
Emergentie, Fysicalisme en Supervenientie

Eerste begeleider: dr. I.E.J. Douven
Tweede begeleider: dr. M.V.P. Slors
Derde begeleider: dr.J.A.M. Bransen

Gerrit van Dalfsen

Universiteit Utrecht
Faculteit Wijsbegeerte

M. Andriessenrade 56
2907 MA Capelle a/d IJssel
tel.: 010/458 4816
email: gdalfsen@wxs.nl
stud.nr: 9173226

Versie: 1.0
Datum: 20-02-2001
Gda11110.doc

Voorwoord

Deze scriptie is het resultaat van een fantastisch filosofisch avontuur. Een avontuur dat opwindend bleef tot het schrijven van de laatste woorden van de conclusie. Een conclusie die niets minder inhoudt dan een ontologische kanteling.

De reis heeft iets langer geduurd dan de nominale tijd die voor een afstudeerscriptie gesteld wordt, maar ik had de reis niet willen missen. De reis heeft geleid door soms onontgonnen land, en de kronkelwegen zijn veelvuldig geweest. Maar de reis heeft schitterende inzichten opgeleverd. Alleen terugkijkend kun je de rechte lijn zien van waar je begonnen bent. Die lijn is in deze scriptie neergelegd.

Ik wil Jan Bransen bedanken voor zijn vasthoudendheid waarmee hij geprobeerd heeft mij tegen mezelf in bescherming te nemen in het kiezen van een onderwerp voor de scriptie. Na de twintigste probleemstelling heeft hij het opgegeven: 'Dit is metafysica. Als je dit wilt, kan ik je niet verder begeleiden'. Die twintigste probleemstelling heb ik nog vele malen opnieuw geformuleerd in het licht van de zich ontwikkelende positie. Maar in de kern is de probleemstelling ongewijzigd gebleven.

Marc Slors wil ik bedanken voor zijn deskundig en stimulerend commentaar op de voorlaatste versie van deze scriptie, waardoor het betoog weer aansluiting vindt bij het gaande debat in het vakgebied.

En Igor Douven wil ik als eerste begeleider bedanken dat hij het mogelijk heeft gemaakt de scriptie te schrijven over een onderwerp, waarover volgens mij meer intellectuele helderheid te geven moest zijn.

Capelle a/d IJssel, 20 februari 2001

Inhoudsopgave

1. Inleiding	7
2. Non-reductie: Emergentie en Superveniëntie	13
2.1. Het Brits emergentisme	13
2.2. De kernpunten van het emergentisme	15
2.2.1. Een fysicalistische ontologie	15
2.2.2. Eigenschapsemergentie	16
2.2.3. De niet afleidbaarheid van emergente eigenschappen.	17
2.2.4. Emergente niveaus	18
2.3. Non-reductie en superveniëntie	19
2.3.1. De superveniëntiedefinities	21
2.3.2. Superveniëntie en emergentie zijn niet synoniem	22
2.4. Non-reductie: Samenvatting en conclusie	23
3. Fysicalisme	24
3.1. Inleiding	24
3.2. Het fysicalistisch dilemma	24
3.3. Het Atomisme	27
4. Het reductie-/non-reductiedebat	34
4.1. Inleiding	34
4.2. Het Reductionisme	35
4.3. Nagels model van theoriereductie	35
4.4. Het mind-body debat over reductie en non-reductie	37
4.5. Kims model van functionele reductie	39
4.6. Het reductie/non-reductiedebat: conclusie	41
5. Emergentie en Reductie	43
5.1. De non-reductieclaim van het emergentisme	43
5.2. Emergente eigenschappen	44
5.2.1. Artificiële structuren	44
5.2.2. Natuurlijke structuren	46
5.3. Semiotische eigenschappen	49
5.4. Conclusie	50
6. Literatuurreferenties	54

1. Inleiding

Met het opgeven van de dualistische en deïstische¹ posities aangaande leven en geest, en met de acceptatie van het materialisme in de tweede helft van de 19-e eeuw, ontstond er een nieuw vraagstuk: hoe zijn leven en geest te relateren aan de materiële wereld waaruit zij voortkomen?

Het materialisme is de opvatting dat alles dat we in deze wereld aantreffen -ook leven en geest- uiteindelijk voortkomt uit materiële condities. Het materialisme wordt vaak synoniem gesteld met fysicalisme². Het fysicalisme is de doctrine dat de wereld "*contains just what a true complete physics would say it contains*" (Crane en Mellor, 1990:186). Een dergelijke ware en complete fysische theorie wordt geacht uitspraken te doen over (a) welke objecten de wereld bevat, en (b) welke relaties er tussen die objecten bestaan. Op grond van die twee soorten uitspraken kan een onderscheid gemaakt worden tussen een *ontologisch* fysicalisme en een *nomologisch* fysicalisme. Het ontologisch fysicalisme beweert dat de kleinste deeltjes fysische entiteiten zijn, en dat alles in deze wereld uiteindelijk is opgebouwd uit deze kleinste fysische entiteiten. Het nomologisch fysicalisme beweert dat de fundamentele wetmatigheden³ fysische wetmatigheden zijn, en dat alles in deze wereld ondergeschikt is aan fysische wetmatigheden.

Emergentisten verwerpen het nomologisch fysicalisme. Emergentisten stellen dat gehelen in deze wereld weliswaar zijn samengesteld uit kleinste deeltjes, maar dat die gehelen vanwege hun samenstelling nieuwe ('emergente'⁴) eigenschappen hebben, nieuw in de zin dat het eigenschappen zijn die de delen niet hebben. Die nieuwe eigenschappen zijn volgens de emergentisten verbonden met eigen wetmatigheden, die niet te reduceren zijn tot fysische wetmatigheden⁵. Emergentisten hebben in de regel geen problemen met het ontologisch fysicalisme, omdat het afwijzen daarvan tot dualisme leidt, waar het emergentisme het dualisme verwerpt.

Sinds de acceptatie van het materialisme als uitgangspunt, inhoudende het afwijzen van de Carthesiaanse geestsubstantie en van *entelechie*⁶ als levenssubstantie, loopt al meer dan 150 jaar een debat tussen reductionisten en non-reductionisten: het debat over de vraag of theorieën over eigenschappen van leven en geest nu wel of niet te reduceren zijn naar fysische theorie. Non-reductionisten putten zich uit in argumentaties die aantonen dat bijvoorbeeld mentale eigenschappen en fysieke eigenschappen niet direct aan elkaar te relateren zijn; reductionisten stellen dat het mogelijk *moet* zijn mentale eigenschappen en fysieke

¹ Het dualisme is de opvatting dat geest een aparte substantie is, onafhankelijk van materiële substantie. Deïsme is de opvatting dat er een godheid is die de bron of de oorzaak van alle dingen is.

² Bijvoorbeeld Pettit, 1993a, p.213. Pettit merkt echter op dat beide termen door verschillende filosofen gebruikt worden om een variëteit aan posities te benoemen.

³ *Nomos* is Grieks voor wet.

⁴ Het Griekse '*emergere*' betekent 'naar boven komen', 'aan het licht komen'.

⁵ De emergentistische positie is te kenschetsen met de uitspraak dat het geheel meer is dan de som van de delen.

⁶ Met entelechie werd 'levenskracht' aangeduid.

eigenschappen aan elkaar te relateren⁷, maar dat het ons momenteel nog aan adequate theorieën ontbreekt om deze relatie te kunnen leggen.

Gevoed vanuit ontwikkelingen ten aanzien van theorieën over niet-lineaire systemen⁸, denk ik dat er een perspectief bestaat voor een derde weg. Een weg tussen het reductionisme en het non-reductionisme in. Dat perspectief werk ik in deze scriptie uit. Die derde weg houdt in dat ik het nomologisch fysicalisme afwijs. Maar dat hoeft niet te leiden tot een positie waarin niets gezegd kan worden over de relatie tussen de verschillende emergentieniveaus, zoals emergentisten stellen.

Non-reductionisten hebben zich steeds sterk gemaakt voor het beargumenteren van de autonomie van de kwalitatieve zijnsniveaus zoals leven en geest ten opzichte van het fysische niveau (en daarmee ook voor de autonomie voor de speciale wetenschappen⁹, zie bijvoorbeeld Fodor,1974). Het perspectief van het reductionistisch fysicalisme, het perspectief van een 'platte'¹⁰ wereld en een sterke relatie met determinisme, is in de ogen van non-reductionisten incommensurabel met het beeld dat zij in hun 'speciale' theorieën van de wereld hebben opgebouwd. Incommensurabel wil zeggen dat de begrippen die de non-reductionisten en die de fysicalisten hanteren niet compatibel zijn, anders gezegd, niet vergelijkbaar zijn, niet op elkaar aansluiten, en niet in elkaar te vertalen zijn. Dit is denk ik een van de redenen waarom er in de laatste 150 jaar weinig toenadering is geweest in de standpunten van fysicalisten en non-reductionisten. In non-reductionistische kamp zijn mensen te onderscheiden als Broad(1926), Davidson(1970), Fodor(1974) en Pettit(1993b). In het reductionistische kamp zijn mensen te onderscheiden als Carnap(1928), Hempel en Oppenheim(1948), en Kim(1992,1998).

De derde weg die ik in deze scriptie uitwerk, biedt uitzicht op de mogelijkheid om de kwalitatieve zijnsniveaus die de non-reductionisten benadrukken, te verklaren in termen van onderliggende niveaus, zonder dat daarmee de kwalitatieve niveauverschillen ontkend worden of gereduceerd worden. Dat betekent dat eigenschappen van leven en geest gerelateerd kunnen worden aan hun materiële inbedding, zonder het speciale van die eigenschappen te ontkennen.

De opzet van de scriptie is als volgt. Ik geef eerst een overzicht van de non-reductionistische positie. De begrippen emergentie en superveniëntie spelen daarin een belangrijke rol. Emergentie is de notie dat gehelen, als samenstel van delen, eigenschappen bezitten die niet aangetroffen worden bij de delen van het geheel, en die ook niet tot de eigenschappen van de delen zijn te reduceren (Emmeche,1997). Het begrip emergentie komen we voor het eerst tegen in *Problems of Life and Mind* van G.H.Lewes (1875). Lewes gebruikt de term 'emergent' voor wat J.S. Mill in zijn *System of Logic* (1843) aanduidt als 'heteropatische effecten'. Heteropatische effecten zijn volgens Mill effecten die zich niet laten beschrijven als som van de effecten van individuele oorzaken. Mill onderscheidt heteropatische effecten van 'homopatische effecten', die wel te beschrijven zijn als de som van de effecten van de individuele oorzaken. Het zijn alleen de homopatische effecten die zich met de in Mills tijd beschikbare natuurwetten adequaat laten beschrijven (Stephan,1992:28).

⁷ Bijvoorbeeld Kim, die stelt dat om de geest te behouden, deze tot zijn materiële basis gereduceerd moet kunnen worden, omdat anders alleen een materie-geest dualisme rest.

⁸ Het betreft hier met name ontwikkelingen op het gebied van artificial life en chaostheorie. Inhoudelijk zullen deze onderwerpen in de scriptie nauwelijks aan de orde komen. Relevante achtergrondliteratuur op dit gebied is o.a. S.A.Kauffman, *The Origins of Order*, New York,1993; S.N.Salthe, *Development and Evolution*, Cambridge,1993; E.Lorenz, *The Essence of Chaos*, London 1993; en S. Levy, *Artificial Life*, NY 1992.

⁹ In de Angelsaksische wereld worden onder 'special sciences' de menswetenschappen verstaan, als tegenover de 'natural sciences', die de fysica, de chemie, de geologie en de biologie omvatten.

¹⁰ Het fysicalisme denkt alles in deze wereld uiteindelijk te kunnen beschrijven in termen van fysische wetmatigheden en eigenschappen, d.w.z. in termen van de fysische basislaag. Daarom de aanduiding 'plat'.

Emergentie speelde een belangrijke rol in wat we de metafysica van de evolutie kunnen noemen. In de tweede helft van de 19-e eeuw worden een aantal belangrijke stappen gezet in het denken over evolutie en ontwikkeling. Darwins *Origin of Species* (1859) is voor deze opvattingen een ijkpunt. Darwin laat als eerste zien dat de ontwikkeling van soorten een continue en een natuurlijk proces is, en dat voor het ontstaan van soorten geen ontwerp vooraf nodig is. Volgens Darwin ontstaan er in de evolutie geen nieuwe eigenschappen ('De natuur maakt geen sprongen'), en dat betekent volgens Darwin dat elk organisme reeds in de kiem alle mogelijke eigenschappen moet bevatten. Zo bevatten volgens Darwin bijvoorbeeld eencelligen in de kiem reeds mentale eigenschappen (Blitz, 1992:11).

Eind 19-e eeuw bestaan er een aantal posities met betrekking tot ontwikkeling en evolutie. Het is Lloyd Morgan (1852-1936) die in zijn werk *Emergent Evolution* (1923) een aantal aspecten ten aanzien van evolutie en ontwikkeling samenbrengt, en daarmee inhoud geeft aan het begrip emergentie zoals we dan nu kennen. Lloyd Morgan doet echter de hiërarchie van emergentieniveaus uitkomen in een goddelijke toestand, en dat zal ongetwijfeld bijgedragen hebben aan de geringe weerklank die zijn model heeft gehad in de wetenschappelijke wereld. Het is dan ook niet Lloyd Morgans werk, maar de verwoording van C.D. Broad (1887-1971) die steeds als referentie wordt gehanteerd naar wat wordt aangeduid als het Brits emergentisme.

Superveniëntie is het tweede belangrijke begrip binnen de non-reductiepositie. Het begrip superveniëntie heeft een belangrijke rol gespeeld in het *mind-body* debat in de tweede helft van de 20-e eeuw. Het begrip is door Davidson (1970) in het *mind-body* debat min of meer terloops geïntroduceerd, en het heeft daarna momentum gekregen. Superveniëntie is een uitdrukking van covariatie tussen twee (verzamelingen van) eigenschappen, zeg S en B, met de toevoeging dat er een relatie van bepaling geldt tussen de eigenschappen S en B. Gezegd wordt dat eigenschap S supervenieert op eigenschap B, indien het noodzakelijk is dat als eigenschap B optreedt, dan ook eigenschap S optreedt. Anders uitgedrukt, als ik een object met eigenschap B één op één reproduceer, dan heeft dat object ook eigenschap S. De relatie is asymmetrisch in die zin dat als een object een superveniënte eigenschap S bezit, het niet noodzakelijk is dat dat object ook eigenschap B bezit. Dus er kunnen meer typen van eigenschap B bestaan, die leiden tot S.

De relatie tussen emergentie en superveniëntie is in de literatuur niet altijd even helder. Soms lijkt er een sterke mate van synonymie te bestaan (bijv. Beckermann, 1992; Humphreys, 1997). Beide begrippen zijn echter duidelijk verschillend. Mijn constatering is dat het begrip superveniëntie ten opzichte van het begrip emergentie weinig toegevoegde waarde lijkt te hebben.

Na de non-reductionistische positie werk ik de fysicalistische positie uit. Ik denk dat de vorderingen in de relativiteitstheorie, de quantumtheorie en de logica in de jaren '20 een belangrijk kantelpunt zijn geweest waardoor de wetenschappelijke aandacht verschoof van het emergentisme naar het fysicalisme. Zo kon men bijvoorbeeld op grond van de nieuwe fysische theorieën empirische verschijnselen wiskundig beschrijven, waar het emergentisme zich alleen in kwalitatieve beschrijvingen kon uiten. Meer nog, met behulp van de quantumtheorie konden chemische eigenschappen verklaard worden, die door Broad nog als emergent gekenmerkt werden.

Ten aanzien van het fysicalisme schetsen Crane en Mellor (1993) een dilemma, dat als we het 'fysieke' beperken tot alleen de microfysische deeltjes, dat dan het fysicalisme niet houdbaar is; maar als we ook hogere orde structuren (bijvoorbeeld biologische of sociale) tot het fysieke rekenen, wordt het nietszeggend breed. In dat geval zou dan ook bijvoorbeeld de psychologie tot de fysica gerekend moeten worden. Ik deel de opvatting van Crane en Mellor dat het nomologisch fysicalisme niet houdbaar is. De fysische wetenschappen vormen geen

'*Theory of Everything*', en ik verwijs ze terug naar het fysische emergentieniveau. Het blijkt dan, dat met het opgeven van het nomologisch fysikalisme het 'uitsluitingsprobleem', dat het *pièce de résistance* van het *mind-body* debat is, vervluchtigt. Het nomologisch fysikalisme roept zijn eigen spook op.

Het fysikalisme wordt vaak verbonden met het reductionisme¹¹. In hoofdstuk 4 ga ik in op het reductie-/non-reductiedebat. Carnap kan met zijn '*Der Logische Aufbau de Welt*' (1928) als een ijkpunt gezien worden in het reductionisme. Volgens Carnap moeten theoretische termen vertaald kunnen worden (d.w.z. gereduceerd kunnen worden) naar logische samenstellingen van elementaire waarnemingsuitspraken. Termen waarbij dat niet mogelijk is, bijvoorbeeld metafysische termen, zijn volgens de Carnap van geen waarde voor de wetenschap. Het programma van de Carnap om theoretische termen te reduceren tot uitsluitend waarnemingstermen bleek echter problematisch, en aan de mogelijkheid ervan wordt tegenwoordig getwijfeld (Zie onder meer Trout,1992:389; Van Brakel en Van den Brink,1988:84). Later werk in de logisch positivistische traditie staat het gebruik van metafysische termen opnieuw toe, maar de eis van logische afleidbaarheid om van reductie te kunnen spreken, blijft een belangrijk kenmerk. Bijvoorbeeld in het Deductief-Nomologische (D-N) verklaringmodel van Hempel en Oppenheim (1948), en later nog in het model voor theoriereductie van Nagel (1961).

Het nomologisch fysikalisme mag dan geen houdbare positie zijn, het ontologisch fysikalisme blijft een geaccepteerd standpunt, ook voor non-reductionisten. Het probleem van de non-reductionisten is dan een relatie te vinden tussen bijvoorbeeld mentale eigenschappen en de materiële implementatie van die eigenschappen, zonder de autonomie te verliezen die aan de mentale eigenschappen wordt toegeschreven. Deze discussie heeft plaats gevonden in het *mind-body* debat in de tweede helft van de twintigste eeuw. De doelstelling van dit debat is in de woorden van Kim(1998:2) '*... to find a way of accommodating the mental within a principled physicalist scheme, while at the same time preserving it as something distinctive - that is, without losing what we value, or find special, in our nature as creatures with minds*'.

In de jaren '70 centreerde het *mind-body* debat zich rond het model van theoriereductie van Nagel. De non-reductionisten baseerden hun positie in belangrijke mate op Nagels reductiemodel. Het is opvallend dat in het *mind-body* debat de emergentistische positie afwezig is, en dat vooral het atomistisch fysikalisme als veronderstelling de boventoon voert. De atomistische veronderstelling plaatst Nagels reductiemodel als een model voor theoriereductie in een onverdiend daglicht. De discussie in het *mind-body* debat heeft Nagels model van reductie achter zich gelaten, en richt zich nu op het idee van functionele reductie.

Functionele reductie houdt in dat een mentale eigenschap gereduceerd kan worden beschouwd tot zijn fysische implementatie, indien aangetoond kan worden dat de causale werking van de mentale eigenschap dezelfde is als de causale werking van de fysische implementatie. Een probleem is dat in de discussie rond functionele reductie het atomistische uitgangspunt nog niet verlaten is. Daardoor is ook functionele reductie behept met de problemen van het atomisme. Het ironische gevolg is dat Kims positie van functionele reductie daardoor een niet-reductieve positie oplevert, hetgeen het tegengestelde is van wat Kim voor ogen heeft.

¹¹ Er zijn vele opvattingen over wat reductionisme inhoudt. Reductie in de sterke vorm van het nomologisch fysikalisme vindt weinig ondersteuning meer. Mildere vormen van reductie zijn bijvoorbeeld theoriereductie, waarin een empirische wet herschreven wordt in termen van een meer fundamentele theoretische wet. Soms wordt de term reductie ook gebruikt indien eigenschappen van verschillende emergente niveaus aan elkaar gerelateerd kunnen worden.

In het laatste hoofdstuk werk ik de derde weg uit, en zal laten zien dat emergente niveaus op een verklarende wijze aan elkaar gerelateerd kunnen worden, *met* behoud van het speciale dat elk emergentieniveau kenmerkt.

In deze scriptie wordt veelvuldig het begrip 'eigenschap' (Engels: *property*) gebruikt. Als ik het over een eigenschap heb, gebruik ik dat in een ontologische betekenis. Een eigenschap is 'eigen' aan een object of een gebeurtenis en bestaat onafhankelijk van onze kennis ervan. Daarnaast gebruik ik het begrip 'predicaat'. Een predicaat is een semantische expressie. Een predicaat is onderdeel van talige uitdrukkingen, die waar of onwaar kunnen zijn. Predicaten komen voor in theorieën, en verwijzen daarbij naar eigenschappen in het ontologisch domein.

Het onderscheid tussen eigenschap en predicaat is niet zonder problemen. Het praten over eigenschappen vindt plaats in talige uitdrukkingen, en deze uitdrukkingen zijn theoriegeladen. Het onderscheid tussen eigenschap en predicaat is vooral een formeel onderscheid, dat niettemin nuttig is omdat hiermee een onderscheid gemaakt kan worden tussen hoe de dingen zijn (hoe problematisch deze uitspraak ook is), en hoe we dingen binnen een theorie aanduiden. Het biedt de mogelijkheid dat we dezelfde eigenschap binnen twee verschillende theorieën met twee verschillende predicaten kunnen aanduiden¹². Deze mogelijkheid speelt een belangrijke rol in deze scriptie.

Eigenschappen zijn nauw verbonden met causaliteit. Ik volg Shoemaker (1980) in zijn these dat de notie van eigenschap verklaard kan worden in de notie van causaliteit. Meer precies: een eigenschap is identiek met zijn causaal vermogen. Een eigenschap betekent dan precies het kunnen teweegbrengen van *dat* causaal effect. Een eigenschap 'rood' houdt bijvoorbeeld in het kunnen weerkaatsen van licht met golflengte van 650 nm, en het absorberen van licht met andere golflengten.

De relatie tussen eigenschap en causaal vermogen speelt in de scriptie een belangrijke rol bij het onderkennen van het causaal vermogen van emergente eigenschappen. De relatie tussen eigenschappen en causaliteit is niet omstreden, hoewel sommige filosofen wel vraagtekens stellen bij de door Shoemaker voorgestelde identiteit tussen eigenschap en causaal vermogen. Het definiëren van een eigenschap als causaal vermogen, is een 'externe' definitie. Daarmee wil ik zeggen dat een eigenschap, als eigenschap van een object, wordt gedefinieerd als een vermogen tot een causaal effect, dat het object vanwege die eigenschap op zijn omgeving, dus extern, kan hebben¹³.

De extern georiënteerde definitie van een eigenschap moet onderscheiden worden van de notie van relationele eigenschap. Dat is een eigenschap die een object kan hebben op grond van enige relatie naar een ander object. Bijvoorbeeld een positie van 2 meter naar links, of een temperatuur van + 10 graden ten opzichte van smeltend water. Relationele eigenschappen staan tegenover intrinsieke eigenschappen. Dat zijn eigenschappen die sterk met de aard van een object samenhangen. Als een intrinsiek eigenschap van een object verandert, dan verandert daarmee het object kenmerkend. Vorm is bijvoorbeeld een intrinsieke eigenschap. Als de vorm van een object verandert, zeggen we dat het object is veranderd. Locatie is een relatieve eigenschap. Als het object twee meter is opgeschoven, zeggen we niet dat daarmee het object is veranderd. Zowel intrinsieke als relationele eigenschappen kunnen extern gedefinieerd worden, d.w.z. in termen van hun causaal vermogen.

¹² Zie ook Armstrong, 1978:57, waar die onderscheid maakt tussen naampredicaten en analysepredicaten, dit zijn semantisch verschillende predicaten, die naar dezelfde eigenschap verwijzen.

¹³ Pettit (1998) noemt deze manier van definiëren van eigenschappen 'functionele' definities.

Tenslotte volgt uit de gepresenteerde notie van eigenschap dat 'reductie' geen categorie is die van toepassing is op eigenschap. Met andere woorden, een eigenschap is alleen identiek met zichzelf, en kan niet gereduceerd worden tot andere eigenschappen.

2. Non-reductie: Emergentie en Superveniëntie

In dit hoofdstuk geef ik een overzicht van de herkomst en de inhoud van de begrippen emergentie en superveniëntie.

2.1. Het Brits emergentisme

Belangrijke wetenschappelijke vorderingen in de tweede helft van de 19-e eeuw doen in de wetenschappelijke wereld van die tijd de overtuiging gemeengoed worden dat leven en geest voorkomen uit materiële voorwaarden. Eind 19-e eeuw zijn er dan een aantal metafysische posities te onderkennen met betrekking tot evolutie en ontwikkeling. De volgende vragen waren daarbij aan de orde: (1) verloopt evolutie continue of sprongsgewijs?; (2) is er sprake van het ontstaan van nieuwe eigenschappen, of van een kwantitatieve ontwikkeling van bestaande eigenschappen die in verschillende dichtheden voorkomen?; en (3) wat is de drijvende kracht is achter evolutionaire ontwikkeling?

Darwins rol met betrekking tot de ontwikkeling van een theorie over evolutie is bekend. Minder bekend is Spencers theorie over evolutionaire ontwikkeling. Beperkt Darwin (1809-1882) met zijn evolutietheorie zich tot de biologische wereld, Spencer (1820-1903) ontwikkelt de these dat het super-organische (d.w.z. sociale organisatie) zich ontwikkelt uit het organische, en dat het organische zich ontwikkelt uit het anorganische. Met name Spencer en Huxley (1825-1895) zijn het die aanzetten geven voor een materialistische ontologie (Zie voor een overzicht Blitz, 1992).

Het is de verdienste van Lloyd Morgan dat hij een aantal van deze ideeën samenbrengt in zijn werk *Emergent Evolution* (1923). De drie kernpunten van dit werk zijn: (a) Het organische (d.w.z. leven) komt voort uit het anorganische: evolutie is een proces van continuïteit en ontwikkeling, meer ontwikkelde organismen komen voort uit minder ontwikkelde organismen en ontwikkeling kent daarbij geen sprongen; (b) in het ontwikkelingsproces ontstaan kwalitatief nieuwe eigenschappen (bijvoorbeeld leven en geest), verbonden met continuïteit in 'resultante' eigenschappen. (Resultante eigenschappen zijn (vector)optelbare eigenschappen, bijvoorbeeld massa.) Op deze manier verbindt Lloyd Morgan de sprongen van het kwalitatief nieuwe met de continuïteit van het materiële. De 'kwalitatief nieuwe' eigenschappen ontleent Lloyd Morgan aan H. Bergsons idee van 'Creatieve Evolutie' (1907). (c) Het derde kernpunt van *Emergent Evolution* is dat de werkelijkheid te analyseren is in niveaus, of lagen, die gekarakteriseerd worden door kenmerkende emergente eigenschappen, die alleen op dat niveau voorkomen.

Volgens Blitz (1992:175) is het Lloyd Morgan die voor het eerst het ontstaan van kwalitatief nieuwe eigenschappen combineert met continue, d.w.z. niet sprongsgewijze ontwikkeling. In zijn these maakt Lloyd Morgan verder gebruik van werk van E.G. Spaulding (1912), waarin die stelt dat *gehelen* eigenschappen kunnen bezitten die 'nieuw' zijn in de zin dat die eigenschappen geen eigenschappen zijn van de *delen* van het geheel. Dergelijke 'nieuwe' eigenschappen worden door Lloyd Morgan *emergent* genoemd.

Als drijvende kracht achter emergente evolutie veronderstelde Lloyd Morgan een supernatuurlijke (deïstische) kracht. Volgens Blitz had Lloyd Morgans theorie sterker gestaan, indien hij zich niet geïmmiteerd had aan teleologische en theologische theses. Het is dan ook niet Lloyd Morgan, maar Broad (1887-1971) die in de latere non-reductiedebatten wordt geciteerd als vertegenwoordiger van het (Brits) emergentisme.

Broad is een psycholoog uit het Oxford van Whitehead en Russell. Broads boek *'The Mind and its Place in Nature'* (1925) is een neerslag van een collegeserie over de relatie tussen de fysische, de biologische en de psychologische wetenschappen. In het eerste deel van zijn boek bespreekt Broad drie theorieën: het 'mechanisme'; het 'vitalisme' en het 'emergentisme'. Broads doel is te onderzoeken welke verklaringen deze theorieën kunnen geven voor het ontstaan van leven en voor het ontstaan van de menselijke geest.

Het vitalisme veronderstelt het bestaan van een aparte substantie die verantwoordelijk geacht wordt voor leven. Deze levenssubstantie of 'entelechie' zou aanwezig zijn in levende organismen, maar niet in dode materie of in afgestorven organismen. De substantie zou niet geïsoleerd of geanalyseerd kunnen worden. Vitalisme houdt een substantiedualisme in, dat niet rijmt met het uitgangspunt van het ontologisch fysicalisme. Op grond van een vergelijking met chemische eigenschappen wijst Broad de mogelijkheid van een entelechie van de hand. Chemische eigenschappen berusten ook niet in een aparte substantie, maar zijn verbonden met de structuur van de chemische stof (Broad, 1925:58).

Ook het 'mechanisme' wordt door Broad afgewezen, als verklaring voor het bestaan van leven of van de menselijke geest. In Broads verwoording is mechanisme een positie waarin gesteld wordt dat (a) alle materiële dingen zijn samengesteld uit een enkele soort fundamentele deeltjes (dit is ook de veronderstelling van het ontologisch fysicalisme, GD), (b) dat er één fundamentele wet is die de interactie van de deeltjes bepaalt (dit is de veronderstelling van het nomologisch fysicalisme) en (c) dat de interactie van twee aggregaten van deeltjes een (vector)compositie is van de paarsgewijze interactie van de samenstellende deeltjes van de aggregaten met elkaar (dit is de veronderstelling van het atomisme, daarover later meer, GD).

'Mechanisme' zou volgens Broad in de vorm van een 'biologisch mechanisme' misschien nog het mechanistisch gedrag van een organisme kunnen verklaren, maar in het verklaren van levensvormen schiet het mechanisme tekort. Organismen zijn doelgericht. Als je uitgaat van 'mechanisme', dan is volgens Broad de veronderstelling van een ontwerper nodig om de doelgerichtheid van organismen te verklaren. Mechanisme leidt daarom volgens Broad tot een deïstische positie die Broad, anders dan Lloyd Morgan, afwijst (Blitz, 1992:118).

Het derde alternatief voor het verklaren van eigenschappen van levende organismen acht Broad plausibel. Dit is de theorie van emergentie. Emergentie houdt volgens Broad in dat de emergente eigenschappen van levende organismen bepaald worden door de aard en de samenhang van de samenstellende delen. Er is geen aparte substantie nodig, waaraan de levensgerelateerde eigenschappen verbonden zijn. Emergentie houdt ook in dat eigenschappen van bijvoorbeeld een biologisch geheel niet afleidbaar zijn van de eigenschappen van de delen van dat geheel. Met emergente eigenschappen zijn karakteristieke (emergente) wetmatigheden verbonden, die kenmerkend zijn voor het betreffende niveau van emergentie¹⁴. Evenals emergente eigenschappen zijn ook emergente wetmatigheden ('intra-ordinale wetten') niet af te leiden van de wetmatigheden van een onderliggend emergentieniveau. Emergente eigenschappen en emergente wetmatigheden kunnen volgens Broad alleen achteraf, empirisch vastgesteld worden, en niet vooraf voorspeld worden op grond van de eigenschappen en de

¹⁴ Emergente eigenschappen en emergente wetten zijn elkaars complement. Ik herinner aan de positie die ik in de inleiding heb ingenomen, dat eigenschappen identiek zijn met hun causaal vermogen. Wetmatigheden zijn precies de regelmatigheden in dat causaal vermogen.

wetmatigheden van de delen. Tenslotte houdt emergentie volgens Broad in dat de wetmatigheden waaronder op een emergentieniveau zich gehelen vormen uit delen ('inter-ordinale wetten'), uniek zijn voor het type organisme dat zich vormt. Ook deze wetmatigheden zijn niet afleidbaar als bijzondere gevallen van algemene wetten (Broad,1925:61-70).

2.2. De kernpunten van het emergentisme

Samenvattend zijn de kernpunten van het emergentisme: (zie ook Kim,1992:122; Emmeche c.s.,1997):

- (1) Een fysicalistische ontologie: Alle entiteiten in deze wereld zijn uiteindelijk aggregaten van basisentiteiten. Basisentiteiten zijn elementaire materiële deeltjes met uitsluitend fysische en niet-emergente eigenschappen.
- (2) Eigenschapsemergentie: Verbonden met de structuur waarin delen een geheel vormen, verkrijgt het geheel eigenschappen die kwalitatief nieuw zijn ten opzichte van de eigenschappen van de delen. Deze kwalitatief nieuwe eigenschappen worden emergente eigenschappen genoemd.
- (3) Emergente eigenschappen zijn niet afleidbaar: (a) Emergente eigenschappen van een geheel zijn niet afleidbaar van de eigenschappen van de delen van dat geheel. (b) De wetmatigheden, waaronder gehelen met emergente eigenschappen zich ontwikkelen uit clusterings van delen, zijn niet afleidbaar van fysische basiswetten.
- (4) De werkelijkheid laat zich analyseren in niveaus van emergentie: Op grond van de kwalitatieve verschillen in emergente eigenschappen kunnen niveaus of lagen onderkend worden, die getypeerd worden door karakteristieke emergente eigenschappen die alleen op dat niveau voorkomen.

In de volgende paragrafen ga ik verder op deze kernpunten in.

2.2.1. Een fysicalistische ontologie

Het belangrijkste van de eerste claim is dat het emergentisme een dualistische ontologie ontkent. Er bestaat geen andere substantie, dan materiële deeltjes. De veronderstelling dat er kleinste deeltjes zijn, die zelf geen emergente eigenschappen hebben, is problematisch voor de overige claims van het emergentisme. Kleinste deeltjes van de soort die in claim (1) benoemd worden, zijn niet in staat gehelen te vormen van de soort die het emergentisme veronderstelt: Een kleinste deeltje is niet meer deelbaar. Als het niet meer deelbaar is, heeft het geen interne structuur, en om die reden heeft het geen emergente eigenschappen. (Dit is ook wat claim (1) zegt.) Als het kleinste deeltje intrinsieke, niet-emergente eigenschappen heeft, dan doet zich het volgende probleem voor. Het veronderstellen van kleinste deeltjes (dus zonder interne structuur) *met* intrinsieke eigenschappen vormt de basis van het causaal uitsluitingsargument (*causal exclusion argument*). Causale uitsluiting houdt in dat één effect geen twee onafhankelijke oorzaken kan hebben. In paragraaf 3.3 ga ik in meer detail op dit argument in. De conclusie van dit argument is, dat veronderstellende kleinste deeltjes met niet-emergente, intrinsieke eigenschappen, emergente eigenschappen epifenomenaal moeten zijn. D.w.z. schijnbaar, bijkomend. Dit is strijdig met claim (3) van het emergentisme.

Als het kleinste deeltje daarentegen geen intrinsieke eigenschappen heeft, kan het alleen nog relationele eigenschappen hebben. Maar een relationele eigenschap heeft alleen causaal effect in een 'waarneming'. D.w.z. een relatie tussen twee objecten moet 'waargenomen'¹⁵ worden om van een relatie te kunnen spreken. Echter het enige object dat als 'waarnemer' in aanmerking komt, is een kleinste deeltje. Ik denk dat het niet veel tegenspraak ontmoet als ik stel dat 'kunnen waarnemen' een intrinsieke eigenschap is van een object. Intrinsieke

¹⁵ 'Waarnemen' bedoel ik hier in een zeer elementaire vorm, d.w.z. een 'gevoeligheid' voor een stand van zaken.

eigenschappen hadden we echter in dit tweede geval juist uitgesloten. Het gevolg daarvan is dat kleinste deeltjes zonder intrinsieke eigenschappen niet tot interactie in staat zijn. Zij kunnen geen gehelen vormen, en er kunnen in dat geval geen emergente eigenschappen ontstaan. Dit is tegenspraak met de tweede emergentistische claim.

De oplossing die het emergentisme kan redden, is de veronderstelling te laten varen dat er kleinste deeltjes bestaan zonder emergente eigenschappen. Tot op heden is de veronderstelling dat er *geen* kleinste deeltjes bestaan nog niet empirisch gefalsificeerd, en er zijn ook geen vooruitzichten, dat de veronderstelling binnenkort gefalsificeerd zal worden. Er zijn filosofen die de mogelijkheid dat er geen kleinste deeltjes bestaan openhouden (zie bijv. Bickhard & Campbell (z.j.), maar ook Pettit (1993a:214)).

De afwijzing van het bestaan van kleinste deeltjes met niet-emergente eigenschappen staat overigens niet op gespannen voet met een fysicalistische ontologie. Ook als de fysische deeltjes emergent zijn, kan gezegd worden dat hogere orde gehelen uiteindelijk zijn opgebouwd uit fysische deeltjes. Ik denk dat we claim (1) van het emergentisme op deze wijze moeten herinterpreteren. Claim (1) is tenslotte de uitdrukking van een materialistische ontologie, niet van reductionistisch fysicalisme.

2.2.2. Eigenschapsemergentie

Eigenschapsemergentie houdt in dat emergente eigenschappen van een geheel voortkomen uit de structuur waarin delen dat geheel vormen. Eigenschapsemergentie staat tegenover substantie-emergentie. Eigenschapsemergentie wil zeggen dat er niet een speciale substantie is, die verantwoordelijk is voor de emergente eigenschap. Een geheel met emergente eigenschappen bestaat, zonder residu, uit zijn samenstellende delen. Met deze claim onderscheidt het Brits emergentisme zich van het substantie-emergentisme, dat emergente eigenschappen toewijst aan een aparte substantie. Met de claim van eigenschapsemergentie bevestigt het Brits emergentisme zijn ontologisch fysicalistisch uitgangspunt.

Structuur is een belangrijke notie voor emergentie, reden om kort over structuur uit te weiden. Structuur in wiskundige zin is ordening. Als een geheel structuur heeft, dan zijn de delen van het geheel onderling geordend. Een geheel kan dan gedefinieerd worden als een geordende verzameling. In een verzameling komt de notie van het geheel naar voren (één verzameling), en de notie van delen (de elementen van de verzameling). In een ordende verzameling hebben de elementen onderlinge relaties.

In deze definitie van een geheel kan de verzameling een verzameling zijn van willekeurige elementen, die toevallig een onderlinge relatie hebben, bijvoorbeeld onderlinge afstand. Als we echter over een geheel praten, dan veronderstellen we een zekere mate van persistentie van dat geheel. Een willekeurige verzameling, waar elementen vrij naar binnen en naar buiten kunnen bewegen, voldoet niet aan het idee van een geheel. Ook mag de afgrenzing van een geheel niet willekeurig zijn. D.w.z. het mag niet aan de willekeur van bijvoorbeeld een externe waarnemer overgelaten worden, of delen tot een geheel behoren of niet. Het criterium of delen samen een geheel vormen, moet daarom uit de geordende verzameling zelf komen. Een kandidaat voor een dergelijk criterium is dat de delen onderlinge binding hebben. Het hebben van een binding sluit aan bij het idee van persistentie van een geheel. Een binding is een bijzondere vorm van relatie, een relatie waarbij het bij wijze van spreken energie kost (in welke opvatting dan ook), om de relatie te verbreken. Als ik het verder over gehelen heb, en over de relatiestructuren waarin de delen deze gehelen vormen, dan heb ik het over verzamelingen van onderling gebonden delen.

Eigenschapsemergentie in combinatie met het afzien van de veronderstelling dat er kleinste deeltjes bestaan, levert een probleem op voor het conventionele (aristotelische) materiebegrip. Aristoteles maakte een onderscheid naar vorm en substantie. Substantie is synoniem met materie, dat zich vormt in de verschillende verschijningsvormen waarin wij de materiële wereld waarnemen. Dit idee is nog steeds de basis voor het conventionele denken over materie. Indien er echter geen kleinste deeltje worden verondersteld, dan is er geen substantie, maar alleen vorm, in de hoedanigheid van structuur. Dat betekent dat, als we over materie willen blijven praten, materialiteit met structuur moeten verbinden.

Als dan op elk emergentieniveau nieuwe structuurvormen ontstaan, dan kan het ontstaan van deze nieuwe structuren opgevat worden als het ontstaan van nieuwe vormen van materie. Er zijn filosofen die deze radicale conclusie trekken, en die ook leven en geest tot vormen van materie rekenen. (Bijvoorbeeld Fink,1990, vgl. Emmeche c.s.,1997:94). Dit idee dat op elke emergentielaag nieuwe vormen van materie ontstaan, strookt echter niet met de veronderstelling dat er één werkelijkheid bestaat. Het kan daarom niet zo zijn dat met elke emergentiestap werkelijkheid wordt toegevoegd. Dit betekent dat met nieuwe emergente niveaus alleen nieuwe structuurvormen worden toegevoegd, en geen nieuwe ontologie. In hoofdstuk 5 kom ik hierop terug.

De opvatting dat niet kleinste deeltjes, maar structuren de ontologische basis vormen, wijkt sterk af van de standaardopvatting van het materiebegrip, waarin verondersteld wordt dat materie alleen uit fysische deeltjes bestaat. De standaardopvatting ten aanzien van de aard van materie is overigens niet strijdig met de opvatting dat er geen kleinste deeltjes bestaan. Het is goed mogelijk de fysische *emergentielaag* als de materiële laag te beschouwen, die de deeltjes bevat waaruit alle gehelen van bovenliggende emergentielagen uitputtend zijn samengesteld, en de emergentielagen beneden de fysische laag buiten beschouwing te laten.

2.2.3. De niet afleidbaarheid van emergente eigenschappen.

Het emergentisme stelt dat emergente eigenschappen van een geheel, kwalitatief nieuw zijn ten opzichte van de eigenschappen van de delen van het geheel, en niet afleidbaar van de eigenschappen van die delen. Broad (1925:59) zegt dit als volgt: '*... emergent theory asserts that there are certain wholes, composed (say) of constituents A, B and C in a relation R to each other; that all wholes composed of constituents of the same kind as A, B and C in relations of the same kind as R have certain characteristic properties; that A, B and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A,B,C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B and C in isolation or in other wholes which are not in the form of R(A,B,C).*' (1925:61).

Deze formulering stelt dat gehelen, die zijn samengesteld uit een bepaald type componenten (in Broads voorbeeld componenten van type A, B en C), en die een bepaald type relatiestructuur R tot elkaar onderhouden, een emergente eigenschap, zeg E bezitten. En alle gehelen die aan de voorwaarde R(A,B,C) voldoen, bezitten emergente eigenschap E. Vervolgens stelt Broad dat E niet *afgeleid* kan worden van de eigenschappen van de delen A,B,C in combinatie met R. Hoe moeten we dit begrijpen? Dit lijkt een tegenstelling.

In Broads formulering lopen drie begripdomeinen door elkaar. Dat een geheel met de specificatie R(A,B,C) een emergente eigenschap E heeft, is een uitspraak in het *ontologisch* domein. De emergente eigenschap E *bestaat* erin dat het een eigenschap is van een geheel met de samenstelling R(A,B,C). Hierbij worden geen uitspraken gedaan over de eigenschappen van A, B of C. Vervolgens doet Broad een uitspraak in het *epistemologisch* domein: als je delen A, B en C in een relatiestructuur R bij elkaar zet, of aantreft, dan heeft dat geheel de emergente eigenschap E. Het gaat hier om verklarende of voorspellende uitspraken, en dat

hoeven nog geen ware uitspraken te zijn. Ook hier wordt niets gezegd over de eigenschappen van A, B of C.

Tenslotte doet Broad een uitspraak in het *semantisch* domein: de eigenschap E is niet af te leiden (te deduceren) van de *eigenschappen* van A, B en C (*in isolatie*), en de relatiestructuur R. Dat wil zeggen er is geen logische relatie te leggen tussen predicaat p_E (de aanduiding voor eigenschap E), de predicaten p_a , p_b en p_c (de aanduiding voor de eigenschappen van A, B en C); en R. Met andere woorden, emergentie bestaat, maar is niet af te leiden.

De door Broad bedekt uitgesproken veronderstelling achter deze claim is dat de eigenschappen van de delen A, B en C, zich binnen de relatiestructuur R wijzigen (Broad spreekt over het 'vermengen' van eigenschappen (1925:61)), en wel zich wijzigen op een manier die specifiek voor R is. In een ander type relatiestructuur, zeg Q, wijzigen de eigenschappen van de delen A,B en C zich anders, waardoor E ook niet is af te leiden uit de eigenschappen van A, B en C in het geheel Q(A,B,C). Vandaar dat Broad zijn uitspraak kwalificeert, dat emergente eigenschappen niet zijn af te leiden van de eigenschappen van de delen in isolatie, noch van de eigenschappen van A, B en C in andere gehelen dan R(A,B,C).

Dat eigenschappen van een object veranderen onder de interactie die het object aangaat, is niet vreemd. Denk bijvoorbeeld aan een plank die doorbuigt als je er op gaat staan. De suggestie die Broad wekt, is dat als je de eigenschappen van A,B en C onder structuur R zou kennen, dat je daarmee de emergente eigenschap wel zou kunnen afleiden. En ik denk dat Broad het daar niet mee oneens zou zijn.

De kwalificatie 'in isolatie' moet gezien worden als de eigenschappen van de delen buiten een relatiestructuur. Op de kwalificatie 'in isolatie' zou tegengeworpen kunnen worden, dat ook in isolatie een object de *counter-factual* eigenschap heeft: 'zou je object A in een structuur R plaatsen samen met objecten B en C, dan zou het eigenschap p_a hebben'. En ook op grond daarvan zou p_E afgeleid kunnen worden. Het is echter zo dat je die counter-factual eigenschap pas kent als het geheel R(A,B,C) al een keer is opgetreden en de eigenschappen p_a , p_b en p_c binnen de relatiestructuur R empirisch zijn vastgesteld.

Het punt waar Broad uiteindelijk op uitkomt is dat er een niet te overbruggen 'gat' zit tussen het nog niet opgetreden zijn van R(A,B,C) en het voorspellen van R(A,B,C). Het probleem is het volgende: eigenschap p_a verandert onder R naar p_a^+ , maar p_a^+ is voorwaardelijk voor het optreden van R. Echter, zo lang je R niet kent, ken je p_a^+ niet, en kun je dus R niet afleiden. De conclusie van de emergentisten is dan dat emergente eigenschappen niet te voorspellen zijn, zolang zij nog niet opgetreden zijn. Het is met betrekking tot dit laatste probleem van emergentie, waar vanuit niet-lineaire theorieën openingen worden geboden.

2.2.4. Emergente niveaus

De vierde claim van het emergentisme is dat de werkelijkheid zich laat analyseren in niveaus van emergentie. De claim houdt in dat op grond van de kwalitatieve verschillen in emergente eigenschappen niveaus of lagen onderkend kunnen worden die getypeerd worden door karakteristieke emergente eigenschappen die alleen op dat niveau voorkomen. Een veel gehanteerde indeling van emergentieniveaus is de indeling naar fysisch, chemisch, biologisch, mentaal en sociaal niveau. (Bijvoorbeeld Emmeche c.s.,1997:112.) Het is opmerkelijk dat Salthe (1993:70) tot de ordening komt: fysisch, chemisch, biologisch, sociaal en mentaal. Deze ordening houdt bijvoorbeeld in dat sociale interactie voorwaarde is voor het bestaan van mentale eigenschappen. Salthes ordening strookt met Pettits (1993b) these dat sociale interactie voorwaarde is om van denken te kunnen spreken.

Er is blijkaar verschil van mening hoe emergente niveaus te ordenen. De ordening van Emmeche is mereologisch, d.w.z. een ordening naar gehelen en hun samenstellende delen. Deze ordening lijkt een probleem op te leveren voor Salthes ordening: mentale eigenschappen

zijn geen eigenschappen van sociale gehelen, maar van sociale individuen. Dit zou tot de conclusie moeten leiden dat *of* mentale eigenschappen geen emergente eigenschappen zijn (en ik denk dat deze conclusie niet acceptabel is, mentale eigenschappen waren het vertrekpunt om over emergentie te gaan denken), *of* dat denken verbijzonderd moet worden van mentale eigenschappen, waarbij denken een eigenschap is van sociale gehelen. Ik neig tot dit tweede alternatief¹⁶. Dat houdt in dat de mereologische ordening het uitgangspunt kan blijven voor ordening van emergente niveaus.

Aan emergentieniveaus worden door sommige emergentisten ontologische status toegekend, bijvoorbeeld door Emmeche c.s.,1997. Deze opvatting is niet zonder problemen. Bestaat bijvoorbeeld al een emergentieniveau, terwijl er nog geen enkel geheel bestaat met de emergente eigenschappen die kenmerkend zijn voor het betreffende emergentieniveau? Indien men deze vraag bevestigend beantwoordt, accepteert men het idee van potentialiteit. Dit leidt dan tot een idealistische ontologie, waarbij alle niveaus van emergentie potentieel reeds bestaan.

Volgens mij zijn de onderscheiden emergentieniveaus tamelijk arbitrair, en afgeleid van de wetenschappen die zich ter bestudering van de typen emergente eigenschappen toevalligerwijs hebben ontwikkeld. Ik denk dat er meer mereologische niveaus zijn dan de soorten wetenschappen weerspiegelen. Sommige typen gehelen zijn direct waarneembaar voor mensen. Bijvoorbeeld stenen, planten en dieren. Maar voor een ongeoefende waarnemer is het al veel lastiger om het onderscheid tussen fysische en chemische eigenschappen te maken. Daarnaast zijn er gehelen die niet als zelfstandig geheel kunnen bestaan, maar bestaan als onderdeel van een groter geheel. Bijvoorbeeld bepaalde molecuulclusters, en bijvoorbeeld organen van organismen. In de categorisering van emergentieniveaus van emergentie zijn deze 'subniveaus' buiten beschouwing gebleven. Het gevolg van het onderkennen van meer mereologische niveaus is dat de 'afstand' tussen de niveaus kleiner wordt, en dat de verschillen tussen de emergente eigenschappen van de opeenvolgende niveaus minder groot zijn dan de verschillen tussen de hoofdniveaus die de emergentisten onderscheiden.

2.3. Non-reductie en superveniëntie

Een positie die een aantal overeenkomsten vertoont met het emergentisme is de positie van het non-reductionistisch fysicalisme in het *mind-body* debat. Dit debat speelde zich met name af in de tweede helft van de 20-e eeuw. In het non-reductionistisch fysicalisme wordt gesteld dat mentale eigenschappen weliswaar voortkomen uit (of 'gerealiseerd worden in') materiële ('fysieke') eigenschappen, maar dat mentale eigenschappen niet te reduceren zijn tot die fysieke eigenschappen.

Eén van de argumentaties is van Davidson. Davidson probeert een positie te ontwikkelen waarin een oplossing geboden wordt voor de schijnbare tegenspraak tussen de volgende claims: (1) Mentale gebeurtenissen interacteren causaal met fysische gebeurtenissen; (2) causaliteit (oorzaak-gevolg) wordt beheerst door strikt deterministische wetmatigheden; (3) er zijn geen strikt deterministische wetten waarmee mentale gebeurtenissen voorspeld en verklaard kunnen worden. Hoe valt dit te rijmen? (Davidson,1970:208). De positie waarin Davidson een oplossing voor dit probleem ontwikkelt, wordt het anomalistisch monisme genoemd. Het anomalistisch monisme houdt in dat alle gebeurtenissen fysiek zijn (dat is het monisme in de positie: er is slechts één fysieke werkelijkheid), maar het ontkent dat mentale gebeurtenissen op

¹⁶ Het argument van Pettit dat sociale interactie voorwaarde is voor 'regelvolgen', waarbij Pettit regelvolgen equivalent stelt met denken, beweegt in deze richting. Hoewel Pettit als individualist denk ik terugschrikt om denken als eigenschap van een sociaal geheel te zien, en niet als eigenschap van een individu (Zie Pettit, 1993b).

basis van fysische wetmatigheden verklaard kunnen worden (dit is het 'anomalisme' van het mentale).

Om zijn positie toe te lichten introduceert Davidson de term superveniëntie. Davidson ontleent de term aan G.E. Moore, een ethisch intuitionist uit het begin van de twintigste eeuw. Moore gebruikt de term superveniëntie om de relatie tussen een beschrijvend en een normatief kenmerk aan te geven, bijvoorbeeld de relatie tussen een handeling van een persoon en het juist zijn van die handeling. Moore stelt dat het beschrijvend kenmerk weliswaar in zeker opzicht het moreel kenmerk bepaalt, maar dat niettemin het moreel kenmerk niet te reduceren is tot dat beschrijvende kenmerk. Davidson neemt het gebruik van de term 'superveniëntie' over naar de relatie tussen mentale en fysische gebeurtenissen, en stelt dat een mentale gebeurtenis wel in zeker opzicht wordt bepaald door een fysische gebeurtenis, maar dat niettemin de mentale gebeurtenis niet te reduceren is tot de fysische gebeurtenis. 'Het in zekere opzicht bepalen' wordt door Davidson uitgelegd als (Moore aanhalend): 'het is onmogelijk voor twee gebeurtenissen om overeen te komen in al hun fysische kenmerken, en te verschillen in hun mentale kenmerken' (Davidson, 1970, 214).

Een tweede argumentatie dat mentale eigenschappen niet te reduceren zijn tot fysieke eigenschappen is het multipele implementatie-argument. Dit argument is met name ontwikkeld door Putnam en door Fodor. Putnam (1967) stelt dat mentale eigenschappen functionele eigenschappen zijn van het organisme, en dat als twee organismen gelijke functionele eigenschappen hebben dat dit niet impliceert dat de fysieke organisatie van beide organismen hetzelfde moet zijn. Dezelfde functionele eigenschap kan op verschillende manieren geïmplementeerd worden. Als voorbeeld wordt hier een computeralgoritme genoemd, dat op verschillende manieren, en in verschillende computertalen geïmplementeerd kan worden. Fodor (1974) werkt dit verder uit, en laat zien dat als een mentale eigenschap óf door eigenschap P, óf door eigenschap P', óf door eigenschap P'', etc. gerealiseerd kan worden, dat er dan geen wetmatige relatie gelegd kan worden tussen een mentale eigenschap van een organisme, en de fysische eigenschappen van dat organisme¹⁷. Het superveniëntiebegrip accommodeert ook de notie van multipele implementatie, reden waarom het begrip superveniëntie in het *mind-body* debat een belangrijke plaats heeft ingenomen.

Met name Kim heeft een belangrijke bijdrage geleverd aan het tot stand komen van wat een standaardformulering van superveniëntie genoemd zou kunnen worden. Kim heeft gepoogd een zodanige definitie van superveniëntie op te stellen, dat de relatie tussen een superveniënte eigenschap en de gerelateerde subveniënte eigenschappen (of superveniëntiebasis) voldoende sterk is om een vorm van bepaling uit te drukken, maar niet sterk genoeg om van een wetmatige relatie te kunnen spreken, op grond waarvan de superveniënte eigenschap reduceerbaar zou zijn tot de superveniëntiebasis.

Superveniëntie is door Kim (1990) gedefinieerd als een *covariatie* tussen twee eigenschappen, in het debat vaak aangeduid als S en P. Covariatie is een symmetrische relatie, maar superveniëntie wordt geacht dat niet te zijn. In een superveniëntierelatie supervenieert de superveniënte eigenschap S op de subveniënte eigenschap P. Als eigenschap P voorkomt, komt eigenschap S ook voor. Eigenschap P is in zekere opzicht bepalend voor eigenschap S en P gaat ontologisch *vooraf* aan eigenschap S. Omgekeerd echter impliceert het bestaan van eigenschap S niet het bestaan van eigenschap P.

Had voor Davidson de superveniëntierelatie betrekking op een relatie tussen beschrijvingen (predicaten), door Kim is de superveniëntierelatie naar het ontologisch domein

¹⁷ Kim (1998:106) bestrijdt Fodors opvatting en stelt dat een reductierelatie: S is reduceerbaar tot $(P \vee P' \vee P'' \vee P''' \dots)$ een acceptabele uitdrukking is, ook indien deze reeks oneindig zou zijn.

getransponeerd, door het als een relatie tussen eigenschappen te definiëren. Een drietal superveniëntiedefinities worden in de literatuur veelal besproken, en ik zal ze hier kort behandelen. Er zijn overigens meer varianten van superveniëntie in omloop¹⁸.

2.3.1. De superveniëntiedefinities

De discussie in het *mind-body* debat heeft zich voornamelijk rond de volgende drie vormen van superveniëntie bewogen: zwakke-, sterke- en globale superveniëntie. Deze varianten worden als volgt geformuleerd (Kim,1990):

Zwakke superveniëntie:

'A weakly supervenes on B if and only if necessarily for any object x and any property F in A: if x has F, then there exists a property G in B such that x has G, and if any y has G, y also has F'.

A en B zijn hierin verzamelingen eigenschappen van een object x, waarbij de verzameling A superveniënt is op B. A is bijvoorbeeld een verzameling mentale eigenschappen en B een verzameling fysieke eigenschappen. De definitie zegt kortweg dat als x een superveniënte eigenschap F heeft, dat x dan ook een subveniënte eigenschap G heeft. De superveniëntiedefinitie accommodeert zowel multiële implementatie, als het anomalistisch monisme. De definitie staat toe dat er meer eigenschappen G bestaan, waarbij eigenschap F als superveniërende eigenschap optreedt (multiële implementatie), en verder zegt de definitie alleen iets over covariatie (Als x G heeft, dan heeft x ook F), maar niets over een strikt causaal verband tussen beide eigenschappen. De superveniëntiedefinitie biedt dus ook ruimte aan het anomalistisch monisme.

De zwakke formulering van superveniëntie staat toe dat, toevallig, voor onze wereld geldt dat $\forall y(Gy \rightarrow Fy)$, maar dat dit in andere mogelijke werelden niet zo is. De zwakke formulering laat de mogelijkheid open dat in een andere mogelijke wereld een object x hersenen heeft als in onze wereld, maar geen mentale eigenschappen; het laat ook de mogelijkheid toe dat in een andere mogelijke wereld stenen mentale eigenschappen hebben. Deze mogelijkheden binnen de definitie van zwakke superveniëntie geven volgens Kim onvoldoende uitdrukking aan onze intuïtie dat het toch juist hersenen zijn die voorwaardelijk zijn voor mentale eigenschappen.

De definitie van sterke superveniëntie geeft beter uitdrukking aan de notie van bepaling:

Sterke superveniëntie:

'A strongly supervenes on B if and only if necessarily for any object x and any property F in A: if x has F, then there exists a property G in B such that x has G, and necessarily if any y has G, y also has F'.

In deze formulering geldt voor alle mogelijke werelden dat als een entiteit eigenschap G heeft, deze ook eigenschap F heeft. De sterke formulering zegt bijvoorbeeld dat in alle mogelijke werelden, als een object hersenen heeft, het object ook mentale eigenschappen heeft. Het probleem van sterke superveniëntie is echter dat de relatie tussen G en F teveel een wetmatige relatie wordt, wat op gespannen voet staat met de eis van non-reductie. Om die reden is het concept van globale superveniëntie ontwikkeld:

Globale superveniëntie:

'A globally supervenes on B if and only if any two worlds w_i and w_j which are B-indiscernible are also A-indiscernible.

¹⁸ Bijv. Kim, 1998:15 praat ook over mereologische superveniëntie, d.w.z. de relatie tussen de delen en een geheel.

Anders dan bij zwakke en sterke superveniëntie heeft globale superveniëntie niet betrekking op de eigenschappen van een concreet object x . Globale superveniëntie houdt in dat bepaalde *soorten* eigenschappen altijd samengaan. Het probleem van globale superveniëntie is dat het toestaat dat in *een* mogelijke wereld, met een voor mentale eigenschappen irrelevant fysisch verschil, in het geheel geen mentale eigenschappen voorkomen, of mentale eigenschappen voorkomen in een volstrekt andere vorm. Kim noemt hierbij als voorbeeld dat als in een mogelijke wereld de ring van Saturnus één molecule meer ammonia bevat dan in onze wereld, het binnen de mogelijkheid van globale superveniëntie valt dat die wereld dan geheel geen mentale eigenschappen zou kennen (Kim,1990).

Volgens Kim (1998:7) was superveniëntie voor non-reductionistisch fysicalisten interessant, omdat het mogelijk een metafysica van non-reductionistisch fysicalisme zou kunnen leveren. Kim laat zien dat superveniëntie alleen een *formulering* is voor het mind-body probleem, en geen oplossing ervan. Zo kan de superveniëntieformulering gebruikt worden door verschillende, onderling strijdige posities ten aanzien van de mind-body relatie. Superveniëntie is een covariatie, d.w.z. een uitspraak (en geen metafysische relatie), over het samen optreden van gebeurtenissen of eigenschappen. Het zegt niets over de *aard* van de relatie tussen een superveniënte eigenschap en zijn superveniëntiebasis.

2.3.2. Superveniëntie en emergentie zijn niet synoniem

In de literatuur is het verschil tussen emergentie en superveniëntie niet altijd even helder. Voor sommige auteurs in het superveniëntiedebat lijken beide begrippen synoniem te zijn. (o.a. Fodor (1974), Yablo(1992),Humpreys(1997), Beckermann (1992). In de voorgaande paragrafen is naar ik hoop duidelijk naar voren gekomen dat emergentie zich in het ontologisch domein afspeelt en dat de emergentierelatie een metafysische relatie is, namelijk de relatie tussen een interactiestructuur van delen en de eigenschap van een geheel dat in die interactiestructuur gevormd wordt. Superveniëntie is een relatie in het semantisch domein. Het is een *uitspraak* over het samen optreden van eigenschappen van dezelfde x , en zegt niets over de ontologische relatie van die eigenschappen. (Behalve dat de superveniëntiebasis ontologisch prior is.)

Als emergentie en superveniëntie verschillende begrippen zijn, is er dan een relatie tussen beide begrippen? Die relatie laat zich uitleggen aan de hand van een uitspraak van Beckermann waar die zegt dat een emergente eigenschap van een geheel supervenieert op de 'fysieke' structuur van dat geheel (1992:103).

Afhankelijk van hoe we 'fysieke' structuur opvatten, is Beckermann wel en niet correct. 'Fysieke' structuur kunnen we op twee manieren opvatten. Ten eerste als de topologische structuur, die ik aansluitend bij Broads voorbeeld beschrijf als $\Sigma(A,B,C|R)$. Dit is de topologische configuratie van de delen A, B en C onder structuur R: d.w.z. delen in hun onderlinge positie. De tweede opvatting van 'fysieke' structuur is de interactiestructuur van het geheel, die ik beschrijf als $\Pi(A,B,C|R)$, d.w.z. de delen A,B en C onder de interactiestructuur R. Π staat hier voor 'product' van de transformatiefuncties A,B en C onder de interactiestructuur R.

Als Beckermann zegt dat een emergente eigenschap E van een geheel supervenieert op de 'fysieke' structuur van dat geheel, dan is hij correct als we over de topologische structuur $\Sigma(A,B,C|R)$ praten. De emergente eigenschap E en de topologische structuur $\Sigma(A,B,C|R)$ zijn verschillende eigenschappen van het geheel x , die samen voorkomen, zonder dat er een causale relatie tussen de twee bestaat. Bijvoorbeeld als het samen voorkomen van de eigenschappen vorm en kleur. Als echter met 'fysieke' structuur de interactiestructuur $\Pi(A,B,C|R)$ bedoeld wordt, dan wordt Beckermanns uitspraak tautologisch, omdat (en ik loop hier vooruit op het

latere betoog in deze scriptie) het causaal vermogen van de eigenschap, *aangeduid* met E en van de interactiestructuur, *aangeduid* met $\Pi(A,B,C|R)$ identiek is, zodat de aanduiding E en de aanduiding $\Pi(A,B,C|R)$ verwijzen naar dezelfde emergente eigenschap¹⁹. En de uitspraak dat een eigenschap op zichzelf supervenieert is tautologisch.

Superveniëntie als covariatie is een tamelijk zwak begrip. Het zegt alleen *dat* er een relatie tussen eigenschappen is, maar het zegt niets over de aard van die relatie. Het zegt hooguit dat het geen wetmatige relatie kan zijn, waardoor reductie is uitgesloten. Het is opmerkelijk dat het begrip emergentie vrijwel afwezig is in het *mind-body* debat. Ook het emergentisme zegt dat de wetmatigheden betreffende emergente eigenschappen niet reduceerbaar zijn tot de wetmatigheden betreffende de eigenschappen van de delen. De verklarende waarden van emergentie is echter beter dan die van superveniëntie, omdat emergentie iets zegt over de relatie tussen delen, de relatiestructuur en de emergente eigenschappen van een aldus gevormd geheel. Is het dan misschien zo dat superveniëntie beter aansluit op de non-reductionistische positie in het *mind-body* debat, omdat het toereikende rechtvaardiging verschaft om de autonome positie van de psychologische wetenschap te handhaven, en de non-reductionistische positie ook niet meer vraagt dan dat?

2.4. Non-reductie: Samenvatting en conclusie

In dit hoofdstuk heb ik de non-reductionistische positie beschreven, met betrekking tot de vraag hoe zaken als leven en geest te integreren zijn in een materialistische positie. In het emergentisme en in het *mind-body* debat is gezocht naar beschrijvingen die een relatie leggen tussen leven, geest en materie, zonder zich te hoeven beperken tot het fysicalisme. Daartoe zijn begrippen als emergentie en superveniëntie ontwikkeld, die een relatie beschrijven tussen leven, geest en materie, zonder dat er sprake is van reductie. Het begrip emergentie heb ik zonder veel bijstelling overgenomen uit het emergentisme. De claim van het bestaan van niet-emergente kleinste deeltjes is problematisch voor het emergentisme, en ik heb de claim bijgesteld naar die van het ontologisch fysicalisme. Dit houdt in dat alles is opgebouwd uit fysische deeltjes, die echter niet de kleinste deeltjes hoeven te zijn, en waarbij ook de fysische eigenschappen emergent zijn.

Het superveniëntiebegrip in het *mind-body* debat heeft qua verklarend vermogen welbeschouwd niet zoveel opgeleverd. Maar misschien was dat ook niet de bedoeling in het non-reductionisme, dat in superveniëntie toereikende rechtvaardiging vond om de autonome positie van de psychologie te handhaven.

¹⁹ In de laatste zin vindt een subtiele overstap plaats van het ontologisch domein naar het semantisch domein. Dit is een gevolg van het feit dat we niet over het ontologisch domein kunnen praten zonder semantische begrippen te gebruiken.

3. Fysicalisme

3.1. Inleiding

Het fysicalisme is de opvatting dat de empirische²⁰ wereld '*contains just what a true complete physics would say it contains*.' (Crane en Mellor, in Pettit, 1993a:213). In de inleiding heb ik reeds het onderscheid gemaakt tussen het ontologisch fysicalisme dat stelt dat alles wat in de empirische wereld bestaat uiteindelijk is opgebouwd uit fysische deeltjes, en het nomologisch fysicalisme dat zegt dat alles in deze wereld ondergeschikt is aan fysische wetmatigheden. Het onderschrijven van het ontologisch fysicalisme impliceert nog niet het accepteren van het nomologisch fysicalisme. Daarom kunnen non-reductionisten een fysicalistische ontologie aanhangen, zonder nomologisch fysicalist te zijn.

3.2. Het fysicalistisch dilemma

Ten aanzien van de fysicalistische positie schetsen Crane en Mellor (1993) het dilemma: indien een 'ware en complete fysica' alles zou beschrijven dat in de empirische wereld bestaat, dan zou de fysica absurd veel omvatten, namelijk alle huidige wetenschappen, dus inclusief biologie, psychologie en sociologie, en dan zouden we niet meer over 'fysica' (als kennis van de natuur) kunnen praten. Indien echter een 'ware en complete fysica' zegt dat er alleen maar (micro)fysische objecten zijn met (micro)fysische eigenschappen, dan is de positie van het fysicalisme niet houdbaar, omdat de empirische wereld objecten bevat die weliswaar een fysische samenstelling hebben, maar die niet uitsluitend fysische eigenschappen bezitten (bijvoorbeeld leven en geest).

Pettit (1993a) probeert een definitie van 'fysicalisme' te formuleren, waarin hij dit dilemma kan ontlopen. Als werkdefinitie voor fysicalisme neemt Pettit de definitie van Crane en Mellor over: 'Het fysicalisme is het standpunt dat de empirische wereld precies dat bevat wat een ware en complete fysica zou zeggen dat het bevat'. Pettit werkt dit uit naar de volgende fysicalistische claims: (1) Er is een empirische wereld bestaande uit microfysische deeltjes, van de soort die de fysica veronderstelt. (2) Alles in deze wereld is uitputtend, zonder rest, samengesteld uit microfysische deeltjes die zelf niet meer samengesteld zijn. (3) Microfysische deeltjes zijn onderworpen aan microfysische wetmatigheden met betrekking tot hun microfysische eigenschappen. (4) Alles in deze wereld is onderworpen aan microfysische wetmatigheden. In de claims (1) t/m (3) kunnen ook non-reductionisten zich vinden. Het is de vierde claim, waarin het kernpunt van het nomologisch fysicalisme naar voren komt.

De basisvraag in het debat tussen nomologisch fysicalisten en non-reductionistisch fysicalisten is hoe zich macroniveau-wetten verhouden tot microniveau-wetten. Op grond van claim (4) kan het niet zo zijn dat macroniveau-wetten 'ruimte' opvullen die de microniveau-wetten overlaten. En evenmin kan het op grond van claim (4) zo zijn dat macroniveau-wetten

²⁰ De term 'empirisch' is afkomstig van Crane en Mellor. De vraag of een '*true and complete physics*' in dat geval zou kunnen zeggen dat er *non-observables* zijn, laat ik in het midden.

onafhankelijk zijn van de microniveau-wetten. In beide gevallen zou niet alles in deze wereld onderworpen zijn aan microfysische wetmatigheden. Als er dan macroniveau-wetten bestaan, waaraan macroniveau objecten onderworpen zijn, en als die macroniveau objecten vervolgens zijn samengesteld uit microniveau objecten, die onderworpen zijn aan microniveau-wetten, hoe zijn die macro- en microniveau-wetten dan gerelateerd?

Volgens Pettit is één mogelijkheid te ontkennen dat er macroniveau-wetten bestaan. Dit is de eliminatistische positie. Deze positie stelt dat er alleen maar microniveau-wetten bestaan. Deze positie heeft dan uit te leggen hoe macroniveau gebeurtenissen, bijvoorbeeld biologische of psychologische gebeurtenissen, verklaard kunnen worden op basis van microniveau-wetten. De andere mogelijkheid is te erkennen dat er macroniveau-wetten bestaan, maar dan moet, indien aan claim (4) vastgehouden wordt, gesteld worden dat macroniveau-wetten afhankelijk zijn van microniveau-wetten, en dat macroniveau regelmatigheden uiteindelijk terug te voeren zijn op microniveau-regelmatigheden. Pettit noemt deze positie 'nomologisch fundamentalistisch'(1993a:219). De nomologisch fundamentalist moet dan concluderen dat macroniveau-wetten epifenomenaal²¹ zijn, en dat alle causale gebeurtenissen onderworpen zijn aan microniveau-wetten.

Beide nomologisch fysicalistische posities vallen ten prooi aan het dilemma dat Crane en Mellor schetsen. Als een '*true complete physics*' zegt dat er alleen maar microniveau-objecten bestaan, dan ontkent dat onze waarneming dat er ook bijvoorbeeld mensen bestaan, d.w.z. macro-objecten, die meer zijn dan de optelling van de samenstellende deeltjes. Een dergelijke fysicalistische positie kan niet juist zijn. Als echter om de fysicalistische positie te redden een '*true complete physics*' moet zeggen dat er ook mensen bestaan, dan moeten alle wetenschappen tot de fysica gerekend worden, en dan wordt fysicalisme betekenisloos. De conclusie die Crane en Mellor dan trekken, is dat claim (4) niet houdbaar is: niet alles in deze wereld is onderworpen aan microfysische wetten.

De positie die Pettit probeert te ontwikkelen is de these dat nomologisch fundamentalisme niet hoeft te leiden tot epifenomenalisme, noch tot reductionisme. Veronderstellend dat claim (4) waar is, waardoor macroniveau-wetten afhankelijk zijn van microniveau-wetten, stelt Pettit dat er praktische belemmeringen kunnen bestaan voor ons mensen, om ooit de precieze afhankelijkheidsrelatie tussen macroniveau-wetten en microniveau-wetten te kunnen analyseren. Pettit stelt dan (1993a:220) dat het voldoende is te veronderstellen dat er een bepaalde 'architectuur' is, waardoor de antecedenten (de begincondities) van macroniveau-wetten, microfysisch gerealiseerd zijn. Deze microfysische configuraties van de begincondities veroorzaken, onder microfysische wetten de microfysische configuraties van de macro-consequensa. Pettit zegt dan dat de macro-oorzaak 'programmeert' voor het macro-effect. 'Programmeren' wil dan zeggen dat alle microfysische voorwaarden ingevuld zijn die op microfysisch niveau leiden tot microfysische configuratie van de consequens. De 'architectuur' zorgt ervoor dat de relatie tussen macro- en microniveau steeds gewaarborgd blijft, d.w.z. dat de macroniveau-wetten afhankelijk blijven van de microniveau-wetten.

Onder andere causale paradigma's²² dan het nomologisch fysicalisme kan dan gedaan worden *alsof* de macro-beginconditie de oorzaak van het effect is, gegeven de macroniveau wetten (P.222). Met deze aanname kunnen we dan praktisch uit de voeten. Binnen de

²¹ Epifenomenalisme wil zeggen dat een beginconditie, die op grond van een macroniveau-wet wordt aangewezen als een oorzaak, in werkelijkheid geen causale relevantie heeft, omdat de causaal relevante antecedenten zich microniveau bevinden.

²² D.w.z. andere wetmatigheden, waarbij Pettit (1993a:220) als voorbeelden noemt: empirische wetmatigheden, counterfactuals en energieoverdrachtsrelaties.

praktische causale paradigma's is er dan geen sprake van epifenomenalisme, noch van reductie.

Pettit probeert de hoorns van het dilemma dat Crane en Mellor schetsen te ontlopen door te zeggen dat er geen hoorns zijn. Pettit zegt: 'Kijk hoe mooi de wereld met de microfysica als basis is geordend, en hoe in de 'architectuur' geregeld is dat de macroniveau-wetten afhankelijk blijven van de microniveau-wetten. Je kunt fysicist zijn, terwijl je je niet druk hoeft te maken over de vraag hoe het macroniveau gerelateerd is aan het microniveau'. 'En als je naar de relatie tussen het macroniveau en het microniveau gevraagd wordt, dan zeg je dat de macroniveau-wetten superveniëren op de microniveau-wetten'. Superveniëntie als schaamlap.

Pettit gaat aan de kern van de zaak voorbij. Die kern is precies de vraag hoe macroniveau antecedenten kunnen 'programmeren' d.w.z. de voorwaarden kunnen scheppen waaronder de micro-causaliteit zijn werk kan doen, en tot de macroniveau effecten leiden. Anders gezegd, hoe werkt die 'architectuur'?

In paragraaf 2.2.2 heb ik verondersteld dat er sprake moet zijn van *bindingen* als deeltjes een configuratie vormen die een zekere persistentie heeft. Als bindingen niet van buitenaf opgelegd worden, en het fysicalisme gaat daar niet van uit, dan moet er een mechanisme zijn dat deeltjes zich onderling binden. Een uitweiding over de fysische voorwaarden waaronder binding mogelijk is, gaat het bestek van deze scriptie te buiten. Een suggestie zou kunnen zijn dat een binding tussen twee deeltjes een toestand oplevert die vanuit energetisch oogpunt gezien voordeliger is. Deze suggestie strookt met de waarneming dat het energie kost om deeltjes uit een binding vrij te maken.

Punt is nu dat wil je de twee deeltjes in hun *gebondenheid* adequaat kunnen beschrijven, dan ontcom je er niet aan een expressie te gebruiken die de gebonden deeltjes *samen*, d.w.z. het geheel uitdrukt. De *beide* deeltjes zijn noodzakelijk voor de binding. Er is geen binding meer als je een van de deeltjes weglaat. In hun relatie tot het geheel bezitten de deeltjes ingeperkte vrijheidsgraden. Dit houdt niet in dat het geheel zijn structuur 'oplegt' aan deeltjes, het houdt alleen in dat het voor een deeltje 'voordeliger' (bijvoorbeeld energetisch voordeliger) is om deel van zo'n structuur uit te maken. Wil nu het fysicalisme een configuratie van gebonden deeltjes adequaat beschrijven, dan moeten zij een begrip introduceren voor dat geheel. Dat betekent dus een uitbreiding van de termen in de fysische theorie. Indien gebonden configuraties samen ook weer gebonden configuraties vormen, herhaalt zich de procedure, en moet de fysica ook een begrip introduceren voor dat grotere geheel. Als we dan na een aantal van deze stappen qua mereologisch niveau bij de gebonden configuratie 'mens' zijn aangekomen, dan moet voor een fysicistische beschrijving van een gebonden configuratie 'mens' ook het begrip 'mens' geïntroduceerd worden. Dit is de tweede hoorn van het fysicistisch dilemma van Crane en Mellon, dat indien het fysicalisme zich uitstrekt tot buiten het terrein van microfysische deeltjes, het de complete wereld moet beschrijven in al zijn niveaus. En dit is niet waar we aan denken als we het over fysicalisme hebben.

Een tegenwerping op deze argumentatie zou kunnen zijn dat wijze waarop microfysische deeltjes betrokken kunnen zijn in de vorming van gehelen en grotere gehelen, etc. dispositionele eigenschappen van microfysische deeltjes zijn. Maar deze tegenwerping ontloopt ook niet de tweede hoorn van het dilemma. Ook in het geval waarin deeltjes dispositionele eigenschappen bezitten, dat zich uit die deeltjes een mens kan vormen, is het nodig in de theorie aangaande die eigenschappen het begrip 'mens' te introduceren. De projectie van de complete werkelijkheid als dispositionele eigenschappen van microdeeltjes lijkt op Leibniz' monade²³, en ik zou deze positie het monadisch fysicalisme willen noemen.

²³ Ook Leibniz heeft zich beziggehouden met de vraag van gehelen en en delen, en met de vraag van continuïteit en kwalitatief nieuwe eigenschappen van gehelen. (Blitz,1992:10).

De conclusie is dat Pettits voorstel voor een fysicalistische werkhypothese volledig voorbijgaat aan dit probleem van het fysicalisme, en dat de enige uitweg uit het probleem de verwerping van claim (4) is. Het nomologisch fysicalisme is niet houdbaar. Niet alles in deze wereld is onderworpen aan microfysische regelmatigheden. Een '*true complete physics*' kan alleen uitspraken doen over de aard van (micro)fysische deeltjes en hun (micro)fysische eigenschappen (en waar deze als fysische eigenschappen 'overerven' naar configuraties van (micro)fysische deeltjes. Bijvoorbeeld een mens heeft nog steeds een fysische eigenschap 'massa').

Zouden we daarentegen de fysica uitspraken laten doen over gehelen met bijbehorende emergente eigenschappen, bijvoorbeeld over organismen, of mensen, dan kunnen we niet meer praten over een fysica, zoals wij in normaal spraakgebruik fysica bedoelen. We zouden dan een andere term voor een dergelijke omvattende theorie moet zoeken. De conclusie is dat de term 'fysicalisme' verwarrend is en niet bijdraagt aan een helder debat. Het ontologisch fysicalisme is een algemeen geaccepteerde positie, en is niet problematisch. Het nomologisch fysicalisme is controversieel en naar mijn mening niet houdbaar. Om de helderheid in het debat te vergroten zou ik willen aanbevelen het ontologisch fysicalisme als 'materialisme' aan te duiden, en het nomologisch fysicalisme als 'fysicalisme'. In deze laatste verbijzonderde aanduiding is het fysicalisme dan geen houdbare positie.

3.3. Het Atomisme

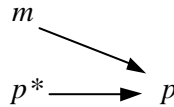
Een positie die verwant is aan het nomologisch fysicalisme is het atomisme. Ook het atomisme zegt dat er kleinste deeltjes bestaan (atomos ~ ondeelbaar). Het specifieke van het atomisme is dat atomen op zichzelf staan, en dat de wijze waarop atomen samenstellingen vormen, volledig uit de atomen zelf verklaard kan worden (zie ook Garfinkel, 1981:454). Het atomisme is verwant aan het individualisme in de sociale wetenschappen. Het individualisme is een positie tegenover het collectivisme. Het individualisme stelt dat het gedrag van individuen in collectieven uit die individuen zelf verklaard kan worden, en wijst het idee af dat een collectief het gedrag van een individu dwingend bepaalt²⁴.

Het atomisme stelt dat atomen reeds alle eigenschappen bezitten, waaronder zij gehelen kunnen vormen, en dat de eigenschappen van die gehelen zijn terug te voeren op (combinaties) van eigenschappen van de atomen. Gehelen zijn daardoor volgens de atomisten geen onafhankelijke entiteiten ten opzichte van de atomen, zoals de emergentisten beweren, maar de eigenschappen van gehelen liggen al in de kiem besloten in de atomen. Het monadisch fysicalisme is een vorm van atomisme, maar niet de vorm die atomisten erkennen. Atomisten onderkennen namelijk qua eigenschappen van hun veronderstelde kleinste deeltjes alleen de microfysische eigenschappen, en wijzen de monadische eigenschappen af.

Het atomisme speelt onderhuids een belangrijke rol in het *mind-body* debat. Het is namelijk de impliciete veronderstelling achter het uitsluitingsprobleem en het ermee samenhangende neerwaartse causatieargument. Het uitsluitingsprobleem vormt een probleem voor de non-reductionistisch fysicalisten. Het uitsluitingsprobleem is gebaseerd op het uitgangspunt dat één effect geen twee onafhankelijke oorzaken kan hebben. Aanleiding voor het probleem is de these dat mentale gebeurtenissen oorzaken kunnen zijn, d.w.z. dat mentale gebeurtenissen fysieke gebeurtenissen kunnen veroorzaken. (Dit is de these van *mental*

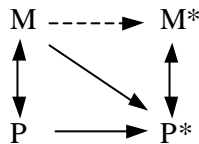
²⁴ Ik vermoed dat het vasthouden aan het idee van atomisme wordt gemotiveerd door de politieke ongewenstheid van het collectivisme. Het idee is dat als men het atomisme zou moeten opgeven, dat dan de weg zou openen voor de mogelijkheid van collectivisme. Het atomisme is niet houdbaar zoals verderop zal blijken. De angst voor het collectivisme betreft echter een ander probleem, dat je niet met het atomisme kunt bezweren. Dat probleem betreft namelijk de vraag welke vrijheid er voor een individu overblijft als de groep waarvan hij deel uitmaakt, voor een andere samenwerkingsstructuur kiest dan de structuur die het betreffende individu voorstaat.

causation.) Het uitsluitingsprobleem laat zich als volgt formuleren (Yablo,1992; Kim,1998:37): Als een mentale gebeurtenis m een fysieke gebeurtenis p kan veroorzaken (mentale veroorzaking), en als voor een fysieke gebeurtenis p de fysieke gebeurtenis p^* causaal voldoende is voor p (fysisch determinisme), en als mentale gebeurtenissen niet één op één te relateren zijn aan fysieke gebeurtenissen (de superveniëntie van het mentale), dan ontstaat het probleem dat er twee onafhankelijke oorzaken zijn voor één gevolg, en dat is in strijd met het uitgangspunt van de causale geslotenheid van het fysieke domein.



Figuur 1 – Het uitsluitingsprobleem

Kim voert dit probleem op als een argument tegen het non-reductionistisch fysicalisme, dat stelt dat er geen reductieve relatie ligt tussen m en p^* . Kim strooit vervolgens nog wat zout in de wond met zijn neerwaartse causatieargument (*downward causation argument*), waarin hij stelt dat wil een mentale eigenschap²⁵ M causale werking kunnen uitoefenen op een mentale eigenschap M^* , dat dit via de superveniëntiebasis van M^* moet plaats vinden, en dat is in het uitsluitingsprobleem precies de relatie van m naar p :



Figuur 2 - het neerwaartse causatieargument

Het neerwaartse causatieargument gaat als volgt (Kim 1992:136): M en M^* zijn mentale c.q. emergente c.q. non-reductieve eigenschappen, die zelfstandig causaal vermogen hebben. Stel dat M causale werking uitoefent op M^* . M en M^* zijn emergente eigenschappen, die geïnstantieerd zijn op hun emergentiebasis P en P^* . Als M direct causaal effect heeft op M^* , dan komt de vraag naar voren, hoe ligt dan de relatie tussen M^* en P^* ? Kan M^* wijzigen zonder dat P^* wijzigt? Dit is in tegenspraak met de superveniëntierelatie die tussen M^* en P^* verondersteld wordt: geen verandering van M^* zonder verandering van P^* . Kims conclusie is dan dat M alleen causale werking op M^* kan uitoefenen, indien het causale werking op P^* uitoefent.

Kims positie naar aanleiding van beide problemen is, dat wil je mentale eigenschappen behouden, d.w.z. tot de materiële wereld blijven rekenen, met andere woorden, wil je niet tot dualisme vervallen, dan moet je een reductieve relatie tussen M en P veronderstellen. Dit is de basis van Kims reductionistische positie, die hij vorm probeert te geven met zijn functionele reductie. In het volgende hoofdstuk kom ik daar op terug.

In een analyse van Kims neerwaartse causatieargument moeten we eerst helder hebben wat Kim met de eigenschappen M en P bedoelt. Wat de mentale eigenschap M inhoudt, lijkt duidelijk genoeg. M is bijvoorbeeld een opvatting die iemand kan hebben, en die we de oorzaak van zijn handelen noemen. Zijn handelen kan zijn het openen van de deur, en dat is een fysieke gebeurtenis. Het zijn dit soort formuleringen op grond waarvan gezegd wordt dat mentale eigenschappen causale werking hebben in het fysieke domein. Over wat vervolgens met de 'fysieke' eigenschap P wordt bedoeld, bestaat over het algemeen minder duidelijkheid.

²⁵ Merk op dat in het neerwaartse causatieargument over eigenschappen en niet over gebeurtenissen gesproken wordt. Gebeurtenissen kunnen opgevat worden als toestandsovergangen van eigenschappen. Zie ook Kim, 1973.

Gaat het hier alleen om de eigenschappen van de betrokken fysische deeltjes, of ook om de fysieke eigenschappen van het organisme? En omvat P dan wel of niet de diverse emergente eigenschappen die bij de onderdelen van het organisme zijn aan te treffen? Kim geeft gelukkig op enkele plaatsen toelichting op wat hij onder P verstaat. Kim verstaat onder de fysieke eigenschap van een systeem *'the total microstructural property of the system'*. Kims aanname is dat *'Any physical system can be exhaustively described in terms of the basic particles that constitute it'*. Kim verbindt zelf aan dit uitgangspunt de term 'atomisme'. (Kim, 1997:46). Elders (1998:82) stelt Kim dat de eigenschappen M en P de eigenschappen van hetzelfde systeem zijn, d.w.z. op één en hetzelfde mereologisch niveau. Kim (1997:46) doet voorkomen dat emergentisten deze atomistische positie ook ondersteunen, maar daarin is Kim niet juist. De emergentisten ondersteunen weliswaar het ontologisch fysicalisme, d.w.z. dat er kleinste deeltjes zijn, maar zijn ondersteunen niet het nomologisch fysicalisme, dat alles in deze wereld ondergeschikt is aan microfysische regelmatigheden. En dat is precies wel de positie die Kim inneemt.

Kim stelt dus dat P de totale microstructurele eigenschap van het systeem O is, dat ook de mentale eigenschap M bezit. Eén mogelijkheid om deze uitspraak te interpreteren is dat ook alle clusterings van microfysische deeltjes naar delen van O (en de subdelen daarvan), en de gerelateerde emergente eigenschappen van deze (sub)delen, verdisconteerd zijn in de microfysische beschrijving. Als we Kim zo interpreteren dat de diverse emergente niveaus die tussen een microfysisch niveau en het systeem O te onderkennen zijn, verdisconteerd zijn in de totale microfysische beschrijving van het systeem O, dan neemt Kim een monadisch-fysicalistische positie in. De eigenschappen van alle te onderkennen deelsystemen van O liggen dan ook al vast als dispositionele eigenschappen van de microdeeltjes. Maar ik denk dat we Kim er niet van kunnen betichten een monadisch fysicalist te zijn. In zijn vele geschriften zijn nergens aanwijzingen naar een dergelijke positie te vinden.

De positie van het monadisch fysicalisme is overigens ook overbodig. De fysicalistische positie van het monadisch fysicalisme ontkent dat er gehelen kunnen ontstaan die nieuwe, emergente eigenschappen hebben met hun bijbehorende (nieuwe) causale vermogens. Dat betekent dat de causale vermogens van gehelen die we in deze wereld aantreffen, in de kiem al aanwezig moeten zijn in de microfysische deeltjes. Maar als de kiem van de causale vermogens van gehelen al aanwezig zijn in de microfysische deeltjes, dan wil dat zeggen dat ook de 'concepten' van de gehelen reeds in de microfysische deeltjes betrokken zijn. Dit is de monadische positie van het monadisch fysicalisme. In deze positie kun je inderdaad zeggen dat er geen nieuwe gehelen met emergente eigenschappen ontstaan, want die liggen reeds besloten in de microfysische deeltjes. Maar dit is de zaak omdraaien. Het monadisch fysicalisme is de positie, die ervan uitgaat dat de werkelijkheid een wilde variëteit aan structuren bevat met bijbehorende causaliteiten, en projecteert die wereld op veronderstelde kleinste deeltjes. Het monadisch fysicalisme zegt dan vervolgens: 'Er zijn alleen maar van die kleinste monadische deeltjes'. Deze projectie is niet economisch. Het monadisch fysicalisme omvat de positie van het emergentisme en voegt daar niets aan toe. Het valt als overbodig onder het scheermes van Ockham.

De positie van het monadisch fysicalisme doet sterk denken aan de opvatting van Darwin, die ik in de inleiding reeds besproken heb. Darwins opvatting was dat mentale eigenschappen reeds in de kiem aanwezig zijn in eencelligen. Darwins positie is door Haeckel (1834-1919) geradicaliseerd, in de these van het panpsychologisme, dat stelt dat alle objecten, organische *en* anorganische zijn uitgerust met een vorm van psychologische kwaliteiten (Blitz, 1992:11).

Kim is geen monadisch fysicist. Dat betekent dat we de interpretatie niet kunnen handhaven dat de diverse emergente niveaus die tussen een microfysisch deeltje en het systeem O te onderkennen zijn, verdisconteerd zijn in de totale microfysische beschrijving van het systeem O. Kim zal zeggen 'Er zijn geen emergente niveaus, er zijn alleen maar mereologische niveaus. Er zijn alleen maar clusterings van deeltjes, maar die clusterings hebben geen zelfstandige causale werking.' Kim opteert dus voor de atomistische positie, waarin de microfysische deeltjes uitsluitend microfysische eigenschappen hebben, en geen monadische eigenschappen.

De consequentie van de atomistische positie is de ontkenning dat gehelen ontstaan door bindingen van delen. Een systeem van gebonden delen is alleen adequaat te beschrijven met een concept van het geheel, en waarbinnen de delen ingeperkte vrijheidsgraden hebben. De atomist ontkent dit idee, en ontkent daarmee ook het bestaan van bindingen. Gehelen zijn in de atomistische opvatting alleen mereologische ophopingen van delen.

De enige manier waarop je dan nog over gehelen als samenstel van delen kunt praten, is door een externe oorzaak te veronderstellen waardoor de delen een geheel vormen. Je hebt dan een externe ontwerper-vormgever nodig. Voor artificiële systemen is dit niet problematisch. Als mensen doen we niet anders. Voor natuurlijke systemen veronderstelt dit echter een Godheid. De conclusie is dan dat het fysicistisch atomisme precies op het tegendeel uitkomt van wat ze wil bereiken. Door te stellen dat er alleen fysische deeltjes zijn met fysische eigenschappen, leidt het atomisme uiteindelijk tot deïsme²⁶.

Wat kan er dan vanuit het emergentisme gezegd worden over het neerwaartse causatieargument en het uitsluitingsprobleem? Het emergentisme stelt dat een eigenschap van een geheel, en het causaal vermogen van die eigenschap, samenhangt met de structuur waarin delen dat geheel vormen. Hoe moeten we ons dan voorstellen dat een eigenschap die op structuur is gebaseerd causale werking heeft?

Een geheel is een netwerk van gebonden delen, waarbij die delen in hun netwerk van verbondenheid precies dat geheel vormen. Dit sluit niet uit dat dezelfde delen in meer overlappende netwerken verbonden zijn. Bij hersencellen is dit zeker het geval. Die zijn verbonden in meer overlappende hersennetwerken. Bij mensen is dit ook het geval, waar die in meer organisatieverbanden tegelijk verbonden zijn, en bij elektronen is dit ook het geval, waar die deel uitmaken van meer atomen in een molecule.

Als de eigenschap van een geheel verandert, dan houdt dat in dat de schikking van zijn interne netwerk verandert. Laten we deze initiële verandering even als gegeven beschouwen. In deze initiële verandering (een gebeurtenis) herschikken de delen zich in hun bindingennetwerk. Dat herschikken wordt niet van 'bovenaf' opgelegd. De initiële toestandsverandering is een gevolg van een toestandsverandering in de omgeving, waarbij de delen zich spontaan en vrij, maar binnen het netwerk van hun bindingen, bewegen naar een toestand die voor de delen het meest 'voordelig' dan wel 'verkieslijk' is. Ik gebruik hier antropomorfe begrippen, maar ik zou ook termen kunnen gebruiken als het streven naar een laatste energieniveau.

Het punt is nu dat als een aantal delen ('knopen') in een netwerk van toestand veranderen, en die delen vormen ook 'knopen' in netwerken van andere gehelen, dan zullen de overige delen in het tweede netwerk hun omgeving veranderd zien, en daarop reageren door zichzelf ook te herschikken. Het effect is dan dat ook de structuur van het tweede geheel zich wijzigt, waarmee de betreffende eigenschap van het tweede geheel is gewijzigd. Conclusie: De wijziging van de eigenschap van het eerste geheel heeft een causaal effect gehad op de

²⁶ Dit is precies ook de conclusie van Broad in zijn behandeling van wat hij noemt het 'mechanisme' (1925:81).

eigenschap van het tweede geheel! En ja, de microstructurele configuratie van beide gehelen is gewijzigd omdat alle deeltjes zich spontaan en vrij in een voor hen meest verkieslijke toestand hebben geplaatst. Dus de superveniëntie-relatie blijft geldig: geen verandering van een macro-eigenschap zonder verandering van de microconfiguratie.

Andersom is niet direct noodzakelijk. Een microconfiguratie kan wijzigen (binnen grenzen) zonder dat de macroconfiguratie wijzigt. Neuronen kunnen zich in mijn hersenen al herschikken, zonder dat dit direct een macro/mentale eigenschap doet veranderen. Eerst wanneer een zekere drempel van het aantal herschikte neuronen binnen een netwerk is overschreden, herschikken zich ook de andere delen in het netwerk, en verandert de macro/mentale eigenschap van het geheel. Om deze reden is *sterke* superveniëntie niet houdbaar in relatie tot emergentie.

Wat betekent dit nu voor Kims neerwaartse causatieargument? Vanuit de emergentistische positie is er geen reden om te stellen dat de mentale eigenschap M neerwaartse causatie moet uitoefenen op de microconfiguratie P*. In het zojuist verwoorde emergentistische argument is het perfect verklaarbaar dat een emergente/mentale eigenschap M directe causale invloed uitoefent op de emergente/mentale eigenschap M*. Kims neerwaartsecausatieargument vormt alleen een probleem indien men Kims atomistisch uitgangspunt hanteert. En verder blijkt op grond van het emergentistische argument ook het uitsluitingsprobleem geen probleem meer te zijn. Er is geen sprake van twee oorzaken en één effect. Er is slechts één oorzaak en één effect. Het uitsluitingsprobleem blijkt gebaseerd op atomistische veronderstellingen en het probleem vervluchtigt indien deze veronderstellingen verworpen worden.

In het voorgaande heb ik het nomologisch fysicalisme als onhoudbaar afgewezen. Het ontologisch fysicalisme daar en tegen is wel een mogelijke positie. Het is een praktische werkhypothese dat er kleinste deeltjes zijn, en deze veronderstelling is werkbaar, zolang we geen situaties tegenkomen waarin nog kleinere deeltjes verondersteld moeten worden. Lange tijd was de werkhypothese houdbaar dat de kleinste deeltjes moleculen waren, tot we geconfronteerd werden met elektriciteit waardoor het nodig werd elektronen te veronderstellen, die kleinere deeltjes zijn dan moleculen. Goed, dan laten we het kleinstedeeltjes-niveau een laag zakken, en kunnen weer verder, enzovoort.

Het atomisme legt de basis van de causaliteit op het niveau van de veronderstelde kleinste deeltjes. Maar als dat kleinstedeeltjes-niveau contingent is, en daartoe moet men concluderen als het kleinstedeeltjes-niveau kan 'opschuiven', dan wordt het argument om daar de basis van causaliteit te beleggen problematisch. Dat het kleinstedeeltjes-niveau kan 'opschuiven' houdt namelijk in dat causaliteit niet intrinsiek is aan het kleinstedeeltjes-niveau. Als we causaliteit daar echter willen positioneren, dat moet dat een projectie zijn op dat kleinstedeeltjes-niveau. Deze projectie van monadische eigenschappen op de veronderstelde kleinste deeltjes, dan wel het afzien van die projectie en alleen te praten over de intrinsieke eigenschappen van de veronderstelde kleinste deeltjes levert de problemen op die boven verwoord zijn.

Waarom heeft het nomologisch fysicalisme zo lang stand gehouden? Ik vermoed dat het standhouden van met name het atomisme te maken heeft met de al eerder gemaakte opmerking dat het atomisme verbonden is met het individualisme, dat gestoeld is op de angst voor collectieven, en dat om die reden het idee van neerwaartse causatie niet acceptabel is²⁷.

²⁷ Deze angst voor collectieven zou nog wel eens te maken kunnen hebben met de Koude Oorlog, en de angst voor het Communisme.

Wat ik heb laten zien is dat het neerwaartse causatieprobleem een probleem is dat het atomisme zelf oproept, en dat vervluchtigt wanneer je het idee van het atomisme opgeeft.

4. Het reductie-/non-reductiedebat

4.1. Inleiding

In de vorige twee hoofdstukken is de emergentistische en de physicalistische positie besproken. In beide besprekingen heb ik gepoogd zoveel mogelijk ontologische termen te gebruiken, maar men ontkomt niet aan het gebruik van de nodige semantische termen. Wel heb ik in de twee vorige hoofdstukken gepoogd zoveel mogelijk de termen reductie en non-reductie uit te bannen, teneinde deze termen in het huidige hoofdstuk de aandacht te geven.

In de inleiding en hoofdstuk 2 is alleen onderscheid gemaakt naar het ontologisch en het semantisch domein. In het ontologisch domein wordt gesproken over het bestaan van zaken. Bijvoorbeeld over de vraag of een emergent geheel als een zelfstandig geheel bestaat, of dat het slechts een mereologische ophoping van delen is. In het semantisch domein wordt gesproken over uitspraken, en over de waarheid daarvan. In het physicalisme zijn beide domeinen gekoppeld, wanneer het zegt, dat 'Er bestaat precies dat in de wereld (ontologie), wat een ware en complete fysica (uitspraken) zegt dat er is'.

In het semantisch domein horen ook begrippen thuis als 'afleiden' en 'deduceren'. In een afleiding blijft waarheid geconserveerd. D.w.z. als de premissen van de afleiding waar zijn, en de afleiding is syntactisch correct, dan zijn ook de conclusies waar.

Ik voeg aan de twee genoemde domeinen een derde toe, en dat is het epistemologisch domein. In het epistemologisch domein gaat het om verklaren en voorspellen. Het epistemologisch domein vormt de koppeling tussen het ontologisch en het semantisch domein. In het epistemologisch domein wordt de relatie gelegd tussen zaken in de wereld en uitspraken over de wereld. Epistemologische uitspraken zijn in eerste instantie gefundeerd in onze waarneming. Waarnemingsuitspraken kunnen in het semantisch domein aan elkaar gerelateerd worden, en daaruit kan blijken dat waarnemingsuitspraken tegenstrijdig aan elkaar zijn. Dit wordt teruggekoppeld naar de waarneming met de vraag 'zie je wel wat je denkt dat je ziet?'. Op die manier worden stelsels van consistente uitspraken ontwikkeld, die theorieën genoemd worden. Een *verklaring* is dan de relatie tussen een waargenomen antecedent, een theorie en een waargenomen consequent. Een goede theorie bewijst zijn waarde in de voorspelling, de voorspelling dat gegeven het optreden van een antecedent, dat onder bepaalde condities de consequent zal volgen.

In dit hoofdstuk gaat het vooral over theorieën, gerelateerd aan de vraag of, en onder welke condities een theorie A te reduceren is tot theorie B, dat wil zeggen of theorie A herschreven kan worden in termen van theorie B.

4.2. Het Reductionisme

Het reductionisme is de positie waarin gesteld wordt dat macroniveau theorieën uiteindelijk beschreven kunnen worden in termen van microniveau theorieën. Het reductionisme vindt zijn startpunt in de positie van de logisch positivisten in de eerste helft van de twintigste eeuw. De opzet van het logisch positivisme was om wetenschappelijke concepten (vaak gerelateerd aan niet direct waarneembare fenomenen, bijvoorbeeld elektrische weerstand), terug te brengen tot verzamelingen van door logische relaties verbonden waarnemingsuitspraken.

Eén van de problemen in deze benadering werd gevormd door zogenaamde dispositionele eigenschappen. Dit zijn eigenschappen die een bepaald effect sorteren in bepaalde omstandigheden²⁸. Bijvoorbeeld de eigenschap 'breekbaar'. Wanneer deze eigenschap in termen van waarnemingsuitspraken vertaald moet worden, ontstaat er een formulering als 'als je er op slaat, breekt het'. Deze uitspraak zegt echter alleen iets over de breekbaarheid van een object in het geval er op geslagen wordt. De uitspraak zegt echter niets over breekbaarheid zolang er niet op geslagen wordt. En dat is nu juist de situatie waarin het begrip 'breekbaar' betekenisvol is. Ook zegt de uitspraak 'als je er op slaat, breekt het' niets over de *reden* waarom het object breekbaar is. Problemen als deze leidden uiteindelijk ertoe dat wetenschapsfilosofen opnieuw metafysische concepten in de wetenschappelijke theorieën toestonden (Trout, 1992:389).

De aandacht van het reductionisme verschoof daarna naar theorie-unificatie. De kern van deze stroming is dat wetenschap ernaar moet streven empirisch gevonden wetmatigheden te herformuleren in termen van meer algemene theorieën. In woorden van Feigl: 'Het samenvatten van een maximum aan feiten en regelmatigheden in termen van een minimum aan theoretische concepten en veronderstellingen.' (vgl. Kitcher, 1981:167). Een voorbeeld van een dergelijke herformulering is de reductie van de empirische gaswet van Boyle/Gay Lussac ($P \cdot V/T = \text{konstant}$) tot de kinetische gastheorie.

Voortbouwend op de logicistische traditie formuleert Nagel in 1961 een aantal criteria waaraan voldaan moet worden om van theoriereductie te kunnen spreken.

4.3. Nagels model van theoriereductie

In het hoofdstuk 11 van zijn *Structure of Science* signaleert Nagel (1961:336) twee ontwikkelingen: aan de ene kant constateert hij dat de wetenschappen er tot dusverre niet in geslaagd zijn een omvattende theorie te ontwikkelen die de diverse domeinen van de natuurwetenschappen integreren; en aan de andere kant constateert hij dat er steeds opnieuw, theorieën die tot dusverre autonoom leken, geïntegreerd worden in een meer algemene theorie.

Nagel geeft de waarschuwing dat er gemakkelijk misverstanden kunnen ontstaan indien termen die in de context van de gereduceerde theorie een bepaalde betekenis hebben, in de context van de reducerende theorie geplaatst worden, zonder zich te vergewissen van de voorwaarden waaronder de reductie tot stand is gekomen. Een voorbeeld is het gelijkstellen van temperatuur en kinetische energie van deeltjes, zonder de voorwaarden te kennen waaronder de wet van Boyle/Gay Lussac gereduceerd kan worden naar de kinetische gastheorie. Ik kom op dit voorbeeld zo dadelijk terug.

Vanwege deze zorgen formuleert Nagel (1961:353) twee noodzakelijke voorwaarden waaraan voldaan moet worden om van een geslaagde theoriereductie te kunnen spreken: (a) er moet voldaan worden aan een conditie van verbindbaarheid (*condition of connectability*); dit wil zeggen dat er een correspondentieregel (*rule of correspondence*)

²⁸ In de opvatting van Shoemaker (1980) zijn eigenschappen per definitie dispositioneel: dat is precies wat ze tot eigenschap maakt: dat ze in bepaalde omstandigheden een bepaald effect sorteren.

moet zijn die termen uit de te reduceren theorie verbindt met termen uit de theorie waarnaar gereduceerd wordt; en (b) er moet voldaan worden aan de conditie van afleidbaarheid (*condition of derivability*). Dat wil zeggen dat met behulp van correspondentieregels de te reduceren theorie analytisch *afgeleid* moet kunnen worden uit de theorie waarnaar gereduceerd wordt.

Correspondentieregels zijn expressies die van buiten de te reduceren theorie ingebracht moeten worden teneinde tot een reductieve beschrijving van de te reduceren theorie te kunnen komen. In een correspondentieregel wordt een term uit de te reduceren theorie verbonden met een term uit de theorie waarnaar gereduceerd wordt. Nagel heeft bij het beschrijven van zijn reductiemodel duidelijk de reductie van de empirische gaswet van Boyle/Gay Lussac naar de kinetische gastheorie voor ogen. De gaswet van Boyle/Gay Lussac zegt dat voor een gas in een gesloten ruimte de volgende empirisch gevonden expressie geldt²⁹:

$$(4.3.1) \quad P.V = K.T$$

(Druk (P) maal volume (V) is een constante (K) maal Temperatuur (T).) In de kinetische gastheorie valt voor een geïdealiseerd gas in een gesloten ruimte af te leiden dat:

$$(4.3.2) \quad P.V = (2/3)U$$

Hierin is U de gesommeerde kinetische energie-inhoud van de gasdeeltjes in de gesloten ruimte. Op grond van de vormovereenkomst van de twee expressies, en omdat beiden hetzelfde systeem beschrijven, wordt nu als correspondentieregel genomen:

$$(4.3.3) \quad U = (2/3)K.T$$

Teruggebracht naar de energie-inhoud per gasdeeltje levert dit de expressie op:

$$(4.3.4) \quad \underline{u} = (2/3)k.T$$

met \underline{u} de gemiddelde kinetische energie van een gasdeeltje en k de constante van Boltzmann. Expressie (4.3.4) wordt vaak als voorbeeld genoemd (-onterecht, daarover zo meer-) van een *eigenschapsreductie*³⁰, waarbij temperatuur wordt gereduceerd naar kinetische energie van gasdeeltjes. De correspondentieregel (4.3.3) legt een relatie tussen termen uit beide bij de reductie betrokken theorieën. Omdat de correspondentieregel termen uit beide theorieën bevat, kan de regel noch uit de ene noch uit de andere theorie *alleen* afgeleid worden. De correspondentieregel moet dus van buiten beide theorieën ingebracht worden (Nagel,1961:354).

Om van een geslaagde reductie te kunnen spreken moet er volgens Nagel naast de conditie van verbindbaarheid tevens voldaan worden aan de conditie van afleidbaarheid. De conditie van afleidbaarheid houdt in dat de te reduceren theorie afgeleid kan worden uit de theorie waarin gereduceerd wordt, plus de correspondentieregels. De conditie van verbindbaarheid is volgens Nagel noodzakelijk maar niet voldoende om afleidbaarheid van de te reduceren theorie te garanderen. Een correspondentieregel is allen voldoende indien deze een equivalentie (*biconditional*) inhoudt tussen een term uit de te reduceren theorie en een term

²⁹ De reductie van de gaswet van Boyle/Gay Lussac wordt in vrijwel elk leerboek thermodynamica beschreven. Bijvoorbeeld in Baart/Van Buuren, Warmte en Strooming, Haarlem 1966,p.116.

³⁰ Bijvoorbeeld Kim(1998:90). Kim stelt daar dat Nagels reductiemodel oorspronkelijk over predicatireductie ging, maar dat later in het reductiedebat gebruikelijk werd om over eigenschappen (*properties*) te praten. Beckermann bijvoorbeeld (1992:107) blijft correct het onderscheid tussen predicaten en eigenschappen hanteren.

uit de theorie waarnaar gereduceerd wordt. Correspondentieregels hoeven niet aan die eis te voldoen, en daarom is additioneel de eis van afleidbaarheid nodig (Nagel, 1961:355n5).

In een geslaagde theoriereductie worden dus termen of predicaten uit verschillende theorieën, die een beschrijving geven van dezelfde eigenschap, met elkaar equivalent gesteld. Theoriereductie speelt zich af in het epistemologisch domein. Predicaatreductie, als voorwaarde voor theoriereductie is dus geen eigenschapsreductie. In dit verband nog een opmerking over de correspondentieregel (4.3.4). Die zegt: 'temperatuur is proportioneel met de *gemiddelde* kinetische energie van gasdeeltjes'. In de aanduiding *gemiddelde* is het macro-geheel van de gesloten ruimte impliciet aanwezig. Je kunt niet over een gemiddelde praten zonder de verzameling waar het een gemiddelde van is. De aanduiding '*gemiddelde* kinetische energie van gasdeeltjes' verwijst naar een macro-eigenschap, die betekenis heeft op het niveau van de gesloten ruimte, maar niet op het niveau van individuele gasdeeltjes. Er is daarom in het geval van de Boyle/Gay Lussac gaswetreductie geen sprake van eigenschapsreductie. Het predicat u verwijst naar een macro-eigenschap van gas in een gesloten ruimte, het predicat v verwijst naar de micro-eigenschap van een gasdeeltje. Het gaat om twee verschillende eigenschappen. Zie ook Garfinkel (1981:457) die stelt dat de thermodynamische reductie van temperatuur geen reductie naar het niveau van individuele gasdeeltjes is, niet breed (h)erkend wordt.

Nagel spreekt ook nog van reductie indien termen uit de te reduceren theorie, en die verwijzen naar ontologisch zelfstandige entiteiten, worden opgenomen in de theorie waarnaar gereduceerd wordt. Dat was zelfs zijn motivatie om zijn model voor theoriereductie op te stellen. De eis die Nagel stelt voor het opnemen van nieuwe termen in de basistheorie is dat er correspondentieregels zijn die termen (of expressies daarvan) uit de basistheorie verbinden met de nieuw op te nemen term. Vervolgens moet de gereduceerde theorie dan analytisch afgeleid kunnen worden uit de uitgebreide basistheorie. De gereduceerde theorie wordt daarmee opgenomen in de uitgebreide basistheorie. Er is dan nog steeds sprake van theoriereductie, want we hadden voor de reductie twee theorieën, en na de reductie hebben we één uitgebreide basistheorie. Die uitgebreide basistheorie bevat dan termen die verwijzen naar objecten op meer emergentieniveaus. Nagel noemt deze vorm van reductie 'heterogene reductie' (1961:342).

Over hoe je dan tot de correspondentieregels komt, houdt Nagel zich op de vlakte. Eén mogelijkheid die hij noemt, is dat het een empirische relatie is. Een andere mogelijkheid is dat er een mathematische relatie geconstrueerd wordt die blijkt te werken, d.w.z. dat daarmee aan de conditie van afleidbaarheid voldaan kan worden (1961:354). Nagel gaat voorbij aan de vraag of je na een aantal geslaagde reducties, waardoor termen met betrekking tot objecten van hogere emergentieniveaus in de basistheorie worden opgenomen, een initieel fysische basistheorie nog wel fysisch kunt blijven noemen.

4.4. Het mind-body debat over reductie en non-reductie

In de tweede helft van de twintigste eeuw is in de *philosophy of mind* gedebatteerd over de vraag of, en in hoeverre, mentale eigenschappen te reduceren zijn tot herseneigenschappen: het *mind-body* debat. Nagels model van theoriereductie heeft hierin een belangrijke rol gespeeld.

Ontwikkelingen in de jaren '50 op het gebied van neurofysiologie en op het gebied van computerwetenschappen deden de idee postvatten dat er een één op één relatie zou kunnen bestaan tussen een mentale eigenschap en een neurofysiologische eigenschap. Deze zogenaamde identiteitstheze, waarin een type-identiteit werd verondersteld tussen een mentaal eigenschapstype en een fysiek eigenschapstype, bleek al snel onhoudbaar. De argumenten

tegen de identiteitstheorie waren het multipele realiseerbaarheidsargument van Putnam en het anomalistisch-monisme argument van Davidson (zie Kim,1998:2).

Het multipele realiseerbaarheidsargument van Putnam houdt in dat mentale eigenschappen functionele eigenschappen zijn, die in verschillende fysieke implementaties gerealiseerd kunnen worden. Een mentaal eigenschapstype en het fysiek eigenschapstype kunnen daarom niet identiek zijn. Davidson praat niet over eigenschappen, maar over gebeurtenissen. Zijn anomalistisch-monisme argument houdt in dat een mentale gebeurtenis weliswaar ook een fysieke gebeurtenis is, maar dat er geen 'strikte' wetmatige relatie bestaat tussen een mentale gebeurtenis en een fysieke gebeurtenis, zoals die wel zouden bestaan tussen fysieke gebeurtenissen onderling. Er bestaan daarom volgens Davidson geen verbindende wetten die de mentale wetmatigheden verbinden met fysieke wetmatigheden.

De conclusie die uit beide problemen getrokken werd, is dat er geen wetmatige relatie bestaat tussen een mentale eigenschap en een fysieke eigenschap, en dat er daarom geen reductie mogelijk is van mentale theorie naar fysieke theorie. Wat niettemin bleef, was de ontologisch fysicalistische overtuiging dat het mentale voortkomt uit het materiële. Mentale en fysieke eigenschappen gaan samen, maar zijn zodanig ongelijksoortig dat er geen wetmatig verband tussen gevonden kan worden. Dit leidde tot de positie van het non-reductionistisch fysicalisme.

Deze positie werd onderbouwd door Fodor onder gebruikmaking van Nagels model van theoriereductie. Volgens Fodor is theoriereductie van psychologische theorie naar fysieke theorie niet mogelijk. Fodors argumentatie is de volgende (Fodor,1974:505; zie ook Beckermann,1992:108; Kim,1998:90):

Theorie T_1 beschrijft de relatie:	$S_1x \rightarrow S_2y$
Theorie T_2 beschrijft de relatie:	$P_1x \rightarrow P_2y$

S_1 en S_2 zijn mentale predicaten; T_1 is de theorie over mentale eigenschappen die gereduceerd moet worden; P_1 en P_2 zijn fysieke predicaten; T_2 is fysieke theorie waarnaar gereduceerd wordt. x en y zijn objecten. T_1 is te reduceren tot T_2 , dan en slechts dan:

$$(4.4.1) \quad S_1x \leftrightarrow P_1x \quad \text{en}$$

$$(4.4.2) \quad S_2x \leftrightarrow P_2x$$

De equivalenties (4.4.1) en (4.4.2) worden door Fodor *brugwetten*³¹ genoemd. Brugwetten leggen een relatie tussen een predicaat uit de te reduceren theorie en een predicaat uit de theorie waarnaar gereduceerd wordt. Indien alle predicaattypen uit de te reduceren theorie equivalent gesteld kunnen worden met predicaattypen uit de basistheorie, dan is de te reduceren theorie te herschrijven in termen van de basistheorie.

Het multipele implementatieargument is erop gebaseerd dat in het geval van mentale eigenschappen, brugwetten in de vorm van (4.1) en (4.2) niet bestaan, maar dat brugwetten in dat geval de vorm hebben:

$$(4.4.3) \quad S_1x \leftrightarrow P_1x \vee P_2x \vee \dots \vee P_nx \vee P_{n+1}x$$

Multipele implementatie wil immers zeggen dat $S_1x \leftrightarrow P_1x$, of dat $S_1x \leftrightarrow P_2x$, etc. Volgens Fodor (1974:511) houdt deze disjunctieve reeks in dat in dit geval geen theoriereductie

³¹ Ik heb niet kunnen achterhalen wie de term 'brugwet' heeft geïntroduceerd. Nagel (1961) gebruikt de term niet, maar spreekt van correspondentieregel, Fodor (1974) spreekt van 'brugwet'.

mogelijk is in de opvatting van reductie zoals die door Nagel in zijn model voor theoriereductie is verwoord, omdat er vanwege de disjunctie geen equivalentie bestaat tussen het predicaat van de te reduceren theorie en een predicaat van de fysieke theorie. Het enige wat volgens Fodor brugwetten in de vorm van (4.4.3) kunnen uitdrukken is een gebeurtenis-identiteit, d.w.z. dat een mentale gebeurtenis ook een fysieke gebeurtenis is, maar geen type-identiteit, hetgeen nodig is voor theoriereductie.

Evenals bij de behandeling van het uitsluitingprobleem in het vorige hoofdstuk, is het van belang helder te hebben naar wat voor soort eigenschappen de predicaten P verwijzen. Fodor stelt dat P fysieke predicaten zijn, onderworpen aan fysieke wetmatigheden. Hij stelt ook dat de equivalentie-operatoren transitief moeten zijn, over de neurologische, biologische, chemische lagen heen. Dus Fodor lijkt het bij 'fysieke' eigenschappen duidelijk over de eigenschappen van microfysische deeltjes te hebben. Verder drukken de brugwetten uit dat de predicaten S en P verwijzen naar eigenschappen van dezelfde x , d.w.z. hetzelfde geheel, en dat betekent dat P verwijst naar de microfysische structuur van object x .

Als we dan de fysica opvatten als beperkt tot de beschrijving van de eigenschappen van alleen microfysische deeltjes, en dat is de gangbare opvatting in het *mind-body* debat, dan ontbreken in deze fysica de noties van gebonden gehele n . Een fysieke beschrijving van object x kan dan niet meer zijn dan de beschrijving van de topologische ordening van de microfysisch deeltjes in x . Fodor is dan correct in zijn conclusie, (die hij via een andere redenering bereikt), dat er geen type-identiteit kan bestaan tussen een mentale (emergente) eigenschap S van object x , en de fysieke somverzameling van de deeltjes van x , want S en P verwijzen naar duidelijk verschillende eigenschappen. Er is dan inderdaad wel een gebeurtenis-identiteit mogelijk, waarbij een verandering van eigenschap S van het object x , tegelijk ook verandering inhoudt van de topologische ordening van de deeltjes in x . En dat P dan een disjunctieve verzameling P_i inhoudt, is ook juist omdat de precieze topologische ordening van de microfysische deeltjes in x niet relevant is voor de eigenschap S. Dus vanuit het emergentisme geredeneerd wordt de conclusie van Fodor ondersteund.

Voor Kim is Fodors conclusie niet acceptabel, omdat het geen reductie oplevert en omdat het niet verklaart. De predicaten S en P verwijzen naar verschillende eigenschappen, en de relatie tussen die eigenschappen is contingent. Er is geen sprake van een 'ontologische versimpeling', hetgeen volgens Kim reductie moet inhouden (1998:96). Verder *verklaart* volgens Kim Nagels model niet, omdat het alleen maar stelt *dat* er een wetmatige relatie bestaat tussen de eigenschappen S en P, maar niet waarom, niet hoe deze relatie tot stand komt. Kim zet deze overweging nog wat zwaarder aan door te stellen Fodors toepassing van Nagels model een toepassing is in de traditie van het Deductief-Nomologisch (D-N) verklaringmodel, een verklaringmodel dat in de wetenschapsfilosofie weinig aanhangers meer telt. Volgens Kim kan daarom een afwijzing van reductie, gebaseerd op de Nageliaanse eisen voor een succesvolle reductie niet beschouwd worden als een belangrijke filosofische bijdrage (Kim 1998,97).

4.5. Kims model van functionele reductie

Kim stelt als alternatief voor Nageliaanse reductie wat Kim noemt 'functionele reductie'. In functionele reductie gaat het erom aan te tonen dat twee predicaten die onderdeel vormen van twee verschillende theorieën, verwijzen naar dezelfde eigenschap. Functionele reductie is de these dat bijvoorbeeld een mentale eigenschap kan worden gereduceerd naar zijn fysieke realisatiebasis, door de mentale eigenschap als functionele eigenschap te beschrijven, d.w.z. te beschrijven in termen van zijn causale werking, en dan vervolgens te zoeken naar een fysieke realisatiebasis die precies die causale werking heeft. Omdat we dan twee referenties hebben (de predicaten M en P, beide *rigid designators*, verwijzend naar dezelfde referent -het

causaal effect-), mogen we concluderen dat M en P naar dezelfde eigenschap verwijzen. Predicaat M is in dat geval te reduceren tot predicaat P (Kim,1998:98).

Het idee van functionele eigenschappen vraagt enige uitleg. Een functionele eigenschap is een bijzonder type *tweede-orde* eigenschap. Een tweede-orde eigenschap van een object is de eigenschap dat een eigenschap van dat object (de 'eerste-orde' eigenschap) aan een bepaalde conditie voldoet. Kim noemt als voorbeeld van een tweede-orde eigenschap 'primaire kleur', dat is de eigenschap van een object dat het een kleur heeft (de eerste-orde eigenschap), *en* dat de voorwaarde ('rood' \vee 'blauw' \vee 'groen') geldt (Kim,1998:20). Dat een object O een *functionele* eigenschap heeft, wil zeggen dat het object O een tweede-orde eigenschap heeft, die erin bestaat dat het object O een eerste-orde eigenschap heeft die aan de voorwaarde voldoet dat die eigenschap een bepaalde causale werking heeft. Een functionele eigenschap is dus een eerste-orde eigenschap met een bepaald causaal effect³².

Het functioneel beschrijven van een mentale eigenschap M van een object O, houdt in dat we van M een bepaald causaal effect beschrijven. Indien we dan bij het object O een fysieke eigenschap P kunnen vinden met dezelfde functionele eigenschap, d.w.z. dat de fysieke eigenschap P hetzelfde causaal effect heeft, dan kunnen we volgens Kim concluderen dat eigenschap M token-identiek is aan eigenschap P.

Wat Kim dus zegt is dat in het geval van een functionele reductie, een token-eigenschap die aangeduid wordt met predicaat M, dezelfde token-eigenschap is als die wordt aangeduid met predicaat P. De stappen die Kim onderneemt in zijn model van functionele reductie zijn evenwel niet wezenlijk anders dan die Nagel beoogt in zijn reductiemodel. Ook in Nagels model gaat het erom dat twee predicaten uit twee verschillende theorieën aan elkaar gelijk gesteld kunnen worden. Alleen Nagel houdt zich op de vlakte ten aanzien van de vraag hoe je er toe komt te kunnen stellen dat twee predicaten naar dezelfde eigenschap verwijzen. Op dit punt geeft Kims procedure van functionele reductie een waardevolle aanvulling.

Dat Kim Nagels model afwijst, en vervolgens onder een andere naam een niet wezenlijk ander reductiemodel opnieuw introduceert, is op zijn minst gezegd vreemd. De gronden waarop Kim Nagels model van theoriereductie afwijst (dat Nagels model een toepassing is van het D-N verklaringmodel), zijn ten opzichte van Nagels model onterecht³³. Het is inderdaad zo dat een brugwet in de vorm zoals door Fodor voorgesteld, een correspondentieregel kan vormen in een Nageliaanse reductie. Maar deze brugwetten zijn niet typisch voor wat door Nagel wordt voorgesteld als een correspondentieregel. Correspondentieregels hebben in Nagels opvatting de karakteristieke vorm van een implicatie of een equivalentie tussen twee predicaten, die naar dezelfde eigenschap verwijzen. En dat is precies ook wat Kim met zijn functionele reductie wil bereiken, namelijk aan te tonen dat twee predicaten uit verschillende theorieën, zeg fysieke theorie en psychologische theorie, naar dezelfde eigenschap verwijzen³⁴.

Brengt Kims model van functionele reductie hem een stap verder, gegeven zijn atomistische positie? Kims atomisme komt opnieuw naar voren in het argument waarin Kim

³² Dit is een wat omslachtige constructie om hetzelfde te zeggen wat Shoemaker in zijn identiteitstheorie zegt m.b.t. de identiteit van een eigenschap en diens causaal vermogen.

³³ Kims kritiek is wel van toepassing op Fodors gebruik van Nagels reductiemodel.

³⁴ Kim is onzorgvuldig in het uit elkaar houden van het ontologisch en het semantisch domein. Zijn aanduiding als $M \leftrightarrow P$ is een semantische uitdrukking en betreft de equivalentie van de predicaten M en P; Zijn aanduiding $M=P$ is een uitdrukking van identiteit tussen twee eigenschappen M en P. In deze laatste uitdrukking speelt ons het feit parten dat we over het ontologisch domein alleen maar kunnen praten in uitspraken. Een aanduiding als 'eigenschap M = eigenschap P' is een tegenspraak omdat een eigenschap alleen identiek kan zijn met zichzelf. Een eigenschap M is niet identiek aan een andere eigenschap P. Wel kun je zeggen: 'De eigenschap die met M wordt aangeduid is identiek aan de eigenschap die met P wordt aangeduid.'

Nagels reductiemodel afwijst. Kim stelt dat Nagels reductiemodel de mogelijkheid openlaat om nieuwe termen voor gehelen op te nemen, waardoor de ontologie van de basistheorie wordt uitgebreid, hetgeen voor Kim niet acceptabel is (1998:97).

In een functionele reductie wordt een mentale eigenschap van een object O functioneel beschreven op grond van zijn causaal effect. Vervolgens moet van het object O een fysische eigenschap gevonden worden die precies dat causaal effect heeft. Over hoe je vaststelt dat een bepaalde fysische eigenschap het betreffende causaal effect heeft, laat Kim zich niet uit. En hierin zit ook precies het probleem voor Kim. Binnen zijn atomistische positie is hij niet in staat een fysische beschrijving geven van een eigenschap die de betreffende mentale causale werking heeft. Omdat Kims atomistische positie hem niet toestaat om noties van gebonden gehelen in zijn basistheorie op te nemen, zal Kim in zijn fysische beschrijving van gehelen niet verder komen dan topologische /mereologische clusterings d.w.z. Σ -beschrijvingen en geen Π -beschrijvingen (zie par. 2.3.2, en in het volgende hoofdstuk ga ik hier verder op in), m.a.w. Kim is vanuit zijn atomistische positie wellicht nog in staat een dood lichaam te beschrijven, maar niet een levend lichaam. En dode lichamen hebben geen mentale eigenschappen. Kim zal daarom in zijn atomisme geen fysische beschrijving van het object O kunnen vinden waarvan de causale werking overeenkomt met de causale werking van de mentale eigenschap M van object O. Dit betekent dat Kim vanuit zijn atomistische positie niet in staat is de causale effecten van mentale eigenschappen te beschrijven. Kims vasthouden aan zijn atomistische positie levert hem uiteindelijk een niet-reductieve positie op. Op basis van het atomistisch fysikalisme is het niet mogelijk een relatie te beschrijven tussen fysische en mentale eigenschappen, omdat het atomisme zelfstandige begrippen voor mentale eigenschappen uitsluit. Kim wordt met zijn atomisme ten aanzien van mentale eigenschappen uiteindelijk tot een dualistische positie gedwongen. Precies het tegenovergestelde van wat Kim wil bereiken!

4.6. Het reductie/non-reductiedebat: conclusie

In dit hoofdstuk zijn de mogelijkheden van en de voorwaarden voor theoriereductie besproken. Theoriereductie houdt in dat een theorie A herschreven kan worden in termen van een theorie B. Voor theoriereductie is predicatireductie nodig, en in veel gevallen is het ook nodig om begrippen en wetmatigheden uit de te reduceren theorie over te nemen in de basistheorie. Nagel blijft ook in dit laatste geval van reductie spreken, omdat we van twee theorieën die onafhankelijk leken, teruggaan naar één theorie.

In het mind-body debat spitst de reductievraag zich toe op de vraag of psychologische theorie te reduceren is tot fysische theorie. De non-reductionist Fodor laat zien, dat indien een Nageliaanse correspondentieregel tussen een predicat uit de reduceren theorie, en een predicat uit de het domein van reductiebasis, geen wetmatige relatie is, dat dan geen reductie mogelijk is. En de relatie tussen een mentale eigenschap en een fysische eigenschap is niet strikt wetmatig.

De reductionist Kim wijst Nagels model van theoriereductie af op grond van Fodors gebruik van dat theoriereductiemodel. Fodors toepassing van 'brugwetten' is echter niet de typische toepassing van Nageliaanse correspondentieregels, als bedoeld door Nagel. Kims methode van functionele reductie is dat nu precies *wel*. De tragiek van Kims positie is dat vanwege zijn atomistische stellingname, hij binnen zijn atomisme nooit die fysische configuratie zal vinden die het causale effect heeft van een mentale eigenschap. De consequentie van het atomisme is namelijk dat kleinste deeltjes alleen als vrij botsende deeltjes beschreven kunnen worden (zoals de ideale gasdeeltjes in de gesloten ruimte van het Boyle/Gay Lussac experiment), en dat die deeltjes geen bindingen kunnen aangaan. Om een binding adequaat te kunnen beschrijven is de notie nodig van een gebonden geheel. Dit is een zelfstandig bestaande

notie, naast die van de individuele delen van het geheel. Het atomisme wijst dergelijke noties van gehelen, buiten die van de kleinste deeltjes af.

Het emergentisme erkent wel het bestaan van gehelen als autonome objecten. Dit behoort tot de kernthesen van het emergentisme. In het volgende hoofdstuk laat ik zien dat emergente gehelen adequaat beschreven kunnen worden als configuratie van *gebonden* delen. Het geheel bestaat hierbij precies *in* de bindingen.

Dit hoofdstuk toont dan een Hegeliaanse dialectiek, waarin de reductionist Kim op grond van zijn atomisme feitelijk een niet-reductieve positie inneemt, en uiteindelijk voor het verklaren van mentale eigenschappen een beroep moet doen op een dualistische of een deïstische veronderstelling; en waar het binnen de (non-reductionistische) emergentistische positie mogelijk is om op Nageliaanse wijze tot theoriereductie te komen. Eén mogelijke vorm van een dergelijke theoriereductie is dat de basistheorie alle vormen van gehelen opneemt, en die aan elkaar relateert. Nagels reductiemodel sluit een dergelijke theorie niet uit. Het resulteert echter in een praktisch onwerkbaar omvangrijke theorie. Indien we echter van zo'n theorie abstraheren, en bezien wat al die verschillende typen gehelen gemeenschappelijk hebben, dan zijn dat precies de bindingen. Een theorie van bindingen is dan de achterliggende en unificerende theorie waarnaar de diverse disciplinespecifieke theorieën reduceren, inclusief de fysica!

De conclusie is dan dat niet de fysica de basistheorie is waarnaar de overigen theorieën gereduceerd moeten worden, maar dat gezocht moet worden alle theorieën betreffende deeltjes, delen en gehelen, *met inbegrip van de fysica*, te reduceren tot een unificerende theorie van bindingen.

5. Emergentie en Reductie

In de voorgaande hoofdstukken zijn emergentie, fysicalisme en reductie besproken. De conclusie is dat de fysica, in de vorm zoals we die nu kennen, niet kan fungeren als de reductiebasis voor emergente eigenschappen. De conclusie van het vorige hoofdstuk reikt zelfs verder, de fysica is zelf kandidaat voor reductie.

In dit hoofdstuk wil ik de reduceerbaarheid van emergente eigenschappen opnieuw beschouwen. Het zal duidelijk zijn dat ik onder reduceerbaarheid niet versta reduceerbaarheid naar een fysische reductiebasis. Als ik over de reduceerbaarheid van emergente eigenschappen praat, dan gaat het om de vraag hoe emergente eigenschappen van een geheel gerelateerd kunnen worden aan de structuur van dat geheel, en de eigenschappen van de delen daarin. Door die relatie van een geheel tot zijn structuur en delen te kennen, kan de notie van dat type geheel in de basistheorie worden opgenomen (die daarmee wordt uitgebreid). De notie van reductie houdt dan in dat we tot één uitgebreide basistheorie komen, waarvan mogelijk geabstraheerd kan worden naar één achterliggende unificerende theorie, in het vorige hoofdstuk benoemd als een theorie van bindingen.

5.1. De non-reductieclaim van het emergentisme

In hoofdstuk 3 is de non-reductieclaim van het emergentisme teruggebracht tot de claim dat de emergente eigenschappen van een geheel $R(A,B,C)$ niet af te leiden zijn, en niet te voorspellen zijn uit de eigenschappen van de delen A, B, C 'in isolatie'. Deze laatste toevoeging wil zeggen dat zolang 'R' niet bekend is, de bekende eigenschappen van A, B en C , als zelfstandige objecten, niet toereikend zijn om tot de mogelijkheid van een relatiestructuur R te concluderen. Het emergentisme als verwoord door Broad biedt wel de ruimte om, a posteriori, als een geheel $R(A,B,C)$ is opgetreden, een empirische wet vast te stellen die zegt, dat als delen A,B,C in een structuur R geplaatst worden, dat dan een geheel met een emergente eigenschap E optreedt.

In hoofdstuk 3 heb ik dit uitgelegd, als dat de delen A, B en C binnen een structuur R eigenschappen hebben, die zij buiten structuur R niet tonen. Hier gebeurt overigens niets magisch. Het is te vergelijken met plank die doorbuigt als je erop gaat staan. Als niemand ooit op een plank gestaan heeft, is die eigenschap niet bekend. A posteriori kan die eigenschap bekend worden, als een eigenschap van planken. Als de R -conditionele eigenschappen van A, B en C a posteriori bekend zijn, dan valt *wel* af te leiden dat indien $R(A,B,C)$ dan E . De nonreductieclaim van het emergentisme komt er op neer dat zolang R -conditionele eigenschappen niet bekend zijn, een structuur R niet afgeleid kan worden³⁵. Volgens het emergentisme zijn de R -conditionele eigenschappen van de delen A,B en C alleen a posteriori vast te stellen.

³⁵ Deze uitspraak is natuurlijk tautologisch, want in een afleiding moeten alle elementaire begrippen al in de premissen besloten liggen, en R is zo'n elementair begrip.

Nu valt er af te dingen op de veronderstelling van Broad (1925:61), dat elk type R zijn eigen karakteristieke R-conditionele eigenschappen voor A, B en C met zich meebrengt. Die veronderstelling is op te maken uit zijn uitspraak dat *'the characteristic properties of the whole R(A,B,C) cannot, ..., be deduced from the most complete knowledge of the properties of A, B and C in isolation or in other wholes which are not in the form of R(A,B,C).'*

In de praktijk blijkt dat R-conditionele eigenschappen van delen niet met elke R verschillen, zoals Broad veronderstelt. Het duidelijkst komt dit tot uiting in de quantum mechanica, waarin atomen alleen discrete R-conditionele eigenschappen bezitten. Op grond daarvan kunnen bijvoorbeeld moleculen ontworpen worden, door delen op basis van hun R-conditionele eigenschappen in zodanige structuren R samen te voegen, dat gehelen ontstaan (moleculen) met gewenste chemische eigenschappen³⁶, eigenschappen die voor Broad nog als emergent golden.

De conclusie is dan dat de non-reductieclaim van het emergentisme alleen geldig is, indien het om atomistische reductie gaat. Het atomisme kan immers geen bindingen tussen delen veronderstellen, en omdat een geheel in de bindingen van de delen bestaat, is geen reductie naar atomistische delen mogelijk. De non-reductieclaim van het emergentisme is niet houdbaar als het gaat om de claim dat iedere R-structuur zijn eigen R-conditionele eigenschappen met zich meebrengt, waardoor een structuur R op grond van eigenschappen van delen in andere structuren dan R niet vooraf 'ontworpen' zou kunnen worden. De onhoudbaarheid van deze claim komt tot uiting in het feit dat moleculen ontworpen en gesynthetiseerd kunnen worden, die niet in de vrije natuur voorkomen. Op grond van Broads claim zou dat niet mogelijk zijn, en zouden we pas moleculen kunnen ontwerpen, nadat die empirisch, dus in de vrije natuur, zouden zijn waargenomen.

5.2. Emergente eigenschappen

In paragraaf 3.3 heb ik in relatie tot het uitsluitingsprobleem laten zien hoe emergente eigenschappen direct causale werking op elkaar kunnen uitoefenen. Het basisidee is dat delen 'knopen' vormen in meer netwerken tegelijk, waarbij elk van die netwerken een verschillend geheel vormt³⁷. Indien een deel zich herschikt in het ene netwerk, houdt dat gelijk ook een herschikking in het andere netwerk in. Zijn er voldoende delen in het andere netwerk die zich op die manier herschikken, dan is precies daardoor een eigenschap van het andere geheel gewijzigd.

Cruciaal voor het begrip van emergentie is, dat gehelen bestaan *in* de bindingen waarin delen een geheel vormen. Ik wil twee vormen van structuren onderkennen waarin delen tot een geheel gebonden worden. De eerste vorm zijn artificiële structuren, die van buitenaf opgelegd worden aan de delen. De tweede vorm zijn natuurlijke structuren, waarin de delen zich spontaan binden.

5.2.1. Artificiële structuren

Ook bij artificiële structuren hebben gehelen eigenschappen die de delen niet hebben. 'Nieuwe' eigenschappen treden dus niet alleen op bij complexe structuren, zoals emergentisten veronderstellen. Bijvoorbeeld heeft een mechanische klok als eigenschap dat de grote wijzer

³⁶ Bijvoorbeeld de 'Designer Drugs', d.w.z. geneesmiddelen die zijn ontworpen om in te grijpen op specifieke chemische transacties.

³⁷ Voor hersenstructuren is dit zeker het geval. Fysische deeltjes worden geacht tot één geheel te behoren, maar fysische deeltjes behorend tot verschillende gehelen zijn niettemin in staat tot het uitwisselen van sub-fysische deeltjes (fotonen, gravitonen etc.). Op grond daarvan kan geconcludeerd worden dat fysische deeltjes in meer netwerken verbonden zijn.

met een eenparige snelheid één maal per uur een omwenteling maakt. Dat is een eigenschap die geen van de andere delen van de klok bezitten. De betreffende eigenschap van de grote wijzer volgt uit de structuur waarin de delen interacteren.

Bij artificiële structuren is het aanschouwelijk dat verschillende structuren leiden tot dezelfde eigenschap van het geheel (m.a.w. multiële implementatie). Een voorbeeld hiervan is een communicatienetwerk. Het volgend voorbeeld toont twee netwerken met een verschillende structuur, maar met dezelfde eigenschap van het geheel, namelijk de eigenschap dat elk punt op de rand van het netwerk met elk punt verbonden kan worden. De reden om voor de ene of de andere structuur te kiezen heeft vooral te maken met efficiëntie, niet met effectiviteit. De topologische eigenschappen van beide netwerken zijn verschillend, de functionele eigenschap, de relatie tussen de invoer en de uitvoer van het netwerk³⁸, is identiek.



Figuur 3 – Gehelen met verschillende structuur en dezelfde eigenschap als geheel.

Het typerende bij artificiële structuren is dat deze ontworpen worden binnen de lineaire bereiken van eigenschappen van de delen. In die gevallen is bekend hoe eigenschappen veranderen onder interactie. Buiten die lineaire bereiken worden de uitkomsten onvoorspelbaar. Bekend zijn de beelden van de door wind in beweging gebrachte, dansende hangbrug, die uiteindelijk instort. Bij deze brug zijn mechanische eigenschappen buiten hun lineaire bereiken geraakt. Daardoor is de brug als geheel niet-lineair d.w.z. chaotisch gedrag gaan vertonen: het dansen.

Broad rekende de eigenschappen van artificiële (mechanische) structuren, bijvoorbeeld die van de klok, niet tot emergente eigenschappen, omdat de eigenschappen van de mechanische structuur zijn af te leiden van het gedrag van de delen, gegeven die structuur³⁹. De opnames van de dansende hangbrug laten echter zien dat wanneer eigenschappen van delen buiten hun lineaire bereiken worden gebracht, dat het gedrag van het geheel niet meer voorspelbaar is. En daarmee is dan aan Broads criteria van emergentie voldaan: het geheel heeft nieuwe eigenschappen die niet tot de eigenschappen van de delen behoren, en het gedrag van die nieuwe eigenschappen van het geheel is niet voorspelbaar op grond van de eigenschappen van de delen. Wat ik hiermee wil aantonen is dat er geen sprong bestaat tussen mechanische, voorspelbare eigenschappen, en emergente, niet voorspelbare eigenschappen. De overgang tussen mechanische en voorspelbare eigenschappen en emergente, niet voorspelbare eigenschappen is een continuüm, waarbij mechanische eigenschappen onvoorspelbaar worden wanneer ze zich buiten hun lineaire bereik bevinden, en zogezegd chaotisch worden⁴⁰.

³⁸ Dit komt overeen met Kims definitie van een functionele eigenschap: *'Functional properties [are] properties specified in terms of their roles as causal intermediaries between sensory inputs and behavioral outputs.'* (1998:19).

³⁹ Een eigenschap van een klok is echter niet uit zijn mechanische structuur af te leiden, en dat is het dat het een klok is. Het 'Klok-zijn' is niet een eigenschap die intrinsiek uit de structuur van het apparaat voortkomt, maar die vanuit de omgeving aan het apparaat wordt toegekend, op grond van de rol die het apparaat speelt in betekenisvolle interacties tussen mensen. 'Klok-zijn' is daarom geen emergente eigenschap. Voor dit type eigenschappen als 'klok-zijn' wil ik de aanduiding semiotische eigenschap reserveren (zie par.5.3, verderop.)

⁴⁰ Niet-lineair gedrag is in het verleden altijd buiten de theorievorming gehouden als 'lastig', 'niet beheersbaar'. Eerst de laatste decennia komt niet-lineaire theorie ter verklaring van chaotisch gedrag meer in de belangstelling. Ik denk te kunnen stellen dat evenwel het lineaire bereik van eigenschappen in de natuur de uitzondering vormt, en dat de niet-lineariteit de norm is.

5.2.2. Natuurlijke structuren

De tweede vorm van structuren die ik wil bespreken zijn natuurlijke structuren, d.w.z. structuren die niet artificiële oorsprong zijn. Dit zijn structuren die niet van buitenaf aan een verzameling van delen opgelegd worden. Een manier waarop natuurlijke structuren kunnen ontstaat, is dat delen zich aan elkaar binden. Indien delen zich niet binden, dan is een geheel van die delen niet stabiel, en dat is wel hetgeen we bij natuurlijke gehelen waarnemen. Delen die zich niet binden kunnen alleen op grond van externe dwang bij elkaar gehouden worden. Een voorbeeld van delen die zich niet binden, is het ideaal gas uit de kinetische gastheorie. Dat wordt alleen bij elkaar gehouden door de gesloten ruimte waar de gasdeeltjes zich in bevinden, d.w.z. door een oorzaak van buiten de gasdeeltjes.

Op wat het 'mechanisme' van binding is, zal ik niet verder ingaan. Er zijn verschillende mogelijkheden voor de oorzaak van binding denkbaar, bijvoorbeeld dat de gebonden delen samen een lager energieniveau innemen, of dat de delen subdeeltjes gezamenlijk hebben. Waar het nu bij een binding om gaat is dat een binding een relatie is, die niet is terug te voeren (te reduceren is) op de delen, onafhankelijk van elkaar. Anders gezegd, als we de tweede delen die gebonden zijn a en b noemen, dan wordt de binding uitgedrukt door $(a \wedge b)$, en dit is ongelijk aan, en niet terug te brengen tot⁴¹, twee ongebonden delen, die zich laten uitdrukken als $(a \vee b)$. De binding bestaat precies in de conjunctie. Je kunt niet van een binding spreken, met betrekking tot één deel, onafhankelijk van het andere deel. De binding vormt een toevoeging op de delen.

Hierop kan tegengeworpen worden dat het mogelijk is dat een object een eigenschap heeft, dat het zich kan binden. Maar in die eigenschap ligt het idee van binding, en dus de relatie met een ander object, al besloten. De mogelijkheid van binding zal pas effect hebben als er een ander object van de geschikte soort 'dicht genoeg in de buurt' is. In dit licht kan de kracht die volgens de eerste wet van Newton uitgeoefend wordt door een massa (de kracht die een massa uitoefent is omgekeerd evenredig met het kwadraat van de afstand), geïnterpreteerd worden als de eigenschap van die massa, dat het zich wil binden met andere massa's. In de wet van Newton is dus al impliciet het idee van binding besloten. In de mogelijkheid van binding zijn al impliciet de soorten gehelen aanwezig die kunnen ontstaan wanneer de betreffende delen zich inderdaad binden.

Zoals gezegd is een geheel een structuur van *gebonden* delen. Als bindingen toevoegingen zouden zijn op de delen, dan zou er sprake zijn van een ontologische toevoeging, indien delen zich binden. Met andere woorden, als twee delen zich binden, zou er sprake zijn van toevoeging op de werkelijkheid. Dit idee, dat bij elke nieuwe binding werkelijkheid toegevoegd zou worden, strookt niet met onze idee dat er één werkelijkheid is, die zich weliswaar ontwikkelt, maar waarop niet toegevoegd wordt. Dit houdt in dat verbondenheid al inherent in de delen aanwezig moet zijn. En dit kan alleen, indien verbondenheid ontologisch voorafgaat aan de delen. Als gezegd wordt dat delen de dispositionele eigenschap bezitten dat zij zich kunnen binden, dan is dat een projectie, van de ontologisch primaire verbondenheid naar de mogelijkheid van binding van delen.

Voor het atomistisch fysicalisme leveren bindingen een ernstig probleem op. Het atomistisch fysicalisme zegt dat alles in deze wereld is samengesteld uit kleinste fysische deeltjes, die uitsluitend ondergeschikt zijn aan fysische wetten. Bindingen vormen geen ontologische toevoegingen op de delen, en moeten dus reeds in de delen vervat zijn. Ten

⁴¹ Ten aanzien van dit punt is ook Armstrong (1978) relevant, die stelt dat er wel conjunctieve universalia kunnen bestaan, maar geen disjunctieve universalia. Zie verder ook Gilbert (1992), die het idee van 'meervoudig subject' introduceert, dat bestaat in de binding tussen personen, en dat niet is terug te brengen tot een verzameling van ongebonden personen.

aanzien van bindingen kan het atomistisch fysicalisme dan drie posities innemen: ten eerste kan ontkend worden dat bindingen bestaan. Dit is geen houdbare positie. Een ontkenning van bindingen zou bijvoorbeeld een ontkenning inhouden van de eerst wet van Newton. De tweede mogelijkheid is de monadische positie, dat bindingen die object kunnen aangaan, al vervat liggen in de kleinste deeltjes. Het atomistisch fysicalisme wijst echter het monadisch fysicalisme af, omdat het atomistisch fysicalisme geen andere gehelen als zelfstandige objecten erkent dan de kleinste deeltjes. De derde mogelijkheid is dan dat de bindingen die delen kunnen aangaan, samenstellingen zijn van bindingen tussen kleinste deeltjes. Deze derde mogelijkheid is de atomistische positie. Maar door de mogelijkheid van bindingen tussen kleinste deeltjes te erkennen, erkent het atomisme ook de mogelijkheid van gehelen (in de vorm van $(a \wedge b)$, niet reduceerbaar tot de delen a of b), groter dan de kleinste deeltjes, een mogelijkheid die in tegenspraak is met het uitgangspunt van het atomisme dat er alleen maar kleinste deeltjes zijn. Het atomisme resulteert dus een in een onhoudbare positie: *Of* het ontkent de mogelijkheid van bindingen, hetgeen in tegenspraak is met hetgeen we alom waarnemen, *of* het erkent de mogelijkheid van bindingen, maar komt dan in tegenspraak met het atomistisch uitgangspunt dat er alleen maar kleinste deeltjes bestaan.

Ik heb gesteld dat 'binding' ontologisch primair moet zijn ten opzichte van gehelen, waarin bindingen geactualiseerd worden. Als we ons dan realiseren dat er geen grond is te veronderstellen dat er *kleinste* deeltjes bestaan, dan vormen niet (kleinste) objecten de primaire werkelijkheid, maar 'binding'. Dit is niets minder dan een *ontologische kanteling* van ons denken. Deeltjes en objecten vormen dan verdichtingen in een ontologisch primaire verbondenheid. Wat we ons moeten voorstellen bij een ontologisch primaire verbondenheid, daar heb ik geen helder beeld bij. Wat nog het meest in de richting komt zijn velden. Deeltjes lijken dan nog het meest op plooien of knopen in een ontologisch primair veld. Grotere gehelen zijn dan voor te stellen als clusterings van plooien en knopen. Het beeld dat zich hierbij opdringt zijn *fractals*, waarbij een initieel ongedifferentieerd element door veelvuldige plooing duizelingwekkende structuren oplevert. Dit idee van ontologische kanteling is coherent met het idee van emergentie. Emergentie houdt de ontwikkeling van nieuwe vormen in, zonder dat er op de ontologie wordt toegevoegd. Het idee van ontologische kanteling biedt tevens zicht op theorie-unificatie, waarbij ook de fysica kandidaat is om gereduceerd te worden tot een achterliggende theorie van bindingen.

Als dan gehelen niets meer zijn dan netwerken van gebonden delen, dan rest nog de vraag hoe (emergente) eigenschappen van gehelen zijn te relateren aan de eigenschappen van gebonden delen. Het idee dat Broad opwerpt, is dat eigenschappen van delen in een netwerkstructuur van gebonden delen (een geheel vormend), anders zijn dan van de delen buiten die netwerkstructuur van gebonden delen (hier even afgezien van de mogelijkheid van dispositionele eigenschappen van de delen). Als voorbeeld van een eigenschap die verandert onder interactie met een ander object, heb ik de plank genoemd die doorbuigt doordat je erop gaat staan.

Als delen bindingen kunnen aangaan, dan is het aannemelijk dat die bindingen onder verschillende spanningen kunnen komen te staan, afhankelijk van de omgeving waarin de delen zich bevinden, en afhankelijk of een deel bindingen aangaat met meer andere delen. Indien dit spanningsverloop lineair is, dan is het geheel als een limiet te beschrijven: Deel a heeft een binding met b , b heeft een binding met c , maar de binding van b met c heeft weer invloed op de binding van b met a , en omgekeerd, etc. Dit is uit te schrijven in een reeks, en reeksen kunnen onder voorwaarden convergeren naar een reekselimiet. Baas (1994:520) laat zien dat limieten eigenschappen kunnen bezitten die de termen van de limietreeks niet bezitten. Hij noemt als voorbeeld de inverse limiet in het Scott-model van λ -calculus: Stel X is een verzameling.

$[X \rightarrow X]$ is een verzameling van afbeeldingen van X op X . In de verzamelingentheorie is de vergelijking $X = [X \rightarrow X]$ niet oplosbaar. Scott construeerde een λ -calculusmodel waarin geldt dat $D_{i+1} = [D_i \rightarrow D_i]$, en $D_i \neq [D_i \rightarrow D_i]$. In de limiet van i naar ∞ geldt echter dat $D_\infty = [D_\infty \rightarrow D_\infty]$ wel oplosbaar is. De limiet als geheel heeft hier dus een eigenschap die de termen in de reeks niet bezitten. Maar dat is precies hoe we emergente eigenschappen hebben gedefinieerd. Dus in een limietuitdrukking kan een relatie gelegd worden tussen een netwerkwerk van gebonden delen (de limiettermen), en een geheel (de limiet), waarbij de reekslimiet eigenschappen bezit, die geen eigenschappen van de delen zijn. Dus de eigenschappen van de reekslimiet kunnen als emergent aangeduid worden.

Ingeval van een lineair spanningsverloop van bindingen onder invloed van de omgeving, is in principe een analytische limietbeschrijving mogelijk, die de relatie beschrijft tussen het netwerk van gebonden delen, en het geheel met zijn eigenschappen. Een lineair spanningsverloop in bindingen is echter meer uitzondering dan regel, waarbij de regel een niet-lineair spanningsverloop is. Analytische limietbeschrijvingen zijn dan niet mogelijk, maar wel complexe limieten in de vorm van 'vreemde aantrekkers'⁴². Een vreemde aantrekker is een eigenschap van een geheel, en is de uitdrukking van de toestandsruimte die bindingen tussen delen van het geheel kunnen innemen.

Dit houdt in dat de delen van het geheel *vrij* zijn zich te positioneren in een door hen meest geprefereerde toestand van binding (gegeven de randvoorwaarden van hun directe omgeving), en dat het geheel een uitdrukking is van de toestandsruimte van de mogelijke bindingen binnen het geheel. Het is dus niet zo dat het geheel zijn 'wil' oplegt aan het gedrag van de delen (hetgeen de angst is waardoor het atomisme gevoed wordt), maar dat het geheel een uitdrukking is van de spontane en vrije organisatie van delen, binnen de mogelijkheden van hun beperktheden. En één van die beperktheden is dat de delen slechts bepaalde typen bindingen kunnen aangaan, passend bij het type geheel: Een mens kan andere typen bindingen aangaan dan een atoom.

Een opvallende vorm waarin gehelen hun eigenschappen aanpassen aan hun rol in grotere gehelen is de functiedifferentiatie en specialisatie. We zien telkens opnieuw dat, onder de gunstige voorwaarden, een verzameling initieel ongedifferentieerde gehelen, zich in verloop van tijd differentieert naar verschillende specialisaties, die op elkaar aansluiten, elkaar aanvullen, en die met elkaar samen een nieuw groter geheel vormen. Uit proeven blijkt bijvoorbeeld dat als bacteriën van een zuivere soort in een bepaalde omgeving geplaatst worden, dat na zoveel generaties telkens een bepaalde -stabiele- verdeling van specialisaties van bacteriën ontstaat, die elk een *niche* invullen in het microklimaat van de proefomgeving. Deze evolutionaire ontwikkeling blijkt in de proefomgeving via sterk deterministische paden te verlopen.

Het zich differentiëren en specialiseren van de delen resulteert in binding, omdat het gespecialiseerde deel voor het uitvoeren van bepaalde functies afhankelijk wordt van andere delen in het geheel, bijvoorbeeld in het aanvoeren van voedingsstoffen. Over wat dan de 'ontwerpparameters' zijn achter een netwerk van bindingen, dan wel de 'drijvende kracht' of de 'gradient', kan ik verder weinig zeggen⁴³. Niettemin kunnen we constateren dat er spontaan netwerken van gebonden delen ontstaan.

⁴² Momenteel ontbreken analytische methoden om de vorm van vreemde aantrekkers wiskundig te beschrijven. Simulatie per computer is wel mogelijk en dit levert de bekende chaos-theoretische diagrammen.

⁴³ Salthe (1993) doet op dit punt een interessante poging, door te stellen dat dissipatieve systemen met structuur weliswaar leiden tot een lagere energetische entropie (ten opzichte van de situatie dat dezelfde deeltjes van het geheel ongeordend zijn), maar tot een hogere informatie-entropie, waardoor er per saldo meer entropie geproduceerd wordt, dan in ongeordende toestand. De entropie gradiënt zou dan de kracht opleveren, die duwt in de richting dat alle mogelijk posities en configuraties ook inderdaad worden ingenomen. In dat geval zou een eindsituatie van maximale entropie bereikt zijn.

Het idee dat een geheel beschreven kan worden als de limiet van een netwerk van gebonden delen, kan toegepast worden in Nagels model van theoriereductie. De limietterm is kandidaat om in een correspondentieregel equivalent gesteld te worden met een term uit een reeds beschikbare theorie waarmee eigenschappen van dat type geheel worden beschreven. De rechtvaardiging voor het equivalent stellen van de limietterm en de term uit de beschikbare theorie kan ontleend worden aan een vaststelling dat een causaal effect dat gerelateerd kan worden aan de theoretische term, hetzelfde causaal effect is dat ook gerelateerd kan worden aan de limietterm. Het causaal effect verwijst naar een causaal vermogen van een eigenschap, en de theoretische term en de limietterm zijn dan equivalente aanduidingen voor dezelfde eigenschap. Kims opvatting in het *mind-body* debat, dat Nagels model van theoriereductie achterhaald en verouderd zou zijn, is niet terecht.

Samengevat: Gehelen bestaan *in* de bindingen tussen de delen die samen het geheel vormen. Niet aan delen en gehelen, maar aan *binding* moet primaire ontologische status toegekend worden. Gehelen en delen kunnen opgevat worden als verdichtingen in een ontologisch primaire achtergrondstructuur van binding. (Emergente) eigenschappen bestaan in de netwerkstructuur van gebonden delen. Eigenschappen zijn identiek met hun causale vermogens. Emergente eigenschappen oefenen directe causale werking uit op andere eigenschappen op hetzelfde emergentieniveau doordat delen knopen vormen in overlappende netwerkwerken van bindingen. En het is mogelijk, anders dan emergentisten beweren, om emergente eigenschappen van gehelen te relateren aan het netwerk van gebonden delen dat het geheel constitueert.

5.3. Semiotische eigenschappen

Eén onderwerp wil ik nog kort belichten, en dat is wat ik semiotische eigenschappen noem. Semiotiek is de wetenschap die zich bezighoudt met signalen en de betekenis ervan. Aan een signaal kunnen vier dimensies onderkend worden: ten eerste de referent, dat waar het signaal naar verwijst; ten tweede de betekenis, dat wat het signaal betekent; ten derde de signaaldrager, dat is meestal een fysieke eigenschap van een object, bijvoorbeeld een rood brandend stoplicht; en tenslotte een waarnemer. Want zonder waarnemers zijn er geen signalen, alleen nog signaaldragers⁴⁴.

Semiotische eigenschappen compliceren het emergentistische plaatje. Zijn emergente eigenschappen verbonden met de interne structuur van een geheel, semiotische eigenschappen zijn verbonden met de externe structuur van een geheel, d.w.z. de grotere interactiestructuur waarin het geheel de rol van deel inneemt. Een semiotische eigenschap supervenieert op de (emergente) eigenschappen van de signaaldrager, waarbij het causaal vermogen van de semiotische eigenschap contingent is ten opzichte van het causaal vermogen van de eigenschappen van de signaaldrager. Met andere woorden, hetzelfde signaal, d.w.z. met dezelfde causale werking, kan gedragen worden door verschillende soorten signaaldragers, met elk hun eigen intrinsieke causale vermogens. Het 'klokzijn' van de klok uit de vorige paragraaf is een semiotische eigenschap, die geïmplementeerd kan zijn in diverse fysische realisatievormen: een veeruurwerk, een slingeruurwerk, een digitaal horloge, en bijvoorbeeld een waterklok. Kan de eenparige rotatie van de grote wijzer worden afgeleid uit de delen van de klok en zijn structuur het 'klokzijn' kan daar niet uit worden afgeleid.

Dat een semiotische eigenschap een eigen causaal vermogen heeft, dat onafhankelijk is van de causale vermogens van de signaaldrager, vraagt enige uitleg. Een semiotische eigenschap veronderstelt het bestaan van een intentionele systemen. Een intentioneel systeem

⁴⁴ Voor een aardige inleiding op de semiotiek, zie Daniel Chandler, z.j.

is een systeem dat zich richt op zijn omgeving, en dat specifiek gedrag ontwikkelt, gerelateerd aan hetgeen het systeem in zijn omgeving waarneemt⁴⁵. Als nu een intentioneel systeem een bepaald type gebeurtenis waarneemt, kan dit resulteren in een specifiek gedrag dat in relatie staat tot dat type gebeurtenis. De gebeurtenis houdt voor het intentioneel systeem een signaal in, wanneer de gebeurtenis in verband staat met andere gebeurtenissen die relevant zijn voor het systeem, en het intentioneel systeem de regel volgt, in reactie op de gebeurtenis die een signaal inhoudt, zeker gedrag te vertonen dat verband houdt met die andere gebeurtenissen. Anders gezegd, een signaal is een gebeurtenis A, die verband houdt met gebeurtenis B, en waarbij het gedrag van een intentioneel systeem S in reactie op A, relevant is in relatie tot gebeurtenis B. Je stopt voor rood licht, dat verband houdt met het doorrijden van auto's uit de kruisende straat, en waardoor een botsing met deze auto's voorkomen wordt.

Het causaal vermogen van een semiotische eigenschap bestaat daarin dat een gebeurtenis betreffende de signaaldrager, resulteert in gedrag van een intentioneel systeem, waarbij het effect van dat gedrag geen fysieke causale relatie heeft met de gebeurtenis aangaande de semiotische eigenschap. Een verklaring voor het ontstaan van de relatie tussen het optreden van gebeurtenis A en het gedrag van het intentionele systeem kan evolutionair zijn: Systemen die onvoldoende het bewuste gedrag vertonen, overleven niet.

Dit ontstaan van nieuwe eigenschappen, verbonden met het voorkomen van intentionele systemen, wordt door Baas (1994:512) '*observational emergence*' genoemd. Het is de vraag of je semiotische eigenschappen ook emergente eigenschappen moet noemen. Emergente eigenschappen van een geheel zijn verbonden met diens interne structuur. Semiotische eigenschappen vloeien voort uit de rol die een geheel inneemt in een grotere interactiestructuur. Ik zou de term emergentie willen reserveren voor eigenschappen van de eerste soort, d.w.z. voortkomend uit de interne structuur van een geheel.

Ik ga hier niet verder in op semiotische eigenschappen. Interessante vragen zijn bijvoorbeeld hoe semiotische eigenschappen zich verhouden tot de mogelijkheid dat delen zich de binden (en tot welke nieuwe typen gehelen dat leidt), en hoe semiotische eigenschappen onderling verbonden kunnen raken, en leiden tot betekenissystemen. Deze vragen openen echter complete nieuwe filosofische vraagstellingen, waar ik op dit punt in de scriptie tot een conclusie wil komen.

5.4. Conclusie

Het doel van de scriptie was het onderzoeken van de relatie tussen geest, leven en de materiële wereld. Het doel was het vinden van een derde weg tussen het reductionistisch fysicalisme en het non-reductionisme. Ik denk dat ik daarin geslaagd ben. Ik heb laten zien hoe niveaus van emergentie aan elkaar gerelateerd kunnen worden, een mogelijkheid die door emergentisten steeds ontkent is geworden. Ik heb ook laten zien dat dit niet leidt tot een reductionistisch fysicalisme. Het reductionistisch fysicalisme is een onhoudbare positie. Niet alles wat in deze wereld voorkomt kan beschreven worden in fysische theorie.

Het belangrijkste argument is dat gehelen bestaan in *gebonden* delen. Een binding voegt toe aan de twee delen in isolatie, en is niet terug te brengen tot de ongebonden delen alleen. Een binding is een conjunctieve eigenschap van de gebonden delen *samen*. Een geheel bestaat alleen *in* de binding.

De meest plausibele mogelijkheid om bindingen in deze wereld plaats te geven is te veronderstellen dat bindingen ontologisch primair zijn. Elk vertrekpunt dat deeltjes ontologisch

⁴⁵ Ik volg Pettits (1993b) definitie van intentionaliteit, waar intentionaliteit bestaat in het volgen van regels. Het is daarbij niet noodzakelijk dat het intentioneel systeem zich bewust is dat het regels volgt. Het regelvolgend gedrag kan genetisch geprogrammeerd zijn, of aangeleerd.

primair stelt levert of een incoherente positie op (atomisme), of moet zich toch overgeven aan het primair zijn van de bindingen (monadisme). Het ontologisch primair stellen van bindingen is een ontologische kanteling ten opzichte van de klassieke kijk op de wereld waarin de deeltjes primair zijn⁴⁶. Niet de (kleinste) objecten vormen de primaire werkelijkheid, maar binding. Objecten als gehelen kunnen opgevat worden als een verdichting en actualisatie van bindingen. Deze conclusie sluit aan bij de opvatting dat er geen kleinste deeltjes bestaan, en dat er uiteindelijk alleen maar structuur bestaat, d.w.z. binding.

Dit is een omgekeerde positie als door het atomisme voorgesteld wordt. Het atomisme stelt dat de wereld is opgebouwd uit kleinste delen, die de dragers zijn van de fundamentele eigenschappen van de werkelijkheid. Het atomisme ontkent het bestaan van gehelen als ontologisch zelfstandige entiteiten. Gehelen zijn volgens het atomisme zonder residu terug te brengen tot een samenstel van atomistische delen. Het atomisme is gedwongen te stellen dat er geen conjunctieve eigenschappen van gebonden delen *samen* kunnen bestaan, omdat dit tot de conclusie moet leiden dat er gehelen bestaan die niet zonder residu tot kleinste delen zijn terug te brengen. Maar als het atomisme de mogelijkheid van bindingen moet ontkennen, moet het ook de mogelijkheid van 'spontane' bindingen ontkennen. D.w.z. dat delen nooit 'uit vrije wil' gehelen kunnen vormen. Het atomisme moet dan, om de structuren te verklaren die wij in deze wereld aantreffen, uiteindelijk een beroep moet doen op een externe oorzaak. Het atomisme leidt dus uiteindelijk tot een dualistische of deïstische positie. Precies waar het atomisme probeert afstand van te nemen! Het atomistisch fysicalisme roept zijn eigen ontkenning op.

In de literatuur wordt emergentie vaak gerelateerd aan complexiteit⁴⁷. Eerst in complexe situaties zou emergentie optreden. Dit is niet het geval. Het is niet zo dat er eerst van emergentie sprake is wanneer er een zekere drempel van complexiteit overschreden wordt. Ik heb laten zien dat ook in niet-complexe situaties er al sprake van is van gehelen die nieuwe eigenschappen hebben. Zelfs bij artificiële, mechanistische structuren is het zo dat een geheel eigenschappen heeft die geen eigenschappen van de delen zijn. Artificiële structuren worden zodanig ontworpen dat het geheel de gewenste eigenschappen heeft. Dit is alleen mogelijk indien de eigenschappen van de delen binnen hun lineaire bereik blijven. Buiten deze lineaire bereiken kunnen de eigenschappen van artificiële structuren niet goed meer voorspeld worden. In dat geval voldoen de betreffende eigenschappen aan de definitie van emergentie: De eigenschappen zijn 'nieuw' ten opzichte van de eigenschappen van de delen, en de nieuwe eigenschappen zijn niet te voorspellen op grond van de structuur en de eigenschappen van de delen in isolatie.

Dus tussen mechanistische structuren en emergente structuren, bevindt zich geen onoverbrugbare kloof, zoals traditionele emergentisten als Lloyd Morgan en Alexander beweren. Dat houdt in dat het in principe mogelijk is een beschrijvende relatie te leggen tussen een emergente eigenschap van een geheel, en het netwerk van gebonden delen, dat het geheel vormt. Bij artificiële structuren binnen het lineaire bereik van de eigenschappen van de delen is de relatie tussen een eigenschap van het geheel en de eigenschappen en de structuurrelaties van de delen nog analytisch te beschrijven. Spontane bindingen waarbij bekend is hoe de eigenschappen van de delen zich aanpassen onder de bindingsstructuur kunnen met limietvormen beschreven worden. En bij gehelen bestaande uit veel deeltjes, kunnen soms statistische beschrijvingen gebruikt worden.

⁴⁶ Ik vermoed dat het idee van een ontologische kanteling naar bindingen aansluit bij het idee van 'snarentheorie' (super string theory).

⁴⁷ Bijvoorbeeld Emmeche (1997). Complexiteit is een problematisch begrip. Het is geen intrinsieke eigenschap van een geheel, maar wordt aan een geheel toegekend door een beschrijvende waarnemer.

In al deze gevallen moeten termen voor de structuurrelaties toegevoegd worden om het geheel te kunnen beschrijven. Met deze termen voor de structuurrelaties worden concepten voor bindingen, d.w.z. de niet-reduceerbare conjunctieve eigenschappen van delen *samen*, in het beschrijvingsmodel ingebracht. De meeste natuurlijke structuurvormen evenwel uit zich als 'vreemde aantrekkers', en de huidige beschrijvende wiskunde is nog maar net bezig met theorievorming over niet-lineaire systemen. Als ik zeg dat deze laatste situatie niet fundamenteel anders is dan de eerdere genoemde gevallen, ben ik mij bewust van de nodige *hand-waiving*, maar vooralsnog zijn er geen tekenen dat in deze laatste gevallen fundamentele nieuwe problemen opdoemen.

In de gevallen dat het mogelijk is een beschrijvende relatie te leggen tussen emergente eigenschappen van een geheel en de structuur en de delen van dat geheel, is Nagels model van theoriereductie toepasbaar. Theoriereductie houdt in dit geval niet in dat de te reduceren theorie volledig beschreven kan worden in termen van de basistheorie. Theoriereductie houdt in dit geval wel in dat de twee theorieën geïntegreerd worden, en dat de nieuwe basistheorie termen bevat voor zowel de delen, de bindingen tussen de delen, het geheel, en theoretische relaties tussen deze termen.

De derde weg die in deze scriptie heb beschreven, geeft zicht op een theorie-unificatie, waarbij vanuit diverse theorieën, inclusief de fysica, geabstraheerd wordt naar een achterliggende theorie van binding, ofwel een theorie van emergentie. J.S. Mill stuitte 150 jaar geleden op het probleem dat niet alle eigenschappen van gehelen zich lang analytische weg lieten beschrijven. Mill noemde die eigenschappen heteropatisch. Mill ging ervan uit dat wanneer theoretische ontwikkeling zou voortgaan, meer en meer heteropatische eigenschappen beschreven zouden kunnen worden op basis van delen en structuur. Emergentie is in Mills opvatting geen ontologisch probleem, maar een epistemologisch probleem. Dat is ook mijn conclusie.

- - -

6. Literatuurreferenties

- Armstrong, D.M. (1978) *A Theory of Universals*, Cambridge 1978
- Baas c.s.(1997) On Emergence and Explanation, N.A.Baas, C.Emmeche, *Intellectica* 25(1997)p.p. 67-83
(<http://www.nbi.dk/~emmeche/copubl/97d.nabce/explemer.htm>)
- Beckermann, A.(1992) *Supervenience, Emergence, and Reduction*, in: A.Beckermann, H.Flohr, J.Kim, *Emergence or Reduction*, Berlin 1992
- Bickhard, M.H., Campbell, D.T.(z.j.) *Emergence*. In: P.B.Andersen, N.O.Finnemann, C.Emmeche, P.V.Christiansen (Eds.) *Emergence and Downward Causation*, Aarhus, nog te verschijnen. Voorpublicatie van het paper:
<http://www.lehigh.edu/~mhb0/emergence.html>
- Blitz, D.(1992) *Emergent Evolution*, Dordrecht 1992
- Brakel, J.van; Brink, J.T.van den (1988) *Filosofie van de Wetenschappen*, Muiderberg 1988
- Broad, C.D.(1925) *Mechanism and Emergentism*, uit: *The Mind and its place in Nature*, London 1925
- Chandler, D.(z.j.) *Semiotics for Beginners*, (<http://www.aber.ac.uk/~dgc/sem01.htm>)
- Crane, T.; Mellor, D.H.(1990) *There is no question of Physicalism*, *Mind* 99(1990)
- Davidson, D.(1970) *Mental Events*, in: Davidson, D., *Essays on Actions and Events*, Oxford 1980
- Emmeche, C.(1997) *Aspects of Complexity in Life and Science*, *Philosophica* 59(1997)1, p.p.41-68
- Emmeche c.s.(1997) *Explaining Emergence: Towards a Ontology of Levels*. C.Emmeche, S.Koppe, F.Stjernfelt, *Journal for General Philosophy of Science*, 28(1)1997, p.83-119
- Garfinkel, A.(1981) *Reductionism*, in: Boyd, Gasper, Trout, *The Philosophy of Science*, Cambridge 1991
- Gilbert, M.(1992) *On Social Facts*, Princeton 1992
- Hempel, C.G.; Oppenheim, P.(1948) *Studies in the Logic of Explanation*, *Philosophy of Science* 15(1948)pp.567-579, in: J.C.Pitt, *Theories of Explanation*, Oxford 1988
- Fodor, J.A.(1974) *Special Sciences, or the disunity of science as a working hypothesis*, in: J.Kim, E.Sosa (eds), *Metaphysics*, Blackwell 1999
- Hellman, G.; Thompson, F. (1975) *Physicalisme: Ontology, determination, and Reduction*, in: J.Kim, E.Sosa, *Metaphysics*, An Anthology, Blackwell 1999.
- Horgan, T., 1993 *From Supervenience to Superdupervenience: Meeting the Demands of a Material world*, *Mind* 1993
- Humphreys, P.(1997) *How Properties Emerge*, *Philosophy of Science*, 64(1997) pp.1-17
- Kauffman, S.A.(1993) *The Origins of Order*, New York, 1993
- Kim, J.(1973) *Causation, Nomic subsumption, and the Concept of Event*, *Journal of Philosophy*, 70(1973)8, p.p. 217-236.
- Kim, J.(1990) *Supervenience as a Philosophical concept*, in: J.Kim, E.Sosa (eds), *Metaphysics*, Blackwell 1999
- Kim, J.(1992) *"Downward Causation" in emergentism and Nonreductive Physicalism*, in: Beckermann e.a, *Emergence or Reduction?*, Berlin 1992
- Kim, J.(1993) *Multiple Realisation and the Metaphysics of reduction*, in: J.Kim, *Supervenience and Mind*, Cambridge 1993.
- Kim, J.(1997) *Explanation, Prediction, and Reduction*, *Intellectica*, 25(1997)2 p.p. 45-57
- Kim, J.(1998) *Mind in a Physical World*, Cambridge 1998
- Kim, J.(1999) *Making Sense of Emergence*, *Philosophical Studies*, 95(1999)1-2, p.p. 3-36
- Kitcher, Ph.(1981) *Explanatory Unification*, in: J.C.Pitt, *Theories of Explanation*, NY 1988
- Kripke, S.(1972) *Naming and Necessity*, Cambridge 1972, 1980
- Nagel, E.(1961) *The Structure of Science*, New York 1961 (ref. 4-e druk 1974)
- Oppenheim, P.; Putnam, H.(1958) *Unity of Science as a working Hypothesis*, in: Boyd, Gasper, Trout, *The Philosophy of Science*, Cambridge 1991
- Pettit, Ph.(1993a) *A Definition of Physicalism*, *Analysis*, 53(1993)4 p.p. 213-223

- Pettit,Ph.(1993b) *The Common Mind*, Oxford, 1993
- Pettit,Ph.(1998) *Defining and defending social holism*, Philosophical Explorations, 1(1998)3
- Putnam,H.(1967) *The Nature of Mental States*, in: W.G.Lycan, Mind and Cognition, Cambridge, 1990
- Salthe,S.N.(1993) *Development and Evolution*, Cambridge 1993
- Shoemaker,S.(1980) *Causality and Properties*, in: J.Kim, E.Soza (eds), Metaphysics, Blackwell 1999
- Stephan,A.(1992) *Emergence - A systematic View on its Historical Facts*, in: Beckermann e.a, Emergence or Reduction?, Berlin 1992
- Trout,J.D.(1991) *Reductionism and the Unity of Science*, in: Boyd,Gasper,Trout, The Philosophy of Science, Cambridge 1991
- Yablo,S.(1992) *Mental Causation*, Philosophical Review 101(1992)p.245-280

Samenvatting voor mijn moeder.

De Bijbel zegt dat God de wereld geschapen heeft, en alles wat daar in is. De Bijbel zegt ook dat God de dieren en de mensen de levensadem geeft, en dat de mensen na hun dood terugkeren naar God. De opvatting van wetenschappers en filosofen in de laatste 150 jaar is dat het leven en de menselijke geest in hun geheel voortkomen uit de materie. Dat wil zeggen dat er volgens deze wetenschappers en filosofen geen levensadem, en geen menselijke geest bestaat, onafhankelijk van de materie.

Een heden ten dage algemeen geaccepteerde gedachte is dat alles in deze wereld bestaat uit natuurkundige deeltjes. Die deeltjes worden atomen genoemd. Alles, stenen, planten, dieren en mensen bestaan uit atomen. Als dan alles in deze wereld uit natuurkundige deeltjes bestaat, kun je dan ook alles in deze wereld verklaren met de natuurkunde? Kun je leven en geest verklaren met de natuurkunde? Is dierkunde en menskunde uiteindelijk niets anders dan natuurkunde? Er zijn wetenschappers en filosofen die dat beweren: Als er uiteindelijk alleen maar atomen bestaan, dan moet de natuurkunde uiteindelijk ook alles in deze wereld kunnen verklaren. Mensen die dit zeggen, noem ik natuurkundig fundamentalisten.

Nee, zeggen andere wetenschappers en filosofen, alles in deze wereld bestaat wel uit atomen, maar je kunt niet alles in deze wereld verklaren met de natuurkunde. Dieren en mensen kunnen meer dan atomen kunnen, en de natuurkunde kan dat niet verklaren. Dat plantencellen, en mensen, meer kunnen dan atomen kunnen, komt volgens deze wetenschappers en filosofen, doordat de delen in een geheel samenwerken. De atomen werken samen en vormen moleculen, de moleculen werken samen en vormen levende cellen, de cellen werken samen en vormen lichamen, en de hersencellen werken samen en vormen menselijke geest. Elk van deze gehelen kan meer dan de delen waaruit het gevormd is. De mensen die dit beweren worden *emergentisten* genoemd. Emergentie is het Griekse woord voor 'naar boven komen'. Bij gehelen komen nieuwe mogelijkheden naar boven, die de delen niet kennen. (Zie ook de Zeeuwse wapenspreuk: *Luctor et Emergo* : Ik worstel en kom boven.) De emergentisten beweren, dat wat een geheel kan, niet is af te leiden uit wat de delen kunnen. Daarom, zeggen de emergentisten, kun je met de natuurkunde nooit verklaren dat er mensen zijn, omdat de natuurkunde alleen maar iets kan zeggen over atomen, en niets over mensen.

De samenwerking tussen de delen, of beter, de binding tussen delen is essentieel om van een geheel te kunnen spreken. Die bindingen tussen de delen zijn precies wat het geheel meer is dan de som van zijn delen. Maar, zeggen dan de natuurkundig fundamentalisten, ik kan toch beschrijven hoe atomen zich kunnen binden, dus ik kan beschrijven hoe gehelen uit atomen ontstaan. De denkfout die de natuurkundig fundamentalisten nu maken, is dat om van een binding te kunnen spreken, je al een idee moet hebben van het geheel van gebonden delen. Je kunt niet van een binding praten, als je maar één deel beschouwt. In het idee van binding ligt

het geheel besloten. Maar dat je al een idee moet hebben van het geheel, voordat je over de binding van delen kunt praten, dat druist in tegen hun gedachte dat er alleen maar atomen zijn, en dat gehelen daarvan zijn af te leiden zijn. Als de natuurkundig fundamentalist niet het bestaan van gehelen in hun eigen recht wil erkennen, dan moet hij zeggen dat er geen bindingen bestaan. Maar dat is in tegenspraak met onze waarneming. We zien dat atomen en andere delen zich wel kunnen binden. De gedachtegang van de natuurkundig fundamentalisten dat atomen zich onderling kunnen binden, maar tegelijk ontkennen ze dat er gehelen in hun eigen recht bestaan. Dit spreekt zichzelf tegen. Het natuurkundig fundamentalisme kan daarom niet waar zijn.

Emergentisten aan de andere kant, beweren dat elk soort geheel (bijvoorbeeld een koe, of een paard) zijn eigen soort bindingen bezit, en dat als je bijvoorbeeld alle cellen van een koe zou kennen, dat je daaruit nooit zou kunnen afleiden dat er ook paarden bestaan. Deze bewering is niet in zijn algemeenheid te handhaven. Er worden bijvoorbeeld nu medicijnen gemaakt met een bepaalde werking, door moleculen te ontwerpen die niet in de vrije natuur voorkomen. Het soort bindingen dat delen kunnen aangaan is niet onbeperkt. Als je dan weet welke bindingen er mogelijk zijn, kun je ook voorspellen onder welke omstandigheden welke gehelen mogelijk zijn. In de praktijk zal dit nog best tegenvallen, maar het is niet onmogelijk, zoals de emergentisten beweren. Dus je kunt iets zeggen welke gehelen zullen ontstaan, als je delen op een bepaalde manier samenbrengt. De voorwaarde is dat je iets over de bindingsmogelijkheden van de delen weet. Maar als je de bindingsmogelijkheden kent, weet je ook al wat voor soort gehelen mogelijk zijn. Bindingsmogelijkheden van delen en soort gehelen zijn elkaars spiegelbeeld.

Het idee van bindingen levert een probleem op voor het idee van het bestaan van kleinste deeltjes. Een binding houdt in dat de delen iets *samen delen*. Maar kleinste deeltjes kunnen niets samen delen, omdat ze ondeelbaar zijn. Kleinste deeltjes kunnen zich dus niet binden. Maar als kleinste deeltjes zich niet kunnen binden, kunnen er ook geen grotere gehelen ontstaan. Maar die bestaan er wel, dus kleinste deeltjes moeten zich kunnen binden, en dat betekent dat kleinste deeltjes ook weer samengesteld moeten zijn. Als je dit doortrekt, moet je concluderen dat er geen kleinste deeltjes bestaan. Het enige dat dan overblijft, *voor* het eerste deeltje, is de mogelijkheid van binding.

Dus *voor* alle gehelen, delen en deeltjes bestaat de mogelijkheid van binding. Dit is een kanteling van 2500 jaar denken over materie, waarin altijd gedacht is dat ondeelbare deeltjes (het Grieks atomos = on-deelbaar) de basis van de werkelijkheid vormden. Mijn conclusie is dat de mogelijkheid van binding de grondslag van de werkelijkheid vormt, en niet de atomen. Hoe moet je je dat nou voorstellen? Dat *voorafgaand aan* elk deel en geheel al de mogelijkheid van binding bestaat?

Ik wil dat illustreren aan de hand van een *fractal*. Een fractal is een wiskundige constructie die er als volgt uit kan zien:

In stap 0 is het alleen een rechte lijn:

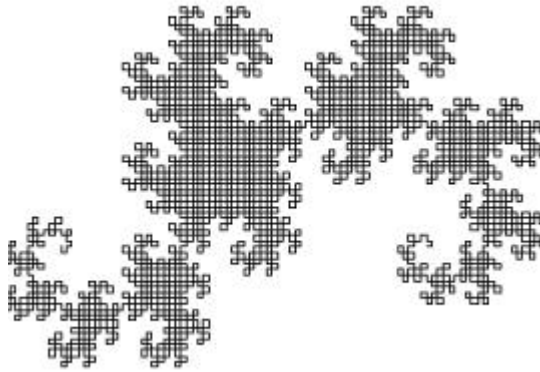


In stap 1 wordt de lijn als volgt geplooid:



In stap 2 wordt elk lijnstuk van stap 1 op dezelfde manier geplooid.

Na vele stappen ontstaat er dan het volgende plaatje:



We zien hier dat een aanvankelijk simpele, rechte lijn door veelvuldig plooien, een complexe structuur oplevert met deeltjes, delen en gehelen, die niettemin nog steeds één lijn vormen. Zo kunnen we ons ook het ontstaan van deeltjes en delen in de werkelijkheid voorstellen.

"Mooi, nu heb je dat allemaal uitgedacht, maar wat kun je er mee?"

De reden om het onderwerp van deze scriptie te kiezen, was mijn constatering dat bedrijfskundige theorieën in de praktijk zo slecht werken. Ondanks alle mooie plannen pakt de praktijk steeds anders uit dan gepland. Hoe komt dat? Een reden is dat de bedrijfskundige theorieën te simpel en te rechtlijnig zijn. Veel bedrijfskundige theorieën werken vanuit de delen, zonder het geheel in ogenschouw te nemen. De conclusies van de scriptie geven onderbouwing aan de stelling dat dit niet voldoende is om een organisatie als geheel te besturen.

Maar ook als je het geheel in ogenschouw zou *willen* nemen, dan is er het probleem dat een individu een organisatie als geheel niet *kan* bevatten, omdat een deel het geheel waar het onderdeel van is, niet kan waarnemen. De conclusie is dan dat de mogelijkheid om een (grote) organisatie te besturen beperkt is. Grote organisaties worden niet bestuurd. Grote organisaties leven van zichzelf.

- - -

Samenvatting voor mijn KPN collega's

De aanleiding voor het onderwerp van de scriptie was mijn constatering dat de mooie bestuurskundige theorieën die ik mijn bedrijfskundestudie heb geleerd, in een complexe organisatie zoals bijvoorbeeld KPN vrijwel onbruikbaar zijn. Reorganisatie volgt op reorganisatie, en het bedrijf verandert, maar je kunt je afvragen of die verandering wel dankzij die reorganisaties gebeurt. De basisgedachte van de traditionele bestuurskundige theorieën is dat je door gerichte beïnvloeding van de delen het geheel kunt veranderen. Maar bestuurlijke ingrepen blijken soms effecten te hebben die tegengesteld zijn aan de doelstelling van de ingreep. En zelden blijkt een reorganisatie volledig het gewenste effect te hebben, reden dat er binnen twee jaar weer een nieuwe reorganisatie volgt. In de tussentijd zorgen de medewerkers op de bedrijfsvloer voor de continuïteit van de primaire processen, die minder beïnvloed worden door de bestuurlijke stormen dan de reorganisatoren zouden wensen. Het lijkt erop dat een complexe organisatie als KPN zijn eigen weg kiest, eerder ondanks dan dankzij reorganisaties. Hoe is dit te verklaren?

Om de reikwijdte van de afstudeerscriptie beperkt te houden, heb ik deze vraag teruggebracht tot zijn metafysische kern: de verhouding tussen gehelen en delen. Metafysica wil zeggen opvattingen over conceptuele zaken als 'eigenschap', 'oorzaak', 'bestaan', 'wetmatigheid', 'identiteit', 'relatie', en ook 'geheel'. De centrale vraag is, wat zijn gehelen, en hoe komt het, dat in voorkomende gevallen gehelen meer zijn dan de som van delen? En als je de samenstelling van een geheel naar zijn delen kent, kun je dan daaruit de eigenschappen van het geheel afleiden? Als dat laatste het geval is, dan kun je het geheel gericht beïnvloeden door de delen te sturen. Deze vraagstelling laat zich transponeren naar de vraag of complexe organisaties te besturen zijn.

De relatie tussen delen en gehelen is onderwerp van het filosofisch debat tussen fysicisten en emergentisten. Beide partijen zijn het eens over de stelling dat alles wat in deze wereld bestaat, uiteindelijk uit fysische deeltjes bestaat. Beide partijen verschillen daarin dat de fysicisten beweren dat dan ook alles in deze wereld, zoals het leven, de menselijke geest en complexe organisaties, uiteindelijk verklaard moet kunnen worden op grond van fysische wetten, terwijl de emergentisten stellen dat je niet alles met de fysica kunt verklaren. Gehelen, beweren de emergentisten, hebben nieuwe eigenschappen, d.w.z. eigenschappen die de delen niet kennen. Die nieuwe eigenschappen brengen eigen wetmatigheden met zich mee, die niet zijn terug te brengen tot fysische wetmatigheden. (Emergentie is afgeleid van het Griekse woord '*emergere*', wat 'boven komend, aan het licht komend' betekent. In gehelen komen nieuwe eigenschappen aan het licht.)

De meerwaarde van het geheel ten opzichte van de som van zijn delen, bestaat volgens de emergentisten in de relatiestructuur, waarin de delen het geheel vormen. In de scriptie ontwikkel ik de argumentatie dat voor natuurlijke gehelen die relatiestructuur uit *bindingen*

moet bestaan. De emergentisten stellen dat elk type geheel zijn eigen type relatiestructuur kent, en dat je, als je het gedrag van delen in het ene type relatie structuur kent, nog niets kunt zeggen over het gedrag van die delen in een ander type relatiestructuur, dus in een ander geheel.

De emergentisten blinken niet uit in helderheid waarom een relatiestructuur van bindingen niet is af te leiden uit de eigenschappen van de delen. Ze zeggen alleen maar dat het niet kan. In de scriptie werk ik een argumentatie uit om deze uitspraak van de emergentisten te onderbouwen. Een binding kan alleen maar bestaan, als er (minimaal) twee delen zijn. Eigenschappen die aan de binding ontspruiten, zijn conjunctieve eigenschappen van de twee delen *samen*, m.a.w. van het geheel. Conjunctief wil zeggen een *en*-relatie. Conjunctieve eigenschappen hebben de bijzonderheid dat ze niet zijn terug te brengen tot eigenschappen van delen onafhankelijk van elkaar. Conjunctieve eigenschappen bestaan daarom additioneel ten opzichte van de eigenschappen van de delen op zichzelf. Dergelijke conjunctieve eigenschappen komen overeen met emergente eigenschappen. Als dan wetmatigheden niets anders zijn dan regelmatigheden ten aanzien van het gedrag van eigenschappen, en als emergente eigenschappen niet terug te brengen zijn tot de eigenschappen van de delen, dan zijn ook de wetmatigheden aangaande die emergente eigenschappen niet terug te brengen tot de wetmatigheden aangaande de delen. Emergente wetmatigheden zijn dus uiteindelijk niet terug te brengen tot fysische wetmatigheden.

Maar, zeggen dan de fysicisten, als ik weet hoe delen zich kunnen binden, dan kan ik toch beschrijven hoe gehelen ontstaan? Het probleem voor de fysicisten hierbij is dat in de mogelijkheid dat delen bindingen kunnen aangaan, de voorwaarde besloten ligt dat er een geheel van gebonden delen kan bestaan. En het *samen* gebonden zijn van de delen kun je niet terugbrengen naar eigenschappen delen onafhankelijk van elkaar, omdat je dan niet meer over gebondenheid kunt praten. Dus als fysicisten over de mogelijkheid van binding van delen praten, dan veronderstellen zij impliciet dat er gehelen bestaan, die niet terug te brengen zijn tot onafhankelijke delen. Als fysicisten de mogelijkheid van bindingen willen erkennen, dan zullen zij in de fysica ook concepten voor moleculen, cellen, organismen, mensen moeten opnemen. Psychologie wordt dan ook een deel van de fysica. Maar dat is strijdig met de these van de fysicisten, dat psychologie uiteindelijk terug te brengen is tot de fysica van de kleinste deeltjes. Als omgekeerd de mogelijkheid van binding wordt ontkend, dan wordt het fysicalisme evenzeer een onhoudbare positie, omdat zonder binding er geen natuurlijke gehelen kunnen ontstaan. De conclusie is dan dat het fysicalisme niet houdbaar is. Niet alles in deze wereld is te verklaren vanuit de fysica.

Dat er gehelen bestaan met eigenschappen die niet terug te brengen zijn tot de eigenschappen van onafhankelijke delen, wil niet zeggen dat er niets gezegd kan worden over de relatie tussen gebonden delen en het geheel, zoals de emergentisten beweren. Het blijkt mogelijk om verklarende relaties te leggen tussen het concept van een geheel, en de delen in hun interactiestructuur. Dit is geen reductieve relatie, omdat het geheel in deze verklarende beschrijving in zijn eigen recht blijft bestaan. Verder is er af te dingen op de claim van de emergentisten dat je op grond van de bindingen die delen aangaan in de ene structuur, niets kunt zeggen over de bindingen die die delen in andere structuren kunnen aangaan. Het blijkt dat de bindingsvariëteit van delen niet onbeperkt is. Als je het repertoire van bindingseigenschappen van delen kent, kun je in principe nieuwe typen gehelen ontwerpen. Dit gebeurt bijvoorbeeld al met betrekking tot het ontwerpen van nieuwe medicijnen. Volgens de emergentisten zou dit niet mogelijk zijn.

Het idee dat gehelen ontstaan, doordat delen zich binden, levert een probleem op voor het bestaan van kleinste deeltjes. In de mogelijkheid dat delen zich kunnen binden, bestaat reeds de mogelijkheid van het geheel. De mogelijkheid van binding gaat vooraf aan elk geheel. Bij kleinste deeltjes kan dat niet het geval zijn, omdat kleinste deeltjes niet meer samengesteld zijn. Maar als kleinste deeltjes niet meer samengesteld zijn, kunnen ze zich ook niet binden, omdat er bij binding sprake moet zijn van een gedeelte iets, en kleinste deeltjes hebben niets om te delen. Dus als er kleinste deeltjes zijn, kunnen er geen gehelen ontstaan. Er bestaan echter gehelen, dus er kunnen geen kleinste deeltjes bestaan. *Voor* alle delen en gehelen moet dan de mogelijkheid van binding bestaan. Dat betekent dat de mogelijkheid van binding moet bestaan *voor* alle materie. Dat laat zich lastig voorstellen, maar het is misschien het best voor te stellen in de hoedanigheid van een veld. Het veld bestaat *voor* de eerste deeltjes. Deeltjes zijn dan te begrijpen als plooiingen van het veld, en gehelen als clusters van die plooiingen, in dat zelfde veld. Dit model biedt een oplossing voor een probleem van de emergentistische positie, dat gehelen nieuw zijn, en een bestaan in hun eigen recht kennen, zonder dat we willen zeggen dat er met emergente eigenschappen nieuwe werkelijkheid ontstaat. Er is één werkelijkheid, en daar wordt niet aan toegevoegd als er nieuwe eigenschappen ontstaan. Vanuit het idee van het geplooid veld, is een nieuw type geheel niets meer dan een extra plooï in het bestaande veld. Vanuit dit perspectief houdt een emergente eigenschap geen nieuwe werkelijkheid in, hetgeen in overeenstemming is met onze intuïtie.

Wat betekenen deze conclusies nu voor het beschouwen van een complexe organisatie als KPN? Ten eerste dat een organisatie bestaat in bindingen tussen mensen. Als je die bindingen aantast, dan tast je vitaliteit van het geheel aan, zoals het geval was bij de reorganisatie van 'KPN Nieuw', twee jaar geleden. Ten tweede dat atomistische besturing faalt. Bij atomistische besturing wordt de relevantie van bindingen genegeerd. Dit levert mechanistische gehelen op, die alleen door randvoorwaarden van buiten het geheel in stand gehouden kunnen worden. Dit soort gehelen zijn weinig flexibel te opzichte van een veranderende omgeving. Bij het besturen van een complexe organisatie moet je het geheel expliciet in ogenschouw nemen. En hierbij doet zich het probleem voor dat een deel niet in staat is het geheel waar het deel van uitmaakt, als geheel waar te nemen. Om een geheel als geheel waar te nemen, heb je als waarnemer een interne structuur nodig van de soort van het geheel dat je wilt waarnemen. Geen mens kan dus een organisatie als geheel waarnemen. En geen mens is dus in staat om een organisatie als geheel te besturen.

- - -