Universitaire wiskunde voor intelligente tieners.

Johan Noldus[1]

April 9, 2017

[1]Johan.Noldus@gmail.com, Relativiteits-groep, departement Wiskunde, Galglaan 2, 9000 Gent.

Contents

Een woordje vooraf.

Deze tekst is geschreven naar aanleiding van de leergierigheid van mijn twaalf jaar oude zoon Eliott, die me gevraagd had een goed wiskunde boek te kopen voor hem. Geheel ontevreden met wat ik op de markt vond heb ik maar besloten zelf iets te schrijven waarmee hij verder kan; het moge duidelijk wezen dat ik niet geloof in de betutteling van onze jeugd en dat zij best in staat is hogere wiskunde te assimileren indien het goed uitgelegd en gemotiveerd is. De bedoeling van deze tekst is hun verbeelding aan te wakkeren en te worstelen met de problemen die de beste wiskundigen slapeloze nachten hebben bezorgd. De tekst is doorspekt met originele oefeningen die lang niet allemaal gemaakt dienen te worden; de lezer moet desalniettemin begrijpen waarom de resultaten van de opgaven juist zijn. Dit boek bevat alle wiskunde die je in een A-programma in de middelbare school ziet alsook veruit de meeste topics die je in een master opleiding aan een doorsnee universiteit te verwerken krijgt en dit op een 150-tal bladzijden. De presentatie is geheel origineel en legt de nadruk op conceptueel denken en fantasie en veel minder op het mechanisch maken van oefeningen; de lezer krijgt een diep gevoel voor de problematiek die voor hem ligt alsook een arsenaal aan ideeen om deze te lijf te gaan.

Chapter 1

Wiskunde, een boeiende taal.

Ieder kind heeft zo zijn eigen woordenschat wat het soms moeilijk maakt dingen tegen elkaar te zeggen. De meest eenvoudige woorden betekenen voor ieder wel iets anders en dat maakt dat we allemaal een eigen interpretatie erop nahouden. Wiskunde is een poging om iets eenduidig te zeggen en zoals zal blijken is dit enorm moeilijk; bijvoorbeeld, wat betekent het dat twee dingen aan elkaar gelijk zijn? Dat is niet zo makkelijk als je wel zou denken. Vermits de wiskunde een eenduidige taal is, is het mogelijk verbanden tussen bepaalde concepten te vinden, deze noemen we stellingen. Deze stellingen zitten diep verborgen in de taal zelf en wij mensen hebben de mogelijkheid deze te ontdekken; het is echter een kunst de concepten op de juiste manier te formuleren zodat de stellingen zo eenvoudig mogelijk bewezen worden. Indien je een bepaald resultaat wenst te bewijzen en je vindt maar geen logische redenering erachter is de kans groot dat je met de verkeerde concepten werkt. Jongeren hebben daar vaak last van, ze willen vanalles zeggen, maar vinden de juiste woorden niet. Dat is een kwestie van oefening en training en in de loop van dit boek zal het erg duidelijk worden dat er erg veel verschillende manieren bestaan om iets te zeggen; de keuzes die we maken, de zogenaamde axioma's zijn daarom erg varierend en kunnen aanleiding geven tot verschillende soorten wiskunde. Inderdaad, ik maak de lezer erop attent dat heel wat mogelijk is en dat je je verbeelding de vrije loop kan laten; soms zal ik van die erg geavanceerde redeneervormen aanhalen en de lezer die daarop verder wil werken moet een gespecialiseerd boek kopen. Mijn doel hier is jongeren op een hoger niveau te tillen op een rustige maar rechtlijnige manier en hun bewust te maken van de rijkdom van de taal. Ik ga ook nog een tweede boek over fysica schrijven indien dit aanslaat gegeven dat je eerst de taal van de natuur moet leren spreken vooraleer je haar wilt bestuderen.

Een bewijs is een "logische" redenering en aldus moet logica een basis onderdeel van de wiskunde vormen. Zonder logica, geen juist bewijs en dus geen stellingen.

De logica die in de wiskunde gebruikt wordt is de meest eenvoudige mogelijk: een stelling of bewering is oftewel goed of fout. Goed betekent dat ze in alle gevallen klopt en fout betekent dat er tenminste één tegenvoorbeeld gevonden kan worden. Vandaar ook dat we in de logica de symbolen $\forall$ en $\exists$ gebruiken die betekenen "voor alle", respectievelijk, "er bestaat". Verder zijn de symbolen $A \wedge B$ en $A \vee B$ van toepassing waarbij de eerste uitspraak waar is als en slechts als ze beide waar zijn en de laatste is waar wanneer tenminste één van de twee waar is. In mensentaal leest dit A én B, vervolgens A of B. De lezer kan nu controleren dat $A \wedge (B \vee C)$ hetzelfde is dan $(A \wedge B) \vee (A \wedge C)$, een formule die gekend staat als de regel van de Morgan. Dit soort logica staat gekend als Booleaanse of klassieke logica waarbij een uitspraak waar of vals is en waar er een absoluut tegengestelde bestaat $\neg A$; dus $\neg A$ is waar als en slechts A vals is. De lezer kan controleren dat $\neg(A \vee B)$ hetzelfde is dan $\neg A \wedge \neg B$; men kan ook wiskunde ontwikkelen waarbij uitspraken gedeeltelijk waar én vals zijn met een bepaalde kans. Dit soort logica wordt quantum logica genoemd, verder bestaat er nog altijd de mogelijkheid om in te bouwen dat een bepaalde uitspraak onbepaald is, wat betekent dat het niet geweten is. Dat soort logica wordt vage logica genoemd en de geinteresseerde lezer kan hier op verder borduren. Een mijlpaal in de geschiedenis van de wiskunde was dat de klassieke logica niet strict bewijsbaar is, waarmee ik bedoel dat er binnen elk axiomatisch stelsel stellingen bestaan die waar zijn maar niet bewezen kunnen worden via een logische redenering. Dit resultaat, dat vooral een groot filosofisch belang heeft, staat gekend als de onvolledigheids stelling van Godel.

Het hoofd idee achter Godel's theorema is simpel en goed gekend in de taalkunde; in de wiskunde wordt namelijk van elke uitspraak verwacht dat je kunt controleren of deze waar al dan niet vals is. Als een uitspraak waar is kun je trachten dit aan te tonen door te onderstellen dat ze vals is en vervolgens op een tegenstrijdigheid botsen, maar compleet nutteloos zijn zinnen die niet aan deze voorwaarde voldoen. Bijvoorbeeld: "deze zin is juist" is van dat type, je kunt niet *bewijzen* of hij juist of vals is alhoewel hij juist is; nog grappiger is de zin "deze zin is vals". Als je veronderstelt dat hij juist is dan kom je op een contradictie vermits hij vals is wat suggereert dat hij vals is. Omgekeerd leidt de veronderstelling dat hij vals is tot de conclusie dat hij juist moet zijn. Dit zijn allemaal zinnen met zelf-referentie of waarbij de eigenschappen van de zin betrekking hebben tot de ganse zin zelf; da's net zoals een verzameling die een element is van zichzelf wat uiteraard verboden is.

Laat het dus duidelijk wezen: aan de grondslagen van de wiskunde kan gesleuteld worden. Zelfs de manier waarop we redeneren is voor verandering vatbaar wat zich in de moderne aanpak naar topos theorie toe uit. Dit is een voorbeeld van quantum wiskunde en de notie van stelling en bewijs is hier geheel verschillend dan het is in de klassieke wiskunde. Vermits deze de meest eenvoudige en directe is vormt zij het onderdeel van dit boek, maar ik zal er dikwijls op wijzen dat bepaalde zaken helemaal niet zo "evident" of vanzelfsprekend zijn. Er zijn wiskundigen die de mening toegedaan zijn dat het ultieme doel is alles zo op te schrijven dat een computer kan controleren of een bewijs juist al dan niet fout is: deze behoren tot de zogenaamde Bourbaki strekking die in Parijs ontstaan is. Alhoewel ik het nut van deze filosofie inzie ben ik haar niet gans genegen: het is namelijk ongelofelijk tijd en plaatsrovend om met dit uiterste detail te werken. Mijn aanpak is praktischer *maar* de lezer moet begrijpen dat vaak het allermoeilijkste is een concept of idee exact te verwoorden: je moet jezelf hierin trainen en in die zin is het *beginnen* te werken met de Bourbaki methode wel aangewezen. Eer je een voldoend hoog niveau van bekwaamheid hebt speelt dit hulpstuk helemaal geen rol meer. Dus ik raad de lezer aan in het begin elke tussenstap in een berekening of bewijs te maken en te verantwoorden in exact detail.

De opbouw van dit boek is vanaf de grond op naar boven: we beginnen met verzamelingenleer, iets wat je allemaal wel gezien hebt in de lagere school en de eerste graad van het middelbaar onderwijs. Dan ontwikkelen we de natuurlijke, gehele, rationale, reeele en complexe getallen: het vinden van een gepaste symbolische notatie hier is veruit het moeilijkst gegeven dat een reeel getal nooit exact neergeschreven kan worden en toch moeten we een manier vinden om te praten over datgene wat niet neergepend kan worden. Indien we voldoende rekenvaardig worden gaan we aan topologie doen wat betekent dat we kaarten gaan leren maken van een ruimte; die kaarten bevatten erg veel informatie wat zich uit in het begrip homologie. homologie bevat vrijwel alle informatie van een ruimte en is niks anders dan een reeks getalletjes: dus we hebben een cijfervormige code gevonden om een bol te onderscheiden van een auto band (torus) en die code is opgebouwd uit eigenschappen van die ruimte. Eigenlijk is het idee achter homologie erg simpel: de rand van een eindig volume is een gesloten oppervlak wat betekent dat het geen rand heeft. Kort samengevat betekent dit dat de rand van de rand leeg is: deze simpele observatie karakteriseert het ganse homologie begrip. Nu dat we vertrokken zijn van erg algemene ruimten en vrij abstracte atlassen, of bedekkingen met kaarten, zoals je de wereldbol kunt bedekken met de kaarten van elk continent is het tijd om te specialiseren. We gaan ons richten naar de zogenaamde lineaire ruimten, dat zijn ruimten waarbij je verplaatsingen v, w met mekaar kunt optellen zodat $v + w = w + v$ en kunt uitrekken of inkrimpen met een getal rv. Op een bol kun je ook twee verplaatsingen na elkaar nemen, maar de lezer moet er zich van overtuigen dat het omwisselen van deze verplaatsingen een ander resultaat geeft. Bovendien heb je op een bol dat helemaal rond gaan hetzelfde is dan de nul verplaatsing: met andere woorden $rv = 0$ voor een zekere $r \neq 0$. Op lineaire ruimten bestaat er een natuurlijke klasse van functies: de zogenaamd lineaire functies welke veralgemeningen vormen van matrices. Deze lineaire functies vormen benaderingen van algemene functies en zijn van groot belang; we bestuderen deze, alsook een paar belangrijke resultaten hieromtrent en presenteren enkele voorbeelden aan de hand van de zogenaamde kwadratische oppervlakken in twee en drie dimensies. Van dan af begint het allemaal en bestuderen we de praktische wiskunde die we in de fysica dagdagelijks gebruiken: hier besteden we heel wat aandacht aan de analytische meetkunde die je beoefent aan de hand van het berekenen van formules. De eerste hoofdstukken van dit boek zijn erg abstract, doch dit mag je niet afschrikken en je moet hier voldoende veel tijd aan besteden. Laat je fantasie maar werken en denk na over hoe jij het zou aanpakken; ik suggereer alleszins verschillende wijzen om deze zaken te benaderen. Vanaf hoofdstuk vijf wordt alles meer praktisch en leren we rekenen op meer concrete ruimten.

Het interessante aan dit concept is dat zaken die in de traditionele benadering doorgaans eerst behandeld worden zoals de exponentiele functie en de sinus en cosinus nu op de laatste plaats komen maar veel diepzinniger en ingenieuzer aangebracht worden. Dus de traditionele moeilijkere rekenoefeningen komen helemaal op de laatste plaats. De bonus hiervan is dat de lezer deze speciale functies onmiddelijk kan uitbreiden naar niet-conventionele getallenstelsels toe zoals de quaternionen en Cliffordgetallen welke erg belangrijk zijn in de moderne fysica. Ik heb het oprechte vermoeden, geput uit mijn eigen ervaringen, dat een dieper conceptueel begrip ook leidt tot een betere rekenkunde. Er zijn natuurlijk heel wat topics die niet in volledige diepte zijn aangeraakt zoals dit het geval is voor getallenleer; dit is echter een bewuste keuze gegeven dat ik in het finaal hoofdstuk nieuwe zienswijzen zal aanraken die compleet coordinaatonafhankelijk zijn. Inderdaad, ik geloof niet in het fetisjisme van de getallentheoreten die menen dat de Riemann hypothese ook maar enig nuttig materiaal kan aanbrengen omtrent de natuurkunde. Veel plezier toegewenst met dit boek dat echt een werkboek moet worden.

Chapter 2

Eerste stappen: verzamelingenleer.

Verzamelingenleer is een erg abstract en moeilijk domein en het is vereist dat je hier de nodige aandacht aan besteed. De aanpak die ik hier zal volgen is een beetje eigenzinnig en iets algemener dan wat gewoonlijk het geval is. Het ongelofelijke voordeel is dan weeral dat je letterlijk ziet wat voor soort veralgemeningen kunnen gebeuren: ik zal ook een paar van die abstracte gevallen presenteren en concrete situaties geven met dewelke ze overeenstemmen. In de wiskunde is het dikwijls zo dat je enkele concrete voorbeelden in je gedachten hebt waarover je wilt praten maar dat de abstractie van die ideeen aanleiding geeft tot situaties waar je voorheen nooit aan gedacht hebt. Bijvoorbeeld, ik kan denken over een verzameling als een zak met spullen in, een tandenborstel, tandpasta, een brood, jam enzovoort, elementen zoals we in de verzamelingenleer zeggen. Maar ik hoef dit niet te doen, misschien zit er in die zak nog wel iets dat ik op zich niet kan pakken maar dat ik er stuksgewijs kan uithalen wanneer ik mijn andere items eruit haal. Bijvoorbeeld, een halve liter water die ik gesmost heb op mijn kaas en brood; verzamelingenleer bevat allemaal zulke details niet en spreekt zich gewoon uit of er een halve liter in mijn zak zit ongeacht of ik het er kan uithalen of niet. Dus misschien is het gewoon in ons voordeel om niet te beginnen met een zak en items, maar gewoon met de handelingen van delen $\cap$ en verenigen $\cup$, de doorsnede en unie respectievelijk. Nu, in standaard verzamelingenleer veronderstellen we wel dat we kunnen groeperen, of in een zak steken, maar dat de zak als dusdanig eigenlijk niet meetelt. Dus we kunnen twee zakken hebben die gemeenschappelijke stukken bevatten zonder dat ik ergens een gat in hoef te maken of zo, een “spookzak”. Direct hierna zullen we het voorbeeld geven van een realistische zak met een cirkelvormig gat in aan de zijkant. Voor de spookzak geldt dat $A \cap B = B \cap A$ en $A \cup B = B \cup A$, we spreken over deze eigenschappen als het commutatief zijn van de doorsnede en unie. Ook hebben we dat $(A \cap B) \cap C = A \cap (B \cap C)$ en hetzelfde voor de unie, wat bekend staat als de associativiteit van de doorsnede, respectievelijk unie. Noteer dus met $A, B, C, \ldots$ verzamelingen, we weten nog niet wat ze zijn maar gaan nu bewerkingen en eigenschappen formuleren; elke verzamelingen theorie $\mathcal{S}$ heeft commutatieve en associatieve bewerkingen $\cap$ en $\cup$ die bovendien voldoen aan $A \cap A = A \cup A = A$. Men eist bovendien dat er een unieke lege verzameling $\emptyset$ bestaat zodat hetvolgende geldig is

$$\begin{aligned} A \cap \emptyset &= \emptyset \\ A \cup \emptyset &= A \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \end{aligned}$$

waar deze laatste regel dezelfde is dan de de-Morgan regel uit de logica. Dit is al heel wat hoor en we onderzoeken enkele gevolgen van deze definities. We zeggen dat A een deelverzameling is van B, $A \subseteq B$ als en slechts als $A \cap B = A$; bovendien is A primitief wanneer elke niet lege deelverzameling A zelf is. In wiskundige taal leest dit dat als $B \subseteq A$ dan moet $B = A$; de lezer controleert dat $A \cap C \neq \emptyset$ een deelverzameling is van A, vermits $A \cap (A \cap C) = (A \cap A) \cap C = A \cap C$, en dus $A = A \cap C \subseteq C$ indien A primitief is. Een primitieve verzameling is zoals een atoom in de fysica: het is ondeelbaar en er bestaat geen kleiner deel dan dat. Dus primitieve verzamelingen kunnen opgevat worden als elementen en we hanteren de notatie, per afspraak, dat $A = \{\widehat{A}\}$. $\widehat{A}$ is dan een element en we schrijven dat $\widehat{A} \in B$ als en slechts als $A \cap B = A$.

Dus, we kunnen elementen definieren vanuit de operaties $\cap, \cup$; ik ben nogmaals niet van elementen vertrokken omdat het niet duidelijk is of ze wel bestaan en dat het niet noodzakelijk zo is dat ze de ganse verzameling uitmaken. Bijvoorbeeld, een oneindige rechte hoeft niet te bestaan uit punten, punten zijn "abstracties" die niet noodzakelijk enig realiteitsgehalte hebben. Desalniettemin, laten we verder onderzoeken wat we kunnen zeggen over de elementen. Als $B \subseteq C$, dan behoort elk element $\widehat{A}$ van B tot C: inderdaad $\widehat{A} \in B$ als en slechts als $A \cap B = A$ en dus $A \cap C = (A \cap B) \cap C = A \cap (B \cap C) = A \cap B = A$ wat dus aantoont dat $A \cap C = A$ en dus $\widehat{A} \in C$. Een ander bewijs is dat $\widehat{A} \in B$ als en slechts als $A \cap B = A$ wat equivalent is aan $(A \cap C) \cap B = A$ en dus $A \cap C \neq \emptyset$ waaruit volgt dat $A \cap C = A$ omdat A primitief is. Dus elementen van deelverzamelingen behoren tot de verzameling zelf. Wat nu betreffende de doorsnede van twee verzamelingen? Eerst tonen we aan dat indien $\widehat{A} \in B, C$ dan is $\widehat{A} \in B \cap C$: dit is waar omdat $A \cap (B \cap C) = (A \cap B) \cap C = A \cap C = A$ en dus $\widehat{A} \in B \cap C$. Aan de andere kant hebben we dat als $\widehat{A} \in B \cap C$ dan is $\widehat{A} \in B, C$ omdat de doorsnede van twee verzamelingen een deelverzameling is van elk. Dus de elementen in de doorsnede van twee vezamelingen zijn precies de gemeenschappelijke elementen van beide verzamelingen. Wat nu betreffende de unie? We tonen aan dat indien $\widehat{A} \in B \cup C$ dan moet $\widehat{A} \in B$ of $\widehat{A} \in C$ vermits $A = A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ en dat daarom tenminste een van die twee niet leeg moet zijn en dus gelijk aan A omdat A primitief is. We hebben ook dat als $\widehat{A} \in B$ dan behoort het tot $B \cup C$ vermits $A \cap (B \cup C) = A \cup (A \cap C)$ en dit is gelijk aan $A \cup A$ of $A \cup \emptyset$ omdat A primitief is. In beide gevallen hebben we dat $A \cap (B \cup C) = A$ vermits $A \cup \emptyset = A = A \cup A$. Dus de elementen in de unie zijn precies die elementen die in tenminste een van de verzamelingen zitten.

Zoals gezegd betekent dit niet dat alle verzamelingen alleen bestaan uit elementen of dat elementen zowiezo bestaan in de eerste plaats. Bijvoorbeeld, veronderstel dat $\mathcal{S}$ bestaat uit $\emptyset, \{1\}, \{1,2\}$, dan is $\{1\}$ primitief, maar $\{1,2\}$ bestaat niet alleen uit elementen. Traditionele verzamelingenleer veronderstelt dat

$$B = \{\widehat{A} | \widehat{A} \in B\}$$

wat betekent dat een verzameling gelijk is aan de collectie van haar elementen. In dat geval hebben we dus net bewezen dat $\cap$ en $\cup$ de gewone operaties van doorsnede en unie zijn. De lezer kan dit nogal allemaal abstract vinden en zeggen "tja, kan ik dit allemaal niet zomaar onderstellen zonder al die afleidingen?". Het simpele antwoord is "neen"; wiskundigen zijn erg zuinig op hun aannames! Waarom een bijkomende zin in de grondwet schrijven wanneer die reeds een gevolg is van de voorgaande zinnen?! De vraag die je je kunt stellen is "jamaar, hoe verzin je al die stellingen?". Hier is het antwoord eenvoudig dat je jezelf altijd de gewenste situatie voor ogen moet houden en "zien" aan wat voor eigenschappen deze voldoet. Het is dan niet meer dan een "oefening" om te controleren of de abstracte theorie hier ook aan voldoet.

Dit zijn lang niet de enige spelregels van verzamelingenleer maar we zullen deze verder stelselmatig opbouwen naargelang de dingen waarover we willen praten meer complex worden. Wiskunde is nogmaals een primitieve taal waarin we de zaken zeer scherp definieren en elke "afwijking" kan dus meer of minder serieuze gevolgen hebben. Laten we even afwijken en fantaseren wat er allemaal kan gebeuren wanneer onze zakken geen spookzakken zijn. Een cruciaal verschil is onmiddelijk dat de doorsnede en unie van "verzamelingen" geen verzamelingen meer zijn vermits het aantal gaten moet meegeteld worden alsook de manier waarop zakken in elkaar gebouwd zitten. Dit leidt tot de operatie van het "in elkaar steken van n-zakken" Δ_n en de lezer merkt op dat we hiervoor de natuurlijke getallen nodig hebben. Deze zijn definieerbaar uit onze huidige axioma's en alsdusdanig kun je toch zeggen dat in een zekere zin verzamelingenleer de meest fundamentele is; we zullen een gelijkaardige situatie tegenkomen in de meetkunde waarbij de vlakke meetkunde de moeder van de (afleidbare) meetkundes blijft.

De natuurlijke getallen n definieren we symbolisch als $n = 1+1+1+1+\ldots+1$ aan de hand van het volgende voorschrift:

$$\begin{aligned} 0 &= \{\emptyset\} \\ n+1 &= \{n, \emptyset\}. \end{aligned}$$

Dus $1 = \{\{\emptyset\}, \emptyset\}$, $2 = \{\{\{\emptyset\}, \emptyset\}, \emptyset\}$ enzovoort. Ik heb de lezer gewaarschuwd dat symbolische notatie dikwijls het moeilijkste onderdeel van de wiskunde is, maar dit is hoe een computer de getallen zou kunnen inlezen. Nu definieren we op dezelfde manier $n+m$ aan de hand van het voorschrift $n+(m+1) = \{(n+m), \emptyset\}$

waarbij $n + 0 = 0 + n = n$ gesteld wordt. Hieruit tonen we dan aan dat $n+m = m+n$ voor elk natuurlijk getal m; dit is al zeker waar voor $m = 0$. Stel dat het ook waar is voor $m = k$, dan tonen we aan dat het geldt voor $m = k+1$; inderdaad, $n + (k + 1) = \{n + k, \emptyset\} = \{k + n, \emptyset\} = \{k + (1 + (n - 1)), \emptyset\} = \{(k+1)+(n-1), \emptyset\} = (k+1)+n$ waar we in de eerste stap de definitie gebruikt hebben, in de tweede stap de veronderstelling dat $k + n = n + k$ en finaal, in de derde stap, de associativiteit van $+$. We stellen dan dat $\mathbb{N}$ de verzameling is van alle natuurlijke getallen, wat een verzamelingen theorie definieert door het nemen van alle deelverzamelingen van $\mathbb{N}$.

De bewerking $+$ beeldt twee natuurlijke getallen af op twee natuurlijke getallen, is associatief en commutatief en heeft 0 als eenheids-element wat betekent dat $0+n = n+0 = n$. Voor elke n kunnen we een inverse $-n$ definieren die voldoet aan $n+(-n) = 0 = (-n)+n$ wat we ook noteren als $n-n = 0$; $n+(-m) = n-m$ is dan een natuurlijk getal als $n > m$ en min een natuurlijk getal als $n < m$. De verzameling van de natuurlijke getallen en hun inverse noemen we de gehele getallen en definieren we als $\mathbb{Z}$. $\mathbb{Z}, +$ noemen we dan een commutatieve groep vermits de bewerking $+$ inwendig is, associatief met eenheidselement en inverse.

Zoals voorheen vermeld begint men met het maken van een onderscheid tussen elementen en verzamelingen; men definieert $\in$ en $\emptyset$ en hieruit leidt men de eerste drie axioma's af van verzamelingenleer. De aanpak hier is algemener in de zin dat we een element definieren als een primitieve verzameling. Zermelo-Frankel verzamelingenleer bevat veel meer axioma's die te doen hebben met oneindigheid zoals het keuze axioma. We bouwen nu stelsmatig verder aan de axioma's van verzamelingenleer. Een vijfde axioma betreft het nemen van het verzameling theoretische verschil

$$B \setminus C = \{\widehat{A} | \widehat{A} \in B \wedge \widehat{A} \notin C\}$$

hetwelke we ook opnemen in onze lijst. In de meetkunde kunnen we niet alleen de unie nemen van twee lijnen of de doorsnede ervan maar ook het zogenaamde Cartesisch product, dat een vlak definieert. Gegeven twee verzamelingen B, C, we definieren het Cartesisch product $B \times C$ als de *verzameling* van alle koppels (x, y) zodat $x \in B$ and $y \in C$ hetgeen een zesde axioma is in $\mathcal{S}$ en aldus is onze verzamelingenleer gesloten met betrekking tot $\times$ waarbij

$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$

en

$$A \times (B \cup C) = (A \times B) \cup (A \times C)$$

en hetzelfde wanneer het tweede item vervangen wordt door het eerste. Het bestaan van Cartesische producten laat ons toe relaties te definieren waarbij een *relatie* R tussewn verzamelingen B en C een deelverzameling is van $B \times C$; in het geval $B = C$ kunnen we meerdere criteria opstellen. Bij notatie zeggen we dat xRy als en slechts als $(x, y) \in R$; we noemen R reflexief als xRx voor alle $x \in B$, symmetrisch als xRy impliceert dat yRx voor alle $x, y \in B$ en finaal transitief als xRy en yRz impliceren dat xRz. Een reflexieve, anti-symmetrische, transitieve relatie is een partiele orde genoteerd door $\prec$ of $\leq$, en een reflexieve, symmetrische en transitieve relatie een equivalentie gewoonlijk genoteerd door $\equiv$. Een equivalentie relatie moet je jezelf indenken als een veralgemening van het gelijkheidsteken; hier betreft het objecten met eenzelfde eigenschap. Je moet dan ook kunnen bewijzen dat een equivalentie relatie op een verzameling A deze doet uiteenvallen in zogenaamde equivalentie-klassen $\overline{x}$ waarbij

$$\overline{x} = \{y \in A | x \equiv y\}.$$

De lezer controleert dat $\overline{x} = \overline{y}$ als en slechts als $x \equiv y$ en de doorsnede $\overline{x} \cap \overline{y} = \emptyset$ anders.

We hebben net de natuurlijke getallen gedefinieerd aan de hand van de bewerking $+$; $\mathbb{N}$ draagt een natuurlijke partiele orde $\leq$ gedefinieerd door $n \leq n$ and $n \leq n+1$ en men neemt de *transitieve sluiting* van dat. Vanuit de natuurlijke getallen hebben we de gehele getallen $\mathbb{Z}$ gedefinieerd en de definitie van $\leq$ kan natuurlijk uitgebreid worden tot $\mathbb{Z}$. We construeren nu de rationale getallen vanuit $\mathbb{Z} \times \mathbb{N}_0$ door middel van de equivalentie relatie $(m, n) \equiv (m', n')$ als en slechts als er bestaat een $k, l \in \mathbb{N}_0$ zo dat $km = lm'$, $kn = ln'$ waar $\mathbb{N}_0 = \mathbb{N} \setminus \{0\}$. De rationale getallen zijn dan gedefinieerd als de equivalentie-klassen onder deze relatie.

De zes axioma's die zover besproken zijn, zijn veruit de belangrijkste van verzamelingenleer; de resterende twee hebben te maken met oneindigheid en zijn veelal toegevoegd om eigenschappen van de natuurlijke getallen te veralgemenen. Ons zevend axioma luidt als volgt: gegeven een verzameling D, de machtsverzameling 2^D van alle deelverzamelingen van D, is een verzameling en behoort tot $\mathcal{S}$. Dit axioma leidt tot de constructie van de ordinale getallen door Cantor. De definitie van een Cartesisch product wordt uitgebreid over zogenaamde "index" verzamelingen; hiervoor hebben we een partiele orde nodig $\prec$. Een index verzameling I is een verzameling uitgerust met een partiele orde $\prec$ zodanig dat voor elke $x, y \in I$ er een $z \in I$ bestaat zodat $x, y \prec z$. Deze voorwaarde is nodig en voldoende indien we unieke limieten willen nemen; als het niet geldig is, dan kunnen verschillende sublimieten bestaan; dus we noteren door

$$\times_{i \in I} A_i = \{(x_i)_{i \in I} | x_i \in A_i\}$$

waar de I-koppels partieel geordend zijn door $\prec$. Finaal, als achtste axioma hebben we het zo-genaamde keuze axioma dat als volgt geformuleerd kan worden: gegeven verzamelingen A_i, $i \in I$ een index verzameling, dan bestaat er een verzameling A gevormd uit elementen in $\cup_{i \in I} A_i$ zodat uit elke A_i precies één x_i tot A behoort. Een andere formulering is dat er een functie f bestaat van I naar $\cup_{i \in I} A_i$ zodat $f(i) \in A_i$. Dus je kunt een verzameling vormen door uit elke verzameling een element te kiezen. Dit keuze axioma heeft erg veel gevolgen en sommige wiskundigen verwerpen het dan ook; het lijkt intuitief correct maar het is een bepaalde manier om het oneindige te bevatten. Ik heb het reeds een aantal malen gezegd: wiskunde op zich is niet bewijsbaar, het is een taal en we moeten enkele "grammaticale" keuzes maken. De lezer moet erg goed nadenken over deze acht spelregels en indachtig zijn dat commutativiteit, associativiteit alsook het vormen van een machtsverzameling de meest eenvoudige of symmetrische spelregels zijn. Elk voorbeeld dat niet aan deze regels voldoet is geconstrueerd vanuit deze ideale situatie; bijvoorbeeld, we zullen later niet commutatieve of associatieve bewerkingen definieren en zien hoe deze vanuit de commutatieve situatie gebouwd kunnen worden. Dit leidt tot niet commutatieve groepen, quantum groepen enzovoort. Dit doet ons een beetje denken aan de Egyptische bouwkunst en hieruit voortvloeiend de Romeinse en Franse symmetrische bouwstijlen: super eenvoudig, prachtig en logisch. Alles wat meer complex is moet vertrekken van dezelfde spelregels. Nogmaals, het is geen schande dat je dit kleine hoofdstukje eens een paar keer moet overpeinzen en ik nodig je uit in te zien dat de symbolische notatie van de natuurlijke getallen de enige mogelijke is.

Chapter 3

Geavanceerde getallenstelsels.

Dit hoofdstukje is al wat praktischer dan het voorgaande; waar we voordien abstracte regels bestudeerd hebben gaan we deze nu aanwenden om concrete berekeningen te leren maken. We beginnen terug met voorstellingen of representaties van de gehele en rationale getallen en later breiden we deze uit naar de reeele getallen toe. Hierna bestuderen we de complexe getallen en reeele quaternionen om finaal te eindigen met algemene Clifford algebra's en octonionen; afgezien van deze laatste zijn ze allen belangrijk in de natuurkunde en hogere rekenkunde. We hebben tot hiertoe $\mathbb{Z}, +$ bestudeerd en vooraleer we overgaan naar hogere getallenstelsels, laat ons even abstracte regels opstellen: $\mathbb{Z}, +$ voldoet aan

- $+$ is inwendig wat betekent dat de som van twee gehele getallen terug een geheel getal is,
- $+$ is associatief wat betekent dat $(n+m)+k = n+(m+k)$, met andere woorden de som van een reeks getallen hangt alleen af van de reeks en niet van de manier waarop je haar opsplitst in paren,
- er bestaat een uniek neutraal element 0 zodat $0+n = n+0 = n$,
- gegeven het neutraal element 0, we hebben dat voor elk element n er een unieke inverse $-n$ bestaat zodat $n+(-n) = (-n)+n = 0$,
- de som is commutatief, wat betekent dat $n+m = m+n$.

Meer algemeen, een vezameling G met bewerking $\star : G \times G \to G$ die aan deze vijf eigenschappen voldoet noemen we een commutatieve groep. Je moet goed begrijpen dat deze regels gegeven zijn in volgorde van belangrijkheid: de tegengestelde of inverse kan niet gedefinieerd worden zonder het neutraal element en de definitie van het neutraal element steunt op de inwendigheid van de optelling. Uiteraard kan je associativiteit laten vallen alsook het bestaan van een unieke inverse (later zullen we hiervan voorbeelden zien); commutativiteit is nog het makkelijkst om te laten vallen en ook hier bestuderen we voorbeelden van. De gehele, rationale, reeele en complexe getallen voldoen allemaal aan deze spelregels voor de optelling; we bestuderen nu de vermenigvuldiging. Voor de natuurlijke getallen definieren we de vermenigvuldiging aan de hand van $n.1 = 1.n = n$ en $n.(m + 1) = n.m + n = (m + 1).n$; we bewijzen nu even dat hieruit volgt dat

$$n.(m + k) = n.m + n.k = (m + k).n$$

voor alle $n, m, k \in \mathbb{N}$. Duidelijk is

$$n.(m+k) = n.((m+k-1)+1) = n.(m+k-1)+n = n.(m+k-2)+n.2 = \ldots = n.m+n.k$$

door het herhaaldelijk toepassen van de elementaire regel. Je toont nu aan dat dezelfde eigenschappen gelden voor alle $n, m, k \in \mathbb{Z}$ door te stellen dat $0.n = n.0 = 0$. 1 is per definitie het eenheidselement voor de vermenigvuldiging, maar elk geheel getal n verschillend van 1 heeft geen inverse in $\mathbb{Z}$ in de zin dat $m.n = n.m = 1$. De vermenigvuldiging is dus inwendig, associatief, commutatief en heeft een eenheidselement. $\mathbb{Z}, +, .$ met deze eigenschappen noemen we een commutatieve ring; meer precies noemen we een verzameling G uitgerust met twee bewerkingen $+, .$ waarbij $G, +$ een commutatieve groep is en . inwendig, associatief is met eenheidselement 1 een ring indien bovendien

$$g.(u + v) = g.u + g.v, \ (u + v).g = u.g + v.g$$

een eigenschap die we de distribuviteit van de vermenigvuldiging ten opzichte van de optelling noemen. Noteer dat deze regel precies dezelfde is dan de regel van de-Morgan uit de verzamelingenleer en logica. Dus + mag je lezen als "of" en . als "en". Je kunt deze abstracties op het eerste zicht als overdreven of onnatuurlijk ervaren; het is niet zo omdat de gehele getallen hieraan voldoen dat dit noodzakelijkerwijs een soort "heilige" regel moet zijn. De lezer zal beter apprecieren dat deze regels wel degelijk belangrijk zin in de zin dat veruit de meeste getallenstelsels die we hieruit kunnen construeren aan die regels voldoen. Vooraleer we overstappen naar de rationale getallen is het goed om na te denken over voorstellingswijzen of representaties van de natuurlijke en gehele getallen; de arabieren hebben hierin het pleit gewonnen. Ze waren slim genoeg om te werken met zogenaamde machten n^m waarbij n, m twee natuurlijke getallen zijn. Per definitie stellen we dat $n^0 = 1$ en $n^{m+1} = n.n^m$ dus $n^1 = n, n^2 = n.n, n^3 = n.n.n$; in het algemeen leest dit dus dat n m maal met zichzelf vermenigvuldigd wordt. Terzake nu, de arabieren hadden begrepen dat je elk natuurlijk getal m op een unieke manier kunt schrijven als

$$m = \sum_k c_k.n^k$$

waar c_k een natuurlijk getal is dat ligt tussen 0 en $n-1$; het bewijs hiervan is erg simpel en een gevolg van de Chinese reststelling. Deze zegt dat elk natuurlijk getal $m = k.n+r$ met r tussen 0 en $n-1$ wat gewoon betekent dat elk getal ligt tussen twee veelvouden van n; een uitspraak die simpelweg waar is. De stelling van de arabieren volgt dan gewoonweg uit

$$m = k_1.n + c_0 = (k_2.n + c_1).n + c_0 = k_2.n^2 + c_1.n + c_0 = \ldots = \sum_{k=0}^{l} c_k.n^k$$

waarbij deze laatste som eindig is omdat op een gegeven moment $k_{l+1} = 0$. Dus een getal kun je voorstellen als $c_l c_{l-1} \ldots c_1 c_0$ en deze code hangt dus af van het zogenaamde basisgetal n. Indien het basisgetal 2 is spreken we over binaire getallen, deze worden gebruikt in de logica en schakelingen in electronische apparaten zoals computers en microgolf ovens. Je moet de volgende oefeningen maken: toon aan dat het getal $3 = 2 + 1$ binair leest als 11 en het getal $7 = 2^2 + 2 + 1$ als 111, speel nog wat verder met binaire codes. De getallen-notatie die het gehaald heeft is de tiendelige, misschien omdat we tien vingers of tien tenen hebben; de basisgetallen zijn dus $0 = \emptyset, 1 = 1, 2 = 1 + 1, 3 = 1 + 1 + 1, 4, 5, 6, 7, 8, 9$. Dus we noteren getallen zoals 923 enzovoort; nu, waarom is zulke notatie extreem handig wat betreft de optelling alsook de vermenigvuldiging? De reden is dat $0 \leq c_k + d_k \leq 2n - 2$ alsook $0 < c_k c_l < n^2$ wat leidt tot de optellings en vermenigvuldigingsregels die je geleerd hebt in de lagere school; een handig begrip hierbij is dat van modulo " mod " waarbij "$n \mod m$" de rest bij deling van n door m. Dit zover de optelling en vermenigvuldiging van natuurlijke getallen; deze regels kunnen triviaal uitgebreid worden naar de gehele getallen toe en je moet hierop enkele oefeningen maken. Dit zijn typisch dingen die je doet in de lagere school; we eindigen deze sectie met de volgende concepten. Een natuurlijk getal $n > 1$ wordt een priemgetal genoemd indien n en 1 haar enige delers zijn wat betekent dat als $n \mod q = 0$ en $0 \leq q \leq n$ dan is $q = 1, n$. Controleer nu dat $2, 3, 5, 7, 11, 13, \ldots$ de eerste 6 priemgetallen zijn. De eerste oefening die je moet maken is dat elk getal kan geschreven worden als een uniek product van priemgetallen, deze ontbinding noemen we de priemontbinding. Bijvoorbeeld, de priemontbindingen van $6, 12, 21$ zijn $6 = 2.3, 12 = 2^2.3, 21 = 3.7$; bereken nu de priemontbindingen van $37, 41, 56$.

We hebben in het vorig hoofdstuk de rationale getallen al abstract gedefinieerd maar hier gaan we verder in op de materie. We vertrokken van $\mathbb{Z} \times \mathbb{N}_0$ en definieerden twee koppels (m, n) en (m', n') als equivalent indien er niet nul natuurlijk getallen k, l bestaan zodat $km' = lm, kn' = ln$. Dit is echter een niet zo handige manier om de zaken voor te stellen en daarom heeft men de notatie

$$\frac{m}{n}$$

geintroduceerd, de equivalentie komt dan neer op de schrappingsregel dat twee gemeenschappelijke factoren boven en onder de streep weggelaten kunnen worden. Meer precies,

$$\frac{km}{kn} = \frac{m}{n}$$

en men definieert de optelling en vermenigvuldiging als volgt

$$\frac{m_1}{n_1} + \frac{m_2}{n_2} = \frac{m_1 n_2 + m_2 n_1}{n_1 n_2}$$
$$\frac{m_1}{n_1}\frac{m_2}{n_2} = \frac{m_1 m_2}{n_1 n_2}.$$

Hieruit leidt je af dat het eenheidselement voor de optelling kan voorgesteld worden als $\frac{0}{n}$ en het eenheidselement voor de vermenigvuldiging als $\frac{n}{n}$ waarbij $n \in \mathbb{N}_0$. Je moet nu controleren, als oefening, dat de optelling en vermenigvuldiging associatief en commutatief zijn en dat aan de distribuviteitsregel voldaan is. Dus $\mathbb{Q}, +, .$ is een ring; de inverse voor de optelling is duidelijk $\frac{-m}{n}$ en je moet een inverse voor de vermenigvuldiging vinden voor elk niet nul element. Met deze bijkomende informatie noemen we $\mathbb{Q}, +, .$ een veld; zover de rekenkundige eigenschappen van de rationale getallen. Nu komen we tot enkele belangrijke inzichten: eerst en vooral merk je op dat de gehele getallen $\mathbb{Z}$ voorgesteld kunnen worden als een rationaal getal $m := \frac{m}{1}$ en dat onder die voorstelling de eigenschappen van optelling en vermenigvuldiging behouden blijven. Meer precies,

$$m + n := \frac{m+n}{1} = \frac{m}{1} + \frac{n}{1}, mn := \frac{mn}{1} = \frac{m}{1}\frac{n}{1}$$

we vertalen dit door te zeggen dat de afbeelding

$$\mathbb{Z} \to \mathbb{Q} : m \to \frac{m}{1}$$

een homomorfisme is. Je moet nu aantonen dat $\mathbb{Q}$ het kleinste veld is dat de gehele getallen omvat (hint: definieer inversen voor de vermenigvuldiging van gehele getallen en neem willekeurige producten). Een tweede belangrijke eigenschap betreft de uitbreiding van de orde relatie $\leq$; we zeggen dat

$$\frac{m}{n} \leq \frac{m'}{n'}$$

als en slechts als $mn' \leq m'n$. Je controleert dat deze definitie geldig is over de ganse equivalentie klasse waarmee ik wil zeggen dat indien $\frac{r}{s} = \frac{m}{n}$ en $\frac{r'}{s'} = \frac{m'}{n'}$ dan is $rs' \leq r's$ als en slechts als $mn' \leq m'n$. Als moeilijke oefening probeer je aan te tonen dat je de rationale getallen niet kunt nummeren volgens orde; met andere woorden, tussen elke twee rationale getallen kun je een ander rationaal getal vinden (hint: neem het midddelpunt).

Deze laatste eigenschap leidt ons tot de constructie van de reeele getallen $\mathbb{R}$; er bestaan verschillende manieren om deze te introduceren, bijvoorbeeld als de metrische sluiting van $\mathbb{Q}$ in de natuurlijke metriek. Vermits we topologie en metrieken pas in het volgende hoofdstuk zullen bestuderen gaan we hier anders te werk; we gebruiken gewoon de orde relatie. Laten we eerst aantonen dat de rationale getallen gaten bevatten; bijvoorbeeld, definieer het getal $\sqrt{2}$ als het uniek positief getal zodat $\sqrt{2}^2 = 2$. We tonen aan dat dit niet rationeel is. Stel dat $\sqrt{2} = \frac{m}{n}$ dan is $2n^2 = m^2$. Nu kun je zien dat dit onmogelijk is omdat 2 een oneven aantal keer voorkomt in de priemontbinding van het linkerlid maar een even aantal in het rechterlid. Dus $\sqrt{2}$ behoort niet tot de rationale getallen en aldus bevat deze gaten; dus hier komt het eerste nut naar voren van de priem-ontbinding. Noteer dat de priemgetallen zijn ten opzichte van de vermenigvuldiging wat de primitieve verzamelingen zijn ten opzichte van de doorsnede; dus het is helemaal niet verwonderlijk dat deze zo belangerijk zijn. De vraag is nu dus hoe je een getal als $\sqrt{2}$ construeert! Dus om even de logica tot hiertoe te herhalen: (a) de natuurlijke getallen zijn geconstrueerd vanuit de optelling (b) de gehele getallen volgenden door inversen ten opzichte van de optelling toe te voegen (c) de rationale getallen volgen uit de gehele getallen door inversen ten opzichte van de vermenigvuldiging te definieren en finaal (d) volgen de reeele getallen uit een compleetheidseigenschap.

De vervollediging definieren we aan de Dedekind procedure: hier neem je stijgende rijen $q_n \leq q_{n+1}$ van rationale getallen q_n die allemaal kleiner zijn dan een bepaald rationaal getal q. We definieren weeral dat twee zulke rijen $(q_n)_{n\in\mathbb{N}}$, $(p_n)_{n\in\mathbb{N}}$ equivalent zijn indien voor elke p_n er een q_m bestaat zodat $q_m \geq p_n$ en omgekeerd. Controleer terug dat deze relatie tussen stijgende rijen een equivalentierelatie is; evenzo kun je werken met dalende rijen die langs onder begrensd zijn. $\mathbb{R}, +, .$ is dan gedefinieerd aan de hand van deze equivalentie-klassen en je moet aantonen dat $+$ en $.$ goed gedefinieerd zijn en dat $\mathbb{R}, +, .$ een veld is. Laten we eens bij wijze van voorbeeld zulke rij construeren voor $\sqrt{2}$:

$$\sqrt{2} = \sqrt{4-2} = 2\sqrt{1 - \frac{1}{2}}$$

en

$$\sqrt{1+z} = 1 + \frac{1}{2}z + \sum_{n=2}^{\infty} (-1)^{n+1} \frac{1.3\ldots(2n-3)}{2^n} z^n$$

en daarom is $\sqrt{2}$ benaderd door de dalende rij

$$q_m = 1 - \frac{1}{4} - \sum_{n=2}^{m} \frac{1.3\ldots(2n-3)}{4^n}$$

van rationale getallen. Uiteraard ken je die formule voor de wortel nog niet en we zullen deze later in dit boek bewijzen. Dus de reeele getallen zijn een compleet en totaal geordend veld dat bovendien aan de Archimedische eigenschap voldoet: het is te zeggen

$$ra \leq rb$$

als en slechts als $a \leq b$ gegeven dat $r > 0$. Nu zou je zeggen dat we klaar zijn met het definieren van getallenstelsels mochten we niet op de volgende uitspraak botsen. Een veelterm is een functie in een reeele veranderlijke x van de vorm

$$P(x) = \sum_{k=0}^{n} a_k x^k$$

waarbij $a_k \in \mathbb{R}$; een wortel of nulpunt van P is een reeel getal c zodat $P(c) = 0$. Indien c een wortel is van $P(x)$ kan men $P(x)$ schrijven als $P(x) = (x - c)Q(x)$ met $Q(x) = \sum_{k=0}^{n-1} b_k x^k$. Toon dit aan! Indien P precies n nulpunten heeft kan deze geschreven worden als

$$P(x) = a_n(x - c_1) \dots (x - c_n).$$

In dat geval noemen we $P(x)$ volledig factoriseerbaar en $(x - c)$ een factor. Nu kan men makkelijk inzien dat niet elke veelterm van die aard factorizeerbaar is: namelijk

$$P(x) = x^2 + 1$$

is altijd strict positief en heeft geen nulpunten. Om deze toch te factorizeren voegt men een getal i toe zodat $i^2 = -1$. In dat geval is

$$P(x) = (x - i)(x + i)$$

wat erop wijst dat we de reeele getallen minstens moeten uitbreiden met i; met andere woorden, definieer

$$z = a + bi$$

met $a, b \in \mathbb{R}$ en

$$\mathbb{C} = \{z | z = a + bi; a, b \in \mathbb{R}\}$$

de verzameling van de complexe getallen. Beschouw vervolgens veeltermen $P(z) = \sum_{k=0}^{n} a_k z^k$ met $a_k \in \mathbb{C}$ dan kan men bewijzen dat $P(z)$ factorizeerbaar is oftewel $P(z)$ heeft n nulpunten. Je merkt op dat het voldoende is te bewijzen dat $P(z)$ minstens een nulpunt c heeft vermits dan $P(z) = (x-c)Q(z)$ en $Q(z)$ is van graad $n - 1$ en heeft terug een nulpunt enzovoorts. Een bewijs van deze stelling vergt hogere wiskunde en we komen daarop terug in hoofdstuk 13. We rusten $\mathbb{C}$ terug uit met een optelling alsook vermenigvuldiging:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

en

$$(a + bi)(c + di) = (ac - bd) + (bc + ad)i.$$

Je toont aan dat $\mathbb{C}, +, .$ een ring is; definieer nu de complex toegevoegde als

$$\overline{z} = a - bi$$

dan is

$$\overline{z}z = z\overline{z} = a^2 + b^2$$

een getal dat alleen nul is indien $a = b = 0$. Dus het getal $\frac{\overline{z}}{z\overline{z}}$ is de inverse voor z voor de vermenigvuldiging en aldus is $\mathbb{C}, +, .$ een veld. Het veld is echter niet Archimedisch gegeven dat de natuurlijke orde gegeven is door $a + bi \leq c + di$ als en slechts als $a < c$ of $a = c$ en $b \leq d$. In dat geval is $1 + 5i > 0$ maar $(1 + 5i)^2 = -24 + 10i < 0$ wat in strijd is met de Archimedische eigenschap. We gaan dieper in op de complexe getallen in hetvolgende hoofdstuk; noteer voor nu dat

$$\overline{z+z'} = \overline{z} + \overline{z'}$$

en

$$\overline{zz'} = \overline{z}\overline{z'}$$

twee eigenschappen die je moet aantonen. Als toepassing van de complexe getallen gaan we nulpunten berekenen van reeele veeltermen $P(x)$ van eerste en tweede graad. Voor een eerste graadsfunctie $ax+b$ is het nulpunt gegeven door $-\frac{b}{a}$ terwijl voor een tweedegraadsfunctie we de nulpunten moeten zoeken van

$$ax^2 + bx + c.$$

Deze laatste kun je herschrijven als

$$a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c$$

en dit is nul als en slechts als

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Indien $b^2 - 4ac \geq 0$ dan is

$$x_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

terwijl in het andere geval

$$x_{\pm} = \frac{-b \pm i\sqrt{4ac - b^2}}{2a}$$

hetgeen expliciet aantoont dat reeele veeltermen van de tweede graad factorizeerbaar zijn over $\mathbb{C}$.

Nu zou je denken dat de meest interessante getallenstelsels bestudeerd zijn doch er is een erg belangrijke klasse van veralgemeningen van de complexe getallen: we bestuderen eerst het meest eenvoudige geval van de reeele quaternionen $\mathbb{RQ}$. Dit getallenstelsel heeft een nauw verwantschap met meetkunde in de drie-dimensionale ruimte, iets waar we later zullen op terug komen. Voorlopig bestuderen we de algebraische eigenschappen van dit getallenstelsel en laten de meetkundige motivatie voor later. $\mathbb{RQ}$ is gegenereerd door twee imaginaire eenheden i, j zodat $i^2 = j^2 = -1$ en $ij + ji = 0, ij = k$; dus, $k^2 = ijij = -i^2j^2 = -1$ en $jk = i, ki = j$. Een algemeen quaternion q is dus van de vorm $q = a + bi + cj + dk$ waarbij $a, b, c, d \in \mathbb{R}$. We definieren terug de quaternion toegevoegde aan de hand van

$$\overline{q} = a - bi - cj - dk$$

en aldus is $q\overline{q} = \overline{q}q = a^2 + b^2 + c^2 + d^2$. De inverse van $q \neq 0$ is dus $\frac{\overline{q}}{q\overline{q}}$ en aldus is $\mathbb{R}\mathbb{Q}, +, .$ een niet commutatieve ring waarbij elk niet nul element een multiplicatieve inverse heeft. Zoals gezegd komt de meetkundige interpretatie later.

Er is een natuurlijke uitbreiding van de quaternionen naar de zogezegde reeele Clifford algebra's; deze zijn gegenereerd door elementen e_i waarvoor geldt dat

$$e_i e_j + e_j e_i = \alpha_{ij}$$

met $\alpha_{ij} = 0$ als $i \neq j$ en $\alpha_{ii} = \pm 2$. We noteren dit met $\mathrm{Cl}(p, q)$ waarbij p het aantal plusjes is en q het aantal minnetjes. De quaternionen zijn dan $\mathrm{Cl}(0, 2)$; men kan aantonen dat met uitzondering van $\mathrm{Cl}(0, 1)$ en $\mathrm{Cl}(0, 2)$, dus de complexe getallen en quaternionen, dat in elk van deze ringen een element kan gevonden worden dat geen inverse heeft voor de vermenigvuldiging. Dit bemoeilijkt een heel aantal zaken maar bijvoorbeeld $\mathrm{Cl}(1, 3)$ wordt veelvuldig gebruikt in de fysica. Tot slot van dit hoofdstuk toon je aan dat elk element van de Clifford algebra geschreven kan worden als

$$q = a + \sum_{n=1}^{p+q} a_{j_1 \ldots j_n} e_{j_1} \ldots e_{j_n}$$

waarbij $j_k \neq j_l$ als $k \neq l$. We zullen later nog meer eigenschappen van Clifford getallen bestuderen maar dit volstaat voor nu. Merk op dat de Clifford getallen niet commutatief zijn voor de vermenigvuldiging en dus meer dan een dimensie vereisen om ze meetkundig voor te stellen; het vermeetkundigen van algebra is een erg belangrijke tak van de wiskunde. Tot slot zijn er ook nog de octonionen, deze vormen een getallenstelsel waarbij de vermenigvuldiging niet associatief noch commutatief is. De geinteresseerde lezer kan hierover alles nalezen in de literatuur.

Zover een eerste inleiding tot getallenleer: alles kan nogal abstract over komen tot hiertoe maar ik heb reeds talloze malen benadrukt dat het neerpennen van een idee in symbolentaal vaak het moeilijkste deel is van wiskunde en dat er veel verschillende manieren bestaan om dit te doen. Finaal een beetje notatie: we hebben reeds gezegd dat de inverse van een getal q ten opzichte van de som gegeven is door $-q$ terwijl dit voor de vermenigvuldiging gelijk is aan q^{-1}. Als je voorheen totaal geen ervaring hebt gehad met deze materie, toon dan aan dat

- $(\frac{5}{4})^{-1} = \frac{4}{5}$,
- $(1 + i)^{-1} = \frac{1-i}{2}$,
- $2^3 = 8$,
- $(\frac{1}{3})^2 = \frac{1}{9}$,
- $\frac{1}{i} = -i$,
- $(1 + 2i)^2 = -3 + 4i$.

Chapter 4

Topologie.

Topologie moet je je indenken als een verfijning van verzamelingenleer; het is te zeggen, we beperken ons tot speciale verzamelingen de zogenaamde open verzamelingen. In de natuur is een open verzameling een abstractie, een denkbeeldig iets wat niet helemaal bestaat. Een open verzameling moet je je indenken als iets met "volume": bijvoorbeeld, op een rechte is de verzameling van alle reeele getallen tussen twee elementen $(a, b) = \{x|a < x < b\}$ open en de natuurlijke lengte hiervan is $b - a$. Een punt is een voorbeeld van een gesloten verzameling maar dit heeft volume of lengte nul. Laten we even verder werken met de open intervallen (a, b): de doorsnede van twee open intervallen is terug een open interval en de unie van open intervallen kan geschreven worden als de disjuncte unie van open intervallen, waarbij disjunct betekent dat ze een lege doorsnede hebben. Kortweg, gegeven een verzameling D, we noemen een verzameling $\tau(D)$ van deelverzamelingen van D een topologie indien

- $\emptyset \in \tau$,
- $A, B \in \tau$ impliceert dat $A \cap B \in \tau$,
- $A_i \in \tau$ impliceert dat $\cup_{i \in I} A_i \in \tau$ voor elke index set I.

Ik benadruk opnieuw dat deze definitie afhangt van de commutativiteit en associativiteit van de doorsnede en unie en dat het daarom mogelijk is een niet commutatieve topologie te definieren door een of beide van de bewerkingen niet commutatief te maken. Dit is echter volledig onbekend gebied en in dit korte hoofdstuk bestuderen we "klassieke" topologieen. Intuitief kun je deze definierende eigenschappen begrijpen als de voorwaarde waaraan een verzameling landkaarten moet voldoen: typisch overlappen zulke kaarten en al wat we vragen is dat de doorsnede ervan terug een kaart is en dat je er willekeurig veel kan bijeen gooien. Er zijn speciale deelverzamelingen $E \subseteq D$ zodat

- *gesloten* als en slechts als $E^c := D \setminus E \in \tau(D)$,

- *compact* als en slechts als voor elke bedekking O_α van E met behulp van open verzamelingen $O_\alpha \in \tau(D)$ er een eindige deelbedekking $O_i; i = 1 \ldots n$ bestaat zodat $E \subseteq \cup_{i=1}^n O_i$.

Dus de compacte verzamelingen zijn deze die altijd bedekt worden door een eindig aantal kaarten zoals bijvoorbeeld een wereldbol: ongeacht hoe klein je ook je kaarten maakt, de bol is bedekt door een eindig aantal van hen. Gegeven een punt $p \in D$, we noemen O een open omgeving van p als en slechts als $p \in O$. Gegeven een punt p, een basis van open omgevingen is gegeven door een aftelbare collectie van open omgevingen O_i van p, zodat voor elke open V die p omvat geldt dat er een index i bestaat zodat $O_i \subseteq V$. Men zou bovendien kunnen eisen dat $O_{i+1} \subseteq O_i$ maar in het vervolg doen we dat niet. Betreffende gesloten verzamelingen X, Y moet je nu controleren dat: (a) $\emptyset, D$ zijn gesloten (b) $X \cup Y$ is gesloten (c) $\cap_{i \in I} X_i$ is gesloten als alle X_i dat zijn. Vezamelingen zoals $\emptyset, D$ welke open en gesloten zijn noemen we clopen. Gegeven een verzameling $B \subseteq D$, de doorsnede van alle gesloten verzamelingen X die B bevatten is gesloten en noemen we de sluiting van B wat we noteren als $\overline{B}$. De sluiting van een verzameling is daarom de kleinste gesloten verzameling die deze verzameling zelf bevat. Met andere woorden, je voegt dus elementen of punten toe die limieten zijn van elementen in B. Meer concreet, we noemen x een limietpunt van een rij $(x_i)_{i \in I}$ als en slechts als voor elke open omgeving $\mathcal{O}$ van x geldt dat er een j bestaat zodat $\forall j \prec i$ geldt dat $x_i \in \mathcal{O}$. Nu toon je aan, gebruik makend van de eigenschappen van een index set dat indien y een ander limiet punt is dat dan de open omgevingen van x en y samenvallen. Dit motiveert de volgende definitie: een topologie is Hausdorff als en slechts als elke twee punten x en y twee disjuncte open omgevingen hebben. Het is te zeggen $x \in \mathcal{O}, y \in \mathcal{V}$ en $\mathcal{O} \cap \mathcal{V} = \emptyset$. Voor Hausdorff topologieen geldt dan dat de limiet van een rij uniek is. We hebben nu het volgende resultaat voor topologieen met een aftelbare basis: een verzameling is gesloten als en slechts als ze al haar limietpunten bevat. We bewijzen dit even: veronderstel dat B gesloten is, en zij $(x_i)_{i \in I}$ een rij in B met limiet punt $x \in D$, dan is $x \in B$ anders bestaat er een open omgeving B^c van x die disjunct is met $(x_i)_{i \in I}$, wat tegen de definitie is van een limiet punt. Omgekeerd, veronderstel dat elk limietpunt van B tot B behoort, dan is B gesloten; indien dit laatste niet zo is, vinden we een $x \in \overline{B} \setminus B$ zodat voor elke basis-open omgeving $\mathcal{O}_n$ van x we een element $x_n \in B \cap \mathcal{O}_n$ vinden en aldus is x een limietpunt van $(x_n)_{n \in \mathbb{N}} \in B$ en aldus behoort het tot B wat een tegenstrijdigheid oplevert. Later geven we een voorbeeld van een compacte verzameling in een niet-Hausdorff topologie met een rij waarvan geen enkele deelrij een limietpunt heeft (indien je zelf wil zoeken, denk in een oneindig aantal dimensies). We gaan nog verdere karakteriseringen van compactheid bedenken in de zogenaamde metrische topologieen, dat zijn diegene die bepaald worden door een afstandsfunctie d.

Zover lijkt de behandeling van topologie erg abstract en weinig nuttig, je kunt zowat elke topologie bedenken waarmee je wil werken. Onze fysische werkelijkheid lijkt bij zeer goede benadering specialistischer gegeven dat we over afstanden en bollen spreken, zoals bijvoorbeeld de cirkel met straal van 10 kilometer rond Brussel gemeten vanaf manneken pis in vogelvlucht uiteraard. Op aarde gaat dit alleen "mis" wanneer je de halve omtrek bewandeld hebt; een grotere wandeling zou economischer geweest zijn mocht je oorspronkelijk in tegengestelde richting vertrokken zijn. Dus op grote afstanden kun je "problemen" krijgen en tegenwoordig denken we in de natuurkunde dat zulke zaken zich ook kunnen voordoen op zeer kleine afstanden; veel kleiner dan de typische schaal van een atoom. Kortweg, een afstandsfunctie $d : X \times X \to \mathbb{R}^+$ op een verzameling X voldoet aan

- $d(x, y) = 0$ als en slechts als $x = y$,

- $d(x, y) = d(y, x)$ voor elke $x, y \in X$,

- $d(x, z) \leq d(x, y) + d(y, z)$ de zogenaamde driehoeksongelijkheid.

Een afstandsfunctie definieert een Hausdorff topologie met aftelbare basis door middel van de open ballen

$$B(x, \epsilon) = \{z | d(x, z) < \epsilon\}$$

een aftelbare basis is dan gegeven door $B(x, \frac{1}{n})$ waarbij $n \in \mathbb{N}_0$. Twee punten x, y gescheiden door een afstand $d(x, y) > 2\epsilon$ kunnen omgeven worden door twee disjuncte bollen $B(x, \epsilon), B(y, \epsilon)$ respectievelijk. Dit is nog altijd een redelijk abstracte voorstelling van zaken gegeven dat we ook nog over hoeken en loodrechte stand willen spreken en dit niet zo eenvoudig is in dit formalisme. Met andere woorden, we hebben nog een verdere specializatie nodig die verder reikt dan de afstands functie alleen. Nochtans kun je op deze manier een redelijk grote rijkdom aan resultaten bekomen door abstract te redeneren met deze drie axioma's waaraan d voldoet. Zulke veralgemening huist erin te stellen dat het afstandsbegrip een "locale" oorsprong heeft; het is te zeggen dat de afstand tussen twee punten in willekeurig kleine stukken gekapt kan worden. Dit leidt tot het begrip van pad-metriek: d is een pad metriek als en slechts als voor elke twee punten x, y er bestaat een z zodat

$$d(x, z) = d(y, z) = \frac{d(x, y)}{2}$$

met andere woorden, elke twee punten hebben een middelpunt. We zullen hier later nog een betere voorstelling van geven.

Laten we nu eens een equivalentie-relatie tussen twee topologische ruimten bestuderen; met andere woorden, wanneer zijn twee topologische ruimten hetzelfde? Om dat te bepalen gaan we nu vergelijkings-middelen bestuderen, het is te zeggen "topologische" afbeeldingen tussen twee topologische ruimten X, Y. Een afbeelding $f : X \to Y$ wordt gedefinieerd aan de hand van een deelverzameling F van het Cartesisch product $X \times Y$; F voldoet aan de eigenschap dat voor elke $x \in X$ er precies een $y \in Y$ bestaat zodat $(x, y) \in F$. y wordt dan genoteerd als $f(x)$ en F is de grafiek van f; in mensentaal betekent dit dat elk element uit X precies een beeld heeft in Y. Betreffende afbeeldingen formuleren we nog de volgende extreme eigenschappen: (a) f is injectief indien $f(x) = f(x')$ impliceert dat $x = x'$ of elke x heeft een verschillend beeld (b) f is surjectief indien voor elke $y \in Y$ er een $x \in X$ bestaat zodat $f(x) = y$ of met andere woorden elke potentiele beeldwaarde is ook effectief gerealiseerd. Finaal noemen we f bijectief als en slects als ze injectief en surjectief is; bijectieve afbeeldingen zijn equivalenties tussen verzamelingen zoals we nu zullen zien. Laat $f : X \to Y$ en $g : Y \to Z$ dan is $g \circ f : X \to Z : x \to g(f(x))$ de samenstelling van deze twee afbeeldingen. Toon nu aan dat $g \circ f$ injectief is als en slechts als g dat is op $f(X)$ en f dat is op X. Toon ook aan dat $g \circ f$ surjectief is als en slechts als g dat is op $f(X)$; finaal, toon aan dat $g \circ f$ een bijectie is als g, f dat zijn. Indien $f : X \to Y$ een bijectie is kunnen we uniek $f^{-1} : Y \to X$ definieren aan de hand van

$$f^{-1}(f(x)) = x$$

of $f^{-1} \circ f = \mathrm{id}_X$ waarbij id_X de eenheids afbeelding is op X. Leidt hieruit af dat

$$f \circ f^{-1} = \mathrm{id}_Y$$

door gebruik te maken van de surjectiviteit van f. Finaal, toon aan dat f^{-1} ook een bijectie is; we zeggen dus dat X en Y equivalent zijn als en slechts als er een bijectie van X naar Y bestaat; gebruik makend van de voorgaande resultaten toon je aan dat deze relatie inderdaad reflexief, symmetrisch en transitief is. Nu komen we terug op topologische equivalenties $f : X \to Y$; een afbeelding f noemen we continue als en slechts als de inverse van elke open verzameling O in Y, $f^{-1}(O)$ terug open is in X. Voor een continue bijectie is f^{-1} ook continue indien $f(V)$ open is in Y. Indien een functie f voldoet aan deze eigenschap noemen we het een open afbeelding. Een voorbeeld van een continue bijectie waarvan de inverse niet continue is, is gegeven door $f : (-1, 1) \to (-1, 0] \times \mathbb{Z}_2 : x \to (-|x|, \theta(x))$ waarbij $|x| = -x$ als $x < 0$ en x als $x \geq 0$. $\theta(x) = 0$ voor $x \leq 0$ en 1 anders; finaal, $\mathbb{Z}_2 = \{0, 1\}$. De topologie die we leggen op $(-1, 0] \times \mathbb{Z}_2$ is deze van $(-1, 0]$ en is dus niet Hausdorff op $\{0, 1\}$. We hebben dat $f((-1, 0)) = (-1, 0) \times \{0\}$ is niet open terwijl $(-1, 0) \times \mathbb{Z}_2$ dat wel is. Een topologische equivalentie is dan gegeven door een bijectie f die continue en open is; controleer terug dat dit inderdaad een equivalentie relatie definieert.

We komen nu even terug op metrische topologieen en in het bijzonder alternatieve bepalingen van compactheid. Een rij $(x_i)_{i \in I}$ noemen we Cauchy als en slechts als voor elke $\epsilon > 0$, er bestaat een i, zodat voor alle $i \prec j, k$ we hebben dat $d(x_j, x_k) < \epsilon$. In mensentaal leest dit: als je maar voldoende ver in de rij kijkt dan liggen de punten willekeurig dicht bij elkaar. Zulke eigenschap suggereert het bestaan van een uniek limietpunt x; een metrische ruimte (X, d) waarvoor elke Cauchy rij een limietpunt heeft noemen we *compleet.* Indien een deelverzameling K compact is, dan hebben we dat elke rij $(x_i)_{i \in I}$ een deelrij heeft met een limiet-punt in K. Het bewijs is simpel, beschouw arbitraire eindige (omwille van de compactheid) overdekkingen met ballen van straal $\frac{1}{n}$, dan vinden we een rij van ballen $B(y_n, \frac{1}{n})$ zodat eindige doorsneden $\cap_{n=1}^{m} B(y_n, \frac{1}{n})$ een oneindig aantal $x_i \in K$ bevat. Dit definieert een deelrij met limietpunt

$$x = \cap_{n=1}^{\infty} B(y_n, \frac{1}{n})$$

in K. Omgekeerd, stel dat elke rij in K een Cauchy deelrij met limietpunt in K heeft, dan is K compact. Kies een overdekking van K met open ballen - zonder beperking van de algemeenheid- $B(y_n, \epsilon_n)$ waarbij $n \in \mathbb{N}$ en stel nu dat er geen enkele eindige deelbedekking bestaat. Noteer dat we ons beperken tot aftelbare overdekkingen wat een kleine beperking van de algemeenheid inhoudt; de geinteresseerde lezer kan uitzoeken of dit enig "wezenlijk" verschil maakt of niet, mij boeit het alleszins niet. Definieer dan $B_m = \cup_{n=1}^{m} B(y_n, \epsilon_n)$, we bekomen dan dat voor elke m er een $m' > m$ bestaat zodat $B_k \cap \overline{B_m}^c \cap K \neq \emptyset$ voor elke $k > m'$. In het bijzonder bekomen we dus een rij (x_m) met de eigenschap dat voor elke m er een $m' > m$ bestaat zodat $x_k \in \overline{B_m}^c$ voor $k \geq m'$. Deze rij kan geen Cauchy deelrij met limietpunt x bevatten vermits $x \in B_m$ voor m voldoende groot wat een tegenstrijdigheid is. We hebben net bewezen dat een verzameling K compact is in een metrische topologie als en slechts als elke rij een Cauchy deelrij met limietpunt in K bevat. Toon nu de volgende eigenschappen aan:

- definieer op $\mathbb{R}$ de functie $d(x, y) = |y - x|$, toon aan dat deze een metriek vormt (makkelijk),
- toon aan dat in de metrische topologie op $\mathbb{R}$, het gesloten interval $[a, b]$ compact is (hint: redeneer vanuit de decimale getalvoorstelling) (moeilijk),
- zij X, Y twee topologische verzamelingen en definieer de product topologie $\tau(X \times Y)$ als de kleinste topologie die $\tau(X) \times \tau(Y)$ bevat, waarbij deze laatste als elementen $U \times V$ heeft met $U \in \tau(X)$ en $V \in \tau(V)$,
- toon aan dat het Cartesisch product van compacte verzamelingen $K_1 \times K_2$ compact is in de product topologie (gemiddeld),
- een metrische ruimte (X, d) is begrensd als en slechts als er een $M > 0$ bestaat zodat $d(x, y) \leq M$ voor alle $x, y \in X$; toon aan dat een compacte ruimte gesloten en begrensd is (makkelijk).

De lezer kan terecht nogmaals uiten dat dit allemaal veel te algemeen is en dat onze wereld specialistischer is in de zin dat bijvoorbeeld lichtstralen buigen en twisten rond elkaar op een continue manier en dat dit gedrag geometrisch bepaald is. Om dit gepast te beschrijven hebbben we de notie nodig van een locale inproduct ruimte wat we zullen bestuderen in hoofdstuk zes: dit leidt tot analytische meetkunde. Noteer hier nog hetvolgende; veronderstel dat $\gamma : [a, b] \to X$ een continue kromme van x naar y is en definieer de lengte $L(\gamma)$ van γ als

$$L(\gamma) = \sup_{a=t_0<t_1<t_2<\ldots<t_{n+1}=b} \sum_{k=0}^{n} d(\gamma(t_k), \gamma(t_{k+1}))$$

waar het sup staat voor het supremum van deze som genomen over alle eindige verdelingen $a = t_0 < t_1 < t_2 < \ldots < t_{n+1} = b$ van het gesloten interval $[a, b]$. Het supremum van een verzameling reeele getallen A is het kleinste getal dat groter of gelijk is aan elk getal $x \in A$. Het supremum wordt ook wel eens de bovengrens genoemd en je moet aantonen dat wat betreft de reeele getallen, het supremum altijd bestaat en uniek is door het getal $+\infty$ toe te voegen. Op een gelijkaardige manier definieer je het infimum of ondergrens en je toont nogmaals aan dat in de reeele getallen het infimum bestaat en uniek is. Wat betreft onze som noteer je dat indien je een interval $[t_k, t_{k+1}]$ in twee breekt door een punt $t_k < t_{k+\frac{1}{2}} < t_{k+1}$ toe te voegen de som verhoogt. Dus een verfijning van de opdeling van het interval $[a, b]$ leidt tot een hogere som door gebruik te maken van de driehoeksongelijkheid.

Nu komen we tot ons hoofd resultaat; een complete metrische ruimte (X, d) definieert een pad-metriek d als en slechts als

$$d(x, y) = \min_{\gamma:[a,b]\to X, \gamma(a)=x, \gamma(b)=y} L(\gamma)$$

in andere woorden wanneer de afstand tussen twee punten gelijk is aan de minimale lengte van een kromme die x met y verbindt. Je bewijst dit zelf: indien d een padmetriek is maak je voortdurend gebruik van de middelpunt eigenschap om zo'n kromme te construeren en je begint met aan te tonen dat steeds $L(\gamma) \geq d(x, y)$ wat simpel volgt uit de driehoeksongelijkheid. Omgekeerd, indien zo'n kromme bestaat heb je uiteraard een middelpunt. Een kromme die de lengte minimaliseert noemen we een geodeet en in een padmetrische ruimte is de lengte van een geodeet gelijk aan de afstand tussen twee punten. Later zullen we nog tot een meer specialistische karakterisatie van geodeten komen wanneer we meer structuur eisen. Nogmaals, met deze primitieve begrippen kun je al een heel eind erg abstract redeneren en vooral Mikhail Gromov en Peter Anderson zijn hier specialisten in; indien je verder over deze erg primitieve metrische ruimten iets wilt leren, moet je hun werk erop naslaan.

Zoals je opmerkt is onze taal nog niet rijk genoeg om te spreken over begrippen zoals loodrechte stand, hoeken en dergelijke alhoewel dit in sommige gevallen mogelijk is. Je leert dan ook dat dit boek specifieker en specifieker gaat worden, dat de taal rijker wordt, maar ook meer complex wat toelaat sterkere verbanden of stellingen te vinden. Compactheid of lokaal compactheid is een belangrijk begrip omdat het de (lokale) topologie "eindig" houdt in een zekere zin; niet lokaal compacte structuren laten vaak sommige wiskundige constructies niet toe omdat de ruimte te groot is, zoals het bijvoorbeeld het geval is voor integralen. We komen nu tot heel speciale bouwstenen: lijnstukken, driehoeken en pyramides alsook hoger dimensionale veralgemeningen hiervan. We zullen deze gebruiken om bepaalde "regelmatige" topologische ruimten te beschrijven en topologisch te karakteriseren: centraal staat hierin het concept van homologie wat op zich weer aanleiding geeft tot een mooie abstracte wiskundige karakterisering.

Chapter 5

Simpliciale homologie.

Waar de vorige drie hoofdstukken erg abstract waren, beginnen we hier te werken met meer tastbare voorwerpen, dingen die we allemaal wel kennen uit het dagdagelijkse leven. We gaan hier op een hoger niveau leren mee omgaan; het is te zeggen op een meer abstracte manier dan gewoonlijk. Dit heeft zo zijn voordelen vermits het ons toelaat met deze objecten te *rekenen*; dit is eigenlijk het hoofdmirakel van abstractie, dat het ons toelaat dingen te doen en in te zien. De topologische ruimten die we gaan bestuderen zijn gemodelleerd aan de hand van de n-dimensionale reeele ruimte

$$\mathbb{R}^n = \times_{i=1}^n \mathbb{R} = \{(x_i)_{i=1}^n | x_i \in \mathbb{R}\}$$

welke de verzameling is van de n-koppels van reeele getallen uitgerust met de product metrische topologie van $\mathbb{R}$. Men kan de notie van een som uitbreiden als

$$(x_i) + (y_i) = (x_i + y_i)$$

alsook heeft men het begrip van de *scalaire* vermenigvuldiging van een reeel getal met een n-koppel of *vector* door

$$r.(x_i) = (rx_i).$$

Meer algemeen, zij R een veld en $G, +$ een commutatieve groep, dan noemen we G een R moduul indien er een scalaire vermenigvuldiging bestaat zodat

$$1.g = g;\ (rs).g = r.(s.g);\ (r+s).g = r.g + s.g;\ r.(g_1 + g_2) = r.g_1 + r.g_2$$

voor alle $r, s \in R$ en $g, g_1, g_2 \in G$. Indien $R = \mathbb{R}$ noemen we het moduul ook wel een vector ruimte. In $\mathbb{R}^n, +$ hebben we de speciale vectoren e_i, die een 1 bevatten op de i-de plaats en nul anders; hiervoor geldt dat

$$\sum_{i=1}^n r_i.e_i = 0$$

als en slechts als alle $r_i = 0$ en bovendien hebben we dat alle vectoren (uniek) geschreven kunnen worden op die manier, dus als $\sum_{i=1}^{n} r_i.e_i$. Indien deze twee eigenschappen gelden voor een verzameling van vectoren $\{v_i | i = 1 \dots m\}$, dan noemen we $\{v_i | i = 1 \dots m\}$ een *basis*. Je merkt op dat we hier twee getallen gebruikt hebben, n voor de e_i en m voor de v_j; het is nu een koud kunstje om aan te tonen dat $n = m$. De reden is devolgende, omdat e_i een basis is kun je v_j uniek schrijven als

$$v_j = \sum_{i=1}^{n} v_j^i e_i$$

en omgekeerd

$$e_i = \sum_{j=1}^{m} e_i^j v_j.$$

Dus,

$$\sum_{i=1}^{n} v_j^i e_i^k = \delta_j^k;\ j, k : 1 \dots m$$

en

$$\sum_{j=1}^{m} e_i^j v_j^l = \delta_i^l;\ i, l : 1 \dots n$$

waarbij $\delta_j^k = 1$ als en slechts als $j = k$ en nul anders. Dit systeem van vergelijkingen is symmetrisch in e en v en dus moet $m = n$ ($m < n$ zou met omwisseling van e en v leiden tot $m > n$ wat een contradictie oplevert). Dus n is een basis *invariant* en noemen we de dimensie van $\mathbb{R}^n, +$. Nu begrijpen we voldoende over reeele vectorruimten om onze speciale bouwsteentjes te gaan definieren; wanneer we deze hebben kunnen we concrete ruimten zoals simpliciale complexen gaan bouwen. Neem de ruimte $\mathbb{R}^{n+1}$ en beschouw een basis $v_i; i = 0 \dots n$, dan is het n simplex $(v_0 v_1 \dots v_n)$ gedefinieerd als de gesloten ruimte

$$(v_0 v_1 \dots v_n) = \{\sum_{i=0}^{n} \lambda_i v_i | \lambda_i \geq 0, \sum_{i=0}^{n} \lambda_i = 1\}.$$

Dit is allemaal een beetje abstract, om een idee te krijgen van hoe deze ruimte eruit ziet nemen we een voorbeeld in $0, 1, 2, 3$ dimensies. Een nul dimensionaal simplex (v_0) is gewoon een punt op een rechte; een één dimensionaal simplex is het lijnstuk $(v_0 v_1)$ in het vlak $\mathbb{R}^2$. Het tweedimensionaal simplex $(v_0 v_1 v_2)$ is een driehoek in $\mathbb{R}^3$ terwijl finaal $(v_0 v_1 v_2 v_3)$ een pyramide is in $\mathbb{R}^4$. In het algemeen is het simplex $(v_0 v_1 \dots v_n)$ convex wat betekent dat het lijnstuk tussen twee punten $x, y \in (v_0 v_1 \dots v_n)$ volledig tot $(v_0 v_1 \dots v_n)$ behoort. Het lijnstuk tussen twee punten x, y is de verzameling

$$\{\lambda x + (1 - \lambda) y | 0 \leq \lambda \leq 1\}.$$

Punten van het simplex die niet op het inwendige van een lijnstuk liggen dat volledig tot het simplex behoort noemen we extremale punten; toon bij wijze van oefening aan dat de enige extremale punten van $(v_0 v_1 \ldots v_n)$ de punten v_i zijn. Men noemt het simplex ook wel eens de convexe omhulling van de punten $\{v_i | i = 0 \ldots n\}$. We weten nu wat een moduul is alsook een simplex wat ons toelaat het begrip van een lineaire operator te definieren. Een afbeelding $A : V \to W$ tussen twee R modulen V, W is lineair als en slechts als

$$A(rv_1 + sv_2) = rA(v_1) + sA(v_2)$$

voor elke $r, s \in R$ en $v_i \in V$. Toon aan dat A injectief is als en slechts als $A(v) = 0$ impliceert dat $v = 0$. Vooraleer we verder gaan, bestuderen we enkele erg speciale ringen $\mathbb{Z}_n$. Deze is gedefinieerd als $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$, met andere woorden de equivalentieklassen bij deling door n:

$$k \equiv l \Leftrightarrow k - l = rn \,\text{voor een bepaalde}\, r \in \mathbb{Z}.$$

De som en product zijn dan gedefinieerd door de equivalentieklasse te nemen van de gewone som en product in $\mathbb{Z}$. Toon aan dat dit goed gedefinieerd is en dat $\mathbb{Z}_n, +, .$ een veld is als en slechts als n een priemgetal is. Bij wijze van voorbeeld heeft men in $\mathbb{Z}_3$ dat $1 + 2 = 0, 2 + 2 = 1$; $\mathbb{Z}_3$ is ook voorgesteld door $\{-1, 0, 1\}$. Nu gaan we een beetje abstracter werken: eigenlijk hebben we helemaal niet nodig dat $v_i \in \mathbb{R}^{n+1}$, dit was eigenlijk alleen maar goed voor de voorstelling of representatie. Vanaf nu zijn v_i, w_j gewoon namen voor punten die niet in een vooraf gegeven lineaire ruimte moeten leven. We definieren het $\mathbb{Z}$ moduul Z_n door het nemen van gehele veelvouden van simpliciale complexen, gedefinieerd door de simplices $(v_0, \ldots, v_n)$ waarbij het omwisselen van v_i en v_j een min-teken veroorzaakt; het is te zeggen, de orientatie draait om. Een simpliciaal complex is gedefinieerd als de topologische ruimte bepaald door een som van simplices waarin elk simplex in de som, of haar orientatie tegengestelde, meerdere malen mag voorkomen. We definieren de randoperator $\partial_n : Z_n \to Z_{n-1}$ als de lineaire operator over $\mathbb{Z}$ die een simplex $(v_0 v_1 \ldots v_n)$ afbeeldt op

$$\partial_n(v_0 v_1 \ldots v_n) = \sum_{i=0}^{n} (-1)^i (v_0 \ldots v_{i-1} v_{i+1} \ldots v_n).$$

Je controleert dan dat $\partial_{n-1}\partial_n S_n = 0$ gegeven een som van simplices S_n dat een simpliciaal complex definieert. We noemen een som T_k van $k = 0 \ldots n$ simplices die behoren tot een van de simplices in S_n gesloten indien $\partial_k T_k = 0$ en exact indien $T_k = \partial_{k+1} T_{k+1}$ for $k = 0 \ldots n-1$. Het is duidelijk dat exacte simplices gesloten zijn en we definieren de bijbehorende $\mathbb{Z}$ modulen $G_k(S_n)$ en $E_k(S_n)$, waarbij $E_n(S_n) = \{0\}$ en $G_0(S_n)$ is $\mathbb{Z}^V$ waarbij V het aantal punten of vertices is in S_n. Omwille van het voorgaande hebben we dat $E_k(S_n) \subseteq G_k(S_n)$ en we definieren de homologie-klassen $H_k(S_n)$ als het quotient

$$H_k(S_n) = \frac{G_k(S_n)}{E_k(S_n)}$$

wat niks anders is dan het $\mathbb{Z}$ moduul van $E_k(S_n)$ equivalentieklassen in $G_k(S_n)$. Hier, twee gesloten simplices T_k, Y_k zijn equivalent wanneer $T_k - Y_k \in E_k(S_n)$

en de lezer noteert dat $H_k(S_n)$ niet afhangt van de niet nul coefficienten van de n simplices in S_n en dus effectief een functie is van het simpliciaal complex gedefinieerd door S_n. Zover de algemene theorie voor simpliciale complexen, we komen nu tot de erg belangrijke deeltheorie van topologische ruimten A die homeomorf zijn aan een simpliciaal complex S_n; hier bestaat de meest belangrijke stap erin aan te tonen dat $H_k(A)$ goed gedefinieerd kan worden vermits homeomorfe simpliciale sommen dezelfde homologie-modulen definieren. De lezer die wil kan trachten dit zelf te bewijzen als een moeilijke oefening maar het statement zelf moet voor iedereen intuitief duidelijk zijn. De dimensie van $H_k(S_n)$ wordt het k-de Betti getal b_k genoemd van het simpliciaal complex S_n; de lezer maakt nu de volgende oefeningen, neem een boloppervlak en toon aan dat $b_2 = 1, b_1 = 0, b_0 = 1$. De twee torus T_2 wordt gevormd door een vierkant te nemen en tegengestelde zijden aan elkaar te plakken, toon hier aan dat $b_2 = 1, b_1 = 2, b_0 = 1$. In het algemeen definieert men het Euler getal van een twee dimensionaal simpliciaal complex S_2 als

$$\chi(S_2) = D - L + V$$

waar D het aantal driehoeken is en L het aantal lijnstukken. Men kan aantonen dat het Euler getal een topologische invariant is; bereken dat het Euler getal van het boloppervlak gelijk is aan $2 = b_2 - b_0 + b_1 = 1 - 0 + 1$ en dat van de torus $0 = 1 - 2 + 1$. In het algemeen toont men aan dat

$$\chi(S_n) := \sum_{i=0}^{n} (-1)^i V_{n-i} = \sum_{i=0}^{n} (-1)^i b_{n-i}$$

waarbij V_i het aantal i dimensionale sub-simplices is. Om aan de berekening van de dimensie van de homologie klassen te beginnen, noteer dat een element van $H_k(S_n)$ correspondeert met een gesloten k dimensionaal oppervlak dat niet tot een punt vervormd kan worden. Wat betreft b_1 op het boloppervlak is het duidelijk dat elke gesloten kromme tot een punt herleid kan worden, terwijl op de twee torus of autoband er twee fundamentele cirkels bestaan. We gaan verder met devolgende oefening: neem twee, twee dimensionale oppervlakken A_2 en B_2 en knip uit beiden een cirkelschijf. Plak nu de overblijfselen aan elkaar langs de rand van de cirkel, dan bekomen we een nieuwe topologische ruimte die we $A_2 \diamond B_2$ noemen. Toon aan dat $\diamond$ associatief en commutatief is met eenheidselement het twee dimensionaal boloppervlak S^2. Bereken dat het Euler getal van het n-voudig crossproduct van T_2 gelijk is aan $2 - 2n$; meer algemeen is

$$\chi(A_2 \diamond B_2) = \chi(A_2) + \chi(B_2) - 2.$$

Later gaan we het begrip varieteit bestuderen en een van de belangrijkste resultaten uit de topologie is dat elke gesloten, compacte, samenhangende en orienteerbare twee dimensionale varieteit homeomorf is met S^2 of een n-voudig product $T_2 \diamond T_2 \diamond \ldots \diamond T_2$. Dit betekent dat gesloten, compacte, samenhangende en orienteerbare twee dimensionale varieteiten volledig gekarakteriseerd zijn, topologisch, door het Euler getal. Voor gesloten varieteiten toont men aan dat $b_{n-i} = b_i$, de zogenaamde Betti dualiteit door een notie van dualiteit $\star$ te definieren zodat $S_n^\star$ homeomorf is aan S_n en $H_k(S_n)$ bijectief wordt afgebeeld op $H_{n-k}(S_n^\star)$. Je kunt jezelf amuseren door $\star$ te vinden als de natuurlijke veralgemening van devolgende operatie op een eendimensionaal S_1: beeldt elke rechte r af op een punt $r^\star$ en elk punt p op een rechte $p^\star$ zodat $r^\star, s^\star$ verbonden worden door $p^\star$ als en slechts als p behoort tot r, s. $S_1^\star$ is een gesloten simpliciaal complex als S_1 dat is, maar kan een verschillend Euler getal hebben indien S_1 geen varieteit is. Dus de varieteits voorwaarde is echt nodig en Betti dualiteit geldt dus niet voor gesloten simpliciale complexen. Toon aan dat twee cirkels die een gemeenschappelijk punt hebben, een voorbeeld leveren van zo'n ruimte. De notie van varieteit is dus wel echt speciaal en ons resultaat in twee dimensies is niet veralgemeenbaar naar hogere dimensies toe; drie dimensionale varieteiten zijn zeer moeilijk te classificeren en het Euler getal volstaat bijlange na niet. Hier stoppen we wat betreft simpliciale homologie; nu gaan we op een hoger plan redeneren. De ganse simpliciale homologie kan samengevat worden als een ketting van operaties $\partial_k : Z_k(S_n) \to Z_{k-1}(S_n)$ met $\partial_0 : Z_0(S_n) \to 0$ en $\partial_{k+1}\partial_k = 0$. Dit noemen we een chain en deze objecten genieten vele mooie abstracte eigenschappen die veel primitiever zijn dan het ruimtelijk concept waarmee we vertrokken zijn. Een beginpunt voor hogere wiskunde dus!

Chapter 6

Lineaire ruimten en operatoren.

De lezer doet er goed aan het pad naar meer specializering in dit boek te overlopen: we zijn vertrokken van verzamelingenleer en daar mag echt erg veel, dan zijn we iets beperkter gaan kijken naar topologie en simpliciale homologie toe om uiteindelijk op de notie varieteit te botsen. Deze heeft erg bijzondere topologische eigenschappen en de volgende twee hoofdstukken zijn gewijd aan het pad naar de varieteit in de meest algemene zin van het woord. Het voordeel van meer gespecifieerde structuren is uiteraard dat je er meer kunt over zeggen, dat ze controleerbaar zijn; hierin anticipeert de wiskundige dikwijls een onbewezen stabiliteitsresultaat, dat structuren die bijna aan die spelregels voldoen ook ongeveer dezelfde eigenschappen genieten. Dat is de werkelijke kracht van subtiele vereenvoudiging, dat zij ons toelaat dingen te begrijpen die niet gelden vanuit een hoger standpunt. Albert Einstein praatte hierover in de termen dat de dingen zo eenvoudig als mogelijk voorgesteld moesten worden maar niet te eenvoudig; te veel beperkingen lopen namelijk het risico dat je veel dingen mist terwijl te weinig beperkingen je dikwijls met het probleem confronteren dat zelfs de meest simpele waarneming niet meer "verklaarbaar" is. Wetenschap is de kunst om deze zeer fijne grens beter en beter af te tasten, stukje per stukje; geen dogmatische zekerheid, die heb je nooit, maar met rede en intelligentie "snuffelen".

In deze filosofie zijn lineaire ruimten te specialistisch maar we willen varieteiten bekijken als iets wat eruit ziet als een lineaire ruimte "op kleine schaal". Dus passen we nu een omgekeerde strategie toe, we keren eerst het "eenvoudige" geval binnenstebuiten vooraleer het monster in haar volle glorie te ontmaskeren. Een lineaire ruimte is een bi-moduul over de reeele of complexe getallen; het woord "bi" slaat op het gegeven dat de scalaire vermenigvuldiging zowel langs links als rechts kan gebeuren en dat beiden gelijk zijn aan elkaar. We noemen de lineaire ruimte eindig dimensionaal wanneer de dimensie over het veld eindig is en oneindig dimensionaal anders. De dimensie is nog steeds goed gedefinieerd als het aantal basis vectoren gegeven dat het bestaan van een basis aangetoond kan worden uit het keuze axioma van de verzamelingenleer.

Een lineaire afbeelding $A : V \to W$ tussen twee lineaire ruimten over hetzelfde veld is een functie die voldoet aan

$$A(r.v + s.w) = r.A(v) + s.A(w)$$

waarbij het punt telkens de scalaire vermenigvuldiging noteert. Evidenterwijs heeft men dat $A(0) = 0$ en dat $A(v) = A(w)$ als en slechts als $A(v - w) = 0$. Dus de zogenaamde kern van A, genoteerd door $\mathrm{Ker}(A) = \{v | A(v) = 0\}$ meet de afwijking van injectiviteit van A; elk beeldpunt $A(w)$ heeft als inverse $w + \mathrm{Ker}(A)$. De kern zelf is een vectorruimte, een zogenaamde deelruimte van V. Op dezelfde manier heeft men dat het zogenaamde beeld

$$\mathrm{Im}(A) = \{A(v) | v \in V\}$$

een deelruimte is van W. Het is nu vanzelfsprekend dat

$$\mathrm{Im}(A) \cong \frac{V}{\mathrm{Ker}(A)}$$

met andere woorden beide ruimten zijn isomorf. Inderdaad,

$$A : \frac{V}{\mathrm{Ker}(A)} \to \mathrm{Im}(A) : w + \mathrm{Ker}(A) \to A(w)$$

is lineair en bijectief wat de betekenis is van een isomorfisme. Een triviaal gevolg van deze stelling is dat

$$\dim(\mathrm{Ker}(A)) + \dim(\mathrm{Im}(A)) = \dim(V)$$

waarbij "dim" staat voor de dimensie. Lineaire afbeeldingen kunnen voorgesteld worden ten opzichte van basissen e_i van V en f_j van W respectievelijk. Dit gaat als volgt

$$A(e_i) = \sum_{j=1}^{m} A_i^j f_j$$

j noemen we de rij-index en i de kolom-index; ten opzichte van een algemene vector $v = v^i e_i$ geeft dit

$$A(v) = \sum_{j=1}^{m} (\sum_{i=1}^{n} A_i^j v^i) f_j.$$

De samenstelling van twee lineaire afbeeldingen $A : V \to W$ en $B : W \to Z$ resulteert dan in het matrix product

$$(BA)_i^j = \sum_{k=1}^{m} B_k^j A_i^k$$

waarbij m de dimensie is van W. Voortaan laten we de sommatie-tekens weg wat bekend staat als de Einstein conventie; dus

$$\sum_{i=1}^{n} A_i^j v^i$$

wordt genoteerd als

$$A_i^j v^i.$$

Een 2×3 matrix, oftewel een matrix met 2 rijen en 3 kolommen wordt voorgesteld als

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

en in een matrix product BA moet de kolom-dimensie van B gelijk zijn aan de rij-dimensie van A. Toon aan als een oefening dat

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 13 \\ 19 & 19 \end{pmatrix}.$$

Bewijs voor jezelf dat indien A, B 2×2 matrices zijn dat dan in het algemeen

$$AB - BA \neq 0$$

waarbij 0 hier staat voor de gepaste nulmatrix. Dit resultaat toont aan dat de matrixvermenigvuldiging in het algemeen niet-commutatief is en aldus hebben we een niet-commutatieve ring geconstrueerd vanuit een veld. Je kan je dus terecht afvragen of onze niet-conventionele, niet-commutatieve getalstelsels zoals de quaternionen en Clifford algebra's voorstelbaar zijn als matrices over de complexe getallen. Het antwoord is ja en er bestaan zelfs voorstellingen in verschillende dimensies: toon aan dat de reeele quaternionen geschreven kunnen worden als

$$q = \begin{pmatrix} a + ib & ic - d \\ ic + d & a - ib \end{pmatrix}$$

waarbij $a, b, c, d \in \mathbb{R}$. Een andere schrijfwijze in termen van de Pauli matrices is

$$q = a.1 + ic.\sigma_1 + id.\sigma_2 + ib.\sigma_3.$$

Nu dat we een aantal essentiele kenmerken van matrixrekening gesnapt hebben komen we tot een niet onbelangrijk vraagstuk: bestaat er een basis e_i in V zodanig dat lineaire operatoren $A : V \to V$ een speciale, uiterst eenvoudige matrixvoorstelling hebben ten opzichte van e_i? Uiteraard is dit wat vaag geformuleerd maar probeer nu eens voor jezelf duidelijk te maken dat een willekeurige $n \times n$ matrix A bijna altijd kan geschreven worden als ODO^{-1} waarbij $OO^{-1} = 1_n = O^{-1}O$ en $D_i^j = \lambda_i \delta_i^j$ met $\delta_i^j = 1$ als en slechts als $i = j$ en nul anders. D is een zogenaamde diagonaal $n \times n$ matrix en λ_i zijn eigenwaarden die voldoen aan

$$A(Oe_i) = \lambda_i(Oe_i)$$

met andere woorden Oe_i is een eigenvector van A met eigenwaarde λ_i. O noemen we een invertiebele of omkeerbare $n \times n$ matrix. De redenering is echt poepsimpel hoor: in het algemeen is bijna elke matrix O invertiebel zodat O effectief n^2 vrijheidsgraden heeft; verder reduceert de afbeelding $O \to ODO^{-1}$ exact n dimensies indien alle λ_i in D verschillend zijn vermits $VDV^{-1} = ODO^{-1}$ impliceert dat $(V^{-1}O)D = D(V^{-1}O)$ zodat $V = OD'$ met D' diagonaal en daarom is een generieke D "orbit" $n^2 - n$ dimensionaal. Vermits het aantal vrijheidsgraden in D ook n is hebben we in totaal n^2 vrijheidsgraden en dus een "generieke" $n \times n$ matrix. Als oefening controleer je elke bewering in deze redenering. Vooraleer we verder gaan, bestuderen we eerst het effect van een basisverandering op de matrixvoorstelling van A. Zij $e'_i = O(e_i)$ dan is

$$A'^i_j e'_i = A(O(e_i)) = A(O^j_i e_j) = A^k_j O^j_i e_k$$

en aldus $A'^i_j = (O^{-1})^i_k A^k_l O^l_j$. Dus, generiek, bestaat er een basis ten opzichte van dewelke de matrixvoorstelling van A diagonaal is. Toon aan dat deze n eigenwaarden uniek zijn alsook de eigenvectoren (op een reeel getal na) insien alle λ_j verschillend zijn. Zijn er uitzonderingen op deze regel? Uiteraard! Bewijs dat de matrix

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

voldoet aan $N^2 = 0$ en daarom niet diagonalizeerbaar is. De redenering gaat als volgt: veronderstel dat dit wel het geval is dan moet elke eigenwaarde nul zijn en aldus $N = 0$ in strijd met het gegeven. Deze uitzondering is botweg de enige ook in hogere dimensies en men kan hieruit de Cartan klassificatie uit afleiden. In twee dimensies leest die dat elke matrix A door geschikte keuze van basis uniek één van de volgende twee voorstellingswijzen heeft:

$$\begin{aligned} A &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \\ A &= \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}. \end{aligned}$$

Indien je dit resultaat, alsook de gepaste uitbreiding naar hogere dimensies toe, graag zou bewijzen wacht dan even tot we de notie van determinant gezien hebben die je toelaat het bestaan van eigenwaarden af te leiden. Een $n \times n$ matrix A kan nog op een andere manier bekeken worden dan de voorstelling van een lineaire afbeelding ten opzichte van een basis; het is te zeggen, men kan A unterpretereb als eeb geordende collectie van n kolom vectoren $A = (v_1, \ldots, v_n)$. Deze zienswijze laat toe A te bekijken als de notatie voor een simplex of *de* meerdimensionale kubus opgespannen door v_i. De determinant $\det(A)$ berekent het volume alsook de orientatie van die kubus indien de basisvectoren e_i loodrecht op elkaar staan. We leiden de vorm van $\det(A)$ af uit de voorwaarden waaraan het volume moet voldoen; ten eerste is det multilineair in de kolommen, het is te zeggen:

$$\det(v_1, \ldots, v_{i-1}, a.v_i + b.w_i, v_{i+1}, \ldots, v_n) =$$

$$a\det(v_1,\ldots,v_{i-1},v_i,v_{i+1},\ldots,v_n)+b\det(v_1,\ldots,v_{i-1},w_i,v_{i+1},\ldots,v_n)$$

alsook nilpotentie wat betekent dat indien $v_i = v_j$ dan moet de determinant nul geven. Deze laatste voorwaarde drukt gewoon uit dat indien bepaalde assen samenvallen dan definieert de matrix een lager dimensionaal geheel. Finaal leggen we de normalizatie conditie op dat $\det(1_n) = 1$. Deze drie voorwaarden leggen de derminant volledig vast: uit de eerste en tweede voorwaarde leidt je af dat de determinant volledig antisymmetrisch is, het is te zeggen $\det(v_1,\ldots,,v_i,\ldots,v_j,\ldots,v_n) = -\det(v_1,\ldots,,v_j,\ldots,v_i,\ldots,v_n)$. Samen met de derde en eerste conditie bekom je dan dat

$$\det(A) = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) A^1_{\sigma(1)}\ldots A^n_{\sigma(n)}$$

waarbij σ een zogezegde permuatie is en sign het teken ervan. Omwille van de anti-symmetrie mag elke index slechts éénmaal voorkomen zodat een permutatie een bijectie is van $\{1,2,\ldots,n\}$ naar zichzelf en sign noteert dan het even of oneven karakter van de omwisselingen die men moet doorvoeren om van de eenheid te komen tot σ. Je toont aan dat de permuaties een niet commutatieve groep vormen S_n met $n.(n-1).(n-2)\ldots 3.2.1$ elementen en we geven het bewijs dat sign goed gedefinieerd is. Dit is een beetje technisch; noteer met (ij) de permutatie die i en j omwisselt van plaats dan is

$$\begin{aligned}(ik)(ij) &= (jk)(ik)\\ (ik)(jl) &= (jl)(ik)\\ (ik)(ij) &= (ij)(jk)\end{aligned}$$

voor verschillende i,j,k,l. Allereerst is het duidelijk dat elke permutatie te schrijven is als een product van dergelijke omwisselingen; gegeven een niet triviaal product van dergelijke omwisselingen voor de eenheid dan is het een koud kunstje om aan te tonen dat deze een even aantal van deze bevat door gebruik te maken van de hierbovenstaande regels. Inderdaad, gegeven $\sigma = l(ik)s(jk)$ waarbij l,s producten zijn van omwisselingen en l geen omwisseling bevat met de index k, toon dan aan dat je deze string kan herschrijven als $\sigma = ls'$ waarbij s' ook niet de index k bevat en evenveel omwisselingen heeft dan s. Op deze manier toon je aan dat σ een even aantal omwisselingen heeft. Hieruit volgt dat twee verschillende producten voor eender welke permutatie altijd verschillen op een even aantal omwisselingen. Dus de functie sign is goed gedefinieerd; toon aan dat de determinant van een 2×2 matrix gegeven is door

$$\det\begin{pmatrix} a & b\\ c & d\end{pmatrix} = ad - bc.$$

Toon nu de volgende eigenschappen aan

- $\mathrm{sign}(\rho\sigma) = \mathrm{sign}(\rho)\mathrm{sign}(\sigma)$
- $\det(AB) = \det(A)\det(B)$.

Dit laatste geldt vermits

$$\begin{aligned}
\det(AB) &= \sum_{\sigma\in S_n} \mathrm{sign}(\sigma)(AB)^1_{\sigma(1)}\dots(AB)^n_{\sigma(n)} \\
&= \sum_{\sigma\in S_n}\sum_{m_1,\dots,m_n} \mathrm{sign}(\sigma)A^1_{m_1}\dots A^n_{m_n}B^{m_1}_{\sigma(1)}\dots B^{m_n}_{\sigma(n)}
\end{aligned}$$

en in dit laatste product moet $m_1,\dots,m_n$ zelf een permutatie zijn, de gevallen waar dit niet zo is tellen op tot nul. Dit is makkelijk in te zien als volgt: veronderstel dat $m_i = m_j$ dan heb je voor elke σ dezelfde term met tegengesteld teken door de permutatie $\sigma(ij)$ te nemen. Dus we hebben dat

$$\begin{aligned}
\det(AB) &= \sum_{\sigma,\rho\in S_n} \mathrm{sign}(\sigma)A^1_{\rho(1)}\dots A^n_{\rho(n)}B^{\rho(1)}_{\sigma(1)}\dots B^{\rho(n)}_{\sigma(n)} \\
&= \sum_{\sigma,\rho\in S_n} \mathrm{sign}(\sigma\rho)A^1_{\rho(1)}\dots A^n_{\rho(n)}B^1_{\sigma(1)}\dots B^n_{\sigma(n)} \\
&= \det(A)\det(B).
\end{aligned}$$

In het bijzonder geldt dan dat $\det(A^{-1}) = (\det(A))^{-1}$. Dus de determinant van $A = (v_1,\dots,v_n)$ is verschillend van nul als en slechts als de v_i een basis vormen als en slechts als A invertiebel is. Toon aan dat de inverse van

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gegeven wordt door

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Bestudeer de determinant voor 3×3 matrices en probeer een gelijkaardige formule te ontwikkelen voor de inverse van een matrix.

Nu komen we terug tot het bestuderen van het Cartan classificatie probleem en in het bijzonder het aantonen van het bestaan van eigenwaarden λ. λ is een eigenwaarde als en slechts als er bestaat een eigenvector v_λ zodat

$$A(v_\lambda) = \lambda v_\lambda.$$

Een andere manier om dit te zeggen is dat de kern van de operator $A - \lambda 1_n$ niet nul is wat dan weer waar is als en slechts als

$$\det(A - \lambda 1_n) = 0.$$

Het is hier dat determinanten nuttig worden omdat deze formule kan geinterpreteerd worden als een nulpuntsvergelijking voor een veelterm van n-de graad. Zoals we weten is deze veelterm volledig factorizeerbaar over $\mathbb{C}$ en hebben we dus in het algemeen n verschillende complexe eigenwaarden wat aantoont dat bijna elke matrix diagonalizeerbaar is; iets wat we al wisten. Werk nu eens uit in twee dimensies wat er kan gebeuren als de twee complexe eigenwaarden samenvallen.

De aandachtige lezer heeft ondertussen al lang opgemerkt dat de determinant van een matrix een basisinvariant is; het is te zeggen

$$\det(O^{-1}AO) = \det(O)^{-1}\det(A)\det(O) = \det(A)$$

wat aantoont dat je peilt naar een eigenschap van de geassocieerde operator. Daarom merk je onmiddellijk op dat elke coefficient in de eigenwaarde veelterm $\det(A - \lambda 1_n)$ een invariant is van de operator. In het bijzonder heb je dus dat de invariant behorende bij de $n-1$ de macht in λ gelijk is aan $(-1)^{n-1}\mathrm{Tr}(A)$ waarbij het spoor Tr gegeven is door

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A^i_i.$$

Controleer bij wijze van oefening op een expliciete manier dat het spoor inderdaad een invariant is en bestudeer de hogere orde invarianten. Tracht deze te schrijven in termen van het spoor; bijvoorbeeld toon aan in twee dimensies dat

$$2\det(A) = (\mathrm{Tr}(A))^2 - \mathrm{Tr}(A^2).$$

Toon aan dat indien je de eigenwaarde veelterm neemt en hierin λ vervangt door de matrix A zelf dat dan de resulterende matrix de nul matrix is. Dit is de stelling van Cayley Hamilton (hint: onderstel eerst dat A diagonalizeerbaar is en gebruik dan de definitie van een eigenwaarde als nulpunt van de eigenwaarde veelterm) en deze leest in twee dimensies als

$$A^2 - \mathrm{Tr}(A)A + \det(A)1_2 = 0.$$

Toon finaal aan dat $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ en dat er daarom geen paar matrices A, B bestaan zodat

$$AB - BA = 1_n$$

een formule die bekend staat als de bosonische Heisenberg relatie. Zoek in twee dimensies daarintegen wel een paar matrices zodat

$$AB + BA = 1_2$$

hetgeen bekend staat als de fermionische Heisenberg relatie. Bosonen hebben dus een oneindig aantal dimensies nodig terwijl Fermionen leven in dimensies $n = 2^d$ waarbij hier $d = 1$. Tot slot definieren we de notie van transpositie A^T en complex toegevoegde $\overline{A}$ van een matrix A

$$(A^T)^i_j = A^j_i, \ (\overline{A})^i_j = \overline{A^i_j}.$$

Toon aan dat

$$(AB)^T = B^T A^T, (A^T)^T = A, (rA + sB)^T = rA^T + sB^T, (A^{-1})^T = (A^T)^{-1}$$

alsook gelijkaardige eigenschappen voor de complex toegevoegde. In het bijzonder is voor

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

$$A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

ofwel rijen en kolommen worden omgedraaid. Toon aan dat

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

voldoet aan $N^T N + N N^T = 1_2, N^2 = 0$ en we noemen N een fermionische creatie-operator. Dit is voldoende voor je eerste kennismaking met dit onderwerp; in hetvolgende hoofdstuk gaan we hier veel dieper op in.

Chapter 7

Hilbert ruimten en enkele belangrijke stellingen.

Tot hiertoe zijn we stilzwijgend geweest over het onderwerp van topologieen op lineaire ruimten alsook op ruimten van lineaire operatoren. De reden is erg simpel, al deze ruimten zijn verzamelingtheoretisch equivalent aan $\mathbb{R}^n$ en op deze zijn alle "natuurlijke" topologieen equivalent aan de product metrische topologie. In een oneindig aantal dimensies bestaan er echter een aantal verschillende natuurlijke kandidaat topologieen welke we verder zullen bestuderen in een vroeg stadium van dit hoofdstuk. We zullen echter beginnen met vlakke meetkunde en bestuderen hoe dit samenhangt met kanstheorie alsook topologie: de filosofie is dat vlakke meetkunde deze concepten voorafgaat in een welbepaalde zin en dat met name continuiteit een relatief begrip wordt dat verabsoluteert in een eindig aantal dimensies gegeven dat alle topologieen daar equivalent of hetzelfde zijn.

Er zijn verschillende manieren om hetgeen wat volgt aan te brengen maar ik geloof dat de presentatie hieronder de beste alsook meest efficiente is. Cruciaal aan vlakke meetkunde is het gegeven dat de ruimte een lineaire ruimte is, dit is niet langer waar wanneer we gekromde meetkunde gaan bestuderen. Dit gegeven heeft er tweeduizend jaar lang voor gezorgd dat we het georienteerd lijnstuk tussen twee punten x, y konden karakteriseren aan de hand van de vector $y - x$ die dan "gebonden" wordt verondersteld in het punt x. Je ziet dat we het woord punt gebruikt hebben in een zin die verschillend is van een vector, maar wiskundig gezien zijn het allemaal vectoren gegeven dat een lineaire ruimte een voorkeurspunt heeft gegeven door het neutraal element 0, ook wel de oorsprong genoemd van deze. Deze gepréfereerde oorsprong is erg lang voorwerp geweest van theologisch en filosofisch debat: is de aarde het centrum van het heelal dat eeuwig stil staat? Of is het de zon of een ander lichaam in de kosmos? Newton en consorten waren de eersten die hier paal en perk aan stelden: zij introduceerden het begrip van een affiene ruimte door translaties toe te laten. Inderdaad, de afbeelding $x \to x + a$ commuteert niet met de optelling vermits

$$(x+y)+a \neq (x+a)+(y+a).$$

Ze laat echter wel het verschil ongemoeid

$$(y+a)-(x+a) = y-x$$

zodat nu vectoren een wezenlijk verschillende status krijgen dan punten. In Newton's wereld staat niks stil en dat was op zich een grote verdienste. Wiskundigen zoals Gauss, Riemann en Cartan zijn nog verder gegaan: de moderne cosmos heeft geen translatie-symmetrie en laat zich dus niet beschrijven in de taal der affiene ruimten. De invoering van deze inzichten in de natuurkunde is de grote verdienste van Albert Einstein en de relativiteitstheorie is bij verre een superieur verklaringsschema wat betreft het universum zoals we het waarnemen. De Euclidische ruimte of een (on)eindig dimensionale vlakke meetkunde is gedefinieerd door een reeele vectorruimte $\mathcal{H}$ alsook een scalair product $\langle v|w\rangle$ waarbij $v, w \in \mathcal{H}$. Het scalair product tussen v en w wordt verondersteld gelijk te zijn aan het product van de georienteerde lengte van de projectie van w op v met de lengte van v. Deze grootheid voldoet, bij het maken van een simpele tekening, aan:

$$\begin{aligned}
\langle v|w\rangle &= \langle w|v\rangle \\
\langle v|aw+bu\rangle &= a\langle v|w\rangle + b\langle v|u\rangle \\
\langle v|v\rangle &\geq 0 \text{ en er is gelijkheid als en slechts als } v=0.
\end{aligned}$$

Een scalair product bepaalt dus de notie van loodrechte stand; het feit dat wij hier op aarde een voorkeursbegrip hebben voor een welbepaalde loodrechte stand is een *fysisch* gegeven. Albert Einstein heeft ontdekt dat dit inherent is aan de beschrijving van het gravitationele veld. Het zou best zo kunnen zijn dat een alien ons gravitationeel veld anders ervaart dan wij en onze vallijnen niet bestempelt als loodrecht op het aardoppervlak. Men kan ook spreken over complexe meetkunden, in dat geval definieert men op exact dezelfde wijze een sesquilineaire vorm waarbij nu $\langle v|w\rangle = \overline{\langle w|v\rangle}$ en de complexe toevoeging als gewoonlijk gegeven is door $\overline{a+bi} = a - bi$.

Bijvoorbeeld, $\mathbb{C}$ is een ééndimensionale Hilbertuimte met scalair product $\overline{v}w$. Zoals gezegd in de inleiding heeft een Hilbertruimte bepaalde topologische eigenschappen; om deze te definieren tonen we eerst aan dat een scalair product canonisch een metriek d definieert. Hiertoe bewijzen we eerst dat de grootheid $||v||$ gegeven door

$$||v|| = \sqrt{\langle v|v\rangle}$$

een zogenaamde norm defnieert wat een veralgemening is van het begrip modulus van een complex getal. We tonen eerst hetvolgend triviaal resultaat aan

$$|\langle v|w\rangle| \leq ||v||||w||$$

wat betekent dat de lengte van de projectie van w of v vermenigvuldigd met de lengte van v kleiner of gelijk is aan het product van de lengtes van v en w, triviaal dus. Het formeel bewijs gaat als volgt:

$$0 \leq ||v + \lambda w||^2 = ||v||^2 + |\lambda|^2 \, ||w||^2 + 2\mathrm{Re}\left(\overline{\lambda}\langle w|v\rangle\right)$$

waarbij $\mathrm{Re}(a + ib) = a$ het reeel deel is van het complexe getal $z = a + bi$. Je controleert dat het reeel deel van het complexe getal z kan geschreven worden als $\frac{1}{2}(z + \overline{z})$ terwijl het imaginaire deel gelijk is aan $-i\frac{1}{2}(z - \overline{z})$. De modulus van een complex getal is gedefinieerd als

$$|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$$

en voldoet aan

$$|z + z'|^2 = |z|^2 + |z'|^2 + (z\overline{z'} + \overline{z}z')$$

waarbij de laatste term op een factor twee na gelijk is aan

$$aa' + bb'$$

en de absolute waarde hiervan is naar boven begrensd door $|a|\,|a'| + |b|\,|b'|$. Het kwadraat van deze laatste uitdrukking is gelijk aan

$$a^2a'^2 + b^2b'^2 + 2\,|a|\,|a'|\,|b|\,|b'| \leq \left(a^2 + b^2\right)\left(a'^2 + b'^2\right) = |z|^2\,|z'|^2$$

en bijgevolg hebben we dat

$$|z + z'|^2 \leq (|z| + |z'|)^2$$

en aldus

$$|z + z'| \leq |z| + |z'|$$

een formule die bekend staat als de driehoeksongelijkheid. Bijgevolg kunnen we terug een metriek definieren op de complexe getallen door

$$d(z, z') = |z - z'|\,.$$

Nu keren we terug naar het bewijs van onze ongelijkheid; je merkt op dat we λ zodanig kunnen kiezen dat

$$\mathrm{Re}\left(\overline{\lambda}\langle w|v\rangle\right) = -\,|\lambda|\,|\langle v|w\rangle|$$

terwijl, in het algemeen, het linkerlid steeds groter of gelijk aan is dan het rechterlid. Daarom hebben we dat

$$0 \leq ||v||^2 + |\lambda|^2\,||w||^2 - 2\,|\lambda|\,|\langle v|w\rangle|$$

wat een veelterm ongelijkheid is van tweede graad in de positieve veranderlijke $|\lambda|$. Aan deze ongelijkheid is voldaan als en slechts als de veelterm hoogstens één nulpunt bevat wat betekent dat de discriminant kleiner of gelijk moet zijn aan nul

$$0 \geq 4\,|\langle v|w\rangle|^2 - 4||v||^2||w||^2$$

wat onze bewering aantoont en gelijkheid geldt als en slechts als $w = -\lambda v$. Bijgevolg,

$$||v+w||^2 \leq ||v||^2 + ||w||^2 + 2\,|\langle v|w\rangle| \leq ||v||^2 + ||w||^2 + 2||v||||w|| = (||v|| + ||w||)^2$$

en dus gehoorzaamt de norm de driehoeksongelijkheid. Bijgevolg heeft elke Hilbertruimte $\mathcal{H}$ een canonische metrische topologie gedefinieerd door

$$d(v,w) = ||v-w||$$

en we eisen dat $\mathcal{H}$ compleet is in deze topologie. Dit axioma is uiterst belangrijk voor de theorie van lineaire operatoren maar vooraleer we hieraan toekomen, laat ons beginnen met een aantal observaties te maken. Twee niet-nul vectoren v, w staan loodrecht op elkaar als en slechts als $\langle v|w\rangle = 0$ en we noemen v genormeerd indien $||v|| = 1$. Door middel van het keuze axioma kan men aantonen dat er een orthonormale basis $(e_i)_{i\in I}$ bestaat die per definitie voldoet aan $\langle e_i|e_j\rangle = \delta_{ij}$ met δ_{ij} gelijk aan 0 als $i \neq j$ en 1 anders. De aandachtige lezer merkt de gelijkenis op met het symbool δ^i_j dat echter een basis invariant is waar δ_{ij} dat niet is. Voor eindig dimensionale Hilbertruimten, met $v = \sum_{i=1}^n v^i e_i$, impliceert dit dat

$$\langle v|w\rangle = \sum_{i,j=1}^n \overline{v^i} w^j \delta_{ij}$$

wat een veralgemening is van het standaard inproduct in driedimensionale Euclidische meetkunde. Toon aan dat onder een basistransformatie $e'_i = O^j_i e_j$ we hebben dat $\delta'_{ij} = \langle e'_i|e'_j\rangle = \overline{O}^k_i O^l_j \delta_{kl}$. Oefening: definieer Hilbertruimten over de quaternionen.

Laat ons eerst enkele algemene constructies bestuderen die doorgevoerd kunnen worden met reeele of complexe Hilbertruimten; de lezer is geadviseerd deze te veralgemenen naar quaternionische Hilbertruimten toe. In de quantum mechanica zijn de volgende twee operaties van belang: het tensorproduct $\otimes$ en de directe som $\oplus$. Gegeven twee Hilbertruimten $\mathcal{H}_i$, het tensorproduct $\mathcal{H}_1 \otimes \mathcal{H}_2$ is opnieuw een Hilbertruimte die als volgt gedefinieerd is: men begint met vectoren $v_1 \otimes v_2$ waar $v_i \in \mathcal{H}_i$ en neemt vervolgens eindige formele sommen $\sum_{i=1}^n z_i v^i \otimes w^i$ hiervan. Vervolgens deelt men de corresponderende lineaire ruimte door de volgende equivalentie relatie

$$\begin{aligned} z(v \otimes w) &\equiv (zv) \otimes w \equiv v \otimes (zw) \\ v \otimes w_1 + v \otimes w_2 &\equiv v \otimes (w_1 + w_2). \end{aligned}$$

We definieren $\mathcal{H}$ als de lineaire ruimte van deze equivalentie klassen en completeren het in de metrische topologie gedefinieerd door het scalair product

$$\langle v_1 \otimes w_1 | v_2 \otimes w_2\rangle := \langle v_1|v_2\rangle\langle w_1|w_2\rangle.$$

Op een gelijkaardige manier kan men de directe som $\mathcal{H}_1 \oplus \mathcal{H}_2$ definieren, maar deze keer leest de equivalentie relatie

$$\begin{aligned} z(v \oplus w) &\equiv (zv) \oplus (zw) \\ v_1 \oplus w_1 + v_2 \oplus w_2 &\equiv (v_1 + v_2) \oplus (w_1 + w_2) \end{aligned}$$

en het scalair product is gedefinieerd door

$$\langle v_1 \oplus w_1 | v_2 \oplus w_2 \rangle := \langle v_1 | v_2 \rangle + \langle w_1 | w_2 \rangle.$$

Je controleert dat een basis voor $\mathcal{H}_1 \otimes \mathcal{H}_2$ gegeven is door $v_i \otimes w_j$ waar v_i een basis is van $\mathcal{H}_1$ en w_j van $\mathcal{H}_2$. Een basis voor $\mathcal{H}_1 \oplus \mathcal{H}_2$ is gegeven door $v_i \oplus 0, 0 \oplus w_j$.

In een eindige vectorruimte definieert een basis dus een scalair product en de afbeelding van de basissen naar de Hilbertruimten toe is surjectief. Basissen die verbonden zijn door een transformatie O waarvoor geldt

$$\overline{O}_i^k O_j^l \delta_{kl} = \delta_{ij}$$

bepalen hetzelfde beeld en omgekeerd, basissen die hetzelfde scalair product bepalen zijn verbonden door zo'n transformatie. Je contoleert bij wijze van oefening dat deze matrices een groep vormen genaamd $U(n)$ in het complexe geval en $O(n)$ in het reeele, de unitaire groep en orthogonale groep in n dimensies. De hierbovenstaande formule leest in matrix taal

$$O^H O = 1$$

waarbij $O^H = (\overline{O})^T = \overline{(O^T)}$. Toon aan dat in twee dimensies de unitaire matrices gegeven zijn door

$$O = \frac{1}{\sqrt{|a|^2 + |b|^2}} \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$$

met $a, b \in \mathbb{C}$. Deze groep heeft dus drie reeele parameters; bepaal een gelijkaardige voorstelling voor $O(2)$. Gegeven lineaire operatoren $A : \mathcal{H}_1 \to \mathcal{H}_3$ en $B : \mathcal{H}_2 \to \mathcal{H}_4$ dan kunnen we operatoren

$$A \oplus B : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_3 \oplus \mathcal{H}_4$$

en

$$A \otimes B : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_3 \otimes \mathcal{H}_4$$

definieren aan de hand van

$$A \oplus B(v_1 \oplus v_2) = A(v_1) \oplus B(v_2)$$

en

$$A \otimes B(v_1 \otimes v_2) = A(v_1) \otimes B(v_2).$$

Denk nu even na en overtuig je van het feit dat het tensorproduct $\otimes$ dient om gescheiden systemen samen te stellen; het is te zeggen functies in n reeele veranderlijken $f_k : (x_1, \ldots, x_n) \to \mathbb{C}$ en m (reeele) veranderlijken $g_k : (y_1, \ldots, y_m) \to \mathbb{C}$ definieren functies in $n+m$ (reeele) veranderlijken door

$$F = \sum_k a_k (f_k \otimes g_k) : \mathbb{R}^{n+m} \to \mathbb{C} : (x_1, \ldots, x_n, y_1 \ldots, y_m) \to \sum_k a_k f_k(x_1, \ldots, x_n) g_k(y_1, \ldots, y_m).$$

Hier mag $\mathbb{R}^{n+m}$ niet bekeken worden in de vectorruimte zin maar in de verzamelingtheoretische zin; in de vectorruimte taal geldt dat $\mathbb{R}^{n+m} = \mathbb{R}^n \oplus \mathbb{R}^m$. Het is een resultaat uit de analyse dat veel functies $F : \mathbb{R}^{n+m} \to \mathbb{C}$ geschreven kunnen worden als $\sum_k a_k (f_k \otimes g_k)$; met andere woorden, we hebben een complexe Hilbertruimte van functies $L_2(\mathbb{R}^{n+m})$ die gelijk is aan $L_2(\mathbb{R}^n) \otimes L_2(\mathbb{R}^m)$.

Je maakt nu de volgende oefeningen: zij $A : V \to V$ en $B : W \to W$ operatoren op eindig dimensionale vectorruimten, toon dan aan dat

$$\mathrm{Tr}(A \oplus B) = \mathrm{Tr}(A) + \mathrm{Tr}(B),\ \mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\mathrm{Tr}(B)$$

en

$$\det(A \oplus B) = \det(A)\det(B),\ \det(A \otimes B) = \det(A)^m \det(B)^n$$

waarbij $n = \dim(V)$ en $m = \dim(W)$. Indien V, W bovendien Hilbertruimten zijn, bewijs dan dat

$$(A \oplus B)^H = A^H \oplus B^H,\ (A \otimes B)^H = A^H \otimes B^H.$$

Toon aan dat de bewerkingen $\oplus, \otimes$ associatief zijn met $\{0\}, \mathbb{C}$ als eenheidselement respectievelijk; noteer met $\otimes_{\mathcal{F}}$ de afbeelding op de ruimte van Hilbertruimten gedefinieerd door $\otimes_{\mathcal{F}}(\mathcal{H}) = \mathcal{H} \otimes \mathcal{F}$. Construeer een $i_{\mathcal{F}}$ zodat $i_{\mathcal{F}} \circ (\otimes_{\mathcal{F}}) = \mathrm{id}$ waarbij id de identiteitstransformatie is. Toon aan dat $\otimes_{\mathcal{F}}$ niet surjectief is tenzij $\mathcal{F} = \mathbb{C}$ wat aantoont dat er geen $p_{\mathcal{F}}$ bestaat zodat $(\otimes_{\mathcal{F}}) \circ p_{\mathcal{F}} = \mathrm{id}$. Doe hetzelfde voor $\oplus_{\mathcal{F}}$ en merk op dat noch $\oplus, \otimes$ commutatief zijn. Hier hebben we dus een voorbeeld van een afgeleide bewerking die niet commutatief is en slechts een linker inverse heeft en geen rechterinverse. Introduceer nu het concept van een anti-Hilbertruimte $\mathcal{F}^{\otimes}$ als formeel rechterinverse voor $\mathcal{F}$, het is te zeggen

$$\mathcal{F} \otimes \mathcal{F}^{\otimes} = \mathbb{C}.$$

In dat geval is $i_{\mathcal{F}}$ gelijk aan $\otimes_{\mathcal{F}^{\otimes}}$ op het beeld van $\otimes_{\mathcal{F}}$. Doe hetzelfde voor $\oplus$ en borduur hier op voort door bijvoorbeeld het concept van niet-commutatieve vectorruimtes te bestuderen. Merk op dat een niet-commutatieve "som" ook verschijnt in Cantor's constructie van de ordinaal getallen. Noteer dat we het concept van een anti-verzameling ook al kunnen definieren hadden voor het Cartesisch product op het niveau van verzamelingenleer. Later zul je begrijpen dat het concept van Hilbertruimte gebruikt wordt in de fysica om gescheiden deeltjes te beschrijven daar waar het concept van anti-Hilbertruimte gebruikt kan worden om deeltjes met elkaar te laten botsen en te verworden tot één nieuw deeltje of energiepakketje. Botsen en niet botsen kunnen weeral approximaties zijn te wijten aan de puntbeschrijving van een deeltje, zoek eens naar een wiskundig concept dat hiertussen ligt!

Met deze kennis kan je ruimschoots problemen uit de standaard vlakke meetkunde oplossen; hier zijn erg sterke stellingen mogelijk die in het algemeen geval helemaal niet geldig zijn omwille van topologische en metrische complicaties. De magie van de vlakke meetkunde schuilt helemaal in de vectorruimte structuur. Bijvoorbeeld, op een boloppervlak snijden elke twee "rechte" lijnen, gedefinieerd als de snijlijn van de bol met een vlak dat door het centrum gaat, mekaar in twee diametraal tegengestelde punten. In het tweedimensionaal vlak daarintegen bestaat er een voorkeursklasse van parallelle rechten gedefinieerd door de eigenschap dat ze mekaar niet snijden. In de driedimensionale Euclidische ruimte noemen we *translaties* van een tweedimensionale deelruimte een vlak, een ééndimensionale een rechte en een nuldimensionale een punt. In de Euclidische ruimte is er maar één nuldimensionale deelruimte en dat is het neutraal element $\{0\}$, ook wel de oorsrong genoemd. Een rechte wordt in parametervorm dus gegeven door $r = \{\lambda.v + a|\lambda \in \mathbb{R},\ v, a \in \mathbb{R}^3\}$ en een vlak door $vl = \{\lambda.v + \mu.w + a|\lambda, \mu \in \mathbb{R},\ v, w, a \in \mathbb{R}^3\}$ waarbij de vectoren v, w altijd genormeerd en orthogonaal gekozen kunnen worden. Een rechte kan altijd geschreven worden als de doorsnede van twee vlakken en een vlak wodt volledig bepaald door een punt en haar loodvector. Om dit op een hoger niveau te begrijpen introduceren we het totaal anti-symmetrisch symbool ϵ_{ijk} waarbij $\epsilon_{123} = 1$ en $\epsilon_{ijk} = \text{sign} \begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix}$ waar dit laatste symbool de permutatie noteert die 1 afbeeldt op i, 2 op j en 3 op k. Dus, in deze notatie,

$$\det(A) = \epsilon_{ijk} A^1_i A^2_j A^3_k$$

voor een 3×3 matrix A. Hier i, j, k zijn indices ten opzichte van een orthonormale basis en alsdusdanig heeft het ϵ symbool een meetkundige betekenis (die van volume). 'Inderdaad, $\delta^{ik}\epsilon_{klm}v^l w^m = (v \times w)^i$ is een vector die per definitie loodrecht staat op de vectoren v, w (gebruik de anti-symmetrie) en δ^{kl} is de inverse van het δ_{ij} symbool,het is te zeggen

$$\delta^{ik}\delta_{kj} = \delta^i_j,\ \ \delta_{ik}\delta^{kj} = \delta^i_j.$$

De lengte in het kwadraat

$$(v \times w)^2 = \epsilon_{lmn}\delta^{li}\epsilon_{ijk}v^m v^j w^n w^k = (\delta_{mj}\delta_{nk} - \delta_{mk}\delta_{nj})v^m v^j w^n w^k$$

wat gelijk is aan

$$v^2 w^2 - (\langle v|w\rangle)^2$$

en dit heeft de meetkundige interpretatie van oppervlakte van het parallellepipedum opgespannen door de vectoren v, w. Dus we hebben expliciet een eenheidsvector geconstrueerd

$$n = \frac{v \times w}{||v \times w||}$$

loodrecht op de tweedimensionale deelruimte opgespannen door v, w en met een orientatie zodat v rechtshandig draait in w. Het vlak

$$vl = \{\lambda.v + \mu.w + a | \lambda, \mu \in \mathbb{R},\ v, w, a \in \mathbb{R}^3\}$$

bestaat dan precies uit de punten x die voldoen aan de vergelijking

$$\langle n | x - a \rangle = 0$$

wat een lineair systeem is in x. Indien $x = (x_1, x_2, x_3)$ is deze vergelijking van de vorm

$$n_1(x_1 - a_1) + n_2(x_2 - a_2) + n_3(x_3 - a_3) = 0.$$

Geef de loodrichting op het vlak $2x - 3y + z - 12 = 0$ en bepaal het punt in het vlak dat het dichtst bij de oorsprong ligt.

Andere belangerijke vergelijkingen betreft de zogenaamde kwadratische vergelijkingen waarvan het meest symmetrische voorbeeld gegeven is door het boloppervlak: dit is gedefinieerd als de plaats van alle punten x die zich op een vaste afstand r van het punt a bevinden. De relevante vergelijking luidt aldus

$$||x - a||^2 = r^2$$

of in coordinaten

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 - a_3)^2 = r^2.$$

Op dezelfde manier is de vergelijking van een cirkel gegeven door

$$(x - a)^2 + (y - b)^2 = r^2.$$

Toon op twee verschillende manieren aan dat de doorsnede van een bol met een vlak leeg is, een punt of een cirkel. Bewijs dat hetzelfde geldt voor de doorsnede van twee bollen. Deze eigenschappen gelden allemaal niet voor de gekromde meetkundes die we later gaan bestuderen. Later zullen we het concept van hoek bestuderen alsook een paar stellingen formuleren voor driehoeken in algemene gekromde ruimten, waarvan het vlakke geval speciaal is. Er gelden erg gesophisticeerde resultaten in de vlakke meetkunde welke de lezer zou willen bestuderen, oude boeken doen hier een uitstekende dienst doch ik ben de mening toegedaan dat het immens veel belangrijker is te verstaan hoe algemene meetkunde "echt" werkt in plaats van een specialist te worden in het oplossen van lineaire en kwadratische vergelijkingen. Toon aan dat het aantal simplices $a(n)$ in hetwelke een n dimensionale kubus onderverdeeld kan worden gelijk is aan $n!$. Dit is makkelijk in te zien door de coefficienten $\lambda_1 \ldots \lambda_n$ te ordenen van klein naar groot; het is te zeggen $\lambda_{i_1} \geq \lambda_{i_2} \geq \lambda_{i_n}$. Er zij zo'n $n!$ mogelijkheden en elke mogelijkheid definieert een n dimensionaal simplex door $\lambda_{i_k} \to \lambda_{i_k} - \lambda_{i_{k+1}}$ voor $k = 1 \ldots n - 1$ en $\lambda_{i_n} \to \lambda_{i_n}$. Vermits de determinant van deze afbeelding 1 is bewaart ze dus het volume wat aantoont dat elke geordende regio het volume heeft van een simplex. Een ander bewijs verloopt via integralen die we later zullen zien; je merkt op dat het volume van een n dimensionaal simplex met lengte r gelijk is aan $r^n b(n)$ en dus is

$$b(n+1) = b(n) \int_0^1 dr r^n = \frac{b(n)}{n+1}$$

en aldus is $b(n) = \frac{1}{n!}$. Daarom is het aantal simplices in een kubus gelijk aan $n!$.

Zover wat betreft de studie van Hilbertruimten, nu komen we tot de behandeling van de theorie van lineaire operatoren alsook topologieen op ruimten van operatoren. Dit onderwerp was erg belangrijk in de oude formulering van quantum mechanica door Heisenberg, Jordan en consorten en berust vooral op een éénmans inspanning door Von Neumann. Laten we eerst een tweetal topologieen bestuderen op een algemene Hilbertruimte $\mathcal{H}$ vooraleer over te stappen naar de ruimte van (begrensde) operatoren. Op $\mathcal{H}$ hebben we reeds de norm topologie bestudeerd en je bewijst nu devolgende twee resultaten:

- Een verzameling in een eindig dimensionale Hilbertruimte is compact in de normtopologie *als en slechts als* ze gesloten en begrensd is; dus we hebben het omgekeerde van een algemener resultaat voor metrische ruimten waar we bewezen hebben dat een compacte ruimte gesloten en begrensd is.

- In een oneindig dimensionale Hilbertruimte met aftelbare orthonormale basis $(e_n)_{n \in \mathbb{N}}$ hebben we dat de eenheidsbol niet compact is in de normtopologie. Hint: toon aan in één lijntje dat de rij $(e_n)_{n \in \mathbb{N}}$ geen convergente deelrij heeft. Een Hilbertruimte met aftelbare basis noemen we separabel of scheidbaar.

We komen nu tot een zwakkere topologie die alle voordelen van het eindig dimensionale geval overhevelt naar het separabele geval. Deze topologie wordt gegenereerd door de lineaire functionalen ω welke lineaire afbeeldingen zijn van $\mathcal{H}$ naar $\mathbb{C}$, een ééndimensionale bril dus waarmee je de Hilbertruimte bekijkt. De ruimte van lineaire functionalen is opnieuw een vectorruimte die we de algebraisch duale noemen; wij zijn echter alleen geinteresseerd in die functionalen die continue zijn met betrekking tot de normtopologie. Deze vormen opnieuw een vectorruimte welke we de topologisch duale $\mathcal{H}^\star$ noemen. Toon aan dat in een eindig dimensionale Hilbertruimte de topologisch en algebraisch duale samenvallen. Een belangrijk resultaat omtrent continue functionalen is dat ze begrensd zijn, het is te zeggen dat

$$|\omega(v)| \leq C||v||$$

voor een bepaalde $C > 0$; omgekeerd toon je aan dat elke begrensde lineaire functionaal continue is in de norm topologie. Ik zal hier een formeel bewijs van geven en je moet goed opletten hoe de argumentatie gemaakt wordt: je zult immers dergelijke redeneringen zelf moeten maken in de hiernavolgende oefeningen. We redeneren als volgt: we tonen aan dat indien de functionaal niet begrensd is dat hij dan ook niet continue is. Veronderstel dus dat er een reeks van genormeerde vectoren v_n bestaat zodat $\omega(v_n) \to \infty$ in de limiet voor n naar ∞. Door het nemen van een deelrij kunnen we onderstellen dat $\omega(v_n) > n^2$ en de rij van vectoren $w_k = \sum_{n=0}^{k} \frac{1}{n^2} v_n$ convergeert naar $w = \sum_{n=0}^{\infty} \frac{1}{n^2} v_n$ welke een eindige norm heeft (omdat de reeks $\sum_{n>0} \frac{1}{n^2}$ convergeert) terwijl $k < \omega(w_k) \to \infty$ wat in strijd is met de continuiteit.

Vermits een continue lineaire functionaal een ééndimensionale bril geeft op een Hilbertruimte moet zij dus overeenstemmen met een projectie op een vector v; het is te zeggen

$$\omega(w) = \langle v|w\rangle$$

met $||v|| < \infty$ en je moet dit bij wijze van oefening aantonen. Geometrisch gezien is dit evident gegeven dat ω volledig bepaald is door haar kern $W = \{w|\omega(w) = 0\}$ en de actie op haar normale loodvector $\frac{v}{||v||}$. Dit motiveert de volgende definitie, de verzamelingen

$$\mathcal{O}_{\epsilon;v_1,\ldots,v_n}(w) = \{w' \,|\, |\langle w - w'|v_i\rangle| < \epsilon \,\text{voor}\, i = 1 \ldots n\}$$

zijn open omgevingen van w in de dimensies bepaald door v_j en vormen een basis voor de *zwakke* of $\star$-topologie.

Open omgevingen van w in de zwakke topologie controleren dus de norm van de projectie nan de verschilvector $w - w'$ opeen eindig dimensionale deelruimte en laten de componenten loodrecht op die deelruimte ongemoeid. Vermits de normtopologie alle dimensies controleert is zij dus sterker in het separabele geval; in het bijzonder is elke open verzameling in de zwakke topoloie open in de norm topologie wat formeel volgt uit de ongelijkheid

$$|\langle w - w'|v_i\rangle| \leq ||w - w'||||v_i||.$$

Het is dus evident dat dezelfde resultaten gelden in de zwakke topologie voor separabele Hilbertruimten dan voor de normtopologie in het eindig dimensionaal geval. In het bijzonder geldt dus dat een verzameling compact is in separabele Hilbertruimten in de zwakke topologie als ze gesloten is in de zwakke topologie en begrensd in de norm. Toon aan dat indien een verzameling begrensd is in de norm dat ze dan gesloten is in de norm topologie als en slechts als ze dit is in de zwakke topologie. Dit is evident omdat de begrensdheid een oneindig aantal dimensies controleert zodat effectief er maar een eindig aantal overschieten en die worden gecontroleerd door de zwakke topologie. Dus de eenheisbol is gesloten en compact in de zwakke topologie wat bekend staat als de stelling van Hahn-Banach. Het omgekeerde is ook waar, een verzameling die compact is in de zwakke topologie is altijd begrensd in de norm.

We draaien nu ons hoofd richting operator topologieen voor lineaire afbeeldingen $A : \mathcal{H} \to \mathcal{H}$ alsook gaan we enkele belangrijke stellingen bewijzen omtrent operatoren die een bijzondere metrische functie hebben zoals bijvoorbeeld de unitaire operatoren uit het vorige hoofdstuk. In het bijzonder vragen we ons af wanneer we een orthonormale basis van eigenvectoren vinden en wat de eventuele beperkingen zijn op de eigenwaarden. Er zijn vele topologieen voorhanden op welbepaalde klassen van operatoren welke allemaal equivalent zijn in het eindig dimensionale geval; we beginnen met de supremum norm-topologie:

$$||A||_{\mathrm{sup}} = \sup_{||v||=1} ||Av||.$$

In het geval deze eindig is, is de operator A begrensd en de ganse theorie van begrensde operatoren is gegoten in het kader van de $C^{\star}$-algebra's. We gaan hier niet verder op in omdat de interessante operatoren uit de fysica allemaal onbegrensd zijn. Twee andere topologieen zijn belangrijk, de sterke en zwakke $\star$ topologie. De eerste wordt gedefinieerd aan de hand van de open omgevingen

$$\mathcal{O}_{\epsilon;v_1,\ldots v_n}(A) = \{B| \; ||(B-A)v_k|| < \epsilon \,\mathrm{voor}\, k = 1\ldots n\}$$

waar de laatste open omgevingen

$$\mathcal{O}_{\epsilon;v_1,\ldots v_n,w_1,\ldots,w_n}(A) = \{B| \; |\langle (B-A)v_k|w_k\rangle| < \epsilon \,\mathrm{voor}\, k = 1\ldots n\}$$

heeft. Je toont aan dat beide topologieen aan de Hausdorff eigenschap voldoen en dat de zwakke-$\star$ topologie zwakker is dan de sterke.

We introduceren eerst enkele belangrijke noties voor lineaire operatoren op separabele Hilbertruimten. Zoals je kunt raden duiken hier enkele subtiliteiten op die we in het voorgaande hoofdstuk niet ontmoet hebben omdat een eindig aantal dimensies immens veel eenvoudiger zijn. Bijvoorbeeld hebben algemene operatoren A een domein $\mathcal{D} \subset \mathcal{H}$, welk we dicht veronderstellen in de norm topologie, waarop A goed gedefinieerd is. De toegevoegde operator $A^{\dagger}$ van A is dan gedefinieerd als volgt: beschouw de deelruimte $\mathcal{D}^{\star}$ van vectoren v zo dat

$$|\langle v|Aw\rangle| < C(v)||w||$$

voor alle $w \in \mathcal{D}$. Dan hebben we dat de functionaal $w \to \langle v|Aw\rangle$ continue uitgebreid kan worden tot gans $\mathcal{H}$ omdat $\mathcal{D}$ dicht is. We weten dan dat er een vector z bestaat zo dat

$$\langle v|Aw\rangle = \langle z|w\rangle$$

en we definieren $A^{\dagger}v = z$ en het is eenvoudig te controleren dat $A^{\dagger}$ een lineaire operator is. Daarom is het domein van $A^{\dagger}$ gelijk aan $\mathcal{D}^{\star}$. De volgende gevallen zijn buitengewoon belangrijk:

- $A = A^{\dagger}$ en $\mathcal{D} = \mathcal{D}^{\star} = \mathcal{H}$ in welk geval de operator zelftoegevoegd is,

- $AA^{\dagger} = A^{\dagger}A$ en $\mathcal{D} = \mathcal{D}^{\star} = \mathcal{H}$ in welk geval de operator normaal is,

- $UU^\dagger = U^\dagger U = 1$ en $\mathcal{D} = \mathcal{D}^\star = \mathcal{H}$ in welk geval de operator unitair is,

- $P^2 = P = P^\dagger$ en $\mathcal{D} = \mathcal{D}^\star = \mathcal{H}$ in welk geval de operator een Hermitische projectie is.

Je controleert dat in het eindig dimensionale geval $A^\dagger = A^H$ en dat unitaire operatoren gewoon veralgemeningen zijn van elementen uit $U(n)$. Bepaal ook het domein van de operator gegeven door $Ae_n = ne_n$ voor $n \in \mathbb{N}$ waar e_m een orthonormale basis is en toon aan dat dit dicht is in $\mathcal{H}$; toon aan dat $\mathcal{D} \subseteq \mathcal{D}^\star$ en dat $A = A^\dagger$ op $\mathcal{D}$. We werken nu naar twee verschillende stellingen toe: de eerste betreft de uitbreiding van een bijzondere klasse van operatoren tot Hermitische operatoren, waar een uitbreiding van een operator een nieuwe operator is met een groter domein die samenvalt met de oude operator op diens domein. Een tweede resultaat zegt dat een normale operator ontbonden kan worden in Hermitische projectie operatoren in de zwakke $\star$ topologie.

Het belang van de eerste stelling ligt letterlijk in de tweede; deze zegt, in het eindig dimensionaal geval, dat elke normale matrix diagonalizeerbaar is en dat de eigenvectoren loodrecht op elkaar staan. Dit laatste aspect is van primordiaal belang om een kans theoretische interpretatie te krijgen van vlakke meetkunde zoals dat gebruikelijk is in de kwantum mechanica. Toon aan, bij wijze van oefening, dat in een eindig aantal dimensies Hermitische operatoren reeele eigenwaarden hebben, unitaire operatoren eigenwaarden die op de eenheidscirkel liggen in het complexe vlak en finaal normale operatoren kunnen eender welke complexe eigenwaarde hebben. Zoals gezegd is er een verband tussen unitaire en zelftoegevoegde operatoren en in die zin is het makkelijker eerst de vraag te stellen naar unitaire uitbreidingen van zogenaamde partiele isometrieen V met een domein $\mathcal{D}$ dat niet noodzakelijk dicht is. Een partiele isometrie is gedefinieerd door de eigenschap dat

$$\langle V(v)|V(w)\rangle = \langle v|w\rangle$$

voor alle $v, w \in \mathcal{D}$. Bij continuiteit kunnen we V dus uitbreiden tot de sluiting $\overline{\mathcal{D}}$ van $\mathcal{D}$ wat resulteert in een lineair homeomorphisme tussen $\overline{\mathcal{D}}$ en $\overline{\mathrm{Im}(V)}$ waar $\mathrm{Im}(V) = \{Vw|w \in \mathcal{D}\}$ het beeld is van V. Het is duidelijk dat alleen in het geval dat het orthogonaal complement $\mathcal{D}^\perp = \{w|\langle w|v\rangle = 0 \; \forall v \in \mathcal{D}\}$ dezelfde dimensie heeft dan het orthogonaal complement $(\mathrm{Im}(V))^\perp$ dat we een uitbreiding kunnen bekomen aan de hand van een unitaire operator $W : \overline{\mathcal{D}}^\perp \to (\mathrm{Im}(V))^\perp$ wat de unitaire operator $U = V \oplus W : \mathcal{H} \to \mathcal{H}$ definieert en deze is duidelijk een unitaire uitbreiding van U. De lezer noteert dat voor een deelruimte W, $W^\perp$ gesloten is in de zwakke en daarom ook in de norm topologie; de deelruimte $W^{\perp\perp} := (W^\perp)^\perp$ is bovendien de zwakke sluiting van W.

Je merkt dus op dat een partiele isometrie veel unitaire uitbreidingen kan hebben indien de dimensies van de orthogonale complementen hetzelfde zijn of helemaal geen indien dit niet het geval is. Nu komen we terug tot het verband tussen

Hermitische en unitaire operatoren; Von Neumann kende de zogenaamde Cayley transformatie tussen Hermitische en unitaire operatoren in eindig dimensionale Hilbert ruimten. Hier wordt een zelftoegevoegde operator A afgebeeldt op

$$U = (A - i1)(A + i1)^{-1}$$

waarbij $(A \pm i1)$ invertiebel is in een eindig aantal dimensies vermits het stelsel $Av = \mp iv$ geen niet nul oplossing heeft. Je ziet dit in door te observeren dat

$$\mp i||v||^2 = \langle v|Av\rangle = \langle Av|v\rangle = \pm i||v||^2$$

wat impliceert dat $v = 0$. Bovendien commuteert $(A + i1)$ met $(A - i1)$ wat tot gevolg heeft dat U unitair is zoals een korte berekening bevestigd. Von Neumann vroeg zich af welke voorwaarden op A gesteld zouden moeten worden zodat U een partiele isometrie is. In dat geval kan men deze uitbreiden tot een unitaire operator die een Hermitische operator definieert aan de hand van de omgekeerde Cayley transformatie:

$$A = -i(U + 1)(U - 1)^{-1}.$$

De operator $A \pm i1$ moet injectief zijn zodat een inverse genomen kan worden wat de voorwaarden $\mathcal{D} \subseteq \mathcal{D}^\star$ en $A = A^\dagger$ op $\mathcal{D}$ suggereert zodat het voorgaande argument nog steeds doorgaat; deze zijn de definierende eigenschappen van een symmetrische operator. In het separabele geval is het echter niet noodzakelijk zo dat $A \pm i1$ surjectief is. De Cayley transformatie is dus een lineaire afbeelding

$$U : \mathrm{Im}(A + i1) \to \mathrm{Im}(A - i1)$$

en we moeten nu drie zaken aantonen : (a) controleer dat U een partiele isometrie is (b) sluit de operator $\mathrm{Im}(A \pm i1)$ en finaal (c) controleer of $\mathrm{Im}(A + i1)^\perp$ en $\mathrm{Im}(A - i1)^\perp$ dezelfdee dimensie hebben. Betreffende (a) merk je op dat

$$\langle U(A+i1)v|U(A+i1)w\rangle = \langle (A-i1)v|(A-i1)w\rangle = \langle Av|Aw\rangle + i\langle v|Aw\rangle - i\langle Av|w\rangle + \langle v|w\rangle$$

en deze laatste uitdrukking is met behulp van de symmetrie van A gelijk aan

$$\langle Av|Aw\rangle + \langle v|w\rangle = \langle (A + i1)v|(A + i1)w\rangle$$

voor alle $v, w \in \mathcal{D}$. Gewoonlijk, in de literatuur, sluit men de operator A vooraleer de Cayley transformatie te maken alhoewel dit niet nodig is; U extendeert triviaal tot een operator

$$U : \overline{\mathrm{Im}(A + i1)} \to \overline{\mathrm{Im}(A - i1)}$$

en we hebben (c) nodig om U uit te breiden tot een unitaire operator op $\mathcal{H}$. Deze laatste voorwaarde kan beter geformuleerd worden als

$$\mathrm{Im}(A \pm i1)^\perp = \mathrm{Ker}(A^\dagger \mp i1).$$

Inderdaad,

$$\langle w|(A \pm i1)v\rangle = 0$$

voor alle $v \in \mathcal{D}$ is equivalent aan $w \in \mathcal{D}^\star$ en

$$\langle (A^\dagger \mp i1)w|v\rangle = 0.$$

Dit is waar als en slechts als $(A^\dagger \mp i1)w = 0$ vermits $\mathcal{D}$ dicht is in $\mathcal{H}$; bij definitie is $\mathrm{Ker}(B) = \{w|Bw = 0\}$ wat het gewenste resultaat oplevert.

We hebben net ons eerste belangrijke resultaat bewezen: symmetrische, dicht gedefinieerde operatoren hebben zelftoegevoegde uitbreidingen als en slechts als de dimensies van $\mathrm{Ker}(A^\dagger \mp i1)$ gelijk zijn aan elkaar. Nu werken we naar ons tweede belangrijk resultaat betreffende normale operatoren A: er bestaat een projectiewaardige maat $dP(z)$ op het complexe vlak zodat in de zwakke $\star$ topologie geldt dat

$$A = \int_{\mathbb{C}} z\, dP(z).$$

Hier komen we voor de eerste maal een integraal tegen, iets wat we in volledig detail in het volgende hoofdstuk gaan bestuderen; we zullen hier slechts raken aan de meest fundamentele definities. Stel nu even dat je dergelijk resultaat wil bekomen dan is de operator $(A - z1)$ niet invertiebel indien $dP(z) \neq 0$; logisch gezien hebben we drie mogelijkheden:

- $(A - z1)$ is niet injectief, noch surjectief; in dit geval behoort z tot het discreet spectrum,

- $(A - z1)$ is niet injectief, maar surjectief; in dit geval behoort z tot het residuaal spectrum,

- $(A - z1)$ is injectief, maar niet surjectief; in dat geval behoort z tot het continue spectrum.

Voor normale operatoren hebben we dat het residuaal spectrum leeg is wat we nu even gaan bewijzen. Noteer dat als A normaal is, dat dan $A_z = A - z1$ ook deze eigenschap heeft; bovendien is A injectief als en slechts $A^\dagger$ het is, wat bewezen kan worden door te stellen dat $Av = 0$ als en slechts als $A^\dagger v = 0$. Let op dat surjectiviteit van A niet impliceert dat $A^\dagger$ surjectief is. Veronderstel dat z tot het residuaal spectrum behoort, dan hebben we dat

$$\langle v|A_z w\rangle = 0$$

voor alle w impliceert dat $v = 0$ omwille van de surjectiviteit van A_z. Maar dan is $\mathrm{Ker}(A_z^\dagger) = \mathrm{Ker}(A_z) = 0$ wat een contradictie is. Dus we hebben bewezen dat het residuaal spectrum leeg is. In het geval dat z tot het discreet spectrum behoort bestaat er een unieke Hermitische projectie operator P_z op

$\text{Ker}(A_z)$. P_z commuteert met A, $AP_z = P_zA = zP_z$ vermits $\langle v|AP_zw\rangle = z\langle v|P_zw\rangle = \langle \overline{z}P_zv|w\rangle = \langle A^\dagger P_zv|w\rangle = \langle v|P_zAw\rangle$ en dezelde commutatie relaties gelden tussen $A^\dagger$ en P_z gegeven dat de laatste Hermitisch is. Bovendien, stel dat $z \neq z'$ en beide behoren tot het discrete spectrum, dan geldt dat $P_zP_{z'} = 0$ wat volgt uit

$$zP_zP_{z'} = AP_zP'_z = z'P_zP_{z'}.$$

Dit lijkt al sterk op het resultaat dat we willen behalen in de zin dat op separabele Hilbertruimten het discreet spectrum hoogstens uit een aftelbaar aantal punten bestaat. We geven het voorbeeld van een begrensde lineaire operator voor dewelke men kan bewijzen dat het spectrum compact is. Gegeven dat $Ae_n = \frac{1}{n}e_n$ waar $n > 0$ en e_m een orthonormale basis: het discreet spectrum is gegeven door $\{\frac{1}{n}|n \in \mathbb{N}_0\}$ en 0 behoort tot het continue spectrum gegeven dat bijvoorbeeld de vector $\sum_{n=1}^{\infty} \frac{1}{n}e_n$ niet behoort tot het beeld van A. Dus het continue spectrum kan "maat nul" hebben en niet bijdragen tot de spectrale decompositie.

Het continue spectrum is duidelijk leeg voor normale operatoren op eindig dimensionale Hilbertruimten en de lezer zou nu makkelijk moeten kunnen aantonen dat

$$A = \sum_{z\in\sigma_d(A)} zP_z$$

waar $\sum_{z\in\sigma_d(A)} P_z = 1$ en $\sigma_d(A)$ noteert het discreet spectrum bestaande uit de eigenwaarden van het vorige hoofdstuk. De volgende notatie is handig: gegeven een éénheidsvector v, definieer door

$$P = vv^\dagger$$

de operator $Pw = v\langle v|w\rangle$. Bewijs dat P een rank één Hermitische projector is en in geval $AP = zP$ hebben we dat $Av = zv$ wat impliceert dat v een eigenvector is. De ganse complexiteit betreffende integralen ligt dus in de behandeling van het continue spectrum in een oneindig aantal dimensies. We zullen de constructie hier niet in haar volste algemeenheid geven vermits dit enkele technische complicaties met zich zou meebrengen die de aandacht alleen maar van het hoofd idee zou afleiden. Noteer ook dat we hier van het fundamenteel theorema van de algebra gebruik gemaakt hebben hetwelk we slechts zullen bewijzen in het hoofdstuk over complexe analyse.

Indien z tot het continue spectrum behoort hebben we, in het bijzonder, dat het beeld van de eenheidsbol onder A_z geen open omgeving van de oorsprong bevat. Anders zou A_z surjectief zijn: dus er bestaat een rij van eenheidsvectoren v_n zodat

$$||A_zv_n|| \to 0$$

in de limiet voor n naar ∞. Dus elementen in het continue spectrum geven aanleiding tot benaderende eigenvectoren. We hebben dat $\text{Im}(A_z)^\perp$ nul is vermits

A_z injectief is en daarom is $\mathrm{Im}(A_z)$ dicht in $\mathcal{H}$. Bovendien hebben we dat als $z \neq z'$ dan

$$\lim_{n,m\to\infty} \langle v_n | w_m \rangle = 0$$

waar $(v_n)_{n\in\mathbb{N}}$ correspondeert met A_z en $(w_n)_{n\in\mathbb{N}}$ met $A_{z'}$ hetgeen lijkt op de eigenschap van Hermitische projectoren voor het discreet spectrum.

We komen nu tot de constructie van de spectrale maat: gegeven een meetbare deelverzameling $\mathcal{O} \subseteq \mathbb{C}$, we definieren $P_{\mathcal{O}}$ als de kleinste Hermitische projectie operator met de eigenschap dat als $z \in \sigma(A) \cap \mathcal{O}$ en $(v_n)_{n\in\mathbb{N}}$ is een reeks van (benaderende) eigenvectoren voor z, dan $||P_{\mathcal{O}}(v_n) - v_n|| \to 0$ in de limiet voor $n \to \infty$. Een meetbare verzameling is geconstrueerd als volgt:

- elke open verzameling is meetbaar,
- het complement van een meetbare verzameling is meetbaar,
- de oneindige unie van meetbare verzamelingen is meetbaar.

We komen terug tot de behandeling van de fijne aspecten van maattheorie in hetvolgende hoofdstuk en geven nu intuitie mee omtrent het verdere verloop van het bewijs. Het is duidelijk dat

$$P_{\mathcal{O}} P_{\mathcal{V}} = P_{\mathcal{O}\cap\mathcal{V}}$$

en je moet dit bewijzen. Gegeven een aftelbare partitie $(B_n)_{n\in\mathbb{N}}$ van $\mathbb{C}$ door middel van meetbare verzamelingen[1] definieren we de sommen

$$A_{(B_n)_{n\in\mathbb{N}}} = \sum_{n=0}^{\infty} z_n P_{B_n}$$

waar $z_n \in B_n$. De integraal is dan gedefinieerd door het verfijnen van de partitie en de rest van het bewijs bestaat erin aan te tonen dat de som convergeert in de zwakke-$\star$ topologie naar de integraal zowel als naar A. Het eerste deel is een delicate oefening, terwijl het laatste terug het keuze axioma benodigt.

Deze uiterst belangrijke stelling, gekend als de spectraal stelling, laat toe functies van normale operatoren te definieren gegeven een meetbare $f : \mathbb{C} \to \mathbb{C}$ en normale operator A. We hebben dat

$$f(A) := \int_{\mathbb{C}} f(z) dP(z)$$

waar we de spectrale decompositie

$$A = \int_{\mathbb{C}} z dP(z)$$

[1] Een partitie voldoet aan de eigenschap dat $B_n \cap B_m = \emptyset$ voor $n \neq m$ en $\cup_{n=0}^{\infty} B_n = \mathbb{C}$.

gebruikt hebben. Er zijn twee belangrijke veralgemeningen van deze stelling: de eerste bestaat erin de complexe getallen te vervangen door de quaternionen $\mathbb{RQ}$ en quaternion bi-modulen te beschouwen met een quaternion waardig scalair product wat een natuurlijke veralgemening is van het begrip complexe Hilbertruimte. De spectraal stelling wordt op een gepaste manier veralgemeend. Een tweede veralgemening bestaat erin de voorwaarde

$$\langle v|v\rangle \geq 0$$

te laten vallen en toe te laten dat deze grootheid negatief wordt. Later bestuderen we eindig dimensionale voorbeelden van zogenaamde Nevanlinna ruimten in het kader van Lorentziaanse meetkunde.

Dit beeindigt onze studie van lineaire ruimten en dito functies; zoals gewoonlijk is er wel wat meer te vertellen maar dit zijn toch de belangrijkste resultaten. We gaan nu over naar de algemene analyse van functies die de basis vormt voor de niet-lineaire meetkunde en cosmologie.

Chapter 8

Meer dimensionale analyse.

In dit hoofdstuk bestuderen we speciale functies $g : \mathbb{R}^n \to \mathbb{R}^m$; in het bijzonder definieren we de differentiaal Dg, vectorvelden, duale velden, algemene tensor velden, de uitwendige afgeleide en algemene integratie theorie. Al deze zaken kunnen veralgemeend worden naar de oneindig dimensionale context waar men de zogenaamde Fréchet afgeleide definieert, maar we vermijden deze situatie met haar inequivalente topologieen. De norm die we gebruiken is de standaard norm geassocieerd aan een reeel scalair product. De coordinaten in $\mathbb{R}^m$ geassocieerd aan een orthonormale basis worden genoteerd door x'^{μ}, x'^{ν} waar deze met betrekking tot een orthornormale basis in $\mathbb{R}^n$ genoteerd worden door x^{α}, x^{β}. Gegeven $g : \mathbb{R}^n \to \mathbb{R}^m$, de notatie $g^{\mu}(x^{\alpha})$ heeft dus een eenduidige betekenis en we beginnen met het definieren van *partiele* afgeleiden $\partial_\alpha = \frac{\partial}{\partial x^\alpha}$. Deze zijn vastgelegd door de voorwaarde dat

$$\lim_{h \to 0} \frac{||f(x + he_\alpha) - f(x) - \partial_\alpha f(x)h||}{h} = 0$$

waar h een reeel getal is. In het geval $n = m = 1$ is er maar één partiele afgeleide die we de afgeleide noemen; toon hetvolgende aan

- $\partial_x x^n = nx^{n-1}$ voor $n \geq 1$,
- $\partial_x c = 0$ voor een constante c,
- $\partial_x y^m = 0$ voor een andere variabele y.

Bewijs nu dat indien alle partiele afgeleiden bestaan in een punt x, dat f dan continue is in dat punt. Meer algemeen, de functie $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentieerbaar als en slechts als er een lineaire afbeelding $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ bestaat zodat

$$\lim_{||h|| \to 0} \frac{||f(x + h) - f(x) - Df(x)h||}{||h||} = 0$$

voor $h \in \mathbb{R}^n$. Het is kinderspel om aan te tonen dat $Df(x) = \partial_\alpha f(x) dx^\alpha$ waar $dx^\alpha(h) = h^\alpha$ en bijgevolg bestaan alle partiele afgeleiden in x als de afgeleide bestaat in datzelfde punt. Het omgekeerde is niet noodzakelijk waar: neem bijvoorbeeld een continue functie in twee veranderlijken met de beperkingen $f(x,0) = x^2, f(0,y) = y^2$ en $f(x,x) = |x|$ voor x, y voldoende dicht tegen nul. Dan bestaan alle partiele afgeleiden maar niet de totale afgeleide zoals je makkelijk kunt verifieren. Men kan hetvolgende theorema bewijzen: als alle partiele afgeleiden bestaan in een *omgeving* van x en continue zijn in die omgeving dan bestaat de totale afgeleide in x. Je kunt proberen dit bewijs zelf te leveren.

Je kunt evidenterwijs meerdere partiele afgeleiden nemen van een functie $f : \mathbb{R}^n \to \mathbb{R}^m$ en men toon hetvolgende belangrijk resultaat aan: als alle tweede orde afgeleiden $\partial_\alpha \partial_\beta f$ bestaan en continue zijn dan hebben we dat

$$\partial_\alpha \partial_\beta f = \partial_\beta \partial_\alpha f$$

en je kunt dit vrij makkelijk bewijzen vanuit de definitie van de partiele afgeleide. Je weet nu dat de matrix

$$H^\alpha_\beta(x) = \delta^{\alpha\gamma} \partial_\gamma \partial_\beta f(x)$$

symmetrisch is voor brave functies en dus heb je een compleet reeel spectrum van eigenwaarden. Definieer nu $\Delta(x)$ als het aantal positieve min het aantal negatieve eigenwaarden van $H^\alpha_\beta(x)$, de zogenaamde deficientie index van de Hessiaan, en noem een punt x kritisch indien $\partial_\alpha f(x) = 0$ voor alle α. Kritische punten en hun deficientie indices spelen een belangrijke rol en dragen topologische informatie zoals we later zullen zien wat zich uit in het werk van Brouwer en Morse. Finaal toon je aan dat de volgende regels gelden

- $\partial_\alpha (f \otimes g)(x) = (\partial_\alpha f) \otimes g(x) + f \otimes (\partial_\alpha g)(x)$,
- $\partial_\alpha (af + bg)(x) = a\partial_\alpha f(x) + b\partial_\alpha g(x)$,
- $\partial_\alpha (f(g(x))) = \partial_{g^\beta(x)} f(g(x)) \partial_\alpha g^\beta(x)$.

De eerste identiteit is gekend als de Leibniz regel, de tweede drukt lineariteit uit en de derde volgt uit de vorige twee.

We bestuderen nu differentieerbare veralgemeningen van topologische homeomorphismen: $g : \mathcal{O} \subseteq \mathbb{R}^n \to \mathcal{V} \subseteq \mathbb{R}^n$ is een C^n diffeomorphisme voor $n \in \mathbb{N}_0$ als en slechts als g is een homeomorphisme en voor elke $k \leq n$ de afgeleiden $D^k g$ en $D^k g^{-1}$ bestaan. In het vervolg bestuderen we hoofdzakelijk C^2 of C^∞ diffeomorphismen maar uitzonderingen kunnen zich voordoen. Het is natuurlijk om de afbeelding $f \circ g : \mathcal{O} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ te beschouwen welke we schrijven als

$$f^\mu(g^\beta(x^\alpha))$$

en we zijn geinteresseerd in het nemen van afgeleiden naar x^α en noteren met $x'^\beta(x^\alpha) = g^\beta(x^\alpha)$. De volgende regel heb je al bewezen

$$\partial_\alpha f(x'^\beta(x^\gamma)) = \partial'_\delta f(x'^\beta(x^\alpha)) \partial_\alpha x'^\delta(x^\gamma)$$

waarbij we Einstein sommatie gebruikt hebben in de δ indices. Vaak wordt deze regel vertaald in

$$\partial_\alpha = \frac{\partial x'^\beta}{\partial x^\alpha}\partial'_\beta$$

hetgeen leidt tot de formule

$$\frac{\partial x'^\beta}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial x'^\gamma} = \delta^\beta_\gamma$$

en idem voor x^α, x'^β omgewisseld. δ^β_γ is de tweede tensor die we ontmoeten en voldoet aan $\delta^\beta_\gamma = 1$ als $\alpha = \gamma$ en 0 anders. Op vectorruimte niveau kun je e_α identificeren met ∂_α zodat de basisvectoren een operationele betekenis verwerven. Definieer daarom

$$dx^\alpha(\partial_\beta) = \delta^\alpha_\beta$$

en vervolgens krijgt dx^α de status van element in $(\mathbb{R}^n)^\star$ de topologisch duale van $\mathbb{R}^n$. Onder een diffeomorphisme hebben we dat

$$dx'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta}dx^\beta$$

zoals het hoort omdat

$$dx'^\alpha(\partial'_\beta) = \delta^\alpha_\beta$$

een invariant is onder locale diffeomorphismen van $\mathbb{R}^n$. We definieren nu vectorvelden als differentiaal operatoren

$$\mathbf{V}(x) = V^\alpha(x)\partial_\alpha$$

zodat onder een diffeomorphisme van $\mathbb{R}^n$ de volgende transformatiewet houdt

$$\mathbf{V}'(x') = V'^\alpha(x'(x))\partial'_\alpha = V^\alpha(x)\partial_\alpha$$

wat impliceert dat

$$V'^\alpha(x'(x))\frac{\partial x^\beta}{\partial x'^\alpha} = V^\beta(x).$$

Evenzo definieren we duale velden

$$\omega = \omega_\alpha dx^\alpha$$

die transformeren als

$$\omega'_\alpha\frac{\partial x'^\alpha}{\partial x^\beta} = \omega_\beta.$$

De vectorvelden $\mathbf{V}, \mathbf{W}$ vormen een Lie-algebra wat betekent dat

$$[\mathbf{V}, \mathbf{W}] = \mathbf{V}\mathbf{W} - \mathbf{W}\mathbf{V} = \left(V^\alpha\partial_\alpha W^\beta - W^\alpha\partial_\alpha V^\beta\right)\partial_\beta$$

transformeert als een vectorveld. We kunnen nu tensor producten van vectoren en éénvormen definieren waarbij we alle vectoren links en éénvormen rechts zetten. Dit leidt tot voorwerpen zoals

$$T^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s}(x)\partial_{\alpha_1}\otimes\ldots\otimes\partial_{\alpha_r}\otimes dx^{\beta_1}\otimes\ldots\otimes dx^{\beta_s}$$

met devolgende transformatiewet

$$T'^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s}(x') = \frac{\partial x'^{\alpha_1}}{\partial x^{\gamma_1}}\ldots\frac{\partial x'^{\alpha_r}}{\partial x^{\gamma_r}}\frac{\partial x^{\delta_1}}{\partial x'^{\beta_1}}\ldots\frac{\partial x^{\delta_s}}{\partial x'^{\beta_s}}T^{\gamma_1\ldots\gamma_r}_{\delta_1\ldots\delta_s}(x).$$

We noemen dit object een (r,s) tensor met r contravariant en s covariant indices.

In de definitie van de determinant van een matrix kwamen we het fenomeen van antisymmetrisatie tegen en associeerden dit met de correcte wijze om volumes te berekenen. Nu gaan we lagere orde "determinanten" definieren die een overeenkomstige betekenis hebben voor lager dimensionale oppervlakten door terug gebruik te maken van antisymmetrizatie. In het bijzonder definieren we dus het "wedge" product van éénvormen dx^α, het heeft de eigenschappen van associativiteit, anti-symmetrie $dx^\alpha\wedge dx^\beta = -dx^\beta\wedge dx^\alpha$ en uiteindelijk $dx^{\alpha_1}\wedge dx^{\alpha_2}\wedge\ldots\wedge dx^{\alpha_k}$ definieert een $(0,k)$ covariante tensor. Vanuit deze definities is het duidelijk dat

$$dx^{\alpha_1}\wedge dx^{\alpha_2}\wedge\ldots\wedge dx^{\alpha_k} = \frac{1}{k!}\sum_{\rho\in S_k}\mathrm{sign}(\rho)dx^{\alpha_{\rho(1)}}\otimes\ldots\otimes dx^{\alpha_{\rho(k)}}$$

en gegeven dat de dimensie van $\mathbb{R}^n$ gelijk is aan n, hebben we dat de ruimte van k-vormen gegeven is door $\left(\begin{array}{c} n \\ k \end{array}\right) = \frac{n!}{(n-k)!k!}$. Gegeven een functie $f:\mathbb{R}^n\to\mathbb{R}$ we definieren de uitwendige afgeleide

$$df = \partial_\alpha f dx^\alpha$$

wat duidelijk coordinaat invariant is. Voor een k vorm

$$\mathbf{A} = A_{\mu_1\ldots\mu_k}dx^{\mu_1}\wedge\ldots\wedge dx^{\mu_k}$$

heeft men dat

$$d\mathbf{A} = \partial_\mu A_{\mu_1\ldots\mu_k}dx^\mu\wedge dx^{\mu_1}\wedge\ldots\wedge dx^{\mu_k}$$

hetgeen coordinaat onafhankelijk is gegeven dat tweede partiele afgeleiden symmetrisch zijn terwijl men het antisymmetrisch deel neemt ten gevolge van het $\wedge$-product. Bijgevolg gedraagt de ganse uitdrukking zich als een totale antisymmetrische $(0,k+1)$ tensor. Noteer dat de rand operator ∂ in simpliciale homologie en de uitwendige afgeleide beide voldoen aan

$$d^2 = 0$$

zoals de lezer kan verifieren. Dus d laat dezelfde constructie toe als die voor ∂ wat resulteert in de-Rahm cohomologie theorie. Je kunt direct aanvoelen waar dit naartoe gaat: als Ω_k een maat is voor kleine k dimensionale oppervlakken, dan is $d\Omega_k$ een maat voor $k+1$ dimensionale oppervlakten die alleen afhangt van de k-dimensionale rand. Dit is precies de inhoud van Stokes theorema dat we later zullen bestuderen in de context van Lebesgue integralen. Stokes theorema kan echter op een meta niveau bewezen worden als volgt (dit bewijs is geheel nieuw en nooit voorheen gepubliceerd bij mijn weten): gegeven een dualiteitsrelatie $\langle \Omega_k, S_k \rangle_k$ tussen k-vormen en k-simpliciale complexen die voldoet aan

- $\langle \Omega_k, S_k \cup T_k \rangle_k = \langle \Omega_k, S_k \rangle_k + \langle \Omega_k, T_k \rangle_k - \langle \Omega_k, S_k \cap T_k \rangle_k$ waarbij $S_k \cap T_k$ gelijkgesteld wordt aan $\emptyset$ indien het een lager dimensionaal simpliciaal complex is,
- $\langle \Omega_k, \emptyset \rangle_k = 0$ en $\langle a\Omega_k + b\Omega'_k, S_k \rangle_k = a\langle \Omega_k, S_k \rangle_k + b\langle \Omega'_k, S_k \rangle_k$,
- er bestaat een lineaire operator D_k die een k-vorm afbeeldt op een $k+1$-vorm zodat $\langle \Omega_k, \partial S_{k+1} \rangle_k = \langle D_k(\Omega_k), S_{k+1} \rangle_{k+1}$.

Je toont nu aan dat aan de derde voorwaarde voldaan is gegeven de eerste twee voorwaarden als en slechts als

$$|\langle \Omega_k, \partial S_{k+1} \rangle_k - \langle D_k \Omega_k, S_{k+1} \rangle_{k+1}| \leq M |\langle \Omega_k, \partial S_{k+1} \rangle_k|^{\frac{k+2}{k}}$$

voor een bepaalde $M > 0$, onafhankelijk van S_{k+1}, in de limiet $S_{k+1} \to \emptyset$. Deze voorwaarde noemen we de micro-rand voorwaarde. Vanuit de definitie van D_k en $\partial^2 = 0$ kun je nu afleiden dat $D_{k+1} D_k = 0$, en vermits d de enige lineaire operator is die men kan bedenken die coordinaat invariant is, moet $D_k = c_k d$ met c_k een reeele constante. Dus een gepaste notie van integraal voldoet aan de micro-rand voorwaarde met $D_k = d$. Er zijn verschillende kandidaten zoals de Riemann, Stieltjes en Lebesgue integraal. De filosofie achter de d operator is dat ze een k-vorm laat kruipen in de $k+1$-de dimensie op een manier zodat het $k+1$ volume alleen afhangt van de k dimensionele rand.

In de rest van dit hoofdstuk introduceren we de Lebesgue integraal die op een meer algemene klasse van functies werkt dan de differentieerbare functies. We behandelen dit op de meest algemene manier voor topologische ruimten X uitgerust met een Hausdorff topologie $\tau(X)$. Banach en Tarski hebben aangetoond dat zulke integratie theorie alleen zinvol is voor meetbare verzamelingen en een lezer geinteresseerd in het argument moet de Banach Tarski paradox erop nalezen. Deze bestaat erin een eenheids sfeer op te delen in stukken die men kan assembleren tot twee eenheidssferen. De stukken zijn uiteraard erg patologisch van aard en dus niet meetbaar. Zouden ze dat wel zijn, dan bekomen we dat het oppervlak van de eenheidsbol nul is, wat niet kan. We herinneren de lezer eraan dat de Borel-Sigma algebra $\mathcal{B}(X)$ gedefinieerd door $\tau(X)$ gegenereerd wordt door de open verzamelingen $A \in \tau(X)$ door het nemen van complementen en oneindige unies en intersecties. Gegeven $\mathcal{B}(X)$, we definieren een maat μ aan de hand van de eigenschappen

- $\mu(A) \geq 0$ voor alle $A \in \mathcal{B}(X)$,
- $\mu(\cup_{n \in \mathbb{N}_0} A_n) = \sum_n \mu(A_n)$ als voor alle $n \neq m$ geldt dat $A_n \cap A_m = \emptyset$,
- $\mu(\emptyset) = 0$.

We noemen de maat niet ontaard als en slechts als $\mu(B) > 0$ voor elke $B \in \tau(X)$, dus open verzamelingen hebben een niet nul volume. De constructie van de Lebesgue integraal is nogal uitgebreid en hangt af van sterke convergentie criteria dewelke ik bekritiseerd heb in het verleden. Een functie $f : X \to \mathbb{R}$ wordt meetbaar genoemd als en slechts als $f^{-1}(C)$, met $C \in \mathcal{B}(\mathbb{R})$, behoort tot $\mathcal{B}(X)$; dus, de inverse van een meetbare verzameling is meetbaar. Je toont aan dat een continue funtie meetbaar is door gebruik te maken van de constructie van meetbare verzamelingen uit open verzamelingen. Lebesgue splitst nu een functie in een positief en negatief deel $f = f^+ - f^-$ met $f^+ = \max\{f, 0\}$ en $f^- = \max\{-f, 0\}$ en de integraal voor meetbare, positieve functies is gegeven door:

$$\int f^{\pm}(x) d\mu(x) = \sup_{\text{partities } (A_n)_{n \in \mathbb{N}}} \sum_n \left(\inf_{x \in A_n} f^{\pm}(x) \right) \mu(A_n).$$

Voor een positieve continue functie $f : \mathbb{R} \to \mathbb{R}^+$ is de integraal niets anders dan de oppervlakte omsloten door de grafiek van f en de x as, mogelijk afgebakend door vertikale lijnen indien de integraal grenzen heeft. Deze oppervlakte wordt berekend door de x-as onder te verdelen in kleine intervalletjes en het minimum van de functie over dit interval te vermenigvuldigen met haar lengte. De integraal is dan de bovenlimiet van deze procedure door de intervalletjes kleiner en kleiner te kiezen zodat het minimum steeds groter wordt. Het verband tussen k-vormen $\Omega_k = \Omega_{\mu_1 \ldots \mu_k} dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_k}$ en maten is dan gegeven door een open verzameling te schrijven als de limiet van een bedekking door k-dimensionale kleine kubussen gedefinieerd door vectoren $v_1 \ldots v_k$ en $\int_{S_k} \Omega_k$ te schrijven als de limiet van $\sum_{\text{kubussen}} \Omega_{\mu_1 \ldots \Omega_{\mu_k}}(x_0) v_1^{\mu_1} \ldots v_k^{\mu_k}$ waarbij x_0 het hoekpunt van de kubus is vanuit hetwelk de vectoren v_j aangrijpen. Duidelijk is deze definitie invariant onder coordinaten transformaties. Toon nu hetvolgende aan voor differentieerbare éénvormen $\Omega_1 = f dx$:

- $\int$ definieert een dualiteitsrelatie tussen éénvormen en ééndimensionale simplices die voldoet aan alle drie voorwaarden (en dus in het bijzonder de micro-rand voorwaarde); in het bijzonder geldt $\int_{[a,b]} df = f(b) - f(a)$,
- de ééndimensionale versie van Stokes theorema suggereert dat d en $\int$ elkaars inverse zijn; toon aan dat $d_x \int_{[a,x]} f(s) ds = f(x) dx$ en dus is $\int$ de rechter inverse van d wanneer men zich beperkt tot éénvormen; de kern van d is de verzameling van de constante functies en dus is $\int$ geen linker inverse wanneer men zich beperkt tot de functies,
- uit $d(fg) = dfg + fdg$ volgt dat $\int_{[a,b]} (g(x) \partial_x f(x) + f(x) \partial_x g(x)) \, dx = f(b)g(b) - f(a)g(a)$ wat de regel van partiele integratie is.

De regel van partiele integratie is erg belangrijk voor het berekenen van veel ééndimensionale integralen. Buiten Stokes theorema is er nog een belangrijke eigenschap van integralen dewelke invariantie is onder de actie van diffeomorphismen wat aantoont dat de dualiteits relatie intrinsiek is. We maken dit precies in het hoofdstuk over meetkunde.

Lebesgue veronderstelde dat de geassocieerde limieten over partities eindig zijn en stelde dat de integraal van f gegeven is door

$$\int f(x)d\mu(x) = \int f^{+}(x)d\mu(x) - \int f^{-}(x)d\mu(x).$$

De uitbreiding naar complexwaardige functies toe is evident en we behandelen nu twee belangrijke stellingen. De eerste is Fubini's theorema hetwelk we continue gebruiken wanneer we integralen berekenen: laat X, Y twee Hausdorff topologische ruimtes zijn uitgerust met een Borel-sigma algebra $\mathcal{B}(X), \mathcal{B}(Y)$, dan kun je $\mathcal{B}(X \times Y)$ definieren vanuit de producttopologie op $X \times Y$ gegenereerd door open vierkanten $A \times B$ met $A \in \tau(X)$ en $B \in \tau(Y)$. Duidelijk is $\mathcal{B}(X) \times \mathcal{B}(Y) \subset \mathcal{B}(X \times Y)$ en hieruit kan je productmaten $\mu \times \nu$ definieren. Zij $f : X \times Y \to \mathbb{C}$ een meetbare functie waarvoor geldt dat $\sup_{y \in Y} \left| \int_X f(x,y)d\mu(x) \right| < \infty$ en $\sup_{x \in X} \left| \int_Y f(x,y)d\nu(y) \right| < \infty$ dan is

$$\left| \int_{X \times Y} f(x,y)d(\mu \times \nu)(x,y) \right| < \infty$$

en bovendien hebben we dat

$$\int_{X \times Y} f(x,y)d(\mu \times \nu)(x,y) = \int_X d\mu(x) \left(\int_Y f(x,y)d\nu(y) \right) = \int_Y d\nu(y) \left(\int_X f(x,y)d\mu(x) \right).$$

Noteer dat het stellen van uniforme grenzen over de partiele integralen noodzakelijk is. Beschouw bijvoorbeeld de ruimte $\mathbb{R} \times \mathbb{R}_0^+$ en de functie $f(x,y) = e^{-x^2 y}$ dan is

$$\int_{\mathbb{R}} e^{-x^2 y} dx = \sqrt{\frac{\pi}{y}}$$

en

$$\int_{\mathbb{R}_0^+} e^{-x^2 y} = \frac{1}{x^2}$$

wat niet uniform begrensd is. Dan heeft men dat

$$\int_{\mathbb{R}_0^+} dy \left(\int_{\mathbb{R}} e^{-x^2 y} dx \right) = \infty = \int_{\mathbb{R}} dx \left(\int_{\mathbb{R}_0^+} e^{-x^2 y} dy \right)$$

waar in de eerste integraal de pool van $\sqrt{\frac{\pi}{y}}$ op $y = +\infty$ verantwoordelijk is voor de divergentie en waar in het laatste geval de pool op $x = 0$ van $\frac{1}{x^2}$ de oorzaak is voor de patologie.

Het tweede theorem is Lebesgues gedomineerde convergentie stelling welke als volgt gaat: veronderstel een rij van meetbare functies f_n die puntsgewijs convergeert naar een functie f zodat $|f_n| \leq g$ voor alle n en g is integreerbaar. Dan,

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$$

en $\left|\int f_n d\mu\right| \leq \int g d\mu$. De bewijzen van deze stellingen zijn nogal saai en evident en de geinteresseerde lezer mag ze opzoeken in een goed boek. Je toont finaal aan dat de Lebesgue integraal voldoet aan de nodige dualiteitsvoorwaarden in elke dimensie zodat de toegevoegde operator van ∂ gegeven is door d.

Chapter 9

Complexe analyse en hypercomplexe analyse.

Met complexe analyse bedoelen we analyse met functies in een complexe veranderlijke; deze is grappig genoeg eenvoudiger dan de analyse van functies in één veranderlijke omdat een complexwaardige functie in één complexe veranderlijke gezien kan worden als een gesloten éénvorm in twee reeele veranderlijken en aldus kunnen we beroep doen op Stokes theorema. Dit is de magie van het getal i wanneer men analyse doet in de complexe veranderlijke $z = x + iy$. Zoals gezegd kan men z zien als een samenstelling van twee reeele veranderlijken x, y wat suggereert over te stappen naar de variabelen $z, \overline{z}$. Men kan dus de partiele differentialen

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\partial_x - i\partial_y\right), \; \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\partial_x + i\partial_y\right)$$

definieren met de eigenschappen

$$\frac{\partial}{\partial z} z = 1, \; \frac{\partial}{\partial z} \overline{z} = 0$$

en omgekeerd. Daarom voldoet een functie in een complexe veranderlijke z, $f(z)$ aan de vergelijking

$$\frac{\partial}{\partial \overline{z}} f(z) = 0$$

tenminste wanneer de eerste partiele afgeleiden naar x, y bestaan. Een complexwaardige functie in z kan geinterpreteerd worden als een eenvorm op $\mathbb{R}^2$ door middel van de formule

$$f(z) = \mathrm{Re} f(z) + i \mathrm{Im} f(z)$$

waar de afbeelding gebeurt in een vast coordinatenstelsel. De transformatie eigenschappen van $f(z)$ onder $z \to z(z')$ komen niet overeen met de gepaste transformatie wetten van de eenvorm wat betekent dat de afbeelding van complexe functies in één complexe veranderlijke naar eenvormen op $\mathbb{R}^2$ niet canonisch is. Dit kan verstaan worden vanuit het voorbeeld $f(z) = z$ en $z = z'^2$; dus

$f(z'^2) = (x'^2 - y'^2) + 2ix'y'$ waar $\frac{\partial x}{\partial x'} = 2x'$, $\frac{\partial y}{\partial x'} = -2y'$, $\frac{\partial y}{\partial x'} = 2y'$ en $\frac{\partial y}{\partial y'} = 2x'$ zodat

$$\mathrm{Re}f'(z') \neq \frac{\partial x}{\partial x'}\mathrm{Re}f(z'^2) \pm \frac{\partial y}{\partial x'}\mathrm{Im}f(z'^2)$$

waar $f'(z') = f(z'^2)$. Desalniettemin, indien men noteert dat $\mathbf{F}(z) = \mathrm{Re}f(z)dx - \mathrm{Im}f(z)dy$, dan bekom je dat de voorwaarde

$$\frac{\partial}{\partial \overline{z}} f(z) = 0$$

equivalent is aan

$$d\mathbf{F}(z) = 0, \partial^{\alpha} F_{\alpha} = 0$$

wat betekent dat de éénvorm gesloten is en nul divergentie heeft. Gebruik makende van $dz = dx + idy$ bekom je dat

$$f(z)dz = \mathbf{F}(z) + i(F_x dy - F_y dx).$$

Het imaginair deel is gesloten omdat $d(F_x dy - F_y dx) = (\partial^{\alpha} F_{\alpha})\, dx \wedge dy = 0$ en dus is deze ganse complexwaardige éénvorm gesloten. Stokes theorema zegt dan dat

$$0 = \int_S d\mathbf{F}(z) + i\, d(F_x dy - F_y dx) = \int_{\partial S} f(z)dz$$

waar S eender welk oppervlak is in $\mathbb{R}^2$ waarop $df(z)$ bestaat. Bijgevolg hangt de integraal van f over elke geslte kromme alleen af van de homologie klasse binnen dewelke f analytisch is wat betekent dat

$$\frac{\partial}{\partial \overline{z}} f(z) = 0.$$

Nu tonen we aan dat indien $f(z)$ analytisch is, dat dan $\frac{\partial}{\partial z} f(z)$ bestaat en terug analytisch is. Duidelijk moeten we alleen aantonen dat de tweede partiele afgeleiden bestaan en continue zijn; in dat geval

$$\frac{\partial}{\partial \overline{z}} \frac{\partial}{\partial z} f(z) = \frac{\partial}{\partial z} \frac{\partial}{\partial \overline{z}} f(z) = 0.$$

Om dat te bewijzen merk je op dat

$$\lim_{|z-a| \to 0} \frac{f(z) - f(a)}{z - a} = \frac{\partial}{\partial a} f(a)$$

wat volgt uit

$$\frac{f(z) - f(a)}{z - a} = \frac{(f(z) - f(a))(\overline{z - a})}{|z - a|^2} = \frac{(\mathrm{Re}f(z) - \mathrm{Re}f(a))(x - b) + (\mathrm{Im}f(z) - \mathrm{Im}f(a))(y - c)}{|z - a|^2}$$

$$\frac{-i(\mathrm{Re}f(z) - \mathrm{Re}f(a))(y - c) + i(\mathrm{Im}f(z) - \mathrm{Im}f(a))(x - b)}{|z - a|^2}$$

$$\sim \frac{\partial}{\partial a} f(a) \frac{|z-a|^2}{|z-a|^2}$$

waar we in de laatste stap $\frac{\partial}{\partial \bar{z}} f(z)$ gebruikt hebben. Door gebruik te maken van Stokes theorema en de limiet te nemen van erg kleine cirkels rond a, toon je aan dat

$$\int_{S^1(a,\epsilon)} \frac{f(z)}{z-a} dz = 2\pi i f(a)$$

waarbij $S^1(a,\epsilon)$ de cirkel met straal ϵ rond a is. Hieruit volgt het dat $f(a)$ een oneindig aantal maal afgeleid kan worden met betrekking tot a en dat alle afgeleiden dus analytisch zijn. Dit resultaat kan makkelijk veralgemeend worden tot

$$\int_\gamma \frac{f(z)}{z-a} dz = 2\pi n i f(a)$$

waarbij γ een gesloten kromme is in een omgeving van a die n keer draait rond a.

Een van de belangrijkste eigenschappen van een analytische functie is dat deze geschreven kan worden als een convergerende machtreeks in een omgeving van a; meer specifiek,

$$f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$$

waarbij het rechterlid eindig is voor $|z-a| < \epsilon$. Het bewijs van deze bewering is vrij makkelijk en volgt uit

$$\int_{S^1(a,\epsilon)} \frac{f(z)}{z-a} dz = 2\pi i f(a).$$

Wanneer men dit n-keer differentieert met betrekking tot a krijgt men

$$n! \int_{S^1(a,\epsilon)} \frac{f(z)}{(z-a)^{n+1}} dz = 2\pi i \left(\frac{\partial}{\partial a}\right)^n f(a)$$

waaruit vgolgt dat

$$\left| \left(\frac{\partial}{\partial a}\right)^n f(a) \right| \leq \frac{n!}{\epsilon^n} \max_{z \in S^1(a,\epsilon)} |f(z)|.$$

Vervolgens toon je aan dat

$$f(b) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\partial}{\partial a}\right)^k f(a)(b-a)^n$$

wat volgt uit

$$f(b)-f(a) = (b-a)\frac{1}{2\pi i} \int_{S^1(a,\epsilon)} \frac{f(z)}{(z-b)(z-a)} = (b-a)\frac{1}{2\pi i} \int_{S^1(a,\epsilon)} \frac{f(z)}{(z-a-(b-a))(z-a)}$$

$$= (b-a)\frac{1}{2\pi i}\int_{S^1(a,\epsilon)}\frac{f(z)}{(z-a)^2}\sum_{n=0}^{\infty}\left(\frac{b-a}{z-a}\right)^n = \sum_{n=0}^{\infty}\frac{(b-a)^{n+1}\left(\frac{\partial}{\partial a}\right)^{n+1}f(a)}{(n+1)!}.$$

Hier hebben we gebruikt dat

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$$

voor $|z| < 1$ zoals je direct kunt verifieren. Convergentie van de reeks volgt dan uit

$$\left|\frac{(b-a)^n\left(\frac{\partial}{\partial a}\right)^n f(a)}{n!}\right| \leq M\left(\frac{|b-a|}{\epsilon}\right)^n$$

zodat de reeks convergeert voor

$$|b-a| < \epsilon.$$

Een complexwaardige functie $f(z)$ wordt meromorf genoemt als en slechts als ze overal analytisch is met uitzondering van een aftelbaar aantal geisoleerde punten a_i zodat $f(z)(z-a_i)^{n_i}$ analytisch is met $\lim_{z\to a_i} f(z)(z-a_i)^{n_i} \neq 0$. We hebben dus dat

$$\int_{\gamma_j} f(z)(z-a_j)^{n_j-1}dz = 2\pi i \lim_{z\to a_j} f(z)(z-a_j)^{n_j}$$

waar γ_j een gesloten kromme is die rond a_j eenmaal windt. In het geval $n_j = 1$ noemen we a_j een pool en $\mathrm{res}(a_j) := \lim_{z\to a_j} f(z)(z-a_j)$ het residu; we hebben in dat geval dus

$$\int_{\gamma} f(z)dz = 2\pi i \sum_{\text{all poles } a_j\in S} m_j \mathrm{res}(a_j)$$

waar m_j het windgetal is van γ rond a_j. Deze formule is van extreem belang voor het berekenen van integralen van meromorfe functies en je maakt hier later een paar oefeningen op.

Je kunt deze analyse uitbreiden naar de hypercomplexe of Cliffordwaardige functies toe in de veranderlijke $x = x_0.1 + x^i.e_i$ waarbij $e_ie_j + e_je_i = 2\epsilon_i\delta_{ij}$ met $\epsilon_i = \pm 1$ en $i : 1\ldots N$. x_0, x_i zijn in dit geval reeele getallen maar kunnen ook complex genomen worden. Cliffordwaardige functies in de veranderlijke x hoeven niet te voldoen aan een differentiaalvergelijking zoals dit het geval is voor de complexe analyse. Inderdaad, het is natuurlijk de Dirac operator $D = e^j\partial_j$ alsook de operatoren $\partial_x = \frac{1}{2N}(N\partial_0 + D)$ en $\partial_{x^\star} = \frac{1}{2N}(N\partial_0 - D)$, waarbij $e_i^\star = -e_i$, te definieren. Je controleert dan onmiddelijk dat $\partial_x(x) = 1, \partial_x(x^\star) = 0$ en omgekeerd voor $\partial_{x^\star}$; dus je kan een functie links monogeen noemen indien

$$\partial_{x^\star}F(x_0, x^je_j) = 0.$$

Dit betekent echter niet dat elke functie van x in aanmerking komt; bijvoorbeeld

$$F(x) = xe_1x$$

is niet links monogeen hetgeen te wijten is aan de niet commutativiteit van de Clifford getallen. Een bijzondere klasse van functies zijn machtreeksen $P(x)$ in x met reeele (complexe) coefficienten, deze gedragen zich net als de analytische functies voor complexe getallen. Het is te zeggen

$$\partial_{x^\star} P(x) = 0$$

en de lezer wordt uitgenodigd te onderzoeken in welke mate Stokes stelling nog van toepassing is.

Finaal toon je als oefening aan dat elke complexe veelterm van n de graad precies n nulpunten heeft; indien dit niet het geval zou zijn, dan bestaat er een veelterm $Q(z)$ zonder nulpunten; bijgevolg is de functie $\frac{1}{Q(z)}$ analytisch op gans het complexe vlak en geldt dus

$$\frac{1}{Q(a)} = \frac{1}{2\pi i} \int_{S^1(a,R)} \frac{dz}{Q(z)(z-a)}.$$

Neem nu de limiet voor R naar oneindig om aan te tonen dat het rechterlid nul is. Dit impliceert dat $Q(a)$ oneindig moet zijn wat niet kan.

Chapter 10

Speciale functies.

De chronologie die we hier zullen aannemen is totaal verschillend van datgene wat je gewend bent in een standaard programma. Bijgevolg zijn de inzichten alsook de methodes die je zult verwerven een pak diepgaander dan standaard het geval is. We beginnen bij de exponentiele functie vermits zij de basis vormt voor de rest; e^x kan geintroduceerd worden op verschillende manieren en we zullen ze allen behandelen. Allereerst is e^{rx} een eigenfunctie van de operator ∂_x; het is te zeggen

$$\partial_x e^{rx} = re^{rx}.$$

De lezer merkt op dat de functie $f_n(rx) = \left(1 + \frac{rx}{n}\right)^n$ voldoet aan

$$\partial_x f_n(rx) = r f_{n-1}(rx)$$

dus het nemen van de limiet voor n naar oneindig produceert e^{rx}. Inderdaad

$$e^{rx} = \lim_{n\to\infty} \left(1 + \frac{rx}{n}\right)^n$$

en je toont aan dat deze uitdrukking convergeert in twee stappen (a) voor positieve rx definieert $f_n(rx)$ een stijgende positieve rij en (b) het supremum is gegeven door

$$e^{rx} = \sum_{n=0}^{\infty} \frac{(rx)^n}{n!}$$

hetgeen eindig is voor elke rx. Toon aan, vanuit de reeksontwikkeling, dat $\partial_x e^{rx} = re^{rx}$. Bovendien heb je dat

$$e^{r(x+y)} = \lim_{n\to\infty} \left(1 + \frac{r(x+y)}{n}\right)^n = \lim_{n\to\infty} \left(\left(1 + \frac{rx}{n}\right)\left(1 + \frac{ry}{n}\right) - \frac{r^2xy}{n^2}\right)^n = e^{rx}e^{ry}$$

hetgeen ook makkelijk volgt uit $\partial_x e^{r(x+y)} = re^{r(x+y)}, \partial_x e^{rx}e^{ry} = re^{rx}e^{ry}$ en $e^{r(0+y)} = e^{ry} = e^{r0}e^{ry}$. Finaal toon je deze eigenschap aan vanuit de reeksontwikkelling; dit laatste suggereert dat de exponentiele functie goed gedefinieerd is voor elke complexe veranderlijke z en voldoet aan

- $e^{z+w} = e^z e^w$,
- $e^0 = 1$,
- $|e^{ix}| = 1$ voor elke $x \in \mathbb{R}$.

Deze laatste eigenschap zegt dat de afbeelding $x \to e^{ix}$ de reeele as op de eenheidscirkel afbeeldt en de eerste formule reveleert dat een positieve x overeenkomt met de booglengte of hoek. Dus

$$e^{ix} = \cos(x) + i \sin(x)$$

waarbij

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

en

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}.$$

Je controleert onmiddelijk dat $\partial_x \sin(x) = \cos(x), \partial_x \cos(x) = -\sin(x)$ en dus $\partial_x^2 \sin(x) = -\sin(x)$ en hetzelfde geldt voor $\cos(x)$. Uit de definitie van de modulus volgt dan onmiddelijk dat

$$\cos^2(x) + \sin^2(x) = 1$$

en de regeltjes van Simpson volgen uit $e^{ix}e^{iy} = e^{i(x+y)}$:

- $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$,
- $\sin(x+y) = \cos(x)\sin(y) + \sin(x)\cos(y)$.

We hebben dus in het algemeen dat elk complex getal voorgesteld kan worden als $z = re^{i\theta}$ waarbij $r \geq 0$ en $\theta = 0 \ldots 2\pi$ met 2π gelijk aan de lengte van de eenheidscirkel. Dit noemen we de polaire decompositie en deze kan veralgemeend worden naar normale operatoren:

$$A = |A|U$$

waarbij $|A| = \sqrt{A^\dagger A}$ en U een partiele isometrie is. De exponentiele functie heeft een omgekeerde gegen door het natuurlijk logarithme $\ln(x)$; dit voldoet aav

$$x + y = \ln(e^{x+y}) = \ln(e^x e^y) = \ln(e^x) + \ln(e^y)$$

en aldus geldt voor een algemeen complex getal dat

$$\ln(z) = \ln(re^{i\theta}) = \ln(r) + i\theta.$$

Je merkt op dat deze definitie een discontunuiteit nalaat op de positieve reele as omdat de limiet langs boven gelijk is aan $\ln(x)$ en langs onder $\ln(x) + i2\pi$.

Dit fenomeen noemen we een singuliere snede wat erop neerkomt dat de functie niet eenduidig continue uitbreidbaar is tot die plaats. Afhankelijk van de keuze van referentiepunt op de cirkel kan deze snede steeds verlegd worden tot bijvoorbeeld de reeele negatieve as. Weg van de snede is ln analytisch zoals de lezer kan verifieren.

Vanuit de exponentiele functie kunnen we vele andere bijzondere functies definieren zoals de alpha, beta en gamma functie. Deze zijn standaard gedefinieerd aan de hand van speciale integralen. Zo is de gamma functie $\Gamma(x)$ voor $x \geq 0$ gedefinieerd als

$$\Gamma(x) = \int_0^\infty e^{-s} s^x ds.$$

Toon aan dat

$$\Gamma(x+1) = (x+1)\Gamma(x)$$

voor $x > 0$ en $\Gamma(1) = 1 = \Gamma(0)$. In het bijzonder is dus $\Gamma(n) = n!$ voor elk natuurlijk getal $n > 0$. Toon aan dat de gamma functie een unieke meromorfe uitbreiding heeft tot $\mathbb{C}$ met polen in $-n$, $n \in \mathbb{N}_0$. Het is niet de bedoeling alle gedetailleerde eigenschappen van die functies te bestuderen en de geinteresseerde lezer kan ze opzoeken in Wikipedia. Punt is dat de basisfunctie gegeven is door de exponentiele.

Er zijn nog heel wat andere speciale functies die (benaderende) eigenfuncties zijn van zogenaamde differentiaaloperatoren; deze laatste zijn Hermitische uitbreidingen van dicht gedefinieerde, symmetrische operatoren op bepaalde functieruimten die gevormd zijn aan de hand van partiele afgeleiden. Meer bepaald zijn deze sommen van operatoren van de vorm

$$W^{i_1 \ldots i_k}(x)\partial_{i_1} \ldots \partial_{i_k}$$

waarbij $k \in \mathbb{N}$. Als voorbeeld heb je in één reele veranderlijke de zogenaamde Laguerre polynomen die eigenfuncties zijn van

$$-\partial_x^2 + \omega^2 x^2$$

met $\omega > 0$. Je kunt deze eenvoudig vinden door de operator $a = i\partial_x + i\omega x$ te definieren op de functieruinte waarvoor het inproduct

$$\langle f|g \rangle = \int \overline{f(x)} g(x) dx$$

goed gesteld is. Je berekent dan door partiele integratie dat $a^\dagger = i\partial_x - i\omega x$ zodat

$$-\partial_x^2 + \omega^2 x^2 = a^\dagger a + \omega$$

en aldus zijn de eigenfuncties van de oorspronkelijke operator gelijk aan deze van $a^\dagger a$. Vindt nu Ψ zodat $a(\Psi) = 0$, dan zijn alle eigenfuncties gegeven door $(a^\dagger)^n(\Psi)$ hetgeen je bewijst uit $\left[a, a^\dagger\right] = 2\omega$. Bereken als oefening ook de corresponderende eigenwaarden. Je moet je in dit soort redeneringen bekwamen, dit kom je vaak tegen in de fysica!

Chapter 11

Cohomologie en het de-Rahm isomorfisme.

Dit kleine hoofdstukje betreft opnieuw een diep inzicht tussen gesloten volumevormen en homologieklassen van simplices. Beter gezegd, we hebben de d operator en een integraal die voldoet aan de micro-randvoorwaarde zodat Stokes theorema voldaan is. We zijn geinteresseerd in het $\mathbb{R}$ moduul C_k van k-gesloten vormen Ω_k die voldoen aan $d\Omega_k = 0$ en die equivalent zijn op een exacte k vorm $d\Omega_{k-1}$; het verschil met standaard homologie theorie is dat d de dimensie laat toenemen daar waar ∂ het doet afnemen. Dit suggereert een dualiteit welke is dat $H_k^\star = C_k$; inderdaad,beschouw de actie van een element $\Omega_k \in C_k$ op een gesloten oppervlak $S_k \in H_k$ gedefinieerd door

$$\widehat{\Omega_k}(S_k) = \int_{S_k} \Omega_k$$

dan tonen we eerst aan dat dit goed gedefinieerd is. In het geval S_k equivalent is aan S'_k gegeven dat beide de rand vormen van T_{k+1} hebben we dat

$$\widehat{\Omega_k}(S_k) - \widehat{\Omega_k}(S'_k) = \int_{T_{k+1}} d\Omega_k = 0$$

waar we hier Stokes theorema gebruikt hebben als wel als de geslotenheid van Ω_k. Ook hebben we dat

$$\widehat{\Omega_k + d\Omega_{k-1}}(S_k) = \widehat{\Omega_k}(S_k) + \int_{\partial S_k} \Omega_{k-1} = \widehat{\Omega_k}(S_k)$$

waar we Stokes theorema gebruikt hebben alsook het gegeven dat S_k geen rand heeft. Dit toont aan dat alles goed gedefinieerd is vermits de uitdrukking niet afhangt van de keuze van representanten van de equivalentieklassen. Nu tonen we aan dat de afbeelding injectief is; duidelijk vergt een niet triviaal gesloten k vorm $\Omega_k \in C_k$ het bestaan van een niet triviaal gesloten oppervlak S_k zodat

$\widehat{\Omega_k}(S_k) \neq 0$. Indien dit niet het geval zou zijn, dan is de integraal over elk k-oppervlak R_k met rand volledig bepaald door de rand wat betekent dat

$$\int_{R'_k} \Omega_k = \int_{R_k} \Omega_k$$

in het geval dat $\partial R_k = \partial R'_k$. Bovendien is de afhankelijkheid van de rand lokaal en additief net zoals dit is voor de integraal. Bijgevolg bestaat er een $k-1$ vorm Ω_{k-1} zodat $d\Omega_{k-1} = \Omega_k$. Finaal toon je surjectiviteit aan door te stellen dat voor elke niet triviale $S_k \in H_k$ er een gesloten, maar niet exacte, Ω_k bestaat zodat $\widehat{\Omega_k}(S_k) = 1$; dit toon je makkelijk aan door aan te tonen dat elke k-vorm Ω_k gedefinieerd op S_k uitgebreid kan worden op een manier zodat aan de vergelijking $d\Omega_k = 0$ voldaan is (in het algemeen bestaan er een oneindig aantal die daar aan voldoen). Dus hebben we bewezen dat de Betti getallen even goed vanuit cohomologie berekend kunnen worden in de plaats van simpliciale homologie. Dit beeindigt het onderwerp van dit hoofdstuk.

Chapter 12

Riemannse en Lorentziaanse meetkunde.

Het doel van dit hoofdstuk is de calculus van hoofdstuk acht uit te breiden naar de zogenaamde varieteiten; dit zijn ruimten die er lokaal uitzien als $\mathbb{R}^n$. In het bijzonder definieren we een C^n varieteit $\mathcal{M}$ als een topologische ruimte die er lokaal uitziet als $\mathbb{R}^n$; meer in het bijzonder bestaat er een bedekking met open verzamelingen $\mathcal{O}_\alpha$ die homeomorf afgebeeld worden $\phi_\alpha : \mathcal{O}_\alpha \to \mathcal{V}_\alpha \subseteq \mathbb{R}^n$ met homeomorphismen ϕ_α op het beeld $\mathcal{V}_\alpha$ welk oftewel een open omgeving is van de oorsprong $\mathcal{W}_\alpha$ of de doorsnede is van deze met de halfruimte $\{x|x_n \geq 0\}$. In het laatste geval bevat $W_\alpha \cap \{x|x_n = 0\}$ een deel van de rand. In het geval $\mathcal{O}_\alpha \cap \mathcal{O}_\beta \neq \emptyset$ heeft men dat $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(\mathcal{O}_\alpha \cap \mathcal{O}_\beta) \to \phi_\beta(\mathcal{O}_\alpha \cap \mathcal{O}_\beta)$ een C^n diffeomorphisme is. We zullen steeds onderstellen dat een varieteit paracompact is, wat betekent dat het bedekt kan worden door een aftelbaar aantal compacte verzamelingen, en niet noodzakelijk samenhangend alhoewel deze onderstelling een standaard vereiste is in het klassiek werk van Hawking en Ellis. $\mathbf{V}$ is een vectorveld op $\mathcal{M}$ als en slechts als voor elke functie $f : \mathcal{M} \to \mathbb{R}$ en kaart $(\phi_\alpha, \mathcal{O}_\alpha)$ op $\mathcal{M}$ geldt dat er een vectorveld $\mathbf{V}_\alpha$ op $\mathcal{V}_\alpha$ bestaat zodat

$$\mathbf{V}(f) = \mathbf{V}_\alpha(f \circ \phi_\alpha).$$

$\mathbf{V}_\alpha$ wordt een lokale voorstelling van $\mathbf{V}$ genoemd in de kaart $(\phi_\alpha, \mathcal{O}_\alpha)$. Hieruit kun je duale velden definieren vanuit hun actie op vectorvelden alsook algemene tensorvelden. Evenzo kunnen we de notie van een k-vorm alsook de uitwendige afgeleide veralgemenen; we komen nu tot de definitie van een push forward, pull back en Lie afgeleide. Een diffeomorphisme $\psi : \mathcal{M} \to \mathcal{M}$ is een homeomorphisme zodat voor elke kaart $(\phi_\alpha, \mathcal{O}_\alpha)$ en kaart $(\phi_\beta, \mathcal{O}_\beta)$ zodat $\psi(\mathcal{O}_\alpha) \cap \mathcal{O}_\beta \neq \emptyset$ geldt dat

$$\phi_\beta \circ \psi \circ \phi_\alpha^{-1} : \mathcal{V}_\alpha \to \phi_\beta \circ \psi(\mathcal{O}_\alpha) \subseteq \mathcal{V}_\beta$$

een C^n differentieerbare afbeelding is en evenzo is ψ^{-1}. Dit is gewoon een afspraak die we maken, dat de graad van afleidbaarheid van diffeomorphismen dezelfde is als die van de "kaart transformaties". Gegeven een diffeomorphisme

$\psi : \mathcal{M} \to \mathcal{M}$ en functie $f : \mathcal{M} \to \mathbb{R}$ we definieren de push forward van f aan de hand van ψ als $(\psi^\star f)(x) = f(\psi^{-1}(x))$. De pullback is dan gedefinieerd als de push forward met ψ^{-1} en wordt genoteerd als $\psi_\star f$. Je veralgemeent de bovenstaande definities voor injectieve, differentieerbare afbeeldingen $\psi : \mathcal{M} \to \mathcal{N}$. Gegeven een vectorveld $\mathbf{V}$ op $\mathcal{M}$, de push forward $\psi^\star \mathbf{V}$ is gedefinieerd aaan de hand van $\psi^\star \mathbf{V}(\psi^\star f)(\psi(x)) = \mathbf{V}(f)(x)$ en evenzo voor de pull back. Op dezelfde manier kunnen we, bij wijze van dualiteit, de push forward en pull back definieren van éénvormen alsook algemene tensorvelden. In een coordinatenvoorstelling leest dit:

$$(\psi^\star \mathbf{V})^\alpha(\psi(x)) = \frac{\partial y^\alpha(\psi(x))}{\partial x^\beta} V^\beta(x)$$

en de evidente veralgemening naar éénvormen en tensorvelden. We gaan nu het “verschil” meten van $\phi^\star \mathbf{V}$ met $\mathbf{V}$; vermits deze definitie coordinaat onafhankelijk moet zijn, zijn slechts infinitisemale verschillen toegelaten. Gegeven de éénparameterfamilie van diffeomorphismen ψ_t zodat $\psi_{t+s} = \psi_t \circ \psi_s$ voor t, s voldoende klein en $\psi_0 = \text{id}$ de identiteitstransformatie. Bijgevolg definieren we de differentiaal

$$\frac{d\psi_s^\star f}{ds}|_{s=0}(x) = \mathbf{V}(f)$$

waar $V^\alpha(x) = \frac{dx^\alpha(\psi_{-s}(x))}{ds}_{|s=0}$. Dit vectorveld bepaalt volledig de diffeomorphismen ψ_s bij middel van

$$\frac{dy^\alpha(\psi_s(x))}{ds} = -V^\alpha(\psi_s(x)).$$

Dus, we hebben een verband tussen vectorvelden en éénparameterfamilies van diffeomorphismen; dit stelt ons in staat de Lie afgeleide van een algemeen tensorvectorveld te definieren

$$\mathcal{L}_{\mathbf{V}}(T)(x) = \lim_{s \to 0} \frac{\psi_s^\star T(x) - T(x)}{s}$$

en we gaan nu de Lie afgeleide definieren van een vectorveld en eenvorm. Bij definitie hebben we dat $\mathcal{L}_{\mathbf{V}}(f)(x) = \frac{d}{ds}|_{s=0} \psi_s^\star f(x) = \mathbf{V}(f)(x)$ en evenzeer

$$(\mathcal{L}_{\mathbf{V}} \mathbf{W})(f)(x) = \frac{d}{ds}|_{s=0}(\psi_s^\star \mathbf{W})(f)(x) = \frac{d}{ds}|_{s=0} \mathbf{W}(\psi_{-s}^\star f)(\psi_{-s}(x)) =$$

$$-\mathbf{W}\mathbf{V}(f)(x) + \mathbf{V}\mathbf{W}(f) = [\mathbf{V}, \mathbf{W}]\,(f)(x).$$

Je merkt op dat

$$d(\psi^\star \Omega) = \psi^\star(d\Omega)$$

voor elke k-vorm Ω, daarom

$$\mathcal{L}_{\mathbf{V}} d = d\mathcal{L}_{\mathbf{V}}.$$

Dit kan ook aangetoond worden met behulp van

$$\mathcal{L}_{\mathbf{V}} = d i_{\mathbf{V}} + i_{\mathbf{V}} d$$

op de ruimte van k-vormen, waarbij $i_{\mathbf{V}}$ de contractie is met een vector $\mathbf{V}$ gedefinieerd als

$$i_{\mathbf{V}}\Omega_{\mu_1\ldots\mu_k}dx^{\mu_1}\wedge\ldots\wedge dx^{\mu_k} = V^{\mu_1}\Omega_{\mu_1\ldots\mu_k}dx^{\mu_2}\wedge\ldots\wedge dx^{\mu_k}.$$

Maattheoretisch heb je dat per definitie $(\psi^\star\mu)(\psi(A)) := \mu(A)$ voor elke maat μ en meetbare verzameling A; daarom

$$\int_{\mathcal{B}} f d\mu = \int_{\psi(\mathcal{B})}(\psi^\star f)(\psi^\star d\mu)$$

wat we de covariantie eigenschap van de integraal noemen. Toon aan dat de definities van $\psi^\star\mu$ en $\psi^\star\Omega$ waarbij Ω een k-vorm is samenvallen en dat de integraal gegeven is door

$$\int_{\mathcal{M}} f\Omega.$$

Tot zover hebben we niks over meetkunde gezegd en ons uitsluitend geconcentreerd op calculus op varieteiten. Met andere woorden, we moeten een lokaal scalair product definieren dat afhangt van de varieteit coordinaten. Dus, we zijn geinteresseerd in $(0,2)$ covariante tensoren $h_{\alpha\beta}$ en $g_{\alpha\beta}$ die locale orthonormale basissen v_a en e_a produceren zodat

$$h_{\alpha\beta}v_a^\alpha v_b^\beta = \delta_{ab}$$

en

$$g_{\alpha\beta}e_a^\alpha e_b^\beta = \eta_{ab}$$

waarbij η_{ab} de zogenaamde Minkowski metriek is gedefinieerd door

$$\eta_{11} = 1, \eta_{ij} = -\delta_{ij}; i,j = 2\ldots n$$

en alle andere componenten verdwijnen. Indien $n = 4$, e_a wordt een tetrad of vierbein genoemd; metrieken zoals δ_{ab} worden Riemanns genoemd en vormen een veralgemening van eindig dimensionale Hilbertruimten terwijl η_{ab} Lorentziaans genoemd wordt en niet-compacte null verzamelingen

$$\eta_{ab}v^a v^b = 0$$

definieert. Onze eerste taak bestaat erin aan te tonen dat deze meetkunden specificaties vormen van de gewoonlijke metrische en Lorentziaanse padmetrieken respectievelijk. In dat geval hadden we dat de kortste of langste lengte van een (tijdachtige) kromme gelijk is aan de afstand of pseudo-metriek respectievelijk. Vooraleer we aan dit resultaat toe zijn, moeten we iets zeggen over de causale structuutr gedefinieerd door η_{ab}. Een vector v^a wordt (a) causaal genoemd als en slechts als $\eta_{ab}v^a v^b \geq 0$ (b) nul of lichtachtig indien $\eta_{ab}v^a v^b = 0$ en (c) ruimtelijk indien $\eta_{ab}v^a v^b < 0$. Bovendien noemen we een causale vector v^a toekomst wijzend met betrekking tot het tetrad e_a als en slechts als $v^0 > 0$. In wat volgt veronderstellen we dat er een globaal goed gedefinieerd tetrad bestaat wat bepaalde topologieeen uitsluit zoals de Mobius strip. Deze laatste is gevormd door een rechthoek te beschouwen, de korte randen vast te nemen en één component ervan 180 graden te draaien en vervolgens met de andere te identificeren. Deze onderstelling op $(\mathcal{M}, g)$ luidt dat deze tijdsorienteerbaar alsook orienteerbaar is.

Gegeven een differentieerbare kromme $\gamma : [a, b] \to \mathcal{M}$, we zeggen dat deze toekomst wijzend en causaal is als en slechts als de raakvector aan elk punt dat is. De lengte van zulke kromme is gegeven door

$$L(\gamma) = \int_a^b \sqrt{g(\dot{\gamma}(s), \dot{\gamma}(s))} ds$$

en evenzo voor de lengte van elke kromme in een Riemannse meetkunde. We definieren $J^+(x)$ als de verzameling van punten y welke aan x verbonden kunnen worden door middel van een toekomstwijzende causale kromme beginnende in x. Gelijkaardig definieren we $J^-(x)$ als de verzameling van alle y welke kunnen verbonden worden aan x door middel van een toekomstwijzende causale kromme die begint in y. We bestuderen nu extremale krommen tussen twee punten x en y die vast zijn. Daarom moeten we een soort "differentiaal" met betrekking tot de kromme nemen en deze gelijk aan nul stellen; het resultaat is

$$0 = \delta L(\gamma) = \int_a^b \frac{1}{2\sqrt{g(\dot{\gamma}(s), \dot{\gamma}(s))}} \left(2 g_{\mu\nu} \delta\dot{\gamma}^\mu(s) \dot{\gamma}^\nu(s) + \partial_\alpha g_{\mu\nu} \dot{\gamma}^\mu(s) \dot{\gamma}^\nu(s) \delta\gamma^\alpha(s)\right)$$

en je moet deze vergelijking verder herschrijven, door gebruik te maken van $\delta\dot{\gamma}^\mu(s) = (\dot{\delta\gamma^\mu(s)})$, tot het reduceert tot de vorm

$$\int_a^b F(\gamma(s), \dot{\gamma}(s), \ddot{\gamma}(s))_\mu \delta\gamma^\mu(s) ds.$$

Door op te merken dat deze uitdrukking moet verdwijnen voor alle $\delta\gamma^\mu(s)$ krijgt men de volgende vergelijkingen

$$\ddot{\gamma}^\mu(s) + \Gamma^\mu_{\nu\alpha} \dot{\gamma}^\alpha(s) \dot{\gamma}^\nu(s) = 0$$

waarbij $\Gamma^\mu_{\nu\alpha} = \frac{1}{2} g^{\mu\kappa} \left(\partial_\nu g_{\kappa\alpha} + \partial_\alpha g_{\nu\kappa} - \partial_\kappa g_{\nu\alpha}\right)$ het zogenaamde Christoffel symbool is. De parametrizatie s wordt de affiene parametrizatie genoemd gegeven dat de hierbovenstaande vergelijking invariant blijft onder transformaties $s \to bs + a$. Gegeven dat de integraal reparametrizatie invariant is alsook invariant onder coordinatentransformaties moet deze vergelijking transformeren als een vector onder coordinaten transformaties en de lezer kan het effect verifieren van een reparametrizatie. We compactificeren deze vergelijking als

$$\dot{\gamma}^\nu \nabla_\nu \dot{\gamma}^\mu = 0$$

waar

$$\nabla_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\nu\kappa} V^\kappa$$

voor elke vector V^μ. De lezer controleert dat $\dot{\gamma}^\mu(s)\partial_\mu = \frac{d}{ds}$ en dat $\nabla_\mu V^\nu$ transformeert als een $(1,1)$ tensor; ook hebben we dat $V^\nu \nabla_\nu W^\mu - W^\nu \nabla_\nu V^\mu - [V,W]^\mu = 0$ voor alle vectoren $\mathbf{V}, \mathbf{W}$. De eerste eigenschap betekent dat de covariante afgeleide een meetkundige betekenis heeft terwijl de laatste zegt dat ze torsie-vrij is wat equivalent is aan

$$\Gamma^\mu_{\nu\kappa} = \Gamma^\mu_{\kappa\nu}.$$

De vergelijking $\dot{\gamma}^\nu \nabla_\nu \dot{\gamma}^\mu = 0$ staat bekend als de geodetische vergelijking en de geassocieerde krommen zijn geodeten. We breiden nu de definitie van de covariante afgeleide uit als volgt

$$\nabla_\mu f = \partial_\mu f$$

of in coordinatenvrije notatie,

$$\nabla f = df, \nabla \mathbf{V} = \nabla_\mu V^\nu dx^\mu \otimes \partial_\nu.$$

Met de notatie

$$\nabla_{\mathbf{V}} = V^\mu \nabla_\mu$$

breiden we de definitie uit

$$\begin{aligned} \nabla_{\mathbf{W}}(\omega(\mathbf{V})) &= (\nabla_{\mathbf{W}}\omega)(\mathbf{V}) + \omega(\nabla_{\mathbf{W}}\mathbf{V}) \\ \nabla_{\mathbf{W}}(S \otimes T) &= (\nabla_{\mathbf{W}} S) \otimes T + S \otimes (\nabla_{\mathbf{W}} T) \end{aligned}$$

waar ω een éénvorm is en S, T algemene tensoren. Vanuit de specifieke vorm van de Christoffel symbolen leest de lezer af dat $\nabla g = 0$ wat betekent dat de metriek covariant constant is. Het is duidelijk dat de geodeten de Lorentziaanse afstand maximalizeren in het geval ze causaal zijn en minimalizeren in het geval van een Riemannse metriek.

We bestuderen nu een aantal tensoren die men kan construeren uit de covariante afgeleide en contracties daarvan welke van primordiaal belang zijn in de geometrische analyse en algemene relativiteitstheorie. Het eerste geval betreft een $(1,2)$ tensorveld

$$\mathbf{T}(\mathbf{V}, \mathbf{W}) = \nabla_{\mathbf{V}} \mathbf{W} - \nabla_{\mathbf{W}} \mathbf{V} - [\mathbf{V}, \mathbf{W}]$$

en de lezer controleert inderdaad dat $\mathbf{T}(\mathbf{V}, \mathbf{W}) = -\mathbf{T}(\mathbf{W}, \mathbf{V})$ en

$$\mathbf{T}(f\mathbf{V} + \mathbf{Z}, \mathbf{W}) = f\mathbf{T}(\mathbf{V}, \mathbf{W}) + \mathbf{T}(\mathbf{Z}, \mathbf{W}).$$

Voor de metrische connectie, gedefinieerd door de Christoffel symbolen, heeft men dat de Torsie tensor verdwijnt en de lezer wordt uitgenodigd connecties met Torsie te bestuderen. Het tweede geval betreft een $(1,3)$ tensor veld genoteerd door

$$\mathbf{R}(\mathbf{V}, \mathbf{W})\mathbf{Z} = \nabla_{\mathbf{V}} \nabla_{\mathbf{W}} \mathbf{Z} - \nabla_{\mathbf{W}} \nabla_{\mathbf{V}} \mathbf{Z} - \nabla_{[\mathbf{V}, \mathbf{W}]} \mathbf{Z}$$

wat terug antisymmetrisch is in $\mathbf{V}, \mathbf{W}$. In componenten leest dit $R_{\mu\nu\alpha}^{\ \ \ \beta}$ en we kunnen indices verhogen en verlagen aan de hand van $g^{\alpha\beta}$ en $g_{\alpha\beta}$ respectievelijk. Vanuit de algemene Jacobi identiteit $[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]] = 0$ voor operatoren, bekomen we, toegepast op de metrische connectie, dat

$$R_{[\mu\nu\alpha]}^{\ \ \ \ \ \ \beta} = 0, \nabla_{[\alpha} R_{\beta\gamma]\kappa}^{\ \ \ \ \ \ \delta} = 0$$

waar de vierkante haakjes staan voor totale anti-symmetrizatie. Deze identiteiten worden de eerste en tweede Bianchi identiteit genoemd; bovendien controleer je door expliciete berekening dat

$$R_{\mu\nu\alpha}^{\ \ \ \ \ \kappa} g_{\kappa\beta} = R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu}.$$

Dit tensor veld staat bekend als de Riemann tensor en is van groot belang in de algemene relativiteitstheorie. De eerste contractie

$$R_{\mu\nu} = R_{\mu\alpha\nu}^{\ \ \ \ \ \alpha}$$

is een symmetrische tensor genaamd de Ricci tensor en haar tweede contractie

$$R = R_{\mu\nu} g^{\mu\nu}$$

is de Ricci scalar. De metrische tensor bepaalt een uniek volume element

$$dV(x) = \sqrt{\det(g_{\mu\nu})} dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n$$

en je controleert dat deze uitdrukking coordinaatonafhankelijk is. Hieruit construeer je de zogenaamde Einstein-Hilbert actie

$$\int dV(x) R(x)$$

de variatie naar de metrische componenten welke de Einstein vergelijkingen produceert.

Ik moedig de lezer aan alle berekeningen te maken, bijvoorbeeld dat de Torsie en Riemann tensoren wel echt tensoren zijn en niet afhangen van afgeleiden van de vectorvelden. De Bianchi identiteiten worden geverifieerd vanuit Jacobi zonder een coordinaten berekening te maken, terwijl de resterende symmetrieen van de Riemann tensor wel degelijk afhangen van het metrische karakter van de connectie en dus een expliciete berekening vergen. Ook is het nuttig de Einstein vergelijkingen te vinden; spendeer hier gerust wat tijd aan vermits tensor calculus erg belangrijk is om te beheersen.

Chapter 13

Dispersie van een bundel geodeten en het opblazen of ineenklappen van bollen.

In dit hoofdstuk gaan we verder in op globale geometrische eigenschappen van Riemannse en Lorentziaanse ruimtetijden. Het studie onderwerp bij uitstek betreft het gedrag van naburige geodeten met betrekking tot elkaar. Gaan ze uiteen of convergeren ze naar een punt; is er rotatie of expansie? Met andere woorden, we gaan de geodetische vergelijking eens nader onder de loep nemen en in het bijzonder bestuderen we willekeurige secties van geodeten, zowel in de Riemannse als Lorentziaanse context. Het is uiteraard mogelijk scherpere resultaten te boeken voor Riemannse meetkunden vermits deze op een natuurlijke manier compacte oppervlakken definieert. De Lorentziaanse meetkunde geniet niet dergelijke eigenschappen en laat alleen toe resultaten te verkrijgen met betrekking tot een tijdachtige kromme of wereldlijn. De Riemannse context is vrij goed begepen doch de Lorentziaanse laat nog heel wat vers onderzoek toe en relatief weinig resultaten zijn hier geboekt. Dus je kan je best uitgenodigd voelen een originele bijdrage tot de wiskunde te leveren en ik zal een paar van die open problemen aanhalen in dit hoofdstuk.

Om te beginnen verifieer je dat voor een geodeet $\gamma(s)$ de lengte van de raakvector $\mathbf{V}(s) = \frac{d}{ds}\gamma(s)$ constant is. Inderdaad,

$$\frac{d}{ds}g(\dot{\gamma}(s), \dot{\gamma}(s)) = (\nabla_{\mathbf{V}(s)}g)(\dot{\gamma}(s), \dot{\gamma}(s)) + 2g(\nabla_{\mathbf{V}(s)}\mathbf{V}(s), \mathbf{V}(s)) = 0$$

omdat de metriek constant is en waarbij ook de geodetische vergelijking gebruikt wordt. In het Lorentziaanse geval kiezen we dus een parametrizatie zodat de lengte één is voor tijdachtige geodeten, nul voor nulgeodeten en min één voor ruimteachtige geodeten. Noteer met $T^{\star}\mathcal{M}_x$ de lineaire ruimte van alle vectoren in x, dan definieren we de afbeelding

$$\exp_x : T\mathcal{M}_x \to \mathcal{M} : v \to \exp_x(v)$$

waarbij $\exp_x(v)$ het eindpunt is van de geodeet met affiene parameterlengte één zodat de raakvector in x gegeven is door v. Duidelijk is

$$\mathrm{D}\exp_x(0)(w) = w$$

waar $w \in TT^\star\mathcal{M}_{(x,0)} \sim T^\star\mathcal{M}_x$ wat gewoon betekent dat in eerste orde de metriek een globale Euclidische of Minkowski metriek is. Definieer $T^\star\mathcal{M} = \cup_{x\in\mathcal{M}} T^\star\mathcal{M}_x$, dan definieren we een topologie dewelke gelijk is aan de producttopologie $\mathcal{O} \times \mathcal{V}$ waar $\mathcal{O} \subseteq \mathcal{M}$ en $\mathcal{V} \subseteq \mathbb{R}^n$. De exponentiele afbeelding is dan een lokaal diffeomorfisme wat betekent dat er een open omgeving $\mathcal{V}$ van 0 in $T^\star\mathcal{M}_x$ bestaat zodat

$$\exp_x : \mathcal{V} \to \exp_x(\mathcal{V})$$

een diffeomorphisme is. Gegeven twee punten x en y dan is het mogelijk dat deze verbonden zijn door meerdere geodeten te wijten aan een niet triviale topologie of door het bestaan van focale punten (denk aan een lens). We leiden nu de geodetische deviatievergelijking af: gegeven een éénparameter familie van geodeten $\gamma(s,t)$ waar $s : a \ldots b$ en $t \in (-\epsilon, \epsilon)$ met als raakvectoren $\mathbf{V} = \left(\frac{\partial}{\partial s}\right)^\star$ en $\mathbf{Z} = \left(\frac{\partial}{\partial t}\right)^\star$ dan is

$$[\mathbf{V}, \mathbf{Z}] = 0$$

en bijgevolg

$$\nabla_V \nabla_{\mathbf{V}} \mathbf{Z} = \nabla_{\mathbf{V}} \nabla_{\mathbf{Z}} \mathbf{V} = \mathbf{R}(\mathbf{V}, \mathbf{Z})\mathbf{V}$$

waar we in de eerste gelijkheid gebruikt hebben dat de connectie torsieloos is, terwijl de geodetische vergelijking $\nabla_{\mathbf{V}}\mathbf{V} = 0$ gebruikt is in de tweede vergelijking. De bovenstaande vergelijking is van de vorm

$$\ddot{\mathbf{Z}} + A(s)\dot{\mathbf{Z}} + B(s)\mathbf{Z} = 0$$

met A, B matrices en het punt duidt op afleiding naar s. Zulke vergelijking is van het Newtoniaanse type met A de frictie matrix en B de oscillatie frequentie ($B > 0$) of expansie factor ($B < 0$) in het kwadraat. Inderdaad, in één dimensie leest dit

$$\ddot{z} + a\dot{z} + bz = 0$$

wat herschreven kan worden als

$$\ddot{h} + (b(s) - \frac{1}{2}\dot{a}(s) - \frac{1}{4}a^2(s))h = 0$$

met $h(s) = e^{\frac{1}{2}\int_0^s a(t)dt} z(s)$. Deze laatste is van het type

$$\ddot{h}(s) + c(s)h(s)$$

en we noemen $c(s)$ minus de expansie in het kwadraat indien het kleiner is dan nul en de oscillatie frequentie in het kwadraat indien het groter is dan nul.

Inderdaad voor c constant zijn de oplossingen $xe^{\sqrt{c}s} + ye^{-\sqrt{c}s}$ en $x\cos(\sqrt{-c}s) + y\sin(\sqrt{-c}s)$ respectievelijk waarbij x, y willekeurige reele constanten zijn. Zoals we straks zullen zien laat deze eenvoudige berekening toe een paar meetkundige eigenschappen te begrenzen.

Langs een geodeet kun je de notie van Fermi-Walker transport definieren van een tensor aan de hand van

$$\nabla_{\mathbf{V}}\mathbf{T} = 0$$

waar $\mathbf{V}$ de raakvector is aan de geodeet. Je kunt makkelijk controleren dat Fermi-Walker transport de uitdrukking

$$g(\mathbf{Z}, \mathbf{W})$$

constant houdt voor elke $\mathbf{Z}, \mathbf{W}$ langs eender welke geodeet. Je ziet hier dat transport van een vector het meetrekken is van een oneindig kleine meetstok langs de beweginskrommen van de waarnemer; dit meetrekken gebeurt krachtenloos en wordt uitsluitend bepaald door de relaties gelegd op ruimtetijd. Dit is dus een erg belangrijk fysisch begrip en een gedodeet kan dus geinterpreteerd worden als een krachtenloze vooruitstuwing doorheen ruimtetijd; Einstein zag dit als de beweging van een *vrije* waarnemer in een gravitationeel veld en dit is hoe de relativiteitstheorie geboren werd.

Een orthonormale basis die een verzameling is van n orthonormale vectorvelden e_a kan dus gezien worden als een lokaal referentiestelsel wat echter geen locaal coordinatenstelsel hoeft te bepalen vermits de vectoren e_a niet commuteren, $[e_a, e_b] \neq 0$. Een Lie algebra is gedefinieerd door te eisen dat

$$[e_a, e_b] = f^c_{ab} e_c$$

voor reeele *constanten* f^c_{ab}. Vermits parallel transport van een orthonormale basis een orthonormale basis is, definieert zij een rotatie of Lorentz transformatie door het getransporteerd referentiestelsel uit te schrijven ten opzichte van het locale. Het is te zeggen, we hebben een transformatie $\Lambda(x, w)^{a'}_b$ waar $w \in T^\star\mathcal{M}_x$ en b een (Lorentz) index is met betrekking tot $e_b(x)$ en a' een (Lorentz) index met betrekking tot $e^{a'}(\exp_x(w))$. Een Lorentz transformatie is een symmetrie van de Minkowski metriek wat betekent dat

$$\Lambda(x, w)^{a'}_b \Lambda(x, w)^{c'}_d \eta_{a'c'} = \eta_{bd}.$$

Je controleert dat deze een continue groep vormen van dimensie $\frac{n(n-1)}{2}$ en dat het 4 gescheiden componenten heeft in geval $n = 4$ welke bepaald worden door de transformaties $1, T, S, ST$ waar T staat voor de tijdsreversie en S voor de ruimte reversie. Het Euclidische geval, gegeven door de rotatie groep, hebben we reeds hiervoor bestudeerd.

Een ander nutig hulpmiddel om geodeten te bestuderen is Synge's functie, een wiskundig object dat de kwadratische vorm

$$\sigma(x,y) = \frac{1}{2}(y-x)^{\mu}(y-x)^{\nu}\eta_{\mu\nu}$$

op Minkowski veralgemeent, waar x, y punten voorstellen in $\mathbb{R}^4$. De definitie luidt dat $\sigma(x,y)$ gelijk is aan *een* extremale waarde van

$$I(x,y) = \frac{1}{2}(t_1 - t_0)\int_{t_0}^{t_1} g_{\mu\nu}\frac{dx^{\mu}(s)}{ds}\frac{dx^{\nu}(s)}{ds}ds$$

waarbij $x^{\mu}(s)$ een kromme is die x met y verbindt. Je berekent dat de extremale waarden gegeven worden door de krommen voor dewelke $x^{\mu}(s)$ een geodeet is. Bovendien is deze uitdrukking invariant onder affine reparametrizaties $s \to as+b$ van de kromme en daarom kan men de variatie $\delta x^{\mu}(s)$ nemen zodanig dat de eindpunten t_0, t_1 vast zijn. Om je op weg te helpen,

$$\delta I(x,y) = (t_1 - t_0)\int_{t_0}^{t_1} g_{\mu\nu}\left(\frac{D}{ds}\frac{dx^{\mu}(s)}{ds}\right)\delta x^{\nu}(s)ds$$

waar

$$\frac{D}{ds} = \nabla_{\frac{d}{ds}}$$

en $\delta x^{\mu}(t_0) = \delta x^{\mu}(t_1) = 0$. Deze variatie verdwijnt als en slechts als

$$\frac{D}{ds}\frac{dx^{\mu}(s)}{ds} = 0$$

wat inderdaad de geodetische vergelijking is. Dus Syge's functie is gelijk aan

$$\sigma(x,y,w) = \frac{1}{2}g(w,w) = \frac{1}{2}\epsilon L^2(x,y,w)$$

waar $w \in T^{\star}\mathcal{M}$, $L(x,y,w)$ gelijk is aan de lengte van de geodeet vertekkende in x met raakvector w en eindpunt $\exp_x(w) = y$. Hier, $\epsilon = 1$ voor tijdachtige geodeten en -1 voor ruimteachtige. Veronderstellende dat w continue varieert wanneer x en y dat doen, besluiten we dat w alleen dient ter aanwijzing van het feit dat verschillende geodeten tussen x en y kunnen bestaan. De lezer controleert dat

$$\sigma(x,y,w)_{,\mu} := \partial^x_{\mu}\sigma(x,y,w) = -w_{\mu} = g_{\mu\nu}w^{\nu}, \sigma(x,y,w)_{,\mu'} := \partial^y_{\mu'}\sigma(x,y,w) = -g_{\mu'\kappa'}\Lambda^{\kappa'}_{\nu}(x,w)w^{\nu}.$$

Hieruit volgt dat

$$2g^{\mu\nu}\sigma(x,y,w)_{,\mu}\sigma(x,y,w)_{,\nu} = \sigma(x,y,w)$$

en idem voor de afgeleiden naar y.

De overblijvende presentatie in dit hoofdstuk betreft Riemannse meetkunde en de oefeningen gaan over originele uitbreidingen van deze stellingen naar

Lorentziaanse meetkunde toe. Laat ons beginnen met een algemene definitie: gegeven twee vectoren v, w, dan is de oppervlakte van het parallelogram opgespannen door v, w gegeven door $g(v,v)g(w,w) - (g(v,w))^2$ en we definieren de sectionele kromming $s(v,w)$ als

$$s(v,w) = \frac{g(R(v,w)v,w)}{g(v,v)g(w,w) - (g(v,w))^2}.$$

Een metriek is van constante sectionele kromming als en slechts als

$$R_{\mu\nu\alpha\beta} = \frac{R}{n(n-1)}(g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha}).$$

Bijgevolg leest de Ricci tensor

$$R_{\mu\nu} = \frac{R}{n}g_{\mu\nu}$$

wat betekent dat elke ruimte van constante sectionele kromming een Einstein ruimte is. Je merkt op dat onder de afbeelding $g \to -g$ de Ricci tensor invariant blijft terwijl de Ricci scalar en sectionele kromming op hun tegengestelde afgebeeld worden. In het bijzonder is dit ook waar voor de Einstein actie in een even aantal dimensies. Het is duidelijk dat indien sectionele kromming alleen van belang is dat we ons kunnen richten op homogene en isotrope modelruimten van constante sectionele kromming. Deze zijn uiteraard uniek; deze modelruimten zijn maximaal symmetrisch waar een symmetrie vertegenwoordigd is door een diffeomorfisme $\psi : \mathcal{M} \to \mathcal{M}$ zodat $\psi^\star g = g$. In het geval we spreken over een éénparameterfamilie ψ_t van diffeomorfismen leidt dit tot

$$\mathcal{L}_{\mathbf{V}}g = 0$$

waarbij $\mathbf{V}$ het genererend vectorveld is. De lezer verifieert dat deze vergelijking herschreven wordt als

$$\nabla_{(\alpha}V_{\beta)} = 0$$

waar $V_\beta = g_{\beta\alpha}V^\alpha$ en de ronde haakjes staan voor symmetrizatie; het is te zeggen

$$Z_{(\alpha_1 \ldots \alpha_n)} = \frac{1}{n!}\sum_{\sigma \in S_n} Z_{\alpha_{\sigma(1)} \cdots \alpha_{\sigma(n)}}.$$

Deze vergelijking staat bekend als Killing's vergelijking en het is makkelijk in te zien dat er $\frac{n(n+1)}{2}$ zijn wat $\frac{n(n-1)}{2} + n$ parameters vrij laat waar n het aantal vrije parameters is in een punt; daarom noemen we een ruimte met $\frac{n(n+1)}{2}$ lineair onafhankelijke Killing velden maximaal symmetrisch. De Killing vergelijking en definitie van de Riemann tensor impliceren dat

$$\nabla_\alpha\nabla_\beta V^\alpha = R_{\alpha\beta\gamma}{}^{\alpha}V^\gamma.$$

Meer algemeen, vanuit

$$\nabla_\alpha\nabla_\beta V_\gamma - \nabla_\beta\nabla_\alpha V_\gamma = -R_{\alpha\beta\gamma}{}^{\kappa}V_\kappa$$

volgt dat

$$\nabla_\alpha \nabla_\beta V_\gamma + \nabla_\beta \nabla_\gamma V_\alpha = -R_{\alpha\beta\gamma}{}^{\kappa} V_\kappa$$

en gebruik makend van de symmetrieen van de Riemann tensor, heeft men

$$\nabla_\alpha \nabla_\beta V_\gamma = R_{\beta\gamma\alpha}{}^{\kappa} V_\kappa .$$

We bestuderen nu de prototype *Riemannse* maximaal symmetrische ruimten met nul, positieve en negatieve sectionele kromming respectievelijk. Deze meetkunden zijn niet uniek en kunnen topologisch van elkaar veschillen: bijvoorbeeld, in twee dimensies, heeft men de vlakke cylinder, torus en vlak. Duidelijk kun je elke cylinder isometrisch inbedden in de ruimte $\mathbb{R}^2 / T_{x,a}$ en idem voor de torus in $\mathbb{R}^2 / \{T_{x,a}, T_{y,b}\}$ waar $T_{x,a}$ een translatie definieert over een afstand $a > 0$ in de x richting. In die zin is het vlak maximaal en je kunt aantonen dat elke vlakke meetkunde door knip en plakwerk van zulke quotienten gedefinieerd kan worden. Deze meetkunden worden vlak, sferisch en respectievelijk hyperbolisch genoemd.

Een, maar niet *de*, "maximale" vlakke ruimte is gegeven door $(\mathbb{R}^n, \delta_{\alpha\beta})$ waar de Kronecker $\delta_{\alpha\beta}$ naar een canonisch coordinatenstelsel refereert. Je wordt uitgenodigd een andere maximale vlakke ruimte te construeren, door twee vlakke ruimten te verbinden door een zogenaamd wormgat; concreet, knip een $n-1$ dimensionale torus uit twee $\mathbb{R}^n$ en verbindt deze aan de hand van een n dimensionale cylinder. Je weet reeds dat de $\frac{n(n+1)}{2}$-dimensionale symmetrie groep van $(\mathbb{R}^n, \delta_{\alpha\beta})$ gegeven is door $SO(n) \times \mathbb{R}^n$ waarbij $SO(n)$ de rotatie groep is bepaald door O^μ_ν zodat

$$O^\alpha_\mu O^\beta_\nu \delta_{\alpha\beta} = \delta_{\mu\nu}$$

en $\mathbb{R}^n$ is de n-dimensionale translatie groep. De actie op $\mathbb{R}^n$ is gegeven door

$$((O, a)x)^\alpha = O^\alpha_\beta x^\beta + a^\alpha$$

en hieruit kun je de groepbewerking expliciet vinden. $SO(n)$ wordt gegenereerd door anti-Hermitische matrices A wat betekent dat $A^\dagger = -A$ wat impliceert dat de groep $\frac{n(n-1)}{2}$ dimensionaal is. Tussen elke twee punten x en y bestaat er exact een geodeet hetwelk het gewone lijnsegment is dat beide punten verbindt. De canonische volumemaat is gegeven door $dx^1 \wedge \ldots \wedge dx^n$ hetgeen herschreven kan worden als

$$r^{n-1} dr \wedge \Omega_{S^{n-1}}$$

waar $r = \sqrt{\sum_{i=1}^n (x^i)^2}$ en $\Omega_{S^{n-1}}$ is de canonische volumemaat op de $n-1$ dimensionale bol welke gesloten is maar niet exact. Dit gegeven kan makkelijk bewezen worden vanuit de eigenschap dat ∂_r loodrecht staat op $S^{n-1}(r)$ alsook de scchalinsformule $d(rx, ry) = r d(x, y)$ voor elke $r > 0$. Daarom hebben ballen van straal r rond een punt x een volume gelijk aan

$$\frac{r^n}{n} \mathrm{Vol}(S^{n-1}).$$

Een andere belangrijke eigenschap die bewezen kan worden is dat de som van hoeken in een geodetische driehoek altijd gelijk is aan π.

Nu komen we tot een maximale modelruimte van positieve sectionele kromming, welke het n dimensionale boloppervlak is van straal r in $(\mathbb{R}^{n+1}, \delta_{\alpha\beta})$; de symmetrie groep is $SO(n+1)$ hetwelk de correcte dimensie is. De afstand tussen twee punten x, y is gegeven door de booglengte van het segment van de cirkel bepaald door x, y en de oorsprong welk gegeven is door

$$\cos(\theta) = \frac{x^{\alpha} y_{\alpha}}{r^2}.$$

waar x, y $n+1$-dimensionale vectoren zijn. Omde meetkunde te begrijpen moet geen enkele berekening gemaakt worden; de meetkunde is immers volledig bepaald door de lengte parameter r; in het bijzonder impliceert dit dat de sfeer een constante sectionele kromming heeft van de vorm $\frac{a}{r^2}$ op dimensionele gronden. Een kleine berekening in twee dimensies onthult dat $a = -1$ en de lezer controleert dat

$$\int_{S^2} R\sqrt{h} = -2\pi\chi(S^2) = -4\pi$$

waarbij we eraan herinneren dat χ het Euler getal is; dit resultaat is geldig voor elke Riemannse metriek op de sfeer. Je lezer wordt uitgenodigd dit te bewijzen als een niet-triviale oefening (hint: toon aan dat de variatie van de dichtheid naar de metriek een totale afgeleide geeft), genaamd het Gauss-Bonnet theorema. Je ziet onmiddelijk in dat de geodetische expansie reduceert tot

$$\frac{d^2}{ds^2} Z = -\frac{1}{r^2} Z$$

wat oplossingen geeft van de vorm $Z(s) = b \sin\left(\frac{s}{r}\right)$ indien $Z(0) = 0$. Dit resultaat impliceert dat de som van de hoeken in een geodetische driehoek groter is dan π vermits geodeten terug convergeren. Noteer dat in de literatuur de Ricci kromming van de sfeer positief is omwille van een andere contractie in de indices en dus aldaar verschijnt een plus teken in het Gauss-Bonnet theorema.

Finaal komen we tot het model van de hyperbolische ruimte $\mathbb{H}^n(r)$, $r > 0$, die we kunnen afleiden uit *minus* $n+1$-dimensionale Minkowski als de ruimte

$$\mathbb{H}^n(r) = \{x | x^{\alpha} x^{\beta} \eta_{\alpha\beta} = r\}.$$

Deze ruimte is maximaal symmetrisch met als symmetrie groep $SO(n, 1)$ wat de $n+1$-dimensionale Lorentz transformaties zijn; deze Riemannse ruimte heeft terug constante sectionele kromming welke gelijk is aan $\frac{1}{r^2}$ zoals een berekening in $n = 2$ aantoont. Geodeten expanderen dus volgens $b \sinh\left(\frac{s}{r}\right)$ en dus te kleiner de straal r des te sneller geodeten van elkaar verwijderen en dus te sneller het volume van een bal groeit. Het volume van zulke bal met straal s is gegeven door

$$V(s) = \int_0^s dt r^{n-1} \left(\sinh\left(\frac{t}{r}\right)\right)^{n-1} V(S^{n-1}).$$

Evenzo is dit in het sferische geval

$$V(s) = \int_0^s dt r^{n-1} \left(\sin \left(\frac{t}{r} \right) \right)^{n-1} V(S^{n-1})$$

en als oefening bereken je deze integralen.

Uit alle overwegingen is het duidelijk dat hetvolgende resultaat waar is: als $(\mathcal{M}, h)$ een n-dimensionale Riemannse varieteit is met sectionele kromming kleiner of gelijk aan $R \in \mathbb{R}$ dan is het volume van eender welke bal met straal r kleiner of gelijk aan dat van dezelfde bal in de modelruimte $\mathbb{H}^n \left(\frac{1}{\sqrt{R}} \right)$ of $S^n \left(\frac{1}{\sqrt{-R}} \right)$ afhankelijk of $R > 0$ of $R < 0$ respectievelijk. De reden is evident: beschouw elke $x \in \mathcal{M}$ en neem de Euclidische sfeer met straal r op de tangente ruimte en warp die aan de hand van de exponentiele map. Dan zie je dat kleine ruimtehoekjes $\delta\Omega$ van radiale geodeten hoogstens een volume hebben van dat in de referentieruimte omdat er in het algemeen verdere compressie en rotatie is. Je kunt dit netjes opschrijven aan de hand van de geodetische expansie vergelijking.

Je kunt nu trachten een dergelijk resultaat te construeren voor zogenaamde Alexandrov sets in de Lorentziaanse meetkunde; voor zover ik weet is dit resultaat onbekend. De Alexandov set $A(x, y)$ voor $y \in J^+(x)$ is gegeven door

$$A(x, y) = J^+(x) \cap J^-(y)$$

het is dus de verzameling van punten die op een causale toekomst georienteerde kromme liggen tussen x en y. Je ziet dus dat er een verband lijkt te bestaan tussen globale krommingseigenschappen en topologie: dit is inderdaad zo en diepere resultaten kunnen geboekt worden (zie Milnor, Perelman). Topologie vanuit het differentieerbaar standpunt leidt tot zeer rijke inzichten zoals de Brouwer en Kakutani vaste puntstellingen, Morse theorie waaruit alle Betti getallen berekend kunnen worden aan de hand van de studie van kritische punten (die plaatsen van “pinshing” aanduiden) enzovoort.

Chapter 14

Gekromde ruimtetijd en het afbuigen en focussen van lichtstralen: een voorproefje van relativiteit.

Dit hoofdstukje vormt een beetje een uitzondering op de rest van dit boek in de zin dat we hier rechtstreeks toepassingen in de fysica aanraken en een korte introductie geven tot de algemene relativiteitstheorie van Einstein. Zoals gezegd kan deze bekomen worden vanuit het actieprincipe

$$\int_{\mathcal{M}} R\sqrt{-g}$$

voor een Lorentziaanse metriek g en varieteit zonder rand $\mathcal{M}$. Indien $\mathcal{M}$ een rand heeft moet de zogenaamde Hawking-York-Gibbons randterm toegevoegd worden om randtermen die verschijnen in de variatie van de Einstein-Hilbert actie te compenseren. Je berekent dat

$$\frac{\delta\sqrt{-g}}{\delta g^{\alpha\beta}} = -\frac{1}{2}\sqrt{-g}g_{\alpha\beta}$$

en

$$\frac{\delta R}{\delta g^{\alpha\beta}} = R_{\alpha\beta} + g^{\gamma\kappa}\frac{\delta R_{\gamma\kappa}}{\delta g^{\alpha\beta}}.$$

Deze laatste tensor $\frac{\delta R_{\gamma\kappa}}{\delta g^{\alpha\beta}}$ is apart symmetrisch onder de uitwisseling van α, β en γ, κ en kan dus niet de Riemann tensor bevatten; bijgevolg is het nul. Desalniettemin is de variatie van de Ricci tensor niet triviaal

$$g^{\gamma\kappa}\delta R_{\gamma\kappa} = \nabla_\alpha V^\alpha(\delta g_{\kappa\gamma})$$

waar het vectoveld $V^\alpha(\delta g_{\kappa\gamma})$ lineair afhangt van covariante afgeleiden van $\delta g_{\kappa\gamma}$. Tracht in te zien dat dit waat *moet* zijn zonder enige berekening te maken (hint: de scalar moet van tweede orde in de afgeleiden zijn en eerste orde in de variatie van het metrisch veld). Bijgevolg hebben we dat de variatie van de Einstein-Hilbert actie levert

$$\int (R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta}) \delta g^{\alpha\beta} \sqrt{-g} + \nabla_\alpha V^\alpha(\delta g_{\kappa\gamma}) \sqrt{-g}.$$

Deze laatste term kan echter herschreven worden in termen van

$$\int \partial_\alpha(\sqrt{-g} V^\alpha) dx^1 \wedge \ldots \wedge dx^n$$

hetgeen gelijkgesteld wordt aan de uitwendige afgeleide van een $n-1$ vorm. Deze verdwijnt echter omwille van Stokes theorema gegeven dat $\partial\mathcal{M} = \emptyset$. De *vacuum* Einstein vergelijkingen zijn dus

$$R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = 0$$

en je wordt uitgenodigd de Hawking-York-Gibbons randactie te bepalen (hint: zoek het begrip externe kromming op gegeven dat infinitisemale informatie van de bulk rond de rand zal meetellen). De Einstein (symmetrische) tensor

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta}$$

voldoet aan

$$g^{\gamma\alpha} \nabla_\gamma G_{\alpha\beta} = 0$$

een behoudswet die rechtstreeks volgt uit de tweede Bianchi identiteit. In de fysica impliceert deze gelijkheid de wet van behoud van de energie en momentum tensor. Het is belangrijk te begrijpen dat deze behoudswet een identiteit is en dus helemaal niet vereist dat de Einstein vergelijkingen voldaan zijn; de wet van behoud van energie en monentum vereist dit wel. Je bekomt nu ook een bewijs van de Gauss-Bonnet stelling in de zin dat in twee dimensies de Einstein tensor identiek verdwijnt: inderdaad $G_{\alpha\beta} g^{\alpha\beta} = (1 - \frac{n}{2}) R$ en $R_{\alpha\beta} = \frac{R}{2} g_{\alpha\beta}$ in twee dimensies impliceren dat $G_{\alpha\beta} = 0$ dus de Einstein vergelijkingen zijn *topologisch* in $n = 2$. Fysici kennen dit resultaat door te stellen dat er geen gravitatie bestaat in twee dimensies; uiteraard is dit niet het geval in hogere dimensies. De vacuum Einstein vergelijkingen kunnen vereenvoudigd worden tot

$$R_{\alpha\beta} = 0$$

wat betekent dat de Ricci tensor nul is. Bewijs dat de behoudswet voor de Einstein tensor een rechtstreeks gevolg is van de invariantie van de actie onder éénparameter groepen van diffeomorphismen; in het bijzonder geldt dat

$$\delta g_{\alpha\beta} = \epsilon \mathcal{L}_{\mathbf{V}} g_{\alpha\beta} = 2\epsilon \nabla_{(\alpha} V_{\beta)}.$$

Dus de behoudswet is een rechtsreeks gevolg van een continue symmetrie, een resultaat dat in het algemeen bekend staat als Emmy Noether's theorema.

De Einstein vergelijkingen zijn bijzonder moeilijk op te lossen alhoewel men onder algemene voorwaarde het bestaan en uniek karakter van een oplossing kan aantonen. Veel makkelijker is het aan post Newtoniaanse fysica te doen; dat is expandeer $g_{\alpha\beta} = \eta_{\alpha\beta} + \epsilon h_{\alpha\beta}$ op een varieteit wiens topologie een globale vlakke Minkowski metriek toelaat en schrijf de Einstein vergelijkingen tot op eerste orde in ϵ waarbij deze laatste een erg klein getal ondersteld wordt. De resulterende vergelijking wordt de graviton vergelijking genoemd.

In Einstein's theorie staat verder de geodetische vergelijking centraal; puntvormige objecten die volledig gekarakteriseerd worden door hun massa bewegen op geodeten. Men zou andere vergelijkingen kunnen bedenken zoals

$$\nabla_{\mathbf{V}}\mathbf{V} + \alpha\mathbf{R}(\nabla_{\mathbf{V}}\mathbf{V}, \mathbf{V})\nabla_{\mathbf{V}}\mathbf{V} = 0.$$

Indien het object extensief is kan men deze vergelijking laten afhangen van de *expansietensor*

$$\nabla_{(\alpha}V_{\beta)}$$

of *rotatietensor*

$$\nabla_{[\alpha}V_{\beta]}$$

zoals je kunt bedenken. Lichtstralen planten zich bij benadering verder op nul geodeten, tenminste toch in de optische benadering (quantum effecten uitgesloten). Je berekent dan ook zaken zoals frequentieverandering door de golfvector parallel te transporteren over de nul geodeet en vervolgens te projecteren op de bewegingsvector van de waarnemer. Concreet, zij k een golfvector aangrijpend in het punt x en y een punt in de toekomst verbonden door een nul geodeet met x. Zij n en m de genormalizeerde raakvectoren van de wereldkrommen van twee meettoestellen in x en y respectievelijk, dan is de frequentie van de golf gemeten in x door het toestel gelijk aan $g(k, n)/\hbar$ waarbij $\hbar$ een universe constante is. De frequentie gemeten in y is dan $g(k', m)/\hbar$ waarbij k' het Fermi-Walker transport is van k langs de nul geodeet tot in y. Opdie manier vindt men roodverschuivingseffecten (een lagere frequentie) omwille van interactie met een gravitationeel veld.

Chapter 15

Verdere abstractie: oefeningen.

Dit hoofdstuk is gans origineel en bevat nieuwe denkwijzen. Je hebt opgemerkt dat we eerst verzamelingenleer hebben opgebouwd, dan getallenleer, operatorenleer, meetkunde gebaseerd op getallen en operatorenleer; deze laatste methoden vergden echter coordinaat afhankelijke representaties omdat we punten voorgesteld hebben door middel van getallen. Om de ganse meetkunde coordinaat onafhankelijk te doen, moeten we dus een aanpak volgen die dichter bij verzamelingenleer ligt. De definitie van continuiteit bijvoorbeeld was reeds van die aard en we geven er nu een voor differentialen. Cruciaal hier is een *meetkundige* notie van transport: neem twee verzamelingen X, Y met reflexieve en symmetrische relaties $R \subset X \times X, T \subset Y \times Y$ respectievelijk die beide topologisch open zijn, dan noemen we

$$\nabla_X : \{(x, y, z) : y, z \in R(x, \cdot)\} \to R : (x, y, z) \to \nabla_{(x,y)}(x, z) \in R(y, \cdot)$$

de getransporteerde van de relatie (x, z) over de relatie (x, y) van x naar y. ∇_X voldoet aan de volgende eigenschap: er bestaat een open omgeving $D \subset \mathcal{Q} \subset R$, waarbij D de diagonaal is, zodat voor elke $(x, w) \in \mathcal{Q}$ er een open omgeving $\mathcal{W}$ van w bestaat en een open omgeving $\mathcal{O} \subset R(x, \cdot)$ zodat de afbeelding

$$T_{\mathcal{O}} : \{(x, y, z) : (x, y), (x, z) \in \mathcal{O}\} \to R : (x, y, z) \to (\pi_2(\nabla_{(x,y)}(x, z)), \pi_2(\nabla_{(x,z)}(x, y))$$

continue en surjectief is op $\mathcal{W} \times \mathcal{W}$. Hier, π_2 is de projectie op de tweede factor. Bovendien definieren we de *commutator* tussen twee elementen van $R(x, \cdot)$ als

$$[(x, w), (x, v)] = ((v, x) \circ P(\nabla_{(x,v)}(x, w)))\sharp(\nabla_{(x,w)}(x, v) \circ (x, w))$$

waarbij $(w, z)\sharp(x, v) = \big(\nabla_{(w,v)}(w, z)\big) \circ (x, v)$ en $P(x, w) = (w, x)$. Merk op dat in een zekere zin we hebben dat onze connectie torsievrij is wat logisch is vermits de commutator een rechtstreeks product is van de vectorruimte structuur van $\mathbb{R}^n$. Vermits deze niet bestaat in het algemeen geval hebben we geen andere keuze dan de commutator te definieren aan de hand van de connectie. We noemen een functie $F : X \to Y$ differentieerbaar in $x \in X$ indien er voor elke open omgeving $\mathcal{V} \subset T(F(x), \cdot)$ er een open omgeving $\mathcal{O} \subset R(x, \cdot)$ bestaat alsook een bi-continue afbeelding $DF_w : \mathcal{O} \to Y$, $w \in \mathcal{O}$ zodat $F(w) = \pi_2(DF_x(x, w))$ en

$$DF_x\left(\left(\nabla_{(x,w)}(x, y)\right) \circ (x, w)\right) = DF_w(\nabla_{(x,w)}(x, y)) \circ DF_x(x, w).$$

Hier heeft men dat de samenstelling $\circ$ op $R\,(T)$ gedefinieerd is als $(w, z) \circ (x, w) = (x, z)$. Werk dit idee verder uit, bestudeer Riemann kromming en ontwikkel een notie van integraal.

Een eenvoudige oefening bestaat erin aan te tonen dat er een oneindig aantal priemgetallen zijn (hint: stel dat er een eindig aantal zijn en neem het product plus een). Moeilijker is uitspraken te doen over de dichteid van het aantal priemgetallen $\frac{a(n)}{n}$ kleiner of gelijk aan n; werk hetvolgende argument uit, zij $b_1, \ldots, b_k$ de eerste k priemgetallen, construeer dan een bovengrens voor $a(b_1 \ldots b_k + 1)$ door het aantal k koppels $(n_1, \ldots, n_k)$ te tellen zodat $\sum_{i=1}^{l} n_i \geq l$ en $\sum_{i=1}^{k} n_i = k$ te tellen waarbij $n_i \in \mathbb{N}$. Uit algemene argumenten haal je dat $\frac{a(n)}{n} \to 0$ en $a(n) \to \infty$ in de limiet voor n naar oneindig; veronderstel nu dat $a(n)$ uitbreidbaar is tot een reele veranderlijke x, en onderstel dat $\dot{a}(x) \to 0$ en $\frac{\dot{a}(x)}{x} \to 0$ in de limiet voor x naar oneindig. Toon dan aan dat $a(x) \sim \frac{x}{\ln(x)}$. Definieer een meromorfe functie met polen in de priemgetallen en bereken contourintegralen.