Inhoudsopgave

Korte Inhoudsbeschrijving

In alle fasen van hun ontwikkeling hebben de natuurkunde en de wiskunde elkaar beïnvloed. In sommige perioden werd er zelfs nauwelijks onderscheid tussen de beide vakken gemaakt, tegenwoordig gaan ze grotendeels hun eigen weg, maar laten zich nog graag wederzijds inspireren waar dat van pas komt.

In de klassieke mechanica is de wederzijdse beïnvloeding altijd buitengewoon intensief geweest en dit gaat tot op de dag van vandaag door. Zo vindt de theorie van stelsels van gewone differentiaalvergelijkingen zoveel toepassingen in de klassieke mechanica, dat het geen verwondering wekt dat deze theorie voor een belangrijk deel historisch voortgekomen is uit mechanische probleemstellingen. Ons eerste doel zal zijn om de wiskundestudenten hiermee kennis te laten maken. Hopelijk treffen ook de natuurkundestudenten, die hier vast al wel van overtuigd zijn, hierbij nog voor hen interessante details aan.

Daarna zullen we zien hoe de differentiaalmeetkundige analyse van functies van meer variabelen het natuurlijke kader levert voor de bespreking van constanten van beweging, zoals energie. En ook voor formuleringen van de bewegingsvergelijkingen, die zich goed gedragen onder willekeurige substituties van variabelen. In meer technische termen betreft dit de variatievergelijkingen van Euler-Lagrange in de raakbundel en de daarmee equivalente Hamilton-vergelijkingen in de coraakbundel. De bespreking van Hamilton-stelsels wordt daarbij het meest doorzichtig als we gebruik maken van de theorie van differentiaalvormen op variëteiten, waarvan we de basisfeiten in kort bestek de revue zullen laten passeren. We passen de theorie onder andere toe op geodeten in (pseudo-)Riemann-variëteiten en stralenbundels uit de geometrische optica. Uitgebreide aandacht zullen we ook besteden aan het verrassende feit dat Hamilton-stelsels een centrale rol spelen bij het oplossen van algemene eerste orde partiële differentiaalvergelijkingen.

Bij een volledige cursus in de klassieke mechanica hoort ook een bespreking van symmetrieën in Hamilton-stelsel en van mechanische systemen met inperkingen, zoals rollende starre lichamen. Voor een grondige behandeling hiervan is echter de kennis van de basisbegrippen over acties van Lie-groepen en Lie-algebra's vereist. Andere wiskundige theorieën met relevante toepassingen in de klassieke mechanica zijn ergodentheorie, chaostheorie en de theorie van singulariteiten. Deze tekst is echter al meer dan genoeg voor een college van twee uur per week gedurende één semester. In latere colleges maak ik misschien nog eens andere keuzen. Zo heeft het historische perspectief hier uitzonderlijk veel ruimte gekregen (voor een collegedictaat althans), dit zou misschien ingeruild kunnen worden voor andere, meer 'actuele' onderwerpen.

1 Bewegingsvergelijkingen

1.1 Versnelling

De eerste opmerking over beweging van mechanische systemen is dat de posities en snelheden van het systeem op een gegeven moment willekeurige waarden kunnen hebben, maar dat als die gegeven zijn, de gehele beweging in de loop van de tijd is vastgelegd.

Is de *positie* van het systeem ten tijde t vastgelegd door de reële coördinaten $x_j(t)$, $1 \leq j \leq n$, dan is de *snelheid* bepaald door de afgeleiden

$$v_j(t) = \frac{dx_j(t)}{dt}, \quad 1 \leq j \leq n. \tag{1.1}$$

Ter vereenvoudiging van de notatie geven we de positie weer met $x(t) \in \mathbf{R}^n$ en de snelheid door middel van de vector $v(t) = dx(t)/dt \in \mathbf{R}^n$. De opmerking dat de beweging is vastgelegd door positie en snelheid op een gegeven moment impliceert dat de coördinaten

$$\frac{dv_j(t)}{dt} = \frac{d^2x_j(t)}{dt^2}, \quad 1 \leq j \leq n, \tag{1.2}$$

van de *versnelling* bij gegeven positie $x = x(t) \in \mathbf{R}^n$, snelheid $v = v(t) \in \mathbf{R}^n$ en tijdstip $t \in \mathbf{R}$ eenduidig bepaald zijn op een manier die karakteristiek is voor het systeem. Preciezer gezegd, er zijn functies a_j van $2n + 1$ variabelen, karakteristiek voor het systeem, met de eigenschap dat voor iedere beweging $t \mapsto x(t)$ van het systeem geldt dat

$$\frac{d^2x_j(t)}{dt^2} = a_j(x(t),\, v(t),\, t), \quad 1 \leq j \leq n. \tag{1.3}$$

In het bovenstaande is (en in het vervolg wordt) aangenomen dat de positie een tweemaal differentieerbare functie is van de tijd, dat wil zeggen voor iedere $1 \leq j \leq n$ is $t \mapsto x_j(t)$ een tweemaal differentieerbare reëelwaardige functie op een interval I in $\mathbf{R}$. Het aantal n van reële coördinaten dat nodig is om de positie van het systeem vast te leggen heet ook wel het *aantal vrijheidsgraden* van het systeem. Men noemt een afbeelding van een reëel interval naar $\mathbf{R}^n$ ook wel een (geparametriseerde) *kromme* in $\mathbf{R}^n$.

1.2 Galilei's Valwet

Één van de eerste voorbeelden is de beschrijving, afkomstig van Galilei [25], van de beweging van een voorwerp onder invloed van zijn zwaarte, waarbij aangenomen wordt dat er geen weerstand is van een omringend medium, zoals de lucht. Hierin is $n = 3$, x_3 is de hoogte van het voorwerp, terwijl x_1 en x_2 coördinaten in het horizontale vlak voorstellen. Volgens Galilei is er een constante g, waarvoor de functies in het rechterlid van (1.3) gegeven worden door $a_1 = a_2 = 0$ en $a_3 = -g$.

Dit leidt tot de volgende conclusies betreffende de beweging. Uit $v_j'(t) = 0$ voor $j = 1, 2$ en iedere $t \in I$ volgt dat er constante c_j zijn met $v_j(t) = c_j$. De conclusie dat een functie constant is als zijn afgeleide iedentiek nul is lijkt evident, maar als men van de functie a priori alleen maar gegeven heeft dat deze differentieerbaar is, dan is dit geen triviaal feit. Het bewijs dat in Analyse A gegeven wordt laat bijvoorbeeld zien dat de volledigheid van het systeem van de reële getallen, de eigenschap dat iedere naar boven begrensde verzameling van reële getallen een reëel getal als *kleinste* bovengrens heeft, een essentiële rol speelt. Ook Galilei worstelde met limietargumenten. Zo staat op p. 26 van [25]: "... But let us remember that we are dealing with infinities and indivisibles, both of which transcend our finite understanding, the former on account of their magnitude, the latter because of their smallness. In spite of this, men cannot refrain from discussing them, even if it must be done in a roundabout way. ..."

Maar dit terzijde. Wat betreft de verticale component, daarvoor krijgen we dat

$$\frac{d}{dt}\left(v_3(t) + g\,t\right) = v_3'(t) + g = 0,$$

dus met hetzelfde argument als boven krijgen we een constante c_3, waarvoor $v_3(t) + g\,t = c_3$, ofwel

$$v_3(t) = c_3 - g\,t.$$

Merk op dat we c_3 kunnen identificeren met $v_3(0)$ als $0 \in I$. Dit is een voorbeeld van het algemene feit dat een primitieve, zo die bestaat, eenduidig bepaald is op een additieve constante na. Doorgaand met primitiveren krijgen we zo nog constanten d_j, $1 \leq j \leq 3$, waarvoor

$$x_j(t) = d_j + c_j\,t, \quad j = 1,\, 2, \tag{1.4}$$

$$x_3(t) = d_3 + c_3\,t - \frac{1}{2}\,g\,t^2. \tag{1.5}$$

Dit leidt tot Galilei's beroemde *paraboolbanen*: de beweging vindt plaats in het verticale vlak $V = d + \mathbf{R}\,c + \mathbf{R}\,e_3$ door het punt d, dat parallel is aan de lineaire ruimte opgespannen door de vector c en de derde basisvector e_3. In dat vlak beschrijft het deeltje een bergparabool, als tenminste $c \neq 0$, omdat t in (1.5) met behulp van (1.4) als $(x_1 - d_1)/c_1$ of $(x_2 - d_2)/c_2$ herschreven kan worden, dus x_3 is langs de baan een tweedegraads veelterm als functie van de horizontale coördinaat in V, met een negatieve tweedegraadscoëfficiënt.

Een interessante observatie is verder dat van het tweede-orde stelsel van differentiaalvergelijkingen (1.3) met $a = -g\,e_3$ alle oplossingen voortgezet kunnen worden tot oplossingen die op de gehele reële as gedefinieerd zijn. Verder zijn deze maximale oplossingen $x(t)$ eenduidig bepaald door $c = v(0) = x'(0)$ en $d = x(0)$, dus door de positie en snelheid op een gegeven tijdstip $t = 0$. We hadden hierin $t = 0$ ook door een willekeurig ander 'begintijdstip' $t = t_0$ kunnen vervangen, omdat $t \mapsto x\,(t - t_0)$ een oplossing is van (1.3), dan en slechts dan als $t \mapsto x(t)$ er een is.

Anders gezegd, de oplossingen van (1.3) zijn eenduidig bepaald door hun positie en snelheid op een gegeven tijdstip. Het uitgangspunt was dat de beweging $t \mapsto x(t)$ van het systeem vastgelegd werd door $x\,(t_0)$ en $x'\,(t_0)$ en oplossing is van (1.3); de bovengenoemde eenduidigheid zegt dat *iedere* oplossing van (1.3), althans voor $a = -g\,e_3$, een beweging van het systeem beschrijft. Deze gelijkheid van de verzameling der mogelijke bewegingen van het systeem en de verzameling van de oplossingen van (1.3) rechtvaardigt de gewoonte om (1.3) *de bewegingsvergelijkingen* te noemen.

Galilei merkte ook op dat de verticale versnelling g *voor alle voorwerpen hetzelfde* is. Voor voorwerpen van hetzelfde materiaal, met dezelfde 'specifieke gravitatie', gaf hij op p. 62, p. 63 van [25] het volgende argument waarom dit eigenlijk niet anders kan. Stel dat g afhangt van de massa m van het voorwerp, dus $g = g(m)$. Neem aan dat $m \mapsto g(m)$ een monotoon niet-dalende functie is; dit wordt gesuggereerd door de ervaring. Hechten we twee voorwerpen van hetzelfde materiaal met massa's m en M met $0 < m < M$ aan elkaar, dan verwachten we dat de snellere vertraagd wordt door de langzamere. Ofwel, dat

$$g\,(M + m) \leq g\,(M)\,.$$

Maar anderzijds is $M+m > M$, dus $g\,(M+m) \geq g\,(M)$. De enige mogelijkheid is dat $g\,(M+m) = g\,(M)$.

Scherp denker als Galilei was, liet hij al redenerend nog de mogelijkheid open dat iedere materiaalsoort een eigen gravitatieversnelling zou kunnen hebben. Maar vervolgens merkte hij op dat talloze proeven met zware materialen van heel verschillende soort, zoals goud, lood, steen, zwaartekrachtsversnellingen hebben die maar heel weinig van elkaar verschillen als men ze door de lucht (die maar weinig 'vat' op zware voorwerpen heeft) naar beneden laat vallen. Hieruit concludeerde hij dat de *waarneming* ons leert dat g ook niet afhangt van de materiaalsoort. Bekend is dat $g = 9,83\,\mathrm{m}\,/\,\mathrm{sec}^2$ aan de polen en $g = 9,78\,\mathrm{m}\,/\,\mathrm{sec}^2$ bij de equator en op onze breedtegraad ongeveer $9,81\,\mathrm{m}\,/\,\mathrm{sec}^2$.

1.3 Wrijving

In Paragraaf 1.2 hebben we de weerstand verwaarloosd die het omringende medium uitoefent op het vallende voorwerp. Deze hangt zeker af van de snelheid van het voorwerp; als men heel precies kijkt dan zou men de stroming van het medium langs het oppervlak van het voorwerp moeten beschrijven, maar dat zou ons hier veel te ver voeren. Bij gegeven oppervlaktestructuur van het voorwerp leidt de weerstand in het medium tot een een afremming, dus een bijdrage aan de versnelling die tegengesteld gericht is aan de snelheid. Verder is het een ervaringsfeit dat het effect van de weerstand omgekeerd evenredig is met de massa m van het voorwerp. Dit leidt tot het volgende model: bij de valversnelling moet opgeteld worden een versnelling van de vorm $-\frac{w(\|x'(t)\|)}{m}\,x'(t)$, waarin voor iedere $v \in \mathbf{R}_{\geq 0}$ het positieve reële getal $w(v)$ de *wrijvingscoëfficiënt* is die het voorwerp ondervindt als $\|x'(t)\| = v$. Hierbij moet bedacht worden dat de versnellingen waarover gesproken wordt vectoren in $\mathbf{R}^3$ zijn; met optellen van versnellingen wordt de vectorsom bedoeld.

Laten we eenvoudigheidshalve aannemen dat de wrijvingscoëfficiënt w constant is, dit is in goede benadering het geval voor niet te grote snelheden in dunne media, zoals lucht. De bewegingsvergelijkingen krijgen dan de gedaante

$$x''(t) = -g\,e_3 - \frac{w}{m}\,x'(t),$$

ofwel

$$m\,x''(t) = -m\,g\,e_3 - w\,x'(t). \tag{1.6}$$

Men noemt de term $-w\,x'(t)$ de *weerstandskracht* die door het medium uitgeoefend wordt op het voorwerp. De factor m heet de *trage massa* van het voorwerp; wil een voorwerp met grotere m dezelfde verandering in de snelheid tengevolge van een bepaalde kracht ondervinden, dan zal deze kracht evenredig groter moeten zijn. Deze terminologie consequent doorvoerend, heet de vector $-m\,g\,e_3$ de *zwaartekracht.* (Het feit dat de versnelling tengevolge van de zwaartekracht onafhankelijk is van de massa van het voorwerp betekent dat de zwaartekracht die op een voorwerp wordt uitgeoefend evenredig moet zijn met de massa van dat voorwerp.) Het is duidelijk dat als er nooit andere versnellingen optraden dan die van het vallen naar de aarde, het eenvoudiger zou zijn geweest om alleen over versnellingen te praten en niet over krachten. We zullen later nog uitgebreid op het krachtbegrip terugkomen: het zal blijken een veel centraler begrip in de mechanica te zijn dan het zich nu nog laat aanzien.

Het is een goed vraagstuk om (1.6) op te lossen. Wij laten de analyse aan de lezer over. Daarbij blijkt dat, net als bij Galilei's valwet, dat er bij willekeurig gegeven beginpositie $x(0)$ en beginsnelheid $x'(0)$ precies één oplossing van (1.6) is; deze is gegeven door de formule

$$x(t) = x(0) + \frac{1 - \mathrm{e}^{-t\,w/m}}{w/m} \left(x'(0) + \frac{m\,g}{w}\, e_3 \right) - \frac{m\,g\,t}{w}\, e_3, \quad t \in \mathbf{R}. \tag{1.7}$$

Newton analyseerde in [68, Book Two, Section I] het probleem (1.6), maar schreef de oplossing niet in de gedaante (1.7), omdat hij niet de beschikking had over de exponentiële functie. Wèl merkte hij op dat de snelheidsvector

$$x'(t) = \mathrm{e}^{-t\,w/m} \left(x'(0) + \frac{m\,g}{w}\, e_3 \right) - \frac{m\,g}{w}\, e_3 \tag{1.8}$$

voor $t \to \infty$ convergeert naar de vector

$$v_\infty = -\frac{m\,g}{w}\, e_3 \tag{1.9}$$

en dat dit precies de snelheid is waarvoor het rechterlid in (1.6) gelijk is aan nul, dat wil zeggen waarvoor de som van de zwaartekracht en de weerstandskracht gelijk is aan nul. Het is een bekend waarnemingsfeit dat de limietsnelheid evenredig is met de massa en omgekeerd evenredig is met de weerstandscoëfficiënt.

Aristoteles en ook latere ballistici gaven een beschrijving van de beweging van voorwerpen die schuin omhoog weggegooid worden, waarin eigenlijk alleen de beginsnelheid en de limietsnelheid figureren: 'het beweegt zó lang tot-ie het zat is en valt dan met constante snelheid recht naar beneden'. Ik vond deze charmante beschrijving in het leerboek van Hamel [28, p. 13]. Deze beschrijving ligt in feite dichter bij de waarneming dan Galilei's paraboolbanen: de limietbeweging na verloop van tijd is gemakkelijker te observeren dan de relatief kortdurende overgang tussen weggooien en vallen. Dit maakt Galilei's analyse, waarin de effecten van de zwaartekracht en de wrijvingskracht van elkaar gescheiden worden, des te opmerkelijker.

1.4 Kepler's Wetten

Voorbeelden zijn leuk, leerzaam, interessant en belangrijk, maar hebben de neiging om, als ze met de nodige details behandeld worden, al gauw veel tijd in beslag te nemen. Al deze kwalificaties zijn van toepassing op de nu volgende bespreking van de bewegingen van (hemel-)lichamen onder invloed van hun wederzijdse gravitationele aantrekking.

Als resultaat van een uitgebreide analyse van de waarnemingen van Tycho Brahe, kwam Johannes Kepler tot de volgende beschrijving van de beweging van de *planeten*, in een orthogonaal coördinatensysteem met de zon in de oorsprong.

I De *eerste wet van Kepler* zegt dat de baan van iedere planeet een ellips is in een vlak V door de oorsprong, met de oorsprong (de zon) als één van de brandpunten van de ellips.

II Voor de *tweede wet van Kepler* beschouwen we het gebied $P(t_0,\, t)$ in het vlak V, dat begrensd wordt door de rechte lijn van de oorsprong naar de positie op een begintijdstip t_0, de baan van de planeet gedurende het tijdsinterval I tussen t_0 en het tijdstip t en tenslotte de rechte lijn van de positie op tijdstip t terug naar de oorsprong. Het gebied $P(t_0,\, t)$ heet ook wel het *perk* dat gedurende het tijdsinterval I wordt doorlopen. De tweede wet van Kepler, ook wel de *perkenwet* genoemd, zegt dat de oppervlakte van $P(t_0,\, t)$ evenredig is met de lengte $t - t_0$ van het tijdsinterval I.

III In de *derde wet van Kepler* tenslotte worden de perioden van de toen bekende planeten Mercurius, Venus, Aarde (toen niet een planeet genoemd), Mars, Jupiter en Saturnus onderling met elkaar vergeleken. De *periode* van een planeet is hierbij gedefinieerd als de tijd die de planeet nodig heeft om zijn ellipsbaan éénmaal te doorlopen. De observatie van Kepler is dat de periode evenredig is met de 3/2-de macht van de lengte van de lange as van de ellips.

Kepler publiceerde de eerste en tweede wet in [44] in 1609 en de derde in [45] in 1619. Newton [68] vertaalde de wetten van Kepler op de volgende manier in conclusies over de versnellingen van de planeten. Daarbij breidde hij de resultaten ook uit tot andere hemellichamen, zoals de kometen en de manen van Jupiter en Saturnus. Deze bleken namenlijk ook aan Kepler's wetten te voldoen, waarbij in het geval van de manen van Jupiter, resp. Saturnus de rol van de zon wordt overgenomen door Jupiter, resp. Saturnus.

Opmerkelijk is dat Newton in [68] geen gebruik maakte van de differentiaal- en integraalrekening, waarvan hij en Leibniz de uitvinders waren. Vermoedelijk om de presentatie van zijn hemelmechanica niet te belasten met discussies over deze nieuwe rekenwijze, gaf Newton de bewijzen zoveel mogelijk in termen van de klassieke Euclidische meetkunde, aangevuld met limietbeschouwingen zoals gebruikt door Archimedes, hetgeen een grote inventiviteit vereiste. Anders dan Newton zal ik in de hieronder volgende afleidingen vrijelijk gebruik maken van de ons bekende wiskundige hulpmiddelen.

Zij $x_1(t)$ en $x_2(t)$ de coördinaten van het hemellichaam op het tijdstip t, ten aanzien van een orthonormale basis in het vlak V, waarvan de richtingen van de basisvectoren in rust zijn met betrekking tot de richtingen van de (vaste) sterren. We nemen weer aan dat $x_1(t)$ en $x_2(t)$ tweemaal differentieerbare functies van t zijn. Dan is de afgeleide naar t van de oppervlakte van het perk gelijk aan $\frac{1}{2}\mu$, waarin

$$\mu = \mu(t) = x_1(t)\,\frac{dx_2(t)}{dt} - x_2(t)\,\frac{dx_1(t)}{dt} \tag{1.10}$$

het *moment om de oorsprong van de snelheid* voorstelt. De tweede wet van Kepler is nu equivalent met de uitspraak dat μ niet van t afhangt, dus een *constante van de beweging* is. Hiermee is op zijn beurt equivalent dat, voor iedere t, $d\mu(t)/dt = 0$. Omdat de termen in $\mu'(t)$ met de producten van de eerste-orde afgeleiden van de $x_j(t)$ tegen elkaar wegvallen, is dit equivalent met de voorwaarde dat

$$x_1(t)\,x_2''(t) - x_2(t)\,x_1''(t) = 0,$$

ofwel de versnellingsvector $x''(t)$ is een scalair veelvoud van de positievector $x(t)$. Anders gezegd: de perkenwet van Kepler is equivalent met de volgende wet van Newton:

De versnelling is in de richting van het lichaam waardoor het bewegende lichaam wordt aangetrokken.

Voor de analyse van de eerste en derde wet van Kepler hebben we een paar feiten nodig betreffende *kegelsneden (ellipsen, parabolen, hyperbolen)* en hun *brandpunten.* Net als Kepler [43] neem ik de brandpuntseigenschap als uitgangspunt.

Laat $t \mapsto x(t)$ een differentieerbare kromme in het vlak zijn die een figuur K doorloopt, met raakvector $x'(t) \neq 0$. Zij a en b twee vast gekozen punten in het vlak. De rechte lijnen van $x(t)$ naar a, resp. b voldoen aan de terugkaatsingswet aan de kromme dat 'hoek van inval = hoek van uitval', als de vectorsom van de richtingsvectoren $\|x(t)-a\|^{-1}\,(x(t)-a)$ en $\|x(t)-b\|^{-1}\,(x(t)-b)$ loodrecht staat op $x'(t)$, ofwel als

$$\frac{\langle x(t)-a,\, x'(t)\rangle}{\|x(t)-a\|}+\frac{\langle x(t)-b,\, x'(t)\rangle}{\|x(t)-b\|}=0.$$

(Hierin is $\langle v,\, w\rangle = \sum_i v_i\, w_i$ het Euclidische inproduct van de vectoren a en b.) Maar dit is equivalent met

$$\frac{d}{dt}\left(\|x(t)-a\|+\|x(t)-b\|\right)=0. \tag{1.11}$$

De conclusie is, dat de figuur K alle stralen uit a terugkaatst naar stralen door b (en vice versa), dan en slechts dan als (1.11) geldt voor alle t, dus dan en slechts dan als er een constante l is met de eigenschap dat voor iedere t geldt dat $x = x(t)$ voldoet aan de vergelijking

$$\|x-a\|+\|x-b\|=l. \tag{1.12}$$

Omgekeerd leidt dit, gegeven de lengte $l > \|a-b\|$, tot de volgende zogenaamde *tuinmansconstructie* van de figuur K. Bevestig een touw ter lengte l met de uiteinden in de punten a en b. Pak het touw in een willekeurig tussenpunt vast en trek het strak. Dit kan op twee manieren gerealiseerd worden met het punt x in het vlak; op deze manier krijgen we alle x in het vlak die aan (1.12) voldoen. De punten a en b heten de *brandpunten* van de figuur K.

Door een verschuiving kunnen we er altijd voor zorgen dan één van de brandpunten, zeg a, in de oorsprong van het coördinatensysteem terecht komt; deze keuze werkt toe naar de formulering van de eerste wet van Kepler. De vergelijking

$$\pm\|x-b\|=l-\|x\| \tag{1.13}$$

is equivalent met

$$\|x\|^2-2\langle x,\, b\rangle+\|b\|^2=\|x-b\|^2=(l-\|x\|)^2=l^2-2\,l\,\|x\|+\|x\|^2,$$

hetgeen op zijn beurt equivalent is met

$$\|x\|=\langle x,\, \varepsilon\rangle+\gamma, \tag{1.14}$$

waarin

$$\varepsilon \;=\; \frac{1}{l}\,b, \tag{1.15}$$

$$\gamma \;=\; \frac{l^2-\|b\|^2}{2l}. \tag{1.16}$$

Kwadrateren van (1.14) geeft dat de punten $x \in K$ voldoen aan de tweedegraads veeltermvergelijking

$$\langle x,\, x\rangle=\left(\langle x,\, \varepsilon\rangle+\gamma\right)^2. \tag{1.17}$$

In het algemeen worden de figuren in het vlak, die bepaald worden door tweedegraads veeltermvergelijkingen, *kwadrieken* of *kegelsneden* genoemd. De tweede naam verwijst naar het feit dat het precies de figuren zijn die verkregen kunnen worden door de standaardkegel ${x_1}^2+{x_2}^2={x_3}^2$ in de ruimte te snijden met een vlak dat niet door de oorsprong gaat. Dit was één van de vele feiten over kwadrieken die Kepler uit het werk van Appolonius [5] had geleerd.

Het plusteken in (1.13) correspondeert met de vergelijking (1.12) voor de *ellips* met brandpunten $a = 0$ en b. We hebben het plusteken in (1.13) dan en slechts dan als $l > \|b\|$, hetgeen met het oog op (1.15) equivalent is met $\|\varepsilon\| < 1$.

Is $\|\varepsilon\| > 1$, dan definieert (1.17) een *hyperbool*. Dit geval correspondeert met de keuze van het minteken in (1.13), welke vergelijking dan één van de takken van de hyperbool definieert; de andere tak krijgen we als we l in (1.13) vervangen door $-l$. De vergelijking (1.13) leidde Kepler [43] tot de volgende constructie van de hyperbool. Bevestig een zeer lang touw met de uiteinden in de punten $a = 0$ en b. Pak het touw in een tussenpunt c vast waarvan de afstand naar a minus de afstand naar b gelijk is aan l. Trek het strak. Knijp nu met de andere hand de stukken touw van c naar a en van c naar b samen, bij een punt d, waarbij de stukken touw van d naar a en van d naar b nog steeds strak blijven en $\|d - a\| - \|d - b\| = l$. De punten d beschrijven de tak van de hyperbool die is bepaald door (1.13) met het minteken.

Het minteken in (1.13) correspondeert met de voorwaarde dat nu het *verschil* van de richtingsvectoren $\|x(t) - a\|^{-1}\,(x(t) - a)$ en $\|x(t) - b\|^{-1}\,(x(t) - b)$ loodrecht staat op $x'(t)$. Dit betekent dat als we de vanuit a aan K teruggekaatste stralen aan de andere kant van K doortrekken, ze allemaal door het punt b gaan Men noemt b ook wel het *virtuele brandpunt* van K bij het brandpunt a van K. Is K' de andere tak van de hyperbool dan is b brandpunt van K', met a als virtueel brandpunt.

Is tenslotte $\|\varepsilon\| = 1$, dan definieert (1.17) een *parabool*. In dit geval is er geen ander brandpunt of virtueel brandpunt b, maar zijn de vanuit de oorsprong aan K teruggekaatste stralen evenwijdig, in de richting van de vector ε.

Tot zover de feiten over brandpunten van kegelsneden. Neem nu aan dat de versnelling in de richting is van de oorsprong (corresponderend met Kepler's tweede wet) en neem aan dat $x = x(t)$ voldoet aan (1.14). Als $\|\varepsilon\| < 1$ dan betekent dit laatste dat aan de eerste wet van Kepler is voldaan. We laten hier echter ook toe dat $\|\varepsilon\| = 1$ of $\|\varepsilon\| > 1$, hetgeen betekent dat $x(t)$ een parabool of hyperbooltak doorloopt, met de oorspong als brandpunt. In het geval dat $\|\varepsilon\| \leq 1$ dan geeft de ongelijkheid van Cauchy-Schwarz dat noodzakelijkerwijze $\gamma \geq 0$ als er een x is die aan (1.14) voldoet. Als $\|\varepsilon\| > 1$ dan correspondeert $\gamma > 0$, resp. $\gamma < 0$ met de situatie dat de oorsprong in het brandpunt, resp. in het virtuele brandpunt ligt van de hyperbooltak.

Substitutie van $x = x(t)$ in (1.14) en differentiatie naar t geeft

$$\|x(t)\|^{-1}\,\langle x(t),\, x'(t)\rangle = \langle x'(t),\, \varepsilon\rangle,$$

waarin we het linkerlid vervolgens lezen als het inproduct van $x'(t)$ met de eenheidsvector $\|x(t)\|^{-1}\,x(t)$. Voor de volgende differentiatie gebruiken we dat

$$\frac{d}{dt}\,\left(\|x(t)\|^{-1}\,x(t)\right) = \|x(t)\|^{-1}\,x'(t) - \|x(t)\|^{-3}\,\langle x(t),\, x'(t)\rangle x(t).$$

Verder geeft (1.10) dat

$$\begin{aligned} \|x\|^2\,x_1' - \langle x,\, x'\rangle\,x_1 &= {x_2}^2\,x_1' - x_2\,x_2'\,x_1 = -\mu\,x_2, \\ \|x\|^2\,x_2' - \langle x,\, x'\rangle\,x_2 &= {x_1}^2\,x_2' - x_1\,x_1'\,x_2 = \mu\,x_1, \end{aligned}$$

dus

$$\frac{d}{dt}\,\left(\|x(t)\|^{-1}\,x(t)\right) = \mu\,\|x(t)\|^{-3}\,J\,(x(t))\,, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{1.18}$$

Hierin is J gelijk aan 'een kwartslag linksom draaien'. Merk op dat $\mu = \langle J(x),\, x' \rangle$.

Tweemaal differentiëren naar t van (1.14) met $x = x(t)$ leidt hiermee tot

$$\mu^2\, \|x\|^{-3} + \|x\|^{-1}\, \langle x,\, x'' \rangle = \langle x'',\, \varepsilon \rangle. \tag{1.19}$$

De perkenwet gaf dat er een reëelwaardige factor $f(t)$ is, waarvoor $x''(t) = f(t)\, x(t)$. Dit substituerend in (1.19) en (1.14) gebruikend, krijgen we

$$0 = \mu^2\, \|x\|^{-3} + \|x\|^{-1}\, f\, \langle x,\, x \rangle - f\, \langle x,\, \varepsilon \rangle = \mu^2\, \|x\|^{-3} + f\, \left(\|x\| - \langle x,\, \varepsilon \rangle \right) = \mu^2\, \|x\|^{-3} + f\, \gamma.$$

De conclusie is hiermee dat

$$x''(t) = -\frac{\mu^2}{\gamma}\, \|x(t)\|^{-3}\, x(t). \tag{1.20}$$

In het geval van ellipsen of parabolen is $\gamma > 0$ en is de versnelling naar de oorsprong toe gericht, dit is ook het geval van een hyperbooltak met de oorsprong in het brandpunt. In het enige overgebleven geval van een hyperbooltak met de oorsprong in het virtuele brandpunt is $\gamma < 0$ en is de versnelling van de oorsprong af gericht; het lichaam wordt in dit geval door de oorsprong *afgestoten*. Merk op dat het rechterlid in (1.20) oneindig groot wordt als $x(t)$ naar de oorsprong gaat; in het voorgaande is stilzwijgend dan ook aangenomen dat steeds $x(t) \neq 0$.

Hiermee hebben we de volgende stelling van Newton [68, p. 56-61] geverifieerd.

Als een tweemaal differentieerbare kromme langs een kegelsnede loopt met versnelling in de richting van het brandpunt, dan is de grootte van de versnelling omgekeerd evenredig met het kwadraat van de afstand tot het brandpunt.

Neem nu omgekeerd aan dat er een constante c is, waarvoor

$$x''(t) = -c\, \|x(t)\|^{-3}\, x(t). \tag{1.21}$$

Dit impliceert dat $x''(t)$ in de richting van $x(t)$ ligt, ofwel Kepler's perkenwet geldt, die op zijn beurt betekent dat $\mu = x_1\, x_2' - x_2\, x_1'$ niet van t afhangt. Uit (1.18) lezen we nu af dat

$$\frac{d}{dt}\left(\|x(t)\|^{-1}\, x(t) \right) = \mu\, \|x(t)\|^{-3}\, J(x(t)) = -\frac{\mu}{c}\, J\left(x''(t) \right),$$

ofwel

$$\|x(t)\|^{-1}\, x(t) = \varepsilon - \frac{\mu}{c}\, J\left(x'(t) \right)$$

voor een vector ε die niet van t afhangt. Inproduct nemen met $x(t)$ leidt tot (1.14) met $x = x(t)$ en $\gamma = \mu^2 / c$. Dit levert de volgende precisering van Newton's stelling.

Stelling 1.1 *Zij $t \mapsto x(t)$ een tweemaal differentieerbare afbeelding van een reëel interval $\mathcal{T}$ naar $\mathbf{R}^2$ waarvoor, voor iedere $t \in \mathcal{T}$, de versnelling $x''(t)$ een veelvoud is van $x(t)$. Dan zijn de volgende uitspraken equivalent.*

a) *Er is een reële constante c met de eigenschap dat voor iedere $t \in \mathcal{T}$ de vergelijking (1.21) geldt.*

b) *Er is een reële constante γ en een constante vector $\varepsilon \in \mathbf{R}^2$, met de eigenschap dat voor iedere $t \in \mathcal{T}$ de vergelijking (1.14) geldt voor $x = x(t)$.*

Hierbij is $\gamma\, c = \mu^2$ met μ als in (1.10) en

$$\varepsilon = \|x(t)\|^{-1}\, x(t) + \frac{\mu}{c}\, J\left(x'(t)\right), \quad t \in \mathcal{T}. \tag{1.22}$$

De implicatie a) $\Longrightarrow$ b) staat ook bij Newton [68, Book I, Sec. III, Prop. XVII]. Daarin liet Newton zien dat er bij iedere positie x en snelheid v een Kepler-beweging $\gamma(t)$ is met $\gamma(0) = x$ en $\gamma'(0) = v$, waarbij ook de parabolische en hyperbolische banen toegelaten zijn. Omdat deze bewegingen voldoen aan (1.21), concludeerde Newton dat *iedere* oplossing van (1.21) gegeven wordt door zo'n Kepler-beweging. Hierbij is stilzwijgend gebruik gemaakt van de eenduidigheid van het beginwaardenprobleem voor (1.21), welke we in Paragraaf 2.10 nader zullen bespreken. Het feit dat Newton in de Principia op diverse plaatsen praatte over *de* beweging bij gegeven versnelling, beginpositie en beginsnelheid, duidt erop dat deze eenduidigheid voor Newton vanzelf sprak.

Om in te zien wat de derde wet van Kepler betekent in termen van de bewegingsvergelijking (1.21), nemen we nu aan dat $\|\varepsilon\| < 1$ en $\gamma > 0$, corresponderend met de beweging op een ellips. Het *centrum* van de ellips is het punt

$$C = \frac{\gamma}{1 - \|\varepsilon\|^2}\, \varepsilon, \tag{1.23}$$

immers substitutie van $x = C + y$ doet (1.17) overgaan in

$$\langle y,\, y\rangle = \frac{\gamma^2}{1 - \|\varepsilon\|^2} + \langle y,\, \varepsilon\rangle^2, \tag{1.24}$$

welke vergelijking invariant is onder de spiegeling $y \mapsto -y$. De *lange*, resp. *korte as* is gedefinieerd als met maximale, resp. minimale afstand van de punten op de ellips tot het centrum. We lezen uit (1.24) af dat

$$\text{lange as} = \lambda \quad = \quad \frac{\gamma}{1 - \|\varepsilon\|^2}, \tag{1.25}$$

$$\text{korte as} = \kappa \quad = \quad \frac{\gamma}{\sqrt{1 - \|\varepsilon\|^2}}. \tag{1.26}$$

Na één periode T is het perk gelijk aan het gebied binnen de ellips, de oppervlakte waarvan gelijk is aan $\pi\, \lambda\, \kappa$. Dit leidt tot de vergelijking

$$\frac{1}{2}\mu\, T = \pi\, \lambda\, \kappa = \pi\, \gamma^2\, \left(1 - \|\varepsilon\|^2\right)^{-3/2},$$

ofwel

$$\text{periode} = T = 2\pi\, \frac{\gamma^2}{\mu}\, \left(1 - \|\varepsilon\|^2\right)^{-3/2} = 2\pi\, \frac{\gamma^{1/2}}{\mu}\, \lambda^{3/2} = 2\pi\, c^{-1/2}\, \lambda^{3/2}, \tag{1.27}$$

waarbij in de laatste identiteit de relatie $\gamma\, c = \mu^2$ is gebruikt. Hieruit lezen we de volgende stelling van Newton af:

Onder aanname van de eerste en de tweede wet van Kepler, is de derde wet van Kepler equivalent met de uitspraak dat voor alle planeten en kometen de constante c in (1.21) dezelfde is. Hetzelfde geldt voor de manen van Jupiter, resp. Saturnus onderling.

Al in zijn jonge jaren had Newton de gedachte gekregen dat de gravitatie van voorwerpen bij het aardoppervlak zich wel eens kon uitstrekken 'tot aan de maan en daar voorbij'. Daarmee sluit de bovenstaande conclusie over de versnellingen van de hemellichamen zeer fraai aan bij Galilei's wet dat alle voorwerpen op aarde dezelfde versnelling ondervinden tengevolge van hun zwaarte. Daarbij is Galilei's verwaarlozing van de weerstand voor de maan en ook voor kunstmanen in extreme mate correct: de lucht wordt bij grotere afstand tot de aarde steeds ijler en is al bij een relatief kleine afstand van het aardoppervlak, zoals 1000 km, totaal verwaarloosbaar. Eenzelfde opmerking geldt voor de planeten omdat de gasdichtheid in de ruimte waar zij doorheen bewegen uiterst gering is.

1.5 Newton's Superpositieprincipe

Het laatste ingrediënt voor Newton's gravitatiemodel is de veronderstelling dat als een lichaam door een aantal lichamen tegelijk aangetrokken wordt, dan is de resulterende versnelling gelijk aan de vectorsom van de versnellingen die het lichaam van ieder van de andere lichamen apart zou krijgen, welke dan gegeven wordt door (1.21) met $x(t)$, resp. c vervangen door $x(t) - x^{(j)}(t)$, resp. c_j waarin $x^{(j)}(t)$ de positie ten tijde t is van het j-de ander lichaam en c_j een positieve constante is die specifiek is voor het j-de lichaam. Newton [68, p. 14] maakte dit plausibel met een argument dat erop neerkomt dat de versnelling hetzelfde is als wanneer de versnellingen afwisselend kort na elkaar zouden werken, met een sterktefactor omgekeerd aan het deel van de tijd dat ieder is ingeschakeld.

Dit *superpositieprincipe* van Newton kan niet wiskundig bewezen worden om de eenvoudige reden dat men voor een samengesteld systeem ieder willekeurig stelsel van bewegingsvergelijkingen op kan stellen, zonder dat daar vanuit de wiskunde bezwaar tegen gemaakt kan worden. De waarnemingen zullen steeds uiteindelijk moeten beslissen in hoeverre een bepaalde aanname geldig is. Newton's aanname over superpositie van de versnellingen wordt bij gravitationele aantrekking van lichamen in grote nauwkeurigheid bevestigd. Hierbij dient opgemerkt te worden dat in het algemeen de bewegingen van meer dan twee lichamen onder invloed van gravitatie al zó ingewikkeld is dat er geen sprake is van een beschrijving in termen van bekende figuren of functies zoals bij Kepler's wetten. Dit maakt dat het superpositieprincipe niet direct uit de beschrijving van de beweging naar voren komt. In plaats daarvan worden de bewegingsvergelijkingen geverifieerd door numerieke berekeningen van de oplossingen, waarna de uitkomsten vergeleken worden met de waarnemingen.

Een groot lichaam opgebouwd denkend uit kleinere onderdelen, komt men hiermee ook op de plausibele hypothese dat de constante c_j evenredig is met de massa m_j van het j-de lichaam, waarbij de evenredigheidsconstante $\mathcal{G}$ onafhankelijk is van het lichaam en de *universele gravitatieconstante* genoemd wordt. Mijn natuurkundeboek van Kronig [50] geeft

$$\mathcal{G} = (6,664 \pm 0,004) \times 10^{-11} \text{ kilogram}^{-1} \text{ meter}^3 \text{ seconde}^{-2} . \tag{1.28}$$

Dit leidt tot het volgende stelsel van bewegingsvergelijkingen voor N lichamen die elkaar gravitationeel aantrekken. Is $x^{(i)}(t)$, resp. m_i de positie ten tijde t, resp. de massa van het i-de lichaam, dan wordt verondersteld dat, voor iedere i, $t \mapsto x^{(i)}(t)$ een tweemaal differentieerbare afbeelding is van het tijdsinterval naar $\mathbf{R}^3$ en dat, voor iedere i en iedere t in het tijdsinterval,

$$\frac{d^2 x^{(i)}(t)}{dt^2} = -\mathcal{G} \sum_{j \neq i} m_j \, \|x^{(i)} - x^{(j)}\|^{-3} \left(x^{(i)}(t) - x^{(j)}(t) \right) . \tag{1.29}$$

Bij het planetenstelsel heeft de zon een zóveel grotere massa dan alle planeten samen, dat het in eerste orde een zeer goede benadering is om de zon in rust in de oorsprong te nemen en ieder van de planeten volgens (1.21) daaromheen te laten draaien. Op deze manier is het stelsel (1.29) in benadering compatibel met Kepler's beschrijving.

Het is in feite gebruikelijker (en later zal het ook voordelen blijken te hebben) om de vergelijkingen net als in (1.6) te herschrijven in termen van krachten. Dat wil zeggen men schrijft

$$m_i \, \frac{d^2 x^{(i)}(t)}{dt^2} = \sum_{j \neq i} \phi^{(i,\,j)}, \tag{1.30}$$

waarin $\phi^{(i,\,j)}$ de *kracht* is die *door het j-de lichaam op het i-de lichaam wordt uitgeoefend*. Bij gravitatie wordt deze kracht gegeven door

$$\phi^{(i,\,j)} = -\mathcal{G}\, m_i\, m_j \, \|x^{(i)}(t) - x^{(j)}(t)\|^{-3} \left(x^{(i)}(t) - x^{(j)}(t) \right). \tag{1.31}$$

Merk weer op dat (1.31) singulier wordt als er een i en j zijn, waarvoor $i \neq j$ en $x^{(i)} = x^{(j)}$, dat wil zeggen als minstens twee van de lichamen *op elkaar botsen*.

Een interessant gevolg van (1.31) is dat

$$\phi^{(i,\,j)} = -\phi^{(j,\,i)}. \tag{1.32}$$

Bij Newton [68, p. 13] is dit *principe van actie is reactie* een postulaat, waaraan hij veel algemenere geldigheid toeschrijft dan alleen voor wederzijdse gravitationele aantrekkingskrachten. Voor gravitatie correspondeert het met de aanname dat de coëfficiënt c in (1.21) evenredig is met de massa van het aantrekkende lichaam.

We benadrukken nogmaals dat men kritiek kan hebben op de argumentatie voor Newton's model, maar dat men er niet omheen kan dat ontelbare waarnemingen de geldigheid van het model met grote nauwkeurigheid hebben bevestigd. Een kleine correctie is er later op aangebracht met *Einstein's relativiteitstheorie*.

1.6 Electrisch Geladen Deeltje

Een deeltje met electrische lading q ondervindt van een electromagnetisch veld een kracht ϕ, die is gegeven door

$$\phi = q\,(E + v \times B)\,. \tag{1.33}$$

Is $x = x(t) \in \mathbf{R}^3$ de positie ten tijde t van het geladen deeltje, dan is $E = E(x,\,t) \in \mathbf{R}^3$, resp. $B = B(x,\,t) \in \mathbf{R}^3$ de *electrische veldsterkte*, resp. *magnetische inductie* in de positie x ten tijde t, terwijl $v = dx(t)/dt \in \mathbf{R}^3$ de snelheid van het geladen deeltje voorstelt. Zie Kronig [50, p. 221, 254].

In (1.33) stelt $v \times B$ het *uitwendige product* of *vectorproduct* van v en B voor. Dit is gedefinieerd door middel van

$$\begin{aligned} (v \times B)_1 &= v_2\, B_3 - v_3\, B_2, \\ (v \times B)_2 &= v_3\, B_1 - v_1\, B_3, \\ (v \times B)_3 &= v_1\, B_2 - v_2\, B_1. \end{aligned} \tag{1.34}$$

Het electromagnetische veld wordt gegeven door de vectorwaardige functies E en B van de positie $x \in \mathbf{R}^3$ en de tijd $t \in \mathbf{R}$. De term $q\,E$ heet de *electrische kracht* die op het deeltje uitgeoefend wordt, terwijl $v \times B$ ook wel de *Lorentz-kracht* heet. In het algemeen levert het bewegende geladen deeltje zelf ook een bijdrage aan het electromagnetische veld, die we bij relatief kleine lading q echter kunnen verwaarlozen. In dat geval kan het gebeuren dat het electromagnetische veld gegeven is en krijgen we voor ons testdeeltje een stelsel bewegingsvergelijkingen van de vorm

$$m \frac{d^2 x(t)}{dt^2} = q \left[E(x(t),\, t) + \frac{dx(t)}{dt} \times B(x(t),\, t) \right]. \tag{1.35}$$

Een interessant aspect van deze bewegingsvergelijkingen is dat nu het rechterlid niet alleen afhangt van de positie $x(t)$, maar ook van de snelheid $dx(t)/dt$ van het deeltje. Heel anders dan bij wrijvingskrachten, staat de Lorentz-kracht loodrecht op de snelheid, dus leidt tot een afbuiging van de baan, in een richting die ook loodrecht staat op B.

Een interessant speciaal geval ontstaat als $B = 0$ en de electrische veldsterkte E opgewekt wordt door een electrische lading ter grootte Q, die bijvoorbeeld in de oorsprong van het coördinatensysteem vastgezet is. De *wet van Coulomb* zegt dat hiervoor

$$E = E(x) = \frac{Q}{4\pi\, \epsilon} \, \|x\|^{-3} \, x, \tag{1.36}$$

dus $\|E\|$ is omgekeerd evenredig met het kwadraat van de afstand. Hierin is ϵ een positieve constante die de *diëlectriciteitsconstante van het medium* genoemd wordt. Zie Kronig [50, p. 223]. In dit geval worden de bewegingsvergelijkingen (1.35) van de vorm (1.21), met

$$c = -\frac{q\, Q}{4\pi\, \epsilon\, m}. \tag{1.37}$$

Hierbij wordt het deeltje met lading q door de lading Q in de oorsprong aangetrokken, resp. afgestoten als q en Q verschillende, resp. dezelfde tekens hebben. Als q en Q verschillende tekens hebben, dan beschrijft $x(t)$ een kegelsnede (ellips, parabool of hyperbool) met de oorsprong als brandpunt. Als q en Q dezelfde tekens hebben, dan beschrijft $x(t)$ een hyperbool met de oorsprong als virtueel brandpunt. In alle gevallen wordt de snelheid waarmee de kegelsnede doorlopen wordt bepaald door Kepler's perkenwet; deze is immers equivalent met de uitspraak dat de versnelling naar de oorsprong gericht is.

1.7 Vraagstukken

Vraagstuk 1.1 Een kogel wordt van het aardoppervlak schuin omhoog gestoten en beweegt volgens Galilei's valwet. De beginsnelheid is daarbij gegeven door de vector $v \in \mathbf{R}^3$. Bereken de tijd die is verlopen als de kogel weer op het aardoppervlak terecht komt. Bereken de afgelegde afstand. Hoe moet de kogel weggestoten worden om bij gegeven waarde c van $\|v\|$ zo ver mogelijk weg te vallen? Bereken de maximale afstand, als functie van c en g. ⊘

Vraagstuk 1.2 Bewijs dat iedere oplossing $x(t)$ van (1.6) voldoet aan (1.8), bijvoorbeeld door differentiatie naar t van

$$\mathrm{e}^{t\, w/m} \left(x'(t) + \frac{m\, g}{w}\, e_3 \right).$$

Bewijs vervolgens dat (1.7) geldt en daarmee dat iedere oplossing $x(t)$ van (1.6) eenduidig bepaald is door de beginpositie $x(0)$ en de beginsnelheid $x'(0)$. Bewijs tenslotte dat voor iedere x_0, $v_0 \in \mathbf{R}^3$

de functie

$$x(t) = x_0 + \frac{1 - \mathrm{e}^{-t\,w/m}}{w/m}\left(v_0 + \frac{m\,g}{w}\,e_3\right) - \frac{m\,g\,t}{w}\,e_3, \quad t \in \mathbf{R}$$

een oplossing is van (1.6) met $x(0) = x_0$ en $x'(0) = v_0$; dit betekent dat de beginpositie en de beginsnelheid willekeurig voorgeschreven kunnen worden voor de oplossingen van (1.6). $\oslash$

Vraagstuk 1.3 Bespreek hoe de oplossingen van (1.6) convergeren naar de oplossingen van Galilei's valwet als $w \to 0$. Geef een schatting voor het verschil tussen de oplossingen, als de beginpositie, resp. beginsnelheid voor $w \neq 0$ gelijk is aan de beginpositie, resp. beginsnelheid voor $w = 0$. $\oslash$

Vraagstuk 1.4 Bewijs dat de horizontale coördinaten van de oplossing $x(t)$ van (1.6) convergeren voor $t \to \infty$ en bepaal de limiet. Bewijs dat voor alle $t \geq 0$ de horizontale afstand tot het beginpunt kleiner blijft dan m/w maal de lengte van de horizontale component van de beginsnelheid. $\oslash$

Vraagstuk 1.5 We onderzoeken het asymptotische gedrag van de oplossingen van (1.6) als $t \to -\infty$. Beschrijf eerst het gedrag van de oplossingen als $x'(0) = -\frac{m\,g}{w}\,e_3$.

Neem nu in het vervolg aan dat $x'(0) \neq -\frac{m\,g}{w}\,e_3$. Bewijs dat $\mathrm{e}^{t\,w/m}\,x(t)$ en $\mathrm{e}^{t\,w/m}\,x'(t)$ voor $t \to -\infty$ convergeren naar vectoren ongelijk aan nul en bepaal deze limietvectoren. Deze convergentie betekent exponentiële groei voor $t \to -\infty$; beargumenteer dat (1.6) geen realistisch model kan zijn als we al te ver in de tijd terugrekenen. Bewijs tenslotte dat de richtingsvectoren $\|x(t)\|^{-1}\,x(t)$ en $\|x'(t)\|^{-1}\,x'(t)$ convergeren voor $t \to -\infty$ en bepaal de limietrichtingen. $\oslash$

Vraagstuk 1.6 Beschouw een stelsel van de vorm (1.30). Bewijs dat het principe van 'actie is reactie', (1.32), impliceert dat

$$\sum_{i,\,j \mid i \neq j} \phi_{i,\,j} = 0. \tag{1.38}$$

Definieer het *zwaartepunt* als

$$z(t) = \frac{1}{M}\sum_i m_i\,x_i(t), \quad M = \sum_i m_i.$$

Bewijs dat (1.38) equivalent is met de uitspraak dat het zwaartepunt een eenparig rechtlijnige beweging beschrijft, dat wil zeggen voor alle t en t_0 in het tijdsinterval geldt dat

$$z(t) = z(t_0) + (t - t_0)\,z'(t_0).$$

$\oslash$

Vraagstuk 1.7 Het *tweelichamenprobleem* betreft twee lichamen die elkaar gravitationeel aantrekken.

a) Schrijf de bewegingsvergelijkingen (1.29) uit voor het twee-lichamenprobleem.

b) Noteer het zwaartepunt met

$$z(t) = \frac{1}{m_1 + m_2} \left(m_1\, x^{(1)}(t) + m_2\, x^{(2)}(t) \right).$$

Bewijs dat $z(t)$ een eenparig rechtlijnige beweging beschrijft.

c) Schrijf

$$y^{(i)}(t) = x^{(i)}(t) - z(t),$$

de relatieve beweging ten opzichte van het zwaartepunt van het i-de lichaam. Bewijs dat

$$\begin{aligned} y^{(1)}(t) &= \frac{m_2}{m_1 + m_2} \left(x^{(1)}(t) - x^{(2)}(t) \right), \\ y^{(2)}(t) &= \frac{m_1}{m_1 + m_2} \left(x^{(2)}(t) - x^{(1)}(t) \right). \end{aligned}$$

Bewijs dat $y^{(i)}(t)$ voldoet aan (1.21) in $\mathbf{R}^3$, met $x(t)$, resp. c vervangen door $y^{(i)}(t)$, resp. een positieve constante c_i. Bereken daarbij c_i.

d) Bewijs, voor de vergelijking (1.21) in $\mathbf{R}^3$, dat het uitwendige product

$$\mu(t) = x(t) \times x'(t)$$

van de positie- en de snelheidsvector een vector in $\mathbf{R}^3$ is die niet van t afhangt. Neem aan dat $\mu \neq 0$; dit is equivalent met de voorwaarde dat $x(t)$ en $x'(t)$ lineair onafhankelijk zijn. Zij V het orthogonale complement van μ. Bewijs dat voor alle t geldt dat $x(t) \in V$.

Ter verklaring van de notatie: er geldt dat $x(t)_3 \equiv 0$ en $x'(t)_3 \equiv 0$ dan en slechts dan als $\mu(t)$ een veelvoud is van de derde basisvector. De scalairwaardige functie in (1.10) is in dit geval gelijk aan dit veelvoud μ_3.

e) Beargumenteer dat de oplossingen van het twee-lichamenprobleem, waarvoor de afstand tussen de lichamen begrensd blijft, beschreven worden door de wetten van Kepler. Bewijs dat

$$\varepsilon(t) = \|x(t)\|^{-1}\, x(t) + \frac{1}{c}\, \mu \times x'(t)$$

een vector in $\mathbf{R}^3$ is die niet van t afhangt. Bewijs dat

$$\langle \mu(t),\, \varepsilon(t) \rangle \equiv 0.$$

f) Neem nu aan dat $x(0) \neq 0$ en $x'(0)$ is een veelvoud van $x(0)$. Bewijs dat $\mu = 0$ en dat er een constante vector ε is met $x(t) = \|x(t)\|\, \varepsilon$. Bewijs dat $\|\varepsilon\| = 1$. Schrijf $\lambda(t) = \|x(t)\|$. Bewijs dat $\lambda(t) > 0$ en dat $\lambda(t)$ voldoet aan de differentiaalvergelijking $\lambda'' = -c\, \lambda^{-2}$.

Bewijs dat als $c > 0$ en bijvoorbeeld $\xi'(0) = 0$, dan is er een eindige $T \in \mathbf{R}$, met de eigenschap dat als $t \to T$, dan convergeert $x(t)$ naar de oorsprong, het singuliere punt van het stelsel (1.21). (Voor meer details, zie Vraagstuk 3.6.)

⊘

Vraagstuk 1.8 Beschouw een lichaam dat door middel van een elastisch koord vastgemaakt is aan de oorsprong. De *wet van Hooke* zegt dat in goede benadering de spanning in het koord evenredig is met de lengte. Dit leidt tot het stelsel van bewegingsvergelijkingen

$$m\,x''(t) = -s\,x(t),$$

waarin s een positieve constante is die de stijfheid van het koord aangeeft. In het mechanische voorbeeld nemen we $x(t) \in \mathbf{R}^3$, maar we kunnen de volgende vragen beantwoorden voor $x(t) \in \mathbf{R}^n$, voor willekeurige n.

a) Bewijs dat de vrijheidsgraden onafhankelijk van elkaar bewegen, in de zin dat voor iedere $1 \le i \le n$ de versnelling $x''(t)_i$ van de $i$de component niet afhangt van de $x(t)_j$ met $j \neq i$.

b) Bewijs dat
$$x(t) = \left(\cos\sqrt{\frac{s}{m}}t\right)\,x(0) + \sqrt{\frac{m}{s}}\,\left(\sin\sqrt{\frac{s}{m}}t\right)\,x'(0).$$
Merk op dat $x(0)$ en $x'(0)$ vectoren zijn, waar we scalaire factoren gewoonlijk *voor de vector* schrijven en niet erachter.

c) Bewijs dat de beweging plaats vindt in de lineaire deelruimte van $\mathbf{R}^n$, die is opgespannen door de vectoren $x(0)$ en $x'(0)$; deze lineaire ruimte heeft dimensie kleiner of gelijk aan twee.

d) Bewijs de volgende conclusies van Newton [68, p. 149]. Als $x(0)$ en $x'(0)$ lineair onafhankelijk zijn, dan is de baan een ellips in een vlak, met de oorsprong in het middelpunt = het beeld van een cirkel onder een lineaire transformatie. (De oorsprong ligt niet in een brandpunt, zoals bij Kepler's planetenbanen.) Alle oplossingen zijn periodiek met een periode die dezelfde is voor alle oplossingen.

e) Bewijs de volgende omgekeerde stelling van Newton [68, p. 53, 54]. Als de beweging plaatvindt op een ellipsoïde met het centrum in de oorsprong en de versnelling is in de richting van de oorsprong, dan is de grootte van de versnelling evenredig met de afstand tot de oorsprong.

   Hint: de ellipsoïde kan beschreven worden als de verzameling der $x \in \mathbf{R}^n$ waarvoor $\langle Q\,x,\, x\rangle = c$, waarin Q een positief definiete symmetrische matrix is en c een positief reëel getal, beiden onafhankelijk van t. Verder is $x''(t) = -f(t)\,x(t)$ voor een reëelwaardige functie f van t. Bewijs nu achtereenvolgens dat $\langle Q\,x,\, x'\rangle = 0$, $f = \langle Q\,x',\, x'\rangle/\langle Q\,x,\, x\rangle$, f is differentieerbaar en $f' = 0$.

⊘

Vraagstuk 1.9 Beschouw het stelsel (1.35), onder de aannamen dat $E = 0$ en $B = (b,\, 0,\, 0)$ voor een positieve constante $b \in \mathbf{R}$.

a) Bewijs dat $x''(t)_1 \equiv 0$ en dat
$$x(t)_1 = x(0)_1 + x'(0)_1\,t.$$

b) Beschouw de vector

$$v(t) = \big(x'(t)_2,\, x'(t)_3\big) \in \mathbf{R}^2.$$

Bewijs dat

$$v'(t) = -c\,J(v(t)), \quad c = \frac{q\,b}{m}$$

en daarmee dat

$$\begin{aligned} v(t)_2 &= v(0)_2\,\cos ct + v(0)_3\,\sin ct, \\ v(t)_3 &= v(0)_3\,\cos ct - v(0)_2\,\sin ct. \end{aligned}$$

c) Bewijs dat $(x(t)_2,\, x(t)_3)$ met constante snelheid op een cirkel in $\mathbf{R}^2$ rondloopt. Bereken het middelpunt en de straal van deze cirkel, in termen van c, $x(0)_2$, $x(0)_3$, $x'(0)_2$ en $x'(0)_3$. Bereken ook de tijd die nodig is om de cirkel éénmaal te doorlopen.

$\oslash$

1.8 Nog Wat Meer Geschiedenis

Richard Cushman maakte mij opmerkzaam op de artikelen van Goldstein [27] over de geschiedenis van de afleiding van de baan uit Newton's bewegingsvergelijkingen (1.21). Newton schijnt zich deze vraag niet te hebben gesteld. In 1710 gaf Jakob Hermann een oplossing, welke hij meedeelde in een brief aan Johann Bernoulli. Deze antwoordde op een wat zure toon dat Hermann toewerkte naar de situatie dat de lange as van de ellips samenvalt met de x-as, door een integratieconstante te onderdrukken. Vervolgens geeft Bernoulli de algemene oplossing met een afleiding die lijkt op het hier gegeven bewijs van het constant zijn van de vector ε in (1.22). Zie [9, p. 469-480].

Laplace behandelde in [55, Partie 1, Livre II, Chap. III] het driedimensionale probleem, door constanten van beweging te bepalen die een veelterm zijn in de snelheidscoördinaten, van de graad ten hoogste twee en met willekeurige functies van de positie als coëfficiënten. Op deze manier vond hij, naast de drie componenten van μ, ook de drie componenten van de vector ε uit Vraagstuk 1.7, e) als constanten van beweging. Laplace besloot dit met een dimensiebeschouwing van de volgende soort. Omdat deze zes functies van de zes positie- en snelheidscoördinaten constant zijn langs de ééndimensionale krommen $(x(t),\, x'(t))$ met $t \in \mathbf{R}$, kunnen ze niet onafhankelijk zijn. Inderdaad hebben we de relatie $\langle \mu,\, \varepsilon \rangle = 0$.

Na Laplace werd het constant zijn van de vector ε talloze malen herontdekt; o.a. door Hamilton [30] in 1845. In de natuurkunde-literatuur staat ε bekend als de *Runge-Lenz-vector*, naar twee van de latere herontdekkers. Goldstein stelt voor om ε de *Laplace-vector* te noemen.

Is C, resp. λ het centrum, resp. de lange as van de ellips, dan heet $e = \|C\|/\lambda$ de *eccentriciteit.* Combinatie van (1.23) en (1.25) geeft dat $e = \|\varepsilon\|$, zoals ook door Hamilton [30] werd opgemerkt. Om deze reden lijkt *eccentriciteitsvector* een natuurlijke naam voor ε te zijn, deze heeft bovendien het voordeel dat alle prioriteitskwesties vermeden worden.

Joop Kolk maakte me opmerkzaam op de historische rol die de zogenaamde *vergelijking van Kepler*, zie (1.39) hieronder, heeft gespeeld in de ontwikkeling van de iteratiemethode van Newton-Raphson, zie (1.41) nóg verderop.

Om de positie $x(t)$ op de ellipsbaan te bepalen als functie van de tijd, ging Kepler [44] als volgt te werk. Zij D dat snijpunt van de lange as met de ellips, waarvoor de oorsprong O ligt op het lijnstuk tussen C en D. Door een verschuiving op de tijdas arrangeren we dat $x(0) = D$ en schrijven $x(t) = E$. Kepler maakte ook gebruik van de lineaire transformatie T in het vlak, die de lange as puntsgewijs vastlaat en de vectoren loodrecht daarop vermenigvuldigt met de factor λ/κ, waarin κ de korte as voorstelt. Onder T gaat de ellips over in de cirkel met middelpunt in C en straal gelijk aan λ. Schrijven we $F = T(E)$, dan heet de hoek $s = \text{hoek}\, DCF$ de *eccentrische anomalie*. Als p de periode is, dan verkreeg Kepler hiervoor de volgende vergelijking:

$$s - e\,\sin s = N := 2\pi\, t/p. \tag{1.39}$$

Voor het bewijs roepen we in herinnering dat de oppervlakte van het perk DOE gedurende het interval $[0,\, t]$ gelijk is aan $\pi\,\lambda\,\kappa\, t/p$. Anderzijds is dit gelijk aan $O_1 - O_2$, waarin O_1, resp. O_2 de oppervlakte is van het perk DCE, resp. de driehoek OCE. Nu voert T het perk DCE, resp. de driehoek OCE over in de cirkelsector DCF, resp. de driehoek OCF, waarvan de oppervlakte gelijk is aan $\frac{1}{2}\,\lambda^2\, s$, resp. $\frac{1}{2}\, OC\,\lambda\,\sin s$. Omdat voor iedere figuur Φ de oppervlakte van $T(\Phi)$ gelijk is aan λ/κ maal de oppervlakte van Φ, krijgen we

$$\frac{\pi\,\lambda\,\kappa\, t}{p} = \frac{\kappa}{\lambda}\left[\frac{1}{2}\,\lambda^2\, s - \frac{1}{2}\, OC\,\lambda\,\sin s\right],$$

hetgeen equivalent is met Kepler's vergelijking (1.39).

Newton [68, p. 112-114] noemde de oplossing van de niet-algebraïsche vergelijking (1.39) een moeilijk probleem. (Het is een aardig Analyse A vraagstuk om te bewijzen dat (1.39) voor iedere $t \in \mathbf{R}$ precies één oplossing $s = s(t) \in \mathbf{R}$ heeft en dat $t \mapsto s(t)$ een differentieerbare functie is.) In plaats van een directe exacte oplossing gaf Newton de oplossing als limiet van een oneindige rij $(s_i)_{i=0}^{\infty}$ van benaderingen. Startend met s_0 worden deze met volledige inductie gedefinieerd door middel van

$$s_{i+1} = s_i + \frac{N - s_i + e\,\sin s_i}{1 - e\,\cos s_i}, \quad i \in \mathbf{Z}_{\geq 0}. \tag{1.40}$$

Newton deelde mee, zonder bewijs, dat als s_0 voldoende dicht bij de exacte oplossing s ligt, dan convergeert de rij s_i naar s als $i \to \infty$ en de convergentie is zó enorm snel dat het zelden nodig is om verder dan de tweede term te gaan.

Voor een algemene vergelijking van de vorm $f(x) = 0$, waarin f een tweemaal continu differentieerbare functie is zonder dat f' nulpunten heeft in het beschouwde interval I, maakt men in de numerieke analyse gebruik van de iteratie

$$x_{i+1} = x_i - \frac{f\,(x_i)}{f'\,(x_i)}. \tag{1.41}$$

Onder de aanname van geschikte schattingen voor $f(x)$, $1/f'(x)$ en $f''(x)$ als $x \in I$, kan men bewijzen dat voor iedere startwaarde $x_0 \in I$ het voorschrift (1.41) leidt tot een oneindige rij $x_i \in I$ die convergeert naar een oplossing $x \in I$ van de vergelijking $f(x) = 0$. Bovendien is de convergentie kwadratisch, in de zin dat er een constante C is met de eigenschap dat $|x_{i+1} - x| \leq C\,|x_i - x|^2$. Anders gezegd, het aantal correcte decimalen verdubbelt ongeveer bij iedere stap.

We zien dat (1.40) equivalent is met (1.41) als we daarin $f(s) = s - e\,\sin s - N$ substitueren. Voor willekeurige veeltermen(!) $f(x)$ werd de iteratie (1.41) in een uitvoerige publicatie uit 1690 door Joseph Raphson besproken. Hij vermeldde daarbij dat hij was geïnspireerd door Newton. Overigens ontbraken zowel bij Newton als bij Raphson bewijzen met schattingen, die kwamen pas veel later.

Eerder, in een brief aan John Smith uit 1675, had Newton de verbetering

$$\widetilde{\xi} = \frac{1}{n}\left[(n-1)\,\xi + \frac{y}{\xi^{n-1}}\right] \tag{1.42}$$

voorgesteld voor een benadering ξ van de n-de machtswortel van y; we herkennen hierin (1.41) met $f(x) = x^n - y$, $x_i = \xi$ en $x_{i+1} = \widetilde{\xi}$. Hierover schreef Newton, zonder bewijs, dat het aantal nauwkeurige decimalen verdubbelt als ξ wordt vervangen door $\widetilde{\xi}$. De verbeteringsprocedure (1.42) voor n-de machtswortels was overigens al in de vijftiende eeuw ook gevonden door de Samarkand'se wiskundige Jamshid al Kāšī, zie [89, p. 207, note (38)]. De huidige naam, *methode van Newton-Raphson*, voor (1.41) ziet er historisch redelijk correct uit; Newton's formule (1.40) suggereert (1.41) veel duidelijker dan (1.42) en Raphson was blijkbaar de eerste om (1.41) als een algemene methode te presenteren.

2 Enkele Wiskundige Hulpmiddelen

Newton [68] liet zijn behandeling van de beweging van lichamen onder invloed van gravitatie, zijn "System of the World", voorafgaan door een uiteenzetting van de wiskundige beginselen, de "Principia Mathematica", die hij daarbij nodig had. In het boek van Abraham en Marsden [3] leidt dezelfde systematische opbouw, maar dan up-to-date gebracht, ertoe dat het boek begint met twee hoofdstukken, van zo'n 150 pagina's, waarin de wiskundige hulpmiddelen die gebruikt gaan worden volledig worden besproken, inclusief de bewijzen.

Wij hebben hier niet de tijd om hetzelfde te doen. Bovendien zou veel hiervan een herhaling van min of meer bekende stof uit de colleges Analyse en Infinitesimaalrekening zijn, waarbij de klassieke mechanica, waar het hier toch om gaat, naar de achtergrond zou verdwijnen. Toch lijkt het me nuttig, zeker voor de niet-wiskundigen die dit college volgen, om de wiskundige termen die regelmatig gebruikt zullen worden even de revue te laten passeren. Ik heb ervoor gekozen dit in een paar gedeelten te doen, telkens pas op het moment dat de desbetreffende begrippen gebruikt gaan worden. Ik zal proberen om de definities en de stellingen volledig te geven, maar voor de bewijzen verwijs ik naar de eerste- en tweedejaarscolleges Analyse. Voor de niet-wiskundigen zitten er hopelijk genoeg herkenningspunten in dat de tekst nog redelijk leesbaar is.

2.1 Topologie

Wij beginnen met het beschouwen van deelverzamelingen van een n-dimensionale Euclidische ruimte, die na keuze van een orthogonaal coördinatensysteeem geïdentificeerd zal worden met $\mathbf{R}^n$. We veronderstellen de lineaire algebra in $\mathbf{R}^n$ als bekend.

De Euclidische afstand van een punt $x \in \mathbf{R}^n$ tot de oorsprong is gegeven door

$$\|x\| = \left(\sum_{i=1}^{n} {x_i}^2 \right)^{1/2} \tag{2.1}$$

en wordt ook wel de *Euclidische norm* van het element $x \in \mathbf{R}^n$ genoemd. Er is een bijbehorend *Euclidisch inproduct*

$$\langle x,\, y \rangle = \sum_{i=1}^{n} x_i,\, y_i, \quad x,\, y \in \mathbf{R}^n \tag{2.2}$$

tussen paren van elementen van $\mathbf{R}^n$, in termen waarvan

$$\|x\| = \langle x,\, x \rangle^{1/2}. \tag{2.3}$$

Een basiseigenschap van het inproduct is, behalve de lineaire afhankelijkheid van ieder van de variabelen x, resp. y, de *ongelijkeid van Cauchy-Schwarz*:

$$|\langle x,\, y \rangle| \leq \|x\| \cdot \|y\|, \quad x,\, y \in \mathbf{R}^n, \tag{2.4}$$

met behulp waarvan afgeleid kan worden dat

$$\|x + y\| \leq \|x\| + \|y\|, \quad x,\, y \in \mathbf{R}^n, \tag{2.5}$$

de zogenaamde *driehoeksongelijkheid* voor de norm.

De *Euclidische afstand* $d(x, y)$ *tussen twee willekeurige punten* $x, y \in \mathbf{R}^n$ wordt gegeven door de norm van de verschilvector:

$$d(x, y) = \|x - y\|, \quad x, y \in \mathbf{R}^n. \tag{2.6}$$

Hiervoor geldt de driehoeksongelijkheid

$$d(x, z) \le d(x, y) + d(y, z), \quad x, y, z \in \mathbf{R}^n. \tag{2.7}$$

Als we ook nog gebruik maken van $\|-v\| = \|v\|$, een speciaal geval van de homogeniteit

$$\|c\,x\| = |c| \cdot \|x\|, \quad c \in \mathbf{R}, \quad x \in \mathbf{R}^n, \tag{2.8}$$

dan is (2.7) equivalent met (2.5); dit verklaart de naam van driehoeksongelijkheid voor (2.5).

Voor iedere $a \in \mathbf{R}^n$ en $r \in \mathbf{R}_{>0}$ is de *open bol, resp. gesloten bol, resp. sfeer om* a *met straal* r gedefinieerd als

$$\begin{aligned} B(a;\, r) &= \{x \in \mathbf{R}^n \mid d(x, a) < r\}, \quad \text{resp.} && (2.9)\\ \overline{B}(a;\, r) &= \{x \in \mathbf{R}^n \mid d(x, a) \le r\}, \quad \text{resp.} && (2.10)\\ S(a;\, r) &= \{x \in \mathbf{R}^n \mid d(x, a) = r\}. && (2.11) \end{aligned}$$

Zij V een deelverzameling van $\mathbf{R}^n$ en a een punt in $\mathbf{R}^n$. Men zegt dat a een *inwendig punt* is van de deelverzameling V, of dat V een *omgeving* is van het punt a, als er een $\epsilon \in \mathbf{R}_{>0}$ is, waarvoor $x \in \mathbf{R}^n$, $d(x, a) < \epsilon$ impliceert dat $x \in V$. Anders gezegd, als er een bol om a is met positieve straal, die in zijn geheel bevat is in V. De verzameling van alle inwendige punten van V heet het *inwendige* van V en wordt met V^{inw} aangeduid.

Een deelverzameling V van $\mathbf{R}^n$ heet *open* (in $\mathbf{R}^n$), als ieder element van V inwendig punt van V is, dat wil zeggen, als $V = V^{\text{inw}}$. Voor iedere deelverzameling V van $\mathbf{R}^n$ geldt dat V^{inw} een open verzameling is. Verder, als $A \subset V$ en A is open, dan is $A \subset V^{\text{inw}}$; in deze zin is V^{inw} de grootste open deelverzameling van V. Voor iedere $a \in \mathbf{R}^n$ en $r > 0$ is $B(a;\, r)$ een open delverzameling van $\mathbf{R}^n$, dit rechtvaardigt de naam 'open bol'. Men kan ook nog opmerken dat $a \in V^{\text{inw}}$ dan en slechts dan als er een open verzameling U is met $a \in U$ en $U \subset V$; op deze manier kan het begrip 'omgeving' geformuleerd worden in termen van het begrip 'open verzameling'.

Eigenschappen van de collectie van open verzamelingen zijn:

i) $\emptyset$ en $\mathbf{R}^n$ zijn open deelverzamelingen van $\mathbf{R}^n$.

ii) Als $(A_i)_{i=1}^k$ een eindige rij van open deelverzamelingen van $\mathbf{R}^n$ is, dan is de doorsnede $\bigcap_{i=1}^k A_i$ een open deelverzameling van $\mathbf{R}^n$.

iii) Is $\mathcal{I}$ een willekeurige indexverzameling, eindig of oneindig, en is $(A_i)_{i\in\mathcal{I}}$ een willekeurige collectie van open deelverzamelingen van $\mathbf{R}^n$, dan is de vereniging $\bigcup_{i\in\mathcal{I}}$ een open delverzameling van $\mathbf{R}^n$.

Soms komt het voor dat we alleen maar elementen beschouwen van een vaste deelverzameling X van $\mathbf{R}^n$, die we als ons 'universum' beschouwen. Een deelverzameling V van X heet *open in* X als er voor iedere $a \in V$ een $r = r(a) \in \mathbf{R}_{>0}$ is, met de eigenschap dat

$$B(a;\, r) \cap X \subset V.$$

Het volgt direct uit de definities dat als U een open deelverzameling is van $\mathbf{R}^n$, dan is $U \cap X$ open in X. Anderzijds, is V open in X, dan is $V = U \cap X$, als we

$$U = \bigcup_{a \in V} B(a;\, r(a))$$

nemen, met $r(a)$ als boven. Omdat U de vereniging is van een collectie van open verzamelingen, is U open. Hiermee zien we dat V open is in X, dan en slechts dan als er een open deelverzameling U van $\mathbf{R}^n$ is, waarvoor $V = U \cap X$.

Noteren we $\mathcal{O}$ voor de collectie van alle verzamelingen die open in X zijn, dan krijgen we de volgende eigenschappen.

i) $\emptyset \in \mathcal{O}$ en $X \in \mathcal{O}$. (Hierin is $\emptyset$ de lege verzameling.)

ii) Als $A_i \in \mathcal{O}$ voor iedere $1 \leq i \leq k$, dan is $\bigcap_{i=1}^{k} A_i \in \mathcal{O}$.

iii) Als $A_i \in \mathcal{O}$ voor iedere $i \in \mathcal{I}$, dan is $\bigcup_{i \in \mathcal{I}} A_i \in \mathcal{O}$.

Is X een willekeurige verzameling, niet noodzakelijkerwijze een deelverzameling van $\mathbf{R}^n$, dan heet een collectie $\mathcal{O}$ van deelverzamelingen van X met deze eigenschappen (i)—(iii) een *topologie in* X. Het paar $(X,\, \mathcal{O})$ heet dan een *topologische ruimte.* De definitie in termen van bollen, die op zijn beurt gedefinieerd zijn in termen van de Euclidische afstand, maakt dat we iedere deelverzameling X van $\mathbf{R}^n$ op kunnen vatten als een topologische ruimte. Is het duidelijk welke topologie bedoeld wordt, dan spreekt men vaak over 'de topologische ruimte X' of 'de ruimte X', in plaats van 'de topologische ruimte $(X,\, \mathcal{O})$'.

In een willekeurige topologische ruimte $(X,\, \mathcal{O})$, met $V \subset X$ en $a \in X$, zegt men dat a een inwendig punt is van V en dat V een omgeving is van a, als er een $U \in \mathcal{O}$ is met $a \in U$ en $U \subset V$. Dit generaliseert de situatie van $X = \mathbf{R}^n$ met de gebruikelijke topologie. Op dezelfde manier noemt men een deelverzameling V van X *open ten aanzien van de topologie* $\mathcal{O}$ als $V \in \mathcal{O}$.

Extreme, maar flauwe, andere voorbeelden van topologische ruimten zijn:

a) X is een willekeurige verzameling en $\mathcal{O}$ is de collectie van *alle* deelverzamelingen van X. Deze $\mathcal{O}$ heet de *discrete topologie* in X, $(X,\, \mathcal{O})$ heet een *discrete topologische ruimte.* Vanwege eigenschap (iii) van een topolgie is $\mathcal{O}$ discreet, dan en slechts dan als voor iedere $a \in X$ geldt dat $\{a\} \in \mathcal{O}$.

b) X is een willekeurige verzameling en $\mathcal{O}$ heeft als enige elementen $\emptyset$ en X. Deze $\mathcal{O}$ heet de *triviale topologie* in X.

Commentaren bij a): In het algemeen heet a een *geïsoleerd punt* van X als $\{a\} \in \mathcal{O}$; dit is equivalent met existentie van een omgeving V van a in X, met de eigenschap dat $V \cap X = \{a\}$. X heet *discreet* dan en slechts dan als X uitsluitend geïsoleerde punten heeft.

Een deelverzameling X van $\mathbf{R}^n$ is een discrete topologische ruimte, dan en slechts dan als er bij iedere $a \in X$ een $r > 0$ is, met de eigenschap dat $B(a;\, r) \cap X = \{a\}$. Dat wil zeggen, als $x \in X$ en $d(x,\, a) < r$, dan is $x = a$. Iedere eindige deelverzameling van $\mathbf{R}^n$ is discreet. Maar ook de verzameling van de $1/k$ met $k \in \mathbf{Z}_{>0}$ is een discrete deelverzameling van $\mathbf{R}$.

Zij $(X,\, \mathcal{O})$ een topologische ruimte, bijvoorbeeld een deelverzameling X van $\mathbf{R}^n$. Zij $V \subset X$ en $a \in X$. Men zegt dat a een *verdichtingspunt* van V is, als er bij iedere omgeving U van a in

X een $x \in V$ is met $x \in U$. Als X een deelverzameling van $\mathbf{R}^n$ is, betekent dit dat er voor iedere $r > 0$ (hoe klein ook) een $x \in V$ is met $d(x, a) < r$. De verzameling van verdichtingspunten van V in X heeft de *afsluiting van V in X* en wordt met $\overline{V}$ aangeduid, soms met $\overline{V} \cap X$ als men wil benadrukken dat men binnen de verzameling X blijft. Men noemt V *gesloten in* X als $\overline{V} \cap X = V$.

De definitie van 'afsluiting' is geformuleerd in termen van de topologie, de collectie van open deelverzamelingen van X en is daarmee een voorbeeld van een *topologisch begrip*. De relatie is echter nog directer. Voor iedere $V \subset X$ heet

$$X \setminus V := \{x \in X \mid x \notin V\} \tag{2.12}$$

het *complement van V in X*. Het is nu een stelling dat

$$\overline{V} = X \setminus \left(U^{\text{inw}}\right) \quad \text{als} \quad U = X \setminus V. \tag{2.13}$$

Alle eigenschappen van inwendigen van verzamelingen en open verzamelingen leiden hiermee tot corresponderende eigenschappen van afsluitingen van verzamelingen en gesloten verzamelingen:

- V is gesloten in X, dan en slechts dan als $X \setminus V$ open is in X, dat wil zeggen $X \setminus V \in \mathcal{O}$.
- $\overline{V} \cap X$ is gesloten in X en bevat V.
- Als W gesloten is in X en $V \subset W$, dan is $\overline{V} \cap X \subset W$; in deze zin is $\overline{V} \cap X$ de kleinste verzameling die V bevat en gesloten is in X.
- X en $\emptyset$ zijn gesloten in X.
- De vereniging van een eindige collectie van gesloten deelverzamelingen is gesloten.
- De doorsnede van een willekeurige collectie van gesloten deelverzamelingen is gesloten.

Waarschuwing: anders dan in het dagelijks taalgebruik is 'open' niet equivalent met 'niet gesloten'. Er zijn talloze verzamelingen die nóch open, nóch gesloten zijn. Er zijn ook verzamelingen die zowel open alsook gesloten zijn, bijvoorbeeld de lege verzameling $\emptyset$ en de hele verzameling X.

Zij $(X, \mathcal{O})$ een topologische ruimte. Een *open overdekking* van X is een collectie $\mathcal{A}$ van open verzamelingen in X, dat wil zeggen $\mathcal{A} \subset \mathcal{O}$, met de eigenschap dat $X = \bigcup_{A \in \mathcal{A}} A$. Dat wil zeggen, voor iedere $x \in X$ is er minstens één $A \in \mathcal{A}$, waarvoor $x \in A$. X heet *compact* als iedere open overdekking $\mathcal{A}$ van X een eindige deeloverdekking $\mathcal{E}$ van X bezit. Dat wil zeggen, er is een eindige deelcollectie $\mathcal{E}$ van $\mathcal{A}$, waarvoor nog steeds $X = \bigcup_{A \in \mathcal{E}} A$. Het is een vrij niet-triviale stelling dat voor een deelverzameling X van $\mathbf{R}^n$ de volgende uitspraken (i)—(iii) equivalent zijn.

i) X is compact.

ii) X is gesloten in $\mathbf{R}^n$ en begrensd.

iii) Bij iedere oneindige rij $\left(x^{(i)}\right)_{i=1}^{\infty}$ in X is er rij $i(j)_{j=1}^{\infty}$ van rangnummers, met $i(j) \to \infty$ als $j \to \infty$ en een $a \in X$ met de eigenschap dat de *deelrij* $y^{(j)} = x^{(i(j))}$ naar a convergeert als $j \to \infty$.

Hierbij betekent X is *begrensd* dat er een constante C is, met de eigenschap dat voor iedere $x \in X$ geldt dat $\|x\| \leq C$. Verder zegt men dat de rij $y^{(j)}$ naar a *convergeert* als $i \to \infty$, notatie

$$\lim_{j\to\infty} y^{(j)} = a,$$

als er bij iedere $\epsilon > 0$ een rangnummer N is, met de eigenschap $\|y^{(j)} - a\| < \epsilon$ zodra $j \geq N$.

Men gebruikt meestal ii) om van een gegeven verzameling na te gaan dat-ie compact is. De overdekkingseigenschap wordt vaak gebruikt om een eigenschap die voor de $A \in \mathcal{A}$ gelden te 'globaliseren' naar de corresponderende eigenschap voor X, in de situatie dat de eigenschap geldt voor de vereniging van eindig veel verzamelingen als ze voor ieder van de verzamelingen geldt. De implicatie ii) $\Rightarrow$ iii) heet de *stelling van Bolzano-Weierstrass* en de implicatie ii) $\Rightarrow$ i) is de *stelling van Heine-Borel*. Waarschuwing: de voorwaarde dat X een deelverzameling van $\mathbf{R}^n$ is, is essentieel. De stelling is bijvoorbeeld niet waar in een oneindig-dimensionale genormeerde lineaire ruimte, bijvoorbeeld in de ruimte van alle continue functies op $[0, 1]$, voorzien van de supremumnorm.

Een topologische ruimte X heet *samenhangend* als X niet gelijk is aan de vereniging van twee disjuncte en niet-lege open deelverzamelingen. Een deelverzameling X van $\mathbf{R}$ is samenhangend, dan en slechts dan als X een interval is.

Men noemt X *boogsamenhangend*, als er voor iedere $p, q \in X$ een continue kromme $\gamma : [a, b] \to X$ van p naar q is, dat wil zeggen waarvoor $\gamma(a) = p$ en $\gamma(b) = q$. Iedere boogsamenhangende ruimte is samenhangend.

X heet *lokaal boogsamenhangend* als iedere $p \in X$ een boogsamenhangende omgeving U in X heeft. Iedere open deelverzameling van $\mathbf{R}^n$ is lokaal boogsamenhangend.

Zij X lokaal boogsamenhangend. Voor iedere $p \in X$ definiëren we $X(p)$ als de verzameling der $q \in X$ waarvoor er een continue kromme in X is die van p naar q gaat. Het is evident dat $p \in X(p)$, dus $X(p)$ is niet leeg. De verzameling $X(p)$ is boogsamenhangend en open in X. Voor iedere $p, q \in X$ zijn de volgende uitspraken equivalent:

i) $q \in X(p)$.

ii) $X(p) \cap X(q) \neq \emptyset$.

iii) $X(p) = X(q)$.

Dit impliceert dat $X \setminus X(p)$ gelijk is aan de vereniging van de open verzamelingen $X(q)$ met $q \notin X(p)$, dus ook $X \setminus X(p)$ is open. Hieruit volgt op zijn beurt dat X samenhangend is dan en slechts dan als X boogsamenhangend is. De verzameling $X(p)$ heet de *samenhangscomponent van p in X*.

2.2 Convergentie

Zij $X \subset \mathbf{R}^n$ en f een afbeelding van X naar $\mathbf{R}^p$, $a \in \overline{X}$, $b \in \mathbf{R}^p$. (Is $p = 1$ dan spreekt men van een reëelwaardige functie. Is $p > 1$, $n = 1$ en X een interval, dan spreekt men van een kromme in de p-dimensionale ruimte.) Men zegt dat $f(x)$ *convergeert naar b als x naar a gaat*, notatie

$$\lim_{x\to a} f(x) = b,$$

als er bij iedere $\epsilon > 0$ een $\delta > 0$ is, met de eigenschap dat als $x \in X$ en $\|x - a\| < \delta$, dan is $\|f(x) - b\| < \epsilon$.

Als $a \in X$, dan zegt men dat f *continu in het punt* a is als $f(x)$ naar $f(a)$ convergeert voor $x \to a$. Dit is equivalent met: bij iedere (open) omgeving V van $f(a)$ in $\mathbf{R}^n$ is er een (open) omgeving U van a in X, met de eigenschap $f(x) \in V$ zodra $x \in U$. Deze laatste eigenschap kan ook geschreven worden als $f(U) \subset V$, ofwel $U \subset f^{-1}(V)$. Hierin is het *beeld* $f(U)$ *van* U *onder* f, resp. het *volledige origineel* $f^{-1}(V)$ *van* V *voor de afbeelding* f gedefinieerd als

$$\begin{aligned} f(U) &:= \{f(x) \mid x \in U\}, \quad \text{resp.} & (2.14)\\ f^{-1}(V) &:= \{x \in X \mid f(x) \in V\}. & (2.15) \end{aligned}$$

Dit leidt tot de volgende generalisatie voor een afbeelding f van een willekeurige topologische ruimte X naar een willekeurige topologische ruimte Y. Men zegt daarvoor dat f continu is in het punt $a \in X$, als er voor iedere omgeving V van $f(a)$ in Y een omgeving U van a in X is, met de eigenschap dat $f(U) \subset V$, ofwel $U \subset f^{-1}(V)$.

De afbeelding $f : X \to Y$ heet *continu* als voor iedere $a \in X$ geldt dat f continu is in het punt a. Uit de definitie volgt bijna direct dat f continu is, dan en slechts dan als voor iedere open deelverzameling V in Y geldt dat $f^{-1}(V)$ open is in X.

Waarschuwing: niet hoeft te gelden dat $f(U)$ open is als U open is. Bijvoorbeeld, de constante afbeelding, die aan iedere $x \in X$ hetzelfde element $c \in Y$ toevoegt, is evident continu. Echter $f(X) = \{c\}$ en dit is alleen een open deelverzameling van Y als c een geïsoleerd punt van Y is. Wel geldt: Is X compact en $f : X \to Y$ continu, dan is de beeldverzameling $f(X)$ ook compact. Als X een begrensde en gesloten deelverzameling is van $\mathbf{R}^n$ en $f : X \to \mathbf{R}^p$ is continu, dan is $f(X)$ een begrensde en gesloten deelverzameling van $\mathbf{R}^p$.

Overgaan op complementen geeft ook dat f continu is, dan en slechts dan als voor iedere gesloten deelverzameling V in Y geldt dat $f^{-1}(V)$ gesloten is in X. Typische toepassingen: Zij f een reëelwaardige functie op X. Dan is, voor iedere $c \in \mathbf{R}$, de verzameling

$$U_{<c} = \{\, x \in X \mid f(x) < c\,\}$$

een open deelverzameling van X, omdat $U_{<c}$ het volledige origineel is van het open interval $]-\infty, c[$ voor de continue functie f. Anderzijds zijn

$$U_{\leq c} = \{x \in X \mid f(x) \leq c\}, \quad \text{resp.} \quad U_c = \{x \in X \mid f(x) = c\}$$

gesloten deelverzamelingen van X, want dit zijn de volledige originelen van de gesloten deelverzameling $]-\infty, c]$, resp. $\{c\}$ van $\mathbf{R}$ voor de continue functie f.

Een afbeelding

$$f : x \mapsto (f_1(x), \dots, f_p(x))$$

van X naar $\mathbf{R}^p$ is continu, dan en slechts dan als ieder van de coördinaatsfuncties f_i een reëelwaardige continue functie op X is. In het algemeen kan echter de continuïteit van een functie van meer variabelen (functie op $\mathbf{R}^n$ met $n > 1$) niet teruggebracht worden tot continuïteit van functies van één variabele.

Wel hebben we een groot aantal rekenregels voor continuïteit, zoals de som, het product en quotiënt (met noemer niet gelijk aan nul) en de samenstelling van continue functies, resp. afbeeldingen is weer continu. Dit maakt dat alle algebraïsche combinaties van elementaire functies, waarvan men de continuïteit ooit eens heeft bewezen, vanzelf weer continu zijn. Zonder die rekenregels zou het verifiëren van de continuïteit, in termen van ϵ-δ-schattingen, zeer bewerkelijk kunnen worden. (Het bewijs van de rekenregels en de continuïteit van de elementaire functies berust overigens wèl op ϵ-δ-schattingen.)

2.3 Differentieerbare Afbeeldingen

Zij U een open deelverzameling van $\mathbf{R}^n$ en f een afbeelding van U naar $\mathbf{R}^p$. Men zegt dat f *differentieerbaar is in het punt* $a \in U$, als er een lineaire afbeelding L van $\mathbf{R}^n$ naar $\mathbf{R}^p$ is, waarvoor $f(x) - f(a)$ benaderd wordt door $L\,(x-a)$ als $x \to a$, in de zin dat

$$\lim_{x \to a} \frac{\|f(x) - f(a) - L\,(x-a)\|}{\|x-a\|} = 0. \tag{2.16}$$

Hierin beperkt men zich bij het limietnemen uiteraard tot de $x \in U$ met $x \neq a$. Uitgeschreven in termen van de definitie van limiet betekent dit dat er bij iedere $\epsilon > 0$ een $\delta > 0$ is, met de eigenschap dat als $x \in U$ en $\|x - a\| < \delta$, dan is

$$\|f(x) - f(a) - L\,(x-a)\| \leq \epsilon\, \|x-a\|. \tag{2.17}$$

Merk op dat als δ voldoende klein is, dan volgt $x \in U$ uit $\|x-a\| < \delta$, omdat U open was verondersteld. Men formuleert (2.16) ook wel als: '$f(x) - f(a)$ is gelijk aan $L\,(x-a)$ plus een restterm die van kleinere orde is dan $\|x-a\|$ als $x \to a$', notatie

$$f(x) - f(a) - L\,(x-a) = o\left(\|x-a\|\right), \quad x \to a. \tag{2.18}$$

Voor iedere $v \in \mathbf{R}^n$ geldt dat

$$L(v) = \lim_{t \to 0} \frac{1}{t}\, \left[f(a + t\,v) - f(a)\right]. \tag{2.19}$$

Dit laat zien dat L eenduidig bepaald is bij gegeven f en a, men noemt dit de *(totale) afgeleide* van f in het punt a en noteert deze met $L = Df(a)$. Voor iedere $v \in \mathbf{R}^n$ heet $Df(a)(v)$, de limiet in (2.19), de *afgeleide van f in de richting van de vector v*. Men noemt $f : U \to \mathbf{R}^p$ *differentieerbaar* als voor iedere $a \in U$ geldt dat f differentieerbaar is in het punt a.

In (2.19) voor v de j-de standaard basisvector in $\mathbf{R}^n$ invullend, zien we dat de differentieerbaarheid van f impliceert dat f partieel differentieerbaar is naar de j-de variabele en dat

$$\frac{\partial f(x)}{\partial x_j}|_{x=a} = Df(a)\,(e_j)\,. \tag{2.20}$$

Omgekeerd is het een stelling dat als f partieel differentieerbaar is naar iedere variabele en de partiële afgeleiden zijn continue functies, dan is f differentieerbaar volgens de bovenstaande definitie. Gebruik makend van de decompositie $v = \sum_j v_j\, e_j$ en de lineariteit van $Df(x)$, krijgen we bovendien dat

$$Df(x)(v) = \sum_{j=1}^{n} v_j \frac{\partial f(x)}{\partial x_j}, \tag{2.21}$$

hetgeen ook wel de *somregel* voor de richtingsafgeleide genoemd wordt. Hiermee is de totale afgeleide uitgedrukt in termen van de partiële afgeleiden. Preciezer: de matrix van de lineaire afbeelding $Df(x)$ heeft op de i-de rij en j-de kolom de partiële afgeleide

$$\left(\partial_j f_i\right)(x) := \partial f_i(x)/\partial x_j$$

van de i-de coördinaatsfunctie f_i naar de j-de variabele staan. Als $p = n$ dan noemt de matrix van $Df(x)$, resp. de determinant van $Df(x)$ ook wel de *Jacobi-matrix*, resp. *Jacobi-determinant* van f in het punt x.

Men noemt $f : U \to \mathbf{R}^p$ *continu differentieerbaar* als f differentieerbaar is en

$$Df : x \mapsto Df(x) : U \to \mathrm{L}\,(\mathbf{R}^n,\, \mathbf{R}^p)$$

is continu. Hierin is $\mathrm{L}\,(\mathbf{R}^n,\, \mathbf{R}^p)$ de ruimte van lineaire afbeeldingen van $\mathbf{R}^n$ naar $\mathbf{R}^p$. Met de gebruikelijke (puntsgewijze) optelling en scalairvermenigvuldiging is $\mathrm{L}\,(\mathbf{R}^n,\, \mathbf{R}^p)$ een lineaire ruimte van dimensie $n\,p$; de afbeelding die aan een lineaire afbeelding zijn matrixcoëfficiënten toevoegt is een lineair isomorfisme van $\mathrm{L}\,(\mathbf{R}^n,\, \mathbf{R}^p)$ naar $\mathbf{R}^{n\,p}$. De bovengenoemde stelling zegt dat f continu differentieerbaar is, dan en slechts dan als voor iedere $1 \leq i \leq p$ en $1 \leq j \leq n$ de functie f_i partieel differentieerbaar is naar de j-de variabele, met continue partiële afgeleide $\partial_j f_i : U \to \mathbf{R}$.

2.4 Hogere Orde Afgeleiden

Met inductie over k definieert men voor iedere $k \in \mathbf{Z}_{>0}$ dat $f : U \to \mathbf{R}^p$ *k keer differentieerbaar is* als f $k-1$ keer differentieerbaar is en $D^{k-1}f$ differentieerbaar is. Hierbij wordt $D^k f$ met inductie over k gedefinieerd als $D^k f = D\left(D^{k-1} f\right)$. Men noemt f k keer continu differentieerbaar als f k keer differentieerbaar is en $D^k f$ continu is. De ruimte van k keer continu differentieerbare afbeeldingen van U naar $\mathbf{R}^p$ wordt met $\mathrm{C}^k\,(U,\, \mathbf{R}^p)$ aangeduid. Voor $k = 0$ kan men dit als de ruimte van continue afbeeldingen van U naar $\mathbf{R}^p$ opvatten.

Voor iedere $x \in U$ is $D^k f(x)$ een element van de lineaire ruimte $\mathrm{L}^k\,(\mathbf{R}^n,\, \mathbf{R}^p)$, die met inductie over k gedefinieerd is als

$$\mathrm{L}^k\,(\mathbf{R}^n,\, \mathbf{R}^p) = \mathrm{L}\left(\mathbf{R}^n,\, \mathrm{L}^{k-1}\,(\mathbf{R}^n,\, \mathbf{R}^p)\right).$$

Als $A \in \mathrm{L}^k\,(\mathbf{R}^n,\, \mathbf{R}^p)$ en $v^{(i)} \in \mathbf{R}^n$, $1 \leq i \leq k$, dan is

$$\widetilde{A}\left(v^{(1)},\, v^{(2)},\, \ldots,\, v^{(k)}\right) := \left(\ldots\left(A\left(v^{(1)}\right)\right)\left(v^{(2)}\right)\ldots\right)\left(v^{(k)}\right) \in \mathbf{R}^p.$$

Anders gezegd, $\widetilde{A}$ is een afbeelding van $(\mathbf{R}^n)^k$ naar $\mathbf{R}^p$. Deze afbeelding is *k-lineair*, in de zin dat voor iedere i en na keuze van de $v^{(j)}$ met $j \neq i$ de afbeelding

$$v^{(i)} \mapsto \widetilde{A}\left(v^{(1)},\, v^{(2)},\, \ldots,\, v^{(k)}\right)$$

een lineaire afbeelding is van $\mathbf{R}^n$ naar $\mathbf{R}^p$. De afbeelding $A \mapsto \widetilde{A}$ is een lineair isomorfisme van $\mathrm{L}^k\,(\mathbf{R}^n,\, \mathbf{R}^p)$ naar de ruimte van k-lineaire afbeeldingen van $(\mathbf{R}^n)^k$ naar $\mathbf{R}^p$. Het is gebruikelijk om A met $\widetilde{A}$ te identificeren en de ruimte $\mathrm{L}^k\,(\mathbf{R}^n,\, \mathbf{R}^p)$ met de ruimte van k-lineaire afbeeldingen van $(\mathbf{R}^n)^k$ naar $\mathbf{R}^p$.

Ter verduidelijking schrijven we het voorgaande uit voor $k = 2$. Voor iedere $x \in U$ is $D^2 f(x)$ een element van de lineaire ruimte $\mathrm{L}^2\,(\mathbf{R}^n,\, \mathbf{R}^p)$, die gedefinieerd is als

$$\mathrm{L}^2\,(\mathbf{R}^n,\, \mathbf{R}^p) = \mathrm{L}\,(\mathbf{R}^n,\, \mathrm{L}\,(\mathbf{R}^n,\, \mathbf{R}^p)).$$

Als $A \in \mathrm{L}^2\,(\mathbf{R}^n,\, \mathbf{R}^p)$ en $v^{(1)},\, v^{(2)} \in \mathbf{R}^n$, dan is

$$\widetilde{A}\left(v^{(1)},\, v^{(2)}\right) := \left(A\left(v^{(1)}\right)\right)\left(v^{(2)}\right) \in \mathbf{R}^p.$$

Anders gezegd, $\widetilde{A}$ is een afbeelding van $(\mathbf{R}^n)^2 = \mathbf{R}^n \times \mathbf{R}^n$ naar $\mathbf{R}^p$. Deze afbeelding is *bilineair*, in de zin dat voor iedere $v^{(2)} \in \mathbf{R}^n$ de afbeelding

$$v^{(1)} \mapsto \widetilde{A}\left(v^{(1)},\, v^{(2)}\right)$$

een lineaire afbeelding is van $\mathbf{R}^n$ naar $\mathbf{R}^p$ en voor iedere $v^{(1)} \in \mathbf{R}^n$ de afbeelding

$$v^{(2)} \mapsto \widetilde{A}\left(v^{(1)},\, v^{(2)}\right)$$

een lineaire afbeelding is van $\mathbf{R}^n$ naar $\mathbf{R}^p$. De afbeelding $A \mapsto \widetilde{A}$ is een lineair isomorfisme van $\mathrm{L}^2\left(\mathbf{R}^n,\, \mathbf{R}^p\right)$ naar de ruimte van bilineaire afbeeldingen van $\mathbf{R}^n \times \mathbf{R}^n$ naar $\mathbf{R}^p$. Het is gebruikelijk om A met $\widetilde{A}$ te identificeren en de ruimte $\mathrm{L}^k\left(\mathbf{R}^n,\, \mathbf{R}^p\right)$ met de ruimte van bilineaire afbeeldingen van $\mathbf{R}^n \times \mathbf{R}^n$ naar $\mathbf{R}^p$.

Een belangrijk feit uit de differentiaalrekening is de *verwisselbaarheid van de differentiatievolgorde* bij herhaalde partiële differentiatie met betrekking tot verschillende variabelen. Een vorm hiervan die vrij zuinig is wat betreft de voorwaarden luidt als volgt. Zij U een open rechthoek om $(a,\, b)$ in $\mathbf{R}^2$, $f \in \mathrm{C}^1\left(U,\, \mathbf{R}\right)$, $\partial_1 f : U \to \mathbf{R}$ partieel differentieerbaar naar de tweede variabele en $\partial_2\partial_1 f : U \to \mathbf{R}$ continu. Dan is $\partial_2 f : U \to \mathbf{R}$ partieel differentieerbaar naar de eerste variabele en $\partial_1\partial_2 f = \partial_2\partial_1 f$.

Bewijs: de aannamen maken dat we kunnen schrijven

$$\begin{aligned} f(x,\, y) &= f(a,\, y) + \int_a^x \partial_1 f(\xi,\, y)\, d\xi \\ &= f(a,\, b) + \int_b^y \partial_2 f(a,\, \eta)\, d\eta \\ &\quad + \int_a^x \partial_1 f(\xi,\, b)\, d\xi + \int_a^x \int_b^y \partial_2\partial_1 f(\xi,\, \eta)\, d\eta\, d\xi. \end{aligned}$$

In de dubbelintegraal mogen we de integratievolgorde verwisselen omdat de integrand een continue functie is van beide variabelen. Daarmee komt de integratie over de tweede variabele als laatste, dus differentiatie naar y geeft dat

$$\partial_2 f(x,\, y) = \partial_2 f(a,\, y) + \int_a^x \partial_2\partial_1 f(\xi,\, y)\, d\xi.$$

Het rechterlid is partieel differentieerbaar naar x, met afgeleide gelijk aan $\partial_2\partial_1 f(x,\, y)$.

Voor $f \in \mathrm{C}^k\left(U,\, \mathbf{R}^p\right)$ volgt hieruit dat als we k partiële differentiaties uitvoeren, waarbij we de j-de keer partieel differentiëren naar de $\iota(j)$-de variabele, dus $\iota : \{1,\, \dots,\, k\} \to \{1,\, \dots,\, n\}$, en $\pi : \{1,\, \dots,\, k\} \to \{1,\, \dots,\, k\}$ is een permutatie van de differentiatievolgorde, dan is is het resultaat gelijk aan k partiële differentiaties waarbij we de j-de keer partieel differentiëren naar de $\iota(\pi(j))$-de variabele. Dit berust op het combinatorische feit dat iedere permutatie een samenstelling is van eindig veel buursverwisselingen. Noteren we voor iedere $1 \leq i \leq n$ met α_i het aantal der j waarvoor $\iota(j) = i$, dan kunnen we de bovengenoemde k-de orde afgeleide in de standaardvorm

$$\partial^\alpha f := \partial_n^{\alpha_n} \circ \ldots \circ \partial_1^{\alpha_1} f$$

brengen. Hierin wordt $\alpha : \{1,\, \dots,\, n\} \to \mathbf{Z}_{\geq 0}$ de *multi-index* van de differentiatie genoemd en de *orde* k van de differentiatie wordt genoteerd met

$$|\alpha| := \sum_{i=1}^n \alpha_i.$$

Een k-lineaire afbeelding $A : (\mathbf{R}^n)^k \to \mathbf{R}^p$ heet *symmetrisch* als voor iedere permutatie $\pi : \{1, \dots, k\} \to \{1, \dots, k\}$ en alle $v^{(i)} \in \mathbf{R}^n$ geldt dat

$$A\left(v^{(\pi(1))}, \dots, v^{(\pi(n))}\right) = A\left(v^{(1)}, \dots, v^{(n)}\right).$$

De ruimte van dergelijke symmetrische A wordt soms met $\mathrm{S}^k\left(\mathbf{R}^n, \mathbf{R}^k\right)$ aangeduid. De verwisselbaarheid van de differentiatievolgorde geeft nu dat als $f \in \mathrm{C}^k\left(U, \mathbf{R}p\right)$ en $x \in U$, dan is $D^k f(x) \in \mathrm{S}^k\left(\mathbf{R}^n, \mathbf{R}^p\right)$.

Is $f \in \mathrm{C}^k\left(U, \mathbf{R}^p\right)$ en is het rechte lijnstuk tussen $a \in \mathbf{R}^n$ en $a + h \in \mathbf{R}^n$ bevat in U, dan is $\phi : t \mapsto f(x + t\,h)$ een k keer continu differentieerbare functie van de reële variabele $t \in [0, 1]$. Hiervoor de *Taylor-ontwikkeling met integraal-restterm* toepassend, krijgen we meteen een n-dimensionale variant hiervan:

$$\begin{aligned} f(a+h) &= \sum_{j=0}^{k} \frac{1}{j!} \frac{d^j f(a+t\,h)}{dt^j}\Big|_{t=0} + R_k(a, h), \\ R_k(a, h) &:= \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \frac{d^k f(a+s\,h)}{ds^k}\Big|_{s=0}^{s=t} dt. \end{aligned}$$

Hierbij kan opgemerkt worden dat alles uitgedrukt is in termen van

$$\frac{d^j f(a+t\,h)}{dt^j} = D^j f(a+t\,h)(h, \dots, h),$$

waarbij de puntjes betekent dat op ieder van de j plaatsen de vector h staat. Met de notaties

$$\begin{aligned} h^\alpha &= \prod_{i=1}^{n} {h_i}^{\alpha_i} \\ \alpha! &= \prod_{i=1}^{n} \alpha_i \end{aligned}$$

kan men deze uitdrukkingen nog herschrijven in termen van de k-de orde partiële afgeleiden $\partial^\alpha f$, door middel van de identiteit

$$\frac{1}{j!} D^j f(x)(h, \dots, h) = \sum_{\alpha||\alpha|=j} \frac{h^\alpha}{\alpha!} \partial^\alpha f(x).$$

De ruimte van willekeurig vaak differentieerbare afbeeldingen $f : U \to \mathbf{R}^p$ wordt met $\mathrm{C}^\infty\left(U, \mathbf{R}^p\right)$ aangeduid. Het is hierbij duidelijk dat

$$\mathrm{C}^\infty\left(U, \mathbf{R}^p\right) = \bigcap_{k=0}^{\infty} \mathrm{C}^k\left(U, \mathbf{R}^p\right).$$

Hiermee hebben we de ruimten $\mathrm{C}^k\left(U, \mathbf{R}^p\right)$ ingevoerd waarbij $k \in \mathbf{Z}_{\geq 0}$ of $k = \infty$. Men noteert dat $k < \infty$ als $k \in \mathbf{Z}_{\geq 0}$ en $\infty \pm l = \infty$ voor iedere $l \in \mathbf{Z}$. Dit maakt een uitspraak mogelijk als: voor iedere $1 \leq k \leq \infty$ geldt dat als $f \in \mathrm{C}^k$, dan is is $Df \in \mathrm{C}^{k-1}$. Voor $k = \infty$ moet deze uitspraak dan gelezen worden als: $f \in \mathrm{C}^\infty$ impliceert dat $Df \in \mathrm{C}^\infty$. Het voordeel van de ruimte C^∞ is dat geen boekhouding van de differentieerbaarheidsgraad bijgehouden hoeft te worden, het nadeel is dat de eis dat de functies in kwestie willekeurig vaak differentieerbaar zijn in veel situaties nogal excessief is. Gemakshalve noemt men een afbeelding tegenwoordig ook wel *glad* als zij willekeurig vaak differentieerbaar is.

2.5 Analytische Afbeeldingen

De afbeelding $f : U \to \mathbf{R}^p$ heet *reëel-analytisch* als f *machtreeksontwikkelbaar* is, in de volgende zin. Bij iedere $a \in U$ is er een $\delta > 0$ en is er voor iedere multi-index α een coëfficiënt $c_\alpha \in \mathbf{R}^p$, met de eigenschap dat als $h \in \mathbf{R}^n$ en $\|h\| \le \delta$, dan is $a + h \in U$ en

$$f(a+h) = \sum_\alpha h^\alpha\, c_\alpha,$$

waarbij de convergentie uniform is voor h in de gesloten bol om de oorsprong met straal gelijk aan δ. Het is een enigszins verrassend feit dat dit impliceert dat $f \in \mathrm{C}^\infty$ en dat voor iedere multi-index α geldt dat

$$c_\alpha = \frac{1}{\alpha!}\partial^\alpha f(a).$$

Anders gezegd, de machtreeks is gelijk aan de Taylor-reeks van f in het punt a.

Omdat we in de machtreeks net zo goed $h \in \mathbf{C}^n$ kunnen substitueren, krijgt men dat iedere reëel-analytische functie uitgebreid kan worden tot een willekeurig vaak complex-differentieerbare functie in een open omgeving van U in $\mathbf{C}^n$. Het is een verrassend feit, dat bewezen kan worden met behulp van de integraalformule van Cauchy, dat omgekeerd een complex-differentieerbare functie f (met continue eerste-orde afgeleiden) machtreeksontwikkelbaar is. Daarmee is een eenmaal complex-differentieerbare functie automatisch willekeurig vaak complex-differentieerbaar en gelijk aan zijn Taylor-reeks. Het is dikwijls een interessante vraag wat de 'maximale' complex-analytische uitbreiding is van een gegeven reëel-analytische functie, we gaan hier echter niet verder op in.

De ruimte van reëel-analytische afbeeldingen van U naar $\mathbf{R}^p$ wordt met $\mathrm{C}^\omega\,(U,\, \mathbf{R}^p)$ genoteerd. Hierbij hanteren we de afspraken dat $k < \infty < \omega$ en $\omega + k = \omega$ als $k \in \mathbf{Z}$.

2.6 De Impliciete-functiestelling

Een belangrijk hulpmiddel voor het onderzoek van vergelijkingen in meer variabelen is de nu volgende zogenaamde *impliciete-functiestelling.* Zij U een open deelverzameling van $\mathbf{R}^n \times \mathbf{R}^p$, $1 \le k \le \omega$, $f \in \mathrm{C}^k\,(U,\, \mathbf{R}^p)$, $\left(x^0,\, y^0\right) \in \Omega$, $f\left(x^0,\, y^0\right) = 0$ en zij tenslotte de $n \times n$-matrix

$$\frac{\partial f_i\left(x,\, y^0\right)}{\partial x_j}\Big|_{x=x^0}, \quad 1 \le i,\, j \le n,$$

niet-singulier. Dan is er een open omgeving X, resp. Y van x^0, resp. y^0 in $\mathbf{R}^n$, resp. $\mathbf{R}^p$ met de eigenschap dat er voor iedere $y \in Y$ precies één $x = \psi(y) \in X$ is, waarvoor $f(x,\, y) = 0$. Verder geldt voor de hierdoor gedefinieerde afbeelding ψ dat $\psi \in \mathrm{C}^k\,(Y,\, \mathbf{R}^n)$.

In deze stelling wordt de vergelijking $f(x,\, y) = 0$ opgevat als een stelsel van n reële vergelijkingen $f_i(x,\, y) = 0$ voor de n onbekenden x_j, $1 \le i,\, j \le n$, met p parameters y_l, $1 \le l \le p$. De stelling zegt dat als we voor speciale waarden ${y^0}_l$ van de parameters oplossingen ${x^0}_j$ hebben en de lineaire benadering in $x = x^0$ van de afbeelding $x \mapsto f\left(x,\, y^0\right)$ is inverteerbaar, dan is voor naburige waarden van de parameter het stelsel eenduidig oplosbaar, waarbij alleen gekeken wordt naar oplossingen in de buurt van x^0. (In het bijzonder is voor $y = y^0$ de oplossing $x = x^0$ een geïsoleerde oplossing: als $f\left(x,\, y^0\right) = 0$ en $x \ne x^0$, dan is $x \notin X$.) Bovendien hangt de oplossing glad af van de parameter, met dezelfde differentieerbaarheidsgraad k als f heeft.

Zij U en V open deelverzamelingen van $\mathbf{R}^n$ en $k \ge 1$. Men zegt dat een afbeelding $\Phi : U \to V$ een C^k-*diffeomorfisme* is van U naar V als Φ bijectief is van U naar V, $\Phi \in C^k\,(U,\, \mathbf{R}^n)$ en

$\Phi^{-1} \in C^k(V, \mathbf{R}^n)$. Een C^k-diffeomorfisme wordt ook wel een *reguliere coördinatentransformatie* of *reguliere substitutie van variabelen* genoemd, waarbij het woord 'regulier' nader gepreciseerd is tot 'C^k'. Men gebruikt de relatie $x = \Psi(y)$ vaak om de variabele x door de variabele y te vervangen, hierbij is $y = \Phi(x)$ de inverse substitutie. Problemen kunnen soms op spectaculaire manier doorzichtig gemaakt worden door gebruik te maken van een goed gekozen substitutie van variabelen. Om deze te vinden is het van belang om te begrijpen hoe eigenschappen van het systeem veranderen onder willekeurige substituties van variabelen; op dit thema zullen we later nog terugkomen.

Zijn Φ en Ψ differentieerbaar, dan is de samengestelde afbeelding $\Psi \circ \Phi$ differentieerbaar en er geldt de *kettingregel voor differentiatie*:

$$D(\Psi \circ \Phi)(a) = D\Psi(\Phi(a)) \circ D\phi(a).$$

Is $\Psi = \Phi^{-1}$, dan krijgen we hieruit dat voor iedere $X \in U$ geldt dat

$$I = DI(x) = D(\Psi \circ \Phi)(x) = D\Psi(\Phi(x)) \circ D\Phi(x).$$

Hieruit lezen we af dat de lineaire afbeelding $D\Phi(x)$ inverteerbaar is, met $D\Psi(\Phi(x))$ als inverse.

Toepassing van de impliciete-functiestelling met $f(x, y) = \Phi(x) - y$ geeft hiervan de volgende omkering, die bekend staat als de *lokale inverse-functiestelling*. Als U open is in $\mathbf{R}^n$, $1 \leq k \leq \omega$, $\Phi \in \mathrm{C}^k(U, \mathbf{R}^n)$, $x^0 \in U$ en $D\Phi\left(x^0\right)$ is inverteerbaar, dan is er een open omgeving X, resp. Y van x^0, resp. $y^0 := \Phi\left(x^0\right)$ in U, resp. $\mathbf{R}^n$, met de eigenschap dat de beperking $\Phi|_X$ van Φ tot X een C^k-diffeomorfisme is van X naar Y.

De *globale inverse-functiestelling* zegt dat als U open is in $\mathbf{R}^n$, $1 \leq k \leq \omega$ en $\Phi \in \mathrm{C}^k(U, \mathbf{R}^n)$, dan is Φ een C^k-diffeomorfisme van U naar een open deelverzameling V van $\mathbf{R}^n$, dan en slechts dan als Φ injectief is en, voor iedere $x \in U$, $D\phi(x)$ bijectief is.

Een klassiek voorbeeld is de substitutie van bolcoördinaten:

$$\Psi(r, \alpha, \theta) = (r \cos\alpha \cos\theta, r \sin\alpha \cos\theta, r \sin\alpha).$$

Er geldt dat $\det D\Psi(r, \alpha, \theta) = r^2 \cos\theta$, dus Ψ is een lokaal diffeomorfisme op een omgeving van (r, α, θ) dan en slechts dan als $r \neq 0$ en $\theta \neq \pm\pi/2$ modulo 2π. De verzameling $U =]0, \infty[\times]-\pi, \pi[\times]-\pi/2, \pi/2[$ is een maximale open verzameling waarop Ψ een diffeomorfisme is; de inperkingen zijn nodig om injectiviteit van $\Psi|_U$ te krijgen. Ψ is verre van injectief op $\mathbf{R}^3$.

De impliciete functiestelling speelt ook een sleutelrol in de beschrijving van deelverzamelingen van $\mathbf{R}^n$ die bepaald worden door vergelijkingen. In de formulering maken we gebruik van afbeeldingen g van een open deelverzameling van $\mathbf{R}^n$ naar $\mathbf{R}^c$ met $n > c$ en van afbeeldingen ψ van een open deelverzameling van $\mathbf{R}^d$ naar $\mathbf{R}^n$ met $d < n$. Nu kan een lineaire afbeelding van $\mathbf{R}^p$ naar $\mathbf{R}^q$ alleen maar bijectief zijn als $p = q$, dus in beide bovenstaande gevallen kan de afbeelding geen inverteerbare afgeleide hebben, dus kan de afbeelding geen diffeomorfisme zijn. Wel kan de afgeleide van g surjectief zijn en die van ψ injectief; dit blijken de effectieve regulariteits voorwaarden te zijn in het geval van verschillende dimensies van definitiegebied en ruimte waar naartoe afgebeeld wordt.

In de volgende discussie maken we gebruik van de volgende feiten uit de lineaire algebra. Zij A een lineaire afbeelding van $\mathbf{R}^p$ naar $\mathbf{R}^q$. Dan is de *rang* van A gedefinieerd als de dimensie van het beeld $A(\mathbf{R}^p)$ van A. ($A(\mathbf{R}^p)$ is een lineaire deelruimte van $\mathbf{R}^q$.) Deze dimensie is gelijk aan het maximale aantal lineair onafhankelijke vectoren onder de $A(e_i)$ waarin de e_i de standaard basisvectoren van $\mathbf{R}^n$ doorlopen. Omdat $A(e_i)$ de i-de kolom is van de matrix van A, is de rang van A gelijk aan het maximale aantal lineair onafhankelijke kolommen in de matrix van A; ook wel de *kolommenrang* van A genoemd. Anderzijds hebben we ook de dimensieformule

$$p = \dim A(\mathbf{R}^p) + \dim\ \ker A,$$

waarin $\ker A$ de *nulruimte* van A voorstelt, de verzameling der $v \in \mathbf{R}^p$ waarvoor $A(v) = 0$. Anders gezegd, $\operatorname{rang} A = p - \dim\ \ker A$. We hebben dat A injectief is, dan en slechts dan als $\ker A = 0$, dan en slechts dan als $\operatorname{rang} A = p$; dit kan uiteraard alleen als $p \leq q$. Anderzijds is A surjectief, dan en slechts dan als $\operatorname{rang} A = q$, dan en slechts dan als $\dim\ \ker A = p - q$; dit kan alleen als $p \geq q$. Tenslotte merken we nog op dat het beeld van A gelijk is aan het orthogonale complement van de nulruimte van de getransponeerde (= gespiegelde) $A^{\mathrm{t}} : \mathbf{R}^q \to \mathbf{R}^p$ van A. Dit geeft:

$$\operatorname{rang} A = p - \dim\ \ker A^{\mathrm{t}} = \operatorname{rang} A^{\mathrm{t}}.$$

Nu zijn de kolommen van A^{t} juist de rijen van A, dus de rang van A is ook gelijk aan het maximale aantal lineair onafhankelijke rijen in de matrix van A; ook wel de *rijenrang* van A genoemd.

Zij U open in $\mathbf{R}^n$, $1 \leq k \leq \omega$, $g \in \mathrm{C}^k(U, \mathbf{R}^c)$, $a \in U$. Men zegt dat g een *submersie* is in het punt a als de lineaire afbeelding $Dg(a) : \mathbf{R}^n \to \mathbf{R}^c$ surjectief is, dat wil zeggen, $\operatorname{rang} Dg(a) = c$. Dit kan alleen als $n \geq c$, hetgeen we in het vervolg veronderstellen. (Als $c = 1$, dan is de voorwaarde equivalent met $Dg(a) \neq 0$.) Men noemt g een *submersie* als voor iedere $a \in U$ geldt dat g een submersie is in het punt a.

De *submersiestelling* zegt dat g een submersie is in het punt a, dan en slechts dan als er een C^k-diffeomorfisme Ψ is van een open deelverzameling V van $\mathbf{R}^n$ naar een open omgeving U_0 van a in U, waarvoor $g \circ \Psi$ gelijk is aan de projectie

$$(y_1, \ldots, y_{n-c}, y_{n-c+1}, \ldots, y_n) \mapsto (y_{n-c+1}, \ldots, y_n) : V \to \mathbf{R}^c$$

op de laatste c coördinaten. (Als $c = n$ dan betekent dit dat g een lokaal diffeomorfisme is.) Voor het bewijs beschouwt men de vergelijking $g(x) = z$, waarbij c van de coördinaten van x, samengevat tot de vector $v \in \mathbf{R}^n$, als onbekenden genomen worden en $z \in \mathbf{R}^c$ en de overige $n - c$ coördinaten van x (weergegeven als $u \in \mathbf{R}^{n-c}$) als parameters worden opgevat. Men gaat na dat er minstens één keuze van de onbekenden is waarvoor de impliciete-functiestelling toegepast kan worden, hetgeen dan leidt tot een oplossing $v = \psi(u, z)$ met $\psi \in \mathrm{C}^k$. Het diffeomorfisme Ψ is nu de afbeelding $(u, z) \mapsto (u, \psi(u, z))$, gevolgd door de omnummering van de coördinaten die de laatste c coördinaten op de plek van de onbekenden brengt.

Is V open in $\mathbf{R}^d$, $1 \leq k \leq \omega$, $\psi \in \mathrm{C}^k\left(V, \mathbf{R}^d\right)$, $b \in V$, dan heet ψ een *immersie in het punt* b als $D\psi(b)$ injectief is. Dit is equivalent met $\ker D\psi(b) = 0$, ofwel $\operatorname{rang} D\psi(b) = d$. Dit kan alleen als $d \leq n$, hetgeen we in het vervolg veronderstellen. (Als $d = 1$ dan is ψ een kromme en is de voorwaarde equivalent met $\psi'(b) \neq 0$.) Men noemt ψ een *immersie* als voor iedere $b \in V$ geldt dat ψ een immersie is in het punt b.

De *immersiestelling* zegt dat ψ een immersie in het punt b is, dan en slechts dan als er een C^k-diffeomorfisme Φ is van een open omgeving U van $a = \psi(b)$ in $\mathbf{R}^n$ naar een open deelverzameling van $\mathbf{R}^n$, benevens een open omgeving V_0 van b in $V \cap \psi^{-1}(U)$, met de eigenschap dat

$$\Phi \circ \psi(y) = (y, 0), \quad y \in V_0.$$

Hierin is 0 de oorsprong in $\mathbf{R}^{n-d}$. (Als $d = n$ dan betekent dit dat ψ een lokaal diffeomorfisme is.) Dit kan bewezen worden door gebruik te maken van een injectieve lineaire afbeelding $C : \mathbf{R}^{n-d} \to$

$\mathbf{R}^n$ met de eigenschap dat $\mathbf{R}^n$ gelijk is aan de som van het beeld van $D\psi(b)$ en het beeld van C. Dit geeft dat de afbeelding

$$\Psi : (y,\, z) \mapsto \psi(y) + C(z)$$

in het punt $(b,\, 0)$ voldoet aan de voorwaarde voor de lokale inverse-functiestelling. Als $\Phi = \Psi^{-1}$ dan geeft $\Psi(y,\, 0) = \psi(y)$ dat $(y,\, 0) = \Phi(\psi(y))$.

Men noemt ψ een *inbedding* als ψ voldoet aan de volgende drie voorwaarden.

a) ψ is een immersie.

b) ψ is injectief.

c) De inverse ψ^{-1} van ψ is een continue afbeelding van de beeldverzameling $\psi(D)$ naar $\mathbf{R}^d$.

Merk op dat b) nodig is om c) te kunnen formuleren. Een voorbeeld van een afbeelding ψ die aan a) en b) voldoet wordt gegeven door $d = 1$, $n = 2$, $D = \,]-\infty,\, 1[$, $\psi(t) = \big(t^2 - 1,\, t^3 - t\big)$. Omdat $\psi(-1) = (0,\, 0)$ en $\psi(t) \to (0,\, 0)$ als $t \uparrow 1$, is ψ^{-1} niet continu in het punt $(0,\, 0)$. Het is instructief om een plaatje te maken van de kromme ψ in het vlak. Anderzijds kan opgemerkt worden dat als $d = n$, dan zegt de globale inverse-functiestelling dat a) en b) samen impliceren dat ψ een C^k-diffeomorfisme is, dus in dit geval volgt c) uit de combinatie van a) en b).

2.7 Deelvariëteiten en Raakruimten

Na deze wat uitgebreide voorbereidingen zijn we nu in staat om de volgende karakterisering van een deelverzameling V van $\mathbf{R}^n$ te formuleren. Zij $1 \leq k \leq \omega$ en $a \in V$. Dan zijn de volgende uitspraken i)—iv) equivalent.

i) Na een eventuele omnummering van de coördinaten in $\mathbf{R}^n$ is er een open omgeving U van a in $\mathbf{R}^n$, een open deelverzameling W van $\mathbf{R}^d$ en een C^k-afbeelding f van W naar $\mathbf{R}^{n-d}$, met de eigenschap dat $V \cap U$ gelijk is aan de grafiek van f, de verzameling der $(w,\, f(w))$ met $w \in W$.

ii) Er is een open deelverzameling D van $\mathbf{R}^d$, een C^k-inbedding ψ van D naar $\mathbf{R}^n$ en een open omgeving U van a in $\mathbf{R}^n$, met de eigenschap dat $V \cap U = \psi(D)$.

iii) Er is C^k-diffeomorfisme Φ, van een open omgeving U van a in $\mathbf{R}^n$ naar een open deelverzameling van $\mathbf{R}^n$, benevens een open deelverzameling Y van $\mathbf{R}^d$, waarvoor $\Phi(V \cap U) = Y \times \{0\}$. Hierin is 0 de oorsprong in $\mathbf{R}^{n-d}$.

iv) Er is een open omgeving U van a in $\mathbf{R}^n$ en een C^k-submersie g van U naar $\mathbf{R}^{n-d}$, waarvoor $V \cap U$ gelijk is aan de verzameling der $x \in U$ met $g(x) = 0$.

De implicatie i) $\Rightarrow$ ii) volgt door $\psi(w) = (w,\, f(w))$ te nemen; omdat ψ gevolgd door de projectie π op de eerste d coördinaten gelijk is aan de identiteit, zien we dat $\psi^{-1} = \pi$ op het beeld van ψ, dit impliceert de continuïteit van ψ^{-1}. De implicatie ii) $\Rightarrow$ iii) volgt uit de immersiestelling, terwijl iii) $\Rightarrow$ iv) volgt door g gelijk te nemen aan Φ gevolgd door de projectie op de laatste $n - d$ coördinaten. Tenslotte volgt iv) $\Rightarrow$ i) door een toepassing van de submersiestelling.

Men zegt nu dat V in het punt a een *d-dimensionale C^k-variëteit* is, als aan een van de equivalente karakteriseringen i)—iv) is voldaan. V heet een d-dimensionale C^k-variëteit, als voor iedere $a \in V$ geldt dat V een d-dimensionale C^k-variëteit in het punt a is. Merk op dat dit een *lokale* eigenschap van de verzameling V is: 'bij iedere $a \in V$ is er een open omgeving U van a in $\mathbf{R}^n$ met de eigenschap dat $V \cap U$ voldoet aan ...'. Ook kan opgemerkt worden dat de verzameling V^{reg} der $a \in V$ waar V een d-dimensionale C^k-deelvariëteit van $\mathbf{R}^n$ een open deelverzameling is van V en dat V^{reg} zèlf een d-dimensionale C^k-deelvariëteit van $\mathbf{R}^n$ is.

Het getal $c = n-d$ heet ook wel de *codimensie* van de variëteit. In termen van de karakterisering iv) is dit het "aantal onafhankelijke vergelijkingen" voor V, waarbij "onafhankelijk" slaat op de voorwaarde dat de rijen $Dg_i(a)$, $1 \leq i \leq c$, in de matrix van $Dg(a)$ lineair onafhankelijk dienen te zijn, wil g een submersie in het punt a zijn. Merk op dat de submersiestelling nog iets meer informatie geeft, namelijk dat als g een submersie in het punt a is, dan is er een open omgeving U, resp V van a, resp, $g(a)$, met de eigenschap dat voor iedere $y \in V$ de verzameling $g^{-1}(\{y\}) \cap U$ een C^k deelvariëteit is van dimensie $d = n - c$. In het algemeen noemt men voor een afbeelding $g : X \to Y$ en een punt $y \in Y$ de verzameling

$$g^{-1}(\{y\}) = \{\, x \in X \mid g(x) = y \,\}$$

de *vezel* van de afbeelding g over het punt y. Is g een C^k submersie in het punt a, dan zijn dus alle vezels van g in een open omgeving van a gelijk aan C^k deelvariëten van dezelfde dimensie d.

Een eenvoudig voorbeeld is de cirkel

$$C = \{x \in \mathbf{R}^2 \mid {x_1}^2 + {x_2}^2 = 1\}$$

in het vlak. De functie $g(x) = {x_1}^2 + {x_2}^2 - 1$ voldoet aan $Dg(x) \neq 0$ als $x \in C$, dus is een submersie, dit geeft dat C een ééndimensionale reëel-analytische deelvariëteit van $\mathbf{R}^2$ is. (Omdat g een veelterm is, is C zelfs wat in de algebraïsche meetkunde een *gladde algebraïsche variëteit* genoemd wordt, we gaan hier echter niet verder op in.) C is lokaal gelijk aan de grafiek van een functie, waarbij we de tweede variabele als analytische functie van de eerste variabele kunnen schrijven in de buurt van de punten $a \in C$ met $a \neq (\pm 1,\, 0)$ en de eerste variabele als functie van de tweede bij de $a \in C$ met $x \neq (0,\, \pm)$. De cirkel wordt overdekt door vier van dergelijke "grafiek-intervallen". De afbeelding $\psi : t \mapsto (\cos t,\, \sin t)$, die correspondeert met het parametriseren van de cirkel met behulp van de booglengte t vanaf het punt $(1,\, 0)$, is een immersie met C als beeld. Deze afbeelding is echter niet injectief Een maximaal open interval I waarvoor $\psi|_I$ injectief is, is bijvoorbeeld $I = \,]-\pi,\, \pi[$. Opgemerkt kan nog worden dat er geen enkele open deelverzameling W van $\mathbf{R}$ is en inbedding $\psi : W \to \mathbf{R}^2$ waarvoor $\psi(W) = C$. De reden is dat C compact is, de continuïteit van $\psi^{-1}|_C$ zou geven dat $W = \psi^{-1}(C)$ een compacte, dus gesloten deelverzameling is van $\mathbf{R}$. Omdat W ook open is en $\mathbf{R}$ samenhangend, zou de conclusie zijn dat $W = \mathbf{R}$. Echter, $\mathbf{R}$ is niet compact.

Het voorbeeld

$$A = \{x \in \mathbf{R}^2 \mid x_1\, x_2 = 0\}$$

van het assenkruis in het vlak is een één-dimensionale analytische variëteit in alle $x \in A$ met $x \neq (0,\, 0)$, terwijl A geen ééndimensionale C^1-variëteit is in het punt $(0,\, 0)$. Een reden is bijvoorbeeld dat daar beide basisvectoren raakvectoren zijn, hetgeen tot de tegenspraak zou leiden dat de raakruimte tweedimensionaal is. Interssant is dat voor iedere $c \in \mathbf{R}$ met $c \neq 0$ de niveauverzameling

$$H = \{x \in \mathbf{R}^2 \mid x_1\, x_2 = c\},$$

een hyperbool in het vlak, een ééndimensionale analytische deelvariëtei van $\mathbf{R}^2$ is die twee samenhangscomponenten heeft, welke de takken van de hyperbool genoemd worden. In het algemeen is op te merken dat iedere C^1-deelvariëteit van $\mathbf{R}^n$ lokaal boogsamenhangend is, dus uiteenvalt in (boog)samenhangscomponenten.

Zij V een d-dimensionale C^1-deelvariëteit van $\mathbf{R}^n$ en zij $a \in V$. Een vector $v \in \mathbf{R}^n$ heet een *raakvector aan* V *in het punt* a als $v = \gamma'(t_0)$ voor een differentieerbare kromme $\gamma : I \to \mathbf{R}^n$ is met $\gamma(I) \subset V$ en $\gamma(t_0) = a$. Hierin is I een open interval in $\mathbf{R}$ en $t_0 \in I$. De verzameling van alle raakvectoren aan V in het punt a wordt met $\mathrm{T}_a V$ aangeduid en heet de *raakruimte aan* V *in het punt* a.

Is ψ, resp g een inbedding, resp. submersie als in de karakterisering ii), resp. iv) van een variëteit, dan hebben we de volgende identiteiten:

$$D\psi(b)\left(\mathbf{R}^d\right) = \mathrm{T}_a V = \ker Dg(a), \quad \psi(b) = a.$$

Hieruit lezen we af dat $\mathrm{T}_a V$ een d-dimensionale *lineaire* deelruimte is van $\mathbf{R}^n$. Voor het bewijs merken we op dat voor iedere $w \in \mathbf{R}^d$ de kromme $\gamma : t \mapsto \psi(b + t\,w)$ differentieerbaar is, afbeeldt naar V en voldoet aan $\gamma(0) = a$, dus $D\psi(b)(w) = \gamma'(0) \in \mathrm{T}_a V$. Anderzijds, is $\gamma(I) \subset V$, dan is $g(\gamma(t)) = 0$ voor iedere $t \in I$, dus $Dg(a)\left(\gamma'(t_0)\right) = 0$. Op deze manier krijgen we inclusies

$$D\psi(b)\left(\mathbf{R}^d\right) \subset \mathrm{T}_a V \subset \ker Dg(a).$$

Omdat de eerste en derde lineaire deelruimte van $\mathbf{R}^n$ beiden dimensie d hebben, zijn ze aan elkaar gelijk, dus ook gelijk aan de tussenliggende $\mathrm{T}_a V$.

2.8 Stelsels van Gewone Differentiaalvergelijkingen

Zij Ω een deelverzameling van $\mathbf{R}^{n+1} = \mathbf{R} \times \mathbf{R}^n$ en $f : \Omega \to \mathbf{R}^n$. Een *oplossing van de differentiaalvergelijking*

$$\frac{dx}{dt} = f(t, x) \tag{2.22}$$

is een differentieerbare kromme γ in $\mathbf{R}^n$, een differentieerbare afbeelding van een open interval I in $\mathbf{R}$ naar $\mathbf{R}^n$, met de eigenschap dat voor iedere $t \in I$ geldt dat $(t, \gamma(t)) \in \Omega$ en

$$\frac{d\gamma(t)}{dt} = f(t, \gamma(t)).$$

Als bovendien $t_0 \in I$ en $\gamma(t_0) = x_0$, dan zegt men dat γ een oplossing is van het *beginwaardeprobleem*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0.$$

Differentiaalvergelijkingen waarbij maar naar één reële variabele gedifferentieerd wordt heten ook wel *gewone* differentiaalvergelijkingen, in contrast met *partiële differentiaalvergelijkingen*, welke vergelijkingen zijn waarin de partiële afgeleiden van de onbekende functie met betrekking tot meer dan één variabele voorkomen. Uitschrijven van de componenten van x, resp. f geeft dat de vergelijking $\frac{dx}{dt} = f(t, x)$ equivalent is met het *stelsel*

$$\frac{dx_i}{dt} = f_i\left(t, x_1, \ldots, x_n\right), \quad 1 \leq i \leq n,$$

van n gewone differentiaalvergelijkingen voor de onbekende functies $t \mapsto x_i(t)$ van de ene reële variabele t.

We formuleren nu een aantal algemene resultaten betreffende de bovenstaande stelsels van gewone differentiaalvergelijkinen, onder de aannamen dat Ω een open deelverzameling is van $\mathbf{R}\times\mathbf{R}^n$, dat voor iedere $t \in \mathbf{R}$ de afbeelding $x \mapsto f(t,\, x)$ van de klasse C^k is en dat alle partiële afgeleiden van f met betrekking tot de x-variabelen, van de orde ten hoogste k, continue functies in Ω definiëren. Hierin is $1 \leq k \leq \omega$.

De *lokale existentiestelling* zegt dat er bij iedere $(t_0,\, x_0) \in \Omega$ een open interval I om t_0 in $\mathbf{R}$ is en een oplossing $\gamma : I \to \mathbf{R}^n$ van het beginwaardeprobleem $\frac{dx}{dt} = f(t,\, x)$, $x\,(t_0) = x_0$.

De *lokale eenduidigheidsstelling* zegt dat als $\delta : J \to \mathbf{R}^n$ een oplossing is van hetzelfde beginwaardeprobleem, dan is er een open interval K om t_0 in $I \cap J$ met de eigenschap dat voor iedere $t \in K$ geldt dat $\delta(t) = \gamma(t)$. De lokale eenduidigheidsstelling impliceert de volgende *globale eenduidigheidsstelling.* Als $\gamma : I \to \mathbf{R}^n$ en $\delta : J \to \mathbf{R}^n$ oplossingen zijn en er is een $t_0 \in I \cap J$ waarvoor $\gamma\,(t_0) = \delta\,(t_0)$, dan geldt voor iedere $t \in I \cap J$ dat $\gamma(t) = \delta(t)$. Immers, zij K de verzameling der $t \in I \cap J$ waarvoor $\gamma(t) = \delta(t)$. K is gesloten in $I \cap J$ omdat γ en δ continu zijn. De lokale eenduidigheidsstelling levert dat K ook open in $I \cap J$. Omdat $I \cap J$ een interval is, is $I \cap J$ samenhangend, dus $K = \emptyset$ of $K = I \cap J$.

De globale eenduidigheidsstelling maakt het mogelijk om af te komen van de kunstmatige meerduidigheid van oplossingen, die ontstaat door het feit dat als $\gamma : I \to \mathbf{R}^n$ een oplossing van het beginwaardeprobleem is, dan geldt voor ieder open interval J met $t_0 \in J \subset I$ dat de beperking van γ tot J ook een oplossing is en strikt gesproken is dit een andere oplossing als $J \neq I$. Als $(t_0,\, x_0)$, dan schrijven we $I\,(t_0,\, x_0)$ voor de vereniging van alle definitie-intervallen I van oplossingen $\gamma : I \to \mathbf{R}^n$ van het beginwaardeprobleem $\frac{dx}{dt} = f(t,\, x)$, $x\,(t_0) = x_0$. Als $t \in I\,(t_0,\, x_0)$ dan is er een oplossing $\gamma : I \to \mathbf{R}^n$ met $t \in I$, definieer $\overline{\gamma}(t) = \gamma(t)$ in dit geval. Is $\delta : J \to \mathbf{R}^n$ een andere oplossing met $t \in J$, dan geeft de globale eenduidigheidsstelling dat $\gamma(t) = \delta(t)$, dus de definitie van de afbeelding $\overline{\gamma} : I(t_0,\, x_0) \to \mathbf{R}^n$ hangt niet af van de keuze van γ. Ook is het duidelijk dat $\overline{\gamma}$ een oplossing is van het beginwaardeprobleem $\frac{dx}{dt} = f(t,\, x)$, $x\,(t_0) = x_0$. $\overline{\gamma}$ is de *maximale oplossing* in de zin dat voor iedere oplossing $\gamma : I \to \mathbf{R}^n$ van hetzelfde beginwaardeprobleem geldt dat $I \subset I\,(t_0,\, x_0)$ en $\gamma = \overline{\gamma}|_I$. Wanneer we in het vervolg spreken over *de* oplossing van het beginwaardeprobleem $\frac{dx}{dt} = f(t,\, x)$, $x\,(t_0) = x_0$, dan zullen we steeds de maximale oplossing $\overline{\gamma}$ bedoelen en we schrijven

$$\overline{\gamma}(t) = \Gamma\,(t,\, t_0,\, x_0)\,, \quad t \in I\,(t_0,\, x_0)\,.$$

We beschrijven nu de algemene conclusies betreffende de afhankelijkheid van de oplossingen van de beginwaarden. Zij

$$D = \{(t,\, t_0,\, x_0) \in \mathbf{R} \times \mathbf{R} \times \mathbf{R}^n \mid t \in I\,(t_0,\, x_0)\}.$$

Dan is D een open deelverzameling van $\mathbf{R} \times \mathbf{R} \times \mathbf{R}^n$, met $(t_0,\, t_0,\, x_0) \in D$ als $(t_0,\, x_0) \in \Omega$. De afbeelding $\Gamma : D \to \mathbf{R}^n$ is continu. Maar er geldt meer: voor iedere $(t,\, t_0)$ is

$$D\,(t,\, t_0) = \{x_0 \in \mathbf{R}^n \mid t \in I\,(t_0,\, x_0)\}$$

een open deelverzameling van $\mathbf{R}^n$ en de afbeelding $\Phi^{(t,\, t_0)} : D\,(t,\, t_0) \to \mathbf{R}^n$, gedefinieerd door

$$\Phi^{t,\, t_0}\,(x_0) := \Gamma\,(t,\, t_0,\, x_0)$$

is van de klasse C^k, waarbij bovendien alle afgeleiden van de orde tot en met k op continue manier van t en t_0 afhangen. $\Phi^{t,\,t_0}$ is de afbeelding die, bij gegeven t_0 en t, de oplossing ten tijde t geeft als functie van de beginwaarde x_0 ten tijde t_0, men noemt $\Phi^{t,\,t_0}$ ook wel de *t_0-t-stroming die bepaald is door het snelheidsveld f.*

De eerste orde afgeleide

$$A(t) = D\left(\Phi^{t,\,t_0}\right)(x_0) \in \mathrm{L}\left(\mathbf{R}^n,\, \mathbf{R}^n\right)$$

voldoet aan het homogene lineaire beginwaardeprobleem

$$\frac{dA}{dt} = L(t) \circ A, \quad A\left(t_0\right) = I$$

in $\mathrm{L}\left(\mathbf{R}^n,\, \mathbf{R}^n\right)$, waarbij

$$L(t) = \frac{\partial f(t,\,x)}{\partial x}\Big|_{x=\Gamma(t,\,t_0,\,x_0)} \in \mathrm{L}\left(\mathbf{R}^n,\, \mathbf{R}^n\right).$$

Men noemt dit de *variatievergelijking met betrekking tot de beginwaarden.*

De eenduidigheid voor het beginwaardeprobleem impliceert dat de oplossing die ten tijde s gelijk is aan $\Phi^{s,\,r}\left(x_0\right)$ dezelfde is als de oplossing die ten tijde r gelijk is aan x_0. Dit leidt tot de *groepseigenschap* dat voor iedere $r,\, s,\, t \in \mathbf{R}$ geldt dat

$$\Phi^{t,\,s} \circ \Phi^{s,\,r} = \Phi^{t,\,r}|_{D(t,\,r)\cap D(s,\,r)}$$

De inperking tot $D(t,\,s) \cap D(s,\,r)$ is nodig want dit is het gemeenschappelijke definitiegebied van de afbeeldingen in het linker-en rechterlid. Als s tussen r en t ligt dan is deze inperking niet nodig, omdat in dat geval $D(t,\,r) \subset D(s,\,r)$.

Een interessant speciaal geval van de groepseigenschap treedt op als $r = t$. In dat geval krijgen we dat

$$\Phi^{t,\,s} \circ \Phi^{s,\,t} = \Phi^{t,\,t}|_{D(s,\,t)} = \text{de identiteit in } D(s,\,t).$$

Eenzelfde relatie geldt met verwisseling van s en t. De conclusie is dat *de stroming $\Phi^{t,\,s}$ een C^k-diffeomorfisme is van $D(t,\,s)$ naar $D(s,\,t)$, met inverse gelijk aan $\Phi^{s,\,t}$.*

Uit de differentiaalvergelijking

$$\frac{\partial \Gamma\left(t,\,t_0,\,x_0\right)}{\partial t} = f\left(t,\,\Gamma\left(t,\,t_0,\,x_0\right)\right)$$

lezen we af dat Γ ook partieel differentieerbaar is naar t, waarbij de afgeleide continu afhangt van $(t,\,t_0,\,x_0) \in D$. Als naast de reeds aangenomen differentieerbaarheidseigenschappen van f bovendien geldt dat $f \in C^{k-1}\left(\Omega,\, \mathbf{R}^n\right)$, dan is $\Gamma \in C^k\left(D,\, \mathbf{R}^n\right)$.

Men noemt het stelsel $\frac{dx}{dt} = f(t,\,x)$ *autonoom* als f niet van t afhangt, dat wil zeggen als er een open deelverzameling U van $\mathbf{R}^n$ is en een afbeelding $\widetilde{f} : U \to \mathbf{R}^n$, waarvoor $\Omega = \mathbf{R} \times U$ en $f(t,\,x) = \widetilde{f}(x)$ voor alle $t \in \mathbf{R}$ en $x \in U$. In dit geval treedt er de vereenvoudiging op dat $\Phi^{t,\,s} = \Phi^{t-s,\,0}$, waarbij $\Phi^{\tau,\,0}$ ook wel genoteerd wordt als Φ^τ en de *stroming na tijd τ* genoemd wordt. Het definitiegebied wordt in dit geval genoteerd met $D(\tau) := D(\tau,\,0)$. De groepseigenschap krijgt nu de gedaante

$$\Phi^b \circ \Phi^a = \Phi^{a+b}\Big|_{D(a+b)\cap D(a)}. \tag{2.23}$$

Verder is Φ^0 gelijk aan de identiteit in U en is Φ^t een C^k-diffeomorfisme van $D(t)$ naar $D(-t)$, met Φ^{-t} als inverse.

Voor autonome stelsels $\frac{dx}{dt} = f(x)$ is het een natuurlijke vraag of er *stationaire* oplossingen zijn, dat wil zeggen oplossingen γ waarvoor $x_0 = \gamma(t)$ niet van t afhangt. Hieruit volgt dat $f(x_0) = f(\gamma(t)) = \frac{d\gamma(t)}{dt} = 0$. Omgekeerd, is x_0 een nulpunt van het vectorveld f, dan geeft $\frac{d\gamma(t)}{dt} = 0 = f(x_0) = f(\gamma(t))$ dat de constante functie $\gamma(t) \equiv x_0$ een oplossing is van $\frac{dx}{dt} = f(x)$ en we zien dat $I(0,\, x_0) = \mathbf{R}$ en $\Phi^t(x_0) = x_0$ v oor iedere $t \in \mathbf{R}$. Dit laat zien dat de nulpunten van f precies corresponderen met de stationaire oplossingen, ofwel de vaste punten van de stroming. Voor de lineaire benadering $A(t) = D\Phi^t(x_0)$ van de stroming in het punt x_0 krijgen we, met de notatie $L = Df(x_0)$, het beginwaardeprobleem $\frac{dA}{dt} = L\, A$, $A(0) = I$, waarvan de oplossing gegeven wordt door

$$A(t) = \mathrm{e}^{t\, L}, \quad t \in \mathbf{R}.$$

Het gedrag hiervan, zoals bijvoorbeeld het asymptotische gedrag voor $t \to \infty$, wordt het beste begrepen door gebruik te maken van de eigenwaarde-ontbinding van L, waarvoor L als complex--lineaire transformatie in $\mathbf{C}^n$ opgevat moet worden als er niet-reële eigenwaarden zijn. De pointe is, dat als $v \in \mathbf{C}^n$, $\lambda \in \mathbf{C}$ en $L\, v = \lambda\, v$, dan is $A(t)\, v = \mathrm{e}^{\lambda\, t}\, v$.

De volgende stap is dat de stroming in een kleine omgeving van het stationaire punt geanalyseerd wordt als een storing van de lineaire stroming $\mathrm{e}^{t\, L}$. Een voorbeeld van een resultaat in deze richting is dat de volgende voorwaarden i)—iii) equivalent zijn.

i) Alle oplossingen van het lineaire stelsel $\frac{dv}{dt} = L\, v$ convergeren naar 0 als $t \to \infty$.

ii) Alle eigenwaarden van L hebben negatief reëel deel.

iii) Er zijn positieve constanten δ, ρ en C met de eigenschap dat als $\|\xi - x_0\| \leq \delta$ dan is $[0,\, \infty[\, \subset I(0,\, \xi)$ en voor iedere $t \geq 0$ geldt dat

$$\|\Phi^t(\xi) - x_0\| \leq C\, \|\xi - x_0\|\, \mathrm{e}^{-\rho\, t}.$$

Men noemt het stationaire punt x_0 *exponentieel asymptotisch stabiel* als aan één van deze (dus aan alle) voorwaarden is voldaan. De voorwaarde iii) is equivalent met een stelsel van algebraïsche ongelijkheden tussen de coëfficiënten van de karakteristieke veelterm van L; dit staat bekend als het *Routh-Hurwitz criterium.* Zie Routh [82], [83, Part II, pp. 221-231], Hurwitz [36] en Marden [62, pp. 134-141]

Nog even doorgaand op het thema van speciale oplossingen met een eenvoudig gedrag, komen we op niet-constante oplossingen $\gamma : I \to \mathbf{R}$ van een autonoom stelsel $\frac{dx}{dt} = f(x)$, waarvoor er een $\omega > 0$ is, waarvoor $0 \in I$, $\omega \in I$ en $\gamma(\omega) = \gamma(0)$. Zonder de algemeenheid te schaden mogen we ω minimaal nemen, dat wil zeggen dat voor iedere $0 < t < \omega$ geldt dat $\gamma(t) \neq \gamma(0)$. Er geldt in dit geval dat $I = \mathbf{R}$ en dat γ *periodiek* is met (minimale) *periode* ω, in de zin dat voor iedere $t \in \mathbf{R}$ geldt dat $\gamma(t + \omega) = \gamma(t)$. Merk op dat dit impliceert dat de baan $B = \gamma(\mathbf{R})$ van γ een compacte, samenhangende, ééndimensionale C^{k+1}-deelvariëteit van $\mathbf{R}^n$ is; B is diffeomorf met een cirkel. Verder is B invariant onder de stroming, dat wil zeggen: voor iedere $t \in \mathbf{R}$ geldt dat $\Phi^t(B) \subset B$. De $x = \gamma(t)$ voor een ω-periodieke oplossing zijn precies de *vaste punten voor* Φ^ω: de punten x in het definitiegebied van f waarvoor $\Phi^\omega(x) = x$. Het onderzoek van het lange-termijn gedrag van de oplossingen in de buurt van de periodieke oplossing γ start met een onderzoek van het gedrag van de machten van $D(\Phi^\omega)(x)$, met $x \in B$.

De complicatie met definitiegebieden hangt samen met het feit dat de maximale oplossingen niet altijd de hele reële as als definitiegebied hebben, bijvoorbeeld omdat ze "in een eindige tijd uit het definitiegebied weglopen". Voorbeeld: zij $m \neq 1$. Het beginwaardeprobleem

$$\frac{dx}{dt} = x^m, \quad x(0) = x_0 > 0,$$

gedefinieerd voor $t \in \mathbf{R}$ en $x \in]0, \infty[$, betreft een autonome diferentiaalvergelijking, waarvan de oplossing $x(t)$ gegeven wordt door

$$\frac{d}{dt}\frac{x(t)^{1-m}}{1-m} = x(t)^{-m}\,\frac{dx(t)}{dt} = 1,$$

dus

$$\frac{x(t)^{1-m}}{1-m} - \frac{{x_0}^{1-m}}{1-m} = t,$$

ofwel

$$x(t) = \left((1-m)\,t + {x_0}^{1-m}\right)^{1/(1-m)}.$$

Als $m > 1$, dan gaat $x(t)$ naar oneindig als $t \uparrow {x_0}^{1-m}/(m-1)$ en de conclusie is dat

$$\begin{aligned} I(0, x_0) &= \left]-\infty, \frac{{x_0}^{1-m}}{m-1}\right[, \\ D(t) &= \begin{cases}]0, \infty[& \text{als} \quad t \leq 0, \\ \left]0, ((m-1)\,t)^{\frac{1}{1-m}}\right[& \text{als} \quad t > 0. \end{cases} \end{aligned}$$

Anderzijds, als $m < 1$, dan convergeert $x(t)$ naar 0 als $t \downarrow -{x_0}^{1-m}/(1-m)$ en de conclusie is dat

$$\begin{aligned} I(0, x_0) &= \left]-\frac{{x_0}^{1-m}}{m-1}, \infty\right[, \\ D(t) &= \begin{cases}]0, \infty, [& \text{als} \quad t \geq 0, \\ \left]0, ((m-1)\,t)^{\frac{1}{1-m}}\right[& \text{als} \quad t < 0. \end{cases} \end{aligned}$$

Dus alleen in het geval $m = 1$, dat wil zeggen het geval van een lineaire vergelijking, zijn de oplossingen globaal gedefinieerd.

De *stelling van Peano* zegt dat continuïteit van f al voldoende is voor een lokale existentiestelling. Het bovenstaande voorbeeld voor $0 < m < 1$ laat echter zien dat continuïteit van f geen voldoende voorwaarde is voor de lokale eenduidigheid van de oplossingen van het beginwaardeprobleem. Immers in de vergelijking

$$\frac{dx}{dt} = |x|^m$$

is het rechterlid nu een continue functie van $x \in \mathbf{R}$. Voor iedere $s \in \mathbf{R}$ definieert

$$x(t) = \begin{cases} ((1-m)(t-s))^{\frac{1}{1-m}} & \text{als} \quad t > s, \\ 0 & \text{als} \quad t \leq s \end{cases}$$

een oplossing en we zien dat er een heel continuüm van verschillende oplossingen $x(t)$ is met $x(0) = 0$, namenlijk alle bovenstaande oplossingen met $s \geq 0$.

In het algemeen geldt nu dat het niet globaal gedefinieerd zijn van oplossingen *altijd* veroorzaakt wordt doordat de oplossing uit het definitiegebied wegloopt, in de volgende zin. Als $T := \sup I(t_0, x_0) < \infty$, dan geldt dat er voor iedere compacte deelverzameling K van Ω een $\epsilon > 0$ is, met de eigenschap dat voor iedere $t \in I(t_0, x_0)$ met $T - \epsilon < t < T$ geldt dat $(t, x(t)) \notin K$. Hierin hebben we geschreven $x(t) = \Gamma(t, t_0, x_0)$. Een soortegelijke uitspraak geldt voor $t \downarrow S$ als $S := \inf I(t_0, x_0) > -\infty$.

Een gevolg van deze stelling is dat als er een compacte deelverzameling K van Ω is en voor iedere $t \in I(t_0, x_0)$ geldt dat $(t, x(t)) \in K$, dan is $I(t_0, x_0) = \mathbf{R}$.

Voor de bewijzen van al deze feiten over stelsels van gewone differentiaalvergelijkingen verwijzen we naar Coddington and Levinson [13, 1.1-7] of naar [20, 2.1-3]. De eerste stap in alle bewijzen is de opmerking dat het beginwaardeprobleem $\frac{dx}{dt} = f(t, x)$, $x(t_0) = x_0$ equivalent is met de integraalvergelijking

$$\gamma(t) = x_0 + \int_{t_0}^{t} f(\tau, \gamma(\tau))\, d\tau.$$

Het grote voordeel hiervan is dat integralen gemakkelijker te schatten zijn dan afgeleiden. Verhelderend is in dit verband ook de opmerking van Robbin [80] dat de lokale theorie gezien kan worden als een toepassing op deze integraalvergelijking van de impliciete-functiestelling in de Banachruimte van continu krommen γ, voorzien van de supremum-norm.

In veel situaties hangt het snelheidsveld f ook nog af van van een aantal, zeg p reële *parameters*, die we samenvatten tot een vector $\epsilon \in \mathbf{R}^p$. Dat wil zeggen, we beschouwen in dit geval een stelsel van de vorm

$$\frac{dx}{dt} = f(t, x, \epsilon), \tag{2.24}$$

waarin nu $f : \Omega \to \mathbf{R}^n$ en $\Omega \subset \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^p$. De afhankelijkheid van de oplossingen van de parameters kan uit de voorgaande theorie over de afhankelijkheid van parameters gehaald worden door over te gaan op het $(n+p)$-dimensionale stelsel in de (x, ϵ)-ruimte, dat gegeven wordt door

$$\begin{aligned} \frac{dx}{dt} &= f(t, x, \epsilon), \\ \frac{d\epsilon}{dt} &= 0. \end{aligned}$$

De aannamen zijn nu dat Ω een open deelverzameling is van $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^p$, dat voor iedere $t \in \mathbf{R}$ de afbeelding $(x, \epsilon) \mapsto f(t, x, \epsilon)$ van de klasse C^k is en dat alle partiële afgeleiden van f met betrekking tot de (x, ϵ)-variabelen, van de orde ten hoogste k, continue functies in Ω definiëren. Hierin is $1 \leq k \leq \omega$. De conclusie is dat de oplossing $t \mapsto \Gamma(t, t_0, x_0, \epsilon)$ alle bovengenoemde eigenschappen heeft, met overal x_0 vervangen door (x_0, ϵ).

Het bovengenoemde stelsel voor (x, ϵ) onderscheidt zich van het algemene stelsel voor (x, ϵ) door de speciale gedaante $\frac{d\epsilon}{dt} = 0$ van de ϵ-component van het snelheidsveld. Dit is equivalent met de uitspraak dat ϵ een constante van beweging is voor de stroming. Dikwijls treden stelsels differentiaalvergelijkingen met parameters op als een limietgeval, voor $\phi \to 0$, van een stelsel waarbij ook de ϵ bewegen volgens een vergelijking van de vorm $\frac{d\epsilon}{dt} = \phi(t, x, \epsilon)$

Ook het optreden van niet-autonome stelsels heeft vaak een dergelijke achtergrond. Hierbij kan men denken aan een autonoom stelsel van de vorm

$$\frac{dx}{dt} = f(x, \epsilon),$$

$$\frac{d\epsilon}{dt} \quad = \quad \phi(\epsilon).$$

Het bijzondere hiervan is dat de beweging van ϵ niet beïnvloed wordt door de positie van x. Is $\epsilon = \epsilon(t)$ een niet-constante oplossing van $\frac{d\epsilon}{dt} = \phi(\epsilon)$, dan leidt substitutie hiervan in $\frac{dx}{dt} = f(x, \epsilon)$ tot de niet-autonome vergelijking

$$\frac{dx}{dt} = F(t, x), \quad F(t, x) := f(x, \epsilon(t)).$$

Omgekeerd kan een niet-autonoom stelsel $\frac{dx}{dt} = f(t, x)$ ook herschreven worden als een autonoom stelsel

$$\begin{aligned} \frac{dx}{dt} &= f(s, x), \\ \frac{ds}{dt} &= 1, \end{aligned}$$

in een ruimte van een één hogere dimensie. Het enige, vrij ondergeschikte punt is dat de voorwaarde dat de natuurlijke aanname hier zou zijn dat $f \in \mathrm{C}^k(\Omega, \mathbf{R}^n)$, hetgeen meer differentieerbaarheid veronderstelt met betrekking tot t dan nodig was in de bovenstaande theorie van niet-autonome stelsels. Met deze reductie naar autonome stelsels in het achterhoofd is het geen al te grote beperking van de algemeenheid als de discussie later af en toe alleen voor autonome stelsels gevoerd wordt.

2.9 Constanten van Beweging

Beschouw een autonoom stelsel $\frac{dx}{dt} = f(x)$ met $f \in \mathrm{C}^1(U, \mathbf{R}^n)$, U open in $\mathbf{R}^n$. Een functie $g : U \to \mathbf{R}$ heet een *constante van beweging* als voor iedere oplossing γ van $x' = f(x)$ geldt dat $g \circ \gamma : t \mapsto g(\gamma(t)) \in \mathbf{R}^p$ een constante functie is. Is g differentieerbaar, dan is $g \circ \gamma$ differentieerbaar, dus constant dan en slechts dan als

$$Dg(\gamma(t))\,(f(\gamma(t)) = Dg(\gamma(t))\,\big(\gamma'(t)\big) = \frac{d}{dt} g(\gamma(t)) \equiv 0.$$

Hieruit lezen we af dat een differentieerbare functie $g : U \to \mathbf{R}$ een constante van beweging is, dan en slechts dan als

$$\sum_{i=1}^{n} \frac{\partial g(x)}{\partial x_i}\, f_i(x) = 0, \quad x \in U.$$

(Voor het “slechts dan” gebruiken we de lokale existentiestelling voor het beginwaardeprobleem.) Dit is een lineaire partiële differentiaalvergelijking voor g, die men soms kan verifiëren voor gegeven f en g. In het algemeen worden partiële differentiaalvergelijkingen als lastiger beschouwd dan gewone differentiaalvergelijkingen en zegt men eerder dat de oplossingen van de partiële differentiaalvergelijking gegeven worden door de differentieerbare functies die constant zijn langs de oplossingsbanen.

Zij $g_j : U \to \mathbf{R}$, $1 \leq j \leq p$ constanten van beweging dan vatten we deze functies samen tot een afbeelding $g : U \to \mathbf{R}^p$. Voor iedere $c \in \mathbf{R}^p$ beschouwen we de *niveauverzameling*

$$U_c := \{x \in U \mid g(x) = c\},$$

de verzameling der $x \in U$ waar $g_j(x) = c_j$, $1 \leq j \leq p$. Het constant zijn van g langs de oplossingen betekent nu dat als $\gamma I \to U$ een oplossing is, $t_0 \in I$, $x_0 \in \gamma(t_0) \in U_c$, dan geldt voor iedere $t \in I$ dat $\gamma(t) \in U_c$; zelfs behoort $\gamma(t)$ tot de boogsamenhangscomponent van x_0 in U_c. In andere woorden, de verzameling U_c is *invariant onder de stroming* in de zin dat voor iedere $t \in \mathbf{R}$ geldt dat

$$\Phi^t(U_c \cap D(t)) \subset U_c.$$

Hierin kunnen we ook U_c vervangen door een boogsamenhangscomponent van U_c.

Neem nu aan dat $g \in \mathrm{C}^k(U, \mathbf{R}^p)$ en zij U^{reg} de verzameling der $x \in U$ waarvoor de $Dg_j(x)$, $1 \leq j \leq p$ lineair onafhankelijk zijn, d.w.z. waar g een submersie is. De verzameling U^{reg} is een open deelverzameling van U en voor iedere $c \in \mathbf{R}^p$ is $U_c^{\mathrm{reg}} := U^{\mathrm{reg}} \cap U_c$ een $(n-p)$-dimensionale C^k-deelvariëteit van $\mathbf{R}^n$ (en bovendien gesloten in U^{reg}). Omdat $g \circ \Phi^t = g$ is ook

$$Dg\left(\Phi^t(x)\right) \circ D\Phi^t(x) = Dg(x),$$

waaruit we aflezen dat $Dg\left(\Phi^t(x)\right)$ surjectief is, dan en slechts dan als $Dg(x)$ surjectief is.

Men zegt dat een deelverzameling V van een verzameling X *invariant* is onder de afbeelding $\Phi : X \to X$, als $\Phi(V) \subset V$. Het bovenstaande geeft daarmee dat U^{reg} invariant is onder Φ^t en daarmee zijn ook de samenhangscomponenten van de variëteiten U_c^{reg} invariant onder de stroming. Tenslotte is ook $U^{\mathrm{sing}} := U \setminus U^{\mathrm{reg}}$ invariant onder de stroming en daarmee ook iedere boogsamenhangscomponent van $U_c^{\mathrm{sing}} := U^{\mathrm{sing}} \cap U_c$. Dit is interessant omdat deze laatste verzamelingen vaak heel “klein” zijn, bestaande uit de uitzonderlijke punten waar de $Dg_j(x)$ lineair afhankelijk zijn.

Een iets algemenere vraag is onder welke voorwaarden een d-dimensionale C^k-deelvariëteit V van U invariant is onder de stroming met snelheidsveld f. Omdat voor iedere $a \in V$ de oplossing $\gamma : I \to \mathbf{R}^n$ van $\frac{dx}{dt} = f(x)$ met $x(0) = a$ voldoet aan $\gamma(t) \in V$ voor alle $t \in I$, volgt uit de definitie van de raakruimte dat noodzakelijkerwijze

$$f(a) = f(\gamma(0)) = \gamma'(0) \in \mathrm{T}_a V.$$

Men zegt dat het vectorveld f *raakt aan* V als voor iedere $a \in V$ geldt dat $f(a) \in \mathrm{T}_a V$; we hebben net geverifieerd dat als V invariant is onder de stroming met snelheidsveld f, dan raakt f aan V.

Ook de omkering hiervan is waar, als tenminste $k \geq 1$. We nemen nu dus aan dat f raakt aan V. Voor iedere $a \in V$ is er een open omgeving U_0 va a in $\mathbf{R}^n$, een open deelverzameling W van $\mathbf{R}^d$ en een C^k-inbedding $\psi : W \to \mathbf{R}^n$, waarvoor $V \cap U_0 = \psi(W)$. Voor iedere $y \in W$, $x = \psi(y)$, is $\mathrm{T}_x V = D\psi(y)\left(\mathbf{R}^d\right)$. Omdat $f(x) \in \mathrm{T}_x V$ en omdat $D\psi(y)$ injectief is, is er precies één $w = w(y) \in \mathbf{R}^d$, waarvoor

$$D\psi(y)\,(w(y)) = f(\psi(y)).$$

Omdat $D\psi \in \mathrm{C}^{k-1}$ krijgen we dat w een continu vectorveld in W definieert. (We kunnen w continu differentieerbaar krijgen door ψ van de vorm $y \mapsto \pi(y, h(y))$ te nemen, waarin π nog een geschikte permutatie van de variabelen voorstelt en $h \in \mathrm{C}^1$.) Zij b het punt in W waarvoor $\psi(b) = a$. Vanwege de existentiestelling van Peano is er een oplossing $\delta(t)$ van het beginwaardeprobleem $\frac{dy}{dt} = w(y)$, $y(0) = b$. Dan voldoet $\gamma(t) := \psi(\delta(t))$ aan

$$\frac{d\gamma(t)}{dt} = D\psi(\delta(t))\left(\delta'(t)\right) = D\psi(\delta(t))\,(w(\delta(t)) = f(\psi(\delta(t))) = f(\gamma(t)),$$

dus γ is een oplossing van het beginwaardeprobleem $\frac{dx}{dt} = f(x)$, $x(0) = a$. Gebruikmakend van de eenduidigheid van de oplossingen van het beginwaardeprobleem voor $\frac{dx}{dt} = f(x)$ concluderen we dat iedere oplossing daarvan die in V start van de vorm $\psi \circ \delta$ is voor een oplossing δ van $\frac{dy}{dt} = w(y)$, dus in het bijzonder in V blijft. Anders gezegd, V is invariant onder de stroming. Merk op dat $\delta \mapsto \psi \circ \delta$ een bijectieve relatie is tussen de oplossingen van $\frac{dy}{dt} = w(y)$ en de oplossingen van $\frac{dx}{dt} = f(x)$ in het open deel $\psi(W)$ van V. In deze zin is de beperking van het n-dimensionale stelsel $\frac{dx}{dt} = f(x)$ tot V op te vatten als een d-dimensionaal stelsel. Merk op dat V niet globaal (wel lokaal) geïdentificeerd hoeft te kunnen worden met een open deelverzameling van $\mathbf{R}^d$

Als voorbeeld geven we de volgende omkering van de karakterisering van de periodieke niet--constante oplossingen van een autonoom stelsel $\frac{dx}{dt} = f(x)$. Zij B een compacte ééndimensionale C^1-deelvariëteit van het definitiegebied U van f, invariant onder de stroming, dat wil zeggen, voor iedere $x \in B$ geldt dat $f(x) \in \mathrm{T}_x B$. We nemen verder aan dat B geen stationaire punten bevat, dus voor iedere $x \in B$ geldt dat $f(x) \neq 0$. Door op een samenhangscomponent over te gaan, mogen we verder aannemen dat B samenhangend is. Voor iedere oplossing γ die in B start krijgen we dat het definitiegebied gelijk is aan $\mathbf{R}$ en dat $\gamma(\mathbf{R})$ open is in B. Anderzijds geeft de groepseigenschap dat twee banen ofwel disjunct, ofwel aan elkaar gelijk zijn. Dit betekent dat B gelijk is aan de vereniging van disjuncte banen, iedere waarvan open in B is. Omdat B samenhangend is, is de conclusie dat voor iedere oplossing γ die in B start geldt dat $\gamma(\mathbf{R}) = B$. Anders gezegd, B is een baan. Als γ injectief (dus bijectief) zou zijn, dan geeft de compactheid van B dat $\gamma^{-1} : B \to \mathbf{R}$ continu is, in tegenspraak met het feit dat $\mathbf{R}$ niet compact is. De conclusie is dat γ niet injectief is, hetgeen op zijn beurt equivalent is met de uitspraak dat de oplossing γ periodiek is en B diffeomorf met een cirkel.

Sommige formuleringen van de bovenstaande resultaten maken bij eerste kennismaking misschien een wat zwaarwichtige of pedante (schoolmeesterachtige) indruk. In de klassieke literatuur wordt weinig of helemaal geen aandacht besteed aan vragen betreffende lokale eenduidigheid van de oplossingen of de differentieerbare afhankelijkheid van de oplossingen van de parameters, maximale definitiegebieden, en dergelijke. Soms krijgt men de indruk dat de resultaten als vanzelfsprekend worden beschouwd en wordt er daarom ook geen aandacht besteed aan de voorwaarden waaronder de resultaten geldig zijn, zoals het maximaal zijn van de rang van een afgeleidenmatrix of het lokaliseren = het beperken van de variabelen tot een geschikte omgeving van speciale waarden. Dit maakt dat de klassieke literatuur vaak veel lichtvoetiger lijkt dan de moderne; ik ben daar wel eens wat jaloers op.

2.10 Tweede-orde Stelsels

De bovenstaande algemeenheden uit de Analyse leiden op het eerste gezicht nog niet tot conclusies voor de oplossingen van stelsels van bewegingsvergelijkingen van de vorm

$$\frac{d^2x(t)}{dt^2} = a\left(t,\, x(t),\, \frac{dx(t)}{dt}\right), \tag{2.25}$$

zoals ingevoerd in (1.3). Immers (2.25) is een stelsel van tweede-orde differentiaalvergelijkingen is, terwijl de theorie gaat over eerste-orde vergelijkingen. Echter, door de snelheden $v(t) = x'(t)$ er als onbekende functies bij te nemen, zien we dat (2.25) equivalent is met het eerste-orde stelsel

$$\frac{dx}{dt} = v, \tag{2.26}$$

$$\frac{dv}{dt} = a\,(t,\, x,\, v)\,. \tag{2.27}$$

De equivalentie betekent dat de afbeelding die aan een tweemaal differentieerbare functie $t \mapsto x(t)$ de differentieerbare functie $t \mapsto \left(x(t),\, \frac{dx(t)}{dt}\right)$ toevoegt een bijectieve afbeelding definieert van de verzameling van alle oplossingen van (2.25) naar de verzameling van alle oplossingen van (2.26), (2.27), waarbij de inverse afbeelding aan de oplossing $t \mapsto (x(t),\, v(t))$ van (2.26), (2.27) de oplossing $t \mapsto x(t)$ van (2.25) toevoegt. Merk op dat de differentieerbaarheid van $t \mapsto x(t)$ en van $t \mapsto v(t)$, in combinatie met (2.26) automatisch geeft dat $t \mapsto x(t)$ tweemaal differentieerbaar is.

De equivalentie van (2.25) met het eerste orde stelsel (2.26), (2.27) leidt nu tot de volgende conclusies voor de oplossingen van (2.25). Deze worden verkregen door de conclusies voor (2.22) toe te passen met x, resp. f vervangen door $(x,\, v)$, resp. de afbeelding $(x,\, v) \mapsto (v,\, a(x,\, v))$.

Zij Ω een open deelverzameling van $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ en neem aan dat a een afbeelding is van Ω naar $\mathbf{R}^n$, met de volgende eigenschappen. Voor iedere $t \in \mathbf{R}$ is $(x,\, v) \mapsto a(t,\, x,\, v)$ een C^k-afbeelding waarvan alle partiële afgeleiden van de orde $\leq k$ continue functies van $(t,\, x,\, v) \in \Omega$ zijn. Hierin is $1 \leq k \leq \omega$. Dan is er bij iedere $(t_0,\, x_0,\, v_0) \in \Omega$ een oplossing γ van (2.25) met $\gamma\,(t_0) = x_0$ en $\gamma'\,(t_0) = v_0$, en deze oplossing is lokaal eenduidig bepaald. De lokale eenduidigheid impliceert dat er een unieke maximale oplossing is, waarvan het definitie-interval met $I\,(t_0,\, x_0,\, v_0)$ wordt aangeduid.

De verzameling D van alle $(t,\, t_0,\, x_0,\, v_0)$ waarvoor $t \in I\,(t_0,\, x_0,\, v_0)$ is een open deelverzameling van $\mathbf{R} \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$, die de verzameling der $(t_0,\, t_0,\, x_0,\, v_0)$ met $(t_0,\, x_0,\, v_0) \in \Omega$ bevat. We noteren $\Gamma\,(t_0,\, t_0,\, x_0,\, v_0) = (\gamma(t),\, \gamma'(t))$, waarbij γ de bovengenoemde oplossing is; de hierdoor gedefinieerde afbeelding $\Gamma : D \to \mathbf{R}^n \times \mathbf{R}^n$ is continu. We noteren $D\,(t,\, t_0)$ voor de verzameling der $(x_0,\, v_0)$ waarvoor $(t,\, t_0,\, x_0,\, v_0) \in D$; dit is een open deelverzameling van $\mathbf{R}^n \times \mathbf{R}^n$. De afbeelding

$$\Phi^{t,\, t_0} : (x_0,\, v_0) \mapsto \Gamma\,(t,\, t_0,\, x_0,\, v_0)$$

is een C^k-diffeomorfisme van $D\,(t,\, t_0)$ naar $D\,(t_0,\, t)$ als inverse en we hebben de groepseigenschap $\Phi^{c,\, b} \circ \Phi^{b,\, a} = \Phi^{c,\, a}$, voorzover de hierin optredende afbeeldingen zijn gedefinieerd. Men noemt $\Phi^{t,\, t_0}$ de *t_0-t-stroming in de $2n$-dimensionale faseruimte van posities en snelheden.*

In het begin van Paragraaf 1 werd gezegd dat het een ervaringsfeit is dat de bewegingen van een mechanisch systeem beschreven kunnen worden door te zeggen dat er bij iedere t_0, x_0, v_0 precies één differentieerbare kromme $t \mapsto x\,(t,\, t_0,\, x_0,\, v_0)$ is met

$$\begin{aligned} x\,(t_0,\, t_0,\, x_0,\, v_0) &= x_0 \quad \text{en} \\ \frac{\partial x\,(t,\, t_0,\, x_0,\, v_0)}{\partial t}\Big|_{t=t_0} &= v_0. \end{aligned}$$

Daarmee definieert

$$\Phi^{t,\, t_0}\,(x_0,\, v_0) = \left(x\,(t,\, t_0,\, x_0,\, v_0)\,,\ \frac{\partial x\,(t,\, t_0,\, x_0,\, v_0)}{\partial t}\right)$$

een afbeelding $\Phi^{t,\, t_0}$ die als t_0-t-stroming in de faseruimte opgevat kan worden.

De volgende stap was de opmerking dat, als de krommen tweemaal differentieerbaar zijn, dan voldoen ze aan en stelsel van de vorm (2.25). We kunnen nu concluderen dat als de daarin voorkomende versnellingsfunctie a voldoende glad is, dan geeft de (lokale) eenduidigheidsstelling dat omgekeerd het stelsel (2.25) de bewegingen al helemeaal vastlegt, zodat er geen informatie verloren gaat in de stap van de stroming naar het stelsel van differentiaalvergelijkingen (2.25). De lokale existentiestelling geeft daarbij dat de aanname dat de bewegingen voldoen aan (2.25) niet leidt tot een tegenspraak voor enige combinatie van beginpositie x_0 en beginsnelheid v_0. Als extra conclusie geeft de algemene theorie verder dat als $a \in \mathrm{C}^k$, dan zijn de stromingen C^k-diffeomorfismen.

Ook alle opmerkingen over constanten van bewegingen kunnen toegepast worden, mits we hiervoor functies van x en v nemen. Een differentieerbare functie $g(x,\, v)$ is een constante van beweging voor het autonome stelsel $x' = v$, $v' = a(x,\, v)$, dan en slechts dan als

$$\sum_{i=1}^{n} v_i \, \frac{\partial g(x,\, v)}{\partial x_i} + \sum_{j=1}^{n} a_j(x,\, v) \, \frac{\partial g(x,\, v)}{\partial v_j} \equiv 0. \tag{2.28}$$

2.11 Vraagstukken

Vraagstuk 2.1 Beschouw de beweging van een ééndimensionaal voorwerp met wrijving:

$$m\, x'' = -w(|x'|)\, x',$$

waarin de wrijvingscoëfficiënt gegeven wordt door een machtswet

$$w(v) = C\, v^{\alpha},$$

waarin C en α reële constanten zijn, $C > 0$. Bewijs dat als $\alpha < 0$, dan komt het voorwerp binnen een eindige tijd tot rust en geldt er geen eenduidigheid voor het beginwaardeprobleem. Merk op dat als $\alpha > -1$, dan is de versnellingsfunctie nog wel continu uit te breiden tot $x' = 0$. $\oslash$

Vraagstuk 2.2 Zij $g(x)$ een differentieerbare constante van beweging voor het autonome stelsel $x' = v$, $v' = a(x,\, v)$, die niet van v afhangt. Bewijs dat g constant is in iedere samenhangscomponent van het definitiegebied van het stelsel en daarom niet interessant is als constante van beweging. $\oslash$

Vraagstuk 2.3 Zij $c > 0$. Beschouw het eerste orde stelsel $x' = v$, $v' = a(x)$ in de open deelverzameling

$$V = \{(x,\, v) \in \mathbf{R}^6 \mid x \neq 0\},$$

van $\mathbf{R}^6$, dat behoort bij het tweede orde stelsel (1.21).

Zij $g : V \to \mathbf{R}^6$ de afbeelding gedefinieerd door

$$g(x,\, v) = \left(x \times v,\, \|x\|^{-1}\, x + c^{-1}(x \times v) \times v\right).$$

Voor iedere $(\mu,\, \epsilon) \in \mathbf{R}^6$, zij $V(\mu,\, \epsilon)$ de verzameling der $(x,\, v) \in V$ waarvoor $g(x,\, v) = (\mu,\, \epsilon)$, de niveauverzameling van de afbeelding g voor het niveau $(\mu,\, \epsilon)$. Zij W de verzameling der $(\mu,\, \epsilon) \in \mathbf{R}^6$, waarvoor $\langle \mu,\, \epsilon \rangle = 0$. Bewijs:

a) $g(V) \subset W$.

b) Zij $(\mu,\, \epsilon) \in W$ en $\mu \neq 0$. Noteer $\gamma := -\|\mu\|^2/c$ en $\mu^{\perp}$ voor het orthogonale complement van μ. Als $(x,\, v) \in V(\mu,\, \epsilon)$ dan is $x \in \mu^{\perp}$ en voldoet aan (1.14). Verder is $v \in \mu^{\perp}$ en $\|x\|^{-1}\, x + c^{-1}\mu \times v = \epsilon$. Omgekeerd is er bij iedere $x \in \mu^{\perp}$ die aan (1.14) voldoet precies één $v \in \mu^{\perp}$ waarvoor $\|x\|^{-1}\, x + c^{-1}\mu \times v = \epsilon$ en voor deze x en v geldt dat $(x,\, v) \in V(\mu,\, \epsilon)$. (Hint: gebruik dat $\epsilon \in \mu^{\perp}$ en dat $v \mapsto \mu \times v$ een bijectieve lineaire afbeelding definieert van $\mu^{\perp}$ naar $\mu^{\perp}$.) Concludeer dat $V(\mu,\, \epsilon)$ gelijk is aan de baan van een oplossing van het eerste orde stelsel in U dat behoort bij het tweede orde stelsel (1.21). Wanneer is $V(\mu,\, \epsilon)$ compact en de bijbehorende oplossing periodiek?

c) Als $V(0,\,\epsilon) \neq \emptyset$, dan is $\|\epsilon\| = 1$. Omgekeerd, is $\|\epsilon\| = 1$, dan bestaat $V(0,\epsilon)$ uit de verzameling van alle $(r\,\xi,\, s\,\xi)$ met $r \in\,]0,\, \infty[$ en $s \in \mathbf{R}$. Beschrijf $V(0,\,\epsilon)$ als een vereniging van banen, van elkaar onderscheiden door de energie van de oplossing.

d) $(\mu,\,\epsilon) \in g(V)$ dan en slechts dan als ($\mu \neq 0$ en $\langle \mu,\, \epsilon\rangle = 0$) of ($\mu = 0$ en $\|\epsilon\| = 1$).

Zie Vraagstuk 5.1 voor een differentiaalmeetkundige beschrijving van de afbeelding g. ⊘

Vraagstuk 2.4 De hieronder volgende *methode van variatie van constanten* is afkomstig van Lagrange [52, t. I, Seconde Partie, Sect. V].

Zij Ω open in $\mathbf{R} \times \mathbf{R}^n$ en $f : \Omega \to \mathbf{R}^n$. Neem verder aan dat I een begrensd interval is in $\mathbf{R}$, U open in $\mathbf{R}^n$, $\Phi \in \mathrm{C}^1\,(I \to U,\, \mathbf{R}^n)$ en dat de afbeelding Φ de volgende eigenschappen heeft:

i) Voor iedere $u \in U$ is $\gamma_u : t \mapsto \Phi(t,\, u)$ een oplossing van $\frac{dx}{dt} = f(t,\, x)$.

ii) Voor iedere $t \in I$ is $\Phi^t : u \mapsto \Phi(t,\, u)$ een C^1-diffeomorfisme van U naar een open deelverzameling X_t van $\mathbf{R}^n$.

iii) Er is een open X deelverzameling van $\mathbf{R}^n$, met de eigenschap dat voor iedere $t \in I$ geldt dat

$$I \times X \subset I \times X_t \subset \Omega.$$

Lagrange merkt op dat men voor Φ^t de t_0-t-stroming met snelheidsveld f mag nemen (deze heeft de gewenste differentieerbaarheidseigenschappen als $f \in \mathrm{C}^1$), maar dat dit niet persé hoeft.

Zij $\widetilde{\gamma} : I \to X$ een oplossing van $\frac{dx}{dt} = f(t,\, x) + s(t,\, x)$, waarin $s(t,\, x)$ een "stoorterm" is. Bewijs:

a) Voor iedere $t \in I$ is er precies één $u = \delta(t) \in U$ met $\widetilde{\gamma}(t) = \Phi(t,\, u)$. Dit definieert een differentieerbare kromme $\delta : I \to U$.

b) $\delta(t)$ voldoet aan de differentiaalvergelijking

$$\frac{d\delta(t)}{dt} = D\Phi^t(\delta(t))^{-1}\, s\,(t,\, \widetilde{\gamma}(t))\,.$$

c) Als $s = 0$ dan is δ constant en er geldt een (lokale) eenduidigheidsstelling voor het beginwaardeprobleem $\frac{dx}{dt} = f(t,\, x)$, $x\,(t_0) = x_0$; althans voor $x_0 \in X$.

d) Neem nu aan dat $\Phi \in \mathrm{C}^2$ met begrensde tweede orde afgeleiden, dat $s \in \mathrm{C}^1$ en dat s en de eerste orde afgeleiden van s van de orde ϵ zijn. Schrijf $\delta\,(t_0) = u_0$. Dan geldt voor alle $t \in I$ dat

$$\delta(t) = u_0 + \int_{t_0}^{t} D\Phi^\tau\,(u_0)^{-1}\; s\,(\tau,\, \Phi^\tau\,(u_0))\; d\tau + \mathcal{O}\left(\epsilon^2\right)$$

en daarmee

$$\widetilde{\gamma}(t) = \Phi^t\,(u_0) + \int_{t_0}^{t} D\Phi^t\,(u_0) \circ D\Phi^\tau\,(u_0)^{-1}\; s\,(\tau,\, \Phi^\tau\,(u_0))\; d\tau + \mathcal{O}\left(\epsilon^2\right).$$

$\oslash$

Vraagstuk 2.5 Beschouw, voor gegeven $a, b \in \mathbf{R}$, de verzameling

$$K_{a,b} := \{\, (x,\, y) \in \mathbf{R}^2 \mid y^2 + x^3 + a\,x + b = 0 \,\}$$

in het vlak. Bepaal de stationaire punten van $f : x \mapsto x^3 + a\,x + b$ en bepaal de waarden van f in deze punten. Bewijs de volgende uitspraken.

a) $K_{a,b}$ is een reëel-analytische ééndimensionale deelvariëteit van $\mathbf{R}^2$, tenzij $a \leq 0$ en

$$\pm\frac{2}{3}\,a\sqrt{-\frac{a}{3}} + b = 0,$$

in welk geval $K_{a,b}$ alleen niet reëel-analytisch is in het punt $x = \pm\sqrt{-a/3}$, $y = 0$.

b) $K_{a,b}$ is samenhangend, tenzij $a < 0$ en

$$4a^3 + 27b^2 \leq 0.$$

Is dit laatste het geval, dan heeft $K_{a,b}$ twee samenhangscomponenten, een onbegrensde in het halfvlak $x \leq 0$ en een compacte in het halfvlak $x \geq 0$. De compacte samenhangscomponent is diffeomorf met een cirkel of met een punt.

Maak een schets van $K_{a,b}$ in de diverse gevallen. $\oslash$

3 Energie

3.1 Één Vrijheidsgraad

Ter inleiding beginnen we met een gedetailleerde analyse van een mechanisch systeem met maar één vrijheidsgraad en een kracht die alleen van de positie afhangt:

$$m\,x'' = \phi(x). \tag{3.1}$$

We nemen slechts aan dat de functie ϕ continu is en gedefinieerd op een open interval I in $\mathbf{R}$. De faseruimte is hier gelijk aan de open deelverzameling $U = I \times \mathbf{R}$ van $\mathbf{R}^2$.

Zij $V : I \to \mathbf{R}$ een primitieve van $-\phi$, dat wil zeggen V is differentieerbaar en voor iedere $x \in U$ geldt dat $\frac{dV(x)}{dx} = -\phi(x)$. De sleutelopmerking is nu dat, met de notatie $x' = v$, de functie

$$H(x,\,v) := \frac{1}{2}\,m\,v^2 + V(x) \tag{3.2}$$

van positie en snelheid een constante van beweging is. Immers,

$$v\,\frac{\partial H(x,\,v)}{\partial x} + \frac{\phi(x)}{m}\,\frac{\partial H(x,\,v)}{\partial v} = v\,\frac{dV(x)}{dx} + \frac{\phi(x)}{m}\,m\,v = v\,\left[-\phi(x) + \phi(x)\right] = 0.$$

Er geldt dat $DH(x,\,v) = \left(\frac{dV(x)}{dx},\,m\,v\right)$, dus $(x,\,v) \in U^{\mathrm{sing}}$ dan en slechts dan als $v = 0$ en $\frac{dV(x)}{dx} = 0$. De laatste vergelijking is equivalent met $\phi(x) = 0$, dus U^{sing} is gelijk aan de verzameling der stationaire punten van het stelsel. We noteren U_c voor de niveauverzameling van H voor het niveau $c \in \mathbf{R}$. Als $1 \leq k \leq \omega$ en $\phi \in \mathrm{C}^{k-1}(I,\,\mathbf{R})$, dan is $H \in \mathrm{C}^k(U,\,\mathbf{R})$ en is $U_c \cap U^{\mathrm{reg}}$ een ééndimensionale C^k-deelvariëteit van U en gesloten in U^{reg}.

Als $v \neq 0$, dan kunnen we, voor gegeven waarde van $c \in \mathbf{R}$, de vergelijking $H(x,\,v) = c$ lokaal eenduidig naar v oplossen, met x en c als parameters. Omdat $H(x,\,v)$ op een kwadratische manier van v afhangt, kunnen we dit hier expliciet doen, met als resultaat dat

$$v = \pm\left(\frac{2}{m}\,(c - V(x))\right)^{1/2}. \tag{3.3}$$

Uit $v \neq 0$ en $H(x,\,v) = c$ volgt dat $V(x) < c$. Nu is

$$I_{<c} := \{x \in I \mid V(x) < c\}$$

een open deelverzameling van $\mathbf{R}$, die uiteenvalt in samenhangscomponenten, ieder waarvan een open interval J in $\mathbf{R}$ is. (Door te kijken naar de lengten van deze intervallen kan men ook nog inzien dat de samenhangscomponenten van iedere open deelverzameling van $\mathbf{R}$ een aftelbare collectie van intervallen vormen.)

Is $\xi : T \to I$ een oplossing, dan is de verzameling der $t \in T$ waarvoor $x = \xi(t) \in J$ een open deelverzameling van T, zij $\widetilde{T}$ een der samenhangscomponenten hiervan. Er geldt voor iedere $t \in \widetilde{T}$ dat $\xi'(t) \neq 0$; omdat γ' continu is in het interval $\widetilde{T}$ krijgen we dat ófwel voor iedere $t \in \widetilde{T}$ geldt dat $\xi'(t) > 0$, ófwel voor iedere $t \in \widetilde{T}$ geldt dat $\xi'(t) < 0$. Omdat $\widetilde{\xi}(t) := \xi(-t)$ een oplossing is als ξ een oplossing is, waarbij de snelheid $\widetilde{\xi}'(t) = -\xi'(-t)$ tegengesteld is, kunnen we ons in eerste instantie beperken tot het geval dat we in (3.3) de positieve wortel gekozen hebben.

De functie ξ is continu en strikt monotoon stijgend, dus een bijectieve afbeelding van $\widetilde{T}$ naar een open interval $\widetilde{J}$ in J, met een continue inverse $\tau : \widetilde{J} \to \widetilde{T}$; dit is *de tijd als functie van de positie.* Omdat de afgeleide steeds ongelijk aan nul is, geeft de inverse-functiestelling (voor functies van één variabele, die is eenvoudiger te bewijzen dan de meerdimensionale) dat τ differentieerbaar is en dat $\frac{d\tau(x)}{dx} = 1/\frac{d\xi(t)}{dt}$ als $x = \xi(t)$, resp. $t = \tau(x)$. Hieruit lezen we meteen af dat

$$\frac{d\tau}{dx} = \left(\frac{2}{m}\,(c - V(x))\right)^{-1/2}.$$

Op zijn beurt is dit equivalent met

$$\tau(x) - t_0 = \int_{x_0}^{x} \left(\frac{2}{m}\,(c - V(\xi))\right)^{-1/2} d\xi, \quad \xi\,(t_0) = x_0; \tag{3.4}$$

een expliciete integraalformule voor de tijd die nodig is om van x_0 naar x te komen. Omgekeerd levert (3.4), voor iedere $x_0 \in J$ en $t_0 \in \mathbf{R}$ en na keuze van het teken een differentieerbare functie $\tau : J \to \mathbf{R}$ op waarvan de afgeleide steeds strikt positief is en waarvan de inverse ξ een oplossing van (3.1) is, gedefinieerd in een interval $\widetilde{T}$ met $\xi(\widetilde{T}) = J$. Hieruit lezen we af dat er een eenduidigheidsstelling voor het beginwaardeprobleem geldt in het gebied $v > 0$, dus daar kunnen we praten over *de* oplossingen met maximaal definitie-interval. We krijgen daarmee dat in het bovenstaande geldt dat $\widetilde{J} = J$.

Definieer $s = \sup J$, waarbij we $s = \infty$ noteren als J niet van boven begrensd is. Er zijn dan de gevallen $s = \infty$, $s \in \mathbf{R} \setminus I$ en $s \in I$. In ieder van deze gevallen treedt weer eeen onderverdeling op die bepaald wordt door het gedrag van de integraal in het rechterlid van (3.4) als $\tau(x) \uparrow s$. Het loont de moeite om in een apart lijstje het gedrag van de oplossingen in ieder van de gevallen op te schrijven. Niet-eenduidigheid van oplossingen van het beginwaardeprobleem treedt alleen op in het geval iiibII) en 'in een eindige tijd weglopen uit het definitiegebied' gebeurt in de gevallen ib) en iib).

i) $s = \infty$. Er zijn dan twee mogelijkheden.

   ia) Als de integraal in het rechterlid van (3.4) *divergeert* voor $x \to \infty$, dan is $\widetilde{T}$ naar boven onbegrensd; de oplossing bestaat voor willekeurig grote tijd en $\xi(t) \uparrow \infty$ als $t \to \infty$.

   ib) Als de integraal in het rechterlid van (3.4) *convergeert* voor $x \to \infty$, naar $\sigma \in \mathbf{R}$, dan is $\sigma = \sup \widetilde{T}$ en er geldt dat $\xi(t) \to \infty$ als $t \uparrow \sigma$. Dit is een drastisch geval van 'uit het definitiegebied van het stelsel weglopen in een eindige tijd'.

ii) Als $s \in \mathbf{R}$ en $s \notin I$, dan is $\sup J = s = \sup I$ en er zijn weer twee mogelijkheden.

   iia) Als de integraal in het rechterlid van (3.4) *divergeert* voor $x \uparrow s$, dan is $\widetilde{T}$, dus ook T, naar boven onbegrensd; de oplossing $\xi(t)$ bestaat voor willekeurig grote t en convergeert monotoon stijgend naar s als $t \to \infty$.

   iib) Als de integraal in het rechterlid van (3.4) *convergeert* voor $x \to \infty$, naar $\sigma \in \mathbf{R}$, dan is $\sigma = \sup \widetilde{T} = \sup T$ en er geldt dat $\xi(t) \uparrow s \notin I$ als $t \uparrow \sigma$. Dit is ook een geval van 'uit het definitiegebied van het stelsel weglopen in een eindige tijd', alleen nu met convergentie naar een eindige waarde.

iii) $s \in I$. In dit geval is $V(s) = c$; we onderscheiden nu de volgende gevallen.

iiia) $V'(s) \neq 0$. Dit betekent precies dat $(s,\,0) \in U^{\mathrm{reg}}$, ofwel $(s,\,0)$ is geen stationair punt van het eerste orde stelsel. Merk op dat $x \in J$ impliceert dat $x < s$ en $V(x) < c$, waaruit we aflezen dat $V'(s) > 0$. We hebben hier een punt $(x,\,v)$ waar $\frac{\partial H(x,v)}{\partial v} = 0$ maar $\frac{\partial H(x,v)}{\partial x} \neq 0$. In de buurt hiervan is U_c niet te beschrijven als 'v is een differentieerbare functie van x', maar wel als 'x is een differentieerbare functie van v', in formule:

$$\begin{aligned} x &= V^{-1}\left(c - \frac{1}{2}\,m\,v^2\right) \\ m\,\frac{dv(t)}{dt} &= \phi \circ V^{-1}\left(c - \frac{1}{2}\,m\,v(t)^2\right). \end{aligned}$$

Als $t \uparrow \sigma$ dan krijgen we dat $x(t) \uparrow s$, dus $V(x(t)) \uparrow V(s) = c$, waaruit volgt dat $v(t) \to 0$ omdat $\frac{1}{2}\,m\,v(t)^2 = c - V(x(t))$. Op zijn beurt krijgen we nu ook dat

$$\frac{dv(t)}{dt} \to \frac{1}{m}\,\phi \circ V^{-1}(c) = \frac{\phi(s)}{m}.$$

Dit impliceert dat als we definiëren $x\,(\sigma) = s$, $v\,(\sigma) = 0$, dan is de functie $t \mapsto (x(t),\,v(t))$ ook differentieerbaar in $t = \sigma$ en voldoet daar aan het stelsel. Door middel van de definities

$$\begin{aligned} x\,(\sigma + h) &= x\,(\sigma - h)\,, \\ v\,(\sigma + h) &= -v\,(\sigma - h) \end{aligned}$$

kunnen we deze oplossing uitbreiden tot een oplossing die gedefinieerd is in een symmetrisch open interval om σ dat T bevat. Daarbij 'keert de beweging om' als $t = \sigma$, $x = s$. Dit is ook de enige oplossing die als voortzetting in aanmerking komt. Immers omdat $v'\,(\sigma) = \phi(s)/m < 0$, en $v\,(\sigma) = 0$ is $v'(t) < 0$ voor $t > \sigma$, dus daar geldt de bovenstaande beschrijving van de oplossing, maar nu startend met het minteken in (3.3).

Opmerking: iedere oplossing van de tweede-orde vergelijking (3.1) is oplossing van de eerste-orde vergelijking (3.3) met $v = \frac{dx}{dt}$, maar de omkering geldt in dit geval *niet*. Immers, de voortzetting met $x(t) = s$ voor alle $t \geq \sigma$ is een oplossing van (3.3), maar niet van (3.1).

iiib) $V'(s) = 0$, ofwel $(s,\,0)$ is een stationair punt van het eerste orde stelsel. We kunnen in dit geval weer twee gevallen onderscheiden.

iiibI) De integraal in het rechterlid van (3.4) *divergeert* als $x \uparrow s$. Dit correspondeert met de situatie dat $\widetilde{T}$ niet naar boven begrensd is en dat $x(t) \uparrow s$ als $t \to \infty$.
Dit is bijvoorbeeld het geval als $\phi \in \mathrm{C}^1$, dus $V \in \mathrm{C}^2$. Immers, er is dan een constante $C > 0$ met de eigenschap dat voor alle x in een omgeving van s geldt dat $c - V(x) \leq C\,(x - s)^2$, dus $(c - V(x))^{-1/2} \geq C^{-1/2}\,|x - s|^{-1}$, waaruit de divergentie volgt.

iiibII) De integraal in het rechterlid van (3.4) *convergeert* als $x \uparrow s$, met een limiet gelijk aan $\sigma - t_0$. In dit geval krijgen we voor $t \uparrow \sigma$ dat $x(t) \uparrow s$, $x'(t) \to 0$ en $x''(t) \to 0$. Er geldt geen eenduidigheid van oplossingen van het beginwaardeprobleem voor (3.3),

omdat we de oplossing kunnen voortzetten met $x(t) = s$ voor alle $t \geq \sigma$, terwijl we echter ook op iedere gewenst moment $t_1 \geq \sigma$ verder kunnen gaan met de omgekeerde beweging zoals die optreedt in het geval iiia).
Deze niet-eenduidigheid treedt bijvoorbeeld op als er constanten $\delta > 0$ $C > 0$, $0 < m < 1$ zijn met de eigenschap dat voor iedere $x \in [s - \delta,\, s]$ geldt dat $\phi(x) \leq -C\,(s - x)^m$. Dit impliceert dat voor iedere $x \in [s - \delta,\, s]$ geldt dat $c - V(x) \geq \frac{C}{m+1}\,(s - x)^{m+1}$, dus

$$(c - V(x))^{-1/2} \leq \left(\frac{m+1}{C}\right)^{1/2} (s - x)^{-(m+1)/2}.$$

Omdat $\frac{m+1}{2} < 1$, concluderen we hieruit dat de integraal in het rechterlid van (3.4) convergeert als $x \uparrow s$.

Soortgelijke conclusies kunnen getrokken worden als we beginnen met de oplossingen met $v < 0$, dat wil zeggen met de keuze van het minteken in (3.3). We laten de formulering van alle gevallen analoog aan de bovenstaande ia) tot en met iiibII) over aan de lezer.

Een belangrijk speciaal geval treedt hierbij op als er $a,\, b \in I$ zijn met $V(x) < c$ als $a < x < b$, $V(a) = c = V(b)$ en $V'(a) \neq 0$ en $V'(b) \neq 0$; dit impliceert dan dat $V'(a) < 0$ en $V'(b) > 0$. Dit betekent dat de oplossingen met $H(x,\, v) = c$ die starten in $[a,\, b]$ steeds *heen en weer blijven slingeren* tussen a en b, waarbij de beweging van b naar a de omgekeerde is van die van a naar b. Deze bewegingen zijn periodiek, met periode gelijk aan

$$\omega = \omega(c) := 2 \int_a^b \left(\frac{2}{m}\,(c - V(x))\right)^{-1/2} dx, \tag{3.5}$$

zoals we direct uit (3.4) aflezen. Deze slingerbeweging correspondeert precies met het geval dat we een compacte samenhangscomponent C van U_c^{reg} hebben, in welk geval we op algemene gronden ook al wisten dat C de baan is van een periodieke oplossing.

Een eenvoudig voorbeeld is

$$\phi(x) = -k\,x,$$

de *wet van Hooke*. Hiervoor kunnen we nemen $V(x) = \frac{1}{2}\,k\,x^2$, dus $H(x,\, v) = \frac{1}{2}\,m\,v^2 + \frac{1}{2}\,k\,x^2$ is een constante van beweging. Voor iedere $c > 0$ is U_c een ellips in het fasevlak en we krijgen een slingerbeweging met periode gelijk aan $2\pi\sqrt{m/k}$, die niet afhangt van c. Dit volgt natuurlijk ook uit het reeds bekende feit dat de oplossingen gegeven worden door $x(t) = A\,\cos\left(\sqrt{\frac{k}{m}}\,t + B\right)$, met A en B constanten.

Voor de beweging van een voorwerp op een kromme K in een verticaal vlak onder invloed van de zwaartekracht, was het constant zijn van H het uitgangspunt van Huygens [37, Part II] voor zijn analyse van de slingerbeweging. Hierbij is $x \in \mathbf{R}$ de *booglengte* langs K gemeten vanuit een basispunt en $V(x) = m\,g\,h(x)$ met $h(x)$ de hoogte van het punt met de coördinaat x. Huygens gaf het principe niet in de vorm van de hierboven vermelde formules, maar zijn beschrijving is equivalent aan het constant zijn van H. Briljant was daarbij zijn vinding dat de slinger *tautochroon* is, dat wil zeggen dat de periode niet afhangt van de uitwijking, als K een *cycloïde* is. Dit is, op een verschuiving en factorvergroting na, de kromme die (niet met booglengte!) geparametriseerd wordt door

$$\xi \mapsto (\xi - \sin\xi,\, \cos\xi). \tag{3.6}$$

Met de ons ter beschikking staande differentiaalrekening is dit niet zo moeilijk om in te zien, maar Huygens had die niet tot zijn beschikking en deed alles in termen van de Euclidische meetkunde en Archimedische limietbeschouwingen.

3.2 Meer Vrijheidsgraden

Voor een systeem met n vrijheidsgraden met een krachtveld ϕ dat alleen van de positie-coördinaten afhangt krijgen we Newton's bewegingsvergelijkingen in de gedaante

$$m_i \frac{d^2 x_i}{dt^2} = \phi_i(x), \quad 1 \le i \le n. \tag{3.7}$$

Hierin is m_i een positieve constante, die de *trage massa van de i-de vrijheidsgraad* genoemd wordt. Verder is $\phi : X \to \mathbf{R}^n$, met X een open deelverzameling van $\mathbf{R}^n$; de faseruimte is hier dus gelijk aan $U = X \times \mathbf{R}^n$.

Aangemoedigd door het geval van één vrijheidsgraad maken we nu de aanname dat ϕ gelijk is aan minus de gradiënt van een functie $V(x)$, dat wil zeggen $V \in \mathrm{C}^1(X,\, \mathbf{R})$ en

$$\phi_i(x) = -\frac{\partial V(x)}{\partial x_i}, \quad 1 \le i \le n,\ x \in X. \tag{3.8}$$

Als $\phi \in \mathrm{C}^1$ dan impliceert dit dat $V \in \mathrm{C}^2$ en de verwisselbaarheid van de tweede orde afgeleiden van V leidt tot de sterk inperkende noodzakelijke voorwaarde voor ϕ dat

$$\frac{\partial \phi_i(x)}{\partial x_j} - \frac{\partial \phi_j(x)}{\partial x_i} = 0, \quad 1 \le i,\, j \le n,\ x \in X. \tag{3.9}$$

Omgekeerd geven $\phi \in \mathrm{C}^1$ en (3.9) dat er lokaal altijd een V is waarvoor (3.8) geldt, deze V is op een additieve constante na eenduidig vastgelegd. Of de lokale $V's$ aan elkaar te passen zijn tot een globale oplossing V van (3.8) blijkt een topologische kwestie; voldoende hiervoor is dat het gebied X zogenaamd *enkelvoudig samenhangend* is, dat wil zeggen iedere gesloten kromme in X is in X samen te trekken tot een punt. We gaan hier niet verder op in, maar vermelden nog dat het krachtveld ϕ *conservatief* genoemd wordt wanneer er een oplossing V van (3.8) in X is.

De pointe is nu dat de functie

$$H(x,\, v) := \frac{1}{2} \sum_{i=1}^{n} m_i\, {v_i}^2 + V(x), \quad x \in X,\, v \in \mathbf{R}^n \tag{3.10}$$

een constante van beweging is voor het stelsel (3.7). Bewijs:

$$\begin{aligned} \frac{d}{dt} H(x(t),\, x'(t)) &= \sum_{i=1}^{n} m_i\, x_i'(t)\, x_i''(t) + \sum_{i=1}^{n} \partial_i V(x(t))\, x_i'(t) \\ &= \sum_{i=1}^{n} \left(m_i\, x_i''(t) + \partial_i V(x) \right)\, x_i'(t) = 0. \end{aligned}$$

De functie H heet de *(totale) energie* van het systeem, waarbij de term

$$T = T(x,\, v) := \sum_{i=1}^{N} \frac{1}{2}\, m_i\, {v_i}^2 \tag{3.11}$$

de *kinetische energie* heet en $V(x)$ de *potentiële energie*. De letters T, V en H zijn de keuzen van Lagrange [52, t. I, p. 268, 288, 290]. Op p.217 schrijft Lagrange: "... et Huyghens ... inventa un principe nouveau, mais indirect, lequel est devenue célèbre depuis, sous le nom de *conservation de forces vives*." Volgens J.-M. Souriau is dit de verklaring voor Lagrange's keuze van de letter H voor de totale energie. Hoewel ik dit een verleidelijk idee vind, kon ik in [52] geen expliciete verklaring voor Lagrange's keuze van de gebruikte letters vinden en zijn er voor T en V geen voor de hand liggende namen van voorgangers. In ieder geval staat vast dat Lagrange met H nooit Hamilton bedoeld kan hebben, want die leefde in een later tijdperk.

Volgens Lagrange [52, t. I, p. 226] werd de wet van behoud van energie in de eerste helft van de 18e eeuw door Johann en Daniel Bernoulli expliciet als een belangrijk principe in de mechanica geformuleerd.

In contrast met het geval van één vrijheidsgraad, geeft het constant zijn van H voor $n > 1$ relatief weinig informatie over de banen. De reden daarvoor is, dat de niveauverzamelingen voor H in U^{reg} $(2n-1)$-dimensionale deelvariëteiten van de $2n$-dimensionale faseruimte zijn, waarmee het stelsel (3.7) gereduceerd wordt tot een eerste-orde stelsel van dimensie $2n-1$. Bijvoorbeeld, als $n = 2$ dan leidt dit tot een driedimensionaal eerste-orde stelsel, hetgeen al voldoende 'ruimte' laat voor het optreden van choatische banen; hopelijk krijgen we nog de gelegenheid om hier later voorbeelden van te zien.

Anderzijds dient opgemerkt te worden dat de *wet van behoud van energie* in de natuurkunde een reikwijdte heeft die ver uitgaat boven het hiervoor beschreven verschijnsel voor klassiek-mechanische systemen. Bij interactie van fysische systemen van allerlei aard blijft de totale energie, de som van kinetische, potentiële, thermische, electromognetische, ... energie onder zeer algemeen geldende voorwaarden constant. Met recht is de wet van behoud van energie daarom als één van de basisprincipes van de klassieke natuurkunde te beschouwen.

3.3 Vraagstukken

Vraagstuk 3.1 Geef een gedetailleerde beschrijving van de oplossingen van de tweede-orde diferentiaalvergelijking $x'' = x^2 - x$. Geef ook een schets in het fasevlak van de niveaukrommen van H. $\oslash$

Vraagstuk 3.2 Beschouw een differentiaalvergelijking van de vorm (3.1) met $\phi \in \mathrm{C}^1$. Neem aan dat $0 \in I$, $\phi\,(0) = 0$, $\phi'\,(0) = -k$, $k > 0$. Bewijs dat er een omgeving U_0 van $(0,\, 0)$ in $I \times \mathbf{R}$ is met de eigenschap dat iedere oplossing met $(x_0,\, v_0) \in U_0 \setminus \{(0,\, 0)\}$ als $x_0 = x\,(t_0)$, $v_0 = x'\,(t_0)$ periodiek is, met periode $\omega = \omega\,(x_0,\, v_0)$. Bewijs dat $\omega\,(x_0,\, v_0)$ convergeert als $(x_0,\, v_0) \to (0,\, 0)$ en bereken de limiet. $\oslash$

Vraagstuk 3.3 Beschouw de zogenaamde *mathematische slinger*, de beweging van een voorwerp onder invloed van de zwaartekracht, op een verticale cirkel met de oorsprong in het middelpunt en met straal gelijk aan R. Als x de booglengte voorstelt vanaf het laagste punt, dan is de hoogte gelijk aan $-R\cos(x/R)$. Maak een schets in het fasevlak van de niveaukrommen van H. Beschrijf de bewegingen en bepaal voor welke beginwaarden van positie en snelheid de beweging periodiek is. Bepaal de limietperiode in de buurt van het evenwichtspunt. Toon aan dat de periode naar ∞ gaat als $H \uparrow m\,g\,R$. Is de mathematische slinger tautochroon? ⊘

Vraagstuk 3.4 Zij $x(\xi)$ de booglengte langs de cycloïde (3.6). Bewijs dat

$$\frac{dx(\xi)}{d\xi} = (2\,(1-\cos\xi))^{1/2}.$$

Differentieer $h(x(\xi)) = \cos\xi$ tweemaal naar ξ en bereken daarmee achtereenvolgens $h'(x(\xi))$ en $h''(x(\xi))$. Verifieer hiermee de stelling van Huygens dat de cycloïde een tautochrone slinger definieert. ⊘

Vraagstuk 3.5 Beschouw het stelsel (1.21) in $\mathbf{R}^3$, waarvoor we het snelheidsmoment μ en de eccentriciteitsvector ε al als constanten van beweging hebben leren kennen. Bewijs dat het krachtveld conservatief is, met potentiële energie gelijk aan $V(x) = -c/\|x\|$. Bewijs dat

$$\|\varepsilon\|^2 = 1 + 2\,\|\mu\|^2\,H/c^2.$$

Omdat voor $\mu \neq 0$ de banen bepaald worden door de waarden van μ en ϵ en H constant is langs de banen, moet H voor $\mu \neq 0$ een functie zijn van μ en ε; dit wordt door de bovenstaande formule fraai bevestigd.

Bewijs dat de elliptische, resp. parabolische, resp. hyperbolische banen corresponderen met $\mu \neq 0$ en $H < 0$, resp. $H = 0$, resp. $H > 0$. Bewijs dat de perioden van de periodieke oplossingen gegeven worden door

$$\omega = 2\pi\,c\,\left(-2H\right)^{-3/2},$$

dus de periode is een functie van de energie.

Bewijs dat voor $\mu = 0$ en $H < 0$ het definitie-interval van de oplossing een eindige lengte heeft, gelijk aan de bovenstaande ω. ⊘

Vraagstuk 3.6 Beschouw de differentiaalvergelijking $\lambda'' = -c\,\lambda^{-2}$ uit onderdeel f) van Vraagstuk 3.6, waarin $c > 0$ een constante is. We beperken ons tot de oplossingen $\lambda(t)$ met $\lambda(0) > 0$, hetgeen impliceert dat $\lambda(t) > 0$ voor alle t in het maximale definitie-interval $]a, b[$ van de oplossing. Hierin is $-\infty \leq a < 0 < b \leq \infty$.

Bewijs dat $h = \frac{1}{2}\,\lambda'(t)^2 - c\,\lambda(t)^{-1}$ een constante van beweging is. Behandel in het vervolg de gevallen $h < 0$, $h = 0$ en $h > 0$ apart. Bereken t in termen van $\lambda(t)$, waarbij c en $\lambda'(0)$ als parameters optreden. Hint: Als $h \gtrless 0$, dan wordt men na wat herschalingen geconfronteerd met het primitiveren van de functie

$$\left(x^{-1} \pm 1\right)^{-1/2} = x\,\left(x \pm x^2\right)^{-1/2} = x\,\left(\pm\left(x \mp \frac{1}{2}\right)^2 \mp \frac{1}{4}\right)^{-1/2},$$

hetgeen na wat nieuwe herschalingen neerkomt op het primitiveren van de functies

$$\pm 2x\,\left(\pm x^2 \mp 1\right)^{-1/2} \quad \text{en} \quad \left(\pm x^2 \mp 1\right)^{-1/2}.$$

De primitieve van de eerste functie is gelijk aan $\left(\pm x^2 \mp 1\right)^{1/2}$, terwijl we voor de tweede gebruik kunnen maken van

$$\int \left(x^2-1\right)^{-1/2}\,\mathrm{d}x = \log\left|x + \left(x^2-1\right)^{1/2}\right|,$$

respectievelijk

$$\int \left(1-x^2\right)^{1/2}\,\mathrm{d}x = \arcsin x.$$

Beschrijf in ieder van de gevallen de oplossingen $\lambda(t)$. Bewijs in het bijzonder de volgende conclusies.

Als $h < 0$, dan is $-\infty < a < b < \infty$ en $\lambda(t) \to 0$ als $t \downarrow a$ en als $t \uparrow b$. Bereken $b - a$ en de maximale waarde van $\lambda(t)$.

Anderzijds, als $h \geq 0$ dan is $\lambda'(0) \neq 0$. Is daarbij $\lambda'(0) > 0$ dan is $a > -\infty$, $b = \infty$ en $\lambda(t) \to 0$ als $t \downarrow a$ en $\lambda(t) \to \infty$ als $t \to \infty$.

De snelheid $w := \sqrt{2c/\lambda(0)}$ heet de *ontsnappingssnelheid*. Bewijs het volgende. Als $\lambda'(0) < w$, dan is $b < \infty$ en $\lambda(t) \to 0$ als $t \uparrow b$. Anderzijds, als $\lambda'(0) \geq w$, dan is $b = \infty$, waarbij $\lambda(t) \uparrow \infty$ en $\lambda'(t) \downarrow \left[\lambda'(0)^2 - w^2\right]^{1/2}$ als $t \to \infty$.

$\oslash$

Vraagstuk 3.7 Bewijs dat het krachtveld (1.31), voor het probleem van N lichamen die elkaar gravitationeel aantrekken, conservatief is en geef een formule voor de potientiële energie. $\oslash$

Vraagstuk 3.8 In de continuümsmechanica heeft een begrensd lichaam met eindige totale massa een *massadichtheid* die gegeven is door een Riemann-integreerbare functie $\rho : \mathbf{R}^3 \to \mathbf{R}_{\geq 0}$ met compacte drager $\operatorname{supp}\rho$. Volgens Newton's superpositieprincipe wordt een deeltje in een positie x, met $x \notin \operatorname{supp}\rho$, gravitationeel door de massaverdeling aangetrokken met een versnelling die gelijk is aan

$$a(x) = -\mathcal{G}\int \|x-y\|^{-3}\,(x-y)\,\rho(y)\,dy, \quad x \notin \operatorname{supp}\rho.$$

Stel nu dat de massaverdeling *rotatie-symmetrisch* is, dat wil zeggen, als $\|x\| = \|y\|$, dan is $\rho(x) = \rho(y)$. Newton [68, p. 193-195] bewees: Als $\|x\| < \|y\|$ voor iedere $y \in \operatorname{supp}\rho$ dan is $a(x) = 0$. En: er is een constante c is met $a(x) = -c\,\|x\|^{-3}\,x$ als $\|x\| > \|y\|$ voor iedere $y \in \operatorname{supp}\rho$.

De volgende elegante afleiding hiervan is afkomstig van Laplace [55, p. 137-139]. Hierbij wordt gebruik gemaakt van de tweede-orde partiële differentiaaloperator

$$\Delta = \sum_{i=1}^{3} \partial_i{}^2 = \operatorname{div}\circ\operatorname{grad},$$

de *Laplace-operator* genoemd. (Eerder had Laplace [55, p. 95, 96] opgemerkt dat $\Delta V = 0$ betekent dat een stroming met snelheidsveld gelijk aan $\operatorname{grad} V$ volume-bewarend is.)

Bewijs achtereenvolgens:

a) $a(x) = -\operatorname{grad} V(x)$, met

$$V(x) := -\mathcal{G}\int \|x-y\|^{-1}\,\rho(y)\,dy, \quad x \notin \operatorname{supp}\rho.$$

b) $W \in \mathrm{C}^2$ en $\Delta W = 0$ als $W(x) = \|x\|^{-1}$, $x \in \mathbf{R}^3 \setminus \{0\}$.

c) $V \in \mathrm{C}^2$ en $\Delta V = 0$ in het complement van $\operatorname{supp} \rho$.

d) Als ρ rotatie-symmetrisch is dan is V rotatie-symmetrisch, dus een functie van de afstand tot de oorsprong.

e) Is $V = f(\|x\|)$ en $\Delta V(x) = 0$ voor $a < \|x\| < b$, dan zijn er constanten A en B waarvoor $f(r) = A + \frac{B}{r}$ als $a < r < b$.

f) Als ρ rotatie-symmetrisch is en $\rho(x) = 0$ zodra $\|x\| < b$, dan is V continu in 0, dus $B = 0$ in e) en daarmee $a(x) = 0$ als $\|x\| < b$.

g) Als ρ rotatie-symmetrisch is en $\rho(x) = 0$ zodra $\|x\| > a$, dan geeft e) dat

$$A + \frac{B}{\|x\|} = V(x) = -\mathcal{G} \int \|x - y\|^{-1} \rho(y)\, dy, \quad \|x\| > a.$$

Een limietbeschouwing voor $\|x\| \to \infty$ geeft nu eerst dat $A = 0$ en vervolgens dat

$$B = -\mathcal{G} \int \rho(x)\, dx.$$

Daarmee is de aantrekking buiten de drager van ρ hetzelfde als wanneer alle massa in de oorsprong geplaatst zou zijn.

$\oslash$

3.4 Abel

Joop Kolk wees mij erop dat de vraag welke functie $x \mapsto V(x)$ een gegeven periodefunctie $c \mapsto \omega(c)$ oplevert als in (3.5) opgelost is door Abel [1], [2]. Impliciet beperkte Abel zich tot het geval dat $V \in \mathrm{C}^1$, dat $V(0) = 0$, $V'(0) = 0$, terwijl verder aangenomen wordt dat $V'(x) > 0$ voor $x > 0$. De tijd $T_+(c)$ die nodig is om van $x = 0$ tot $x = b$, met $b > 0$ en $V(b) = c$, te geraken is dan gelijk aan

$$T_+(c) = \int_0^b \left(\frac{2}{m} (c - V(x)) \right)^{-1/2} dx.$$

Abel loste hieruit de functie $x \mapsto V(x)$, $x \geq 0$ op in termen van de functie $c \mapsto T_+(c)$. Als we ook nog eisen dat V *symmetrisch* is om de oorsprong, dan is $\omega(c) = 4\, T_+(c)$ en krijgen we V in termen van ω. Laten we echter asymmetrische V toe, dan geeft het gedeelte van de beweging met $x \leq 0$ een tijd $T_-(c)$ en krijgen we $\omega(c) = 2\ (T_+(c) + T_-(c))$. In dit geval kunnen we $V(x)$ voor $x \leq 0$ willekeurig kiezen, hetgeen een functie $c \mapsto T_-(c)$ oplevert en vervolgens Abel's oplossing voor $V(x)$, $x \geq 0$ zó inrichten dat we gewenste gedaante voor $c \mapsto \omega(c)$. Natuurlijk moet hierbij dan wèl 'ruimte' voor $T_+(c)$ overblijven, dat wil zeggen dat $V(x)$ voor $x \leq 0$ zó moet zijn dat $\omega(c) - 2\, T_-(c) > 0$. Dit laat echter in ieder geval zien dat V verre van eenduidig is als we asymmetrische V toelaten.

De eerste stap van Abel was om $y = V(x)$ als nieuwe variabele te gebruiken, waarmee hij (3.4) verkreeg in de vorm waarbij in de integraal de substitutie van variabelen $x = V^{-1}(y)$ is toegepast. Voor $x_0 = 0$ en $x = V^{-1}(c)$ leidt dit tot een integraalvergelijking van de vorm

$$f(c) = \int_0^c \frac{g(y)}{\sqrt{c - y}}\, dy, \tag{3.12}$$

waarin $f(c) = \sqrt{\frac{2}{m}}\, T_+(c)$ en $g(y) = 1/V'\left(V^{-1}(y)\right)$. Merk op dat $g(y)$ continu is voor $Y > 0$. Verder, als $V \in \mathrm{C}^2$ en $k := V''(0) > 0$, dan is

$$V'(x) \sim k\,x, \quad y \sim k\,x^2, \quad x \downarrow 0,$$

dus

$$V'\left(V^{-1}(y)\right) \sim (2\,k\,y)^{1/2}, \quad y \downarrow 0;$$

dit maakt dat $g(y)$ absoluut integreerbaar is bij $y = 0$.

Het rechterlid in (3.12) hangt op lineaire wijze af van de functie g; daarmee is (3.12) een *lineaire integraalvergelijking* voor g. Het idee is om, als we hieruit g op kunnen lossen in termen van f, we vervolgens V zullen bepalen als oplossing van de eerste-orde differentiaalvergelijking

$$g(V(x))\, V'(x) = 1, \quad V(0) = 0,$$

ofwel $V(x) = G^{-1}(x)$ als G een primitieve is van g.

In [1] loste Abel de vergelijking (3.12) op met behulp van reeksen; hiervoor moeten echter onnodig sterke eisen opgelegd worden aan de functies f en g. Abel's oplossing in [2] is algemener en ging als volgt. Hij vermenigvuldigde beide leden van (3.12) met $1/\sqrt{z-c}$, integreerde over $c \in\,]0,\, z[$ en verwisselde de integratievolgorde in de dubbelintegraal:

$$\begin{aligned}\int_0^z \frac{f(c)}{\sqrt{z-c}}\, dc &= \int_0^z \left(\int_0^c \frac{g(y)}{\sqrt{(c-y)(z-c)}}\, dy \right) dc \\ &= \int_0^z g(y) \left(\int_y^z \frac{dc}{\sqrt{(c-y)(z-c)}} \right) dy = \pi \int_0^z g(y)\, dy.\end{aligned}$$

Hierin zijn de substituties van variabelen $c = y + u\,(z-y)$ en $u = \frac{1}{2} + \frac{1}{2}\,v$ gebruikt om te concluderen dat

$$\int_y^z \frac{dc}{\sqrt{(c-y)(z-c)}} = \int_0^1 \frac{du}{\sqrt{u(1-u)}}\, du = \int_{-1}^1 \frac{dv}{\sqrt{1-v^2}}\, dv = \pi.$$

De conclusie is daarmee dat als $g(y)$ continu is voor $y > 0$ en absoluut integreerbaar bij $y = 0$, dan impliceert (3.12) dat

$$G(z) := \int_0^z g(y)\, dy = \frac{1}{\pi} \int_0^z \frac{f(z)}{\sqrt{z-c}}\, dc.$$

Hiermee is g eenduidig bepaald in termen van f door de integraalvergelijking (3.12).

Abel paste dit vervolgens toe op het probleem van de tautochrone symmetrische slinger: daarvoor wordt geëist dat $f(z)$ een constante is, hetgeen impliceert dat $G(z)$ een constant veelvoud is van $z^{1/2}$, waarmee $V(x)$ een constant veelvoud is van x^2, hetgeen op zijn beurt correspondeert met de cycloïde van Huygens.

Opmerkelijk is dat Abel niet slechts de integraalvergelijking (3.12) oploste, maar voor iedere constante α met $0 < \alpha < 1$ bewees dat

$$f(c) = \left(I^{\alpha}(g)\right)(c) := \frac{1}{\Gamma(\alpha)} \int_0^c g(y)\, (c-y)^{\alpha-1}\, dy \tag{3.13}$$

impliceert dat $G(y) = \left(I^{1-\alpha}(f)\right)(y)$. De vergelijking (3.13) heet de *integraalvergelijking van Abel.* Vaak wordt hierbij alleen maar het geval (3.12) besproken. (Het geval dat $\alpha = 1/2$, met de toepassing in de klassieke mechanica.)

Liouville [58] generaliseerde de identiteit $I^{1-\alpha} \circ I^{\alpha} = I^{1}$ van Abel tot de identiteit

$$I^{\beta} \circ I^{\alpha} = I^{\beta+\alpha}, \tag{3.14}$$

voor alle positieve reële getallen α en β. In het artikel wordt Abel niet vermeld. Dit resultaat werd door Riemann [78] als student herontdekt, zonder verwijzing naar Abel of Liouville. Riemann merkte op dat (3.14) het zinvol maakt om de parameter α in I^{α} als een exponent op te vatten en $(I^{\alpha} g)(y)$ te interpreteren als de *fractionele afgeleide* $\left(\frac{d}{dy}\right)^{-\alpha} g(y)$. Verder merkte hij op dat hierin ook complexe exponenten toegelaten kunnen worden. (Daarbij is de integraal in (3.13) voor een begrensde en continue functie g absoluut convergent als aangenomen wordt dat het reële deel van α positief is.) De theorie met complexe exponenten is zeer volledig uitgewerkt door M. Riesz [79]. Als van de theorie van distributies gebruik gemaakt mag worden, dan kan dit in veel korter bestek uitgelegd worden, zie bijvoorbeeld [17, §13] of [18].

4 Lagrange's Mechanica

4.1 Substituties van Positievariabelen

Lagrange [52, t. I, 2. Partie, Sect. IV] stelde zich de vraag hoe de vergelijkingen van Newton (3.7) eruit gaan zien na een willekeurige substitutie van variabelen $x = \Psi(t, y)$ met $x \in \mathbf{R}^n$ en $y \in \mathbf{R}^p$. Hij legde daarbij de nadruk op het woord 'willekeurige', omdat pas ná de substitutie gekeken kan worden of het stelsel er eenvoudiger uit is gaan zien bij geschikte keuze van Ψ. We nemen aan dat $\Psi \in \mathrm{C}^2$, zodat een tweemaal differentieerbare kromme $t \mapsto \delta(t)$ door middel van

$$\gamma(t) = \Psi(t, \delta(t)) \tag{4.1}$$

overgevoerd wordt in een tweemaal differentieerbare functie $t \mapsto \gamma(t)$.

In eerste instantie keek Lagrange naar het ietwat eenvoudiger geval dat $p = n$ en dat de substitutie $x = \Psi(t, y)$ niet expliciet van de tijde t afhangt, maar we zullen zien dat veel van de beweringen geldig blijven voor willekeurige dimensies p en voor tijdsafhankelijke substituties.

Een rechttoe-rechtaan berekening geeft nu dat

$$\begin{aligned} \gamma_h'(t) &= \frac{\partial \Psi_h(t, y)}{\partial t}|_{y=\delta(t)} \\ &\quad + \sum_{i=1}^{p} \frac{\partial \Psi_h(t, y)}{\partial y_i}|_{y=\delta(t)} \cdot \delta_i'(t) \end{aligned} \tag{4.2}$$

en vervolgens

$$\begin{aligned} \gamma_h''(t) &= \frac{\partial^2 \Psi_h(t, y)}{\partial t^2}|_{y=\delta(t)} \\ &\quad +2 \sum_{i=1}^{p} \frac{\partial^2 \Psi_h(t, y)}{\partial t \partial y_i}|_{y=\delta(t)} \cdot \delta_i'(t) + \sum_{i,\, j=1}^{p} \frac{\partial^2 \Psi_h(t, y)}{\partial y_j \partial y_i}|_{y=\delta(t)} \cdot \delta_i'(t)\, \delta_j'(t) \\ &\quad + \sum_{i=1}^{p} \frac{\partial \Psi_h(t, y)}{\partial y_i}|_{y=\delta(t)} \cdot \delta_i''(t). \end{aligned} \tag{4.3}$$

Zelfs in het geval dat de substitutie niet expliciet van t afhangt, dat wil zeggen als $x = \Psi(y)$ voor een C^2-diffeomorfisme Ψ, dan zien we dat $\gamma''(t)$ niet eenvoudigweg het beeld is van $\delta''(t)$ onder de lineaire afbeelding $\partial\Psi/\partial y$, maar dat er nog kwadratische termen in de snelheid $\delta'(t)$ bijkomen, zodra de tweede-orde afgeleiden van $y \mapsto \Psi(t, y)$ niet identiek gelijk aan nul zijn. Anders gezegd, het enige geval waarvoor de uitdrukking voor $\gamma''(t)$ geen kwadratische termen in $\delta'(t)$ bevat, treedt op als voor iedere t de afbeelding $y \mapsto \Psi(t, y)$ een zogenaamde *affiene transformatie* is, een samenstelling van een lineaire transformatie en een translatie.

Lagrange's eerste opmerking was dat het linkerlid in (3.7) uitgedrukt kan worden in termen van de kinetische energiefunctie, zoals die gedefinieerd is in (3.11). We hebben de formule

$$p_i := m_i\, v_i = \frac{\partial T(v)}{\partial v_i} \tag{4.4}$$

voor de i-de component van de zogenaamde *impulsvector* p. Hieruit volgt dat het linkerlid in (3.7) gegeven wordt door

$$\frac{dp_i(t)}{dt} = m_i\, \gamma_i''(t) = \frac{d}{dt} \frac{\partial T(v)}{\partial v_i}|_{v=\gamma'(t)}. \tag{4.5}$$

Nu leidt een substitutie van variabelen $x = \Psi(t,\, y)$ na substitutie van (4.2) in $T(v)$ tot een kinetische energiefunctie

$$S(t,\, y,\, w) = T(v), \quad v := \frac{\partial\Psi(t,\, y)}{\partial t} + \frac{\partial\Psi(t,\, y)}{\partial y} \cdot w, \tag{4.6}$$

waarbij de leidraad is dat moet gelden dat

$$S(t,\, \delta(t),\, \delta'(t)) = T(\gamma'(t))$$

als (4.1) geldt. De nieuwe kinetische energiefunctie hangt in het algemeen niet alleen af van de nieuwe snelheidsvector w, maar ook van de de nieuwe positie y en de tijd t, waarbij de y-afhankelijkheid alleen dan niet optreedt als de substitutie van variabelen een affiene transformatie is. Lagrange's ontdekking was nu dat we een eenvoudige transformatieformule voor het linkerlid in (3.7) krijgen, als we van de vector $\frac{d}{dt}\frac{\partial S}{\partial w}$ nog de vector $\frac{\partial S}{\partial y}$ aftrekken. Preciezer, er geldt de volgende algemene stelling.

Stelling 4.1 *Zij U een open deelverzameling van $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ en L een reëelwaardige C^1-functie in U, waarvoor $(t,\, x,\, v) \mapsto \frac{\partial L(t,x,v)}{\partial v}$ een C^1-afbeelding is van U naar $\mathbf{R}^n$. Zij verder Ψ een C^2-afbeelding van een open deelverzameling Ω van $\mathbf{R} \times \mathbf{R}^p$ naar $\mathbf{R}^n$, zij V de verzameling der $(t,\, y,\, w)$ waarvoor $(t,\, x,\, v) \in U$ als*

$$x = \Psi(t,\, y), \quad \text{en} \quad v = \frac{\partial\Psi(t,\, y)}{\partial t} + \frac{\partial\Psi(t,\, y)}{\partial y} \cdot w. \tag{4.7}$$

Tenslotte definiëren we $M(t,\, y,\, w) = L(t,\, x,\, v)$ als x en v gegeven zijn door (4.7). Met deze definities is V een open deelverzameling van $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ en M een reëelwaardige C^1-functie in V, waarvoor $(t,\, y,\, w) \mapsto \frac{\partial M(t,y,w)}{\partial w}$ een C^1-afbeelding is van V naar $\mathbf{R}^n$.

Zij nu $\delta : I \to \mathbf{R}^p$ een tweemaal differentieerbare kromme, waarin I een open interval in $\mathbf{R}$ is en voor iedere $t \in I$ geldt dat $(t,\, \delta(t),\, \delta'(t)) \in V$; schrijf $\gamma(t) := \Psi(t,\, \delta(t))$. Definieer

$$[L]_h^\gamma(t) := \frac{d}{dt}\left(\frac{\partial L(t,\, \gamma(t),\, v)}{\partial v_h}|_{v=\gamma'(t)}\right) - \frac{\partial L(t,\, x,\, \gamma'(t))}{\partial x_h}|_{x=\gamma(t)} \tag{4.8}$$

en analoog $[M]_i^\delta(t)$ met L, resp. γ vervangen door M, resp. δ. Dan geldt de transformatieformule

$$[M]_i^\delta(t) = \sum_{h=1}^{n} [L]_h^\gamma(t) \cdot \frac{\partial\Psi_h(t,\, y)}{\partial y_i}|_{y=\delta(t)}, \quad 1 \le i \le p. \tag{4.9}$$

In het geciteerde werk van Lagrange ziet de uitspraak er veel eenvoudiger uit, omdat de variabelen in de diverse functies niet uitgeschreven worden. In de notatie wordt zelfs geen onderscheid gemaakt tussen de functies L, resp. M: omdat deze functies dezelfde waarden aannemen in $(t,\, \gamma(t),\, \gamma'(t))$, resp. $(t,\, \delta(t),\, \delta'(t))$, worden ze als 'dezelfde grootheid' beschouwd en daarom met dezelfde letter aangeduid.

Gevolg *Het tweede orde stelsel van differentiaalvergelijkingen $[L]^\gamma = 0$ voor de kromme γ is equivalent met het tweede orde stelsel van differentiaalvergelijkingen $[M]^\delta = 0$ voor de kromme δ.*

4.2 Een Variatieformule

Stelling 4.1 kan bewezen worden door een directe berekening, die in de hier toegestane algemeenheid nogal bewerkelijk is en niet veel inzicht verschaft. In plaats daarvan zullen wij in navolging van Lagrange laten zien dat de grootheid $[L]_h^{\gamma}(t)$, $1 \leq h \leq n$ in (4.8) optreedt als integrand in een *variatieformule*, waaruit Stelling 4.1 dan meteen volgt.

Om dit uit te leggen, fixeren we a, $b \in I$ met $a < b$ en beschouwen we de zogenaamde *actie-integraal*

$$\mathcal{I}(L, \gamma) := \int_a^b L(t, \gamma(t), \gamma'(t))\, dt \tag{4.10}$$

van L langs $t \mapsto (t, \gamma(t), \gamma'(t))$ als functie van de kromme γ. Anders gezegd, $\mathcal{I}$ wordt gezien als een scalairwaardige functie op de (oneindig-dimensionale) ruimte van krommen γ. We nemen vervolgens aan dat $\gamma = \gamma_0$, waarin $\epsilon \mapsto \gamma_\epsilon$ een differentieerbare afbeelding is van een omgeving J van 0 in $\mathbf{R}$, naar de ruimte van krommen in $\mathbf{R}^n$. Preciezer, we nemen aan dat

$$\gamma_\epsilon(t) = \Gamma(t, \epsilon), \quad a \leq t \leq b, \quad \epsilon \in J, \tag{4.11}$$

waarin $\Gamma : [a, b] \times J \to \mathbf{R}^n$ een C^1-afbeelding is waarvoor ook

$$(t, \epsilon) \mapsto \frac{\partial \Gamma(t, \epsilon)}{\partial t}$$

een C^1-afbeelding is van $[a, b] \times J$ naar $\mathbf{R}^n$. Vanwege de stelling over de verwisselbaarheid van de differentiatievolgorde is dan ook $t \mapsto \frac{\partial \Gamma(t, \epsilon)}{\partial \epsilon}$ differentieerbaar en is

$$\frac{\partial^2 \Gamma(t, \epsilon)}{\partial \epsilon \partial t} = \frac{\partial^2 \Gamma(t, \epsilon)}{\partial t \partial \epsilon}$$

We vervangen nu γ door γ_ϵ in (4.10), differentiëren naar ϵ onder het integraalteken, verwisselen de differentiatievolgorde in de gemengde tweede-orde afgeleide van Γ die daarbij verschijnt en voeren tot slot een partiële integratie uit met betrekking tot de integratievariabele t. Noteren we

$$\widetilde{\gamma}(t) := \frac{\partial \Gamma(t, \epsilon)}{\partial \epsilon}|_{\epsilon=0}, \tag{4.12}$$

dan is het resultaat van de bovengenoemde operaties dat

$$\frac{d}{d\epsilon} \mathcal{I}(L, \gamma_\epsilon)|_{\epsilon=0} = -\int_a^b \sum_{h=1}^n [L]_h^{\gamma}(t) \cdot \widetilde{\gamma}_h(t)\, dt + \left[\sum_{h=1}^n \frac{\partial L(t, \gamma(t), v)}{\partial v_h}|_{v=\gamma'(t)} \cdot \widetilde{\gamma}_h(t) \right]_{t=a}^{t=b}. \tag{4.13}$$

De gebruikelijke toepassing van de variatieformule (4.13) gaat als volgt. Veronderstel dat we ons beperken tot een deelruimte van krommen die bepaald zijn door voorwaarden in de randpunten $t = a$, resp. $t = b$, die zó gekozen zijn dat voor de richting $\widetilde{\gamma}$ waarin we γ variëren automatisch geldt dat de bijdrage van de randtermen in (4.13) gelijk aan nul is. Een voorbeeld hiervan is dat men het begin- resp. eindpunt $\gamma(a)$, resp. $\gamma(b)$ fixeert in punten x_0, resp. x_1, in welk geval $\widetilde{\gamma}(a) = 0$ en $\widetilde{\gamma}(b) = 0$. Een ander voorbeeld is het geval van de zogenaamde *periodieke randvoorwaarden*, waarbij men eist dat $\Gamma(t, \epsilon)$ en $\frac{\partial \Gamma(t, \epsilon)}{\partial t}$ allebei voor $t = b$ dezelfde waarde aannemen als voor $t = a$, zodat we een gesloten kromme in de faseruimte krijgen. Dit impliceert dat $\widetilde{\gamma}(a) = \widetilde{\gamma}(b)$ en ook in dit geval is het verschil van de randtermen eveneens gelijk aan nul. De uitspraak is nu dat in een dergelijke deelruimte van krommen de functie $\mathcal{I}$ een stationair punt heeft in γ, dat wil zeggen dat (4.13) gelijk aan nul is voor alle als hierboven toegestane variaties $\widetilde{\gamma}$ van γ, dan en slechts dan als

$$[L]_h^\gamma(t) = 0, \quad a \leq t \leq b, \quad 1 \leq h \leq n. \tag{4.14}$$

Dat (4.14) impliceert dat (4.13) gelijk aan nul is voor alle toegestane variaties $\widetilde{\gamma}$ is evident. Voor het bewijs van de noodzakelijkheid van (4.14) kan men bijvoorbeeld Γ zó kiezen dat

$$\widetilde{\gamma}_h(t) = \phi(t)\,[L]_h^\gamma(t),$$

waarin $\phi(t)$ een reëelwaardige continue functie is met $\phi(t) > 0$ als $a < t < b$ en $\phi(a) = \phi(b) = 0$. Daarmee wordt de integrand in (4.13) gelijk aan

$$\Phi(t) := \phi(t) \sum_{h=1}^{n} [L]_h^\gamma(t)^2, \tag{4.15}$$

hetgeen een continue functie is van t die nergens negatief is, terwijl de integraal hiervan gelijk is aan nul. Als $\Phi(t) > 0$, dan geeft de continuïteit dat er een interval I om t is met de eigenschap dat voor iedere $s \in I$ geldt dat $|\Phi(s) - \Phi(t)| < |Phi(t)/2$, hetgeen impliceert dat $\Phi(s) > \Phi(t)/2$. Omdat $\Phi(s) \geq 0$ voor iedere $s \notin I$, leidt dit tot de conclusie dat de integraal groter of gelijk is aan $\Phi(t)/2$ maal de lengte van I, in tegenspraak met de aanname dat de integraal gelijk is aan nul. Dit leidt tot de conclusie dat $\Phi(t) = 0$, hetgeen met het oog op $\phi(t) > 0$ impliceert dat $[L]_h^\gamma(t) = 0$ voor iedere $1 \leq h \leq n$. De continuïteit van $t \mapsto [L]_h^\gamma(t)$ gebruikend krijgen we dat (4.14) ook geldt voor $t = a$ en voor $t = b$.

De vergelijkingen (4.14), resp. de functie L heten de *variatievergelijkingen van Euler-Lagrange*, resp. de *Lagrange-functie*.

4.3 Bewijs van Lagrange's Stelling

We starten het bewijs van Stelling 4.1 met δ in te bedden in een één-parameterfamilie $\delta_\epsilon(t) = \Delta(t,\, \epsilon)$ als boven, met γ, resp. γ_ϵ, resp. Γ vervangen door δ, resp. δ_ϵ, resp. Δ. Door bijvoorbeeld te nemen

$$\Delta(t,\, \epsilon) = \delta(t) + \epsilon\, \widetilde{\delta}(t)$$

voor een willekeurige $\widetilde{\delta} \in \mathrm{C}^2\left([a,\, b]\,,\, \mathbf{R}^p\right)$, zien we dat hierbij

$$\widetilde{\delta}(t) = \frac{\partial \Delta(t,\, \epsilon)}{\partial \epsilon}|_{\epsilon=0}$$

een willekeurige C^2-functie is. Schrijf vervolgens

$$\gamma_\epsilon(t) := \Gamma(t,\, \epsilon) := \Psi(t,\, \Delta(t,\, \epsilon)), \quad a \leq t \leq b, \quad \epsilon \in J.$$

Deze variatie van γ voldoet aan de voorwaarden en we hebben de relatie

$$\widetilde{\gamma}_h(t) := \frac{\partial \Gamma_h(t,\, \epsilon)}{\partial \epsilon}|_{\epsilon=0} = \sum_{i=1}^{p} \frac{\partial \Psi_h(t,\, y)}{\partial y_i}|_{y=\delta(t)} \cdot \widetilde{\delta}_i(t), \quad 1 \leq h \leq n. \tag{4.16}$$

die $\widetilde{\gamma}$ geeft in termen van $\widetilde{\delta}$.

De definitie van M impliceert nu dat

$$\mathcal{I}\left(M,\, \delta_\epsilon\right) = \mathcal{I}\left(L,\, \gamma_\epsilon\right), \quad \epsilon \in J. \tag{4.17}$$

De afgeleide naar ϵ in $\epsilon = 0$ van het linker- en rechterlid van (4.17) uitschrijvend met behulp van (4.13), daarin (4.16) substituerend en opmerkend dat de gelijkheid tussen het linker- en rechterlid geldt voor iedere C^2-kromme $\widetilde{\delta}$, krijgen we met eenzelfde argument als voor (4.14), dat (4.9) geldt.

4.4 Lagrange's Krachtvelden

Zij S de kinetische energie als functie van de nieuwe variabelen $(t,\, y,\, w)$, gedefinieerd in (4.6). Neem verder aan dat de krommen γ en δ aan elkaar gerelateerd zijn door middel van de substitutie (4.1). Stelling 4.1 met L, resp. M vervangen door T, resp. S geeft nu dat

$$[S]_i^\delta(t) = \sum_{h=1}^{n} [T]_h^\gamma(t) \cdot \frac{\partial \Psi_h(t,\, y)}{\partial y_i}|_{y=\delta(t)}, \quad 1 \leq i \leq p,$$

waarbij $[T]_h^\gamma(t) = m_h\, \gamma_h''(t)$, zie (4.5). Zij nu $\phi_h(t,\, x,\, v)$ de h-de component van een krachtenveld. Definieer

$$\psi_i(t,\, y,\, w) := \sum_{h=1}^{n} \phi_h(t,\, x,\, v) \cdot \frac{\partial \Psi_h(t,\, y)}{\partial y_i}, \quad 1 \leq i \leq p, \tag{4.18}$$

waarin $x = \Psi(t,\, y)$ en $v = \frac{\partial \Psi(t,y)}{\partial t} + \frac{\partial \Psi(t,y)}{\partial y} \cdot w$. Met deze notaties krijgen we:

Gevolg 4.2 *Als γ een oplossing is van de bewegingsvergelijkingen van Newton $m_h\, \gamma'' = \phi_h$, $1 \leq h \leq n$, dan voldoet δ aan het tweede-orde stelsel*

$$[S]_i^\delta(t) = \psi_i(t,\, \delta(t),\, \delta'(t)), \quad 1 \leq i \leq p. \tag{4.19}$$

Als $\frac{\partial \Psi(t,y)}{\partial y}$ surjectief is, dan geldt ook de omkering.

De transformatieformule (4.18) voor het krachtenveld is *niet* hetzelfde als die voor snelheidsvectoren, zelfs niet als we ons beperken tot substituties $x = \Psi(y)$ die niet van de tijd afhangen. Daarvoor geldt dat

$$\gamma_h'(t) = \sum_{i=1}^{p} \frac{\partial \Psi_h(y)}{\partial y_i}|_{y=\delta(t)} \cdot \delta_i'(t),$$

afgekort tot $\gamma' = D\Psi\, \delta'$, terwijl ψ gelijk is aan het beeld van ϕ onder de *getransponeerde* van de lineaire afbeelding $D\Psi$. In de terminologie die Ricci en Levi-Civita in het begin van deze eeuw geïntroduceerd hebben, transformeren snelheidsvectoren *contravariant* onder substituties van variabelen en krachten *covariant*.

Het linkerlid in (4.19) hangt op lineaire wijze af van $\delta''(t)$, via de tweede-orde afgeleidenmatrix van S met betrekking tot w. Hieruit zien we dat (4.19) herschreven kan worden als een expliciet tweede-orde stelsel, een stelsel waarin $\delta''(t)$ gegeven wordt in termen van $(t,\, \delta(t),\, \delta'(t))$, dan en slechts dan als deze tweede-orde afgeleidenmatrix niet-singulier is. Nu is

$$\frac{\partial^2 S}{\partial w^2} = \left(\frac{\partial \Psi}{\partial y}\right)^{\mathrm{t}} \circ \frac{\partial^2 T}{\partial v^2} \circ \frac{\partial \Psi}{\partial y}$$

en $\frac{\partial^2 T}{\partial v^2}$ is niet-singulier, dus $\frac{\partial^2 S}{\partial w^2}$ is niet-singulier dan en slechts dan als $\frac{\partial \Psi}{\partial y}$ injectief is. Dit laatste kan alleen als $p \leq n$. Is $p = n$, dan is $\frac{\partial \Psi}{\partial y}$ bijectief en is $y \mapsto \Psi(t, y)$ een lokaal diffeomorfisme; in dit geval zegt Gevolg 4.2 dat de bewegingsvergelijkingen van Newton equivalent zijn met het expliciete tweede-orde stelsel 4.19, welke de *bewegingsvergelijkingen van Lagrange* genoemd worden. Een voor de hand liggende generalisering is dat men hierin een willekeurige 'kinetische energiefunctie' S toelaat (waarvoor de tweede-orde afgeleidenmatrix met betrekking tot de snelheden niet-singulier is) en een willekeurig 'krachtenveld' ψ.

4.5 Hamilton's Principe

Neem nu aan dat we in het autonome (tijdsonafhankelijke) geval verkeren, dat het krachtenveld ψ in de bewegingsvergelijkingen van Lagrange (4.19) alleen van de positie y afhangt en conservatief is, in de zin dat er een 'potentiële energiefunctie' $W = W(y)$ is, waarvoor $\psi_i(y) = -\frac{\partial W(y)}{\partial y_i}$, $1 \leq i \leq p$. In dit geval kunnen we (4.19) herschrijven als $[Z]_i^\delta(t) = 0$, $1 \leq i \leq p$, als we noteren

$$Z(y, w) = S(y, w) - W(y). \tag{4.20}$$

Maar dit betekent precies dat δ een stationaire kromme is voor de actie-integraal van de functie Z. Deze uitspraak heet het *principe van Hamilton.* Hoewel Lagrange de Euler-Lagrange-vergelijkingen zeer goed kende, komt in [52] de uitspraak niet voor dat de bewegingsvergelijkingen van de klassieke mechanica equivalent zijn met het hebben van een stationaire kromme van de actie-integraal van Z. Misschien vond Lagrange dat het *verschil* van de kinetische en de potentiële energie niet voldoende fysische betekenis heeft, in vergelijking met de *som* = de totale energie van het systeem.

We merken nog op dat Lagrange de kinetische, resp. potentiële energie steeds met dezelfde letter T, resp. V noteerde, ook als hij deze beschouwde als functie van de nieuwe positie- en snelheidsvariabelen y en w. Het is gebruikelijk om $L = T - V$ te schrijven. Dan luidt Hamilton's principe dat $[L] = 0$ in het geval dat het krachtenveld conservatief is. Lagrange's stelling 4.1 impliceert dat Hamilton's principe in ieder coördinatensysteem geldig is. De voorwaarde dat het krachtenveld conservatief is hangt daarbij ook niet af van het gebruikte coördinatensysteem: is het krachtenveld conservatief in één coördinatensysteem, dan is het conservatief in ieder ander coördinatensysteem.

In onze notaties ziet het bovenstaande er als volgt uit. S is gelijk aan de kinetische energie, maar nu beschouwd als functie van de nieuwe positie- en snelheidsvariabelen y en w. Hierbij is $v = \mathrm{D}\Psi(y) \cdot w$ als $x = \Psi(y)$. Newton's bewegingsvergelijkingen $[T] = \phi$ krijgen in de nieuwe coördinaten de vorm $[S] = \psi$, waarin $\psi = \mathrm{D}\Psi(y)^{\mathrm{t}} \cdot \phi(x, v)$ het getransformeerde krachtenveld is, beschouwd als functie van y en w. Is het krachtenveld ϕ conservatief met potentiële energiefunctie $V = V(x)$, dan is ψ conservatief met potentiële energiefunctie gelijk aan $W = W(y) = V(\Psi(y))$. Stelling 4.1 impliceert dat de vergelijking $[L] = 0$ equivalent is met $[Z] = 0$, waarin $Z = S - W$ het verschil is van de kinetische en de potentiële energie, maar nu beschouwd als functie van de nieuwe positie- en snelheidsvariabelen y en w,

Een voor de hand liggende generalisatie van een conservatief mechanisch systeem is nu het tweede-orde stelsel dat bepaald wordt door de Euler-Lagrange-vergelijkingen (4.14) voor een willekeurige functie $L(t, x, v)$. Om een expliciet stelsel te krijgen, moeten we voor L eisen dat de tweede-orde afgeleidenmatrix met betrekking tot de snelheid niet-singulier is.

Deze relatie tussen variatieproblemen en klassieke mechanica werkt twee kanten op. Voor variatieproblemen betekent het dat de aandacht gericht wordt op de stroming in de faseruimte die door de Euler-Lagrange-vergelijkingen, nu immers opgevat als bewegingsvergelijkingen, geïnduceerd wordt.

Voor de klassieke mechanica betekent het dat soms existentie van oplossingen die aan bepaalde randvoorwaarden voldoen volgt uit de existentie van stationaire punten voor de actie-integraal, beschouwd als functie op de ruimte R der krommen die aan de randvoorwaarden voldoen. Bijvoorbeeld, zij K de verzameling der krommen in R waarvoor de actie-integraal kleiner of gelijk is aan een gegeven bovengrens C, die strikt groter is dan het infimum van de actie-integraal op R. Als K compact is met betrekking tot een geschikte topologie in de ruimte der krommen en als de actie-integraal continu is met betrekking tot deze topologie, dan kunnen we op de gebruikelijke manier concluderen dat de actie-integraal zijn minimum aanneemt in een $\gamma \in K$; dit is dan meteen een stationaire kromme voor de actie-integraal. Het probleem zit hem hier in de zinsnede 'compact met betrekking tot een geschikte topologie in de ruimte der krommen', want R is een oneindig--dimensionale ruimte en daarin is 'compact' een veel sterkere eis dan 'begrensd en gesloten'. Wij gaan hier niet verder in op de functionaalanalyse die nodig is om deze existentietheorie van de grond te brengen.

4.6 Constanten van Beweging

Het is een voor de hand liggende vraag of er voor algemene Euler-Lagrange-vergelijkingen $[L]^{\gamma}(t) = 0$ ook een constante van beweging is, die uitgedrukt kan worden in de Lagrange-functie L. Als we definiëren

$$H(t,\, x,\, v) := \sum_{i=1}^{n} \frac{\partial L(t,\, x,\, v)}{\partial v_i} \cdot v_i - L(t,\, x,\, v), \tag{4.21}$$

dan is gemakkelijk na te gaan dat voor iedere oplossing γ van de Euler-Lagrange-vergelijkingen geldt dat

$$\frac{d}{dt} H(t,\, \gamma(t),\, \gamma'(t)) = -\frac{\partial L(t,\, x,\, v)}{\partial t}|_{x=\gamma(t),\, v=\gamma'(t)}. \tag{4.22}$$

Hieruit lezen we meteen af:

Gevolg 4.3 *De functie H, die is gedefinieerd in (4.21), is een constante van beweging voor de Euler-Lagrange-vergelijkingen met Lagrange-functie L, dan en slechts dan als $L(t,\, x,\, v)$ niet expliciet van t afhangt.*

Andere constanten van beweging kunnen gevonden worden door toepassing van de volgende

Stelling 4.4 (E. Noether) *Zij $w(x)$ een vectorveld in de positieruimte met de eigenschap dat voor een éénparameter familie $x \mapsto \phi^s(x)$ van transformaties met*

$$\frac{\mathrm{d}}{\mathrm{d}s} \phi^s(x)|_{s=0} = w(x)$$

geldt dat

$$\frac{\mathrm{d}}{\mathrm{d}s} L\left(\phi^s(x),\, \mathrm{D}\phi^s(x) \cdot v\right)|_{s=0} = 0. \tag{4.23}$$

Dan is de functie

$$I_w(x,\, v) := \langle w(x),\, \frac{\partial L(x,\, v)}{\partial v} \rangle,$$

de w-component van de impuls, *een constante van beweging voor de oplossingen van de Euler-Lagrange-vergelijkingen gedefinieerd door de functie L.*

Bewijs Als $\gamma(t)$ een C^2 kromme in X is, dan geldt voor $\gamma_s(t) := \phi^s(\gamma(t))$ dat

$$\left(\gamma_s(t),\, \gamma_s'(t)\right) = \left(\phi^s(\gamma(t)),\, \mathrm{D}\phi^s(\gamma(t)) \cdot \gamma'(t)\right).$$

Hiermee geeft (4.23) dat

$$\begin{aligned} 0 &= \int_a^b \frac{\mathrm{d}}{\mathrm{d}s}|_{s=0}\, L\left(\gamma_s(t),\, \gamma_s'(t)\right)\, \mathrm{d}t \\ &= \frac{\mathrm{d}}{\mathrm{d}s}|_{s=0} \int_a^b L\left(\gamma_s(t),\, \gamma_s'(t)\right)\, \mathrm{d}t \\ &= -\int_a^b [L]^\gamma(t) \cdot w(\gamma(t))\, \mathrm{d}t + I_w\left(\gamma(b),\, \gamma'(b)\right) - I_w\left(\gamma(a),\, \gamma'(a)\right), \end{aligned}$$

waarbij we in de derde identiteit de variatieformule (4.13) hebben gebruikt. Dus, als $[L]^\gamma = 0$, dan is $I_w\left(\gamma(b),\, \gamma'(b)\right) = I_w\left(\gamma(a),\, \gamma'(a)\right)$, ofwel I_w is een constante van beweging. $\square$

In het algemeen, dus zonder de aanname dat $[L]^\gamma = 0$, leidt differentiatie met betrekking tot b in de voorgaande integraalformule tot

$$\frac{\mathrm{d}}{\mathrm{d}t} I_w\left(\gamma(t),\, \gamma'(t)\right) = \frac{\mathrm{d}}{\mathrm{d}s} L\left(\gamma_s(t),\, \gamma_s'(t)\right)|_{s=0} + \langle w(\gamma(t)),\, [L]^\gamma(t)\rangle. \tag{4.24}$$

Deze formule is equivalent met de variatieformule (4.13). Zij zegt dat voor een willekeurig vectorveld w de tijdsafgeleide van de w-component van de impuls gelijk is aan de afgeleide van L in het linkerlid van (4.23) plus de w-component van $[L]^\gamma$, de 'reactiekracht' die optreedt als γ geen oplossing is van de Euler-Lagrange-vergelijkingen.

De formule (4.24) en de conclusie van Stelling 4.4 zijn afkomstig van Emmy Noether [70]. Zij formuleerde haar resultaten voor een veel algemenere klasse van variatieproblemen, met integratie over meer dan één variabele (in plaats van t) en met een functie L waarin ook hogere orde afgeleiden van γ zijn toegestaan.

Merk op dat aan (4.23) is voldaan indien voor alle s, x, v geldt dat

$$L\left(\phi^s(x),\, \mathrm{D}\phi^s(x) \cdot v\right) = L(x,\, v).$$

Men zegt in dit geval dat L *invariant is onder de door de transformaties ϕ^s geïnduceerde transformaties in de faseruimte*, de positie-snelheidsruimte der $(x,\, v)$. Alle klassiek bekende stellingen over behoud van impuls-, resp. impulsmomentcoördinaten uit de mechanica zijn toepassingen van de stelling van Noether.

4.7 Homogene Functies van de Snelheid

Een functie $T(x,\, v)$ heet *homogeen van de graad m* met betrekking tot v als voor iedere $\tau > 0$ geldt dat $T(x,\, \tau\, v) = \tau^m\, T(x,\, v)$. Voor differentieerbare functies $v \mapsto T(x,\, v)$ is dit equivalent met de *differentiaalvergelijking van Euler*

$$\sum_{h=1}^{n} \frac{\partial T(x,\, v)}{\partial v_h} \cdot v_h = m\, T(x,\, v), \tag{4.25}$$

zoals gemakkelijk is na te gaan. Hieruit lezen we af dat als $L(x, v) = T(x, v) - V(x)$ en $v \mapsto T(x, v)$ is homogeen van de graad twee, zoals het geval is voor de klassieke kinetische energiefunctie, dan is de constante van beweging $H(x, v)$ gelijk aan $T(x, v) + V(x)$, de totale energie.

Is $v \mapsto T(x, v)$ homogeen van de graad twee en tweemaal continu differentieerbaar in $\mathbf{R}^n$, dan is er een eenduidig bepaalde symmetrische matrix $m_{ij}(x)$, waarvoor

$$T(x, v) = \frac{1}{2} \sum_{i,j=1}^{n} m_{ij}(x)\, v_i\, v_j. \tag{4.26}$$

Dit volgt door uitschrijven van de tweede-orde Taylor-ontwikkeling in $v = 0$, waarbij de homogeniteit ertoe leidt dat de restterm gelijk is aan nul. De eenduidigheid volgt uit de opmerking dat $m_{ij}(x)$ gelijk is aan de tweede-orde afgeleidenmatrix van T met betrekking tot v. Als boven opgemerkt vormen de Euler-Lagrange-vergelijkingen een expliciet tweede-orde stelsel, dan en slechts dan als deze matrix niet-singulier is. Een x-afhankelijke niet-singuliere symmetrische matrix heet ook wel een *pseudo-Riemann-structuur* in de positieruimte, men spreekt van een *Riemann-structuur* als de matrices positief definiet zijn, hetgeen correspondeert met $T(x, v) > 0$ zodra $v \neq 0$. Als T de interpretatie heeft van kinetische ergie, dan heet $m_{ij}(x)$ de (tensor van de) *trage massa*, omdat het de factor is die aangeeft hoeveel energie het kost om het systeem in beweging te brengen.

De Euler-Lagrange-vergelijkingen met T als Lagrange-functie beschrijven de *geodeten* van de (pseudo-)Riemann-structuur. Deze begrippen spelen een belangrijke rol in de differentiaalmeetkunde. Als $n = 4$ en de matrix $m_{ij}(x)$ heeft drie negatieve en één positieve eigenwaarde, dan spreken we van een *Lorentz-structuur* en zijn we in het kader van Einstein's relativiteitstheorie.

4.8 Stelsels met Inperkingen

Zij nu $p < n$ in de substitutie $x = \Psi(y, t)$, $y \in \mathbf{R}^p$, $x \in \mathbf{R}^n$. Dan is, voor iedere t, de afbeelding $y \mapsto \Psi(t, y)$ een immersie van een open deelverzameling van $\mathbf{R}^p$ naar $\mathbf{R}^n$, waarvan lokaal het beeld een p-dimensionale deelvariëteit van $\mathbf{R}^n$ is. In het algemeen zal deze afbeelding oplossingen van (4.19) *niet* overvoeren in oplossingen van de bewegingsvergelijkingen van Newton. In veel mechanische systemen komt het echter voor dat de posities op een deelvariëteit gehouden worden, door *reactiekrachten* die op gaan treden zodra het systeem van de deelvariëteit af beweegt. Het is nu een aantrekkelijke hypothese dat het *ingeperkte systeem* (engels: *system with constraints*) voldoet aan de bewegingsvergelijkingen van Lagrange in (4.19). In het geval van een conservatief krachtenveld schrijft men de kinetische energie minus de potentiële energie uit in de coördinaten y, resp. w voor positie en snelheid van het ingeperkte systeem en neemt men als bewegingsvergelijkingen $[Z]_i^\delta(t) = 0$, $1 \leq i \leq p$, met $Z(y, w) = S(y, w) - W(y)$.

Of dit werkelijk het correcte mathematische model is kan alleen beslist worden door het experiment of een gedetailleerde analyse van de aard van de reactiekrachten. In veel situaties blijken echter de bewegingsvergelijkingen van Lagrange een goed model te zijn voor het gedrag van ingeperkte systemen.

4.9 Continuümsmechanica

Een d-dimensionaal continu medium in de uitgangspositie heeft een continuüm aan punten x, die in de $\mathbf{R}^d$ kunnen variëren, de gebruikelijke dimensie hier is $d = 3$. Zij $u = u(t, x) \in \mathbf{R}^d$ de positie van het massapunt x ten tijde t. Het ligt voor de hand om hierbij aan te nemen dat het medium niet samengeklapt kan worden; een manier om dit uit te drukken is dat we aannemen dat $x \mapsto u(t, x)$ een C^r-diffeomorfisme is, waarin $r \geq 1$ nader gespecifieerd zal worden, afhankelijk van wat we nodig denken te hebben. Hiermee wordt de configuratieruimte (= de ruimte der posities) gelijk aan de ruimte van alle diffeomorfismen van $\mathbf{R}^d$ naar $\mathbf{R}^d$. Omdat ieder diffeomorfisme gegeven wordt door d reëelwaardige functies van d reële variabelen, en dergelijke functieruimten altijd oneindigdimensionaal zijn, worden we hier geconfronteerd met het probleem van een oneindigdimensionale configuratieruimte. De bewegingsvergelijkingen kunnen daarmee niet geschreven worden als een stelsel van tweede-orde gewone differentiaalvergelijkingen voor een eindig aantal van positiecoördinaten.

De bewegingsvergelijkingen in Lagrange's vorm kunnen echter wél op een voor de hand liggende manier gegeneraliseerd worden. Veronderstel dat het medium in het punt $x \in \mathbf{R}^d$ een *massadichtheid* $\rho(x)$ heeft. Noteren we $v(x) = \partial u(t, x)/\partial t$ voor de snelheid van het mediumpunt x, dan is de kinetische energie van het medium gelijk aan

$$T = T(v) := \int_{\mathbf{R}^d} \tfrac{1}{2}\rho(x)\, \|v(x)\|^2 \, \mathrm{d}x, \tag{4.27}$$

de integraal over alle x van de *kinetische energiedichtheid* $\rho(x)\, \|v(x)\|^2/2$ (met $\rho(x)$ als een dichtheidsfactor). Om dit op te kunnen schrijven is het nodig dat v een kwadratisch-integreerbare functie van x is. Merk op dat we hier in de relatief eenvoudige situatie verkeren dat de kinetische energie niet expliciet van de tijd of de positie afhangt, maar alleen van de snelheid. Anderzijds is, net als de positie u, de snelheid een afbeelding van $\mathbf{R}^d$ naar $\mathbf{R}^d$, dus zeker geen element van een eindigdimensionale lineaire ruimte.

Laten we nu aannemen dat op analoge manier er een potentiële energiefunctie is, die verkregen wordt door een potentiële energiedichtheidsfunctie te integreren over $x \in \mathbf{R}^3$. Hierbij nemen we aan dat de *potentiële energiedichtheid* W in het punt x niet alleen afhangt van de positie $u(x) = u(t, x)$ die het mediumpunt x inneemt, maar ook van deformaties van het lichaam in dat punt. Deze deformaties worden beschreven door de Taylor-ontwikkeling van de afbeelding u in het punt x. Laten we eenvoudigheidshalve aannemen dat de potentiële energiedichtheid afhangt van de Taylor-ontwikkeling van u tot en met de orde 1, dat wil zeggen van de positie $u(x)$ en van de eerste orde afgeleide $\mathrm{D}u(x)$ van u in het punt x, maar niet van hogere orde afgeleiden van u in het punt x. Dit leidt tot een potentiële energiefunctie

$$V = V(u) := \int_{\mathbf{R}^d} W(x,\, u(x),\, \mathrm{D}u(x))\, \mathrm{d}x, \tag{4.28}$$

waarin de potentiële energiedichtheidsfunctie W een reëelwaardige functie is op $\mathbf{R}^d \times \mathbf{R}^d \times \mathbf{R}^{d^2}$, waarin de laatste $\mathbf{R}^{d^2}$ staat voor de ruimte van lineaire afbeeldingen van $\mathbf{R}^d$ naar $\mathbf{R}^d$, geïdentificeerd met de ruimte van $d \times d$-matrices.

Deze aanpak is in de continuümsmechanica geïntroduceerd door Green (1855) en Thomson (1863), zie Thomson and Tait [85, Appendix C]. Daar wordt de extra aanname gemaakt dat de potentiële energiedichtheid wat betreft $A = \mathrm{D}u(x)$ alleen afhangt van de manier waarop het inproduct door deze lineaire transformatie verandert. Anders gezegd, de potentiële energiedichtheid is een functie van de symmetrische matrix $A^* \circ A$. Verder wordt alleen gekeken naar kleine uitwijkingen vanuit een evenwichtssituatie en daarvoor wordt de energiedichtheid vervangen door zijn Taylor-ontwikkeling tot en met de orde twee. Schrijven we $A = I + B$, met B klein, dan leidt dit tot een kwadratische vorm q in de symmetrische matrix $S = B^* + B$. De ruimte van symmetrische

$d \times d$-matrices heeft dimensie gelijk aan $d' = d(d+1)/2$, de ruimte van kwadratische vormen daarop heeft dimensie gelijk aan $d'' = d'(d'+1)/2$. Dit betekent dat er in de energiedichtheid d'' *elasticiteitscoëffiënten* zullen figureren. Voor $d = 3$ krijgen we $d' = 6$ en $d'' = 21$. In [85, Section 673] staat de opmerking dat "... an untenable theory (Boscovich's), falsely worked out by mathematicians, has led to relations among the coefficients of elasticity which experiment has proved to be false". (Één van die wiskundigen was Cauchy en het aantal relaties was zes.) Tenslotte kan nog vermeld worden dat als het medium *isotroop* is, dat wil zeggen $q(U\,S\,U^{-1}) = q(S)$ voor iedere orthogonale transformatie U, dan zijn er twee coëfficiënten a en b, waarvoor $q(S) = a\,(\text{spoor}\,S)^2 + b\,\text{spoor}(S^2)$. Voor $d = 3$ geeft dit de nogal drastisch reductie van 21 naar twee coëfficiënten.

Schrijven we $L(u, v) := T(v) - V(u)$ dan kunnen we kijken hoe de bewegingsvergelijkingen eruit gaan zien in de vorm van Hamilton's principe. Daartoe schrijven we op de *actie-integraal*

$$\begin{aligned}\mathcal{I}(\gamma) &:= \int_a^b L(\gamma(t), \gamma'(t))\,\mathrm{d}t \\ &= \int_a^b \int_{\mathbf{R}^d} \left\{ \tfrac{1}{2}\rho(x)\,\|\frac{\partial u(t, x)}{\partial t}\|^2 - W(x, u(t, x), \frac{\partial u(t, x)}{\partial x}) \right\} \mathrm{d}x\,\mathrm{d}t. \qquad (4.29)\end{aligned}$$

Merk op dat dit een integraal over de $(d+1)$-dimensionale (t, x)-ruimte wordt. Hiervan moeten we de variatie gelijk aan nul stellen, waarin we $u(t, x)$ variëren met $u(t, x) + \epsilon\,\phi(t, x)$, waarin $\phi(a, x) = \phi(b, x) = 0$. Schrijven we dit uit dan krijgen we dat voor iedere voldoend gladde afbeelding $\phi : \mathbf{R}^{d+1} \to \mathbf{R}^d$ (de eerste orde variatie in u) moet gelden dat

$$\begin{aligned}\int_{\mathbf{R}^{d+1}} \Bigg\{ \rho(x) \sum_i \frac{\partial u_i(t, x)}{\partial t}\,\frac{\partial \phi_i(t, x)}{\partial t} - \sum_i \frac{\partial W(x, u, A)}{\partial u_i}\Big|_{u=u(t, x),\, A=\partial u(t, x)/\partial x}\,\phi_i(t, x) \\ - \sum_{i, j} \frac{\partial W(x, u, A)}{\partial A_{i, j}}\Big|_{u=u(t, x),\, A=\partial u(t, x)/\partial x}\,\frac{\partial \phi_i(t, x)}{\partial x_j} \Bigg\}\,\mathrm{d}x\,\mathrm{d}t = 0. \qquad (4.30)\end{aligned}$$

Is u tweemaal continu differentieerbaar, dan kunnen we de afgeleiden van ϕ met behulp van partiële integratie overgooien en concluderen dat (4.30) voor alle ϕ geldt, dan en slechts dan als u voldoet aan het volgende stelsel van partiële differentiaalvergelijkingen.

$$\begin{aligned}&\frac{\partial}{\partial t}\,\rho(x)\,\frac{\partial u_i(t, x)}{\partial t} + \frac{\partial W(x, u, A)}{\partial u_i}\Big|_{u=u(t, x),\, A=\partial u(t, x)/\partial x} \\ &- \sum_j \frac{\partial}{\partial x_j}\left[\frac{\partial W(x, u(t, x), A)}{\partial A_{i, j}}\Big|_{u=u(t, x),\, A=\partial u(t, x)/\partial x}\right] = 0, \quad 1 \le i \le d. \qquad (4.31)\end{aligned}$$

Dit stelsel van partiële differentiaalvergelijkingen (van hetzelfde type als de *golfvergelijking*) kunnen we als Newton's vergelijkingen interpreteren door $\rho\,\partial u_i^2/\partial t^2$ als de i-de component van een versnellingsdichtheid te interpreteren, waarbij dan

$$-\frac{\partial W}{\partial u_i} + \sum_j \frac{\partial}{\partial x_j}\frac{\partial W}{\partial A_{ij}}$$

als de i-de component van een krachtdichtheid gezien wordt.

Lezers die met distributies vertrouwd zijn herkennen in (4.30) een *distributionele (zwakke)* vorm van het stelsel (4.31), waarbij de distributie in het linkerlid van (4.31) getest wordt met de testfunctie ϕ. Het voordeel van (4.30) is dat daarin voor de functie u alleen maar de existentie van de eerste orde afgeleiden geëist hoeft te worden, terwijl we in (4.31) de tweede orde afgeleiden nodig hebben. We gaan in dit college niet verder in op de theorie van stelsels van partiële differentiaalvergelijkingen zoals (4.31). Deze heeft heel wat voeten in de aarde. We volstaan met de mededeling dat de vergelijkingen (4.30), resp. (4.31), verkregen door toepassing van Hamilton's principe op de actie-integraal, in de fysica geaccepteerd worden als het correcte stelsel van bewegingsvergelijkingen wanneer de potentiële energiedichtheid alleen afhangt van x, $u(x)$ en $\partial u(t,\, x)/\partial x$. Hangt de potentiële energiedichtheid af van de Taylor-ontwikkeling van $x \mapsto u(t,\, x)$ tot en met de orde p, dan krijgt men een stelsel van partiële differentiaalvergelijkingen van de orde 2 in t, maar met afgeleiden tot en met de orde $2p$ met betrekking tot de x-variabelen. Soortgelijke variatieprincipes worden toegepast in de quantumveldentheorie.

Een klassiek leerboek, waarin onder ander de gebruikelijke terminologie is te vinden, is het boek van Love [60], de paragrafen 66, 68, 69, 115 liggen het dichtst bij het bovenstaande. Voor meer over de geschiedenis, zie [66].

4.10 Vraagstukken

Vraagstuk 4.1 Een *draaiing* of *rotatie* in $\mathbf{R}^3$ is een éénparameter familie $t \mapsto \mathrm{e}^{t\,Q}$ van lineaire transformaties, waarin Q een antisymmetrische lineaire afbeelding van $\mathbf{R}^3$ naar $\mathbf{R}^3$ is. Er geldt dat $d\,\mathrm{e}^{t\,Q}/dt = Q\,\mathrm{e}^{t\,Q} = \mathrm{e}^{t\,Q}$ en voor iedere $t \in \mathbf{R}$ is $\mathrm{e}^{t\,Q}$ een orthogonale lineare transformaties. De antisymmetrische afbeelding Q is ook te schrijven als

$$Q(x) = q \times x, \quad x \in \mathbf{R}^3,$$

voor een eenduidig bepaalde vector $q \in \mathbf{R}^3$.

Zij $T(t,\, x,\, v) = \frac{1}{2}\, m\, \langle v,\, v \rangle$ de gebruikelijke kinetische energie van een deeltje met trage massa $m > 0$ en beschouw de tijdsafhankelijke substitutie van variabelen $x = \mathrm{e}^{t\,Q}\, y$.

a) Bereken de kinetische energie $S(t,\, y,\, w)$ als functie van de tijd, de nieuwe positie y en de nieuwe snelheidsvector w. Ga na dat deze niet expliciet van t afhangt, dus dat we $S(y,\, w)$ mogen schrijven in plaats van $S(t,\, y,\, w)$.

b) Bereken $[S]^{\delta}(t)$. Laat zien dat

$$[S]^{\delta}(t) = m\,\delta''(t) + 2m\,Q\,\delta'(t) + m\,Q^2\,\delta(t).$$

Hierin heet minus de tweede term de *Coriolis-kracht* en minus de derde term de *centrifugale kracht*. De Coriolis-kracht werkt op dezelfde manier als een magnetisch veld op een geladen deeltje.

c) Neem nu aan dat op het deeltje een krachtveld werkt met een potentiële energie $V = V(x)$ die rotatie-symmetrisch is. Geef de bewegingsvergelijkingen in de nieuwe coördinaten, door de Euler-Lagrange-vergelijkingen op te schrijven voor de Lagrange-functie

$$Z(y,\, w) = S(y,\, w) - V(y).$$

Laat zien dat

$$K(y,\, w) = \sum_{i=1}^{3} \frac{\partial Z(y,\, w)}{\partial w_i} \cdot w_i - Z(y,\, w),$$

een constante van beweging is.

d) Bereken de totale energie $E = T + V$ als functie van y en w. Schrijf $E - K$ als functie van x en v en laat zien dat $E - K = I_Q$, de Q-component I_Q van de impuls, ook wel het *impulsmoment* van de draaiïng genoemd. Bewijs nogmaals, maar nu met behulp van de stelling van Noether, dat K een constante van beweging is.

$\oslash$

Vraagstuk 4.2 Geef de bewegingsvergelijkingen van Lagrange voor de beweging van een voorwerp op een kromme K in een verticaal vlak onder invloed van de zwaartekracht. Hierbij zal het handig blijken te zijn om de kromme K *met booglengte* te parametriseren, dat wil zeggen met behulp van een continu differentieerbare afbeelding $\gamma : I \to K$ van een interval I naar K, met de eigenschap dat voor iedere $s \in I$ geldt dat $\|\gamma'(s)\| = 1$. Dit heeft tot gevolg dat als $\mapsto \gamma(s(t))$ een beweging op K voorstelt, dan wordt de kinetische energie gegeven door de eenvoudige formule $T = m\,\frac{ds(t)}{dt}^2/2$. Als K niet met booglengte geparametriseerd wordt, dan komt er nog een s-afhankelijke factor in de kinetische energie, die de bewegingsvergelijkingen aanmerkelijk compliceert.

Laat zien dat deze equivalent zijn met de uitspraak dat we, bij de ontbinding van de zwaartekracht in een vector die raakt aan K en één die loodrecht staat op K, de eerste gebruiken om de versnelling mee te bepalen. Men zegt in dit geval dat de component loodrecht op K opgeheven wordt door de reactiekracht die het deeltje op K houdt. $\oslash$

Vraagstuk 4.3 Een koord ter lengte l beweegt op de reële as met het linker uiteinde in $x = x(t)$. Het is omgeslagen in het punt $x + y$, waarbij $0 < y = y(t) < l$, waarmee het andere uiteinde terecht komt in $(x + y) - (l - y)$. Verder nemen we aan dat het koord een homogene massadichtheid ρ heeft, waarmee de massa van het niet-omgeslagen, resp. omgeslagen stuk gelijk is aan $\rho\, y$, resp. $\rho\,(l - y)$. Het koord is willekeurig buigzaam bij het omslagpunt, maar niet rekbaar.

Bereken de kinetische energie T van het koord, als functie van x, y, $\dot{x} = \frac{\mathrm{d}x(t)}{\mathrm{d}t}$ en $\dot{y} = \frac{\mathrm{d}y(t)}{\mathrm{d}t}$. Geef de bewegingsvergelijkingen volgens Lagrange, aannemend dat er verder geen krachten op het koord werken. Merk op dat er behalve de T nog een impulscomponent p als constante van beweging optreedt. Bewijs dat $c = y(t)\,(l - y(t))\,y'(t)^2$ constant is en bereken c in termen van T en p. Laat zien dat ófwel y constant is, ófwel y' heeft een strikt positieve ondergrens, ófwel y' heeft een strikt negatieve bovengrens. Als y niet constant is, dan is de oplossing gedefinieerd op een begrensd interval $I =]a, b[$. Als $y' > 0$ dan geldt dat $y(t) \uparrow l$ en $y'(t) \to \infty$ als $t \uparrow b$ en $y(t) \downarrow 0$ en $y'(t) \to \infty$ als $t \downarrow a$. Als $y' < 0$ dan geldt dat $y(t) \downarrow 0$ en $y'(t) \to -\infty$ als $t \uparrow b$ en $y(t) \uparrow l$ en $y'(t) \to -\infty$ als $t \downarrow a$. Bereken de tijd die nodig is om van $y = y_-$ naar $y = y_+$ te komen. (Hint: $z\,\left(1 - z^2\right)^{1/2} + \arcsin z$ is een primitieve van de functie $2\left(1 - z^2\right)^{1/2}$.) Bereken de lengte van het tijdsinterval I in termen van ρ, l, T en p.

Dit voorbeeld kan gezien worden als een sterk vereenvoudigd model voor een zweep. Het verschijnsel dat $y'(t) \to \pm\infty$ verklaart de grote snelheid die optreedt bij het afglijden van het korte gedeelte. Het bereiken van de geluidssnelheid veroorzaakt een scherpe knal, het ‘klappen van de zweep’.

Zij $p = 0$. Zij T_1, resp. T_2 de kinetische energie van het niet-omgeslagen, resp. het omgeslagen deel van het koord ter lengte y, resp. $l - y$. Bewijs dat $T_2/T_1 = y/(l - y)$, hetgeen naar 0, resp. ∞ gaat als $y \downarrow 0$, resp. $y \uparrow l$. Dit impliceert dat alle kinetische energie in het korte eind zit als $t \downarrow a$, resp. $t \uparrow b$.

$\oslash$

Vraagstuk 4.4 Neem aan dat de massadichtheid ρ constant is en dat de potentiële energiedichtheid van de vorm

$$W(A) = \frac{1}{2}\,\lambda\;(\mathrm{spoor}\,A)^2$$

is, waarin λ een andere positieve constante is. Definieer

$$U(t,\,x) = \sum_{j=1}^{n} \frac{\partial u_j(t,\,x)}{\partial x_j},$$

de *divergentie* van het vectorveld $x \mapsto u(t,\,x)$. Laat zien dat (4.31) impliceert dat U voldoet aan de scalaire golfvergelijking

$$\rho\,\frac{\partial^2 U}{\partial t^2} = \lambda\,\Delta U,$$

waarin

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial {x_j}^2}$$

de *Laplace-operator* voorstelt. $\oslash$

5 Differentiaalmeeetkunde

Tot nu toe zijn snelheidsvectoren, versnellingen en krachtvelden steeds ingevoerd in termen van coördinaten, waarbij we in Hoofdstuk 4 bekeken hebben hoe de formules transformeren onder willekeurige substituties van de positiecoördinaten. Nu is het gebruik van coördinaten in de meetkunde een relatief late ontwikkeling: in de klassieke Euclidische meetkunde redeneert men met objecten als punten, rechte lijnen, vlakken, afstanden en dergelijke op een directe wijze, zonder eerst de ruimte door middel van coördinaten met een $\mathbf{R}^n$ te hebben geïdentificeerd. In de differentiaalmeetkunde is het ook één van de thema's om begrippen te introduceren waarmee op een coördinaatonafhankelijke manier geredeneerd kan worden. Omdat in de klassieke mechanica sedert Galilei en Newton van afgeleiden gebruik gemaakt wordt, willen we in ieder geval over snelheidsvectoren, krachtvelden en afgeleiden van transformaties op een coördinaatonafhankelijke manier kunnen praten. Bij eerste kennismaking doet de theorie nogal abstract aan, maar toch is dit het natuurlijke kader waarin bijna alle differentiaalmeetkundige beschouwingen tegenwoordig worden gegeven. Dit is ook het geval voor een toenemend deel van de beschouwingen in de theoretische natuurkunde. We geven hier een korte samenvatting van de basisbegrippen. Voor meer details, zie bijvoorbeeld [19] of de leerboeken die in de literatuurlijst van [19] zijn genoemd.

5.1 Variëteiten

Zij X een verzameling. Een *n-dimensionale kaart of coördinatisering* voor X is een bijectieve afbeelding κ, van een deelverzameling X_κ van X, naar een open deelverzameling V_κ van $\mathbf{R}^n$. Voor $x \in X_\kappa$ heten $\kappa(x)_j$, $1 \leq j \leq n$ de *coördinaten* van het punt x in de kaart κ, de verzameling X_κ heet de bijbehorende *coördinaatomgeving* in X. De reëelwaardige functies $\kappa_j : x \mapsto \kappa(x)_j$ op X_κ heten de *coördinaatsfuncties*, het geven van een n-dimensionale kaart is dus equivalent met het geven van een n-tal reëelwaardige functies, met de eigenschap dat er een open deelverameling V_κ van $\mathbf{R}^n$ is, waarvoor geldt dat $y \in V_\kappa$ dan en slechts dan als er precies één $x \in X$ is met $\kappa_j(x) = y_j$ voor iedere $1 \leq j \leq n$.

Als de $x \in X$ de *toestanden van een fysisch systeem* voorstellen dan kan men de κ_j als *metingen* interpreteren. Dat κ injectief is betekent dan dat een toestand in X_κ eenduidig is vastgelegd door de n meetresultaten $\kappa_j(x)$, $1 \leq j \leq n$. Dat $V_\kappa := \kappa\left(X_\kappa\right)$ open in $\mathbf{R}^n$ is, betekent dat iedere kleine storing van de meetresultaten weer de meetresultaten vormen van een toestand in X_κ van het systeem. Dat X_κ niet de gehele X hoeft te zijn, correspondeert met de situatie meetinstrumenten niet altijd voor alle toestanden van het systeem bruikbaar zijn, ze hebben vaak een bepaald natuurlijk bereik.

De naam 'kaart' is uit de aardrijkskunde afkomstig, waarbij gedeelten van het aardoppervlak afgebeeld worden op gewoonlijk een rechthoek in het vlak. De afbeelding wordt hier gesuggereerd door de markeringen in de rechthoek die toegevoegd worden aan bepaalde markante plaatsen op het aardoppervlak. Dergelijk kaarten werden ook gemaakt en gebruikt door mensen die (nog) geen voorstelling van het totale aardoppervlak als boloppervlak hadden.

Het idee van lokale coördinatiseringen wordt echter ook in de wiskunde in talloze situaties toegepast. Wij zijn dit al tegengekomen als X een n-dimensionale C^k-deelvariëteit van $\mathbf{R}^N$ is. Als $N = n$ dan betekent dit dat X een open deelverzameling van $\mathbf{R}^n$ is en dan kunnen we $X_\kappa = X = V_\kappa$ nemen en κ gelijk aan de identiteit. Als $n < N$ dan hebben we bij iedere $x \in X$ een open omgeving U van x in $\mathbf{R}^N$, een open deelverzameling V van $\mathbf{R}^n$ en een C^k-inbedding $\psi : V \to \mathbf{R}^N$, met $\psi(V) = X \cap U$. In dit geval is ψ bijectief van V naar $\psi(V)$, met de inverse $\kappa : \psi(V) \to V$ als coördinatisering; in dit geval is $X_\kappa = X \cap U$ en $V_\kappa = V$. Is $X \cap U$ de grafiek in $\mathbf{R}^n \times \mathbf{R}^{N-n}$ van een C^k-afbeelding $h : V \to \mathbf{R}^{N-n}$, dan is $\psi(y) = (y,\, h(y))$, dus κ is gelijk aan de beperking tot $X \cap U$ van de projectie $x = (y,\, z) \mapsto y$.

Een *n-dimensionale* C^r*-atlas* voor X is een collectie $\mathcal{A}$ van n-dimensionale kaarten voor X met de volgende twee eigenschappen.

i) Voor iedere $x \in X$ is er een $\kappa \in \mathcal{A}$ waarvoor $x \in X_\kappa$. Anders gezegd, de kaartomgevingen in X voor de kaarten uit de atlas vormen een overdekking van X.

ii) Als $\kappa \in \mathcal{A}$ en $\lambda \in \mathcal{A}$, dan is
$$V_{\kappa,\,\lambda} := \kappa\,(X_\kappa \cap X_\lambda)$$
een open deelverzameling van $\mathbf{R}^n$ en $\lambda \circ \kappa^{-1}$ is een C^r-afbeelding van $V_{\kappa,\,\lambda}$ naar $\mathbf{R}^n$.

Merk op dat we in ii) de rol van κ en λ ook kunnen verwisselen, hetgeen leidt tot de conclusie dat $\kappa \circ \lambda^{-1}$, welke afbeelding de inverse is van $\lambda \circ \kappa^{-1}$, een C^r-afbeelding is van $V_{\lambda,\,\kappa}$ naar $\mathbf{R}^n$. Daarmee is $\lambda \circ \kappa^{-1}$ een C^r-diffeomorfisme van $V_{\kappa,\,\lambda}$ naar $V_{\lambda,\,\kappa}$, met $\kappa \circ \lambda^{-1}$ als inverse. Men noemt $\lambda \circ \kappa^{-1}$ de *verkaarting* of de *coördinatentransformatie van de kaart κ naar de kaart λ.*

In het geval van een C^r-deelvariëteit X van $\mathbf{R}^N$ voldoen de $\kappa = \psi^{-1}$, met ψ een C^r-inbedding van een open deelverzameling V van $\mathbf{R}^n$ naar een open deel van X, aan i) en ii), dus vormen deze een C^r-atlas voor X.

Zij $\mathcal{A}$ een C^r-atlas voor X. Een deelverzameling U van X heet *open met betrekking tot de atlas* $\mathcal{A}$, als voor iedere $\kappa \in \mathcal{A}$ geldt dat $\kappa\,(U \cap X_\kappa)$ een open deelverzameling van $\mathbf{R}^n$ is. De collectie van deze open verzamelingen is een topologie in X, die de *door de atlas $\mathcal{A}$ in X gedefinieerde topologie* heet.

In het geval van een C^r-deelvariëteit X van $\mathbf{R}^N$ is de door de hierboven vermelde atlas gedefinieerde topologie gelijk aan de topologie die X als deelverzameling van $\mathbf{R}^N$ al had.

Een topologie $\mathcal{O}$ in een verzameling X heet *Hausdorff's* als er voor iedere $x,\, y \in X$ met $x \neq y$ verzamelingen $A,\, B \in \mathcal{O}$ zijn, waarvoor $x \in A$, $y \in B$ en $A \cap B = \emptyset$. De gebruikelijke topologie in $\mathbf{R}^N$ is Hausdorff's, dus die in een deelverzameling X van $\mathbf{R}^N$ is dat ook. Het is nu een kleine schrik dat in het algemeen de atlastopologie niet Hausdorff's hoeft te zijn. Als voorbeeld nemen we de 'splitsing van één spoor naar twee sporen', die als volgt kan worden gedefinieerd. We nemen $X = L \cup R_+ \cup R_-$, waarin
$$L = \{(x,\, 0) \in \mathbf{R}^2 \mid x < 0\}, \quad R_\pm = \{(x,\, \pm 1) \in \mathbf{R}^2 \mid x \geq 0\}.$$
Als kaarten nemen we de bijectieve afbeeldingen $\kappa_\pm$ van $L \cup R_\pm$ naar $\mathbf{R}$, gedefinieerd door $\kappa_\pm(x,\, \pm 1) = x$, $x \in \mathbf{R}$. Deze vormen een reëel-analytische atlas voor X. Echter, als $A_\pm$ een omgeving is van $(0,\, \pm 1)$ in X ten aanzien van de atlastopolgie, dan is er een altijd een $\epsilon > 0$ met de eigenschap dat $]-\epsilon,\, 0[\ \times\ \{0\} \subset A_+ \cap A_-$.

Het is niet moeilijk om na te gaan dat de atlastopologie Hausdorff's is, dan en slechts dan als voor iedere $\kappa,\, \lambda \in \mathcal{A}$ er géén rij $x(j) \in X_\kappa \cap X_\lambda$ is, waarvoor de rij $\kappa(x(j))$ voor $j \to \infty$ convergeert naar een punt van $V_\kappa \setminus V_{\kappa,\,\lambda}$ terwijl de rij $\lambda(x(j))$ voor $j \to \infty$ convergeert naar een punt van $V_\lambda \setminus V_{\lambda,\,\kappa}$.

Een kaart λ voor X heet *compatibel met de atlas* $\mathcal{A}$, als voor iedere $\kappa \in \mathcal{A}$ geldt dat de boven ingevoerde verzamelingen $V_{\kappa,\,\lambda}$ en $V_{\lambda,\,\kappa}$ open in $\mathbf{R}^n$ zijn en $\lambda \circ \kappa^{-1}$ een C^r-diffeomorfisme is van $V_{\kappa,\,\lambda}$ naar $V_{\lambda,\,\kappa}$. Men noemt de n-dimensionale C^r-atlas $\mathcal{B}$ voor X compatibel met $\mathcal{A}$ als iedere $\lambda \in \mathcal{B}$

compatibel is met $\mathcal{A}$. Zij nu $\mathcal{M}$ de collectie van *alle* kaarten λ voor X die compatibel zijn met $\mathcal{A}$. Dan is $\mathcal{M}$ een n-dimensionale C^r-atlas voor X die compatibel is met $\mathcal{A}$. Deze atlas is *maximaal* in de zin dat als λ compatibel is met $\mathcal{M}$, dan is $\lambda \in \mathcal{M}$. Is $\mathcal{B}$ een atlas, resp. maximale atlas die compatibel is met $\mathcal{A}$, dan is $\mathcal{B} \subset \mathcal{M}$, resp. $\mathcal{B} = \mathcal{M}$. Hiermee is $\mathcal{M}$ *de* maximale atlas die compatibel is met $\mathcal{A}$.

Een *n-dimensionale* C^r*-variëteit* is nu gedefinieerd als een paar $(X,\, \mathcal{A})$ waarin X een verzameling is en $\mathcal{A}$ een maximale n-dimensionale C^r-atlas voor X, met de eigenschap dat de door $\mathcal{A}$ in X gedefinieerde topologie Hausdorff's is. Als we de laatste eis laten vallen, dan spreekt men van een *niet-Hausdorff'se n-dimensionale* C^r*-variëteit.* Verder spreekt men, als het duidelijk is welke atlas bedoeld wordt, over *de variëteit X* in plaats van over 'de variëteit $(X,\, \mathcal{A})$'.

Zoals hiervoor is beschreven, kunnen we met een kleinere atlas starten en daar dan de maximale atlas bij nemen. Bij het invoeren van een variëteit beperkt men zich vaak tot de beschrijving van een minimale atlas; de 'spoorsplitsing' was hiervan een (niet-Hausdorff's) voorbeeld. Het geeft vervolgens maximale flexibiliteit om over te gaan op de maximale atlas. Ook maakt de eenduidigheid van de maximale atlas dat we verlost zijn van de enorme keuzevrijheid in compatibele atlassen.

Er zijn diverse constructies om bij gegeven variëteiten nieuwe variëteiten te maken. Zo wordt iedere open deelverzameling U van X een n-dimensionale C^r-variëteit, door als kaarten te nemen de beperkingen to $U \cap X_\kappa$ van de $\kappa \in \mathcal{A}$.

Een wellicht wat interessantere constructie is die van het *Cartesische product* $X \times Y$ van de n-dimensionale C^r-variëteit $(X,\, \mathcal{A})$ en de p-dimensionale variëteit $(Y,\, \mathcal{B})$. Hiervoor neemt men de kaarten

$$\kappa \times \lambda : (x,\, y) \mapsto (\kappa(x),\, \lambda(y)) : X_\kappa \times Y_\lambda \to \mathbf{R}^n \times \mathbf{R}^p,$$

waarin $\kappa \in \mathcal{A}$ en $\lambda \in \mathcal{B}$. Deze vormen een (niet-maximale!) $(n+p)$-dimensionale C^r-atlas voor $X \times Y$, met Hausdorff'se atlastopologie. Met de bijbehorende maximale atlas wordt dan $X \times Y$ een $(n+p)$-dimensionale C^r-variëteit.

Herhaald toepassen van deze definitie leidt met volledige inductie over d tot een d-voudig Cartesisch product

$$X_1 \times \ldots \times X_{d-1} \times X_d = (X_1 \times \ldots \times X_{d-1}) \times X_d.$$

Een bekend voorbeeld is de *d-dimensionale torus T^d*, die onstaat als we voor iedere X_i de cirkel C nemen. Denken we C als geparametriseerd door een periodieke variabele (een hoekvariabele), dan is T^d de natuurlijke toestandsruimte voor multi-periodieke verschijnselen, beschreven door d hoekvariabelen.

5.2 Afbeeldingen tussen Variëteiten

Is X, resp. Y een C^r-variëteit van dimensie n, resp. p, zeggen we voor een afbeelding $f : X \to Y$ dat $f \in \mathrm{C}^r$ als f continu is met betrekking tot de atlastopologieën en voor iedere kaart κ voor X en iedere kaart λ voor Y geldt dat $\lambda \circ f \circ \kappa^{-1}$ een C^r-afbeelding is van

$$V = \kappa\left(X_\kappa \cap f^{-1}\left(Y_\lambda\right)\right)$$

naar $\mathbf{R}^p$. Analoog zegt men dat f differentieerbaar is in het punt $a \in X$ als de bovenstaande afbeelding van $\mathbf{R}^n$ naar $\mathbf{R}^p$ differentieerbaar is in het punt $\kappa(a)$. Uiteraard is hierin de uitdrukking $f \in \mathrm{C}^r$ een afkorting voor f is *r keer continu differentieerbaar*, resp. *willekeurig vaak differentieerbaar = glad*, resp. *reëel-analytisch*, als $1 \leq r < \infty$, resp. $r = \infty$, resp. $r = \omega$.

Merk op dat de continuïteit van f impliceert dat V een open deelverzameling is van $\mathbf{R}^n$, de definitie zegt niets anders dan dat 'f in lokale coördinaten gelijk is aan een C^r-afbeelding van een open deelverzameling van $\mathbf{R}^n$ naar $\mathbf{R}^p$'. Merk op dat het hierbij voldoende is om de voorwaarde te verifiëren voor κ, resp. λ in *een* (minimale) compatibele atlas voor X, resp. Y, de kettingregel geeft dat dan vanzelf aan de voorwaarde is voldaan voor alle kaarten in de maximale atlas voor X, resp. Y.

Deze opmerking is vooral relevant als X een open deelverzameling is van $\mathbf{R}^n$ en/of Y een open deelverzameling is van $\mathbf{R}^p$, in welk geval we kunnen volstaan met de identiteit als enige kaart in X, resp. Y. Als beiden open deelverzamelingen van een Euclidische ruimte zijn, dan komt de definitie van $f \in \mathrm{C}^r$ overeen met de oude. Zelfs als één van beide een lager-dimensionale deelvariëteit is van een Euclidische ruimte, dan is de de definitie echter al nieuw.

Is X een open interval in $\mathbf{R}$, dan is $f \in \mathrm{C}^r$ dan en slechts dan als f continu is en voor iedere kaart λ in een deelatlas voor Y geldt dat $\lambda \circ f$ een C^r-afbeelding is van $X \cap f^{-1}(Y_\lambda)$ naar $\mathbf{R}^p$; men noemt in dit geval f een C^r*kromme in* Y.

Is anderzijds $Y = \mathbf{R}$, dan is f een reëelwaardige functie op X en we krijgen dat deze C^r is, dan en slechts dan als f continu is en voor iedere kaart κ in een deelatlas voor X geldt dat $f \circ \kappa$ een reëelwaardige C^r-functie is op de open deelverzameling $V_\kappa = \kappa(X_\kappa)$ van $\mathbf{R}^n$.

5.3 Raakvectoren

Zij $1 \leq r \leq \omega$, X een n-dimensionale C^r-variëteit en $a \in X$. Net als voor deelvariëteiten van een Euclidische ruimte willen we de raakvectoren aan X in het punt a identificeren met de snelheidsvectoren $\gamma'(0)$ van differentieerbare krommen γ in X met $\gamma(0) = a$. We hebben hierboven net gedefinieerd wat het betekent dat γ een differentieerbare kromme in X is, namelijk dat voor iedere kaart $\kappa \in \mathcal{A}$ met $a \in X_\kappa$ de afbeelding $\kappa \circ \gamma$ een differentieerbare kromme in $\mathbf{R}^n$ is. Daarvan is de snelheidsvector

$$v_\kappa := (\kappa \circ \gamma)'(0) \in \mathbf{R}^n$$

een vector in $\mathbf{R}^n$, die echter van de keuze van de kaart κ afhangt en wel op de volgende manier. Is ook $\lambda \in \mathcal{A}$ en $a \in X_\lambda$, dan is

$$\lambda \circ \gamma = \left(\lambda \circ \kappa^{-1}\right) \circ (\kappa \circ \gamma),$$

dus toepassing van de de kettingregel op de samenstelling van de coördinatentransformatie $\Psi := \lambda \circ \kappa^{-1}$ in $\mathbf{R}^n$ met de kromme $\kappa \circ \gamma$ in $\mathbf{R}^n$ geeft dat

$$v_\lambda = \mathrm{D}\left(\lambda \circ \kappa^{-1}\right)(\kappa(a)) \cdot v_\kappa \tag{5.1}$$

Uitgeschreven in coördinaten betekent dit dat

$$v_{\lambda,\,h} = \sum_{i=1}^{n} \frac{\partial \Psi_h(y)}{\partial y_i}\Big|_{y=\kappa(a)} \cdot v_{\kappa,\,i}. \tag{5.2}$$

In de terminologie van Ricci en Levi-Civita, *snelheidsvectoren transformeren contravariant onder substituties van variabelen.*

Wanneer men een grootheid heeft die afhangt van een of andere keuze, dan is een standaardmanier om zich van de plicht om te kiezen te bevrijden om de grootheid te beschouwen als een functie van deze keuze, zonder daarbij deze keuze in te perken. Dit leidt tot de volgende definitie. Zij $\mathcal{A}_a$

de collectie van alle $\kappa \in \mathcal{A}$ waarvoor $a \in X_\kappa$. Een *raakvector aan X in het punt a* is gedefinieerd als een afbeelding

$$v : \kappa \mapsto v_\kappa : \mathcal{A}_a \mapsto \mathbf{R}^n$$

zodanig dat voor iedere $\kappa, \lambda \in \mathcal{A}_a$ de transformatieformule (5.1) geldt. De verzameling van alle raakvectoren aan X in het punt a heet de *raakruimte aan X in het punt a* en wordt met $\mathrm{T}_a X$ aangeduid.

De transformatieformule (5.1) impliceert dat als $\kappa \in \mathcal{A}_a$, dan ligt voor iedere $\lambda \in \mathcal{A}_a$ de vector v_λ vast in termen van de vector v_κ. Anders gezegd, de afbeelding

$$\mathrm{D}_a \kappa : v \mapsto v_\kappa : \mathrm{T}_a X \to \mathbf{R}^n \tag{5.3}$$

is injectief. Zij is ook surjectief omdat voor iedere $w \in \mathbf{R}^n$ de formule

$$v_\lambda := \mathrm{D}\left(\lambda \circ \kappa^{-1}\right)(\kappa(a)) \cdot w$$

een $v \in \mathrm{T}_a V$ definieert (gebruik de kettingregel voor het bewijs!), waarvoor $v_\kappa = w$.

We kunnen de bijectieve afbeelding $\mathrm{D}_a \kappa : \mathrm{T}_a X \to \mathbf{R}^n$ gebruiken om $\mathrm{T}_a X$ met $\mathbf{R}^n$ te identificeren en in het bijzonder om van $\mathrm{T}_a X$ een n-dimensionale lineaire ruimte te maken met de optelling en scalairvermenigvuldiging gedefinieerd door

$$\begin{aligned} v +_\kappa w &:= \mathrm{D}_a \kappa^{-1} \left(\mathrm{D}_a \kappa(v) + \mathrm{D}_a \kappa(w)\right), \\ c \cdot_\kappa v &:= \mathrm{D}_a \kappa^{-1} \left(c \cdot \mathrm{D}_a \kappa(v)\right). \end{aligned}$$

In principe hangt deze definitie af van de keuze van $\kappa \in \mathcal{A}_a$. Echter, uit (5.1) lezen we af dat

$$\mathrm{D}_a \lambda \circ (\mathrm{D}_a \kappa)^{-1} = \mathrm{D}\left(\lambda \circ \kappa^{-1}\right)(\kappa(a)) \tag{5.4}$$

en het feit dat dit een lineaire afbeelding is van $\mathbf{R}^n$ naar $\mathbf{R}^n$ is equivalent met de uitspraak dat $+_\lambda = +_\kappa$ en $\cdot_\lambda = \cdot_\kappa$.

Hiermee wordt voor iedere $a \in X$ de raakruimte $\mathrm{T}_a X$ een n-dimensionale lineaire ruimte, waarbij voor iedere $\kappa \in \mathcal{A}_a$ de afbeelding $\mathrm{D}_a \kappa$ een lineair isomorfisme is van $\mathrm{T}_a X$ naar $\mathbf{R}^n$. De raakruimten $\mathrm{T}_a X$ worden daarbij voor verschillende $a \in X$ als *disjuncte* lineaire ruimten opgevat, hetgeen ze volgens de definitie inderdaad zijn.

Als X een open deelverzameling is van $\mathbf{R}^n$, dan correspondeert dit met het gebruik om een vector $v \in \mathbf{R}^n$, als die de rol vervult van raakvector aan $\mathbf{R}^n$ in het punt a, te tekenen als een pijltje met het voetpunt in a en de punt in $a + v$. Anderzijds gebruikt men echter dikwijls de afbeelding (5.3), met κ gelijk aan de identiteit, om alle raakruimten aan X met $\mathbf{R}^n$ te identificeren. Anders gezegd, in $\mathbf{R}^n$ kunnen we kiezen: ofwel raakvectoren in verschillende punten als verschillend beschouwen, ofwel 'parallelle' raakvectoren met elkaar identificeren.

Voor iedere differentieerbare kromme γ in X met $\gamma(0) = a$ definieert $v_\kappa = (\kappa \circ \gamma)'(0)$ nu een element van $\mathrm{T}_a X$, dat met $\gamma'(0)$ wordt aangeduid. In formule:

$$\mathrm{D}_a \kappa\left(\gamma'(0)\right) = \gamma'(0)_\kappa = (\kappa \circ \gamma)'(0).$$

Dit kan gelezen worden als een kettingregel voor de afgeleide van $\kappa \circ \gamma$. Het als disjunct opvatten van de raakruimten in verschillende punten betekent dat we de snelheidsvectoren $\gamma'(0)$ en $\delta'(0)$ van de krommen γ en δ in X nooit met elkaar identificeren als $\gamma(0) \neq \delta(0)$, zelfs niet als ze beiden gelijk zijn aan het nulelement in hun respectievelijke raakruimten. Dit onder het motto dat $\gamma(t)$ en $\delta(t)$ pas op elkaar gaan lijken voor t dichtbij 0 als in eerste instantie $\gamma(0 = \delta(0)$, ze hebben vervolgens nog grotere overeenkomst als in dat geval óók nog geldt dat $\gamma'(0) = \delta'(0)$.

5.4 De Raakafbeelding

Zij X en Y differentieerbare variëteiten en f een afbeelding van X naar Y, die differentieerbaar is in het punt $a \in X$. Dan is er een eenduidig bepaalde afbeelding $\mathrm{T}_a f$ van $\mathrm{T}_a X$ naar $\mathrm{T}_{f(a)} Y$, de *raakafbeelding van f in het punt a* genaamd met de eigenschap dat voor iedere differentieerbare kromme γ in X met $\gamma(0) = a$ geldt dat

$$(f \circ \gamma)'(0) = \mathrm{T}_a f\left(\gamma'(0)\right). \tag{5.5}$$

Is κ, resp. λ een kaart in een omgeving van a, resp. $f(a)$ in X, resp. Y, dan is

$$\mathrm{T}_a f = \left(\mathrm{D}_{f(a)} \lambda\right)^{-1} \circ \mathrm{D}\left(\lambda \circ f \circ \kappa^{-1}\right)(\kappa(a)) \circ \mathrm{D}_a \kappa, \tag{5.6}$$

waaruit we ook aflezen dat $\mathrm{T}_a f$ een *lineaire* afbeelding is van de lineaire ruimte $\mathrm{T}_a X$ naar de lineaire ruimte $\mathrm{T}_{f(a)} Y$.

De *kettingregel* voor differentieerbare afbeeldingen tussen variëteiten luidt nu als volgt. Zij X, Y, Z differentieerbare variëteiten, $f : X \to Y$ differentieerbaar in het punt $a \in X$ en $g : Y \to Z$ differentieerbaar in het punt $f(a) \in Y$. Dan is $g \circ f$ differentieerbaar in het punt a en

$$\mathrm{T}_a(g \circ f) = \mathrm{T}_{f(a)}\, g \circ \mathrm{T}_a f. \tag{5.7}$$

Dit volgt door toepassing van de kettingregel op

$$\mu \circ (g \circ f) \circ \kappa^{-1} = \left(\mu \circ g \circ \lambda^{-1}\right) \circ \left(\lambda \circ f \circ \kappa^{-1}\right)$$

en (5.6) te gebruiken. Hierin zijn κ, resp. λ, resp. μ kaarten voor X, resp. Y, resp. Z. Merk op dat het disjunct houden van de raakruimten in verschillende punten hier helpt bij het onthouden van de regel dat g gedifferentieerd moet worden in het punt $f(a)$: $\mathrm{T}_b\, g$ kan alleen uitgevoerd worden na $\mathrm{T}_a f$ als de vectoren $\mathrm{T}_a f \cdot v \in \mathrm{T}_{f(a)} Y$ in $\mathrm{T}_b Y$ liggen, hetgeen alleen maar kan als $b = f(a)$.

Alle begrippen die in Hoofdstuk 2 in open deelverzamelingen van een Euclidische ruimte zijn ingevoerd kunnen naar variëteiten gegeneraliseerd worden. Een C^r*-diffeomorfisme* van een C^r--variëteit X naar een C^r-variëteit Y is een bijectieve afbeelding f van X naar Y, waarvoor zowel $f \in \mathrm{C}^r$ als $f^{-1} \in \mathrm{C}^r$. Dit impliceert dat voor iedere $x \in X$ de raakafbeelding $\mathrm{T}_x f$ een bijectieve lineaire afbeelding is van $\mathrm{T}_x X$ naar $\mathrm{T}_{f(x)} Y$, met $\mathrm{T}_{f(x)}\left(f^{-1}\right)$ als inverse. Omgekeerd, is $f \in \mathrm{C}^r(X, Y)$ en is in ieder punt van X de raakafbeelding van f bijectief, dan is er bij iedere $a \in X$ een open omgeving U van a in X en een open deelverzameling V van Y, met de eigenschap dat $f|_U$ een C^r-diffeomorfisme is van U naar V. Dit impliceert dat $f(X)$ een open deelverzameling is van Y. De globale inverse functiestelling voor variëteiten zegt dat als f bovendien injectief is, dan is f een C^r-diffeomorfisme van X naar $f(X)$.

Men noemt $f \in \mathrm{C}^r(X, Y)$ een C^r*-immersie*, resp. C^r*-submersie*, als in ieder punt van X de raakfbeelding van f injectief, resp. surjectief is. Men noemt f een C^r*-inbedding* als f een immersie is, injectief en f^{-1} een continue afbeelding is van $f(X)$ naar X. Voor een deelverzameling V van X krijgen we nu dat de volgende uitspraken equivalent zijn:

a) Er is een d-dimensionale C^r-variëteit D en een C^r-inbedding ψ van D naar X, waarvoor $\psi(D) = V$.

b) Bij iedere $a \in V$ is er een open omgeving U van a in X, een open deelverzameling D van $\mathbf{R}^d$ en een C^r-inbedding ψ van D naar X, waarvoor $\psi(D) = V \cap U$.

c) Bij iedere $a \in V$ is er een open omgeving U van a in X en een C^k-diffeomorfisme Φ van U naar een open deelverzameling van $\mathbf{R}^n$, benevens een open deelverzameling Y van $\mathbf{R}^d$, waarvoor $\Phi(V \cap U) = Y \times \{0\}$. Hierin is 0 de oorsprong in $\mathbf{R}^{n-d}$.

d) Bij iedere $a \in V$ is er een open omgeving U van a in $\mathbf{R}^n$ en een C^k-submersie g van U naar $\mathbf{R}^{n-d}$, waarvoor $V \cap U$ gelijk is aan de verzameling der $x \in U$ met $g(x) = 0$.

De inverses van de inbeddingen ψ in b) vormen een atlas voor V, waarmee V een d-dimensionale C^r-variëteit wordt en de identiteit van V naar X een C^r-inbedding is van V naar X. Dit bewijst b) $\Rightarrow$ a).

Men noemt V een *d-dimensionale C^r-deelvariëteit van X* als aan één van de equivalente voorwaarden a) — d) is voldaan. Het is duidelijk dat de definitie met de vroegere overeenkomt als $X = \mathbf{R}^n$. Een toepassing van d) is: als g een C^r-submersie is van X naar een $(n-d)$-dimensionale C^r-variëteit Y, dan is $g(X)$ een open deelverzameling van Y en voor iedere $y \in g(X)$ is het volledige origineel $g^{-1}(\{y\})$ een d-dimensionale C^r-deelvariëteit van X. Men noemt het volledige origineel $g^{-1}(\{y\})$ ook wel de *vezel* van de afbeelding g over het punt $y \in Y$.

5.5 De Duale Ruimte van de Raakruimte

Als $\gamma : I \to X$ een differentieerbare kromme is van een interval I naar de variëteit X, dan zal men gewoonlijk voor iedere $t \in I$ werken met de snelheidsvector $\gamma'(t) \in \mathrm{T}_{\gamma(t)} X$ en niet met de lineaire afbeelding $\mathrm{T}_t\,\gamma$ van $\mathrm{T}_t\,\mathbf{R}$ naar $\mathrm{T}_{\gamma(t)} X$. Tussen $\gamma'(t)$ en $\mathrm{T}_t\,\gamma$ hebben we de gecompliceerd ogende relatie

$$\gamma'(t) = \mathrm{T}_t\,\gamma\,(1_t)\,,$$

waarin $1_t = \iota'(t) \in \mathrm{T}_t\,\mathbf{R}$ als $\iota(t) = t$. Men kan ook te consequent zijn.

Substantiëler is de volgende opmerking. Zij f een differentieerbare reëelwaardige functie op de n-dimensionale differentieerbare variëteit X. Als we de raakruimten van $\mathbf{R}$ allen met $\mathbf{R}$ identificeren, dan definieert voor iedere $x \in X$ de raakafbeelding $\mathrm{T}_x\,f$ een lineaire afbeelding van $\mathrm{T}_x\,X$ naar $\mathbf{R}$, die de *differentiaal* $\mathrm{d}f(x)$ *van f in het punt x* genoemd wordt.

Voor een willekeurige lineaire ruimte E wordt een lineaire afbeelding $\epsilon : E \to \mathbf{R}$ een *lineaire vorm op E* genoemd. De ruimte van lineaire vormen op E heet de *duale ruimte van E* en wordt met E^* aangeduid. Is F een andere lineaire ruimte en A een lineaire afbeelding van E naar F dan definieert

$$F^* \ni \phi \mapsto \phi \circ A \in E^* \tag{5.8}$$

een lineaire afbeelding van F^* naar E^*, die de *duale* of *getransponeerde* of *gespiegelde* van A genoemd wordt en met A^* of A^{t} wordt aangeduid. In formule:

$$A^*(\phi) = \phi \circ A, \quad \phi \in F^*$$

ofwel

$$A^*(\phi)(e) = \phi(A(e)), \quad \phi \in F^*, \quad e \in E.$$

De notatie met A^* benadrukt de afhankelijkheid van ϕ in $\phi(A(e))$.

Als $E = \mathbf{R}^n$ dan levert

$$y \mapsto (x \mapsto \langle x,\, y\rangle) : \mathbf{R}^n \to (\mathbf{R}^n)^*$$

een lineair isomorfisme op van $\mathbf{R}^n$ naar $(\mathbf{R}^n)^*$, dat gebruikt wordt om $\mathbf{R}^n$ met zijn duale ruimte te identificeren. Als A een lineaire afbeelding is van $\mathbf{R}^n$ naar $\mathbf{R}^p$, dan wordt A^* met deze identificatie een lineaire afbeelding van $\mathbf{R}^p$ naar $\mathbf{R}^n$ en de matrix van A^* is dan gelijk aan de gespiegelde van de matrix van A.

Keren we terug naar de raakruimte $\mathrm{T}_x\, X$ en zijn duale. Een kaart κ levert een lineair isomorfisme $\mathrm{D}_x\, \kappa$ van $\mathrm{T}_x\, X$ naar $\mathbf{R}^n$, waarvan de getransponeerde $(\mathrm{D}_x\, \kappa)^*$ een lineair isomorfisme is van $\mathbf{R}^n = (\mathbf{R}^n)^*$ naar $(\mathrm{T}_x\, X)^*$. Daarmee is $(\mathrm{D}_x\, \kappa)^* \circ \mathrm{D}_x\, \kappa$ weliswaar een lineair isomorfisme van $\mathrm{T}_x\, X$ naar $(\mathrm{T}_x\, X)^*$, maar dit is zeker niet onafhankelijk van de keuze van de kaart κ. Om precies te zijn, als λ een andere kaart is, dan krijgen we hetzelfde isomorfisme, dan en slechts dan als

$$\mathrm{D}\left(\lambda \circ \kappa^{-1}\right)(\kappa(x))$$

een orthogonale lineaire transformatie in $\mathbf{R}^n$ is, hetgeen voor de coördinatentransformatie $\lambda \circ \kappa^{-1}$ een hele sterke inperking is.

Naar analogie met de definitie van raakvectoren in termen van de transformatieformule (5.1) kunnen we een lineaire vorm $\xi \in (\mathrm{T}_a\, X)^*$ identificeren met de afbeelding $\kappa \mapsto \xi_\kappa \in \mathbf{R}^n$, waarin

$$(\mathrm{D}_a\, \kappa)^*\,(\xi_\kappa) = \xi.$$

Hiervoor geldt de transformatieformule

$$\xi_\kappa = \mathrm{D}\left(\lambda \circ \kappa^{-1}\right)(\kappa(a))^* \cdot \xi_\lambda \tag{5.9}$$

In de terminologie van Ricci en Levi-Civita: lineaire vormen op de raakruimten transformeren *covariant* onder substituties van variabelen en niet contravariant, zoals snelheidsvectoren dat doen. Dit heeft direct te maken met het verschijnsel dat de totale afgeleiden van een reëlwaardige functie f van n variabelen onder de substitutie van variabelen $x = \Psi(y)$ covariant transformeert, zoals blijkt uit de formule

$$\mathrm{d}(f \circ \Psi)(y) = \mathrm{d}f(x) \circ \mathrm{D}\,\Psi(y) = \mathrm{D}\,\Psi(y)^*\,(\mathrm{d}f(x))\,,$$

die in coördinaten uitgeschreven de vorm

$$\frac{\partial(f \circ \Psi)(y)}{\partial y_j} = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} \cdot \frac{\partial \Psi_i(y)}{\partial y_j}, \quad x = \Psi(y)$$

krijgt.

5.6 De Raakbundel

Als de positieruimte een n-dimensionale variëteit X is, dan wordt de positie-snelheidsruimte, de faseruimte waarin de bewegingsvergelijkingen als een eerste orde stelsel van gewone differentiaalvergelijkingen geschreven werden, gelijk aan de vereniging over $x \in X$ van de raakruimten $\mathrm{T}_x\, X$. Men noemt deze vereniging de *raakbundel* $\mathrm{T}\, X$ van X. De afbeelding $\pi = \pi_{\mathrm{T}\, X}$ die aan $v \in \mathrm{T}_x\, X$ het 'basispunt' x toevoegt is een surjectieve afbeelding van $\mathrm{T}\, X$ naar X, met $\mathrm{T}_x\, X$ als *vezel* over x. Het is gebruikelijk om een raakvector $v \in \mathrm{T}\, X$ aan te duiden met $(x,\, v)$ als $v \in \mathrm{T}_x\, X$, dat wil zeggen als $x = \pi(v)$.

Is κ een kaart voor X, dan definieert

$$\mathrm{T}\,\kappa : (x,\, v) \mapsto (\kappa(x),\, \mathrm{D}_x\, \kappa \cdot v) \tag{5.10}$$

een bijectieve afbeelding van $\pi^{-1}(X_\kappa)$ naar de open deelverzameling $V_\kappa \times \mathbf{R}^n$ van $\mathbf{R}^n \times \mathbf{R}^n = \mathbf{R}^{2n}$. Men noemt $\mathrm{T}\,\kappa$ de *door κ geïnduceerde kaart voor* $\mathrm{T}\,X$. Uit de transformatieformule (5.4) lezen we af dat de $\mathrm{T}\,\kappa$ met $\kappa \in \mathcal{A}$ een (niet-maximale!) $2n$-dimensionale C^{r-1}-atlas voor $\mathrm{T}\,X$ vormen. Hiermee wordt $\mathrm{T}\,X$ een $2n$-dimensionale C^{r-1}-variëteit en de projectie $\pi : \mathrm{T}\,X \to X$ naar de basis is een C^{r-1}-submersie. De vezels $\mathrm{T}_x\,X$ zijn n-dimensionale C^{r-1}-deelvariëteiten van $\mathrm{T}\,X$, tegelijkertijd hebben deze de structuur van een n-dimensionale lineaire ruimte.

Zijn X en Y C^r-variëteiten en is $f \in \mathrm{C}^r(X,\,Y)$ dan definieert de raakafbeelding een C^{r-1}-afbeelding

$$\mathrm{T}\,f : (x,\,v) \mapsto \mathrm{T}_x\,f \cdot v \tag{5.11}$$

van $\mathrm{T}\,X$ naar $\mathrm{T}\,Y$. Deze heeft de bijzondere eigenschap dat zij de vezel over x afbeeldt naar de vezel over $f(x)$, in formule:

$$\pi_{\mathrm{T}\,Y} \circ \mathrm{T}\,f = f \circ \pi_{\mathrm{T}\,X}. \tag{5.12}$$

Verder is de beperking tot de vezel over x een lineaire afbeelding, namelijk $\mathrm{T}_x\,f$.

5.7 Vectorvelden en Stromingen

Een *(raak)vectorveld in* X is een afbeelding f die aan iedere $x \in X$ een raakvector $f(x)$ aan X in het punt x toevoegt. Dat wil zeggen, f is een afbeelding van X naar $\mathrm{T}\,X$, met de eigenschap dat voor iedere $x \in X$ geldt dat $f(x) \in \mathrm{T}_x\,X$, ofwel $\pi(f(x)) = x$. Nog weer anders gezegd, $f : X \to \mathrm{T}\,X$ en $\pi \circ f$ is gelijk aan de identiteit in X.

Nu we van $\mathrm{T}\,X$ een C^{r-1}-variëteit hebben gemaakt, kunnen we voor iedere $k \leq r-1$ spreken over een C^k-*vectorveld in* X, dat is een vectorveld in X dat als afbeelding van X naar $\mathrm{T}\,X$ een C^k-afbeelding is.

Zij f een vectorveld in X. Een kromme $\gamma : I \to X$ in X, met I een open interval in $\mathbf{R}$, heet een *oplossing van het beginwaardeprobleem*

$$x' = f(x), \quad x(t_0) = x_0, \tag{5.13}$$

als $\gamma(t_0) = x_0$ en voor iedere $t \in I$ geldt dat

$$\gamma'(t) = f(\gamma(t)). \tag{5.14}$$

Hierin is aangenomen dat $t_0 \in I$ en $x_0 \in X$. Merk op dat (5.14) een zinvolle vergelijking is, omdat zowel het linker- als het rechterlid elementen zijn van $\mathrm{T}_{\gamma(t)}\,X$.

Omdat in lokale coördinaten (5.13) een beginwaardeprobleem is voor een stelsel van gewone differentiaalvergelijkingen in een open deelverzameling van $\mathbf{R}^n$, kunnen we lokale theorie daarvan meteen overplanten naar willekeurige variëteiten. Dit geeft, als f een C^k-vectorveld is met $k \geq 1$, een lokale existentie- en eenduidigheidsstelling voor de oplossingen, benevens C^k-afhankelijkheid van de beginwaarden. Op dezelfde manier als voor stelsels in een open deel van $\mathbf{R}^n$, impliceert dit op zijn beurt een globale eenduidigheidsstelling en krijgen we unieke maximale oplossingen. Voor iedere $t \in \mathbf{R}$ hebben we de $0,\,t$-stroming Φ^t, hetgeen een C^k-diffeomorfisme is van een open deelverzameling X_t van X naar een open deelverzameling X_{-t} van X, met Φ^{-t} als inverse. Ook geldt de stelling dat als I het definitie-interval is van een maximale oplossing γ en $s = \sup I < \infty$, dan loopt $\gamma(t)$ voor $t \uparrow s$ uit X weg, in de zin dat er bij iedere compacte deelverzameling K van X een $\epsilon > 0$ is met de eigenschap dat $\gamma(t) \notin K$ zodra $t \in I$ en $t > s - \epsilon$. In het bijzonder, is de variëteit X compact (hetgeen niet zelden het geval is!) dan is iedere oplossing op de hele $\mathbf{R}$ gedefinieerd en is voor iedere $t \in \mathbf{R}$ de $0,\,t$-stroming een C^k-diffeomorfisme van X naar X.

5.8 Tweede-orde Stelsels

Echter, de bewegingsvergelijkingen van de klassieke mechanica vormen een tweede-orde stelsel van gewone differentiaalvergelijkingen in de positieruimte, hetgeen correspondeert met een eerste orde stelsel in de positie-snelheidsruimte. Generaliseren we de positieruimte tot een n-dimensionale C^r- -variëteit X met $r \geq 2$, dan wordt de positie-snelheidsruimte de $2n$dimensionale C^r-variëteit $\mathrm{T}\,X$. De bewegingsvergelijkingen corresponderen dan met een differentiaalvergelijking $v' = f(v)$, waarin f een vectorveld is dat gedefinieerd is in een open deelverzameling V van $\mathrm{T}\,X$. Echter, de speciale vorm $x' = v$, zie (2.26), van het eerste gedeelte van het stelsel in lokale coördinaten maakt dat niet ieder vectorveld in $\mathrm{T}\,X$ op kan treden als f behorend bij een tweede-orde stelsel in de positieruimte.

Om de voorwaarde voor f in coördinaatinvariante vorm te formuleren, merken we op dat een algemeen vectorveld f in V gedefinieerd was als een afbeelding $f : V \to \mathrm{T}\,(\mathrm{T}\,X)$, met de eigenschap dat $\pi_{\mathrm{T}(\mathrm{T}\,X)} \circ f$ gelijk is aan de identiteit in V. Echter, we hebben in $\mathrm{T}\,(\mathrm{T}\,X)$ niet alleen de projectie $\pi_{\mathrm{T}(\mathrm{T}\,X)}$ naar $\mathrm{T}\,X$, maar ook nog de raakfbeelding $\mathrm{T}\,(\pi_{\mathrm{T}\,X})$ van de projectie $\pi_{\mathrm{T}\,X} : \mathrm{T}\,X \to X$. In lokale coördinaten krijgen we voor een kromme $t \mapsto (x(t),\, v(t))$ in $\mathrm{T}\,X$ dat $\pi_{\mathrm{T}\,X}$ deze afbeeldt naar $x(t)$, dus $\mathrm{T}\,(\pi_{\mathrm{T}\,X})$ beeldt de raakvector $(x'(t),\, v'(t))$ in het punt $(x(t),\, v(t))$ af naar de raakvector $x'(t)$ in het punt $x(t)$. In formule:

$$\mathrm{T}\,(\pi_{\mathrm{T}\,X})\,\big((x(t),\, v(t)),\, (x'(t),\, v'(t))\big) = \big(x(t),\, x'(t)\big)\,.$$

Anderzijds hebben we dat

$$\pi_{\mathrm{T}(\mathrm{T}\,X)}\,\big((x(t),\, v(t)),\, (x'(t),\, v'(t))\big) = (x(t),\, v(t))\,.$$

De vergelijking $x' = v$ betekent daarmee precies dat f alleen maar waarden aanneemt in de deelverzameling

$$\mathrm{T}^{(2)}\,X := \{w \in \mathrm{T}\,(\mathrm{T}\,X) \mid \mathrm{T}\,(\pi_{\mathrm{T}\,X})\,(w) = \pi_{\mathrm{T}(\mathrm{T}\,X)}(w)\}. \tag{5.15}$$

van de raakbundel van $\mathrm{T}\,X$.

We kunnen de verzameling $\mathrm{T}^{(2)}\,X$ ook karakteriseren in termen van tweede orde afgeleiden van krommen $\gamma : I \to X$. Immers, daarvoor is $\gamma' : I \to \mathrm{T}\,X$, dus $(\gamma')' : I \to \mathrm{T}\,(\mathrm{T}\,X)$ en het bovenstaande geeft dat $(\gamma')'\,(I) \subset \mathrm{T}\,(\mathrm{T}\,X)$. Anderzijds zien we in lokale coördinaten ook dat ieder element van $\mathrm{T}^{(2)}\,X$ ook inderdaad optreedt als $(\gamma')'\,(t)$ voor een geschikte tweemaal differentieerbare kromme $\gamma : I \to X$, dus $\mathrm{T}^{(2)}\,X$ kan geïdentificeerd worden met *de verzameling van tweede orde Taylor-ontwikkelingen van krommen in* X, maar dan coördinaat-invariant. Merk op dat $\mathrm{T}^{(2)}\,X$ een $3n$-dimensionale C^{r-2}-deelvariëteit is van de $4n$-dimensionale variëteit $\mathrm{T}\,(\mathrm{T}\,X)$; dit is de goede dimensie want de tweede-orde Taylor-ontwikkeling van γ in het punt t wordt in lokale coördinaten vastgelegd door het geven van

$$(\gamma(t),\, \gamma'(t),\, \gamma''(t)) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n.$$

Verzuchting: dit zijn eigenlijk allemaal trivialiteiten!

5.9 De Coraakbundel

We hebben in Hoofdstuk 4 gezien dat de grootheid $[L]$ covariant transformeert onder substituties van positie-variabelen. Daarmee is $[L]$ te beschouwen als een element van de duale ruimte $(\mathrm{T}_x\,X)^*$ van de raakruimte $\mathrm{T}_x\,X$. De bewegingsvergelijking $[T] = \phi$ van Lagrange maakt dat de daarin verschijnende kracht ϕ ook een lineaire vorm op de raakruimte is.

Analoog aan de raakbundel van X wordt nu de *coraakbundel* $\mathrm{T}^* X$ van X gedefinieerd als de vereniging, over alle $x \in X$, van de lineaire ruimten $(\mathrm{T}_x X)^*$. Iedere kaart κ voor X induceert een kaart $\widetilde{\kappa}$ voor $\mathrm{T}^* X$, die is gedefinieerd door

$$\widetilde{\kappa}\,(x,\, \eta \circ \mathrm{D}_x\, \kappa) = (\kappa(x),\, \eta)\,. \tag{5.16}$$

Anders gezegd, $\eta \in \mathbf{R}^n$ wordt uit $\xi \in (\mathrm{T}_x X)^*$ verkregen door daarop toe te passen: de getransponeerde van de inverse van $\mathrm{D}_x\, \kappa =$ de inverse van de getransponeerde van $\mathrm{D}_x\, \kappa$.

Met deze kaarten wordt $\mathrm{T}^* X$ een $2n$-dimensionale C^{r-1}-variëteit. De afbeelding $\pi_{\mathrm{T}^* X}$ die aan $\xi \in (\mathrm{T}_x X)^*$ het basispunt $x \in X$ toevoegt is een surjectieve C^{r-1}-submersie van $\mathrm{T}^* X$ naar X, met de n-dimensionale lineaire ruimte $(\mathrm{T}_x X)^*$ als *vezel* over $x \in X$.

Een *differentiaalvorm in* X is nu een afbeelding ϕ die aan iedere $x \in X$ een lineaire vorm $\phi(x)$ op $\mathrm{T}_x X$ toevoegt. Dat wil zeggen, ϕ is een afbeelding van X naar $\mathrm{T}^* X$, waarvoor $\pi_{\mathrm{T}^* X} \circ \phi$ gelijk is aan de identiteit in X. Voorbeeld: een tijdsonafhankelijk krachtveld in X. Ander voorbeeld: als f een differentieerbare reëelwaardige functie in X is, dan is $\mathrm{d}f$ een differentiaalvorm in X.

Dit laatste voorbeeld leidt tot de natuurlijke vraag, voor welke differentiaalvormen ϕ er een reëelwaardige functie f bestaat, waarvoor $\mathrm{d}f = \phi$. Men noemt ϕ *exact* als dit het geval is. In het geval van een krachtveld, dan noemt men ϕ ook wel *conservatief* en de functie $V = -f$ is een bijbehorende *potentiële energiefunctie.*

In Hoofdstuk 2 is hierop het volgende antwoord gegeven: als ϕ continu differentieerbaar is, dan is in lokale coördinaten de exactheid van ϕ equivalent met de voorwaarde dat

$$\alpha_{ij}(x) := \frac{\partial \phi_i(x)}{\partial x_j} - \frac{\partial \phi_j(x)}{\partial x_i} = 0, \quad 1 \leq i,\, j \leq n.$$

Wij geven nu een coördinaat-invariante interpretatie van de antisymmetrische $n \times n$-matrix $\alpha_{ij}(x)$ die hier in het linkerlid staat. De eerste stap is dat we voor ieder paar van vectoren $u,\, v \in \mathbf{R}^n$ schrijven

$$\mathrm{d}\phi(x)(u,\, v) := \sum_{i,\, j=1}^{n} \alpha_{ij}(x)\, u_j\, v_i = [\mathrm{D}\, \phi(x) \cdot u] \cdot v - [\mathrm{D}\, \phi(x) \cdot v] \cdot u. \tag{5.17}$$

Dit definieert een afbeelding $\mathrm{d}\phi(x)$ van $\mathbf{R}^n \times \mathbf{R}^n$ die *bilineair* is en *antisymmetrisch*; dit laatste betekent dat de uitdrukking in zijn tegengestelde overgaat bij verwisselen van u en v. Men kan dit vergelijken met een inproduct, dat een symmetrische bilineaire vorm is (bovendien positief definiet).

Passen we nu een substitutie van variabelen $x = \Psi(y)$ toe, dan transformeert de differentiaalvorm ϕ covariant, dat wil zeggen dat deze in de y-coördinaten gegeven wordt door

$$\chi_j(y) = \sum_{i=1}^{n} \phi_i(\Psi(y))\, \frac{\partial \Psi_i(y)}{\partial y_j}.$$

Nemen we nu aan dat $\Psi \in \mathrm{C}^2$ dan krijgen we, gebruikmakend van de symmetrie van de tweede-orde afgeleidenmatrix van de functie Ψ_i, dat

$$\mathrm{d}\chi(y)(u,\, v) = \mathrm{d}\phi(\Psi(y))\, (\mathrm{D}\, \Psi(y) \cdot u,\, \mathrm{D}\, \Psi(y) \cdot v)\,. \tag{5.18}$$

Dit betekent dat $\mathrm{d}\phi(x)$ coördinaatinvariant gedefinieerd is als een antisymmetrische bilineaire vorm op $\mathrm{T}_x X \times \mathrm{T}_x X$. Men noemt $\mathrm{d}\phi(x)$ de *uitwendige afgeleide van* ϕ *in het punt* x.

De ruimte van antisymmetrische bilineaire vormen op $\mathrm{T}_x X \times \mathrm{T}_x X$ wordt met $\bigwedge^2 (\mathrm{T}_x X)^*$ aangeduid, dit is een lineaire ruimte, met dimensie gelijk aan $\frac{1}{2}n(n-1)$. De vereniging over alle $x \in X$ van deze ruimten wordt met $\bigwedge^2 \mathrm{T}^* X$ genoteerd en we hebben een projectie $\pi : \bigwedge^2 \mathrm{T}^* X \to X$ naar de basisruimte X. Net als bij de coraakbundel induceren de kaarten voor X een atlas voor $\bigwedge^2 \mathrm{T}^* X$, waarmee $\bigwedge^2 \mathrm{T}^* X$ een C^{r-1}-variëteit wordt met dimensie gelijk aan $n + \frac{1}{2}n(n-1)$. Een *differentiaalvorm van de graad twee* of kortweg een *tweevorm* is nu gedefinieerd als een afbeelding ω die aan iedere $x \in X$ een antisymmetrische bilineaire vorm op $\mathrm{T}_x X \times \mathrm{T}_x X$ toevoegt. Dat wil zeggen, ω is een afbeelding van X naar $\bigwedge^2 \mathrm{T}^* X$ met de eigenschap dat $\pi \circ \omega$ gelijk is aan de identiteit in X.

Voorbeeld: is ϕ een continu differentieerbare differentiaalvorm in X, dan is de uitwendige afgeleide

$$\mathrm{d}\phi : x \mapsto \mathrm{d}\phi(x)$$

van ϕ een tweevorm in X. Als $\phi = \mathrm{d}f$ voor een differentieerbare functie f, dan is $f \in \mathrm{C}^2$ en $\mathrm{d}\phi = 0$. Omgekeerd is er bij iedere $a \in X$ een open omgeving U van a in X met de eigenschap dat als $\mathrm{d}\phi = 0$ in U, dan is er een C^2-functie f in U, waarvoor $\phi|_U = \mathrm{d}f$. Men noemt een continu differentieerbare differentiaalvorm ϕ *gesloten* als $\mathrm{d}\phi = 0$. Het bovenstaande geeft dat iedere exacte differentiaalvorm gesloten is en dat lokaal ook de omkering geldt.

5.10 De Uitwendige Algebra

Functies, resp. differentiaalvormen worden ook wel differentiaalvormen van de graad nul, resp. één genoemd. Differentiaalvormen van de graad nul, één en twee passen in een algemener kader van differentiaalvormen van willekeurige graad. Voor de invoering daarvan hebben we wat meer multilineaire algebra nodig.

Zij E een n-dimensionale lineaire ruimte. Een *p-vorm in E* is een p-lineaire reëelwaardige functie op $E^p = E \times \ldots E$, die *antisymmetrisch* is de zin dat voor iedere permutatie π van de rangnummers $1, \ldots, p$ geldt dat

$$\omega\left(v_{\pi(1)}, \ldots, v_{\pi(p)}\right) = (\operatorname{sgn} \pi)\, \omega\left(v_1, \ldots, v_p\right). \tag{5.19}$$

Hierin is $\operatorname{sgn} \pi$ het teken van de permutatie π. De ruimte van p-vormen in E wordt met $\bigwedge^p E^*$ aangeduid, dit is een lineaire deelruimte van de ruimte van alle reëelwaardige functies op E^p.

Als $\omega \in \bigwedge^p E^*$ en $\nu \in \bigwedge^q E^*$, dan is het *uitwendige product van ω en ν* de $(p+q)$-vorm $\omega \wedge \nu$ in E die is gedefinieerd door

$$(\omega \wedge \nu)\left(v_1, \ldots, v_{p+q}\right) = \frac{1}{p!\,q!} \sum_{\pi} (\operatorname{sgn} \pi)\, \omega\left(v_{\pi(1)}, \ldots, v_{\pi(p)}\right) \nu\left(v_{\pi(p+1)}, \ldots, v_{\pi(p+q)}\right). \tag{5.20}$$

Hierbij dient opgemerkt te worden dat voor iedere opdeling $\{1, \ldots, p+q\} = I \cup J$ met $\#(I) = p$ en $\#(J) = q$ alle termen met $\pi(i) \in I$ als $1 \le i \le p$ en $\pi(j) \in J$ als $p+1 \le j \le p+q$ aan elkaar gelijk zijn. Omdat dit $p!\,q!$ termen zijn, kunnen we (5.20) dus ook beschouwen als een som over alle partities I, J van termen zonder voorfactoren.

Het uitwendige product is *associatief*, in de zin dat als $\omega \in \bigwedge^p E^*$, $\nu \in \bigwedge^q E^*$, $\mu \in \bigwedge^r E^*$, dan is

$$\omega \wedge (\nu \wedge \mu) = (\omega \wedge \nu) \wedge \mu. \tag{5.21}$$

Verder is

$$\omega \wedge \nu = (-1)^{p\,q}\, \nu \wedge \omega, \tag{5.22}$$

dus het uitwendige product is *commutatief* zodra de graad van een van de beide factoren even is en *anticommutatief* als de graad van beide factoren oneven is. Tenslotte, is $\epsilon_1, \ldots, \epsilon_n$ een basis van E^*, dan vormen de vectoren

$$\epsilon_{i(1)} \wedge \epsilon_{i(2)} \wedge \ldots \wedge \epsilon_{i(p)}, \quad \text{met } i(1) < i(2) < \ldots < i(p) \tag{5.23}$$

een basis van $\bigwedge^p E^*$. In het bijzonder is

$$\dim \bigwedge^p E^* = \begin{pmatrix} n \\ p \end{pmatrix}. \tag{5.24}$$

Dit impliceert dat $\bigwedge^p E^* = \{0\}$ als $p > n$ en dat $\bigwedge^p E^*$ en $\bigwedge^{n-p} E^*$ dezelfde dimensie hebben. In deze formules hanteert men ook nog de afspraak dat $\bigwedge^0 E^* := \mathbf{R}$. Tenslotte noemt men de directe som

$$\bigwedge E^* \tag{5.25}$$

ook wel de *uitwendige algebra* van E^*, deze vormt een lineaire ruimte over $\mathbf{R}$ met dimensie gelijk aan

$$\dim \bigwedge E^* = \sum_{p=0}^{n} \begin{pmatrix} n \\ p \end{pmatrix} = 2^n.$$

Is A een lineaire afbeelding van E naar een eindig-dimensionale lineaire ruimte F en $\omega \in \bigwedge^p F^*$, dan definieert

$$(A^*\omega)\,(v_1, \ldots, v_p) = \omega\,(A \cdot v_1, \ldots, A \cdot v_p)\,, \quad v_i \in E, \tag{5.26}$$

een element $A^*\omega \in \bigwedge^p E^*$, de *door A naar E teruggetrokken p-vorm* genoemd. Dit definieert een lineaire afbeelding $A^* : \omega \mapsto A^*\omega$ van $\bigwedge^p F^*$ naar $\bigwedge^p E^*$. Merk op dat we hierbij geen restricties hebben opgelegd aan de dimensies van E en F en aan de graad p. Het uitwendige product transformeert onder terugtrekken zoals verwacht:

$$A^*(\omega \wedge \nu) = (A^*\omega) \wedge (A^*\nu), \quad \omega \in \bigwedge^p F^*,\ \nu \in \bigwedge^q F^*. \tag{5.27}$$

Verder, is B een lineaire afbeelding van F naar een eindig-dimensionale lineaire ruimte G, dan geldt dat

$$A^*(B^*\omega) = (B \circ A)^*\omega, \quad \omega \in \bigwedge^p G^*. \tag{5.28}$$

Deze eigenschappen van terugtrekken volgen direct uit de definitie.

Interessant is hier nog het speciale geval dat $E = F$ en $p = n = \dim E$. In dit geval is de ruimte der n-vormen in E ééndimensionaal, hetgeen impliceert dat er voor iedere $A \in \mathrm{L}(E, E)$ een reëel getal $\delta(A)$ is, met de eigenschap dat $A^*\omega = \delta(A)\,\omega$ voor iedere n-vorm ω. Men gaat gemakkelijk na uit de definitie dat $\delta(A) = \det A$; men kan dit ook heel goed opvatten als de definitie van de determinant. Omdat het n-dimensionale volume van $A(P)$ gelijk is aan $|\det A|$ maal het n-dimensionale volume van P, wordt een n-vorm in E ook wel een *(georiënteerde) volumevorm* genoemd. Het woord georiënteerd wordt toegevoegd om eraan te herinneren dat A een oriëntatiebehoudende lineaire transformatie is, dan en slechts dan als $\det A > 0$; als $\det A < 0$ dan krijgt de n-vorm tegengesteld teken terwijl volumes van gebieden altijd positief zijn.

Is $v \in E$ en $\omega \in \bigwedge^p E^*$, dan definieert

$$(\mathrm{i}_v\,\omega)\,(v_1, \ldots, v_{p-1}) = \omega(v, v_1, \ldots, v_{p-1}) \tag{5.29}$$

een $(p-1)$-vorm $\mathrm{i}_v\,\omega$ in E, die het *inwendige product van* ω *met* v genoemd wordt. Formules:

$$\mathrm{i}_v(\omega\wedge\nu) = (\mathrm{i}_v\,\omega)\wedge\nu + (-1)^p\omega\wedge(\mathrm{i}_v\,\nu)\,,\quad \omega\in\bigwedge^p E^*,\ \nu\in\bigwedge^q E^*, \tag{5.30}$$

terwijl voor een lineaire afbeelding A van E naar F geldt dat

$$\mathrm{i}_v\,(A^*\omega) = A^*\,(\mathrm{i}_{A\,v}\,\omega)\,,\quad v\in E,\ \omega\in\bigwedge^p F^*. \tag{5.31}$$

5.11 Differentiaalvormen van Willekeurige Graad

Zij X een n-dimensionale C^r-variëteit, met r voldoende groot. Een *differentiaalvorm van de graad* p *in* X of kortweg een p*-vorm in* X is een afbeelding ω die aan iedere $x\in X$ een p-vorm in $\mathrm{T}_x\,X$ toevoegt. Duiden we de vereniging, over alle $x\in X$, van de lineaire ruimten $\bigwedge^p(\mathrm{T}_x\,X)^*$ aan met $\bigwedge^p\mathrm{T}^*\,X$, dan induceren de kaarten voor X op natuurlijke wijze kaarten voor $\bigwedge^p\mathrm{T}^*\,X$, waarmee dit een C^{r-1}-variëteit wordt. Als π weer de projectie naar de basis X is, dan is een p-vorm een afbeelding ω van X naar $\bigwedge^p\mathrm{T}^*\,X$, met de eigenschap dat $\pi\circ\omega$ gelijk is aan de identiteit in X.

Alle algebraïsche bewerkingen voor p-vormen in een lineaire ruimte kunnen puntsgewijs, voor iedere $x\in X$, uitgevoerd worden voor p-vormen. Zo is het uitwendige product van de p-vorm ω en de q-vorm ν de $(p+q)$-vorm $\omega\wedge\nu$ in X die is gedefinieerd door

$$(\omega\wedge\nu)_x := \omega_x\wedge\nu_x,\quad x\in X. \tag{5.32}$$

Hierbij heb ik me aangewend om het basispunt onderaan te hangen, om de raakvectoren waar de p-vorm op werkt erachter te kunnen schrijven. Is ω, resp v een p-vorm, resp. vectorveld in X, dan wordt het inwendige product van ω met v de $(p-1)$-vorm $\mathrm{i}_v\,\omega$ in X die is gedefinieerd door

$$(\mathrm{i}_v\,\omega)_x := \mathrm{i}_{v(x)}\,\omega_x,\quad x\in X. \tag{5.33}$$

Het is duidelijk dat daarbij alle algebraïsche identiteiten voor uit- en inproducten van p–vormen puntsgewijs geldig blijven.

Is Φ een differentieerbare afbeelding van X naar een variëteit Y, dan wordt voor iedere p-vorm ω in Y de *door* Φ *teruggetrokken* p*-vorm* $\Phi^*\omega$ *in* X gedefinieerd door:

$$(\Phi^*\omega)_x := (\mathrm{T}_x\,\Phi)^*\ \omega_{\Phi(x)},\quad x\in X. \tag{5.34}$$

Formules:

$$\begin{aligned}
\Phi^*(\omega\wedge\nu) &= (\Phi^*\omega)\wedge(\Phi^*\nu)\,, &(5.35)\\
\Phi^*\,(\Psi^*\omega) &= (\Psi\circ\Phi)^*\omega, &(5.36)\\
\Phi^*\,(\mathrm{i}_v\,\omega) &= \mathrm{i}_{\Phi^*v}\ \Phi^*\omega. &(5.37)
\end{aligned}$$

Hierbij is in (5.36) Ψ een differentieerbare afbeelding van Y naar een varëteit Z en is ω een p-vorm in Z. In (5.37) hebben we de notatie

$$(\Phi^*v)(x) := (\mathrm{T}_x\,\Phi)^{-1}\cdot v(\Phi(x)) \tag{5.38}$$

gebruikt; waarbij we hebben aangenomen dat de raakafbeeldingen van Φ bijectief zijn, dus dat Φ een lokaal diffeomorfisme is.

In een coördinatensysteem correspondeert $\mathrm{d}x_i$, waarin x_i opgevat wordt als de i-de coördinaatsfunctie $x \mapsto x_i$, met de i-de standaard basisvector in $(\mathbf{R}^n)^*$. Als κ een kaart is voor X, dan definieert het terugtrekken door κ een isomorfisme van de ruimte van de p-vormen in V_κ naar de ruimte van de p-vormen in X_κ. Als $\omega = \kappa^* \widetilde{\omega}$ voor de p-vorm $\widetilde{\omega}$ in V_κ, dan zijn er eenduidig bepaalde reëelwaardige functies $\omega_{i(1),\,\ldots,\,i(p)}$ in V_κ, voor $1 \leq i(1) < \ldots < i(p) \leq n$, met de eigenschap dat

$$\widetilde{\omega} = \sum_{1 \leq i(1) < \ldots < i(p) \leq n} \omega_{i(1),\,\ldots,\,i(p)}(x)\ \mathrm{d}x_{i(1)} \wedge \ldots \wedge \mathrm{d}x_{i(p)}. \tag{5.39}$$

Er geldt dat $\omega \in \mathrm{C}^k$, $k \leq r-1$, dan en slechts dan als alle $\omega_{i(1),\,\ldots,\,i(p)}$ C^k-functies in V_κ zijn. Men gebruikt (5.39) om concrete berekeningen uit te voeren voor p-vormen in termen van reëelwaardige functies van n reële variabelen.

5.12 Uitwendige Afgeleide

De uitwendige afgeleide kan nu gegeneraliseerd worden tot p-vormen, op de volgende manier. Er is een afbeelding d, die aan iedere differentieerbare p-vorm ω in X een $(p+1)$-vorm $\mathrm{d}\omega$ toevoegt en die voor iedere kaart κ voor X gegeven wordt door de formule

$$\widetilde{\mathrm{d}\omega} = \sum_{1 \leq i(1) < \ldots < i(p) \leq n} \mathrm{d}\omega_{i(1),\,\ldots,\,i(p)} \wedge \mathrm{d}x_{i(1)} \wedge \ldots \wedge \mathrm{d}x_{i(p)}, \quad x \in V_\kappa, \tag{5.40}$$

als ω is gegeven door (5.39). Het rechterlid van (5.40) kan in de gedaante (5.39) gebracht worden, met p vervangen door $p+1$ en met andere coëfficiënten, door in (5.40) te substitueren

$$\mathrm{d}\omega_{i(1),\,\ldots,\,i(p)} = \sum_{j=1}^{n} \frac{\partial \omega_{i(1),\,\ldots,\,i(p)}(x)}{\partial x_j}\ \mathrm{d}x_j,$$

te gebruiken dat $\mathrm{d}x_j \wedge \mathrm{d}x_{i(1)} \wedge \ldots \wedge \mathrm{d}x_{i(p)} = 0$ als $j \in \{i(1),\,\ldots,\,i(p)\}$, dat

$$\mathrm{d}x_j \wedge \mathrm{d}x_{i(1)} \wedge \ldots \wedge \mathrm{d}x_{i(p)} = (-1)^l\ \mathrm{d}x_{i(1)} \wedge \ldots \wedge \mathrm{d}x_{i(l)} \wedge \mathrm{d}x_j \wedge \mathrm{d}x_{i(l+1)} \wedge \ldots \wedge \mathrm{d}x_{i(p)}$$

als $i(l) < j < i(l+1)$, $0 \leq l \leq p$, en tenslotte alle coëfficiënten voor dezelfde $(p+1)$-vorm

$$\mathrm{d}x_{j(1)} \wedge \ldots \wedge \mathrm{d}x_{j(p+1)}$$

met $1 \leq j(1) < j(2) < \ldots j(p) < j(p+1)$ te verzamelen.

Hiermee werkt de uitwendige afgeleide op de coëfficiënten van de differentiaalvormen als een eerste orde lineaire partiële differentiaaloperator. Voor $p = 0$ is $\mathrm{d}f$ de totale afgeleide van f en voor $p = 1$ is $\mathrm{d}\omega$ gelijk aan de vroeger gedefinieerde uitwendige afgeleide van de differentiaalvorm ω. Is ν een differentieerbare q-vorm, dan geldt dat

$$\mathrm{d}(\omega \wedge \nu) = \mathrm{d}\omega \wedge \nu + (-1)^p \omega \wedge \mathrm{d}\nu. \tag{5.41}$$

Is ω tweemaal continu differentieerbaar, dan is

$$\mathrm{d}(\mathrm{d}\omega) = 0. \tag{5.42}$$

Omgekeerd impliceren (5.41), (5.42), en de lineariteit van de operator d, dat (5.40) geldt. Dit betekent dat de operator d volledig vastgelegd is door de eigenschappen dat zij lineair is en aan (5.41) en (5.42) voldoet. Anders gezegd, men kan d ook *definiëren* als de unieke lineaire operator d met de eigenschappen (5.41) en (5.42).

Tenslotte, is Φ een differentieerbare afbeelding van X naar een differentieerbare variëteit Y en is ω een differentieerbare p-vorm in Y, dan is

$$\Phi^*(\mathrm{d}\omega) = \mathrm{d}(\Phi^*\omega)\,. \tag{5.43}$$

Men zegt: *uitwendige differentiatie commuteert met terugtrekken door willekeurige afbeeldingen.*

5.13 Lie-afgeleide

Is v een C^1-vectorveld, dan hebben we de differentieerbare éénparameterfamilie van stromingen na tijd t, die we met een wat gedurfde notatie aanduiden met $\mathrm{e}^{t\,v}$. Voor een differentieerbare p-vorm ω in X definieert men nu de *Lie-afgeleide* $\mathcal{L}_v\omega$ *van* ω *naar het vectorveld* v door middel van:

$$\mathcal{L}_v\omega := \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{t\,v}\right)^*\,\omega|_{t=0}. \tag{5.44}$$

Een motivering voor deze definitie is de stelling dat $\mathcal{L}_v\omega = 0$, dan en slechts dan als ω invariant is onder de stroming met snelheidsveld gelijk aan v, in de zin dat

$$\left(\mathrm{e}^{t\,v}\right)^*\,\omega = \omega, \quad t \in \mathbf{R}.$$

Het 'dan' is evident, voor het bewijs van het 'slechts dan' schrijft men de afgeleide van $\left(\mathrm{e}^{t\,v}\right)^*\,\omega$ naar t als de afgeleide naar h van

$$\left(\mathrm{e}^{(t+h)\,v}\right)^*\,\omega = \left(\mathrm{e}^{h\,v}\circ\mathrm{e}^{t\,v}\right)^*\,\omega = \left(\mathrm{e}^{t\,v}\right)^*\left(\left(\mathrm{e}^{h\,v}\right)^*\,\omega\right)$$

in $h = 0$. Gebruikmakend van het feit dat Φ^* een continue lineaire afbeelding is, geeft dit dat

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{t\,v}\right)^*\,\omega = \left(\mathrm{e}^{t\,v}\right)^*(\mathcal{L}_v\,\omega)\,,$$

hetgeen identiek gelijk aan nul is als $\mathcal{L}_v\omega = 0$.

Uit (5.35) volgt dat

$$\mathcal{L}_v(\alpha\wedge\beta) = (\mathcal{L}_v\alpha)\wedge\beta + \alpha\wedge(\mathcal{L}_v\beta)\,, \tag{5.45}$$

terwijl (5.43) leidt tot

$$\mathrm{d}(\mathcal{L}_v\omega) = \mathcal{L}_v\,(\mathrm{d}\omega)\,. \tag{5.46}$$

De volgende zogenaamde *homotopieformule* legt een verband tussen Lie-afgeleiden, uitwendige differentiatie en inproducten:

$$\mathcal{L}_v\omega = \mathrm{d}(\mathrm{i}_v\,\omega) + \mathrm{i}_v\,(\mathrm{d}\omega)\,. \tag{5.47}$$

Het bewijs kan gegeven worden met volledige inductie naar de graad. Als ω een functie f is dan geldt (5.47) met $\mathrm{i}_v\,f = 0$. Lokaal kunnen we ω schrijven in de gedaante (5.39), waarbij opgemerkt kan worden dat

$$\mathrm{d}\kappa_{i(1)}\wedge\ldots\wedge\mathrm{d}\kappa_{i(p)} = \mathrm{d}\big(\kappa_{i(1)}\;\mathrm{d}\kappa_{i(2)}\wedge\ldots\wedge\mathrm{d}\kappa_{i(p)}\big)\,,$$

waarbij we hebben gebruikt dat $\mathrm{d} \circ \mathrm{d} = 0$. Omdat $\mathcal{L}_v$ lokaal werkt en een lineaire operator is, leidt dit ertoe toe dat het voldoende is om (5.47) te bewijzen voor het geval dat $\omega = f\ \mathrm{d}\mu$, waaarin f een functie is en μ een differentiaalvorm van de graad $p-1$. De inductieveronderstelling geeft dat (5.47) geldt met ω vervangen door μ. Dit geeft dat $\mathcal{L}_v(f\ \mathrm{d}\mu)$ gelijk is aan

$$
\begin{aligned}
&(\mathcal{L}_v f)\ \mathrm{d}\mu + f\,\mathcal{L}_v\,\mathrm{d}\mu = (\mathrm{i}_v\,\mathrm{d}f)\ \mathrm{d}\mu + f\ \mathrm{d}(\mathcal{L}_v\mu) = (\mathrm{i}_v\,\mathrm{d}f)\ \mathrm{d}\mu + f\ \mathrm{d}(\mathrm{d}(\mathrm{i}_v\,\mu) + \mathrm{i}_v\,\mathrm{d}\mu)\\
&= (\mathrm{i}_v\,\mathrm{d}f)\ \mathrm{d}\mu + f\ \mathrm{d}\mathrm{i}_v\,\mathrm{d}\mu = (\mathrm{i}_v\,\mathrm{d}f)\ \mathrm{d}\mu - \mathrm{d}f\wedge\mathrm{i}_v\,\mathrm{d}\mu + \mathrm{d}(f\ \mathrm{i}_v\,\mathrm{d}\mu)\\
&= \mathrm{i}_v\,(\mathrm{d}f\wedge\mathrm{d}\mu) + \mathrm{d}\,\mathrm{i}_v\,(f\ \mathrm{d}\mu) = \mathrm{i}_v\,\mathrm{d}(f\ \mathrm{d}\mu) + \mathrm{d}\,\mathrm{i}_v\,(f\ \mathrm{d}\mu)\,.
\end{aligned}
$$

5.14 Het Lemma van Poincaré

Dit lemma is een lokale omkering van de identiteit (5.42), die zegt dat als α *exact* is, dat wil zeggen er is een ω met $\alpha = \mathrm{d}\omega$, dan is α *gesloten*, dat wil zeggen $\mathrm{d}\alpha = 0$. Een elegant bewijs van deze stelling, voor differentiaalvormen van willekeurige graad, kan gegeven worden door gebruik te maken van het idee van *homotopie*.

In de meest algemene vorm betreft dit een C^1-familie van C^1-afbeeldingen ϕ_t van een C^r-variëteit Y naar een C^r-variëteit X, geparametriseerd door een reële parameter $t \in I := [0,\,1]$. Hierin is $r \geq 2$. Schrijven we $\Phi(t,\,y) = \phi_t(y)$ dan betekent dit dat Φ een C^1-afbeelding is van $I\times Y$ naar X en dat, in lokale coördinaten, $\frac{\partial\Phi(t,\,y)}{\partial y}$ differentieerbaar is naar t en dat

$$
(t,\,y) \mapsto \frac{\partial^2\Phi(t,\,y)}{\partial t\,\partial y}
$$

continu is. We roepen in herinnering dat de stelling over verwisselbaarheid van de differentiatievolgorde in deze situatie geeft dat $\frac{\partial\Phi(t,\,y)}{\partial t}$ differentieerbaar is naar y en dat

$$
\frac{\partial^2\Phi(t,\,y)}{\partial y\,\partial t} = \frac{\partial^2\Phi(t,\,y)}{\partial t\,\partial y}.
$$

We noteren verder $\iota_t(y) = (t,\,y)$, dit definieert voor iedere $t \in I$ een C^r-afbeelding $\iota_t : Y \to I\times Y$. Tenslotte is $\partial/\partial t$ een notatie voor het vectorveld $I\times Y$ waarvoor

$$
\mathrm{i}_{\partial/\partial t}\,\mathrm{d}f = \frac{\partial f}{\partial t}
$$

voor iedere differentieerbare functie f in $I\times Y$; in lokale coördinaten is $\partial/\partial t$ constant gelijk aan $(1,\,0)$. De stroming van $\partial/\partial t$ na tijd h is de verschuiving $\tau_h : (t,\,y) \mapsto (t+h,\,y)$. Opmerkend dat

$$
\iota_{t+h} = \tau_h \circ \iota_t,
$$

krijgen we nu de volgende rij van identiteiten.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\,{\phi_t}^*\omega &= \frac{\mathrm{d}}{\mathrm{d}t}\,(\Phi\circ\iota_t)^*\,\omega = \frac{\mathrm{d}}{\mathrm{d}t}\,{\iota_t}^*\Phi^*\omega\\
&= \frac{\mathrm{d}}{\mathrm{d}h}\,{\iota_{t+h}}^*\Phi^*\omega\Big|_{h=0} && (5.48)\\
&= \frac{\mathrm{d}}{\mathrm{d}h}\,(\tau_h\circ\iota_t)^*\,\Phi^*\omega\Big|_{h=0} = {\iota_t}^*\mathcal{L}_{\partial/\partial t}\,(\Phi^*\omega)\\
&= {\iota_t}^*\circ\left[\mathrm{d}\circ\mathrm{i}_{\partial/\partial t} + \mathrm{i}_{\partial/\partial t}\circ\mathrm{d}\right]\circ\Phi^*\omega\\
&= \mathrm{d}\circ{\iota_t}^*\circ\mathrm{i}_{\partial/\partial t}\circ\Phi^*\omega + {\iota_t}^*\circ\mathrm{i}_{\partial/\partial t}\circ\Phi^*\circ\mathrm{d}\,\omega. && (5.49)
\end{aligned}
$$

Een essentiële stap hierin was het gebruik van de formule (5.47) voor de Lie-afgeleide, deze toepassing op homotopiëen is de reden voor de naam 'homotopieformule' voor (5.47). Schrijven we nu

$$H\alpha = \int_0^1 \iota_t{}^* \circ \mathrm{i}_{\partial/\partial t} \circ \Phi^* \alpha \, dt, \tag{5.50}$$

dan is H een afbeelding die aan differentiaalvormen van de graad q differentiaalvormen van de graad $q-1$ toevoegt. Door integratie van $\frac{d}{dt}\,\phi_t{}^*\omega$ over t van $t=0$ naar $t=1$ krijgen we dat

$$\phi_1{}^*\omega - \phi_0{}^*\omega = \mathrm{d}(H\omega) + H(\mathrm{d}\omega). \tag{5.51}$$

In het bijzonder, is ω gesloten, dan $\phi_1{}^*\omega - \phi_0{}^*\omega$ exact, dit staat bekend als het *homotopieprincipe*.

De variëteit X heet *samentrekbaar* als er een homotopie als boven bestaat met $Y = X$, $\phi_0(X) = \{a\}$ voor een punt $a \in X$ en ϕ_1 gelijk aan de identiteit in X. Men noemt in dit geval Φ een *samentrekking* (engels: retraction) van X, in zichzelf, naar het punt a. In dit geval geldt voor iedere $x \in X$ dat $\mathrm{T}_x\,\phi_0 = 0$, hetgeen impliceert dat $\phi_0{}^*\omega = 0$ voor iedere differentiaalvorm ω van graad groter of gelijk aan één. De conclusie is daarmee:

Stelling 5.1 *Zij X een samentrekbare C^{k+1}-variëteit, $k \geq 1$ en ω een gesloten C^k-differentiaalvorm van de graad p, met $p \geq 1$. Dan is er een C^k-differentiaalvorm ν van de graad $p-1$, waarvoor $\omega = \mathrm{d}\nu$.*

Opgemerkt kan worden dat als $p = 1$ dan is $\nu \in \mathrm{C}^{k+1}$. Voor willekeurige p kan men ook ν zó kiezen dat de differentieerbaarheidsgraad van ν één meer is dan die van ω. Echter dit geldt voor Hölder- of Sobolev-differentieerbaarheidsgraden en in het bewijs maakt men gebruik van een Laplace-operator werkend op differentiaalvormen; deze technieken zouden ons hier veel te ver voeren.

Voor ons van meer belang is in dit stadium de opmerking dat in een willekeurige variëteit X er bij iedere $a \in X$ een open omgeving U van a in X is en een samentrekking van U, in zichzelf, naar het punt a. Men kan voor U bijvoorbeeld een bol om de oorsprong nemen in een lokaal coördinatensysteem met a als oorsprong. Dan definieert $\phi_t(x) = t\,x$ een samentrekking naar de oorsprong, de *radiële samentrekking* genaamd. Hiermee is het lemma van Poincaré bewezen.

Hiermee besluiten we dit overzicht van de differentiaalmeetkunde. Historische opmerking: hoewel veel van de begrippen al in de achttiende- en negentiende-eeuwse wiskunde min of meer expliciet voorkomen, kreeg de differentiaalmeetkunde haar huidige vorm pas in deze eeuw, in de handen van onder anderen Ricci, Levi-Civita, Poincaré, Élie Cartan, Hermann Weyl, Whitney. Sedert de 1930'er jaren kan deze theorie als standaard beschouwd worden.

5.15 Vraagstukken

Vraagstuk 5.1 Zij g de afbeelding die in Vraagstuk 2.3 is gedefinieerd, zie ook Vraagstuk 1.7. Zij

$$W' = \{(\mu,\, \epsilon) \in \mathbf{R}^6 \mid \mu \neq 0,\ \langle \mu,\, \epsilon \rangle = 0\}.$$

en zij V' de verzameling der $(x,\, v) \in \mathbf{R}^6$ waarvoor x en v lineair onafhankelijke vectoren in $\mathbf{R}^3$ zijn. Bewijs:

a) W' is een 5-dimensionale analytische deelvariëteit van $\mathbf{R}^6$.

b) V' is een open deelverzameling van $\mathbf{R}^6$ en g is een analytische submersie van V' naar W'. Hint: zij $(\widetilde{x}, \widetilde{v}) \in \ker \mathrm{D}g(x, v)$, $(x, v) \in V'$. Bewijs dat $\widetilde{x} \times v + x \times \widetilde{v} = 0$ en leid hieruit af dat $\widetilde{x} = \alpha\, x + \beta\, v$, $\widetilde{v} = \gamma\, x + \delta\, v$ voor zekere $\alpha, \beta, \gamma, \delta \in \mathbf{R}$ en dat $\alpha + \delta = 0$. Verder dat

$$\|x\|^{-1}\, \widetilde{x} - \|x\|^{-3}\, \langle x, \widetilde{x}\rangle\, x + c^{-1}\, (x \times v) \times \widetilde{v} = 0.$$

Leidt hieruit af dat $(\widetilde{x}, \widetilde{v})$ een veelvoud is van de vector $\big(v, -c\, \|x\|^{-3}\, x\big)$.

c) $g(V') = W'$ en de vezels van g in V' zijn 1-dimensionale analytische deelvariëteiten van $\mathbf{R}^6$, gelijk aan de banen van de oplossingen van het tweelichamenprobleem die in V' starten. Deze zijn óf diffeomorf met de cirkel óf diffeomorf met $\mathbf{R}$, al naar gelang de energie < 0 of ≥ 0 is.

d) Het complement $C := V \setminus V'$ is een 4-dimensionale analytische deelvariëteit van $\mathbf{R}^6$. g definieert een surjectieve analytische submersie van C naar de 2-dimensionale analytische deelvariëteit

$$S := \{(0, \epsilon) \mid \|\epsilon\| = 1\}$$

van $\mathbf{R}^6$. De vezels van $g|_C$ zijn 2-dimensionale analytische deelvariëteiten van C en allemaal diffeomorf met $\mathbf{R}^2$.

$\oslash$

Vraagstuk 5.2 Zij $2 \leq r \leq \omega$, U een open deelverzameling van $\mathbf{R}^n$ en g een C^r-afbeelding van U naar $\mathbf{R}^c$. Schrijf $d = n - c$. Zij X de verzameling der $x \in U$ waarvoor $g(x) = 0$ en neem aan dat g een submersie is in de punten van X. Dit impliceert zoals bekend dat X een d-dimensionale C^r-deelvariëteit is van $\mathbf{R}^n$.

Definieer de afbeelding $g' : U \times \mathbf{R}^n \to \mathbf{R}^c \times \mathbf{R}^c$ door middel van

$$g'(x, v) = (g(x), \mathrm{D}g(x) \cdot v)$$

en zij X' de verzameling der $y \in U \times \mathbf{R}^n$ waarvoor $g'(y) = 0$. Bewijs dat g' een C^{r-1} afbeelding is en een submersie in de punten van X'. Bewijs dat X' een $2d$-dimensionale deelvariëteit is van $\mathbf{R}^{2n}$. Geef een C^{r-1}-diffeomorfisme Ψ van X' naar $\mathrm{T}\, X$, waarvoor $\pi \circ \Psi$ gelijk is aan de beperking tot X' van de projectie $(x, v) \mapsto x$ en met de eigenschap dat, voor iedere $x \in X$, de afbeelding $v \mapsto \Psi(x, v)$ een lineair isomorfisme is van $\ker \mathrm{D}g(x)$ naar $\mathrm{T}_x\, X$. Hint: maak gebruik van de identiteit ι, beschouwd als afbeelding van X naar $\mathbf{R}^n$. Bewijs dat voor iedere $x \in X$ de afbeelding $\mathrm{T}_x\, \iota$ een bijectieve lineaire afbeelding is van $\mathrm{T}_x\, X$ naar $\ker \mathrm{D}g(x)$. $\oslash$

Vraagstuk 5.3 Ga na dat als X de positie-ruimte is, de functie L in Hoofdstuk 4 coördinaat--invariant gedefinieerd is als een reëelwaardige functie in een open deelverzameling U van $\mathbf{R} \times \mathrm{T}\, X$. En dat $[L]$ een afbeelding is van een open deelverzameling van $\mathbf{R} \times \mathrm{T}^{(2)}\, X$ naar $\mathrm{T}^*\, X$. Tenslotte dat een krachtveld ϕ een afbeelding is van U naar $\mathrm{T}^*\, X$ en dat γ een oplossing is van de bewegingsvergelijkingen van Lagrange, dan en slechts dan als

$$[T](t, \gamma(t), \gamma'(t), \gamma''(t)) = \phi(t, \gamma(t), \gamma'(t)).$$

Deze notatie is bedoeld als coördinaat-onafhankelijk, maar ligt dicht aan tegen de notatie in lokale coördinaten. $\oslash$

Vraagstuk 5.4 In de 'vector calculus' is 'grad' een operator die aan functies vectorvelden toevoegt, 'rot' een operator die aan een vectorveld in $\mathbf{R}^3$ een vectorveld in $\mathbf{R}^3$ toevoegt en 'div' een operator die aan een vectorveld een functie toevoegt. Zoek de definities hiervan op. We gaan aantonen dat al deze operatoren een interpretatie hebben in termen van de uitwendige afgeleide.

a) Zij v een vectorveld in een open deelverzameling V van $\mathbf{R}^n$. Met het vectorveld v kunnen we ook de éénvorm

$$\alpha_v := \sum_{j=1}^{n} v_j \, \mathrm{d}x_j$$

associëren. Bewijs dat voor iedere differentieerbare functie f in V geldt dat

$$\mathrm{d}f = \alpha_{\mathrm{grad}\, f}.$$

b) Het is ook gebruikelijk om met het vectorveld v in de open deelverzameling V van $\mathbf{R}^n$ de $(n-1)$-vorm $\nu = \nu_v$ in V te associëren, die is gedefinieerd door

$$\nu = \nu_v := \sum_{j=1}^{n} v_j (-1)^{j-1} \, \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_{j-1} \wedge \mathrm{d}x_{j+1} \wedge \ldots \wedge \mathrm{d}x_n.$$

Zij

$$\omega = \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n$$

de standaard volumevorm in V. Bewijs dat $\nu = \mathrm{i}_v \, \omega$ en dat

$$\mathrm{d}\nu = (\mathrm{div}\, v)\, \omega = \mathcal{L}_v \omega.$$

Hiermee krijgt de divergentie van v een dubbele interpretatie: één in termen van de uitwendige afgeleide van de corresponderende $(n-1)$-vorm $\nu = \nu_v$ en één in termen van de Lie-afgeleide naar v van de standaard volumevorm.

c) Zij nu $n = 3$. Bewijs dat

$$\mathrm{d}\,(\alpha_v) = \nu_{\mathrm{rot}\, v}.$$

d) Voor $n = 2$ krijgen we dat $\alpha_v = \nu_{\mathrm{J}\, v}$, waarin $w = \mathrm{J}\, v$ het 'een kwartslag rechtsom gedraaide vectorveld v' is, gedefinieerd door $w_1 = v_2$, $w_2 = -v_1$. Bewijs dat in dit geval

$$\mathrm{d}\,(\alpha_v) = (\mathrm{div}(\mathrm{J}\, v))\, \omega.$$

Voor $n = 2$ ziet men ook wel de formule $\mathrm{rot}\, v = \mathrm{div}(\mathrm{J}\, v)$. Bespreek deze formule.

$\oslash$

6 Geodeten

6.1 (Pseudo-)Riemann'se Variëteiten

Zij X een n-dimensionale C^r-variëteit, $r \geq 3$. Een *Riemann-structuur in X* is een afbeelding g, die aan iedere $x \in X$ een inproduct g_x in $\mathrm{T}_x X$ toevoegt. Vervangen we hierin het woord 'inproduct' door 'niet-gedegenereerde symmetrische bilineaire vorm', dan heet g een *pseudo-Riemann-structuur* in X. Een *(pseudo-)Riemann-variëteit* is een paar (X, g), waarin X een variëteit is en g een (pseudo-)Riemann-structuur in X.

In een kaart κ voor X, met inverse $\mu = \kappa^{-1} : V_\kappa \to X_\kappa$, wordt g gegeven door de bilineaire vorm

$$g_\kappa(x) := (\mathrm{T}_x \, \mu)^* \, g_{\mu(x)}$$

in $\mathbf{R}^n$, die van het punt $x \in V_\kappa$ afhangt. De coëfficiënten hiervan ten aanzien van de standaardbasis e_i in $\mathbf{R}^n$ worden gegeven door

$$g^\kappa_{ij}(x) = g^\kappa(x) \, (e_i, \, e_j) := g_{\mu(x)} \, (\mathrm{T}_x \, \mu \, (e_i) \, , \, \mathrm{T}_x \, \mu \, (e_j)) \, . \tag{6.1}$$

Men noemt g van de klasse C^k, $k \leq r - 1$, als deze coëfficiënten reëelwaardige C^k-functies van $x \in V_\kappa$ zijn Deze voorwaarde is onafhankelijk van de keuze van de kaart, immers als λ een andere kaart is en we schrijven $\phi = \lambda \circ \kappa^{-1}$, dan is

$$g^\kappa = \phi^* \, g^\lambda,$$

hetgeen uitgeschreven betekent dat

$$g^\kappa_{ij}(x) = \sum_{k,\, l=1}^{n} g^\lambda_{kl}(\phi(x)) \, \frac{\partial \phi_k(x)}{\partial x^i} \, \frac{\partial \phi_l(x)}{\partial x^j}. \tag{6.2}$$

Traditioneel schrijft men ook wel g_{ij} in plaats van g^κ_{ij}.

Riemann introduceerde het idee, om in algemene variëteiten de raakruimte in ieder punt x te voorzien van een x-afhankelijk inproduct, in zijn 'Habilitationsvortrag' (1854) [78, p. 272-287]. Hij vond dat er lokale coördinaten zijn waarvoor de coëfficiënten van het inproduct niet van x afhangen, dan en slechts dan als een grootheid, die nu de *krommingstensor van Riemann* genoemd wordt, identiek gelijk aan nul is. Hij vermeldde dit in zijn Habilitationsvortrag zonder formules, die hij vervolgens neerschreef in [78, p. 391-404], een antwoord uit 1861 op een prijsvraag van de Akademie van Parijs. Wél de Habilitationsvortrag maar niet het antwoord op de prijsvraag kennende, behandelde Christoffel [12] de vraag wanneer twee pseudo-Riemann-structuren door middel van een diffeomorfisme in elkaar overgevoerd kunnen worden op een veel grondiger manier. We gaan hier niet verder op deze vraag in, maar beperken ons in plaats daarvan tot het beschrijven van de zogenaamde *geodeten* van een pseudo-Riemann-structuur als de bewegingen van een klassiek mechanisch systeem met 'kinetische energie' gegeven in termen van de pseudo-Riemann-structuur en potentiële energie gelijk aan nul. Vanwege dit laatste spreekt men ook van een *vrij deeltje in een pseudo-Riemann-variëteit.*

6.2 Kortste Verbindingswegen

We beginnen met de bespreking van geodeten in het geval dat g een Riemann-structuur is. In dat geval noteren we

$$\|v\|_x := g_x(v,\, v)^{1/2}, \quad x \in X,\; v \in \mathrm{T}_x\, X \tag{6.3}$$

voor de *norm* of de *lengte* van de raakvector v in het punt x, met betrekking tot de Riemann--structuur g. Voor een continu differentieerbare kromme $\gamma : [a,\, b] \to X$ (of iets algemener, een continue en stuksgewijs continu differentieerbare kromme) wordt de *lengte* van γ gedefinieerd als

$$l(\gamma) = l_g(\gamma) := \int_a^b \|\gamma'(t)\|_{\gamma(t)}\, \mathrm{d}t. \tag{6.4}$$

Voor $x,\, y \in X$ wordt de *afstand* $d(x,\, y) = d_g(x,\, y)$ tussen x en y gedefinieerd als het infimum van de $l(\gamma)$'s, waarbij γ varieert over alle C^1 krommen met $\gamma(a) = x$ en $\gamma(b) = y$. We nemen hierbij aan dat X samenhangend is, men zou $d(x,\, y) = \infty$ kunnen nemen als x en y in verschillende samenhangscomponenten liggen. Het is niet moeilijk om na te gaan dat d voldoet aan de driehoeksongelijkheid $d(x,\, z) \leq d(x,\, y) + d(y, z)$, dat $d(x,\, x) = 0$ en $d(x,\, y) = d(y,\, x)$. We zullen later bewijzen dat er een open omgeving Ω in $X \times X$ is van de *diagonaal*

$$\mathrm{diag}(X) = \{\, (x,\, y) \in X \times X \mid y = x\, \},$$

met de eigenschap dat er voor iedere $(x,\, y) \in \Omega$ precies één C^1-kromme $\gamma : [a,\, b] \to U$ is met $\gamma(a) = x$, $\gamma(b) = y$, $\|\gamma'(t)\|_{\gamma(t)} \equiv 1$ en $d(x,\, y) = l(\gamma) = b - a$. Dat wil zeggen, voor $(x,\, y) \in \Omega$ wordt het infimum van de $l(\gamma)$ aangenomen voor een kromme van x naar y. Dit impliceert dat $d(x,\, y) > 0$ als $(x,\, y) \in \Omega$ en $x \neq y$. We zullen zien dat ook $d(x,\, y) > 0$ als $(x,\, y) \notin \Omega$. We krijgen daarmee dat voor iedere $x,\, y \in X$ met $x \neq y$ geldt dat $d(x,\, y) > 0$, waarmee ook de laatste eigenschap is geverifieerd opdat d een metriek in X is. Bovendien zullen we aantonen dat de topologie in X, die door de metriek wordt gedefinieerd, gelijk is aan de variëteitstopologie van X.

De lengte in (6.4) is een actie-integraal van de vorm (4.10), met Lagrange-functie L gegeven door $L(x,\, v) = \|v\|_x$. Als $d(x,\, y) = l(\gamma)$ voor een kromme γ van x naar y, dan is $[L]_\gamma = 0$. Echter, de Lagrange-functie L heeft een paar nadelen. Ten eerste is deze niet differentieerbaar naar v in de punten $(x,\, v)$ met $v = 0$, dus we beperken ons in de bestudering van de Euler-Lagrange variatievergelijking $[L]_\gamma(t) \equiv 0$ tot krommen waarvoor $\gamma'(t) \neq 0$ voor iedere $t \in [a,\, b]$. Zie (4.8) voor de definitie van $[L]_\gamma(t)$.

Er is echter nog een tweede, ernstiger probleem. De bijbehorende impuls wordt gegeven door

$$\xi = \frac{\partial L(x,\, v)}{\partial v} = \|v\|_x{}^{-1}\, g_x(v). \tag{6.5}$$

Hiervoor geldt dat

$$g_x{}^{-1}(\xi,\, \xi) = 1.$$

Anders gezegd, ξ heeft lengte gelijk aan één met betrekking tot het inproduct $g_x{}^{-1}$ in $(\mathrm{T}_x\, X)^*$. Verder geldt: als $\frac{\partial L(x,v)}{\partial v} = \xi$ en $w = r\, v$ voor een $r > 0$, dan is ook $\frac{\partial L(x,w)}{\partial w} = \xi$. Deze onbepaaldheid is gerelateerd aan het feit dat de tweede-orde afgeleidenmatrix van $L(x,\, v)$ met betrekking tot v de veelvouden van v als nulruimte heeft: de functie L voldoet niet aan de voorwaarde van Legendre. Op zijn beurt betekent dit dat de tweede orde afgeleide $\gamma''(t)$ van γ niet eenduidig bepaald is door de Euler-Lagrange-vergelijking $[L]_\gamma(t) = 0$: ieder veelvoud van $\gamma'(t)$ mag bij $\gamma''(t)$ opgeteld worden om een nieuwe oplossing te krijgen. Anders gezegd, de Euler-Lagrangevergelijkingen definiëren in dit geval *geen expliciet* tweede orde stelsel van differentiaalvergelijkingen, in de zin dat $\gamma''(t)$ een voorgeschreven functie is van t, $\gamma(t)$ en $\gamma'(t)$.

De bovenstaande onbepaaldheid hangt direct samen met het feit dat de lengte van een kromme dezelfde blijft bij een willekeurige herparametrisering van de kromme:

Lemma 6.1 *Is* $\gamma \in \mathrm{C}^1([a, b], X)$ *en is* φ *een stijgend* C^1*-diffeomorfisme van* $[\alpha, g]$ *naar* $[a, b]$, *dan is* $l(\gamma) = l(\gamma \circ \varphi)$.

Bewijs De substitutie van variabelen $t = \varphi(\tau)$ in de integraal geeft

$$\begin{aligned} l(\gamma) &= \int_a^b \|\gamma'(t)\|_{\gamma(t)}\, dt = \int_\alpha^g \|\gamma'(\varphi(\tau))\|_{\gamma(\varphi(\tau))} \cdot \varphi'(\tau)\, \mathrm{d}\tau \\ &= \int_\alpha^g \|\varphi'(\tau) \cdot \gamma'(\varphi(\tau))\|\, \mathrm{d}\tau = \int_\alpha^g \|(\gamma \circ \varphi)'(\tau)\|\, \mathrm{d}\tau = l(\gamma \circ \varphi). \end{aligned}$$

□

Men kan deze onbepaaldheid ook gebruiken, daarmee de onbepaaldheid eliminerend, om over te gaan op krommen die *door booglengte geparametriseerd* zijn, dat wil zeggen, waarvoor $\|\gamma'(t)\| \equiv 1$. De naam wordt daarbij verklaard door de identiteit

$$l\left(\gamma|_{[a,\,t]}\right) \equiv t - a,$$

welke equivalent is met de voorwaarde dat $\|\gamma'(t)\| \equiv 1$. En vervolgens kan men opmerken dat voor dergelijke krommen de Euler-Lagrange-vergelijking $[L]_\gamma(t) \equiv 0$ equivalent is met de Euler-Lagrange-vergelijking $[T]_\gamma(t) \equiv 0$, waarin

$$T(x, v) = T_g(x, v) := \frac{1}{2}\, g_x(v, v) = \frac{1}{2}\, \|v\|_x{}^2 \tag{6.6}$$

de *kinetische energie* voorstelt, die gedefinieerd is in termen van de Riemann-structuur g. Merk op dat T een constante van beweging is voor het bijbehorende Euler-Lagrange stelsel, omdat

$$H_T(x, v) := \frac{\partial T(x, v)}{\partial v} \cdot v - T(x, v) = 2T(x, v) - T(x, v) = T(x, v) \tag{6.7}$$

en het linkerlid hierin is een constante van beweging, zie (4.22). Voor het bewijs van de equivalentie van $[L] = 0$ met $[T] = 0$ merken we op dat $T = \frac{1}{2}L^2$ geeft dat $\mathrm{d}T = L\ \mathrm{d}L$. Is $L(\gamma(t), \gamma'(t)) \equiv 1$, dan geeft dit dat $[L]_\gamma(t) \equiv [T]_\gamma(t)$. Omgekeerd, als $[T]_\gamma(t) \equiv 0$, dan is $T(\gamma(t), \gamma'(t))$ constant als functie van t en geeft $\mathrm{d}L = (2T)^{-1/2}\ \mathrm{d}T$ dat $[L]_\gamma(t) \equiv 0$. De parametrisering met booglengte wordt hier gekarakteriseerd door het niveau van de kinetische energie gelijk aan $1/2$ te nemen. De oplossingen op de andere energieniveau's representeren de oplossingen van $[L]_\gamma(t) \equiv 0$ die geparametriseerd worden door een constante maal de booglengte.

6.3 Geodeten in Pseudo-Riemann-variëteiten

Een extra bonus van de overgang van de lengtefunctie naar de kinetische energie is dat de bijbehorende Euler-Lagrange-vergelijkingen daarvoor ook goed gedefinieerd zijn voor een pseudo-Riemann--structuur, waarbij de niet-gedegenereerdheid van g_x ervoor zorgt dat het stelsel expliciet is en de snelheids-impuls-afbeelding een diffeomorfisme is van $\mathrm{T}\,X$ naar $\mathrm{T}^*\,X$.

Definitie 6.1 Zij $(X,\, g)$ een pseudo-Riemannvariëteit. De *geodeten in* X zijn de oplossingen γ van de Euler-Lagrange-vergelijking $[T]_\gamma(t) \equiv 0$, waarin de functie T op $\mathrm{T}\,X$ is gedefinieerd door (6.6). De bijbehorende stroming in $\mathrm{T}\,X$ heet de *geodetische stroming in* $\mathrm{T}\,X$.

Voorbeeld 6.1 Zij X een n-dimensionale C^r-deelvariëteit van $\mathbf{R}^N$. Als $x \in X$ en $v,\, w \in \mathrm{T}_x\,X$, dan kunnen v en w ook als element van $\mathbf{R}^N$ gezien worden. Definiëren we $g_x(v,\, w) = \langle v,\, w\rangle$ dan is g een C^{r-1} Riemann-structuur in X, die de *door de standaard Riemann-structuur van* $\mathbf{R}^N$ *geïnduceerde Riemann-structuur in* X genoemd wordt.

Men kan dit ook als volgt formuleren. Zij ι de identiteit, beschouwd als afbeelding van X naar $\mathbf{R}^N$, de standaard inbedding van X in $\mathbf{R}^N$. Dan is $\mathrm{T}\iota$ een C^{r-1}-afbeelding van $\mathrm{T}\,X$ naar $\mathrm{T}\,\mathbf{R}^N = \mathbf{R}^N \times \mathbf{R}^N$, de standaard inbedding van $\mathrm{T}\,X$ in $\mathbf{R}^N \times \mathbf{R}^N$. Zij $T_{_N}(x,\, v) = \langle v,\, v\rangle/2$ de standaard kinetische energie, beschouwd als functie op $\mathbf{R}^N \times \mathbf{R}^N$. Dan is de kinetische energie T van de geïnduceerde Riemann-structuur g gelijk aan $T_{_N} \circ \mathrm{T}\iota$. Omdat g eenduidig bepaald is door T, is hiermee g vastgelegd.

Vanwege Gevolg 4.2 worden de geodeten γ in X nu gekarakteriseerd door de eigenschap dat voor iedere t de beperking van $[T_{_N}]_\gamma(t)$ tot $\mathrm{T}_{\gamma(t)}\,X$ gelijk is aan nul. Dat wil zeggen,

$$\langle \gamma''(t),\, w\rangle = 0, \quad w \in \mathrm{T}_{\gamma(t)}\,X, \tag{6.8}$$

ofwel: *de versnelling staat loodrecht op de raakruimte aan* X.

Het geval van oppervlakken in de drie-dimensionale Euclidische ruimte, dus $n = 2$ en $N = 3$, werd diepgaand geanalyseerd door Gauss [26], waarbij (6.8) al door Euler is opgemerkt. $\oslash$

Voorbeeld 6.2 In Einstein's relativiteitstheorie wordt het gravitatieveld gegeven door een pseudo-Riemann-structuur in de vierdimensionale positie-tijdsruimte, waarbij de tekens $-,\, -,\, -,\, +$ zijn, in de zin dat iedere raakruimte een basis e_j heeft, waarvoor $g_x\,(e_i,\, e_j) = 0$ als $i \neq j$, $g_x\,(e_i,\, e_i) = -1$ als $1 \leq i \leq 3$ en $g_x\,(e_4,\, e_4) = c^2$. Het is in de relativiteitstheorie overigens gebruikelijk om e_0 in plaats van e_4 te schrijven. Volgens de relativiteitstheorie beweegt een testdeeltje zich in het gravitatieveld langs een geodeet, waarbij $g_x(w,\, w) > 0$ correspondeert met een 'ruimte-achtige beweging met eindige energie' en $g_x(w,\, w) = 0$ een 'lichtdeeltje' voorstelt. Hierin is w een raakvector aan de positie-tijdsruimte. Zie bijvoorbeeld Landau and Lifschitz [53, Ch. 10].

Interessant is het nog om hierbij te zien hoe Einstein een Lagrange-functie L introduceerde in analogie met de norm van de snelheid. Schrijf δt voor de vierde component van de raakvector w en δx voor het 'positiedeel'. Neem

$$L = \alpha\, g(v,\, v)^{1/2} = \alpha\, \left(c^2\,(\delta t)^2 - \|\delta x\|^2\right)^{1/2},$$

waarin α een nader te bepalen constante is. Parametriseren we de kromme met t, dan wordt $\delta t = 1$ en $\delta x = v$ de snelheid in de positie-ruimte. Taylor-ontwikkeling in $v = 0$ geeft dat

$$L = \alpha\, c - \tfrac{\alpha}{2c}\, v^2 + \mathcal{O}\left(\|v\|^4\right)$$

als $v \to 0$. Omdat bij kleine snelheden, de 'niet-relativistische limiet', de energie moet toenemen als $\frac{1}{2} m \|v\|^2$, is de conclusie dat $\alpha = -m\, c$, ofwel

$$L = -m\, c^2 \left(1 - \frac{\|v\|^2}{c^2}\right)^{1/2}.$$

Voor $v = 0$ geeft dit $L = -m\, c^2$. De impuls wordt in dit geval gegeven door

$$p = \frac{\partial L}{\partial v} = m \left(1 - \frac{\|v\|^2}{c^2}\right)^{-1/2} v.$$

De snelheids-impulsafbeelding is hier een diffeomorfisme van de begrensde open verzameling van de v met $\|v\| < c$ naar de gehele $\mathbf{R}^3$. Zie bijvoorbeeld Landau and Lifschitz [53, Ch. 1].

Als we het gravitatieveld opgewekt denken door materie, waarvan de beweging op zijn beurt weer bepaald wordt door het gravitatieveld, dan wordt het begrijpelijk dat niet iedere pseudo-Riemann--structuur met tekens $-, -, -, +$ toegelaten is als een gravitatiestructuur. In de (algemene) relativiteitstheorie wordt nog een stelsel van (niet-lineaire) partiële differentiaalvergelijkingen opgeschreven waaraan de coëfficiënten van g moeten voldoen, de zogenaamde *Einstein-vergelijkingen*. Zie bijv. Landau and Lifschitz [53, p. 299]. Ze kunnen als evolutievergelijkingen voor g opgevat worden en spelen een centrale rol in de cosmologie, zie bijv. Weinberg [87, Ch. 7]. Als dynamisch systeem in een oneindig-dimensionale ruimte van pseudo-Riemann-structuren heeft het eigenschappen die het doen lijken op een Hamilton-systeem (de definitie van Hamilton-systemen wordt in het volgende hoofdstuk gegeven), zie bijvoorbeeld Abraham and Marsden [3, Sec. 5.5.10]. Differentiaalmeetkundigen herkennen in de Einstein-vergelijkingen de zogenaamde *Ricci-kromming* van g als één van de termen, maar erkennen dat hiermee de differentiaalmeetkundige betekenis van de Einstein-vergelijkingen nog niet meteen zo duidelijk is, zie Besse [10]. Voorshands vallen de Einstein-vergelijkingen buiten ons kader van klassieke mechanische systemen met eindig veel vrijheidsgraden, terwijl de beweging van testdeeltjes er juist helemaal binnen valt omdat het daarbij gaat om de geodeten voor de als gegeven veronderstelde pseudo-Riemann-structuur g. $\oslash$

In lokale coördinaten zien de Euler-Lagrange-vergelijkingen voor de geodeten er als volgt uit: $\frac{dx}{dt} = v$ en

$$\sum_{i=1}^{n} g_{ik}(x) \frac{dv^i}{dt} + \sum_{i,\,j=1}^{n} \frac{\partial g_{ik}(x)}{\partial x^j} v^i\, v^j = \frac{1}{2} \sum_{i,\,j=1}^{n} \frac{\partial g_{ij}(x)}{\partial x^k} v^i\, v^j, \quad 1 \leq k \leq n. \tag{6.9}$$

Opmerking 6.1 De notatie met bovenindices voor x en v en benedenindices voor (multi)lineaire vormen op de raakruimte is standaard in de (pseudo-)Riemann'se differentiaalmeetkunde. Verder gebruikt men ook nog vaak de *sommatieconventie van Einstein*, waarbij het somteken weggelaten wordt en gesommeerd wordt over indices die in paren voorkomen. De notatie met bovenindices kan (alleen) tot verwarring aanleiding geven als men de grootheden als machten zou kunnen interpreteren; dat gevaar valt hier wel mee. Einstein's sommatieconventie kan tot verwarring leiden als men een term apart wil bekijken waarvoor twee van de indices aan elkaar gelijk zijn. Ik vind het niet erg en eigenlijk wel zo duidelijk om de somtekens erbij te schrijven. $\oslash$

De tweede term in het linkerlid van (6.9) is een kwadratische vorm in v, die geschreven kan worden als $\sum_{i,j} s_{ij}\, v^i\, v^j$ met een *symmetrische* matrix s_{ij}, die dan gelijk is (moet zijn) aan

$$s_{ij} = \frac{1}{2}\,\frac{\partial g_{ik}(x)}{\partial x^j} + \frac{1}{2}\,\frac{\partial g_{jk}(x)}{\partial x^i}.$$

(Deze matrix hangt ook nog af van k.) De uitdrukking voor dv/dt in termen van x en v kan gevonden worden door gebruik te maken van de inverse matrix van $g_{ij}(x)$, die genoteerd wordt als

$$\left(g_x{}^{-1}\right)_{ij} = g^{ij}(x). \tag{6.10}$$

Het resultaat is:

$$\frac{dv^l}{dt} + \sum_{i,\,j=1}^{n} \Gamma^l_{ij}(x)\, v^i\, v^j = 0, \tag{6.11}$$

waarin

$$\Gamma^l_{ij}(x) := \sum_{k=1}^{n} g^{lk}(x)\,\frac{1}{2}\left\{\frac{\partial g_{ik}(x)}{\partial x^j} + \frac{\partial g_{jk}(x)}{\partial x^i} - \frac{\partial g_{ij}(x)}{\partial x^k}\right\} \tag{6.12}$$

de zogenaamde *Christoffel-symbolen* zijn.

In (6.12) zijn de Christoffel-symbolen uitgedrukt in termen van de eerste orde partiële afgeleiden van de coëfficiënten van de pseudo-Riemann-structuur. Het is een enigszins verrassende opmerking van Christoffel [12] dat ook een omkering geldt:

$$\frac{\partial g_{jk}(x)}{\partial x_i} = \sum_{l=1}^{n} g_{kl}(x)\,\Gamma^l_{ij}(x) + \sum_{l=1}^{n} g_{jl}(x)\,\Gamma^l_{ki}(x). \tag{6.13}$$

Voor het bewijs schrijven we

$$\Gamma_{k,\,ij} := \sum_{l=1}^{n} g_{kl}(x)\,\Gamma^l_{ij}(x) = \frac{1}{2}\{\frac{\partial g_{ik}(x)}{\partial x^j} + \frac{1}{2}\,\frac{\partial g_{jk}(x)}{\partial x^i} - \frac{\partial g_{ij}(x)}{\partial x^k}\}.$$

De symmetrie van g gebruikend, concluderen we dat

$$\Gamma_{k,\,ij} + \Gamma_{j,\,ki} = \frac{\partial g_{jk}(x)}{\partial x^i},$$

dus dat (6.13) geldt.

Het feit dat dv/dt een kwadratische functie van v is heeft tot gevolg dat als γ een geodeet in X is en r een reële constante, dan is ook $t \mapsto \gamma(r\,t)$ een geodeet. Dit is een speciaal geval van de volgende algemene opmerking.

Lemma 6.2 *Voor een stelsel van de vorm* $d^2x/dt^2 = a(x,\, dx/dt)$ *en constante* $r \in \mathbf{R}$ *zijn de volgende uitspraken equivalent.*

a) *Als* γ *een oplossing is, dan is* $t \mapsto \gamma(r\,t)$ *een oplossing.*

b) *Voor alle* $(x,\, v)$ *in het definitiegebied* D *van* a *geldt dat* $(x,\, r\,v) \in D$ *en* $a(x,\, r\,v) = r^2\, a(x,\, v)$.

Bewijs Dit volgt uit $\frac{d\gamma(r\,t)}{dt} = r\,\gamma'(r\,t)$ en $\frac{d^2\gamma(r\,t)}{dt^2} = r^2\,\gamma''(r\,t)$. $\square$

Opgemerkt kan nog worden dat als voor iedere $r > 0$ geldt dat $a(x,\, r\,v) = r^2\,a(x,\, v)$ en $v \mapsto a(x,\, v)$ is een tweemaal differentieerbare afbeelding in $v = 0$, dan is $v \mapsto a(x,\, v)$ een homogene veelterm van de graad twee, dus

$$a^l(x,\, v) = -\sum_{i,\,j=1}^{n} \Gamma^l_{ij}(x)\,v_j\,v_j, \tag{6.14}$$

voor eenduidig bepaalde $\Gamma^l_{ij}(x)$ met $\Gamma^l_{ij}(x) = \Gamma^l_{ji}(x)$. Immers, Taylor-ontwikkeling tot en met de orde twee geeft dat

$$a^l(x,\, v) = a^l(x,\, 0) + \sum_{i=1}^{n} b^l_i(x)\,v^i + \sum_{i,\,j} c^l_{i,\,j}(x)\,v^i\,v^j + R^l(x,\, v), \tag{6.15}$$

waarin b^l_i en c^l_{ij} de eerste, resp. tweede orde afgeleiden van a^l in $v = 0$ voorstellen en de restterm voldoet aan de voorwaarde dat $R^l(x,\, v)/\|v\|^2$ naar nul convergeert als $v \to 0$, $v \neq 0$. Vervang nu in (6.15) overal v door $r\,v$ en vul in het linkerlid de homogeniteit van a in. Dit geeft

$$r^2\,a^l(x,\, v) = a^l(x,\, 0) + r\sum_{i=1}^{n} b^l_i(x)\,v^i + r^2\sum_{i,\,j} c^l_{i,\,j}(x)\,v^i,\, v^j + r^2\,\|v\|^2\,\|r\,v\|^{-2}\,R^l(x,\, r\,v), \tag{6.16}$$

Laten we hierin $r \downarrow 0$ gaan, dan krijgen we dat $a^l(x,\, 0) = 0$. Delen we nu het linker-en rechterlid in (6.16) door r en laten we $r \downarrow 0$ gaan, dan krijgen we dat $b^l_i(x) = 0$. Delen we tenslotte het linker-en rechterlid in (6.16) door r^2, dan zien we dat $\|r\,v\|^{-2}R^l(x,\, r\,v)$ enerzijds niet van r afhangt en anderzijds naar nul convergeert als $r \downarrow 0$. Dit kan alleen als $R^l(x,\, v) = 0$, dus (6.14) geldt met $\Gamma^l_{ij}(x) = -c^l_{ij}(x)$.

Als gevolg van deze homogeniteit geldt voor alle oplossingen van $x'' = a(x,\, x')$ die door een gegeven punt gaan en waarvan de banen dezelfde raaklijn in het punt hebben dat ze dezelfde baan hebben, in iedere richting door het punt gaat één baan. Om deze reden noemt men een een tweede-orde stelsel $x'' = a(x,\, x')$ met versnellingsfunctie a van de vorm (6.14) een *sproei*. Worden daarbij de Christoffel-symbolen gegeven door (6.12), voor een pseudo-Riemann-structuur g, dan spreekt men van een *geodetische sproei*.

Opmerking 6.2 Het begrip sproei (engels: spray) is afkomstig van Ambrose, Palais en Singer [4] en heeft enige bekendheidheid gekregen door de uiteenzetting in Lang [54, Ch. IV, §3]. Niet iedere sproei is geodetisch (met betrekking tot een pseudo-Riemann-structuur) en een pseudo-Riemann'se geodetische sproei is niet altijd van een Riemann-structuur afkomstig. Ook is de (pseudo-)Riemann-structuur niet altijd eenduidig bepaald door de sproei (omgekeerd natuurlijk wel). Op deze kwesties gaan we hier niet verder in. $\oslash$

6.4 De Exponentiële Afbeelding

Voor een sproei $x'' = a(x,\, x')$ met a als in (6.14) en C^1-coëfficiënten $\Gamma^l_{ij}(x)$, zien we dat alle $(x,\, v)$ met $v = 0$ nulpunten zijn van het bijbehorende vectorveld in de faseruimte $\mathrm{T}\,X$. De collectie van de $(x,\, 0) \in \mathrm{T}\,X$ met $x \in X$ heet ook wel de *nulsnede* N van de raakbundel $\mathrm{T}\,X$ van X, dit is een n-dimensionale C^{r-1}-deelvariëteit van $\mathrm{T}\,X$ en de afbeelding $x \mapsto (x,\, 0)$ is evident een diffeomorfisme van X naar N, met $\pi|_N$ als inverse. Het is nogal bijzonder dat een vectorveld zo'n grote variëteit van nulpunten heeft, zelfs Hamilton-stelsels hebben gewoonlijk slechts geïsoleerde nulpunten.

Definitie 6.2 Zij G^t de stroming na tijd t in de faseruimte $\mathrm{T}\,X$ die is gedefinieerd door de sproei $x'' = a(x,\, x')$, met definitiegebied $D(t)$. Voor iedere $x \in X$, $v \in \mathrm{T}_x\,X \cap D(t)$ definiëren we

$$\exp_x(v) = \pi\left(G^1(x,\, v)\right). \tag{6.17}$$

De afbeelding $\exp_x$ van $\mathrm{T}_x\,X \cap D(t)$ naar X heet de *exponentiële afbeelding met middelpunt in x*. De afbeelding

$$\exp : (x,\, v) \mapsto (x,\, \exp_x(v))$$

van $D(t)$ naar $X \times X$ heet de *collectieve exponentiële afbeelding* van de sproei $x'' = a(x,\, x')$.

Lemma 6.3 *Zij $x'' = a(x,\, x')$ een sproei in een variëteit X, waarbij het bijbehorende vectorveld in de faseruimte gedefinieerd is in een open omgeving van de nulsnede N in $\mathrm{T}\,X$ en van de klasse C^k is met $1 \leq k \leq r-1$. Zij D het definitiegebied van de stroming na tijd 1. Dan is D een open omgeving van N in $\mathrm{T}\,X$ en de collectieve exponentiële afbeelding is een C^k-afbeelding van D naar $X \times X$.*

Voor iedere $x \in X$ is $\mathrm{T}_0 \exp_x = I =$ de identiteit in $\mathrm{T}_x\,X$ en er is een open omgeving E_x van 0 in $D \cap \mathrm{T}_x\,X$, waarvoor de beperking van $\exp_x$ tot E_x een C^k-diffeomorfisme is van E_x naar een open omgeving U_x van x in X. Voor iedere $v \in \mathrm{T}_x\,X$ is $t \mapsto \exp_x(t\,v)$ de oplossing van $x'' = a(x,\, x')$ met $x(0) = x$ en $x'(0) = v$.

Bij iedere $x_0 \in X$ is er een open omgeving E van $(x_0,\, 0)$ in $\mathrm{T}\,X$, waarvoor de beperking tot E van de collectieve exponentiële afbeelding een C^k-diffeomorfisme is van E naar een open omgeving U van $(x_0,\, x_0)$ in $X \times X$.

Bewijs De eerste uitspraak is een gevolg van het feit dat G^1 een C^k-afbeelding is van D naar $\mathrm{T}\,X$.

In lokale coördinaten is de linearisatie van het stelsel in het nulpunt $(x,\, 0)$ van de vorm

$$\frac{d(\delta x)}{dt} = \delta v, \quad \frac{d(\delta v)}{dt} = 0.$$

Hiervan worden de oplossingen gegeven door $\delta v(t) = \delta v(0)$ (constant) en $\delta x(t) = \delta x(0) + t\,\delta v(0)$. Hieruit lezen we af dat de totale afgeleide van exp in het punt $(x,\, 0)$ gegeven wordt door

$$(\delta x,\, \delta v) \mapsto (\delta x + \delta v,\, \delta v). \tag{6.18}$$

Door hierin $\delta x = 0$ in te vullen zien we ook dat de totale afgeleide van $\exp_x$ in $v = 0$ gelijk is aan de identiteit. De tweede en de laatste uitspraak volgen door een toepassing van de impliciete functiestelling op de vergelijking $\exp_x(v) = \exp(x,\, v) = y$, waarbij v in de buurt van 0, resp. $(x,\, v)$ in de buurt van $(x_0,\, 0)$ als parameters worden opgevat.

Tenslotte, is γ een geodeet met $\gamma(0) = x$ en $\gamma'(0) = v$, dan is $\delta : t \mapsto \gamma(c\,t)$ ook een geodeet. Hiervoor geldt dat $\delta(0) = x$ en $\delta'(0) = c\,v$. Invullen van $t = 1$ geeft dat voor iedere c geldt dat $\gamma(c) = \exp_x(c\,v)$. □

In de notatie van Lemma 6.3 is de inverse van $\exp_x$ een C^k-diffeomorfisme van U_x naar E_x. Keuze van een basis in $\mathrm{T}_x X$ leidt tot een lineaire coördinatisering $\mathrm{T}_x X \to \mathbf{R}^n$, door die achter ${\exp_x}^{-1}$ te schakelen krijgen we een kaart in U_x. Deze kaart heet het *normale coördinatensysteem met middelpunt* x (en ten aanzien van de keuze van de basis in $\mathrm{T}_x X$). Voor oppervlakken X in $\mathbf{R}^3$ gebruikte Gauss [26, Art. 15] de afbeelding ${\exp_x}^{-1}$, gevolgd door poolcoördinaten in $\mathrm{T}_x X$ ten aanzien van een orthonormale basis in $\mathrm{T}_x X$. Gauss beperkte daarbij de beschouwing tot dat deel van X dat hiermee gecoördinatiseerd wordt, maar gaf geen bewijs dat dit een omgeving van het punt x vormt. Ik weet niet of hij dit vanzelfsprekend vond of dat hij zich eenvoudigweg de vraag niet stelde.

Lemma 6.4 *De coördinaten $x_1, \ldots, x_n$ in een omgeving S van de oorsprong in $\mathbf{R}^n$ zijn normaal dan en slechts dan als voor alle $x \in S$ geldt dat*

$$\sum_{i,\, j=1}^{n} \Gamma_{ij}^{l}(x)\, x^i\, x^j = 0, \quad 1 \leq l \leq n. \tag{6.19}$$

Is dit het geval, dan geldt voor iedere $1 \leq i,\, j,\, l \leq n$ dat $\Gamma_{ij}^{l}(0) = 0$. In het geval van de geodetische sproei voor een pseudo-Riemann-structuur g geldt dat in normale lokale coördinaten alle eerste orde partiële afgeleiden in de oorsprong van alle coëfficiënten van g gelijk aan nul zijn.

Bewijs In normale coördinaten zijn de geodeten door de oorsprong de krommen $\gamma : t \mapsto t\, x$. De vergelijking (6.11) voor $t = 1$ wordt dan (6.19).

Omgekeerd, is aan (6.19) voldaan dan geeft daarin vervangen van x door $t\, x$ en delen door t^2 dat $t \mapsto t\, x$ voldoet aan (6.11), hetgeen betekent dat we een normaal coördinatensysteem hebben.

Toepassen van $\partial^2 / \partial x_p \partial x_q$ op (6.19) in $x = 0$ en gebruikmaken van de symmetrie van de Christoffel-symbolen leidt tot $0 = \Gamma_{pq}^{l}(0) + \Gamma_{qp}^{l}(0) = 2\, \Gamma_{pq}^{l}(0)$. De laatste bewering volgt door toepassing van (6.13). $\square$

Het volgende lemma is, voor oppervlakken in $\mathbf{R}^3$, afkomstig van Gauss [26, Art. 15]. In normale coördinaten vervangen we x door 0 en v door x en krijgen we dat

$$\sum_{i=1}^{n} g_{ij}(0)\, x^i = \sum_{i=1}^{n} g_{ij}(x)\, x^i, \quad 1 \leq j \leq n. \tag{6.20}$$

Anders gezegd: $g_0(x,\, w) = g_x(x,\, w)$ voor alle $w \in \mathbf{R}^n$.

Lemma 6.5 *Zij $x \in X$ en $v \in \mathrm{T}_x X$ in het definitiegebied van $\exp_x$. Schrijf $\varepsilon = \exp_x$. Dan is*

$$g_x(v,\, w) = g_{\varepsilon(v)} \left(\mathrm{T}_v\, \varepsilon(v),\, \mathrm{T}_v\, \varepsilon(w) \right), \quad w \in \mathrm{T}_x X. \tag{6.21}$$

Bewijs Schrijf $\gamma = \gamma_v$ voor de geodeet met $\gamma(0) = x$ en $\gamma'(0) = v$. We hebben na (6.6) al opgemerkt dat de kinetische energie een constante van beweging is, dus $T(x,\, v) = T(\gamma_v(t),\, \gamma_v'(t))$, hetgeen impliceert dat

$$T(x,\, v) = \int_0^1 T(\gamma_v(t),\, \gamma_v'(t))\, dt.$$

Differentiatie hiervan naar v in de richting van $w \in \mathrm{T}_x X$ geeft volgens (4.13) dat

$$g_x(v,\, w) = g_{\gamma(1)}\left(\gamma'(1),\, \mathrm{T}_v \exp(w)\right), \tag{6.22}$$

au waarbij we hebben gebruikt dat $[T]_\gamma = 0$. Hierbij hebben we gebruikt dat de afgeleide van $T(x,\, v)$ met betrekking tot v in de richting van w gelijk is aan $g_x(v,\, w)$. Anders gezegd, de impuls in $(x,\, v)$ is gelijk aan de lineaire vorm $w \mapsto g_x(v,\, w)$. Verder is de randterm bij $t = 0$ gelijk aan nul omdat $\gamma_v(0) = x$ niet van v afhangt. Tenslotte is de randterm bij $t = 1$ gelijk aan de impuls in $(\gamma(t),\, \gamma'(t))$ losgelaten op de afgeleide van $\gamma_v(1) = \epsilon(v)$ met betrekking tot v in de richting van de vector w, dus gelijk aan het rechterlid in (6.22). (Dit argument is analoog aan het bewijs van Stelling 4.4.)

De formule (6.21) volgt nu uit (6.22) door op te merken dat $\gamma_v(t) = \exp_x(t\, v) = \epsilon(t\, v)$, zie Lemma 6.3. □

6.5 De Injectiviteitsstraal

We beperken ons nu tot de beschouwing van een Riemann-variëteit $(X,\, g)$. Voor iedere $x \in X$ en $r \in \mathbf{R}_{\geq 0}$ noteren we met $B_x(r)$ de verzameling der $v \in \mathrm{T}_x X$ waarvoor $g_x(v,\, v) < r^2$, dit is de open bol in $\mathrm{T}_x X$ met straal r en de oorsprong als middelpunt, met betrekking tot de Riemann-structuur g. Verder definiëren we $r(x)$ als het supremum van de $r \in \mathbf{R}_{\geq 0}$ met de volgende eigenschappen:

a) De bol $B_x(r)$ is bevat in het definitiegebied D_x van $\exp_x$.

b) De beperking van $\exp_x$ tot $B_x(r)$ is injectief.

c) Voor iedere $v \in B_x(r)$ is $\mathrm{T}_v \exp_x$ bijectief.

Het getal $r(x)$ heet de *injectiviteitsstraal* in het punt x. Men moet erop bedacht zijn dat deze vaak eindig is. Dat wil zeggen dat minstens één van de volgende drie dingen kan gebeuren:

a) Een geodeet die in x start loopt in een begrensde tijd uit iedere compacte deelverzameling van X weg.

b) $\exp_x$ is niet injectief.

c) Er is een $v \in \mathrm{T}_x X$ waarvoor $\mathrm{T}_v \exp_x$ niet bijectief is.

Als toepassing van het begrip injectiviteitsstraal kunnen we de de laatste uitspraak in Lemma 6.3 nu in een sterkere vorm gieten.

Lemma 6.6 *Zij $(X,\, g)$ een Riemann-variëteit en zij E het inwendige van de verzameling van alle $(x,\, v) \in \mathrm{T}\, X$ met $v \in B_x(r(x))$. Zij ε de beperking tot E van de collectieve exponentiële afbeelding. Dan is E een open omgeving van de nulsnede N in $\mathrm{T}\, X$ en is ε een C^k-diffeomorfisme van E naar een open omgeving U in $X \times X$ van de $\mathrm{diag}(X)$ = de verzameling der $(x,\, y) \in X \times X$ waarvoor $y = x$.*

Bewijs Dat E een omgeving is van N in $\mathrm{T}\, X$ volgt uit de laatste uitspraak van Lemma 6.3.

Als $(x,\, v) \in E$, $(x',\, v') \in E$ en $\exp(x,\, v) = \exp(x',\, v')$ dan geeft een beschouwing van de eerste factor dat $x = x'$ en vervolgens geeft een beschouwing van de tweede factor dat $\exp_x(v) = \exp_x(v')$.

Omdat $v,\, v' \in B_x(r(x))$ geeft eigenschap b) van de injectiviteitsstraal nu dat $v = v'$. De conclusie is dat ε injectief is.

In lokale coördinaten krijgen we voor iedere $(x,\, v) \in D$ dat er een lineaire afbeelding $C = C(x,\, \xi)$ is, waarvoor

$$\mathrm{T}_{(x,\, v)} \exp(\delta x,\, \delta v) = (\delta x,\, C\, \delta x + \mathrm{T}_v \exp_x(\delta v))\,.$$

Is deze uitdrukking gelijk aan nul dan geeft een beschouwing van de eerste factor dat $\delta x = 0$ en vervolgens een beschouwing van de tweede factor dat $\mathrm{T}_v \exp_x(\delta v) = 0$. Als $(x,\, v) \in E$ dan geeft eigenschap c) van de injectiviteitsstraal dat $\delta v = 0$. De conclusie is dat voor iedere $(x,\, v) \in E$ de lineaire afbeelding $\mathrm{T}_{(x,\, v)}\, \varepsilon$ injectief is, dus bijectief.

De globale inverse functiestelling geeft nu dat $U := \exp(E) = \varepsilon(E)$ een open deelverzameling is van $X \times X$ en dat ε een C^k-diffeomorfisme is van E naar U. Verder is $\mathrm{diag}(X) = \exp(N) \subset \exp(E) = U$. $\square$

Lemma 6.5 leidt in het geval van een Riemann-structuur tot de conclusie dat, althans lokaal, de geodeten inderdaad de kortste krommen tussen punten zijn.

Lemma 6.7 *Zij g een Riemann-structuur. In normale coördinaten waarin $g_{ij}(0) = \delta_{ij}$ geldt voor iedere continue en stuksgewijs continu differentieerbare kromme $\gamma : [a,\, b] \to \mathbf{R}^n$ met $\gamma(a) = 0$, dat $\|\gamma(b)\| \le l(\gamma)$. Er geldt gelijkheid, dan en slechts dan als er een monotoon niet-dalende functie $\mu : [a,\, b] \to [0,\, 1]$ is met $\gamma(t) = \mu(t)\, \gamma(b)$. Dat wil zeggen, dan en slechts dan als na herparametrisering γ gelijk is aan een geodeet.*

Bewijs Zij x in het beeldgebied van de normale coördinaten, $x \neq 0$. Iedere $v \in \mathbf{R}^n$ kan dan geschreven worden als $v = \alpha\, x + w$, waarin $\alpha \in \mathbf{R}$ en $g_x(x,\, w) = 0$. Uit $g_x(x,\, v) = \alpha\, g_x(x,\, x)$ lezen we af dat we hiertoe moeten nemen $\alpha = g_x(x,\, v)/g_x(x,\, x)$ en met deze keuze van α zien we dat $g_x(x,\, w) = 0$ als $w = v - \alpha\, x$. Voor deze α en w geldt vervolgens dat

$$g_x(v,\, v) = \alpha^2\, g_x(x,\, x) + g_x(w,\, w) = g_x(x,\, v)^2/g_x(x,\, x) + g_x(w,\, w),$$

dus

$$g_x(v,\, v)^{1/2} \ge g_x(x,\, v)/g_x(x,\, x)^{1/2} = \langle x,\, v\rangle/\|x\|.$$

met gelijkteken dan en slechts dan als $w = 0$ en $g_x(x,\, v) \ge 0$, ofwel dan en slechts dan als er een $\lambda \ge 0$ is met $v = \lambda\, x$. De identiteit aan het eind volgt uit Lemma 6.5.

Als $\gamma(b) = 0$ dan is de uitspraak flauw, we nemen daarom nu aan dat $\gamma(b) \neq 0$. Zij $\widetilde{a}$ het supremum van de verzameling der $t \in [a,\, b]$ waarvoor $\gamma(t) = 0$. Dan is $a \le \widetilde{a} < b$, $\gamma(\widetilde{a}) = 0$ en voor iedere $\widetilde{a} < t \le b$ geldt dat $\gamma(t) \neq 0$. Zij $\widetilde{\gamma}$ de beperking van γ tot $[\widetilde{a},\, b]$. Daarvoor krijgen we dat

$$\begin{aligned}\|\gamma(b)\| &= \int_{\widetilde{a}}^{b} \frac{d\|\gamma(t)\|}{dt}\, dt = \int_{\widetilde{a}}^{b} \frac{\langle \gamma(t),\, \gamma'(t)\rangle}{\|\gamma(t)\|}\, dt \\ &\le \int_{\widetilde{a}}^{b} g_{\gamma(t)}(\gamma'(t),\, \gamma'(t))^{1/2}\, dt = l(\widetilde{\gamma}),\end{aligned}$$

met gelijkheid, dan en slechts dan als er een niet-negatieve functie $\lambda(t)$ is (gedefinieerd in ieder deelinterval waarin $\widetilde{\gamma}$ continu differentieerbaar is), waarvoor $\gamma'(t) = \lambda(t)\, \gamma(t)$. Omdat de lengte van γ gelijk is aan de lengte van $\gamma^0 := \gamma|_{[a,\, \widetilde{a}]}$ plus die van $\widetilde{\gamma}$, krijgen we dat $\|\gamma(b)\| \le l(\gamma)$, met gelijkteken dan en slechts dan als $\|\gamma(b)\| = l(\widetilde{\gamma})$ en $l\left(\gamma^0\right) = 0$. Deze laatste voorwaarde betekent dat $\gamma(t) = 0$ voor $a \leq t \leq \widetilde{a}$.

Componentsgewijs oplossen van de differentiaalvergelijking $\gamma'(t) = \lambda(t)\,\gamma(t)$ met voorgeschreven waarde in $t = b$ geeft dat

$$\gamma(t) = \mu(t)\,\gamma(b), \quad \mu(t) = \mathrm{e}^{-\int_t^b \lambda(s)\,ds}, \quad \widetilde{a} < t \leq b.$$

Het is hier misschien nog interessant om op te merken dat $\mu(t) \downarrow 0$ als $t \downarrow \widetilde{a}$, hetgeen betekent dat de integraal van λ over $[t,\, b]$ naar oneindig moet gaan als $t \downarrow \widetilde{a}$, dus dat $\lambda(t)$ zeker niet begrensd kan zijn voor $t \downarrow \widetilde{a}$. □

Gevolg 6.8 *Zij $(X,\, g)$ een Riemann-variëteit, $x \in X$ en $r(x)$ de injectiviteitsstraal in het punt x. Is $v \in \mathrm{T}_x\, X$ en $\|v\|_x < r(x)$ dan is $d\left(x,\, \exp_x(v)\right) = \|v\|_x$ = de lengte van de geodeet $t \mapsto \exp_x(t\, v)$, $0 \leq t \leq 1$. Is anderzijds $y \in X \setminus \exp_x\left(B_x(r(x))\right)$, dan is $d(x,\, y) \geq r(x)$.*

Bewijs Als γ een kromme is die in x start en $\exp_x\left(B_x(r(x))\right)$ verlaat, dan volgt uit Lemma 6.5 dat $l(\gamma) \geq r(x)$. Is anderzijds $v \in B_x(r(x))$ en is γ een kromme van x naar $\exp_x(v)$ die helemaal in $\exp_x\left(B_x(r(x))\right)$ verloopt, dan geeft Lemma 6.5 dat $l(\gamma) \geq \|v\|_x$, waarbij gelijkheid optreedt als $\gamma(t) = \exp_x(t\, v)$. □

Een deelverzameling U van X heet *open in X met betrekking tot de metriek d* als er bij iedere $x \in U$ een $\delta > 0$ is met de eigenschap dat als $y \in X$ en $d(x,\, y) < \delta$, dan is $y \in U$. Uit de laatste uitspraak in Lemma 6.8 zien we dat, als we $\delta \leq r(x)$ kiezen, dan is $y \in \exp_x\left(B_x(r(x))\right)$ en de eerste uitspraak geeft vervolgens dat de verzameling der $y \in X$ met $d(x,\, y) < \delta$ gelijk is aan $\exp_x\left(B_x(\delta)\right)$. Dit geeft dat U open is met betrekking tot de geodetische afstand, dan en slechts dan als U een open is in een kaart. Anders gezegd:

Gevolg 6.9 *In een Riemann-variëteit is de topologie die is gedefinieerd door de geodetische afstand gelijk aan de variëteitstopologie.*

6.6 Verder Onderzoek

De differentiaalmeetkundige theorie van (pseudo-)Riemann-variëteiten gaat verder met de introductie van begrippen als covariante differentiatie, connectie, kromming.

De *covariante afgeleide* van een vectorveld v in de richting van het vectorveld u is gedefinieerd als het vectorveld $\nabla_u v$ dat in lokale coördinaten gegeven wordt door

$$\nabla_u v(x)^i = \sum_{j=1}^{n} \frac{\partial v^i(x)}{\partial x^j}\, u^j(x) + \sum_{j,\,k=1}^{n} \Gamma^i_{jk}\, u^j(x)\, v^k(x). \tag{6.23}$$

De pointe hier is dat de „gewone” afgeleide van v in de richting van u, gegeven door de eerste som in het rechterlid van (6.23), niet contravariant transformeert onder substituties van variabelen, maar dat het gehele rechterlid van (6.23) dat wel doet.

De *Riemann'se krommingstensor* in het punt x is gedefinieerd als de afbeelding $R(x) : \mathrm{T}_x\, X \times \mathrm{T}_x\, X \times \mathrm{T}_x\, X \to \mathrm{T}_x\, X$ met de geigenschap dat voor ieder drietal van vectorvelden $u,\, v,\, w$ geldt dat

$$\left(\nabla_u \nabla v w - \nabla_v \nabla u w - \nabla_{[u,\,v]} w\right)(x) = R(x)\left(u(x),\, v(x),\, w(x)\right). \tag{6.24}$$

Hierin is $[u,\, v]$, de *Lie haakjes van u en v*, het vectorveld dat in lokale coördinaten gegeven wordt door

$$[u,\, v]\,(x)^i = \sum_{j=1}^{n} \frac{\partial v^i(x)}{\partial x^j}\, u^j(x) - \sum_{j=1}^{n} \frac{\partial u^i(x)}{\partial x^j}\, v^j(x). \tag{6.25}$$

Het rechterlid in (6.25 transformeert contravariant onder substituties van variabelen, hoewel ieder van de sommen daarin dat niet doet. De pointe van (6.24) is dat het linkerlid alleen maar afhangt van de waarden $u(x)$, $v(x)$, $w(x)$ van de vectorvelden u, v, w in het punt x en niet van hun eerste of tweede orde afgeleiden in het punt x, zoals a priori verwacht zou mogen worden.

Uit de formules is het duidelijk dat $R(x) = 0$ als er een lokaal systeem van coördinaten bestaat waarin alle partiële afgeleiden van de g_{ij} van de eerste en de tweede orde in het punt x gelijk aan nul zijn. Omgekeerd is het een stelling dat als $R(x) = 0$ voor alle x, dan is er een lokaal systeem van coördinaten waarin de $g_{ij}(x)$ constant zijn als functie van x. Men noemt de pseudo-Riemann'se structuur g *vlak* als dat laatste het geval is, ofwel als $R \equiv 0$. Voor meer details, zie [19, §6.1], Helgason [33, Ch. 1], Kobayashi and Nomizu [47], in opklimmende graad van grondigheid.

Een interessante relatie tussen kromming en de dynamische eigenschappen van de geodetische stroming is bijvoorbeeld het volgende. Zij $(X,\, g)$ een Riemann-variëteit waarvoor de zogenaamde sectionele kromming *negatief* is en neem verder aan dat X *compact* is. Dan 'roert de geodetische stroming op de lange termijn (voor $t \to \infty$) de niveauvariëteit $T = c > 0$ in zeer sterke zin door elkaar'. Zie Arnol'd and Avez [7, Ch. 3, §13 en p. 76], voor de precieze beweringen en aanduidingen van de bewijzen.

Er is ook veel onderzoek gedaan naar de existentie van periodieke geodeten op compacte Riemann-variëteiten, waarvan de existentie aangetoond wordt door een combinatie van variationele en (algebraïsch-) topologische technieken in lussenruimten. Zie bijvoorbeeld Klingenberg [46].

6.7 Vraagstukken

Vraagstuk 6.1 Zij X een n-dimensionale C^r-deelvariëteit van $\mathbf{R}^{n+m}$, $a \in X$. Kies een orthonormale basis $e_1, \ldots, e_n$ in $\mathrm{T}_a\, X$ en vul die aan tot een orthonormale basis van $\mathbf{R}^{n+m}$. Gebruik in $\mathbf{R}^{n+m}$ de coördinaten ten aanzien van deze basis. Bewijs dat er een open omgeving U van a in $\mathbf{R}^{n+m}$ is, een open omgeving V van $b = (a_1, \ldots, a_n)$ in $\mathbf{R}^n$ en een C^r-afbeelding $h : V \to \mathbf{R}^m$, waarvoor

$$X \cap U = \{(y,\, z) \in \mathbf{R}^{n+m} \mid y \in V,\, z = h(y)\}.$$

Verder is $\mathrm{D}\, h(a) = 0$. Gebruik $y \in V$ als coördinaten in X. Bewijs dat de Riemann-structuur gegeven wordt

$$g_y(v,\, w) = \langle v,\, w\rangle + \langle \mathrm{D}\, h(y)\, v,\, \mathrm{D}\, h(y)\, w\rangle.$$

Bewijs dat alle eerste orde afgeleiden van alle coëfficiënten van g in het punt $y = a$ gelijk aan nul zijn. Verifieer hiermee de uitspraak dat $\gamma(t)$ een geodeet is in X, dan en slechts dan als voor iedere t geldt dat $\gamma''(t)$ loodrecht staat op $\mathrm{T}_{\gamma(t)}\, X$. ⊘

Vraagstuk 6.2 Zij g een reëelwaardige C^r-functie in $\mathbf{R}^{n+1}$, $X = \{x \in \mathbf{R}^{n+1} \mid g(x) = 0\}$ en $\operatorname{grad} g(x) \neq 0$ voor iedere $x \in X$. Bewijs dat een tweemaal differentieerbare kromme $\gamma(t)$ een geodeet is in X, dan en slechts dan als $g(\gamma(0)) = 0$, $\langle \operatorname{grad} g(\gamma(0)),\, \gamma'(0)\rangle = 0$ en voor iedere t geldt dat

$$\gamma''(t) = -\frac{\mathrm{D}^2\, g(\gamma(t))\, (\gamma'(t),\, \gamma'(t))}{\|\operatorname{grad} g(\gamma(t))\|^2}\, \operatorname{grad} g(\gamma(t)).$$

Hint voor het 'dan' gedeelte: beschouw $\mathrm{T}\,X$ als de deelvariëteit van $\mathbf{R}^{n+1} \times \mathbf{R}^{n+1}$ bestaande uit de $(x,\, v)$ waarvoor $g(x) = 0$ en $\langle \operatorname{grad} g(x),\, v\rangle = 0$. Beschouw het vectorveld w, gedefinieerd in de open deelverzameling van $\mathbf{R}^{n+1} \times \mathbf{R}^{n+1}$ van de x waarvoor $\operatorname{grad} g(x) \neq 0$, door middel van $w(x,\, v) = (v,\, c \operatorname{grad} g(x))$, waarin

$$c = -\,\frac{\mathrm{D}^2 g(x)\,(v,\, v)}{\|\operatorname{grad} g(x)\|^2}.$$

Bewijs dat w raakt aan $\mathrm{T}\,X$, dus de stroming van w laat $\mathrm{T}\,X$ invariant. $\oslash$

Vraagstuk 6.3 Zij $S = \{x \in \mathbf{R}^{n+1} \mid \langle x,\, x\rangle = 1\}$ de n-dimensionale sfeer met straal 1 en middelpunt in de oorsprong in $\mathbf{R}^{n+1}$. Beschouw $\mathrm{T}\,S$ als de deelvariëteit van $\mathbf{R}^{n+1} \times \mathbf{R}^{n+1}$ bestaande uit de $(x,\, v)$ waarvoor $\langle x,\, x\rangle = 1$ en $\langle x,\, v\rangle = 0$. Bewijs dat de geodetische stroming gegeven wordt door

$$G^t(x,\, v) = \big(\cos(\|v\|\, t)\, x + \|v\|^{-1}\, \sin(\|v\|\, t)\, v,\; -\|v\|\, \sin(\|v\|\, t)\, x + \cos(\|v\|\, t)\, v\big)\,.$$

Bewijs dat voor iedere $x \in S$ de exponentiële afbeelding een analytisch diffeomorfisme is van

$$\{v \in \mathrm{T}_x\, S \mid \|v\| < \pi\}$$

naar $S \setminus \{-x\}$. En voor iedere $k \in \mathbf{Z}_{\geq 1}$ een analytisch diffeomorfisme van

$$\{v \in \mathrm{T}_x\, S \mid k\,\pi < \|v\| < (k+1)\,\pi\}$$

naar $S \setminus \{x,\, -x\}$. Wat is het beeld van de sfeer

$$\{v \in \mathrm{T}_x\, S \mid \|v\| = k\,\pi\}$$

onder de exponentiële afbeelding als $k \in \mathbf{Z}$? $\oslash$

Vraagstuk 6.4 Zij g een pseudo-Riemann-structuur in een variëteit X. Een differentieerbare afbeelding $\Phi : X \to X$ heet een *isometrie* met betrekking tot g, als voor iedere $x \in X$, $u,\, v \in \mathrm{T}_x\, X$ geldt dat

$$g_{\Phi(x)}\,(\mathrm{T}_x\, \Phi(u),\, \mathrm{T}_x\, \Phi(v)) = g_x(u,\, v).$$

Bewijs dat deze voorwaarde equivalent is met: de afbeelding $\mathrm{T}\,\Phi : \mathrm{T}\,X \to \mathrm{T}\,X$ laat de functie $L(x,\, v) := \frac{1}{2}\, g_x(v,\, v)$ invariant.

Neem nu in het vervolg aan dat Φ een C^2-isometrie is. Bewijs, door gebruik te maken van de interpretatie van de Euler-Lagrange-vergelijking als variatievergelijking, dat $\Phi \circ \gamma$ een geodeet is als γ een geodeet is. $\oslash$

Vraagstuk 6.5 Zij $H = \{z \in \mathbf{C} \mid \operatorname{Im} z > 0\}$ het *complexe bovenhalfvlak*, hetgeen we tegelijkertijd beschouwen als een open deelverzameling van $\mathbf{R}^2$ door $z = x + i\, y \in \mathbf{C}$ te identificeren met $(x,\, y) \in \mathbf{R}^2$. Met $\mathrm{SL}(2,\, \mathbf{R})$ duidt men de groep aan van alle $A \in \mathrm{L}\left(\mathbf{R}^2,\, \mathbf{R}^2\right)$ met $\det A = 1$. Voor iedere $A \in \mathrm{SL}(2,\, \mathbf{R})$ definiëren we de *gebroken lineaire transformatie* Φ_A door middel van

$$\Phi_A(z) = \frac{a\, z + b}{c\, z + d} \quad \text{als} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{6.26}$$

Bewijs de volgende uitspraken. Hierbij wordt het woord 'isometrie' gebruikt in de zin van Vraagstuk 6.4.

a) Φ_A is een complex-differentieerbare afbeelding van H naar $\mathbf{C}$ met afgeleide gelijk aan $(c\,z + d)^{-2}$. Er geldt dat $\Phi_A(H) \subset H$ en $\Phi_{A\,B} = \Phi_A \circ \Phi_B$. Φ_A is bijectief van H naar H, met inverse gelijk aan Φ_B, waarin $B = A^{-1}$. Φ_A is een complex-analytisch diffeomorfisme van H naar H.

b) Bij iedere $z \in H$ en $v \in \mathbf{C}$ met $v \neq 0$ is er een precies één $A = A(z,\, v) \in \mathrm{SL}(2,\, \mathbf{R})/\{\pm 1\}$ met $\Phi_A(z) = i$ en $\Phi'_A(z)\, v$ gelijk aan een positief veelvoud van i.

c) Zij g een Riemann-structuur in H met de eigenschap dat voor iedere $A \in \mathrm{SL}(2,\, \mathbf{R})$ de transformatie Φ_A een isometrie is. Dan is g eenduidig bepaald door g_i, door middel van de formule

$$g_z(v,\, v) = g_i\left(\Phi'_A(z)\, v,\, \Phi'_A(z)\, v\right), \tag{6.27}$$

waarin $A \in \mathrm{SL}(2,\, \mathbf{R})$, $\Phi_A(z) = i$.

d) De formule (6.27) leidt tot een Riemann-structuur g in H met de eigenschap dat voor iedere $A \in \mathrm{SL}(2,\, \mathbf{R})$ de transformatie Φ_A een isometrie is, dan en slechts dan als

$$g_i(v,\, v) = g_i\left(\Phi'_A(i)\, v,\, \Phi'_A(i)\, v\right)$$

voor iedere $A \in \mathrm{SL}(2,\, \mathbf{R})$ met $\Phi_A(i) = i$. Dit is equivalent met de voorwaarde dat er een constante $c > 0$ is met $g_i(v,\, v) = c\,|v|^2$. We kiezen in het vervolg $c = 1$. Hiermee krijgen we dan:

$$g_z(v,\, v) = (\mathrm{Im}\, z)^{-2}\, |v|^2, \quad z \in H,\, v \in \mathbf{C}.$$

e) De spiegeling om de imaginaire as $S : x + i\,y \mapsto -x + i\,y$ is een isometrie. Is γ de geodeet met $\gamma(0) = i$ en $\gamma'(0) = i$ dan is $\delta = S \circ \gamma$ een geodeet met $\delta(0) = i$ en $\delta'(0) = i$. Er geldt dat $\delta = \gamma$, ofwel $\gamma(t) = y(t)\, i$ voor een positieve reëelwaardige functie $y(t)$. De voorwaarde dat de g-lengte van de snelheidsvector constant is en dus gelijk aan 1, impliceert dat y moet voldoen aan de differentiaalvergelijking $y'/y = 1$ en de conclusie is dat $\gamma(t) = \mathrm{e}^t\, i$.

f) Iedere geodeet δ die met booglengte is geparametriseerd is van de vorm $\Phi_A \circ \gamma$, met γ als in e) en $A \in \mathrm{SL}(2,\, \mathbf{R})$ als in (6.26). Als $c \neq 0$, resp. $d \neq 0$ dan geldt

$$\lim_{t \to \infty} \delta(t) = \frac{a}{c}, \quad \text{resp.} \lim_{t \to -\infty} \delta(t) = \frac{b}{d}.$$

Als $c \neq 0$ en $d \neq 0$ dan is de baan van δ gelijk aan de halve cirkel in H met middelpunt gelijk aan het punt $m = (a/c + b/d)/2$ op de reële as. Als $c = 0$ dan is de baan een verticale halflijn in H met voetpunt in b/d. Als $d = 0$ dan is de baan een verticale halflijn met voetpunt in a/c.

g) Bij ieder paar van punten z en w in H is er precies één $A \in \mathrm{SL}(2,\, \mathbf{R})/\{\pm 1\}$ waarvoor: $\Phi_A(z) = i$ en er is een $y \geq 1$ met $\Phi_A(w) = y\, i$. De geodetische afstand is gelijk aan $\log y$, waarin $y = \frac{\tau + \sqrt{\tau^2 - 4}}{2}$ en

$$\tau = \frac{(\mathrm{Re}\, z - \mathrm{Re}\, w)^2 + (\mathrm{Im}\, z)^2 + (\mathrm{Im}\, w)^2}{\mathrm{Im}\, z\ \mathrm{Im}\, w}.$$

Hint: Schrijf A als in (6.26) met $a,\, b,\, c,\, d \in \mathbf{R}$. Beschouw de vergelijking $a\,z + b = i\,(c\,z + d)$ als twee reële vergelijkingen. Los hieruit b en d op in termen van a, c, $\mathrm{Re}\, z$, $\mathrm{Im}\, z$. Vul dit in in $a\,w + b = i\,y\,(c\,w + d)$ en beschouw het resultaat als een stelsel van twee lineaire vergelijkingen voor de onbekenden a, c. Dit stelsel heeft alleen een oplossing $(a,\, c) \neq (0,\, 0)$ als de determinant van de bijbehorende coëfficiëntenmatrix gelijk is aan nul. Dit leidt tot een kwadratische vergelijking voor y. Bepaal hiervan de oplossing $y \geq 1$, in termen van $\operatorname{Re} z$, $\operatorname{Im} z$, $\operatorname{Re} w$, $\operatorname{Im} w$.

$\oslash$

Vraagstuk 6.6 Zij gegeven een sproei in een open omgeving X van 0 in $\mathbf{R}^n$, met Christoffel-symbolen van de klasse C^k. We noteren met $\|x\|$ de Euclidische norm van $x \in \mathbf{R}^n$ en met $B(r)$ de verzameling der $x \in \mathbf{R}^n$ met $\|x\| < r$. We willen aantonen dat als $r > 0$ voldoende klein is, dan zijn alle paren van punten in $B(r)$ verbonden door eenduidig bepaalde geodeten die in $B(r)$ verlopen, bovendien hangen deze geodeten op een C^k manier af van de begin- en eindpunten. Bewijs daartoe:

a) Zij U een open omgeving van $(0,\, 0)$ in $X \times X$ als in Lemma 6.3, met x_0 vervangen door 0. Er is een $r_0 > 0$ met de eigenschap dat als $0 < r \leq r_0$ dan is $B(r) \times B(r) \subset U$.

b) Voor iedere $0 < \epsilon < 1$ is er een $r > 0$ met de eigenschap dat als γ een geodeet is en $\gamma(t) \in B(r)$, dan is

$$\frac{d^2}{dt^2}\, \|\gamma(t)\|^2 \geq (1 - \epsilon)\, \|\gamma'(t)\|^2.$$

c) Er is een $r_0 > 0$ met de eigenschap dat voor iedere $0 < r \leq r_0$ het volgende geldt. Is γ een geodeet in X, $a < b$, $\gamma(a) \in B(r)$ en $\gamma(b) \in B(r)$, dan geldt voor iedere $a \leq t \leq b$ dat $\gamma(t) \in B(r)$.

d) Voor iedere $\delta > 0$ is er een $r > 0$ met de eigenschap dat als γ een geodeet in $B(r)$ is, dan gelden de schattingen

$$\|\gamma''(t)\| \leq \delta, \quad 2r > \|\gamma(1) - \gamma(0)\| \geq \|\gamma'(0)\| - \frac{\delta}{2}.$$

e) Er is een $r_0 > 0$ met de eigenschap dat voor iedere $0 < r \leq r_0$ de volgende eigenschappen gelden.

   i) $B(r) \subset X$.

   ii) Voor iedere $x \in B(r)$ en $y \in B(r)$ is er precies één $v = v(x,\, y) \in \mathbf{R}^n$ waarvoor de geodeet met $\gamma(0) = x$ en $\gamma'(0) = v$ de eigenschap heeft dat $[0,\, 1]$ bevat is in het definitie-interval van γ, $\gamma\,([0,\, 1]) \subset B(r)$ en $\gamma(1) = y$.

   iii) De afbeelding $(x,\, y) \mapsto v(x,\, y)$ is C^k.

$\oslash$

Vraagstuk 6.7 Zij $g_{ij}(x)$ een pseudo-Riemann-structuur in een open deelverzameling van $\mathbf{R}^n$ met bijbehorende Christoffel-symbolen $\Gamma_{ij}^k(x)$. Zij $\widetilde{\Gamma}_{ij}^k(x)$ de Christoffel-symbolen van een sproei die een stroming $\widetilde{G}^t$ in de faseruimte induceert. Schrijf

$$\begin{aligned} \Gamma_{hij}(x) &= \sum_{l=1}^{n} g_{hl}(x)\, \Gamma_{ij}^l(x), \\ \widetilde{\Gamma}_{hij}(x) &= \sum_{l=1}^{n} g_{hl}(x)\, \widetilde{\Gamma}_{ij}^l(x). \end{aligned}$$

Bewijs dat $\widetilde{G}^t$ de functie $T = \frac{1}{2}\, g_x(v,\, v)$ invariant laat, dan en slechts dan als $\Delta_{hij} = \widetilde{\Gamma}_{hij} - \Gamma_{hij}$ voldoet aan

$$\Delta_{hij} + \Delta_{jhi} + \Delta_{ijh} \equiv 0.$$

Laat zien dat in het geval $n = 2$ deze voorwaarde equivalent is met de existentie van functies $f_1(x)$ en $f_2(x)$, waarvoor

$$\begin{aligned}
&\Delta_{111} = \Delta_{222} = 0, \\
&-\frac{1}{2}\,\Delta_{211} = \Delta_{112} = \Delta_{121} = f_1, \\
&-\frac{1}{2}\,\Delta_{122} = \Delta_{221} = \Delta_{112} = f_2.
\end{aligned}$$

⊘

Vraagstuk 6.8 Zij g een pseudo-Riemann-structuur en zij v en w differentieerbare vectorvelden op de variëteit M. Definieer de reëelwaardige functie $f = g(v,\, w)$ op M door middel van $f(x) = g_x(v(x),\, w(x))$ voor iedere $x \in M$. Bewijs dat

$$\mathrm{d}f(x) \cdot u = g_x\left(\left(\nabla_u v\right)(x),\, w(x)\right) + g_x\left(v(x),\, \left(\nabla_u w\right)(x)\right). \tag{6.28}$$

Als men het symbool ∇_u opvat als een 'differentiatie in de richting van de vector u' (en wel de covariante differentiatie), dan zou men met het oog op de somregel nog een term verwachten waarin g_x met betrekking tot x in de richting van u (covariant) gedifferentieerd wordt. Omdat deze term in (6.28) ontbreekt, zegt men ook wel dat de covariante afgeleide van g gelijk is aan nul, ofwel de pseudo-Riemann-structuur g is *covariant constant.* ⊘

6.8 Het Complexe Bovenhalfvlak

In Vraagstuk 6.5 wordt het *bovenhalfvlak van Poincaré* beschreven, één van de modellen die Poincaré [73] gaf voor het niet-Euclidische vlak, dat in het begin van de 19e eeuw was gevonden door *Lobachevski en Bolyai* en omschreven kan worden als een vlak met een Riemann-structuur met kromming constant gelijk aan -1.

Een *Riemann-oppervlak* X is gedefinieerd als een complex ééndimensionale complex-analytische variëteit. (Merk op dat dit iets anders is dan een oppervlak met een Riemann-structuur, hetgeen men terecht zou kunnen denken.) X is ook een reëel twee-dimensionale reëel-analytische variëteit, waarbij de vermenigvuldiging met i in de raakruimten een orientatie in X definieert. Ieder compact en georienteerd oppervlak X is diffeomorf met een sfeer met g handvaten, men noemt g het *geslacht* van X. Het is nu een interessant feit dat ieder Riemann-oppervlak X geïdentificeerd kan worden met de ruimte H/Γ van de Γ-banen in H, waarin Γ een geschikte discrete ondergroep van $\mathrm{SL}(2,\, \mathbf{R})/\{\pm 1\}$ vormt.

Als eerste stap van het bewijs construeert men bij X de zogenaamde *universele overdekking* $(\widetilde{X}, p)$ van X. Hierin is $\widetilde{X}$ een enkelvoudig samenhangende variëteit en $p : \widetilde{X} \to X$ is een lokaal diffeomorfisme, waarmee de complexe structuur in X teruggetrokken kunnen worden naar een complexe structuur in $\widetilde{X}$. Bij een gegeven basispunt $\widetilde{x_0} \in \widetilde{X}$ is er bij iedere $\widetilde{x} \in \widetilde{X}$ met $p(\widetilde{x}) = p(\widetilde{x_0})$ precies één diffeomorfisme $g : \widetilde{X} \to \widetilde{X}$, waarvoor $p \circ \gamma = p$ en $p(\widetilde{x_0}) = \widetilde{x}$. Deze g vormen een groep

Γ van transformaties in $\widetilde{X}$, waarvoor X gezien kan worden als de ruimte $\widetilde{X}/\Gamma$ van Γ-banen. Verder zijn de $g \in \Gamma$ complex-analytische diffeomorfismen van $\widetilde{X}$.

Volgens de *uniformiseringsstelling* van Poincaré [76] en Koebe [48] is iedere enkelvoudig samenhangend Riemann-oppervlak M complex-analytisch diffeomorf met ófwel de Riemann-sfeer $S = \mathbf{C} \cup \{\infty\}$, ófwel met $\mathbf{C}$, ofwel met H. We passen dit toe op $M = \widetilde{X}$. Nu impliceert $g \geq 1$ dat Γ oneindig is en dat sluit het geval dat $\widetilde{X} = S$ uit. Verder, ieder complex-analytisch diffeomorfisme van $\mathbf{C}$ is van de vorm $z \mapsto a\,z + b$ met $a,\, b \in \mathbf{C}$, $a \neq 0$. Dit betekent dat de elementen van Γ alleen maar translaties kunnen zijn, dus $X = \mathbf{C}/\Gamma$ is diffeomorf met een reëel tweedimensionale torus, dus $g = 1$. Anders gezegd, de voorwaarde dat $g \geq 2$ impliceert dat we $\widetilde{X}$ als complex--analytische variëteit met H mogen identificeren. Verder krijgen we dat Γ een ondergroep van $\mathrm{SL}(2,\, \mathbf{R})/\{\pm 1\}$ is, omdat ieder complex-analytisch diffeomorfisme g van H van de vorm $g = \Phi_A$ is voor een $A \in \mathrm{SL}(2,\, \mathbf{R})$. Dit laatste feit heeft, in tegenstelling tot de uniformiseringsstelling, een eenvoudig bewijs:

Bewijs De fractionele lineaire transformatie $\psi : z \mapsto \frac{z-i}{z+i}$ beeldt H af op de schijf $D = \{z \in \mathbf{C} \mid |z| < 1\}$, $f = \psi \circ g \circ \psi^{-1}$ is een complex-analytisch diffeomorfisme van D. Door samenstelling met een geschikte fractionele lineaire transformatie kunnen we ervoor zorgen dat $f(0) = 0$. We gaan bewijzen dat dit impliceert dat er een constante c is met $f(z) = c\,z$. Teruglezen geeft dan dat g een fractionele lineaire transformatie is. Omdat daarvoor ook moet gelden dat $g(\mathbf{R} \cup \{\infty\}) = \mathbf{R} \cup \{\infty\}$, krijgen we dat $g = \Phi_A$ voor een $A \in \mathrm{SL}(2,\, \mathbf{R})$.

Omdat $f(0) = 0$ is de functie $h(z) = f(z)/z$ complex-analytisch in D. Verder geeft $|f(z)| < 1$ dat $|h(z)| \leq 1/r$ als $|z| = r < 1$, vanwege het maximum-modulus principe geeft dit dat $|h(z)| \leq 1/r$ als $|z| \leq r < 1$. Hieruit volgt dat voor iedere $z \in D$ geldt dat $|h(z)| \leq 1$, ofwel $|f(z)| \leq |z|$. Dezelfde conclusie geldt met f vervangen door f^{-1}, dus $|z| = |f^{-1}(f(z))| \leq |f(z)|$. Dus voor iedere $z \in D$ geldt dat $|f(z)| = |z|$, ofwel $|h(z)| = 1$. Maar nu geeft het maximum-modulus principe dat de functie h constant is. $\square$

Opmerking: De uitspraak, dat $|f(z)| \leq |z|$ als $f : D \to D$ complex-analytisch is en $f(0) = 0$, is afkomstig van Schwarz [84]. Dit werd door Carathéodory [11] het *lemma van Schwarz* genoemd en gebruikt om het bovenstaande bewijs te geven.

Een interessant gevolg van de uniformiseringsstelling is nog het volgende. Men noemt twee (pseudo-)Riemann-structuren g en $\widetilde{g}$ *conform*, als er een positieve functie λ is, waarvoor $\widetilde{g} = \lambda\, g$. Stelling: bij iedere Riemann-structuur g in een compact georienteerd oppervlak X met geslacht ≥ 2 een met g conforme Riemann-structuur in X is, waarvan de kromming constant gelijk aan -1 is. Bewijs: voor iedere $x \in X$ is er een éénduidig bepaalde lineaire afbeelding J_x in $\mathrm{T}_x X$, die 'een kwartslag draaien in de positieve richting' voorstelt. Dat wil zeggen ${J_x}^* g_x = g_x$, ${J_x}^2 = -1$ en als $v \in \mathrm{T}_x X$, $v \neq 0$, dan is het paar v, $J_x(v)$ positief georienteerd. Door J_x op te vatten als vermenigvuldiging met i wordt van $\mathrm{T}_x X$ een complex ééndimensionale lineaire ruimte over $\mathbf{C}$ gemaakt en we kunnen hiermee praten over complex-analytische functies op X. Het is een (niet-triviaal) feit dat X een atlas heeft die bestaat uit complex-analytische functies, gedefinieerd in open deelverzamelingen van X. Hiermee is X een Riemann-oppervlak, dat we vanwege het voorgaande met H/Γ kunnen identificeren. Het feit dat g en $g_{{}_H}$ dezelfde complexe structuur in X definiëren impliceert dat $g_{{}_H}$ conform is met g; anderzijds is de kromming van $g_{{}_H}$ constant gelijk aan -1.

De uniformiseringsstelling heeft als historische voorganger *afbeeldingsstelling van Riemann* [78, Dissertation 1851, p. 3-45, Art. 21]. Deze zegt dat als M een enkelvoudig samenhangende, begrensde en open deelverzameling is van $\mathbf{C}$ met een gesloten kromme K als rand, dan is er een complex-analytisch diffeomorfisme f is van M naar de open eenheidsschijf D in $\mathbf{C}$, met een continue uitbreiding naar K. Daarbij ligt f vast door te eisen dat $f(p) = 0$ en $f(k) = 1$, waarin p, resp. k een gekozen punt in M resp. K zijn. Dit correspondeert met de stelling dat als $f : D \to D$ een complex--analytisch diffeomorfisme is met $f(0) = 0$, dan is $f(z) = c\,z$ voor een constante c met $|c| = 1$. Alle bewijzen van de uniformiseringsstelling gaan uit van Riemann's idee om een reëelwaardige harmonische functie u in $M \setminus \{p\}$ te vinden met $u(z) \sim \log|z - p|$ als $z \to p$ en $u(z) \to 0$ als z uit alle compacte deelverzamelingen van M wegloopt. Is v de bijbehorende geconjugeerde harmonische functie dan is $f = \mathrm{e}^{u+i\,v}$ de gezochte afbeelding van M naar D. Zie bijvoorbeeld Farkas en Kra [23, IV.4].

Voor hoger-dimensionale variëteiten met constante kromming is een schat aan informatie te vinden in het boek van Wolf [90].

7 Hamilton-stelsels

7.1 De Snelheids-impuls-afbeelding

Lagrange [52, p. 309] merkte op dat we de oplossing van de bewegingsvergelijkingen kunnen beschouwen als functie van de tijd, beginpositie en beginsnelheid, maar dat we hierin de beginsnelheid ook kunnen vervangen door de impuls $p = \frac{\partial T(t,x,v)}{\partial v}$ op het begintijdstip. Lagrange werkte dit uit voor de variatievergelijkingen die onststaan bij storingen in de potentiële energie-functie. Hamilton ontdekte dat in het algemeen deze transformatie de bewegingsvergelijkingen, wanneer die gegeven zijn als de Euler-Lagrange-vergelijkingen (4.14) voor een functie $L = L(t, x, v)$, een verrassend elegante vorm geeft. Wij volgen hier de presentatie van Jacobi [42], die bekender was en beter begrepen werd dan Hamilton's publicaties over de klassieke mechanica.

Zij U een open deelverzameling van $\mathbf{R} \times \mathrm{T}\, X$ en zij L een reëelwaardige functie in U. We nemen aan dat $k \geq 1$, $L \in \mathrm{C}^k$ en dat $(t, x, v) \mapsto \frac{\partial L(t,x,v)}{\partial v}$ een C^k-afbeelding is.

Voor iedere $(t, x) \in \mathbf{R} \times X$ is

$$U_{(t,\, x)} := \{v \in \mathrm{T}_x\, X \mid (t,\, x,\, v) \in U\}$$

een open deelverzameling van de lineaire ruimte $\mathrm{T}_x\, X$. Voor iedere $v \in U(t,\, x)$ kunnen we de impuls $p = \frac{\partial L(t,x,v)}{\partial v}$ opvatten als een lineaire vorm op $\mathrm{T}_x\, X$. Dit leidt tot de *snelheids-impuls-afbeelding* $\Phi = \Phi_L : U \to \mathbf{R} \times \mathrm{T}^*\, X$, die is gedefinieerd door

$$\Phi(t,\, x,\, v) := \left(t,\, x,\, \frac{\partial L(t,\, x,\, v)}{\partial v}\right). \tag{7.1}$$

Merk op dat $\Phi \in \mathrm{C}^k$.

De volgende aanname is dat voor iedere $(t,\, x) \in \mathbf{R} \times X$ de afbeelding $v \mapsto \frac{\partial L(t,x,v)}{\partial v}$, van $U(t,\, x)$ naar $(\mathrm{T}_x\, X)^*$, injectief is en dat in iedere $v \in U(t,\, x)$ de totale afgeleide hiervan bijectief is. Deze laatste voorwaarde betekent in termen van een lineair coördinatensysteem in $\mathrm{T}_x\, X$ dat de tweede-orde afgeleidenmatrix van L met betrekking tot de snelheden een niet-singuliere symmetrische $n \times n$-matrix is, ofwel dat

$$\det \left(\frac{\partial^2 L(t,\, x,\, v)}{\partial v_i \partial v_j}\right)_{i,\, j=1}^{n} \neq 0, \quad (t,\, x,\, v) \in U. \tag{7.2}$$

(7.2) heet de *voorwaarde van Legendre* voor de functie L.

Merk op dat aan alle voorwaarden is voldaan als

$$L(t,\, x,\, v) = \frac{1}{2} \sum_{i,\, j}^{n} m_{ij}(t,\, x)\, v_i\, v_j + \sum_{i=1}^{n} \mu_i(t,\, x)\, v_i + c(t,\, x),$$

waarin m_{ij}, μ_i en c reëelwaardige C^k-functies zijn, $m_{ij} = m_{ji}$ en voor iedere $(t,\, x)$ de symmetrische $n \times n$ matrix $m_{ij}(t,\, x)$ niet-singulier is. (Anders gezegd, voor iedere $(t,\, x)$ is $v \mapsto L(t,\, x, v)$ een tweedegraads-veelterm, met de extra eigenschap dat de coëfficiënten van de tweedegraadstermen een niet-singuliere matrix vormen.) Immers in dit geval wordt de snelheids-impuls-afbeelding gegeven door de affiene afbeelding

$$v \mapsto m(t,\, x) \cdot v + \mu(t,\, x),$$

welke injectief is, dan en slechts dan als de matrix $m(t, x)$ niet-singulier is; verder is de Jacobimatrix van deze afbeelding gelijk aan $m(t, x)$.

Als L gelijk is aan het verschil van de kinetische energie en de potentiële energie, in een willekeurig tijdsafhankelijk coördinatensysteem, dan zijn we in deze situatie met $m_{ij}(t, x)$ gelijk aan een positief-definiete symmetrische matrix; zo'n matrix is niet-singulier.

De voorwaarden impliceren dat in ieder punt van U de raakafbeelding van Φ bijectief is en dat Φ injectief is. Vanwege de globale inverse-functie-stelling krijgen we dat $P := \Phi(U)$ een open deelverzameling is $\mathbf{R} \times \mathrm{T}^* X$ en dat Φ een C^k-diffeomorfisme is van U naar P.

7.2 De Legendre-transformatie

We gaan nu, in een kaart voor X, onderzoeken aan welk stelsel van differentiaalvergelijkingen de kromme $\delta : t \mapsto (x(t),\, p(t))$ voldoet als $\gamma : t \mapsto (x(t),\, v(t))$ een oplossing is van de Euler-Lagrange-vergelijkingen (4.14) en $\delta = \Phi \circ \gamma$, dat wil zeggen

$$p(t) = \frac{\partial L(t,\, x(t),\, v)}{\partial v}|_{v=v(t)}.$$

Hierbij maken we gebruik van de reëelwaardige functie h in P die is gedefinieerd als $h = H \circ \Phi^{-1}$, met H als in (4.21), op coordinaat-invariante manier geschreven als

$$H(t,\, x,\, v) := \frac{\partial L(t,\, x,\, v)}{\partial v} \cdot v - L(t,\, x,\, v). \tag{7.3}$$

Merk op dat, als $L(x,\, v) = T(x,\, v) - V(x)$, waarin $v \mapsto T(x,\, v)$ homogeen is van de graad twee, dan geeft Euler's differentiaalvergelijking dat

$$\frac{\partial L(t,\, ,\, v)}{\partial v} \cdot v = \frac{\partial T(x,\, v)}{\partial v} \cdot v = 2T.$$

Dus $H = 2T - (T - V) = T + V$, ofwel H is in dit geval gelijk aan de totale energie.

Zij $v = v(t,\, x,\, p)$ de oplossing van de vergelijking $p = \frac{\partial L(t,x,v)}{\partial v}$, dit is het v-deel van $\Phi^{-1}(t,\, x,\, p)$. Dan is

$$h(t,\, x,\, p) = \sum_{i=1}^{n} p_i\, v_i(t,\, x,\, p) - L(t,\, x,\, v(t,\, x,\, p)), \tag{7.4}$$

hetgeen impliceert dat

$$\frac{\partial h(t,\, x,\, p)}{\partial p_j} = v_j(t,\, x,\, p) + \sum_{i=1}^{n} p_i\, \frac{v_i(t,\, x,\, p)}{\partial p_j} - \sum_{i=1}^{n} \frac{\partial L(t,\, x,\, v)}{\partial v_i}|_{v=v(t,\,x,\,p)}\, \frac{v_i(t,\, x,\, p)}{\partial p_j},$$

waaruit we aflezen dat

$$\frac{\partial h(t,\, x,\, p)}{\partial p_j} = v_j(t,\, x,\, p), \quad 1 \leq j \leq n. \tag{7.5}$$

Een soortgelijke berekening geeft dat

$$\frac{\partial h(t,\, x,\, p)}{\partial x_j} = -\frac{\partial L(t,\, x,\, v)}{\partial x_j}|_{v=v(t,\,x,\,p)}.$$

Hiermee worden de Euler-Lagrange-vergelijkingen $\frac{dx}{dt} = v$, $\frac{d}{dt}\frac{\partial L}{\partial v} = \frac{\partial L}{\partial x}$ equivalent met het eerste orde stelsel

$$\left.\begin{aligned} \frac{dx_j}{dt} &= \frac{\partial h(t,x,p)}{\partial p_j}, \\ \frac{dp_j}{dt} &= -\frac{\partial h(t,x,p)}{\partial x_j}, \end{aligned}\right. \quad 1 \leq j \leq n. \tag{7.6}$$

Met andere woorden, het snelheidsveld in de $(x,\, p)$-ruimte is een uitdrukking in termen van de totale afgeleide van de functie h. Het is echter *niet* gelijk aan de gradiënt van h, vanwege de verwisseling van de q- en de p-componenten en het minteken in de p-component. Men noemt (7.6) een *(tijdsafhankelijk) Hamilton-stelsel* gedefinieerd door de *Hamilton-functie* h.

De transformatie die aan de functie L de functie $h = h_L$ toevoegt staat bekend als de *Legendre--transformatie*, althans in de situatie zonder de parameters t en x. Volgens de Encyclopaedia of Mathematics [32, Vol. 5, p. 394] gaat de Legendre-transformatie, van een functie L van v naar een functie h van p, voor $n = 1$ terug tot Leibniz en werd voor algemene n beschouwd door Euler in 1776, vóór Legendre in 1789 zijn definitie gaf.

Voor strikt convexe functies $v \mapsto L(t,\, x,\, v)$, gedefinieerd in een convexe open deelverzameling $U(t,\, x)$ van $\mathbf{R}^n$, kan de Legendre-transformatie h ook gedefinieerd worden als

$$h(t,\, x,\, p) = \max\{\langle p,\, v\rangle - L(t, x,\, v) \mid v \in U(t,\, x)\},$$

omdat $v = v(t,\, x,\, p)$ precies het stationaire punt is van $v \mapsto \langle p,\, v\rangle - L(t,\, x,\, v)$. In dit geval heet h ook wel de *geconjugeerde functie van L*. De geconjugeerde functie $p \mapsto h(t,\, x,\, p)$ is dan ook strikt convex en we hebben de ongelijkheid

$$\langle p,\, v\rangle \leq L(t,\, x,\, v) + h(t,\, x,\, p),$$

geldig voor alle $v \in U(t,\, x)$ en $p \in P(t,\, x)$, waarin $P(t,\, x)$ de (open) verzameling der p voorstelt, waarvoor $(t,\, x,\, p) \in P$.

Lagrange schreef in [52, p. 310] de variatie-vergelijkingen, bij storingen van de potentiële energie, als een tijdsafhankelijk Hamilton-stelsel voor de positie-impulscoördinaten. Echter de Legendre--transformatie komt in [52] niet voor; Lagrange schreef de totale energie alleen in de gedaante $H = T + V$ en niet in de gedaante (7.3). Zo miste Lagrange, ik zou willen zeggen op een haar na, de constatering dat in het algemeen de Euler-Lagrange-vergelijkingen overgaan in een Hamilton-stelsel voor de positie-impulscoördinaten.

Voordat we verder gaan, merken we nog even op dat (7.5) impliceert dat de inverse Ψ van Φ gegeven wordt door

$$\Psi(t,\, x,\, p) = (t,\, x,\, v) \quad \text{met} \quad v = \frac{\partial h(t,\, x,\, p)}{\partial p}. \tag{7.7}$$

Merk op dat hierin $\partial h(t,\, x,\, p)/\partial p$ een lineaire vorm is op $(\mathrm{T}_x\, X)^*$. Is E een eindigdimensionale lineaire ruimte, dan is de afbeelding die aan iedere $v \in E$ toevoegt de lineaire vorm

$$\operatorname{test} v : \xi \mapsto \xi(v)$$

op E^* een bijectieve lineaire afbeelding van E naar $(E^*)^*$, die men gebruikt om E met $(E^*)^*$ te identificeren. (Men schrijft dan ook v in plaats van $\operatorname{test} v$.) Op deze manier wordt $\partial h(t,\, x,\, p)/\partial p$ als een element van $\mathrm{T}_x\, X$ gezien.

Omdat Ψ een C^k-diffeomorfisme is van V naar U, krijgen we dat voor iedere t en x de afbeelding $p \mapsto \frac{\partial h(t,x,p)}{\partial p}$ een C^k-diffeomorfisme is van $P(t, x)$ naar $U(t, x)$. Dit impliceert dat deze afbeelding injectief is en dat voor iedere (t, x, p) de tweede-orde afgeleidenmatrix van $p \mapsto h(t, x, p)$ niet-singulier is.

Zij nu omgekeerd h een reëelwaardige C^k functie in een open deelverzameling P van $\mathbf{R} \times \mathrm{T}^* X$, die aan de laatste twee voorwaarden voldoet, dat wil zeggen waarvoor (7.7) een C^k-diffeomorfisme $\Psi = \Psi_h$ definieert van P naar een open deelverzameling U van $\mathrm{T}\, X$. Definieer de functie

$$l(t, x, p) = \frac{\partial h(t, x, p)}{\partial p} \cdot p - h(t, x\, p) \tag{7.8}$$

en schrijf $L = l \circ \Psi^{-1}$, ofwel in lokale coördinaten:

$$L(t, x, v) = \sum_{i=1}^{n} v_i\, p_i(t, x, v) - h(t, x, p(t, x, v)),$$

waarin $p(t, x, v)$ gedefinieerd is als de oplossing p van de vergelijking $\frac{\partial h(t,x,p)}{\partial p} = v$. Anders gezegd, $L = L_h$ is de Legendre-transformatie van h.

We lezen hieruit af dat

$$\begin{aligned} \frac{\partial L(t, x, v)}{\partial v} &= p(t, x, v), \\ \frac{\partial L(t, x, v)}{\partial x} &= -\frac{\partial h(t, x, p)}{\partial x}\Big|_{p=p(t,x,v)}. \end{aligned}$$

Is $\delta(t) = (t, x(t), p(t))$ een oplossing van het Hamilton-stelsel (7.6) met Hamilton-functie gelijk aan h, dan voldoet $\gamma(t) := \Psi(\delta(t)) = (t, x(t), v(t))$ aan

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial h}{\partial p} = v, \\ \frac{d}{dt}\frac{\partial L}{\partial v} &= \frac{dp}{dt} = -\frac{\partial h}{\partial x} = \frac{\partial L}{\partial x}, \end{aligned}$$

dus aan de Euler-Lagrange-vergelijkingen met Lagrange-functie gelijk aan L.

Dit leidt tot een volledige equivalentie tussen enerzijds Euler-Lagrange-vergelijkingen met een Lagrange-functie L waarvoor de snelheids-impulsafbeelding $v \mapsto \frac{\partial L(t,x,v)}{\partial v}$ een C^k-diffeomorfisme definieert en anderzijds Hamilton-stelsels met Hamilton-functie h waarvoor de impuls-snelheidsafbeelding $p \mapsto \frac{\partial h(t,x,p)}{\partial p}$ een C^k-diffeomorfisme definieert. Daarbij is h de Legendre-transformatie van L en L de Legendre-transformatie van h. In formule: $h = h_L$ dan en slechts dan als $L = L_h$ en in deze situatie zijn Φ_L en Ψ_h elkaars inverse.

7.3 De Kanonieke Tweevorm

De Euler-Lagrange-vergelijkingen met Lagrange-functie L zijn coördinaat-invariant gedefinieerd in een open deelverzameling U van $\mathbf{R} \times \mathrm{T}\, X$. De snelheids-impuls-afbeelding Φ_L is coördinaatinvariant gedefinieerd als afbeelding van een open deelverzameling van U naar een open deelverzameling P van $\mathbf{R} \times \mathrm{T}^* X$. Ten derde is de Legendre-transformatie $h = h_L$ coördinaat-invariant gedefinieerd in P. Hieruit volgt dat, althans voor functies h in een open deelverzameling P van $\mathbf{R} \times \mathrm{T}^* X$ waarvoor

Ψ_h een C^k-diffeomorfisme is, het Hamilton-stelsel met Hamilton-functie h, dat in lokale coördinaten in $T^* X$ afkomstig van een kaart in X is gegeven door (7.6), in feite coördinaat-invariant gedefinieerd is. Dat wil zeggen, als we een t-afhankelijke substitutie van variabelen $\widetilde{\psi}$ in $T^* X$ uitvoeren die afkomstig is van een t-afhankelijke susbtitutie van variabelen ψ in X, dan gaat (7.6) over in het Hamilton-stelsel met Hamilton-functie $\widetilde{h} = h \circ \widetilde{\psi}$. Men kan dit ook door nagaan door middel van een directe berekening, waarbij dan blijkt dat de voorwaarde dat Ψ_h een diffeomorfisme is voor deze bewering niet nodig is.

In plaats van deze directe berekening geven wij nu een coördinaat-invariante karakterisering van Hamilton-stelsels die afkomstig is van Lie [57]. Lie behandelde dit in het iets algemenere kader van de zogenaamde *contact-transformaties*. Dit is misschien één van de oorzaken waarom Lie's ideeën relatief laat ingang vonden in de klassieke mechanica.

Het uitgangspunt is de opmerking dat in de coraakbundel $T^* X$ van een willekeurige variëteit een differentiaalvorm τ van de graad één gedefinieerd is door middel van de formule

$$\tau_{(x,\,\xi)} = \xi \circ T_{(x,\,\xi)}\, \pi, \quad x \in X \;\; \xi \in (T_x\, X)^* . \tag{7.9}$$

Hierin is $\pi = \pi_{T^* X} : T^* X \to X$ de projectie naar de basis, dus $T_{(x,\,\xi)}\, \pi$ is een lineaire afbeelding van $T_{(x,\,\xi)}\,(T^* X)$ naar $T_x\, X$. De samenstelling van $T_{(x,\,\xi)}\, \pi$ met de lineaire vorm ξ op $T_x\, X$ levert de lineaire vorm $\tau_{(x,\,\xi)}$ op $T_{(x,\,\xi)}\,(T^* X)$.

In lokale coördinaten in $T^* X$ afkomstig van een kaart in X is $\pi(x,\, \xi) = x$ en we zien dat, in het punt $(x,\, \xi)$,

$$\tau = \sum_{i=1}^{n} \xi_i \; \mathrm{d}x_i. \tag{7.10}$$

De definitie (7.9) is echter evident coördinaat-invariant.

De differentiaalvorm τ wordt gekarakteriseerd door een zeer merkwaardige eigenschap. Zij g een differentieerbare differentiaalvorm in X, een differentieerbare afbeelding van X naar $T^* X$, waarvoor $\pi \circ g$ gelijk is aan de identiteit in X. Als we in (7.9) $\xi = g(x)$ invullen en het linker-en rechterlid laten volgen op $T_x\, g$ en we gebruiken dat $\pi \circ g$ gelijk is aan de identiteit in X, dan krijgen we voor iedere $x \in X$ dat

$$(g^* \tau)\,(x) = \tau_{g(x)} \circ T_x\, g = g(x) \circ T_{g(x)}\, \pi \circ T_x\, g = g(x) \circ T_x(\pi \circ g) = g(x).$$

Anders gezegd,

$$g = g^* \tau, \tag{7.11}$$

ofwel iedere differentieerbare differentiaalvorm g van de graad één in X kan worden verkregen door de ene differentiaalvorm τ in $T^* X$ terug te trekken door middel van de afbeelding $g : X \to T^* X$. Om deze reden zouden we τ de *tautologische éénvorm in* $T^* X$ kunnen noemen.

Omdat τ coördinaat-invariant gedefinieerd is, is zijn uitwendige afgeleide, de twee-vorm

$$\sigma := \mathrm{d}\tau \tag{7.12}$$

in $T^* X$ dat ook. Men noemt σ de *kanonieke twee-vorm in* $T^* X$. Merk op dat $\mathrm{d} \circ \mathrm{d} = 0$ geeft dat $\mathrm{d}\sigma = 0$, dus σ is een gesloten twee-vorm in $T^* X$ (de definitie impliceert dat σ exact is).

In lokale coördinaten afkomstig van een kaart in X wordt σ gegeven door

$$\sigma = \sum_{i=1}^{n} \mathrm{d}\xi_i \wedge \mathrm{d}x_i, \tag{7.13}$$

zoals direct volgt uit (7.10) en (5.40).

Lemma 7.1 *De kanonieke twee-vorm σ in $P := \mathrm{T}^* X$ is niet-gedegenereerd, in de volgende zin. Als $p \in P$, $u \in \mathrm{T}_p P$ en voor iedere $v \in \mathrm{T}_p P$ geldt dat $\sigma_p(u, v) = 0$, dan is $u = 0$.*

Bewijs We geven een bewijs in lokale coördinaten afkomstig van een kaart voor X, gedefinieerd in een omgeving van p. We lezen dan uit (7.10) en (5.20) af dat

$$\sigma_p(u, v) = \sum_{i=1}^{n} \mathrm{d}\xi_i(u) \cdot \mathrm{d}x_i(v) - \mathrm{d}\xi_i(v) \cdot \mathrm{d}x_i(u); \tag{7.14}$$

neem aan dat dit voor iedere $v \in \mathbf{R}^{2n}$ gelijk is aan nul. Nu vormen de $\mathrm{d}x_i(v)$ en $\mathrm{d}\xi_i(v)$ de coördinaten van v. Er is dus een vector $v = e_i$ waarvoor $\mathrm{d}x_i(v) = 1$ en alle andere coördinaten van v gelijk aan nul zijn; dit geeft dat $\mathrm{d}\xi_i(u) = 0$. Anderzijds geeft invullen van de vector $v = f_j$ waarvoor $\mathrm{d}\xi_i(v) = 1$ en alle andere coördinaten van v gelijk aan nul zijn dat $\mathrm{d}x_j(u) = 0$. De conclusie is dat alle coördinaten van u gelijk aan nul zijn, dus $u = 0$. □

Lemma 7.2 *Een vectorveld v, gedefinieerd in een open deelverzameling van $\mathrm{T}^* X$, is gelijk aan het snelheidsveld van het Hamilton-stelsel met Hamilton-functie f, dan en slechts dan als*

$$\mathrm{i}_v\, \sigma = -\,\mathrm{d}f. \tag{7.15}$$

Bewijs De vergelijking (7.15) in het punt $p = (x, \xi)$ betekent in lokale coördinaten afkomstig van een kaart in X dat voor iedere $u \in \mathbf{R}^{2n}$ de uitdrukking (7.14) gelijk is aan

$$\mathrm{d}f_p(u) = \sum_{i=1}^{n} \frac{\partial f(x, \xi)}{\partial x_i}\, \mathrm{d}x_i(u) + \frac{\partial f(x, \xi)}{\partial \xi_i}\, \mathrm{d}\xi_i(u).$$

Dit is het geval dan en slechts dan als voor iedere $1 \leq i \leq n$ geldt dat

$$\begin{aligned} \mathrm{d}x_i(v) &= \tfrac{\partial f(x,\xi)}{\partial \xi_i} \quad \text{en} \\ \mathrm{d}\xi_i(v) &= -\tfrac{\partial f(x,\xi)}{\partial x_i}. \end{aligned} \tag{7.16}$$

Maar dit betekent dat v gelijk is aan het snelheidsveld van het Hamilton-stelsel met Hamilton-functie f, zie (7.6) met h, resp. p vervangen door f, resp. ξ. □

Merk op dat we in Lemma 7.2 het vectorveld v en de functie h nog van t af mogen laten hangen, omdat het voor iedere t een algebraïsche identiteit betreft waarbij niet naar t gedifferentieerd wordt. Eenvoudigheidshalve beperken we ons in het vervolg van dit hoofdstuk tot het autonome geval.

Opmerking 7.1 De notatie in de literatuur is niet uniform. Lie [57] gebruikte net als hier de letters x_j, resp. p_j voor de positie-, resp. impulscoördinaten. Jacobi [42] handhaafde de p_j voor de impulscoördinaten, maar gaf de positiecoördinaten aan met q_j in plaats van x_j, vermoedelijk om de symmetrische manier waarop ze in een Hamilton-stelsel optreden te benadrukken. Deze notatie is zeer gebruikelijk geworden, met als verder uitbouw $\dot{q}_j$ in plaats van v_j voor de snelheidscoördinaten.

De hier ook gebruikte notatie van ξ_j voor de coördinaten van een lineaire vorm op $\mathrm{T}_x X$, als x_j de positiecoördinaten zijn, heeft het voordeel dat als we behalve de positie x ook een andere positie y willen beschouwen, we de lineaire vormen op de raakruimte in het punt y met de corresponderende griekse letter η aan kunnen duiden.

Als kanonieke twee-vorm van $\mathrm{T}^* X$ ziet men ook vaak $-\,\mathrm{d}\tau$ in plaats van $\mathrm{d}\tau$, dit heeft het voordeel dat in (7.15) geen minteken meer optreedt. Verder is ω een veelgebruikte letter voor de kanonieke twee-vorm. Tenslotte ziet men vaak vectorvelden aangeduid met X, een notatie die hier in botsing zou komen met notatie van X voor de positieruimte. Omdat de Hamilton-functie vaak aangeduid wordt met H, ziet men dan de notatie X_H voor het Hamilton-veld waarvan de Hamilton-functie gelijk is aan H. In de oudere literatuur komen geen speciale notaties voor de kanonieke twee-vorm of Hamilton-velden voor, omdat men alles steeds in coördinaten, afkomstig van een kaart in de positieruimte, uitschreef.

In Lemma 7.3 zullen we zien dat de Hamilton-vectorvelden lokaal gekarakteriseerd kunnen worden als de vectorvelden waarvan de stromingen de kanonieke symplectische vorm σ invariant laten. Het is historisch zeer opmerkelijk dat Lagrange weliswaar de bewegingsvergelijkingen niet als een Hamilton-stelsel schreef, maar wél de invariantie van de symplectische vorm ontdekte. Preciezer gezegd, hij introduceerde, uitgeschreven in lokale coördinaten, een tweevorm ρ in de raakbundel $\mathrm{T}\,X$. Vervolgens bewees hij dat ρ invariant is onder de stroming in $\mathrm{T}\,X$, waarvan de bewegingsvergelijkingen in Lagrange's vorm $[L] = 0$ staan. Nu herkent men gemakkelijk dat Lagrange's tweevorm ρ gelijk is aan $\Phi^*\sigma$, waarin $\Phi = \Phi_L$ de snelheids-impulstransformatie is en σ de kanonieke tweevorm in de coraakbundel $\mathrm{T}^* X$ voorstelt. Lagrange's uitspraak is equivalent met de invariantie van σ onder de stroming van het Hamilton-vectorveld H_h, omdat onder de snelheids-impulsafbeelding de bewegingsvergelijking $[L] = 0$ overgaat in het Hamilton-stelsel (7.6), zie Paragraaf Legss.

Dit resultaat van Lagrange is des te opmerkelijker, omdat tweevormen in Lagrange's tijd bepaald nog niet gebruikelijk waren en bovendien de tweevorm ρ er aanmerkelijk ingewikkelder uitziet dan de kanonieke tweevorm van de coraakbundel. $\oslash$

7.4 Symplectische Algebra

Is E een eindig-dimensionale lineaire ruimte, wordt een antisymmetrische bilineaire vorm σ in E een *symplectische vorm* genoemd als σ niet-gedegenereerd is in de zin dat als $u \in E$ en voor iedere $v \in E$ geldt dat $\sigma(u,\, v) = 0$, dan is $u = 0$. Anders gezegd, als de lineaire afbeelding $\sigma : u \mapsto \mathrm{i}_u\,\sigma$ van E naar E^* injectief is. Omdat E en E^* dezelfde dimensie hebben, is dit equivalent met de voorwaarde dat $\sigma : E \to E^*$ bijectief is. Dit is analoog aan de identificatie van E met E^* die bekend is als we in het bovenstaande de vorm σ vervangen door een inproduct in E. Een *symplectische lineaire ruimte* is gedefinieerd als een paar $(E,\, \sigma)$, waarin E een lineaire ruimte is en σ een symplectische vorm in E.

Er zijn echter ook grote verschillen met inproductruimten. Voor iedere deelverzameling L van E kunnen we net als bij inproducten het *orthogonale complement* L^σ *van* L *met betrekking tot* σ definiëren als de verzameling der $v \in E$ met de eigenschap dat voor iedere $u \in L$ geldt dat $\sigma(u,\, v) = 0$. Anders gezegd, L^σ is de doorsnede van de nulruimten in E der $\sigma(u) \in E^*$ met $u \in L$. Daarmee is de verzameling L^σ een lineaire deelruimte van E, met codimensie gelijk aan de dimensie van de lineaire ruimte opgespannen door $\sigma(L)$ = de dimensie van de lineaire ruimte opgespannen door L, omdat $\sigma : E \to E^*$ een lineair isomorfisme is. Verder is het evident dat $L \subset (L^\sigma)^\sigma$, dus als

L een lineaire deelruimte is van E, dan geeft de gelijkheid van de dimensies dat $L = (L^\sigma)^\sigma$. Men noemt een lineaire deelruimte L van E *isotroop* met betrekking tot σ als $L \subset L^\sigma$. Het opmerkelijke verschil met een inproduct is nu dat de antisymmetrie van σ betekent dat voor iedere $u \in E$ geldt dat $\sigma(u,\, u) = 0$, ofwel iedere één-dimensionale lineaire deelruimte van E is isotroop.

Is L isotroop dan is

$$\dim L \leq \dim L^\sigma = \dim E - \dim L,$$

dus $2 \dim L \leq \dim E$. Geldt hierin het $<$-teken, dan is er een $u \in L^\sigma$ met $u \notin L$. De antisymmetrie van σ gebruikend, krijgen we dat $L' := L + \mathbf{R}\, u$ isotroop is en $\dim L' = \dim L + 1$. Dit uitbreidingsproces voortzettend, zien we dat iedere isotrope lineaire deelruimte L van E bevat is in een maximale isotrope lineaire deelruimte $\widetilde{L}$ van E; dat wil zeggen $L \subset \widetilde{L} = \widetilde{L}^\sigma$, hetgeen ook impliceert dat $2 \dim \widetilde{L} = \dim E$. In het bijzonder zien we dat E noodzakelijkerwijze even-dimensionaal is. De maximale isotrope lineaire deelruimten van E hebben allen dezelfde dimensie, gelijk aan $\frac{1}{2} \dim E$, en worden de *Lagrange-vlakken* in de symplectische lineaire ruimte $(E,\, \sigma)$ genoemd. (De naam 'Lagrange-vlak' hangt samen met de naam 'Lagrange-variëteiten' die in Hoofdstuk 8 zullen worden besproken.)

7.5 Hamilton-stelsels in Symplectische Variëteiten

Zij M een C^r-variëteit, $r \geq 1$. Een *symplectische vorm in M* is gedefinieerd als een C^r-tweevorm σ in M met de volgende twee eigenschappen.

a) σ is gesloten, dat wil zeggen $\mathrm{d}\sigma = 0$.

b) Voor iedere $m \in M$ is σ_m niet-gedegenereerd, dat wil zeggen een symplectische vorm in $\mathrm{T}_m\, M$.

Een *symplectische variëteit* is een paar $(M,\, \sigma)$, waarin M een eindig-dimensionale differentieerbare variëteit is en σ een symplectische vorm in M. In Paragraaf 7.4 hebben we gezien dat dit impliceert dat M even-dimensionaal is.

Voorbeeld: iedere open deelverzameling M van $\mathrm{T}^*\, X$ is, met de beperking tot U van de kanonieke tweevorm, een symplectische variëteit, zie Lemma 7.1. Omgekeerd zegt het *lemma van Darboux* dat als $(M,\, \sigma)$ een symplectische variëteit is, dan is er bij iedere $m \in M$ een kaart $(x_1,\, \ldots,\, x_n,\, \xi_1,\, \ldots,\, \xi_n)$ in een open omgeving U van m in M, met de eigenschap dat σ in U gegeven wordt door (7.13). Zo'n kaart heet een *Darboux-kaart* voor $(M,\, \sigma)$. In Paragraaf 7.14 geven we een bewijs van het lemma van Darboux. We hebben vroeger al gezien dat als M de coraakbundel is van een variëteit X, dan is iedere kaart in $\mathrm{T}^*\, X$ die afkomstig is van een kaart voor X een Darboux-kaart.

Zij $(M,\, \sigma)$ een symplectische variëteit en zij f een differentieerbare reëelwaardige functie in M. Het *Hamilton-veld* H_f *gedefinieerd door de (Hamilton-)functie* f is het eenduidig bepaalde vectorveld v in M dat voldoet aan (7.15), in formule:

$$\mathrm{H}_f(m) = -\,(\sigma_m)^{-1}\,(\mathrm{d}f(m))\,, \quad m \in M.$$

Vanwege Lemma 7.2 is dit equivalent met de voorwaarde dat v in iedere Darboux-kaart gelijk is aan het snelheidsveld van het Hamilton-stelsel met Hamilton-functie gelijk aan f.

Lemma 7.3 *Zij $(M,\, \sigma)$ een symplectische C^r-variëteit, $r \geq 2$. Voor een C^{r-1}-vectorveld v in M zijn de volgende uitspraken* a) — d) *equivalent.*

a) *Voor iedere* $t \in \mathbf{R}$ *is* $\left(\mathrm{e}^{t\,v}\right)^* \sigma = \sigma$.

b) $\mathcal{L}_v \sigma = 0$.

c) *De éénvorm* $\theta := -\,\mathrm{i}_v\,\sigma$ *is gesloten.*

d) *Bij iedere* $m \in M$ *is er een open omgeving* U *van* m *in* M *en een reëelwaardige* C^r*-functie* f *in* U *met de eigenschap dat* $v = \mathrm{H}_f$ *in* U.

Bewijs De equivalentie van a) en b) is besproken bij (5.44). Gebruikmakend van de homotopieformule (5.47) voor de Lie-afgeleide en het gegeven dat $\mathrm{d}\sigma = 0$, krijgen we dat

$$\mathcal{L}_v \sigma = \mathrm{d}(\mathrm{i}_v\,\sigma) \tag{7.17}$$

en dit impliceert de equivalentie van b) en c). Tenslotte berust de equivalentie van c) en d) op het feit dat een C^{-1} éénvorm θ in een variëteit gesloten is, dan en slechts dan als deze lokaal gelijk is aan de afgeleide van een C^r-functie f. □

Lemma 7.3 karakteriseert, althans lokaal, de Hamilton-velden als precies die vectorvelden waarvan de stromingen de symplectische vorm invariant laten. Hiermee spelen Hamilton-velden een centrale rol in de symplectische differentiaalmeetkunde. Dat ze ook gerelateerd zijn aan Euler-Lagrange-vergelijkingen (en daarmee aan de klassieke mechanica) maakt ze extra interessant.

In het algemeen, zijn $(M,\,\sigma)$ en $(N,\,\rho)$ symplectische variëteiten van dezelfde dimensie, dan heet een C^r-afbeelding $\Phi : M \to N$ een *symplectische transformatie* als $\Phi^*\rho = \sigma$. Zijn M en N coraakbundels en zijn de symplectische vormen de bijbehorende kanonieke tweevormen, dan noemt men een symplectische transformatie ook wel een *kanonieke transformatie*. Lemma 7.3 zegt dat, althans lokaal, de Hamilton-velden precies die vectorvelden zijn waarvan de stromingen uit symplectische transformaties bestaan. In een coraakbundel: ... uit kanonieke transformaties bestaan.

Opmerking 7.2 De vraag of het vectorveld v met $\mathcal{L}_v\sigma = 0$ ook globaal een Hamilton-veld is, komt neer op de vraag of er bij de gesloten éénvorm θ een differentieerbare reëelwaardige functie in M is waarvoor $\mathrm{d}f = \theta$ in de hele variëteit M. Voor de bespreking van dit probleem beperken we ons tot de samenhangscomponent in M van een gegeven punt $m_0 \in M$. Dit betekent dat er voor iedere $m \in M$ een C^1-kromme $\gamma : [a,\,b] \to M$ is met $\gamma(a) = m_0$ en $\gamma(b) = m$. Als $\mathrm{d}f = \theta$ dan krijgen we dat

$$\int_\gamma \theta := \int_a^b \theta_{\gamma(t)}\left(\frac{d\gamma(t)}{dt}\right) dt = \int_a^b \frac{d}{dt} f(\gamma(t))\,dt = f(m) - f(m_0),$$

waarmee f op een additieve constante na eenduidig is bepaald in termen van θ. Het feit dat $\mathrm{d}\theta = 0$ is equivalent met de uitspraak dat $\int_\gamma \theta = \int_{\widetilde{\gamma}} \theta$ als $\widetilde{\gamma}$ een voldoend kleine storing is van γ en hetzelfde begin-en eindpunt heeft. Anders gezegd, $\int_\gamma \theta$ blijft gelijk bij een homotopie van γ waarbij begin-en eindpunt vastgehouden worden. We krijgen een globale oplossing f van de vergelijking $\mathrm{d}f = \theta$, dan en slechts dan als $\int_\gamma \theta = \int_{\widetilde{\gamma}} \theta$ voor *ieder* paar van krommen γ, $\widetilde{\gamma}$ met hetzelfde begin- en eindpunt. Op zijn beurt is dit equivalent met: $\int_\delta \theta = 0$ voor iedere kromme δ met $\delta(a) = \delta(b) = m_0$.

De geslotenheid van θ betekent dat $\int_\delta \theta = \int_{\widetilde{\delta}} \theta$ als $\widetilde{\delta}$ homotoop is met δ. De homotopieklasse van δ wordt met $[\delta]$ aangeduid. De homotopieklassen van gesloten krommen in M, startend en eindigend in m_0, vormen een groep, die de *fundamentaalgroep* $\pi_1(M,\, m_0)$ *van* M *met basispunt* m_0 genoemd wordt. Het bovenstaande geeft nu dat er een eenduidig bepaald homomorfisme $[\theta]$ is van $\pi_1(M,\, m_0)$ naar $\mathbf{R}$, met de eigenschap dat

$$[\theta]([\delta]) = \int_\delta \theta$$

voor iedere kromme δ met $\delta(a) = \delta(b) = m_0$. De homomorfismen van $\pi_1(M,\, m_0)$ naar $\mathbf{R}$ vormen een lineaire ruimte, die met $\mathrm{H}^1(M,\, \mathbf{R})$ aangeduid wordt en de *cohomologiegroep van* M *in dimensie* 1*, met coëfficiënten in* $\mathbf{R}$ genoemd wordt. Het is een gevolg van een algemene *stelling van de Rham*, dat $\theta \mapsto [\theta]$ een surjectieve lineaire afbeelding is, van de lineaire ruimte der gesloten continu differentieerbare differentiaalvormen in M, naar $\mathrm{H}^1(M,\, \mathbf{R})$. Het bovenstaande zegt verder $\theta = \mathrm{d}f$ voor een functie f, dan en slechts dan als $[\theta] = 0$. De conclusie is dat *ieder* vectorveld v met $\mathcal{L}_v\sigma = 0$ een Hamilton-veld is, dan en slechts dan als $\mathrm{H}^1(M,\, \mathbf{R}) = 0$. Dit is bijvoorbeeld het geval als M *enkelvoudig samenhangend* is, dat wil zeggen als iedere gesloten kromme in M samentrekbaar is naar een punt. Maar $\mathrm{H}^1(M,\, \mathbf{R}) = 0$ geldt ook al als er voor iedere gesloten kromme een positief geheel getal k is, waarvoor de kromme $k\,\delta$, die ontstaat k keer de kromme δ k te doorlopen, samentrekbaar is tot een punt. Er zijn voorbeelden van variëteiten die hieraan voldoen en niet enkelvoudig samenhangend zijn. $\oslash$

7.6 Kritieke Punten van Hamilton-stelsels

Een eerste eigenschap van een Hamilton-veld v die meteen opvalt is dat de nulpunten gelijk zijn aan de kritieke punten (= stationaire punten) van de Hamilton-functie f, de punten $m \in M$ waar $\mathrm{d}\, f(m) = 0$.

Een punt $m \in M$ is een nulpunt van v dan en slechts dan als m een *rustpunt* is van de v--stroming, dat wil zeggen voor iedere $t \in \mathbf{R}$ geldt dat $\mathrm{e}^{t\,v}(m) = m$. Dan is $L(t) := \mathrm{T}_m\left(\mathrm{e}^{t\,v}\right)$ een lineaire afbeelding van $\mathrm{T}_m\, M$ naar zichzelf, deze hangt differentieerbaar af van t en in lokale coördinaten krijgen we dat

$$A := L'(0) = \mathrm{D}\, v(m).$$

Dit laat zien dat $\mathrm{D}\, v(m)$ coördinaat-invariant gedefinieerd is als een lineaire afbeelding van $\mathrm{T}_m\, M$ naar zichzelf; omdat $\frac{dL(t)}{dt} = A \circ L(t)$ en $L(0) = I$ krijgen we dat $L(t) = \mathrm{e}^{t\,A}$. Men noemt A de *linearisering van het vectorveld in het punt* m en $L(t)$ de *linearisering van de stroming van* v.

Anderzijds geeft differentiatie van $-\,\mathrm{i}_v\, \sigma = \mathrm{d}\, f$ in het punt m in lokale coördinaten dat

$$Q := -\sigma_m \circ A = \mathrm{D}^2\, f(m).$$

Enerzijds bevestigt dit het bekende feit dat voor een stationair punt m van een C^2-functie f de tweede-orde afgeleide $\mathrm{D}^2\, f(m)$ een coördinaat-invariant gedefinieerde symmetrische bilineaire vorm in $\mathrm{T}_m\, M$ is en anderzijds geeft $Q = Q^*$ dat

$$\sigma_m \circ A = (\sigma_m \circ A)^* = A^* \circ {\sigma_m}^* = -A^* \circ \sigma_m.$$

Deze laatste eigenschap volgt ook uit het feit dat de v-stroming de symplectische vorm invariant laat. Immers, dit geeft dat

$$\sigma_m = \left(\left(\mathrm{e}^{t\,v}\right)^* \sigma\right)_m = L(t)^* \sigma_m,$$

hetgeen uitgeschreven betekent dat voor iedere u en v in $\mathrm{T}_m\,M$ geldt dat

$$\sigma_m(L(t)\,u,\,L(t)\,v) = \sigma_m(u,\,v). \tag{7.18}$$

Differentiatie hiervan naar t in $t=0$ geeft dat

$$\sigma_m(A\,u,\,v) + \sigma_m(u,\,A\,v) = 0, \quad u,\,v \in \mathrm{T}_m\,M, \tag{7.19}$$

ofwel $\sigma_m \circ A + A^* \circ \sigma_m = 0$. Men gaat verder nog gemakkelijk na dat omgekeerd (7.18) volgt uit (7.19).

Is $(E,\,\sigma)$ een symplectische lineaire ruimte, dan noemt men een lineaire afbeelding $L : E \to E$ een *symplectische lineaire transformatie*, als $L^*\sigma = \sigma$. Omdat dit impliceert dat $\ker L \subset \ker \sigma = 0$, zien we dat dit impliceert dat L bijectief is. Men noemt een lineaire afbeelding $A : E \to E$ een *infinitesimaal-symplectische lineaire transformatie* als $\sigma \circ A + A^* \circ \sigma = 0$.

Het bovenstaande argument geeft dat A infinitesimaal-symplectisch is, dan en slechts dan als voor iedere $t \in \mathbf{R}$ de afbeelding $\mathrm{e}^{t\,A}$ een symplectische transformatie is. Verder, is m een nulpunt van $v = \mathrm{H}_f$, dus een stationair punt van f, dan is

$$A = Dv(m) = -{\sigma_m}^{-1} \circ \mathrm{D}^2 f(m)$$

een coördinaat-invariant gedefinieerde infinitesimaal-symplectische lineaire transformatie in $\mathrm{T}_m\,M$, met $\mathrm{e}^{t\,A}$ gelijk aan de symplectische lineaire transformatie $\mathrm{T}_m\left(\mathrm{e}^{t\,v}\right)$.

Lemma 7.4 *Zij m een rustpunt van een continu differentieerbaar Hamilton-veld v, zij $A = Dv(m)$: $\mathrm{T}_m\,M \to \mathrm{T}_m\,M$ de coördinaat-invariant gedefinieerde linearisatie van v in het punt m. Is $\lambda \in \mathbf{C}$ een eigenwaarde van A met algebraïsche multipliciteit gelijk aan k, dan is ook $-\lambda$ eigenwaarde van A met algebraïsche multipliciteit gelijk aan k. Verder heeft de eigenwaarde 0 een even algebraïsche multipliciteit. Er treden de volgende mogelijkheden op voor de eigenwaarden λ.*

a) $\lambda = 0$ *met even algebraïsche multipliciteit.*

b) *Een positieve reële eigenwaarde λ en zijn tegengestelde $-\lambda$, met dezelfde algebraïsche multiplipliciteit.*

c) $\lambda = i\,\omega$ *met ω positief reëel en zijn tegengestelde = complex-geconjugeerde $-i\,\omega$, met dezelfde algebraïsche multipliciteit.*

d) $\lambda_{++} = \rho + i\,\omega$, $\lambda_{+-} = \rho - i\,\omega$, $\lambda_{--} = -\rho - i\,\omega$ *en* $\lambda_{-+} = -\rho + i\,\omega$, *met ρ en ω positef reëel; alle vier deze eigenwaarden met dezelfde algebraïsche multipliciteit.*

Bewijs We gebruiken dat voor de infinitesimaal-symplectische lineaire afbeelding A geldt dat

$$\sigma \circ A \circ \sigma^{-1} = -A^*,$$

ofwel A is door middel van σ geconjugeerd aan $-A^*$. Dit geeft voor iedere $\lambda \in \mathbf{C}$ dat

$$\sigma \circ (A - \lambda I) \circ \sigma^{-1} = \sigma \circ A \circ \sigma^{-1} - \lambda\,I = -A^* - \lambda\,I = -(A + \lambda\,I)^*.$$

Nu geldt dat $\det\left(\sigma \circ B \circ \sigma^{-1}\right) = \det B$, $\det C^* = \det C$ en $\det(-D) = (-1)^{2n} \det D = \det D$, dus we krijgen dat

$$\det(A - \lambda\, I) = \det(A + \lambda\, I), \quad \lambda \in \mathbf{C}. \tag{7.20}$$

Dit betekent dat de karakteristieke veelterm $p(\lambda)$ van A symmetrisch is in de zin dat $p(\lambda) = p(-\lambda)$, hetgeen op zijn beurt betekent dat ieder nulpunt $-\lambda$ dezelfde multipliciteit heeft als λ.

Omdat het totaal aantal nulpunten gelijk is aan $2n$, dus even is, en de nulpunten ongelijk aan nul in tegengestelde paren optreden, heeft het overblijvende nulpunt $\lambda = 0$ even multipliciteit.

Voor de lijst a) — d) combineren we het tot nu toe verkregen resultaat met het feit dat A reëel is, hetgeen impliceert dat ook de complex-geconjugeerde eigenwaarde $\overline{\lambda}$ dezelfde algebraïsche multipliciteit heeft als de eigenwaarde λ.

$\square$

Opgemerkt kan nog worden dat d) alleen kan optreden als $\dim M = 2n \geq 4$, dus als het aantal n van vrijheidsgraden van het mechanische systeem minstens gelijk aan twee is.

Voor een nog meer gedetailleerd begrip van de mogelijkheden zou men kunnen vragen naar een standaardvorm voor een de matrix van een infinitesimaal-symplectische lineaire transformatie, ten aanzien van een basis waarop σ_m in de standaarvorm (7.13) staat. Dit zou hier echter teveel tijd vergen.

De raakruimte $\mathrm{T}_m\, M$ is gelijk aan de directe som van lineaire deelruimten T_-, T_0 en T_+, waarbij $A(T_-) \subset T_-$, $A(T_0) \subset T_0$, $A(T_+) \subset T_+$ en van de beperking van A tot T_-, resp. T_0, resp. T_+ zijn de reële delen van de eigenwaarden negatief, resp. gelijk aan nul, resp. positief. De T_+-component van oplossingen van het gelineariseerde stelsel groeit exponentieel voor $t \to \infty$ als deze component niet gelijk aan nul is. Dit geeft dat de oplossing alleen begrensd kan zijn voor $t \geq 0$ als deze in $T_- + T_0$ ligt.

Voor het gedrag van de oplossingen $m(t)$ van het niet-lineaire stelsel in een omgeving van het evenwichtspunt m heeft men de volgende conclusie. Er is een C^{r-1}-deelvariëteit S van M met $m \in S$, $\mathrm{T}_m\, S = T_- + T_0$ en S is lokaal invariant onder de stroming, in de zin dat er een (kleine) open omgeving U van m in M is met de eigenschap dat als een oplossing in S start, dan kan deze alleen uit S weglopen door uit U weg te lopen. Verder geldt voor iedere oplossing $m(t)$ waarvoor $m(t) \in U$ voor alle $t \geq 0$, dat $m(0) \in S$. Anders gezegd, alle oplossingen die niet in S starten lopen uit U weg (waarbij de afstand tot m net als bij het lineaire stelsel exponentieel toeneemt). De variëteit S heet ook wel de *centrum-stabiele variëteit* van het evenwichtspunt m. Voor een bewijs van de existentie hiervan, zie bijvoorbeeld Hirsch, Pugh and Shub [35]. Een gevolg is dat het evenwichtspunt m alleen maar stabiel kan zijn in de zin dat alle oplossingen die dicht bij m starten ook dicht bij m blijven, als $T_+ = \{0\}$, dat wil zeggen als A geen eigenwaarden heeft met positief reëel deel. (De uitspraak dat de oplossingen die niet in S starten uit U weglopen is echter veel sterker.)

Voor een Hamilton-stelsel betekent dit dat het evenwichtspunt alleen maar stabiel kan zijn als alleen a) of c) voorkomen. Interessant is hierbij nog het volgende verschijnsel van 'behoud van stabiliteit onder storingen'. Het is bekend dat de eigenwaarden van A op continue manier afhangen van (de matrixcoëfficiënten van) A. Hebben we een paar van eigenwaarden als in c) met multipliciteit gelijk aan één, dan kunnen deze eigenwaarden bij kleine storingen niet van de imaginaire as af, omdat we dan in het geval d) zouden komen en er minstens vier eigenwaarden dicht bij het oorspronkelijke zouden verschijnen. Anders gezegd, de verzameling van de infinitesimaal-symplectische lineaire afbeeldingen A, waarvan alle eigenwaarden zuiver-imaginair en enkelvoudig zijn, vormen een open deelverzameling van de verzameling van alle infinitesimaal-symplectische lineaire afbeeldingen.

7.7 Constanten van Beweging

Lemma 7.5 *De Hamilton-functie f is een constante van beweging voor het Hamilton-stelsel met Hamilton-functie gelijk aan f.*

Bewijs Dit volgt direct uit de opmerking dat, voor iedere $m \in M$,

$$\mathrm{d}f(m)\left(\mathrm{H}_f(m)\right) = -\sigma_m\left(\mathrm{H}_f(m),\, \mathrm{H}_f(m)\right) = 0.$$

□

Is het Hamilton-stelsel afkomstig uit de klassieke mechanica, dan is f gelijk aan de totale energie, beschouwd als functie van positie en impuls en hebben we de *wet van behoud van energie* teruggevonden. Lemma 7.5 kan dus beschouwd worden als een generalisatie van de wet van behoud van energie.

In de open deelverzameling

$$M^{\mathrm{reg}} := \{m \in M \mid \mathrm{d}f(m) \neq 0\}$$

zijn de niveauverzamelingen $M_c := f^{-1}(\{c\})$ van f gesloten $(2n-1)$-dimensionale C^r-deelvariëteiten van de $2n$-dimensionale variëteit M. De rustpunten van het stelsel zijn precies de punten in $M^{\mathrm{sing}} := M \setminus M^{\mathrm{reg}}$; als $m \in M^{\mathrm{sing}}$ en $c = f(m)$, dan vertoont M_c gewoonlijk in het punt m singulier gedrag.

Als $\mathrm{d}f(m) = 0$ en de symmetrische bilineaire vorm $\mathrm{D}^2 f(m)$ op $\mathrm{T}_m M$ is niet-gedegenereerd, dan zegt het *Morse-lemma* dat er een C^{r-2}-kaart κ in een open omgeving M_κ van m in M is, waarvoor $\kappa(m) = 0$ en

$$f = f(m) + \frac{1}{2}\left(\sum_{j=1}^{s} {\kappa_j}^2 - \sum_{j=s+1}^{2n} {\kappa_j}^2\right)$$

in M_κ. (Een bewijs van het Morse-lemma staat in het Analyse C-dictaat. Waarschuwing: κ is gewoonlijk geen Darboux-kaart.) We zien dat in M_κ alleen m een singulier punt is. Is $1 \leq s \leq 2n-1$, dan ziet de niveauverzameling voor het niveau $f(m)$ eruit als een kegel met top in het punt m en is daarmee niet gelijk aan een differentieerbare variëteit in het punt m. De naburige niveauverzamelingen zien er dan uit als gladde hyperboloïden.

Is echter $s = 2n$, resp. $s = 0$, corresponderend met de uitspraak dat f een lokaal minimum, resp. lokaal maximum heeft in het punt m, dan bestaat de niveauverzameling uit alleen maar het punt m. Voor een niveau een klein beetje groter, resp. kleiner dan $f(m)$ ziet de niveauverzameling in M_κ eruit als een kleine sfeer om het punt m, terwijl voor niveau kleiner, resp. groter dan $f(m)$ de niveauverzameling in M_κ leeg is. Omdat f een constante van beweging is, betekent dit dat voor alle oplossingen van het H_f-stelsel die op een kleine sfeer om m (in de kaart κ) starten voor alle $t \in \mathbf{R}$ op deze sfeer blijven.

Men zegt dat een rustpunt m van een stelsel *Lyapunov-stabiel* is als er bij iedere omgeving U van m een omgeving V van m is, met de eigenschap dat iedere oplossing die in V start vervolgens voor alle $t \geq 0$ in U blijft. Het bovenstaande geeft dat m een Lyapunov-stabiel rustpunt is van H_f, als $\mathrm{D}^2 f(m)$ positief definiet is of negatief definiet. Het is gemakkelijk in te zien dat dit impliceert dat alle eigenwaarden van A zuiver imaginair zijn. Echter, de omkering hoeft niet waar te zijn: het rustpunt m kan Lyapunov-stabiel zijn en daarmee alle eigenwaarden van A zuiver imaginair, terwijl toch $1 \leq s \leq 2n-1$.

Zijn f en g twee differentieerbare reëelwaardige functies in M, dan definieert men de *Poisson-haakjes van f en g* als de reëelwaardige functie

$$\{f,\, g\} := \mathrm{i}_{\mathrm{H}_f}(\mathrm{d}g) = \mathcal{L}_{\mathrm{H}_f}(g)\,, \tag{7.21}$$

de afgeleide van g in de richting van het Hamilton-veld met Hamilton-functie gelijk aan f. Hiervoor hebben we:

Lemma 7.6 *Voor iedere $m \in M$ is*

$$\{f,\, g\}(m) = \sigma_m\left(\mathrm{H}_f(m),\, \mathrm{H}_g(m)\right), \tag{7.22}$$

terwijl in Darboux-coördinaten de Poisson-haakjes worden gegeven door

$$\{f,\, g\}(x,\, \xi) = \sum_{j=1}^{n} \frac{\partial f(x,\, \xi)}{\partial \xi_j}\,\frac{\partial g(x,\, \xi)}{\partial x_j} - \frac{\partial f(x,\, \xi)}{\partial x_j}\,\frac{\partial g(x,\, \xi)}{\partial \xi_j}. \tag{7.23}$$

De Poisson-haakjes zijn antisymmetrisch in de zin dat

$$\{f,\, g\} = -\{g,\, f\}. \tag{7.24}$$

Tenslotte, de volgende uitspraken a) — c) *zijn equivalent:*

a) *g is een constante van beweging voor de H_f-stroming.*

b) $\{f,\, g\} = 0$.

c) *f is een constante van beweging voor de H_g-stroming.*

Bewijs De formule (7.22) volgt uit (7.21) en $\mathrm{d}g = -\,\mathrm{i}_{\mathrm{H}_g}\,\sigma$. Daarbij is ook de antisymmetrie van de bilineaire vorm σ_m gebruikt; deze levert met het oog op (7.22) ook de antisymmetrie van de Poissonhaakjes. De equivalentie tussen a) en b) volgt direct uit de definitie (7.21), terwijl b) equivalent is met $\{g,\, f\} = -\{f,\, g\} = -0 = 0$, dus met c). $\square$

Men zegt dat de $\mathrm{H_g}$-stroming Ψ^t een *éénparameter symmetriegroep* voor de Hamilton-functie f is, als f invariant is onder de transformaties Ψ^t, dat wil zeggen voor iedere $t \in \mathbf{R}$ geldt dat $f \circ \Psi^t = f$. De implicatie c) $\Rightarrow$ a) in Lemma 7.6 krijgt daarmee de volgende vorm:

Gevolg 7.7 Is de H_g-stroming een éénparameter symmetriegroep voor de functie f, dan is g een constante van beweging voor het Hamilton-stelsel met Hamilton-functie gelijk aan f.

Opmerking 7.3 Zij L een voldoende vaak differentieerbare functie op $\mathrm{T}\,X$ waarvoor de snelheids--impuls-afbeelding Φ een diffeomorfisme is van $\mathrm{T}\,X$ naar $\mathrm{T}^*\,X$ en zij f de Legendre transformatie van L, de Hamilton-functie van het vectorveld in $\mathrm{T}^*\,X$ dat via de Φ correspondeert met het Euler-Lagrange stelsel in $\mathrm{T}\,X$.

Zij nu w een voldoende vaak differentieerbaar vectorveld in de variëteit X, met stroming ϕ^s. Dan hebben we de geïnduceerde stroming $\mathrm{T}\,\phi^s$ in de raakbundel $\mathrm{T}\,X$ van X, terwijl in Paragraaf 7.10 een geïnduceerde stroming $\widetilde{\phi^s}$ in $\mathrm{T}^*\,X$ wordt ingevoerd. Deze laatste is de stroming van een Hamilton-vectorveld, met Hamilton-functie i_w gedefinieerd door $i_w(x,\,\xi) = \langle w(x),\,\xi\rangle$, zie Stelling 7.11. Merk op dat $\Phi^* i_w = I_w$, de w-component I_w van de impuls die is ingevoerd in Stelling 4.4. Met deze notaties zijn de volgende vier uitspraken equivalent.

a) L is invariant onder de geïnduceerde stroming $\mathrm{T}\,\phi^s$ in $\mathrm{T}\,X$.

b) I_w is een constante van beweging voor het Euler-Lagrange-stelsel $[L] = 0$ in $\mathrm{T}\,X$.

c) i_w is een constante van beweging voor het Hamilton-stelsel in $\mathrm{T}^*\,X$ met Hamilton-functie gelijk aan f.

d) f is invariant onder de geïnduceerde stroming $\widetilde{\phi^s}$ in $\mathrm{T}^*\,X$.

De equivalentie van a) met b) volgt uit het bewijs van de stelling van Noether, Stelling 4.4. De equivalentie van b) met c) volgt uit $I_w = \Phi^* i_w$, gecombineerd met het feit dat Φ oplossingen van $[L] = 0$ overvoert in oplossing van het Hamilton-stelsel in $\mathrm{T}^*\,X$ met Hamilton-functie gelijk aan f. Tenslotte volgt de equivalentie van c) met d) uit de equivalentie van a) met c) in Lemma 7.6, met $g = i_w$.

Vanwege deze relatie met de stelling van Noether wordt Gevolg 7.7 ook wel *Noether's principe voor Hamilton-systemen* genoemd.

De equivalentie tussen a) met d) is niet zo direct duidelijk, omdat alleen maar geldt dat $\Phi^* f = L$ als $L(x,\,v)$ homogeen van de graad twee is als functie van v. Echter, de voorwaarde a) impliceert dat ook $\Phi * f$ invariant is onder de geïnduceerde stroming $\mathrm{T}\,\phi^s$ in $\mathrm{T}\,X$ en dat $\Phi \circ \mathrm{T}\,\phi^s = \widetilde{\phi^s} \circ \Phi$. $\oslash$

7.8 De Jacobi-identiteit

Veronderstel dat v en w beiden continu differentieerbare vectorvelden zijn op een variëteit X. Analoog aan (5.44) kunnen we de *Lie-afgeleide* $\mathcal{L}_v w$ *van het vectorveld* w *naar het vectorveld* v, meer bekend als de *Lie-haakjes* $[v,\,w]$ *van* v *en* w, definiëren door middel van

$$[v,\,w] := \mathcal{L}_v w := \frac{d}{dt}\left(\mathrm{e}^{t\,v}\right)^*\,w|_{t=0}, \tag{7.25}$$

waarin het teruggetrokken vectorveld $\left(\mathrm{e}^{t\,v}\right)^*\,w$ gedefinieerd is als in (5.38), met Φ, resp. v vervangen door $\mathrm{e}^{t\,v}$, resp. w. Omdat bij het invullen in (5.38) de variabele t op twee plaatsen voorkomt, krijgen we bij het uitwerken van de differentiatie naar t in $t = 0$ twee termen. In lokale coördinaten leidt dit tot de volgende formule

$$[v,\,w](x) = \mathrm{D}\,w(x)\cdot v(x) - \mathrm{D}\,v(x)\cdot w(x), \quad x \in X. \tag{7.26}$$

Hoewel v en w in de definitie (7.25) heel verschillende rollen lijken te spelen, laat (7.26) zien dat

$$[v,\,w] = -[w,\,v], \tag{7.27}$$

dat wil zeggen de Lie-haakjes van vectorvelden zijn *antisymmetrisch* onder versisseling van de vectorvelden. Ook kan nog opgemerkt worden dat de Lie-haakjes een bilineaire afbeelding $(v,\, w) \mapsto [v,\, w]$ van $\mathrm{V}^l(X) \times \mathrm{V}^l(X)$ naar $\mathrm{V}^{l-1}(X)$ definiëren. Hierin is $\mathrm{V}^l(X)$ de lineaire ruimte van de l keer continu differentieerbare vectorvelden op X.

Er is nog een andere leerzame manier om tegen de Lie-haakjes aan te kijken, namelijk door in plaats van naar het vectorveld v, gedefinieerd als een afbeelding van X naar $\mathrm{T}\,X$, te kijken naar de differentiaaloperator $v_{\mathrm{op}} : f \mapsto \mathcal{L}_v f = \mathrm{i}_v\, \mathrm{d}f$ van 'differentiatie van functies in de richting van het vectorveld v', in lokale coördinaten gegeven door

$$\big(v_{\mathrm{op}}(f)\big)\,(x) = \sum_{j=1}^{n} v_j(x)\, \frac{\partial f(x)}{\partial x_j}. \tag{7.28}$$

Dit is een lineaire afbeelding van $\mathrm{C}^k(X)$ naar $\mathrm{C}^{k-1}(X)$, als $v \in \mathrm{V}^{k-1}(X)$. Vervangen we in (5.37) de afbeelding Φ, resp. het vectorveld v door $\mathrm{e}^{t\,v}$, resp. w en differentiëren we de zo ontstane identiteit naar t in $t = 0$, dan krijgen we

$$\mathcal{L}_v\,(\mathrm{i}_w\,\omega) = \mathrm{i}_{\mathcal{L}_v w}\,\omega + \mathrm{i}_w\,(\mathcal{L}_v\omega)\,. \tag{7.29}$$

Dit toepassend op $\omega = \mathrm{d}f$ en gebruikend dat $\mathcal{L}_v\,\mathrm{d}f = \mathrm{d}\,\mathcal{L}_v f$, zie (5.46), krijgen we als conclusie dat

$$[v,\, w]_{\mathrm{op}} = \left[v_{\mathrm{op}},\, w_{\mathrm{op}}\right], \tag{7.30}$$

waarbij de *commutator* $[A,\, B]$ van de lineaire afbeeldingen A en B is gedefinieerd als

$$[A,\, B] := A \circ B - B \circ A. \tag{7.31}$$

In woorden: als we vectorvelden opvatten als differentiaaloperatoren, dan correspondeert de Lie-haakjes van twee vectorvelden met de commutator van de bijbehorende operatoren. De afbeelding $v \mapsto v_{\mathrm{op}}$ is injectief, immers door in (7.28) voor f de j-de coördinaatsfunctie te nemen krijgen we $v_j(x)$, dus als $v_{\mathrm{op}} = w_{\mathrm{op}}$, dan geldt voor iedere x en iedere j dat $v_j(x) = w_j(x)$. Hiermee kunnen we de antisymmetrie van de Lie-haakjes ook zien als een gevolg van de evidente antisymmetrie van de commutator: uit (7.30) en de anti-symmetrie van de commutator volgt dat $[v,\, w]_{\mathrm{op}} = -[w,\, v]_{\mathrm{op}} = (-[w,\, v])_{\mathrm{op}}$, waarna de injectiviteit van $v \mapsto v_{\mathrm{op}}$ geeft dat $[v,\, w] = -[w,\, v]$.

Voor commutatoren van lineaire afbeeldingen A, B, C geldt ook de zogenaamde *Jacobi-identiteit*

$$[[A,\, B],\, C] + [[B,\, C],\, A] + [[C,\, A],\, B] = 0, \tag{7.32}$$

hierin zijn de tweede en de derde term in het linkerlid uit de eerste term onstaan door *cyclisch verwisselen* van A, B en C. De formule (7.32) kan door uitschrijven van de commutatoren gemakkelijk nagegaan worden, maar dit geeft nog geen inzicht hoe men op deze formule gekomen is. Nogmaals (7.30) en de injectiviteit van $v \mapsto v_{\mathrm{op}}$ gebruikend, zien we dat ook de Lie-haakjes van $u,\, v,\, w \in \mathrm{V}^2(X)$ voldoen aan de Jacobi-identiteit

$$[[u,\, v],\, w] + [[v,\, w],\, u] + [[w,\, u],\, v] = 0. \tag{7.33}$$

Het uitschrijven van alle termen hierin met behulp van (7.26) zou al heel wat ruimte in beslag nemen; hiermee kan (7.33) ook gezien worden als een gevolg van de verwisselbaarheid van de differentiatievolgorde bij tweede orde afgeleiden.

Stelling 7.8 *Zijn f, g en h tweemaal continu differentieerbare functies op de symplectische variëteit $(M,\,\sigma)$, dan geldt de Jacobi-identiteit voor de Poisson-haakjes:*

$$\{\{f,\,g\},\,h\}+\{\{g,\,h\},\,f\}+\{\{h,\,f\},\,g\}=0, \quad f,\,g,\,h\in \mathrm{C}^2(M). \tag{7.34}$$

Bewijs Er geldt

$$\mathrm{i}_{[\mathrm{H}_f,\,\mathrm{H}_g]}\,\sigma=\mathcal{L}_{\mathrm{H}_f}\left(\mathrm{i}_{\mathrm{H}_g}\,\sigma\right)=\mathcal{L}_{\mathrm{H}_f}(-\,\mathrm{d}g)=-\,\mathrm{d}\,\mathcal{L}_{\mathrm{H}_f}g=-\,\mathrm{d}\{f,\,g\},$$

ofwel

$$[\mathrm{H}_f,\,\mathrm{H}_g]=\mathrm{H}_{\{f,\,g\}}\,. \tag{7.35}$$

In de afleiding is achtereenvolgens gebruik gemaakt van (7.25), (7.29), $\mathcal{L}_{\mathrm{H}_f}\sigma=0$, (5.46) en (7.21).

Als we nu het linker- en rechterlid van (7.35) als differentiaaloperator laten werken op de functie h dan krijgen we de identiteit

$$\{f,\,\{g,\,h\}\}-\{g,\,\{f,\,h\}\}=\{\{f,\,g\},\,h\}. \tag{7.36}$$

Als we nu op een paar plaatsen de antisymmetrie (7.24) van de Poisson-haakjes (zie Lemma 7.6) gebruiken, dan zien we dat (7.36) equivalent is met (7.34). Daarmee is (7.35) equivalent met (7.34) voor alle functies h. □

Gebruikmakend van de formule (7.23) in Darboux-coördinaten kan de Jacobi-identiteit (7.34) ook gezien worden als een gevolg van de symmetrie van de tweede orde partiële afgeleiden. Deze berekening is niet moeilijk, al verschijnen er wel heel wat termen die vervolgens weer allemaal tegen elkaar wegvallen.

Het volgende resultaat staat bekend als de *stelling van Poisson*.

Gevolg 7.9 *Zijn g en h constanten van beweging voor het Hamilton-stelsel dat is gedefinieerd door de functie f, dan is $\{g,\,h\}$ ook een constante van beweging voor H_f.*

Bewijs We moeten aantonen dat $\{f,\,\{g,\,h\}\}=0$ als $\{f,\,g\}=0$ en $\{f,\,h\}=0$. Dit volgt meteen uit (7.34) en de antisymmetrie van de Poisson-haakjes. □

Opmerking 7.4 De identiteit (7.34) komt onder andere voor op p. 41 in het posthuum in 1862 verschenen artikel [40] van Jacobi, dat hij in de 1830-er jaren heeft geschreven. Daarin vestigde Jacobi ook de aandacht op de stelling van Poisson (= Gevolg 7.9), die hier meteen uit volgt. De naam 'Poisson-haakjes' is vermoedelijk hierom in zwang gekomen, al kwamen de Poisson-haakjes ook al bij Lagrange [52, p. 315] voor.

Geïnspireerd door de identiteit (7.34) voor de Poisson-haakjes, werd de identiteit (7.33) voor vectorvelden naar voren gebracht door Lie [56, Kap. 5, §26], zie ook Lie [57, Kap. 7, §44, 45]. Daarbij sloeg de naam 'Jacobi-identiteit' in die tijd nog uitsluitend op (7.34).

Lie introduceerde ook een algemene lineaire ruimte L met een bilineaire afbeelding

$$(u,\,v)\mapsto [u,\,v]:L\times L\to L,$$

die voor iedere $u, v, w \in L$ voldoet aan (7.33). Zo'n paar $(L, [\cdot, \cdot])$ heet tegenwoordig een *Lie--algebra*. Daarin heten $[u, v]$ de *Lie-haakjes van* $u, v \in L$ en (7.33) voor $u, v, w \in L$ heet de *Jacobi-identiteit* in de Lie-algebra L.

In deze terminologie geeft de theorie van deze paragraaf dat $\mathrm{V}^\infty(X)$, resp. de ruimte $\mathrm{L}(E, E)$ van lineaire afbeeldingen van een vectorruimte E naar zichzelf, resp. $\mathrm{C}^\infty(M)$ een Lie-algebra is ten aanzien van de Lie-haakjes van vectorvelden, resp. de commutator van lineaire afbeeldingen, resp. de Poisson-haakjes van functies. Daarbij zegt (7.25), resp. (7.35) dat de afbeelding $v \mapsto v_{\mathrm{op}}$, resp. $f \mapsto \mathrm{H}_f$ een zogenaamd *homomorfisme van Lie-algebra's* is van $\mathrm{V}^\infty(X)$ naar $\mathrm{L}\left(\mathrm{C}^\infty(X), \mathrm{C}^\infty(X)\right)$, resp. van $\mathrm{C}^\infty(M)$ naar $\mathrm{V}^\infty(M)$.

In de quantummechanica zijn de observabelen lineaire operatoren in een Hilbert-ruimte, waarbij de commutator, bijvoorbeeld in Heisenberg's onzekerheidsrelatie, een belangrijke rol speelt. De theorie van deze paragraaf vormt daarom een belangrijke stimulans in de discussie over de relatie tussen de quantummechanica en de klassieke mechanica, waarbij de laatste in het raamwerk van Hamilton-stelsels beschouwd wordt. Daarbij zegt men dat de functies op de faseruimte (de symplectische variëteit M) de observabelen zijn in de klassieke mechanica en dat in de relatie tussen de quantummechanica en de klassieke mechanica er een verband zou moeten zijn tussen de operatoren in de Hilbert-ruimte en de functies op de faseruimte, waarbij de commutator van operatoren correspondeert met de Poisson-haakjes van de functies. Eigenlijk verwacht men hier alleen een asymptotische relatie, omdat alleen op grote (macroscopische) schaal de klassieke mechanica een goed substituut is voor de quantummechica. $\oslash$

7.9 Volumebehoud

We besluiten deze invoering van Hamilton-stelsels met de opmerking dat $\mathcal{L}_v\sigma = 0$ met inductie over k impliceert dat $\mathcal{L}_v\left(\sigma^k\right) = 0$. Hierin is de $2k$-vorm σ^k met inductie over k gedefinieerd door middel van:

$$\sigma^{k+1} := \sigma \wedge \sigma^k = \sigma^k \wedge \sigma.$$

Als $k > n$, dan is de graad van σ^k groter dan de dimensie van de variëteit, dus dan is $\sigma^k = 0$ en dan levert dit geen interessante informatie op. Echter, voor $k = n$ dan levert de formule (7.13) voor σ in Darboux-coördinaten dat

$$\begin{aligned}\sigma^n &= n!\, \mathrm{d}\xi_1 \wedge \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}\xi_n \wedge \mathrm{d}x_n \\ &= n!\,(-1)^{n(n+1)/2}\, \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n \wedge \mathrm{d}\xi_1 \wedge \ldots \wedge \mathrm{d}\xi_n.\end{aligned}$$

Hieruit lezen we af dat de *kanonieke volume-vorm*

$$\omega := (-1)^{n(n+1)/2}\, \frac{1}{n!}\, \sigma^n, \tag{7.37}$$

die in Darboux-coördinaten gelijk is aan de standaard volumevorm in $\mathbf{R}^n \times \mathbf{R}^n$, nergens gelijk aan nul is. (Terzijde: de existentie van een nergens verdwijnende continue volumevorm in een variëteit is equivalent met de oriënteerbaarheid van de variëteit; de kanonieke volumevorm van een symplectische variëteit (M, σ) leidt tot een oriëntatie van M.) Hieruit volgt uiteraard dat voor iedere $1 \le k \le n$ de $2k$-vorm σ^k nergens gelijk aan nul is.

De volgende conclusie staat bekend als de *stelling van Liouville voor Hamilton-stelsels*:

Lemma 7.10 *Zij ω de kanonieke volume-vorm (7.37) van een symplectische variëteit (M, σ) en zij f een C^2-functie in M. Dan is ω invariant onder de stroming van het Hamilton-stelsel met Hamilton-functie gelijk aan f.*

In Darboux-coördinaten, waar het Hamilton-veld v de gedaante (7.16) heeft, kan men dit verifiëren door na te rekenen dat de divergentie van v gelijk is aan nul:

$$\operatorname{div} v = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial \xi_i} + \sum_{j=1}^{n} \frac{\partial}{\partial \xi_j} \frac{-\partial f}{\partial x_j} = 0.$$

7.10 Geïnduceerde Kanonieke Transformaties

Is Φ een C^2-diffeomorfisme van de variëteit X naar de variëteit Y, dan induceert dit een C^1-diffeomorfisme $\widetilde{\Phi}$ van $\mathrm{T}^* X$ naar $\mathrm{T}^* Y$, dat is gedefinieerd door middel van

$$\widetilde{\Phi}\,(x,\, \eta \circ \mathrm{T}_x \Phi) = (\Phi(x),\, \eta)\,, \quad x \in X, \quad \eta \in \left(\mathrm{T}_{\Phi(x)}\, Y\right)^*. \tag{7.38}$$

Een speciaal geval hiervan vormen de geïnduceerde kaarten voor coraakbundels, zie (5.16).

Zij $\pi_X : \mathrm{T}^* X \to X$ en $\pi_Y : \mathrm{T}^* Y \to Y$ de projecties naar de bases. Men gaat gemakkelijk na dat

$$\pi_Y \circ \widetilde{\Phi} = \Phi \circ \pi_X \tag{7.39}$$

en dat

$$\widetilde{\Phi}^* \tau_{\mathrm{T}^* Y} = \tau_{\mathrm{T}^* X}. \tag{7.40}$$

De afbeelding $\widetilde{\Phi}$ is zelfs gekarakteriseerd door deze twee identiteiten: is $\widetilde{\Phi}$ een differentieerbare afbeelding van $\mathrm{T}^* X$ naar $\mathrm{T}^* Y$ die aan (7.39) en (7.40) voldoet, dan geldt (7.38).

Nemen we de uitwendige afgeleide van (7.40) en gebruiken we (5.43) met Φ vervangen door $\widetilde{\Phi}$, dan krijgen we dat

$$\widetilde{\Phi}^* \sigma_{\mathrm{T}^* Y} = \sigma_{\mathrm{T}^* X}, \tag{7.41}$$

ofwel $\widetilde{\Phi}$ is een kanonieke transformatie van $\mathrm{T}^* X$ naar $\mathrm{T}^* Y$.

Lang niet iedere kanonieke transformatie is geïnduceerd door een transformatie van de basis. Voor een willekeurige gladde functie f op $\mathrm{T}^* X$ bestaat de stroming van het Hamilton-vectorveld H_f uit kanonieke transformaties Φ^t, die echter in het algemeen niet de vezels van de coraakbundel in vezels van de coraakbundel overvoeren, dat wil zeggen, ze voldoen niet aan (7.39) met $\widetilde{\Phi}$ vervangen door Φ^t. Preciezer, de functies f op $\mathrm{T}^* X$, waarvoor de stroming van H_f geïnduceerd is door een stroming in de basis zijn gekarakteriseerd door de eigenschap dat voor iedere $x \in X$ de beperking van f tot de vezel $(\mathrm{T}_x X)^*$ een lineaire vorm op $(\mathrm{T}_x X)^*$, dus van de vorm $\xi \mapsto \langle v(x),\, \xi \rangle$ voor een $v(x) \in \mathrm{T}_x X$. We formuleren dit in de vorm van een stelling.

Stelling 7.11 *Zij v een C^2 vectorveld in de C^2-variëteit X, met stroming ϕ^t. Zij $\widetilde{\phi^t}$ de geïnduceerde kanonieke transformaties in $\mathrm{T}^* X$, gegeven door (7.40) met Φ vervangen door ϕ^t. Dan is $t \mapsto \widetilde{\phi^t}$ de stroming van het Hamilton-vectorveld H_I, met Hamilton-functie $I = I_v$ op $\mathrm{T}^* X$ gegeven door*

$$I_v(x,\, \xi) = \langle v(x),\, \xi \rangle, \quad x \in X, \quad \xi \in \left(\mathrm{T}_x\, X\right)^*, \tag{7.42}$$

de v-component van de impuls, beschouwd als functie op de coraakbundel.

Bewijs Er geldt (7.40) met Φ vervangen door ϕ^t. Differentiatie hiervan naar t in $t = 0$, gecombineerd met de homotopieformule (5.47), geeft dat

$$0 = \mathcal{L}_{\widetilde{v}}\tau = \mathrm{d}(\mathrm{i}_{\widetilde{v}}\,\tau) + \mathrm{i}_{\widetilde{v}}\,(\mathrm{d}\tau) = \mathrm{d}(\mathrm{i}_{\widetilde{v}}\,\tau) + \mathrm{i}_{\widetilde{v}}\,\sigma,$$

waarin $\widetilde{v}$ het snelheidveld van de éénparameter familie $t \mapsto \widetilde{\phi^t}$ van transformaties in T^*X voorstelt. Maar dit betekent met het oog op (7.15) dat $\widetilde{v} = \mathrm{H}_f$ met

$$f := \mathrm{i}_{\widetilde{v}}\,\tau,$$

hetgeen equivalent is met (7.42). □

Men noemt de *impulsfunctie* I_v ook wel de *v-component van de impuls*, omdat de $\xi \in (\mathrm{T}_x X)^*$ impulsvectoren heten. Bijvoorbeeld, als $X = \mathbf{R}^n$ en v is een constant vectorveld in $\mathbf{R}^n$, waarvan de stroming bestaat uit de translaties $x \mapsto x + t\,v$, $t \in \mathbf{R}$, dan is de bijbehorende impulsfunctie gegeven door $(x, \xi) \mapsto \langle v, \xi\rangle$, de '$v$-coördinaat van de impuls ξ'.

Ander voorbeeld: als $X = \mathbf{R}^3$ en $q \in \mathbf{R}^3$ is een gegeven vector, dan definieert

$$v(x) = q \times x,$$

het uitproduct of vectorproduct van q met x, een vectorveld in $\mathbf{R}^3$ waarvan de stroming bestaat uit draaiïngen om de oorsprong met $\mathbf{R}\,q$ als as en $\|q\|$ als draaisnelheid. De bijbehorende impulsfunctie is gegeven door

$$I(x, \xi) = \langle q \times x, \xi\rangle = \langle q, x \times \xi\rangle. \tag{7.43}$$

Hierin is ξ de impulsvector en $x \times \xi$ heet het *moment* om de oorsprong van de impuls ξ bij de positie x. Dus (7.43) is de 'q-coördinaat van het impulsmoment om de oorsprong'. Vanwege dit voorbeeld wordt de functie in (7.42) in de literatuur ook wel de *momentfunctie* van een willekeurig vectorveld v genoemd.

7.11 De Geodetische Stroming in de Coraakbundel

Is g een willekeurige pseudo-Riemann-structuur in X met Lagrange-functie T gegeven door

$$T(x, v) = \frac{1}{2}\,g_x(v, v),$$

dan is de bijbehorende snelheids-impulsafbeelding Φ de afbeelding van $\mathrm{T}\,X$ naar T^*X die aan $v \in \mathrm{T}_x X$ toevoegt het element van $(\mathrm{T}_x X)^*$ dat gelijk is aan

$$\text{de lineaire vorm} \quad w \mapsto g_x(v, w) \quad \text{op} \quad \mathrm{T}_x X. \tag{7.44}$$

De niet-gedegenereerdheid van g geeft dat Φ een C^k-diffeomorfisme is van $\mathrm{T}\,X$ naar T^*X. Met het oog op (7.44) ligt het voor de hand om $\Phi = g$ te schrijven. Verder is $H_T = T$, cf. (6.7), dus de Legendre-transformatie van T is de functie h op T^*X die is gegeven door

$$h(x, \xi) = h_g(x, \xi) := \frac{1}{2}\,{g_x}^{-1}(\xi, \xi), \quad x \in X,\ \xi \in (\mathrm{T}_x X)^*. \tag{7.45}$$

De snelheids-impulsafbeelding g voert de geodetische stroming in $\mathrm{T}\,X$ over in de stroming van het Hamilton-stelsel in T^*X met Hamilton-functie h, deze laatste wordt ook wel de *geodetische stroming in* T^*X genoemd.

Een equivalente formulering hiervan is: de snelheids-impulsafbeelding g trekt de kanonieke symplectische vorm σ van $\mathrm{T}^* X$ terug naar een symplectische vorm $g^*\sigma$ in $\mathrm{T}\, X$. (Deze hangt af van de keuze van de pseudo-Riemann-structuur g.) Dan is de geodetische stroming in $\mathrm{T}\, X$ gelijk aan de stroming van het Hamilton-vectorveld met betrekking tot $g^*\sigma$ en met Hamilton-functie gelijk aan T. Dit impliceert dat T een constante van beweging is (hetgeen we bij (6.7) al opgemerkt hadden), maar ook dat de geodetische stroming de symplectische vorm $g^*\sigma$ invariant laat en daarmee ook de volumevorm $\frac{\pm 1}{n!}\, (g^*\sigma)^n$ in $\mathrm{T}\, X$. In de differentiaalmeetkundige literatuur wordt opmerkelijk weinig aandacht besteed aan het Hamilton'se aspect van de geodetische stroming.

Het feit dat de exponentiële afbeelding een lokaal diffeomorfisme is, zie Lemma 6.3, impliceert op sterke wijze dat de geodetische stroming in $\mathrm{T}^* X$ de vezels niet in vezels overvoert, dus deze bestaat niet uit kanonieke transformaties die geïnduceerd zijn door transformaties in X. Dit hadden we ook kunnen concluderen uit het feit dat de Hamilton-functie (7.45) op iedere vezel een kwadratische functie is, dus geen lineaire functie, zie Stelling 7.11.

7.12 Het Beperkte Drielichamenprobleem

Het N-lichamenprobleem (met $N \geq 2$) heeft geen rustpunten, omdat ieder lichaam dat zich 'aan de buitenkant' bevindt een 'naar binnen' gerichte versnelling ondervindt, die niet gelijk aan nul is.

Dit kan als volgt formeel worden bewezen. Beschouw voor iedere i en j met $i \neq j$ het orthogonale complement $N_{i,\, j}$ in $\mathbf{R}^3$ van de vector $x^{(i)} - x^{(j)}$, dit is een twee-dimensionale lineaire deelruimte van $\mathbf{R}^3$. De *vereniging* N van de $N_{i,\, j}$ (niet de som!) over de eindige collectie van alle i en j met $i \neq j$ is een deelverzameling van $\mathbf{R}^3$ die niet gelijk is aan $\mathbf{R}^3$. Zij $\xi \in \mathbf{R}^3 \setminus N$ en zij i een rangnummer waarvoor $j \mapsto \langle x^{(j)},\, \xi \rangle$ maximaal is. Omdat voor iedere $j \neq i$ geldt dat $\xi \notin N_{i,\, j}$, betekent dit dat

$$\langle x^{(i)} - x^{(j)},\, \xi \rangle > 0, \quad j \neq i.$$

Uit de bewegingsvergelijking (1.29) lezen we af dat $d^2 x^{(i)}/dt^2$ een lineaire combinatie is van de vectoren $x^{(i)} - x^{(j)}$ met $j \neq i$, met strikt negatieve coëficiënten. Dit geeft dat

$$\langle \frac{d^2 x^{(i)}}{dt^2},\, \xi \rangle < 0.$$

In het bijzonder is de versnelling van het i-de deeltje niet gelijk aan nul.

Echter, in meedraaiende coördinatensystemen kunnen wél evenwichten optreden, zoals in het zogenaamde *beperkte drie-lichamenprobleem*, dat we nu zullen beschrijven. Om te beginnen nemen we $N = 3$ en merken op dat m_3 niet in de vergelijking (1.29) voor $i = 3$ voorkomt. Laten we in de vergelijkingen (1.29) voor $i = 1,\, 2$ de massa m_3 naar nul convergeren, dan convergeren deze vergelijkingen naar die voor het twee-lichamenprobleem, waarbij de versnellingen van de eerste twee lichamen niet meer afhangen van de positie van het derde lichaam. Van nu af aan wordt aangenomen dat $m_3 = 0$ in de bewegingsvergelijkingen (1.29).

Als in Vraagstuk 1.7 elimineren we de beweging van het zwaartepunt $z(t)$ door de substitutie $x^{(i)}(t) = z(t) + y^{(i)}(t)$. Dit geeft voor $i = 1,\, 2$ dat $y^{(i)}(t)$ voldoet aan (1.21) met $x(t)$, resp. c vervangen door $y^{(i)}(t)$, resp. c_i, waarin

$$c_1 = {m_2}^3\; (m_1 + m_2)^{-2}\; \mathcal{G}, \quad c_2 = {m_1}^3\; (m_1 + m_2)^{-2}\; \mathcal{G}.$$

De derde aanname is dat de eccentriciteit ε voor de oplossingen van het tweelichamenprobleem voor $y^{(1)}$ en $y^{(2)}$ gelijk is aan nul. Dit betekent dat $y^{(i)}$ met constante hoeksnelheid rondloopt op een cirkel met middelpunt in de oorsprong en straal gelijk aan r_i. Daarbij is $m_1\, y^{(1)} + m_2\, y^{(2)} = 0$, dus $m_1\, r_1 = m_2\, r_2$. Verder lezen we uit (1.27) af dat de periode gelijk is aan

$$T = 2\pi\, {c_i}^{-1/2}\, {r_i}^{3/2},$$

zodat

$$y^{(i)}(t) = \mathrm{e}^{\omega\, t\, J}\ y^{(i)}(0),$$

waarin J de matrix is van een kwartslag draaien in het vlak van de beweging en

$$\omega := \frac{2\pi}{T} = \left(\frac{m_2}{r_1}\right)^{3/2}\ (m_1 + m_2)^{-1}\ \mathcal{G}^{1/2} \tag{7.46}$$

de hoeksnelheid van de draaiïng is. Merk op dat hierin $m_2/r_1 = m_1/r_2$.

Met deze beweging van het eerste en tweede lichaam gegeven, wordt de positie $y = y(t)$ van het derde lichaam ten opzichte van het zwaartepunt beschreven door de bewegingsvergelijking

$$\frac{d^2y}{dt^2} = -\mathcal{G} \sum_{i=1}^{2} m_i\, \|y - y^{(i)}(t)\|\ \left(y - y^{(i)}(t)\right).$$

Op een orthonormale basis waarbij de eerste twee basisvectoren in het vlak van de eerste twee lichamen ligt, zien we dat $d^2y_3/dt^2 = 0$ als $y_3 = 0$, dus als $y_3(0) = 0$ en $y_3{}'(0) = 0$, dan geldt voor iedere t dat $y_3(t) = 0$. De vierde en laatste aanname voor het beperkte drielichamenprobleem is nu dat de beweging van het derde lichaam zich afspeelt in het vlak van de eerste twee lichamen. Bij gegeven beweging van de eerste twee lichamen krijgen we hiermee een tijdsafhankelijk (niet-autonoom) systeem met twee vrijheidsgraden.

Door middel van een paar herschalingen kan men dit herschrijven als een stelsel waarin maar één parameter voorkomt, de *massaverhouding* $\mu := m_1/\,(m_1 + m_2)$. Hierbij mogen we aannemen dat $m_1 \leq m_2$, dus $0 < \mu \leq 1/2$.

Om te beginnen past men een herschaling van de tijd toe, door middel van de substitutie $\omega\, t = s$. Schrijven we $y(t) = a(s) = a(\omega\, t)$, dan is $y''(t) = \omega^2\, a''(s)$ en we krijgen de bewegings- vergelijkingen

$$a'' = -\left(\frac{r_1}{m_2}\right)^3\ (m_1 + m_2)^2 \sum_{i=1}^{2} m_i\, \|a - a^{(i)}(s)\|\ \left(a - a^{(i)}(s)\right),$$

waarin nu

$$a^{(i)}(s) = \mathrm{e}^{s\, J}\ a^{(i)}(0), \quad a^{(i)}(0) = y^{(i)}(0).$$

Hiermee is zowel de hoeksnelheid van de draaiïng gelijk aan 1 gemaakt alsook de gravitatie-constante $\mathcal{G}$ uit de vergelijkingen verdwenen.

Vervolgens herschaalt men de afstand tot de oorsprong, door middel van de substitutie $a = (r_1 + r_2)\ b$. Schrijven we

$$a^{(i)}(s) = (r_1 + r_2)\ b^{(i)}(s),$$

dan wordt hiermee de afstand tussen $b^{(1)}(s)$ en $b^{(2)}(s)$ constant gelijk aan 1. De bewegingsvergelijkingen gaan hierbij over in

$$b'' = -\left(\frac{r_1}{r_1 + r_2}\right)^3\ \frac{(m_1 + m_2)^2}{{m_2}^3} \sum_{i=1}^{2} m_i\, \|b - b^{(i)}(s)\|\ \left(b - b^{(i)}(s)\right).$$

Opmerkend dat

$$\frac{r_1 + r_2}{r_1} = 1 + \frac{r_2}{r_1} = 1 + \frac{m_1}{m_2} = \frac{m_2 + m_1}{m_2} = \frac{1}{1 - \mu},$$

kunnen we de bewegingsvergelijkingen vereenvoudigen tot

$$b'' = -\mu \|b - b^{(1)}(s)\|^{-3} \left(b - b^{(1)}(s)\right) - (1 - \mu)\|b - b^{(2)}(s)\|^{-3} \left(b - b^{(2)}(s)\right)$$

Kiezen we tenslotte de eerste basisvector e_1 in de richting van $b^{(1)}(0)$, dan is

$$b^{(i)}(s) = \mathrm{e}^{s\,J}\, b^{(i)}(0),$$

met

$$b^{(1)}(0) = (1 - \mu)\, e_1, \quad b^{(2)}(0) = -\mu\, e_1.$$

Hiermee zijn de bewegingsvergelijkingen in de beloofde standaardvorm gebracht, met de massaverhouding μ als de enige overgebleven parameter.

De bewegingsvergelijkingen zijn Euler-Lagrange-vergelijkingen met Lagrange-functie gelijk aan

$$\frac{1}{2}\, \|b'\|^2 - V(s,\, b),$$

waarin

$$V(s,\, b) = -\frac{\mu}{\|b - b^{(1)}(s)\|} - \frac{1 - \mu}{\|b - b^{(2)}(s)\|}$$

een tijdsafhankelijke 'potentiële energie-functie' voorstelt. Om een autonoom stelsel te krijgen, gebruiken we een 'meedraaiend coördinatensysteem', dat wil zeggen we maken de tijdsafhankelijke substitutie van variabelen

$$b(s) = \mathrm{e}^{s\,J}\, x(s).$$

Dit geeft $b' = \mathrm{e}^{s\,J}\,(J\,x + x')$ en omdat $\mathrm{e}^{s\,J}$ de Euclidische norm behoudt krijgen we voor $x(s)$ de Euler-Lagrange-vergelijkingen met de tijdsonafhankelijke Lagrange-functie

$$L(x,\, x') = \frac{1}{2}\, \|J\,x + x'\|^2 - V(x), \tag{7.47}$$

$$V(x) = -\frac{\mu}{\|x - (1 - \mu)\, e_1\|} - \frac{1 - \mu}{\|x + \mu\, e_1\|}. \tag{7.48}$$

De impuls is in dit geval gelijk aan

$$\xi = \frac{\partial L}{\partial x'} = x' + J\,x.$$

Dit substituerend in

$$H(x,\, x') = \langle x',\, \xi \rangle - L(x,\, x')$$

krijgen we de Hamilton-functie

$$h(x,\, \xi) = \frac{1}{2}\, \|\xi\|^2 - \langle J\,x,\, \xi \rangle + V(x), \tag{7.49}$$

ten aanzien waarvan de bewegingsvergelijkingen in de termen van positie- en impuls- coördinaten een Hamilton-stelsel vormen. De term $V(x)$ in (7.49) is gedefinieerd in (7.48) en is gelijk aan de potentiële energie van een lichaam dat gravitationeel aangetrokken wordt door een lichaam met massa μ in $(1-\mu)\, e_1$ en een lichaam met massa $1-\mu$ in $-\mu\, e_1$. Vanwege de term $-\langle J\, x,\, \xi\rangle$ heeft de functie h niet de interpretatie van een totale energie. Echter, omdat de bewegingsvergelijkingen een Hamilton-stelsel vormen met Hamilton-functie gelijk aan h, is h wèl een constante van beweging.[1] Het stelsel heeft twee vrijheidsgraden, dus we krijgen een stroming in een vier-dimensionale faseruimte, die zich voor iedere constante c afspeelt in de drie-dimensionale analytische deelvariëteit waar h gelijk is aan c, althans als we ons beperken tot de open deelverzameling waar $\mathrm{d}\, h \neq 0$.

De evenwichtspunten van het stelsel voor $x(s)$ zijn precies de oplossingen van het oorspronkelijke stelsel, waarvoor het derde lichaam net als de eerste twee lichamen met constante hoeksnelheid gelijk aan ω rondloopt op een cirkel met middelpunt in de oorsprong. Anders gezegd, waarvoor het derde lichaam constante afstand houdt tot ieder van de eerste twee lichamen. Deze oplossingen corresponderen met de stationaire punten van de functie h; dat wil zeggen met

$$\xi - J\, x = 0, \quad J\, \xi + \frac{\partial V(x)}{\partial x} = 0,$$

ofwel $\xi = J\, x$ en, omdat $J^2 = -1$, $x = \frac{\partial V(x)}{\partial x}$. De vergelijking voor x is equivalent met het stelsel

$$\begin{aligned} x_1 &= \mu\, \|x - (1-\mu)\, e_1\|^{-3}\, (x_1 - 1 + \mu) + (1-\mu)\, \|x + \mu\, e_1\|^{-3}\, (x_1 + \mu), && (7.50)\\ x_2 &= \mu\, \|x - (1-\mu)\, e_1\|^{-3}\, x_2 + (1-\mu)\, \|x + \mu\, e_1\|^{-3}\, x_2. && (7.51) \end{aligned}$$

Dit stelsel heeft de volgende oplossingen.

a) $x_2 = 0$, waarmee is voldaan aan (7.51). Het rechterlid van (7.50) is een monotoon dalende functie van x_1 omdat de functie $y \mapsto |y|^{-3} y = y^{-2}\, \mathrm{sgn}\, y$ dat is. Verder is voor ieder van de drie intervallen $]-\infty, -\mu[$, $]-\mu, 1-\mu[$, $]1-\mu, \infty[$ in het linker, rep. rechter eindpunt het linkerlid in (7.50) kleiner, resp. groter dan het rechterlid in (7.50). Hieruit volgt dat, als $x_2 = 0$, de vergelijking (7.50) in ieder van de bovengenoemde intervallen precies één oplossing heeft.

b) $x_2 \neq 0$. In dit geval is (7.51) equivalent met

$$1 = \mu\, \|x - (1-\mu)\, e_1\|^{-3} + (1-\mu)\, \|x + \mu\, e_1\|^{-3}.$$

Dit geeft dat in (7.50) alle termen met de factor x_1 tegen elkaar wegvallen en we zien dat (7.50) equivalent is met $\|x - (1-\mu)\, e_1\| = \|x + \mu\, e_1\|$. Dit gebruikend is op zijn beurt (7.51) dan equivalent met $\|x + \mu\, e_1\| = 1$. Dit leidt tot twee oplossingen, waarbij x, $(1-\mu)\, e_1$ en $-\mu\, e_1$ een gelijkzijdige driehoek vormen met zijden gelijk aan 1.

De evenwichtspunten van het beperkte tweelichamenprobleem in meebewegende coördinaten in a) heten de *collineaire oplossingen* en werden voor het eerst gevonden door Euler [22]. Die in b) heten de *equilaterale oplossingen*, deze werden gevonden door Lagrange [51]. Deze vijf evenwichtsoplossingen werden echter vooral bekend door de bespreking ervan door Laplace [55, Livre X, Chap. VI].

[1] Dat h een constante van beweging is staat in een notitie [39] van Jacobi uit 1836; hij deed er verder niets mee en schreef in de inleiding dat dit tot de opmerkingen behoort die hun bestaansrecht niet ontlenen aan het feit dat ze zo diepzinnig zijn ... . Liouville [59] merkte op dat het constant zijn van h een speciaal geval is van de versie van Coriolis [14] van de wet van behoud van energie voor bewegingen in meedraaiende coördinaten.

Voor iedere $0 < \mu \leq 1/2$ zijn de collineaire oplossingen instabiel, in de zin dat er altijd eigenwaarden λ van de linearisatie zijn met positief reëel deel. Abraham and Marsden [3, p. 683] schrijven de stelling in deze algemeenheid toe aan Plummer [72]. Tot de voorgeschiedenis hiervan behoort dat Liouville [59] opmerkte dat als de aarde en de maan op een rechte lijn vanuit de zon zouden liggen op afstanden met een verhouding van ongeveer $1 : 1,01$, dan zou het altijd volle maan zijn. Hij legde vervolgens uit waarom het belangrijk is om de stabiliteit te bepalen van de linearisatie bij speciale oplossingen zoals rustpunten, en bewees tenslotte dat deze bij de hierboven beschreven configuratie van zon, aarde en maan instabiel is.

In contrast hiermee vond Routh [81] voor de equilaterale evenwichtspunten dat er een getal $0 < \mu_{\mathrm{Routh}} < 1/2$ is, met de eigenschap dat alle eigenwaarden van de linearisatie zuiver imaginair zijn, dan en slechts dan als $0 < \mu \leq \mu_{\mathrm{Routh}}$. Het getal μ_{Routh} heet sedertdien de *kritieke massaverhouding van Routh.* Voor het systeem waarbij het eerste lichaam gelijk is aan Jupiter en tweede lichaam gelijk is aan de zon, is aan deze voorwaarde voldaan. Zoekend bij de equilaterale oplossingen vond men een groep van planetoïden, de *Trojanen* genaamd, waarvan de eerste, Achilles genaamd, in 1906 werd ontdekt.

Het beperkte drielichamenprobleem werd door Poincaré, in zijn meesterwerk [75] over de klassieke mechanica, bij iedere methode die hij ontwikkelde als voorbeeld gebruikt om de methode mee te illustreren. In het bijzonder had hij in [74, §22] gevonden dat er geen constante van beweging is, onafhankelijk van de functie h, die een analytische functie is van (x, ξ) en μ, voor μ in een omgeving van $\mu = 0$. Dit resultaat, dat geldig is voor 'de meeste problemen' omdat het op een algemeen geldige redenering berust, ging volledig in tegen de optimistische verwachtingen die men tot dan toe had over de oplosbaarheid van de bewegingsvergelijkingen in de klassieke mechanica.

Het beperkte drielichamenprobleem kwam opnieuw in de belangstelling toen het mogelijk werd om kunstmanen te lanceren; daarvoor is het interessant om oplossingen te vinden die 'uit zichzelf', dus zonder veel met raketkracht bij te sturen, dicht bij het gewenste hemellichaam komen. Ik heb mijn historische verwijzingen voor het beperkte drielichamenprobleem uit van der Meer [64, Ch. VI]. Zie ook Abraham and Marsden [3, Sec. 10.2]

7.13 Andere Vectorvelden in de Coraakbundel

Wanneer de bewegingen nog een extra krachtveld $\phi(x, v) \in (\mathrm{T}_x X)^*$ als storing ondervinden, dan dienen de bewegingsvergelijkingen $[L] = 0$ te worden vervangen door $[L] = \phi$. Schrijven we $\psi(x, \xi) = \phi(x, v)$ als $\xi = \frac{\partial L(x, v)}{\partial v}$, dan kunnen we ψ als een vectorveld in $\mathrm{T}^* X$ opvatten met de speciale eigenschap dat ψ in ieder punt (x, ξ) raakt aan de vezel, ofwel door $\mathrm{T}_{(x, \xi)}\, \pi$ naar nul wordt afgebeeld. Het snelheidsveld van de beweging in $\mathrm{T}^* X$ is dan gelijk aan $\mathrm{H}_h + \psi$, de som van het Hamilton-vectorveld met Hamilton-functie gelijk aan de Legendre-transformatie h van L en het raakvecorveld ψ aan de vezels dat de storing vertegenwoordigt. In lokale coördinaten in $\mathrm{T}^* X$ afkomstig van een kaart voor X zien de bewegingsvergelijkingen eruit als:

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial f(x, \xi)}{\partial \xi_i}, \\ \frac{d\xi_i}{dt} &= -\frac{\partial f(x, \xi)}{\partial x_i} + \psi(x, \xi), \end{aligned} \tag{7.52}$$

vergelijk dit met (7.16).

Een interessant voorbeeld van een vectorveld in $\mathrm{T}^* X$ dat raakt aan de vezels is het *Euler--vectorveld* E, dat in ieder punt (x, ξ) gelijk is aan ξ, beschouwd als raakvector aan $(\mathrm{T}_x X)^*$. In lokale coördinaten afkomstig van een kaart voor X is $\mathrm{E}(x,\,\xi) = (0,\,\xi)$. De naam is afkomstig van het feit dat voor een differentieerbare functie f in $\mathrm{T}^*\,X$ geldt dat $\xi \mapsto f(x,\,\xi)$ homogeen van de graad m is, dan en slechts dan als T voldoet aan de differentiaalvergelijking $\mathrm{i}_{\mathrm{E}}\,\mathrm{d}f = m\,f$ van Euler.

Het Euler-vectorveld voldoet aan de vergelijking

$$\mathrm{i}_{\mathrm{E}}\,\sigma = \tau. \tag{7.53}$$

Men kan (7.53) ook opvatten als de definitie van E, omdat $v \mapsto \mathrm{i}_v\,\sigma$ een bijectieve afbeelding is van de ruimte van vectorvelden naar de ruimte van éénvormen. Een direct gevolg van (7.53) is dat $\mathcal{L}_{\mathrm{E}}\sigma = \sigma$, dus

$$\left(\mathrm{e}^{t\,\mathrm{E}}\right)^*\sigma = \mathrm{e}^t\,\sigma.$$

7.14 Bewijs van het Lemma van Darboux

Als eerste stap bewijzen we dat een symplectische vorm in een lineaire ruimte E, die opgevat kan worden als een tweevorm in $\mathbf{R}^{2n}$ met constante coëfficiënten, een *lineaire* Darboux-kaart heeft.

We beginnen daartoe met de keuze van een Lagrange-vlak L, gevolgd door de keuze van een basis $e_1, \ldots, e_n$ in L. Vervolgens kiezen we met inductie over k vectoren ϵ_k, waarvoor

$$\begin{aligned} \sigma\left(\epsilon_k,\, e_j\right) &= 0 \qquad \text{als} \qquad j \neq k, \\ \sigma\left(\epsilon_k,\, e_k\right) &= 1, \\ \sigma\left(\epsilon_k,\, \epsilon_j\right) &= 0 \quad \text{als} 1 \leq j \leq k-1. \end{aligned} \tag{7.54}$$

Deze vergelijkingen impliceren dat de vectoren $e_1, \ldots, e_n, \epsilon_1, \ldots, \epsilon_k$ lineair onafhankelijk zijn. Deze uitspraak met k vervangen door $k-1$ geeft dat de $\sigma\,e_i$ en $\sigma\,\epsilon_j$, met $1 \leq i \leq n$ en $1 \leq j \leq k-1$, lineair onafhankelijk zijn in E^*. Dus de afbeelding, die aan $\epsilon_k \in E$ de rij van linkerleden in (7.54) toevoegt, is surjectief naar $\mathbf{R}^{n+k-1}$; dit maakt dat (7.54) een oplossing $\epsilon_k \in E$ heeft. Dit proces stopt voor $k = n$, waarna we een basis $e_1, \ldots, e_n, \epsilon_1, \ldots, \epsilon_n$ in E hebben verkregen. Noteren we de coördinaten ten aanzien van deze basis met $x_1, \ldots, x_n, \xi_1, \ldots, \xi_n$, dan krijgt σ ten aanzien van deze coördinaten de vorm (7.13).

Voor de volgende stap mogen we aannemen dat het punt $m \in M$, in de omgeving waarvan we de Darboux-coördinaten willen construeren, gelijk is aan de oorsprong in een gegeven coördinatensysteem. Verder kunnen we vanwege het voorgaande een lineaire substitutie van variabelen toepassen die er voor zorgt dat σ_0 de gedaante (7.13) heeft. Hierin wordt σ_0 beschouwd als een symplectische vorm in $\mathbf{R}^n \times \mathbf{R}^n$ met constante coëfficiënten, terwijl σ een symplectische vorm is, gedefinieerd in een open omgeving U van de oorsprong, met variabele coëffiënten, van de klasse C^{r-1}, waarbij de coëfficiënten in de oorsprong gelijk zijn aan die van σ_0.

Schrijf $\sigma_{(t)} = \sigma_0 + t\,(\sigma - \sigma_0)$, dus $\sigma_{(0)} = \sigma_0$ en $\sigma_{(1)} = \sigma$, dit is het 'rechte lijnstuk van σ_0 naar σ in de ruimte van tweevormen'. De volgende stap is dat we proberen om een homotopie ϕ_t als boven te vinden, waarvoor ϕ_0 gelijk is aan de identiteit in U en voor iedere $0 \leq t \leq 1$ geldt dat

$$\alpha_t := {\phi_t}^*\,\sigma_{(t)} = \sigma_0. \tag{7.55}$$

Hierin $t = 1$ invullend, zien we dat de substitutie van variabelen ϕ_1 de symplectische vorm σ_1 terugtrekt naar σ_0, waarmee het lemma van Darboux zou zijn bewezen.

Omdat ϕ_0 gelijk is aan de identiteit is $\alpha_0 = \sigma_0$; daarmee wordt het voor alle $0 \leq t \leq 1$ gelden van (7.55) equivalent met het voor alle $0 \leq t \leq 1$ gelijk aan nul zijn van

$$\frac{d}{dt}\alpha_t = \left(\frac{d}{dt}{\phi_t}^*\right)\,\sigma_{(t)} + {\phi_t}^*\,\frac{d}{dt}\,\sigma_{(t)} = {\phi_t}^*\left(\mathrm{d}\circ\mathrm{i}_{v_t}\,\sigma_{(t)} + \sigma - \sigma_0\right),$$

ofwel

$$\mathrm{d} \circ \mathrm{i}_{v_t}\, \sigma_{(t)} + \sigma - \sigma_0 = 0, \quad 0 \leq t \leq 1. \tag{7.56}$$

Hierin hebben we (5.49) gebruikt, met ω gelijk aan de gesloten differentiaalvorm $\sigma_{(t)}$. Verder is v_t het t-afhankelijke vectorveld waarvoor

$$\frac{d}{dt}\phi_t(x) = v_t\left(\phi_t(x)\right), \tag{7.57}$$

dat wil zeggen, waarvoor geldt dat

$$\iota_t{}^* \circ \mathrm{i}_{\partial/\partial t} \circ \Phi^* = \mathrm{i}_{v_t}\,.$$

Omdat σ en σ_0 gesloten zijn is $\sigma - \sigma_0$ gesloten. Door U samentrekbaar te nemen kunnen we garanderen dat er een éénvorm ν is, van de klasse C^{r-1}, met $\sigma - \sigma_0 = \mathrm{d}\nu$ in U. Hiermee krijgen we dat aan (7.56) is voldaan, als

$$\mathrm{i}_{v_t}\, \sigma_{(t)} = -\nu, \quad 0 \leq t \leq 1. \tag{7.58}$$

Nu is $\sigma_{(t)}$ in de oorsprong gelijk aan σ_0, dus door U in te perken tot een voldoend kleine bol om de oorsprong kunnen we er bovendien voor zorgen dat voor iedere $0 \leq t \leq 1$ de tweevorm $\sigma_{(t)}$ niet-gedegenereerd is (een symplectische vorm) in U. Dit impliceert dat (7.58) precies één oplossing v_t heeft, een tijdsafhankelijk vectorveld van de klasse C^{r-1}. Als ϕ_t de 0-t-stroming is met v_t als snelheidsveld, dan krijgen we (7.57) en $\phi_0(x) = x$, dus de inleiding van het bewijs teruglezend krijgen we dat $\phi_1{}^*\sigma = \sigma_0$.

Dit completeert het bewijs van het lemma van Darboux. Dit bewijs is afkomstig van Weinstein [88, p. 333, 334]. Het principe om een object in een gewenste gedaante te krijgen door middel van een homotopie, waarvan het snelheidsveld verkregen wordt als oplossing van een lineaire algebraïsche vergelijking, is bij mijn weten voor het eerst zo expliciet naar voren gebracht door Moser [65]. Deze methode is sindsdien in zeer veel situaties heel efficiënt gebleken. (Zo kan men er ook een redelijk kort bewijs mee geven van het lemma van Morse.)

7.15 Vraagstukken

Vraagstuk 7.1 Zij $L = L(x,\, v) = T(x,\, v) - V(x)$, met $T(x,\, v) = \frac{1}{2}\, m(x)(v,\, v)$ en, voor iedere $x \in X$, $m(x)$ een inproduct in $\mathrm{T}_x\, X$. Aangenomen wordt dat $L \in \mathrm{C}^2$. Bewijs dat

$$h(x,\, p) = \frac{1}{2}\, m(x)^{-1}(p,\, p) + V(x), \quad p \in \left(\mathrm{T}_x\, X\right)^*,$$

waarin we $m(x)$ identificeren met de lineaire afbeelding $v \mapsto p$ en

$$m(x)^{-1}(p,\, p) = p\left(m(x)^{-1}(p)\right).$$

Bewijs dat $(x,\, p)$ een evenwichtspunt is, dan en slechts dan als $p = 0$ en $\mathrm{d}V(x) = 0$. Schrijf in dit geval $Q = \mathrm{D}^2\, V(x)$. Bewijs dat er een orthonormale basis e_j, $1 \leq j \leq n$, in $\mathrm{T}_x\, X$ is met betrekking tot het inproduct $m(x)$, waarop Q diagonaalgedaante heeft, met reële getallen Q_j, $1 \leq j \leq n$, op de diagonaal. Bewijs dat, in lokale coördinaten waarvoor de e_j corresponderen met de standaard basisvectoren, de linearisering van het Hamilton-stelsel in het evenwichtspunt gegeven wordt door

$\frac{dx_j}{dt} = p_j$, $\frac{dp_j}{dt} = -Q_j\, x_j$, $1 \le j \le n$. Bewijs dat de eigenwaarden van de linearisering van het Hamilton-veld in het evenwichtspunt óf reëel óf zuiver imaginair zijn, dus dat het na Lemma 7.4 genoemde geval d) hier niet op kan treden. Preciezer: $\lambda_j = \pm\sqrt{-Q_j}$ als $Q_j < 0$ en $\lambda_j = \pm i\,\sqrt{Q_j}$ als $Q_j > 0$. ⊘

Vraagstuk 7.2 Zij A een inverteerbare lineaire transformatie van $\mathbf{R}^n$ naar $\mathbf{R}^n$. Bewijs dat de geïnduceerde kanonieke transformatie $\widetilde{A}$ in $\mathrm{T}^*\,\mathbf{R}^n = \mathbf{R}^n \times \mathbf{R}^n$ gegeven wordt door

$$\widetilde{A} : (x,\, \xi) \mapsto \big(A(x),\, (A^*)^{-1}(\xi)\big)\,.$$

Onderzoek de eigenwaarden van $\widetilde{A}$, in termen van de eigenwaarden van A. Geef voor een willekeurige lineaire transformatie

$$L : (x,\, \xi) \mapsto (A(x) + B(\xi),\, C(x) + D(\xi))$$

de vergelijkingen waar de $n \times n$-matrices A, B, C, D aan moeten voldoen, opdat L een kanonieke transformatie is. (Het geval $L = \widetilde{A}$ correspondeert met $B = 0$ en $C = 0$.) Vergelijk in het geval $n = 1$ de verzameling van de $\widetilde{A}$ met de verzameling van alle lineaire kanonieke transformaties. ⊘

Vraagstuk 7.3 Bewijs dat voor de Lagrange-punten van het beperkte drielichamenprobleem, de eigenwaardevergelijking de volgende gedaante heeft:

$$\lambda^4 + \lambda^2 + \frac{27}{4}\,\mu\,(1 - \mu) = 0.$$

Bereken de kritieke massaverhouding van Routh en bewijs dat de eigenwaarden in het (instabiele) geval dat $\mu_{\text{Routh}} < \mu \le 1/2$ de configuratie d) in Lemma 7.4 vertonen. ⊘

Vraagstuk 7.4 De volgende constructie is de basis van de zogenaamde *methode van de laatste factor* van Jacobi [42, p. 71-85]. Zij $1 \le k \le \omega$ en zij f een C^k-functie in een d-dimensionale C^k-variëteit M. Neem aan dat voor iedere $m \in M$ geldt dat $\mathrm{d}f(m) \neq 0$. Dit impliceert dat, voor iedere $c \in \mathbf{R}$, $M_c := \{m \in M \mid f(m) = c\}$ een gesloten en $(d-1)$-dimensionale C^k-deelvariëteit is van M. Verder is ω een C^{k-1}-volumevorm in M, dat wil zeggen een C^{k-1}-differentiaalvorm van de graad d. Bewijs achtereenvolgens:

a) Bij iedere $m \in M$ is er een open omgeving U van m in M en een C^{k-1}-vectorveld v in U, waarvoor $\mathrm{i}_v\, \mathrm{d}f = 1$ in U.

b) $\omega = (\mathrm{d}f) \wedge (\mathrm{i}_v\, \omega)$ in U.

c) Zij $\iota_c : M_c \to M$ de identiteit in M_c, beschouwd als afbeelding van M_c naar M. Als θ een $(d-1)$-vorm is in U waarvoor $\omega = \mathrm{d}f \wedge \theta$ in U, dan is ${\iota_c}^*\theta = {\iota_c}^*\,(\mathrm{i}_v\, \omega)$ in $M_c \cap U$.

d) Er is een éénduidig bepaalde C^{k-1}-differentiaalvorm ν van de graad $d-1$ in M_c (een volumevorm in M_c), met de eigenschap dat als U open is in M, θ een $(d-1)$-vorm in U en $\omega = \mathrm{d}f \wedge \theta$ in U, dan is $\nu = {\iota_c}^*\theta$ in $M_c \cap U$. Men schrijft soms ook wel $\nu = \omega/\, \mathrm{d}f$ in M_c, al is deze notatie een beetje riskant.

e) Is $k \ge 2$, u een C^1-vectorveld in M, $\mathcal{L}_u\omega = 0$ en $\mathcal{L}_u f = 0$, dan raakt u aan M_c en is $\mathcal{L}_u\nu = 0$.

f) Is σ een symplectische vorm in M en ω de corresponderende kanonieke volumevorm, dan is de volumevorm $\omega / \mathrm{d}f$ in M_c invariant onder de stroming van het Hamilton-veld H_f in M_c.

$\oslash$

Vraagstuk 7.5 Zij E het Euler-vectorveld, dat is gedefinieerd door $\mathrm{E}(x, \xi) = (0, \xi)$. Zij $v = \mathrm{H}_f - c\,\mathrm{E}$, waarin c een positieve constante voorstelt en $-c\,\mathrm{E}$ geïnterpreteerd wordt als een wrijvingsterm. Bewijs:

a) $\left(\mathrm{e}^{t\,v}\right)^* \sigma = \mathrm{e}^{-c\,t}\,\sigma$. Verder, is ω de kanonieke volumevorm, dan is $\left(\mathrm{e}^{t\,v}\right)^* \omega = \mathrm{e}^{-n\,c\,t}\,\omega$. Anders gezegd, de stroming in de coraakbundel is *volumeverkleinend*, met een exponentiële factor $\mathrm{e}^{-n\,c\,t}$.

b) Is $f(x, \xi) = T(x, \xi) + V(x)$, waarin de kinetische energie $T(x, \xi)$ als functie van ξ een positief definiete kwadratische vorm is, dan geldt voor iedere oplossing $m(t) = (x(t), \xi(t))$ van het stelsel $\mathrm{d}m/\,\mathrm{d}t = v(m)$ dat $\frac{d}{dt} f(m(t)) = -2c\,T(m(t))$. Bewijs dat $t \mapsto f(m(t))$ strikt monotoon dalend is, tenzij $t \mapsto m(t)$ constant is (een evenwichtsoplossing).

c) Als m_0 een evenwichtspunt is waar $\mathrm{D}^2 f\,(m_0)$ positief definiet is, dan is dit evenwichtspunt *asymptotisch Lyapunov-stabiel*, in de volgende zin. Ten eerste is het evenwichtspunt Lyapunov-stabiel. En ten tweede is er een open omgeving U van m_0 in $\mathrm{T}^*\,X$ met de eigenschap dat als $m(t)$ een oplossing is met $m(0) \in U$, dan convergeert $m(t)$ naar m_0 als $t \to \infty$.

$\oslash$

Vraagstuk 7.6 Beschouw voor het beperkte drielichamenprobleem de functie

$$W(x) := -\frac{1}{2}\,\|x\|^2 + V(x),$$

met $V(x)$ als in (7.48). Bewijs dat $W(x) \leq h$, waarin h de constante van beweging is, gedefinieerd in (7.49). Dit is de *ongelijkheid van Hill* [34], die hij als volgt gebruikte. Zij $X_{\leq h}$ de verzameling der $x \in \mathbf{R}^2$ waarvoor $W(x) \leq h$ en zij C een samenhangscomponent van $X_{\leq h}$. Is $t \mapsto (x(t), \xi(t))$ een oplossing en $x(0) \in C$, dan geldt voor iedere t dat $x(t) \in C$.

Bewijs dat als h voldoende klein (groot negatief) is, dan heeft $X_{\leq h}$ drie (disjuncte) samenhangscomponenten, één daarvan is een compacte omgeving van $(1-\mu)\,e_1$, de tweede is een compacte omgeving van $-\mu\,e_1$ en de derde is het complement in $\mathbf{R}^2$ van een begrensde open deelverzameling van $\mathbf{R}^2$. Men kan bewijzen dat dit optreedt zodra h kleiner is dan de kleinste kritieke waarde van W.

Project: maak, met behulp van een computer, plaatjes van $X_{\leq h}$ voor diverse waarden van h. Bekijk daarbij in het bijzonder wat er gebeurt als h een kritieke waarde van W passeert. (De opmerking, dat de topologie van $X_{\leq h}$ alleen verandert als h een kritieke waarde van de functie W passeert, is één van de basisideeën van de zogenaamde *Morse theorie*.) $\oslash$

Vraagstuk 7.7 Zij $2 \leq r \leq \omega$ en zij M een C^r-*oppervlak*, dat wil zeggen een tweedimensionale C^r-variëteit. Een *oppervlaktevorm* in M is een C^r-tweevorm σ, met de eigenschap dat voor iedere $m \in M$ geldt dat $\sigma_m \neq 0$. Bekend is dat M oriëntieerbaar is, dan en slechts dan als er een oppervlaktevorm in M is. Bewijs:

a) σ is een oppervlaktevorm in M $\Leftrightarrow$ σ is een symplectische vorm in M.

b) Is σ een oppervlaktevorm in M en f een reëelwaardige C^r-functie in M, dan is de stroming van H_f oppervlaktebewarend en laat de niveaukrommen van de functie f invariant. Is C een compacte samenhangscomponent in M^{reg} van een niveaukromme, dan is de oplossing van het Hamilton-stelsel die in C start periodiek en is C diffeomorf met een cirkel.

c) Neem aan dat we in de situatie van b) zijn en dat M compact is. Dan zijn alle oplossingen, waar f niet gelijk is aan een kritieke waarde van f, periodiek.

d) Neem aan dat we in de situatie van c) zijn. Dan is M niet diffeomorf met een open deelverzameling van een coraakbundel. Welke voorbeelden kent U van compacte oriënteerbare oppervlakken? Neem Uw favoriete voorbeeld van een oppervlak in $\mathbf{R}^3$ en beschrijf de oplossingen van het Hamilton-stelsel voor de hoogtefunctie, de functie die aan $x \in \mathbf{R}^3$ de derde coördinaat x_3 toevoegt.

$\oslash$

Vraagstuk 7.8 Beschouw een mechanisch systeem met n vrijheidsgraden waarvan de kinetische energie gelijk is aan $\frac{1}{2}\,\|x'\|^2$ en de potentiële energie tijdsafhankelijk is, maar van de vorm $V(t,\, x) = W\left(\mathrm{e}^{-t\,A}\, x\right)$, waarin A een antisymmetrische matrix is. Substitueer $x(t) = \mathrm{e}^{t\,A}\, y(t)$. Bewijs de *stelling van Coriolis*, dat de grootheid

$$H = H(y,\, y') := \frac{1}{2}\|y'\|^2 - \frac{1}{2}\|A\,y\|^2 + W(y)$$

een constante van beweging is. Bewijs dat

$$-\frac{1}{2}\|A\,y\|^2 + W(y) \leq H,$$

als generalisatie van de ongelijkheid van Hill. Bewijs: als $\mathrm{d}\,W(0) = 0$ en de symmetrische matrix $A^2 + \mathrm{D}^2\,W(0)$ is positief definiet, dan is $y = y' = 0$ een stabiel rustpunt. Waarschuwing: $\|A\,y\|^2 = -\langle A^2\,y,\, y\rangle$, dus de symmetrische matrix A^2 is *negatief* semidefiniet. Dit maakt dat het in het algemeen niet voldoende is dat $\mathrm{D}^2\,W(0)$ positief definiet is. $\oslash$

Vraagstuk 7.9 Beschouw het Hamilton-systeem met Hamilton-functie gelijk aan de totale energie

$$h(x,\, \xi) := \frac{1}{2m}\,\|\xi\|^2 + V(x), \quad x,\, \xi \in \mathbf{R}^3.$$

Hierin is V een C^2 functie op $\mathbf{R}^3$. Bewijs dat de stelling van Noether tot de volgende conclusies leidt.

a) Als $v \in \mathbf{R}^3$ en V is invariant onder de translaties $x \mapsto x + t\,v$, $t \in \mathbf{R}$, dan is de v-component van de impuls een constante van beweging.

b) Als $w \in \mathbf{R}^3$ en V is invariant onder alle draaiïngen om de as $\mathbf{R}\,w$, dan is de w-component van het impulsmoment een constante van beweging.

$\oslash$

Vraagstuk 7.10 Zij g een C^2 pseudo-Riemann-structuur op een variëteit X. Men noemt een vectorveld k in X een *Killing-vectorveld* voor g, als de stroming ϕ^t van k uit isometriën van $(X,\, g)$ bestaat. Dat wil zeggen $T \circ \mathrm{T}\,\phi^t = T$, als $T(x,\, v) = g_x(v,\, v)/2$, $x \in X$, $v \in \mathrm{T}_x\, X$.

Bewijs: als k een Killing-vectorveld is voor g, dan is de functie $I = I_k$ op $\mathrm{T}\,X$, gedefinieerd door

$$I(x,\, v) = g_x(v,\, k(x)), \quad x \in X, \quad v \in \mathrm{T}_x\, X,$$

een constante van beweging voor de geodetische stroming in $\mathrm{T}\,X$. $\oslash$

Vraagstuk 7.11 Beschouw, voor een vectorveld v op een variëteit X, de afbeelding $\mathcal{L}_v : \omega \mapsto \mathcal{L}_v\omega$ als een lineaire operator in de ruimte van differentiaalvormen op X van een willekeurige graad. Bewijs dat $\mathcal{L}_{[u,\, v]} = [\mathcal{L}_u,\, \mathcal{L}_v]$. $\oslash$

Vraagstuk 7.12 Zij L een symplectische lineaire transformatie in de symplectische lineaire ruimte $(E,\, \sigma)$. Bewijs dat dit betekent dat $\sigma \circ L \circ \sigma^{-1}$ gelijk is aan de inverse van L^*. Bewijs dat hieruit volgt dat er voor de eigenwaarden μ van L de volgende mogelijkheden zijn.

a) $\mu = \pm 1$, ieder met even algebraïsche multipliciteit. (Hint: gebruik dat $\det L = 1$ omdat L de georiënteerde volumevorm σ^n invariant laat.)

b) Een reële eigenwaarde μ met $|\mu| > 1$, waarbij $1/\mu$ ook een eigenwaarde van L is met dezelfde algebraïsche multipliciteit.

c) Een niet-reële eigenwaarde μ met $|\mu| = 1$, waarbij $\overline{\mu} = 1/\mu$ ook een eigenwaarde is met dezelfde algebraïsche multipliciteit.

d) Vier complexe eigenwaarden μ, $1/\mu$, $\overline{\mu}$, $1/\overline{\mu}$, niet reëel en met absolute waarde niet gelijk aan 1, alle vier met dezelfde algebraïsche multipliciteit.

Merk op dat d) niet op kan treden als $\dim E = 2$.

Laat zien dat het bovenstaande een gevolg is van Lemma 7.4, indien $L = \mathrm{e}^A$, waarin A een infinitesimaal-symplectische transformatie in E is.

Zij tenslotte v een Hamilton-vectorveld in de symplectische variëteit M. Neem aan dat er een speciaal punt $m \in M$ is en een $\omega > 0$, waarvoor $\mathrm{e}^{\omega\, v}(m) = m$. (Dit betekent dat de oplossingskromme $t \mapsto \mathrm{e}^{t\, v}(m)$ periodiek is met periode ω.) Bewijs dat het bovenstaande van toepassing is op de lineaire transformatie $L = \mathrm{T}_m\, (\mathrm{e}^{\omega\, v})$ in $\mathrm{T}_m\, M$. $\oslash$

Vraagstuk 7.13 In de terminolgie van Stelling 7.11, bewijs dat

$$I_{[u,\, v]} = \{I_u,\, I_v\}\,.$$

Leid de Jacobi-identiteit voor Lie-haakjes van vectorvelden af uit de Jacobi-identeit voor Poisson-haakjes van functies. $\oslash$

8 Hamilton-Jacobi-theorie

In dit hoofdstuk laten we zien hoe een algemene eerste-orde partiële differentiaalvergelijking opgelost kan worden door gebruik te maken van de stroming van een Hamilton-stelsel.

8.1 Lagrange-variëteiten

Om geen boekhouding van differentieerbaarheidsgraden hoeven bij te houden, werken we in dit hoofdstuk in de gladde = C^∞ categorie. Deze inperking is overdreven sterk, maar leidt tot een aangename vereenvoudiging in de notaties.

Zij X een gladde n-dimensionale variëteit en zij g een gladde differentiaalvorm in X, dat wil zeggen een gladde afbeelding van X naar $\mathrm{T}^* X$ waarvoor $\pi \circ g$ gelijk aan de identiteit in X. Hieruit lezen we af dat g een gladde inbedding is van X naar $\mathrm{T}^* X$ en het beeld $g(X)$ een n-dimensionale gesloten en gladde deelvariëteit is van $\mathrm{T}^* X$. Verder is, voor iedere $x \in X$,

$$\mathrm{T}_{(x,\, g(x))}\, (g(X)) = \mathrm{T}_x\, g\, (\mathrm{T}_x\, X)\,. \tag{8.1}$$

Nemen we de uitwendige afgeleide in de formule (7.11) dan krijgen we dat

$$\mathrm{d}g = \mathrm{d}(g^*\tau) = g^*\,(\mathrm{d}\tau) = g^*\sigma. \tag{8.2}$$

Hieruit lezen we af dat de differentiaalvorm g gesloten is, dat wil zeggen dat $\mathrm{d}g = 0$, dan en slechts dan als $g^*\sigma = 0$, hetgeen equivalent is met de voorwaarde dat voor iedere $\lambda \in \Lambda := g(X)$ geldt:

$$\sigma_\lambda(u,\, v) = 0, \quad u,\, v \in \mathrm{T}_\lambda\, \Lambda. \tag{8.3}$$

In de terminologie van Hoofdstuk 7 betekent (8.3) dat $\mathrm{T}_\lambda\, \Lambda$ een isotrope lineaire deelruimte is van de symplectische lineaire ruimte $\mathrm{T}_\lambda\, (\mathrm{T}^* X)$, ten aanzien van de symplectische vorm σ_λ. Verder is $\mathrm{T}_\lambda\, \Lambda$ n-dimensionaal, dus zelfs een Lagrange-vlak in $\mathrm{T}_\lambda\, (\mathrm{T}^* X)$.

In het algemeen heet een deelvariëteit Λ (van willekeurige dimensie $d \leq n$) van een $2n$--dimensionale symplectische variëteit $(M,\, \sigma)$ *isotroop* als voor iedere $\lambda \in \Lambda$ geldt dat $\mathrm{T}_\lambda\, \Lambda$ een isotrope lineaire deelruimte is van $\mathrm{T}_\lambda\, M$. Is bovendien Λ n-dimensionaal, dan heet Λ een *Lagrange-variëteit* in M.

Opmerking 8.1 De naam 'Lagrange-variëteit' is geïntroduceerd door Maslov [63, p. 115]. Zijn motivering was dat de symplectische vorm, in termen waarvan de isotropie is geformuleerd, al bij Lagrange [52, t. I] voorkwam. De daarmee direct samenhangende naam 'Lagrange-vlak' voor de raakruimte aan een Lagrange-variëteit is afkomstig van Arnol'd [6]. Deze namen zijn intussen gemeengoed geworden.

Hoewel Lagrange artikelen heeft geschreven over eerste-orde partiële differentiaalvergelijkingen, heeft hij bij mijn weten het begrip van n-dimensionale isotrope deelvarëteiten van $\mathrm{T}^* X$ niet gebruikt. Deze werden wèl expliciet ingevoerd door Lie [57, §30], maar zonder er een speciale naam aan te geven. Lie vermeldde daarbij dat Engel de naam 'Verein' invoerde voor een (niet noodzakelijk isotrope) deelvariëteit van $\mathrm{T}^* X$, ter onderscheiding van een willekeurige variëteit die in het duits een 'Mannigfaltigkeit' heet. ⊘

We krijgen nu de volgende lokale karakterisering van 'grafieken $\mathrm{d}\phi(X)$ van afgeleiden van functies ϕ'.

Lemma 8.1 *Zij Λ een gladde deelvariëteit van $\mathrm{T}^* X$ Dan zijn de volgende uitspraken equivalent.*

a) *Bij iedere $\lambda \in \Lambda$ is er een open omgeving U van λ in $\mathrm{T}^* X$, een open omgeving V van $x = \pi(\lambda)$ in X en een reëelwaardige en gladde functie ϕ in V, waarvoor $\Lambda \cap U = \Lambda_\phi := (\mathrm{d}\phi)(V)$.*

b) *Λ is een Lagrange-variëteit in $\mathrm{T}^* X$ en voor iedere $\lambda \in \Lambda$ is de doorsnede van $\mathrm{T}_\lambda \Lambda$ met de raakruimte in λ aan de vezel gelijk aan nul.*

Bewijs We mogen aannemen dat $\dim \Lambda = n$, want dit volgt zowel uit a) alsook uit b).

We onderzoeken eerst de voorwaarde dat er bij iedere $\lambda \in \Lambda$ een open omgeving U van λ in $\mathrm{T}^* X$ is, een open omgeving van $x = \pi(\lambda)$ in X en een gladde differentiaalvorm g in V, waarvoor $\Lambda \cap U = g(V)$. Is dit het geval dan lezen we uit het feit dat $\pi \circ g$ gelijk is aan de identiteit in X af dat $\mathrm{T}_\lambda \pi \circ \mathrm{T}_x g$ gelijk is aan de identiteit in $\mathrm{T}_x X$, hetgeen impliceert dat

$$\mathrm{T}_\lambda (\pi|_\Lambda) = \mathrm{T}_\lambda \pi|_{\mathrm{T}_\lambda \Lambda}$$

een injectieve, dus bijectieve lineare afbeelding is van $\mathrm{T}_\lambda \Lambda$ naar $\mathrm{T}_x X$. Op zijn beurt is de injectiviteit van deze lineaire afbeelding equivalent met de voorwaarde dat de doorsnede van $\mathrm{T}_\lambda \Lambda$ met $\ker \mathrm{T}_\lambda \pi$ gelijk is aan nul, waarbij de laatste lineaire ruimte gelijk is aan de raakruimte in λ aan de vezel.

Is omgekeerd deze doorsnede gelijk aan nul dan is $\mathrm{T}_\lambda (\pi|_\Lambda)$ bijectief, dus geeft toepassing van de lokale inverse-functiestelling op $\pi|_\Lambda$ dat er een open omgeving U van λ in $\mathrm{T}^* X$ is en een open omgeving U van x in X, waarvoor $\pi|_{\Lambda \cap U}$ een glad diffeomorfisme is van $\Lambda \cap U$ naar V. Voor de inverse g hiervan geldt dan dat g een gladde afbeelding is van V naar $\mathrm{T}^* X$, $\pi \circ g$ is gelijk aan de identiteit in V en $g(V) = \Lambda \cap U$.

Bij (8.2) hebben we gezien dat Λ een Lagrange-variëteit is, dan en slechts dan als $\mathrm{d}g = 0$. Verder is bij (5.17) besproken dat de voorwaarde $\mathrm{d}g = 0$ lokaal equivalent is met de existentie van een differentieerbare functie ϕ met $g = \mathrm{d}\phi$, deze functie ϕ is op een additieve constante na eenduidig bepaald en automatisch glad. $\square$

Opmerking 8.2 Zij b') de voorwaarde b) met daaraan toegevoegd de voorwaarde toe dat voor iedere $x \in X$ de variëteit Λ precies één punt gemeen heeft met de vezel $(\mathrm{T}_x X)^*$. Deze toevoeging levert dat $\pi|_\Lambda$ een glad diffeomorfisme van Λ naar X is en de conclusie is dat b') equivalent is met de voorwaarde dat $\Lambda = g(X)$ voor een gesloten en gladde differentiaal-vorm g in X. In Opmerking 7.2 is besproken dat iedere gesloten g gelijk is aan $\mathrm{d}\phi$ voor een differentieerbare functie ϕ, dan en slechts dan als $\mathrm{H}^1(X, \mathbf{R}) = 0$. Dit leidt tot de volgende globale versie van Lemma 8.1.

Zij $\mathrm{H}^1(X, \mathbf{R}) = 0$. Dan zijn de volgende uitspraken voor een gladde deelvariëteit Λ van $\mathrm{T}^* X$ equivalent.

a') Er is een reëelwaardige gladde functie ϕ in X waarvoor $\Lambda = \Lambda_\phi := (\mathrm{d}\phi)(X)$.

b') Λ is een Lagrange-variëteit in $\mathrm{T}^* X$, die iedere vezel in precies één punt en dwars snijdt.

In toepassingen gebeurt het vaak dat de variëteit Λ die we construeren wel een Lagrange-variëteit is, maar één die niet steeds voldoet aan de eis dat de vezels dwars, resp. in één punt, gesneden worden. $\oslash$

8.2 Een Eerste-orde Partiële Differentiaalvergelijking

Zij f een reëelwaardige gladde functie, gedefinieerd in een open deelverzameling P van $\mathrm{T}^* X$. We onderzoeken nu een eerste-orde partiële differentiaalvergelijking van de vorm

$$f(x,\, \mathrm{d}\phi(x)) = 0. \tag{8.4}$$

Een *oplossing van* (8.4) is per definitie een reëelwaardige differentieerbare functie ϕ, gedefinieerd in een open deelverzameling U van X, met de eigenschap dat voor iedere $x \in U$ geldt dat $(x,\, \mathrm{d}\phi(x)) \in P$ en dat (8.4) geldt.

Lemma 8.2 *Zij Λ een C^1 Lagrange-variëteit in een symplectische variëteit $(P,\, \sigma)$ en zij $f|_\Lambda = 0$. Dan geldt voor iedere $\lambda \in \Lambda$ dat $\mathrm{H}_f(\lambda) \in \mathrm{T}_\lambda \Lambda$. In het bijzonder is Λ invariant onder de stroming van het Hamilton-veld met Hamilton-functie gelijk aan f. Is ϕ een C^2 oplossing van (8.4), dan is de grafiek van $\mathrm{d}\phi$ invariant onder de H_f-stroming.*

Bewijs Uit $f|_\Lambda = 0$ volgt dat voor iedere $\lambda \in \Lambda$ geldt dat $\mathrm{T}_\lambda \Lambda \subset \ker \mathrm{d}f(\lambda)$. Dit impliceert dat

$$\mathrm{H}_f(\lambda) \in (\ker \mathrm{d}f(\lambda))^\sigma \subset (\mathrm{T}_\lambda \Lambda)^\sigma = \mathrm{T}_\lambda \Lambda.$$

Hierin betekent de bovenindex σ het orthogonale complement met betrekking tot de symplectische vorm $\sigma = \sigma_\lambda$ in $\mathrm{T}_\lambda P$ en de laatste identiteit geeft weer dat $\mathrm{T}_\lambda \Lambda$ een Lagrange-vlak in $\mathrm{T}_\lambda P$ is.

De tweede conclusie volgt uit het algemene principe dat een deelvariëteit V invariant is onder de stroming met snelheidsveld v, dan en slechts dan als het vectorveld v raakt aan V, zoals besproken tegen het eind van Hoofdstuk 2.

De laatste conclusie volgt door het voorgaande toe te passen met $\Lambda = \Lambda_\phi$. $\square$

De laatste conclusie in Lemma 8.2 maakt dat Λ_ϕ eenduidig vastligt in de richting van de H_f-oplossingskrommen. Maar in de $(n-1)$-dimensionale richting dwars daarop is Λ_ϕ nog vrij, met dien verstande dat het bevat moet zijn in de verzameling waar $f = 0$ en dat het isotroop moet zijn. De relatie tussen de Lagrange-variëteit in $f^{-1}(\{0\})$ en zijn isotrope $(n-1)$-dimensionale deelvariëteit die dwars ligt ten opzichte van H_f wordt beschreven in het volgende lemma.

Lemma 8.3 *Zij $(P,\, \sigma)$ een gladde, $2n$-dimensionale symplectische variëteit en zij f een reëelwaardige en gladde functie in P. Zij Λ_0 een $(n-1)$-dimensionale isotrope en gladde deelvariëteit van P met de eigenschap dat de beperking van f tot Λ_0 gelijk is aan nul en dat voor iedere $\lambda \in \Lambda_0$ geldt dat*

$$\mathrm{H}_f(\lambda) \notin \mathrm{T}_\lambda \Lambda_0. \tag{8.5}$$

Zij $\Phi^t : D(t) \to P$ de stroming na tijd t van het Hamilton-vectorveld met Hamilton-functie gelijk aan f. We noteren C voor de verzameling der $(\lambda,\, t) \in \Lambda_0 \times \mathbf{R}$ waarvoor $\lambda \in \Lambda_0 \cap D(t)$ en voor iedere $(\lambda,\, t) \in C$ schrijven we $\Phi(\lambda,\, t) = \Phi^t(\lambda)$. Dan is Φ een gladde immersie van C naar P, $f \circ \Phi = 0$ en $\Phi^*\sigma = 0$. Is B een open deelverzameling van C waarvoor $\Phi|_B$ een inbedding is, dan is $\Lambda := \Phi(B)$ een gladde Lagrange-variëteit in P met de eigenschappen dat $\Lambda_0 \subset \Lambda$ en $f|_\Lambda = 0$.

Bewijs Er geldt voor iedere t dat $\left(\Phi^t\right)^* f = f$, dus ook $\left(\Phi^t\right)^* \mathrm{d}f = \mathrm{d}f$. Omdat ook $\left(\Phi^t\right)^* \sigma = \sigma$, krijgen we dat $\left(\Phi^t\right)^* \mathrm{H}_f = \mathrm{H}_f$, hetgeen uitgeschreven betekent dat voor iedere $p \in P$ geldt dat

$\mathrm{T}_\Phi^t\left(\mathrm{H}_f(p)\right)=\mathrm{H}_f\left(\Phi^t(p)\right)$. Dit toepassend voor $p=\lambda\in\Lambda_0$ krijgen we voor iedere $\delta\lambda\in\mathrm{T}_\lambda\,\Lambda_0$ en $\delta t\in\mathbf{R}$ dat

$$\mathrm{T}_{(\lambda,\,t)}\,\Phi(\delta\lambda,\,\delta t)=\mathrm{T}_\lambda\,\Phi^t(\delta\lambda)+\delta t\ \mathrm{H}_f\left(\Phi^t(\lambda)\right)=\mathrm{T}_\lambda\,\Phi^t\left[\delta\lambda+\delta t\ \mathrm{H}_f(\lambda)\right]. \tag{8.6}$$

Omdat $\mathrm{T}_\lambda\,\Phi^t$ injectief is, is (8.6) gelijk aan nul alleen als $\delta\lambda+\delta t\ \mathrm{H}_f(\lambda)=0$, hetgeen vanwege (8.5) alleen kan gebeuren als $\delta\lambda=0$ en $\delta t=0$. Hiermee is bewezen dat Φ een immersie is.

Omdat $\left(\Phi^t\right)^*\sigma=\sigma$ zien we uit (8.6) ook dat als $v_i=(\delta\lambda_i,\,\delta t_i)$, $i=1,2$, dan is

$$\begin{aligned}(\Phi^*\sigma)_\lambda\,(v_1,\,v_2) &= \sigma_\lambda\,(\delta\lambda_1+\delta t_1\ \mathrm{H}_f(\lambda),\,\delta\lambda_2+\delta t_2\ \mathrm{H}_f(\lambda))\\ &= \sigma_\lambda\,(\delta\lambda_1,\,\delta\lambda_2)+t_2\,\sigma_\lambda\,(\delta\lambda_1,\,\mathrm{H}_f(\lambda))-t_1\,\sigma_\lambda\,(\delta\lambda_2,\,\mathrm{H}_f(\lambda)),\end{aligned} \tag{8.7}$$

omdat het symplectische product van $\mathrm{H}_f(\lambda)$ met zichzelf gelijk aan nul is. De isotropie van Λ_0 geeft dat de eerste term in (8.7) gelijk aan nul is. Anderzijds geeft $f=0$ in Λ_0 dat $\mathrm{d}f(\lambda)=0$ in $\mathrm{T}_\lambda\,\Lambda_0$, dus voor iedere $\delta\lambda\in\mathrm{T}_\lambda\,\Lambda_0$ geldt dat

$$\sigma_\lambda\,(\delta\lambda,\,\mathrm{H}_f(\lambda))=\mathrm{d}f(\lambda)(\delta\lambda)=0.$$

Hiermee is bewezen dat $\Phi^*\sigma=0$.

Tenslotte geeft $\left(\Phi^t\right)^*f=f$ en $f=0$ in Λ_0 dat voor iedere $(\lambda,\,t)\in C$ geldt dat

$$f(\Phi(\lambda,\,t))=f\left(\Phi^t(\lambda)\right)=f(\lambda)=0.$$

Is $\Phi|_B$ een inbedding, dan is $\Lambda:=\Phi(B)$ een gladde, n-dimensionale deelvariëteit van P met

$$\Lambda_0=\Phi\left(\Lambda_0\times\{0\}\right)\subset\Lambda.$$

Verder is Λ isotroop, dus een Lagrangevariëteit, omdat $\Phi^*\sigma=0$. Tenslotte is Λ bevat in $f^{-1}\{0\})$ omdat $f\circ\Phi=0$. □

De vrijheid van keuze van de isotrope deelvariëteit Λ_0 kan gebruikt worden om de oplossing ϕ willekeurig voor te schrijven op een $(n-1)$-dimensionale deelvariëteit S van X, analoog aan de beginvoorwaarde bij een eerste-orde gewone differentiaalvergelijking. Ter voorbereiding daarvan maken we een paar opmerkingen over de relatie tussen de coraakbundel van S en de coraakbundel van X.

Lemma 8.4 *Zij S een gladde deelvariëteit van X, definieer $\mathrm{T}^*_S\,X$ als de verzameling der $p\in\mathrm{T}^*\,X$ met $\pi(p)\in S$, dat wil zeggen de verzameling der $(x,\,\xi)$ met $x\in S$ en $\xi\in(\mathrm{T}_x\,X)^*$. Zij $\iota:\mathrm{T}^*_S\,X\to\mathrm{T}^*\,X$ de beperking tot $\mathrm{T}^*_S\,X$ van de identiteit in $\mathrm{T}^*\,X$ en zij $\rho:\mathrm{T}^*_S\,X\to\mathrm{T}^*\,S$ de afbeelding die, voor iedere $x\in S$, aan de lineaire vorm ξ op $\mathrm{T}_x\,X$ toevoegt zijn beperking tot de lineaire deelruimte $\mathrm{T}_x\,S$ van $\mathrm{T}_x\,X$. Dan geldt dat*

$$\iota^*\sigma_{\mathrm{T}^*\,X}=\rho^*\sigma_{\mathrm{T}^*\,S}. \tag{8.8}$$

Is Σ een gladde Lagrange-variëteit in $\mathrm{T}^*\,S$, dan is $\rho^{-1}(\Sigma)$ een gladde Lagrange-variëteit in $\mathrm{T}^*\,X$.

Bewijs Het is duidelijk dat voor de tautologische éénvormen in $\mathrm{T}^*\,X$ en $\mathrm{T}^*\,S$ de identiteit

$$\iota^*\tau_{\mathrm{T}^*\,X}=\rho^*\tau_{\mathrm{T}^*\,S}$$

geldt. Door hiervan de uitwendige afgeleide te nemen en te gebruiken dat de uitwendige afgeleide commuteert met het terugtrekken door middel van differentieerbare afbeeldingen, krijgen we (8.8).

Voor de tweede conclusie beginnen we met op te merken dat ρ een gladde submersie is van $\mathrm{T}_S^* X$ naar $\mathrm{T}^* S$ en dat de dimensie van de vezel van ρ gelijk is aan de codimensie $\dim X - \dim S$ van S in X. Hieruit volgt dat $\Sigma' := \rho^{-1}(\Sigma)$ een gladde deelvariëteit is van $\mathrm{T}_S^* X$, dus van $\mathrm{T}^* X$ en dat de dimensie van $\rho^{-1}(\Sigma)$ gelijk is aan $\dim \Sigma + (\dim X - \dim S) = \dim S + (\dim X - \dim S) = \dim X$. Verder volgt uit $\rho(\Sigma') \subset \Sigma$ dat, voor iedere $p \in \Sigma'$, de ruimte $\mathrm{T}_p\, \rho\, (\mathrm{T}_p\, \Sigma')$ bevat is in de isotrope lineaire deelruimte $\mathrm{T}_{\rho(p)}\, \Sigma$ van $\mathrm{T}^* S$. Vanwege (8.8) krijgen we daarom dat voor iedere $u,\, v \in \mathrm{T}_p\, \Sigma'$ geldt dat

$$\sigma_p(u,\, v) = \sigma_{\rho(p)}\, (\mathrm{T}_p\, \rho(u),\, \mathrm{T}_p\, \rho(v)) = 0,$$

dus Σ' is isotroop. □

Stelling 8.5 *Zij f een gladde reëelwaardige functie gedefinieerd in een open deelverzameling P van $\mathrm{T}^* X$. Zij $(x_0,\, \xi_0) \in P$ zodat $f\,(x_0,\, \xi_0) = 0$. Neem aan dat S een gladde, $(n-1)$-dimensionale deelvariëteit is van X met $x_0 \in S$ en dat*

$$\frac{\partial f\,(x_0,\, \xi)}{\partial \xi}|_{\xi=\xi_0} \notin \mathrm{T}_{x_0}\, S. \tag{8.9}$$

Zij ψ een gladde reëelwaardige functie op S met $\mathrm{d}\psi\,(x_0)$ gelijk aan de beperking van ξ_0 tot $\mathrm{T}_{x_0}\, S$. Dan is er een open omgeving U van x_0 in X met de eigenschap dat er in U precies één gladde oplossing ϕ van (8.4) is waarvan de beperking tot $S \cap U$ gelijk is aan de beperking tot $S \cap U$ van ψ en waarvoor $\mathrm{d}\phi\,(x_0) = \xi_0$.

Bewijs Toepassing van Lemma 8.4 met $\Sigma = \Lambda_\psi$ geeft dat $\Sigma' := \rho^{-1}\,(\Lambda_\psi)$ een gladde Lagrange--variëteit in $\mathrm{T}^* X$ is, met $p_0 := (x_0,\, \xi_0) \in \Sigma'$. We noteren met $(\mathrm{T}_{x_0}\, S)^\perp$ voor de verzameling der lineaire vormen ξ op $\mathrm{T}_{x_0}\, X$ waarvoor $\xi(v) = 0$ voor iedere $v \in \mathrm{T}_{x_0}\, S$. Dan betekent de voorwaarde (8.9) dat de beperking tot $(\mathrm{T}_{x_0}\, S)^\perp$ van de afgeleide in ξ_0 van $\xi \mapsto f\,(x_0,\, \xi)$ niet gelijk aan nul is. Vanwege de impliciete functiestelling geeft dit dat er een open omgeving Ω, resp. S_0 van p_0, resp. x_0 in $\mathrm{T}^* X$, resp. S is, met de eigenschap dat er voor iedere $x \in S_0$ precies één $\xi = \xi(s) \in (\mathrm{T}_s\, X)^*$ is, waarvoor $(x,\, \xi(s)) \in \Omega$,

$$\xi|_{\mathrm{T}_s\, S} = \mathrm{d}\psi(s) \tag{8.10}$$

en

$$f(s,\, \xi) = 0. \tag{8.11}$$

(De eenduidigheid hierin impliceert dat $\xi\,(x_0) = \xi_0$.) Verder is $\xi(s)$ een gladde functie van s, hetgeen impliceert dat $s \mapsto (s,\, \xi(s))$ een gladde inbedding is van S_0 naar een gladde, $(n-1)$-dimensionale deelvariëteit Λ_0 van $\mathrm{T}^* X$. Hierin betekent (8.10) dat $(s,\, \xi) \in \Sigma'$, dus (8.10) en (8.11) samen betekenen dat in Ω de doorsnede van Σ' met $f^{-1}(\{0\})$ gelijk is aan de gladde, $(n-1)$-dimensionale deelvariëteit Λ_0 van $\mathrm{T}^* X$. Vanwege Lemma 8.4 is Σ' isotroop, dus de deelvariëteit Λ_0 van Σ' is ook isotroop.

De voorwaarde (8.9) kan ook gelezen worden als

$$\mathrm{T}_{p_0}\, \pi\, (\mathrm{H}_f\, (p_0)) \notin \mathrm{T}_{x_0}\, S = \mathrm{T}_{p_0}\, \pi\, (\mathrm{T}_{p_0}\, \mathrm{T}_S^* X)\,,$$

hetgeen impliceert dat $\mathrm{H}_f(p_0) \notin \mathrm{T}_{p_0}\mathrm{T}^*_S X$, waaruit op zijn beurt volgt dat $\mathrm{H}_f(p_0)$ niet bevat is in de lineaire deelruimte $\mathrm{T}_{p_0}\Lambda_0$ van $\mathrm{T}_{p_0}(\mathrm{T}^*_S X)$. Door Λ_0 te vervangen door de doorsnede van Λ_0 met een voldoend kleine omgeving van p_0 in $\mathrm{T}^* X$, krijgen we dat voor iedere $\lambda \in \Lambda_0$ geldt dat $\mathrm{H}_f(\lambda) \notin \mathrm{T}_\lambda \Lambda_0$.

Is nu Φ als in Lemma 8.3 met Λ_0 vervangen door Λ_0 en is B gelijk aan een voldoend kleine open omgeving van $(p_0, 0)$ in C, dan is de beperking van ϕ tot B een inbedding en is $\Lambda = \Phi(B)$ een gladde Lagrange-variëteit, bevat in de nulpuntsverzameling van f en waarvoor Λ een open omgeving van p_0 in Λ_0 bevat. De raakruimte in p_0 aan Λ is gelijk aan de som van $\mathrm{T}_{x_0}\Lambda_0$ en $\mathbf{R}\cdot \mathrm{H}_f(p_0)$, deze wordt door $\mathrm{T}_{p_0}\pi$ afgebeeld naar de som van $\mathrm{T}_{x_0} S$ en de scalaire veelvouden van het linkerlid in (8.9), dus $\mathrm{T}_{p_0}\pi$ is surjectief van $\mathrm{T}_{p_0}\Lambda$ naar $\mathrm{T}_{x_0} X$. Toepassing van Lemma 8.1 geeft nu dat, als we B voldoende klein nemen, er een open omgeving U van x_0 in X is en een gladde functie $\phi : U \to \mathbf{R}$, waarvoor $\Lambda = \Lambda_\phi$.

De eigenschap dat $f = 0$ op Λ betekent hierbij dat ϕ een oplossing is van (8.4). Anderzijds betekent $\Lambda_0 \subset \Lambda$ en $\rho(\Lambda_0) \subset \Lambda_\psi$ dat voor iedere $s \in S \cap U$ geldt dat

$$\mathrm{d}\psi(s) = \mathrm{d}\phi(s)|_{\mathrm{T}_s S} = \mathrm{d}(\phi|_S)(s).$$

Vervang S door de samenhangscomponent van x_0 in $S \cap U$. Dan impliceert het bovenstaande dat er een constante c is met de eigenschap dat voor iedere $s \in S$ geldt dat $\phi(s) = \psi(s) + c$. Nu verandert $\mathrm{d}\phi$, dus Λ_ϕ, niet als we bij ϕ een constante optellen, er is daarom precies één keuze van ϕ, waarvoor $\phi = \psi$ in S. $\square$

Opmerking 8.3 Als $t \mapsto (x(t), \xi(t))$ een oplossingskromme is van het Hamilton-stelsel waarvoor $x(0) \in S$, dan is

$$\frac{\mathrm{d}\phi(x(t))}{\mathrm{d}t} = (\mathrm{d}\phi)(x(t)) \cdot \frac{\mathrm{d}x(t)}{\mathrm{d}t} = \xi(t) \cdot \left.\frac{\partial f(x(t), \xi)}{\partial \xi}\right|_{\xi=\xi(t)}.$$

In combinatie met met $\phi(x(0)) = \psi(x(0))$ leidt dit tot de formule

$$\phi(x(T)) = \psi(x(0)) + \int_0^T \xi(t) \cdot \left.\frac{\partial f(x(t), \xi)}{\partial \xi}\right|_{\xi=\xi(t)} \mathrm{d}t$$

voor de oplossing ϕ van de partiële differentiaalvergelijking $f(x, \mathrm{d}\phi(x)) = 0$, in termen van de beginfunctie ψ op S. Stelling 8.5 zegt dat lokaal tot een eenduidig bepaalde gladde oplossing ϕ van het beginwaardenprobleem leidt.

Globaal kunnen er allerlei problemen optreden. Bijvoorbeels kan het gebeuren dat er een punt x is waarvoor er twee verschillende oplossingen $(x(t), \xi(t))$ en $(\widetilde{x}(t), \widetilde{\xi}(t))$ van het Hamilton-stelsel zijn, en bijbehorende tijdstippen T en $\widetilde{T}$, waarvoor $x(0) \in S$, $\widetilde{x}(0) \in S$, $x(T) = x = \widetilde{x}(\widetilde{T})$, terwijl de bovenstaande uitdrukkingen voor $\phi(x(T))$ en $\phi(\widetilde{x}(\widetilde{T}))$ verschillend zijn. In dit geval hebben we twee verschillende, concurrerende waarden voor $\phi(x)$, ofwel is er geen eenduidigheid van de oplossing. Wij gaan niet verder in op de interessante vraag hoe men in toepassingen soms toch kan komen tot een keuze die gedicteerd wordt door extra eisen die in de toepassing gesteld worden, dit kan dan leiden tot oplossingen $\phi(x)$ die geen gladde functies van x zijn. $\oslash$

De methode van het oplossen van (8.4), door gebruik te maken van de oplossingen van het Hamilton-stelsel met Hamilton-functie gelijk aan f, staat in de literatuur bekend als de *Hamilton-Jacobi-theorie*.

Stelling 8.5 is een *lokale* existentie- en eenduidigheidsstelling voor oplossingen van een eerste--orde partiële differentiaalvergelijking van de vorm (8.4). Obstakels voor het bestaan van globale oplossingen zijn, in termen van het H_f-uitstroomsel Λ van Λ_0:

i) Het kan gebeuren dat Λ de grafiek is van een gesloten differentiaalvorm g in X, maar dat er geen differentieerbare functie ϕ in X is met $g = \mathrm{d}\phi$. In dit geval definieert g een element van $\mathrm{H}^1(X, \mathbf{R})$ dat niet gelijk aan nul is. Dit kan dus niet optreden als $\mathrm{H}^1(X, \mathbf{R}) = 0$, dus in het bijzonder niet als X enkelvoudig samenhangend is. Zie Opmerking 7.2.

ii) $\pi : \Lambda \to X$ zou overal bijectieve raakafbeelding kunnen hebben, zonder bijectief te zijn. Is Λ samenhangend en is π een zogenaamde overdekking, dan kan dit niet optreden als X enkelvoudig samenhangend is.

iii) Er zijn $\lambda \in \Lambda$ waarvoor $\mathrm{T}_\lambda \pi$ niet bijectief is van $\mathrm{T}_\lambda \Lambda$ naar $\mathrm{T}_x X$, $x = \pi(\lambda)$. Dit zijn de punten waar de raakruimte aan Λ 'verticaal gedraaid' is, dat wil zeggen een doorsnede van positieve dimensie heeft met de raakruimte $\ker \mathrm{T}_\lambda \pi$ aan de vezel $(\mathrm{T}_x X)^*$ over x van de coraakbundel. Hiervoor is er geen enkele open omgeving Λ' van λ in Λ waarvoor er een ϕ is met $\Lambda' = \Lambda_\phi$. Deze situatie correspondeert met het veel voorkomende verschijnsel dat de π-beelden van de H_f-oplossingskrommen in Λ in X een zogenaamde *brandfiguur* gaan vormen.

Opmerking 8.4 Tot de gegevens in Stelling 8.5 behoren niet alleen de functie ψ in S welke met de oplossing ϕ in S moet overeenstemmen, maar ook nog de $\xi_0 \in (\mathrm{T}_{x_0} X)^*$ met

$$f(x_0, \xi_0) = 0 \quad \text{en} \quad \xi_0|_{\mathrm{T}_{x_0} S} = \mathrm{d}\psi(x_0), \tag{8.12}$$

waarvoor de afgeleide van de oplossing ϕ in het punt x_0 gelijk moet zijn aan ξ_0. In het algemeen is, bij gegeven ψ en x_0, de ξ_0 die aan (8.12) voldoet niet eenduidig bepaald. Bijvoorbeeld, als $X = \mathbf{R}^n$, $S = \mathbf{R}^{n-1} \times \{0\}$ en $f(x, \xi) = c^2 \|\xi\|^2 - 1$, dan kan alleen aan de voorwaarden van de stelling worden voldaan als $\|\mathrm{d}\psi(x_0)\| < 1/c$ en in dit geval zijn er twee keuzen van ξ_0, namelijk

$$\xi_0^\pm = \left(\mathrm{d}\psi(x_0), \pm\left[c^{-2} - \|\mathrm{d}\psi(x_0)\|^2\right]^{1/2}\right). \tag{8.13}$$

Voor ieder van de keuzen $\xi_0 = \xi_0^\pm$ in (8.13) krijgen we een lokale oplossing $\phi^\pm$ van (8.4) met $\phi^\pm|_S = \psi$. Merk op dat $\phi^+ = \phi^-$ in S, maar dat $\mathrm{d}\phi^+(x_0) \neq \mathrm{d}\phi^-(x_0)$, dus buiten S stemmen de oplossingen zeker niet overeen. $\oslash$

8.3 Hoogfrequente Golven

De meeste partiële differentiaalvergelijkingen uit de mathematische fysica, zoals de Laplace-vergelijking, de golfvergelijking of de Maxwell-vergelijkingen, kunnen *niet* opgelost worden door reductie tot stelsels gewone differentiaalvergelijkingen.

Toch kan de Hamilton-Jacobi-theorie ook daaraan een bijdrage leveren, zij het op een wat indirecte manier. Neem bijvoorbeeld de *golfvergelijking*

$$\Box u := \frac{\partial^2 u}{\partial t^2} - \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2} = 0. \tag{8.14}$$

We proberen oplossingen van de vorm

$$u(x,\, t) = e^{i\omega\,(t-\phi(x))}\, a(x),$$

waarin $\phi(x)$ en $a(x)$ nader te bepalen reëelwaardige functies van x zijn en $\omega \in \mathbf{R}$. De exponentiële factor is een harmonische trilling als functie van t, met ω als *frequentie*. De *fase* van de trilling is constant gelijk aan c langs het oppervlak $\phi(x) = t - c$. De factor $a(x) \neq 0$ is een *amplitude*-factor. We proberen dit type oplossingen voor hoge frequentie, dat wil zeggen voor grote waarden van ω.

We krijgen nu

$$\begin{aligned}\Box u = \; & e^{i\omega\,(t-\phi(x))} \\ & \cdot[\omega^2 \cdot (\sum_{j=1}^{n}(\partial_j\phi(x))^2 - 1)\cdot a(x) \\ & + i\omega\left(2\sum_{j=1}^{n}\partial_j\phi(x)\cdot\partial_j a(x) + \sum_{j=1}^{n}\partial_j^2\phi(x)\cdot a(x)\right) \\ & - \sum_{j=1}^{n}\partial_j^2 a(x)].\end{aligned}$$

Dit groeit zelfs kwadratisch als functie van ω voor $\omega \to \infty$, tenzij de fasefunctie voldoet aan de niet-lineaire eerste-orde partiële differentiaalvergelijking

$$\sum_{j=1}^{n}(\partial_j\phi(x))^2 - 1 = 0, \tag{8.15}$$

welke vergelijking in de geometrische optica ook wel de *eikonaalvergelijking* genoemd wordt.

Hebben we (8.15) opgelost met behulp van de Hamilton-Jacobi-theorie en daarmee $\phi(x)$ vastgelegd, dan kunnen we vervolgens de lineaire term in ω gelijk aan nul krijgen, door de amplitude-factor $a(x)$ te laten voldoen aan de lineaire partiële differentiaalvergelijking

$$2\sum_{j=1}^{n}\partial_j\phi(x)\cdot\partial_j a(x) + \sum_{j=1}^{n}\partial_j^2\phi(x)\cdot a(x) = 0. \tag{8.16}$$

Om de aard van deze vergelijking te doorzien, beschouwen we de oplossingen $x(s)$ van het stelsel gewone differentiaalvergelijkingen

$$\frac{dx_j}{ds} = 2\partial_j\phi(x), \;\; 1 \leq j \leq n. \tag{8.17}$$

De oplossingskrommen hiervan staan loodrecht op de 'golffronten' $\phi(x) = \text{constant}$. Met de notatie

$$c(x) = \sum_{j=1}^{n}\partial_j^2\phi(x)$$

is (8.16) dan equivalent met

$$\frac{d}{ds}a(x(s)) + c(x(s))\cdot a(x(s)) = 0,$$

een lineaire eerste-orde gewone differentiaalvergelijking voor de functie $s \mapsto a(x(s))$. Dit betekent dat we $a(x)$ vrij kunnen voorschrijven in een $(n-1)$-dimensionale variëteit die de oplossingen $s \mapsto x(s)$ dwars snijdt. Kiezen we de drager daarvan in een kleine omgeving van een gegeven punt x_0, dan is de drager van $a(x)$ geconcentreerd in een dunne buis om de oplossing $x(s)$ met $x(0) = x_0$. Men noemt (8.16) de *transportvergelijking* voor de amplitude $a(x)$ langs de oplossingen $x(s)$ van (8.17).

Met deze methode worden geen oplossingen u gevonden van de vergelijking $\Box u = 0$, alleen hoogfrequente oplossingen van de asymptotische vergelijking '$\Box u$ blijft begrensd voor $\omega \to \infty$'. Op zich zijn deze echter al zeer interessant; bovendien kunnen ze gebruikt worden als bouwstenen voor de constructie van hoogfrequente oplossingen van de exacte vergelijking $\Box u = 0$.

De methode is uit te breiden tot veel algemenere lineaire partiële differentiaalvergelijkingen. De fasefunctie $\phi(x)$ voldoet dan aan een eerste-orde niet-lineaire partiële differentiaalvergelijking van de vorm (8.4), die we met de Hamilton-Jacobi-theorie oplossen. De amplitude $a(x)$ moet dan vervolgens voldoen aan een vergelijking van de vorm

$$\sum_{j=1}^{n} \frac{\partial f}{\partial \xi_j}(x,\, \mathrm{d}\phi(x)) \cdot \frac{\partial a}{\partial x_j}(x) + c(x) \cdot a(x),$$

voor een geschikte functie $c(x)$. Dit is een transportvergelijking langs de oplossingen $x(s)$ van

$$\frac{dx_j}{ds} = \frac{\partial f}{\partial \xi_j}(x, \mathrm{d}\phi(x)).$$

Dit betekent dat $x(s)$ de projectie in X is van díe oplossingen $(x(s),\, \xi(s))$ van het Hamilton-stelsel met Hamilton-functie gelijk aan f, die in de variëteit Λ_ϕ lopen. Lemma 8.1 liet daarbij zien dat Λ_ϕ invariant is onder de H_f-stroming.

Voor meer details, inclusief een bespreking van de hoogfrequente asymptotische oplossingen bij brandfiguren, zie bijvoorbeeld [15].

8.4 Hamilton's Karakteristieke Functie

In Stelling 8.5 hebben we stroming van een Hamilton-stelsel in $\mathrm{T}^* X$ gebruikt om een functie ϕ in X te vinden die voldoet aan de partiële differentiaalvergelijking (8.4); dit was gebaseerd op de invariantie onde de H_f-stroming van de grafiek Λ_ϕ van $\mathrm{d}\phi$ in $\mathrm{T}^* X$. Er is verrassenderwijze ook een omkering hiervan: de stroming van een Hamilton-stelsel kan beschreven worden in termen van een reëelwaardige functie, welke voldoet aan een eerste-orde partiële differentiaalvergelijking waarin de Hamilton-functie voorkomt.

Om dit uit te leggen, beginnen we met de algemene situatie dat $(M,\, \sigma_M)$ en $(N,\, \sigma_N)$ symplectische variëteiten zijn. Een symplectische afbeelding van $(M,\, \sigma_M)$ naar $(N,\, \sigma_N)$ was in Paragraaf 7.5 gedefinieerd als een differentieerbare afbeelding $\Phi : M \to N$ met de eigenschap dat $\Phi^* \sigma_N = \sigma_M$. Merk op dat als $x \in M$, $v \in \mathrm{T}_x\, M$ en $\mathrm{T}_x\, \Phi(v) = 0$, dan geldt voor iedere $w \in \mathrm{T}_x\, M$ dat

$$\sigma_{M,\, x}(v,\, w) = \sigma_{N,\, \Phi(x)}\left(\mathrm{T}_x\, \Phi(v),\, \mathrm{T}_x\, \Phi(w)\right) = 0,$$

hetgeen op grond van de niet-gedegenereerdheid van $\sigma_{M,\, x}$ impliceert dat $v = 0$. Dit betekent dat een symplectische afbeelding automatisch een immersie is. Hebben we bovendien dat $\dim M = \dim N$, hetgeen we in het vervolg steeds zullen aannemen, dan is de conclusie dat een gladde symplectische afbeelding lokaal een glad diffeomorfisme is.

De grafiek $\mathrm{G}(\Phi)$ van Φ, de verzameling der $(m, \Phi(m)) \in M \times N$ met $m \in M$, is een gladde deelvariëteit van $M \times N$ en $\dim \mathrm{G}(\Phi) = \dim M = \frac{1}{2} \dim(M \times N)$. Wat de dimensie betreft is het daarom denkbaar dat de grafiek van Φ een Lagrange-variëteit is in $M \times N$, met betrekking tot een geschikte symplectische vorm in $M \times N$.

Lemma 8.6 *Zij (M, σ_M) en (N, σ_N) gladde symplectische variëteiten van dezelfde dimensie en zij Φ een gladde afbeelding van M naar N. Dan is Φ een symplectische afbeelding van (M, σ_M) naar (N, σ_N), dan en slechts dan als de grafiek van Φ een gladde Lagrange-variëteit in $M \times N$ is met betrekking tot de symplectische vorm $(-\sigma_M) \oplus (\sigma_N)$ in $M \times N$.*

Is M, resp. N een open deelverzameling van $\mathrm{T}^* X$, resp. $\mathrm{T}^* Y$, dan is de 'gespiegelde grafiek'

$$\mathrm{G}^{-}(\Phi) := \{((x, y), (-\xi, \eta)) \in \mathrm{T}^*(X \times Y) \mid (y, \eta) = \Phi(x, \xi)\} \tag{8.18}$$

van Φ een Lagrange-variëteit in $\mathrm{T}^*(X \times Y)$, dan en slechts dan als Φ een kanonieke transformatie is.

Bewijs De tweevorm $\sigma := (-\sigma_M) \oplus (\sigma_N)$ in $M \times N$ is als volgt gedefinieerd: voor iedere $m \in M$, $n \in N$, $u, \widetilde{u} \in \mathrm{T}_m M$, $v, \widetilde{v} \in \mathrm{T}_n N$ geldt dat

$$\sigma_{(m,n)}((u, v), (\widetilde{u}, \widetilde{v})) = -\sigma_{M,m}(u, \widetilde{u}) + \sigma_{N,n}(v, \widetilde{v}). \tag{8.19}$$

De raakruimte aan $\mathrm{G}(\Phi)$ in het punt (m, n), met $n = \Phi(m)$ is gelijk aan de verzameling der (u, v) met $u \in \mathrm{T}_m M$, $v \in \mathrm{T}_n N$ en $v = \mathrm{T}_x \Phi(u)$. Vullen we in (8.19) in $n = \Phi(m)$, $v = \mathrm{T}_x \Phi(u)$ en $\widetilde{v} = \mathrm{T}_x \Phi(v)$, dan zien we dat $\Phi^* \sigma_N$ in het punt m gelijk is aan σ_M, dan en slechts dan als $\mathrm{T}_{(m,n)} \mathrm{G}(\Phi)$ isotroop is voor σ.

Voor de tweede uitspraak merken we op dat de spiegeling $S : (x, \xi) \mapsto (x, -\xi)$ in de vezels van $\mathrm{T}^* X$ een 'antikanonieke transformatie' in $\mathrm{T}^* X$ is, in de zin dat

$$S^* \sigma_{\mathrm{T}^* X} = -\sigma_{\mathrm{T}^* X}.$$

We passen nu de eerste uitspraak toe met $M = \mathrm{T}^* X$, $\sigma_M = -\sigma_{\mathrm{T}^* X}$, $N = \mathrm{T}^* Y$, $\sigma_N = \sigma_{\mathrm{T}^* Y}$ en met Φ vervangen door $\Phi \circ S$. Daarbij merken we op dat Φ een kanonieke transformatie is, dan en slechts dan als $\Phi \circ S$ een symplectische afbeelding is van (M, σ_M) naar (N, σ_N). Tenslotte merken we op dat $\mathrm{G}^{-}(\Phi) = T(\mathrm{G}(\Phi \circ S))$, waarin de 'transpositie'

$$T : ((x, \xi), (y, \eta)) \mapsto ((x, y), (\xi, \eta))$$

van $\mathrm{T}^* X \times \mathrm{T}^* Y$ naar $\mathrm{T}^*(X \times Y)$ een symplectische afbeelding is waarbij in $\mathrm{T}^* X \times \mathrm{T}^* Y$ de directe som van de kanonieke symplectische vormen wordt gebruikt. $\square$

Combinatie van de tweede uitspraak in Lemma 8.6 met Lemma 8.1 leidt nu tot:

Gevolg 8.7 *Zij X en Y gladde n-dimensionale variëteiten, M een open deelverzameling van $\mathrm{T}^* X$ en $\Phi : M \to \mathrm{T}^* N$ een gladde kanonieke transformatie. Met de notatie $y(x, \xi) = \pi(\Phi(x, \xi))$ nemen we verder aan dat $(x_0, \xi_0) \in M$ en dat de lineaire afbeelding*

$$A := \frac{\partial y\,(x_0, \xi)}{\partial \xi}|_{\xi = \xi_0}$$

injectief is. Dan is er een open omgeving M_0 van $(x_0,\, \xi_0)$ in M, een open omgeving U van $(x_0,\, y_0)$, $y_0 = y\,(x_0,\, \xi_0)$, in $X \times Y$ en een gladde reëelwaardige functie ϕ in U, met de eigenschap dat voor $(x,\, \xi) \in M_0$ de vergelijking $(y,\, \eta) = \Phi(x,\, \xi)$ equivalent is met $(x,\, y) \in U$, $\xi = -\frac{\partial\phi(x,y)}{\partial x}$ en $\eta = \frac{\partial\phi(x,y)}{\partial y}$.

Bewijs We schrijven in lokale coördinaten

$$\Phi(x,\, \xi) = (y(x,\, \xi),\, \eta(x,\, \xi)).$$

en $\eta_0 = \eta\,(x_0,\, \xi_0)$. Dan is de raakruimte in $((x_0,\, y_0)\,,\, (y_0,\, \eta_0))$ aan $\mathrm{G}^-(\Phi)$ gelijk aan de verzameling der $((\delta x,\, \delta y),\, (-\delta\xi,\, \delta\eta))$, waarvoor

$$\delta y = \frac{\partial y}{\partial x}\,\delta x + \frac{\partial y}{\partial \xi}\,\delta\xi$$

en

$$\delta\eta = \frac{\partial \eta}{\partial x}\,\delta x + \frac{\partial \eta}{\partial \xi}\,\delta\xi,$$

waarbij alle afgeleiden genomen worden in het punt $(x_0,\, \xi_0)$. De raakafbeelding van de projectie naar $X \times Y$ is injectief op deze raakruimte aan $\mathrm{G}^-(\Phi)$, dan en slechts dan als $\delta x = 0$ en $\delta y = 0$ impliceren dat $\delta\xi = 0$ en $\delta\eta = 0$. Nu geeft $\delta x = 0$ en $\delta y = 0$ dat

$$0 = \frac{\partial y}{\partial \xi}\,\delta\xi,$$

dus de injectiviteit van A geeft dat $\delta\xi = 0$ en dit geeft op zijn beurt dat $\delta\eta = 0$. □

Een functie ϕ als in Gevolg 8.7 heet een *voortbrengende functie* voor de kanonieke transformatie Φ. Wanneer Φ gelijk is aan de stroming van een Hamilton-stelsel dat via een Legendre-transformatie verkregen is uit een Euler-Lagrange-stelsel, dan definieert de actie-integraal, de integraal van de Lagrange-functie langs de oplossingen, een voortbrengende functie:

Stelling 8.8 *Zij L als in Stelling 4.1 waarbij L aan de Legendre-voorwaarde voldoet en zij $\Psi : D \to \mathrm{T}\,X$ de a, b-stroming in $\mathrm{T}\,X$ van het Euler-Lagrange-stelsel, waarin D een open deelverzameling is van $\mathrm{T}\,X$. Zij $(x_0,\, v_0) \in D$ en neem aan dat de raakafbeelding in v_0 van $v \mapsto \pi\,(\Psi\,(x_0,\, v))$ bijectief is. Schrijf $y_0 = \pi\,(\Psi\,(x_0,\, v_0))$. Dan is er een open omgeving U, resp. V van $(x_0,\, v_0,\, b)$, resp. $(x_0,\, y_0,\, b)$ in $\mathrm{T}\,X \times \mathbf{R}$, resp. $X \times X \times \mathbf{R}$, met de eigenschap dat voor iedere $(x,\, v,\, t) \in U$ de oplossing $\gamma = \gamma_{x,\,v}$ van het Euler-Lagrange stelsel met $\gamma(a) = x$ en $\gamma'(a) = v$ gedefinieerd is in $[a,\, t]$ en dat de afbeelding*

$$(x,\, v,\, t) \mapsto (x,\, \gamma_{x,\,v}(t),\, t)$$

een diffeomorfisme is van U naar V.

Voor $(x,\, y,\, t) \in V$ schrijven we $\gamma_{x,\,y,\,t} = \gamma_{x,\,v}$ als $\gamma_{x,\,v}(t) = y$ en definiëren we

$$\phi(x,\, y,\, t) := \int_a^t L(s,\, \gamma_{x,\,y,\,t}(s),\, \gamma'_{x,\,y,\,t}(s))\, ds. \tag{8.20}$$

Dan is ϕ een differentieerbare reëelwaardige functie in V,

$$\frac{\partial\phi(x,\, y,\, t)}{\partial x} = -\frac{\partial L(a,\, x,\, v)}{\partial v}|_{\gamma_{x,\,v}(t)=y}, \tag{8.21}$$

$$\frac{\partial\phi(x,\,y,\,t)}{\partial y} = \frac{\partial L(t,\,y,\,w)}{\partial w}|_{w=\gamma'_{x,\,v}(t),\,y=\gamma_{x,\,v}(t)} \tag{8.22}$$

en tenslotte

$$\frac{\partial\phi(x,\,y,\,t)}{\partial t} = -h\left(t,\,y,\,\frac{\partial\phi(x,\,y,\,t)}{\partial y}\right), \tag{8.23}$$

waarin h de Legendre-transformatie van L voorstelt.

Is bovendien L tijdsonafhankelijk, dan voldoet ϕ ook nog aan de partiële differentiaalvergelijking

$$h\left(y,\,\frac{\partial\phi(x,\,y,\,t)}{\partial y}\right) - h\left(x,\,-\frac{\partial\phi(x,\,y,\,t)}{\partial x}\right) = 0. \tag{8.24}$$

Bewijs De eerste uitspraak is een directe toepassing van de impliciete-functiestelling.

De vergelijkingen (8.21) en (8.22) volgen uit (4.13), waarin b, resp. a, resp. γ vervangen worden door t, resp. s, resp. $\gamma_{x,\,y,\,t}$ en waarbij we gebruiken dat $[L]_\gamma = 0$. (Dit is in feite een generalisering van het bewijs van Lemma 6.5.)

Op dezelfde manier krijgen we dat de partiële afgeleide van ϕ naar t gelijk is aan

$$L(t,\,y,\,w) + \frac{\partial L(t,\,y,\,w)}{\partial w}\cdot\frac{\partial\gamma_{x,\,y,\,u}(t)}{\partial u}|_{u=t} = L(t,\,y,\,w) - \frac{\partial L(t,\,y,\,w)}{\partial w}\cdot w = -H(t,\,y,\,w),$$

waarbij we hebben gebruikt dat differentiatie van $\gamma_{x,\,y,\,t}(t) \equiv y$ naar t impliceert dat

$$\frac{\partial\gamma_{x,\,y,\,u}(t)}{\partial u}|_{u=t} + \gamma'_{x,\,y,\,t}(t) = 0.$$

Vervolgens leidt de vervanging van de snelheid w door de impuls in (8.22) tot (8.23).

Voor (8.24) merken we op dat de snelheids-impulsafbeelding de Euler-Lagrange-stroming conjugeert met de stroming in $\mathrm{T}^*\,X$ van het Hamilton-stelsel met Hamilton-functie gelijk aan h en dat daarvoor h een constante van beweging is. Daarbij voert de stroming in $\mathrm{T}^*\,X$ de impuls in (8.21) over in de impuls in (8.22). □

De vergelijking (8.23) kan, voor iedere x, gelezen worden als een eerste-orde partiële differentiaalvergelijking voor ϕ als functie van y en t, terwijl (8.24) voor iedere t een eerste-orde partiële differentiaalvergelijking is voor ϕ als functie van x en y.

Stelling 8.8 vindt zijn oorsprong in het werk [29] van Hamilton over de meetkunde van stralenbundels. Hij liet zien dat de stralenbundels uit de meetkundige optica gelijk zijn aan de projecties in X van de H_f-oplossingskrommen die in Λ_ϕ verlopen, waarbij ϕ een oplossing is van de partiële differentiaalvergelijking (8.4). Daarmee wordt een stralenbundel bepaald door het geven van maar één enkele reëelwaardige functie ϕ, die Hamilton de *karakteristieke functie* van de stralenbundel noemde. Dit was voor Hamilton de belangrijkste conclusie van zijn analyse. Zie Hoofdstuk 8.6 voor een bespreking van de geometrische optica in termen van Lagrange-variëteiten in $\mathrm{T}^*\,X$. Later, in [31], paste Hamilton dezelfde ideeën toe in de mechanica.

Jacobi [41], [42, 31. Vorlesung] reageerde daarop met de opmerking dat in het algemeen het oplossen van partiële differentiaalvergelijkingen lastiger is dan het oplossen van stelsels van gewone differentiaalvergelijkingen. *Wij* zouden zeggen: voor het laatste hebben we in ieder geval algemene stellingen over lokale existentie en eenduidigheid, afhankelijkheid van beginwaarden en parameters en, last but not least, numerieke benaderingsmethoden. Daarom benadrukte Jacobi dat een partiële differentiaalvergelijking van de vorm (8.4) opgelost kan worden door gebruik te maken van de oplossingen van het Hamilton-stelsel met Hamilton-functie gelijk aan f.

Lie merkte daarbij op, in een reeks van artikelen uit 1872-78, dat dit berust op de eigenschap dat de H_f-stroming zowel de functie f als de kanonieke tweevorm σ invariant laat. Lie zag daarbij een analogie tussen de groep van transformaties die σ en f invariant laten en de permutatiegroepen die Galois introduceerde bij zijn onderzoek aan hogere graads veeltermvergelijkingen. De ideeën van Lie over partiële differentiaalvergelijkingen zijn in een meer toegankelijke vorm uitgewerkt in het boek van Engel en Faber [21].

8.5 Nog Iets Algemenere Vergelijkingen

Een algemene eerste-orde partiële differentiaalvergelijking voor een reëelwaardige differentieerbare functie ψ in een variëteit Y is een vergelijking van de vorm

$$g\left(y,\, \psi(y),\, \mathrm{d}\psi(y)\right) = 0, \quad y \in Y. \tag{8.25}$$

Lezen we hierin $(y,\, \mathrm{d}\psi(y))$ als een element van $\mathrm{T}^*\, Y$ en $z = \psi(y) \in \mathbf{R}$, dan is g een functie die (na omschikken van $\psi(y)$ en $\mathrm{d}\phi(y)$ in een (meestal open) deelverzameling van $\mathrm{T}^*\, Y \times \mathbf{R}$ is gedefinieerd. Jacobi [42, p. 237] stelde voor om zo'n vergelijking te reduceren tot een vergelijking van de vorm (8.4) in de variëteit $X = Y \times \mathbf{R}$, door te schrijven $x = (y,\, z)$,

$$f(y,\, z;\, \eta,\, \zeta) = g\left(y,\, \zeta,\, \frac{1}{z}\, \eta\right)$$

en op te merken dat als ψ een oplossing is van (8.25), dan definieert $\phi(y,\, z) := z\, \psi(y)$ een oplossing ϕ van (8.4). Hij liet echter niet zien hoe omgekeerd oplossingen van (8.4) tot oplossingen van (8.25) leiden.

Een variant op Jacobi's idee, dat wel direct oplossingen van (8.25) uit oplossingen van (8.4) oplevert, is het idee van Lie [57, p. 108-110] om te definiëren

$$f(y,\, z;\, \eta,\, \zeta) = g\left(y,\, z,\, -\frac{1}{\zeta}\, \eta\right).$$

Hiervoor geldt dat als ϕ een gladde oplossing is van (8.4) met $\partial\phi(y,\, z)/\partial z \neq 0$ en c een reële constante is, dan is de vergelijking $\phi(y,\, z) = c$ lokaal equivalent met $z = \psi(y)$ voor een gladde functie ψ en

$$\mathrm{d}\psi(y) = -\frac{1}{\partial\phi(y,\, z)/\partial z}\, \frac{\partial\phi(y,\, z)}{\partial y},$$

dus is ψ een oplossing van (8.25). Zelfs krijgen we zo een familie van oplossingen van (8.25), namelijk voor iedere c definieert de niveauvariëteit van ϕ voor het niveau c een oplossing van (8.25).

Wat hier achter zit is dat de grafiek van ψ gezien wordt als een hyperoppervlak in X, waarvan de normaalbundel in $\mathrm{T}^*\, X$ een Lagrange-variëteit is die bevat is in de nulpuntsverzameling van de functie f. Lie [57] formuleerde dit in termen van de zogenaamde *contactstructuur* in de bundel van hypervlakken, een onderwerp waar we hier niet verder op in zullen gaan, maar dat nauw gerelateerd is aan de symplectische structuur van $\mathrm{T}^*\, X$.

8.6 Stralenbundels

In de geometrische optica voor inhomogene media is er in een open deelverzameling X van $\mathbf{R}^3$ een positieve reëelwaardige functie c gegeven en zijn de lichtstralen gedefinieerd als de geodeten met betrekking tot de Riemann-structuur waarvoor $\|v\|_x = \frac{\|v\|}{c(x)}$. Dat wil zeggen, $\|v\|_x = 1$ correspondeert met de uitspraak dat de Euclidische lengte van v gelijk is aan $c(x)$, daarom heeft $c(x)$ de interpretatie van de *lokale lichtsnelheid*. In de klassieke geometrische optica neemt men c constant, met een sprong bij passage van grensoppervlakken. Dit leidt tot de vereenvoudiging dat de geodeten rechte lijnen worden die met constante snelheid worden doorlopen.

De stralen uitgaande van een gegeven punt $p \in X$ zijn gelijk aan de oplossingen γ van het Euler-Lagrange-stelsel met Lagrange-functie gelijk aan $\frac{1}{2}\,\|v\|^2/c(x)^2$, waarvoor $\gamma(0) = p$ en $\gamma'(0)$ de verzameling der vectoren v met $\|v\|_p = 1$ doorloopt. De snelheids-impuls-afbeelding voert de corresponderende krommen $t \mapsto (\gamma(t),\, \gamma'(t))$ in $\mathrm{T}\,X$ over in de oplossingskrommen van het Hamilton-stelsel met Hamilton-functie $h = \frac{1}{2}\,c(x)^2\,\|\xi\|^2$, die starten op de sfeer $\|\xi\| = 1/c(p)$ in $(\mathrm{T}_p\,X)^*$. Omdat de vezels van de coraakbundel Lagrange-variëteiten zijn, is deze sfeer isotroop en het Hamilton-vectorveld raakt niet aan de sfeer, dus vanwege Lemma 8.3 vullen deze oplossingskrommen, wanneer we het definitiegebied beperken tot een deelverzameling B als in het lemma genoemd, een Lagrange-variëteit Λ in $\mathrm{T}^*\,X$, met $h = 1/2$ in Λ. Als c constant is, dan is deze conclusie ook op andere manieren heel gemakkelijk in te zien, voor variabele $c(x)$ is het bovenstaande argument het eenvoudigste.

Ter afkorting zullen we in het vervolg de bundel van lichtstralen, die de projecties in X zijn van de oplossingskrommen van het Hamilton-stelsel die in een Lagrange-variëteit Λ verlopen, de *Λ-bundel* noemen. Een bundel heet een *Lagrange-bundel* als er een Lagrange-variëteit Λ is waarvoor de bundel gelijk is aan de Λ-bundel. Is ϕ een differentieerbare reëelwaardige functie in de open deelverzameling U van X waarvoor $\Lambda = \Lambda_\phi$, dan is er voor iedere $x \in U$ precies één straal γ van de Λ-bundel en tijdstip t met $\gamma(t) = x$. Daarvoor geldt dat

$$\gamma'(t) = \frac{\partial h(x,\, \xi)}{\partial \xi} = c(x)^2\,\xi = c(x)^2\ \mathrm{grad}\,\phi(x), \tag{8.26}$$

waaruit we aflezen dat de straal loodrecht staat op het oppervlak $\phi =$ constant. Men zou de stralenbundel een *normalenbundel* kunnen noemen als er door ieder punt een oppervlak gaat waar alle stralen loodrecht op staan.

Het feit dat $h = 1/2$ in Λ betekent dat ϕ voldoet aan de partiële differentiaalvergelijking

$$c(x)^2\,\|\,\mathrm{grad}\,\phi(x)\|^2 - 1 = 0. \tag{8.27}$$

Dit is de eikonaalvergelijking van (8.15), maar nu met de factor $c(x)^2$ erin. Uit (8.26) en (8.27) volgt dat

$$\frac{d\phi(\gamma(t))}{dt} = \langle \mathrm{grad}\,\phi(x),\, \gamma'(t)\rangle = c(x)^2\,\|\,\mathrm{grad}\,\phi(x)\|^2 = 1,$$

dus $\phi(\gamma(t)) = t$ plus een constante. De parameter t van de lichtstraal heet de *optische weglengte* en we krijgen dat, op een additieve constante na, de functie ϕ gelijk is aan de optische weglengte naar het oppervlak waar de stralen loodrecht op staan.

Als een lichtstraal aankomt in een punt p van een spiegelend oppervlak S in X, met snelheidsvector v dwars op S, dat wil zeggen met $v \notin \mathrm{T}_p\,S$, dan betekent de terugkaatsingswet 'hoek van inval is gelijk aan hoek van uitval' voor de bijbehorende covector $\xi \in (\mathrm{T}_p\, X)^*$ dat de covector ξ' van de gereflecteerde straal voldoet aan

$$\xi'|_{\mathrm{T}_p\, S} = \xi|_{\mathrm{T}_p\, S}, \quad h(p,\, \xi') = \frac{1}{2} = h(p,\, \xi). \tag{8.28}$$

Dit betekent dat de afbeelding $\xi \mapsto \xi'$ de twee oplossingen ξ van $h(p,\, \xi) = 0$, met voorgeschreven beperking van ξ tot $\mathrm{T}_p\, S$, met elkaar verwisselt, zoals in Opmerking 8.4.

Hebben we een bundel van lichtstralen die de Λ-bundel is voor een Lagrange-variëteit Λ, dan geeft Lemma 8.4 dat het beeld onder ρ van de isotrope variëteit

$$\Lambda_0 := \Lambda \cap \mathrm{T}_S^*\, X$$

een Lagrange-variëteit Σ in $\mathrm{T}^*\, S$ is. De afbeelding $\xi \mapsto \xi'$ beeldt Λ_0 af naar een variëteit Λ_0' die op dezelfde manier als Λ_0 bij Σ hoort als in Lemma 8.4, maar nu met de andere keuze van de oplossing van $f = h - 1/2 = 0$. Bijgevolg is Λ_0' een isotrope variëteit en is, vanwege Lemma 8.3, het H_f-uitstroomsel van Λ_0' een Lagrange-variëteit Λ', waarin $h = 1/2$. Conclusie: de teruggekaatste bundel is opnieuw een Lagrange-bundel. Opgemerkt kan ook worden dat als $\Lambda = \Lambda_\phi$ en $\Lambda' = \Lambda_{\phi'}$, dan betekent de terugkaatsingswet (8.28) dat de beperkingen van ϕ en ϕ' tot S dezelfde afgeleide hebben, dus aan elkaar gelijk genomen kunnen worden.

Neem nu aan dat S de gemeenschappelijke rand is van twee disjuncte open deelverzamelingen ('halfruimten') X_+ en X_- van X. Zij verder $c(x) = c_\pm(x)$ voor $x \in X_\pm$, waarbij ieder van de functies $c_\pm$ en zijn afgeleiden een continue uitbreiding heeft tot de afsluiting $X_\pm \cup S$ van $X_\pm$ in X, waarbij echter deze functies in S niet overeen hoeven te stemmen. In deze situatie zegt men dat $c(x)$ een *gladde sprong* maakt langs S.

Het voorschrift voor de doorgang van lichtstralen van X_- naar X_+ is dat als γ_- de vanuit X_- op S invallende lichtstraal is dan is de naar X_+ uitvallende lichtstraal γ_+ bepaald door de volgende voorwaarden, als $\gamma_-(0) = \gamma_+(0) = p \in S$:

a) $\gamma_-'(0$ en $\gamma_+'(0)$ prikken beiden in de richting van X_+.

b) Is $\xi_\pm$ de bij $v_\pm = \gamma_\pm'(0)$ behorende covector, dat wil zeggen

$$v_\pm = \frac{\partial h_\pm(p,\, \xi)}{\partial \xi}|_{\xi = \xi_\pm} = c_\pm(p)^2\, \xi^\pm,$$

dan geldt dat

$$\xi_+|_{\mathrm{T}_p\, S} = \xi_-|_{\mathrm{T}_p\, S}, \quad h_-(p,\, \xi_+) = \frac{1}{2} = h_+(p,\, \xi_-). \tag{8.29}$$

De identiteit betreffende het h-niveau betekent dat als

$$w_\pm = c_\pm(p)\, \xi_\pm = c_\pm(p)^{-1}\, v_\pm,$$

dan is $\|w_\pm\| = 1$, dus w_-, resp. w_+ is de richtingsvector van de invallende, resp. uitvallende straal. In termen hiervan betekent (8.29) dat

$$w_+|_{\mathrm{T}_p\, S} = \frac{c_+(p)}{c_-(p)} w_-|_{\mathrm{T}_p\, S},$$

waarin we de *brekingswet van Snellius* herkennen: de hoek van de uitvallende straal met S is groter dan die van de invallende straal als $c_+(p) < c_-(p)$. We nemen in het vervolg aan dat w_- en w_+ bestaan en niet aan S raken. Is de invallende bundel gelijk aan de Λ_--bundel voor een Lagrange-variëteit Λ_-, dan geeft Lemma 8.4 dat het beeld onder ρ van de isotrope variëteit

$$\Lambda_{-,0} := \Lambda_- \cap \mathrm{T}^*_S X$$

een Lagrange-variëteit Σ in $\mathrm{T}^* S$ is. De afbeelding $\xi_- \mapsto \xi_+$ beeldt $\Lambda_{-,0}$ af naar een variëteit $\Lambda_{+,0}$ die op dezelfde manier als Λ_0 bij Σ hoort als in Lemma 8.4, maar nu met $f = h_-$ vervangen door $f = h_+$.

Bijgevolg is $\Lambda_{+,0}$ een isotrope variëteit en is, vanwege Lemma 8.3, het H_{h_+}-uitstroomsel van $\Lambda_{+,0}$ een Lagrange-variëteit Λ_+, waarin $h_+ = 1/2$. Conclusie: de gebroken bundel is opnieuw een Lagrange-bundel. Opgemerkt kan ook worden dat als $\Lambda_\pm = \Lambda_{\phi_\pm}$ dan betekent de brekingswet (8.29) dat de beperkingen van ϕ_- en ϕ_+ tot S dezelfde afgeleide hebben, dus aan elkaar gelijk genomen kunnen worden.

De functie ϕ moet voldoen aan de partiële differentiaalvergelijking (8.27) als het een fasefunctie is als in Paragraaf 8.3, maar nu met de standaard golfoperator vervangen door

$$\Box := \frac{\partial^2}{\partial t^2} - c(x)^2 \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}.$$

Met het oog op de hoogfrequente asymptotische oplossingen van de vergelijking $\Box u =$ begrensd van de vorm $u(x,\, t) = \mathrm{e}^{i\tau\,(t-\phi(x))}$ worden de oppervlakken waar de stralen loodrecht op staan, de oppervlakken $\phi =$ constant, ook wel de *golffronten* van de stralenbundel genoemd. Voor deze hoogfrequente golven ziet het er al helemaal zeer natuurlijk uit om te eisen dat bij terugkaatsing en breking de fasefuncties continu aansluiten.

De eerste beschrijving van stralenbundels in termen van de golffronten werd gegeven door Huygens [38]. Daarbij dient vermeld te worden dat, hoewel hij het over superpositie van verstoringen had, hij zijn theorie niet formuleerde in termen van functies $u(x,\, t)$ en er zijn dan ook geen fasefuncties, amplitudes of golfvergelijkingen in [38] te bekennen.

De stelling dat normalenbundels in normalenbundels overgaan bij terugkaatsing en breking aan oppervlakken werd voor het eerst geformuleerd door Malus [61]. Hij werkte in het geval van constante c en geloofde ten onrechte dat deze eigenschap alleen geldt als de invallende bundel uit één punt afkomstig is. Zie Atzema [8, p. 31]. Uit dit proefschrift heb ik ook geleerd dat later de *stelling van Malus*, zoals de uitspraak desondanks is gaan heten, van betere bewijzen zijn voorzien door Gergonne (1823), Quetelet (1825) en Hamilton [29]. De algemenere uitspraken, dat iedere bundel afkomstig uit één punt een Lagrange-bundel is en dat Lagrange-bundels in Lagrange-bundels overgaan bij terugkaatsing en breking, kan opgevat worden als een moderne versie van de stelling van Malus. Terzijde: bij passage door brandfiguren blijft de bundel een Lagrange-bundel, dus het is ook waar dat als een normalenbundel een brandpunt passeert, dan is het daarna opnieuw een normalenbundel.

Hamilton maakte in zijn bewijs gebruik van de oplossingen ϕ van (8.27). Hij liet zien dat als deze bij terugkaatsing, resp. breking op het oppervlak S continu aansluiten, dan voldoen de normalenbundels van de niveau-oppervlakken van ϕ aan de gebruikelijke terugkaatsings- resp. brekingswet. Voor Hamilton was het daarbij een belangrijke pointe dat een normalenbundel beschreven kan worden in termen van één enkele functie ϕ, die hij de *karakteristieke functie* van de bundel noemde, waarbij iedere functie ϕ genomen mag worden die voldoet aan de partiële differentiaalvergelijking 8.27.

De interpretatie van ϕ als fasefunctie van hoogfrequente golven kwam pas nadat Fresnel [24] met zijn buigingsproeven overtuigend aangetoond had dat het voorstellen van licht als hoogfrequente golven de eenvoudigste verklaring geeft voor de waargenomen intensiteiten. In [16] staan nog wat meer opmerkingen over de relaties tussen geometrische optica, golfoptica en lineaire partiële differentiaalvergelijkingen van het type van de golfvergelijking.

8.7 Vraagstukken

Vraagstuk 8.1 Een (inhomogene) *lineaire en eerste-orde* partiële differentiaalvergelijking

$$\langle v(x),\, \mathrm{d}\phi(x)\rangle = a(x) \tag{8.30}$$

is veel gemakkelijker op te lossen dan de algemene eerste-orde partiële differentiaalvergelijking (8.4). Hierbij nemen we aan dat v, resp. a een C^k-vectorveld, resp. C^k-functie in X is, met $k \geq 1$.

a) Is γ een oplossing van

$$\frac{d\gamma(t)}{dt} = v(\gamma(t)),$$

leid dan een inhomogene lineaire gewone differentiaalvergelijking af voor $t \mapsto \phi(\gamma(t))$ en los deze op in termen van $\phi(\gamma(0))$.

b) Schrijf $\Phi(x,\, t) = \gamma(t)$ als $\gamma(0) = x$. Het definitiegebied D van Φ is een open deelverzameling van $X \times \mathbf{R}$, $X \times \{0\} \subset D$ en Φ is een C^k-afbeelding van D naar X. Bewijs dit.

c) Zij S een $(n-1)$-dimensionale C^k-deelvariëteit van X, $x_0 \in S$, $T > 0$, $(x_0,\, T) \in D$. Neem verder aan dat $v\,(x_0) \notin \mathrm{T}_{x_0} S$ en dat voor iedere $0 < t \leq T$ geldt dat $\Phi\,(x_0,\, t) \notin S$. Bewijs dat er een open omgeving S_0 van x_0 in S is en een open interval I om $[0,\, T]$, waarvoor $D_0 := S_0 \times I \subset D$ en $\Phi_0 := \Phi|_{D_0}$ is een C^k-diffeomorfisme van D_0 naar een open deelverzameling U van X. Bewijs dat er bij iedere C^k-functie ψ in S_0 precies één differentieerbare oplossing ϕ van (8.30) is met $\phi|_{S_0} = \psi$. Geef een formule voor de oplossing in termen van ψ, a en de inverse van Φ_0.

d) Doorloop voor dit geval het bewijs van Stelling 8.5, dat wil zeggen identificeer Λ_0, de oplossingskrommen van het Hamilton-stelsel, Λ en de relatie tussen Λ en ϕ.

⊘

Vraagstuk 8.2 Zij Λ een Lagrange-variëteit in $\mathrm{T}^* X$. Schrijf $\tau_\Lambda = \iota_\Lambda^* \tau$ als ι_Λ de identiteit voorstelt, beschouwd als afbeelding van Λ naar $\mathrm{T}^* X$.

a) Bewijs dat er bij iedere $\lambda_0 \in \Lambda$ een open omgeving U van λ_0 in Λ is en een differentieerbare functie χ in U, met de eigenschap dat $\mathrm{d}\chi = \tau_\Lambda$ in U.

b) Als $(x_1,\, \ldots,\, x_n)$ lokale coördinaten in X zijn met bijbehorende lokale coördinaten $(x_1,\, \ldots,\, x_n;\, \xi_1,\, \ldots,\, \xi_n)$ in $\mathrm{T}^* X$ en $(\lambda_1,\, \ldots,\, \lambda_n)$ zijn lokale coördinaten in Λ, dan kunnen de x_j en ξ_j in Λ opgevat worden als functies van de variabelen λ_i. Bewijs dat in deze termen de functie χ bepaald is door het stelsel van differentiaalvergelijkingen

$$\frac{\partial \chi(\lambda)}{\partial \lambda_i} = \sum_{j=1}^{n} \xi_j(\lambda) \, \frac{\partial x_j(\lambda)}{\partial \lambda_i}, \quad 1 \leq i \leq n.$$

c) Bewijs dat als $\Lambda = \Lambda_\phi$, dan is $\chi = \phi \circ \pi$ plus een lokaal constante functie (in het gemeenschappelijke definitiegebied).

d) Zij $X = \mathbf{R}^2$, $f(x, \xi) = \frac{1}{2} \, \|\xi\|^2 - \frac{1}{2}$, Λ_0 gelijk aan de verzameling der

$$(\cos s, \, \sin s; \, -\sin s, \, \cos s)$$

met $s \in \mathbf{R}$. Bewijs dat het H_f-uitstroomsel gelijk is aan de verzameling Λ der

$$(\cos s - t \, \sin s, \, \sin s + t \, \cos s; \, -\sin s, \, \cos s)$$

met $s, \, t \in \mathbf{R}$. Bepaal de functies χ uit b), als functie van $(s, \, t)$.

e) Bewijs dat, alwaar de Λ uit d) gelijk is aan Λ_ϕ, moet gelden dat $\phi(x) = \chi(s, \, t)$ plus een constante, waarbij $x_1 = \cos s - t \, \sin s$ en $x_2 = \sin s + t \, \cos s$. Bewijs dat als x buiten de eenheidscirkel ligt, dan zij er twee combinaties $(s, \, t)$ die hieraan voldoen. Bepaal eerst t uit de vergelijking voor ${x_1}^2 + {x_2}^2$ en bepaal vervolgens $\cos s$ en $\sin s$ door de twee lineaire vergelijkingen daarvoor op te lossen. Beschijf deze combinaties $(s, \, t)$ ook in termen van de twee raaklijnen aan de eenheidscirkel die door het punt x gaan. Anderzijds, als x binnen de eenheidscirkel ligt, dan zijn er geen $(s, \, t)$. Geef een beschrijving van het (singuliere) gedrag van de functie $\phi(x)$ als x van buiten naar de eenheidscirkel toegaat.

$\oslash$

Vraagstuk 8.3 Zij f een gladde reëelwaardige functie in een open deelverzameling P van $\mathrm{T}^* X$, $(x_0, \, \xi_0) \in P$ en

$$\frac{\partial f \, (x_0, \, \xi)}{\partial \xi}|_{\xi = \xi_0} \neq 0.$$

Schrijf $\Phi(p, \, t) = \gamma(t)$ als γ een oplossing is van het H_f-stelsel met $\gamma(0) = p$. Bewijs de volgende uitspraken.

Het definitiegebied D van Φ is een open deelverzameling van $P \times \mathbf{R}$ met $P \times \{0\} \subset D$ en Φ is een gladde afbeelding van D naar P.

Er een open omgeving U van ξ_0 in $(\mathrm{T}_{x_0} X)^*$ en een open interval I om 0 in $\mathbf{R}$, met de volgende eigenschappen:

i) De verzameling Λ_0 der $\xi \in U$ met $f \, (x_0, \, \xi) = 0$ is een gladde, $(n-1)$-dimensionale en isotrope deelvariëteit van $\mathrm{T}^* X$.

ii) De beperking Φ_0 van Φ tot $\Lambda_0 \times I$ is een gladde inbedding van $\Lambda_0 \times I$ naar P en het beeld Λ is een gladde Lagrange-variëteit in $\mathrm{T}^* X$, waarop $f = 0$.

$\oslash$

Vraagstuk 8.4 Zij β een gladde Riemann-structuur in X is en zij Λ de Lagrange-variëteit die is gedefinieerd in Vraagstuk 8.3, met

$$f(x, \xi) = \frac{1}{2}\,{\beta_x}^{-1}(\xi, \xi) - \frac{1}{2}.$$

Beschrijf Λ in termen van geodeten. Bewijs de volgende uitspraken.

a) In normale coördinaten met x_0 in het middelpunt geldt, voor $x \neq 0$, x in een voldoende kleine open omgeving van 0, dat $(x, \xi) \in \Lambda$ dan en slechts dan als $\xi = \|x\|^{-1}\, x$.

b) Er is een open omgeving V van x_0 in X met de eigenschap dat $\phi : x \mapsto d\,(x_0, x)$ een gladde functie is in $V \setminus \{0\}$ en dat
$$\Lambda_\phi = \Lambda \cap \pi^{-1}\,(V \setminus \{0\})\,.$$

c) De functie ϕ in b) voldoet aan de partiële differentiaalvergelijking
$${\beta_x}^{-1}\,(\mathrm{d}\phi(x),\, \mathrm{d}\phi(x)) = 1, \quad x \in V,\ x \neq 0.$$

d) Is $\psi = \frac{1}{2}\,\phi^2$, dan is ψ glad in V en voldoet aan de partiële differentiaalvergelijking
$${\beta_x}^{-1}\,(\mathrm{d}\psi(x),\, \mathrm{d}\psi(x)) = 2\,\psi(x), \quad x \in V.$$

$\oslash$

Vraagstuk 8.5 Formuleer de theorie van stralenbundels voor geodeten met betrekking tot een willekeurige pseudo-Riemann-structuur. Beschrijf daarbij in het bijzonder wat er gebeurt bij terugkaatsing en breking. $\oslash$

Vraagstuk 8.6 Bewijs dat iedere één-dimensionale deelvariëteit van een symplectische variëteit isotroop is. Beargumenteer dat iedere stralenbundel in een vlak een normalenbundel is en dat daarmee de stelling van Malus in het vlak flauw is. $\oslash$

Vraagstuk 8.7 Beschouw de raaklijnenbundel aan de eenheidscirkel in het vlak $\mathbf{R}^2$, die gegeven wordt door aan iedere $s \in \mathbf{R}/2\pi\mathbf{Z}$ toe te voegen de rechte lijn

$$l_s : t \mapsto (\cos s - t\, \sin s,\, \sin s + t\, \cos s).$$

a) Bewijs dat door ieder punt $(x, y) \in \mathbf{R}^2$ met $x^2 + y^2 < 1$, resp. $x^2 + y^2 = 1$, resp. $x^2 + y^2 > 1$ geen, resp. één, resp. twee stralen van de bundel gaan.

b) Vind de differentieerbare krommen γ die loodrecht op de l_s staan door de differentieerbare functies $s \mapsto t(s)$ te bepalen, waarvoor
$$\langle \gamma'(s),\, l_s'(t(s))\rangle = 0 \quad \text{als} \quad \gamma : s \mapsto l_s(t(s)).$$

c) Vergelijk het resultaat met de uitkomst van Vraagstuk 8.2, d).

d) Zij γ_0 de kromme in het vlak die gedefinieerd is door

$$\gamma_0(s) := (\cos s + s\ \sin s,\ \sin s - s\ \cos s), \quad s \in \mathbf{R}.$$

Bewijs dat γ_0 loodrecht staat op de l_s en er voor iedere kromme γ uit b) een draaiïng R in het vlak om de oorsprong is en een translatie T in de parametrisering van de kromme, waarvoor $\gamma = R \circ \gamma_0 \circ T$.

e) Bewijs dat

$$\gamma_0(s) = (1,\ 0) + \left(\frac{1}{2}\,s^2,\ \frac{1}{3}\,s^3\right) + \mathcal{O}\left(s^4\right) \quad \text{als} \quad s \to 0.$$

f) Maak en schets van het beeld van γ_0 in het vlak.

$\oslash$

Referenties

[1] N.H. Abel: Solutions de quelques problèmes à l'aide d'intégrales définies. pp. 11-27 in *Œuvres Complètes*, t. I, Grøndahl & Søn, Christiania, 1881. Franse vertaling van een artikel in het noors in *Magazin for Naturvidenskabene* **2** (1823).

[2] N.H. Abel: Auflösung einer mechanischen Aufgabe. *Crelle's J. f. d. reine u. angew. Mathematik* **1** (1826) 153-157. Franse vertaling: pp. 97-101 in *Œuvres Complètes*, t. I, Grøndahl & Søn, Christiania, 1881.

[3] R. Abraham and J.E. Marsden: *Foundations of Mechanics*. The Benjamin/Cummings Publ. Co., Reading, Mass., 1987.

[4] W. Ambrose, R.S. Palais and I.M. Singer: Sprays. *Acad. Brasileira de Ciencias 32* (1960) 163-178.

[5] Appolonius van Perga: *Over de Kegelsneden*. Omstreeks 200 v. Chr., Alexandrië. Vertaling uit het grieks: *Les Coniques d'Appolonius de Perge*. Albert Blanchard, Paris 1963.

[6] V.I. Arnol'd: On the characteristic class entering in the quantization condition. *Functional Analysis and its Applications* **1** (1967) 1-13.

[7] V.I. Arnol'd and A. Avez: *Ergodic Problems of Classical Mechanics.* W.A. Benjamin. Inc., New York, Amsterdam, 1968.

[8] E.J. Atzema: *The Structure of Systems of Lines in 19th Century Geometrical Optics: Malus' Theorem and the Description of the Infinitely Thin Pencil.* Proefschrift, Rijksuniversiteit te Utrecht, 5 april 1993.

[9] Johannis Bernoulli: *Opera Omnia*. Herduk van de uitgave uit 1742, G. Olms, Hildesheim, 1968.

[10] A.L. Besse: *Einstein Manifolds*. Springer-Verlag, Berlin, etc. 1987.

[11] C. Carathéodory: Untersuchungen über die konformen Abbildungen von festen und veränderlichen Gebieten. *Mathematische Annalen* **72** (1912) 107-144.

[12] E.B. Christoffel: Ueber die Transformationen der homogenen Differentialausdrücke zweiten Grades. *J. f.d. reine u. angew. Math.* **70** (1869) 46–70, 241–245.

[13] E.A. Coddington and N. Levinson: *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company, Inc., New York, Toronto, London, 1955.

[14] G. Coriolis: Sur le principe des forces vives dans les mouvements relatifs des machines. *J. de l'École Polytechnique* **21** (1850) 268-297. Voorgelezen Acad. Sci. 6 juni 1831.

[15] J.J. Duistermaat: Oscillatory integrals, Lagrange immersions, and unfoldings of singularities. *Comm. Pure Appl. Math.* **27** (1974) 207-281.

[16] J.J. Duistermaat: Huygens' principle for linear partial differential equations. pp. 273-297 in *Huygens' Principle 1690-1990, Theory and Applications*, ed. H. Blok, H.A. Ferwerda, H.K. Kuiken, North-Holland, Amsterdam, 1992.

[17] J.J. Duistermaat: *Distributies.* Mathematisch Instituut Universiteit Utrecht, 1991 of later.

[18] J.J. Duistermaat: M. Riesz's family of operators. *Nieuw Archief voor Wiskunde* **9** (1991) 93-101.

[19] J.J. Duistermaat: *Differentiaalrekening op Variëteiten.* Mathematisch Instituut, Universiteit Utrecht, 1992.

[20] J.J. Duistermaat en W. Eckhaus: *Analyse van Gewone Differentiaalvergelijkingen.* Epsilon Uitgaven, Utrecht, 1995.

[21] F. Engel und K. Faber: *Die Liesche Theorie der Partiellen Differentialgleichungen Erster Ordnung.* Teubner, Berlin, 1935.

[22] L. Euler: De motu rectilineo trium corporum se mutuo attrahentium (E327). *Novi Comm. Ac. Soc. Petr.* **11** (1765) 144-152. Ook in: *Opera Omnia* II, **25**, 281-289.

[23] H.M. Farkas and I. Kra: *Riemann Surfaces.* Springer-Verlag, New York, etc., 1980, 1992.

[24] A. Fresnel: Mémoire sur la diffraction de la lumière. Couronné par l'Académie des Sciences en 1819. pp. 247-382 in *Œuvres Complètes d'Augustin Fresnel.* Paris, 1886. Johnson Reprint Corporation, New York, 1965.

[25] Galileo Galilei: *Dialogues concerning Two New Sciences.* Vertaling van het boek in het italiaans dat in 1638 bij Elsevier in Antwerpen werd uitgegeven. Dover Publications, New York.

[26] C.F. Gauss: Disquisitiones generales circa superficies curvas. *Comment. Soc. Gött., Math. Kl.* **6** (1828) 99-146 = *Werke*, Band 4, pp. 217-258. Duitse vertaling door A. Wangerin: *Allgemeine Flächentheorie.* Ostwald's Klassiker der exakten Wissenschaften **5**. Engelmann, Leipzig, 1889.

[27] H. Goldstein: Prehistory of the "Runge-Lenz vector". *Amer. J. Physics*, **43** (1975) 1123-1124. More on the prehistory of the Laplace or Runge-Lenz vector. ibid. **44** (1976) 737-738.

[28] G. Hamel: *Theoretische Mechanik.* Springer-Verlag, Berlin-Göttingen-Heidelberg, 1949.

[29] W.R. Hamilton: Essay on the theory of systems of rays, with three supplements, *Trans. R. Irish Academy* **15** (1828) 69-174, **16** (1830) 4-62, **16** (1831) 85-92, **17** (1837) 1-144.

[30] W.R. Hamilton: On the application of the method of quaternions to some dynamical questions. *Proc. Roy. Irish Acad.* **3** (1847) Appendix xxxvi-l (Communicated on 14 and 21 July 1845) = pp. 441-448 in: *The Mathematical Papers of Sir William Rowan Hamilton, vol III: Algebra.* Cambridge University Press, 1967.

[31] W.R. Hamilton: On a general method in dynamics. *Phil Trans. R. Soc. London* (1834) 247-308, (1835) 95-144.

[32] M. Hazewinkel (ed.): *Encyclopedia of Mathematics.* Translation of the Soviet Mathematical Encyclopaedia. Kluwer Academic Publishers, 1990.

[33] S. Helgason: *Differential Geometry and Symmetric Spaces.* Academic Press, New York, San Fransisco, London, 1962.

[34] G. Hill: Researches in lunar theory. *Amer. Math. J.* **1** (1878) 5-26, 129-147, 245-260.

[35] M.W. Hirsch, C.C. Pugh and M. Shub: *Invariant Manifolds.* Lecture Notes in Mathematics 583. Springer-Verlag, Berlin, etc., 1977.

[36] A. Hurwitz: Ueber die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt. *Math. Annalen* **46** (1895) 273-284 = *Mathematische Werke*, vol. 2, pp. 533-545.

[37] Christiaan Huygens: *Horologium Oscillatorium.* F. Muget, Paris 1673. Herdruk: Culture et Civilisation, Bruxelles, 1966. Duitse vertaling: Ostwald's Klassiker de Exacten Wissenschaften, No. 192, Wilhelm Engelmann Verlag, Leipzig, 1913. Engelse vertaling: *The Pendulum Clock.* The Iowa State Univerity Press, AMES, 1986. Zie ook: E.J. Dijksterhuis: De ontdekking van het tautochronisme der cycloïdale valbeweging. Een bijdrage tot de 300e herdenking van den geboortedag van Christiaan Huygens op 14 April 1929. *Euclides* **5** (1928 - 1929) 193-208.

[38] C. Huygens: *Traité de la Lumière.* Van der Aa, Leyden, 1690. Engelse vertaling: *Treatise on Light*, 1690, herdukt door Dover Publications, New York, 1962. Nederlandse vertaling door Dieuwke Eringa: *Verhandeling over het Licht.* Epsilon Uitgaven, Utrecht, 1990.

[39] C.G.J. Jacobi: Sur le mouvement d'un point et sur un cas particulier du problème des trois corps. *Gesammelte Werke*, IV. Band, Berlin 1886. Herduk Chelsea Publ. Cy., New York, 1969.

[40] C.G.J. Jacobi: Nova methodus, equationes differentiales partiales primi ordinis inter numerum variabilium quemcuque propositas integrandi. Crelle Journal für die reine und angewandte Mathematik, **60** (1862) 1-181. Ook in: *Gesammelte Werke*, V. Band, pp. 1-189, Berlin 1890. Herduk Chelsea Publ. Cy., New York, 1969.

[41] C.G.J. Jacobi: Über die Reduction der Integration der partiellen Differentialgleichungen erster Ordnung zwischen irgend einer Zahl Variabeln auf die Integration eines einzigen Systemes gewḧnlicher Differentialgleichungen. *Crelle's J. f. d. reine u. angew. Math.* **17** (1837) 97-162. Ook in: *Gesammelte Werke*, 4. Band, 1886. Herdruk: Chelsea Publ. Cy., New York, 1969.

[42] C.G.J. Jacobi: *Vorlesungen über Dynamik.* Gehalten an der Universität Königsberg im Wintersemester 1842-1843 und nach einem von C.W. Borchardt ausgearbeiteten Hefte. Verlag G. Reimer, Berlin 1881. Reprint Chelsea Publ. Co., New York, N.Y., 1969.

[43] J. Kepler: *De Coni Sectionibus.* Deel IV, 4 van "Ad Vitellionem Paralipomena quibus Astronomiae Pars Optica Traditur". Claude Marne, Frankfurt, 1604. Vertaling uit het latijns: *Über die Kegelschnitte.* Buchhandlung Thomas Dittert, Baunatal, 1990.

[44] J. Kepler: *Astronomia Nova* (1609). Vertaling uit het latijns: *Neue Astronomie.* Oldenbourg, München, Berlin, 1929.

[45] J. Kepler: *Harmonice Mundi* (1619). Vertaling uit het latijns: *Weltharmonik.* Oldenbourg, München, Berlin, 1939.

[46] W. Klingenberg: *Lectures on Closed Geodesics.* Springer-Verlag, Berlin, Heidelberg, New York, 1978.

[47] S. Kobayashi and K. Nomizu: *Foundations of Differential Geometry, I, II.* Interscience Publishers, New York, London, 1963, 1969.

[48] P. Koebe: Zur Uniformisierung der beliebiger analytischen Kurven. *Göttinger Nachr.* (1907) 191-210.

[49] L. Kronecker: Näherungsweise ganzzahlige Auflösung linearer Gleichungen. Monatsber. Kónigl. Preuss. Akad. Wiss. Berlin (1884) 1179-1193, 1271-1299 = Werke, band 3:1, p. 47-109, Teubner Verlag, Leipzig 1899.

[50] R. Kronig (red.): *Leerboek der Natuurkunde*. Scheltema en Holkema, Amsterdam, 1958.

[51] J. Lagrange: Essai d'une nouvelle méthode pour resoudre le problème des trois corps. *Prix de l'Academie* (1772), in 1777 gepubliceerd bij Panekoucke, Paris. Ook in *Œuvres*, vol. 6, Gauthier-Villars, Paris, 1873.

[52] J.-L. Lagrange: *Mécanique Analytique.* Eerste druk 1788 in Parijs. Herdruk bij A. Blanchard, Paris, 1965.

[53] L. Landau and E. Lifschitz: *The Classical Theory of Fields.* (Vertaald uit het Russisch.) Pergamon Press, 1951.

[54] S. Lang: *Differential Manifolds.* Addison-Wesley, Reading, Mass., 1972.

[55] P.S. Laplace: *Traité de Mécanique Celeste*. J.B.M. Duprat, Paris 1798. Herdruk Chelsea, New York, 1966.

[56] S. Lie u. F. Engel: *Theorie der Transformationsgruppen I.* Teubner Verlag, Leipzig u. Berlin, 1888, reprinted in 1930.

[57] S. Lie u. F. Engel: *Theorie der Transformationsgruppen II.* Teubner Verlag, Leipzig u. Berlin, 1890, reprinted in 1930.

[58] J. Liouville: Note sur quelques intégrales définies. *J. de Math. Pures et Appl.* **4** (1839) 225-235.

[59] J. Liouville: Mémoire sur un cas particulier du problème des trois corps. *J. de Math. Pures et Appl.*, 2ème Série, **1** (1856) 248-264.

[60] A.E.H. Love: *A Treatise on the Mathematical Theory of Elasticity.* Cambridge University Press, 1944.

[61] E. Malus: Mémoire sur l'optique. *J. de l.École Polytechnique* **14** (1808) 1-44, 84-129.

[62] M. Marden: *The Geometry of the Zeros of a Polynomial in a Complex Variable.* Math. Surveys III, Amer. Math. Soc., 1949.

[63] V.P. Maslov: *Théorie des Perturbations et Méthodes Asymptotiques.* Dunod, Gauthiers-Villars, Paris, 1972. Vertaald uit de russische uitgave van de Universiteit van Moskou, 1965.

[64] J.J. van der Meer: *The Hamiltonian Hopf Bifurcation.* Proefschrift, Universiteit van Utrecht, 1985. = Lecture Notes in Mathematics 1160. Springer-Verlag, Berlin, etc., 1985.

[65] J. Moser: On the volume elements on a manifold. *Trans. Amer. Math. Soc.* **120** (1965) 286-294.

[66] C.H. Müller und A. Timpe (1906): Die Grundgleichungen der mathematische Elastizitätstheorie. pp. 1-54 in: *Encyklopädie der Mathematischen Wissenschaften, IV.4.*

[67] J. von Neumann: Zur Theorie der Darstellungen kontinuerlicher Gruppen. Sitzungsber. Preuss. Akad. Wiss. Phys.-math. Kl. (1927) 76-90 = Collected Works, Vol. I, p. 134-148. Pergamon Press, Oxford, 1961.

[68] I. Newton: *Mathematical Principles of Natural Philosophy and System of the World.* Vertaling van de latijnse "Principia" , die in 1686 in Londen werd gepubliceerd. University of California Press, Berkeley, vierde druk, 1960.

[69] *The Mathematical Papers of Isaac Newton*, vol. IV. Edited by D.T. Whiteside. Cambridge University Press, 1971.

[70] E. Noether: Invarianten beliebiger Differentialausdrücke, resp. Invariante Variationsprobleme, *Nachr. Ges. Wiss. Göttingen* (1918) 37-44, resp. 237-257 = *Gesammelte Abhandlungen*, Springer-Verlag, 1983, pp. 240-247, resp. 248-270.

[71] J. Palis Jr.: Vector fields generate few diffeomorphisms. Bull. Amer. Math. Soc. **80** (1974) 503-505.

[72] H.C. Plummer: Instability of the Lagrange collinear solutions. *Monthly Not. Roy. Astron. Soc.* **62** (1901) 6-12.

[73] H. Poincaré: Théorie de groupes Fuchsiennes. *Acta Math.* **1** (1882) 1-62.

[74] H. Poincaré: Sur le problème des trois corps et les équations de la dynamique. (Mémoire couronné du prix de Sa Majesté le Roi Oscar II.) *Acta mathematica* **13** (1890) 5-271.

[75] H. Poincaré: *Méthodes Nouvelles de la Mécanique Céleste*, vol. I-III. Gauthier-Villars, Paris, 1892, 1893, 1900. Herdrukt in: Dover Publications, New York, 1957.

[76] H. Poincaré: Sur l'uniformisation des fonctions analytiques. *Acta Math.* **31** (1907) 1-63.

[77] H. Poincaré: *Papers on Fuchsian Functions.* (Vertaling in het engels door J. Stilwell.) Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1985.

[78] B. Riemann: Versuch einer allgemeinen Auffassung der Integration und Differentiation. (1847, Nachlass.) pp. 353-366 in: ed. H. Weber: *Gesammelte Mathematische Werke.* 1876, 1892, 9102. Dover reprint, 1953.

[79] M. Riesz: L'intégrale de Riemann-Liouville et le problème de Cauchy. *Acta Mathematica* **81** (1949) 1-223.

[80] J. Robbin: On the existence theorem for differential equations. *Proc. Amer. Math. Soc.* **19** (1968) 1005-1006.

[81] E.J. Routh: On Laplace's three particles, with a supplement on the stability of steady motion. Proc. London Math. Soc. **6** (1875) 86-97. Opgenomen als pp. 145-157 in: *Stability of Motion.* Taylor & Francis Ltd., London, 1975.

[82] E.J. Routh: *Stability of a Given State of Motion.* Essay, London, 1877. Opgenomen als pp. 19-138 in: *Stability of Motion.* Taylor & Francis Ltd., London, 1975.

[83] E.J. Routh: *Dynamics of a System of Rigid Bodies.* 6th edition, London, 1905.

[84] H.A. Schwarz: *Zur Theorie der Abbildung.* Programm der eidgenössischen polytechnischen Schule in Zürich, 1869-70 = *Gesammelte Mathematische Abhandlungen*, Band 2, Springer, Berlin, 1890.

[85] W. Thomson and P.G. Tait: Treatise on Natural Philosophy, Part II, New edition. Cambridge University Press, 1883.

[86] K. Weierstrass: Über eine Gattung reell periodischer Functionen. *Königl. Akad. de. Wiss.* **185** (1866) 97-115 = *Werke*, 2. Band, 1-18.

[87] S. Weinberg: *Gravitation and Cosmology.* Wiley Interscience, 1972.

[88] A. Weinstein: Symplectic manifolds and their Lagrangian submanifolds. *Adv. in Math.* **6** (1971) 329-346.

[89] D.T. Whiteside: Patterns of mathematical thought in the later seventeenth century. *Archive for History of Exact Sciences*, **1** (1961) 179-388.

[90] J.A. Wolf: *Spaces of Constant Curvature.* McGraw-Hill, New York, etc., 1967.

Index

D

E

L

M

N