Vectoren en Tensoren; extra document voor

Deel 3: Elektromagnetisme, Relativiteit
Deel 4: Gravitatie volgens Einstein

Inhoudsopgave

1 Inleiding

Dit document bevat extra uitleg met betrekking tot de begrippen, vector (contravariant en covariant) en tensor.

Het begrijpen van, en kunnen werken met, deze begrippen is essentieel voor het wiskundig begrijpen van de Algemene relativiteitstheorie.

Laat u niet afschrikken door het eerste hoofdstuk! Dat zijn wat inleidende overdenkingen. Het enige dat u daarvan moet onthouden is het idee dat in de omgeving van een punt in de ruimte er meerdere coördinaatsystemen kunnen zijn.

Opmerking: we hebben het in dit document vaak over 'afbeeldingen'. Dit is de wiskundige benaming voor een object f dat aan elk element van een verzameling A een element van de verzameling B toevoegt. Notatie: $f : A \to B$. Voorbeeld: $f(x, y) = x^2 + 3y$ is een afbeelding en $f : \mathbb{R}^2 \to \mathbb{R}$. In die gevallen waarin de elementen getallen zijn, spreken we natuurlijk ook over functies. En $\mathbb{R}$ stelt de verzameling van alle reële getallen voor.

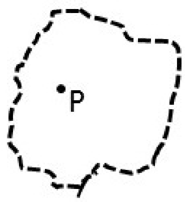
2 Differentieerbare ruimten

2.1 Inleiding

De begrippen 'vector' en 'tensor' maken deel uit van een deelgebied van de wiskunde dat 'differentiaalmeetkunde' heet. Centraal hierin staat het concept van een 'differentieerbare variëteit' (Engels: differential manifold). Maar wij doen niet zo deftig en noemen het gewoon een *differentieerbare ruimte* of zelfs kortweg een 'ruimte'. We zullen zien dat uit de definitie volgt dat zo'n ruimte een *dimensie* heeft. We kunnen dus spreken van een 1-dimensionale ruimte (voorbeelden: een rechte lijn, een cirkel), of een 2-dimensionale ruimte (voorbeelden: het platte vlak, een boloppervlak), en het algemeen: een n-dimensionale ruimte. We zullen zien dat er in de Algemene relativiteitstheorie sprake is van een 4-dimensionale ruimte (de 'ruimtetijd').

2.2 Definitie

Een wat losse definitie: een *differentieerbare ruimte* is een verzameling $\mathcal{M}$ (de $\mathcal{M}$ van manifold), waarop lokaal coördinaten kunnen worden aangebracht.



Figuur 1: Omgeving

De elementen van zo'n verzameling noemen we *punten*, tenzij het de 4-dimensionale ruimte van de Algemene relativiteitstheorie betreft, dan noemen we de elementen ook wel *gebeurtenissen*. Het woord lokaal wil zeggen dat die coördinaten alleen maar in een omgeving van een punt gedefinieerd hoeven te zijn. Dat roept dan natuurlijk weer de vraag op wat de definitie van het woord 'omgeving' is in dit verband (zie figuur 1). We gaan daar hier niet te diep op in, maar volstaan met de opmerking dat de formeel wiskundige definitie in feite er nog vanuit gaat dat de verzameling $\mathcal{M}$ al een zogenaamde topologische ruimte is. Daarin zijn de begrippen *open deelverzameling* en *omgeving* al gedefinieerd. Het begrip open deelverzameling komt overeen (is een generalisatie van) dit begrip zoals we dat op de reële rechte $\mathbb{R}$ (=de verzameling van de reële getallen) kennen:

$$(3,5) := \{x \in \mathbb{R} \mid 3 < x < 5\} \text{ is een open interval}$$

$$[3,5) := \{x \in \mathbb{R} \mid 3 \leqslant x < 5\} \text{ is niet open}$$

Maar, zoals gezegd, deze topologische overwegingen zijn verder niet van belang. We geven nu dus een definitie iets meer precies is (zij het dat we niet precies weten wat een topologische ruimte is):

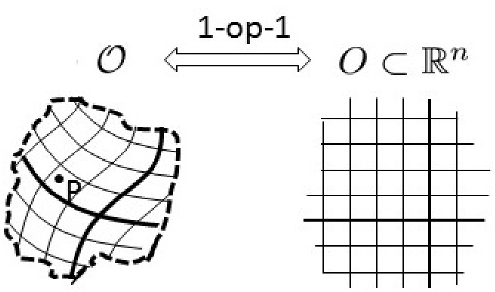
Definitie 1: **Een topologische ruimte $\mathcal{M}$ heet een n-dimensionale *differentieerbare ruimte* als er voor elk punt P er minstens één zogenaamde *kaart* bestaat. Dat is een omgeving $\mathcal{O}$ van P tezamen met een 1-op-1 relatie van $\mathcal{O}$ met een open deelverzameling $O \subset \mathbb{R}^n$. De afbeeldingen die ontstaan door verschillende kaarten te combineren zijn differentieerbaar. Zie hoofdstuk 3 voor meer details.**

2.3 Coördinaat kaarten

Met andere woorden: rond elk punt kunnen we coördinaten aanbrengen. Dit betekent dus ook dat we de hele ruimte kunnen overdekken met kaarten. Stel dat dit met één kaart zou kunnen. Dat betekent dat die ruimte 'equivalent' is met een open deelverzameling van $\mathbb{R}^n$. Is dat dan niet altijd zo? Met andere woorden: is het niet allemaal dikdoenerij wat er in dit hoofdstuk staat? Nee! Het gewone 2-dimensionale boloppervlak, precies gedefinieerd door:

$$\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$

(Dit is *2-dimensionaal*!) kan niet met één kaart worden overdekt (wel met twee) en dat is een tamelijk diepe topologische stelling.



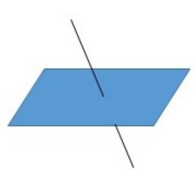
Figuur 2: Coördinaten kaart

Denk niet dat het stelsel van coördinaten dat wij op de aardbol gebruiken (lengte en breedtegraden) deze stelling tegenspreekt, want deze zijn eigenlijk alleen buiten de beide polen goed gedefinieerd. Wel is het zo dat als we er maar 1 punt uit weglaten (bijvoorbeeld de zuidpool) dan kan het wel met één kaart (maar dat is geen handige).

Bij het woord 'differentieerbare ruimte' kunnen we natuurlijk het beste de genoemde voorbeelden (platte vlak, boloppervlak) in het achterhoofd houden. Wel belangrijk is om op te merken, dat we daarbij aan de 'maagdelijke' oppervlakken moeten denken, die nog niet 'bevuild' zijn met coördinaten. Deze laatste hebben we nu eenmaal nodig als we willen gaan rekenen.

2.4 Tegenvoorbeeld

Tenslotte is het een goed gebruik om bij wiskundige definities, naast voorbeelden, ook een tegenvoorbeeld te geven. Een voorbeeld van een topologische ('whatever that may be'!) ruimte die *geen* differentieerbare ruimte is, verkrijgen we als volgt: stel u in de driedimensionale ruimte een vlak voor en een lijn die dit vlak snijdt, maar er niet in ligt. Zie figuur 3.



Figuur 3: Tegenvoorbeeld

Dan is de vereniging van deze twee wel een topologische ruimte, maar het snijpunt van vlak en lijn is een problematisch punt. Men kan laten zien dat het onmogelijk is hier een kaart omgeving omheen te maken (wat zou immers de dimensie moeten zijn? 1 of 2?). De precieze bewijzen van dit soort zaken zijn niet eenvoudig, maar intuïtief is het een en ander hopelijk wel duidelijk.

Voor de algemene relativiteitstheorie zoals we die in de cursus gaan behandelen zijn alleen de lokale eigenschappen van de ruimte van belang. Dit hoofdstuk was dus alleen bedoeld als achtergrondinformatie.

3 Coördinaat transformaties

3.1 Overlappende kaarten

In de definitie van een differentieerbare ruimte (Definitie 1 op pagina 4) is sprake van 'De afbeeldingen die ontstaan door verschillende kaarten te combineren zijn differentieerbaar'. In dit hoofdstuk gaan we bekijken wat dat precies inhoudt.

Figuur 4: Overlap

Stel we hebben twee kaarten (=coördinaatsystemen) rondom een punt P dan hebben we dus twee omgevingen van het punt P. Op de doorsnede van deze twee omgevingen zijn de coördinaatsystemen allebei gedefinieerd. Aangezien dit 1-op-1 afbeeldingen zijn, krijgen we dus een 1-op-1 afbeelding van een open deel van $\mathbb{R}^n$ naar een (ander) open deel van $\mathbb{R}^n$. We noemen de coördinaten van het ene system $(x^1, x^2, \ldots, x^n)$ (let op de boven indexen, dit zijn *geen* exponenten). Van het andere: $(x'^1, x'^2, \ldots, x'^n)$. [1] Het feit dat er sprake is van één-op-één afbeeldingen betekent dat we de x'^i-coördinaten kunnen uitdrukken in de x^i-coördinaten (en andersom!). Met ander woorden: x'^i is een functie van n variabelen: $x^1, x^2, \ldots, x^n$:

$$x'^i = x'^i(x^1, x^2, \ldots, x^n) \tag{1}$$

Het is nu duidelijk wat 'De afbeeldingen die ontstaan door verschillende kaarten te combineren zijn differentieerbaar' zal moeten betekenen: de n functies x'^i, zoals in (1), zijn differentieerbaar! Dat betekent dus dat we een n bij n matrix M van functies hebben:

$$M^i{}_j := \frac{\partial x'^i}{\partial x^j} \tag{2}$$

3.2 Poolcoördinaten

Tijd voor een voorbeeld. Voor de ruimte nemen we het platte vlak. De standaardcoördinaten noemen we (x, y). Als ander stelsel gaan we poolcoördinaten

[1] Bij het handmatig opschrijven van datgene dat komen gaat is het vaak handiger om de nieuwe coördinaten y^i te noemen in plaats van x'^i.

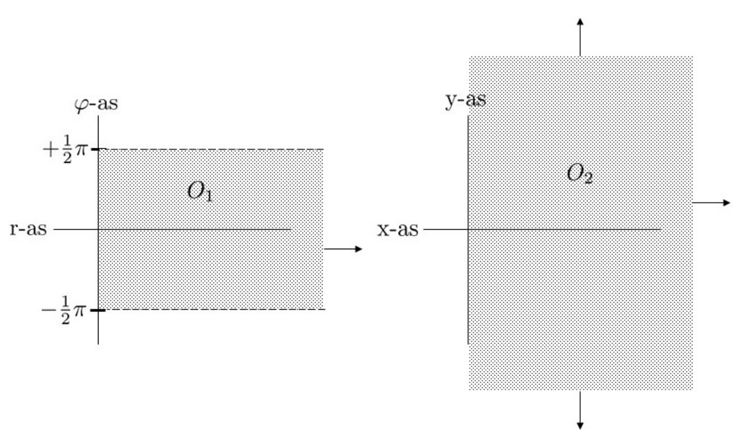
(r,φ) bestuderen. De (x,y) spelen de rol van (x'^1, x'^2) en (r,φ) de rol van (x^1, x^2). De functies uit (1) kunnen nu geschreven worden als:

$$x = r\cos\varphi \tag{3}$$
$$y = r\sin\varphi \tag{4}$$

Dit zijn dus de coördinaattransformaties. Zie ook figuur 7 op pagina 17.

Nu was er steeds sprake van 1-op-1 afbeeldingen. Als we naar de formules (3) en (4) kijken, dan zien we dat x en y een goed gedefinieerde waarde krijgen voor alle waarden van r en φ. Met andere woorden: deze formules definiëren een afbeelding van $\mathbb{R}^2 \to \mathbb{R}^2$, maar deze is geenszins 1-op-1. Als bijvoorbeeld $r = 0$, dan wordt, ongeacht de waarde van φ alles op het punt $(x,y) = (0,0)$ afgebeeld. Zelfs als we ons zouden beperken tot $r > 0$, dan nog is de afbeelding niet 1-op-1, immers φ en $\varphi + 2\pi$ worden op hetzelfde punt afgebeeld. Laten we eerst eens proberen of we inverse formules voor (3) en (4) kunnen opstellen. We zien dat $x^2 + y^2 = (r\cos\varphi)^2 + (r\sin\varphi)^2 = r^2\cos^2\varphi + r^2\sin^2\varphi = r^2(\cos^2\varphi + \sin^2\varphi) = r^2$, dus $r = \sqrt{x^2+y^2}$. Verder zien we dat $\frac{y}{x} = \frac{r\sin\varphi}{r\cos\varphi} = \frac{\sin\varphi}{\cos\varphi} =: \tan\varphi$. Uit dit alles volgt dus dat we de volgende inverse transformatie kunnen proberen:

$$r = \sqrt{x^2+y^2} \tag{5}$$
$$\varphi = \arctan\frac{y}{x} \tag{6}$$



Figuur 5: Twee open gebieden

Aan formule (6) zien we dat we in ieder geval $x \neq 0$ moeten nemen. We zijn echter op zoek naar gebieden die 1-op-1 op elkaar worden afgebeeld (en die hoeven niet maximaal te zijn) en maken het ons even makkelijk door $x > 0$ te nemen. Dan zijn de formules (5) en (6) ook goed gedefinieerd (5 was dat sowieso al). Uiteindelijk zien we dus dat de formules (3), (4), (5) en (6) een 1-op-1 verband definiëren tussen de volgende twee open gebieden van $\mathbb{R}^2$:

$$O_1 = \{(r, \varphi) \mid r > 0 \text{ en } -\frac{1}{2}\pi < \varphi < \frac{1}{2}\pi\}$$
$$O_2 = \{(x, y) \mid x > 0\}$$

We hebben in dit voorbeeld dus heel precies twee gebieden bepaald waar de transformatie 1-op-1 is. Kunt u in figuur 5 deze transformatie visualiseren?

Voor de Algemene Relativiteitstheorie zoals we die gaan behandelen is dat eigenlijk niet zo van belang, omdat, zoals eerder gezegd, dan alleen de lokale eigenschappen van de ruimte een rol spelen.

3.3 Inverse transformatie matrix

We gaan nu terug naar de algemene transformatie van formule (1). Aangezien deze een inverse heeft als volgt dat x^j ook gezien kan worden als functie van de variabelen $x'^1, x'^2, \ldots, x'^n$:

$$x^j = x^j(x'^1, x'^2, \ldots, x'^n) \tag{7}$$

Als we dit combineren met formule (1) ($x'^i = (x^1, x^2, \ldots, x^n)$) dan krijgen we:

$$x^j = x^j(x'^1(x^1, x^2, \ldots, x^n), x'^2(x^1, x^2, \ldots, x^n), \ldots, x'^n(x^1, x^2, \ldots, x^n)) \tag{8}$$

Nu gaan we formule (8) aan beide kanten differentiëren naar x^i. De linkerkant wordt dan 1 als $i = j$ en 0 als $i \neq j$. Dit vatten we samen met het symbool δ_i^j. De rechter kant differentiëren kan met de kettingregel. We krijgen op deze manier:

$$\delta_i^j = \frac{\partial x^j}{\partial x'^k} \frac{\partial x'^k}{\partial x^i} \tag{9}$$

Hierbij is de Einstein sommatie conventie toegepast. Het rechterlid van (9) kan gezien worden als een matrixvermenigvuldiging en het linker lid representeert de eenheidsmatrix (matrix met allemaal 1-en op de diagonaal en 0-en daarbuiten). In feite zegt formule (9) dus dat de matrixen $\frac{\partial x^j}{\partial x'^i}$ en $\frac{\partial x'^i}{\partial x^j}$ elkaars inverse zijn, als matrix. Eigenlijk is dat wel logisch. De afbeeldingen die de x'-en toevoegen aan de x-en en omgekeerd, waren immers elkaars inverse. Formule (9) zegt nu dat dit ook 'infinitesimaal' het geval is.

3.4 Inverse transformatie poolcoördinaten

Nu is formule (9) even abstract als belangrijk, dus laten we even kijken hoe dit in ons voorbeeld van de poolcoördinaten uitpakt. De 2 bij 2 matrix $\frac{\partial x'^i}{\partial x^j}$ kunnen we aan de hand van de formules (3) en (4) als volgt bepalen:

$$M = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \tag{10}$$

Uit de formules (5) en (6) moeten we een dergelijke matrix voor de inverse afbeelding kunnen afleiden. De afgeleiden van die functies zijn echter wat lastiger. Uit (5), was $r = \sqrt{x^2+y^2}$, leiden we af:

$$\frac{\partial r}{\partial x} = \frac{1}{2}\frac{1}{\sqrt{x^2+y^2}}2x = \frac{x}{\sqrt{x^2+y^2}}$$

Analoog kunnen we $\frac{\partial r}{\partial y}$ vinden. Voor het differentiëren van formule (6) hebben we de afgeleide van arctan nodig. Als u die vergeten bent (en dat bent u waarschijnlijk net zoals ik), dan moet u 'm opnieuw afleiden. Eerst de afgeleide van tan bepalen en dan even herinneren (of opnieuw afleiden) hoe dat ook weer zat met de afgeleide van een inverse functie. Als u daar allemaal geen zin in hebt dan kijkt u op `http://www.voorbijeinstein.nl/html/wiskunde_differentieren_510.htm`. En dan zult u zien dat

$$\frac{\mathrm{d}}{\mathrm{d}u}\arctan u = \frac{1}{1+u^2}$$

daarmee kunnen we via de kettingregel uit (6), was $\varphi = \arctan\frac{y}{x}$, uitrekenen dat

$$\frac{\partial \varphi}{\partial y} = \frac{1}{x}\frac{1}{1+(y/x)^2} = \frac{x}{x^2+y^2}$$

en ook

$$\frac{\partial \varphi}{\partial x} = -\frac{y}{x^2}\frac{1}{1+(y/x)^2} = -\frac{y}{x^2+y^2}$$

Samenvattend:

$$\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} \tag{11}$$

Volgens formule 9 moeten de rechtermatrices van (10) en (11) elkaars inverse zijn. Om dat te checken moeten we in één van beide de variabelen weer vervangen, teneinde ze vergelijkbaar te maken. Laten we door middel van de transformatie formules (3) en (4) x en y in (11) weer vervangen door r en φ. Het rechterlid van (11) wordt dan:

$$\begin{pmatrix} \frac{r\cos\varphi}{r} & \frac{r\sin\varphi}{r} \\ -\frac{r\sin\varphi}{r^2} & \frac{r\cos\varphi}{r^2} \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{\sin\varphi}{r} & \frac{\cos\varphi}{r} \end{pmatrix} \tag{12}$$

Nu gaan we dus de matrices aan de rechterkant van (10) en (12) met elkaar vermenigvuldigen en hopen het beste ervan:

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{\sin\varphi}{r} & \frac{\cos\varphi}{r} \end{pmatrix} \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} =$$

$$\begin{pmatrix} \cos^2\varphi + \sin^2\varphi & (\cos\varphi)(-r\sin\varphi) + (\sin\varphi)(r\cos\varphi) \\ (-\frac{\sin\varphi}{r})(\cos\varphi) + (\frac{\cos\varphi}{r})(\sin\varphi) & (-\frac{\sin\varphi}{r})(-r\sin\varphi) + (\frac{\cos\varphi}{r})(r\cos\varphi) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Pffff ... Het klopt!

3.5 Een lineaire transformatie

We bekijken nog een ander, eenvoudiger, voorbeeld van een transformatie. Voor de ruimte nemen we weer het platte vlak. We hebben de standaard coördinaten (x, y) en definiëren nieuwe coördinaten (u, v) door:

$$u = y \tag{13}$$

$$v = -2x \tag{14}$$

Deze transformatie oogt wel heel simpel en is dat ook. Het is een soort draaiing over 90^0 plus een uitrekking met een factor 2. Maar dat is niet zo relevant. We kunnen (13) en (14) ook schrijven als:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix} \tag{15}$$

Als we een transformatie kunnen schrijven met een matrix, zoals in (15) dan noemen we hem *lineair*. Het bijzondere is dan dat de matrix $\frac{\partial x'^i}{\partial x^j}$, zoals genoemd in (2) nu overal hetzelfde is, dus niet afhangt van (x, y), noch van (u, v).

4 Scalairen en vectoren

4.1 Scalairen en Scalar velden

Een scalar is simpelweg een getal dat op één of andere wijze bij een punt P van de ruimte hoort. We kunnen bijvoorbeeld denken aan de temperatuur in dat punt. Of, en dat voorbeeld zullen we hier vaker gebruiken, de hoogte ter plekke op een twee dimensionale landschapskaart. Wanneer, zoals in dit voorbeeld, we in meerdere punten zo'n waarde hebben, dan spreken we van een scalarveld. Wiskundig gesproken is dat dus simpelweg een afbeelding van de ruimte $\mathcal{M}$ (of een deel daarvan) naar $\mathbb{R}$. We kunnen nu eenvoudig definiëren wanneer zo'n veld *differentieerbaar* is: namelijk wanneer de afbeelding, uitgedrukt in coördinaten, differentieerbaar is [2]. We noemen het dan ook wel eens een 'glad' scalarveld. Denk aan een mooi glad heuvellandschap.

Scalairen en scalar velden blijken later de eenvoudigste vorm van tensor (-velden) te zijn, namelijk van rang 0.

4.2 Contravariante Vectoren

Als P een punt is van een ruimte, dan kennen we het begrip *vector*. Intuïtief gesproken:

Definitie 2: Een *vector* in het punt P is een pijl die begint in P en die een bepaalde richting en lengte heeft.

Het is (hopelijk) duidelijk dat deze definitie nog niet echt exact is. Bovendien (of juist daarom) geeft hij weinig houvast over wat je met een vector kan doen en hoe je er mee kan rekenen. Verder zal blijken dat het begrip 'lengte' pas exact kan worden gemaakt na de introductie van de metrische tensor in hoofdstuk 5.

Maar deze definitie 2 ondersteunt de mentale voorstelling. In het voorbeeld van het boloppervlak moet men ook denken aan een rechte pijl. Deze steekt dus uit de bol en ligt in het raakvlak aan de bol in het punt P. Zie figuur 6.

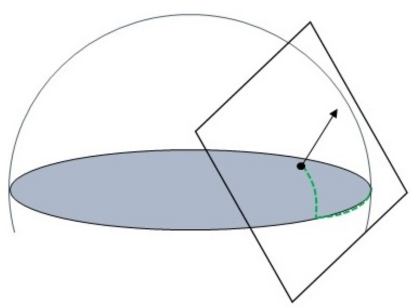
Laten we, om tot een wat werkbaarder definitie te komen eens kijken hoe zo'n pijl ontstaat. We nemen daartoe een geparametriseerde kromme. Dat is een afbeelding die aan een parameter t steeds een punt toevoegt. Denk aan

[2] Uit paragraaf 3.1 volgt, dat als de functie in één kaart differentieerbaar is, dan is ie dat in allemaal

een baan van een deeltje door de ruimte. Wanneer we coördinaten kiezen, dan ziet dat er dus uit als:

$$t \to (x^1(t), \ldots, x^n(t))$$

Dit (de $x^i(t)$) zijn uiteraard weer steeds afleidbare functies van t. Als we ons t als de tijd voorstellen, dan zijn de getallen $(\dot{x}^1(t), \ldots, \dot{x}^n(t))$ (met, zoals gebruikelijk, $\dot{x}^i(t) := \frac{\mathrm{d}}{\mathrm{d}t}x^i(t)$) een 'vector' van getallen, die de snelheid van het deeltje representeert.



Figuur 6: Raakvlak en-vector

Snelheid is dan ook een typisch voorbeeld van een vector. In dit geval is het zelfs een vectorveld langs de kromme, omdat in elk punt van de kromme zo'n vector is gedefinieerd. Laten we nu eens kijken wat er gebeurt als we een ander coördinatensysteem introduceren, zoals in formule (1). Dan kunnen wij ook in dit stelsel zo'n 'vector' van n getallen $(\dot{x}'^1(t), \ldots, \dot{x}'^n(t))$ uitrekenen, maar ook het verband. Immers omdat $x'^i = x'^i(x^1(t), x(t)^2, \ldots, x(t)^n)$ volgt uit de kettingregel dat:

$$\dot{x}'^i = \frac{\mathrm{d}x'^i}{\mathrm{d}t} = \frac{\partial x'^i}{\partial x^j}\frac{\mathrm{d}x^j}{\mathrm{d}t} = \frac{\partial x'^i}{\partial x^j}\dot{x}^j \tag{16}$$

Als we (16) even samenvatten, waarbij we $V'^i = \dot{x}'^i$ en $V^j = \dot{x}^j$ nemen, dan staat er:

$$V'^i = \frac{\partial x'^i}{\partial x^j}V^j \tag{17}$$

Formule (17) gaat nu de centrale rol spelen in onze werkdefinitie van een vector:

Definitie 3: Een *(contravariante) vector* in het punt P is een afbeelding die aan elk coördinatensysteem rond P een serie van n getallen toevoegt, zodanig dat die getallen bij een coördinaten transformatie transformeren volgens formule (17). Hierbij is $(V^1, \ldots, V^n)$ de serie die bij de coördinaten $(x^1, \ldots, x^n)$ hoort, en $(V'^1, \ldots, V'^n)$ de serie die bij de coördinaten $(x'^1, \ldots, x'^n)$ hoort.

Het is dus duidelijk dat zo'n raakvector $(\dot{x}^1(t), \ldots, \dot{x}^n(t))$ aan een kromme voldoet aan deze definitie. Het woord 'contravariant' gebruiken we om dit type te onderscheiden van een 'covariante' vector die we in paragraaf 4.3 gaan beschrijven.

Veel boeken (en ook Einstein!) introduceren de contravariante vector en de transformatie daarvan met behulp van de infinitesimale verplaatsing $\mathrm{d}x^i$. Dit is inderdaad wel de beste manier om de transformatie (17) te onthouden. Immers de transformatie van $\mathrm{d}x'^i$ naar $\mathrm{d}x^i$ laat zich heel natuurlijk opschrijven:

$$\mathrm{d}x'^i = \frac{\partial x'^i}{\partial x^j}\mathrm{d}x^j \tag{18}$$

Dit is echter wiskundig een beetje bedenkelijk [3] . Natuurlijk heeft de uitdrukking (18) veel te maken met formule (16). Immers als we beide kanten van (18) 'delen door' $\mathrm{d}t$ dan komt er:

$$\frac{\mathrm{d}x'^i}{\mathrm{d}t} = \frac{\partial x'^i}{\partial x^j}\frac{\mathrm{d}x^j}{\mathrm{d}t} \text{ En dat is: } \dot{x}'^i = \frac{\partial x'^i}{\partial x^j}\dot{x}^j \text{ (zie (16))}$$

Zoals we over één vector in een punt spreken, kunnen we ook een *vectorveld* hebben. Dan is in elk punt van de ruimte $\mathcal{M}$ (of weer een deel daarvan) een vector gedefinieerd. Voorbeelden van globale (overal gedefinieerde) vectorvelden zijn: een elektrisch veld, het gravitatieveld, een windsnelheidsveld, etc. De raakvectoren langs een kromme zijn een voorbeeld van een vectorveld dat alleen maar langs de kromme is gedefinieerd.

Het is ook goed om te bedenken, dat de coördinaten (in een punt) zelf natuurlijk ook 'een serie van n getallen' $(x^1, \ldots, x^n)$ vormen. En dat ook hiervoor geldt dat 'aan elk coördinatensysteem rond P een serie van n getallen wordt toegevoegd'. Toch is dit in het algemeen geen 'vector'. Immers de transformatie voldoet niet altijd aan de transformatie regel (17). Zie bijvoorbeeld de transformatie formules (3) en (4) voor de poolcoördinaten. In beperkte zin spreken we in dit geval overigens wel eens over vectoren. We hebben het immers over 'plaats vectoren'. In dat geval spelen de coördinaten van het punt dus wel de rol van een vector. We doen dit echter alleen als we alleen maar *lineaire* transformaties bekijken. Dan is formule (17) namelijk wel geldig. De matrix $\frac{\partial x'^i}{\partial x^j}$ is dan overal constant (hangt niet af van de coördinaten). Zie ook paragraaf 3.5.

[3] Uitdrukking (18) *heeft* een precieze wiskundige betekenis, maar deze is meestal niet bekend aan de degenen die hem gebruiken

Voor een iets abstractere maar coördinaatvrije definitie zoals wiskundigen die hanteren zie de bijlage (7).

4.3 Covariante vectoren

Zie ook Collier 5.3.6. Als u de inhoud van dit hoofdstuk goed begrijpt, dan is het belangrijkste werk gedaan. Daarna is het niet meer zo moeilijk om de rest ook te doorgronden. U begrijpt dan namelijk het verschil tussen boven- en onder indexen.

We beginnen natuurlijk weer met een (het standaard-) voorbeeld. Stel we hebben een scalair veld ϕ. Dat is dus een functie op de ruimte. Dat wil zeggen: aan elk punt van de ruimte wordt een getal gekoppeld. Zie hoofdstuk 4.1, ook voor voorbeelden. Het voorbeeld dat we hier zullen hanteren is dat van een hoogtefunctie op een 2-dimensionaal landschap. Zo gauw we coördinaten $(x^1, \ldots, x^n)$ hebben dan wordt ϕ *in dit coördinatensysteem* dus een functie van deze variabelen: $\phi = \phi(x^1, \ldots, x^n)$.

Voorbeeld:

$$\phi(x, y) := 9 - x^2 - y^2 \tag{19}$$

Dit is een mooie paraboolvormige ronde berg in het landschap met zijn top in $(x, y) = (0, 0)$ op een hoogte van 9.

Nu kunnen we (eventueel weer in één punt P, maar ook globaal) de zogenaamde 'gradiënt' van deze scalar functie bepalen. Dat zijn gewoon de afgeleiden van ϕ naar de variabelen x^i (en we hebben hier dus al 4 notatie wijzen voor!):

$$\frac{\partial\phi}{\partial x^i} \text{ of } \frac{\partial}{\partial x^i}\phi \text{ of } \partial_i\phi \text{ of zelfs } \phi_{,i} \tag{20}$$

Voor de duidelijkheid gebruiken we in het vervolg even de eerste notatiewijze.

In het voorbeeld (19) hebben we:

$$\left(\frac{\partial\phi}{\partial x} \quad , \quad \frac{\partial\phi}{\partial y}\right) = (-2x, -2y) \tag{21}$$

We hebben dus weer een serie van n getallen $(\frac{\partial\phi}{\partial x^1}, \ldots, \frac{\partial\phi}{\partial x^n})$ toegevoegd aan een punt, tezamen met een coördinatenstelsel. Mmm... Dat klinkt als een vector. Laten we dus eens even kijken hoe deze getallen transformeren als we nieuwe coördinaten $(x'^1, \ldots, x'^n)$ hebben, die een functie zijn van de oude, dus $(x'^i = x'^i(x^1, \ldots, x^n))$. Dan hebben we:

$$\phi = \phi(x'^1, \ldots, x'^n) = \phi(x'^1(x^1, \ldots, x^n), \ldots, x'^n(x^1, \ldots, x^n)) \tag{22}$$

Als we nu (22) een beide zijden differentiëren naar x^j en daarbij op het rechterlid de kettingregel toepassen dan krijgen we:

$$\frac{\partial \phi}{\partial x^j} = \frac{\partial \phi}{\partial x'^i} \frac{\partial x'^i}{\partial x^j} \tag{23}$$

We voeren nu de volgende afkortingen in:

$$A_j := \frac{\partial \phi}{\partial x^j} \text{ en } A'_i := \frac{\partial \phi}{\partial x'^i} \tag{24}$$

Dan leest (23) dus als:

$$A_j = A'_i \frac{\partial x'^i}{\partial x^j} \tag{25}$$

Bij de transformatie formule voor de vectortransformatie (17), stond het accent-object aan de linkerkant en was uitgedrukt in de getallen van de serie zonder accenten. We willen daarom (25) zodanig masseren dat A'_i alleen aan de linkerkant komt. We kunnen dat doen door beide zijden te vermenigvuldigen met de inverse matrix van $\frac{\partial x'^i}{\partial x^j}$. Dat kunnen we doen door te vermenigvuldigen met $\frac{\partial x^j}{\partial x'^k}$ en sommeren over j:

$$A_j \frac{\partial x^j}{\partial x'^k} = A'_i \frac{\partial x'^i}{\partial x^j} \frac{\partial x^j}{\partial x'^k} = A'_i \delta^i_k = A'_k$$

dus (zet A'_k links en vervang de index k door i):

$$A'_i = \frac{\partial x^j}{\partial x'^i} A_j \tag{26}$$

Laten we dit eens vergelijken met de formule (17) voor de vectortransformatie, die we hier even herhalen:

$$V'^i = \frac{\partial x'^i}{\partial x^j} V^j \tag{27}$$

Wat dan opvalt zijn twee verschillen:

1. In (26) wordt voor de serie getallen een onderindex gebruikt en in (27) een bovenindex;

2. Bij (27) wordt de 'normale' transformatie matrix $\frac{\partial x'^i}{\partial x^j}$ gebruikt, maar bij (26) de inverse daarvan: $\frac{\partial x^j}{\partial x'^i}$ (let op de plaats van het accentje!).

Punt 1 is tamelijk onbelangrijk. Het is niet meer dan een handige notatie conventie, die wel vooruitloopt op het belangrijkere verschil, namelijk punt 2. Dit is een echt inhoudelijk verschil! Daarom noemen we $(A_1, \ldots, A_n)$ een covariante vector, ook wel een 1-vorm genoemd. Samenvattend:

Contravariante vectoren transformeren door middel van de gewone transformatiematrix, terwijl covariante vectoren (1-vormen) transformeren met hulp van de inverse daarvan.

We geven dan nu ook de formele definitie:

Definitie 4: **Een *covariante vector* of *1-vorm* in het punt P is een afbeelding die aan elk coördinatensysteem rond P een serie van n getallen toevoegt, zodanig dat die getallen bij een coördinaten transformatie transformeren volgens formule (26). Hierbij is $(A_1, \ldots, A_n)$ de serie die bij de coördinaten $(x^1, \ldots, x^n)$ hoort, en $(A'_1, \ldots, A'_n)$ de serie die bij de coördinaten $(x'^1, \ldots, x'^n)$ hoort.**

Let op de overeenkomst én het verschil met definitie 3 op pagina 12.

Nu kan ik me voorstellen, dat u nog niet echt 'een gevoel' hebt voor het verschil tussen deze twee typen vectoren. Dat komt ook doordat in het geval van bijvoorbeeld de gradiënt van de gravitatie potentiaal (wat het versnellingsveld wordt) men het gewoon weer over een vectorveld heeft.

5 Metriek op een ruimte

5.1 Afstand en lengte

Het is goed om zich te realiseren dat we tot nu toe bij onze definities het begrip 'afstand' of 'lengte' nog helemaal niet zijn tegengekomen. OK, het stond wel in definitie 2 op bladzijde 11, maar dat was een intuïtieve definitie en nog geen serieuze. In feite is het echt zo dat een differentieerbare ruimte niet vanzelf een afstandsbegrip heeft. We moeten dit expliciet postuleren. Stel dus dat we zo'n maagdelijke differentieerbare ruimte hebben. We gaan ervan uit dat we op één of andere manier de afstand tussen twee punten kunnen berekenen. We benaderen dit vanuit het standpunt dat als de twee punten 'vlak bij elkaar' liggen, dit in een coördinatenstelsel kan gebeuren.

Laten we maar weer even simpel beginnen. In het gewone platte vlak, met onze gewone (x, y)-coördinaten, weten we dat de afstand tussen twee punten P en Q, met coördinaten (x_P, y_P), respectievelijk (x_Q, y_Q) is:

$$\Delta s = \sqrt{\Delta x^2 + \Delta x^2}$$

waarbij $\Delta x = x_Q - x_Q$ en $\Delta y = y_Q - y_Q$. Dit is de stelling van Pythagoras en we gaan er dus stiekem vanuit dat de coördinaatassen loodrecht op elkaar staan.

5.2 Het lijnelement

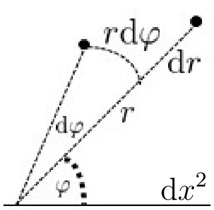
In het geval dat de punten dichtbij elkaar liggen kunnen we schrijven:

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 \tag{28}$$

Een uitdrukking voor $\mathrm{d}s^2$ heet een *lijnelement* en formule (28) geeft het standaardlijnelement voor het platte vlak.

5.3 Het lijnelement in poolcoördinaten

Het is nu heel interessant wat er gebeurt als we formules (3) en (4) voor de poolcoördinaten transformatie gaan gebruiken.



Figuur 7

We hebben:

$$\text{Formule (3): } x = r\cos\varphi \Rightarrow \mathrm{d}x = \cos\varphi\,\mathrm{d}r - r\sin\varphi\,\mathrm{d}\varphi$$

$$\text{Formule (4): } y = r\sin\varphi \Rightarrow \mathrm{d}y = \sin\varphi\,\mathrm{d}r + r\cos\varphi\,\mathrm{d}\varphi$$

Hieruit volgt:

$$\mathrm{d}x^2 = \cos^2\varphi\,\mathrm{d}r^2 - 2r\cos\varphi\sin\varphi\,\mathrm{d}r\mathrm{d}\varphi + r^2\sin^2\varphi\,\mathrm{d}\varphi^2 \tag{29}$$

en

$$\mathrm{d}y^2 = \sin^2\varphi\,\mathrm{d}r^2 + 2r\cos\varphi\sin\varphi\,\mathrm{d}r\mathrm{d}\varphi + r^2\cos^2\varphi\,\mathrm{d}\varphi^2 \tag{30}$$

Nu combineren we (28) en de som van (29) en (30) om een uitdrukking voor $\mathrm{d}s^2$ in termen van r en φ te verkrijgen:

$$\mathrm{d}s^2 = (\cos^2\varphi + \sin^2\varphi)\mathrm{d}r^2 + r^2(\cos^2\varphi + \sin^2\varphi)\mathrm{d}\varphi^2$$

(De $2r\cos\varphi\sin\varphi\,\mathrm{d}r\mathrm{d}\varphi$ term valt weg!) en dus:

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2 \tag{31}$$

Formule (31) is dus het lijnelement van het platte vlak in poolcoördinaten. Dit lijnelement kunnen we ook aflezen uit figuur 7.

5.4 De metrische matrix

In het algemeen kunnen we een lijnelement als volgt schrijven:

$$\mathrm{d}s^2 = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j \tag{32}$$

Hierbij zijn de n^2 elementen g_{ij} functies van de variabelen $(x^1, \ldots, x^n)$ en zij vormen dus een n bij n matrix. Deze is symmetrisch: $g_{ij} = g_{ji}$. We geven voorbeelden:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad : \quad \text{het standaard lijnelement plat vlak; zie (28)} \tag{33}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \quad : \quad \text{het lijnelement in poolcoördinaten; zie (31)} \tag{34}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & \cos^2\vartheta \end{pmatrix} \quad : \quad \text{bolcoördinaten} \tag{35}$$

$$\begin{pmatrix} 1 & 2v \\ 2v & 1+4v^2 \end{pmatrix} \quad : \quad \text{scheve parabool coördinaten; zie hieronder} \tag{36}$$

$$\begin{pmatrix} \frac{1}{x^2} & 0 \\ 0 & \frac{1}{x^2} \end{pmatrix} \quad : \quad \text{het Poincaré halfvlak} \tag{37}$$

We lichten nog even het ontstaan van (36) toe. We gaan over op (u, v)-coördinaten volgens de formules:

$$x = u + v^2 \text{ en } y = v \text{ dan: } \mathrm{d}x = \mathrm{d}u + 2v\mathrm{d}v \text{ en } \mathrm{d}y = \mathrm{d}v$$

en dus:

$$\begin{aligned} \mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = (\mathrm{d}u + 2v\mathrm{d}v)^2 + \mathrm{d}v^2 &= \mathrm{d}u^2 + 4v\,\mathrm{d}u\,\mathrm{d}v + 4v^2\mathrm{d}v^2 + \mathrm{d}v^2 \\ &= \mathrm{d}u^2 + 4v\,\mathrm{d}u\,\mathrm{d}v + (1+4v^2)\mathrm{d}v^2 \end{aligned}$$

Matrix (36) is een voorbeeld van een niet constante en niet-diagonale metriek.

5.5 De transformatie van metriek

Het laatste van de vorige paragraaf was een voorbeeld hoe we een metrische matrix verkregen, door een bestaande te onderwerpen aan een coördinaten transformatie. We gaan dat nu in het algemeen doen. We herhalen het algemene lijnelement (32):

$$\mathrm{d}s^2 = g_{kl}\,\mathrm{d}x^k\,\mathrm{d}x^l \tag{38}$$

We hebben hier bewust de dummy variabelen i en j even vervangen door k en l met het oog op het verkrijgen van formule (43).

Nu hebben we ook andere coördinaten $(x'^1, \ldots, x'^n)$ en we gaan ervan uit dat de afstand tussen twee dichtbij elkaar liggende punten natuurlijk niet afhangt van de coördinaatkeuze dus we hebben ook:

$$\mathrm{d}s^2 = g'_{ij}\,\mathrm{d}x'^i\,\mathrm{d}x'^j \tag{39}$$

Verder kunnen we weer x'^i uitdrukken in $(x^1, \ldots, x^n)$ via een coördinaten transformatie:

$$x'^i = x'^i(x^1, \ldots, x^n) \tag{40}$$

Maar uit (40) volgt dat:

$$\mathrm{d}x'^i = \frac{\partial x'^i}{\partial x^k}\mathrm{d}x^k \tag{41}$$

We schrijven (41) nog een keer op, maar met andere indices:

$$\mathrm{d}x'^j = \frac{\partial x'^j}{\partial x^l}\mathrm{d}x^l \tag{42}$$

Als we nu (41) en (42) in vullen in (39) dan krijgen we:

$$\mathrm{d}s^2 = g'_{ij}\,\frac{\partial x'^i}{\partial x^k}\mathrm{d}x^k\,\frac{\partial x'^j}{\partial x^l}\mathrm{d}x^l = g'_{ij}\,\frac{\partial x'^i}{\partial x^k}\,\frac{\partial x'^j}{\partial x^l}\mathrm{d}x^k\mathrm{d}x^l$$

Combineren we dit met (38) dan krijgen we:

$$g_{kl}\,\mathrm{d}x^k\,\mathrm{d}x^l = g'_{ij}\,\frac{\partial x'^i}{\partial x^k}\,\frac{\partial x'^j}{\partial x^l}\mathrm{d}x^k\mathrm{d}x^l \tag{43}$$

waaruit volgt dat:

$$g_{kl} = g'_{ij}\,\frac{\partial x'^i}{\partial x^k}\,\frac{\partial x'^j}{\partial x^l} \tag{44}$$

Nu hebben we al eerder gezien dat we het accent object (hier g'_{ij}) graag aan de linkerkant van de transformatie vergelijking willen hebben. We kunnen dat voor elkaar krijgen door in (44) beide kanten twee keer met de inverse transformatie matrix te vermenigvuldigen. We houden dan over:

$$g'_{ij} = \frac{\partial x^k}{\partial x'^i} \frac{\partial x^l}{\partial x'^j} g_{kl} \tag{45}$$

Hier zien we dus de transformatie wet voor de metriek. We noemen dit een covariante tensor van rang 2.

Laten we dit nogmaals in definitie vangen. Stel we hebben een matrix S_{kl} die als volgt transformeert:

$$S'_{ij} = \frac{\partial x^k}{\partial x'^i} \frac{\partial x^l}{\partial x'^j} S_{kl} \tag{46}$$

Dan hebben we de volgende definitie:

Definitie 5: Een *covariante tensor van rang 2* in het punt P is een afbeelding die aan elk coördinatensysteem rond P een serie (matrix) van n^2 getallen toevoegt, zodanig dat die getallen bij een coördinaten transformatie transformeren volgens formule (46). Hierbij is $S_{\mu\nu}, \mu = 1, \ldots, n, \nu = 1, \ldots, n$ de serie die bij de coördinaten $(x^1, \ldots, x^n)$ hoort, en $S'_{\mu\nu}, \mu = 1, \ldots, n, \nu = 1, \ldots, n$ de serie (matrix) die bij de coördinaten $(x'^1, \ldots, x'^n)$ hoort.

We zien dus dat de metriek matrix g_{kl} inderdaad nu officieel een covariante tensor van rang 2 is. Het is misschien wel goed om even te kijken naar 'de praktijk'. Vaak definiëren we een concreet voorbeeld door coördinaten te geven en daar bij de functies g_{kl}, oftewel het lijnelement. We hebben dat hierboven een aantal malen gezien. Bij het gewone platte vlak kiezen we coördinaten x en y en via de aanname dat de assen loodrecht op elkaar staan en dat we in de beide richtingen dezelfde meetlat gebruiken komen we tot het lijnelement (28) ($\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$). Ook bij het Poincaré halfvlak (een opgave) gingen we zo te werk. We poneerden gewoon het lijnelement, zij het dat het in dit geval wat uit de lucht kwam vallen. Bij het boloppervlak tenslotte ging het iets anders. Deze ruimte zelf is gedefinieerd al een tweedimensionale ruimte ingebed in de driedimensionale standaard platte ruimte. De metriek wordt vervolgens 'geërfd' van die omringende ruimte.

Deze metrische tensor speelt wel een centrale rol in het geheel. Het is de basis van de Riemann differentiaal meetkende. Want we hebben het volgende:

Definitie 6: Een *Riemann ruimte* is een Differentieerbare ruimte (zie definitie 1 op pagina 4) waarop een metriek is gedefinieerd. Een metriek is een symmetrische covariante tensor van rang 2.

Uit de theorie van symmetrische matrices weten we dat we via lineaire transformaties zo'n matrix altijd in diagonaalvorm kunnen brengen. Zelfs zodanig dat op de diagonaal alleen maar 1, -1 of 0 komt te staan. Let wel: dit geldt voor de tensor op één plek. Dus niet voor het hele tensorveld tegelijk. Dit betekent bijvoorbeeld concreet, dat het mogelijk is op bij één gegeven punt op aarde coördinaten aan te brengen die daar (en in de omgeving bijna) orthonormaal zijn. Voor de ART leidt dit tot de bewering dat rondom één gebeurtenis een coördinatensysteem is aan te brengen, dat in die gebeurtenis precies 'Lorentz' achtig is. Denk aan de vallende lift!

Meestal eisen we dat er geen nullen mogen staan op de diagonaal. Het gaat dan om het aantal 1-en en -1-en. Als op de diagonaal alleen maar 1-en staan, dan spreken we over een *positief definiete metriek*. Een echte afstand, zeg maar. In dit geval spreken we over 'echte' Riemann ruimten. Dit waren namelijk de ruimtes die Riemann zelf voor ogen had. Hij kon niet bevroeden dat de meest interessante toepassing van zijn theorie, de Algemene relativiteitstheorie, juist een zogenaamde *pseudo Riemann ruimte* nodig zou hebben. Namelijk een 4-dimensionale Riemann ruimte waarvan de diagonaal vorm van de metriek [1,-1,-1,-1] (of en dat is hetzelfde: [-1,1,1,1]) is!

6 Tensoren in het algemeen

6.1 De definitie

We zijn nu toe aan de algemene definitie van een tensor (in één punt P) en een tensorveld. Daartoe zetten we de gegeven definities nog eens op een rijtje:

Defnr	Naam	Transformatie	Type
3	Contravariante vector	$V'^i = \frac{\partial x'^i}{\partial x^j} V^j$	$\binom{1}{0}$
4	Covariante vector	$A'_i = \frac{\partial x^j}{\partial x'^i} A_j$	$\binom{0}{1}$
5	Covariante tensor rang 2	$S'_{ij} = \frac{\partial x^k}{\partial x'^i} \frac{\partial x^l}{\partial x'^j} S_{kl}$	$\binom{0}{2}$
	'Mixed' Tensor rang 3	$S'^k_{ij} = \frac{\partial x^l}{\partial x'^i} \frac{\partial x^m}{\partial x'^j} \frac{\partial x'^k}{\partial x^n} S^n_{lm}$	$\binom{1}{2}$

Misschien ziet u het patroon en bent u dus rijp voor de algemene definitie van een tensor, die de definities 3, 4 en 5 vervangt (generaliseert).

Definitie 7: Een *Tensor van type* $\binom{p}{q}$ *en rang* $r = p + q$ in het punt P is een afbeelding die aan elk coördinatensysteem rond P een complex van n^r getallen toevoegt, zodanig dat die getallen bij een coördinaten transformatie transformeren volgens formule (47). Hierbij is $T^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q}$ het complex dat bij de coördinaten $(x^1, \ldots, x^n)$ hoort, en $T'^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q}$ het complex dat bij de coördinaten $(x'^1, \ldots, x'^n)$ hoort.

De genoemde transformatie formule is heel indrukwekkend en ziet er als volgt uit:

$$T'^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q} = \frac{\partial x'^{\alpha_1}}{\partial x^{\mu_1}} \frac{\partial x'^{\alpha_2}}{\partial x^{\mu_2}} \cdots \frac{\partial x'^{\alpha_p}}{\partial x^{\mu_p}} \frac{\partial x^{\nu_1}}{\partial x'^{\beta_1}} \frac{\partial x^{\nu_2}}{\partial x'^{\beta_2}} \cdots \frac{\partial x^{\nu_q}}{\partial x'^{\beta_q}} T^{\mu_1\mu_2\ldots\mu_p}_{\nu_1\nu_2\ldots\nu_q} \tag{47}$$

Heel kort vat men dit wel eens samen als: Een tensor is een object dat transformeert volgens formule (47). Of nog korter: 'Een tensor is een object dat transformeert als een tensor', maar dat heeft natuurlijk een hoog 'een eend is een dier dat kwaakt als een eend'-gehalte.

6.2 Het belang

Het idee van tensoren is dat ze iets vertellen wat 'in zekere zin' onafhankelijk is van de coördinaten. Het beste is dat te zien aan de metrische tensor zoals besproken in hoofdstuk 5. De afstand tussen twee naburige punten is immers iets coördinaatonafhankelijks.

Het grote belang van tensoren is dat ze een uitermate geschikt wiskundig hulpmiddel zijn om natuurwetten in uit te drukken. Deze mogen immers niet afhankelijk zijn van een toevallig gekozen coördinaatsysteem. En uit de transformatie formules (47) volgt dat als een tensorvergelijking waar is in één coördinaatsysteem, dan is ie dat in alle!

6.3 De operaties

Om met tensoren te kunnen werken zijn de rekenregels van belang. In figuur

	Operatie	*Type wijziging*	*R*	*Voorbeeld(-en)*
O	Optellen	$\binom{p}{q}$-type + $\binom{p}{q}$-type $\rightarrow$ $\binom{p}{q}$-type	$-$	$T_{\alpha\beta}{}^{\gamma} = A_{\alpha\beta}{}^{\gamma} + B_{\alpha\beta}{}^{\gamma}$
V	Vermenigvuldigen	$\binom{p}{q}$-type $\cdot$ $\binom{n}{m}$-type $\rightarrow$ $\binom{p+n}{q+m}$-type	$\uparrow$	$T_{\alpha\beta\delta}{}^{\gamma\epsilon} = A_{\alpha\beta}{}^{\gamma} B_{\delta}{}^{\epsilon}$
C	Contractie	$\binom{p}{q}$-type $\rightarrow$ $\binom{p-1}{q-1}$-type	$\downarrow$	$T_{\beta} = A_{\gamma\beta}{}^{\gamma}$
D	Differentiëren*	$\binom{p}{q}$-type $\rightarrow$ $\binom{p}{q+1}$-type	$\uparrow$	$T_{\alpha\beta}{}^{\gamma} = \partial_{\alpha} A_{\beta}{}^{\gamma}\ (= \partial_{x^{\alpha}} A_{\beta}{}^{\gamma} = \frac{\partial A_{\beta}{}^{\gamma}}{\partial x^{\alpha}})$
V komt vrijwel altijd gecombineerd met C voor				$T^{\alpha} = A^{\beta\alpha} B_{\beta}$
Verhogen en verlagen index		In aanwezigheid van metrische tensor $g_{\mu\nu}$ $\binom{p}{q}$-type $\rightarrow$ $\binom{p-1}{q+1}$-type of $\binom{p+1}{q-1}$-type	$-$	$T_{\alpha\beta\gamma} = g_{\gamma\nu} T_{\alpha\beta}{}^{\nu} \quad T_{\alpha\beta}{}^{\nu} = T_{\alpha\beta\gamma} g^{\gamma\nu}$

Opmerkingen: Een scalar veld is een $\binom{0}{0}$-type tensor

$g^{\mu\nu}$ is de inverse matrix van $g_{\mu\nu}$ dus: $g^{\mu\alpha} g_{\alpha\nu} = \delta^{\mu}_{\nu}$

*D levert alleen een tensor in vlakke ruimtes

*D werkt alleen op een tensor**veld**.

Figuur 8: Tensor rekening

8 staan alle operaties opgesomd, waarmee we met behulp van tensoren weer nieuwe tensoren kunnen maken. Het optellen van tensoren kennen we al van vectoren. Misschien vraagt u zich af: is er nog niet iets vergeten in figuur 8? Bij vectoren hadden we toch ook de operatie dat we die kunnen vermenigvuldigen met een getal? Wees gerust, ook dat staat hierboven. Een scalar is immers ook een tensor, dus dit valt dit onder de vermenigvuldiging van tensoren.

We moeten officieel natuurlijk wel nagaan of het vermenigvuldigen en optellen van (de componenten van-) tensoren wel weer tot tensoren leidt volgens de definitie. Maar dat is vrij makkelijk te bewijzen.

Een andere belangrijke operatie is de ‘contractie’ (’verjüngung’ bij Einstein). Het is illustratief om daarvan, met een voorbeeld, even te laten zien dat dat tot een nieuwe tensor leidt en waarom dat zo is. Laten we eens kijken naar een $\begin{pmatrix}1\\2\end{pmatrix}$ tensor S^n_{lm} kijken. Deze transformeert als volgt:

$$S'^k_{ij} = \frac{\partial x^l}{\partial x'^i}\frac{\partial x^m}{\partial x'^j}\frac{\partial x'^k}{\partial x^n}S^n_{lm} \tag{48}$$

Laten hiervan *de* (er is er maar één) contravariante index n eens contraheren met de tweede covariante index m. Oftewel aan de linkerkant van (48) gaan we de indexen k en j contraheren. Praktisch betekent dat, dat we in (48) de j overal door een k vervangen:

$$S'^k_{ik} = \frac{\partial x^l}{\partial x'^i}\frac{\partial x^m}{\partial x'^k}\frac{\partial x'^k}{\partial x^n}S^n_{lm} \tag{49}$$

Nu gebeurt er iets moois, immers $\frac{\partial x^m}{\partial x'^k}$ en $\frac{\partial x'^k}{\partial x^n}$ zijn elkaar inverse, dus $\frac{\partial x^m}{\partial x'^k}\frac{\partial x'^k}{\partial x^n} = \delta^m_n$. Vergelijking (49) wordt dus:

$$S'^k_{ik} = \frac{\partial x^l}{\partial x'^i}\delta^m_n S^n_{lm} \tag{50}$$

En volgens de definitie van δ^m_n staat hier:

$$S'^k_{ik} = \frac{\partial x^l}{\partial x'^i}S^n_{ln} \tag{51}$$

hetgeen dus bewijst dat S^n_{ln} een $\begin{pmatrix}0\\1\end{pmatrix}$ tensor, oftewel een covariante vector (1-vorm) is.

Deze contractie is nu misschien nog wat abstract, maar hij komt heel vaak voor, vooral ook in combinatie met vermenigvuldiging. Van dit laatste geven we een voorbeeld. Stel we hebben een vector V^μ en een 1-vorm (covariante vector) A_ν. Deze kunnen we eerst vermenigvuldigen $= V^\mu A_\nu$, hetgeen een tensor van rang 2 (type: $\begin{pmatrix}1\\1\end{pmatrix}$) oplevert. Passen we hierop nu de contractie toe dan krijgen we: $= V^\mu A_\mu$ (μ is nu een dummy index) en dit is een tensor van rang 0 (een scalar). We kunnen dit voorbeeld nog iets verder concreet maken. Stel V^μ is een raakvector aan een kromme $(t \to (x^1(t), \ldots, x^n(t)))$ en A_ν is de gradiënt van een scalarveld ϕ. Dus:

$$V^\mu = \dot{x}^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \text{ en } A_\nu = \phi_{,\nu} = \partial_\nu \phi = \frac{\partial \phi}{\partial x^\mu} \tag{52}$$

6.4 Op en neer halen indexen

Als we in het bijzonder de operaties vermenigvuldigen en contraheren toepassen bij een willekeurige tensor en de metrische tensor dan krijgen we een operatie die bekend staat als het 'neerhalen van een index'. Een voorbeeld:

$$T_{lij} := g_{lk} S^k_{ij} \tag{53}$$

Hier hebben we dus van een $\binom{1}{2}$-tensor S^k_{ij} een $\binom{0}{3}$-tensor T_{lij} gefabriceerd (beide van rang 3!). We zeggen dat we de index k 'naar beneden' hebben gehaald.

Het naar boven halen van indices kan ook. Daarvoor gebruiken we de inverse van g_{ij}: g^{ij}.

Als we naar vergelijking (53) kijken dan schrijven we dat ook wel:

$$S_{lij} := g_{lk} S^k_{ij} \tag{54}$$

want we vinden dat het linker lid en het rechterlid van (54) eigenlijk verschillende vormen van dezelfde tensor van rang 3 zijn.

We besluiten met het toepassen van al deze zaken op de zogenaamde Riemann krommingstensor. In eerste instantie ontstaat die als een $\binom{1}{3}$-tensor (van rang 4 dus):

$$R^a{}_{bcd} \tag{55}$$

We kunnen deze in zichzelf contraheren en krijgen dan de Ricci tensor:

$$R_{bd} := R^a{}_{bad} \tag{56}$$

Verder kunnen we de a-index naar beneden halen:

$$R_{abcd} := g_{ar} R^r{}_{bad} \tag{57}$$

en dit is makkelijk in verband met het formuleren van de (vele) symmetrieën.

Tenslotte hebben we de krommingsscalar:

$$R := g^{ab} R_{ab} \tag{58}$$

7 Bijlage: Coördinaatvrije definitie vector

7.1 Inleiding

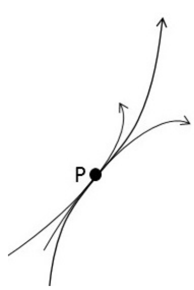
De inhoud van deze bijlage is niet nodig voor het begrijpen van de stof van de cursus. Het geeft een kijkje in de wereld van wiskundigen en de manier waarop zij dit soort definities aanpakken. En, hoewel wellicht wat abstracter, staat die eigenlijk dichter bij datgene waar het echt omgaat. Laat ik uitleggen wat ik hiermee bedoel.

Als we ons de gegeven voorbeelden van ruimtes weer voor de geest halen (het platte vlak, de 3-dimensionale ruimte ons heen, een boloppervlak) dan zijn dat eigenlijk hele mooie maagdelijke objecten. Dat wil zeggen ze zijn nog niet 'bevuild' door, door de mens gecreeërde, coördinaatsystemen. De definitie van een scalarfunctie op zo'n ruimte is nog mooi onafhankelijk van de coördinaatkeuze. Het is gewoon een functie ϕ van $\mathcal{M}$ naar $\mathbb{R}$, notatie: $\phi : \mathcal{M} \to \mathbb{R}$. Alleen als we willen definiëren wat het betekent dat zo'n functie *differentieerbaar* is maken we even gebruik van het feit dat we lokaal coördinaten kunnen aanbrengen. Zie paragraaf 4.1.

Dus scalar velden hebben we redelijk coördinaatvrij kunnen definiëren. Dit blijkt ook voor tensoren te kunnen. Dat begint bij gewone (contravariante) vectoren en dat is wat we in deze bijlage zullen schetsen.

7.2 Eerste poging: via equivalente krommen

Stel we hebben een punt $P \in \mathcal{M}$.



Figuur 9: Equivalente krommen

We hebben al gezien dat als we een differentieerbare kromme $k : \mathbb{R} \to \mathcal{M}$ hebben die door P gaat (dus bijvoorbeeld $P = k(0)$), dan hebben we een raakvector aan die kromme in P. De uitspraak in de vorige zin is coördinaatonafhankelijk. Dit idee zouden we kunnen gebruiken om tot een coördinaatvrije definitie van het begrip vector te komen. Dat gaat ongeveer zo: we kunnen een equivalentie relatie definiëren op de verzameling van alle krommen door P. Daarvoor grijpen we 'even' terug op coördinaten. Twee krommen zijn namelijk equivalent, als ze, bij een coördinaat keuze tot dezelfde rij van afgeleiden $\frac{\mathrm{d}k^i}{\mathrm{d}t}$ leiden. De equivalentieklassen die zo ontstaan zijn dan *per definitie* de vectoren in P. We werken dit niet verder uit want er is een betere manier en die werken we uit in de volgende paragraaf.

7.3 Beter: via scalar velden

Het betere idee komt als volgt tot stand. Als we weer even het intuïtieve idee van een vector in een punt P in gedachten nemen: een pijl met een bepaalde grootte die een richting aangeeft. Wanneer we nu een scalarveld hebben dat in een omgeving van P is gedefinieerd, dan kunnen we ons voorstellen dat er een getal is dat aangeeft hoe hard dit scalarveld verandert in de richting van onze pijl.

Definitie 8: Een *vector* in het punt P is een afbeelding v die aan elk scalarveld dat in een omgeving van P is gedefinieerd een getal $v(f) \in \mathbb{R}$ toevoegt en dat voldoet aan de voorwaarden 1, 2 en 3 hieronder.

Stel $\lambda \in \mathbb{R}$ en stel f en g zijn scalar functies. We kunnen functies bij elkaar optellen en vermenigvuldigen met een getal.

1. $v(f+g) = v(f) + v(g)$
2. $v(\lambda f) = \lambda v(f)$
3. $v(fg) = v(f)g + fv(g)$

De eerste twee punten zeggen dat de operatie *lineair* is. In het derde punt herkennen we de productregel voor differentiëren. Met andere woorden: een vector is een soort differentiaaloperator.

Dit is natuurlijk wel aardig (en lekker abstract) maar we willen natuurlijk wel verbinding maken met definitie 3 op pagina 12 en zelfs het liefst zien dat deze twee definities equivalent zijn. Laten we gewoon eens een coördinaten system $(x^1, \ldots, x^n)$ in een omgeving van P kiezen. En stel dat v een vector is volgens onze nieuwe abstracte definitie 8. Volgens definitie 3 zou er nu een serie getallen $(V^1, \ldots, V^n)$, de componenten van de vector, moeten bestaan. Is dat het geval? Kunnen we met onze abstracte vector (een operator op scalar functies) nu n getallen fabriceren? Jawel hoor! Heel voor de hand ligt namelijk om deze operator v toe te passen op de functies (ja dat zijn het!) x^i. We krijgen dan dus n getallen $V^i = v(x^i)$. We zouden verder kunnen gaan en bewijzen dat deze componenten dan transformeren als formule (17), maar dat gaan we hier niet doen.

Wel willen we nog even kijken naar het idee dat, op het moment dat we een coördinaten system $(x^1, \ldots, x^n)$ hebben gekozen er ook n vectorvelden bestaan. Op slide 12 van les 2 hebben we die genoteerd als $\vec{e}_1, ..., \vec{e}_n$. Hoe zit dat met de nieuwe definitie? Ook dat gaat heel soepel: we hebben immers ineens n operatoren die aan een scalarfunctie f een getal toevoegen, namelijk

$$\frac{\partial}{\partial x^i} : f \to \frac{\partial f}{\partial x^i} \tag{59}$$

en het is vrij makkelijk in te zien dat elk van deze operatoren voldoet aan definitie 8. (59) is immers per definitie een afgeleide en voldoet dus aan alle 3 de voorwaarden. En we noteren die vectoren ook als $\frac{\partial}{\partial x^i}$ als in (59), of zelfs ∂_i.

7.4 Abstracte raakruimte

Nu we toch bezig zijn gaan we nog één stapje verder. We merken eerst op dat als v en w vectoren zijn in de zin van definitie 8 en $\lambda \in \mathbb{R}$ dan kunnen we twee nieuwe vectoren $v + w$ en λv volgens definitie 8 definiëren. We moeten dan dus aangeven wat deze met een scalarfunctie f doen:

$$\begin{aligned} (v + w)(f) &= v(f) + w(f) \\ (\lambda v)(f) &= \lambda v(f) \end{aligned} \tag{60}$$

Nu hebben de wiskundigen al eerder iets gedefinieerd wat ze een (abstracte) vectorruimte noemen. Dat is een verzameling 'dingen' die je kunt op tellen en die je kunt vermenigvuldigen met een getal. Precies dus de zaken die we in (60) te pakken hebben. Met andere woorden: de verzameling van alle vectoren in een punt P (volgens definitie 8) van een differentieerbare ruimte $\mathcal{M}$ vormen een vectorruimte. Men noemt dat dan de *raakruimte* aan $\mathcal{M}$ in het punt P, notatie: T_P (de T van 'tangent'). En ik wed dat alle wiskundigen hierbij stiekem denken aan zo'n raakvlak aan de bol in een punt van de bol. Trouwens, vandaar de naam natuurlijk ook.

7.5 Differentiëren van een vectorveld

Als allerlaatste merken we nog op dat ook het covariant (absoluut) differentiëren natuurlijk heel fraai coördinaatvrij kan worden gedefinieerd. Het idee is dat, als we een vector v in een punt P hebben (kunnen we noteren als

$v \in T_P$!) en een vectorveld X, dat we dan dit vectorveld ook in de richting van v kunnen differentiëren. De formele definitie gaat als volgt: een *affiene connectie* (zo heet dat nu eenmaal) is een afbeelding die aan elke vector v en aan elk vectorveld X een vector toevoegt, notatie $\nabla_v X$ zodanig dat de volgende eigenschappen gelden:

$$\begin{aligned} \nabla_{\lambda v+\mu w} X &= \lambda \nabla_v X + \mu \nabla_w X \\ \nabla_v (X+Y) &= \nabla_v X + \nabla_v Y \\ \nabla_v (fX) &= v(f)X + f\nabla_v (X) \end{aligned}$$

waarbij Y ook een vectorveld is en w ook nog een vector en λ en μ getallen. In de derde herkennen we weer de productregel.

Hoe ziet dit eruit in een coördinatenstelsel? Welnu, dan hebben we dus vectorvelden ∂_i zoals in paragraaf 7.3 beschreven. We kunnen dus een vectorveld $\nabla_{\partial_i}\partial_j$ maken. En hierbij horen dus componenten. Met andere woorden we kunnen dit schrijven als $a^k_{\ ij}\partial_k$. Maar wacht eens even! Zouden die $a^k_{\ ij}$ niet precies de Christoffelsymbolen zijn? Natuurlijk. Dus we hebben:

$$\nabla_{\partial_i}\partial_j = \Gamma^k_{\ ij}\partial_k$$