Pijlers van ons wereldbeeld 1: Relativiteitstheorie

najaar 2009

Prof. Dr. R.H.P. Kleiss
IMAPP
Radboud Universiteit, Nijmegen

Contents

1 Newtoniaanse mechanica **4**
- 1.1 Ruimte en coördinaten . . . 4
- 1.2 Tijd . . . 6
- 1.3 Mechanische grootheden . . . 6
- 1.4 Behoudswetten . . . 9
- 1.5 De wetten van Newton . . . 11
- 1.6 De Galilei-transformaties . . . 14
- 1.7 Behoud van impuls en energie . . . 16

2 De lichtsnelheid als natuurconstante **20**
- 2.1 Licht als golfbeweging in de æther . . . 20
- 2.2 Lichtgolven als electromagnetisch verschijnsel . . . 21
- 2.3 Ætherwind? . . . 23
- 2.4 De Lorentz-Fitzgerald contractie . . . 24
- 2.5 Einstein's postulaat . . . 25
- 2.6 De lichtklok: tijdsdilatatie . . . 26
- 2.7 De lichtsnelstheid . . . 28
- 2.8 Experimentele waarneming van tijdsdilatatie . . . 29
- 2.9 Wederom de Lorentzcontractie . . . 30
- 2.10 "Meanwhile, in a distant galaxy..." . . . 30

3 De tijdruimte van de relativiteitstheorie **32**
- 3.1 De Minkowski-ruimte . . . 32
- 3.2 De Lorentz-transformaties . . . 33
- 3.3 De Lorentz-transformatie als pseudo-rotatie . . . 36
- 3.4 Samenstellen van snelheden . . . 38
- 3.5 De 'echte afstand' in de Minkowski-ruimte . . . 40
- 3.6 Symmetrieën van ruimte, symmetrieën van de natuur . . . 41
- 3.7 Causaliteit en de lichtkegel . . . 43
- 3.8 De metriek van de Minkowski-ruimte . . . 45

4 Relativistische mechanica **47**
- 4.1 De eis van covariantie . . . 47
- 4.2 De eigentijd van een puntdeeltje . . . 47
- 4.3 De vier-snelheid . . . 48
- 4.4 De vier-impuls . . . 49

4.5 Relatie met de Newtoniaanse grootheden: $E = mc^2$ 50
4.6 Feiten en misvattingen rond $E = mc^2$ 51
4.7 Voorwerpen met lichtsnelheid 53
4.8 Het gewicht van nat licht 54
4.9 Antimaterie 55

5 Schets van de algemene relativiteitstheorie 59
5.1 Algemene coördinatensystemen, algemene metrieken 59
5.2 Het probleem van de zwaartekracht 60
5.3 De wiskundige artillerie van Einstein's theorie 61
5.4 Zwarte gaten 64

1 Newtoniaanse mechanica

Om de positie van Einstein's relativiteitstheorie in de fysica van de twintigste eeuw te kunnen appreciëren moeten wij haar stellen tegenover de 'oude' mechanica, die van Galilei en Newton. Daarom zal eerst deze kort behandeld worden.

1.1 Ruimte en coördinaten

Naar het zich laat aanzien leven wij in een universum dat een continuum van drie ruimtelijke dimensies vormt[1]. Dit betekent dat de positie van ieder punt in de ruimte kan worden vastgelegd met behulp van

1. een *oorsprong*, dat wil zeggen de keuze van een centraal punt van waaruit de posities van alle andere worden vastgelegd;

2. een *coördinatensysteem*, dat wil zeggen een voorschrift over hoe vanuit de oorsprong een punt bereikt kan worden als de coördinaten van dat punt gegeven zijn;

3. een groep van *drie* coordinaten, dat wil zeggen de informatie die in het coördinatensysteem 'ingevoerd' moet worden om het bedoelde punt te bereiken. De oorsprong heeft altijd de coordinaten $(0, 0, 0)$.

Een eenvoudig voorbeeld van een coördinatensysteem is het Cartesiaanse systeem, waarin vanuit de oorsprong drie onderling loodrechte assen worden gekozen; de coördinaten van een punt ten opzichte van de oorsprong zijn dan de drie afstanden langs de assen. De *afstand* tussen twee punten A en B wordt dan als volgt gevonden: laat de coördinaten van de twee punten (in hetzelfde coördinatensysteem!) gegeven worden door

$$\text{punt A:} \quad \vec{x}_A = (x_A, y_A, z_A) \quad , \quad \text{punt B:} \quad \vec{x}_B = (x_B, y_B, z_B) \quad . \tag{1}$$

De afstand $r(AB)$ tussen A en B wordt dan gevonden uit de (drie-dimensionele) regel van Pythagoras:

$$r(AB)^2 = (x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2 \quad . \tag{2}$$

[1] Hoewel in het dagelijks leven niet betwijfeld, is dit in feite een fysisch postulaat! In geavanceerde theorieën die proberen de wereld te beschrijven op het niveau van het allerkleinste détail (10^{-33} centimeter) wordt rekening gehouden met het bestaan van vijf of zelfs tien ruimtelijke dimensies. Ook wordt door sommigen betwijfeld of de ruimte er op dit niveau nog steeds uitziet als een continuum, zonder 'pixel' structuur.

Deze regel geeft nooit een negatief resultaat: we zien ook onmiddellijk dat indien $r(AB) = 0$ de punten A en B dezelfe coördinaten moeten hebben, en A en B dus in feite identiek zijn.

Niet ieder coördinatensysteem is Cartesisch: een aansprekend twee-dimensioneel voorbeeld is het coordinatensysteem van lengte- en breedtegraden op de aardbol. De oorsprong is dan het snijpunt van de meridiaan van Greenwich met de evenaar. De afstand tussen twee punten (gedefinieerd als de lengte van de korste weg over het aardoppervlak tussen de twee punten) wordt dan gegeven door een regel die veel ingewikkelder is dan die van Pythagoras. Laat van punt A de lengte- en breedtegraad l_A respectivelijk b_A zijn, en voor punt B hebben we de corresponderende l_B en b_B. De afstand $r(AB)$ wordt dan gegeven door

$$r(AB) = \frac{C}{360}\,\Omega \quad , \tag{3}$$

waarbij $C \approx 40000$ km de omtrek van de aarde is en de hoek Ω (in graden, liggend tussen 0° en 180°) wordt gegeven door

$$\cos(\Omega) = \sin(b_A)\sin(b_B) + \cos(b_A)\cos(b_B)\cos(l_A - l_B) \quad . \tag{4}$$

Dit voorschrift ziet er inderdaad veel ingewikkelder uit dan het Cartesiaanse[2], zelfs als we buiten beschouwing laten dat in werkelijkheid de aarde niet zuiver bolvormig is.

De punten in de drie-dimensionele ruimte zijn *objectief*, dat wil zeggen ieder punt, samen met wat zich daar afspeelt, heeft een eigen identiteit onafhankelijk van ieder coordinatensysteem. Dit betekent dat in principe *ieder* coördinatensysteem gebruikt zou kunnen worden om de fysica te beschrijven, immers het coördinatensysteem is een constructie van, en voor, de waarnemer. Aan de andere kant is een vergelijking van resultaten tussen verschillende waarnemers alleen op een zinvolle manier mogelijk als er een methode bestaat om de coördinaten van ieder gegeven punt in een bepaald coördinatensysteem om te rekenen in de coördinaten van *hetzelfde* punt in een ander coördinatensysteem. Zo'n 'omrekenings-

[2]Merk echter op, dat voor punten A en B dichtbij de oorsprong, waarvoor l en b klein zijn, de relatie

$$\Omega^2 = (b_B - b_A)^2 + (l_B - l_A)^2$$

in zeer goede benadering geldt: in die buurt is het coördinatensysteem dus niet zo heel erg verschillend van een Cartesiaans systeem.

voorschrift' wordt een *coördinatentransformatie* genoemd. Tussen ieder tweetal acceptabele coördinatensystemen moet een coördinatentransformatie bestaan. Bovenstaande laat natuurlijk onverlet dat in een gegeven situatie het ene coördinatensysteem *gemakkelijker* is dan het andere. In veel toepassingen heeft een Cartesiaans systeem daarom de voorkeur.

1.2 Tijd

Natuurkunde houdt zich bezig met relaties tussen verschijnselen in ruimte en tijd. De combinatie van een gegeven punt in de ruimte en een gegeven moment in de tijd wordt een *gebeurtenis* genoemd. Men denke hierbij aan een bepaald objectief verschijnsel dat zich daar afspeelt: bijvoorbeeld het op elkaar botsen van twee deeltjes, het afgaan van een flitslampje, etcetera[3]. De localisatie van een gebeurtenis vergt dus een aanvullende keuze van een oorsprong in de tijdrekening (het 'uur nul') en het aangeven van tijdcoördinaten van punten. Laat ons aannemen dat tijd gemeten wordt met een *gelijkmatig lopende* klok. Als gebeurtenis A op tijdstip t_A plaatsvindt, en gebeurtenis B op tijdstip t_B, dan wordt de afstand in de tijd[4] tussen A en B gegeven door $t(AB)$, waarbij

$$t(AB)^2 = (t_B - t_A)^2 \quad . \tag{5}$$

In het geval van een ongelijkmatig lopende klok is de tijdsafstand een ingewikkelder functie van de tijdstippen. Ook zo'n klok kan in principe gebruikt worden, mits ook voor deze tijds-coördinaten een coördinatentransformatie naar een gelijkmatig lopende klok voorhanden is, dat wil zeggen het moet bekend zijn op precies welke wijze de klok onregelmatig loopt. Om de beschrijving van fysische processen eenvoudig te houden streeft men er algemeen naar om gelijkmatig lopende klokken te gebruiken: *"Time is defined in such a way that physics looks simple "* (Wheeler).

1.3 Mechanische grootheden

De Newtoniaanse mechanica houdt zich bezig met het beschrijven en relateren van *mechanische* grootheden. Van alle aspecten van een verschijnsel worden er

[3] De concentratie van een fysische gebeurtenis op één punt in ruimte en tijd is natuurlijk een idealisatie, maar voor veel toepassing is deze gerechtvaardigd.

[4] We doelen hier op het aantal tijdseenheden dat de gebeurtenissen scheidt, een niet-negatief getal: dit verklaart de kwadratische aard van de hier gegeven definitie.

slechts een beperkt aantal geselecteerd om te worden bestudeerd. We nemen aan dat het gebruikte coördinatensysteem Cartesiaans is, met een gelijkmatig lopende klok. De volgende grootheden gelden als elementair:

- **Lengte**. De drie afstanden vanaf de oorsprong tot een gegeven punt worden (in principe) gemeten met meetlatten, en uitgedrukt in *meters* (symbool m). De plaats van een punt wordt vaak aangegeven met $\vec{x}$ zoals hiervoor al gebeurde. De *vector* die van een punt A naar een punt B wijst is het verschil van hun coördinaten dat wil zeggen $\vec{x}_B - \vec{x}_A$.

- **Tijd**. De tijdsafstand vanaf de tijds-oorsprong tot een gegeven moment wordt (in principe) gemeten met een klok, en uitgedrukt in *seconden* (symbool s).

- **Massa**. De massa van een voorwerp is een maat voor de 'hoeveelheid materie' bevat in dat voorwerp. In een gegeven zwaartekrachtsveld is de massa van een voorwerp evenredig met het gewicht van dat voorwerp. Massa wordt uitgedrukt in *kilogrammen* (symbool kg).

Naast deze elementaire grootheden spelen een aantal samengestelde grootheden een rol in de mechanica. De belangrijkste hiervan zijn

- **Snelheid**. Bekijk een bewegend puntdeeltje[5]. Laat op een tijdstip t de positie van het deeltje $\vec{x}$ zijn, en laat op een infinitesimaal later tijdstip, $t + dt$, de positie veranderd zijn met een infintesimaal afstandje $d\vec{x}$:

$$\vec{x}(t) = (x, y, z) \quad , \quad \vec{x}(t + dt) = \vec{x}(t) + d\vec{x} = (x + dx, y + dy, z + dz) \ . \tag{6}$$

  De momentane snelheid op tijdstip t is dan per definitie

$$\vec{v} = \frac{d\vec{x}}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \equiv (v_x, v_y, v_z) \ . \tag{7}$$

[5]Het concept 'puntdeeltje' is een idealisatie. In het algemeen bedoelt de fysica met een puntdeeltje een materieel voorwerp dat zo klein of eenvoudig is dat de positie van het voorwerp op elk gegeven tijdstip volledig bepaald wordt door een enkele vector $\vec{x}$. Wat precies als een puntdeeltje beschouwd kan worden hangt dus van de context af: voor een kosmoloog die zich met de evolutie van het universum op grote schaal bezighoudt is een heel melkwegstelsel een puntdeeltje, terwijl voor de deeltjesfysicus een puntdeeltje wordt gevormd door een enkel quark, dat zich bevindt binnen een proton, in een atoom, in een preparaat, in een laboratorium, op een planeet, in een zonnestelsel, binnen diezelfde melkweg.

De ruimtelijke snelheid van het deeltje is dus een drie-dimensionele vector uitgedrukt in eenheden m/s. De *grootte* van de snelheid, $|\vec{v}|$, is geen vector maar een enkel getal, gegeven door

$$|\vec{v}|^2 = {v_x}^2 + {v_y}^2 + {v_z}^2 \quad . \tag{8}$$

- **Versnelling**. Evenals posities kunnen ook de snelheden zelf veranderlijk zijn in de tijd. Als de snelheid op een tijdstip t gelijk is aan $\vec{v}$, en op tijdstip $t + dt$ gelijk aan $\vec{v} + d\vec{v}$, dan noemt men de versnelling op tijdstip t

$$\vec{a} = \frac{d\vec{v}}{dt} \quad . \tag{9}$$

  Ook dit is weer een driedimensionele vector, met eenheid m/s^2.

- **Impuls**. Er bestaat een primitieve notie van 'hoeveelheid beweging' opgesloten in een bepaald voorwerp. Het moge duidelijk zijn dat van twee voorwerpen met gelijke massa het snellere 'meer beweging' bevat dan het langzamere, en dat bij gelijke snelheid een zwaarder voorwerp 'meer beweging' bevat dan een lichter voorwerp. Dit leidt tot de definitie van impuls van een voorwerp met massa M en snelheid $\vec{v}$:

$$\vec{p} = M\vec{v} = (Mv_x, Mv_y, Mv_z) \quad . \tag{10}$$

  Impuls wordt dus uitgedrukt in (kg m)/s, en is een driedimensionele vector.

- **Kinetische energie**. Een andere maat voor 'hoeveelheid beweging', die zich ook schikt naar de primitieve notie als boven, is de kinetische energie, die alleen afhangt van de massa en de grootte van de snelheid:

$$E_{\text{kin}} = \frac{1}{2}M|\vec{v}|^2 \quad . \tag{11}$$

  De eenheid hiervan is dus (kg m^2)/s^2, deze wordt ook wel de Joule genoemd.

- **Kracht**. Er bestaat ook een primitieve notie van 'hoeveelheid moeite' benodigd om beweging te beinvloeden. Het moge duidelijk zijn dat de beweging van zware voorwerpen moeilijker te beinvloeden valt dan van lichte voorwerpen, en dat de benodigde moeite groter is naarmate de beoogde snelheidsverandering groter is. De *definitie* van een kracht $\vec{F}$ stellen we nog even uit; maar in ieder geval is kracht een drie-dimensionele vector (aangezien snelheid en versnelling dat ook zijn), en zij wordt uitgedrukt in eenheden van (kg m)/s^2, ofwel Newton.

- **Draai-impuls**. Hoewel voor dit college minder relevant[6], noemen wij hier de 'hoeveelheid draaiing' van een systeem. Laat een puntdeeltje met massa M een cirkelbeweging uitvoeren rond een middelpunt, op afstand r, met snelheid $|\vec{v}|$. De grootte van de draai-impuls $\vec{L}$ wordt dan gegeven door

$$|\vec{L}| = M\ r\ |\vec{v}| \tag{12}$$

en heeft eenheid (kg m^2)/s.

1.4 Behoudswetten

Men kan zich afvragen waarom de definities van impuls, kinetische energie en draai-impuls de gekozen vorm hebben. Immers, ook de combinatie $M^{2.7}|\vec{v}|^9\vec{v}$ voldoet aan de primitieve notie van 'hoeveelheid beweging'. De achterliggende reden is als volgt. Allereerst zijn alle drie de genoemde definities *lineair* in de massa, en dit maakt het mogelijk ze samen te stellen. Als bijvoorbeeld twee puntdeeltjes met massa's M_1 en M_2 dezelfde snelheid $\vec{v}$ hebben, zijn hun afzonderlijke impulsen gelijk aan $\vec{p}_1 = M_1\vec{v}$ en $\vec{p} = M_2\vec{v}$. Men kan echter ook de twee puntdeeltjes *tezamen* beschouwen als één systeem, van massa $M \equiv M_1+M_2$, en het gezamenlijke systeem heeft dan impuls $\vec{p} = M\vec{v}$, hetgeen precies gelijk is aan $\vec{p}_1 + \vec{p}_2$. Ook als de snelheden verschillend zijn, kunnen we spreken over de *totale* impuls, energie en draai-impuls van een systeem als de *som* van deze grootheden, genomen over alle afzonderlijke componenten.

Nu geldt in zeer veel fysische processen dat van een systeem de totale impuls, de totale kinetische energie, en de totale draai-impuls *behouden* zijn, hetgeen wil zeggen dat deze totalen in de tijd niet veranderen[7]. Deze behoudswetten leggen dus beperkingen op aan de mogelijke einduitkomsten van een fysisch proces. Voor andere, alternatieve, definities van 'hoeveelheid beweging' of 'draaiing' gelden niet dergelijke nuttige behoudswetten.

Het moet worden opgemerkt dat, in uiterste aanleg, de behoudswetten *empirische* feiten zijn, geschraagd door een veelheid aan waarnemingen. Zij zijn *waargenomen* eigenschappen van ons universum. Echter, wij kunnen deze eigenschappen relateren aan dieper liggende eigenschappen. Howel dit hier niet bewezen zal worden, gelden de volgende relaties

[6] In de *quantummechanica* speelt draai-impuls een bijzonder belangrijke rol.

[7] De wet van behoud van kinetische energie geldt in het geval er geen kinetische energie wordt omgezet in inwendige energie, bijvoorbeeld door opwarming of chemische reacties. Voor elementaire deeltjes, die geen inwendige structuur hebben, geldt energiebehoud altijd.

- De wet van behoud van impuls volgt als we aannemen dat de oorsprong van het coördinatensysteem in de ruimte vrij gekozen mag worden zonder dat dit de natuurwetten beïnvloedt. Dat wil zeggen, de drie-dimensionale ruimte bevat niet een bevoorrecht *punt*, met bijzondere eigenschappen.

- De wet van behoud van energie volgt als we aannemen dat de oorsprong van het coördinatensysteem in de tijd vrij gekozen mag worden zonder dat dit de natuurwetten beïnvloedt. Dat wil zeggen, de tijd bevat niet een bevoorrecht *moment*, met bijzondere eigenschappen.

- De wet van behoud van draai-impuls volgt als we aannemen dat de coördinaat-assen van het Cartesiaanse coördinatensysteem in de ruimte vrij gekozen mogen worden zonder dat dit de natuurwetten beïnvloedt. Dat wil zeggen, de drie-dimensionele ruimte bevat niet een bevoorrechte *richting*, met bijzondere eigenschappen.

De mechanische behoudswetten kunnen dus geïnterpreteerd worden als aanwijzingen dat ons universum *homogeen* en *isotroop* is.

In gevallen waar wel sprake is van een bevoorrechte grootheid gelden de behoudswetten niet zonder meer. Als voorbeeld kunnen we de beweging van de aarde in het zonnestelsel nemen, die zich afspeelt onder invloed van de door de zon uitgeoefende zwaartekracht. Laat ons ons de zon voorstellen als vastgeprikt in het middelpunt van het zonnestelsel. Aangezien de door de zon uitgeoefende zwaartekracht in alle *richtingen* even groot is, geldt de wet van behoud van draai-impuls, en blijft de draai-impuls van de aarde constant (deze 'perkenwet' werd door Kepler *empirisch* gevonden, en eerst later door Newton *afgeleid*). De zwaartekracht hangt echter wel af van de *afstand* tot de zon. De wet van behoud van impuls geldt dus *niet*, en inderdaad verandert de impuls van de aarde voortdurend (de verandering van grootte is klein, maar de *richting* van de snelheid verandert, met 30 graden per maand). Laten we echter de notie van een vastgeprikte zon los, dan is ook het bevoorrechte punt verdwenen, en inderdaad is de totale impuls van het systeem 'zon+aarde' constant[8]. Merk op dat we in dit geval niet alleen de kracht moeten beschouwen die de zon op de aarde uitoefent, maar ook de kracht die de aarde op de zon uitoefent. Aangezien het zon-aarde systeem ook energie bevat in de vorm van gravitationele energie, die uitgewisseld kan worden met kinetische energie, is de kinetische energie *niet* behouden, echter weer wel de som van kinetische en gravitationele energie.

[8]Hierbij hebben we voor het gemak alle andere planeten buiten beschouwing gelaten.

1.5 De wetten van Newton

De Newtoniaanse mechanica kan geformuleerd worden in enkele hoofdwetten. Deze zijn in zekere zin *postulaten* die slechts door waarnemingen getoetst kunnen worden; zij zijn echter bijzonder succesvol gebleken in het klassieke regime, waar de beschouwde systemen betrekkelijk groot zijn (vergeleken met het niveau waarop de quantummechanica van belang wordt) en traag bewegen (in vergelijking met de lichtsnelheid).

1. **Eerste wet**. Als op een puntdeeltje geen kracht wordt uitgeoefend, volhardt het in zijn hoeveelheid beweging. Dat wil zeggen, als op tijdstip t de impuls $\vec{p}$ bedraagt, en op tijdstip $t + dt$ de waarde $\vec{p} + d\vec{p}$ heeft, zal bij afwezigheid van krachten gelden
$$d\vec{p} = 0 \tag{13}$$
Hierbij is het volgende van belang: het coördinatensysteem moet Cartesiaans zijn, met gelijkmatige klokken. Als dit niet het geval is, geldt de eerste wet niet. Wel moet het mogelijk zijn een coördinatentransformatie naar een toepasselijk Cartesiaans stelsel (met gelijkmatige klok) te maken. In feite is deze wet dus eigenlijk een keuze voor een coördinatensysteem en we mogen haar omkeren: als ons op een of anderen manier *gegarandeerd* is dat op een deeltje geen kracht werkt, kunnen we een Cartesiaans coördinatensysteem met gelijkmatige klok kiezen waarin de eerste wet van Newton geldt. Zo'n stelsel (inclusief keuze van oorsprong en coördinaat-assen) heet een *inertiaalstelsel* [9]. In feite zijn er oneindig veel mogelijke inertiaalstelsels, alle verbonden door coördinatentransformaties: deze zullen we nog nader bekijken. Hoewel inertiaalstelsels niet inherent superieur zijn aan andere mogelijke coördinatensystemen, hebben zij de prettige eigenschap dat zij de eerste wet mogelijk maken en beweging, in ieder geval in afwezigheid van krachten, eenvoudig is[10].

2. **Tweede wet**. Laat ons aannemen dat we een inertiaalstelsel gebruiken. Als de impuls van een deeltje in het (infinitesimale) tijdsinterval dt *wel* ver-

[9] Het in het bovenstaande besprolen voorbeeld van het zon-aarde systeem met de zon 'vastgeprikt' in de oorsprong van het coördinatensysteem kan dus anders gezien worden, namelijk als de keuze van een coördinatensysteem waarin de oorsprong meebeweegt met het middelpunt van de zon. Dit is dus *geen* inertiaalstelsel

[10] Merk op dat ook wrijving als een kracht telt. Het is dit feit dat de eerste wet van Newton doet verschillen van die van Aristoteles, die zegt dat in afwezigheid van krachten een puntdeeltje in *rust* is.

andert, relateert de tweede wet deze verandering aan de kracht die op het deeltje wordt uitgeoefend, als volgt:

$$\vec{F} = \frac{d\vec{p}}{dt} \tag{14}$$

Niettegenstaande dat er een primitieve notie van 'kracht' bestaat, is deze tweede wet in feite de *definitie* van mechanische kracht: als er een verandering van impuls wordt waargenomen, dan *moet* er een kracht werkzaam zijn. De vraag naar de *oorsprong* van deze kracht is weer een geheel andere. Een laatste punt: de tweede wet wordt ook wel geformuleerd als $\vec{F} = M\vec{a}$. Deze formulering geldt echter alleen in het geval de massa M van het deeltje constant is in de tijd. In bijvoorbeeld het geval van raket-aandrijving is dit echter niet zo: als de massa van het puntdeeltje in het tijdsinterval verandert met dM luidt de correcte formulering

$$\vec{F} = M\vec{a} + \frac{dM}{dt}\vec{v} \ .$$

Overigens is de eerste vorm van deze tweede wet de vorm die Newton zelf reeds gekozen had.

3. **Derde wet**. Deze behelst in feite het behoud van impuls. Bekijk twee deeltjes die onderling wisselwerken maar neem aan dat van buitenaf geen kracht op het gecombineerde systeem wordt uitgeoefend. Dan zal de totale impuls van het systeem constant blijven. De impulsverandering van deeltje 1 en die van deeltje 2, beide in tijdsinterval dt, moeten elkaar dus opheffen:

$$d\vec{p}_1 = -d\vec{p}_2 \tag{15}$$

met andere woorden de kracht $\vec{F}_{2\to 1}$, die op deeltje 1 wordt uitgeoefend (door deeltje 2) en de kracht $\vec{F}_{1\to 2}$, die op deeltje 2 wordt uitgeoefend (door deeltje 1) zijn ook tegengesteld:

$$\vec{F}_{2\to 1} = \frac{d\vec{p}_1}{dt} = -\ \frac{d\vec{p}_2}{dt} = -\ \vec{F}_{1\to 2} \ . \tag{16}$$

Waar de eerste twee wetten eigenlijk definities zijn (van inertiaalstelsel resp. kracht), is de derde wet er een met een fysisch postulaat, namelijk impulsbehoud.

4. **Vierde wet**. De tweede wet zegt ons niet meer dan dat een impulsverandering gerelateerd is aan een kracht: *als* we de kracht kennen, is de impulsverandering (en daarmee de beweging) in principe berekenbaar. De vierde wet postuleert wat de kracht is in het geval van gravitatie. Beschouw twee (punt)deeltjes met massa M_1 en M_2, op onderlinge afstand r. De gravitatiewet van Newton zegt dan dat de twee massa's elkaar aantrekken met een kracht die langs hun verbindingslijn ligt, ter grootte van

$$|\vec{F}_{\text{grav}}| = G_N \, \frac{M_1 M_2}{r^2} \quad . \tag{17}$$

Het feit dat deeltje 1 altijd een kracht ondervindt die naar deeltje 2 gericht is houdt in dat de zwaartekracht van deeltje 2 in alle richtingen even sterk is, en daarmee is draai-impuls behouden. Dit houdt onder andere in dat, voor een systeem van twee deeltjes die alleen gravitationeel wisselwerken, de onderlinge beweging van de deeltjes altijd in een enkel plat vlak plaatsvindt (in het aarde-zon systeem is dit het vlak van de eclipta). Het feit dat de kracht evenredig is met r^{-2} is verantwoordelijk voor het waargenomen feit dat de planeten in ellipsbanen rond de zon bewegen (in feite, in *kegelsneden*). De constante G_N is noodzakelijk, zoals al te zien uit het feit dat de eenheid van kracht (kg m)/s^2 is, en de combinatie $M_1 M_2/r^2$ de eenheid kg^2/m^2 heeft. Uit experiment blijkt de numerieke waarde van G_N:

$$G_N \approx 6.67\ 10^{-11}\ \frac{\text{m}^3}{\text{kg s}^2} \quad . \tag{18}$$

Dit is een fundamentele natuurconstante: haar waarde is in feite vrij *onnauwkeurig* bepaald, vanwege het feit dat zwaartekracht de *zwakste* onder de fundamentele natuurkrachten is[11].

Een ander, niet onbelangrijk aspect van deze vierde wet is dat de *tijd* er niet in voorkomt. In ons zonnestelsel wordt de beweging van de (lichte) aarde niet alleen geregeerd door de zon maar ook door de planeet Jupiter, de zwaarste der planeten[12]. De beweging van de aarde op enig moment

[11] De dagelijkse ervaring suggereert het tegendeel; dit ligt echter aan het feit dat bijvoorbeeld de aarde een grote massa heeft, terwijl haar electrische lading (die een electrische kracht oplevert) vrijwel te verwaarlozen is.

[12] In feite oefenen alle hemellichamen invloed op elkaar uit, onder invloed van de vierde wet; wij hebben het zonnestelsel hier vereenvoudigd tot het zon-Jupiter systeem met een relatief heel lichte aarde.

wordt dus ook beïnvloed door de onderlinge stand van de zon en Jupiter *op datzelfde moment*. De beweging van Jupiter rond de zon werkt dus *ogenblikkelijk* door in die van de aarde: met andere woorden, de effecten van zwaartekracht verplaatsen zich met *oneindige* snelheid.

1.6 De Galilei-transformaties

Zoals we gezien hebben, onderscheiden de inertiaalstelsels zich van andere coördinatensystemen in het feit dat in inertiaalstelsels de eerste wet van Newton geldt. Dit leidt natuurlijk tot de vraag op welke wijze deze inertiaalstelsels met elkaar verbonden zijn, dat wil zeggen: wat is de coördinatentransformatie tussen het ene inertiaalstelsel en het andere? We kunnen dit als volgt formuleren. Laat ons ons twee coördinatensystemen voorstellen, beiden Cartesisch met gelijkmatig lopende klokken; we hebben nog niet aangenomen dat Newton's eerste wet geldt in deze systemen, dus het zijn nog geen inertiaalstelsels. Wel nemen we aan dat homogeniteit en isotropie van de ruimte gelden, dus de oorsprong van beide systemen is vrij te kiezen zonder dat dit de beschrijving van de fysica beïnvloedt. Laten we de coördinaten van een gegeven gebeurtenis in het ene coördinatensysteem (systeem 1) aanduiden met

$$\vec{x} = (x, y, z) \quad \text{en} \quad t \;\;,$$

en de coördinaten van *dezelfde* gebeuretins in het andere coördinatensysteem (systeem 2) met

$$\vec{x}' = (x', y', z') \quad \text{en} \quad t' \;\;.$$

Een coördinatentransformatie is dan het recept om van de beschrijving van fysica in termen van systeem 1 over te stappen naar de beschrijving van *dezelfde* fysica in termen van systeem 2; dat wil zeggen, we zoeken de vorm van $\vec{x}'$ en t' in termen van $\vec{x}$ en t. De homogeniteit en isotropie laat ons toe om de inertiaalstelsels zo te kiezen dat de oorsprongen samenvallen:

$$(\vec{x} = 0 \;,\; t = 0) \quad \leftrightarrow \quad (\vec{x}' = 0 \;,\; t' = 0) \;\;. \tag{19}$$

Analoog hebben we snelheden $\vec{v}$ en $\vec{v}'$ in de twee systemen, en evenzo de versnellingen $\vec{a}$ en $\vec{a}'$. Aangezien er geen bevoorrechte richting is, mogen we ook de Cartesisiche coordinaatassen zo kiezen dat ze op $t = t' = 0$ met elkaar samenvallen: dat wil zeggen dat voor de ruimtelijke coördinaten van iedere gebeurtenis geldt

$$t = t' = 0 \quad \Rightarrow \quad \vec{x} = \vec{x}' \;\;. \tag{20}$$

Laat ons nu aannemen dat we met twee inertiaalstelsels te doen hebben, en bekijk een deeltje waarop geen kracht werkt. In dat geval hebben we

$$\vec{a} = \frac{d\vec{v}}{dt} = 0 \quad \text{en} \quad \vec{a}' = \frac{d\vec{v}'}{dt'} = 0 \ \ . \tag{21}$$

We maken nu de cruciale aanname dat **de klokken in beide inertiaalstelsels even snel lopen**:

$$dt = dt' \ \ . \tag{22}$$

Dit is op het eerste gezicht zeer redelijk: toch is het dit punt waarop de relativiteitstheorie van de Newtoniaanse mechanica afwijkt. Onder deze aanname hebben we voor ons vrije deeltje

$$d\vec{v} = 0 \quad \text{en} \quad d\vec{v}' = 0 \ \ . \tag{23}$$

Dit betekent dat de vectoren $\vec{v}$ en $\vec{v}'$, die per slot in principe ook functies van t resp. t' zijn, slechts een constant verschil kunnen hebben. Laat ze dan verschillen met een constante vector

$$\vec{v}_0 = (v_0^x, v_0^y, v_0^z) \tag{24}$$

We hebben dan

$$\vec{v} = \vec{v}' + \vec{v}_0 \ \ , \tag{25}$$

ofwel

$$\vec{v}' = \vec{v} - \vec{v}_0 \ \ , \tag{26}$$

ofwel, meer expliciet (met $dt = dt'$):

$$\frac{dx'}{dt} = \frac{dx}{dt} - v_0^x \quad , \quad \frac{dy'}{dt} = \frac{dy}{dt} - v_0^y \quad , \quad \frac{dz'}{dt} = \frac{dz}{dt} - v_0^z \quad . \tag{27}$$

Dit geeft ons het gevraagde verband tussen $\vec{x}$ en $\vec{x}'$:

$$\vec{x}' = \vec{x} - t\vec{v}_0 \ \ . \tag{28}$$

Dit is de *Galilei-transformatie*: heel expliciet luidt deze

$$\begin{aligned} x' &= x - t\, v_0^x \ \ , \\ y' &= y - t\, v_0^y \ \ , \\ z' &= z - t\, v_0^z \ \ , \\ t' &= t \ \ . \end{aligned} \tag{29}$$

Merk op dat inderdaad de oorsprongen $\vec{x} = 0, t = 0$ en $\vec{x}' = 0, t' = 0$ samenvallen. De fysische inhoud van deze transformaties wordt duidelijk als we bekijken hoe de ruimtelijke oorsprong van systeem 2, $\vec{x}' = 0$, er uitziet in termen van systeem 1:

$$\vec{x}' = 0 \quad \Rightarrow \quad \vec{x} = t\,\vec{v}_0 \ . \tag{30}$$

Deze ruimtelijke oorsprong beweegt zich dus met uniforme snelheid $\vec{v}_0$ in systeem 1. Zonder verlies van algemeenheid stelt men vaak dat deze snelheid langs de x-as gericht is (als alle richtingen equivalent zijn, zoals aangenomen, mag dit): dan hebben we $v_0^x = v_0, v_0^y = 0, v_0^z = 0$. De Galilei transformatie geldt *tussen inertiaalstelsels*, onafhankelijk van wat er zich precies in die inertiaalstelsels afspeelt! De upshot van de Galilei transformatie is dus de volgende: als een deeltje, al dan niet aan een kracht onderworpen, een snelheid $\vec{v}$ heeft in systeem 1, dan is zijn snelheid op datzelfde moment in systeem 2 gegeven door

$$\vec{v}' = \vec{v} - \vec{v}_0 \ . \tag{31}$$

Omdat $\vec{v}_0$ constant is, zijn de tijdsafgeleiden van de snelheden, dat wil zeggen de versnellingen, en dus de krachten, gelijk in beide inertiaalstelsels.

Een laatste opmerking: we hebben inertiaalstelsels bekeken waarin op $t = t' = 0$ de assen samenvallen. Dit is niet het meest algemene geval: twee inertiaalstelsels die ten opzichte van elkaar in rust zijn, maar waarin de assen van het ene systeem gekanteld zijn ten opzichte van het andere systeem, vertonen beide de eerste wet van Newton, *nits* de kanteling van de assen vast is, dat wil zeggen niet van de tijd afhangt. Aangezien deze kanteling geen bijzondere fysische relevantie heeft wordt dit geval meestal buiten beschouwing gelaten. Zij kan wel worden betrokken, maar slechts ten koste van ingewikkelder wiskunde. Een andere situatie is die van een *roterend* coördinatensysteem: een ieder die wel eens op een draaimolen heeft gezeten weet dat in zo'n systeem de eerste wet van Newton *niet* geldt, en het is dan ook geen inertiaalsysteem[13].

1.7 Behoud van impuls en energie

Een natuurwet is alleen nuttig als ze voldoende algemeen is. We moeten daarom onderzoeken of de wetten van behoud van impuls en energie, bovengenoemd,

[13]Merk op dat er tussen een inertiaalstelsel en een roterend, bewegend coördinatensysteem natuurlijk wel een coördinatentransformatie bestaat: haar vorm is echter behoorlijk ingewikkeld, en voor dit college niet interssant.

dezelfde vorm hebben in twee inertiaalstelsels die door Galilei transformaties verbonden zijn. Laat ons de volgende situatie bekijen: een aantal *inkomende* puntdeeltjes, genummerd $i = 1, 2, \ldots, n$, komen bij elkaar en hebben wisselwerking: na enige tijd komen een aantal *uitgaande* puntdeeltjes, genummerd $j = 1, 2, \ldots, k$ uit deze wisselwerking tevoorschijn[14]. Laat ons de massa's van de inkomende deeltjes aangeven met M_i, en die van de uitgaande deeltjes met M_j. Het is mogelijk dat het aantal deeltjes, en hun massa's gelijk blijven (zodat dus $n = k$ en $M_i = M_j$ wanneer $i = j$), maar dit hoeft niet.

Ten eerste bekijken we nu deze situatie in coördinatensysteem 1. Laat ons de snelheden van de inkomende deeltjes aangeven met $\vec{v}_i$, en die van de uitgaande deeltjes met $\vec{v}_j$. De *totale* impuls van het systeem vóór de wisselwerking is de som van de afzonderlijke impulsen:

$$\vec{p}_{in} = \sum_{i=1}^{n} M_i \vec{v}_i \quad . \tag{32}$$

en die van het systeem ná de wisselwerking is

$$\vec{p}_{uit} = \sum_{j=1}^{k} M_j \vec{v}_j \quad . \tag{33}$$

Behoud van impuls zegt dus dat, in termen van systeem 1,

$$\vec{p}_{in} = \vec{p}_{uit} \quad \Rightarrow \quad \sum_{i=1}^{n} M_i \vec{v}_i = \sum_{j=1}^{k} M_j \vec{v}_j \quad . \tag{34}$$

Laat ons nu bezien hoe dit er uitziet in termen van systeem 2. Voor elk deeltje moet de regel (**??**) gelden , dus $\vec{v}_i' = \vec{v}_i - \vec{v}_0$ en $\vec{v}_j' = \vec{v}_j - \vec{v}_0$. Daarmee wordt het impulsbehoud zoals waargenomen in systeem 1 vertaald naar de termen van systeem 2 als volgt:

$$\sum_{i=1}^{n} M_i(\vec{v}_i' - \vec{v}_0) = \sum_{j=1}^{k} M_j(\vec{v}_j - \vec{v}_0) \quad . \tag{35}$$

[14]We kunnen hierbij denken aan een aantal moleculen die botsen, en daarbij ingrediërende atomen uitwisselen: bijvoorbeeld, de chemische reactie $2H_2 + O_2 \quad \rightarrow \quad 2H_2O$, waarbij dus $n = 3$ en $k = 2$: anderzijds mogen we ons ook twee sterren indenken, die in elkaars nabijheid materiaal uitwisselen

Met, uiteraard,

$$\vec{p}\,'_{in} = \sum_{i=1}^{n} M_i \vec{v}\,'_i \quad , \quad \vec{p}\,'_{uit} = \sum_{j=1}^{K} M_j \vec{v}\,'_j \tag{36}$$

kunnen we dit schrijven als

$$\vec{p}\,'_{in} - \vec{v}_0 \sum_{i=1}^{n} M_i = \vec{p}\,'_{uit} - \vec{v}_0 \sum_{j=1}^{k} M_j \quad . \tag{37}$$

We nemen nu aan **dat de totale massa behouden is**, dat wil zeggen

$$\sum_{i=1}^{n} M_i = \sum_{j=1}^{k} M_j \quad . \tag{38}$$

Hoewel niet expliciet in de wetten van Newton bevat, is dit een noodzakelijke, en op het eerste gezicht ook redelijke aanname[15]. Onder deze aanname hebben we

$$\vec{p}\,'_{in} = \vec{p}\,'_{uit} \quad , \tag{39}$$

en we concluderen: *als* in systeem 1 impulsbehoud geldt, en *als* massa behouden is, *dan* geldt impulsbehoud ook in systeem 2. Merk op dat dit *niet* het bewijs is dat impuls behouden is, maar *wel* het bewijs dat de behoudswet *invariant* is onder Galilei transformaties, of ze nu empirisch juist is of niet.

Voor de kinetische energie kunnen we eenzelfde beschouwing uitvoeren. Hierbij nemen we aan dat vóór de wisselwerking de inkomende deeltjes als vrij te beschouwen zijn, dat wil zeggen dat hun energie zuiver kinetisch is en er geen energie bevat is in enige wisselwerking tussen de deeltjes; zo ook ná de wisselwerking. In dat geval geldt voor de totale (kinetische) energie vóór de wisselwerking, in termen van systeem 1,

$$E_{in} = \sum_{i=1}^{n} \frac{1}{2} M_i |\vec{v}_i|^2 \quad , \tag{40}$$

en voor de totale energie ná de wisselwerking:

$$E_{uit} = \sum_{j=1}^{k} \frac{1}{2} M_j |\vec{v}_j|^2 \quad . \tag{41}$$

---

[15] Dit gaat er van uit dat massa geschapen noch vernietigd kan worden: in de klassieke mechanica is dit inderdaad het geval, zoals bijvoorbeeld ons wordt geleerd in de scheikunde. Wel kan massa her-verdeeld worden over de puntdeeltjes, en dit is waar de scheikunde zich mee bezig houdt.

Behoud van energie is dan de bewering $E_{in} = E_{uit}$, oftewel

$$\sum_{i=1}^{n} \frac{1}{2} M_i |\vec{v}_i|^2 = \sum_{j=1}^{k} \frac{1}{2} M_j |\vec{v}_j|^2 \quad . \tag{42}$$

De overgang naar de beschrijving in termen van systeem 2 gaat als volgt:

$$\begin{aligned} |\vec{v}|^2 &= \vec{v} \cdot \vec{v} = (\vec{v}' - \vec{v}_0) \cdot (\vec{v}' - \vec{v}_0) \\ &= \vec{v}' \cdot \vec{v}' - 2\vec{v}' \cdot \vec{v}_0 + \vec{v}_0 \cdot \vec{v}_0 = |\vec{v}'|^2 - 2\vec{v}' \cdot \vec{v}_0 + |\vec{v}_0|^2 \quad , \end{aligned} \tag{43}$$

zodat E_{in} dan de volgende vorm krijgt:

$$\begin{aligned} E_{in} &= \sum_{i=1}^{n} \frac{1}{2} M_i |\vec{v}_i'|^2 - \vec{v}_0 \cdot \left(\sum_{i=1}^{n} M_i \vec{v}_i' \right) + \frac{1}{2} |\vec{v}_0|^2 \sum_{i=1}^{n} M_i \\ &= E_{in}' - \vec{v}_0 \cdot \vec{p}_{in}' + \frac{1}{2} |\vec{v}_0|^2 \sum_{i=1}^{n} M_i \quad . \end{aligned} \tag{44}$$

Zo ook hebben we natuurlijk

$$E_{uit} = E_{uit}' - \vec{v}_0 \cdot \vec{p}_{uit}' + \frac{1}{2} |\vec{v}_0|^2 \sum_{i=j}^{n} M_j \quad . \tag{45}$$

Samen met de wet van impulsbehoud en die van behoud van massa komen we dus op

$$E_{in}' = E_{uit}' \quad , \tag{46}$$

zodat: *als* impuls behouden is, en *als* massa behouden is, en *als* in systeem 1 energie behouden is, dan is zij ook in systeem 2 behouden[16].

[16]Het moge duidelijk zijn dat ook voor draai-impuls een behoudswet geformuleerd kan worden die invariant is onder Galilei transformaties. Gezien de ingewikkelder definitie van draai-impuls in termen van posities, massa's en snelheden laten we deze hier achterwege: volsta het op te merken dat ook het behoud van draai-impuls volgt uit de aanname van impulsbehoud en behoud van massa.

2 De lichtsnelheid als natuurconstante

De Post-Newtoniaanse mechanica zoals ontwikkeld door Einstein, Minkowski en anderen berust op het postulaat van de lichtsnelheid als fundamentele constante van de natuur. De semi-quantitatieve consequenties van dit postulaat worden hieronder behandeld.

2.1 Licht als golfbeweging in de æther

De snelheid van het licht is naar alledaagse maatstaven zodanig groot dat het geen verwondering hoeft te wekken dat tot voor enige eeuwen het geenszins duidelijk was dat licht überhaupt een snelheid *heeft*: in een wereld beheerst door afstanden als die van de ene stad tot de andere, en snelheden als die van paarden, postduiven, en in extreme gevallen kanonskogels, is het een goede benadering dat licht zich ogenblikkelijk verplaatst, over willekeurige afstanden[17]. In de achttiende en negentiende eeuw kwam de technologie ter beschikking die een meting van de lichtsnelheid[18] mogelijk maakte, in overeenstemming met de moderne waarde:

$$c = 299792458\ \frac{\mathrm{m}}{\mathrm{s}}\ \ . \tag{47}$$

Ook over de aard van het licht bestond onduidelijkheid: waar Newton meende dat licht uit 'licht'deeltjes bestond, opperde Huygens dat licht een golfverschijnsel was[19]. In de negentiende eeuw was de consensus dat het pleit beslecht was in het voordeel van de golf-hypothese, met name door de interferentie-experimenten van Young (1802). Het bestuderen en meten van interferentie-effecten van licht nam gedurende deze eeuw een hoge vlucht.

Een conceptueel probleem was echter het volgende. Naar analogie van oppervlaktegolven in water, geluidsgolven in lucht, water, metaal etcetera, leek het duidelijk dat de golfbeweging, die licht was, er een moest zijn in een *medium*, een stof die als drager van deze trillingen dienst doet. Wat nu was de aard van dit

[17]Experimenten zoals ondernomen door Galilei, met assistenten op verschillende heuveltoppen, uitgerust met lantarens en snelle reflexen, waren achteraf gezien natuurlijk tot mislukken gedoemd.

[18]Hier en in het vervolg zullen wij met 'lichtsnelheid' bedoelen de snelheid van licht*signalen* in het *vacuum*. In een medium bewegen lichtsignalen zich langzamer; in lucht is de lichtsnelheid echter maar zeer weinig kleiner dan in vacuum.

[19]De tegenstelling tussen deze twee theorieën is maar schijn; zij worden verzoend in de *quantummechanica*. Het was Einstein die in zijn verklaring van het foto-electrisch effect het deeltjeskarakter van licht aannemelijk maakte; deze verklaring, en niet de relativiteitstheorie, was de aanleiding om hem de Nobelprijs toe te kenne.

medium? Het kon geen der bekende stoffen zijn, immers dan zou licht van verre sterren ons nooit kunnen bereiken door het vacuüm van de ruimte heen. Men poneerde dus dat de ruimte (ook dat wat wij met 'vacuüm' aanduiden) gevuld was met een nieuw, onbekend medium, waaraan de naam *æther* werd gegeven[20].

Het moge duidelijk zijn dat de æther een bijzonder materiaal is. Stoffelijke voorwerpen bewegen zich zonder wrijving door de æther; immers, anders zouden de planeten, geleidelijk door wrijving afgeremd, vroeger of later in de zon storten. De æther moet dus wel bijzonder ijl zijn. Anderzijds is de snelheid van golven in een medium afhankelijk van de eigenschappen van dit medium: de snelheid van het geluid in onze relatief onsamenhangende atmosfeer is veel lager dan die onder het zeeoppervlak[21], omdat water dichter en steviger is: in metaal of steen is ze nog weer veel hoger. Een snelheid van 300000 kilometer per seconde veronderstelt een æther van ongekende starheid.

2.2 Lichtgolven als electromagnetisch verschijnsel

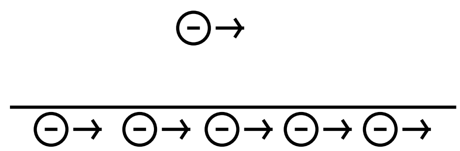
In de tweede helft van de negentiende eeuw werd duidelijk dat licht een electromagnetisch verschijnsel was. Deze worden beschreven door de wetten geformuleerd door Maxwell, die voor het electromagnetisme dezelfde status hebben als die van Newton voor de mechanica: met hun hulp kunnen electromagnetische verschijnselen nauwkeurig verklaard en voorspeld worden. De Maxwell vergelijkingen zijn geformuleerd in termen van electrische en magnetische velden die geacht worden zich af te spelen in een materieel medium of in vacuüm (dat wil zeggen, in de æther). Zij laten de mogelijkheid toe dat electrische en magnetische velden gezamenlijk een trilling uitvoeren, die zich dan met snelheid c voortplant. In de experimenten van Hertz werd aangetoond dat zulke electromagnetische golven inderdaad opgewekt kunnen worden. Daarmee was duidelijk geworden dat zichtbaar licht onderdeel uitmaakt van een veel breder spectrum aan electromagnetische straling: zichtbaar licht is toevallig met het blote oog waarneembaar,

[20]Aan het klassiek-Griekse systeem van Empedokles, met de vier elementen aarde, water, vuur em lucht (ἄηρ, de 'beneden-lucht' zoals men die ademt), voegde Aristotles een vijfde toe, waaruit de sferen der hemellichamen waren opgebouwd: de æther (ἄιθηρ, de 'boven-lucht', het materiaal van het uitspansel).

[21]Overigens is deze hogere snlheid *niet* de reden dat walvissen over afstanden van vele honderden kilomters met elkaar kunnen communiceren door te toeteren en te fluiten: de werkelijke reden is dat onze oceanen essentieel twee-dimensionaal zijn (veel langer en breder dan dieper) waardoor de geluidsintensiteit niet afneemt als $1/r^2$ maar als $1/r$.

andere kleuren zijn dat voor ons niet[22]. Niettegenstaande hun succes leken de Maxwell vergelijkingen een ernstig mankement te hebben: zij zijn niet invariant onder Galilei transformaties.

Dit kan eenvoudig worden ingezien in de situatie zoals hieronder geschetst. We beschouwen een metaaldraad, in rust in een zeker coördinatensysteem. Door de metaaldraad laten we een electrische stroom lopen: deze wordt gevormd door electronen die door de draad stromen als water door een tuinslang. Hun snelheid hoeft niet groot te zijn[23]. De draad bevat evenveel negatieve lading (in de electronen) als positieve (in het metaal-rooster van de draad) en is dus electrisch neutraal. Buiten de draad beweegt zich, parallel eraan, een ander electron, met dezelfde snelheid. De electronen in de draad vormen een electrische stroom, die een magneetveld buiten de draad opwekt: door dit magneetveld beweegt zich het enkele electron. De wetten van Maxwell poneren dat een electrisch geladen deeltje dat door een magneetveld beweegt een kracht ondervindt, de zogenaamde Lorentzkracht. In dit geval is deze evenredig met de snelheid van het deeltje, en naar de draad toe gericht: het electron wordt dus door de draad aangetrokken. Dit effect is eenvoudig te bestuderen en wordt op grote schaal gebruikt in bijvoorbeeld electromotoren.



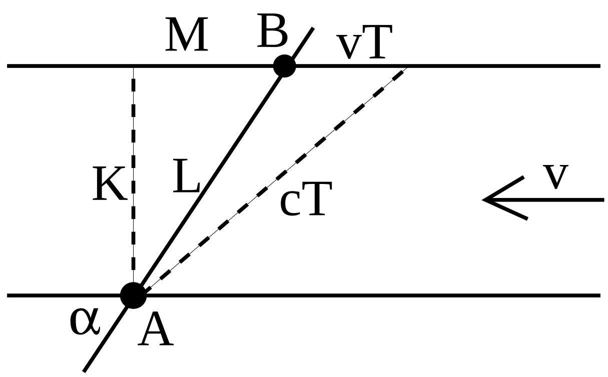
Laat ons nu een Galilei-transformatie uitvoeren: we stappen over op een coördinatensysteem dat precies met het electron mee beweegt. In dit systeem is het losse electron in rust, en kan dus geen Lorentzkracht ondervinden (aangezien deze met de snelheid evenredig moet zijn). Echter, versnellingen en daarmee krachten zijn Galilei-invariant. Het is dus niet duidelijk wat in *dit* coördinatensysteem de aantrekkende kracht veroorzaakt. De relativiteitstheorie kan, zoals we zullen zien, deze paradox oplossen.

[22]Veel insecten kunnen kleuren waarnemen in het *ultraviolet*, dat wil zeggen kleuren die in de regenboog 'voorbij het violet' liggen. Andere diersoorten zijn weer op andere manieren gevoelig voor electromagnetisme: zo oriënteren postduiven zich via het aardmagnetische veld, en hebben haaien zintuigen die waarnemen hoe een prooi het omringende magnetisch veld met zijn aanwezigheid verstoort.

[23]Gewoonlijk enige centimeters per seconde.

2.3 Ætherwind?

Als de æther niet direct voelbaar is, hoe is haar bestaan dan aan te tonen? Een mogelijke aanpak is de volgende. Als licht een golfbeweging in de æther is, dan moet de bewegingstoestand van het licht afhangen van de bewegingstoestand van de æther – *ten opzichte van de waarnemer*. Aangezien wij ons op een roterende aardbol bevinden, die eenmaal 's jaars een cirkelbaan rond de zon beschrijft, die op haar beurt weer meeroteert in de melkweg, die (enzovoorts), zal een aardse waarnemer zich in het algemeen niet in rust bevinden ten opzichte van de æther[24]. We voeren nu het volgende experiment uit. Laat een lichtsignaal een weg afleggen van een punt A naar een punt B en terug. Ten opzichte van de waarnemer waait er een ætherwind waarop het licht 'meedrijft' tijdens zijn beweging. Onderstaande figuur maakt dit aanschouwelijk.



De af te leggen weg is L. Het licht beweegt met snelheid c ten opzichte van de æther, die ten opzichte van de waarnemer beweegt met snelheid v. In de tijd T die het licht nodig heeft drijft het dus af over een afstand vT. De regel van Pythagoras geeft ons twee relaties:

$$K^2 + M^2 = L^2 \quad \text{en} \quad K^2 + (M + vT)^2 = (cT)^2 \ , \tag{48}$$

waaruit we kunnen concluderen dat

$$(c^2 - v^2)T^2 - 2MvT - L^2 = 0 \ , \tag{49}$$

[24]Het valt niet op beginselgronden uit te sluiten dat het universum zo geconstrueerd is dat de æther in rust is ten opzichte van het aardoppervlak – een wereldbeeld,echter, dat ons onmiddelijk terugvoert naar de dagen van Ptolemæus.

ofwel

$$\begin{aligned} T &= \frac{1}{c^2 - v^2}\left(Mv + \sqrt{M^2v^2 + L^2(c^2 - v^2)}\right) \\ &= \frac{L}{c^2 - v^2}\left(v\cos\alpha + \sqrt{c^2 - v^2(\sin\alpha)^2}\right) \quad ;. \end{aligned} \tag{50}$$

Hierbij hebben we gebruikt dat α de hoek is tussen het stuk AB en de richting van de ætherwind. Om de totale tijd, benodigd voor het traject $A \to B \to A$ af te leggen, moeten we het resultaat voor v en dat voor $-v$ bij elkaar optellen, hetgeen leidt tot

$$T_{ABA} = 2L\frac{\sqrt{c^2 - v^2(\sin\alpha)^2}}{c^2 - v^2} \quad . \tag{51}$$

Merk op dat dit altijd groter is dan 1: het effect van de ætherwind is dus altijd een tijdsverlies. Als $v = 0$ vinden we $T_{ABA} = 2L/c$ naar behoren; als v klein is ten opzichte van c is het antwoord in goede benadering

$$T_{ABA} \approx \frac{2L}{c}\left(1 + \frac{v^2}{c^2}\left(1 - \frac{(\sin\alpha)^2}{2}\right)\right) \quad . \tag{52}$$

Door nu lichtsignalen in gelijke fase van een punt te laten vertrekken in verschillende richtingen en door spiegels naar het vertrekpunt terug te laten keren, kan men proberen de kleine verschillen in vertraging door middel van interferentie-effecten waar te nemen. Voor v denke men aan de snelheid van de aarde in haar baan roind de zon, zo'n 40 kilometer per seconde. Het behoeft geen betoog dat dit een precisie-experiment van hoge orde is; toch was dit in de tweede helft van de negentiende eeuw zonder meer uitvoerbaar. Michelson en Morley, die als eersten dit effect probeerden te meten, vonden echter dat de vertraging, zo zij er al was, *in alle richtingen gelijk was!*

2.4 De Lorentz-Fitzgerald contractie

Een mogelijke uitweg uit het prbleem van de ogenschijnlijk onveranderlijke terugkeertijden T_{ABA} werd geopperd door Fitzgerald en Lorentz. Laat ons voor het gemak aannemen dat we het Michelson-Morley experiment uitvoeren voor twee waarden van α, namelijk 0^o ('tegen de wind in/met de wind in de rug') en 90^o ('dwars op de wind'). In het eerste geval bedraagt de vertraging een factor $1/(1 - v^2/c^2)$, in het tweede een factor $1/\sqrt{1 - v^2/c^2}$, zoals onmiddelijk uit vergelijking (**??**) valt af te lezen. Laat ons nu eens aannemen dat er *een of ander* mechanisme bestaat dat voorwerpen die met snelheid v door de æther bewegen, daardoor afgeplat worden met een factor $\sqrt{1 - v^2/c^2}$. Dit zou dan precies het ontbreken van een effect in het Michelson-Morley experiment verklaren. Het probleem bij deze hypothese is, enerzijds, dat we hiermee toegeven dat de æther wel degelijk wisselwerking met de materie moet hebben (ook al is bij lage snelheden het effect natuurlijk bijzonder klein), anderzijds dat een dergelijk mechanisme niet zonder meer voorhanden is.

Eerst toen in de eerste decennia van de twintigste eeuw de structuur van materie zodanig was opgehelderd dat het duidelijk was dat de atomen, waaruit alle materie is opgebouwd, werden beheerst door de wetten van het electromagnetisme, waarbij het atoom wordt opgevat als een mini-zonnestelsel geregeerd door electrische in plaats van gravitationele aantrekking[25], werd het contractie-mechanisme geïdentificeerd. Als de atoomkern in rust is, is zijn electromagnetisch veld in alle richtingen gelijk, en zal een klassiek electron een cirkelbaan kunnen beschrijven. Een dergelijk atoom kan men dus als een bol zien. Als echter de kern beweegt, wordt zijn electromagnetisch veld vervormd, en het electron zal dan een ellipsvormige baan beschrijven, die afgeplat is met, inderaad, een factor $\sqrt{1 - v^2/c^2}$. Bewegende voorwerpen worden dus inderdaad korter, in overeenstemming met de Lorentz-Fitzgerald hypothese. Tegenwoordig zien wij deze contractie echter liever als een gevolg van de Lorentz-transformaties, waarover later meer.

2.5 Einstein's postulaat

Indien de æther zich zodanig pervers gedraagt dat het onmogelijk lijkt haar bestaan op enige wijze aan te tonen, mag men zich gevoeglijk afvragen of dit bestaan nog wel nodig of zelfs maar gewenst is. Inderdaad verwierp Einstein het bestaan van de æther. De constantheid van de lichtsnelheid komt hiervoor in de plaats: Einstein postuleerde eenvoudigweg dat het hebben van snelheid c een eigenschap van licht is. Vraagt men zich af ten opzichte van welk coördinatensysteem deze snelheid is gedefinieerd, dan luidt het antwoord dat, bij het ontbreken van een voorkeurs-systeem zoals een coördinatensysteem dat in rust is ten opzichte van

[25] Op zijn beurt geeft dit beeld ook weer aanleiding tot problemen, zoals het gebrek aan stabiliteit van een dergelijk systeem. Deze kunnen alleen worden opgelost in een quantummechanische beschrijving van het atoom: binnen de *klassieke* fysica heeft dit model dus een twijfelachtige status.

de æther (die immers niet bestaat), deze snelheid dezelfde is in *alle* systemen[26]. Het spreekt voor zich dat deze boude bewering op gespannen voet staat met de dagelijkse ervaring. Immers, een snel voertuig kan men proberen in te halen door er snel achteraan te bewegen, zodat de relatieve snelheid kleiner wordt: voor licht geldt dit, naar beweerd wordt, *niet*. Zo ook zal, als een snelle raket een lichtsignaal vooruit zendt, dit signaal zich met snelheid c van de raket verwijderen (vanuit het standpunt van de raketvaarders); een waarnemer op aarde echter zal voor het licht dezelfde snelheid c meten, zonder dat het licht van de raket 'een zetje mee krijgt'. Zoals zo vaak in de fundamentele natuurkunde is bij deze tegenspraak met de op dagelijkse praktijk gebaseerde intuïtie het deze laatste die moet wijken: de meesten onder ons hebben eenvoudigweg niet genoeg ervaring met bewegingen bij extreem hoge snelheid om een adequate intuïtie te hebben ontwikkeld.

Het moge duidelijk zijn dat de klassieke, Newtoniaanse mechanica moet worden aangepast om ook situaties waarin snelheden in de buurt van c optreden correct te beschrijven. Als een voorbeeld kan de Galilei- transformatie dienen: hierin worden snelheden eenvoudig bij elkaar opgeteld, zie vergelijking (**??**), en deze kan dus niet onverkort gehandhaafd blijven. Wel is het natuurlijk zo dat bij 'lage' snelheden[27] de successen van het Newtoniaanse model moeten worden gereproduceerd. De precieze vorm van deze *relativistische mechanica* zal nog nader worden behandeld: eerst zullen we een aantal effecten op semi-quantitatieve manier bekijken.

2.6 De lichtklok: tijdsdilatatie

Aangenomen dat wij ons hebben laten overtuigen van het *empirische* feit van een lichtsnelheid die voor alle waarnemers gelijk is aan c, kunnen we komen tot de constructie van een *lichtklok*. Deze is niets anders dan een holle glazen buis van één meter lengte, met aan een kant een handvat, uitgerust met een drukknop. Als op de knop gedrukt wordt gaat een flitslampje aan de kant van het handvat over, en beweegt zich een lichtsignaaltje met snelheid c naar het andere uiteinde van de buis. We hebben hier een 'klok' die precies één keer 'tikt'[28], te weten in een

[26] Hiermee bedoelt men de *lokale* snelheid, dat wil zeggen de snelheid gemeten over 'kleine' afstanden.

[27] 'Laag' in de zin van: veel kleiner dan c.

[28] met behulp van spiegeltjes kan men desgewenst een klok construeren waarin het licht dan heen en weer kaatst, en de klok meerdere keren tikt: voor ons betoog is dit echter niet nodig.

tijdsduur van

$$\tau = \frac{1\ \mathrm{m}}{c} = \frac{1}{3\ 10^8}\mathrm{s}\ \ . \tag{53}$$

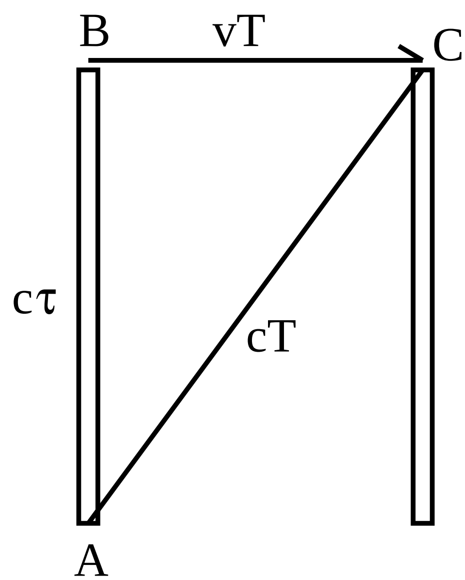
Deze klok is ideaal in die zin, dat wij weten dat ze altijd goed loopt (voor de bezitter van de klok wel te verstaan). We voeren nu het volgende gedachten-experiment uit. Waarnemer W_1 beweegt met snelheid v langs een andere waarnemer, W_2. W_1 is uitgerust met een lichtklok en houdt deze 'verticaal', loodrecht op de bewegingsrichting. Nu laat waarnemer W_1 zijn lichtklok tikken. Zij beschrijft[29] de gebeurtenis als volgt:

"*de lichtklok staat voor mij stil. Het licht reist van* A *naar* B*. De lengte van mijn lichtklok is* $c\tau$ *en het licht is na een tijd* τ *aangekomen.*"

Zij heeft natuurlijk gelijk. Waarnemer W_2 beschrijft de gebeurtenis daarentegen als volgt:

"*de lichtklok beweegt voor mij met snelheid* v*. Op het moment dat het licht, na een tijd* T*, is aangekomen, heeft dit uiteinde zich verplaatst, en wel over een afstand* vT*. Het bevindt zich dan ter plekke* C*. Het licht heeft dus de diagonale weg* AC *afgelegd, en deze afstand is* cT*".

Ook zij heeft gelijk. De situatie is hieronder weergegeven.



[29]Een bedenkelijke valkuil wordt geopend door gebruik van de term 'waarnemer X *ziet* deze of gene gebeurtenis'. Aangezien licht met eindige snelheid reist, gaat het *zien* van een gebeurtenis altijd gepaard met enige vertraging, tenzij de waarnemer ter plekke van de gebeurtenis is. Het is beter te veronderstellen dat de waarnemers op alle relevante plaatsten detectie-apparatuur hebben geïnstalleerd, die dan, eventueel veel later, de waarnemingen rapporteren aan de waarnemers.

Wederom met behulp van Pythagoras kunnen we dus stellen dat

$$(c\tau)^2 + (\nu T)^2 = (cT)^2 \quad , \tag{54}$$

ofwel dat T en τ gerelateerd zijn als volgt:

$$T = \frac{\tau}{\sqrt{1 - \nu^2/c^2}} \quad . \tag{55}$$

De onvermijdelijke conclusie is dat T, dat wil zeggen de tijd die waarnemer W_2 meet voor de duur van één tik van W_1's lichtklok, is toegenomen met een factor $1/\sqrt{1-\nu^2/c^2}$. Dit effect is dat van de zogenaamde *tijds-dilatatie*: **bewegende klokken lopen langzamer.**

Een aantal opmerkingen zijn hier toepasselijk. In de eerste plaats is het natuurlijk niet alleen W_1's klok die langzamer loopt (waargenomen door W_2): ook W_1's biologische en andere processen verlopen langzamer, zodat er geen sprake van is dat W_1's klok op een of andere manier zou gaan 'achterlopen' op W_1's eigen tijdsbesef. In de tweede plaats zet de tijdsdilatatie Newton's idee van absolute tijd op losse schroeven: net zo min als het universum een absoluut gedefinieerde ruimte bevat, bevat het enig 'intrinsiek', absoluut gedefinieerd tijdsverloop. In de derde plaats, ook W_2 kan zich voorzien van een tijdsklok, en beide waarnemers zullen concluderen dat de klok van de ander trager loopt dan die van haarzelf[30]!

2.7 De lichtsnelstheid

De tijdsklok kan ook gebruikt worden onder extremere omstandigheden. Zou W_1 er in slagen de lichtsnelheid te bereiken, dan zou W_2 waarnemen dat het lichtsignaal roerloos op de bodem van W_1's klok blijft liggen, aangezien elke opwaartse beweging een langere weg (en dus een langere tijd) zou vergen. Zoals vergelijking (**??**) al aangeeft, staan met lichtsnelheid bewegende klokken stil.

Een zelfs nog extremere situatie zou zich voordoen als W_1's snelheid, zoals gemeten door W_2, de lichtsnelheid c zou overtreffen. Indien dit het geval zou zijn, zou W_2 opmerken dat het lichtsignaal niet in staat is de klok bij te houden, en het zou dus roemloos 'uit de lichtklok vallen'. Dit is natuurlijk niet wat W_1 zou

[30]Enige paradox in deze situatie is maar schijn: men moet wel bedenken dat vertrek en aankomst van de lichtsignalen in de klokken plaatsvinden op *verschillende plaatsen* op het onderlinge traject van W_1 en W_2. Als W_1 en W_2 hun experiment zo regelen dat de aankomst van het signaal van bijvoorbeeld W_2's klok plaatsvindt op dezelfde positie als waarin het vertrok, zijn de waarnemers en hun klokken ten opzichte van elkaar in rust, en is er natuurlijk geen tijdsdilatatie.

zien: dit is een echte fysische tegenspraak waaruit we concluderen dat, in ieder geval voor materiële voorwerpen, snelheden groter dan c niet mogelijk zijn[31]. De lichtsnelheid kan dus met reden de *snelstheid* genoemd worden.

Terzijde moet hier worden opgemerkt dat vaak, en zoals wij zien abusievelijk, gemeend wordt dat klokken die sneller dan het licht bewegen *achteruit* zouden lopen: een romantische misvatting.

2.8 Experimentele waarneming van tijdsdilatatie

Hoezeer zij ook op het eerste gezicht ongerijmd lijkt, de tijdsdilatatie is desalniettemin een werkelijk fysisch effect. Door nauwkeurige klokken met hoge snelheid te laten reizen is ze wel degelijk aangetoond. Een wellicht nog sprekender voorbeeld is het geval van het muon-deeltje. Dit elementaire subatomaire deeltje heeft wel-begrepen eigenschappen, waaronder die van een eindige levensduur: wanneer, in versneller-experimenten, of in het vrije veld, een muondeeltje *in rust* geproduceerd wordt, zal het na een levensduur[32] van $t_m \approx 2\ 10^{-6}$ s verdwijnen, om plaats te maken voor (meestal) een drietal andere elementaire deeltjes[33]. In de bovenste lagen van de atmosfeer worden in grote getale muondeeltjes geproduceerd door de inslag van hoog-energetische kosmische straling in de moluceulen van de bovenste dampkring: dit gebeurt op enige tientallen kilometers boven het aardoppervlak. Deze muonen jagen dan met hoge snelheid naar beneden. Zelfs als zij met lichtsnelheid zouden reizen, zouden ze in een gemiddelde tijd t_m een afstand van $ct_m \approx 600$ meter kunnen afleggen om daarna te verdwijnen. Ondanks dit bereiken ze gemakkelijk het aardoppervlak, waar ze eenvoudig gedetecteerd kunnen worden[34].

De verklaring voor deze ogenschijnlijke ongerijmdheid ligt in de tijdsdilatatie: het muon vormt met zijn eindige levensduur een prachtig subatomair klokje[35], en het verloop van de 'eigen tijd' van het muon is vertraagd in ons coördinatensysteem. Aangezien muonen vaak een hoge snelheid hebben, zo'n 99.9% van c,

[31]Merk op dat vergelijking (**??**) voor $v > c$ geen reële oplossingen heeft, hetgeen op zich al een aanwijzing van ernstige problemen is.

[32]In feite is deze de *gemiddelde* levensduur.

[33]Een negatief geladen muon vervalt in een electron, een muon-neutrino en een electron-antineutrino; een positief geladen muon vervalt in een positron (dat wil zeggen een anti-electron), een electron-neutrino en een muon-antineutrino.

[34]Uw hele leven lang hagelt er gemiddeld een muon per seconde dwars door uw lichaam heen de grond in. Blijkbaar, en gelukkig, is het DNA van het aardse leven over het algemeen goed bestand tegen dit kosmische bombardement.

[35]Dat, inderdaad, precies één keer tikt!

halen ze het regelmatig tot aan het zeeniveau[36].

2.9 Wederom de Lorentzcontractie

Zoals in de vorige paragraaf besproken, kan een snel muon vanuit de bovenste atmosfeer het aardoppervlak bereiken dankzij de tijdsdilatatie-factor $1/\sqrt{1-v^2/c^2}$. Een tegen-argument hiervoor zou de volgende redenering kunnen zijn. Laat ons ons eens in het muon verplaatsen. Vanuit zichzelf gezien staat het stil, en is het het aardoppervlak met de daarop gelegen atmosfeer die met grote snelheid op het muon afstormt. Echter, het muon in rust ondergaat geen tijdsdilatatie! Hoe is het dan mogelijk dat de aarde snel genoeg het muon bereikt heeft?

Het enige antwoord hierop is de Lorentz-contractie au sérieux te nemen: de op het muon afstormende luchtlagen zijn door hun hoge snelheid gecontraheerd, en het muon neemt dus geen atmosfeer van tientallen kilometers dikte waar, maar een van slechts een paar honderd meter: en inderdaad heeft de Lorentz contractiefactor precies de juiste waarde. Dientengevolge zullen we ook in een met het muon meebewegend coördinatensysteem de juiste conclusie trekken.

We merken op dat in het geval van het muon-deeltje de *fysische geburtenissen* zelf ondubbelzinnig zijn: in beide systemen nemen we enerzijds de productie van het muon, en anderzijds het inslaan van het muon in de grond waar. De *verklaring* van het feit dat dit mogelijk is is in de twee coördinatensystemen echter verschillend – een in de fysica veel voorkomende situatie.

2.10 “Meanwhile, in a distant galaxy...”

Newton veronderstelde dat er een absolute tijd bestaat, waardoor het zinvol is te spreken over het gelijktijdig optreden van gebeurtenissen. Einstein zag deze veronderstelling als de crux van het verschil tussen de Newtoniaanse en de relativische mechanica: gelijktijdigheid van ruimtelijk gescheiden gebuertenissen is geen houdbaar begrip, zoals we zullen aantonen in het volgende gedachteexperiment (door Einstein zelf zo beschreven). Waarnemer W_1 zet zich gedurende een hevige onweersbui precies in het midden van een rijdende treinwagon, terwijl waarnemer W_2 langs de spoorbaan staat. Op het moment dat W_1 en W_2 elkaar passeren (en waarover zij het dus eens zijn), slaat de bliksem in op de twee uiteinden van de treinwagon. Waarnemer W_2 rapporteert hierover:

[36]Op grotere hoogte boven het aardoppervlak neemt het aantal muonen inderdaad flink toe.

"De blikseminslagen hadden plaats toen beide uiteinden van de treinwagon precies even ver bij mij vandaan waren. De lichtsignalen van beide inslagen reisden naar mij toe met dezelfde snelheid c*, en kwamen tegelijkertijd bij mij aan. Daarom vonden ook de blikseminslagen tegelijk plaats."*

Zij heeft natuurlijk gelijk. Voor waarnemer W_1 ligt de situatie anders. Tijdens de periode dat de lichtsignalen naar haar op weg waren is zijzelf ook verplaatst door de beweging van de trein. Het signaal van de voorkant van de wagon bereikt haar dus eerder dan dat van de achterkant, en zij rapporteert:

"De blikseminslagen hadden plaats aan beide uiteinden van de treinwagon, die de hele tijd precies even ver bij mij vandaan waren. De lichtsignalen van beide inslagen reisden naar mij toe met dezelfde snelheid c*. Het signaal van de voorkant van de wagon kwam eerder aan dan dat van de achterkant, een daarom sloeg de bliksem aan de voorkant eerder in dan aan de achterkant."*

Ook zij heeft gelijk – in *haar* coördinatensysteem[37]. Wij moeten concluderen dat gelijktijdigheid in de relativiteitstheorie *geen* zinvol begrip is. Zoals Einstein al voorvoelde, is het precies dit loslaten van gelijktijdigheid dat grote offers van onze intuïtie vergt.

Uiteraard is het boven beschreven effect bijzonder klein waar het niet-relativistische treinwagons betreft, en de Nederlandse Spoorwegen zullen geen relativistische correcties in de treinplanning behoeven. Als de ruimtelijke afstanden tussen twee gebeurtenissen echter zeer groot worden, is gelijktijdigheid wel degelijk problematisch. Wanneer ik U te voet op straat tegemoetloop terwijl U stil staat, ligt een gebeurtenis in een ver melkwegstelsel, die in Uw coördinatensysteem op *datzelfde* moment plaats heeft, voor mij nog maanden in de toekomst, terwijl het voor de wandelaar die U vanaf de andere kant nadert al even zovele maanden verleden tijd is!

[37] Merk hierbij op dat Lorentzcontractie van de treinwagon in dit geval geen rol speelt: het middelpunt van een gecontraheerde wagon is ook dat van een ongecontraheerde wagon. De lichtsignalen doen er, zoals gemeten door W_2, inderdaad korter over dan het geval zou zijn voor een stilstaande treinwagon, die immers langer zou zijn; maar waar het hier om gaat is dat de twee signalen er voor W_1 *niet even lang* over doen.

3 De tijdruimte van de relativiteitstheorie

We zullen de hiervoor geschetste, op het eerste gezicht tegen-intuïtieve effecten van de relativiteitstheorie een meer quantitatief en formeel karakter geven. De aanzet hiertoe is de constructie van de Minkowski-ruimte. Daarna zullen we de Lorentz-transformaties afleiden, en tenslotte de geëigende metriek van de Minkowski-ruimte postuleren.

3.1 De Minkowski-ruimte

In de Newtoniaanse mechanica worden ruimte en tijd op verschillende manieren behandeld. Immers, ruimtelijke afstanden en temporele afstanden worden gemeten in *verschillende eenheden*, en kunnen daarom niet zonder meer gecombineerd worden aangezien het resultaat zou afhangen van de keuze van eenheden in de beide grootheden[38]. De door Einstein gepostuleerde[39] universaliteit van de lichtsnelheid in vacuüm geeft ons *wel* een ondubbelzinnige manier om tijd en ruimte in elkaar uit te drukken: een lichtjaar is de afstand, door het licht in een jaar afgelegd[40]; een lichtmeter is de tijd, nodig om licht een afstand van een meter te laten afleggen[41]. Ook hebben we gezien dat, wanneer we beschrijvingen vergelijken die gegeven worden in inertiaalstelsels die ten opzichte van elkaar met grote snelheden bewegen, ruimte en tijd op onverwachte manieren lijken te vervloeien en vervormen. Het licht daarom voor de hand om een overstap te maken van de Newtoniaanse ruimte en Newtoniaanse tijd naar een geünificeerde arena voor de fysica, de zogenaamde *tijdruimte*. Deze wordt als volgt gedefinieerd. Laat, in een gegeven inertiaalstelsel, een gebeurtenis beschreven worden door een ruimtelijke vector $\vec{x}$ en een tijdstip t. We kunnen deze drie-dimensionele vector en het enkele getal combineren in een vierdimensionele *Minkowski-vector*, ook wel *vier-vector* genoemd, door t met c te vermenigvuldigen en aan de drie-vector toe te voegen.

[38] Laat ons aannemen dat ruimte en tijd zonder meer zouden worden opgeteld. In dat geval zou de som 'één jaar + één kilometer' het resultaat 2 hebben als we als tijdseenheid het jaar, en als afstandseenheid de kilometer zouden kiezen; als we echter tijdsintervallen in dagen, en afstanden in meters zouden meten zou het resultaat 1365 zijn.

[39] En, niet minder belangrijk, tot op heden door geen enkel experimenteel resultaat tegengesproken!

[40] Een lichtjaar is ongeveer 10^{16} kilometer.

[41] De 'tik' van de lichtklok, ongeveer 1/300000000-ste seconde.

Dit wordt meestal aangegeven door een Griekse index voor de vector te gebruiken:

$$\left.\begin{array}{l}\vec{x} = (x^1, x^2, x^3) \\ \\ t\end{array}\right\} \quad \to \quad x^\mu = (ct, x^1, x^2, x^3) \quad . \tag{56}$$

Het is conventioneel om de tijdscomponent ct aan te duiden met x^0. De zogenaamde Lorentz-index μ neemt dus de waarden $0, 1, 2, 3$ aan. In deze geünificeerde beschrijving ligt het voor de hand om, analoog aan het geval van kanteling van de coordinaat-assen van het inertiaalstelsel die tot govlg hebben dat de componenten van $\vec{x}$ zich onderling vermengen, er ook transformaties kunnen zijn die de x^0-component met andere mengen. Wij zullen deze hieronder nader bezien.

3.2 De Lorentz-transformaties

Zoals gezegd is het onmiddellijk duidelijk dat de Galilei-transformaties niet geldig kunnen zijn in het relativisitische régime, gezien de constantheid van c. Laat ons opnieuw de situatie bezien, die aanleiding geeft de coördinatentransformatie tussen twee inertiaalstelsels te bepalen. Wij stellen ons twee waarnemers W_1 en W_2 voor, elk uitgerust met een eigen inertiaalstelsel, S_1 resp. S_2. Voor het gemak nemen we aan dat de drie ruimtelijke assen van de inertiaalstelsels parallel aan elkaar zijn; dat de twee coördinatensystemen met eenparige snelheid ten opzichte van elkaar bewegen langs de onderlinge x-as; en dat de *tijden* in de twee systemen zo gedefinieerd zijn dat de oorsprongen O_1 en O_2 van de inertiaalstelsels samenvallen: dat wil zeggen dat beide waarnemers tijdstip nul definiëren als het moment waarop de ruimtelijke oorsprongen elkaar passeren. We geven de coördinaten die waarnemer 1 hanteert aan zonder, en die welke waarnemer 2 hanteert, mét een accent.

Laat nu de oorsprong O_2 bewgen met een snelheid v langs de positieve $x-as$ (dat wil zeggen, de positieve x-as van S_1). Dan zal de oorsprong O_1 met een snelheid $-v$ bewegen langs de x'-as (dat wil zeggen, met snelheid v langs de *negatieve* x-as van S_2. Dat dit zo is volgt uit de aanname dat er geen bevoorrecht inertiaalstelsel bestaat: immers, als de twee snelheden niet precies gelijk waren in absolute waarde, zou het mogelijk zijn onderscheid te maken, en (in principe) door middel van een competitie dat inertiaalstelsel te identificeren dat ten opzichte van alle andere 'het langzaamst' beweegt. Uit hetzelfe principe valt af te leiden dat afstanden in de y- en z-richting niet mogen veranderen onder de overgang van S_1 naar S_2. We komen tot de conclusie

$$y' = y \quad , \quad z' = z \quad . \tag{57}$$

We komen tot de volgende relatie tussen (ct) en x, en (ct') en x':

$$\begin{aligned} (ct') &= A\,(ct) + B\,x \ , \\ x' &= C\,(ct) + D\,x \ , \end{aligned} \tag{58}$$

waarbij de grootheden A, B, C en D te bepalen zijn: zij mogen niet van de viervectoren afhangen, maar uiteraard wel van v. Er moet nog worden opgemerkt dat deze transformatie *lineair* is. Dit is nodig omdat anders de notie van het samenstellen van vectoren niet meer zinvol zou zijn: wij eisen dat de getransformeerde van de samenstelling van twee vectoren gelijk is aan de samenstelling van de twee getransformeerde vectoren, en dit maakt de lineariteit onontkoombaar. We gebruiken nu de volgende informatie om de vier grootheden te vinden.

1. De oorsprong O_2 heeft $x' = 0$ voor W_2; in het systeem van W_1 echter geldt voor O_2 dat $x = vt$. Door dit in te vullen in de tweede vergelijking van (**??**) komen we op

$$0 = C(ct) + D(vt) \quad \Rightarrow \quad C = -\frac{v}{c}D \ . \tag{59}$$

2. De oorsprong O_1 heeft coordinaat $x = 0$ voor W_1: voor W_2 wordt de x'-coordinaat gegeven door $x' = -vt'$. Invullen in (**??**) geeft dan

$$(ct') = A(ct) \text{ en } -vt' = C(ct) \quad \Rightarrow \quad C = -\frac{v}{c}A \ , \tag{60}$$

   waaruit volgt dat $A = D$.

3. De lichtsnelheid moet in beide inertiaalstelsels gelijk zijn aan c. Laat op $t = t' = 0$ een lichtsignaal in de positieve x-richting worden uitgezonden. Als enige tijd later de waarnemers de positie van het lichtsignaal bepalen, bijvoorbeeld de positie van die gebeurtenis waarbij het lichtsignaal door een detector wordt geabsorbeerd, moet W_1 voor deze gebeurtenis vinden dat $x = (ct)$, en W_2 op haar beurt moet vinden $x' = (ct)'$. Wederom invullen in **??**) geeft nu

$$(ct') = (A + B)(ct) = (C + D)(ct) \ , \tag{61}$$

   waarmee nu vastligt dat $A = D$ en ook $B = C = -(v/c)A$. De vorm van de coördinatentransformatie wordt dus nu gegeven door

$$\begin{aligned} (ct') &= A\left((ct) - \frac{v}{c}x \right) \ , \\ x' &= A\left(-\frac{v}{c}(ct) + x \right) \ . \end{aligned} \tag{62}$$

4. We maken nu de aanname dat A alleen van $|\vec{v}|$ mag afhangen. Als dit niet het geval was zou dit betekenen dat de tijdruimte enige richtings-voorkeur bevat, in strijd met de aanname van isotropie. Laat ons nu een tweede coördinatentransformatie uitvoeren, die vanuit (ct') en x' de coördinaten geeft in een derde inertiaalstelsel dat beweegt met snelheid $-v$ langs de x'-as van S_2. Aangezien $A(|\vec{v}|) = A(|-\vec{v}|)$ moet dus voor dit derde inertiaalstelsel (met dubbel-geaccenteerde coordinaten) gelden:

$$(ct'') = A\left((ct') + \frac{v}{c}x' \right) \ . \tag{63}$$

Het $+$-teken komt uiteraard van de vervanging van v door $-v$. Drukken we dit nu weer uit in termen van (ct) en x, dan vinden we

$$\begin{aligned} (ct'') &= A\left(A\left((ct) - \frac{v}{c}x\right) + \frac{v}{c}A\left(-\frac{v}{c}(ct) + x\right)\right) \\ &= A^2\left(1 - \frac{v^2}{c^2}\right)(ct) \ . \end{aligned} \tag{64}$$

Nu merken we op dat, in feite, systeem S_3 niets anders is dan S_1! Daarom moeten t'' en t aan elkaar gelijk zijn, en we vinden

$$A = \frac{1}{\sqrt{1 - (v^2/c^2)}} \ . \tag{65}$$

Hiermee hebben de vorm van de coördinatentransformatie tussen de systemen S_1 en S_2 gevonden. Vaak worden de notaties

$$\gamma = \frac{1}{\sqrt{1 - (v^2/c^2)}} \quad , \quad \beta = \frac{v}{c} \tag{66}$$

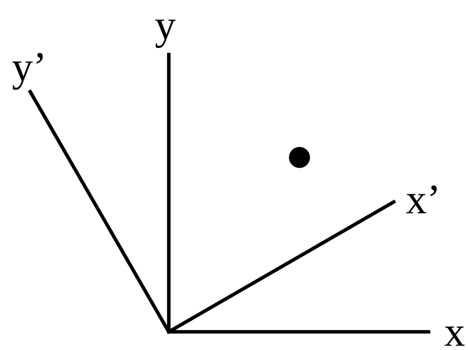
gebruikt, zodat deze coördinatentransformatie die de *Lorentz-transformatie* genoemd wordt, de volgende vorm krijgt:

$$\begin{aligned} (x')^0 &= \gamma x^0 - \beta\gamma x^1 \ , \\ (x')^1 &= -\beta\gamma x^0 + \gamma x^1 \ , \\ (x')^2 &= x^2 \ , \\ (x')^3 &= x^3 \ . \end{aligned} \tag{67}$$

Zoals eenvoudig valt te verifi"eren, is de uitdrukking van x^μ in x'^μ dezelfde mits β vervangen wordt door $-\beta$. Uiteraard zijn ingewikkelder transformaties mogelijk. Zo is het mogelijk de tijdstippen $t = 0$ en $t' = 0$ anders te kiezen, de onderlinge beweging van de inertiaalstelsels in een meer algemene richting te kiezen, of de coordinaat-assen van de twee systemen niet aan elkaar parallel te kiezen. De transformatie-formules worden daarmee ingewikkelder maar verschaffen geen nieuwe informatie: wij zullen het meer algemene geval dan ook niet behandelen. Wel merken we op dat in alle gevallen we moeten aannemen dat beide coördinatensystemen werkelijk inertiaalstelsels zijn, zodat hun ondxerlinge snelheid eenparig moet zijn[42]. In dat geval is iedere coördinatentransformatie tussen twee inertiaalstelsels samen te stellen uit een verschuiving van de oorsprong (zodat bij alle coordinaten een vast bedrag wordt opgeteld), een rotatie van de ruimtelijke assen, en een Lorentztransformatie.

3.3 De Lorentz-transformatie als pseudo-rotatie

. Laat ons ons twee stelsels van ruimtelijke coordinaat-assen voorstellen, die de oorsprong gemeen hebben, maar waarin de coordinaat-assen van het eerste systeem overgaan in die van het tweede systeem onder een rotatie over een hoek θ in het xy-vlak, zoals hieronder aangegeven.



De coördinaten van het aangegeven punt in de twee systemen zijn gerelateerd door de bekende formule

$$\begin{aligned} x' &= \cos(\theta)\, x + \sin(\theta)\, y \;, \\ y' &= -\sin(\theta)\, x + \cos(\theta)\, y \;. \end{aligned} \tag{68}$$

[42]Natuurlijk zijn nog algemenere relaties tussen coördinatensystemen mogelijk, bijvoorbeeld twee systemen die onderling *niet* eenparig bewegen: daarmee echter betreden wij het terrein van de *algemene* relativiteitstheorie.

Een vergelijking met de eerste twee relaties in (**??**) dringt zich hier op. We kunnen deze vergelijking enigzins formeel duidelijker maken door middel van het invoeren van hyperbolische trogonometrische functies: we mogen stellen dat er een hoek η bestaat zodanig dat

$$\gamma = \cosh(\eta) \quad , \quad \beta\gamma = \sinh(\eta) \ . \tag{69}$$

Deze hoek η wordt de *rapheid* genoemd[43]; ze wordt gegeven door

$$\eta = \frac{1}{2}\log\left(\frac{1+\beta}{1-\beta}\right) = \frac{1}{2}\log\left(\frac{c+\nu}{c-\nu}\right) \ . \tag{70}$$

De vergelijkingen (**??**) verkrijgen nu de vorm

$$\begin{aligned} (x')^0 &= \quad \cosh(\eta)\, x^0 - \sinh(\eta)\, x^1 \ , \\ (x')^1 &= -\sinh(\eta)\, x^0 + \cosh(\eta)\, x^1 \ , \end{aligned} \tag{71}$$

hetgeen de gelijkenis met de rotatie-formule pregnant maakt: een Lorentztransformatie is daarmee te zien als *een soort van* rotatie in een vlak opgespannen door de tijd-as en een van de ruimte-assen. Net zoals een rotatie over een hoek θ_1, gevolgd door nog een rotatie (in hetzelfde vlak!) over een hoek θ_2 kan worden gezien als een enkele rotatie over een hoek $\theta_1 + \theta_2$, is ook een Lorentztransformatie met een rapheid η_1 gevolgd door nog een tweede (in dezelfde richting!) met een rapheid η_2, gelijk aan een enkele Lorentztransformatie over een rapheid $\eta_1 + \eta_2$.

Er is echter een belangrijk verschil: waar een rotatie over 360 graden de ruimtelijke assen weer op hun oorspronkelijke posities brengt, zodat een rotatie over 360 graden hetzelfde is als in het geheel geen rotatie, geldt een dergelijke terugkeer niet voor Lorentztransformaties. In het wiskundig jargon wordt dit uitgedrukt door te zeggen dat het geheel van rotaties een *compacte* groep vormt, en het geheel van Lorentztransformaties een *niet-compacte* groep. De formele overeenkomst tussen rotaties en Lorentztransformaties moet ons niet de ogen doen sluiten voor het feit dat, hoewel c ons de mogelijkheid biedt tijd en ruimte op *formele* wijze tezamen behandelen, er fysisch wel degelijk verschil is tussen tijd en ruimte; de overeenkomsten en verschillen zijn eenvoudigweg anders (en subtieler) dan in de Newtoniaanse mechanica.

[43] In het Engels wordt hier gesproken van *rapidity*: de term is echter voor het eerst in het Nedxerlands gebruikt, en wel door Lorentz.

3.4 Samenstellen van snelheden

In de Newtoniaanse mechanica geldt, zoals we hebben gezien, de Galilei-regel voor het samenstellen van snelheden: als W_1 beweegt met snelheid v_1 ten opzichte van W_2, die zelf weer beweegt[44] met snelheid v_2 ten opzichte van waarnemer W_3, dane beweegt W_1 met snelheid v_1+v_2 ten opzichte van W_3. Laat ons nu deze situatie evalueren in de meer correcte relativistische beschrijving. De drie waarnemers hebben elk hun eigen tijds-coördinaat en x-coördinaat, die we zullen aangeven met corresponderende indices: ook zullen we aannemen dat de oorsprongen van de drie coördinatensystemen samenvallen. De ruimtelijke oorsprong van W_1's coördinatensysteem heeft natuurlijk $x_1 = 0$. Als W_1's klok tijd t_1 aanwijst, heeft haar ruimtelijke oorsprong de volgende coördinaten in het systeem van W_2:

$$ct_2 = \cosh(\eta_1)\ ct_1 \quad , \quad x_2 = \sinh(\eta_1)\ ct_1 \ \ , \tag{72}$$

waarbij, zoals boven, de rapheid en de snelheid v_1 (dat wil zeggen de snelheid van W_1 in het coördinatensysteem van W_2) gegeven zijn door

$$\eta_1 = \frac{1}{2}\log\left(\frac{c+v_1}{c-v_1}\right) \quad . \tag{73}$$

Om de snelheid van W_1 in het coördinatensysteem van W_2 te vinden moeten we nu de Lorentztransformatie uitvoeren die tijdruimte-coördinaten in het systeem van W_2 uitdrukt in die van het systeem van W_3. Hiertoe gebruiken we de rapheid

$$\eta_2 = \frac{1}{2}\log\left(\frac{c+v_2}{c-v_2}\right) \quad . \tag{74}$$

We vinden dus

$$\begin{aligned}
ct_3 &= \cosh(\eta_2)\ ct_2 \ + \ \sinh(\eta_2)\ x_2 \\
&= \Big(\cosh(\eta_1)\cosh(\eta_2) + \sinh(\eta_1)\sinh(\eta_2)\Big)\ ct_1 \\
&= \cosh(\eta_1+\eta_2)\ ct_1 \ \ , \\
x_3 &= \sinh(\eta_2)\ ct_2 \ + \ \cosh(\eta_2)\ x_2 \\
&= \Big(\cosh(\eta_1)\sinh(\eta_2) + \sinh(\eta_1)\cosh(\eta_2)\Big)\ ct_1 \\
&= \sinh(\eta_1+\eta_2)\ ct_1 \ \ .
\end{aligned} \tag{75}$$

[44] We nemen hier aan dat de snelheden v_1 en v_2 bedie parallel zijn aan de x-as: in het meer algemene geval is de regel natuurlijk ingewikkeolder.

We zien dus dat het in feite niet de snelheden zijn die worden opgeteld, maar de rapheden (zoals in de vorige paragraaf reeds vermeld): in het coördinatensysteem van W_3 beweegt het punt $x_1 = 0$ met een rapheid η die word gegeven door

$$\eta = \eta_1 + \eta_2 \quad . \tag{76}$$

Voor de *snelheid* v van het punt $x_1 = 0$ in het coördinatensysteem van W_3 geldt dus

$$\begin{aligned} \frac{1}{2} \log\left(\frac{c+v}{c-v}\right) &= \frac{1}{2} \log\left(\frac{c+v_1}{c-v_1}\right) + \frac{1}{2} \log\left(\frac{c+v_2}{c-v_2}\right) \\ &= \frac{1}{2} \log\left(\frac{(c+v_1)(c+v_2)}{(c-v_1)(c-v_2)}\right) \quad . \end{aligned} \tag{77}$$

Hieruit vinden we de relativistisch correcte formule voor het ‘optellen’ van snelheden:

$$v = \frac{v_1 + v_2}{1 + (v_1 v_2 / c^2)} \quad . \tag{78}$$

Als we de lichtsnelheid c oneindig groot nemen krijgen we het Galilei-resultaat $v = v_1 + v_2$, zodat deze regel bij niet-relativistische snelheden een uitstekende benadering blijft. Bij grote snelheden is het resultaat anders. Laat ons proberen de lichtsnelheid te bereiken door in een raket te stappen die (ten opzichte van de achterblijvers) beweegt met snelheiud $c/2$, en dan vanuit deze raket in een tweede raket plaats te nemen, die met snelheid $c/2$ ten opzichte van de eerste raket vertrekt. Ten opzichte van de achterblijvers hebben we dan een snelheid $4c/5$, dus minder dan c. Na nog een derde overstap hebben we een snelheid $13c/14$, en een vierde stap geeft een snelheid $40c/41$ ten opzichte van de achterblijvers. Nog steeds is c niet bereikt, terwijl de Galilei-regel voor ons al een snelheid $2c$ zou hebben voorspeld[45]! In feite kan geen enkele meer-traps poging ons de lichtsnelheid doen bereiken. Aan de andere kant, als we $v_2 = c$ invullen in vergelijking (**??**) is het resultaat altijd $v = c$, ongeacht de waarde van v_1. Dit wil zeggen dat een lichtsignaal afgevuurd vanuit een raket (dat dus snelheid c heeft ten opzichte van deze), ook voor de achterblijvers snelheid c heeft. De samenstellings-formule (**??**) is dus consistent met de universaliteit van de lichtsnelheid.

[45] De hierbij gebruikte berekeningen luiden $(1/2 + 1/2)/(1 + 1/2 \times 1/2) = 4/5$, $(1/2 + 4/5)/(1 + 1/2 \times 4/5) = 13/14$, en $(1/2 + 13/14)/(1 + 1/2 \times 13/14) = 40/41$.

3.5 De 'echte afstand' in de Minkowski-ruimte

Het concept van een afstand in de ruimte tussen twee punten, en in de tijd tussen twee momenten, speelt in de fysica een belangrijke rol. Ook dit concept moet worden aangepast aan de eisen van een Minkowski-tijdruimte waarin tijd en ruimte onder Lorentz-transformaties met elkaar worden vervlochten, net zoals ruimtelijke rotaties de ruimtelijke coördinaten onderling mengt. Laat ons twee gebeurtenissen, A en B bezien, met tijdruimte-coördinaten x_A, y_A, z_A, t_A en x_B, y_B, z_B, t_B. In de Newtoniaanse mechanica zijn er dan twee afstanden gegeven: de spatiële afstand (in vierkante meters),

$$r(AB)^2 = (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 \ , \tag{79}$$

en de temporele afstand (in vierkante seconden),

$$t(AB)^2 = (t_A - t_B)^2 \ . \tag{80}$$

Deze definities hebben, zoals reeds besproken, de aantrekkelijke eigenschap dat ze onveranderd blijven wanneer de ruimtelijke oorsprong of de temporele oorsprong anders worden gekozen, en ook is de ruimtelijke afstand invariant onder rotaties van de coördinaat-assen.

In de Minkowski-ruimte kunnen we de natuurconstante c gebruiken om temporele afstanden om te rekenen in vierkante meters, waarna zij met de spatiële afstanden gecombineerd kunnen worden. In de Minkowski-ruimte is er tussen twee gebeurtenissen A en B dus één enkele afstand gegeven, en wel

$$s(AB)^2 = c^2 t(AB)^2 - r(AB)^2 \ . \tag{81}$$

De op deze wijze gedefinieerde afstand heeft het voordeel dat zij ook invariant is onder Lorentz-transformaties, zoals we nu zullen laten zien. Voor het gemak nemen we voor A de oorsprong, en voor B het punt met coordinaten $x^\mu = (ct, x, y, z)$ Laten we nu een coördinatentransformatie uitvoeren naar een systeem met dezelfde oorsprong, maar dat beweegt met rapheid η langs de x-as. De gebeurtenis B heeft in dit nieuwe systeem de coordinaten $(x')^\mu = (ct', x', y', z')$, waarbij

$$ct' = \cosh(\eta)ct - \sinh(\eta)x \ , \ \ x' = -\sinh(\eta)ct + \cosh(\eta)x \ , \ \ y' = y \ , \ \ z' = z \ . \tag{82}$$

We vinden onmiddelijk dat

$$(ct')^2 - (x')^2 = \left(\cosh(\eta)^2 - \sinh(\eta)^2\right)((ct^2) - x^2) = (ct)^2 - x^2 \ , \tag{83}$$

zodat inderdaad $s(AB)^2$ invariant is onder de Lorentz-transformatie.

We moeten hier opmerken dat het min-teken in de definitie (**??**) noodzakelijk is om de invariantie onder Lorentz-transformaties te bewerkstelligen. De definitie $c^2t(AB)^2 + r(AB)^2$ zou wellicht æsthetisch aantrekkelijker zijn, maar *klaarblijkelijk* heeft de natuur niet hiervoor gekozen[46]. De afstand $s(AB)^2$ moet worden gezien als de 'echte afstand' tussen twee gebuertenissen, die de fysische correlaties tussen de gebuertenissen regeert.

Aan de andere kant heeft de afstand (**??**) tegen-intuïtieve eigenschappen. Zo is eenvoudig in te zien dat $s(AB)^2 = 0$ voor twee gebeurtenissen A en B die door een licht-signaal met elkaar verbonden kunnen worden – hetgeen echter *niet* betekent dat A en B samenvallen. Zelfs zijn er gebeurtenissen A en B waarvoor $s(AB)^2$ negatief is, en $s(AB)$ bijgevolge niet reëel!

3.6 Symmetrieën van ruimte, symmetrieën van de natuur

De eigenschappen van de Minkowski-ruimte liggen in feite geheel bevat in de afstands-definitie (**??**). Zoals we hebben gezien is de afstand $s(AB)^2$ invariant onder zowel ruimtelijke rotaties als onder Lorentz-transformaties, alsmede onder een verschuiving van de positie (in plaats en tijd) van de oorsprong. Zulke coördinatentransformaties die een gegeven afstandsmaat invariant laten, worden *symmetrieën van de ruimte* genoemd. Daarnaast kunnen we het *empirische feit* opmerken dat de natuurwetten[47] invariant zijn onder deze transformaties. Deze *symmetrieën van de natuur* kunnen gezien worden als een rechtvaardiging van de keuze van (**??**) als afstandsmaat.

Het is verleidelijk om uit het bovenstaande een algemeen principe af te leiden, dat zou inhouden dat elke symmetrie van de ruimte automatisch een symmetrie van de natuur inhoudt. Helaas is dit niet het geval! Er bestaan nog andere transformaties in de Minkowski-ruimte waaronder $s(AB)^2$ niet verandert. In de eerste plaats is er de *spiegeling*: als we alle ruimtelijke coördinaten vervangen door hun tegengestelde, verandert $r(AB)^2$ daardoor niet. Bekijk nu een systeem waarin een zekere mate van draaiing plaatsvindt: een voorbeeld wordt gegeven door veel elementaire deeltjes, die altijd een zekere hoeveelheid intrinsieke draaiing bevatten,

[46]Men zou eventueel de tijds-coördinaten t kunnen vermenigvuldigen met $i = \sqrt{-1}$, waardoor het relatieve min-teken in de definitie (**??**) zou komen te vervallen. Dit is echter een puur cosmetische ingreep die geen enkel voordeel oplevert in de formulering van fysische theorieën. In de moderne literatuur wordt dit dan ook vrijwel niet meer gedaan.

[47]Met name geldt dit voor de natuurwetten op het meest fundamentele niveau, dat van de elementaire deeltjes.

de zogenaamde *spin*. Onder spiegeling wordt de draairichting omgekeerd: als spiegeling, naast een symmetrie van dce ruimte, ook een symmetrie van de natuur was, zou dit inhouden dat alle natuurwetten linksom- en rechtsom-draaiende deeltjes op gelijke wijze beïnvloeden. Voor de wetten van het electromagnetisme, en ook die van de sterke kernkrachten, geldt dit inderdaad: linksom- en rechtsom-draaiende electronen nemen op gelijke wijze aan electromagnetische interacties deel. Voor de zogenaamde *zwakke* kracht, verantwoordelijk voor radio-activiteit (en het branden van de zon!) geldt dit echter *niet*. Voor het zogenaamde neutrino-deeltje, dat noch electromagnetische, noch sterke interacties, maar alleen zwakke interacties kent[48], is deze asymmetrie zelfs extreem: voor zover bekend, *bestaan* linksom-draaiende neutrino's in het geheel niet! De spiegel-symmetrie is dus *geen* symmetrie van de natuur[49].

Een andere symmetrie van de Minkowski-ruimte is de *tijdsomkeer*: als we alle tijd-coördinaten laten verkeren in hun tegendeel, blijft $t(AB)^2$, en daarmee ook $r(AB)^2$, invariant. Op elementair niveau blijkt ook deze symmetrie niet door de natuur gerespecteerd, alhoewel het niet eenvoudig is dit experimenteel aan te tonen omdat de effecten van deze 'tijdsomkeer-schending' klein en subtiel zijn. Ook tijdsomkeer is geen symmetrie van de natuur.

Laat ons deze paragraaf afsluiten met een quasi-filosofische opmerking. De rotaties en Lorentz-transformaties, waaronder de natuur symmetrisch is[50], onderscheiden zich van de spiegeling en tijdsomkeer op twee manieren. In de eerste plaats bestaan er rotaties over willekeurig kleine hoeken, en Lorentz-transformaties met willekeurig kleine snelheden. Zulke transformaties onderscheiden zich slechts inifinitesimaal van in het geheel geen transformatie[51]. Daarentegen bestaat er niet zoiets als een 'infinitesimale tijdsomkeer': tijdsomkeer en spiegeling zijn alles-of-niets procedures. Er bestaat in de drie-dimensionele ruimte *geen* combinatie van rotaties die er toe leidt dat alle cooördinaat-assen omgekeerd worden. In de tweede plaats zijn tijdsomkeer, waaronder de tijd zelf van richting verandert, en spiegeling, waardoor we als het ware 'through the looking-glass' stappen, naar het

[48]We laten de – op het niveau van elementaire deeltjes verwaarloosbare – zwaartekracht hier buiten beschouwing.

[49]De ontdekking van deze asymmetrie in de zwakke interacties kwam voor vele natuurkundigen als een grote schok: zo natuurlijk komt het ons voor dat de natuur 'zo symmetrisch mogelijk' moet zijn.

[50]Dit standpunt kan natuurlijk altijd door experiment ontkracht worden, maar anno Domini 2005 is hiervoor geen enkele aanwijzing.

[51]In het jargon heet het dat deze transformaties 'continu met de eenheidstransformatie verbonden zijn'.

schijnt geen transformaties die wij fysiek kunnen uitvoeren. De natuur lijkt ons hier een boodschap te geven: symmetrieën van de ruimte die geen symmetrieën van de natuur zijn, zijn voor ons 'onbereikbaar'.

3.7 Causaliteit en de lichtkegel

Laat ons twee gebeurtenissen A en B bezien. Er zijn drie mogelijke gevallen te onderscheiden:

- $s(AB)^2 > 0$. Dit betekent $r(AB)^2 < c^2t(AB)^2$, hetgeen inhoudt dat een signaal dat eenparig en rechtlijnig tussen A en B beweegt dit moet doen met een snelheid kleiner dan c. Laat ons aannemen dat $t_B > t_A$ in een zeker inertiaalstelsel. Het kan eenvoudig gechecked worden dat in dat geval B *later* plaatsvindt dan A in *elk* inertiaalstelsel. We kunnen dus ondubbelzinnig zeggen: B komt *na* A. Het is dan ook mogelijk een inertiaalstelsel te vinden waarin $r(AB)^2 = 0$, zodat de afstand tussen A en B zuiver temporeel is.

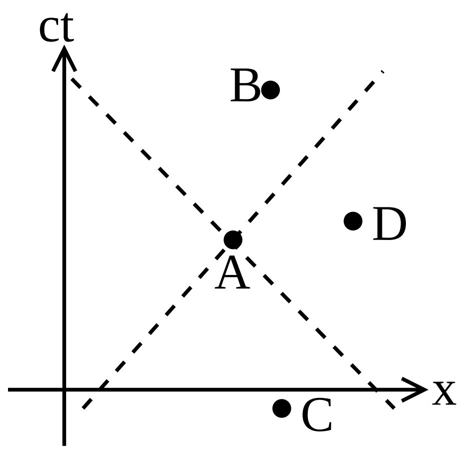
- $s(AB)^2 < 0$. Dit betekent dat $r(AB)^2 > c^2t(AB)^2$, zodat enig signaal tussen A en B een snelheid groter dan c zou moeten hebben. Nu is er een inertiaalstelsel te vinden waarin $t(AB)^2 = 0$: in dit systeem vinden A en B tegelijkertijd plaats, op verschillende locaties. Nu is het eenvoudig een inertiaalstelsel te vinden waarin $t_B > t_A$: in dit systeem komt B *na* A. Anderzijds is het even eenvoudig een inertiaalstelsel te vinden waarin $t_B < t_A$: in dit tweede systeem komt B *voor* A. De vraag welke van de twee gebeurtenissen het eerst optrad is dus niet ondubbelzinnig te beantwoorden (zoals we al eerder gezien hadden).

- $s(AB)^2 = 0$. In dit geval kan er een lichtsignaal lopen tussen A en B. Laat ons ons voorstellen dat een waarnemer W met lichtsnelheid beweegt tussen A en B. Vanuit elk ander systeem dan dat van W zelf staat W's klok stil. W zal dus concluderen dat A en B (en alle andere gebeurtenissen op haar traject) alle gelijktijdig plaatsvinden, en op dezelfde locatie (namelijk, die van W zelf).

We kunnen deze overwegingen gebruiken in verband met het begrip *causaliteit*: dit is het principe dat *een gevolg ná zijn oorzaak komt*[52]. We zien dat gebeurtenis

[52]Ook dit is in feite een natuurkundig vooroordeel. Aristoteles nam wel degelijk het bestaan van een *causa finalis* aan, maar in het moderne wereldbeeld is het niet erg duidelijk hoe een dergelijke doel-oorzaak te identificeren zou moeten zijn zonder aan levenloze systemen een 'aandrang' of een 'doelgerichtheid' toe te kennen.

A alleen dan een oorzaak kan zijn van gebeurtenis B als $s(AB)^2$ niet negatief is, aangezien alleen dan ondubbelzinnig is vast te stellen of A voor B optreedt. Met andere woorden, A kan alleen een fysische invloed uitoefenen op B als een signaal van A naar B met snelheid c of kleiner mogelijk is.



We komen op deze manier tot een indeling van de Minkowski-ruimte. Deze is schematisch gegeven in de afbeelding. Vanuit gebeurtenis A kunnen we in het (x, ct) vlak lijnen tekenen die overeenkomen met de positie in plaats en tijd van lichtsignalen uitgezonden vanuit A (de omhoog wijzende stippellijnen) en lijnen die de posities van lichtsignalen aangeven die in A aankomen (de omlaag wijzende stippellijnen). Gebeurtenis B kan vanuit A bereikt worden met snelheid kleiner dan c, en daarom *kan* A invloed op B uitoefenen. Vanuit gebeurtenis C kan een signaal met snelheid kleiner dan c bij A aankomen, en daarom *kan* A invloed ondervinden van C. We zeggen dat B binnen de *voorwaartse lichtkegel* van A valt, en dat C binnen de *achterwaardse lichtkegel* van A valt. Gebeurtenis D kan niet met A verbonden worden door een signaal met snelheid c of minder, en daarom kunnen A en D geen fysische invloed op elkaar hebben. In het getekende schema is het tijdstip van D later dan dat van A, maar zoals gezegd kan met ditzelfde schema tekenen in een ander inertiaalstelsel, waarin het tijdstip van D voor dat van A valt. In elk inertiaalstelsel echter bevindt B zich binnen de voorwaartse, en D zich binnen de achterwaartse, lichtkegel van A. We zien dat, ondanks de 'relativiteit van gelijktijdigheid', de begrippen 'oorzaak' en 'gevolg' nog steeds opgeld doen, echter niet overal in de Minkowski ruimte[53].

[53]Merk op dat D zich bevindt binnen de voorwaartse lichtkegel van C, en binnen de achterwaartse lichtkegel van B; daarom kan C invloed uitoefenen op B zowel via A als via D: de gebeurtenissen A en D zij echter causaal onafhankelijk van elkaar.

In eenvoudige termen kunnen we zeggen dat de voorwaartse lichtkegel van A de 'toekomst' van A is; dat de achterwaartse lichtkegel het 'verleden' van A is; en dat het overige gebied het 'elders' van A is.

3.8 De metriek van de Minkowski-ruimte

Laat ons ons twee gebeurtenissen voorstellen, dat wil zeggen twee punten A en B in de tijdruimte, die zeer dicht in elkaars buurt liggen. We zullen schrijven:

$$t_B = t_A + dt \ , \ \ x_B = x_A + dx \ , \ \ y_B = y_A + dy \ , \ \ z_B = z_A + dz \ ; \tag{84}$$

hierbij worden de coördinaten-separaties dt, dx, dy en dz als infinitesimaal klein voorgesteld. Zoals al boven besproken, is de 'echte' afstand tussen A en B in de tijdruimte dan

$$(ds)^2 = c^2(dt)^2 - \Big((dx)^2 + (dy)^2 + (dz)^2\Big) \quad . \tag{85}$$

We kunnen deze afstand aangeven op een manier die zich leent voor verdere generalizatie (die ook werkelijk toegpast wordt, in de *algemene* relativiteitstheorie): als we de Lorentz-indices μ invoeren wordt de coördinaten-separatie

$$dx^\mu \quad : \quad dx^0 = c\,dt \ , \ dx^1 = dx \ , \ dx^2 = dy \ , \ dx^3 = dz \ , \tag{86}$$

en we kunnen de 'echte' afstand dan schrijven als volgt:

$$(ds)^2 = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \ g_{\mu\nu} \, dx^\mu \, dx^\nu \ , \tag{87}$$

waarbij het object $g_{\mu\nu}$ gedefinieerd is als volgt:

$$g_{\mu\nu} = \begin{cases} 1 & \text{indien} \quad \mu = \nu = 0 \\ -1 & \text{indien} \quad \mu = \nu \text{ gelijk aan 1,2, of 3} \\ 0 & \text{indien} \quad \mu \neq \nu \end{cases} \tag{88}$$

Ook kunnen we, voor elke vier-vector A^μ, het object A_μ invoeren, door te definiëren

$$A_\mu = \sum_{\nu=0}^{3} \ g_{\mu\nu} \, A^\nu \ , \tag{89}$$

waarop we ook kunnen schrijven

$$(ds)^2 = \sum_{\mu=0}^{3} dx_\mu \, dx^\mu \quad . \tag{90}$$

Merk op dat

$$dx_0 = dx^0 \quad , \quad dx_k = -dx^k \;\; (k = 1, 2, 3) \quad . \tag{91}$$

Overigens is het een gebruikelijke conventie om, wanneer in een uitdrukking een Lorentz-index zowel als boven- als als beneden-index voorkomt, automatisch te impliceren dat over die index gesommeerd wordt. Merk op dat daarmee de *naam* van de Lorentz index onbelangrijk wordt, immers over alle waarden die zij kan aannemen wordt gesommeerd. Met deze *Eintein conventie* kunnen we dus schrijven

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu = dx_\mu dx^\mu = dx_\alpha dx^\alpha = dx^\beta dx_\beta \quad . \tag{92}$$

Het moge duidelijk zijn dat het in feite het object $g_{\mu\nu}$ is dat de vorm van de ‘echte’ afstand tussen twee nabijgelegen punten bepaalt. Dit object wordt de *metriek* genoemd, in dit geval de metriek van de Minkowski-ruimte.

4 Relativistische mechanica

We zullen nu onderzoeken hoe de vertrouwde Newtoniaanse mechanica aangepast moet worden om in het reine te komen met de eisen van de relativiteitstheorie.

4.1 De eis van covariantie

Een van de primaire eisen die gesteld moeten worden aan potentiële kandidaten voor de begrippen van impuls en energie is dat zulke objecten op een begrijpelijke manier veranderen onder de symmetrieën van de Minkowski ruimte, dat wil zeggen hun gedrag onder rotaties en Lorentz transformaties moet eenvoudig zijn. Om die reden zullen wij eisen dat de begrippen van snelheid, impuls en energie de vorm hebben van vier-vectoren, net zoals de vier-vector x^μ die de separatie tussen de oorsprong en een zekere gebeurtenis aangeeft. Deze eis, dat de mechanische grootheden transformeren *op gelijke wijze* als vier-vectoren, wordt de eis van *covariantie* genoemd[54]

Een voorbeeld van het nut van de eis van covariantie is het volgende. Veronderstellen we dat een covariante definitie van impuls is gegeven. In één bepaald coördinatensysteem zal dan de wet van impulsbehoud de vorm hebben:

$$(p_{\text{voor}})^\mu = (p_{\text{na}})^\mu \quad . \tag{93}$$

Laat ons nu transformeren naar een ander inertiaalstelsel. Aangezien de impulsviervector op dezelfde eenvoudige wijze transformeert als elke andere viervector, is op voorhand duidelijk dat ook in het nieuwe coördinatensysteem zal gelden:

$$(p'_{\text{voor}})^\mu = (p'_{\text{na}})^\mu \quad . \tag{94}$$

Hiermee is, zoals in de Newtoniaanse mechanica, niet aangetoond dat impulsbehoud geldt: maar wel, dat *als* dit behoud geldt in enig inertiaalstelsel, het ook (en automatisch) geldt in ieder ander inertiaalstelsel dat door een symmetrie van de Minkowski-ruimte uit het eerste verkregen kan worden.

4.2 De eigentijd van een puntdeeltje

De mechanica houdt zich bezig met de verandering van grootheden zoals positie, impuls etcetera in de loop van de tijd. Daartoe zullen we tijds-afgeleiden moeten

[54]In feite heeft het begrip ‘covariantie’ een meer precieze betekenis, en bestaat er ook een ‘contravariantie’. Voor onze doeleinden is dit meer verfijnde begrip niet nodig.

kunnen definiëren, dat wil zeggen de infinitesimale verandering van een grootheid gedurende een infinitesimaal tijdsinterval. Nu hebben we gezien dat voor waarnemers die met verschillende snelheden bewegen, de tijd met verschillende snelheden verstrijkt: welke definitie van tijd moeten we dan aannemen? Gelukkig bestaat op deze vraag wel degelijk een ondubbelzinnig antwoord. Laat ons ons voorstellen dat met een puntdeeltje, waarvan wij de mechanische eigenschappen willen beschrijven, een klok meebeweegt: deze klok en het ountdeeltje zijn dan op ieder moment in rust ten opzichte van elkaar. Iedere waarnemer die het ountdeeltje op enig moment op zekere plaats waarneemt, zal dan ook de stand van de meebewegende klok kunnen registreren: *en over de tijd die deze klok aanwijst, moeten alle waarnemers het noodzakelijkerwijs eens zijn* (niettegenstaande het feit dat de klokken van de waarnemers zelf niet noodzakelijkerwijs gelijk lopen met de meebewegende klok). Deze tijd wordt de *eigentijd* van het puntdeeltje genoemd[55]. Wij zullen dit nu nader preciseren. Laat ons aannemen dat in ons inertiaalstelsel het puntdeeltje beweegt met snelheid $\vec{v}$: dit wil zeggen dat in een tijdsinterval dt, zoals door onze klok gemeten, het deeltje een afstand $d\vec{x} = dt\ \vec{v}$ aflegt, in *ons* coördinatensysteem. De 'echte' afstand die het deeltje heeft afgelegd is uiteraard

$$(ds)^2 = c^2(dt)^2 - |d\vec{x}|^2 = (c^2 - |\vec{v}|^2)\ (dt)^2 \quad . \tag{95}$$

In het systeem waarin het deeltje in rust is, is dezelfde afstand afgelegd, maar *alleen* in temporele richting, door het verstrijken van de eigentijd, die we aangeven met τ. Daarom is

$$(ds)^2 = c^2(d\tau)^2 \quad . \tag{96}$$

Dit geeft ons de relatie tussen dt en $d\tau$:

$$d\tau = dr\ \sqrt{1 - |\vec{v}|^2/c^2} \quad , \tag{97}$$

hetgeen niets anders is dan de reeds gevonden formule voor tijdsdilatatie. We concluderen dat indien het deeltje voor twee verschillende waarnemers snelheid $\vec{v}_1$ en snelheid $\vec{v}_2$ heeft, de verstreken tijd dt_1 en dt_2 voor de twee waarnemers niet noodzakelijkerwijze gelijk zijn: maar de waarde van de verstreken *eigen*tijd $d\tau$ is voor beide waarnemers wel degelijk hetzelfde.

4.3 De vier-snelheid

Laat, in een gegeven inertiaalstelsel, een puntdeeltje snelheid $\vec{v}$ hebben. In een tijdsinterval dt legt het deeltje een ruimtelijk interval $\vec{v}\ dt$ af: in totaal vindt er

[55] Merk op dat het natuurlijk niet nodig is dat ieder deeltje *werkelijk* vergezeld gaat van een klok: hier gaat het om het principe.

dus een verplaatsing *in de Minkowski-ruimte* plaat ter grootte

$$dx^\mu = (c\,dt, v^1 dt, v^2 dt, v^3 dt) = (c, v^1, v^2, v^3)\ dt \quad , \tag{98}$$

waar v^k $(k = 1, 2, 3)$ de drie componenten van de snelheid zijn. Zouden we de relativistische versie van de snelheid, de *vier-snelheid*, definiëren als

$$u^\mu = \frac{dx^\mu}{dt} \quad ,$$

dan zou deze geen viervector zijn aangezien dt onder een coördinatentransformatie ook zou veranderen. Een betere definitie is voorhanden middels de eigentijd:

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{1}{\sqrt{1 - |\vec{v}|^2/c^2}}\frac{dx^\mu}{dt} = \frac{1}{\sqrt{1 - |\vec{v}|^2/c^2}}(c, v^1, v^2, v^3) \quad . \tag{99}$$

Per constructie is deze snelheid *wel* een zich goed gedragende vier-vector.

Een aantrekkelijk facet van deze definitie is dat het relativistisch kwadraat van de viersnelheid vast ligt:

$$u^\mu\, u_\mu = \frac{1}{1 - |\vec{v}|^2/c^2}\left(c^2 - |\vec{v}|^2\right) = c^2 \quad . \tag{100}$$

In zekere zin bewgen alle deeltjes met dezelfde ‘absolute waarde’ van de snelheid!

4.4 De vier-impuls

De relativistische definitie van een acceptabele maat voor hoeveelheid beweging wordt gegeven door directe generalizatie van de Newtoniaanse definitie van impuls: naar analogie van $\vec{p} = m\vec{v}$ schrijven we

$$p^\mu = m\, u^\mu \quad . \tag{101}$$

Ook deze is automatisch een goede vier-vector. Wij hebben, in termen van expliciete componenten,

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - |\vec{v}|^2/c^2}} \quad , \tag{102}$$

$$p^0 = \frac{mc}{\sqrt{1 - |\vec{v}|^2/c^2}} \quad . \tag{103}$$

Eén vraag dringt zich nu aan ons op. De impuls heeft, in de Newtoniaanse mechanica, drie componenten. Wat dan is de betekenis van de component p^0, die immers ook een hoeveelheid beweging aangeeft? Het antwoord ligt voor de hand: p^0 moet gerelateerd zijn aan de relativistische formulering van kinetische energie! Passen we de dimensie aan, dan vinden we dat we kunnen schrijven

$$E = cp^0 = \frac{mc^2}{\sqrt{1 - |\vec{v}|^2/c^2}} \quad . \tag{104}$$

We zien dat energie en impuls in de relativiteitstheorie bijzonder veel inniger verbonden zijn dan in de Newtoniaanse mechanica: zij treden tevoorschijn als de vier componenten van een enkel object, de vier-impuls.

Uit het resultaat $u^\mu u_\mu = c^2$ leiden we onmiddelijk af dat

$$p^\mu p_\mu = m^2 c^2 \quad , \tag{105}$$

en dit bedrag is hetzelfde in alle coördinatensystemen. Een laatste opmerking in dit verband: de massa van een deeltje wordt geacht een intrinsieke eigenschap van het deeltje te zijn, dat onafhankelijk moet zijn van het gekozen inertiaalstelsel, en daarmee onafhankelijk van de waargenomen snelheid van het deeltje. In het verleden werd wel gewerkt met snelheids-afhankelijke massa: men definieerde

$$m(\vec{v}) = \frac{m}{\sqrt{1 - |\vec{v}|^2/c^2}} \quad ,$$

zodat de impuls de vorm

$$\vec{p} = m(\vec{v})\, \vec{v}$$

kreeg. Tegenwoordig wordt deze beschrijving doorzien voor wat zij is: een kosmetische poging de Newtoniaanse formulering aan een inherent niet-Newtoniaanse werkelijkheid op te dringen.

4.5 Relatie met de Newtoniaanse grootheden: $E = mc^2$

Uiteraard moeten de relativistische grootheden van hoeveelheid beweging een aanwijsbare relatie met hun Newtoniaanse versies hebben. Laat ons daartoe de limiet bezien waarin c bijzonder groot is ten opzichte van v. Voor de impuls $\vec{p}$ vinden we dan onmiddelijk[56]

$$\vec{p} = m\vec{v} + \mathcal{O}(|\vec{v}|^3/c^2) \quad , \tag{106}$$

[56]Het hier gebruikte symbool $\mathcal{O}$ duidt de grootte van de verwaarloosde termen aan.

hetgeen de Newtoniaanse uitdrukking is. Voor de energie ligt het minder eenvoudig: we kunnen schrijven

$$E^2 = c^2(p^0)^2 = c^2\left(p^\mu p_\mu + |\vec{p}|^2\right) = c^2|\vec{p}|^2 + m^2c^4 \quad , \tag{107}$$

zodat we, voor $|\vec{v}| \ll c$, komen tot

$$E^2 \approx m^2c^4 + m^2c^2|\vec{v}|^2 \approx \left(mc^2 + \frac{1}{2}m|\vec{v}|^2\right)^2 \quad . \tag{108}$$

We mogen dus schrijven

$$E \approx mc^2 + \frac{1}{2}m|\vec{v}|^2 + \mathcal{O}(|\vec{v}|^4/c^2) \quad . \tag{109}$$

De tweede term in deze uitdrukking is de bekende Newtoniaanse uitdrukking voor de kinetische energie; maar de eerste term heeft geen tegenhanger in de Newtoniaanse mechanica[57]. Desalniettemin is deze term noodzakelijk om een covariante vier-impuls te krijgen, en Einstein poneerde dan ook: **ieder materieel voorwerp met massa m in rust bevat een kinetische energie** mc^2. Dit is de beroemde 'massa-energie equivalentie' welker formule $E = mc^2$ een icoon van de twintigste-eeuwse natuurwetenschap is geworden[58]. In de Newtoniaanse mechanica geldt behoud van massa, en daarmee blijft de totale rust-energie van een systeem onveranderd: dit verklaart waarom de rust-energie in de Newtoniaanse mechanica geen rol speelt.

4.6 Feiten en misvattingen rond $E = mc^2$

Een aantal populaire misvattingen rond de beroemde formule verdienen hier kort aandacht.

[57]Een dergelijke tegenhanger zou buiten het blikveld van de Newtoniaanse mechanica vallen, aangezien de lichtsnelheid c daar geen bijzondere betekenis heeft.

[58]Deze roem is waarschinlijk eerder te danken aan de compactheid van de formule dan aan enig algemeen begrip bij het publiek! Overigens dient te worden opgemerkt dat, aangezien de eenheid van energie (J) het product van massa (kg) en vierkante snelheid (m^2/s^2) is, het loutere *bestaan* van een natuurconstante met de dimensie van snelheid reeds de *mogelijkheid* van een dergelijke relatie impliceert: uiteraard betekent dit niet dat $E = mc^2$ op grond daarvan automatisch correct moet zijn.

- $E = mc^2$ betekent *niet* dat materie en energie 'hetzelfde' zijn! Immers, materie is een substantie, en energie een eigenschap, begrippen met een verschillende epistemologische en ontologische status. Zij wil slechts zeggen dat aan de massa, een eigenschap van een zekere hoeveelheid materie, een andere eigenschap, namelijk energie gerelateerd is.

- $E = mc^2$ is niet een relatie die alleen van belang is bij kernreacties! *Elke* reactie waarbij energie vrijkomt (zoals het afsteken van een lucifer, de verwerking van voedsel tijdens de spijsvertering, etcetera) gaan gepaard met massa-verlies; en omgekeerd: het opwarmen van een pannetje water maakt het water zwaarder. Uiteraard is, in het dagelijks leven, de massa-verandering miniem omdat c naar de maatstaven van huis, tuin en keuken zo groot is.

- In de praktijk is het niet (of nauwelijks) mogelijk een aanmerkelijk deel van massa in energie om te zetten. Dit ligt aan het feit dat voor materiële voorwerpen talrijke behoudswetten gelden: zo is het niet mogelijk een electron in energie om te zetten omdat de electrische lading van het electron behouden moet blijven. Ook voor protonen en neutronen gelden dergelijke behoudswetten. Het best haalbare is, onder normale omstandigheden, het herschikken van de bestanddelen van materie, waarbij zogenaamde bindingsenergie kan vrijkomen[59]; dit uiteraard gepaard gaand met klein massaverlies. Opmerkingen als 'iedere gram materie komt overeen met een 22-kiloton bom' zijn dan ook op hun, louter rhetorische, waarde te schatten.

- 'Energie' kan niet 'vrij' voorkomen, maar moet altijd gedragen worden door deeltjes. Eventueel kunnen dit fotonen zijn, die aangezien zij geen rustmassa hebben, de meest efficiënte dragers van energie zijn. Een dergelijk proces treedt op in de wederzijdse annihilatie van materie en anti-materie, waarover later meer.

- De energie in $E = mc^2$ is wel degelijk *bewegings*-energie, ondanks haar benaming van rust-energie. Immers, ieder voorwerp dat ruimtelijk in rust is, beweegt nog steeds temporeel, en wel met snelheid c — en gelijk op met alle andere voorwerpen die in rust zijn.

Tot slot valt een aardige illustratie van de formule op te merken. Het valt eenvoudig te meten dat de zon op elke vierkante meter aardoppervlak waar zij lood-

[59] Hierbij valt de denken aan bijvoorbeeld de verbranding van waterstof: $2H_2 + O_2 \rightarrow 2H_2O$.

recht boven staat een hoeveelheid vermogen instraalt van ongeveer anderhalf kilowatt: dat wil zeggen 1500 Joule per seconde. Uiteraard straalt de zon dit vermogen in alle richtingen uit. Aangezien de aarde zich op 150 miljoen kilometer van de zon bevindt, straalt de zon aan licht een vermogen uit dat overeenkomt met een massaverlies van enige miljoenen tonnen massa per seconde[60]! Gelukkig is onze zon wel zo zwaar dat het vele malen de huidige leeftijd van ons universum zal duren voordat de zon 'opgesoupeerd' zal zijn[61].

4.7 Voorwerpen met lichtsnelheid

Zoals reeds eerder aangegeven, stelt een eventuele sneller-dan-licht beweging ons voor conceptuele problemen, met name vanwege de causaliteit. We merken nu op dat, indien de snelheid van een voorwerp die van het licht benadert, de energie als functie van de snelheid zeer snel toeneemt: in de limiet $|\vec{v}| \to c$ gaat de energie zelfs naar oneindig. De conclusie is dat er een oneindige hoeveelheid energie nodig is om een voorwerp met $m > 0$ tot de lichtsnelheid te versnellen: toevoeren van extra energie heeft een steeds kleinere snelheids-toename tot gevolg[62].

Een andere situatie is die van een object dat 'van zichzelf' reeds de lichtsnelheid heeft. In dat geval kunnen we de energie-vergelijking (**??**) toepassen in de vorm

$$m^2c^4 = E^2\left(1 - \frac{|\vec{v}|^2}{c^2}\right) \quad . \tag{110}$$

Wij zien dat $|\vec{v}| = c$ wel degelijk mogelijk is, *mits* $m = 0$. Anderzijds geldt ook dat als $m = 0$, de snelheid c moet zijn om enige energie mogelijk te maken. Wij concluderen, tweeledig:

- Massaloze deeltjes bewegen met snelheid c (in ieder coördinatensysteem!)
- Elk deeltje dat met lichtsnelheid beweegt is massaloos.

[60]Het echte verlies is groter, aangezien de zon ook energie verliest in de vorm van bijvoorbeeld neutrino-deeltjes.

[61]Reeds lang vóór die tijd zal de zon uitgedoofd zijn wegens andere processen in haar evolutie.

[62]Het verschil tussen de snelheid en c kan natuurlijk willekeurig klein worden. Als voorbeeld zij op te merken dat in de LEP-deeltjesversneller die van 1989 tot 2001 op CERN te Genève actief geweest is, er electronen werden versneld tot een energie van 10^{11} electronvolt, hetgeen ongeveer 100000 maal hun rust-energie is: zullke electronen hebben een snelheid die slechts een paar *centimeter* per seconde van die van het licht verschilt.

Een voorbeeld van voorwerpen die met lichtsnelheid reizen, is uiteraard het licht zelf! Inderdaad bestaat licht[63] uit massaloze deeltjes, de fotonen.

Men zou zich kunnen afvragen waarin twee fotonen van elkaar kunnen verschillen, aangezien zij altijd dezelfde massa (0) en dezelfde snelheid (c) hebben. Echter, twee fotonen kunnen wel degelijk een verschillende *energie* hebben. Het is interessant op te merken dat niet de relativiteitstheorie, maar de *quantummechanica* dit verschil een interpretatie geeft: fotonen van verschillende energie hebben een verschillende *kleur*! Een foton met een energie van 2 electronvolt is rood, terwijl een foton van 3 electronvolt blauw is[64].

4.8 Het gewicht van nat licht

Laat ons de volgende situatie bezien. Een bundel rode fotonen, met een energie van 2 electronvolt per foton, valt vanuit vacuum in op een vat met water. In het water is de snelheid van de fotonen lager dan in vacuum: dit uit zich in het bekende effect van lichtbreking. De *brekingsindex* n van water is de verhouding van de lichtsnelheden. Deze brekingsindex bedraagt ongeveer 1,33, dat wil zeggen

$$n = \frac{c_{\text{in vacuum}}}{c_{\text{in water}}} = 1.33 \quad . \tag{111}$$

De *kleur* van het licht blijft onder water onveranderd, en daarmee de energie van de fotonen[65]. Aangezien E dus onveranderd is, maar v kleiner is, *moet* het foton in water een rustmassa m hebben! Uit vergelijking (**??**) volgt de waarde van deze massa, uitgedrukt als rust-energie:

$$mc^2 = E\sqrt{-\frac{|\vec{v}|^2}{c^2}} = E\sqrt{1 - 1/n^2} \approx 1.3 \text{ eV} \quad . \tag{112}$$

Het grootste gedeelte van de foton-energie gaat dus in de rustmassa zitten!

[63] In feite, iedere vorm van electromagnetische straling.

[64] De kleur van licht wordt bepaald door de frequentie van de lichtgolven: de rol van de quantummechanica is dat zij de energie van het foton en de frequentie aan elkander relateert.

[65] Dit valt op te maken uit de volgende overweging. Als het licht weer uit het water treedt, hefet het zijn oorspronkelijke kleur. De *intesniteit* kan zijn afgenomen doordat sommige fotonen zijn geabsorbeerd in het water: maar die fotonen die uittreden hebben de oorspronkelijke energie. Nu zijn er allerlei processen denkbaar die aan een foton energie zouden kunnen onntrekken, maar het laat zich toch slecht voorstellen dat dergelijke processen aan de uittredende fotonen precies de juiste hoeveelheid energie zouden teruggeven!

Een paar kanttekeningen zijn hier toepasselijk. In de eerste plaats wordt lichtbreking meestal besproken op het grensvlak van lucht en water, en niet op het grensvlak van vacuum en water[66]. De brekingsindex van lucht ligt echter bijzonder dicht bij 1, zodat dit geen principieel verschil maakt. Ten tweede geldt dat, aangezien licht in water met snelheid kleiner dan c beweegt, deze snelheid *wel degelijk* afhangt van het gebruikte inertiaalstelsel. Dit is inderdaad aangetoond door licht-propagatie in zeer snel stromend water te onderzoeken. Ten derde is de brekingsindex voor rood en blauw licht maar weinig verschillend. Dit houdt in dat de rustmassa van fotonen van verschillende kleur in water verschillend is: voor een blauw foton komt de rust-energie in water op 2 electronvolt.

Het bovenstaande effect kan als een curiositeit van de relativiteitstheorie gezien worden: het is echter in wezen hetzelfde als dat waarmee in de theorie van elementaire deeltjes deze deeltjes geacht worden aan hun rustmassa te komen. Dit is het zogenaamde *Higgs mechanisme*. In dit beeld doordingt een medium, het *Higgs veld* de gehele ruimte. De verschillende elementaire deeltjes, wier 'oer'-massa's alle gelijk zijn aan nul, bewegen zich door dit Higgs veld, en door interacties met dit veld ondergaan zij een vertraging soortgelijk aan dat van licht in water: zodoende krijgen zij een effectieve rustmassa. Overigens moet het bestaan van het Higgs veld nog altijd aangetoond worden. Wie zich hierbij afvraagt of nu niet de æther alsnog weer op slinkse wijze in het wereldbeeld is teruggebracht, heeft met deze bedenking gelijk. Wel is het zo dat de theoretische beschrijving van de wereld der elemantaire deeltjes inclusief Higgs veld volledig relativistisch is: er bestaat niet zoiets als een inertiaalstelsel dat in rust is ten opzichte van het Higgs veld.

4.9 Antimaterie

Eén van de meest tot de verbeelding sprekende voorspellingen van de relativisitsche natuurkunde is het bestaan van anti-materie. Aangezien de precieze argumenten die aan deze voorspelling ten grondslag liggen afkomstig zijn uit de relativistische quantummechanica, kunnen we hier slechts een beperkte redenering geven. Laat ons ons twee tijdruimte-punten voorstellen, A (dat we voor het gemak in de oorsprong plaatsen) en B, met coördinaten $x^\mu = (ct, \vec{x})$. Laat ons ook aannemen dat er een deeltje zich beweegt tussen A en B, met vier-impuls

[66] Dit laatste zou ook in de practijk niet goed te verwezenlijken zijn aangezien water, aan vacuum blootgesteld, onmiddelijk verkookt.

$p^\mu = (E/c, \vec{p})$. Uit hoofde van covariantie weten we dat het product

$$\Phi \equiv x^\mu p_\mu = Et - \vec{p} \cdot \vec{x} \qquad (113)$$

invariant is, en dus voor alle waarnemers dezelfde waarde heeft. De relativisitsche quantummechanica nu stelt dat dit product een eigen fysische betekenis heeft[67].

Laat ons eerst aannemen dat het deeltje, zoals gebruikelijk, positieve energie E heeft, en van A naar B beweegt: dan is $t > 0$ zodat het deeltje, eveneens zoals gebruikelijk, 'van verleden naar toekomst' beweegt. De massa-realtie (**??**),

$$p^\mu p_\mu = E^2/c^2 - |\vec{p}|^2 = m^2c^2 \ , \qquad (114)$$

heeft echter nog een andere oplossing voor de energie, namelijk de negatieve waarde $-E$. Wij kunnen nu *dezelfde* waarde krijgen voor Φ (die in principe alle informatie over de propagatie van het deeltje door de tijdruimte bevat) door ook t van teken te veranderen:

$$\Phi\rfloor_{E\to -E, t\to -t} = (-E)(-t) - \vec{p} \cdot \vec{x} = \Phi \ . \qquad (115)$$

De interpretatie van dit inzicht is het volgende: *de relativisitsche quantummechanica laat deeltjes met negatieve energie toe, mits ze 'van toekokmst naar verleden' bewegen.* Laat ons zo'n geval bekijken: het deeltje beweegt, met *negatieve energie*, van A naar B, waarbij B in het *verleden* van A ligt, aangezien nu $t < 0$.

Zowel negatieve energie als beweging naar het verleden zijn concepten die in flagrante tegenspraak lijken met de dagelijkse practijk. We kunnen echter ons realiseren dat we in feite hier van doen hebben met een conceptueel gezichtsbedrog! Immers, als het deeltje *negatieve* energie van A naar B voert, is dit niet te onderscheiden van een deeltje dat *positieve* energie van B naar A voert, dat wil zeggen precies het soort geval dat wij het normale achten! Wel is hierbij het volgende op te merken: indien het deeltje niet alleen energie transporteert maar ook bijvoorbeeld electrische lading, keert ook deze om: als we een electron (met *negatieve* lading) van A naar B beweegt, is de 'aantrekkelijker' interpretatie er een van een deeltje met *positieve* electrische lading dat van B naar A beweegt. De relativistische quantummechanica voorspelt dus dat er voor ieder type deeltje een *antideeltje* bestaat, met dezelfde massa, maar tegengestelde eigenschappen als lading[68].

[67]Meer nauwkeurig: $x^\mu p_\mu$ is het faseverschil van de golffunctie van het deeltje tussen de twee punten A en B, uitgedrukt in eenheden van de constante van Planck-Dirac, $\hbar$.

[68]Voor neutrale deeltjes is het mogelijk dat het deeltje en het antideeltje in feite identiek zijn. Een voorbeeld hiervan is het foton. De vraag of een deeltje identiek aan het antideeltje moet in principe experimenteel beantwoord worden: wel zien we dat deze identiteit alleen mogelijk is voor electrisch neutrale deeltjes.

De eerste voorspelling van deze aard was die van het anti-electron, tegenwoordig *positron* genoemd. Dit deeltje werd voor het eerst aangetoond in kosmische straling. Tegenwoordig is de deeltje-antideeltje relatie een van de fundamenten van de deeltjes-fysica[69].

Een van de belangrijkste gevolgen van het bestaan van anti-materie is dat, waar de massa van een enkel electron niet als energie kan optreden vanwege het behoud van electrische lading, dit niet hoeft te gelden voor de combinatie materie-antimaterie, vanwege de precies tegengestelde eigenschappen van de twee deeltjes: immers, de ladingen van electron en positron heffen elkaar precies op. Dit maakt een efficient proces mogelijk waarbij materie en antimaterie elkaar annihileren:

$$\text{electron} + \text{positron} \quad \to \quad \text{twee fotonen} \quad . \tag{116}$$

In dit proces kan niet slechts één enkel foton worden geproduceerd, omdat dit niet in overeenstemming te brengen is met het behoud van vier-impuls. Als wij ons het electron en positron beide in rust voorstellen, hebben elk van de twee fotonen precies de rust-energie van het electron of positron: deze bedraagt 511000 electronvolt, en de fotonen zijn dan ook niet die voor zichtbaar licht, maar vallen in het spectrum der *gamma-stralen*. Deze doordringen gemakkelijk objecten als het menselijk lichaam, reden waarom het proces (**??**) gebruikt wordt in diagnostiek onder de naam van *positron-emissie-tomografie* (PET). Ook het omgekeerde proces is mogelijk:

$$\text{twee fotonen} \quad \to \quad \text{electron} + \text{positron} \tag{117}$$

kan optreden mits aan behoud van energie en momentum wordt voldaan. In het algemeen kan men stellen dat, in het geval genoeg energie op een enkel punt geconcentreerd wordt (bijvoorbeeld in hoog-energetische deeltjesbotsingen), deze energie vaak leidt tot de creatie van deeltje- antideeltje paren. Het aanmaken van antimaterie is in dat opzicht betrekkelijk eenvoudig[70].

Tenslotte: het zou aantrekkelijk zijn om de annihilatie van materie aan antimaterie te zien als een bron van schone energie (immers, alleen fotonen worden geproduceerd), waarbij één gram materie en één gram antimaterie (bestaande uit de corresponderende antideeltjes) samen het equivalent vormen van een 44-kiloton bom! Helaas kost het aanmaken van antimaterie uiteraard *meer* energie dan er

[69] In technische termen heet dit de CPT-symmetrie: mocht deze ooit door experiment weerlegd worden, dan zal dat een gebeurtenis zijn van de orde van het niet-bestaan van de æther.

[70] Dat wil zeggen, conceptueel. De nodige technologie is natuurlijk minder triviaal.

bij de annihilatie vrij kan komen, en slechts microscopische hoeveelheden antimaterie kunnen geproduceerd worden. Daarnaast is het opslaan van antimaterie inherent moeilijk en uiterst onveilig; en de geproduceerde fotonen zouden alle bijzonder energetisch zijn, en daarmee niet goed toe te passen in de gebruikelijke methodes van energie-opwekking zoals stoomturbines[71] en dergelijke.

[71] Ook in kerncentrales wordt de verkregen energie, uiteindelijk, gebruikt om stoomturbines aan te drijven.

5 Schets van de algemene relativiteitstheorie

In dit gedeeltje zullen we een schets geven van de voornaamste aspecten van de algemene relativiteitstheorie. Vanwege haar wiskundige ingewikkeldheid zal deze bespreking oppervlakkig blijven.

5.1 Algemene coördinatensystemen, algemene metrieken

Eén van de principes van de speciale relativiteitstheorie is dat van de equivalentie van inertiaalstelsels: elk inertiaalstelsel is in principe even geschikt om fysische processen te beschrijven als elk ander. Men kan zich echter afvragen waarom zich te beperken tot inertiaalstelsels, die zich alleen van willekeurige coördinatensysteem onderscheiden in het feit dat in een inertiaalstelsel de eerste wet van Newton geldt. De algemene relativiteitstheorie stelt dan ook zonder meer dat, in principe, *ieder* coördinatensysteem geschikt is om fysische processen te beschrijven. Wel vereist dit dat de natuurwetten geschreven worden in een vorm die invariant is onder *willekeurige* coördinatentransformaties: een formidabele opgave! Een van de gevolgen van het bezien van willekeurige coördinatensystemen is dat de 'echte' afstand tussen twee punten niet langer gegeven wordt door de simpele Minkowski vorm: in plaats daarvan gebruiken we de metrische tensor in zijn algemene vorm: twee punten waarvan de coördinaten verschillen met een infinitesimaal bedrag dx^μ worden geacht een onderlinge 'echte afstand' te hebben ter grootte

$$(ds)^2 = g(x)_{\mu\nu}\, dx^\mu\, dx^\nu \quad , \tag{118}$$

waarbij nu de metrische tensor $g_{\mu\nu}$ een meer algemene vorm kan hebben, die ook nog van plaats tot plaats kan verschillen[72]. De Minkowski-metriek kan nu een heel ongewone en ingewikkelde gedaante krijgen, afhankelijk van het gekozen coördinatensysteem. Het onderzoeken van de transformatie-eigenschappen van objecten onder willekeurige coördinatentransformaties vormt een gebied van de zogenaamde *differentiaalmeetkunde*. Het moge duidelijk zijn dat het niet altijd eenvoudig is om in te zien dat twee verschillend uitziende metrieken in feite dezelfde ruimte beschrijven in termen van twee verschillende coördinatensystemen!

[72] Het is eenvoudig te zien dat de metrische tensor wel symmetrisch kan worden genomen in zijn indices: dat wil zeggen, $g(x)_{\mu\nu} = g(x)_{\nu\mu}$.

5.2 Het probleem van de zwaartekracht

Zoals we hebben gezien is de gravitatiewet van Newton niet in overeenstemming te brengen met de relativiteitstheorie omdat haar werking onmiddelijk is, over willekeurig grote afstanden. Een aanpassing is dus noodzakelijk. Daarbij past nog een overweging. De zwaartekracht dat een voorwerp ondervindt, is volgens Newton evenredig met zijn massa: de tweede wet van Newton zegt dus dat de *versnelling* die het voorwerp ondervindt precies *onafhankelijk* moet zijn van zijn massa! Alle voorwerpen die in een gegeven zwaartekrachtsveld op enig moment dezelfde positie en snelheid hebben, zullen dus *identieke* banen doorlopen[73].

De positie van Einstein in deze was, zoals gebruikelijk, er een van robuuste eenvoud. Als alle trajecten die voorwerpen, onder invloed van zwaartekracht, afleggen niet afhangen van de aard van de voorwerpen, moet dat liggen aan het feit dat deze trajecten niet door de voorwerpen zelf bepaald worden, maar slechts door de ruimte waardoorheen zij bewegen. Indien de baan van een deeltje gekromd is (zoals planeetbanen dat zijn), moet dat veroorzaakt worden doordat *de ruimte zelf gekromd is*. Het is betrekkelijk eenvoudig een mentaal beeld te vormen van een twee-dimensioneel oppervlak dat gekromd ligt in de drie-dimensionele ruimte. Een vier-dimensionele gekromde tijdruimte gaat ons inbeeldingsvermogen wellicht te boven, niet echter ons wiskundig vermogen. De differentiaalgeometrie stelt ons in staat, de krommings-eigenschappen van een ruimte te destilleren uit haar metrische tensor. Merk op dat we hier een dubbele taak hebben: ten eerste, de krommingseigenschappen te bepalen uit de metriek, en ten tweede ons daarbij niet te laten misleiden door de enorme vrijheid aan keuzes van een coördinatensysteem...

Einstein postuleert verder het volgende: een puntdeeltje waarop geen andere krachten dan de zwaartekracht werkzaam zijn, beweegt zich in ‘een rechte lijn’ met ‘eenparige’ snelheid voort. Nu is natuurlijk een strikt rechte lijn niet in een gekromde ruimte in te passen, en daarom spreken wij liever van een *geodeet*: dit is de lijn in een gekromd oppervlak die de korste route vormt tussen twee punten, gemeten in termen van de ‘echte afstand’ s. Voor een waarnemer die met het deeltje meebeweegt is de ‘echte afstand’ de eigentijd, dus beweging over een geodeet is die beweging die ‘de minste tijd vergt’. Als een waarnemer een deeltje over een niet-rechtlijnige baan, of niet-eenparig, ziet bewegen is dit het gevolg

[73]Merk op dat deze consequentie van Newton’s wetten inhoudt dat ook *licht* aan de zwaartekracht onderhevig is: immers, het moet voor ons niet uitmaken of een foton een massa van 10^{-100} gram heeft, of 10^{-2468} gram, of massa gelijk nul: in feite is de bekende bovengrens op de gemeten fotonmassa nogal wat royaler, zo’n 10^{-70} gram.

van het feit dat het coördinatensysteem van deze waarnemer geen inertiaalstelsel is. In die zin is de zwaartekracht een zogenaamde *schijn-kracht*, waarvan de ‘middelpuntvliedende kracht’ bij draaibewegingen een ander voorbeeld vormt.

Uiteraard moeten we onder ogen zien, dat de zwaartekracht uitgeoefend *door* een voorwerp wel degelijk van de massa van dat voorwerp afhangt. De gravitatiewet van Newton moet dus vervangen worden door een andere, die laat zien *hoe* de (tijd)ruimte gekromd wordt door massa’s.

Einstein slaagde er in een vergelijking te vinden die de laatste twee overwegingen combineert, in een vorm die covariant is onder willekeurige coördinatentransformaties. Inmiddels is deze vergelijking door talrijke waarnemingen grondig beproefd, en tot op heden staat zij nog fier overeind.

Eén van de eerste verificaties van de algemene relativiteitstheorie was de waarneming van de afbuiging van licht, afkomstig van vaste terren, dat vlak langs het zonneoppervlak scheert. Omdat onder normale omstandigheden de zon zelf dit licht uiteraard overstemt, kan zo’n waarneming alleen plaatsvinden tijdens een totale zonsverduistering. Hierbij moet worden opgemerkt dat, zoals gezien, ook de Newtoniaanse mechanica een dergelijke afbuiging voorspelt, zodra men accepteert dat licht uit deeltjes bestaat. Het verschil met Einstein’s voorspelling is veeleer kwantitatief: de relativiteitstheorie voorspelt een afbuiging die precies twee maal de Newtoniaanse waarde bedraagt. Het moment waarop, in 1919, deze waarde inderdaad gevonden werd verhief Einstein onmiddelijk tot sterrenstatus. Overigens kan het worden opgemerkt dat een zwaartekrachtsveld dus als een optische ‘lens’ kan werken. Inmiddels zijn talrijke gevallen van ‘gravitational lensing’ aangetroffen, waarin verschillende kopieën van lichtbronnen vlak naast elkaar te zien zijn omdat het licht op zijn weg naar de aarde op verschillende manieren door zeer zware objecten is afgebogen, en de aarde dus bereikt vanuit verschillende (schijnbare) richtingen.

5.3 De wiskundige artillerie van Einstein’s theorie

Ter illustratie geven we hier de verschillende vergelijkingen die ten grondslag liggen aan de algemene relativiteitstheorie. De eigenschappen van de ruimte worden geheel bepaald door de metriek $g(x)_{\mu\nu}$, die zoals gezegd in het algemeen van x (en dus ook van de tijdscomponent!) afhangt.

De *Christoffel-symbolen* beschrijven de manier waarop de metriek van punt tot punt verandert: laten we de expliciete afhankelijkheid van de metriek van x in onze notatie buiten beschouwing, dan worden de Christoffel-symbolen gegeven, met gebruik van de Einstein sommatie, door

$$\Gamma^{\alpha}{}_{\mu\nu} = g^{\alpha\beta}\left(\frac{\partial}{\partial x^{\nu}} g_{\beta\mu} + \frac{\partial}{\partial x^{\mu}} g_{\beta\nu} - \frac{\partial}{\partial x^{\beta}} g_{\mu\nu}\right) \quad . \tag{119}$$

Hierbij is $g^{\alpha\beta}$ de contravariante metrische tensor, waarvoor geldt

$$g^{\alpha\beta} g_{\beta\sigma} = \begin{cases} 1 & \text{if } \alpha = \sigma \\ 0 & \text{if } \alpha \neq \sigma \end{cases} \quad . \tag{120}$$

De afgeleiden zijn zogenaamde partiële afgeleiden, dat wil zeggen $\partial / partial x^{0}$ is de afgeleide naar x^{0} waarbij de andere coördinaten vast worden gehouden. Merk op dat in onze vier-dimenionele ruimte er in principe 64 Christoffel symbole te berekenen zijn.

Laat de coordinaat van een deeltje op eigentijd *s* gegeven worden door $x(s)^{\mu}$: dan is de vergelijking voor een geodeet-baan de volgende:

$$\frac{d^{2}}{ds^{2}} x(s)^{\mu} + \Gamma^{\mu}{}_{\nu\rho}\left(\frac{d}{ds} x(s)^{\nu}\right)\left(\frac{d}{ds} x(s)^{\rho}\right) = 0 \quad . \tag{121}$$

In het eenvoudige geval van de Minkowski ruimte en een Cartesisch coördinatensysteem is de metriek constant, en zijn de Christoffel symbolen gelijk aan nul. De geodeet-vergelijking wordt daan eenvoudig $d^{2} x(s)^{\mu} / (ds)^{2} = 0$, inderdaad de vergelijking voor eenparige rechtlijnige beweging!

De krommingseigenschappen van de ruimte worden belichaamd in de zogenaamde *Riemann-Christoffel tensor*:

$$R^{\beta}{}_{\nu\rho\sigma} = \frac{\partial}{\partial x^{\rho}} \Gamma^{\beta}{}_{\nu\sigma} - \frac{\partial}{\partial x^{\sigma}} \Gamma^{\beta}{}_{\nu\rho} + \Gamma^{\alpha}{}_{\nu\sigma} \Gamma^{\beta}{}_{\alpha\rho} - \Gamma^{\alpha}{}_{\nu\rho} \Gamma^{\beta}{}_{\alpha\sigma} \quad . \tag{122}$$

Deze tensor (met 256 componenten!) heeft de volgende eigenschap: alle componenten zijn gelijk aan nul dan en slechts dan wanneer de ruimte feitelijk vlak is, ongeacht het gekozen coördinatensysteem. Daarom wordt zij ook wel de krommings-tensor genoemd. Een compacter object dat eveneens krommingsinformatie bevat is de *Ricci tensor*, met 16 componenten:

$$R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu} \quad . \tag{123}$$

Ook is er de *Gauss kromming*:

$$R = g^{\mu\nu} R_{\mu\nu} \tag{124}$$

De Gauss kromming hangt in het geheel niet van het gekozen coördinatensysteem af. Dit wil zeggen dat wanneer wij bijvoorbeeld op een boloppervlak leven, de kromtestraal van dit boloppervlak berekend kan worden beginnend vanuit een willekeurig coördinatensysteem: in alle gevallen is de straal r van de bol dan $r = 1/\sqrt{R}$. De *Einstein tensor*, tenslotte, wordt gegeven door

$$E_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R \, g_{\mu\nu} \quad , \tag{125}$$

en dit is de tensor die in de vergelijkingen van de algemene relativiteitstheorie optreedt als beschrijving van de krommingseigenschappen van de ruimte.

Wat betreft de puntmassa's in Einstein's vergelijking: laat ons een puntdeeltje bezien met massa m, en vier-snelheid $u^\mu = dx(s)^\mu/ds$ in enig coördinatensysteem. We vormen dan de *stress-energie tensor* $T^{\mu\nu}$ als volgt: buiten de onmiddelijke nabijheid van het deeltje is deze nul, en gesommeerd over de onmiddelijke nabijheid van het deeltje[74] levert zij het bedrag

$$T^{\mu\nu} = m \, u^\mu \, u^\nu \quad . \tag{126}$$

Als er meerdere puntdeeltjes zijn, moeten we de resultaten voor $T^{\mu\nu}$ uiteraard sommeren. Merk op dat $T^{\mu\nu}$ natuurlijk ook van x afhangt.

De Einstein vergelijking heeft nu de bedrieglijk simpele vorm

$$E^{\mu\nu} + \frac{8\pi G_N}{c^4} T^{\mu\nu} = 0 \quad . \tag{127}$$

Hierbij is G_N de gravitatieconstante van Newton. Deze vergelijking *opschrijven* is één ding, haar *oplossen* weer iets heel anders!

Een aantal opmerkingen zijn hier toepasselijk. Ten eerste: de combinatie

$$\frac{8\pi G_N}{c^4} \approx 2 \times 10^{-43} \, \frac{\mathrm{s}^2}{\mathrm{kg\ m}} \tag{128}$$

is bijzonder klein in termen van onze alledaagse maatstaven. Dit houdt in dat een zeer kleine kromming van de ruimte reeds grote effecten kan hebben op de beweging van de materie. In het geval waarin de kromming slechts zeer weinig van nul verschilt (en de riumte dus *bijna* vlak is) kunnen we uit (**??**) de gravitatiewet van Newton afleiden, tenminste voor puntdeeltjes met relatief lage snelheden.

[74] In feite introduceren wij hier de 'Dirac-δ-verdeling'.

Ten tweede, de combinatie van tensoren in $E^{\mu\nu}$ is bijzonder listig door Einstein gekozen: het valt namelijk eenvoudig af te leiden dat vergelijking (**??**) impliceert dat puntdeeltjes bewgen langs geodeten, voorzover zij niet aan andere, niet-gravitationele krachten onderworpen zijn[75]. In feite kan men het volgende laten zien. De Christoffel symbolen bevatten eerste afgeleiden van de metriek, en de krommingstensoren bevatten daarom tweede afgeleiden. In het algemeen worden tweede afgeleiden gezien als de hoogste die in een accpetabele fysische theorie kunnen voorkomen[76]. De tensor $E^{\mu\nu}$ is nu de *enige* tensor die hoogstens tweede afgeleiden bevat, en daarbij de geodetische beweging van ountdeeltjes impliceert, *én* de Minowski-ruimte toelaat als oplossing in het geval van ledige ruimte, waarvoor $T^{\mu\nu}$ uiteraard nul is en dus ook $E^{\mu\nu}$.

De bovengenoemde wiskundige objecten zijn hier uiteraard gegeven om te dienen als illustratie van de theorie. Het oplossen in concrete gevallen is, het moge duidelijk zijn, een technisch zeer zware opgave, en het is dan ook slechts in een beperkt aantal situaties mogelijk een eenvoudig uitziende oplossing te presenteren.

5.4 Zwarte gaten

Vrijwel onmiddelijk na de publicatie van de Einstein vergelijking (**??**) werd door Schwarzschild een eenvoudige oplossing gevonden. Deze zullen we nu kort bespreken, aangzien zij een van de weinige direct inzichtelijke is. We bezien een coördinatensysteem met de gewone tijd-coordinaat t, en ruimtelijke coördinaten in de vorm van bol-coördinaten r, θ en ϕ. Dat wil zeggen, in termen van een Cartesisch systeem hebben we

$$x^0 = ct \quad , \quad x^1 = r\sin(\theta)\sin(\phi) \quad , \quad x^2 = r\sin(\theta)\cos(\phi) \quad , \quad x^3 = r\cos(\theta) \quad . \tag{129}$$

In de gewone Minkowski-ruimte zou de metriek dan gegeven worden door

$$(ds)^2 = (c\ dt)^2 - (dr)^2 - r^2(d\theta)^2 - r^2\sin(\theta)^2(d\phi)^2 \quad . \tag{130}$$

Schwarzschild onderzocht de mogelijkheid van een oplossing die van r afhangt, maar niet van de hoeken θ en ϕ, een zogenaamde *bolsymmetrische* oplossing, en

[75]Gelukkig zijn zij dit meestal wel: het menselijk lichaam kan niet door gravitatiekrachten bijeengehouden worden.

[76]Theoriën met hogere afgeleiden vereisen voor hun oplossing meer beginvoorwaarden dan de gebruikelijke positie en snelheid. Dit is geen absoluut obstakel maar drijft de fysische ingewikkeldheid van de theorie naar een geheel andere grootteorde.

wel in de ledige riumte, waarin $T^{\mu\nu} = 0$. Deze symmetrie reduceert het probleem enigzins, zodat een exacte oplossing gevonden kan worden. In het algemene geval luidt deze oplossing:

$$(ds)^2 = \left(1 - \frac{R_s}{r}\right)(c\ dt)^2 - \left(1 - \frac{R_s}{r}\right)^{-1}(dr)^2 - r^2(d\theta)^2 - r^2\sin(\theta)^2(d\phi)^2 \quad . \tag{131}$$

De zogenaamde *Schwarzschild straal* R_s karakteriseert deze oplossing. Voor $R_s = 0$ hebben we de Minkowski-metriek. Op grote afstand $r \gg R_S$, is de kromming van de ruimte maar klein, en kunnen we vergelijken met de Newtoniaanse limiet: deze houdt in dat we moeten aannemen dat zich op het punt $r = 0$ een puntmassa bevindt. De Schwarzschild straal hangt van deze massa m af volgens

$$R_s = \frac{2G_N}{c^2}m \quad . \tag{132}$$

Deze straal is vaak klein: voor de zon bedraagt ze 2.9 kilometer, voor de aarde is zij 0.9 centimeter, en voor een proton neemt ze de miniscule waarde 2.4×10^{-52} centimeter aan.

Laat ons de Schwarzschild oplossing eens nader bekijken. Ten eerste moet worden opgemerkt dat zij in essentie de *enige* bolsymmetrische oplossing is. Ten tweede verwachten we, zo geen problemen, dan toch minstens merkwaardige verschijnselen indien r de waarde R_s aanneemt. Aangezien de Schwarzschild oplossing slechts geldt in de ledige ruimte, is er geen reden haar *au sérieux* te nemen in het inwendige van de zon of de aarde. Het is echter wel degelijk voorstelbaar, en dit wordt door astronomische waarnemingen ook geïndiceerd, dat er voorwerpen bestaan waarvan de *fysische* afmeting kleiner is dan de Schwarzschild straal R_s. Uiteraard zal een dergelijk voorwerp een extreem grote dichtheid hebben: maar bijvoorbeeld, grote uitgebrande sterren kunnen een dergelijke dichtheid bereiken wanneer zij onder invloed van hun eigen zwaartekracht ineenstorten. In dat geval reikt de Schwarzschild straal tot buiten het voorwerp, en is $r = R_s$ gelegen in de ledige ruimte rond het voorwerp. Laat ons een puntdeeltje bezien dat van buitenaf ($r > R_s$) naar binnen valt. Het valt betrekkelijk eenvoudig te berekenen dat *de tijd* t*, benodigd om de waarde* $r = R_s$ *te bereiken, oneindig groot is!*. Met andere woorden, van buitenaf kan de Schwarzschild straal niet doorbroken worden. Ook is het niet mogelijk enig signaal met lichtsnelheid of minder vanuit $r < R_s$ naar buiten te sturen: vandaar de benaming *zwart gat* voor een dergelijk object. De situatie is geheel anders wanneer we de vertreken *eigentijd* van het deeltje als maatstaf nemen: deze is *eindig* op het moment dat $r = R_s$ bereikt wordt.

Een waarnemer die in een zwart gat valt zal zelf vaststellen dat zij er eindig lang over doet: maar inmiddels is de rest van het universum verdwenen wanneer zij de Schwarzschild straal doorbreekt — er is geen weg terug.

Er lest nog, op te merken dat zeer onlangs een langlopende serie waarnemingen van sterren in het centrum van onze melkweg een aantal sterren heeft geïdentificeerd die met zeer hoge snelheid zeer nauwe banen doorlopen rond een bepaald punt. Uit de sterrebanen kan men opmaken dat zich op dat punt een massa moet bevinden van meer dan twee miljoen keer die van onze zon, dat wil zeggen met een Schwarzschild straal van enige miljoenen kilometers. Van dit massieve object is geen straling waar te nemen. De conclusie dat zich hier een zwart gat bevindt, ligt dan ook voor de hand!

Appendix: gewone en hyperbolische trigonometrische functies

Voor de volledigheid vermelden we hier de voornaamste eigenschappen van de trigonometrische functies. Het meest vertrouwd zijn waarschijnlijk de gewone trigonometrische functies, $\sin(x)$ en $\cos(x)$. Voor deze gelden de volgende relaties:

$$\sin'(x) = \cos(x) \quad , \quad \cos'(x) = -\sin(x) \quad , \tag{133}$$

en

$$\begin{aligned} \cos(x)^2 + \sin(x)^2 &= 1 \quad , \\ \cos(x)\cos(y) - \sin(x)\sin(y) &= \cos(x+y) \quad , \\ \sin(x)\cos(y) + \cos(x)\sin(y) &= \sin(x+y) \quad . \end{aligned} \tag{134}$$

Verder is $\cos(x)$ even, en $\sin(x)$ oneven, in x. Voor de hyperbolische trigonometrische functies, $\sinh(x)$ en $\cosh(x)$, gelden analoge relaties:

$$\sinh'(x) = \cosh(x) \quad , \quad \cosh'(x) = \sinh(x) \quad , \tag{135}$$

en

$$\begin{aligned} \cosh(x)^2 - \sinh(x)^2 &= 1 \quad , \\ \cosh(x)\cosh(y) + \sinh(x)\sinh(y) &= \cosh(x+y) \quad , \\ \sinh(x)\cosh(y) + \cosh(x)\sinh(y) &= \sinh(x+y) \quad ; \end{aligned} \tag{136}$$

eveneens is $\cosh(x)$ even, en $\sinh(x)$ oneven, in x. Verder hebben we

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \quad , \quad \tanh(x) = \frac{\sinh(x)}{\cosh(x)} \quad . \tag{137}$$

We kunnen een en ander ook uitdrukken in exponentiële functies, met zoals gewoonlijk $i \equiv \sqrt{-1}$:

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i} \quad , \quad \cos(x) = \frac{e^{ix} + e^{-ix}}{2} \quad , \tag{138}$$

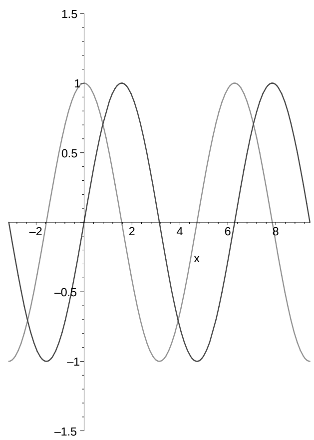
en

$$\sinh(x) = \frac{e^{x} - e^{-x}}{2} \quad , \quad \cosh(x) = \frac{e^{x} + e^{-x}}{2} \quad . \tag{139}$$

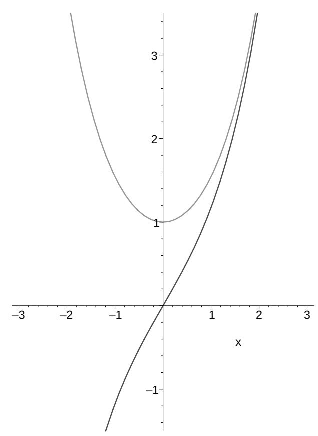
Hieruit valt tevens de in de tekst gebruikte vorm

$$\tanh(x) = \frac{1}{2} \log\left(\frac{1+x}{1-x}\right) \tag{140}$$

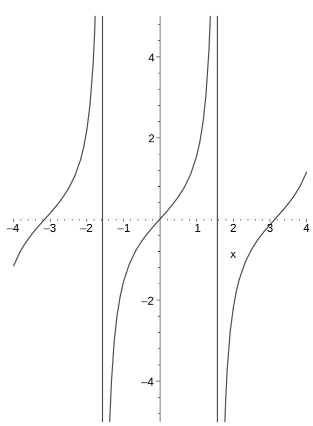
af te leiden.



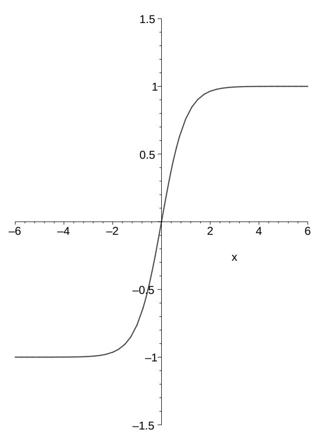
sin(x) en cos(x)



sinh(x) en cosh(x)



tan(x)



tanh(x)