Het Levend Heelal
aantekeningen 2003

Vincent Icke ◇ icke@strw.LeidenUniv.nl

1. Doelstelling en samenvatting

Waarin wordt aangekondigd waar het allemaal over gaat. Er staat hier veel in om te onthouden. Maar onthoud vooral dit: astrofysica is zware kost! Hou vol.

Het doel van het deze inleiding op het college *Het Levend Heelal* bestaat uit drie delen: ten eerste, een basis geven voor het begrijpen van wat zich in het Heelal afspeelt; ten tweede, aanwijzingen geven voor de delen van wis- en natuurkunde die hierbij een rol spelen; ten derde, een raamwerk vormen voor de meer gedetailleerde colleges en practica die de rest van de cursus uitmaken.

Het gaat hierbij in eerste instantie om de klassieke mechanica, en in het bijzonder de Newtonse theorie van baanberekeningen ("hemelmechanica"). Die wordt toegepast op statische situaties (vorm en structuur van planeten en sterren), op stationaire situaties (cirkelbanen), en op dynamische systemen (planeetbanen, clusters van sterren, sterrenstelsels, en het Heelal). Bij de behandeling van al deze zaken wordt geleidelijk duidelijk welke wiskundige technieken nodig zijn, en waarom. In het bijzonder zijn dat algebra, lineaire algebra, en analyse.

De formule-nummers van zeer belangrijk materiaal zijn in de tekst met een schoppenaas ♠ aangegeven, bijvoorbeeld alsvolgt:

$$E = \gamma m c^2 \tag{1.1 ♠}$$

Het is aan te bevelen om eerst het *globale gedrag* van de variabelen in zo'n formule te onthouden, en dan pas de (eventuele) constanten. Bijvoorbeeld, in de formule

$$T = \frac{G\mu}{3k}\frac{M}{R} \tag{1.2 ♠}$$

is de relatie tussen T, M en R het belangrijkst:

$$T \propto \frac{M}{R} \tag{1.3}$$

Deze weergave met een 'evenredig' symbool $\propto$ moet gelezen worden als

$$T = \text{constante} \times \frac{M}{R} \tag{1.4}$$

en uiteindelijk, na invullen van de constanten, als Eq.(0.2) .

Paragrafen die beginnen met **Ga nu zelf het volgende na** zijn bedoeld als oefening om te zien of de lezer de stof begrijpt; tevens geven deze passages een aanwijzing voor het soort vragen die op een tentamen zouden kunnen voorkomen.

Hieronder volgt een overzicht van de samenvattingen van alle hoofdstukken uit deze notities.

Als een buitenaards wezen onze planeet zou bezoeken, en ons zou vragen wat wij over het Heelal te weten zijn gekomen sinds we uit de bomen klommen, dan is het kortste antwoord: het Heelal bestaat uit deeltjes, ruimte en tijd. Op basis van deze schijnbaar eenvoudige zin kunnen we de hele natuurkunde reconstrueren, en alle sterrenkundige toepassingen daarvan. In dit college beperken wij ons tot de 'eerste levensbehoeften': de klassieke mechanica, een deel van de quantummechanica, stukjes thermodynamica, en relativiteitstheorie.

De regels van de mechanica zijn van zeer groot belang in de sterrenkunde. Het is dus essentieel om te weten waar die regels vandaan komen. Het blijkt dat de 'wetten' van de beweging voortkomen uit bepaalde regelmatigheden in de Natuur, 'symmetrie' genoemd. Dat zijn: homogeniteit van tijd en ruimte, en relativiteit van de snelheid. Uit deze symmetrieën leidt men de bewegingsvergelijkingen af. Bij deze afleiding blijkt dat wiskundige analyse niet gemist kan worden.

Bij iedere symmetrie hoort een grootheid die niet verandert. In de klassieke mechanica zijn dat de energie en de impuls. De klassieke bewegingsvergelijking heeft in het rechterlid nul, in het geval dat er geen uitwendige krachten werken. Dit leidt tot behoud van impuls. De vergelijking kan in alle gevallen eenmaal worden geïntegreerd. Zo ontstaat een vorm waaruit blijkt dat een bepaalde grootheid bij de beweging hetzelfde blijft. Dit is de behoudswet voor de energie.

Gewapend met de vergelijkingen van de klassieke mechanica gaan we het Heelal proberen te begrijpen. Het uiteindelijke doel is, te berekenen hoe astronomische dingen zich gedragen. Het eenvoudigst is, om daarbij te beginnen met dingen in evenwicht, zoals planeten en sterren. Eerst bezien wij de structuur van planeten, in evenwicht tussen enerzijds hun zwaartekracht en anderzijds de druk van de materie waaruit zij zijn gebouwd. Hier berekenen we de sterkte van de materie uit een simpele schatting van materiaalsterkte.

Nu we eenmaal weten dat een voldoende grote massa bolvormig is, kunnen we schatten wat de inwendige structuur van zo'n bol is. Wij doen dat door de bol in twee delen te splitsen: een schil en een pit, met allebei dezelfde dikte en massa. Wij berekenen een verband tussen de massa, straal, dichtheid, temperatuur en andere eigenschappen van de bol. Later zullen wij zien dat deze schattingen verrassend realistisch zijn.

Nu gaan we de structuurformule van een zelf-graviterende bol uitbreiden naar zware dingen zoals sterren. Natuurlijk bestaat een planeet of een ster niet uit een simpele schil en een pit daarbinnen. Op basis van precies dezelfde aanpak kunnen we het veel beter doen: we nemen niet twee schillen, maar stapelen er oneindig veel op. Uiteraard leidt dat tot een differentiaalvergelijking voor de structuur van de bol. De oplossing hiervan is erg ingewikkeld, maar een schatting van de bijbehorende fysische grootheden levert een uitkomst die sterk lijkt op het twee-schillenmodel. In de hier gegeven beschrijving zijn sterren in evenwicht doordat er in hun binnenste een zeer hoge temperatuur heerst, dat wil zeggen snelle chaotische bewegingen van de atomen. Later komen wij een soortgelijke aanpak tegen bij de structuur van sterrenstelsels, die in evenwicht zijn door de chaotische bewegingen van hun sterren.

Nu gaan we over naar de dynamica: de snelheden zijn nu niet meer nul. De krachten worden niet in evenwicht gehouden door een tegengestelde kracht. In alles wat volgt zullen wij meestal werken met krachten die naar een vast punt zijn gericht. Uit de afleiding van de bewegingsvergelijking weten we dat alleen een deeltje waarop geen krachten werken, een rechte baan doorloopt met constante snelheid. Als de baan gekromd is, is er sprake van versnelling, en wij moeten weten hoe groot die is. Wij volgen de berekening van Huygens. We stellen vast dat elke kromming plaatselijk op twee manieren kan worden benaderd: door een cirkelboog en door een parabool. Het stukje cirkel wordt doorlopen met een hoeksnelheid en een kromtestraal. Door die af te beelden op de parabool vinden we welke versnelling overeenkomt met dat stukje van de baan. Zo'n versnelling noemen we 'centrifugaal'.

Een centrale kracht heeft een bijzondere eigenschap: de kracht is steeds naar een vast punt gericht. Deze afwijking van het algemene geval betekent dat de beweging van een deeltje in zekere zin in zijn vrijheid is beperkt. Bij zo'n beperking hoort een behouden grootheid, een 'bewegingsconstante'. In het geval van beweging onder invloed van een centrale kracht is dat het impulsmoment. Dat is het uitproduct van de impulsvector en de vector van de hoeksnelheid. In een vlak reduceert dit tot het product van de massa, de snelheid en de kromtestraal van de baan. Het behoud van impulsmoment is verantwoordelijk voor de tweede wet van Kepler, de 'perkenwet'.

Als voorbereiding op de algemene oplossing van de bewegingsvergelijkingen staat de cirkelbaan centraal. De reden is dat in veel sterrenkundige toepassingen de kracht die op het deeltje werkt naar een vast punt is gericht. In het geval van zo'n centrale kracht kan de hoeksnelheid van de cirkelbaan eenvoudig worden gevonden. In het planetenstelsel is dat de Derde Wet van Kepler.

Uit het verband tussen massa, straal en temperatuur van bollen zien wij dat in het binnenste van een voorwerp als de Zon de temperatuur van de orde van 10 miljoen kelvin is. Onder die omstandigheden zijn de electronen geheel los van de atoomkernen. Die kernen botsen tegen elkaar en zo kan er kernfusie optreden, wardoor de ster kan blijven stralen. Om te berekenen hoe dat gaat moeten we weten hoe twee protonen zo dicht bij elkaar kunnen komen dat de 'sterke kernkracht' ze bijeen bindt. Dit is een toepassing van de quantummechanica. We volgen de methode van Gamow om te schatting bij welke energie kernreacties in een ster het best verlopen.

De straling die in het binnenste van een ster wordt opgewekt ontsnapt uiteindelijk door het oppervlak. De temperatuur daar is wel veel lager dan in het centrum van de ster, maar toch nog vele duizenden kelvin. Om de energie die de ster verliest (de lichtkracht) te koppelen aan de temperatuur gebruiken wij de regel van Stefan-Boltzmann.

In het geval van het eenvoudige twee-schillenmodel zagen wij voor het eerst dat er een verband bestaat tusen de karakteristieke grootheden van een zelfgraviterende bol, zoals de massa, de straal en de temperatuur. Omdat de lichtkracht een functie is van de oppervlakte en de temperatuur, is er een verband tussen de massa en de lichtkracht van een ster.

2. Deeltjes, ruimte en tijd

Als een buitenaards wezen onze planeet zou bezoeken, en ons zou vragen wat wij over het Heelal te weten zijn gekomen sinds we uit de bomen klommen, dan is het kortste antwoord: het Heelal bestaat uit deeltjes, ruimte en tijd. Op basis van deze schijnbaar eenvoudige zin kunnen we de hele natuurkunde reconstrueren, en alle sterrenkundige toepassingen daarvan. In dit college beperken wij ons tot de 'eerste levensbehoeften': de klassieke mechanica, een deel van de quantummechanica, stukjes thermodynamica, en relativiteitstheorie.

Het Heelal bestaat uit deeltjes, ruimte en tijd. Het samenspel daarvan vormt de gehele (astro)fysica. Een complete natuurkunde van deze drie ("theorie van alles") is er nog niet. We behelpen ons met enerzijds theorieën van deeltjes die zich in een vooraf gegeven ruimte bewegen (quantummechanica, veldentheorie), anderzijds met ruimte-tijdstructuren waarin deeltjes passief aanwezig zijn (algemene relativiteitstheorie).

Uit dagelijkse ervaring met materie weten we dat er deeltjes zijn. Er bestaan chemische `elementen` (H, O, S,... heel weinig soorten!), die in zeer weinig en zeer specifieke combinaties kunnen voorkomen. Bij het analyseren van (zuivere) stoffen vinden we nooit samenstellingen van het type $H_{2.4}O_{0.9}$, maar altijd (kleine) gehele getallen: H_2O, H_2SO_4, en dergelijke `quantumgetallen`. Uit de genetica weten we, dat erfelijke eigenschappen als herkenbare brokstukken worden doorgegeven (Mendel-wetten).

De beschrijving van deeltjes maakt gebruik van `complexe amplituden`, getallen van het type $e^{i\phi}$. Omdat we kunnen schrijven $e^{i\phi} = \cos\phi + i \sin\phi$ vertonen deeltjes bepaalde soorten van gedrag die aan golven doen denken (bijvoorbeeld interferentie), maar het is onjuist om te zeggen 'deeltjes zijn golven'. Door dit golf-achtige gedrag kan een deeltje maar zeer weinig vaste toestanden aannemen, vandaar dat die quantumgetallen optreden, en vandaar dat deeltjes en moleculen *identiek* kunnen zijn en niet, zoals grote dingen (muizen en mensen) maar zo'n beetje op elkaar lijken.

Ruimte en tijd kunnen worden gemeten met dezelfde maat, namelijk de seconde, omdat de Natuur ons een absolute snelheid biedt: de lichtsnelheid is (locaal) onder alle omstandigheden hetzelfde (experiment van Michelson en Morley). Zo is de afstand tot de Maan 1.1 seconde, tot de Zon 8 minuten, tot de Andromeda Nevel 2 miljoen jaar.

Deze en verwante merkwaardigheden zijn te ingewikkeld voor dit college, en worden later uitgewerkt, bijvoorbeeld in de colleges quantummechanica, veldentheorie en relativiteitstheorie.

3. Symmetrie en klassieke mechanica

De regels van de mechanica zijn van zeer groot belang in de sterrenkunde. Het is dus essentieel om te weten waar die regels vandaan komen. Het blijkt dat de 'wetten' van de beweging voortkomen uit bepaalde regelmatigheden in de Natuur, 'symmetrie' genoemd. Dat zijn: homogeniteit van tijd en ruimte, en relativiteit van de snelheid. Uit deze symmetrieën leidt men de bewegingsvergelijkingen af. Bij deze afleiding blijkt dat wiskundige analyse niet gemist kan worden.

Om erachter te komen welke vergelijking(en) gebruikt kunnen worden om beweging van deeltjes te beschrijven moeten we eerst zien welke fysica moet worden ingebouwd. Op het eerste gezicht zou je denken dat de vergelijking voor de baan van een deeltje een *algebraïsche* vorm is, bv. de vergelijking van de parabool

$$y = a + bt + ct^2 \tag{3.1}$$

waarin y zoiets als de hoogte van een deeltje en t de tijd. Maar het blijkt dat het niet zo gaat. Dat komt doordat de absolute positie en de absolute snelheid van deeltjes in ons Heelal blijkbaar geen meetbare grootheden zijn: aan niets is af te lezen wat de ruimtelijke coördinaten van een deeltje zijn. Dat is geen 'principe' of zo; in de studeerkamer, afgesloten van de werkelijkheid, zouden we best een heelal kunnen verzinnen waarin deeltjes een soort inwendige kilometerteller hebben waarop je de plaats ervan kunt aflezen. Maar het Heelal waarin wij wonen werkt niet zo. In onze natuur geldt de `homogeniteit van de ruimte`: als je bij de positie van een deeltje een willekeurig vast getal optelt, verandert er niets. Dit komt neer op een globale verschuiving van het Heelal. Met andere woorden, er geldt blijkbaar een soort 'relativiteit van de ruimte': de waarde van een coördinaat als zodanig is niet waarneembaar. In formule luidt deze invariantie

$$\vec{r} \quad \Longrightarrow \quad \vec{r} + \vec{a} \tag{3.2}$$ ♠

Dus is niet de absolute positie van een deeltje van belang, maar de *relatieve* plaats, in het bijzonder de verandering van de plaats in de loop van de tijd: de snelheid. Dus alleen het verschil in positie telt, vandaar dat de bewegingsvergelijking geen algebraïsche vergelijking voor de positie $\vec{r}$ is maar een *differentiaalvergelijking*[*1] voor de verandering van de plaats, $d\vec{r}$, in een klein intervalletje van tijd, dt. Die verandering van plaats heeft een naam, de `snelheid`:

$$\vec{v} \equiv \lim_{\Delta t \to 0} \frac{\vec{r}(t+\Delta t) - \vec{r}(t)}{(t+\Delta t) - t} = \frac{d\vec{r}}{dt} \tag{3.3}$$ ♠

Merk op dat hierin niet alleen de homogeniteit van de ruimte is gebruikt (omdat we het *verschil* van twee posities bepalen), maar ook de homogeniteit van de tijd (omdat we door een tijdsverschil delen, en niet door een absolute tijdswaarde).

Nu zou je denken dat dan tenminste de vergelijking voor de snelheid v van een deeltje een algebraïsche vorm zou kunnen zijn, bijvoorbeeld

$$v = a + bt + ct^2 \tag{3.4}$$

maar ook dat is niet het geval. Dat komt doordat de absolute snelheid van deeltjes in ons Heelal blijkbaar geen meetbare grootheid is: aan niets is af te lezen wat de ruimtelijke snelheid van een deeltje is. Ook dat is geen 'principe'; in de studeerkamer kunnen we best een heelal verzinnen waarin deeltjes een soort inwendig wijzertje hebben waarop je de snelheid van het deeltje kunt aflezen. Maar het Heelal waarin wij wonen werkt niet zo. In onze natuur geldt de `Galilei-Huygens symmetrie`, dit is de invariantie

$$\vec{v} \quad \Longrightarrow \quad \vec{v} + \vec{w} \tag{3.5}$$ ♠

[*1] Zie bijvoorbeeld W.T. van Horssen, *Differentiaalvergelijkingen*, Epsilon Uitgaven, Utrecht 1993; J. Grasman, *Wiskundige methoden toegepast*, Epsilon Uitgaven, Utrecht 1992.

Als je bij de snelheid van een deeltje een willekeurige vaste waarde $\vec{w}$ optelt, verandert er niets. Met andere woorden, er geldt blijkbaar een soort 'relativiteit van de snelheid': de waarde van een snelheid als zodanig is niet waarneembaar. Dus is niet de absolute snelheid van een deeltje van belang, maar de *relatieve* snelheid, in het bijzonder de verandering van de snelheid in de loop van de tijd: de `versnelling`

$$\vec{a} \equiv \lim_{\Delta t \to 0} \frac{\vec{v}(t + \Delta t) - \vec{v}(t)}{(t + \Delta t) - t} = \frac{d\vec{v}}{dt} \tag{3.6}$$

Omdat alleen het verschil in snelheid telt, is de bewegingsvergelijking geen algebraïsche vergelijking voor de snelheid $\vec{v}$, maar een *differentiaalvergelijking* voor de verandering van de snelheid, $d\vec{v}$.

Hogere symmetrieën zijn er blijkbaar niet, want we kunnen een versnelling wèl absoluut meten (experiment met draaiende emmer). Ook dat is iets wat ons door de Natuur wordt voorgeschoteld, en waarvan we nog niet weten wat de diepere achtergrond is. Omdat er blijkbaar niet zoiets bestaat als een 'relativiteit van versnelling' is de bewegingsvergelijking een tweede-orde differentiaalvergelijking, de klassieke `bewegingsvergelijking`

$$\frac{d^2\vec{r}}{dt^2} = \frac{d\vec{v}}{dt} = \vec{a} = \text{“iets”} \tag{3.7} \spadesuit$$

waarin “iets” de versnelling ten gevolge van uitwendige invloeden. Dat kan van alles zijn; in dit college zullen wij zeer zwaar leunen op het voorschrift van de Newtonse zwaartekracht.

Om deze tweede-orde differentiaalvergelijking te kunnen oplossen, moeten wij dus zeven getallen invoeren: de beginpositie (drie stuks), de beginsnelheid (ook 3), en de begintijd. Die laatste doet er niet toe als de uitwendige versnelling niet van de tijd afhangt. Pas als we al deze gegevens hebben gekozen, kunnen we (in principe, tenminste) de vergelijkingen oplossen. Dat is dus heel iets anders dan een algebraïsche functie uitrekenen! Toch hebben veel mensen zo'n spoorrails-beeld in hun hoofd als ze aan bewegingen denken. We zien dat aardig terug in rare uitspraken van het type “het ruimtevoertuig is uit zijn baan geraakt.”

Een stel eenvoudige gevallen kunnen wij meteen oplossen. De meest voor de hand liggende is 'iets is niets':

$$\frac{d^2\vec{r}}{dt^2} = 0 \tag{3.8}$$

met de oplossing

$$\vec{v} = \text{constant} = \vec{v}_0 \ ; \qquad \vec{r} = \vec{r}_0 + \vec{v}_0 t \tag{3.9}$$

Met andere woorden, *een deeltje waarop geen kracht werkt, beweegt met constante snelheid.* Let op: snelheid is een vector, dus zowel de richting als de grootte van de snelheid blijven constant! Dit noemt men wel de `wet van de traagheid`. Die komt er dus uitrollen dankzij de symmetrieën die we in de bewegingsvergelijking hebben ingebouwd.

Oefening.

Van de klassieke oudheid tot de Middeleeuwen vond men niet de rechtlijnige beweging, maar de cirkelbeweging de 'meest ideale'. Als je dit opvat in de zin van “een deeltje waarop geen kracht werkt, beweegt met constante snelheid *op een cirkel*” dan zegt dit iets over een soort vervanger van de Galilei-Huygens symmetrie. Probeer daarvoor eens een formulering te vinden.

De volgende vorm van de bewegingsvergelijking veronderstelt dat de uitwendige versnelling constant is:

$$\frac{d^2\vec{r}}{dt^2} = \vec{g} \tag{3.10}$$

De oplossing hiervan is

$$\vec{v} = \vec{v}_0 + \vec{g}t \tag{3.11}$$

Met andere woorden: bij constante versnelling neemt de snelheid vanaf een gegeven beginwaarde $\vec{v}_0$ lineair toe (of af) met de tijd t. Dit is een `eenparig versnelde beweging`. De vergelijking voor $\vec{v}$ is weer een differentiaalvergelijking voor $\vec{x}$, die we oplossen als

$$\vec{r} = \vec{r}_0 + \vec{v}_0 t + \frac{1}{2}\vec{g}t^2 \tag{3.12}$$

Als wij even aannemen dat $\vec{g}$ langs de y-as gericht is, vinden we de beroemde paraboolbaan

$$y - y_0 = \frac{v_y}{v_x}(x - x_0) + \frac{g}{2v_x^2}(x - x_0)^2 \tag{3.13}$$

Galilei heeft daar zeer veel experimenteel onderzoek naar gedaan.

Wij zien hier het optreden van de `integratieconstanten` $\vec{v}_0$ en $\vec{r}_0$. Die komen er in omdat we zagen dat de ruimte homogeen is, dus de absolute positie van een voorwerp doet er niet toe in de bewegingsvergelijking; de ijking $\vec{x} = \vec{x}_0$ op tijd $t = 0$ moeten we dus achteraf vastleggen. Op soortgelijke manier moet ook de beginsnelheid $\vec{v}_0$ worden gegeven, omdat Galilei-Huygens symmetrie zegt dat absolute snelheden er niet toe doen. Dus hebben we een tweede-orde differentiaalvergelijking, die dan ook twee (vector-)constanten nodig heeft voor de oplossing.

4. Energie en impuls

Bij iedere symmetrie hoort een grootheid die niet verandert. In de klassieke mechanica zijn dat de energie en de impuls. De klassieke bewegingsvergelijking heeft in het rechterlid nul, in het geval dat er geen uitwendige krachten werken. Dit leidt tot behoud van impuls. De vergelijking kan in alle gevallen eenmaal worden geïntegreerd. Zo ontstaat een vorm waaruit blijkt dat een bepaalde grootheid bij de beweging hetzelfde blijft. Dit is de behoudswet voor de energie.

Om de bewegingsvergelijking op te lossen moeten we dus twee integratie-stappen ondernemen. Dat is niet altijd eenvoudig; sterker nog, in de meeste gevallen gaat het niet en moeten we numerieke methoden gebruiken. Maar één stap is altijd uit te voeren, namelijk het vinden van de energie-integraal. Kijken we eerst naar één dimensie. Dan hebben we

$$\frac{dv}{dt} = a \tag{4.1}$$

Dit kunnen we even anders schrijven met de truc (wiskundig behoeft dit nog wel wat onderbouwing!)

$$\frac{dv}{dt} = \frac{dr}{dr} \times \frac{dv}{dt} = \frac{dr}{dt}\frac{dv}{dr} = a \tag{4.2}$$

Nu weten we dat, per definitie, $dr/dt = v$ en dus

$$v\,\frac{dv}{dr} = \frac{1}{2}\frac{dv^2}{dr} = a \tag{4.3}$$

en zodoende

$$\frac{1}{2}v^2 = \mathcal{E} + \int a\,dr \tag{4.4}$$

De constante $\mathcal{E}$ is de energie per massa-eenheid. We kunnen dan schrijven

$$\frac{1}{2}m\,v^2 - \int ma\,dr = \text{constant} \equiv E \tag{4.5}$$ ♠

Deze constante is de 'eerste integraal' van het mechanische systeem, een grootheid die tijdens de beweging onveranderd blijft. Men beschrijft de vorm Eq.(4.5) weleens met de woorden 'kinetische energie plus potentiële energie is constant'. De hier gegeven schets is wat kort door de bocht en er zijn hier veel meer en veel nauwkeuriger dingen over te zeggen, maar daarvoor wordt verwezen naar de colleges klassieke mechanica. Voor de toepassingen hier is de vorm Eq.(4.5) bijna altijd voldoende.

Uit Eq.(4.1) kunnen wij nog een behoudswet afleiden, die geldt voor een systeem van deeltjes waarop geen netto kracht werkt. Dit is bijvoorbeeld het geval wanneer de krachten alleen maar tussen de deeltjes onderling werken. Voor elk deeltje, met rangnummer i, geldt

$$m_i \frac{dv_i}{dt} = m_i a_i = F_i \tag{4.6}$$

Nu sommeren we over alle deeltjes, en vinden dan

$$\sum_i m_i \frac{dv_i}{dt} = \sum_i F_i = F_{\text{tot}} \tag{4.7}$$

In een afgesloten systeem is per definitie de totale kracht F_{tot} nul, en dus is de totale impuls P, gedefinieerd door

$$P \equiv \sum_i m_i v_i = \text{constant} \tag{4.8} \spadesuit$$

constant. Merk op dat wij hier stiekem verondersteld hebben dat alle m_i constant zijn, maar men kan bewijzen dat dit niet hoeft, en dat toch de behoudswet Eq.(4.8) geldt. Bij dit bewijs, dat het makkelijkst te leveren is met de 'Lagrange-methode', wordt gebruik gemaakt van de translatie-symmetrie van de ruimte.

5. Beschrijven van kansen

Hoewel dit college gaat over astrofysica, en dus nauwelijks verwijst naar waarnemingen en meettechnieken, is het dienstig om hier het een en ander te vertellen over de theorie van metingen en meetfouten. Deze is geheel gebaseerd op de waarschijnlijkheidsrekening: ook fouten hebben hun wetten.

Het is de bedoeling te zien hoe we kansen in getallen kunnen uitdrukken, en ermee rekenen. Als je zeker weet dat iets niet gebeurt, noemen we de kans op die gebeurtenis 0. Als je zeker weet dat iets wèl gebeurt, noemen we de kans op die gebeurtenis 1. Dat getal mag van alles wezen, bijvoorbeeld 8 of 100, als je je maar aan de eenmaal gemaakte keuze houdt. Vaak gebruikt men 100, en noemt de kans dan een 'percentage'. Soms gebruikt men 1.000.000, en noemt de kans een ppm (*parts-per-million*). Maar de waarde 1 rekent het handigst. Bij deze keuze kunnen we de kans of waarschijnlijkheid P (voor *probabilitas*) uitrekenen, door het aantal malen N_A dat een bepaalde gebeurtenis A voorvalt te delen door het totale aantal gebeurtenissen N van welk type dan ook. Zo is de kans dat een willekeurig iets gebeurt, een getal van 0 tot en met 1:

$$P_A = \frac{N_A}{N} \tag{1}$$

Hierbij veronderstellen we wel dat we te maken hebben met gebeurtenissen die elkaar uitsluiten, en die gezamenlijk alle mogelijkheden overdekken; bijvoorbeeld of een persoon man of vrouw is, of dat een van de zes zijden van een dobbelsteen bovenkomt. Dan is de som van die kansen gelijk aan 1, want er gebeurt altijd *iets* (alle mogelijkheden zijn overdekt) en er gebeuren nooit twee dingen tegelijk (de mogelijkheden sluiten elkaar uit). Als $\diamondsuit$ staat voor de gebeurtenis 'ruitenaas trekken uit een compleet spel kaarten', dan is

$$P_\diamondsuit = \frac{1}{52} = 0,01923 = 1,92\,\% \tag{2}$$

Als $\clubsuit$ staat voor de gebeurtenis 'een klaverkaart trekken' dan is

$$P_\clubsuit = \frac{13}{52} = 0,25 \tag{3}$$

Zo kunnen we talloze kansberekeningen baseren op eenvoudige uitgangspunten. Merk op dat we steeds veronderstellen dat de gebeurtenissen *onderling onafhankelijk* zijn en dat zij *elkaar uitsluiten.* Deze schijnbaar onschuldige voorwaarde heeft verregaande gevolgen. Stel bijvoorbeeld dat ik wil weten wat de kans is op de gebeurtenis 'een aas trekken of een klaverkaart'. Er zijn 13 klaveren en 4 azen, maar... één van die klaveren is klaveraas! Dus is de kans daarop niet $(13+4)/52$ maar $(13+3)/52$, een verschil van 3,7% waar een gladde gokker zijn voordeel mee kan doen. In praktische gevallen zijn zulke afhankelijkheden (de technische term is *voorwaardelijke waarschijnlijkheid*) aan de orde van de dag, en aan het eind van dit stuk komen we daarop terug. Je kunt er heel goed mee rekenen, maar het is een kwestie van voortdurend oppassen.

Als twee gebeurtenissen A en B onafhankelijk zijn en elkaar uitsluiten, dan is de kans op A *of* B gegeven door

$$P_{A|B} = \frac{N_A + N_B}{N} = \frac{N_A}{N} + \frac{N_B}{N} = P_A + P_B \tag{4}$$

Dit noemen we de *optelregel voor onafhankelijke kansen.* De kans op A *èn* B ligt iets ingewikkelder, want het is een kwestie van mogelijke combinaties tellen. Vanwege dat verraderlijk korte 'èn' gaat het om *paren* van gebeurtenissen. Stel dat van een totaal N gebeurtenissen, N_A maal geval A voorkomt. En stel dat van de M gebeurtenissen, M_B maal B gebeurt. Als ik die combineer, heb ik $N \times M$ mogelijkheden voor paren, waarvan er $N_A \times M_B$ paren A *èn* B zijn. Dus is de kans

$$P_{A\&B} = \frac{N_A \times M_B}{N \times M} = \frac{N_A}{N} \times \frac{M_B}{M} = P_A \times P_B \tag{5}$$

Dit is de *productregel voor samengestelde kansen.* Deze regel is de basis voor de 'vertakkingen' die we dadelijk zullen tegenkomen.

Doet-ie het of doet-ie het niet?

In het hier beschreven voorbeeld gebruiken we een 'ding' dat in twee toestanden kan verkeren: het werkt of het werkt niet. Het ding bestaat uit onderdelen; een onderdeel is heel of het is stuk. Het ding werkt alleen als alle onderdelen heel zijn. We gaan berekenen wat de kans is dat een ding werkt, als we weten wat de kans is dat een onderdeel stuk is.

Stel dat het ding uit één onderdeel bestaat. Laat K de kans zijn dat het onderdeel stuk is. Dan is de kans dat het ding niet werkt ook K. De kans dat het ding wel werkt is $1-K$, want de som van de kansen moet 1 zijn.

Stel nu dat het ding uit *twee* onderdelen bestaat. Voorlopig nemen we aan dat de kans dat een onderdeel stuk is, altijd K is. Dat is niet erg realistisch, want bijvoorbeeld de accu van een auto of de banden gaan veel sneller stuk dan de cylinderwanden, maar het rekent makkelijker. Aan het eind van dit verhaal komen we hierop terug.

Vertakkingen en kansen

We rekenen* nu alsvolgt. In een fractie K van alle gevallen is onderdeel 1 stuk; de kans is $1-K$ dat het nog heel is. Hetzelfde geldt voor onderdeel 2. Hoe groot is nu de kans dat het ding werkt? We kunnen dit zien als een boom, die zich vertakt met takken ter dikte K en $1-K$, of een rivier, of zoiets. Elke tak vertakt zich weer, *in dezelfde verhouding*: als een tak een dikte D heeft, vertakt hij zich in een twijg met dikte $D \times K$ en een andere met dikte $D \times (1-K)$. Het aantal lagen van vertakking is gelijk aan het aantal onderdelen in het ding.

* Het komt hier en daar voor dat een berekening of bewijs wat lang of lastig is. In dat geval wordt verwezen naar aanhangsels aan het eind van dit stuk, gemerkt met ♠.

De tak met dikte K splitst zich in de verhouding $(K) : (1-K)$, dus de takken krijgen na twee splitsingen een dikte $K \times K = K^2$ en $K \times (1-K) = K - K^2$. De tak $(1-K)$ splitst zich in $(1-K) \times K = K - K^2$ en $(1-K) \times (1-K) = (1-K)^2 = 1 - 2K + K^2$. Tellen we alles op, dan komt er

$$K^2 + K - K^2 + K - K^2 + 1 - 2K + K^2 = 1 \tag{6}$$

dus we hebben een totale kans 1, dat klopt: alle boomtakken samen zijn net zo dik als de stam, ofwel er stroomt net zoveel water door de zijrivieren als uiteindelijk in zee uitkomt.

Alleen de tak waarlangs geen enkel onderdeel stuk is, levert een ding op dat werkt. Dus de kans dat een ding met twee onderdelen werkt, is $1-K$ tot de macht 2,

$$P_0 = (1-K)^2 = 1 - 2K + K^2 \tag{7}$$

We zullen deze kans P_0 noemen, de kans-op-nul-fouten. De kans dat het ding niet werkt noemen we P_f, en omdat de som van alle mogelijke kansen gelijk aan 1 moet zijn, is

$$P_f = 1 - P_0 \tag{8}$$

Een voorbeeld. Als de kans dat één onderdeel stuk is, gelijk is aan 0,1 (tien procent), dan is

$$P_0 = (1-0,1)^2 = 1 - 0,2 + 0,01 = 0,81 \tag{9}$$
$$P_f = 1 - P_0 = 0,19 \tag{10}$$

Benadering: Het is vaak handig en snel om gebruik te maken van het feit dat K in practische gevallen meestal klein is (je wilt immers met betrouwbare onderdelen werken). Als dat het geval is, dan zien we direct dat 1 veel groter is dan K, zodat K veel groter is dan K^2, die weer veel groter is dan K^3, enzovoorts. We gebruiken het symbool "$\simeq$" als we bedoelen 'is bij benadering gelijk aan'. Dan volgt uit het bovenstaande, voor *kleine* waarde van K (zeg b.v. kleiner dan 0,1; hoe kleiner K, hoe nauwkeuriger de schatting)

$$P_0 \simeq 1 - 2K \tag{11}$$
$$P_f \simeq 2K \tag{12}$$

Algemene vertakkingen

Stel nu dat het ding uit N onderdelen bestaat. Dan kun je een boom of rivier tekenen die zich $N-1$ maal vertakt. Je ziet direct dat de kans dat het ding werkt wordt gevonden langs slechts één tak: namelijk die waarlangs elk onderdeel heel is. Dus

$$P_0 = (1-K) \times (1-K) \times (1-K) \times \cdots \times (1-K) = (1-K)^N \tag{13}$$

Je moet dus het getal $1-K$ steeds met zichzelf vermenigvuldigen totdat je N factoren bij elkaar hebt.

Laten we eens zien wat dat oplevert. Boven hadden we al $N = 2$, dat wil zeggen $P_0 = (1-K)^2$, en voor hogere N komt er

$$(1-K)^3 = 1 - 3K + 3K^2 - K^3 \tag{14}$$
$$(1-K)^4 = 1 - 4K + 6K^2 - 4K^3 + K^4 \tag{15}$$
$$(1-K)^5 = 1 - 5K + 10K^2 - 10K^3 + 5K^4 - K^5 \tag{16}$$

Kijk even naar de cijfers vóór de machten van K. Die vormen een speciale figuur, als we de gevallen $N = 0, 1, 2, 3 \cdots$ van boven naar beneden opstapelen:

$$\begin{array}{lccccccccccccc}
N=0 & & & & & & & 1 & & & & & & \\
N=1 & & & & & & 1 & & 1 & & & & & \\
N=2 & & & & & 1 & & 2 & & 1 & & & & \\
N=3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
N=4 & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
N=5 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
N=6 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & & & &
\end{array} \tag{17}$$

Dat heet de *driehoek van Pascal.* We zullen hem nog tegenkomen (zie ♣**1**). Je kunt elke volgende rij eenvoudig berekenen door een 1 aan weerszijden te zetten (op de stippeltjes uiterst links en rechts), en het getal in een tussenruimte te vinden door de twee getallen erboven op te tellen. Zie bijvoorbeeld de rij voor $N = 5$: die staat zo omdat $1 + 4 = 5$ en $4 + 6 = 10$.

Gebruiken we nu de benadering dat K klein is, dan zien we meteen aan de driehoek dat

$$P_0 \simeq 1 - NK \tag{18}$$
$$P_f \simeq NK \tag{19}$$

Voorbeeld: als $K = 0,01$ (een procent), en $N = 20$, dan is

$$P_0 \simeq 1 - 20 \times 0,01 = 0,8 \tag{20}$$

Dus zelfs als de onderdelen redelijk betrouwbaar zijn (99%) gaat de kans dat het ding werkt snel omlaag als het uit veel onderdelen bestaat (80% bij 20 onderdelen).

Een belangrijke conclusie uit (18,19) is, dat de kans dat een ding niet werkt kan worden verkleind door *de betrouwbaarheid van de onderdelen te vergroten* (K kleiner maken), of door *het aantal onderdelen te verkleinen* (N kleiner maken), en dat *beide methoden evenveel aan het resultaat bijdragen* omdat P_f gelijk is aan het *product* van K en N. In de praktijk is het soms makkelijker om het aantal onderdelen te verkleinen dan om de betrouwbaarheid per onderdeel te vergroten.

Omgekeerd kunnen we de kans dat een onderdeel stuk is, uitrekenen als we weten wat de kans is dat een ding niet werkt, mits we ook weten uit hoeveel onderdelen het bestaat. We bepalen P_f uit waarnemingen, bijvoorbeeld door te kijken hoeveel dingen uit een partij van een miljoen niet werken. Zo berekenen we

$$K \simeq \frac{P_f}{N} \tag{21}$$

Als bijvoorbeeld een ding in 10 procent van de gevallen niet blijkt te werken, is $P_f = 0,1$. Weten we ook N, zeg $N = 50$, dan is

$$K \simeq \frac{0,1}{50} = 0,002 \tag{22}$$

dus in dit voorbeeld is de kans dat een onderdeel stuk is, 2 promille.

Een willekeurig ding stuk

Tot dusver zijn we ervan uitgegaan 'dat een ketting zo sterk is als de zwakste schakel'. Maar in de praktijk zijn kettingproblemen eigenlijk zeldzaam. Als je in een muur een enkele baksteen stukslaat, blijft de muur vrijwel altijd staan. Laten we dus eens kijken naar 'aftakeling' inplaats van 'niet werken'. We bezien een ding dat altijd min of meer werkt, maar des te slechter naarmate er meer onderdelen stuk zijn.

In het bovenstaande berekenden we de kans dat er *geen enkel* onderdeel stuk is. In (8, 13) zagen we dat die kans gegeven wordt door

$$P_0 = (1 - K)^N \simeq 1 - NK \tag{23}$$

Op soortgelijke manier zien we, dat de kans dat *alle* onderdelen stuk zijn gelijk is aan

$$P_N = K^N \tag{24}$$

Wat is nu de kans P_1 dat precies één onderdeel stuk is? laten we eerst het geval $N = 2$ bekijken. Uit de boomstructuur zien we direct dat er twee mogelijkheden zijn, namelijk

$$\begin{array}{ll} (1-K)K & \text{onderdeel 1 werkt, 2 is stuk} \\ K(1-K) & \text{onderdeel 2 werkt, 1 is stuk} \end{array} \tag{25}$$

Samen hebben we dus

$$P_1 = 2K(1-K) = 2K - 2K^2 \tag{26}$$

want elk van deze twee boomtakken levert dezelfde bijdrage.

In het algemene geval, dus met N onderdelen, is de kans langs één enkele boomtak

$$P_1 = K(1-K)^{N-1} \tag{27}$$

dat wil zeggen, een factor K voor de kans dat een onderdeel stuk is, en een factor $1-K$ voor elk van de $N-1$ onderdelen die nog heel zijn. Omdat elk van de N onderdelen stuk kan gaan, zijn er N takken die het resultaat (27) geven, en dus

$$P_1 = NK(1-K)^{N-1} \tag{28}$$

In benadering voor kleine K komt er

$$\begin{aligned} P_1 &= NK(1-K)^{N-1} \simeq NK\,(1-(N-1)K) \\ &= NK - N(N-1)K^2 \end{aligned} \tag{29}$$

Voor kleine K is dit dus vrijwel gelijk aan de eerder gevonden P_f. Dat is niet zo vreemd: als er maar een heel kleine kans is dat een enkel onderdeel stuk is, is de kans dat er twee stuk zijn dienovereenkomstig kleiner. Het exacte verschil is

$$\begin{aligned} P_f - P_1 &= 1 - (1-K)^N - NK(1-K)^{N-1} \\ &= 1 - (1-K)^{N-1}(1-K+NK) \end{aligned} \tag{30}$$

Hoeveel wegen leiden naar Rome?

Nog steeds uitgaande van de veronderstelling dat alle onderdelen hetzelfde zijn, of althans dezelfde kans hebben om stuk te zijn, vragen we ons af: wat is de kans dat niet slechts één, maar twee of meer onderdelen stuk zijn? In het algemeen komt dat neer op het beantwoorden van de vraag: op hoeveel manieren kunnen M van de N onderdelen stuk zijn? Als bijvoorbeeld $N = 2$, heb je de volgende mogelijkheden. Ofwel alles is stuk; dat kan maar op een enkele manier. Ofwel alle onderdelen zijn nog heel; ook dat kan maar op één manier. Ofwel er is precies 1 onderdeel stuk; dat kan het eerste onderdeel zijn of het tweede, dus 2 manieren. Kortom:

$$\begin{array}{cccccc} N=0 & & & 1 & & \\ N=1 & & 1 & & 1 & \\ N=2 & 1 & & 2 & & 1 \\ & \text{beide} & & \text{een} & & \text{geen} \\ M= & 2 & & 1 & & 0 \end{array} \tag{31}$$

Doen we hetzelfde spelletje met $N = 3$, dan komt er

$$\begin{array}{cccccccc} N=0 & & & & 1 & & & \\ N=1 & & & 1 & & 1 & & \\ N=2 & & 1 & & 2 & & 1 & \\ N=3 & 1 & & 3 & & 3 & & 1 \\ & \text{alle} & & & & & & \text{geen} \\ M= & 3 & & 2 & & 1 & & 0 \end{array} \tag{32}$$

We zien nu direct in, dat *het aantal manieren Δ^N_M waarop M van de N onderdelen stuk kunnen zijn, wordt gegeven door de driehoek van Pascal.* Voorwaarde is steeds dat we ons er niet om bekommeren welke onderdelen dat precies zijn; elke combinatie van M niet-werkende onderdelen wordt geacht hetzelfde resultaat te hebben.

Hoe komt het dat we Δ hier weer terugvinden? De reden is de bijzondere eigenschap van de driehoek van Pascal, namelijk dat elk getal wordt gevormd door de som van de twee getallen erboven. Zien we de driehoek als een boom of rivier, met de stam (of de monding) bij de eerste 1, dan herkennen we weer het verhaal uit bovengenoemde voorbeelden. Omlaag werkend in de driehoek hebben we op elke 'vork van de boomtakken' de keuze: doet-ie het, of doet-ie het niet? Zo ja, dan nemen we de vertakking naar (bijvoorbeeld) links, zo nee, dan nemen we de weg naar rechts. De kans op 'onderdeel is stuk', overeenkomend met 'vork naar rechts', is K; de kans op 'onderdeel is niet stuk', ofwel 'vork naar links', is $1-K$. Zodoende is de kans dat M willekeurige onderdelen stuk zijn

$$P^N_M = \Delta^N_M \, K^M (1-K)^{N-M} \tag{33}$$

Dit is de *kansverdeling van Bernoulli.* Hierin is Δ^N_M het aantal manieren waarop M van de N onderdelen stuk kunnen zijn, dus de Pascal-factor. Voor $M = 0$ krijgen we formule (8) weer terug voor de kans dat geen enkel onderdeel stuk is. Als $M = N$ vinden we de uitdrukking (24) voor de kans dat *alle* onderdelen stuk zijn. Dat hoort ook zo, want de uiterste getallen van de driehoek zijn altijd gelijk aan 1: $\Delta^N_0 = \Delta^N_N = 1$.

De kansverdeling van Bernoulli

Er zit echter nog een adder onder het gras. We hadden immers in het begin afgesproken dat *de som van alle kansen is 1.* Voor de driehoek van Pascal is de som van de getallen op een rij absoluut niet gelijk aan 1; integendeel, een beetje werk toont aan dat die som gegeven wordt door $1, 2, 4, 8, 16, \cdots$, dat wil zeggen

$$\text{som van rij } N \text{ is } 2^N \tag{34}$$

Bovendien zien we, dat de grootste waarde die op een rij bereikt wordt, de pan uitrijst wanneer N steeds groter wordt. We kunnen dat proberen te beteugelen door te delen door die factor 2^N:

$$\begin{array}{ccccccccccccccc}
 & & & & & & & 1{,}0 & & & & & & & \\
 & & & & & & 0{,}5 & & 0{,}5 & & & & & & \\
 & & & & & 0{,}25 & & 0{,}5 & & 0{,}25 & & & & & \\
 & & & & 0{,}13 & & 0{,}38 & & 0{,}38 & & 0{,}13 & & & & \\
 & & & 0{,}06 & & 0{,}25 & & 0{,}38 & & 0{,}25 & & 0{,}06 & & & \\
 & & 0{,}03 & & 0{,}16 & & 0{,}31 & & 0{,}31 & & 0{,}16 & & 0{,}03 & & \\
 & 0{,}02 & & 0{,}09 & & 0{,}23 & & 0{,}31 & & 0{,}23 & & 0{,}09 & & 0{,}02 & \\
0{,}01 & & 0{,}05 & & 0{,}16 & & 0{,}27 & & 0{,}27 & & 0{,}16 & & 0{,}05 & & 0{,}01
\end{array} \tag{35}$$

Deze driehoek voldoet natuurlijk *niet* meer aan de regel dat een getal gevonden wordt door de twee bovenliggende bij elkaar op te tellen!

Maar in het geval van de boven gevonden kansverdeling van Bernoulli,

$$P^N_M = \Delta^N_M \, K^M (1-K)^{N-M} \tag{36}$$

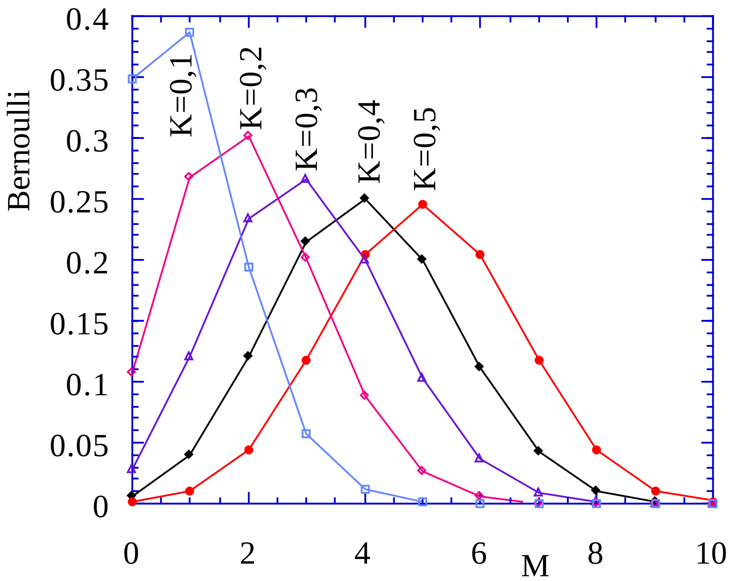
hebben we geluk. Het blijkt namelijk dat hiervoor *wel* geldt dat de som van de kansen 1 is (voor nadere toelichting, zie ♣**1**). Een grafiek van deze verdeling voor $N = 10$ achtereenvolgende 'worpen', en voor een kans K per worp die loopt van 0,1 tot 0,5 is gegeven in onderstaande figuur. Laten we bijvoorbeeld nemen $K = 0,5$ ('kruis of munt'), dan vinden we bij $N = 7$ de rij voor $M = 0$ tot en met $M = 7$,

$$P^N_M = \Delta^N_M \left(\frac{1}{2}\right)^N = 2^{-N} \Delta^N_M \tag{37}$$

0,0078	0,055	0,164	0,273	0,273	0,164	0,055	0,0078
M=7	M=6	M=5	M=4	M=3	M=2	M=1	M=0

(38)

Alles gaat mis, zeg je wel eens, maar je weet dat dat eigenlijk overdreven is, en (38) laat zien waarom. Zelfs als er slechts 7 dingen mis kunnen gaan, en we nemen aan dat de kans dat iets mis gaat 0,5 bedraagt, is de kans kleiner dan 8 promille dat werkelijk *alles* in de soep loopt. Daar staat tegenover dat de kans net zo klein dat er *niets* aan de hand is, dus 'Met mij gaat alles goed' is waarschijnlijk evenmin waar.



Wedden dat...

Wanneer je een lot koopt waarop fl. 25 wordt uitgekeerd als het wordt getrokken, en de kans op zo'n trekking is 10%, verwacht je gemiddeld fl. 2,50 te krijgen. Betaal je meer dan een knaak voor zo'n lot, dan word je beetgenomen. Dat is uiteraard bij *alle* loterijen het geval. Twee gulden zou een voordelige prijs zijn voor dat lot, want per trekking houd je er *gemiddeld* twee kwartjes aan over. Het is verlakkerij om te zeggen 'ik heb zo-en-zoveel gulden gewonnen', tenzij je daarvan eerst je totale inleg aftrekt. Niemand doet dat, om niet als ezel te boek te staan, want gemiddeld moet je dan bij alle bekende loterijen steeds van verlies spreken.

Hieruit zien we dat de 'waarde' van een kans gegeven wordt door *het product van de uitkomst en de kans daarop*:

$$W = M \times P_M \tag{39}$$

Wat is nu de te verwachten waarde als we alle mogelijke trekkingen in aanmerking nemen? Deze grootheid V, de *verwachtingswaarde* van de kansverdeling P, vinden we door (39) te sommeren over alle toegelaten waarden van M:

$$V = \sum_M M\, P_M \tag{40}$$

(zie ♣**2** voor de betekenis van het somteken Σ).

Laten we de verwachtingswaarde eens uitrekenen in een zeer eenvoudig geval: het werpen met een dobbelsteen. In dat geval is de kans op een uitkomst van 1 tot en met 6 altijd gelijk, namelijk 1/6 tenzij er met de dobbelsteen geknoeid is. Dus vinden we

$$V = \sum_{i=1}^{6} i \times P_i = \sum_{i=1}^{6} \frac{i}{6} = \frac{1+2+3+4+5+6}{6} = 3,5 \tag{41}$$

Stel dat je een spel speelt waarbij je evenveel guldens krijgt uitgekeerd als er ogen bij een worp bovenkomen, dan moet je per spel hoogstens f. 3,50 inzetten, anders ga je de boot in.

In het geval van de Bernoulli-verdeling (36) kunnen we de verwachtingswaarde uitrekenen volgens bovenstaand recept:

$$\begin{aligned} P_M^N &= \Delta_M^N \, K^M (1-K)^{N-M} \\ &= \frac{N(N-1)\cdots(N-M+1)}{1\times 2\times 3\times\cdots\times M} K^M (1-K)^{N-M} \end{aligned} \tag{42}$$

$$V = \sum_{M=0}^{N} M \, P_M^N = NK \tag{43}$$

(zie voor het bewijs ♣**3**). We vinden dus dat het verwachte aantal onderdelen dat stuk is, wordt gegeven door het product van het aantal N der onderdelen en de kans K dat een onderdeel stuk is. De bijzondere rol van het product NK kwamen we al veel eerder tegen (bijvoorbeeld in (18) en (19)). We zien nu dat dat geen toeval is, maar berust op een exacte eigenschap van de Bernoulli-kansverdeling. Ook hier concluderen we dus weer: je kunt de verwachtingswaarde verlagen (en dus de prestaties van het 'ding' verhogen) door het aantal onderdelen N waaruit het ding bestaat kleiner te maken, of de kans K dat een onderdeel stuk is te verkleinen, of beide. Wegens de productregel (43) werkt elk van deze alternatieven even goed. In de praktijk hangt het er natuurlijk van af wat het makkelijkst te bereiken is; meestal is het verkleinen van N de aangewezen weg.

Als je een ding koopt dat uit N onderdelen bestaat, en er is een kans K dat een willekeurig onderdeel stuk is, verwacht je dat gemiddeld NK onderdelen stuk zullen zijn. Voorbeeld: een huis bestaat uit minstens een half miljoen onderdelen: spijkers, buizen, stukken betonijzer, planken, loodslabben enzo. Laat de kans dat een willekeurig onderdeel stuk is, zeer klein zijn, zeg 1 op een miljoen. Dan verwachten we dat een opgeleverd huis $500.000/1.000.000 = 0,5$ kapotte onderdelen bevat. De *vereniging eigen huis* beweert dat elk nieuw opgeleverd huis 38 gebreken vertoont. Zo te zien doet de bouw-industrie het dus geweldig goed, namelijk 76 maal beter dan verwacht! Natuurlijk is deze berekening onzin: een kromme spijker ergens in een vloer heeft niet dezelfde gevolgen als een gescheurde loodslab. Ook is de kans op een kromme spijker niet dezelfde als de kans op een scheur in het lood. Op dit soort problemen komen we aan het eind van het verhaal terug.

Strooiing

Het getal V is een *verwachtings*waarde: er is geen garantie dat elke trekking uit een kansverdeling steeds de waarde V heeft. Sterker nog, dat kan soms helemaal niet: bij het gooien van een dobbelsteen is de verwachtingswaarde 3,5 – een getal dat je bij het dobbelen nooit boven zult zien komen.

Er is dus bij elke trekking een zeker verschil tussen de getrokken waarde en de verwachtingswaarde. De grootte van dat verschil heet *strooiing*, soms ook wel *spreiding* genoemd. We zullen deze strooiing aan het eind van dit verhaal tegenkomen in de vorm van het getal 'sigma', σ. Om de spreiding S te berekenen vinden we eerst de verwachtingswaarde van de variabele M, en vergelijken die met de verwachtingswaarde van *het kwadraat van* M. Dat doen we omdat een afwijking van V zowel positief kan zijn (als we dobbelen en gooien 6, is de afwijking $6 - 3,5 = 2,5$) als negatief (als we 1, 2 of 3 gooien). Door te kwadrateren tellen beide afwijkingen even sterk mee. Dus we nemen

$$S^2 = \sum_i i^2 \, P_i - V^2 = \sum_i i^2 \, P_i - \left(\sum_i i \, P_i\right)^2 \tag{44}$$

In het geval van de dobbelsteen vinden we

$$\begin{aligned} S^2 &= \frac{1}{6}(1+4+9+16+25+36) - 3,5^2 = 2,9166\ldots \\ S &= \sqrt{2,91667} = 1,708 \end{aligned}$$

Je moet dus niet raar opkijken als je bij het dobbelen een paar punten van de waarde 3,5 vandaan zit. Merk overigens op dat de strooiing S *niet* gelijk is aan het verschil tussen de verwachtingswaarde 3,5 en de uiterste waarde (6 of 1) die je kunt gooien! Het verschil is *kleiner*, een feit waar een beroepsgokker zijn voordeel mee kan doen.

De strooiing voor de Bernoulli-verdeling berekenen we uit

$$\sum_{M=0}^{N} M^2 \, P_M^N = NK(1-K) + N^2K^2 \tag{45}$$

(voor het bewijs, zie ♣**4**). Gebruiken we hierbij de verwachtingswaarde (43), dan vinden we

$$S^2 = NK(1-K) + N^2K^2 - V^2 = NK(1-K) \tag{46}$$

$$S = \sqrt{NK(1-K)} \tag{47}$$

Hieruit zien we iets zeer belangrijks: *de strooiing neemt slechts toe met de wortel uit het aantal onderdelen.* Als we naïef te werk gaan, zouden we dat niet hebben verwacht: de verwachtingswaarde V is recht evenredig met N, dus verdubbelen van het aantal onderdelen verdubbelt de verwachtingswaarde van het aantal falende delen. Doch de strooiing is *niet* recht evenredig met N, maar slechts met $\sqrt{N}$. Als we dus viermaal zoveel onderdelen gebruiken, neemt de strooiing slechts toe met een factor 2. Deze regel geldt veel algemener dan hier beschreven, en heet de *wet van de grote getallen.*

De Wet van De Moivre en het getal σ (sigma)

We zagen al eerder dat de kans dat M van de N onderdelen stuk zijn, wanneer alle onderdelen dezelfde kans K hebben om te falen, gegeven wordt door de Bernoulli-verdeling

$$P_M^N = \frac{N(N-1)\cdots(N-M+1)}{1\times 2\times 3\times\cdots\times M}\, K^M(1-K)^{N-M} \tag{48}$$

Wanneer we P_M^N tekenen, door bij een gegeven waarde van N te nemen $M = 0, 1, 2, 3, \cdots, N$, dan zien we dat de zo gevormde grafiek steeds minder van vorm verandert naarmate N groter wordt. Door de wiskundige De Moivre is bewezen dat de vorm die zo uiteindelijk ontstaat, in het geval $K = 0,5$, gegeven wordt door de *exponentiële functie* (zie ♣**5**)

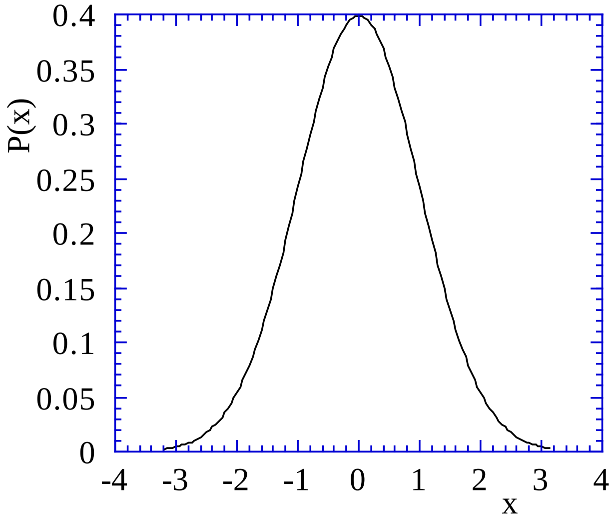
$$P(x) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-x^2/2\sigma^2} \tag{49}$$

Hierin is nu x een continue variabele, die de rol van de discrete waarden van M heeft overgenomen.

Later werd dit opnieuw gedaan door Gauss, en naar hem wordt dit vaak de ‘Gauss-functie’ genoemd (‘eer wie ere toekomt’ gaat in de wetenschap niet altijd op). De De Moivre-functie heet in de Angelsaksische literatuur *Gauss curve*, *normal distribution*, of ook wel *bell curve* (vanwege de klokvorm).

Door de constructie die we boven hebben gevolgd, kan deze kromme alsvolgt gelezen worden. Aan het ene uiterste ($x = -\infty$) staat de kans dat van een oneindig aantal mogelijkheden, *alles* fout gaat. Die kans is nul, want de exponent van een negatief oneindig getal is nul. Aan het andere uiterste ($x = +\infty$) staat de kans dat, van een oneindig aantal mogelijkheden, alles *goed* gaat. Die kans is eveneens nul. In het midden (bij $x = 0$) is de kansverdeling maximaal. Door vergelijken met de redeneringen die we tot dusver hebben gevolgd, zien we meteen dat hierbij de ene helft van alle dingen goed gaat, de andere helft fout. Het maximum bij $x = 0$ in de De Moivre-functie is de wiskundige uitdrukking van het gevoelen *’t kan vriezen, ’t kan dooien.*

Uit de vorm van de functie (36) zien we direct dat de kans op een bepaalde afwijking van het gemiddelde (bij $x = 0$) bepaald wordt door de verhouding x/σ. Het getal σ heet de *strooiing* of ook wel *standaarddeviatie*. In de wandeling zegt men wel 'een kans van twee sigma' als men de waarde van $P(x)$ bedoelt op het punt dat $x = 2\sigma$.



Vanwege het kwadraat in de exponent, neemt de functie $P(x)$ razendsnel af naarmate x/σ verder van nul af ligt. De kans dat een bepaald resultaat meer dan 2σ van het gemiddelde afwijkt is slechts 2%. Het 'aantal sigma' wordt wel kortweg gebruikt om de kwaliteit van een bepaald product of resultaat aan te duiden.

Wanneer iemand dus beweert dat de nagestreefde kwaliteitsnorm *six sigma* is, grenst dat aan opschepperij, want het betekent dat de kans op een fout van de orde van een op 10 miljoen is. Alleen in het geval van uiterst uitgebreide fout-correctieprocedures (zoals gebruikt in computer- en communicatiesoftware) kan dat ten naaste bij bereikt worden. Een voorbeeld van wat zoiets betekent: over de *Space Shuttle* werd oorspronkelijk beweerd dat de kans op een fatale fout tijdens de lancering ongeveer *five sigma* was, iets van de orde van 1 fout op 100.000 lanceringen. Dat lijkt een begrijpelijk getal, maar – zoals Feynman opmerkte – als je het nagaat, betekent dit dat je slechts één ongeluk verwacht als je *een ruimteveer per dag gedurende driehonderd jaar achtereen* zou lanceren. Zo uitgedrukt zie je meteen dat die foutkans grootspraak is, zoals op tragische wijze bleek bij de ontploffing van het ruimteveer *Challenger*. Dat was ruwweg de vijftigste lancering, overeenkomend met een kans van $2\,\sigma$.

Slotopmerking

Tenslotte moet nog worden opgemerkt dat het hele sigma-verhaal uitgaat van de sterk geïdealiseerde situatie zoals boven geschetst. In werkelijkheid zijn natuurlijk lang niet alle foutkansen gelijk aan dat ene getal K, en – erger nog – de kans op een bepaalde fout is meestal niet onafhankelijk van de andere mogelijke foutbronnen. Dan krijgen we te maken met *voorwaardelijke waarschijnlijkheid*, en gaat het sigma-sprookje niet door.

—◇—

Aanhangsel met meer wiskundige uitleg

♣1. Pascal-getallen

De driehoek van Pascal wordt gevormd door aan een rij getallen steeds een nieuwe toe te voegen volgens een bepaald recept. We beginnen met de rij genummerd $N = 0$; die rij bestaat uit het getal 1. Elke volgende rij heeft één getal meer dan de vorige, en elke rij begint en eindigt met 1. Daarmee ligt de rij voor $N = 1$ vast. De rij voor $N = 2$ is de eerste die een 'open plek' in het midden heeft. Elke open plek van een rij wordt gevuld met het getal dat gelijk is aan de som van de twee getallen in de rij erboven die ter linker- en ter rechterzijde staan. Zodoende komt er

$$\begin{array}{lccccccccccccc}
N=0 & & & & & & & 1 & & & & & & \\
N=1 & & & & & & 1 & & 1 & & & & & \\
N=2 & & & & & 1 & & 2 & & 1 & & & & \\
N=3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
N=4 & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
N=5 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
N=6 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
 & \cdots & \cdots & \cdots & & \cdots & & \cdots & & \cdots & \cdots & & \cdots &
\end{array}$$

We willen nu uiteraard weten wat de algemene formule is voor de N-de regel van de driehoek, want het is teveel gedoe om het steeds helemaal uit te rekenen. Als we het elke keer weer van voren af aan willen doen, moeten we steeds met het Pascal-recept een hele pyramide bouwen van de vorm

$$\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 1{+}6 & & 6{+}15 & & 15{+}20 & & 20{+}15 & & 15{+}6 & & 6{+}1 & & 1 \\
1 & & =7 & & =21 & & =35 & & =35 & & =21 & & =7 & & 1
\end{array}$$

Op de N-de rij staan de Pascal-getallen Δ_M^N voor $M = 0$ tot en met $M = N$. Na enig prutsen vermoeden we dat

$$\Delta_M^N = \frac{N(N-1)(N-2)\cdots(N-M+1)}{1 \times 2 \times 3 \times \cdots \times M}$$

Proberen we het eens voor $N = 4$ en $M = 3$, dan zien we dat het klopt:

$$\Delta_3^4 = \frac{4 \times 3 \times 2}{1 \times 2 \times 3} = 4$$

Het algemene bewijs berust op het feit dat we uit regel N altijd die voor $N+1$ kunnen uitrekenen. Tellen we, volgens het Pascal-voorschrift, het $(M-1)$-de getal van regel N op bij het M-de getal van diezelfde regel, dan vinden we

$$\begin{aligned}
\mathcal{P}_M^{N+1} =& \frac{N(N-1)(N-2)\cdots(N-(M-1)+1)}{1 \times 2 \times 3 \times \cdots \times (M-1)} \\
&+ \frac{N(N-1)(N-2)\cdots(N-M+1)}{1 \times 2 \times 3 \times \cdots \times M}
\end{aligned}$$

Nu brengen we dit geheel op één noemer door in de eerste term de teller en de noemer met M te vermenigvuldigen. Brengen we de gemeenschappelijke factor buiten haakjes, dan kunnen we schrijven

$$\mathcal{P}_M^{N+1} = \frac{N(N-1)\cdots(N+1-M+1)}{1 \times 2 \times 3 \times \cdots \times M} \times (M + (N-M+1))$$

We zien direct dat dit precies hetzelfde is als de formule voor Δ_M^N, maar nu met $N+1$ inplaats van N, want in de laatste factor tussen haakjes valt M weg tegen $-M$ zodat we $N+1$ overhouden, en deze factor zetten we links:

$$\mathcal{P}_M^{N+1} = \frac{(N+1)N(N-1)\cdots((N+1)-M+1)}{1 \times 2 \times 3 \times \cdots \times M}$$

Hiermee is bewezen dat het M-de Pascal-getal op de N-de regel voor alle M en N gegeven is door

$$\Delta_M^N = \frac{N(N-1)\cdots(N-M+1)}{1 \times 2 \times \cdots \times M}$$

De Pascal-getallen Δ zijn van groot belang omdat zij in doorlopende vermenigvuldigingen ('machten') voorkomen. De algemene vorm van (8) is, voor een product van n factoren,

$$(a+b) \times (a+b) \times (a+b) \cdots \times (a+b) = (a+b)^n$$

Uitschrijven levert het volgende resultaat:

$$(a+b)^n = \Delta_0^n a^0 b^n + \Delta_1^n a^1 b^{n-1} + \Delta_2^n a^2 b^{n-2} + \cdots + \Delta_n^n a^n b^0$$

waarin Δ_m^n het m-de Pascal-getal op rij n voorstelt. Zo'n oneindige som wordt afgekort met de notatie (zie ♠**2**)

$$(a+b)^n = \sum_{m=0}^{n} \Delta_m^n a^m b^{n-m}$$

(merk op dat $a^0 = 1$ en $a^1 = a$). Door deze uitdrukking te vergelijken met de formule voor de kansverdeling van Bernoulli, zien we dat deze sommeert tot 1:

$$\sum_{M=0}^{N} P_M^N = \sum_{M=0}^{N} \Delta_M^N K^M (1-K)^{N-M} = (K + (1-K))^N = 1 \tag{50}$$

♣2. Sommering

De hoofdletter sigma, Σ, wordt gebruikt als afkorting om de som van een rij getallen aan te geven. Zo wordt de som van n getallen

$$a_1 + a_2 + a_3 + \cdots + a_n$$

afgekort weergegeven door

$$\sum_{i=1}^{n} a_i$$

hetgeen te lezen is als *de som van alle getallen a_i waarin i de waarden van 1 tot en met n aanneemt.* Zo schrijven we bijvoorbeeld

$$\sum_{i=0}^{n} a_i = a_0 + a_1 + a_2 + a_3 + \cdots + a_n$$

en, voor een oneindige som,

$$\sum_{i=0}^{\infty} a_i = a_0 + a_1 + a_2 + a_3 + \cdots$$

♣3. Verwachtingswaarde van de Bernoulli-kansverdeling

De verwachtingswaarde V van een discrete kansverdeling P is gedefinieerd als

$$V = \sum_i i\, P_i$$

In het geval van de Bernoulli-verdeling moeten we dus berekenen

$$\begin{aligned}
V &= \sum_{M=0}^{N} \Delta_M^N \, M \, K^M (1-K)^{N-M} \\
&= \sum_{M=1}^{N} K^M (1-K)^{N-M} \frac{N(N-1)\cdots(N-M+1)}{1\times 2\times 3\cdots\times(M-1)}
\end{aligned}$$

We halen nu een factor NK links van het somteken en schrijven

$$\begin{aligned}
V &= NK \sum_{M=1}^{N-1} K^{M-1}(1-K)^{N-M} \times \\
&\qquad \frac{(N-1)(N-2)\cdots(N-M+1)}{1\times 2\times 3\cdots\times(M-1)} \\
&= NK \sum_{M=1}^{N-1} K^{M-1}(1-K)^{(N-1)-(M-1)} \times \\
&\qquad \frac{(N-1)(N-2)\cdots((N-1)-(M-1)+1)}{1\times 2\times 3\cdots\times(M-1)}
\end{aligned}$$

Nu zien we iets aardigs: de som over M is *precies* van de vorm gegeven in ♣**1**, met dien verstande dat $m = M-1$ en $n = N-1$. We concluderen dus dat

$$V = NK\,(K + (1-K))^{N-1} = NK$$

hetgeen we wilden bewijzen. We gebruikten deze truc ook al in ♣**1** . De Bernoulli-verdeling is dus, dankzij de bijzondere eigenschappen van de Pascal-getallen, een zeer nette functie.

♣4. Strooiing van de Bernoulli-kansverdeling

Ook als we de strooiing van de Bernoulli-vorm uitrekenen kunnen we gebruik maken van de kunstgreep uit ♣**3**. We moeten berekenen

$$\sum_{M=0}^{N} K^M (1-K)^{N-M} \, M^2 \frac{N(N-1)\cdots(N-M+1)}{1\times 2\times 3\cdots\times M}$$

Meteen een factor M van het kwadraat in de teller wegstrepen tegen de M in de noemer, heeft weinig zin omdat we dan nog met een M in de teller blijven zitten. We gebruiken de truc

$$M^2 = M(M-1+1) = M + M(M-1)$$

De eerste term M behandelen we zoals in ♣**3**; dat levert een waarde NK op, zoals we al zagen. In de tweede term $M(M-1)$ kunnen we zowel een factor M als een factor $M-1$ wegstrepen tegen $M\times(M-1)$ in de noemer. Dan houden we precies dezelfde vorm over als in ♣**3**, behalve dan dat we inplaats van NK een factor $NK\times(N-1)K$ links van het somteken halen. Zo vinden we

$$\begin{aligned}
&\sum_{M=0}^{N} K^M (1-K)^{N-M} \, M^2 \ldots = \\
&\qquad NK + NK(N-1)K = NK(1-K) + N^2K^2
\end{aligned}$$

We wisten al dat de verwachtingswaarde NK is, en dus

$$\begin{aligned}
S^2 &= NK(1-K) + N^2K^2 - (NK)^2 = NK(1-K) \\
S &= \sqrt{NK(1-K)}
\end{aligned}$$

De strooiing neemt dus slechts met de *wortel* uit N toe. Hoe groter het aantal onderdelen N, hoe zekerder je bent dat het aantal te verwachten fouten NK is.

♣5. Exponentiële functie

De *exponentiële functie* is gedefinieerd door de oneindige som

$$\mathrm{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{1 \times 2 \times 3 \times \cdots \times n}$$

Nemen we hierin $x = 1$, dan blijkt dat het *exponentiële grondgetal* e de waarde heeft van

$$\mathrm{e}^1 = \mathrm{e} = \sum_{n=0}^{\infty} \frac{1}{1 \times 2 \times 3 \times \cdots \times n} = 2,71828182846\ldots$$

De formule van De Moivre geeft zo'n exponentiële functie, maar dan niet met de exponent x maar met $-x^2/2\sigma^2$. Deze vorm wordt bereikt door van de variabele M over te gaan op

$$x = (M - V)\frac{\sigma}{\sqrt{NK(1-K)}}$$

dat wil zeggen een variabele die de verwachtingswaarde nul heeft (omdat we NK van M hebben afgetrokken), en een strooiing σ (omdat we door $\sigma\sqrt{NK(1-K)}$ hebben gedeeld).

6. De Boltzmann-verdeling

In een gas in evenwicht heeft niet elk deeltje dezelfde energie. Afgezien van systematische bewegingen (stroming) hebben de deeltjes een *thermische energie*, die overeenkomt met een willekeurige beweging met een bepaalde snelheid. De energieën zijn verdeeld volgens een bepaalde wiskundige vorm, de Gauss-kromme. Zo'n distributie heet een Boltzmann-verdeling.

Beschouw een klein deelvolume in een groot vat waarin een gas zit. Neem het volume zo klein dat de verschillen in de toestand van de deeltjes van de ene kant van het volume naar de andere gemiddeld te verwaarlozen zijn. Dat wil zeggen dat de *gradiënten* onbelangrijk zijn: het gas in het kleine volume is *homogeen.*

Wij bepalen nu de gemiddelde snelheid van de deeltjes, door alle snelheden (pijltjes) bij elkaar op te tellen en te delen door het aantal deeltjes in het volume. Die gemiddelde snelheid, de 'windsnelheid', trekken we van alle snelheden af zodat we met de wind meewaaien. Dan hebben de deeltjes een gemiddelde snelheid nul, maar dat betekent niet dat alle snelheden ook exact nul zijn. Er is een zekere kansverdeling: de meeste deeltjes zullen wel niet zo hard gaan, want de deeltjes botsen vaak tegen elkaar aan, maar soms zal het wel voorkomen dat een deeltje wat harder of langzamer gaat dan gemiddeld.

Die kansverdeling kunnen wij uitrekenen met het volgende argument. Als iedere botsing die een deeltje ondergaat, een zekere kans oplevert dat het deeltje een hoeveelheid ΔE aan energie wint, wat is dan de totale kans P dat het deeltje een energie E heeft?

Als het deeltje k botsingen ondergaat waarbij elke keer een hoeveelheid ΔE gewonnen wordt, is de uiteindelijke energie

$$E = k\,\Delta E \tag{6.1}$$

maar natuurlijk is niet ieder deeltje zo 'gelukkig' dat het k maal achter elkaar 'wint'. Er moeten ook verliezers zijn, wegens behoud van energie. Het is het eenvoudigst om te kijken wat de kans is dat de 'winnende reeks' van het deeltje *niet* wordt onderbroken. We hebben

$$\text{kans op verdergaan} = 1 - \text{kans op stoppen} = 1 - p\,\Delta E \tag{6.2}$$

met zekere constante p, die we later zullen bepalen. Doen we dit k maal, dan is de totale kans P gelijk aan het product van alle afzonderlijke kansen, en dus

$$P = (1 - p\,\Delta E)^k \tag{6.3}$$

Uit (1.1) weten wij dat $\Delta E = E/k$, en dus

$$P = \left(1 - \frac{p}{k}E\right)^k \tag{6.4}$$

In de limiet voor zeer veel stapjes ($k \to \infty$) wordt dit

$$P = \mathrm{e}^{-pE} \tag{6.5}$$

De constante p is een maat die de gemiddelde energie van de deeltjes normaliseert; immers, in (1.5) moet de exponent dimensieloos zijn, dus p moet zich gedragen als $[p] = 1/[\text{energie}]$. Door een historisch toeval was er reeds een naam voor de gemiddelde energie per deeltje: de *temperatuur* T. Wij zouden dus kunnen zeggen $p = 1/T$, maar helaas gebruikte men bij stoommachines en locomotieven geen MKSA-eenheden. Dus moest er een extra constante bij komen om van een energiemaat (joule) naar temperatuur (kelvin) over te stappen. Zodoende wordt

$$p = \frac{1}{kT} \tag{6.6}$$

waarin k de *constante van Boltzmann*. Omdat wij bovendien weten dat $E = \frac{1}{2}mv^2$, vinden we tenslotte de *Boltzmann-verdeling*

$$P = \mathrm{e}^{\frac{-mv^2}{2kT}} \tag{6.7}$$

7. Speciale relativiteitstheorie

Bij nader inzien blijkt de Galilei-Huygens symmetrie niet exact te zijn. Daarvoor in de plaats komt Lorentz symmetrie, die de lichtsnelheid invariant laat. De wiskundige vorm van deze symmetrie wordt afgeleid naar analogie van de klassieke bewegingsvergelijkingen. Een paar eenvoudige toepassingen zijn: de tijddilatatie, het optellen van snelheden, en beweging met een constante versnelling.

In hoofdstuk 3 werden de bewegingsvergelijkingen van de klassieke mechanica afgeleid op grond van de Galilei-Huygens symmetrie. Maar bij nader inzien blijkt, dat onze natuur niet exact aan die symmetrie voldoet. Uit de experimenten van Michelson en Morley (1887-1888) volgt namelijk dat het licht *niet* gehoorzaamt aan die symmetrie. Als licht beweegt met een snelheid c, zouden we volgens Huygens verwachten dat die snelheid slechts relatief is; meebewegen met diezelfde snelheid (en dat kan, door een Galilei-Huygens-transformatie) zou dan een effectieve snelheid $c - c = 0$ opleveren. Maar het Michelson-Morley experiment (dat de afgelopen eeuw in vele varianten, met steeds grotere precisie is herhaald) toont aan dat *de lichtsnelheid niet afhangt van de bewegingstoestand van de bron of de ontvanger.* De lichtsnelheid is niet relatief, maar absoluut: de lichtsnelheid is `invariant`.

Dat is totaal strijdig met Galilei-Huygens symmetrie, en daarmee komt de klassieke mechanica op losse schroeven te staan. Het gaat er nu om, een mechanica te bedenken waarin c invariant is. Je zou zo denken dat een dergelijke constructie 'absoluutheidstheorie' of zoiets zou worden genoemd, maar de naam is 'relativiteitstheorie' geworden (Einstein, de bedenker van de theorie, vond die naam ook maar niks). Om in vogelvlucht *2 te zien hoe deze werkt, gaan we eerst na hoe een G-H-transformatie wordt geformuleerd. Wij beperken ons even tot één ruimte-dimensie. Laat x de positie zijn van een deeltje op tijd t. Het deeltje wordt dus beschreven met de coördinaten (x, t) in een stelsel $\mathcal{K}$. Stel dat een waarnemer in een ander stelsel, $\mathcal{K}'$, met een snelheid w beweegt ten opzichte van $\mathcal{K}$. De coördinaten in $\mathcal{K}'$ zijn (x', t') en volgens Huygens geldt dat

$$x' = x + wt \tag{7.1}$$

$$t' = t \tag{7.2}$$

Een snelheid is $v = dx/dt$, en dus volgt hieruit dat

$$v' = \frac{dx'}{dt'} = \frac{dx'}{dt} = \frac{dx}{dt} + w = v + w \tag{7.3}$$

zoals verwacht. Het is duidelijk dat er, onder de transformatie Eq.(7.3) , nooit een invariante snelheid kan bestaan.

De vraag is nu: wat moet er in de plaats komen voor Eq.(7.1,2) zodat er wèl een invariante snelheid is? Dit kunnen wij beantwoorden door ons eerst te realiseren dat het bestaan van zo'n bijzondere snelheid betekent dat *ruimte en tijd iets met elkaar te maken hebben.* Immers, er geldt

$$\text{snelheid} = \frac{\text{ruimte}}{\text{tijd}} \tag{7.4}$$

dus meters per seconde, kilometers per uur, en dergelijke. Als een snelheid invariant is, dan moeten teller en noemer samenspannen om ervoor te zorgen dat het quotiënt steeds hetzelfde blijft. Ruimte en tijd hebben dus iets met elkaar te maken, sterker nog, ze kunnen met een en dezelfde maat worden gemeten, door als maat voor de afstand de reistijd van het licht te nemen (omdat c invariant is, is dat een prima manier). De afstand van de Aarde naar de Maan is 1.3 seconde, naar de Zon 8.3 minuten, naar de Andromeda Nevel 2 miljoen jaar.

Blijkbaar stelt de Natuur ons de eis: "Ontwerp een mechanica waarin c onveranderd blijft." De symmetrie die c invariant maakt heet `Lorentz symmetrie`. Hieronder zullen we zien op welke manier je zoiets aanpakt, als illustratie van het thema *een formule is er niet in de eerste plaats om in je rekenmachine te proppen, maar om naar je hand te zetten en te analyseren.*

Omdat tijd en ruimte in deze zin 'hetzelfde zijn', ligt het voor de hand om de transformatie Eq.(7.1,2) uit te breiden tot een symmetrische vorm:

$$x' = L_{xx}x + L_{xt}ct \tag{7.5}$$

$$ct' = L_{tx}x + L_{tt}ct \tag{7.6}$$

Dit is eigenlijk de belangrijkste stap van de hele behandeling. Uit het experimentele feit dat c invariant is, concluderen wij dat ruimte en tijd met dezelfde maat te meten zijn, en staan wij toe dat een term evenredig met x verschijnt in de transformatieregel voor de tijd t. Dat betekent iets heel merkwaardigs: namelijk, dat tijd *relatief* kan zijn, omdat het nu niet langer gegarandeerd is dat $t' = t$. Gegeven de mogelijkheid dat x en t in een mengvorm optreden, is het verstandig om ze met dezelfde maat te meten. De beste manier daarvoor is zulke eenheden te kiezen dat $c \equiv 1$, maar dat is onder sterrenkundigen helaas niet gebruikelijk. Daarom schrijven we in Eq.(7.5,6) ct inplaats van t. Eq.(7.5,6) komen in de plaats van Eq.(7.1,2) .

Het gaat er uiteraard om, de matrix L te bepalen. Laten we eerst eens zien hoe zoiets gaat in een wat minder exotisch geval: draaiingen. Bij een rotatie van coördinaten (x, y) naar (x', y') hebben we

$$x' = R_{xx}x + R_{xy}y \tag{7.7}$$

$$y' = R_{yx}x + R_{yy}y \tag{7.8}$$

Omdat bij een draaiing alle lengtes hetzelfde blijven, hebben wij

$$x^2 + y^2 = x'^2 + y'^2 = r^2 \tag{7.9}$$

Merk op dat het opleggen van de invariantie Eq.(7.9) een voorschrift is voor het afstandsrecept in de ruimte (Pythagoras)! In een ander type ruimte zou dat wel eens helemaal anders kunnen zijn.

---

*2 Uitgebreide uitleg is te vinden in E.F. Taylor & J.A. Wheeler, *Spacetime phycics,* Freeman, New York 1966; D. Bohm, *The special theory of relativity,* Routledge, London, 1996; A.P. French, *Special relativity,* Chapman & Hall, London, 1997.

Oefening.

Een cirkel is de verzameling van alle punten met een vaste afstand tot een gegeven punt. Teken de cirkels die overeenkomen met het afstandsrecept Eq.(7.9) , met $r^2 = |x| + |y|$ en met $r^2 = x^2 - y^2$.

Passen we Eq.(7.9) toe op Eq.(7.7,8) dan krijgen wij de eisen

$$R_{xx}^2 + R_{yx}^2 = 1 \ ; \quad R_{xy}^2 + R_{yy}^2 = 1 \ ; \quad R_{xx}R_{xy} + R_{yx}R_{yy} = 0 \tag{7.10}$$

Hieraan wordt voldaan door Eq.(7.7,8) te schrijven als

$$x' = x \cos\phi - y \sin\phi \tag{7.11}$$

$$y' = x \sin\phi + y \cos\phi \tag{7.12}$$

waarin ϕ de draaiingshoek.

Op soortgelijke manier gaan we nu de coëfficiënten vinden in Eq.(7.5,6) . De eerste eis die we stellen is, dat de beweging van het licht in de coördinatenstelsels $\mathcal{K}$ en $\mathcal{K}'$ hetzelfde is:

$$x = \pm ct \tag{7.13}$$

$$x' = \pm ct' \tag{7.14}$$

Met behulp van Eq.(7.5,6) volgt hieruit, als we eerst het geval $x = +ct$ nemen, dat

$$ct' = (L_{xx} + L_{xt})\, ct \tag{7.15}$$

$$ct' = (L_{tx} + L_{tt})\, ct \tag{7.16}$$

en dus

$$L_{xx} + L_{xt} = L_{tx} + L_{tt} \tag{7.17}$$

$$L_{xx} - L_{xt} = -L_{tx} + L_{tt} \tag{7.18}$$

waarin de tweede regel op dezelfde manier is afgeleid als de eerste, maar dan voor licht dat de andere kant opgaat, $x = -ct$. *Merk op dat er van deze hele truc niets terecht zou komen als c niet invariant was*, want dan zouden we hebben $x' = \pm c't'$ en schoten we niets op. Uit Eq.(7.17,18) volgen de eerste twee voorwaarden waaraan de transformatiematrix L moet voldoen:

$$L_{xx} = L_{tt} \tag{7.19}$$

$$L_{xt} = L_{tx} \tag{7.20}$$

Hoe komen we nu aan verdere voorwaarden voor L? Het uiteindelijke doel is immers om alle vier de componenten van L dwingend voor te schrijven uit de eis dat c invariant blijft. Merk op dat het helemaal niet voor de hand ligt dat dat ook echt kan! We gaan net zo te werk als in Eq.(7.3) en schrijven

$$\frac{v'}{c} = \frac{dx'}{c\,dt'} = \frac{L_{xx}dx + L_{tx}c\,dt}{L_{tx}dx + L_{xx}c\,dt} = \frac{L_{xx}v + L_{tx}c}{L_{tx}v + L_{xx}c} \tag{7.21}$$

Vervolgens stellen we vast dat voor snelheden ver beneden die van het licht, de GH-symmetrie wel degelijk goed werkt. Dus eisen wij dat, in de limiet voor $c \to \infty$, Eq.(7.21) het resultaat geeft dat $v' = v$, en dus

$$\frac{v}{c} = \frac{v'}{c} \simeq \frac{L_{tx}c}{L_{xx}c} = \frac{L_{tx}}{L_{xx}} \tag{7.22}$$

Er blijft dus nog maar één grootheid te bepalen over, namelijk L_{xx}, en wij kunnen Eq.(7.5,6) schrijven als

$$x' = L_{xx}(x + vt) \tag{7.23}$$

$$ct' = L_{xx}(\frac{v}{c}x + ct) \tag{7.24}$$

Door Eq.(7.23,24) van elkaar af te trekken komt er

$$x' - ct' = L_{xx}(1 - \frac{v}{c})(x - ct) \tag{7.25}$$

Nu komt het slotstuk: in de transformatie die door L wordt beschreven, beweegt het coördinatenstelsel $\mathcal{K}'$ met een snelheid v ten opzichte van $\mathcal{K}$. Uiteraard moet L zo zijn, dat een terugtransformatie met een snelheid $-v$ weer de oorspronkelijke toestand oplevert. Dus hebben we dat naast Eq.(7.25) moet gelden

$$x - ct = L_{xx}(1 + \frac{v}{c})(x' - ct') \tag{7.26}$$

Door substitutie van Eq.(7.25) in Eq.(7.26) komt er

$$L_{xx}^2 \left(1 - \frac{v^2}{c^2}\right) = 1 \tag{7.27}$$

waarmee tenslotte de hele matrix L is vastgelegd:

$$L = \frac{1}{\sqrt{1 - (v/c)^2}} \begin{pmatrix} 1 & v/c \\ v/c & 1 \end{pmatrix} \tag{7.28}$$

Dit is de matrix van de `Lorentztransformatie`, die in de plaats komt van de Galilei-Huygens transformatie Eq.(7.1,2) .

Deze transformatie kan uiteraard ook worden uitgeschreven in componenten. Het is gebruikelijk om dat wat af te korten, door gebruik te maken van de definities

$$\beta \equiv \frac{v}{c} \tag{7.29}$$

$$\gamma \equiv \frac{1}{\sqrt{1 - \beta^2}} \tag{7.30}$$

waarmee de Lorentztransformatie een plezierig symmetrische vorm aanneemt:

$$x' = \gamma(x + \beta ct) \tag{7.31}$$

$$ct' = \gamma(\beta x + ct) \tag{7.32}$$

Oefening.

Bewijs expliciet dat L in de limiet voor $c \to \infty$ inderdaad de Galilei-Huygens-transformatie oplevert. Doe dit door een Taylor-ontwikkeling toe te passen op de wortelvorm in Eq.(7.28) .

Oefening.

Bewijs met behulp van Eq.(7.31,32) dat de grootheid $s^2 = c^2t^2 - x^2$ invariant is onder Lorentztransformaties. Men noemt s het `interval`. De invariantie van het interval is analoog aan de invariantie van de straal r van de cirkel in Eq.(7.9) .

De Lorentztransformatie is de basis van de relativistische mechanica. Hier laten we daarvan maar een zeer klein stukje zien. Ten eerste een stelling over het optellen van snelheden in één dimensie. Laat een deeltje een snelheid u hebben in het stelsel $\mathcal{K}$ en laat $\mathcal{K}'$ met een snelheid v bewegen ten opzichte van $\mathcal{K}$. Met welke snelheid beweegt het deeltje gezien vanuit $\mathcal{K}'$? Passen wij Eq.(7.21) toe op u dan zien we (met behulp van Eq.(7.28)) dat

$$\frac{u'}{c} = \frac{dx + v\,dt}{(v/c)dx + c\,dt} = \frac{u + v}{uv/c + c} \tag{7.33}$$

en zodoende

$$u' = \frac{u + v}{1 + uv/c^2} \tag{7.34}$$

Oefening.

Bewijs uit Eq.(7.34) dat de lichtsnelheid de grootst mogelijke snelheid is, door in te vullen $v = c$.

Uit Eq.(7.34) kunnen wij ook nog een vergelijking afleiden voor relativistische beweging onder invloed van een constante kracht. Dit is een nuttige oefening, omdat het laat zien hoe je moet oppassen met uitspraken over snelheden in een relativistische omgeving. Evenals boven beperken wij ons hier tot één ruimte-dimensie.

Stel dat wij een ruimteschip waarnemen dat ten opzichte van ons coördinatenstelsel $\mathcal{K}$ een snelheid v heeft. Aan boord van het ruimteschip gebruikt men stelsel $\mathcal{K}'$, waarin $v' = 0$, want ten opzichte van zichzelf staat het schip stil. De stoker van het ruimteschip gooit er een schepje op, en bereikt daarmee een kleine versnelling $\delta v'$. Met 'klein' wordt uiteraard bedoeld $|\delta v'| \ll c$: in de klassieke mechanica heeft zo'n bewering geen zin, omdat er geen absolute snelheid bestaat en je dus ook niet van 'snel' of 'langzaam' kunt spreken! Gezien door ons verandert de snelheid van het schip van v naar $v + \delta v$. Door de transformatie $\mathcal{K}' \to \mathcal{K}$ in Eq.(7.34) vinden we het verband tussen $\delta v'$ en δv:

$$\delta v' = (v + \delta v)' = \frac{(v + \delta v) - v}{1 - v(v + \delta v)/c^2} \tag{7.35}$$

De veranderingen δ worden als zeer klein beschouwd. Wij maken gebruik van de benadering

$$\frac{1}{1 + \epsilon} \simeq 1 - \epsilon \quad ; \qquad \epsilon \ll 1 \tag{7.36}$$

en vinden dan voor Eq.(7.35)

$$\delta v' \simeq \frac{1}{1 - v^2/c^2}\, \delta v \tag{7.37}$$

Stel nu dat de stoker er per tijdseenheid $\delta t'$ een schepje op doet. Vanwege de Lorentztransformatie weten wij dat, gezien vanaf een vast ruimtelijk punt, de tijd transformeert volgens de formule van de `tijddilatatie`

$$t' = t\,\sqrt{1 - (v/c)^2} \tag{7.38}$$

en zodoende is

$$\frac{dv'}{dt'} = \left(1 - \frac{v^2}{c^2}\right)^{-3/2} \frac{dv}{dt} \tag{7.39}$$

Wanneer wij dv'/dt' invullen als een constante a, krijgen we de bewegingsvergelijking

$$\frac{dv}{dt} = a\,\left(1 - \frac{v^2}{c^2}\right)^{3/2} \tag{7.40}$$

Die komt dus in de plaats voor de klassieke vorm $F = ma$.

Het treft dat Eq.(7.40) exact opgelost kan worden. Om de berekeningen gemakkelijker te maken, drukken wij alle snelheden uit in eenheden van c door de transformatie

$$\beta \equiv \frac{v}{c} \tag{7.41}$$

De oplossing van Eq.(7.40) is dan

$$\frac{at}{c} \equiv \tau = \int \left(1 - \beta^2\right)^{-3/2} d\beta = \frac{\beta}{\sqrt{1 - \beta^2}} \tag{7.42}$$

Hierin is ook ook de tijd dimensieloos gemaakt. Door inversie van Eq.(7.42) wordt de oplossing voor β

$$\beta = \frac{\tau}{\sqrt{1 + \tau^2}} \tag{7.43}$$

We zien nu direct dat in het begin $\tau \approx 0$, als het ruimteschip nog maar net begint met versnellen,

$$\beta = \tau \; ; \qquad \text{en dus} \qquad v = at \tag{7.44}$$

hetgeen precies overeenkomt met het klassieke geval. Als $\tau \to \infty$ daarentegen, hebben we

$$\beta = 1 \; ; \qquad \text{en dus} \qquad v = c \tag{7.45}$$

Ook een oneindig lang volgehouden versnelling brengt ons niet boven de lichtsnelheid uit!

Tenslote berekenen wij nog welke afstand het ruimteschip aflegt. Omdat $v = dx/dt$ hebben we uit Eq.(7.43)

$$\frac{a}{c^2}\frac{dx}{d\tau} = \frac{\tau}{\sqrt{1+\tau^2}} \tag{7.46}$$

met als oplossing

$$\frac{a}{c^2}x = \int \frac{\tau}{\sqrt{1+\tau^2}}\,d\tau = \int \frac{1}{2\sqrt{1+\tau^2}}\,d\tau^2 = \sqrt{1+\tau^2} - 1 \tag{7.47}$$

waarin de integratieconstante zo gekozen is dat het ruimteschip vertrekt van $x = 0$ op $\tau = 0$.

Oefening.

Bereken twee limietgevallen voor Eq.(7.47) , namelijk $\tau \ll 1$ en $\tau \gg 1$. Laat zien dat het eerste geval overeenkomt met het klassiek-mechanische $x = \frac{1}{2}at^2$ en het tweede geval met het extreem-relativistische $x = ct$.

Het leven aan boord gaat gewoon door, en we zien pas merkwaardigheden wanneer wij bezien hoe de tijd bij ons verloopt vergeleken met die welke wij (niet zij!) aflezen op de klok van het schip. Wegens Eq.(7.38) is

$$\tau' = \tau\,\sqrt{1-\beta^2} \tag{7.48}$$

en dus, gebruik makend van Eq.(7.42) ,

$$d\tau' = \frac{1}{\sqrt{1+\tau^2}}\,d\tau \tag{7.49}$$

De oplossing hiervan is

$$\tau' = \int \frac{1}{\sqrt{1+\tau^2}}\,d\tau = \log\left(\tau + \sqrt{1+\tau^2}\right) \tag{7.50}$$

Opnieuw bekijken we de limietgevallen. Als de versnelling begint, $\tau \simeq 0$, is

$$\tau' \simeq \log(\tau + 1) \simeq \tau \tag{7.51}$$

hetgeen weer precies is wat we klassiek verwachten. Als $\tau \to \infty$ daarentegen, komt er

$$\tau' \simeq \log(\tau + \sqrt{\tau^2}) = \log(2\tau) \tag{7.52}$$

Dus: ten gevolge van de tijddilatatie *zien wij de tijd aan boord van het ruimteschip veel langzamer lopen dan bij ons!*

Oefening.

Reken de versnelling a uit, in de veronderstelling dat de tijdseenheid overeenkomend met $\tau = 1$ precies 1 jaar is. Hoe groot is deze a in vergelijking met de versnelling g van de zwaartekracht aan het oppervlak van de Aarde? Zou het comfortabel zijn aan boord van zo'n ruimteschip?

Oefening.

De afstand van de Zon tot het centrum van de Melkweg is ongeveer 32000 jaar. Als een ruimteschip die afstand overbrugt met de versnelling a uit de vorige som, hoeveel ouder zijn de bemanningsleden dan volgens de boordklok wanneer zij, naar onze tijdrekening, 32000 jaar gereisd hebben? Doe dezelfde berekening voor een sterrenstelsel op een afstand van 10 miljard jaar. Wat zegt dit over de bereikbaarheid (in principe!) van verre plaatsen in ons Heelal?

Tot besluit nog een stuk over de energie van relativistische bewegingen. De arbeid die een kracht F verricht is gelijk aan de kracht vermenigvuldigd met de weglengte. Een klein stukje dx van de weg komt dus overeen met een verandering van de energie E:

$$dE = F\,dx \tag{7.53}$$

In bovenstaande vergelijkingen Eq.(7.43,46) zagen wij, dat

$$dx = \frac{c^2}{a}\frac{\tau\,d\tau}{\sqrt{1+\tau^2}} = \frac{c^2}{a}\beta\,d\tau \tag{7.54}$$

Merk op dat dit niet alleen voor het raket-voorbeeld geldt, maar *algemeen* is, omdat we slechts infinitesimale veranderingen bezien. Gebruiken we Eq.(7.53) in Eq.(7.54) dan komt er

$$dE = mc^2\beta\,d\tau \tag{7.55}$$

Met behulp van Eq.(7.42) is dit te schrijven als

$$dE = mc^2\frac{\beta\,d\beta}{(1-\beta^2)^{3/2}} \tag{7.56}$$

Enig geploeter met integratie laat zien dat dan

$$dE = mc^2\,d\frac{1}{\sqrt{1-\beta^2}} \tag{7.57}$$

Omdat dit voor elk infinitesimaal stukje van de baan zo is, hoeven we niets te veronderstellen over F, en dus geldt algemeen dat

$$E = \frac{mc^2}{\sqrt{1-\beta^2}} \tag{7.58}$$

Dit is Einsteins beroemde massa-energieformule. Dus niet $E = mc^2$, zoals amateurs zeggen! De formule Eq.(7.58) is veel algemener dan dat. In het bijzondere geval van de beweging onder invloed van een constante kracht F (constant zoals gezien door de raket) hebben we uit Eq.(7.43,47,58)

$$E = mc^2\sqrt{1+\tau^2} = mc^2 + Fx \tag{7.59}$$

Hoe realistisch is het idee van zo'n voertuig, waarmee wij in principe overal in het Heelal kunnen komen binnen een redelijke eigen tijd? Hoe je zoiets moet bouwen is onbekend, maar in principe is er geen bezwaar tegen. Een mogelijke tegenwerping is, dat de 'brandstof' voor zo'n raket van buiten moet komen. Maar wij weten dat de ruimte tussen de sterren niet leeg is, en de vraag is dus: kan een ruimtevoertuig voldoende materie verzamelen om zichzelf daarmee te versnellen? We bekijken eerst het geval dat de raket reeds relativistisch beweegt, dus $v \approx c$. Dan hebben we wegens Eq.(7.42,59)

$$E = mc^2\tau = amct = amx \tag{7.60}$$

Als het ruimteschip een doorsnede D heeft, kan het over een afstand x een hoeveelheid massa M opscheppen die wordt gegeven door

$$M = \rho D x \tag{7.61}$$

waarin ρ de massadichtheid van het medium waar de raket doorheen beweegt. Hieruit volgt een schatting voor D. Als $\tau = 1$ overeen komt met een jaar, dan is wegens Eq.(7.42)

$$\frac{c}{a}\tau = \frac{c}{a} = 1 \text{ jaar} = 3.156 \times 10^7 \text{ s} \tag{7.62}$$

en dus is de versnelling van zo'n raket

$$a = 9.51 \text{ m s}^{-2} \tag{7.63}$$

hetgeen prettig dicht bij de aardse waarde 9.8 ligt, zodat wij ons aan boord zeer goed zouden voelen (wegens het equivalentie-principe is er geen verschil tussen een versnelling en de zwaartekracht). De energie die we uit de opgeveegde materie kunnen halen is Mc^2, en dus vinden we uit Eq.(7.60,61) dat

$$amx = \rho D x c^2 \quad \rightarrow \quad D = \frac{am}{\rho c^2} \tag{7.64}$$

Nu vullen we een paar schattingen in. Laat de massa m van de raket een miljoen ton zijn (wie weet?) en laat de gewenste versnelling gegeven zijn door Eq.(7.63) . De minimale dichtheid in het Heelal is de gemiddelde massadichtheid tussen de sterrenstelsels. Deze is ongeveer een waterstofatoom per kuub, ofwel

$$\rho = 10^{-26} \text{ kg m}^{-3} \tag{7.65}$$

Zodoende wordt

$$D = \frac{9.51 \times 10^9}{10^{-26} c^2} = 1.06 \times 10^{19} \text{ m}^2 \tag{7.66}$$

De doorsnede van de raket moet dus ongeveer de wortel hieruit zijn, en dat is ruim 3 miljoen kilometer. De baan van de Maan heeft een straal van 384,000 km, de baanstraal van Mercurius is 57.9 miljoen kilometer. Dus de waarde in Eq.(7.66) is niet absurd groot, hoewel duidelijk niet binnen ons technisch bereik. De toestand wordt iets beter als we in plaats van Eq.(7.65) de gemiddelde dichtheid van de interstellaire materie nemen. Deze is ongeveer een waterstofatoom per kubieke centimeter, dus een miljoen maal groter dan Eq.(7.65) zodat D overeenkomstig kleiner wordt:

$$D = 1.06 \times 10^{13} \text{ m}^2 \tag{7.67}$$

hetgeen betekent dat een raket met een 'schep' van ruim 3000 km doorsnede volstaat. Dat is kleiner dan de straal van Aarde (6371 km). Zoals Ya.B. Zel'dovich bij zulke gelegenheden zei: *It is possible, but it is difficult.*

Een iets nauwkeuriger beschouwing leert, dat de schatting Eq.(7.64) een ondergrens is. Immers, wanneer wij Eq.(7.59) in het algemeen gebruiken (dus niet alleen in het geval $v \approx c$) vinden we in plaats van Eq.(7.64)

$$D = \frac{ma}{\rho c^2} + \frac{m}{\rho x} \tag{7.68}$$

Aan het begin van de reis is $x = 0$, dus om te starten hebben we een veel groter schep-oppervlak nodig. De raket moet dus op de een of andere manier gelanceerd worden. Ook dat hoeft geen probleem te zijn. In de buurt van een ster is ρ vele malen groter dan in de interstellaire ruimte; voor een sterrewind met constante snelheid is $\rho \propto r^{-2}$, waarin r de afstand tot de ster. Samen met Eq.(7.68) zien we daaruit dat het lanceren van zo'n denkbeeldige raket vanuit een planetenstelsel in principe haalbaar is.

Oefening.

Laat met behulp van Eq.(7.62-64) zien dat een raket van het beschreven type zijn eigen massa opveegt in precies 1 jaar, dus over een afstand van 1 lichtjaar.

De afstand tot het centrum van de Melkweg is 32,000 jaar, dus om die reis te maken moet de raket evenzoveel maal zijn eigen massa opvegen. Ter vergelijking: een auto van 500 kg die 'een-op-tien' loopt, kan 5000 km afleggen bij het verstoken van zijn eigen gewicht aan brandstof. Rijdt zo'n auto 32,000 maal zijn eigen gewicht op, dan is de afgelegde afstand 160 miljoen kilometer. Dat is bijna precies de afstand tussen de Aarde en de Zon. Met een snelheid van 200 km/h zou de auto hier 91 jaar over doen. De relativistische raket bereikt het centrum van de Melkweg in iets meer dan 11 jaar (zie Eq.(7.50) voor $\tau = 32,000$). Het is aardig dat autorijden binnen ons Zonnestelsel in deze opzichten vergelijkbaar is met relativistisch door de Melkweg rossen.

8. Enkele nuttige boeken

De volgende boeken zijn allemaal goed en nuttig. Boeken gemerkt met ♢ zijn speciaal aanbevolen voor mensen met weinig bèta-achtergrond. Suggesties voor aanvullingen zijn welkom, uitsluitend per epost, op adres `icke@strw.LeidenUniv.nl` Speciaal aanbevolen voor dit college is het boek van Zeilik & Gregory.

Abell, G.O., & Singer, B., ♢ *Science and the paranormal*, Scribners, New York 1983

Achterberg, A., *Van oerknal via niets tot straling en stof*, Epsilon, Utrecht 1994

Allen, R.H., ♢ *Star names,* Dover, New York 1963

Atkins, P.W., ♢ *Chemische reacties*, Natuur & Techniek, Maastricht 1991

Bak, P., ♢ *How Nature works,* Oxford Univ. Press, 1997

Bakich, M.E., *The Cambridge guide to the constellations,* Cambridge U. Press, 1995

Ball, P., ♢ *The self-made tapestry,* Cambridge Univ. Press, 1999

Beatty, J.K., Petersen, C.C., & Chaikin, A., *The new solar system,* Cambridge Univ.Press, 1999

Begelman, M.C., & Rees, M.J., *Gravity's fatal attraction,* Freeman, New York, 1995

Blocksma, M., & Van Maanen, H., ♢ *De schaal van Richter en andere getallen,* Bert Bakker, Amsterdam 1991

Bohm, D., *The special theory of relativity,* Routledge, London, 1996

Broer, H., *Meetkunde en fysica,* Epsilon, Utrecht, 1999

Dehling, H.G., & Kalma, J.N., *Kansrekening,* Epsilon, Utrecht, 1995

Duistermaat, J.J., & Eckhaus, W., *Analyse van gewone differentiaalvergelijkingen,* Epsilon, Utrecht, 1995

Dyson, F., *Origins of life*, Cambridge Univ. Press, 1985

Eigen, M., *Steps towards life*, Oxford Univ. Press, 1996

Enzensberger, H.M., ♢ *De telduivel*, Bezige Bij, Amsterdam 1998

Feynman, R.P., ♢ *QED,* Princeton Univ. Press, 1985

French, A.P., *Special relativity,* Chapman & Hall, London, 1997

Grasman, J., *Wiskundige methoden toegepast*, Epsilon,Utrecht, 1992

Harrison, E.R., ♢ *Cosmology,* Cambridge Univ. Press, 1999

Hogan, C.J. ♢ *Het kleine boek van de oerknal,* Bert Bakker, Amsterdam 2000

Hogben, L., ♢ *Mathematics for the million,* Merlin Press, Rendlesham 1997

't Hooft, G., ♢ *De bouwstenen van de schepping,* Prometheus, Amsterdam 1992

Horssen, W.T. van, *Differentiaalvergelijkingen,* Epsilon,Utrecht, 1993

Hulspas, M., & Nienhuys, J.W., ♢ *Tussen waarheid en waanzin,* Scheffers, Utrecht 1997

Icke, V., ♢ *The force of symmetry,* Cambridge Univ. Press, 1997

Kaku, M., ♢ *Hyperruimte,* Contact, Amsterdam 1995

Knip, K., ♢ *Alledaagse wetenschap,* Contact, Amsterdam 2000

Mandelbrot, B.B., ♢ *The fractal geometry of nature,* Freeman, San Francisco 1982

Minnaert, M.G.J. ♢ *De natuurkunde van 't vrije veld*, Thieme, Zutphen 1968

Morrison, P., & Morrison, P., ♢ *Machten van Tien,* Natuur & Techniek, Maastricht 1984

Paulos, J.A., ♢ *Een wiskundige leest de krant,* Bert Bakker, Amsterdam 1995

Rees, M.J. ♢ *Zes getallen,* Contact, Amsterdam 2000

Schilling, G., ♢ *De salon van god,* Wereldbibliotheek, Amsterdam 1993

Schilling, G., ♢ *Wat was er voor de oerknal?* Aramith, Amsterdam 1995

Shu, F.H., *The physics of astrophysics. I. Radiation,* Univ. Science Books, Sausalito, 1991

Shu, F.H., *The physics of astrophysics. II. Gas dynamics,* Univ. Science Books, Sausalito, 1992

Taylor, E.F., & Wheeler, J.A., *Spacetime phycics,* Freeman, New York 1966

Temme, N.M., *Speciale functies in de mathematische fysica*, Epsilon, Utrecht, 1990

Tijms, H., *Spelen met kansen,* Epsilon, Utrecht, 1999

Walker, J.A., ♢ *The flying circus of physics,* Wiley, New York 1977

Weinberg, S., ♢ *The first three minutes,* Basic Books, New York 1977

Weinberg, S., ♢ *Elementaire deeltjes,* Natuur & Techniek, Maastricht 1993

Wheeler, J.A., *A journey into gravity and spacetime*, Freeman, New York, 1990

Zeilik, M., & Gregory, S.A., *Introductory astronomy and astrophysics,* Saunders,Fort Worth,1998

Physical constants for Introductory Astrophysics

Physics Today, August 1996, BG9-BG16

Quantity	**Symbol**	**Value**	**Units**
speed of light	c	299792458	m s^{-1}
vacuum permeability	μ_0	$4\pi \times 10^{-7}$	N A^{-2}
vacuum permittivity	ϵ_0	$1/\mu_0 c^2$	F m^{-1}
Newtonian gravity constant	G	6.6726×10^{-11}	m^{3}kg^{-1}s^{-2}
Planck constant	h	6.626076×10^{-34}	J s
modified Planck constant	$\hbar$	1.0545727×10^{-34}	J s
electron charge	e	1.6021773×10^{-19}	C
proton mass	m_p	1.672623×10^{-27}	kg
neutron mass	m_n	1.674929×10^{-27}	kg
electron mass	m_e	9.1093897×10^{-31}	kg
Thomson cross section	σ_T	6.6652462×10^{-29}	m^2
Boltzmann constant	k	1.380658×10^{-23}	J K^{-1}
Stefan-Boltzmann constant	σ	5.67051×10^{-8}	W m^{-2}K^{-4}
solar mass	$M_\odot$	1.989×10^{30}	kg
solar luminosity	$L_\odot$	3.85×10^{26}	W
solar radius	$R_\odot$	6.96×10^8	m
mean Earth radius	$R_\oplus$	6.371×10^6	m
Earth mass	$M_\oplus$	5.98×10^{24}	kg
semimajor axis of Earth orbit	AU	1.496×10^{11}	m
reference Hubble parameter	h	100	km s^{-1}Mpc^{-1}
		3.24149×10^{-18}	s^{-1}
Hubble parameter	H_0	65	km s^{-1}Mpc^{-1}
		2.107×10^{-18}	s^{-1}
parsec	pc	3.085×10^{16}	m
year	yr	3.1558×10^7	s
		9.4609×10^{15}	m
	1 km s^{-1} = 1.023 pc Myr^{-1}		

Begrippenlijst

Inhoud