vervangen we x_i voorkomend in het argument van een functie, door $R_{ij}\,x_j$, waarbij R_{ij} de matrix is die hoort bij de betreffende draaiing. We kunnen dit als volgt samenvatten,

$$\psi(\vec{x}) \xrightarrow{R} \psi'(\vec{x}) = \psi(R\,\vec{x}) \xrightarrow{S} \psi''(\vec{x}) = \psi(R\,S\,\vec{x}). \tag{2.15}$$

Nu gebruiken we dat $\mathcal{D}$, de differentiaaloperator die voorkomt in de vergelijking (2.1), rotatie-invariant is. Dat betekent *niet* dat de oplossingen ψ rotatie-invariant moeten

[3] Dit wordt wel een *passieve* transformatie genoemd.

zijn. Bijvoorbeeld, een mogelijke oplossing kan een golffunctie zijn behorend bij een deeltje bewegend van oost naar west. Na een rotatie over $\pi/2$ beweegt het deeltje van noord naar zuid, en dat is een andere situatie met een andere bijbehorende golffunctie. Echter omdat $\mathcal{D}$ rotatie-invariant is, moet gelden dat elke oplossing ψ onder een rotatie overgaat in een andere oplossing ψ'. Dus, door rotatie over $\pi/2$ wordt de golffunctie die behoort bij een deeltje bewegend van oost naar west, overgevoerd in een andere golffunctie die behoort bij een van noord naar zuid bewegend deeltje, en beide golffuncties zijn oplossingen van dezelfde vergelijking. Nog iets anders geformuleerd: een mogelijke fysische situatie (beschreven door een golffunctie ψ_1 gehoorzamend aan een bepaalde vergelijking) gaat onder rotatie over in een andere mogelijke fysische situatie, beschreven door ψ_2. Meer konkreet, als er deeltjes kunnen bestaan die van oost naar west reizen dan kunnen er ook deeltjes zijn die van noord naar zuid gaan. Ware dit niet zo dan zou er geen rotatie-invariantie van de natuur zijn.

Laten ψ_1 en ψ_2 nu verschillende oplossingen zijn van (2.1), al dan niet samenhangend via een rotatie,

$$\mathcal{D}\psi_1 = 0, \qquad \mathcal{D}\psi_2 = 0. \tag{2.16}$$

Omdat de vergelijking (2.1) lineair is in ψ, geldt dat iedere lineaire combinatie $\lambda\,\psi_1 + \mu\,\psi_2$ ook een oplossing moet zijn,

$$\mathcal{D}\,(\lambda\,\psi_1 + \mu\,\psi_2) = \lambda\,\mathcal{D}\psi_1 + \mu\,\mathcal{D}\psi_2 = 0. \tag{2.17}$$

In het algemeen: zijn $\psi_1, \ldots, \psi_n$ oplossingen van de vergelijking in (2.1) dan is ook iedere lineaire combinatie

$$\lambda_1\,\psi_1 + \lambda_2\,\psi_2 + \cdots + \lambda_n\,\psi_n \tag{2.18}$$

een oplossing van (2.1)

Ten aanzien van het gedrag onder rotaties kunnen we nu twee situaties onderscheiden. Of de golffunctie ψ is rotatie-invariant, d.w.z. ψ gaat onder rotatie over in zichzelf,

$$\psi'(\vec{x}) = \psi(\vec{x}) \iff \psi(\vec{x}\,') = \psi(\vec{x}), \tag{2.19}$$

ofwel er zijn verzamelingen van (lineair onafhankelijke) oplossingen $\psi_1, \ldots, \psi_n$, die onder rotaties overgaan in lineaire combinaties van zichzelf. Voor wat betreft de tweede mogelijkheid heeft men bijvoorbeeld de verzameling van oplossingen van deeltjes bewegende in alle mogelijke richtingen. In dit geval bevat de verzameling $\psi_1, \ldots, \psi_n$ een oneindig aantal oplossingen. Met het oogmerk complicaties samenhangende met de oneindigheid van genoemde verzameling te vermijden kunnen we ons beperken tot deeltjes in rust, of tenminste met verwaarloosbaar kleine impuls. Een deeltje in rust gaat onder een rotatie in zichzelf over, maar de interne structuur kan eventueel veranderen. De verzameling van golffuncties welke met elkaar samenhangen door rotatie bevat dan i.h.a. een eindig aantal oplossingen. Als het deeltje zich in de grondtoestand bevindt dan is de bijbehorende golffunctie gewoonlijk rotatie-invariant; de verzameling bevat dan dus slechts één golffunctie. Bevindt het deeltje zich in een geëxciteerde toestand, dan kunnen verschillende geëxciteerde toestanden in elkaar overgaan via een rotatie.

Zij nu gegeven zo'n verzameling $\psi_1, \ldots, \psi_n$ van golffunckties die in elkaar over gaan onder rotaties. Dat betekent dat ψ_1 na rotatie over gaat in een lineaire combinatie van $\psi_1, \ldots, \psi_n$,

$$\psi_1'(\vec{x}) \equiv \psi_1(R\,\vec{x}) = d_{11}\,\psi_1(\vec{x}) + d_{12}\,\psi_2(\vec{x}) + \cdots + d_{1n}\,\psi_n(\vec{x}), \tag{2.20}$$

en een soortgelijk resultaat geldt voor $\psi_2, \ldots \psi_n$. Algemeen kunnen we dus schrijven

$$\psi_A' = \sum_B d_{AB}\,\psi_B. \qquad (A, B = 1, \ldots, n) \tag{2.21}$$

De coefficienten d_{AB} vormen een matrix $D(R)$, zodat

$$\Psi'(\vec{x}) = \Psi(R\,\vec{x}) = D(R)\,\Psi(\vec{x}), \tag{2.22}$$

waarbij we de golffuncties $\psi_1, \ldots, \psi_n$ aangeduid hebben als een kolomvector Ψ.

De matrices $D(R)$ in (2.21-22) hangen samen met de draaiingsmatrices in die zin dat er bij elke draaiing R in de 3-dimensionale ruimte een matrix $D(R)$ bestaat welke de oplossingen ψ_A overvoert in lineaire combinaties van dezelfde oplossingen. Echter we kunnen verder gaan. Een bepaalde rotatie kan ineens uitgevoerd worden, of ook verkregen worden als het resultaat van enkele na elkaar uitgevoerde rotaties. Hoe het ook zij, het eindresultaat moet hetzelfde zijn. Bijvoorbeeld, een rotatie over $\pi/2$ kan ook gezien worden als een rotatie over π gevolgd door een rotatie over $-\pi/2$. Het feit dat het eindresultaat hetzelfde moet zijn betekent dat de matrices $D(R)$ bepaalde producteigenschappen moeten bezitten. Om dit af te leiden, beschouwen we weer twee achtereenvolgende rotaties, R en S (zie (2.15)). Bij R behoort een matrix $D(R)$ en bij S hoort een matrix $D(S)$. In formules uitgedrukt,

$$\begin{aligned} \Psi(R\,\vec{x}) &= D(R)\,\Psi(\vec{x}), \\ \Psi(S\,\vec{x}) &= D(S)\,\Psi(\vec{x}). \end{aligned} \tag{2.23}$$

Vanzelfsprekend hoort bij de gecombineeerde draaiing $R\,S$ een matrix $D(R\,S)$, zodat geldt

$$\Psi(R\,S\,\vec{x}) = D(R\,S)\,\Psi(\vec{x}). \tag{2.24}$$

Maar we kunnen $\Psi(R\,S)$ ook bepalen met behulp van (2.23),

$$\Psi(R\,S\,\vec{x}) = D(R)\,\Psi(S\,\vec{x}) = D(R)\,D(S)\,\Psi(\vec{x}). \tag{2.25}$$

Daarom moet gelden dat

$$D(R)\,D(S) = D(R\,S). \tag{2.26}$$

De matrices $D(R)$ moeten in hun vermenigvuldigingsregels dus dezelfde structuur hebben als de matrices R. Als dit het geval is zeggen we dat de matrices $D(R)$ een representatie vormen van de groep van draaiingen in drie dimensies.

Meer algemeen: een verzameling van matrices vormt een *representatie* van een groep, indien geldt:

1. bij elk element a van de groep behoort een matrix A
2. bij het product van twee elementen behoort het product van de corresponderende matrices, d.w.z. als bij a, b en c de matrices A, B en C behoren dan moet, als $c = ab$, tevens gelden $C = AB$.

We hebben nu het volgende resultaat gevonden: Onder draaiingen in de driedimensionale ruimte moeten de golffuncties van elk fysisch systeem transformeren volgens lineaire afbeeldingen welke een representatie vormen van de groep van draaiingen in drie dimensies.

Als eenvoudig voorbeeld kunnen we de drie functies

$$\psi_1(\vec{x}) = x_1, \qquad \psi_2(\vec{x}) = x_2, \qquad \psi_3(\vec{x}) = x_3, \tag{2.27}$$

nemen. Onder een draaiing transformeren deze drie functies met een matrix $D(R)$ die gelijk is aan R. Aan de conditie (2.26) is dus triviaal voldaan.

Echter, de bovenstaande konklusie hoeft niet altijd geheel juist te zijn. Volgens de quantummechanica beschrijven twee golffuncties die slechts verschillen in een factor met absolute waarde 1, dezelfde fysische situatie. De golffuncties ψ en $\mathrm{e}^{i\lambda}\psi$ beschrijven dus dezelfde fysische situatie voor reële λ. Daarmee is dus de mogelijkheid van een zekere meerduidigheid in de definitie van de matrices $D(R)$ aanwezig. De voorwaarde (2.26) kan daarom in principe vervangen worden door een zwakkere voorwaarde

$$D(R_1)\, D(R_2) = \exp\left[i\alpha(R_1, R_2)\right] D\left(R_1\, R_2\right), \tag{2.28}$$

waarbij α een (reële) fasehoek is die afhangt van R_1 en R_2. Matrices $D(R)$ die voldoen aan (2.28) met een niet-triviale fasefactor vormen een z.g. *projectieve representatie*. Projectieve representaties zijn inderdaad van belang in de fysika. Het blijkt dat er situaties bestaan waarbij er bij elke matrix R van de draaiingsgroep *twee* matrices $D(R)$ en $D'(R)$ bestaan, die een fasefactor, om precies te zijn een factor -1, verschillen. Er geldt dus $D'(R) = -D(R)$. Dit is toegestaan omdat de golffuncties ψ en $-\psi$ dezelfde fysische situatie beschrijven. Deze meerduidigheid impliceert dat de relatie (2.26) slechts op een teken na bepaald is, zodat de fasehoek in (2.28) gelijk kan zijn aan 0 of π. Deeltjes beschreven door golffuncties welke transformeren volgens een dergelijke projectieve representatie, hebben geen analogon in de klassieke mechanica. Voorbeelden van zulke deeltjes zijn het electron, het proton en het neutron. Hun bijbehorende golffunkties zullen dus op een meer ingewikkelde manier transformeren dan zoals beschreven door (2.12). We zullen hier later op terugkomen.

Er is nog een belangrijke eis waaraan de matrices $D(R)$ moeten voldoen, en die hangt samen met de fysische interpretatie van de quantummechanika. Hierin wordt een belangrijke rol toegekend aan het *inproduct*, dat aan twee golffuncties ψ_1 en ψ_2 een complex getal toevoegt, aangeduid door $\langle\psi_1 \mid \psi_2\rangle$, dat voldoet aan de volgende relaties

$$\begin{aligned} \langle\psi \mid \psi\rangle &\geq 0, \\ \langle\psi \mid \psi\rangle &= 0, \quad \text{dan en slechts dan als } \mid\psi\rangle = 0, \\ \langle\psi_1 \mid \lambda\,\psi_2 + \mu\,\psi_3\rangle &= \lambda\,\langle\psi_1 \mid \psi_2\rangle + \mu\,\langle\psi_1 \mid \psi_3\rangle, \quad \text{voor ieder complex getal } \lambda \text{ en } \mu, \\ \langle\psi_1 \mid \psi_2\rangle^* &= \langle\psi_2 \mid \psi_1\rangle. \end{aligned} \tag{2.29}$$

Voor golffuncties die afhangen van één coördinaat, is een dergelijk inproduct gedefinieerd door

$$\langle \psi_1 \mid \psi_2 \rangle = \int_{-\infty}^{\infty} \mathrm{d}x \; \psi_1^*(x) \, \psi_2(x), \tag{2.30}$$

maar voor ons doel is de precieze definitie van het inproduct niet van belang.

Volgens de quantummechanika moet de absolute waarde van het inproduct geïnterpreteerd worden als een waarschijnlijkheid. Om dit wat nauwkeuriger te zeggen, neem aan dat de toestand van het systeem beschreven wordt door $\mid \psi\rangle$. De waarschijnlijkheid dat men bij een meting het systeem aantreft in een toestand $\mid \varphi\rangle$ wordt dan gegeven door $|\langle \varphi \mid \psi\rangle|^2$. Laten we nu het systeem, inklusief het meetapparaat, een rotatie ondergaan. Volgens (2.22) veranderen de toestanden dan volgens

$$\mid \psi\rangle \to D \mid \psi\rangle, \qquad \mid \varphi\rangle \to D \mid \varphi\rangle. \tag{2.31}$$

De bijbehorende verandering van het inproduct is dan

$$\langle \varphi \mid \psi\rangle \longrightarrow \langle \varphi \mid D^\dagger D \mid \psi\rangle \tag{2.32}$$

Echter, als de natuur invariant is onder rotaties, dan mag de waarschijnlijkheid geässocieerd met het inproduct, *niet* veranderen na de draaiing. De twee inproducten in (2.30) moeten dus gelijk zijn. Aangezien deze gelijkheid moet gelden voor alle mogelijke paren van toestanden $\mid \psi\rangle$ en $\mid \varphi\rangle$, kunnen we konkluderen dat de matrices D moeten voldoen aan de volgende conditie

$$D^\dagger D = \mathbf{1}, \tag{2.33}$$

m.a.w. D moet een *unitaire* matrix zijn.[4] Omdat dit moet gelden voor iedere matrix $D(R)$ overeenkomend met een draaiing, geldt deze eis voor de gehele representatie. In deze context zullen we dus uitsluitend geïnteresseerd zijn in *unitaire representaties.*

[4]De conditie is dat de absolute waarde van het inproduct niet verandert, dus het is voldoende als $D^\dagger D$ gelijk is aan de eenheid, tot op een fasefactor na. Echter, $D^\dagger D$ is een hermitische, positief-definiete matrix, zodat we kunnen konkluderen dat de fasefaktor gelijk moet zijn aan 1.

3. De groep van draaiingen in drie dimensies

Een draaiing in een driedimensionale ruimte kan worden voorgesteld door een 3×3 matrix van reële getallen. Doordat bij een draaiing van een stel vectoren de onderlinge hoeken dezelfde blijven is de betreffende matrix *orthogonaal.* Deze orthogonale matrices vormen de groep $O(3)$. Uit de eis $RR^\dagger = \mathbf{1}$ leiden we af dat $\det(R) = \pm 1$. Beperken we ons tot de orthogonale matrices met $\det(R) = +1$ dan spreken we van de groep $SO(3)$.

Zo'n draaiing in de driedimensionale ruimte wordt volledig bepaald door de draaias en de hoek waarover wordt gedraaid. De draaias kan bijvoorbeeld worden gespecificeerd m.b.v. een driedimensionale vector $\vec{\alpha}$; de grootte van de draaiing kan dan worden aangegeven door de lengte van deze vector gelijk te kiezen aan de hoek waarover wordt gedraaid (in radialen). Omdat draaiingen over hoeken die 2π in grootte verschillen, identiek zijn, kunnen we de draaias-vectoren $\vec{\alpha}$ beperken tot een driedimensionale bol met straal π. Dit geeft een natuurlijke parametrisatie van alle draaiingen. Ieder punt in deze parameterbol komt dus overeen met een mogelijke rotatie: de draaias wordt gegeven door de lijn door dit punt en het middelpunt van de bol, en de hoek waarover wordt gedraaid (volgens een linkse schroef bijvoorbeeld) varieert van 0 tot π (de draaiingen over hoeken van $-\pi$ tot 0 hangen dan samen met de tegengesteld gerichte vector). Twee tegengestelde vectoren op de rand van de bol, dus, $\vec{\alpha}$ en $-\vec{\alpha}$ met $|\vec{\alpha}| = \pi$, beschrijven dezelfde rotatie, de één over een hoek π en de ander over een hoek $-\pi$ om dezelfde draaias. Echter, afgezien van deze identifikatie van diametraal gelegen punten op het boloppervlak, beschrijven twee verschillende punten van de parameterbol altijd twee verschillende rotaties.

Uit het voorafgaande is duidelijk dat rotaties geparametriseerd kunnen worden in termen van drie onafhankelijke parameters, namelijk de componenten van de vectoren $\vec{\alpha}$, en verder dat de rotaties op continue wijze afhangen van deze parameters. Om die afhankelijkheid verder te bestuderen, beschouwen we infinitesimale transformaties, m.a.w. rotaties overeenkomende met $|\vec{\alpha}| \approx 0$. Laten we ons eerst beperken tot rotaties om de z-as, zodat $\vec{\alpha} = (0, 0, \alpha)$. De bijbehorende rotatie volgt uit

$$\begin{aligned} x &\to \cos\alpha\, x + \sin\alpha\, y, \\ y &\to \cos\alpha\, y - \sin\alpha\, x, \\ z &\to z. \end{aligned} \tag{3.1}$$

Dit leidt tot een bijbehorende matrix $R(\alpha)$, gelijk aan

$$R(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.2}$$

De draaiing over een hoek α kan ook gezien worden als zijnde het resultaat van n successieve draaiingen over een hoek α/n. Voor zeer grote waarde van n zal de draaiing over de kleine hoek α/n slechts infinitesimaal van de eenheid verschillen; verwaarlozen we termen van orde $(\alpha/n)^2$, dan vinden we voor de bijbehorende 3×3 matrix,

$$R(\alpha/n) \;=\; \begin{pmatrix} 1 & \alpha/n & 0 \\ -\alpha/n & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathrm{O}\left(\frac{\alpha^2}{n^2}\right)$$

$$= \mathbf{1} + \frac{\alpha}{n}\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathrm{O}\left(\frac{\alpha^2}{n^2}\right) \tag{3.3}$$

Het is mogelijk om de oorspronkelijke eindige rotatie over een hoek α weer te reconstrueren, door de n-de macht te nemen van (3.3),

$$R(\alpha) = [R(\alpha/n)]^n = \left[\mathbf{1} + \frac{\alpha}{n}T + \mathrm{O}\left(\frac{\alpha^2}{n^2}\right)\right]^n, \tag{3.4}$$

waarbij de matrix T gegeven wordt door

$$T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.5}$$

In de limiet $n \to \infty$ verwachten we dat we de termen van orde $1/n^2$ kunnen verwaarlozen; voorts maken we gebruik van de formule

$$\mathrm{e}^A = \lim_{n\to\infty}\left(\mathbf{1} + \frac{1}{n}A\right)^n. \tag{3.6}$$

Dit leidt tot

$$R(\alpha) = \exp(\alpha T). \tag{3.7}$$

De exponent van deze matrix kunnen we uitwerken door gebruik te maken van de reeksontwikkeling

$$\mathrm{e}^A = \sum_{n=0}^{\infty}\frac{1}{n!}A^n. \tag{3.8}$$

Vervolgens merken we op

$$T^{2n} = (-)^n \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad (n \geq 1) \tag{3.9}$$

waaruit onmiddelijk volgt dat $T^{2n+1} = (-)^n T$ voor $n \geq 0$. Met behulp hiervan kunnen we de exponentiatie uitvoeren door de even en oneven machten afzonderlijk te behandelen. Dit geeft aanleiding tot

$$\begin{aligned} \exp(\alpha T) &= \mathbf{1} + \sum_{n=1}^{\infty}\frac{(-)^n\alpha^{2n}}{(2n)!}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \sum_{n=0}^{\infty}\frac{(-)^n\alpha^{2n+1}}{(2n+1)!}T \\ &= \mathbf{1} + (\cos\alpha - 1)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \sin\alpha\, T, \end{aligned} \tag{3.10}$$

hetgeen inderdaad precies in overeenstemming is met de oorspronkelijke matrix (3.2).

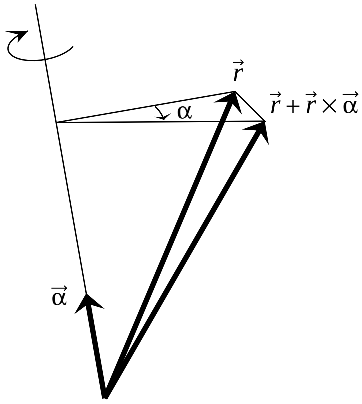

Figure 2: Infinitesimale draaiing met draaias $\vec{\alpha}$ van een vector $\vec{r}$.

We kunnen de relatie tussen eindige en infinitesimale transformaties zoals gegeven door (3.7), ook voor meer algemene draaiingen beschouwen. Voor draaiingen over een kleine hoek wordt bij iedere $\vec{r}$ een kleine vector opgeteld welke loodrecht staat op de vector $\vec{r}$ en de draaias, en waarvan de lengte gelijk is aan het product van de draaihoek en de afstand van $\vec{r}$ tot de draaias (zie figuur). Deze kleine vector is precies het uitproduct van $\vec{r}$ en de draaias-vector $\vec{\alpha}$ (waarbij dus aangenomen is dat $|\vec{\alpha}| \approx 0$), zodat

$$\vec{r} \to \vec{r} + \vec{r} \times \vec{\alpha} + \mathrm{O}\left(|\vec{\alpha}|^2\right), \tag{3.11}$$

Voor een willekeurige draaias-vector $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ kunnen we daarom schrijven

$$\begin{aligned} x &\to x + \alpha_3\, y - \alpha_2\, z + \mathrm{O}\left(|\vec{\alpha}|^2\right), \\ y &\to y + \alpha_1\, z - \alpha_3\, x + \mathrm{O}\left(|\vec{\alpha}|^2\right), \\ z &\to z + \alpha_2\, x - \alpha_1\, y + \mathrm{O}\left(|\vec{\alpha}|^2\right), \end{aligned} \tag{3.12}$$

We kunnen infinitesimale draaiingen dus als volgt schrijven,

$$R(\vec{\alpha}) = \mathbf{1} + i\Big(\alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3\Big) + \mathrm{O}\left(|\vec{\alpha}|^2\right), \tag{3.13}$$

waarbij we een factor i hebben ingevoerd om aan te sluiten bij de gebruikelijke notaties, en de hermitische matrices L_1, L_2 en L_3 gedefinieerd zijn door

$$\begin{aligned} L_1 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \\ L_2 &= \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \\ L_3 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{3.14}$$

Bovenstaand resultaat kan worden samengevat in één formule door gebruik te maken van de volledig-antisymmetrische epsilon-tensor,

$$(L_i)_{jk} = -i\epsilon_{ijk}. \tag{3.15}$$

Inderdaad gaan we gemakkelijk na dat

$$\begin{aligned}(L_1)_{23} = -\,(L_1)_{32} = -i\epsilon_{123} = -i,\\ (L_2)_{31} = -\,(L_2)_{13} = -i\epsilon_{231} = -i,\\ (L_3)_{12} = -\,(L_3)_{21} = -i\epsilon_{312} = -i.\end{aligned} \tag{3.16}$$

We kunnen $R(\vec{\alpha})$ nu weer opgebouwd denken uit n successieve draaiingen met draaias $\vec{\alpha}/n$,

$$\begin{aligned} R(\vec{\alpha}) &= [R(\vec{\alpha}/n)]^n \\ &= \left[\mathbf{1} + \frac{1}{n}\Big(i\alpha_1 L_1 + i\alpha_2 L_2 + i\alpha_3 L_3\Big) + \mathrm{O}\left(\frac{|\vec{\alpha}|^2}{n^2}\right)\right]^n. \end{aligned} \tag{3.17}$$

Door gebruik te maken van (3.4) vinden we dan de volgende uitdrukking in de limiet $n \to \infty$,

$$R(\vec{\alpha}) = \exp\left(i\sum_k \alpha_k L_k\right). \tag{3.18}$$

We kunnen de juistheid van (3.18) ook nog op een andere manier bewijzen. Allereerst merken we op dat voor draaiingen met dezelfde draaias, maar niet dezelfde draaihoek, de volgende productregel geldt

$$R(s\vec{\alpha})\,R(t\vec{\alpha}) = R((s+t)\vec{\alpha}), \tag{3.19}$$

waarbij s en t reële getallen zijn. De draaiingen $R(s\vec{\alpha})$ met dezelfde draaias definiëren een commuterende ondergroep van de volledige groep van draaiingen. Dit is eenvoudig in te zien: de matrices $R(s\vec{\alpha})$ definiëren een groep (voor een vaste vector $\vec{\alpha}$ en variabele s) waarvoor geldt dat het resultaat van een vermenigvuldiging niet afhangt van de volgorde in het product,

$$R(s\vec{\alpha})\,R(t\vec{\alpha}) = R(t\vec{\alpha})\,R(s\vec{\alpha}). \tag{3.20}$$

Deze ondergroep is de groep $SO(2)$, de groep van tweedimensionale draaiingen (de draai-as blijft immers onveranderd onder deze rotaties, alleen de componenten loodrecht op de draaias worden geroteerd). Met behulp van (3.19) kunnen we nu eenvoudig de volgende differentiaalvergelijking afleiden voor $R(s\vec{\alpha})$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}R(s\vec{\alpha}) &= \lim_{\Delta\to 0}\frac{R((s+\Delta)\vec{\alpha}) - R(s\vec{\alpha})}{\Delta} \\ &= \lim_{\Delta\to 0}\frac{R(\Delta\vec{\alpha}) - \mathbf{1}}{\Delta}R(s\vec{\alpha}) \\ &= \left(i\sum_k \alpha_k L_k\right) R(s\vec{\alpha}), \end{aligned} \tag{3.21}$$

waarbij we eerst gebruik gemaakt hebben van (3.19) en vervolgens van (3.13). Het is nu eenvoudig in te zien dat de oplossing van deze differentiaalvergelijking precies gegeven wordt door (3.18).

Ten overvloede zullen we nogmaals nagaan dat de matrices (3.18) draaiingen representeren, door te bewijzen dat deze matrices orthogonaal zijn en determinant 1 hebben, ofwel dat aan de volgende relaties is voldaan,

$$\widetilde{R(\vec{\alpha})} = [R(\vec{\alpha})]^{-1} = R(-\vec{\alpha}), \qquad \det R(\vec{\alpha}) = 1, \tag{3.22}$$

Het bewijs volgt uit de volgende eigenschappen voor een algemene matrix A,

$$\widetilde{(\mathrm{e}^{A})} = \mathrm{e}^{\widetilde{A}}, \qquad \det\left(\mathrm{e}^{A}\right) = \mathrm{e}^{\mathrm{Sp}\,A} \tag{3.23}$$

Hieruit volgt dat de matrices (3.18) voldoen aan (3.22) mits de matrix $i\sum_k \alpha_k L_k$ reëel, antisymmetrisch en spoorloos is. Hieraan blijkt inderdaad te zijn voldaan; uit de definities (3.14) volgt dat $i\sum_k \alpha_k L_k$ in feite de meest algemene reële, antisymmetrische, spoorloze 3×3 matrix representeert.

We kunnen de bovenstaande vraag echter ook omkeren: kunnen *alle* draaiingen ook in de vorm (3.18) geschreven worden? Het antwoord op deze vraag is minder eenvoudig te geven. In principe kan de exponentiatie in (3.18) weer expliciet worden uitgevoerd via de machtreeksrepresentatie (3.8), en het resultaat kan worden vergeleken met de meest algemene draaiingsmatrix. Het zal dan blijken dat de vraag een bevestigend antwoord heeft: alle draaiingen kunnen inderdaad geschreven worden in de vorm (3.18). Dit echter niet voor alle groepen het geval. Voor z.g. *niet-compacte* groepen bestaan er elementen die niet in een dergelijke exponentiële vorm geschreven kunnen worden, alhoewel het mogelijk is ze te schrijven als een product van een eindig aantal van dergelijke e-machten. Deze groepen worden niet-compact genoemd, omdat het parametervolume niet compact is. De draaiingsgroep, waar alle mogelijke groepselementen gedefinieerd zijn in termen van de parameters α_k die beperkt kunnen worden tot een bol met straal π, is een compacte groep. Binnen het kader van onze diskussie zullen niet-compacte groepen geen rol spelen, maar dergelijke groepen zijn niet onbelangrijk in de fysika. Zo is de Lorentz-groep, de groep bestaande uit alle Lorentz-transformaties, een voorbeeld van een niet-compacte groep.

Uit het voorafgaande zal duidelijk zijn dat de matrices L_k, geassocieerd met de infinitesimale transformaties, een belangrijke rol spelen en, voor compacte groepen althans, de groepselementen volledig bepalen d.m.v. de exponentiatie (3.18). Vandaar dat deze matrices de *generatoren* van de groep worden genoemd. Hoewel onze diskussie zich heeft afgespeeld binnen de context van de draaiingsgroep, is het bovenstaande van toepassing op iedere *Lie-groep*[5]: een groep waarvan de elementen analytisch afhangen van een eindig aantal parameters, in dit geval α_1, α_2, α_3. Voor het geval de groepselementen de vorm aannemen van matrices, moet dus gelden dat ieder van de matrixelementen een differentieerbare functie is van de parameters.[6] Het aantal lineair-onafhankelijke parameters

[5] naar de Noorse wiskundige Sophus Lie, 1842-1899

[6] Dit is duidelijk het geval voor de draaiingsgroep. In het algemene geval kan de bovenstaande eis iets worden afgezwakt; voor een algemene Lie-groep is het voldoende als de elementen tweevoudig differentieerbaar zijn naar de parameters.

definieert de *dimensie* van de Lie-groep, niet te verwarren met de dimensie van de matrices die we beschouwen.[7] Het aantal lineair-onafhankelijke generatoren moet vanzelfsprekend ook gelijk zijn aan de dimensie van de groep.

Een van de meest essentiële ingrediënten van een groep, is zijn vermenigvuldigingsstruktuur, volgens welke het product van twee rotaties, $R(\vec{\alpha})$ en $R(\vec{\beta})$, weer een draaiing moet zijn,

$$R(\vec{\alpha})\,R(\vec{\beta}) = R(\vec{\gamma}), \tag{3.24}$$

waarbij $\vec{\gamma}$ dus afhangt van $\vec{\alpha}$ en $\vec{\beta}$. De precieze afhankelijkheid legt de vermenigvuldigingsstruktuur van de groep vast. Het feit dat er een dergelijke $\vec{\gamma}(\vec{\alpha},\vec{\beta})$ moet bestaan, heeft ook implikaties voor de producten van de generatoren. Om deze af te leiden, ontwikkelen we (3.24) naar α en β,[8]

$$\begin{aligned}
\mathrm{e}^{i\vec{\alpha}\cdot\vec{L}}\,\mathrm{e}^{i\vec{\beta}\cdot\vec{L}} &= \Big(\mathbf{1} + i\alpha_k L_k + \mathrm{O}(\alpha^2)\Big)\Big(\mathbf{1} + i\beta_l L_l + \mathrm{O}(\beta^2)\Big) \\
&= \mathbf{1} + i(\alpha+\beta)_k\, L_k - \alpha_k\beta_l\, L_k L_l + \mathrm{O}(\alpha^2) + \mathrm{O}(\beta^2) \\
&= \mathbf{1} + i(\alpha+\beta)_k\, L_k - \tfrac{1}{2}(\alpha+\beta)_k(\alpha+\beta)_l\, L_k L_l \\
&\quad -\tfrac{1}{2}\alpha_k\beta_l\,[L_k, L_l] + \mathrm{O}(\alpha^2) + \mathrm{O}(\beta^2).
\end{aligned} \tag{3.25}$$

De eerste drie termen herkennnen we als het begin van de machtreeks van $\exp(i(\vec{\alpha}+\vec{\beta})\cdot\vec{L})$. Als de vierde term gelijk zou zijn aan nul, m.a.w. als de matrices L_k en L_l commuteren, dan geldt inderdaad dat $\gamma_k = \alpha_k + \beta_k$. Echter, het zal blijken dat de generatoren van de draaiingsgroep *niet* commuteren. Aangezien we de rechterkant van de vergelijking in ieder geval weer moeten kunnen schrijven in de vorm van de machtreeks voor $\exp(i\vec{\gamma}\cdot\vec{L})$, zullen de commutatoren van de generatoren geschreven moeten kunnen worden als een lineaire kombinatie van de generatoren. Met andere woorden, er moet gelden dat

$$[L_i, L_j] = c_{ij}^k\, L_k, \tag{3.26}$$

waarbij de constanten c_{ij}^k de *struktuurconstanten* worden genoemd, omdat ze in belangrijke mate de vermenigvuldigingsstruktuur van de groep bepalen. Merk op dat, voor hermitische generatoren L_k, de struktuurconstanten puur imaginair zijn.

Voordat we verder gaan zullen we eerst verifiëren of de generatoren (3.14) voldoen aan de eis (3.26). Na expliciete matrixvermenigvuldiging vinden we dat dit inderdaad het geval is,

$$[L_1, L_2] = iL_3, \qquad [L_2, L_3] = iL_1, \qquad [L_3, L_1] = iL_2, \tag{3.27}$$

ofwel

$$[L_i, L_j] = i\epsilon_{ijk}\, L_k. \tag{3.28}$$

[7] Voor de draaiingsgroep in drie dimensies is de dimensie van de groep en van de matrices gelijk aan 3. Dit is een toevallige omstandigheid: de dimensie van de draaiingsgroep in d dimensies is gelijk aan $\frac{1}{2}d(d-1)$.

[8] Met de notatie $\vec{\alpha}\cdot\vec{L}$ bedoelen we $\alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3$. In (3.25) maken we ook gebruik van de sommatieconventie: over herhaalde indices wordt altijd gesommeerd, ook als er geen expliciet som-teken aanwezig is. Dus $\alpha_k L_k \equiv \sum_k \alpha_k L_k$. Vanaf nu zullen we regelmatig van deze conventie gebruik maken.

Gebruikmakend van (3.26) kunnen we nu het volgende resultaat afleiden voor $\vec{\gamma}(\vec{\alpha}, \vec{\beta})$,

$$\gamma_k = \alpha_k + \beta_k + \tfrac{i}{2} c^k_{mn}\, \alpha_m \beta_n + \mathrm{O}(\alpha^2) + \mathrm{O}(\beta^2). \tag{3.29}$$

In principe kunnen we de hogere-orde bijdragen bepalen door middel van iteratie; zo vinden we bijvoorbeeld

$$\gamma_k = \alpha_k + \beta_k + \tfrac{i}{2} c^k_{mn}\, \alpha_m \beta_n - \tfrac{1}{12} \left(\alpha_m \alpha_n \beta_p + \beta_m \beta_n \alpha_p \right) c^k_{mq} c^q_{np} + \cdots. \tag{3.30}$$

Het feit dat alle termen in deze iteratie uitgedrukt kunnen worden in de struktuurconstanten volgt uit de Campbell-Baker-Hausdorff formule, die de logaritme van $(\exp A \exp B)$ uitdrukt in een machtreeks bestaande uit herhaalde commutatoren van de matrices A en B (zie appendix D). De vermenigvuldigingsstruktuur van de groep wordt dus bepaald door de struktuurconstanten (althans voor die groepselementen die zich bevinden in een eindige omgeving van de identiteit).

Stel nu eens dat we matrices A_k kunnen vinden, ongelijk aan de matrices L_k, die voldoen aan dezelfde commutatierelatie (3.26) als de L_k. In dat geval kunnen we door middel van exponentiatie bijbehorende groepselementen definiëren, die dezelfde vermenigvuldigingsregels hebben als de elementen van de oorspronkelijke groep. Met andere woorden, we vinden op deze wijze een representatie van de groep. Anderzijds kunnen we voor iedere representatie van de groep de bijbehorende generatoren konstrueren, middels de infinitesimale transformaties, die dan moeten voldoen aan de commutatierelaties (3.26) met dezelfde struktuurconstanten. Er bestaat dus een direkte relatie tussen de groepsrepresentaties en de representaties van (3.26). (In meer wiskundige termen: de generatoren L_k samen met de commutatierelaties (3.26) definiëren een algebra, die de *Lie-algebra* wordt genoemd. Matrices A_k met dezelfde commutatierelaties definiëren dan een *representatie* van de Lie-algebra.)

Men kan eenvoudig nagaan dat de struktuurconstanten ook aan bepaalde relaties moeten voldoen. Dit volgt uit de z.g. *Jacobi-identiteit*, die geldt voor iedere drie matrices A, B en C,

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0. \tag{3.31}$$

Deze identiteit kan worden bewezen door de commutatoren volledig uit te schrijven; we krijgen dan 12 termen die paarsgewijs tegen elkaar wegvallen. Gebruikmakend van de Jacobi-identiteit met $A = L_i$, $B = L_j$ en $C = L_k$, leiden we de volgende vergelijking voor de struktuurconstanten af,

$$c^m_{ij}\, c^n_{mk} + c^m_{jk}\, c^n_{mi} + c^m_{ki}\, c^n_{mj} = 0, \tag{3.32}$$

waarbij we gebruik gemaakt hebben van (3.26). De vergelijking (3.32) wordt ook aangeduid als de Jacobi-identiteit. Voor de draaiingsgroep impliceert dit de volgende vergelijking voor de ϵ-tensoren,

$$\epsilon_{ijm}\, \epsilon_{mkn} + \epsilon_{jkm}\, \epsilon_{min} + \epsilon_{kim}\, \epsilon_{mjn} = 0, \tag{3.33}$$

waar we later nog herhaaldelijk gebruik van zullen maken. Het bewijs van (3.33) volgt uit de identiteit

$$\epsilon_{ijm}\, \epsilon_{mkl} = \delta_{ik}\, \delta_{jl} - \delta_{il}\, \delta_{jk}, \tag{3.34}$$

die eenvoudig te bewijzen is (bijvoorbeeld door bepaalde waarden te kiezen voor twee van de indices).

Vergelijking (3.32) heeft nog een andere konsekwentie. Laten we n $n \times n$ matrices C_i definiëren volgens

$$(C_i)_j{}^k \equiv -c_{ij}^k, \tag{3.35}$$

waarbij n de dimensie is van de Lie-groep. We kunnen (3.32) dan schrijven als

$$(c_{ij}^m\, C_m)_k{}^n + (C_j C_i)_k{}^n - (C_i C_j)_k{}^n = 0\,, \qquad \text{ofwel} \qquad C_i\, C_j - C_j\, C_i = c_{ij}^k\, C_k\,. \tag{3.36}$$

Dit zijn precies dezelfde commutatierelaties als waarmee we de struktuurconstanten aanvankelijk gedefinieerd hebben in (3.26). (De matrices C_i definiëren dus een representatie van de Lie-algebra gebaseerd op (3.26).) Via exponentiatie van de matrices C_i kunnen we dan een groep definiëren die dezelfde vermenigvuldigingseigenschappen heeft (tenminste in een eindige omgeving van de identiteit) als de oorspronkelijke Lie-groep. Of, met andere woorden, we kunnen op deze wijze altijd een representatie vinden van de Lie-groep bestaande uit $n \times n$ matrices, waarbij n de dimensie van de Lie-groep is. Deze representatie wordt de *geadjungeerde representatie* genoemd.

Wanneer we de bovenstaande beweringen toepassen op het geval van de draaiingsgroep, wacht ons een teleurstelling. Aangezien in dit geval geldt dat $c_{ij}^k = i\epsilon_{ijk}$, zijn de matrices C_i precies gelijk aan de matrices L_i (zie 3.15), en we hebben dus de oorspronkelijke driedimensionale draaiingen teruggevonden. De geadjungeerde representatie is dus precies gelijk aan de oorspronkelijke groep. Dit is echter meer uitzondering dan regel zoals we later zullen zien.

4. Meer over representaties

In het vorige hoofdstuk hebben we de eigenschappen besproken van de groep van drie-dimensionale draaiingen. Nu keren we weer terug naar de mogelijke representaties van deze groep. Allereerst merken we op dat we uitgaande van een gegeven representatie, bijvoorbeeld gegeven door de matrices D werkende op golffuncties die we samengevoegd hebben in een kolomvector ψ, door herdefinitie van ψ een andere representatie kunnen verkrijgen. Bijvoorbeeld, herdefinieer ψ in golffuncties $\hat{\psi}$ volgens

$$\hat{\psi} = U\psi. \tag{4.1}$$

Onder draaiingen transformeert $\hat{\psi}$ dan volgens

$$\hat{\psi} \to \hat{\psi}' = \hat{D}\hat{\psi}, \tag{4.2}$$

waarbij $\hat{D}$ gegeven wordt door

$$\hat{D} = U D U^{-1}. \tag{4.3}$$

Zowel de oorspronkelijke matrices D als de matrices $\hat{D}$ definiëren een representatie van de draaiingsgroep, maar dergelijke representaties willen we niet beschouwen als essentieel van elkaar verschillend. Vandaar dat representaties die met elkaar samenhangen volgens (4.3), *equivalente representaties* worden genoemd. We kunnen dan tevens een belangrijk resultaat uit de representatietheorie formuleren:

Alle (eindig-dimensionale) representaties van eindige of compacte groepen zijn unitair. Hiermee bedoelen we dat iedere representatie via een herdefinitie (4.3) unitair gekozen kan worden, zodat voor alle matrices D behorende tot de representatie geldt dat $D^\dagger = D^{-1}$.

Tot op dit moment hebben we voornamelijk gesproken over een bepaalde representatie van de draaiingsgroep, namelijk de representatie gedefinieerd door draaiingen van de drie-dimensionale vector $\vec{x} = (x_1, x_2, x_3)$. Er is een voor de hand liggende manier om grotere representaties te construeren. Namelijk beschouw twee vectoren, $\vec{x}$ en $\vec{y}$, die beide transformeren onder rotaties. Deze vectoren kunnen bijvoorbeeld de posities beschrijven van twee verschillende deeltjes. Samen vormen zij een zes-dimensionale vector $\vec{z} = (x_1, x_2, x_3, y_1, y_2, y_3)$, die onder rotaties transformeert als

$$\vec{z} \to \vec{z}' = D\vec{z}, \tag{4.4}$$

waarbij de matrix D als volgt ontbonden kan worden in 3×3 matrices,

$$D = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix}. \tag{4.5}$$

Een dergelijke representatie wordt *reducibel* genoemd, omdat de zes-dimensionale ruimte onder rotaties wordt opsplitst in twee invariante drie-dimensionale deelruimtes. Deze reducibele zes-dimensionale representatie kan dus beschouwd worden als de (z.g. directe) som van twee drie-dimensionale representaties, en we schrijven

$$\mathbf{6} = \mathbf{3} \oplus \mathbf{3}. \tag{4.6}$$

Het zal nu duidelijk zijn dat representaties die *geen* deelruimte invariant laten, en dus niet te schrijven zijn in blokdiagonale vorm zoals in (4.5), *irreducibele representaties* worden genoemd.

Andere representaties kunnen verkregen worden door zogenaamde *product-representaties* te vormen. Beschouwen we bijvoorbeeld een systeem van twee (vrije) deeltjes met golffuncties $\psi_1(\vec{x})$ en $\psi_2(\vec{y})$, waarbij $\vec{x}$ en $\vec{y}$ de coördinaten zijn van de twee deeltjes. De golffuncties $\Psi(\vec{x},\vec{y})$ van het gehele systeem bestaan dan uit alle mogelijke producten van golffuncties ψ_1 en ψ_2, respectievelijk behorende bij deeltje 1 en 2. Dit noemen we een *tensor-product*, hetgeen wordt aangeven door

$$\Psi = \psi_1 \otimes \psi_2. \tag{4.7}$$

Onder draaiingen van zowel $\vec{x}$ als $\vec{y}$ transformeert Ψ nu onder de draaiingsgroep, maar de bijbehorende representatie is gecompliceerder dan die behorende bij de afzonderlijke golffuncties ψ_1 en ψ_2. Het is meestal zo dat zo'n product-representatie niet irreducibel is, en ontbonden kan worden in een aantal afzonderlijke representaties die wel irreducibel zijn. Laten we dit verschijnsel eerst demonstreren aan de hand van het volgende voorbeeld. Laat drie mogelijke functies ψ_1 gegeven zijn door de coördinaten x_i en drie mogelijke functies ψ_2 door de coördinaten y_i. Zowel ψ_1's als ψ_2's transformeren dus volgens de drie-dimensionale representatie van de draaiingsgroep. De product-representatie werkt op alle mogelijke produkten van ψ_1 en ψ_2, en we kunnen daarom negen onafhankelijke funkties onderscheiden,

$$T_{ij}(\vec{x},\vec{y}) = x_i\, y_j, \tag{4.8}$$

die onder draaiingen transformeren volgens

$$T_{ij} \to T'_{ij} = R_{ii'} R_{jj'}\, T_{i'j'}. \tag{4.9}$$

Deze negen-dimensionale representatie is echter niet irreducibel. Zo transformeren het symmetrische en het antisymmetrische deel van T_{ij}, gedefinieerd als $T_{(ij)} \equiv \frac{1}{2}(T_{ij} + T_{ji})$, en $T_{[ij]} \equiv \frac{1}{2}(T_{ij} - T_{ji})$, afzonderlijk onder rotaties. Dit volgt rechtstreeks door het (anti)symmetrische deel van T'_{ij} te bekijken, en op te merken dat de (anti)symmetrie in i en j van $\frac{1}{2}(R_{ii'}R_{jj'} \pm R_{ji'}R_{ij'})$ de (anti)symmetrie in i' en j' impliceert. Vandaar dat geldt

$$T_{(ij)} \to T'_{(ij)} = R_{ii'} R_{jj'}\, T_{(i'j')}, \qquad T_{[ij]} \to T'_{[ij]} = R_{ii'} R_{jj'}\, T_{[i'j']}. \tag{4.10}$$

Het antisymmetrische deel van T_{ij} bevat drie onafhankelijke componenten die transformeren als een drie-dimensionale representatie van de draaiingsgroep. Het symmetrische deel van T_{ij} bevat zes componenten, die echter *niet* transformeren volgens een irreducibele representatie. Dit volgt onmiddelijk uit het feit dat het spoor van T_{ij} gelijk is aan

$$T_{ii} = \vec{x}\cdot\vec{y}, \tag{4.11}$$

en dus invariant is onder draaiingen. We kunnen dus konkluderen dat T_{ij} ontbonden kan worden in drie onafhankelijke tensoren[9],

$$T_{ij} \to \begin{cases} T = \vec{x} \cdot \vec{y} \\ T_i = \epsilon_{ijk} x_j y_k \\ S_{ij} = x_i y_j + x_j y_i - \frac{2}{3}\delta_{ij}(\vec{x} \cdot \vec{y}) \end{cases} \tag{4.12}$$

Merk op dat we het epsilon-symbool gebruikt hebben om het antisymmetrische deel van T_{ij} weer als een drie-dimensionale vector $\vec{T}$ te schrijven (dit is niets anders dan het uitprodukt $\vec{x} \times \vec{y}$). Voorts hebben we het symmetrische deel S_{ij} spoorloos gemaakt door toevoeging van de extra term evenredig met δ_{ij}. Het gevolg hiervan is dat S_{ij} slechts *vijf* ondeling onafhankelijke componenten bevat. Onder draaiingen gaan de bovenstaande termen over in uitdrukkingen van hetzelfde type; zo transformeren de vijf onafhankelijke componenten van S_{ij} onder elkaar.[10] We kunnen dus schrijven voor het product van twee drie-dimensionale representaties,

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}, \tag{4.13}$$

waarbij de representaties worden aangeduid door hun dimensie (we gaan hier gemakshalve voorbij aan het feit dat er mogelijk niet-equivalente representaties kunnen bestaan van gelijke dimensie).

De procedure die we in dit voorbeeld gevolgd hebben, berust op twee elementen; allereerst gebruiken we dat de symmetrie-eigenschappen van tensoren niet veranderen onder de transformaties, en voorts maken we gebruik van het bestaan van twee invariante tensoren, te weten

$$T_{ij} = \delta_{ij}, \qquad T_{ijk} = \epsilon_{ijk}. \tag{4.14}$$

Met een invariante tensor bedoelen we dat de tensor niet verandert onder de groepstransformaties zoals die werken volgens de indexstructuur van de tensor, dus

$$T_{ijk\cdots} \to T'_{ijk\cdots} = R_{ii'} R_{jj'} R_{kk'} \cdots \; T_{i'j'k'\cdots} = T_{ijk\cdots} \tag{4.15}$$

Inderdaad voldoen de beide tensoren δ_{ij} en ϵ_{ijk} aan (4.15); immers, aan de vergelijking

$$R_{ii'} R_{jj'}\, \delta_{i'j'} = \delta_{ij} \tag{4.16}$$

is voldaan omdat de R_{ij} orthogonale matrices zijn, en

$$R_{ii'} R_{jj'} R_{kk'}\; \epsilon_{i'j'k'} = \det R\, \epsilon_{ijk} = \epsilon_{ijk} \tag{4.17}$$

[9]In de tweede vergelijking wordt weer sommatieconventie gebruikt, zie voetnoot op blz. 19

[10]We kunnen voor ieder van deze representaties de matrices $D(R)$ aangeven die gedefinieerd zijn in hoofdstuk 2. Voor de eerste representatie in (4.12) geldt dat $D(R) = 1$. In de tweede representatie hebben we 3×3 matrices $D(R)$ die gelijk zijn aan de matrix R. Voor de derde representatie hebben we 5×5 matrices $D(R)$. De indices hiervan komen overeen met de symmetrisch spoorloze indexparen (ij). De matrices $D(R)$ kunnen worden geschreven als

$$D(R)_{(ij)\,(kl)} = \frac{1}{2}\,(R_{ik}\, R_{jl} + R_{il}\, R_{jl}) - \frac{1}{3}\delta_{ij}\, \delta_{kl}.$$

geldt omdat $\det R = 1$ voor de draaingsmatrices R_{ij}. Voor iedere gegeven tensor $T_{ijk\cdots}$ kunnen we nu de indices contraheren met behulp van de invariante tensoren. Het is dan duidelijk dat aldus gecontraheerde tensoren een invariante deelruimte zullen opspannen, m.a.w. ze zullen onder draaiingen transformeren in tensoren die op dezelfde wijze gevormd zijn. Bijvoorbeeld, zij $T_{ijk\cdots}$ een tensor transformerend als

$$T_{ijk\cdots} \to T'_{ijk\cdots} = R_{ii'} R_{jj'} R_{kk'} \cdots \, T_{i'j'k'\cdots}. \tag{4.18}$$

Vormen we nu de tensor

$$\hat{T}_{klm\cdots} \equiv \delta_{ij} \, T_{ijklm\cdots} \tag{4.19}$$

met twee indices minder. Het is nu eenvoudig te controleren door gebruik te maken van (4.16), dat $\hat{T}$ transformeert volgens

$$\hat{T}_{klm\cdots} \to \hat{T}'_{klm\cdots} = R_{kk'} R_{ll'} R_{mm'} \cdots \, \hat{T}_{k'l'm'\cdots}, \tag{4.20}$$

en op soortgelijke wijze kunnen we nagaan dat contracties met een of meerdere δ- en ϵ-tensoren, gecontraheerde tensoren geven die invariante deelruimtes opspannen. Voor het eerdervermelde voorbeeld kunnen we de ontwikkeling geven door

$$T_{ij} = \frac{1}{2}\epsilon_{ijk} \left(\epsilon_{klm} \, T_{lm}\right) + \frac{1}{2}\left(T_{ij} + T_{ji} - \frac{2}{3}\delta_{ij} T_{kk}\right) + \frac{1}{3}\delta_{ij} \, T_{kk}, \tag{4.21}$$

waarbij de eerste term ook geschreven kan worden als $\frac{1}{2}(T_{ij} - T_{ji})$ met behulp van de identiteit

$$\epsilon_{ijk} \, \epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \,, \tag{4.22}$$

en de tweede term in (4.21) is zodanig geconstrueerd dat hij spoorloos is:

$$\delta_{ij} \left(T_{ij} + T_{ji} - \frac{2}{3}\delta_{ij} T_{kk}\right) = 0 \,. \tag{4.23}$$

5. Ladderoperatoren

Laten we een representatie beschouwen van de draaiingsgroep, gegenereerd door hermitische matrices I_1, I_2 en I_3, die aan dezelfde commutatieregels gehoorzamen als L_1, L_2 en L_3, gegeven in (3.14),

$$[I_1, I_2] = iI_3, \qquad [I_2, I_3] = iI_1, \qquad [I_3, I_1] = iI_2, \tag{5.1}$$

ofwel

$$[I_i, I_j] = i\epsilon_{ijk}\, I_k. \tag{5.2}$$

We eisen dat de matrices I_i hermitisch zijn,

$$I_i^\dagger = I_i \ , \tag{5.3}$$

zodat de matrices $\exp(i\alpha_k I_k)$ unitair zijn. Uitgaande van deze gegevens willen we nu alle irreducibele matrices I_i bepalen die aan deze eigenschappen voldoen. Op deze wijze zijn we dan in staat alle (eindig-dimensionale, unitaire) representaties te bepalen van de draaiingsgroep.

Definieer daartoe eerst de lineaire combinaties

$$I_\pm = I_1 \pm iI_2 \ , \tag{5.4}$$

zodat $(I_\pm)^\dagger = I_\mp$, en

$$[I_\pm, I_3] = [I_1, I_3] \pm i[I_2, I_3] = -iI_2 \mp I_1 = \mp I_\pm \ . \tag{5.5}$$

Dus geldt voor iedere toestand $\mid \psi\rangle$,

$$I_3 \Big(I_+ \mid \psi\rangle \Big) = I_+(I_3 + 1) \mid \psi\rangle. \tag{5.6}$$

Een *Casimir operator* is een combinatie van operatoren voor een representatie die zodanig is dat alle generatoren ermee commuteren. Als de representatie irredubibel is, moet iedere Casimir operator een veelvoud zijn van de eenheidsmatrix. In het geval van de draaiingen in drie dimensies hebben we zo'n Casimir operator:

$$\vec{I}^{\,2} \equiv I_1^2 + I_2^2 + I_3^2 \ . \tag{5.7}$$

We leiden af uit (5.2):

$$[\vec{I}^{\,2}, I_3] = 0. \tag{5.8}$$

Omdat $\vec{I}^{\,2}$ en I_3 twee hermitische commuterende matrices zijn, kunnen we $\vec{I}^{\,2}$ en I_3 tegelijkertijd op diagonaalvorm brengen met reële eigenwaarden. Bovendien moeten de eigenwaarden van $\vec{I}^{\,2}$ positief zijn, omdat geldt dat

$$\langle\psi \mid \vec{I}^{\,2} \mid \psi\rangle = \Big| I_1 \mid \psi\rangle \Big|^2 + \Big| I_2 \mid \psi\rangle \Big|^2 + \Big| I_3 \mid \psi\rangle \Big|^2 \geq 0. \tag{5.9}$$

Het blijkt handig te zijn om de eigenwaarden van $\vec{I}^2$ te schrijven als $l(l+1)$, waarbij $l \geq 0$.

Beschouw nu een toestand $\mid l,m\rangle$ die een eigentoestand is van $\vec{I}^{\,2}$ en I_3, met eigenwaarden $l(l+1)$ en m,

$$\vec{I}^{\,2} \mid l,m\rangle = l(l+1) \mid l,m\rangle, \qquad I_3 \mid l,m\rangle = m \mid l,m\rangle. \tag{5.10}$$

Uit (5.6) en (5.7) leiden we af dat

$$\begin{aligned} I_3\big(I_+ \mid l,m\rangle\big) &= (m+1)\big(I_+ \mid l,m\rangle\big), \\ \vec{I}^2\big(I_+ \mid l,m\rangle\big) &= l(l+1)\big(I_+ \mid l,m\rangle\big). \end{aligned} \tag{5.11}$$

Substitueren we $I_+ \mid l,m\rangle = \mid \psi\rangle$, dan geldt

$$I_3 \mid \psi\rangle = (m+1) \mid \psi\rangle, \qquad \vec{I}^2 \mid \psi\rangle = l(l+1) \mid \psi\rangle, \tag{5.12}$$

m.a.w. $\mid \psi\rangle$ is een nieuwe eigenvector van I_3 en $\vec{I}^2$ met eigenwaarden $m' = m+1$, en $l' = l$, tenzij

$$\mid \psi\rangle \equiv I_+ \mid l,m\rangle = 0. \tag{5.13}$$

Voorts vinden we

$$\begin{aligned} \langle\psi \mid \psi\rangle &= \langle l,m \mid I^- I^+ \mid l,m\rangle \\ &= \langle l,m \mid I_1^2 + I_2^2 + i[I_1,I_2] \mid l,m\rangle \\ &= \langle l,m \mid I_1^2 + I_2^2 - I_3 \mid l,m\rangle \\ &= \langle l,m \mid \vec{I}^2 - I_3(I_3+1) \mid l,m\rangle, \end{aligned} \tag{5.14}$$

waarbij we gebruikt hebben dat $I_+^\dagger = I_-$. Met behulp van (5.10) vinden we dus

$$\langle\psi \mid \psi\rangle = \Big[l(l+1) - m(m+1)\Big]\,\langle l,m \mid l,m\rangle. \tag{5.15}$$

Als we nu aannemen dat $\mid l,m\rangle$ een genormeerde toestand is (dus $\langle l,m \mid l,m\rangle = 1$), dan kunnen we $\mid \psi\rangle$ schrijven als een genormeerde toestand $\mid l,m+1\rangle$ maal een evenredigheidsfaktor die wordt gegeven door (5.15). Deze faktor is bepaald op een fasefaktor na, die we absorberen in de definitie van $\mid l,m+1\rangle$. Op deze wijze konkluderen we dat

$$equation I_+ \mid l,m\rangle = \sqrt{l(l+1) - m(m+1)}\;\mid l,m+1\rangle. \tag{5.16}$$

De operator I_+ geeft dus aanleiding tot toestanden met een steeds hogere eigenwaarde van I_3:

$$\mid l,m\rangle \xrightarrow{I_+} \mid l,m+1\rangle \xrightarrow{I_+} \mid l,m+2\rangle \xrightarrow{I_+} \mid l,m+3\rangle \xrightarrow{I_+} \text{ etc.} \tag{5.17}$$

Vandaar dat I_+ een “ladder-operator” of “stap-operator” wordt genoemd. We zijn echter geïnteresseerd in eindige matrices I_i, zodat de reeks (5.17) ergens moet afbreken. Volgens (5.16) gebeurt dit alleen indien in de reeks (5.17) een toestand voorkomt waarvoor de eigenwaarde van I_3 gelijk wordt aan l. Dit vereist weer dat de oorspronkelijke eigenwaarde m van de toestand waarmee we begonnen zijn, een geheel getal moet verschillen van l. De noodzaak hiervan volgt in feite al uit (5.15): omdat $\langle\psi \mid \psi\rangle$ en $\langle l,m \mid l,m\rangle$ positief zijn moet gelden dat $l(l+1)-m(m+1)\geq 0$, ofwel dat $-l-1\leq m\leq l$. Om te zorgen dat de reeks (5.16) afbreekt wanneer m groter dreigt te worden dan l, moeten we eisen dat $l-m$ een positief geheel getal is. We vinden daarom de situatie

$$\mid l,m\rangle \xrightarrow{I_+} \mid l,m+1\rangle \xrightarrow{I_+} \cdots\cdots \xrightarrow{I_+} \mid l,l\rangle \tag{5.18}$$

waarbij de vektor $\mid l,l\rangle$ met de hoogste eigenwaarde van I_3 voldoet aan

$$I_+ \mid l,l\rangle = 0. \tag{5.19}$$

We kunnen nu eenvoudig nagaan dat de matrix I_- ook een ladder-operator is, maar een die aanleiding geeft tot kleinere eigenwaarden van I_3. Uitgaande van een toestand $\mid l,m\rangle$ kunnen we nu toestanden konstrueren met steeds kleinere eigenwaarden van I_3:

$$\text{etc.} \xleftarrow{I_-} \mid l,m-3\rangle \xleftarrow{I_-} \mid l,m-2\rangle \xleftarrow{I_-} \mid l,m-1\rangle \xleftarrow{I_-} \mid l,m\rangle \tag{5.20}$$

Herhalen van dezelfde bewerkingen als voor I_+ laat zien dat, voor $\mid\psi\rangle = I_- \mid l,m\rangle$,

$$\langle\psi \mid \psi\rangle = \Big[l(l+1)-m(m-1)\Big]\,\langle l,m \mid l,m\rangle. \tag{5.21}$$

zodat moet gelden dat $l(l+1)-m(m-1)\geq 0$, ofwel $-l\leq m\leq l+1$. Aangezien we willen dat de reeks (5.20) afbreekt moet er een toestand zijn met minimale eigenwaarde $m=-l$, zodat

$$I_- \mid l,-l\rangle = 0. \tag{5.22}$$

Analoog aan (5.16) kunnen we nu afleiden, na een geschikte keuze van de fasefaktor in de definitie van de toestand $\mid l,m-1\rangle$, dat moet gelden

$$I_- \mid l,m\rangle = \sqrt{l(l+1)-m(m-1)}\,\mid l,m-1\rangle. \tag{5.23}$$

Uitgaande van een gegeven toestand $\mid l,m\rangle$ hebben we nu $l-m$ toestanden gekonstrueerd met I_3-eigenwaarden $m{+}1, m{+}2,\ldots,l$ en $l{+}m$ toestanden met I_3-eigenwaarden $m{-}1, m{-}2,\ldots,-l$. Op deze wijze vinden we dus $1+(l-m)+(l+m)=2l+1$ toestanden. Vandaar dat $2l+1$ een geheel getal moet zijn, zodat l, en daarom ook m, heeltallig of halftallig is.

Bovenstaande argumenten zijn nog niet voldoende om te bewijzen dat we nu alle toestanden gevonden hebben. In principe zou het mogelijk zijn dat na het achtereenvolgens toepasssen van een I_+ en een I_- operator, we een toestand krijgen met dezelfde eigenwaarden van $\vec{I}^{\,2}$ en I_3, die echter niet gelijk is aan de oorspronkelijke toestand. Men kan algemeen bewijzen dat dit niet mogelijk is. Dat de toestanden $\mid l,m\rangle$ met $-l\leq m<l$ en multipliciteit 1 inderdaad voldoende zijn, kunnen we eenvoudig nagaan door te verifiëren dat aan de commutatierelaties

$$[I_\pm, I_3] = \mp I_\pm, \qquad [I_+, I_-] = 2I_3, \tag{5.24}$$

is voldaan wanneer we gebruik maken voor (5.10), (5.16) en (5.23),

$$\begin{aligned}(I_\pm I_3 - I_3 I_\pm) \mid l,m\rangle &= m\sqrt{l(l+1)-m(m\pm 1)} \mid l,m\pm 1\rangle \\ &\quad -(m\pm 1)\sqrt{l(l+1)-m(m\pm 1)} \mid l,m\pm 1\rangle \\ &= \mp\sqrt{l(l+1)-m(m\pm 1)} \mid l,m\pm 1\rangle \\ &= \mp I_\pm \mid l,m\rangle, \end{aligned} \tag{5.25}$$

$$\begin{aligned}(I_+ I_- - I_- I_+) \mid l,m\rangle &= \sqrt{l(l+1)-(m-1)m}\,\sqrt{l(l+1)-(m-1)m} \mid l,m\rangle \\ &\quad -\sqrt{l(l+1)-(m+1)m}\,\sqrt{l(l+1)-(m+1)m} \mid l,m\rangle \\ &= 2m \mid l,m\rangle \\ &= 2I_3 \mid l,m\rangle. \end{aligned} \tag{5.26}$$

Samenvattend vinden we dus dat een irreducibele representatie van I_1, I_2, I_3 gekarakteriseerd wordt door een getal l, en werkt op een ruimte van $2l+1$ toestanden $\mid l,m\rangle$ waarvoor geldt

$$\begin{aligned}\vec{I}^{\,2} \mid l,m\rangle &= l(l+1) \mid l,m\rangle, \\ I_3 \mid l,m\rangle &= m \mid l,m\rangle, \\ I_\pm \mid l,m\rangle &= \sqrt{l(l+1)-m(m\pm 1)} \mid l,m\pm 1\rangle, \end{aligned} \tag{5.27}$$

met $m = -l,\ -l+1,\ -l+2,\cdots,l-2,\ l-1,l$. Zowel l als m zijn heel- of halftallig.

We geven nu als voorbeeld de representaties voor $l = \frac{1}{2}, 1$ en $\frac{3}{2}$:

- Voor $l = {}^1\!/_2$ vinden we een tweedimensionale representatie. We onderscheiden de toestanden $\mid {}^1\!/_2, {}^1\!/_2\rangle$ en $\mid {}^1\!/_2, -{}^1\!/_2\rangle$, waarvoor geldt, volgens (5.27),

$$\begin{aligned} I_+ \mid {}^1\!/_2, -{}^1\!/_2\rangle &= \mid {}^1\!/_2, {}^1\!/_2\rangle, \\ I_+ \mid {}^1\!/_2, {}^1\!/_2\rangle &= 0, \\ I_- \mid {}^1\!/_2, {}^1\!/_2\rangle &= \mid {}^1\!/_2, -{}^1\!/_2\rangle, \\ I_- \mid {}^1\!/_2, -{}^1\!/_2\rangle &= 0. \end{aligned} \tag{5.28}$$

  Op deze wijze vinden we de matrices

$$I_3 = \begin{pmatrix} {}^1\!/_2 & 0 \\ 0 & -{}^1\!/_2 \end{pmatrix}, \qquad I_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad I_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{5.29}$$

  De matrices I_1, I_2 en I_3 die hieruit volgen zijn precies gelijk aan de matrices $\frac{1}{2}\tau_i$ die we nader zullen bestuderen in hoofdstuk 5.

- Voor $l = 1$ vinden we een driedimensionale representatie. We onderscheiden de toestanden $\mid 1,1\rangle$, $\mid 1,0\rangle$ en $\mid 1,-1\rangle$, waarvoor geldt, volgens (5.27),

$$\begin{aligned} I_+ \mid 1,-1\rangle &= \sqrt{2} \mid 1,0\rangle, \\ I_+ \mid 0\rangle &= \sqrt{2} \mid 1,1\rangle, \\ I_+ \mid 1,1\rangle &= 0, \end{aligned}$$

$$\begin{aligned} I_- \mid 1,1\rangle &= \sqrt{2} \mid 1,0\rangle, \\ I_- \mid 1,0\rangle &= \sqrt{2} \mid 1,-1\rangle, \\ I_- \mid -1,-1\rangle &= 0. \end{aligned} \tag{5.30}$$

Op deze wijze vinden we de matrices

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad I_+ = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}, \qquad I_- = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}. \tag{5.31}$$

De matrices I_1, I_2 en I_3 zijn gelijk aan de matrices L_i, maar in een andere (complexe) basis waarvoor L_3 diagonaal is.

- Voor $l = \frac{3}{2}$ vinden we een vierdimensionale representatie. We onderscheiden de toestanden $\mid \frac{3}{2}, \frac{3}{2}\rangle$, $\mid \frac{3}{2}, {}^1\!/_2\rangle$, $\mid \frac{3}{2}, -{}^1\!/_2\rangle$ en $\mid \frac{3}{2}, -\frac{3}{2}\rangle$, waarvoor geldt, volgens (5.27),

$$\begin{aligned} I_+ \mid \tfrac{3}{2}, -\tfrac{3}{2}\rangle &= \sqrt{3} \mid \tfrac{3}{2}, -{}^1\!/_2\rangle, \\ I_+ \mid \tfrac{3}{2}, -{}^1\!/_2\rangle &= 2 \mid \tfrac{3}{2}, {}^1\!/_2\rangle, \\ I_+ \mid \tfrac{3}{2}, {}^1\!/_2\rangle &= \sqrt{3} \mid \tfrac{3}{2}, \tfrac{3}{2}\rangle, \\ I_+ \mid \tfrac{3}{2}, \tfrac{3}{2}\rangle &= 0. \end{aligned} \tag{5.32}$$

Op deze wijze vinden we de matrices

$$I_3 = \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & {}^1\!/_2 & 0 & 0 \\ 0 & 0 & -{}^1\!/_2 & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}, \qquad I_+ = \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.33}$$

De matrix I_- kan op soortgelijke wijze worden afgeleid uit (5.27), of kan direkt uit I_+ verkregen worden door hermitische conjugatie: $I_- = I_+^\dagger$.

6. De groep $SU(2)$

In het vorige hoofdstuk hebben we representaties geconstrueerd van de drie-dimensionale draaiingsgroep die oneven-dimensionaal zijn. In dit hoofdstuk zullen onderzoeken of er ook andere representaties kunnen voorkomen. We beschouwen daartoe eerst de unitaire 2×2 matrices X met determinant 1:

$$X^\dagger = X^{-1}, \qquad \det X = 1. \tag{6.1}$$

Dergelijke matrices kunnen worden geparametriseerd in termen van twee complexe parameters a en b,

$$X = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{6.2}$$

met $|a|^2 + |b|^2 = 1$. Matrices van de vorm (6.2) met algemene a en b worden *quaternionen* genoemd. Quaternionen vormen een getallenlichaam, net zoals de reële en complexe getallen. De quaternionen met norm 1 (d.w.z. met parameters a en b die voldoen aan $|a|^2 + |b|^2 = 1$) vormen de elementen van de groep $SU(2)$, de groep van unitaire 2×2 matrices met determinant 1.[11] Het is niet moeilijk om in te zien dat de matrices X die voldoen aan (6.1), een groep vormen. Immers, indien X_1 en X_2 beiden voldoen aan (6.1), dan voldoet ook $X_3 = X_1 X_2$ aan (5.1) en behoort dus tot de groep. Verder voldoen de eenheidsmatrix en de inverse matrix aan (6.1) en behoren dus tot de groep, terwijl associativiteit voor de matrixvermenigvuldiging vanzelfsprekend is.

Aangezien complexe getallen a en b die voldoen aan $|a|^2 + |b|^2 = 1$, uitgedrukt kunnen worden in drie reële parameters, hebben we weer te maken met een drie-dimensionale Liegroep, net zoals de rotatiegroep die we hiervoor bekeken hebben. Dit feit suggereert dat de groep $SU(2)$ iets te maken heeft met de draaiingsgroep.

Iedere 2×2 matrix kan worden ontbonden in de de eenheidsmatrix $\mathbf{1}$ en de volgende drie matrices[12],

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.3}$$

Deze τ-matrices voldoen aan de volgende productregel

$$\tau_i \, \tau_j = \delta_{ij} \mathbf{1} + i\epsilon_{ijk} \, \tau_k. \tag{6.4}$$

We kunnen de matrices X nu ontbinden in de matrices τ_i en de eenheidsmatrix $\mathbf{1}$

$$X = c_0 \, \mathbf{1} + ic_i \, \tau_i. \tag{6.5}$$

Vergelijken we dit met de parametrizatie (6.2), dan is duidelijk dat de coëfficiënten c_0 en c_i reëel moeten zijn, terwijl de conditie $|a|^2 + |b|^2 = 1$ impliceert dat

$$c_o^2 + c_i^2 = 1. \tag{6.6}$$

[11] Op soortgelijke wijze vormen de complexe getallen met norm 1 de groep U(1), die bestaat uit alle fasefactoren $\exp i\alpha$.

[12] Ook wel Pauli-matrices genoemd, en vaak aangeduid met σ_i.

$SU(2)$ matrices kunnen dus algemeen geschreven worden in de vorm (6.5) met coëfficiënten die voldoen aan de konditie (6.6).

Laten we nu de infinitesimale $SU(2)$ matrices bekijken

$$X = \mathbf{1} + iB + \mathrm{O}(B^2), \tag{6.7}$$

waarbij de (infinitesimale) matrix B zodanig gekozen moet worden dat X voldoet aan (5.1). Dit leidt tot de volgende restricties

$$\begin{aligned} \Big(\mathbf{1} + iB + \mathrm{O}(B^2)\Big)^\dagger &= \Big(\mathbf{1} + iB + \mathrm{O}(B^2)\Big)^{-1} = \mathbf{1} - iB + \mathrm{O}(B^2), \\ \det\Big(\mathbf{1} + iB + \mathrm{O}(B^2)\Big) &= 1 + \mathrm{Sp}\, B + \mathrm{O}(B^2) = 1, \end{aligned} \tag{6.8}$$

ofwel

$$B^\dagger = B, \qquad \mathrm{Sp}\, B = 0. \tag{6.9}$$

Aangezien iedere hermitische spoorloze 2×2 matrix geschreven kan worden als een lineaire combinatie van de τ-matrices met reële coëfficiënten, kunnen we infinitesimale $SU(2)$-transformaties dus schrijven in de vorm

$$X(\vec{\alpha}) = \mathbf{1} + i\alpha_i \frac{\tau_i}{2} + \mathrm{O}(\alpha^2). \tag{6.10}$$

Deze infinitesimale transformaties kunnen vervolgens geëxponentieerd worden tot een eindige transformatie volgens de limietprocedure uiteengezet in hoofdstuk 3,

$$X(\vec{\alpha}) = \lim_{n\to\infty} \left\{\mathbf{1} + i\frac{\alpha_i}{n}\frac{\tau_i}{2}\right\}^n = \exp\left(i\alpha_i \frac{\tau_i}{2}\right). \tag{6.11}$$

We hebben de parameters α_i zo gekozen dat de generatoren van de $SU(2)$ transformaties gegeven zijn door de matrices $\frac{1}{2}\tau_i$. De reden hiervoor is dat deze matrices voldoen aan dezelfde commutatierelaties als de generatoren L_i van de draaiingsgroep. Dit volgt direkt uit de produktregel (6.4),

$$\left[\frac{\tau_i}{2}, \frac{\tau_j}{2}\right] = i\epsilon_{ijk}\,\frac{\tau_k}{2}. \tag{6.12}$$

Deze eigenschap betekent dat de de groep $SU(2)$ dezelfde struktuurconstanten, en dus dezelfde productstruktuur, heeft als de draaiingsgroep. Dit geldt aleen voor groepselementen in een eindige omgeving van de identiteit, maar niet noodzakelijk voor de groep als geheel. Inderdaad zullen we zien dat de $SU(2)$ matrices in de strikte zin geen representatie van de draaiingsgroep vormen, maar slechts een *projectieve* representatie.

Het is mogelijk om de exponentiële uitdrukking voor $X(\vec{\alpha})$ te berekenen, gebruikmakend van de producteigenschap van de τ-matrices. Dit gaat als volgt

$$\begin{aligned} X(\vec{\alpha}) &= \mathrm{e}^{i\alpha_i \frac{\tau_i}{2}} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{i\alpha_j\tau_j}{2}\right)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n)!}\left(\frac{i\alpha_j\tau_j}{2}\right)^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!}\left(\frac{i\alpha_j\tau_j}{2}\right)^{2n+1}, \end{aligned} \tag{6.13}$$

waar we, in de laatste regel, afzonderlijk sommeren over de even en oneven machten van $(i\alpha_j\tau_j)$. Nu merken we op dat

$$(i\alpha_j\tau_j)^2 = -\alpha_j\alpha_k\,\tau_j\,\tau_k = -\alpha^2\,\mathbf{1}, \tag{6.14}$$

waarbij we gebruik gemaakt hebben van (6.4), en α gedefinieerd hebben als

$$\alpha = \sqrt{\alpha_1^2+\alpha_2^2+\alpha_3^2}. \tag{6.15}$$

Uit (6.14) volgt onmiddelijk dat

$$(i\alpha_j\tau_j)^{2n} = (-)^n\,\alpha^{2n}\,\mathbf{1}, \qquad (i\alpha_j\tau_j)^{2n+1} = (-)^n\,\alpha^{2n}\,(i\alpha_j\tau_j), \tag{6.16}$$

zodat we (6.13) kunnen schrijven als

$$\begin{aligned} X(\vec{\alpha}) &= \left\{\sum_{n=0}^{\infty}\frac{(-)^n}{(2n)!}\left(\frac{\alpha}{2}\right)^{2n}\right\}\mathbf{1} + \left\{\sum_{n=0}^{\infty}\frac{(-)^n}{(2n+1)!}\left(\frac{\alpha}{2}\right)^{2n+1}\right\}\left(\frac{i\alpha_j\tau_j}{\alpha}\right), \\ &= \cos\frac{\alpha}{2}\,\mathbf{1} + i\sin\frac{\alpha}{2}\,\frac{\alpha_j\tau_j}{\alpha}. \end{aligned} \tag{6.17}$$

Dit resultaat komt inderdaad overeen met de decompositie (6.5) waarbij we de volgende uitdrukkingen vinden voor de coëfficiënten c_0 en c_i,

$$c_o = \cos\frac{\alpha}{2}, \qquad c_i = \frac{\alpha_i}{\alpha}\,\sin\frac{\alpha}{2}. \tag{6.18}$$

In hoofdstuk 3 hebben we vastgesteld dat de draaiingen geparametriseerd kunnen worden door vectoren $\vec{\alpha}$ die liggen in een bol met straal $\alpha=\pi$. De richting van $\vec{\alpha}$ geeft de draaias aan en de lengte α de hoek waarover gedraaid wordt. Aangezien draaiingen over $+\pi$ en $-\pi$ radialen gelijk zijn, hebben we vastgesteld dat

$$R(\vec{\alpha}) = R(-\vec{\alpha}), \qquad \text{indien} \qquad \alpha=\pi. \tag{6.19}$$

Zoals volgt uit (6.12) kunnen de elementen van $SU(2)$ geparametriseerd worden door dezelfde vectoren $\vec{\alpha}$, echter, om alle elementen $X(\vec{\alpha})$ te parametriseren moeten we de straal van de bol twee keer zo groot nemen en dus gelijk aan 2π. Twee tegengesteld gerichte vectoren, $\vec{\alpha}$ en $\vec{\alpha}\,'$, in deze bol waarvoor geldt dat $\alpha+\alpha'=2\pi$, geven aanleiding tot dezelfde draaiing,

$$R(\vec{\alpha}\,') = R(\vec{\alpha}), \tag{6.20}$$

omdat sprake is van dezelfde draaias en draaiingshoeken die 2π van grootte verschillen. De bijbehorende $SU(2)$ elementen, $X(\vec{\alpha}\,')$ en $X(\vec{\alpha})$, zijn echter elkaars tegengestelde,

$$X(\vec{\alpha}\,') = -X(\vec{\alpha}). \tag{6.21}$$

Dit laatste volgt uit (6.17), (6.18) en het feit dat $\cos\frac{\alpha'}{2} = -\cos\frac{\alpha}{2}$ en $\sin\frac{\alpha'}{2} = \sin\frac{\alpha}{2}$.

Het bovenstaande betekent dat $SU(2)$ elementen strikt genomen geen representatie vormen van de draaiingsgroep, maar een *projectieve* representatie. Immers, voor het product van draaiingen

$$R(\vec{\alpha})\,R(\vec{\beta}) = R(\vec{\gamma}), \tag{6.22}$$

met $\alpha,\ \beta,\ \gamma \leq \pi$, hebben we het product van de bijbehorende $SU(2)$ matrices

$$X(\vec{\alpha})\,X(\vec{\beta}) = \pm X(\vec{\gamma}), \tag{6.23}$$

waarbij het teken aan de rechterkant van de vergelijking afhangt van $\vec{\alpha}$ en $\vec{\beta}$.[13]

Laten we nu representaties bekijken van de groep $SU(2)$. We laten de $SU(2)$ matrices werken in een abstracte vectorruimte met *complexe* coördinaten φ^α met $\alpha = 1,\ 2$, zodat[14]

$$\varphi^\alpha \to \varphi^{\alpha\prime} = X^\alpha_{\ \beta}\,\varphi^\beta, \tag{6.24}$$

De complex-geconjugeerde vectoren transformeren dan volgens

$$\varphi^*_\alpha \to \varphi^{*\prime}_\alpha = (X_\alpha^{\ \beta})^*\,\varphi^*_\beta, \tag{6.25}$$

We hebben hier een belangrijke nieuwe notatie ingevoerd: de indices staan soms bovenaan (superscripts), en soms onderaan (subscripts). We doen dit om aan te geven dat vectoren met indices bovenin, zoals in (6.24), anders transformeren dan vectoren met onder-indices, zoals (6.25). Bovendien beperken we de sommatieconventie alleen tot die gevallen waar een bovenindex gelijkgesteld wordt met een onderindex:

$$\phi_\alpha \psi^\alpha \equiv \sum_{\alpha=1}^{2} \phi_\alpha \psi^\alpha\ . \tag{6.26}$$

Het blijkt echter dat de representatie corresponderend met de matrices X^* equivalent is met de oorspronkelijke representatie. Om dit te bewijzen, beschouw de vergelijking

$$\begin{aligned} \tau_2\,X\tau_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^* & b^* \\ -b & a \end{pmatrix} \\ &= X^*. \end{aligned} \tag{6.27}$$

Omdat $\tau_2^2 = \mathbf{1}$, volgt dat de twee representaties gegeven door (6.24) en (6.25) equivalent zijn volgens de definitie gegeven in (4.3). Het bestaan van bovenstaande relatie is gemakkelijker te begrijpen door op te merken dat

$$X^\alpha_{\ \alpha'}\,X^\beta_{\ \beta'}\,\epsilon^{\alpha'\beta'} = \det X\,\epsilon^{\alpha\beta} = \epsilon^{\alpha\beta}. \tag{6.28}$$

Gebruikmakend van $\epsilon^{\alpha\beta} = i(\tau_2)^{\alpha\beta}$ kunnen we (6.28) als een matrixvergelijking schrijven,

$$X\,i\tau_2\,\tilde{X} = i\tau_2. \tag{6.29}$$

Aangezien $\tilde{X} = (X^\dagger)^* = (X^{-1})^*$, volgt onmiddellijk vergelijking (6.27). Merk echter op dat we hier in τ_2 de indices naar boven hebben gezet. Het is zo dat in (6.28) de tensor $\varepsilon^{\alpha\beta}$ beide indices bovenin moet hebben. Alleen $\varepsilon^{\alpha\beta}$, $\varepsilon_{\alpha\beta}$ en δ^α_β zijn invariante tensoren.

[13]Anderzijds kunnen we wel stellen dat de drie-dimensionale draaiingen een representatie vormen van de groep $SU(2)$.

[14]De coördinaten φ^α hebben dus geen voor de hand liggende fysische betekenis

We beschouwen nu analytische functies f van deze twee coördinaten. Maak je de Taolor-reeksontwikkeling van zo'n functie bij de oorsprong, dan vormen de Taylorreeks termen in de N^{e} orde homogene, symmetrische polynomen van graad N. Uiteraard is N een niet-negatief, geheel getal. Schrijf

$$Y^{\alpha_1\alpha_2\cdots\alpha_N} = \varphi^{\alpha_1}\,\varphi^{\alpha_2}\cdots\varphi^{\alpha_N}. \tag{6.30}$$

Onder $SU(2)$ transformeren deze polynomen als volgt

$$Y^{\alpha_1\alpha_2\cdots\alpha_N} \to Y^{\alpha_1\alpha_2\cdots\alpha_N\,\prime} = X^{\alpha_1}_{\ \alpha'_1}\,X^{\alpha_2}_{\ \alpha'_2}\cdots X^{\alpha_N}_{\ \alpha'_N}\,Y^{\alpha'_1\alpha'_2\cdots\alpha'_N}. \tag{6.31}$$

Anders dan in het geval van de rotatiegroep kunnen we geen groep-invariante contracties uitvoeren met δ-tensoren, aangezien

$$X^{\alpha}_{\ \alpha'}\,X^{\beta}_{\ \beta'}\,\delta^{\alpha'\beta'} \neq \delta^{\alpha\beta}, \tag{6.32}$$

dit omdat X i.h.a. niet orthogonaal is, maar unitair. De enige deltafunctie die we mogen gebruiken is er een met één bovenindex en één onderindex: δ^{α}_{β}. We bekijken ook niet de complex-geconjugeerde coördinaten φ^*, aangezien die transformeren onder een equivalente representatie zoals we al eerder hebben vastgesteld. Gezien het voorafgaande verwachten we dat de tensoren $Y^{\alpha_1\alpha_2\cdots\alpha_N}$ volgens irreducibele representaties van $SU(2)$ zullen transformeren. De onafhankelijke polynomen van dit type zijn volledig gekarakteriseerd door het aantal indices p_1 te specificeren dat gelijk is aan 1 (het aantal overgebleven indices, $p_2 = N - p_1$, moet dan gelijk zijn aan 2), en we tellen dus

$$\sum_{p_1=0}^{N} = N+1 \tag{6.33}$$

onafhankelijke polynomen van graad N.

Vervolgens kunnen we de $SU(2)$ generatoren werkende op functies van de coördinaten φ schrijven als differentiaaloperatoren. Dit leidt tot

$$L_i^{SU(2)} = -\frac{1}{2}(\tau_i)^{\alpha}_{\ \beta}\,\varphi^{\beta}\,\frac{\partial}{\partial\varphi^{\alpha}}, \tag{6.34}$$

zodat infinitesimale $SU(2)$ transformaties van functies $\psi(\varphi)$ geschreven kunnen worden als

$$\begin{aligned}\psi(\varphi) \to \psi'(\varphi) &= \left(\mathbf{1} - i\vec{\alpha}\cdot\vec{L}^{\,SU(2)} + \mathrm{O}(\alpha^2)\right)\psi(\varphi) \\ &= \psi(\varphi) + \frac{i}{2}\alpha_j\,(\tau_j)^{\alpha}_{\ \beta}\,\varphi^{\beta}\,\frac{\partial\psi(\varphi)}{\partial\varphi^{\alpha}} + \mathrm{O}(\alpha^2)\ . \end{aligned}\tag{6.35}$$

Merk overigens op dat de index α in $\frac{\partial}{\partial\phi^{\alpha}}$ behandeld wordt als een *onder*-index.

Gebruikmakend van (6.34) kunnen we nu de Casimir operator $(\vec{L}^{SU(2)})^2$ schrijven als een differentiaaloperator,

$$(L_i^{SU(2)})^2 = \frac{1}{4}\sum_i\left(\tau^{\alpha}_{i\,\beta}\,\varphi^{\beta}\,\frac{\partial}{\partial\varphi^{\alpha}}\right)\left(\tau^{\gamma}_{i\,\delta}\,\varphi^{\delta}\,\frac{\partial}{\partial\varphi^{\gamma}}\right)$$

$$
\begin{aligned}
&= \frac{1}{4}\big(-\delta^\alpha_\beta\,\delta^\gamma_\delta + 2\,\delta^\alpha_\delta\,\delta^\gamma_\beta\big)\,\varphi^\beta\,\frac{\partial}{\partial\varphi^\alpha}\,\varphi^\delta\,\frac{\partial}{\partial\varphi^\gamma} \\
&= -\frac{1}{4}\,\varphi^\alpha\,\frac{\partial}{\partial\varphi^\alpha}\,\varphi^\gamma\,\frac{\partial}{\partial\varphi^\gamma} + \frac{1}{2}\,\varphi^\beta\,\frac{\partial}{\partial\varphi^\alpha}\,\varphi^\alpha\,\frac{\partial}{\partial\varphi^\beta} \\
&= \frac{1}{4}\left(\varphi^\alpha\frac{\partial}{\partial\varphi^\alpha}\right)^2 + \frac{1}{2}\left(\varphi^\alpha\frac{\partial}{\partial\varphi^\alpha}\right). \qquad (6.36)
\end{aligned}
$$

Het is eenvoudig in te zien dat de laatste twee regels van (6.36) aan elkaar gelijk zijn door alle afgeleiden rechts van de coördinaten te schrijven. De overgang van de eerste naar de tweede regel is minder triviaal. Daar hebben we gebruik gemaakt van de identiteit

$$
\sum_i (\tau_i)^\alpha{}_\beta\,(\tau_i)^\gamma{}_\delta = -\delta^\alpha_\beta\,\delta^\gamma_\delta + 2\,\delta^\alpha_\delta\,\delta^\gamma_\delta. \qquad (6.37)
$$

Een handige manier om deze vergelijking af te leiden is gebruik te maken van het feit dat iedere 2×2 matrix A ontbonden kan worden in de eenheidsmatrix en de drie τ-matrices

$$
A^\alpha{}_\beta = \frac{1}{2}\mathbf{1}^\alpha_\beta\,\mathrm{Sp}(A) + \frac{1}{2}(\tau_i)^\alpha{}_\beta\,\mathrm{Sp}(A\,\tau_i) \qquad (6.38)
$$

Als we nu de afgeleide nemen naar $A_{\delta\gamma}$ van deze vergelijking, die lineair is in de matrixelementen van A, dan vinden we de identiteit (6.37). Vanzelfsprekend kunnen we de geldigheid van (6.37) ook rechtstreeks verifiëren voor gegeven waarden van de indices α, β, γ en δ.

Laten we nu de operator (6.36) werken op de polynomen $Y^{\alpha_1\alpha_2\cdots\alpha_{2s}}$. Gebruikmakend van het feit dat

$$
\left(\varphi^\alpha\frac{\partial}{\partial\varphi^\alpha}\right) Y^{\alpha_1\alpha_2\cdots\alpha_N} = N\,Y^{\alpha_1\alpha_2\cdots\alpha_{2s}}, \qquad (6.39)
$$

vinden we direkt het resultaat

$$
(L_i^{SU(2)})^2\,Y^{\alpha_1\alpha_2\cdots\alpha_N} = \tfrac{1}{2}N(\tfrac{1}{2}N+1)\,Y^{\alpha_1\alpha_2\cdots\alpha_N}. \qquad (6.40)
$$

We herkennen nu de representaties ℓ van hoofdstuk 5, als we schrijven $\ell = s$, $s = \tfrac{1}{2}N$ We zijn er dus in geslaagd om $(2s+1)$-dimensionale representaties van $SU(2)$ te konstrueren, waarbij s heel- of halftallig kan zijn. De eigenwaarde van de Casimir-operator in deze representaties, is volgens (6.40) gelijk aan $s(s+1)$. In hoofdstuk 5 is bewezen dat hiermee *alle* representaties van $SU(2)$ zijn gevonden.

We verwachten dat, voor heeltallige waarden van s, de representaties samenvallen met de representaties van de draaingsgroep die we gevonden hebben in het vorige hoofdstuk. Dit blijkt inderdaad het geval te zijn. Om dit in te zien, beschouw de tensor $Y^{\alpha\beta}$, kwadratisch in de coördinaten φ,

$$
Y^{\alpha\beta} = \varphi^\alpha\,\varphi^\beta. \qquad (6.41)
$$

Er zijn drie onafhankelijke tensoren van dit type, die we als volgt kunnen schrijven,

$$
\xi_i = \varphi^\alpha\varepsilon_{\alpha\gamma}\,(\tau_i)^\gamma{}_\beta\,\varphi^\beta\ . \qquad (6.42)
$$

waarbij we gebruik hebben gemaakt van het feit dat iedere symmetrische 2×2 matrix ontbonden kan worden in de onafhankelijke symmetrische matrices $\varepsilon\tau_i$. Onder $SU(2)$ transformeren de grootheden ξ_i als

$$\begin{aligned} \xi_i \to \xi_i' &= \varphi^{\alpha'}\, X^{\alpha}{}_{\alpha'}\, \varepsilon_{\alpha\gamma}\, \tau_{i\,\beta}^{\gamma}\, X^{\beta}{}_{\beta'}\, \varphi^{\beta'} \\ &= \varphi^{\alpha'}\, \varepsilon_{\alpha'\gamma'} (X^{-1})^{\gamma'}{}_{\gamma} (\tau_i X)^{\gamma}{}_{\beta'}\, \varphi^{\beta'}, \end{aligned} \tag{6.43}$$

waarbij we gebruik hebben gemaakt van de transformatieregel voor $\varepsilon_{\alpha\beta}$, die net is als voor $\varepsilon^{\alpha\beta}$, zie (6.28) . Men kan nu bewijzen dat

$$X^{-1}(\vec{\alpha})\, \tau_i\, X(\vec{\alpha}) = R(\vec{\alpha})_{ij}\, \tau_j, \tag{6.44}$$

zodat de tensoren ξ_i dus op dezelfde wijze transformeren als de coördinaten x_i in hoofdstuk 4. We zullen de juistheid van (5.43) verifieren voor infinitesimale transformaties. Dan geldt

$$\begin{aligned} X^{-1}(\vec{\alpha})\, \tau_i\, X(\vec{\alpha}) &\approx \Big(\mathbf{1} - \frac{i}{2}\alpha_j\, \tau_j + \mathrm{O}(\alpha^2)\Big)\; \tau_i\, \Big(\mathbf{1} + \frac{i}{2}\alpha_k\, \tau_k + \mathrm{O}(\alpha^2)\Big) \\ &= \Big(\mathbf{1} + \frac{i}{2}\alpha_j\; [\tau_i, \tau_j] + \mathrm{O}(\alpha^2)\Big) \\ &= \tau_i + \epsilon_{ijk}\, \tau_j\, \alpha_k + \mathrm{O}(\alpha^2), \end{aligned} \tag{6.45}$$

hetgeen inderdaad dezelfde vorm aanneemt als infinitesimale rotaties werkend op de coördinaten x_i.

7. Spin en hoekverdeling

In de voorafgaande hoofdstukken hebben we gezien hoe de mogelijke golffunkties van een rotatie-invariant systeem geklassificeerd kunnen worden volgens representaties van de draaiingsgroep. Deze representaties worden gekarakteriseerd door een heel- of halftallig getal, l of s genoemd. We kennen dit verschijnsel al van het waterstofatoom, waar de mogelijke golffunkties afhangen van een heeltallige l, die het totale impulsmoment definieert van de bijbehorende toestand. De generatoren, L_k, zoals we die gedefinieerd hebben in de vorige hoofdstukken (i.h.b. in (4.29)), korrespondeert met de impulsmoment-operator. Dit volgt uit het feit dat de quantummechanische operator behorende bij de impuls, werkend op golffunkties $\psi(\vec{r})$ de vorm heeft

$$\vec{p} = \frac{\hbar}{i}\frac{\partial}{\partial\vec{r}}\,. \tag{7.1}$$

De operator behorende bij het baanimpulsmoment, $\vec{L} \equiv \vec{r} \times \vec{p}$, neemt nu dezelfde vorm aan als de generatoren (4.29), afgezien van een faktor $\hbar$. Volgens (4.50) volgt daarom dat de totale baanimpuls van de toestanden gegeven wordt door de eigenwaarde-vergelijking

$$\vec{L}^2 \mid \psi\rangle = l(l+1)\,\hbar^2 \mid \psi\rangle. \tag{7.2}$$

Echter, het is ook mogelijk dat deeltjes een *intrinsiek* impulsmoment hebben, een bijdrage tot het impulsmoment die te maken heeft met de interne struktuur van een deeltje en niet met zijn (zwaartepunts)beweging. In principe is het mogelijk dit te begrijpen door aan te nemen dat het deeltje eindige afmetingen heeft, zodat de rotatie om het zwaartepunt verantwoordelijk is voor deze extra bijdrage aan het impulsmoment, die ook wel *spin* wordt genoemd. Deze situatie is bekend uit de klassieke mechanika (een voorbeeld is de roterende tol). De bovenstaande interpretatie van het intrinsieke impulsmoment is niet vrij van problemen, maar voor ons betoog speelt dit geen rol. Ons uitgangspunt is dat een intrinsiek impulsmoment een natuurlijke plaats heeft binnen het formalisme dat we in de vorige hoofdstukken hebben geïntroduceerd. Het is dan aan het experiment om te beslissen of de Natuur inderdaad van deze mogelijkheid gebruik maakt. Voor deeltjes met spin is het natuurlijk niet langer vanzelfsprekend dat alleen representaties optreden met heeltallige waarden van l. Inderdaad blijkt dat ook deeltjes met halftallige spin in de Natuur voorkomen, zoals het elektron, het proton en het neutron.

Uit de representatietheorie volgt dat een deeltje met spin s in $2s+1$ verschillende toestanden kan voorkomen.[15] Vandaar dat de golffunktie $2s+1$ verschillende komponenten heeft. Onder rotaties gaan die verschillende toestanden in elkaar over volgens de bijbehorende representatie van de draaiingsgroep. Het feit dat deeltjes in verschillende spintoestanden voorkomen, heeft natuurlijk implikaties voor het experiment. Veel van de experimenten waaruit het bestaan van spin, alsmede de daarmee samenhangende eigenschappen, kunnen worden afgeleid, maken gebruik van het door de spin geïnduceerde

[15] Een uitzondering moeten we maken voor *massaloze* deeltjes, zoals het foton, die altijd met de lichtsnelheid bewegen. Voor massaloze deeltjes zijn er slechts twee verschillende toestanden mogelijk; spin wordt in dat geval op een andere wijze gedefinieerd.

magnetisch moment. De bekende experimenten van Stern en Gerlach zijn hier een voorbeeld van. Een andere meting waaruit het bestaan van verschillende spintoestanden kan blijken, is die van de hoekverdeling in verstrooiingsexperimenten. In dit hoofdstuk zullen we dit demonstreren aan de hand van de verstrooiing van protonen aan een koolstofkern.

Protonen hebben spin $s = \frac{1}{2}$ zodat we te maken hebben met twee verschillende soorten protonen, die we kunnen beschrijven met een twee-componentige golffunktie. In de praktijk schrijven we deze golffunktie vaak als een produkt van een funktie die het ruimtelijke deel beschrijft, en een twee-componentige spinor. De twee verschillende protonen worden dan door de onafhankelijke golffunkties $\psi_\uparrow$ en $\psi_\downarrow$ beschreven,

$$\psi_\uparrow(\vec{r}) = f(\vec{r}) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \psi_\downarrow(\vec{r}) = f(\vec{r}) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.3}$$

Bovenstaande toestanden zijn eigentoestand van de generator $L_3^{SU(2)}$, die geassocieerd is met draaiingen om de z-as. Voor de eerste toestand zeggen we daarom dat de spin gericht staat langs de positieve z-as ($S_z = \frac{1}{2}\hbar$), voor de tweede toestand is de spin tegengesteld gericht ($S_z = -\frac{1}{2}\hbar$). Onder rotaties kunnen deze twee toestanden in elkaar overgaan. Bijvoorbeeld, een rotatie om de x-as van π radialen werkt op de spinoren door middel van de matrix (zie (5.17))

$$X = i\tau_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \tag{7.4}$$

zodat de beide toestanden (6.3) in elkaar worden overgevoerd.

Beschouwen we nu de situatie van een proton dat verstrooid wordt aan een koolstofkern (C^{12}). Laat het proton invallen langs de postieve x-as op een koolstofkern in de oorsprong van het coördinatenstelsel. Koolstofkernen zijn (in goede benadering) bolsymmetrisch, en we kunnen nu de verstrooiing van protonen met spin “op” (d.w.z. $S_z = +\frac{1}{2}\hbar$) vergelijken met de verstrooiing van protonen met spin ”neer” ($S_z = -\frac{1}{2}\hbar$). Het blijkt nu dat de ene soort protonen bij voorkeur afwijkt in de richting van de negatieve y-as, en de andere soort in de richting van de positieve y-as. Dit volgt uit de *hoekverdeling*, die de waarschijnlijkheid meet voor het proton om te verstrooien over een hoek θ, in dit geval gedefinieerd als de hoek tussen de z-as en het verstrooiingsvlak (zie figuur). Het blijkt dat de hoekverdeling op gevoelige wijze afhangt van de spin van het inkomende deeltje (we nemen aan dat het deeltje waaraan we verstrooien, in dit geval de koolstofkern, geen spin heeft en bolsymmetrisch is); hoe groter de spin van het inkomende deeltje is, des te gecompliceerder zal de hoekverdeling zijn. Wat we hier zullen aantonen is dat de hoekverdeling in feite volledig wordt bepaald door de eis van rotatie-invariantie, afgezien van enkele constanten.

Laat ons nu de quantummechanische toestanden bespreken voor en na de botsing.[16]

[16]De inkomende toestanden worden nu gekarakteriseerd door aan te geven hoe de golffunktie er uit ziet voor $t \to -\infty$, m.a.w. voor asymptotische tijden lang voordat de verstrooiing heeft plaatsgevonden. Op soortgelijke wijze worden de uitgaande toestanden gespecificeerd aan de hand van gedaante van de golffunktie voor $t \to \infty$, m.a.w. lang nadat de verstrooiing heeft plaatsgevonden. Toestanden op deze wijze gespecificeerd worden “in”- en “uit”-toestanden genoemd. Om het inproduct te berekenen moeten we de bijbehorende golffunkties gebruiken, genomen op hetzelfde tijdstip. Hoewel deze aspekten conceptueel van groot belang zijn, spelen ze hier geen belangrijke rol.

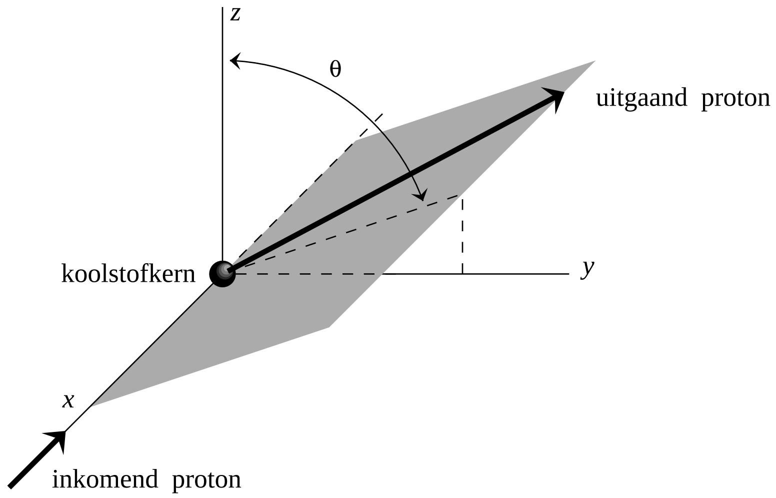

Figure 3: Protonverstrooiing aan een koolstofkern. De protonen vallen in langs de x-as en worden verstrooid onder een hoek θ. Deze hoek is gedefinieerd als de hoek tussen de z-as en het verstroooiingsvlak, het vlak waarin het inkomende en uitgaande proton zich bewegen.

Vóór de botsing hebben we te maken met een inkomend proton waarvan de impuls is bepaald door de experimentele condities, terwijl we de spintoestand naar wens kunnen kiezen; de koolstofkern bevindt zich in zijn grondtoestand. De inkomende toestanden die we kunnen bekijken worden daarom aangeduid met $|\uparrow\rangle$ en $|\downarrow\rangle$, overeenkomende met de proton-golffunkties (6.3). De mogelijke toestanden na de verstrooiing zijn veel moeilijker te karakteriseren. Het proton is van impuls en spin veranderd, en ook de koolstofkern zal i.h.a. niet meer in rust zijn en mogelijk in een geëxciteerde toestand zijn geraakt. Echter, afgezien van de verstrooiingshoek θ, die bepaald wordt door het vlak waarin het inkomende en uitgaande proton zich bewegen, is het experiment niet in staat om alle details betreffende het koolstof-protonsysteem te meten, of in ieder geval zijn we niet in deze details geïnteresseerd. De mogelijke uitgaande toestanden zullen we aangeven met $|\ \psi_{\{\rho\}}(\theta)\rangle$, waarbij $\{\rho\}$ de grootheden aanduidt die samenhangen met de details van de toestand die we *niet* zullen, of kunnen meten.

De waarschijnlijkheid dat, na verstrooiing met een proton met spin "op", er een uitgaande toestand gemeten wordt gekarakteriseerd door $\{\rho\}$ en θ, is gelijk aan $|\langle\psi_{\{\rho\}}(\theta)\ |\uparrow\rangle|^2$. Omdat we niet geïnteresseerd zijn in de grootheden $\{\rho\}$, is de kans om een verstrooiingshoek θ te meten onafhankelijk van de waarden voor $\{\rho\}$, gegeven door[17]

$$f_1(\theta) = \sum_{\{\rho\}} \left|\langle\psi_{\{\rho\}}(\theta)\ |\uparrow\rangle\right|^2. \tag{7.5}$$

Op soortgelijke wijze vinden we dat de waarschijnlijkheid om, na verstrooiing van een proton met spin "neer", een uitgaand proton te meten met verstrooiingshoek θ, gelijk is

[17]Hiermee bedoelen we dat de kans om een verstrooiingshoek te meten tussen θ en $\theta + \mathrm{d}\theta$ gelijk is aan $f_1(\theta)\,\mathrm{d}\theta$. Het is voor ons van ondergeschikt belang hoe de expliciete sommatie over $\{\rho\}$ moet worden uitgevoerd.

aan

$$f_2(\theta) = \sum_{\{\rho\}} \left| \langle \psi_{\{\rho\}}(\theta) \mid \downarrow \rangle \right|^2. \tag{7.6}$$

Onder een rotatie over π radialen om de x-as gaat de hoek θ over in $\theta + \pi$, terwijl de twee spintoestanden (6.3) van het inkomende proton in elkaar overgaan. Vandaar dat rotatie-invariantie impliceert dat moet gelden

$$f_1(\theta) = f_2(\theta + \pi). \tag{7.7}$$

We kunnen echter nog meer zeggen. Beschouw een inkomend proton met spin gekarakteriseerd door

$$\mid \psi(\mathbf{a}) \rangle = a_1 \mid \uparrow \rangle + a_2 \mid \downarrow \rangle. \tag{7.8}$$

Op de zelfde wijze als in (6.5) en (6.6) leiden we dan af dat de hoekverdeling $f(\theta)$ gegeven wordt door

$$\begin{aligned} f(\theta) &= \sum_{\{\rho\}} \left| \langle \psi_{\{\rho\}}(\theta) \mid \psi(\mathbf{a}) \rangle \right|^2 \\ &= \sum_{\{\rho\}} \left| a_1 \langle \psi_{\{\rho\}}(\theta) \mid \uparrow \rangle + a_2 \langle \psi_{\{\rho\}}(\theta) \mid \downarrow \rangle \right|^2 \\ &= f_1(\theta) \left| a_1 \right|^2 + f_2(\theta) \left| a_2 \right|^2 + f_{12}(\theta)\, a_1^* a_2 + f_{12}^*(\theta)\, a_1 a_2^*, \end{aligned} \tag{7.9}$$

waarbij f_1 en f_2 gedefinieerd zijn in (6.5) en (6.6), terwijl de mengterm gegeven wordt door

$$f_{12}(\theta) = \sum_{\{\rho\}} \langle \psi_{\{\rho\}}(\theta) \mid \downarrow \rangle \, \langle \psi_{\{\rho\}}(\theta) \mid \uparrow \rangle^*. \tag{7.10}$$

We kunnen (6.9) ook in matrix-notatie schrijven,

$$f(\theta) = \sum_{\alpha,\beta=1,2} a_\alpha^* \, F(\theta)_{\alpha\beta} \, a_\beta, \tag{7.11}$$

waarbij de 2×2 matrix $F(\theta)$, gedefinieerd door

$$F(\theta) = \begin{pmatrix} f_1(\theta) & f_{12}(\theta) \\ f_{12}^*(\theta) & f_2(\theta) \end{pmatrix}, \tag{7.12}$$

hermitisch en positief-definiet is.

Beschouw nu een rotatie van de gehele experimentele opstelling over een hoek ϕ om de x-as, zodat de verstrooiingshoek θ overgaat in $\theta + \phi$. Onder deze rotatie transformeren vectoren m.b.v. de matrix

$$R(\phi) = \mathrm{e}^{i\phi L_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix}, \tag{7.13}$$

terwijl de toestanden $|\uparrow\rangle$ en $|\downarrow\rangle$ in elkaar overgaan m.b.v. de matrix

$$X(\phi) = \mathrm{e}^{i\phi\tau_1/2} = \begin{pmatrix} \cos\frac{\phi}{2} & i\sin\frac{\phi}{2} \\ i\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}, \tag{7.14}$$

Dit betekent dat de twee-componentige vector $\mathbf{a} = (a_1, a_2)$ die de toestand $|\,\psi(\mathbf{a})\rangle$ karakteriseert, transformeert volgens de gespiegelde van $X(\phi)$,

$$\mathbf{a} \to \mathbf{a}' = \widetilde{X}(\phi)\,\mathbf{a}. \tag{7.15}$$

Na de rotatie van de gehele experimentele opstelling moet gelden dat de waarde van de hoekverdeling gemeten voor $\theta + \phi$ en behorende bij een inkomende protontoestand gekarakteriseerd door $\mathbf{a}'$, gelijk moet zijn aan de oorspronkelijke waarde, d.w.z. de hoekverdeling gemeten bij een hoek θ en behorende bij een inkomende protontoestand gekarakteriseerd door de vector $\mathbf{a}$. Dit betekent dat moet gelden dat

$$\widetilde{X}^{\dagger}(\phi)\,F(\theta+\phi)\,\widetilde{X}(\phi) = F(\theta). \tag{7.16}$$

Deze vergelijking bepaalt de θ afhankelijkheid van de hoekverdeling. Neem namelijk $\theta = 0$ in (6.16), zodat we deze vergelijking kunnen schrijven als (we gebruiken $X(\phi)^{\dagger} = X(\phi)^{-1}$)

$$F(\phi) = \widetilde{X}(\phi)\,F(0)\,\widetilde{X}(\phi)^{-1}. \tag{7.17}$$

Het uitwerken van deze matrixvergelijking voor de afzonderlijke componenten geeft dan aanleiding tot (we vervangen ϕ nu weer door θ)

$$f_1(\theta) = f_1\cos^2\tfrac{\theta}{2} + f_2\sin^2\tfrac{\theta}{2} - i(f_{12} - f_{12}^*)\cos\tfrac{\theta}{2}\sin\tfrac{\theta}{2}, \tag{7.18}$$

$$f_2(\theta) = f_2\cos^2\tfrac{\theta}{2} + f_1\sin^2\tfrac{\theta}{2} + i(f_{12} - f_{12}^*)\cos\tfrac{\theta}{2}\sin\tfrac{\theta}{2}, \tag{7.19}$$

$$f_{12}(\theta) = f_{12}\cos^2\tfrac{\theta}{2} + f_{12}^*\sin^2\tfrac{\theta}{2} - i(f_1 - f_2)\cos\tfrac{\theta}{2}\sin\tfrac{\theta}{2}, \tag{7.20}$$

waarbij $f_1 \equiv f_1(0)$, $f_2 \equiv f_2(0)$ and $f_{12} \equiv f_{12}(0)$. Bovenstaand resultaat kan herschreven worden als

$$f_1(\theta) = \tfrac{1}{2}(f_1+f_2) + \tfrac{1}{2}(f_1-f_2)\cos\theta - \tfrac{i}{2}(f_{12}-f_{12}^*)\sin\theta, \tag{7.21}$$

$$f_2(\theta) = \tfrac{1}{2}(f_1+f_2) - \tfrac{1}{2}(f_1-f_2)\cos\theta + \tfrac{i}{2}(f_{12}-f_{12}^*)\sin\theta, \tag{7.22}$$

$$f_{12}(\theta) = \tfrac{1}{2}(f_{12}+f_{12}^*) + \tfrac{1}{2}(f_{12}-f_{12}^*)\cos\theta - \tfrac{i}{2}(f_1-f_2)\sin\theta. \tag{7.23}$$

De θ-afhankelijkheid van de hoekverdelingen volgt dus geheel uit rotatie-invariantie. De funkties $f_1(\theta)$ en $f_2(\theta)$ hebben precies één maximum en één minimum voor het interval $-\pi \leq \theta \leq \pi$ en geven dus inderdaad aanleiding tot één voorkeursrichting bij verstrooiing, in overeenstemming met wat we al eerder beschreven hebben.

Het is mogelijk om de hoekverdelingen nog verder te specificeren door gebruik te maken van nog een symmetrie. In zeer goede benadering is de natuur namelijk ook invariant onder ruimtelijke *spiegelingen*. Onder een spiegeling t.o.v. de oorsprong, P, veranderen de coördinaten in hun tegengestelde,

$$\vec{r} \xrightarrow{\mathcal{P}} -\vec{r}. \tag{7.24}$$

Tweemaal spiegelen geeft de identiteit, zodat we hier te maken hebben met een eindige groep bestaande uit twee elementen, de spiegelingstransformatie en de identiteit. Door de spiegelingen aan de groep toe te voegen wordt $SO(3)$ dus uitgebreid tot $O(3)$, zie begin Hoofdstuk 3.

De generatoren van de draaiingsgroep veranderen niet onder een spiegeling, zoals direkt blijkt uit (4.29): onder spiegelingen transformeert $\vec{L}$ als een pseudo-vector. Dit volgt ook direkt uit het feit dat een spiegeling t.o.v. de oorsprong commuteert met rotaties, zoals eenvoudig is na te gaan. Hieruit volgt weer dat een dergelijke spiegeling werkend op irreducibele representaties van de draaiingsgroep, evenredig moet zijn met de identiteit (dit volgt weer uit het lemma van Schur (zie de diskussie volgend op (4.41)). Vandaar dat de spin van een toestand niet zal veranderen onder een spiegeling, en de toestand hoogstens een minteken kan oplopen.

Beschouwen we nu weer de verstrooiing van protonen op koolstofkernen, en het effect van de $O(3)$ matrix

$$P_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{7.25}$$

Na deze spiegeling hebben we te maken met dezelfde experimentele situatie, want de impuls van het inkomende proton is niet veranderd. Alleen de hoek θ is nu vervangen door $\pi - \theta$. Daar de hoekverdelingen $f_1(\theta)$ en $f_2(\theta)$ niet moeten veranderen onder deze substitutie moet gelden:

$$f_1(\theta) = f_1(\pi - \theta), \quad f_2(\theta) = f_2(\pi - \theta). \tag{7.26}$$

Dit impliceert dat $f_1 = f_2$.

Merk op dat we P_z kunnen schrijven als het product van de spiegeling $P = -\mathbf{1}$ en een draaiing over π radialen rond de z-as. De spiegeling P doet niets met de spin-toestanden, maar de draaiing korrespondeert met

$$\mathrm{e}^{\frac{1}{2} i\pi\tau_3} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i\tau_3. \tag{7.27}$$

Bijgevolg geldt (naar analogie met (6.16)) dat $\tau_3 \, F(\pi - \theta) \, \tau_3 = F(\theta)$, hetgeen nu ook impliceert dat $f_{12}^* = -f_{12}$.

8. Isospin

Zoals we in hoofdstuk 1 hebben uiteengezet, laten elementaire deeltjes van ongeveer dezelfde massa zich groeperen in z.g. *multipletten*. Bijvoorbeeld,

$$\begin{aligned} \text{nukleon} - \text{doublet}: &\quad (p, n), \\ \text{pion} - \text{triplet}: &\quad (\pi^+, \pi^0, \pi^-), \\ \Delta - \text{quadruplet}: &\quad (\Delta^{++}, \Delta^+, \Delta^0, \Delta^-), \\ \text{enz.} & \end{aligned} \tag{8.1}$$

Deze ordening is te verklaren door aan te nemen dat de natuur in goede benadering invariant is onder z.g. *isospinrotaties*. Isospinrotaties vormen weer een groep, waarvan de elementen zijn op te vatten als rotaties in een "interne" driedimensionale ruimte. De toestanden behorende bij de isospinmultipletten transformeren volgens (irreducibele) representaties van deze groep. Het begrip isospin is daarom analoog aan de gebruikelijke spin samenhangend met ruimtelijke draaiingen.

De krachten die de nukleonen binden in de atoomkernen, de z.g. sterke krachten, zijn invariant onder isospinrotaties. Echter, de electromagnetische krachten zijn dit niet. Vandaar dat de electrische lading binnen één multiplet niet hetzelfde is. Behoud van electrische lading is wel exact. Dit correspondeert met behoud van isospin-"impulsmoment" in de 3-richting, aangezien de elektrische lading gegeven wordt door de Gell-Mann-Nishijima relatie

$$Q = I_3 + \tfrac{1}{2}Y, \tag{8.2}$$

waarbij Y de "hyperlading" is die voor iedere toestand binnen één multiplet dezelfde waarde heeft. We gaan eenvoudig na dat $Y = 1$, 0 en 1 voor respectievelijk het nukleon-, het pion- en het Δ-multiplet.

Laten we nu het verval van een Δ^{++} bekijken,

$$\Delta^{++} \longrightarrow p\,\pi^+ \tag{8.3}$$

Voor de deeltjes voorkomend in dit verval geldt

$$I_3 \mid \Delta^{++}\rangle = \tfrac{3}{2} \mid \Delta^{++}\rangle, \qquad I_3 \mid p\rangle = {}^1\!/_2 \mid p\rangle, \qquad I_3 \mid \pi^+\rangle = \mid \pi^+\rangle, \tag{8.4}$$

zodat

$$I_3\Big(\mid p\rangle \mid \pi^+\rangle\Big) = \Big(I_3 \mid p\rangle\Big) \mid \pi^+\rangle + \mid p\rangle\Big(I_3 \mid \pi^+\rangle\Big) = \tfrac{3}{2} \mid p\rangle \mid \pi^+\rangle. \tag{8.5}$$

Bovenstaande formule volgt uit het feit dat de generatoren I_i voor een representatie bestaande uit producten van toestanden, gegeven worden door de som van de generatoren werkende op de afzonderlijke toestanden. Soortgelijke formules gelden dus ook voor I_+ en I_-.

Nu vallen Δ-deeltjes voornamelijk uiteen in een nukleon en een pion. Vanwege ladingsbehoud vervalt Δ^{++} daarom volledig volgens (8.3). T.a.v. de andere Δ-deeltjes is de situatie minder eenvoudig. Laten we beginnen met de zes verschillende nukleon-piontoestanden te geven, die allen een eigentoestand zijn van de I_3-operator,

$$
\begin{aligned}
I_3\Big(\,|\,p\rangle\,|\,\pi^+\rangle\Big) &= \tfrac{3}{2}\,|\,p\rangle\,|\,\pi^+\rangle, \\
I_3\Big(\,|\,p\rangle\,|\,\pi^0\rangle\Big) &= \tfrac{1}{2}\,|\,p\rangle\,|\,\pi^0\rangle, \\
I_3\Big(\,|\,p\rangle\,|\,\pi^-\rangle\Big) &= -\tfrac{1}{2}\,|\,p\rangle\,|\,\pi^-\rangle, \\
I_3\Big(\,|\,n\rangle\,|\,\pi^+\rangle\Big) &= \tfrac{1}{2}\,|\,n\rangle\,|\,\pi^+\rangle, \\
I_3\Big(\,|\,n\rangle\,|\,\pi^0\rangle\Big) &= -\tfrac{1}{2}\,|\,n\rangle\,|\,\pi^0\rangle, \\
I_3\Big(\,|\,n\rangle\,|\,\pi^-\rangle\Big) &= -\tfrac{3}{2}\,|\,n\rangle\,|\,\pi^-\rangle.
\end{aligned}
\tag{8.6}
$$

De Δ-deeltjes kunnen echter slechts in bepaalde lineaire combinaties van deze toestanden uiteenvallen. Dit heeft te maken met het feit dat de product-toestanden (8.6) uiteenvallen in twee verschillende representaties van de isospin-groep. Dit blijkt onmiddelijk uit de eigenwaarden van I_3: de eigenwaarden $\pm\frac{3}{2}$ komen slechts eenmaal voor terwijl de eigenwaarden $\pm\frac{1}{2}$ tweemaal voorkomen. De bijbehorende toestanden moeten daarom overeenkomen met de toestanden behorende bij een isospin-$\frac{3}{2}$ en een isospin-$\frac{1}{2}$ representatie. We kunnen daarom de volgende decompositie opschrijven van de product-representatie (analoog aan (4.13)),

$$
\mathbf{2}\otimes\mathbf{3}=\mathbf{2}\oplus\mathbf{4}, \tag{8.7}
$$

waarbij $\mathbf{2}$, $\mathbf{3}$ en $\mathbf{4}$ de 2-, 3- en 4-dimensionele representaties aangeven behorende bij isospin $\frac{1}{2}$, 1 en $\frac{3}{2}$.

Als deze vervalsprocessen invariant zijn onder isospindraaiingen, dan kunnen we het verval van de andere Δ-deeltjes relateren aan het verval van het Δ^{++}-deeltje. Onder de operator I_- verandert de Δ^{++} toestand volgens

$$
I_-\,|\,\Delta^{++}\rangle = I_-\,|\,\tfrac{3}{2},\tfrac{3}{2}\rangle = \sqrt{3}\,|\,\tfrac{3}{2},\tfrac{1}{2}\rangle = \sqrt{3}\,|\,\Delta^{+}\rangle, \tag{8.8}
$$

waarbij we gebruik gemaakt hebben van (5.26). Voor de nukleon-pion producttoestand waarin de Δ^{++} vervalt, vinden we

$$
\begin{aligned}
I_-\Big(\,|\,p^+\rangle\,|\,\pi^+\rangle\Big) &= \Big(I_-\,|\,p^+\rangle\Big)\,|\,\pi^+\rangle + |\,p^+\rangle\Big(I^-\,|\,\pi^+\rangle\Big) \\
&= |\,n\rangle\,|\,\pi^+\rangle + \sqrt{2}\,|\,p^+\rangle\,|\,\pi^0\rangle.
\end{aligned}
\tag{8.9}
$$

Uit isospin-invariantie volgt dat $|\,\Delta^+\rangle$, die verkregen kan worden m.b.v. een isospinrotatie werkend op $|\,\Delta^{++}\rangle$, vervalt in de toestand die verkregen wordt door dezelfde rotatie toe te passen op $|\,p^+\,|\,\pi^+\rangle$; deze toestand is de genormeerde toestand die verkregen is in (8.9),

$$
|\,\psi\rangle = \frac{1}{\sqrt{3}}\Big(\,|\,n\rangle\,|\,\pi^+\rangle + \sqrt{2}\,|\,p^+\rangle\,|\,\pi^0\rangle\Big). \tag{8.10}
$$

Vandaar dat de vervalswaarschijnlijkheden voor de twee mogelijke vervallen van Δ^+ voldoen aan

$$
\Gamma(\Delta^+\to n\,\pi^+) : \Gamma(\Delta^+\to p\,\pi^0) = 1:2\,, \tag{8.11}
$$

hetgeen geheel in overeenstemming is met de waarnemingen.

Om de vervalswijzen van Δ^0 te bestuderen laten we nogmaals I_- werken,

$$\begin{aligned} I_- \mid \Delta^+\rangle &= 2 \mid \Delta^0\rangle, \\ I_-\Big(\mid n\rangle \mid \pi^+\rangle\Big) &= \sqrt{2} \mid n\rangle \mid \pi^0\rangle, \\ I_-\Big(\mid p\rangle \mid \pi^0\rangle\Big) &= \mid n\rangle \mid \pi^0\rangle + \sqrt{2} \mid p\rangle \mid \pi^-\rangle, \end{aligned} \tag{8.12}$$

zodat

$$I_- \left(\frac{\mid n\rangle \mid \pi^+\rangle + \sqrt{2} \mid p\rangle \mid \pi^0\rangle}{\sqrt{3}} \right) = 2 \left(\frac{\sqrt{2} \mid n\rangle \mid \pi^0\rangle + \mid p\rangle \mid \pi^-\rangle}{\sqrt{3}} \right). \tag{8.13}$$

Dus geldt dat de vervalswaarschijnlijkheden voor de twee mogelijke vervalswijzen van Δ^0, zich verhouden volgens

$$\Gamma(\Delta^0 \to n\,\pi^0) : \Gamma(\Delta^0 \to p\,\pi^-) = 2 : 1\,. \tag{8.14}$$

Zo zien we aan een voorbeeld hoe vervalsverhoudingen van 2 : 1 bij elementaire deeltjes kunnen volgen uit de invariantie onder transformaties behorende tot een Lie-groep.

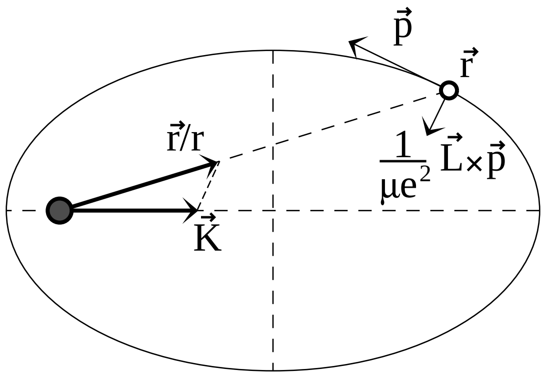

Figure 4: Constructie van de Runge-Lenz-vector $\vec{K}$.

9. Het waterstofatoom

In dit hoofdstuk geven we een groepen-theoretische afleiding van het energiespectrum van het waterstofatoom. De Hamiltoniaan van het waterstofatoom heeft de vorm

$$H = \frac{\vec{p}^{\,2}}{2\mu} - \frac{e^2}{r}, \tag{9.1}$$

waarbij de eerste term de kinetische energie voorstelt (met μ de gereduceerde massa van het elektron-proton systeem), en de tweede term de Coulomb-potentiaal. De klassieke bewegingsvergelijking behorende bij (9.1) is

$$\dot{\vec{p}} = -e^2 \frac{\vec{r}}{r^3}\,, \tag{9.2}$$

waarbij de impuls $\vec{p}$ gegeven wordt door

$$\vec{p} = \mu\,\dot{\vec{r}}. \tag{9.3}$$

Zoals bekend bestaan er voor het waterstofatoom een aantal behouden grootheden, d.w.z. grootheden die konstant blijven in de tijd. De meest bekende zijn de energie gegeven door (9.1) en de komponenten van de impulsmoment-vector,

$$\vec{L} = \vec{r} \times \vec{p}. \tag{9.4}$$

Inderdaad volgt direkt uit (9.2) en (9.3) dat

$$\frac{\mathrm{d}}{\mathrm{d}t}H = \frac{\vec{p}\cdot\dot{\vec{p}}}{\mu} + \frac{e^2}{r^3}\,\vec{r}\cdot\dot{\vec{r}} = 0, \tag{9.5}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{L} = \dot{\vec{r}} \times \vec{p} + \vec{r} \times \dot{\vec{p}} = 0. \tag{9.6}$$

Het laatste resultaat volgt uit het feit dat $\dot{\vec{r}}$ en $\vec{p}$ en $\vec{r}$ en $\dot{\vec{p}}$ evenwijdig zijn volgens (9.2) en (9.3). Er is echter ook nog een andere vector die konstant blijft. Dit is de z.g.

Runge-Lenz-vector, die gedefinieerd is door

$$\vec{K} = \frac{1}{\mu e^2}\,\vec{L}\times\vec{p} + \frac{\vec{r}}{r}\,. \tag{9.7}$$

Het behouden zijn van $\vec{K}$ volgt weer uit (9.2) en (9.3), maar het bewijs is wat minder eenvoudig. Eerst bewijzen we dat

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\vec{K} &= \frac{1}{\mu e^2}\,\vec{L}\times\dot{\vec{p}} + \frac{\dot{\vec{r}}}{r} - (\dot{\vec{r}}\cdot\vec{r})\,\frac{\vec{r}}{r^3} \\ &= -\frac{1}{\mu}\,\frac{\vec{L}\times\vec{r}}{r^3} + \frac{r^2\,\vec{p} - (\vec{p}\cdot\vec{r})\vec{r}}{\mu\, r^3},\end{aligned} \tag{9.8}$$

waarbij we in de eerste regel gebruik hebben gemaakt van (9.5) en in de tweede regel van (9.2) en (9.3). Vervolgens laten we zien dat (9.7) gelijk is aan nul door (9.4) te substitueren, zodat

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{K} = \frac{1}{\mu\, r^3}\left\{-(\vec{r}\times\vec{p})\times\vec{r} + r^2\,\vec{p} - (\vec{r}\cdot\vec{p})\,\vec{r}\right\}. \tag{9.9}$$

Gebruik makend van de relatie

$$(\vec{a}\times\vec{b})\times\vec{c} = (\vec{c}\cdot\vec{a})\,\vec{b} - (\vec{c}\cdot\vec{b})\,\vec{a}, \tag{9.10}$$

laten we vervolgens zien dat (9.9) gelijk is aan nul, zodat $\vec{K}$ inderdaad een konstante vector is.

De Runge-Lenz-vector heeft nog een aantal andere belangwekkende eigenschappen, die rechtstreeks volgen uit de definities en elementaire relaties uit de vector-analyse, namelijk

$$\vec{K}\cdot\vec{L} = 0, \tag{9.11a}$$

$$\begin{aligned}\vec{K}^2 &= \left(\frac{1}{\mu e^2}\right)^2 (\vec{L}\times\vec{p})^2 + \frac{2}{\mu\, e^2\, r}(\vec{L}\times\vec{p})\cdot\vec{r} + 1 \\ &= \left(\frac{1}{\mu e^2}\right)^2 \vec{L}^2\,\vec{p}^{\,2} - \frac{2}{\mu\, e^2\, r}\vec{L}^{\,2} + 1 \\ &= \frac{2H}{\mu e^4}\,\vec{L}^2 + 1.\end{aligned} \tag{9.11b}$$

Voor de ellipsbaan van een klassiek deeltje geeft de Runge-Lenz vector $\vec{K}$ in feite de richting aan van de *lange as van de ellips*, zie de illiustratie in Figuur 4. De lengte van $\vec{K}$ is de *excentriciteit* ε van de ellips.

In de quantummechanica zijn $\vec{r}$ en $\vec{p}$ operatoren, die voldoen aan de Heisenberg commutatie-relaties,

$$\begin{aligned}[p_i, r_j] &= -i\hbar\,\delta_{ij}, \\ [r_i, r_j] &= [p_i, p_j] = 0.\end{aligned} \tag{9.12}$$

In de coördinaten-representatie beschrijven we toestanden in termen van golffunkties van de coördinaten r_i. De coördinaat-operator werkt dan triviaal op de golffunktie en vermenigvuldigt deze met de waarde van de bijbehorende coördinaat, terwijl de impulsoperatoren worden gegeven door

$$\vec{p} = \frac{\hbar}{i}\frac{\partial}{\partial\vec{r}}\,. \tag{9.13}$$

Vanzelfsprekend zijn H, $\vec{L}$ en $\vec{K}$ nu ook operatoren, maar in de definitie van $\vec{K}$ doet zich een probleem voor, aangezien $\vec{p}$ en $\vec{L}$ niet commuteren,

$$\begin{aligned} [L_i, p_j] &= i\hbar\, \epsilon_{ijk}\, p_k, \\ [L_i, r_j] &= i\hbar\, \epsilon_{ijk}\, r_k. \end{aligned} \tag{9.14}$$

Dit resultaat volgt rechtstreeks uit (9.4) en (9.12) met behulp van de commutatorrelatie

$$[A, BC] = [A, B]\, C + B\, [A, C]. \tag{9.15}$$

Gezien (9.14) is het niet duidelijk in welke volgorde we de operatoren $\vec{L}$ en $\vec{p}$ moeten schrijven in de definitie (9.7) van $\vec{K}$.[18] We maken nu de volgende keuze voor de quantummechanische definitie van $\vec{K}$:

$$\vec{K} = \frac{1}{2\mu e^2}\Big(\vec{L}\times\vec{p} - \vec{p}\times\vec{L}\Big) + \frac{\vec{r}}{r}\,. \tag{9.16}$$

Het is duidelijk dat wanneer $\vec{L}$ en $\vec{p}$ commuteren, deze definitie overeenkomt met (9.7). Bovendien is de operator $\vec{K}$ hermitisch in deze vorm.

In de quantummechanica geldt voor behouden grootheden dat ze commuteren met de Hamiltoniaan. Laten we dit eerst eens nagaan voor de operatoren $\vec{L}$. We hebben al eerder de commutatierelatie bepaald tussen $\vec{L}$ en $\vec{p}$. Op dezelfde wijze stellen we vast dat

$$[L_i, r_j] = i\hbar\, \epsilon_{ijk}\, r_k. \tag{9.17}$$

Met behulp van (9.15) kunnen we dan eenvoudig afleiden dat

$$[L_i, H] = 0, \tag{9.18}$$

$$[L_i, L_j] = i\hbar\, \epsilon_{ijk}\, L_k. \tag{9.19}$$

De vergelijking (9.18) geeft aan dat L_i een behouden grootheid is. De vergelijking (9.19) laat zien dat de impulsmomentoperatoren dezelfde commutatierelaties hebben als de generatoren van de drie-dimensionale draaiingsgroep.

Ter illustratie geven we de expliciete afleiding van (9.19):

$$\begin{aligned} [L_i, L_j] &= \epsilon_{jkl}\, [L_i, r_k\, p_l] \\ &= \epsilon_{jkl}\, ([L_i, r_k]\, p_l + r_k\, [L_i, p_l]) \\ &= i\hbar\, \epsilon_{jkl}\, (\epsilon_{ikm}\, r_m\, p_l + \epsilon_{ilm}\, r_k\, p_m) \\ &= i\hbar\, (\epsilon_{jlk}\, \epsilon_{imk} + \epsilon_{jmk}\, \epsilon_{ikl})\, r_m\, p_l \\ &= i\hbar\, \epsilon_{ijk}\, \epsilon_{kml}\, r_m\, p_l\ , \end{aligned} \tag{9.20}$$

[18] De lezer kan eenvoudig nagaan dat dit probleem zich niet voordoet in de definities (9.1) en (9.4) van H en $\vec{L}$.

waarbij we achtereenvolgens gebruik gemaakt hebben van (9.4), (9.15), (9.14) en (9.17); vervolgens hebben we indices hernoemd en gebruik gemaakt van de identiteit (3.33) voor ϵ-tensoren.

Op soortgelijke wijze kunnen nu ook commutatoren bepalen waarin de operatoren $\vec{K}$ optreden. Eerst stellen we vast, door gebruik te maken van (9.13), dat geldt

$$\left[\vec{p}, \frac{1}{r}\right] = \frac{\hbar}{i} \left(\frac{\partial}{\partial \vec{r}} \frac{1}{r} - \frac{1}{r} \frac{\partial}{\partial \vec{r}} \right) = i\hbar \, \frac{\vec{r}}{r^3}. \tag{9.21}$$

Vervolgens leiden we af dat

$$[L_i, K_j] = i\hbar \, \epsilon_{ijk} \, K_k \tag{9.22}$$

(een eigenschap die het gevolg is van het feit dat de K_i een vector vormen),

$$[K_i, p_j] = \frac{i\hbar}{\mu e^2} \, (p_i p_j - p^2 \, \delta_{ij}) + \frac{i\hbar}{r^3} \, (r^2 \delta_{ij} - r_i r_j), \tag{9.23}$$

$$\left[K_i, \frac{1}{r}\right] = \frac{i\hbar}{2\mu e^2} \, \epsilon_{ijk} \, \frac{L_j r_k + r_k L_j}{r^3}. \tag{9.24}$$

Met behulp van (9.23) en (9.24) is het nu niet moeilijk om te bewijzen dat $\vec{K}$ commuteert met de Hamiltoniaan,

$$[\vec{K}, H] = 0. \tag{9.25}$$

Tenslotte kunnen we uit de Jacobi-identiteit

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \tag{9.26}$$

konkluderen dat ook $[K_i, K_j]$ met de Hamiltoniaan moet commuteren, en dus een behouden grootheid is. Uit berekening volgt dan dat dat de commutator niet leidt tot een nieuwe behouden grootheid, maar tot een product van twee al bekende behouden grootheden,[19]

$$[K_i, K_j] = i\hbar \; \epsilon_{ijk} \; L_k \left(\frac{-2H}{\mu e^4} \right). \tag{9.27}$$

Aangezien L_k en H commuteren, is de volgorde van deze operatoren niet van belang.

We hebben nu alle relevante commutatoren bepaald. Vervolgens moeten we nagaan of de klassieke relaties (9.11) geldig blijven in het geval dat $\vec{K}$ en $\vec{L}$ operatoren zijn. Voor de eerste vergelijking vinden we,

$$\vec{K} \cdot \vec{L} = \vec{L} \cdot \vec{K} = \frac{1}{2\mu e^2} \left\{ \vec{L} \cdot (\vec{L} \times \vec{p}) - \vec{L} \cdot (p \times \vec{L}) \right\} + \frac{\vec{L} \cdot \vec{r}}{r}. \tag{9.28}$$

Aangezien $\vec{L} \cdot \vec{r} = 0$ (dit geldt ook in de quantum-mechanica zoals men gemakkelijk nagaat), dragen alleen de eerste twee termen bij, evenredig met

$$\epsilon_{ijk} \, L_i \, (L_j p_k - p_j L_k) = \epsilon_{ijk} \, (2 \, L_i \, L_j \, p_k - i\hbar \epsilon_{jkl} \, L_i \, p_l) \, , \tag{9.29}$$

[19] Op algemene gronden kan men bewijzen dat het waterstofatoom niet meer dan 5 *onafhankelijke* behouden grootheden kent. Hiervoor kunnen we bijvoorbeeld de componenten van de impulsmoment-vector en de twee coördinaten van de Runge-Lenz-vector, die loodrecht staat op $\vec{L}$, kiezen.

waar we gebruik hebben gemaakt van (9.14) om de p_i rechts van de operatoren L_i te schrijven. Met behulp van (9.19) schrijven we de eerste term evenredig met $\epsilon_{ijk}\,\epsilon_{ijl}\,L_l\,p_k \propto \vec{L}\cdot\vec{p} = 0$; de tweede term verdwijnt om dezelfde reden, zodat we hetzelfde resultaat vinden als voorheen,

$$\vec{K}\cdot\vec{L} = 0. \tag{9.30}$$

De vergelijking (9.11b) is in het quantummechanische geval niet meer geheel juist, en het verwisselen van operatoren geeft aanleiding tot extra termen. Het juiste resultaat, dat enige berekening vergt, neemt de volgende vorm aan,

$$\vec{K}^2 = \frac{2H}{\mu e^4}(\vec{L}^{\,2} + \hbar^2) + 1. \tag{9.31}$$

In de limiet $\hbar \to 0$ komt dit resultaat overeen met (9.11b).

Met behulp van de bovenstaande resultaten bepalen we nu het energiespectrum van het waterstofatoom. We beschouwen eigenfuncties van H, die voldoen aan

$$H \mid \psi\rangle = E \mid \psi\rangle, \tag{9.32}$$

waarbij we zullen aannemen dat de energie E negatief is, $E < 0$, zodat we te maken hebben met gebonden toestanden. Omdat de operatoren L_i en K_i commuteren met de Hamiltoniaan, zijn de toestanden $L_i \mid \psi\rangle$ en $K_i \mid \psi\rangle$ ook oplossingen van de Schrödingervergelijking (9.32). Voor wat betreft de operatoren H, $\vec{L}$ en $\vec{K}$ kunnen we ons daarom beperken tot de deelruimte van toestanden $\mid \psi\rangle$ met dezelfde energie E.

Definieer nu de volgende twee lineaire combinaties van de operatoren $\vec{L}$ en $\vec{K}$,[20]

$$\vec{L}^{\pm} = \frac{1}{2}\left(\vec{L} \pm \sqrt{\frac{\mu\, e^4}{-2E}}\,\vec{K}\right). \tag{9.33}$$

Deze operatoren commuteren met de Hamiltoniaan,

$$[L_i^{\pm}, H] = 0. \tag{9.34}$$

Met behulp van (9.19), (9.22) en (9.27) bewijzen we de volgende commutatierelaties

$$\left[L_i^+, L_j^+\right] = i\hbar\,\epsilon_{ijk}\,L_k^+, \tag{9.35}$$

$$\left[L_i^-, L_j^-\right] = i\hbar\,\epsilon_{ijk}\,L_k^-, \tag{9.36}$$

$$\left[L_i^+, L_j^-\right] = 0. \tag{9.37}$$

Gezien deze commutatierelaties kunnen we $\vec{L}^+$ en $\vec{L}^-$ opvatten als de generatoren van twee afzonderlijke draaiingsgroepen, die onderling commuteren en op de toestanden werken volgens

$$\mid \psi\rangle \to \mid \psi'\rangle = \mathrm{e}^{i\vec{\alpha}\cdot\vec{L}^+}\,\mathrm{e}^{i\vec{\beta}\cdot\vec{L}^-} \mid \psi\rangle, \tag{9.38}$$

[20]Het feit dat E negatief is, maakt dat het argument van de wortel in deze definities positief is.

waarbij de parameters $\vec{\alpha}$ en $\vec{\beta}$ de twee groepen beschrijven gegenereerd door respectievelijk $\vec{L}^+$ en $\vec{L}^-$.[21]

Als $\mid \psi\rangle$ voldoet aan de Schrödingervergelijking (9.32), dan voldoen ook de getransformeerde toestanden $\mid \psi'\rangle$ aan deze vergelijking met dezelfde eigenwaarde E. Al deze toestanden moeten gekarakteriseerd kunnen worden als representaties van de twee draaiingsgroepen (deze productgroep wordt wel geschreven als $SU(2)\otimes SU(2)$). De toestanden moeten dus transformeren volgens een $(2s_+ + 1)$-dimensionale (irreducibele) representatie van de eerste draaiingsgroep. Wanneer we met een element van de tweede groep werken op deze representatie, dan vinden we een identieke representatie van de eerste groep, die eventueel gelijk kan zijn aan de oorspronkelijke representatie (daarmee bedoelen we dat ieder van de toestanden van de oorspronkelijke representatie gelijk blijft tot op een uniforme faktor na. Dit laatste volgt uit het lemma van Schur). Onder de tweede groep worden de representaties van de eerste groep in zijn geheel in elkaar overgevoerd. Vanzelfsprekend geldt ook het omgekeerde: onder de eerste groep gaan representaties van de tweede groep in hun geheel in elkaar over. We hebben daarom te maken met een tensorproduct van een representatie van de eerste groep met een representatie van de tweede groep. Kiezen we een $(2s_+ + 1)$-dimensionale representatie van de eerste, en een $(2s_- + 1)$-dimensionale representatie van de tweede groep, dan verkrijgen we op deze manier $(2s_+ + 1)(2s_- + 1)$ toestanden die we aangeven met $\mid s_+, s_-\rangle$, en die voldoen aan de eigenwaarde-vergelijkingen

$$
\begin{aligned}
\left(\vec{L}^+\right)^2 \mid s_+, s_-\rangle &= s_+(s_+ + 1)\,\hbar^2 \mid s_+, s_-\rangle, \\
\left(\vec{L}^-\right)^2 \mid s_+, s_-\rangle &= s_-(s_- + 1)\,\hbar^2 \mid s_+, s_-\rangle,
\end{aligned} \tag{9.39}
$$

waarbij $s_\pm = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots$. Echter, we beschikken over een extra conditie voor de generatoren $\vec{L}^\pm$, die volgt uit (9.30), namelijk

$$
\left(\vec{L}^+ + \vec{L}^-\right)\cdot\left(\vec{L}^+ - \vec{L}^-\right) \propto \vec{L}\cdot\vec{K} = 0, \tag{9.40}
$$

ofwel

$$
\left(\vec{L}^+\right)^2 = \left(\vec{L}^-\right)^2. \tag{9.41}
$$

Hieruit concluderen we dat moet gelden

$$
s_+ = s_- = {}^1\!/_2(n-1), \tag{9.42}
$$

met $n = 1, 2, \ldots$ zodat we te maken hebben met n^2 toestanden $\mid s_+, s_-\rangle$.

Tenslotte kunnen we nu de Hamiloniaan, en dus de energie E, uitdrukken in termen van de operatoren $\vec{L}^\pm$ middels (9.31). We vinden dan

$$
-\left(\vec{L}^+ - \vec{L}^-\right)^2 = \frac{\mu\, e^4}{2E}\,\vec{K}^2 = \left(\vec{L}^+ + \vec{L}^-\right)^2 + \hbar^2 + \frac{\mu\, e^4}{2E}, \tag{9.43}
$$

[21] We wijzen er nogmaals op dat deze twee groepen onderling commuteren. De twee ge-exponentieerde operatoren in (9.38) kunnen dus verwisseld worden.

ofwel

$$-2\left(\vec{L}^{+}\right)^{2}-2\left(\vec{L}^{-}\right)^{2}=-2(n-1)(1/2(n-1)+1)\,\hbar^{2}=\hbar^{2}+\frac{\mu\,e^{4}}{2E}. \tag{9.44}$$

Dit impliceert dat de gebonden toestanden n^2-voudig gedegenereerd zijn met energie

$$E_n=-\frac{\mu\,e^{4}}{2\,\hbar^{2}\,n^{2}}. \tag{9.45}$$

De impulsmoment-operatoren worden gegeven door de operatoren $\vec{L}^{+}+\vec{L}^{-}$. Deze lineaire combinatie werkt op identieke wijze op de beide representaties in het tensorproduct dat de toestanden $|s_+,s_-\rangle$ definieert. Onder ruimtelijke rotaties, gegenereerd door de operatoren $\vec{L}^{+}+\vec{L}^{-}$, transformeren de toestanden $|s_+,s_-\rangle$ daarom als een tensorproduct van twee representaties van de gebruikelijke draaiingsgroep. Dergelijke product-representaties zijn niet langer irreducibel, zoals we al eerder hebben gezien. Uit de product-eigenschappen van representaties kunnen we bepalen welke waarden van het impulsmoment voorkomen bij gegeven n; dit volgt uit de representaties die voorkomen in het product van twee n-dimensionale representaties,

$$\mathbf{n}\otimes\mathbf{n}=\mathbf{2n{-}1}\oplus\mathbf{2n{-}3}\oplus\cdots\oplus\mathbf{3}\oplus\mathbf{1}, \tag{9.46}$$

waarbij $\mathbf{n}$ de n-dimensionale representatie aangeeft van de draaiingsgroep (die dus een toestand beschrijft met spin-$1/2(n-1)$). We zullen de bovenstaande decompositie niet bewijzen; voor $n=3$ hebben we dit al gedaan in (4.13), en een soortgelijk product hebben we bepaald in (8.7).

Samenvattend kunnen we dus zeggen dat de gebonden toestanden van het waterstofatoom bepaald worden door het *hoofdquantumgetal* n, hetgeen gelijk kan zijn aan $n=1,2,\ldots$. Voor gegeven n hebben we n^2 toestanden met impulsmomenten $l=0,1,\ldots,n-1$, en energie E_n gegeven door (9.45).

Referenties : O. Klein, Z. für Physik **22** (1924) 109; W. Lenz, Z. für Physik **24** (1924) 197; W. Pauli, Z. für Physik **36** (1926) 336.

isospin-multiplet	spin	pariteit	isospin	S	massa (MeV)
pionen, $\pi^{\pm}$, π^0	0	$-$	1	0	139.6, 135
kaonen, K^0, K^+	0	$-$	$1/2$	1	497.7, 493.7
anti-kaonen, K^-, $\overline{K^0}$	0	$-$	$1/2$	-1	493.7, 497.7
eta, η^0	0	$-$	0	0	547.3
eta$'$, η'	0	$-$	0	0	$\approx$ 960
nucleonen, n, p	$1/2$	$+$	$1/2$	0	939.5, 938.3
delta, Δ^-, Δ^0, Δ^+, Δ^{++}	$3/2$	$+$	$3/2$	0	$\approx$ 1235
sigma, Σ^-, Σ^0, Σ^+	$1/2$	$+$	1	-1	1197.4, 1192.6, 1189.4
sigma*, Σ^{*-}, Σ^{*0}, Σ^{*+}	$3/2$	$+$	1	-1	$\approx$ 1385
lambda, Λ^0	$1/2$	$+$	0	-1	1115.7
ksi, Ξ^-, Ξ^0	$1/2$	$+$	$1/2$	-2	1321.3, 1315
ksi*, Ξ^{*-}, Ξ^{*0}	$1/2$	$+$	$1/2$	-2	$\approx$ 1530
omega, Ω^-	$3/2$	$+$	0	-3	1672.4

Table 2: Diverse isospin-multipletten en hun massa's.

10. De groep $SU(3)$

De Lie-groep $SU(3)$ heeft een belangrijke rol gespeeld bij de totstandkoming van de quarktheorie. Bij de bestudering van de *hadronen* (de sterk wisselwerkende subatomaire deeltjes) was het opgevallen dat deze lijken te passen in een schema van *multipletten* dat ingewikkelder is dan de isospin-multipletten die we al gezien hebben. Verschillende isospin multipletten leken binnen grotere kaders te passen.

Nadat men zich realiseerde dat hiervoor wellicht weer één of meerdere Lie-groepen verantwoordelijk zouden kunnen zijn, ging men diverse mogelijkheden proberen. Hieruit kwam $SU(3)$ naar voren. Thans weten we waarom juist deze groep van toepassing is. Diep verborgen in deze subatomaire deeltjes zitten hun bouwstenen, de *quarks*. Drie soorten daarvan treden op. We noemen ze u, d en s. Hiervan vormen u en d een isospin-doublet. Zij zijn dus een $I = 1/2$ representatie van $SU(2)$. Het derde quark, s, is een iso-singlet. We schrijven de golffuncties van de quarks als

$$|\,\psi\,\rangle = \begin{pmatrix} u \\ d \\ s \end{pmatrix} \ , \tag{10.47}$$

en we gaan ervan uit dat de fysische eigenschappen van de deeltjes (bij benadering) invariant zijn onder de *unitaire* transformaties van deze 3-vectoren. Daarom vatten we (10.47) op als de *fundamentele representatie* van $SU(3)$, waarin 3 staat voor de dimensie van de vectoren, U staat voor "unitair", en S voor "speciaal": we beperken ons tot matrices met determinant 1.

Beschouwen we de transformatyies die alleen u en d in elkaar overvoeren maar s invariant laten, dan hebben we de subgroep $SU(2)$ van $SU(3)$. Bij $SU(2)$ hebben we gezien dat als je representaties probeert te vormen uitgaande van representaties met heeltallige

ℓ, je altijd representaties krijgt waarin ℓ weer heeltallig is. Dit zijn alle representaties die het element

$$-\mathbb{I} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \ , \tag{10.48}$$

op de eenheidsmatrix $+\mathbb{I}$ afbeelden.

Iets soortgelijks hebben we bij $SU(3)$. Deze groep heeft drie speciale elementen die met alle andere elementen van de groep commuteren (we noemen deze subgroep het *centrum* van de groep):

$$\mathbb{I} \ ; \qquad e^{\frac{2\pi i}{3}}\mathbb{I} = \begin{pmatrix} e^{\frac{2\pi i}{3}} & & \\ & e^{\frac{2\pi i}{3}} & \\ & & e^{\frac{2\pi i}{3}} \end{pmatrix} \ ; \qquad e^{\frac{-2\pi i}{3}}\mathbb{I} = \begin{pmatrix} e^{\frac{-2\pi i}{3}} & & \\ & e^{\frac{-2\pi i}{3}} & \\ & & e^{\frac{-2\pi i}{3}} \end{pmatrix} \ . \tag{10.49}$$

We onderscheiden nu twee soorten representaties van $SU(3)$: de *reguliere* en de *exotische* representaties. Reguliere representaties zijn die welke de drie elementen (10.49) van het centrum alle op de eenheid afbeelden. Exotische representaties doen dit niet. De elementaire representatie zelf is exotisch. De representatie die gevormd wordt door de 3×3 matrices van de $SU(3)$ transformaties U, de *geadjungeerde representatie*, is wel regulier, want deze transformeren volgens:

$$U \to U' = X\,U\,X^{-1} \ , \tag{10.50}$$

en als we voor X één van de matrices (10.49) nemen, dan zien we dat U niet verandert.

Schrijven we $U = e^{iT}$, dan volgt unitariteit van U als we eisen dat T hermitisch moet zijn: $T = T^\dagger$, en het feit dat $\det(U) = 1$ correspondeert met de eis $\mathrm{Tr}\,(T) = 0$. De 3×3-matrices hebben 9 complexe elementen. Hermiticiteit van T beperkt dit aantal tot 9 *reële* parameters. Uit $\mathrm{Tr}\,(U) = 0$ volgt dat hiervan 8 onafhankelijk zijn. De geadjungeerde representatie spant dus een 8-dimensionale reële ruimte op. Net als in hoofdstuk 6, vgl. (6.41), schrijven we deze represenatie als $Y^{\alpha\beta}$, maar nu kunnen de indices α, β drie waarden aannemen. Nemen we de elementen waarbij α en β beide 1 of 2 zijn, dan hebben we de driedimensionale $I = 1$ representatie **3** van $SU(2)$. Als $\alpha = 1,2$ en $\beta = 3$ dan hebben we de complexe, fundamentele **2** representatie van $SU(2)$, en we vinden die weer terug als $\alpha = 3$ en $\beta = 1,2$, en voorts is er op de diagonaal nog een **1**. Omdat de **2** representatie complex is tellen we die dubbel in:

$$\mathbf{8}_{\,\mathrm{SU(3)}} = (\mathbf{3} + \mathbf{2} + \mathbf{2} + \mathbf{1})_{\,\mathrm{SU(2)}} \tag{10.51}$$

De andere representaties van $SU(3)$ vinden we door symmetrische of antisymmetrische combinaties van de fundamentele represenatie te vermenigvuldigen. We krijgen zo "tensoren" $Y^{\alpha_1\alpha_2\cdots}_{\beta_1\beta_2\cdots}$, en hieruit destilleren we de irreducibele combinaties, enerzijds door symmetriseren of antisymmetriseren, anderzijds door te vermenigvuldigen met *invariante tensoren.* Hiervan zijn er, als we alleen fundamentele indices gebruiken (zie later) maar drie: δ^α_β, $\varepsilon_{\alpha\beta\gamma}$ en $\varepsilon^{\alpha\beta\gamma}$. De laatste twee zijn invariant, net als de $\varepsilon_{\alpha\beta}$ en $\varepsilon^{\alpha\beta}$ in $SU(2)$, vanwege de eis $\det(X) = 1$ voor de transformatiematrices X in $SU(3)$. We kunnen deze $\varepsilon_{\alpha\beta\gamma}$

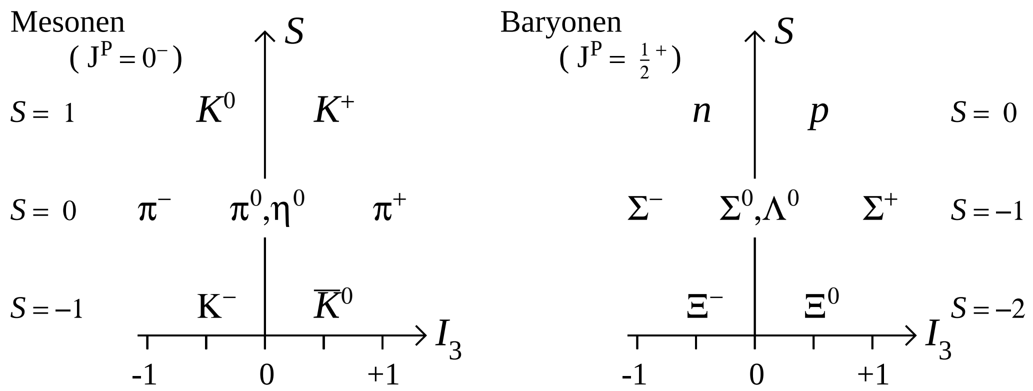

Figure 5: Het octet van mesonen met spin 0 en pariteit 1, en het baryon-octet met spin $^1\!/_2$ en pariteit +

tensoren gebruiken om bij iedere representatie alle indices bovenin (of juist onderin) te krijgen, dus alleen daar hoeven we naar te kijken.

De representaties gegenereerd door ϕ^α en ϕ_α zijn niet equivalent, zoals in $SU(2)$. We noemen deze twee fundamentele representaties $\mathbf{3}$ en $\bar{\mathbf{3}}$. De $\bar{\mathbf{3}}$ kunnen we dus ook schrijven als een antisymmtrische $Y^{\alpha\beta}$. De symmetrische representatie $Y^{\alpha\beta}$ is een complexe $\mathbf{6}$ representatie. De eerste niet-exotische representatie, na de $\mathbf{8}$, die we tegenkomen is een $\mathbf{10}$, gevormd door de volledig symmetrische tensoren $Y^{\alpha\beta\gamma}$. Dit *decuplet* is opgebouwd uit $SU(2)$ representaties als volgt:

$$\mathbf{10}_{\,\mathrm{SU(3)}} = (\mathbf{4}+\mathbf{3}+\mathbf{2}+\mathbf{1})_{\,\mathrm{SU(2)}}\ . \tag{10.52}$$

Alle niet-exotische representaties zijn gemakkelijk te vinden: het aantal boven-indices minus het aantal onder-indices moet dan een veelvoud zijn van 3.

In de hadronfysica kan men deeltjes van een bepaald type rangschikken naar isospin I_3 en een nieuw quantumgetal S genaamd “vreemdheid” (strangeness). De patronen die men verkreeg zijn geschetst in Figuren 5 en 6.

We herkennen er twee octetten en een decuplet in. Schrijven we het meson-octet als $M^\alpha_{\ \beta}$, dan zien we dat de Vreemdheid S correspondeert met het aantal onderindices die de waarde 3 aannemen minus het aantal bovenindices die de waarde 3 aannemen. Thans weten we dat we hier te maken hebben met golffuncties van het type $|\psi^\alpha\cdot\psi_\beta\rangle$ waarbij $|\psi^\alpha\rangle$ de quarks van het type α aangeven en $|\psi_\beta\rangle$ antiquarks van het type β. Hier corresponeren $\alpha = 1$ en 2 met het isodoublet u en d, en $\alpha = 3$ met het s-quark, dat nu vreemdheid -1 heeft gekregen. Een negende meson, met $M^\alpha_{\ \beta} = \delta^\alpha_\beta$ bestaat ook. Dit is het η', waarvan de massa wat verder afligt van die van de andere. Wat heel duidelijk opvalt is dat alle representaties van $SU(3)$ die we als hadronen tegenkomen in de natuur reguliere representaties zijn. De exotische representatie lijkt verboden te zijn.

De baryonen schrijven we alle als $B^{\alpha\beta\gamma}$. Het octet, dat te schrijven is als

$$B^{\alpha\beta\gamma} = \varepsilon^{\alpha\beta\kappa} B^\gamma_{\ \kappa}\ ; \qquad B^\alpha_{\ \alpha} = 0\ , \tag{10.53}$$

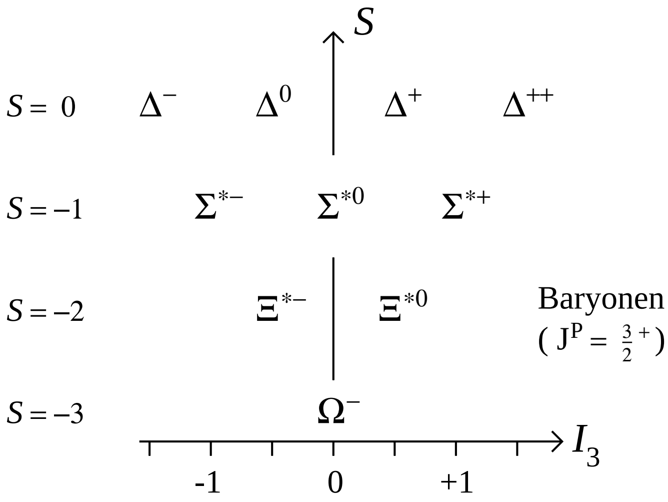

Figure 6: Het decuplet van de baryonen met spin ${}^3\!/_2$ en pariteit $P = +$.

kan ook gedefiniëerd worden door:

$$B^{\alpha\beta\gamma} = -B^{\beta\alpha\gamma} \ ; \qquad \varepsilon_{\alpha\beta\gamma} B^{\alpha\beta\gamma} = 0 \ . \tag{10.54}$$

$SU(3)$ symmetrie is zwaarder gebroken dan isospin, want de massa's van de deeltjes in verschillende isospin-multipletten liggen veel verder uit elkaar dan in één isospin-multiplet (zie Tabel 2). Toch is deze symmetrie zeer belangrijk. M. Gell-Mann ontdekte het decuplet toen het nog niet compleet was, en kon toen het nog ontbrekende deeltje Ω^- voorspellen.

We kunnen het baryon-decuplet en het baryon-octet met elkaar in verband brengen. Beschouw namelijk de basis van 'quark-toestanden' waarbij we nu ook hun spin in rekening nemen. Ieder quark mag $S_z = +1$ of -1 hebben. Als we de zes quark-toestanden die je dan krijgt in elkaar zouden mogen transformeren, hebben we de Lie-groep $SU(6)$. De symmetrische representatie $|\psi^{ABC}\rangle$ van $SU(6)$ heeft $6{\cdot}7{\cdot}8/3! = 56$ elementen. De indices van $SU(6)$ zijn ieder dus opgebouwd uit een $SU(3)$ index α en een $SU(2)$ index voor de spin, die we σ zullen noemen:

$$A = (\alpha, \sigma_1) \ , \qquad B = (\beta, \sigma_2) \ , \qquad C = (\gamma, \sigma_3) \ . \tag{10.55}$$

Nemen we alle elementen die symmetrisch zijn in α, β, en γ, dan zien we gemakkelijk in dat die dan ook symmetrisch moeten zijn bij verwisseling van de gewone spin indices σ_1, σ_2 en σ_3. Dus die hebben spin ${}^3\!/_2$, precies als ons baryonische decuplet. Ieder deeltje in het decuplet heeft 4 spin toestanden, dus hier zijn $10 \cdot 4 = 40$ basis-elementen. Er blijven $56 - 40 = 16$ elementen over, precies het octet met spin ${}^1\!/_2$, omdat ieder lid van het octet 2 spin-toestanden heeft.

$SU(6)$ is nog minder goed behouden dan $SU(3)$, getuige het feit dat de massa's van het baryon-decuplet aanzienlijk groter zijn dan die in het baryon-octet.

Een vraag die hierbij nog niet is beantwoord, is hoe het mogelijk is, dat quarks met spin ${}^1\!/_2$ in een *symmetrische* golffunctie kunnen zitten; als ze spin ${}^1\!/_2$ hebben dan moeten het immers fermionen zijn, en daarvoor zou je juist een volledig *anti*symmetrische golffunctie verwachten, vanwege Pauli's uitsluitingsprincipe. Het antwoord hierop bleek te zijn dat quarks nog een intern quantumgetal hebben, *kleur* (color), dat ook weer drie waarden kan aannemen. De kleur-theorie verklaart de krachten die er heersen tussen de quarks, en waarom deze juist op zodanige wijze bij elkaar blijven zitten dat ze alleen niet-exotische representaties van $SU(3)$ vormen.

Om de generatoren van een $SU(3)$ transformatie X te beschrijven gebruiken we de hermitische matrices λ_a, $\alpha = 1, \cdots, 8$:

$$X = \exp(\tfrac{1}{2} i\theta^a \lambda_a) \ . \tag{10.56}$$

We hebben hier een generalisatie van de $SU(2)$ matrices τ_i:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & \cdot \\ 1 & 0 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 0 & -i & \cdot \\ i & 0 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & \cdot \\ 0 & -1 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$$
$$\lambda_4 = \begin{pmatrix} 0 & \cdot & 1 \\ \cdot & \cdot & \cdot \\ 1 & \cdot & 0 \end{pmatrix} \qquad \lambda_5 = \begin{pmatrix} 0 & \cdot & -i \\ \cdot & \cdot & \cdot \\ i & \cdot & 0 \end{pmatrix} \qquad \lambda_6 = \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & 0 & 1 \\ \cdot & 1 & 0 \end{pmatrix}$$
$$\lambda_7 = \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & 0 & -i \\ \cdot & i & 0 \end{pmatrix} \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{10.57}$$

De λ_a zijn genormeerd volgens

$$\mathrm{Tr}\,(\lambda_a \lambda_b) = 2\delta_{ab} \ , \tag{10.58}$$

en ze voldoen aan de commutatierelaties

$$[\tfrac{1}{2}\lambda_a,\, \tfrac{1}{2}\lambda_b] = i f_{abc} (\tfrac{1}{2}\lambda_c) \ . \tag{10.59}$$

De lijst van structuurconstanten f_{abc} is

$$\begin{aligned} f_{123} &= 1 \\ f_{147} = f_{246} = f_{257} = f_{345} = f_{516} = f_{637} &= \tfrac{1}{2} \\ f_{458} = f_{678} &= \tfrac{1}{2}\sqrt{3} \ ; \end{aligned} \tag{10.60}$$

de overige volgen uit de volledige antisymmetrie van deze coëfficiënten:

$$f_{abc} = f_{bca} = -f_{bac} \ , \qquad \text{etc.} \tag{10.61}$$

f_{abc} en δ_{ab} zijn niet de enige invariante tensoren van de **8**-representatie. Beschouwen we namelijk de *anti*-commutatoren van de λ's, dan vinden we

$$\{\tfrac{1}{2}\lambda_a,\, \tfrac{1}{2}\lambda_b\} = \tfrac{1}{3}\delta_{ab} + d_{abc}(\tfrac{1}{2}\lambda_c) \ . \tag{10.62}$$

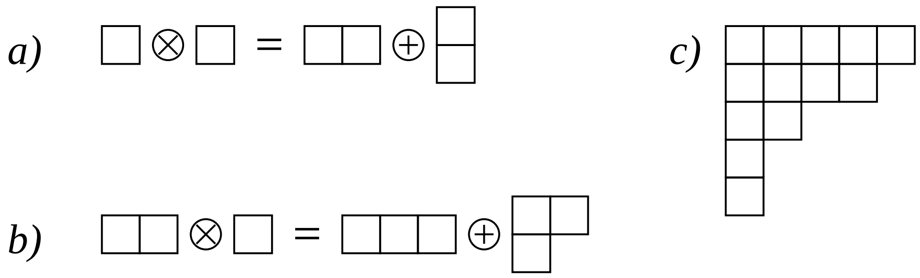

Figure 7: Young tableaux

De anticommutator is gedefiniëerd als $\{A,\, B\} = A\,B + B\,A$. De d_{abc} zijn helemaal symmetrisch:

$$d_{abc} = d_{bca} = d_{bac}\ , \qquad \text{etc.} \tag{10.63}$$

en je kunt er ook een lijst van maken:

$$\begin{aligned} d_{118} = d_{228} = d_{338} = -d_{888} &= \tfrac{1}{\sqrt{3}} \\ d_{146} = d_{157} = d_{256} = d_{344} = d_{355} &= 1/2 \\ d_{247} = d_{366} = d_{377} &= -1/2 \\ d_{448} = d_{558} = d_{668} = d_{778} &= -\tfrac{1}{2\sqrt{3}} \end{aligned} \tag{10.64}$$

De tensoren δ_{ab}, f_{abc} en d_{abc} zijn de enige onafhankelijke invariante tensoren in de 8-representatie.

11. Representaties van $SU(N)$; Young tablaux

We hebben al even gezien hoe $SU(6)$ de kop opstak. Grotere Lie-groepen komen ook voor in de fysica. Er zijn algemeen geldige methoden om hun representaties te vinden. Allereerst besluiten we alle onder-indices naar boven te verhuizen m.b.v. de invariante temsor $\varepsilon^{\alpha_1\alpha_2\cdots\alpha_N}$.

De fundamentele representatie, corresponderend met de N-dimensionale 'spinor' ϕ^α in $SU(N)$, geven we aan met een hokje. Een produkt van twee zulke representaties wordt door twee hokjes weergegeven, zie Figuur 7a. Is de representatie symmetrisch, dan zetten we de hokjes naast elkaar, is zij antisymmetrisch, dan onder elkaar. In Fig. 7b zien we hoe grotere representaties zijn op te bouwen. We stellen ons voor dat in ieder hokje een index moet staan. Hebben we een ingewikkelde representatie, zoals in Fig. 7c, dan is de bijbehorende tensor symmetrisch onder alle permutaties van de indices binnen één horizontale rij hokjes. In vertikale richting is de regel wat ingewikkelder; deze wordt volledig bepaald door de eis dat alle representaties die we toestaan onafhankelijk van elkaar moeten zijn, zodat tensoren die dooe verschillende Young tableaux worden beschreven altijd loodrecht op elkaar moeten staan. Het blijkt dan dat we op de volgende regel uitkomen:

i Alle onafhankelijke, irreducibele representaties zijn van elkaar te onderscheiden doordat ze verschillende Young tableaux hebben, en ieder Young tableau stelt een irreducibele representatie voor, mits:

ii Iedere horizontale rij in een Young tableau mag niet meer hokjes bevatten dan de rij erboven, en iedere kolom mag niet meer hokjes bevatten dan de kolom links ervan.

iii Er kunnen niet meer dan $N-1$ rijen zijn, dus de kolommen mogen niet meer dan $N-1$ hokjes bevatten.

Een typisch voorbeeld is het Young tableau van Fig. 7c.

Appendix A

Overzicht van enkele eigenschappen van matrices

Zij gegeven een $n \times n$ matrix A met elementen a_{ij}:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

De matrix-elementen a_{ij} zijn (eventueel complexe) getallen. Zij verder nog gegeven een $n \times n$ matrix B met elementen b_{ij}.

Definitie: Het product van de matrixes A en B is de matrix C met matrix-elementen c_{ij} gegeven door

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{A.1}$$

We schrijven:

$$C = AB. \tag{A.2}$$

Toelichting: (A1) definieert de matrix-elementen c_{ij} als de a_{ij} en b_{ij} gegeven zijn. Aangezien de a_{ij} en b_{ij} gewone getallen zijn doet het er niet toe in welke volgorde ze geschreven worden. Dit geldt niet voor producten van matrices, waar in het algemeen $C = AB \neq BA$. Immers $C' = BA$ is een matrix met elementen c'_{ij} gegeven door

$$c'_{ij} = \sum_{k} b_{ik} a_{kj}. \tag{A.3}$$

We zeggen dat de matrix-vermenigvuldiging *niet-commutatief* is. Zij is wèl *associatief:* $A(BC) = (AB)C$.

Definitie: De commutator van twee matrices A en B wordt gegeven door

$$[A, B] \equiv AB - BA. \tag{A.4}$$

Men zegt dat A en B commuteren als de commutator nul is.

Definitie: De som, respectievelijk, het verschil van twee matrices A en B is een matrix C met matrix-elementen c_{ij} gegeven door:

$$c_{ij} = a_{ij} + b_{ij}, \qquad \text{respectievelijk,} \qquad c_{ij} = a_{ij} - b_{ij}. \tag{A.5}$$

Notatie:

$$C = A + B, \qquad \text{respectievelijk} \qquad C = A - B$$

Uit (A.5) volgt $A + B = B + A$.

Definitie: Het product van een (eventueel complex) getal x en een matrix A is een matrix B met de elementen:

$$b_{ij} = x a_{ij}. \tag{A.6}$$

Notatie:

$$B = xA \tag{A.7}$$

Uit (A6) volgt:

$$B^2 = x^2 A^2.$$

Immers

$$\sum_k b_{ik} b_{kj} = x^2 \sum_k a_{ik} a_{kj}.$$

Definitie: De éénheidsmatrix $\mathbf{1}$ is gedefinieerd als

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

en geeft $\mathbf{1}A = A\mathbf{1} = A$ voor iedere matrix A.
Bij elke $n \times n$ matrix behoort een (eventueel complex) getal, de determinant van die matrix. De matrix is (ontwikkeling naar i-de rij):

$$\mathrm{Det}(A) = \sum_{k=1}^{n} a_{ik} (-1)^{k+i} \mathrm{Det}(A(i,k)), \tag{A.8}$$

waar $A(i,k)$ de matrix is verkregen uit A door de i-de rij en k-de kolom te schrappen. De determinant van een 0×0 matrix is 1.

Er zijn vele eigenschappen van determinanten eenvoudig af te leiden: Als $C = AB$ dan is

$$\mathrm{Det}(C) = \mathrm{Det}(AB) = \mathrm{Det}(A)\,\mathrm{Det}(B). \tag{A.9}$$

Als $C = xA$ met x een getal en A een $n \times n$ matrix

$$\mathrm{Det}(C) = \mathrm{Det}(xA) = x^n \mathrm{Det}(A). \tag{A.10}$$

Verder: de inverse van een matrix A is een matrix B zodanig dat $BA = AB = \mathbf{1}$. Notatie: $B = A^{-1}$. De matrix-elementen zijn (merk op: $A(j,i)$, niet $A(i,j)$):

$$b_{ij} = \frac{(-1)^{i+j} \mathrm{Det}(A(j,i))}{\mathrm{Det}(A)}. \tag{A.11}$$

Hieruit blijkt dat een matrix A een inverse heeft als $\mathrm{Det}(A) \neq 0$. Een andere definitie van de determinant maakt gebruik van z.g. ϵ-symbolen, gedefinieerd door:

$\epsilon_{i_1 \cdots i_n} = 0,$ als $i_1 \ldots i_n$ twee of meer gelijke indices bevat
$\epsilon_{i_1 \cdots i_n} = \pm 1$ als $i_1 \ldots i_n$ géén gelijke indices bevat. In dat geval geldt
$\epsilon_{i_1 \cdots i_n} = 1$ als $i_1 \ldots i_n$ een even permutatie is van $123 \ldots n$
$\epsilon_{i_1 \cdots i_n} = -1$ als $i_1 \ldots i_n$ een oneven permutatie is van $123 \ldots n$. In het algemeen geldt
$\epsilon_{123 \cdots n} = +1.$

De determinant van een matrix A is dan gelijk aan

$$A_{i_1 j_1} A_{i_2 j_2} \cdots A_{i_n j_n} \, \epsilon_{j_1 \cdots j_n} = \mathrm{Det}(A) \, \epsilon_{i_1 \cdots i_n}.$$

Tenslotte: als $\widetilde{A}$ en A^{-1}, respectievelijk, de gespiegelde en inverse van A zijn dan is

$$\widetilde{AB} = \widetilde{B}\,\widetilde{A}, \qquad (AB)^{-1} = B^{-1}A^{-1}. \tag{A.12}$$

De eerste formule volgt uit de definitie van matrix product (A.1). De juistheid van de tweede formule blijkt uit $\mathbf{1} = (AB)\,(AB)^{-1} = AB\,B^{-1}A^{-1} = \mathbf{1}$.

Appendix B

Differentiatie van matrices

Matrix-elementen kunnen functies van een of meerdere variabelen zijn. Deze functies kunnen differentieerbaar zijn. In dat geval kan de afgeleide van de matrix gedefinieerd worden.

Definitie: De afgeleide van een matrix A is een matrix met als elementen de afgeleiden van de matrix-elementen van A.
Dus de matrix A' met

$$A'(x) = \frac{\mathrm{d}A(x)}{\mathrm{d}x} \tag{B.1}$$

heeft als matrix-elementen

$$a'_{ij}(x) = \frac{\mathrm{d}a_{ij}(x)}{\mathrm{d}x}. \tag{B.2}$$

De afgeleide van een matrix zal in de regel niet commuteren met de matrix zelf. Afgezien van dit feit zijn de formules bekend uit de differentiaalrekening ook hier geldig en volgen uit de definitie (B.1). Bijvoorbeeld,

$$\frac{\mathrm{d}(A+B)}{\mathrm{d}x} = \frac{\mathrm{d}A}{\mathrm{d}x} + \frac{\mathrm{d}B}{\mathrm{d}x}, \tag{B.3}$$

en

$$\frac{\mathrm{d}(AB)}{\mathrm{d}x} = \frac{\mathrm{d}A}{\mathrm{d}x}B + A\frac{\mathrm{d}B}{\mathrm{d}x} \tag{B.4}$$

Het laatste resultaat volgt uit de kettingregel

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_k a_{ik} b_{kj} = \sum_k a'_{ik} b_{kj} + \sum_k a_{ik} b'_{kj} \tag{B.5}$$

Op soortgelijke wijze

$$\frac{\mathrm{d}A^n}{\mathrm{d}x} = A^{n-1}A' + A^{n-2}A'A + \cdots + A'A^{n-1} \tag{B.6}$$

Merk op dat dit *niet* altijd gelijk is aan $n\,A^{n-1}\,A'$, omdat A en A' niet noodzakelijkerwijs commuteren.

De afgeleide van de inverse van een matrix A kan als volgt gevonden worden. Omdat $A^{-1}A = \mathbf{1}$ geldt

$$O = \frac{\mathrm{d}}{\mathrm{d}x}(A^{-1}A) = \frac{\mathrm{d}A^{-1}}{\mathrm{d}x}A + A^{-1}\frac{\mathrm{d}A}{\mathrm{d}x},$$

ofwel

$$\frac{\mathrm{d}A^{-1}}{\mathrm{d}x} = -A^{-1}\frac{\mathrm{d}A}{\mathrm{d}x}A^{-1}. \tag{B.7}$$

Appendix C

Functies van matrices

Functies gebaseerd op optelling en vermenigvuldiging kunnen ook voor matrices gedefinieerd worden. In het bijzonder de functie e^x is van belang. Hiervoor geven we twee mogelijke definities.

Zij gegeven een matrix A. De matrix e^A is gedefinieerd door:

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!}\,A^k, \tag{C.1}$$

waarbij $A^0 = \mathbf{1}$. Een tweede definitie is

$$e^A = \lim_{m\to\infty}\left(\mathbf{1} + \frac{1}{m}A\right)^m. \tag{C.2}$$

Het bewijs dat (C.1) = (C.2) gaat precies zoals voor gewone exponentialen, namelijk door toepassen van de binomiaal ontwikkeling van (C.2).

Formule (C.2) leidt direct tot een interessante formule:

$$\mathrm{Det}(e^A) = \mathrm{Lim}_{m\to\infty}\left\{\mathrm{Det}(\mathbf{1} + \frac{1}{m}A)\right\}^m. \tag{C.3}$$

We berekenen de determinant tussen haakjes met verwaarlozing van termen $1/m^2$ en kleiner. Als voorbeeld bekijken we een 2×2 matrix. Dan:

$$\mathbf{1} + \frac{1}{m}A = \begin{pmatrix} 1+\frac{a_{11}}{m} & \frac{a_{12}}{m} \\ \frac{a_{21}}{m} & 1+\frac{a_{22}}{m} \end{pmatrix}$$

Als we termen van orde $1/m^2$ en hoger verwaarlozen in de berekening van de determinant dan kunnen de niet-diagonaal elementen verwaarloosd worden (in een determinant zijn er geen termen die slechts één niet diagonaal-element als factor hebben). In deze benadering wordt de determinant eenvoudig het product van de diagonaal-elementen:

$$\begin{aligned} \mathrm{Det}\left(\mathbf{1} + \tfrac{1}{m}A\right) &\approx \left(1 + \tfrac{a_{11}}{m}\right)\left(1 + \tfrac{a_{22}}{m}\right) \\ &\approx 1 + \tfrac{1}{m}\left(a_{11} + a_{22}\right) + \mathrm{O}\left(\tfrac{1}{m^2}\right). \end{aligned}$$

In het algemeen vindt men:

$$\mathrm{Det}\left(\mathbf{1}+\frac{1}{m}A\right)=1+\frac{1}{m}Sp(A)+\mathrm{O}\left(\frac{1}{m^2}\right)$$

waar $Sp(A)$ het spoor van A is, d.w.z. de som van alle diagonaalelementen:

$$Sp(A)=\sum_k a_{kk}.$$

Dit invullen in (C.3) geeft

$$\mathrm{Det}(e^A)=\lim_{m\to\infty}\left\{1+\frac{1}{m}Sp(A)+\mathrm{O}\left(\frac{1}{m^2}\right)\right\}^m=e^{Sp(A)},$$

zodat geldt

$$\mathrm{Det}(e^A)=e^{Sp(A)}. \tag{C.4}$$

Appendix D

De Campbell-Baker-Hausdorff formule

De CBH-formule zegt dat er een matrix C bestaat, zodat

$$e^Ae^B=e^C, \tag{D.1}$$

met $C=A+B+{}^1\!/_2[A,B]$ + herhaalde commutatoren van A en B.
Onder herhaalde commutatoren verstaan we uitdrukkingen van het type

$$\begin{aligned}[[A,B],B] &= [A,B]B-B[A,B]\\ &= ABB-BAB-BAB+BBA,\end{aligned}$$

$$[[B,[A,B]],B],\qquad [[[A,B],B],B],\qquad \text{enz.}$$

De numerieke factoren welke als coefficienten van de herhaalde commutatoren optreden zijn helaas niet eenvoudig te vinden. Voor onze doeleinden is het voldoende in te zien dat C inderdaad van de vorm (D.1) is. We zullen dit bewijzen en tevens de coefficienten van de multipele commutatoren met twee B's en één A of twee A's en één B vinden.

Allereerst voeren we een variabele x in en beschouwen

$$e^{xA}e^{xB}=e^{C(x)}. \tag{D.2}$$

Dus de x-afhankelijkheid van het linkerlid is eenvoudig en expliciet in (D.2), echter C zal een gecompliceerde functie van x zijn. In feite, de reeksontwikkeling van $C(x)$ naar x zal tevens de ontwikkeling van C naar aantallen factoren A en/of B zijn. We proberen dus C als functie van x te vinden. Dit doen we door (D.2) links en rechts te differentieren naar x, en dat geeft dan een vergelijking waaruit C in de vorm van een machtreeks opgelost kan worden. Eerst moeten we echter nog enkele formules bewijzen.

Beschouw

$$H(y) = e^{-yF} G\, e^{yF} \tag{D.3}$$

waarin y een variable is en F en G van y onafhankelijke matrices. H kan in de vorm van een machtreeks geschreven worden:

$$H = H_0 + yH_1 + \frac{y^2}{2!}H_2 + \frac{y^3}{3!}H_3 + \cdots . \tag{D.4}$$

Hierin zijn de factoren 2!, 3!, ... toegevoegd om de tussenresultaten straks te vergemakkelijken. We zullen de H_n berekenen. Differentiatie van (D.3) naar y geeft:

$$\frac{\mathrm{d}H(y)}{\mathrm{d}y} = \left(\frac{\mathrm{d}e^{-yF}}{\mathrm{d}y}\right) G\, e^{yF} + e^{-yF} G\, \frac{\mathrm{d}e^{yF}}{\mathrm{d}y}. \tag{D.5}$$

Met behulp van

$$\frac{\mathrm{d}}{\mathrm{d}y} e^{yF} = F\, e^{yF} = e^{yF}\, F \tag{D.6}$$

vinden we

$$\frac{\mathrm{d}H(y)}{\mathrm{d}y} \;=\; -F e^{-yF} G e^{yF} + e^{-yF} G e^{yF} F = [H, F]\,. \tag{D.7}$$

Dit vullen we in in de machtreeks (D.4):

$$H_1 + y\, H_2 + \frac{y^2}{2!}H_3 + \cdots \;=\; [H_0, F] + y[H_1, F] + \frac{y^2}{2!}[H_2, F] + \cdots . \tag{D.8}$$

Aangezien $H_0 = G$, impliceert het gelijkstellen van de coëfficiënten voor de verschillende machten van y :

$$H_n = [H_{n-1}, F] = [[[\cdots[G, F], F], F]\cdots F]. \tag{D.9}$$

Invullen van (D.4) geeft voor $y = 1$:

$$\begin{aligned} e^{-F} G\, e^{F} = G \;&+\; [G, F] + \frac{1}{2!}[[G, F], F] + \\ &+\; \frac{1}{3!}[[[G, F], F], F] + \cdots . \end{aligned} \tag{D.10}$$

We voeren een verkorte notatie in:

$$H_n = \{G, F^n\}\ , \tag{D.11}$$

en, nog korter:

$$e^{-F} G e^{F} = \{G, e^{F}\}\ , \tag{D.12}$$

waarmee we bedoelen dat e^F eerst door zijn machtreeks moet worden vervangen. Laat nu F een functie van een variabele x zijn. Beschouw

$$e^{-yF(x)} \frac{\mathrm{d}}{\mathrm{d}x} e^{yF(x)}. \tag{D.13}$$

Op gelijke wijze als boven vindt men:

$$\begin{aligned} e^{-F(x)}\frac{\mathrm{d}}{\mathrm{d}x}e^{F(x)} &= F' + \frac{1}{2!}[F',F] + \frac{1}{3!}[[F',F],F] + \cdots \\ &= \{F', \frac{e^F-1}{F}\}\ . \end{aligned} \tag{D.14}$$

In feite is (D.14) wat men krijgt door $G = \frac{\mathrm{d}}{\mathrm{d}x}$ te zetten in (D.10). Immers dan is:

$$[G,F] = \frac{\mathrm{d}}{\mathrm{d}x}F - F\frac{\mathrm{d}}{\mathrm{d}x} \tag{D.15}$$

Hierin wordt verondersteld dat $\mathrm{d}/\mathrm{d}x$ naar rechts werkt op wat er eventueel nog rechts van de commutator $[G,F]$ staat.
Volgens de kettingregel is

$$\frac{\mathrm{d}}{\mathrm{d}x}F = \frac{\mathrm{d}F}{\mathrm{d}x} + F\frac{\mathrm{d}}{\mathrm{d}x} = F' + F\frac{\mathrm{d}}{\mathrm{d}x}, \tag{D.16}$$

dus

$$[G,F] = F'. \tag{D.17}$$

Uit (D.14) volgt:

$$\frac{\mathrm{d}}{\mathrm{d}x}e^{F(x)} = e^{F(x)}\{F', \frac{e^F-1}{F}\} \tag{D.18}$$

We keren nu terug maar (D.2). Differentiatie naar x geeft met toepassing van (D.7):

$$e^{xA}Ae^{xB} + e^{xA}Be^{xB} = e^{C(x)}\{C', \frac{e^C-1}{C}\} \tag{D.19}$$

of

$$e^{xA}e^{xB}\left(e^{-xB}Ae^{xB} + e^{-xB}Be^{xB}\right) = e^{C(x)}\{C', \frac{e^C-1}{C}\}. \tag{D.20}$$

De factoren buiten haakjes vallen tegen elkaar weg, vanwege (D.2). Vanwege $Be^{xB} = e^{xB}B$ geeft de tweede term binnen haakjes precies B. Op de eerste term kunnen we (D.12) toepassen. Er komt zo:

$$\{C', \frac{e^C-1}{C}\} = \{A, e^{xB}\} + B\ . \tag{D.21}$$

Uit deze vergelijking kan C in de vorm van een machtreeks opgelost worden. De kracht van onze abstracte notatie blijkt als we de rechterkant van deze vergelijking even H noemen:

$$C' = \{H, \frac{C}{e^C-1}\} = \{H, 1 - \frac{C}{2} + \frac{C^2}{12} + \cdots\} = H - {}^1\!/\!_2[H,C] + \frac{1}{12}[[H,C],C] + \cdots\ , \tag{D.22}$$

waarbij we dus de machtreeks van $\frac{C}{e^C-1} = \frac{1}{1+{}^1\!/\!_2 C + \frac{1}{6}C^2 + \cdots}$ gebruikt hebben.

Schrijven we

$$\begin{aligned}
C &= C_0 + xC_1 + \frac{x^2}{2!}C_2 + \frac{x^3}{3!}C_3 + \cdots , \\
C' &= C_1 + xC_2 + \frac{x^2}{2!}C_3 + \cdots , \\
H &= H_0 + xH_1 + \frac{x^2}{2!}H_2 \;=\; A + B + x[A,B] + \frac{x^2}{2!}[[A,B],B] + \cdots ,
\end{aligned} \tag{D.23}$$

dan kunnen we nu gelijke machten van x in (D.22) en (D.23) vergelijken. Allereerst:

$$C_0 = C(0) = 0. \tag{D.24}$$

De termen onafhankelijk van x geven

$$C_1 = H_0 = A + B\,, \tag{D.25}$$

en de termen lineair in x:

$$C_2 = H_1 - {}^1\!/_2[H_0, C_1] = [A, B]\,. \tag{D.26}$$

En zo voortgaande kunnen we alle andere coëfficiënten C_n vinden:

$$\begin{aligned}
{}^1\!/_2 C_3 &= {}^1\!/_2 H_2 - {}^1\!/_4[H_0, C_2] - {}^1\!/_2[H_1, C_1]\,, \\
C_3 &= [[A,B],B] - {}^1\!/_2[A+B,[A,B]] - [[A,B],A+B] \\
&= {}^1\!/_2[A[A,B]] + {}^1\!/_2[[A,B],B]\,,
\end{aligned} \tag{D.27}$$

(omdat $[B,[A,B]] = -[[A,B],B]$).

Steeds vinden we zo dat de volgende coefficient C_i te schrijven is in termen van multipele commutatoren welke de voorgaande $C_j, j < i$ bevatten. Aangezien de C_j zelf multipele commutatoren zijn volgt dan hetzelfde voor C_i.

Invullen van (D.24), (D.25), (D.26) en (D.27) in de reeksontwikkeling (D.23) geeft, voor $x = 1$:

$$e^A e^B = e^C \tag{D.28}$$

$$C = A + B + {}^1\!/_2[A,B] + \tfrac{1}{12}[A,[A,B]] + \tfrac{1}{12}[[A,B],B] + \cdots \ . \tag{D.29}$$

Het is een aardige uitdaging de hier gevolgde methode om de reeks van commutatoren voor C te vinden nog verder te stroomlijnen.

Appendix E

Complex inproduct, unitaire en hermitische matrices

Zij gegeven in een n-dimensionale lineaire ruimte met een basis $e_1 \cdots e_n$ die orthonormaal is met betrekking tot een bepaald inproduct. Elke vector in deze ruimte kan geschreven worden als een lineaire combinatie van de $e_1 \cdots e_n$. De coefficienten zijn i.h.a. complexe getallen.

Beschouw nu 2 vectoren α en β,

$$\begin{aligned} \alpha &= \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n, \\ \beta &= \beta_1 e_1 + \beta_2 e_2 + \cdots + \beta_n e_n. \end{aligned}$$

De coefficienten α_i en β_i zijn dus getallen. Het inproduct (α, β) is een complex getal gegeven door

$$(\alpha, \beta) = \alpha_1^* \beta_1 + \alpha_2^* \beta_2 + \cdots + \alpha_n^* \beta_n. \tag{E.1}$$

Merk op dat geldt

$$(\beta, \alpha) = (\alpha, \beta)^*. \tag{E.2}$$

Er geldt (axioma's voor inproduct)

$$(\alpha, \alpha) \geq 0, \text{ gelijkteken alleen als } \alpha = 0, \tag{E.3}$$

$$(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma), \tag{E.4}$$

$$(\alpha, x\beta) = x(\alpha, \beta) \qquad x = \text{getal} \tag{E.5}$$

$$(x\alpha, \beta) = x^*(\alpha, \beta) \tag{E.6}$$

Hierin is $\beta + \gamma$ een vector met componenten $\beta_i + \gamma_i$, en $x\beta$ is een vector met componenten $x\beta_i$.
Men kan de zaak ook omkeren: gelden bovenstaande eigenschappen dan volgt uit (E.1) uit de orthonormaliteit van de basisvectoren:

$$(e_i, e_j) = \begin{cases} 1 & \text{als } i = j \\ 0 & \text{als } i \neq j \end{cases} \tag{$E.7$}$$

Zij nu verder gegeven een matrix A. Deze matrix definieert een afbeelding van de ruimte in zichzelf:

$$\alpha' = A\alpha, \tag{E.8}$$

hetgeen symbolisch is voor

$$\alpha_i' = \sum_k a_{ik} \alpha_k. \tag{E.9}$$

Bij elke vector α wordt dus door A een beeld α' gedefinieerd. Onder welke voorwaarden te stellen aan A is het inproduct invariant, in de zin dat $(\alpha, \beta) = (\alpha', \beta')$ voor willekeurige

α en β?
Door uitschrijven in componenten is dit eenvoudig af te leiden:

$$\begin{aligned}(\alpha', \beta') &= \sum_k \alpha_k'^* \beta_k' = \sum_{klm} (a_{kl}\alpha_l)^* a_{km}\beta_m \\ &= \sum_{k,l,m} \alpha_l^* a_{kl}^* a_{km}\beta_m.\end{aligned}$$

Als nu $\widetilde{A}$ de gespiegelde van A is dan is

$$(\widetilde{a})_{lk} = a_{kl} \tag{E.10}$$

en als $A^\dagger$ de gespiegelde en complex geconjugeerde van A is:

$$(a^\dagger)_{lk} = a_{kl}^* \qquad \text{en als } B = A^\dagger A \tag{E.11}$$

dan is blijkbaar

$$\begin{aligned} b_{lm} &= \sum_k a_{lk}^\dagger a_{km} \\ &= \sum_k a_{kl}^* a_{km}\end{aligned}$$

Dus:

$$(\alpha', \beta') = \sum_{l,m} \alpha_l^* b_{lm}\beta_m \tag{E.12}$$

met $\beta = A^\dagger A$. Is nu β'' de vector verkregen uit β door B er op toe te passen:

$$\beta'' = B\beta \tag{E.13}$$

d.w.z. volgens (E.9):

$$\beta_l'' = \sum_m b_{lm}\beta_m \tag{E.14}$$

en (E.12) is niets anders dan het inproduct (α, β''), dus $(\alpha', \beta') = (\alpha, \beta'')$. Dit moet gelijk zijn aan (α, β) voor alle α en β. Dat kan alleen als $\beta'' = \beta$ voor alle β, en dus moet B de eenheidsmatrix zijn.
Het inproduct is dus invariant onder afbeelding met een matrix A als geldt:

$$A^\dagger A = \mathbf{1}. \tag{E.15}$$

ofwel $A^\dagger = A^{-1}$. Een matrix die voldoet aan deze voorwaarde heet *unitair*. In reële ruimten is $A^* = A$, en dus $A^\dagger = \widetilde{A} = A^{-1}$.
Een korte symbolische weergave van de bovenstaande afleiding is:

$$(\alpha', \beta') = (A\alpha, A\beta) = (\alpha, A^\dagger A\beta). \tag{E.16}$$

Voor inproducten geldt dus:

$$(A\alpha, \beta) = (\alpha, A^\dagger\beta) \tag{E.17}$$

voor willekeurige matrix A.

Een eigenvector van een matrix A is een vector zodanig dat

$$A\alpha = \gamma\alpha \tag{E.18}$$

d.w.z.

$$\sum a_{ik}\alpha_k = \gamma\alpha_i. \tag{E.19}$$

γ is de eigenwaarde van A behorende bij de eigenvector α. Onder welke voorwaarden te stellen aan A zijn de γ reëel?
Beschouw het inproduct $(\alpha, A\alpha) = \gamma(\alpha, \alpha)$. Nu is (α, α) reëel en ongelijk nul als α niet nul is. Reële γ betekent kennelijk dat $(\alpha, A\alpha)$ reëel is. D.w.z.:

$$(\alpha, A\alpha) = (\alpha, A\alpha)^* = (A\alpha, \alpha) = (\alpha, A^\dagger\alpha). \tag{E.20}$$

Kennelijk is het voldoende als $A^\dagger = A$. Een matrix A met $A^\dagger = A$ heet hermitisch. Dus:

$$A = \text{hermitisch, d.w.z. } A^\dagger = A \longrightarrow \text{alleen reële eigenwaarden.} \tag{E.21}$$