Analyse: van $\mathbb{R}$ naar $\mathbb{R}^n$

Tom Koornwinder Jan Wiegerinck

10 januari 2014

Inhoudsopgave

Deel I

Aanvullingen op Ross

Dit eerste deel van de syllabus is een aanvulling op het boek van Kenneth A. Ross, *Elementary Analysis: The Theory of Calculus.* Bij de eerste vijf hoofdstukken uit het boek vind je hier aanvullingen en aanvullende opgaven: sommige om de theorie toe te lichten, andere meer van het niveau dat je op het tentamen kunt verwachten. Het boek is niet erg ambitieus in zijn opgaven, en huppelt af en toe over een moeilijk punt heen. Voor studenten wiskunde zijn een goede grondslag van de analyse en bijpassende opgaven noodzakelijk om verder in het vak te komen. Deze syllabus helpt daarbij.

Verwijzingen in **vet** zijn steeds naar definities, stellingen etc. in het boek van Ross.

Aantekeningen bij de hoofdstukken die in het college "analyse op de lijn" behandeld werden, zijn voor dit college natuurlijk van minder belang.

Jan Wiegerinck

1 Reële getallen en de supremumeigenschap

1.1 Bij paragraaf 4

We zullen *de volledigheid* van $\mathbb{R}$, dat is het **Completeness Axiom 4.4**, dus de eigenschap dat iedere niet-lege naar boven begrensde verzameling in $\mathbb{R}$ een *kleinste bovengrens* of *supremum* in $\mathbb{R}$ heeft, heel vaak gebruiken (zie **pp. 20–22, Definitions 4.2** en **4.3**). Deze eigenschap is een definiërende eigenschap van $\mathbb{R}$ in de volgende zin: Er bestaat (op isomorfe kopieën na) precies één *ordered field* (= geordend lichaam) met eigenschap **4.4**. Dit lichaam noemen we $\mathbb{R}$. Het heeft alle eigenschappen die we intuïtief al aan de reële getallen toekenden. Echter, door deze nieuwe opzet kunnen we de eigenschappen voorzover die niet in de definitie zijn opgenomen, ook bewijzen! In **§6** wordt een volledig, geordend lichaam geconstrueerd. De stellingen dat geordende lichamen met eigenschap **4.4** allemaal geordend-lichaamsisomorf zijn, vallen buiten het bestek van dit college.

Ga zelf na dat je $\mathbb{R}$ een kopie van $\mathbb{Q}$ bevat. Het is belangrijk dat je je realiseert dat *deel*verzamelingen van $\mathbb{R}$ eigenschap **4.4** *niet* hoeven te hebben. Neem maar de deelverzameling $\mathbb{Q}$: er zijn begrensde deelverzamelingen A van $\mathbb{Q}$ waarvoor géén element van $\mathbb{Q}$ bestaat dat de kleinste bovengrens van A is; we geven zo meteen een voorbeeld.

Merk op dat in **Definition 4.3** niet wiskundig wordt uitgelegd wat 'kleinste' betekent, maar de betekenis is als in het dagelijks taalgebruik: 'kleinste' betekent 'kleiner dan alle andere', wat hetzelfde is als 'kleiner dan of gelijk aan alle'. Een definitie van kleinste bovengrens is dus:

Definitie 1.1. Als $(X, \leq)$ een totaal geordende verzameling is, dat wil zeggen een verzameling met de eigenschappen **O1–O3**, en $A \subset X$, dan heet $M \in X$ de *kleinste bovengrens van* A, of, als je misverstand wilt voorkomen in het geval $X \subset Y$ voor een of andere Y, de *kleinste bovengrens van* A *in* X, als (1) voor alle $x \in X$ geldt: $x \in A \Rightarrow x \leq M$ (dus als M een *bovengrens van* A is) en (2) voor elke bovengrens $Q \in X$ van A geldt: $M \leq Q$.

Voorwaarde (2) hierboven is, net als dat boven **§4 Example 3** in het boek gebeurde, ook te formuleren als

$$\text{voor alle } Q \in X \text{ geldt: } Q < M \implies \exists x \in A: \ x > Q. \tag{1.1}$$

Bijvoorbeeld $A = \{x \in \mathbb{Q} : x^2 < 2\}$. Dit is een verzameling rationale getallen die naar boven begrensd is, want 2 is een bovengrens. Neem nu $X := \mathbb{R}$ in definitie 1.1, dan is het supremum van A in $\mathbb{R}$ gelijk aan $\sqrt{2}$. Maar met $X := \mathbb{Q}$ lukt het niet: er bestaat volgens aanvullende opgave 1.3 geen rationaal getal $M \in \mathbb{Q}$ dat aan definitie 1.1 voldoet.

In het vervolg zullen we altijd $X = \mathbb{R}$ nemen, tenzij expliciet anders vermeld staat.

Iets over gaten. Op **p. 19 §4 rr. 2–3**:

'[The completeness axiom] is the axiom that will assure us that $\mathbb{R}$ has no "gaps".'

en **p. 22 rr. 7–9**:

'[...] $\sup S$ exists. This is not an accident. Otherwise there would be a "gap" between the set S and the set of its upper bounds.'

wordt het bestaan van 'gaten' equivalent verklaard met het bestaan van niet-lege naar boven begrensde deelverzamelingen die geen supremum hebben. Natuurlijk heeft geen enkele verzameling in zichzelf beschouwd gaten. Gaten zie je pas als je een verzameling binnen een grotere verzameling beschouwt en gaten zijn per definitie geen onderdeel van de verzameling maar van de grotere verzameling. Het is het eenvoudigst om de equivalentie dan als definitie van gat te interpreteren. Er treedt een gat op voor iedere naar boven begrensde deelverzameling van $\mathbb{Q}$ die geen supremum heeft in $\mathbb{Q}$, het gat is dan het supremum in $\mathbb{R}$. Als je alleen maar $\mathbb{Q}$ kende en van de reële rechte met z'n $\sqrt{2}$ en zo nooit gehoord had, kun je zo aan het niet bestaan van suprema zien dat er 'gaten' in $\mathbb{Q}$ als deel van $\mathbb{R}$ zitten.

1.2 Aanvullende opgaven

1.2. Bewijs dat definitie 1.1 en het alternatief gegeven met behulp van (1.1) equivalent zijn.

1.3. Bewijs dat $A = \{x \in \mathbb{Q} : x^2 < 2\}$ geen supremum in $\mathbb{Q}$ heeft.

1.4. Bewijs de opmerking in het boek (**p. 21 r.–13**) dat een deelverzameling van $\mathbb{R}$ ten hoogste één supremum en één infimum kan hebben (als afgesproken: alles in $\mathbb{R}$).

1.5. Nu naar de algemene situatie van definitie 1.1: wat je bewijs van 1.4 hopelijk laat zien, eventueel na een kleine aanpassing, is dat *in een totaal geordende verzameling* X een verzameling $A \subset X$ ten hoogste één supremum kan hebben, met andere woorden, dat als zo'n M bestaat, deze uniek bepaald is. Laten we dit supremem $\sup_X(A)$ noemen. Veronderstel nu dat X met behoud van ordening een deelverzameling is van een of andere totaal geordende verzameling Y, en $A \subset X, A \neq \emptyset$. Bewijs of weerleg:
(1) als $\sup_Y(A)$ bestaat, bestaat $\sup_X(A)$;
(2) als $\sup_X(A)$ bestaat, bestaat $\sup_Y(A)$;
(3) als $\sup_X(A)$ en $\sup_Y(A)$ bestaan, zijn ze gelijk;
(4) de implicaties (2) en (3) zijn waar als $X = \mathbb{Q}$ en $Y = \mathbb{R}$.

1.6. Van welke verzamelingen in **Exercise 4.1** bestaat een supremum in $\mathbb{Q}$?

1.7. Definieer het product $A \cdot B$ van begrensde deelverzamelingen A en B van $\mathbb{R}$ als $A \cdot B := \{a \cdot b : a \in A, b \in B\}$. Geef een voorbeeld waaruit blijkt dat niet altijd geldt $\sup A \cdot B = \sup A \cdot \sup B$. Geef ook een redelijke voorwaarde die garandeert dat dit wel het geval is. Bewijs je uitspraak.

1.8. Veronderstel dat A en B begrensde verzamelingen van $\mathbb{R}$ zijn met $B \subset A \subset \mathbb{R}$. Bewijs dat $\sup B \leq \sup A$ en $\inf B \geq \inf A$. Laat zien dat gelijkheid kan optreden, ook als $A \neq B$.

1.9. Veronderstel dat A en B begrensde verzamelingen van $\mathbb{R}$ zijn met $B \subset A \subset \mathbb{R}$ waarvoor geldt:

$$\forall a \in A, n \in \mathbb{N} \quad \exists b \in B \text{ met } a \leq b + 1/n.$$

Bewijs dat $\sup A = \sup B$.

1.10. Bepaal

$$\sup\{\sup S : S \subset (0,1) \cap \mathbb{Q}\}.$$

1.11. Laat zien dat het open interval $(0,1)$ niet geschreven kan worden als $\cup_{\alpha \in A} I_\alpha$ waar A een indexverzameling is die uit twee of meer elementen bestaat en de $I_\alpha = (a_\alpha, b_\alpha) \subset (0,1)$ onderling disjuncte, niet-lege intervallen zijn.

1.12. Laat zien dat voor ieder niet-negatief getal $y \in \mathbb{R}$ de vergelijking $x^2 = y$ een unieke niet-negatieve oplossing in $\mathbb{R}$ heeft. We noemen die $\sqrt{y}$. Waarom geeft je recept geen oplossing voor negatieve y?

In de volgende opgaven wordt gesuggereerd hoe nu y^α te definiëren voor positieve α en y, en te controleren dat de gebruikelijke eigenschappen gelden. Het is een elementaire benadering, maar zeker niet de snelste, zie ook **§37.1**.

1.13. (vervolg) Laat zien dat voor ieder getal $y \in \mathbb{R}$ en iedere $n = 1, 2, \ldots$ de vergelijking $x^{2n-1} = y$ een unieke oplossing $\sqrt[2n-1]{y}$ in $\mathbb{R}$ heeft, en voor ieder niet-negatief getal $y \in \mathbb{R}$ en iedere $n = 1, 2, \ldots$ de vergelijking $x^{2n} = y$ een unieke niet-negatieve oplossing $\sqrt[2n]{y}$ in $\mathbb{R}$ heeft. Wat is er nu tegen om bijvoorbeeld te definiëren $(-8)^{2/6} = \sqrt[6]{(-8)^2}$?

1.14. (vervolg) Dus even geen negatieve y, en ook maar meteen niet $y = 0$, want anders geeft die straks nog moeilijkheden. Laat zien dat voor alle positieve $y \in \mathbb{R}$ de gebruikelijke definitie van $y^{p/q}$ $(p, q \in \mathbb{N})$ een functie $r \in \mathbb{Q}^+ \mapsto y^r \in \mathbb{R}$ op de positieve rationale getallen definieert.

1.15. (vervolg) Breid de definitie uit opgave 1.14 uit tot niet-positieve rationale exponenten.

1.16. (slot) Hoe zou je nu y^α voor $y > 0$, $\alpha \in \mathbb{R}$ definiëren? En hoe zou je laten zien dat de gewone rekenregels als $y^\alpha \cdot y^\beta = y^{\alpha+\beta}$, $\alpha, \beta \in \mathbb{R}$, nog gelden?

2 Rijen

2.1 Bij paragraaf 7 en 8

Het is noodzakelijk dat je een formeel bewijs kunt geven van uitspraken als $\lim \frac{1}{n^2} = 0$. Daarbij is het belangrijk onderscheid te maken tussen twee uitspraken

1. Voor iedere $\varepsilon > 0$ is er een $N \geq 0$ zodat $n \in \mathbb{N}$, $n > N$ impliceert $|1/n^2 - 0| < \varepsilon$.
2. Zoek voor iedere $\varepsilon > 0$ de kleinste $N \in \mathbb{N}$ zodat $n \geq N$ impliceert $|1/n^2 - 0| < \varepsilon$.

Natuurlijk zal in beide gevallen N van ε afhangen. Je moet in het bewijs dan aangeven hoe je N kiest. Meestal is geval 1 veel makkelijker te doen (als je N hebt, kun je ook $N + 7$ nemen), er is veel keus. In geval 2 heb je geen keus, vaak is de waarde van N dan op een ingewikkelde manier afhankelijk van ε. Uitspraak 1 is wat je nodig hebt om het bestaan van een limiet te bewijzen. Uitspraak 2 doet eigenlijk "te veel", maar het kan voor andere doelen handig zijn om die kant op te gaan.

Voorbeeld 2.1 (Example 8.1). Het formele bewijs in het boek doet (2). Als je het voorbeeld met (1) doet gaat het zo. Zij $\varepsilon > 0$. Als $\varepsilon > 1$ kies je $N = 1$ en $n > 1$ impliceert $|1/n^2 - 0| < 1 < \varepsilon$. Als nu $0 < \varepsilon < 1$, kies $N = 1/\varepsilon$. Dan geldt voor $n > N$

$$|1/n^2 - 0| = 1/n^2 < \varepsilon^2 \leq \varepsilon,$$

omdat $0 < \varepsilon < 1$. Als $1/n^2 < \varepsilon$ voor $n > N$, is voor iedere $\eta > \varepsilon$ ook $1/n^2 < \eta$ voor $n > N$. Het behandelen van het geval $\varepsilon > 1$ is daarmee eigenlijk overbodig.

Voorbeeld 2.2 (**7.3b en 7.3 p**). We beginnen met **7.3.b**. Merk op dat $b_n = \frac{1+3/n^2}{1-3/n^2}$ dus net als in voorbeeld 2 raden we dat $\lim b_n = 1$. Nu een bewijs. Neem $\varepsilon > 0$. We willen het verschil tussen b_n en 1 schatten;

$$|b_n - 1| = \left| \frac{n^2 + 3 - n^2 + 3}{n^2 - 3} \right| \leq \frac{6}{(n-2)(n+2)} \leq \frac{6}{n+2}, \quad \text{als } n > 3. \tag{2.1}$$

De laatste uitdrukking is kleiner dan ε als $n > 1/(6\varepsilon)$ (en $n > 3$). Kies dus $N = \max\{1/(6\varepsilon), 3\}$, dan is voor $n > N$ $|b_n - 1| < \varepsilon$.

Nu **7.3.p** Laat $b_n = \frac{2^{n+1}+5}{2^n-7} = \frac{2+5/(2^n)}{1-7/(2^n)}$ zodat we raden dat $\lim b_n = 2$. Nu weer een bewijs. Zij $\varepsilon > 0$. Er geldt

$$|b_n - 2| = \left| \frac{12}{2^n - 7} \right| \leq \frac{12}{n-7}, \quad \text{als } n > 7. \tag{2.2}$$

Dit is kleiner dan ε als $n > 7 + 12/\varepsilon$. We kiezen dus $N = 7 + 12/\varepsilon$.

Bij **Example 3 (e), p.35**. Met calculus vind je eenvoudig:

$$\begin{aligned} \lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x &= \lim_{y\to 0} (1+y)^{\frac{1}{y}} \\ &= \lim_{y\to 0} e^{\frac{\log(1+y)}{y}} = e^{\lim_{y\to 0} \frac{\log(1+y)}{y}} = e. \end{aligned} \tag{2.3}$$

We gebruikten de substitutie $y = 1/x$ en herkenden het differentiaalquotient in de laatste limiet. Het manipuleren met limieten zal in **Chap. 3** gerechtvaardigd worden. **Example 3e, p.35** volgt.

Je kunt deze limiet ook als *definitie* van e nemen.

2.2 Opgaven bij paragraaf 7 en 8

2.3. Geef bewijzen voor de limieten in **7.3.j, 7.3.r, 7.3.t.**

2.3 Bij paragraaf 9

Bij **Basic Examples 9.7**. Bij deze *standaardlimieten*, die je uit het hoofd moet weten(!), horen er nog enkele. Sommigen zijn in Calculus behandeld.

1. De limiet 2.3 $\lim_{n\to\infty}(1+1/n)^n = e$;
2. $\forall p > 0 :\ \lim_{n\to\infty} n^{-p}\log n = 0$;
3. $\forall p \in \mathbb{R} :\ \lim_{n\to\infty} n^p e^{-n} = 0$.

Bij **Example 5**. In het voorbeeld wordt weer een precieze M bepaald, dat is niet nodig. Bijvoorbeeld Zij $M > 0$. Dan

$$\sqrt{n} + 7 > \sqrt{n} > M, \quad \text{voor } n > N = M^2.$$

M.a.w. $\lim_{n\to\infty} \sqrt{n} + 7 = \infty$.

2.4 Bij paragraaf 10

Een belangrijke stelling is **Stelling 10.2** omdat deze je in staat stelt te bewijzen dat een rij convergeert zonder met de definitie van convergentie te werken. Je kunt dit gebruiken in opgave 10.8 t/m 10.12.

In dit soort opgaven wordt een rij inductief gedefinieerd. De eerste paar a_j zijn gegeven en verder is gegeven hoe a_{n+1} afhangt van de vorige termen. Dan kan heel eenvoudig zijn:

$$a_{n+1} = a_n/2 + 1, \tag{2.4}$$

of ingewikkeld:

$$a_{n+1} = a_n^2/4 + a_{n-1}/3 + 3n/(8n+7). \tag{2.5}$$

De vraag is meestal te bewijzen dat de limiet bestaat, en deze te bepalen. Het stramien van de oplossing is als volgt.

Stap 0. Bepaal de mogelijke limieten. Noem $L = \lim_{n\to\infty} a_n$. Je weet niet dat die limiet bestaat, maar we nemen even aan dat we dat al bewezen hebben.

Dan ook $L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} a_{n-1}$. Hoe je dit gebruikt laat het volgende voorbeeld zien. Er zijn vaak maar eindig veel waarden van L mogelijk. Onderzoek nu of de rij stijgend of dalend is. In het stijgende geval ga je nu:

Stap 1. *Bewijzen* dat de rij stijgend (althans niet dalend) is. Inductie kan daarbij helpen.

Stap 2. *Bewijzen* dat de rij naar boven begrensd is. Vaak is het handig voor de bovengrens de waarde van de limiet die het meest waarschijnlijk is, te kiezen. Ook hier kan inductie van pas komen.

Stap 3. Met stap 1 en 2 kun je **Stelling 10.2** gebruiken. Dan weet je dat de limiet bestaat. Vaak is maar één waarde mogelijk, bijvoorbeeld omdat de andere waarden al kleiner zijn dan de beginwaarde, of groter dan de bovengrens die je bewezen hebt. Daarmee is dan de opgave opgelost.

Voorbeeld 2.4. De rij gedefinieerd door (2.5) met $a_1 = 0, a_2 = 0$.

Stap 1. Bepalen van de mogelijke limiet(en). Stel dat de limiet bestaat en noem deze L. Dan

$$\begin{aligned} L &= \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \left(a_n^2/4 + a_{n-1}/3 + 3n/(8n+7)\right) \\ &= \lim_{n\to\infty} a_n^2/4 + \lim_{n\to\infty} a_{n-1}/3 + \lim_{n\to\infty} 3n/(8n+7) = L^2/4 + L/3 + 3/8. \end{aligned} \tag{2.6}$$

Hier staat een vierkantsvergelijking in L met oplossingen $L = \frac{8\pm\sqrt{10}}{6}$.

Uit (2.5) vind je $a_3 = 6/23$, $a_4 > 9/31 > 6/23$. Je krijgt het idee dat de rij monotoon stijgend is vanaf a_2, en alle waarden tot nu liggen onder $L = \frac{8-\sqrt{10}}{6}$. Dat is een goede gok voor een bovengrens.

Stap 2. Bewijs dat de rij monotoon niet dalend is. Dit gaat met inductie. We weten dat $a_1 = a_2 = 0$. Stel nu dat de $a_{j-1} \leq a_j$ voor $j \leq n$ $(n \geq 2))$. Dan is

$$a_{n+1} = a_n^2/4 + a_{n-1}/3 + 3n/(8n+7) \geq a_{n-1}^2/4 + a_{n-2}/3 + 3(n-1)/(8(n-1)+7) = a_n$$

want $a_{n-1} \leq a_n$, $a_{n-2} \leq a_{a_n-1}$, en $3(n-1)/(8(n-1)+7) \leq 3n/(8n+7)$. Dus de rij is niet dalend.

Stap 3. Bewijs dat de rij naar boven begrensd is. Hier moet je niet zomaar een bovengrens kiezen. Je ziet zo dat als je probeert te bewijzen dat alle termen kleiner dan 501 zijn door de inductie aanname $a_n < 501$, dan kan $a_n = 500$ en zal $a_{n+1} \geq 25000/4$ dus dat werkt niet. Handig is de waarschijnlijke limiet $L = \frac{8-\sqrt{10}}{6}$ te proberen als bovengrens. Dan gaat het met inductie zo: a_1 en a_2 zijn kleiner dan L. Stel dat $a_j \leq L$ voor $j \leq n$ $(n \geq 2))$. Dan is

$$a_{n+1} = a_n^2/4 + a_{n-1}/3 + 3n/(8n+7) \leq L^2/4 + L/3 + 3/8 = L.$$

We hebben bewezen dat alle termen $\leq \frac{8-\sqrt{10}}{6}$.

Er volgt nu met **Stelling 10.2** dat de rij (a_n) een limiet heeft en omdat alle $a_n \leq \frac{8-\sqrt{10}}{6}$ is $\frac{8+\sqrt{10}}{6}$ niet de limiet, dus $\lim_{n\to\infty} a_n = \frac{8-\sqrt{10}}{6}$.

Opmerkingen 2.5. De methode is ook bruikbaar als de a_n om een limiet heen slingeren, bijvoorbeeld (a_{2n}) is stijgend en a_{2n+1} dalend, dan pas je de methode op die stijgende en dalende deelrijen toe.

Het bewijs dat de rij monotoon en begrensd is, kan best lastiger zijn.

2.5 *Opgaven bij paragraaf 10*

2.6. Zij $a > 0$ en $x_0 > 0$. Voor $n \geq 0$ definieren we $x_{n+1} = \frac{x_n^2+a}{2x_n}$ Bewijs dat de rij (x_n) convergeert en bepaal de limiet.

Opmerking 2.7. De rij (x_n) is als volgt tot stand gekomen. Beschouw de functie $f(x) = x^2 - a$. Gegeven het punt x_n is x_{n+1} het snijpunt van de x-as met de raaklijn aan de grafiek van f in het punt $(x_n, f(x_n))$. Je ziet snel uit een tekening, dat wanneer je met een positieve x_0 start, de rij (x_n) naar het snijpunt van de grafiek van f met de positieve x-as convergeert. Deze manier om numeriek een nulpunt van een functie te vinden, door de a_n net zo lang uit te rekenen, tot je ziet dat er(bijna) geen verandering meer in optreedt, heet de methode van *Newton-Raphson*. De methode werkt vaak, maar niet altijd, en voor vrij algemene functies. Bovendien is het vaak een snelle methode.

2.8. De rij (a_n) is gegeven door $a_1 = \frac{1}{2}$, $a_{n+1} = \sqrt{\frac{(1+a_n^2)}{2}}$, $(n \geq 1)$. Laat zien dat de rij convergeert en bepaal de limiet.

2.9. De rij (a_n) is gegeven door $a_1 = 3$, $a_{n+1}^3 = 3a_n + 2$, $(n \geq 1)$. Onderzoek of de rij convergeert. Zo ja, bepaal de limiet.

2.10. Zij $a_n = (1 + 1/n)^n$. Laat zien (zonder te gebruiken dat je weet dat de limiet bestaat) dat de rij a_n naar boven begrensd is en stijgt.

2.11. (tentamen 2010) Gegeven is een begrensde rij (s_n). Laat

$$X := \{x \in \mathbb{R} : \text{Er zijn maar eindig veel termen van de rij met } s_n \leq x.\}.$$

Geef de definitie van $\sup X$ en van $\liminf_{n\to\infty} s_n$. Bewijs dat $\sup X = \liminf_{n\to\infty} s_n$.

2.6 Bij paragraaf 11

Theorem 11.3 en de stelling van Bolzano-Weierstrass (**Theorem 11.5**) worden vaak wat anders behandeld. Eerst een ander bewijs van de stelling van Bolzano-Weierstrass.

2.6.1 Stelling. Iedere begrensde rij reële getallen $(c_n)_n$ heeft een deelrij die convergeert.

Bewijs. Omdat de rij begrensd is zijn er een $a_0, b_0 \in \mathbb{R}$ zodat voor iedere n geldt $c_n \in [a_0, b_0]$. Nu zal voor tenminste één van de intervallen $[a_0, \frac{a_0+b_0}{2}]$ en $[\frac{a_0+b_0}{2}, b_0]$ gelden dat er oneindig veel van de c_n in bevat zijn. Noem dit interval $[a_1, b_1]$. Weer zal voor tenminste één van de intervallen $[a_1, \frac{a_1+b_1}{2}]$ en $[\frac{a_1+b_1}{2}, b_1]$ gelden dat er oneindig veel van de c_n in bevat zijn. We gaan zo door en vinden een rij gesloten intervallen $[a_m, b_m]$ met de volgende eigenschappen.

- Voor alle $m = 1, 2, \ldots$ $[a_m, b_m] \subset [a_{m-1}, b_{m-1}]$.
- Ieder interval $[a_m, b_m]$ bevat oneindig veel van de c_n.
- $b_m - a_m = 2^{-m}(b_0 - a_0)$.
- $\forall p, q \in \mathbb{N} \;\; a_p < b_q$.

Merk nu op dat de rij getallen (a_m) naar boven begrensd is, en niet dalend, dus wegens **Theorem 10.2**, $\lim_{n\to\infty} a_n = a$ bestaat, terwijl de rij $(b_n)_n$ naar beneden begrensd is en niet stijgt, dus $\lim_{n\to\infty} b_n = b$ bestaat ook. Verder is $b - a = \lim_{n\to\infty} b_n - a_n = 0$. Maak nu een deelrij c_{m_j} als volgt met inductie. Kies m_0. Stel dat m_k gekozen is, kies dan $m_{k+1} > m_k$ zodanig dat $c_{m_k} \in [a_k, b_k]$. Dit kan want er liggen oneindig veel c_n in het interval $[a_k, b_k]$. De rij (c_{m_k}) convergeert naar a vanwege de insluitstelling (**Exercise 8.5)** $\square$

2.7 Opgaven bij paragraaf 11

2.12. Laat zien dat een convergente rij een monotone deelrij heeft.

2.13. Laat (a_n) een begrensde rij zijn. Bewijs rechtstreeks uit de definitie van lim sup en het bestaan van suprema in $\mathbb{R}$, dat er een deelrij a_{n_j} is met $\lim_{j\to\infty} a_{n_j} = \limsup_{n\to\infty} a_n$.

2.8 Opgaven bij Paragraaf 14

2.14. Onderzoek of de volgende reeksen convergeren of divergeren.

$$a \quad \sum_{n=1}^{\infty} \frac{1}{n\sqrt[n]{n}}, \qquad b \quad \sum_{n=1}^{\infty} (-1)^n \left(\sqrt{1+\log n} - \sqrt{\log n}\right).$$

2.15. Van de reeks $\sum_{n=1}^{\infty} a_n$ is gegeven dat $\lim_{n\to\infty} a_n = 0$. Van de reeks partieelsommen (s_N), $s_N = \sum_{n=1}^{N} a_n$, is gegeven dat $\lim_{N\to\infty} s_N s_{N+1} = p$ voor zekere $p > 0$.

- Laat zien dat de rij (s_N) begrensd is;
- Bewijs dat $\sum_{n=1}^{\infty} a_n$ convergeert;
- Wat kun je zeggen over de som van de reeks?

2.16. Van een rij positieve, reële getallen (a_n) is gegeven dat $\sum_1^{\infty} a_n^2$ convergeert. Bewijs:

$$\sum_1^{\infty} a_n \text{ convergeert} \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} (\sqrt{1+a_n} - 1) \text{ convergeert.}$$

2.9 Bij paragraaf 15

Een nuttige techniek bij het onderzoeken van convergentie van reeksen is het *Partieel Sommeren.* Dit proces is nauw verwant aan het bekende partieel integreren Ingredienten zijn een rij $(a_n)_1^N$, een rij $(b_n)_1^N$ en de rij van partieelsommen $s_n = a_1 + \ldots a_n$, terwijl $s_0 = 0$. Er geldt.

2.17 Lemma.

$$\sum_{n=1}^{N} a_n b_n = \sum_{n=1}^{N-1} s_n(b_n - b_{n+1}) + s_N b_N - s_0 b_1. \tag{2.7}$$

Bewijs.

$$\begin{aligned}\sum_{n=1}^{N} a_n b_n &= \sum_{n=1}^{N} (s_n - s_{n-1}) b_n = \sum_{n=1}^{N} s_n b_n - \sum_{n=0}^{N-1} s_n b_{n+1} \\ &= \sum_{n=1}^{N-1} s_n(b_n - b_{n+1}) + s_N b_N - s_0 b_1.\end{aligned} \tag{2.8}$$

□

2.9.1 Stelling (Alt Theorem 15.3 (Leibnitz))**.** Als (c_n) monotoon niet stijgend is naar 0, dan convergeert $\sum_{n=1}^{\infty} (-1)^n c_n$.

Bewijs. We sommeren partieel met $a_n = (-1)^n$ en $b_n = c_n$. We vinden s_n is 0 of -1 al naar gelang n even of oneven is. Dan

$$\sum_{n=1}^{N} (-1)^n c_n = \sum_{n=1}^{N-1} s_n(c_n - c_{n+1}) + s_N c_N - s_0 b_1. \tag{2.9}$$

Merk nu op dat $\lim_{N\to\infty} s_N c_N = 0$, omdat s_N begrensd is. Verder is de tweede som in (2.9) absoluut convergent. Inderdaad, de partieelsommen van de absolute waarden zijn begrensd:

$$\sum_{n=1}^{N-1} |s_n(c_n - c_{n+1})| < \sum_{n=1}^{N-1} (c_n - c_{n+1}) = c_1 - c_N \leq c_1. \tag{2.10}$$

We gebruikten dat (c_n) monotoon niet stijgend is. De conclusie is dat het rechterlid in (2.9) convergeert als n naar oneindig gaat, en het linkerlid dus ook. □

2.10 Opgaven bij paragraaf 15

2.18. Onderzoek of de reeks
$$\sum_{n=4}^{\infty} \frac{(-1)^n}{\log \log n}$$
convergeert.

2.19. Bewijs de stelling van Dirichlet:

2.10.1 Stelling. Laat $(a_n)_n$ een monotoon niet stijgende rij positieve getallen zijn met limiet 0 en $(b_n)_n$ een rij getallen in $\mathbb{R}$ (of $\mathbb{C}$) met begrensde partieelsommen, i.e.
$$\exists M > 0 \forall N \in \mathbb{N} |s_N| = \left| \sum_{n=1}^{N} b_n \right| < M.$$
Dan convergeert $\sum_{n=1}^{\infty} a_n b_n$.

Pas het idee van het alternatieve bewijs van de Stelling van Leibnitz toe.

2.20.
- Laat z een complex getal zijn met $|z| = 1$, Bereken $\sum_{n=1}^{N} z^n$ en laat zien dat als $z \neq 1$
$$\left| \sum_{n=1}^{N} z^n \right| \leq \frac{2}{|1-z|}.$$
- Laat zien dat voor $|z| = 1$, $z \neq 1$ de som $\sum_{n=1}^{\infty} \frac{z^n}{n}$ convergeert.
- Leidt hieruit af dat voor alle $\theta \in \mathbb{R}$ $\sum_{n=0}^{\infty} \frac{\sin(n\theta)}{n}$ convergeert.

3 Continuïteit

3.1 Bij paragraaf 18

3.1. Laat zien dat er geen continue functie op $\mathbb{R}$ bestaat die elke waarde in $\mathbb{R}$ precies twee keer aanneemt. Aanwijzing: Veronderstel dat f zo'n functie is en laat $c \in \mathbb{R}$. Dan zijn er x_1, x_2 met $f(x_1) = f(x_2) = c$. Laat zien dat f op het interval $[x_1, x_2]$ precies één extreme waarde M (maximum of minimum) aanneemt in precies één punt $x_3 \in (x_1, x_2)$. Laat x_4 het tweede punt zijn met $f(x_4) = M$. Beschouw nu f op het interval dat van x_3 naar x_4 loopt.

3.2. Construeer (=teken) een functie op $\mathbb{R}$ die elke waarde precies 3 keer aanneemt.

3.2 Bij paragraaf 20

Limiet is een basisbegrip, meer nog dan continuïteit. **Definition 20.1** is een wat ongelukkige keus. Op pagina 151 vóór **Theorem 20.6** wordt vermeld dat deze definitie kan worden herschreven waarbij het gebruik van rijen wordt vermeden. Omdat in algemenere situaties rijen helemaal niet gebruikt kunnen worden om limieten te definiëren, kan men beter meteen maar de volgende definitie geven.

Definitie 3.3. Zij f gedefinieerd op $S \subset \mathbb{R}$ en $a \in \mathbb{R}$. Zij $L \in \mathbb{R}$. We zeggen

$$\lim_{\substack{x \to a \\ x \in S}} f(x) = L$$

als $\forall \varepsilon > 0 \; \exists \delta > 0$ zodanig dat

$$x \in S \text{ en } 0 < |x - a| < \delta \text{ impliceert } |f(x) - L| < \varepsilon.$$

Het is hierbij belangrijk dat wat er met f in $x = a$ gebeurt, geen rol speelt. Het kan dat $f(a)$ niet gedefinieerd is of een andere waarde heeft dan de limiet.

Voorbeeld 3.4.

- $f(x) = 0$ op $\mathbb{R} \setminus \{0\}$. Dan $\lim_{x \to 0} f(x) = 0$.
- Nu definiëren we f op $\mathbb{R}$ door $f(x) = 0$ op $\mathbb{R} \setminus \{0\}$ en $f(0) = 23,7589$. Dan $\lim_{x \to 0} f(x) = 0$.

Opmerking 3.5. Het boek is er nogal slordig mee het punt $x = a$ uit te sluiten.

1. **Definition 20.1**. Men moet $a \in \mathbb{R} \setminus S$ kiezen, of op andere wijze uitsluiten dat $x_n = a$ voor sommige n.
2. **Theorem 20.4** In het bewijs, weer vermijden $x_n = a$.
3. **Theorem 20.5** Idem
4. **Theorem 20.6** Zelfde opmerking voor wat betreft de mogelijke waarden van x_n bovenaan pagina 152.

Het volgende lemma legt verband met continuïteit.

3.6 Lemma. Zij f gedefinieerd op $S \subset \mathbb{R}$ en $a \in \mathbb{R}$. Zij $L \in \mathbb{R}$. Equivalent zijn

$$\lim_{\substack{x \to a \\ x \in S}} f(x) = L \tag{3.1}$$

en de functie $\tilde{f}$ gedefinieerd op $S \cap a$ door

$$\tilde{f}(x) = \begin{cases} L, \text{ als } x = a; \\ f(x) \text{ als } x \in S \setminus \{a\} \end{cases} \tag{3.2}$$

is continu in a.

Bewijs. Als $\tilde{f}$ continu is in a geldt $\forall \varepsilon > 0\ \exists \delta > 0$ zodanig dat

$$x \in S \text{ en } |x - a| < \delta \text{ impliceert } |\tilde{f}(x) - \tilde{f}(a)| < \varepsilon.$$

Met het oog op de definitie van $\tilde{f}$, (3.2), geldt dan ook $\forall \varepsilon > 0\ \exists \delta > 0$ zodanig dat

$$x \in S \text{ en } 0 < |x - a| < \delta \text{ impliceert } |f(x) - L| < \varepsilon.$$

Omgekeerd, als (3.1) geldt is $\tilde{f}$ volgens (3.2) ondubbelzinnig gedefinieerd. Er geldt dan

$$x \in S \text{ en } 0 < |x - a| < \delta \text{ impliceert } |\tilde{f}(x) - L| < \varepsilon.$$

En omdat $L = \tilde{f}(a)$ geldt ook

$$x \in S \text{ en } |x - a| < \delta \text{ impliceert } |\tilde{f}(x) - \tilde{f}(a)| < \varepsilon,$$

Dus $\tilde{f}$ is continu in a. □

Met dit Lemma zijn veel van de stellingen in paragraaf 20 eenvoudig uit corresponderende stellingen over continuïteit af te leiden.

Voorbeeld 3.7. Alternatief bewijs van **Theorem 20.5** Beschouw de functie $\tilde{f}$ gedefinieerd door $\tilde{f}(a) = L$, $\tilde{f}(x) = f(x)$ $(x \in S \setminus \{a\})$. Dan is $\tilde{f}$ continu in a volgens het lemma. We kunnen nu Theorem 17.5 toepassen. Dit geeft $g \circ \tilde{f}$ is continu in a. Met het lemma volgt dan

$$\lim_{\substack{x \to a \\ x \in S}} g \circ f(x) = \lim_{\substack{x \to a \\ x \in S}} g \circ \tilde{f}(x) = g \circ \tilde{f}(a) = g(L).$$

3.3 Bij paragraaf 21

In metrische ruimten kan een definitie van continuïteit makkelijk aan de hand van limieten gegeven worden. Eerst de definitie van limiet.

Definitie 3.8. Laten (X, d) en (Y, d') twee metrische ruimten zijn. Zij $a \in X$ en zij $f\colon X \setminus \{a\} \to Y$. We zeggen dat

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

voor zekere $b \in Y$ als er bij elke $\varepsilon > 0$ een $\delta > 0$ bestaat zo dat voor elke $x \in X$ geldt dat:

$$0 < d(x, a) < \delta \quad \Longrightarrow \quad d'(f(x), b) < \varepsilon. \tag{3.3}$$

In deze definitie kan f op X inclusief het punt a gedefinieerd zijn, terwijl $f(a) \neq b$. Als we echter continuïteit van f in a willen definiëren dan gaat het er juist om dat $f(a) = b$, want:

Definitie 3.9. Een afbeelding $f\colon X \to Y$ heet *continu* in a desda $\lim_{x\to a} f(x) = f(a)$. Verder heet f *continu* op X als f continu is in ieder punt van x. We noemen f continu op X als f in ieder punt van x continu is.

In de opgaven laat je zien dat dit equivalent is met de definitie van continuïteit in **Definition 21.1**. Deel 1 van de volgende stelling is **Theorem 20.5** voor metrische ruimten. Deel 3 bewijs je makkelijk direct met **Theorem 21.3**.

3.3.1 Stelling. Laat (X, d_1), (Y, d_2), $Z, d_3)$ metrische ruimte zijn. Veronderstel $f\colon X \to Y$, $g\colon Y \to Z$ en $a \in X$ gegeven

1. Veronderstel dat $\lim_{x\to a} f(x) = y \in Y$ bestaat, en dat g continu is in y. Dan bestaat $\lim_{x\to a} g \circ f(x)$ en deze limiet is gelijk aan $g(y)$.
2. Als f continu in a en g continu in $f(a)$ dan is $g \circ f$ continu in a.
3. Als f en g beide continu zijn op X respectievelijk Y, dan is $g \circ f$ continu.

Bewijs. Voor 1 kopieert men het bewijs van **Theorem 20.5** 2 volgt direct uit 1 en 3 direct uit 2. □

Opmerking 3.10. Het zal nu duidelijk zijn dat het begrip *metriek* gemodelleerd is op de absolute waarde in $\mathbb{R}$; $d(x, y) = |x - y|$ is het leidende voorbeeld. Stellingen over continuïteit in analyse op de lijn hebben een analogon voor metrische ruimte als alleen positiviteit, symmetrie en driehoeksongelijkheid in het bewijs gebruikt worden. Je kunt als het ware een bewijs in de metrische settting vinden, door domweg $|x - y|$ door $d(x, y)$ te vervangen.

3.4 Opgaven bij paragraaf 21

3.11. Veronderstel dat f_1, f_2 functies van een metrische ruimte (S, d) naar $\mathbb{R}^n$ met de gebruikelijke metriek zijn. Zij $a \in S$ en veronderstel dat $\lim_{x\to a} f_1(x) = L_1$, $\lim_{x\to a} f_2(x) = L_2$. Bewijs in deze situatie de uitspraken van **Theorem 20.4**.

Concludeer hieruit dat als f_1 en f_2 continu zijn, ook $f_1 + f_2$ en $f_1 f_2$ continu zijn, en verder dat f_1/f_2 continu is op $\{x : f_2(x) \neq 0\}$.

3.12. Bewijs Uitspraak 3 in Stelling 3.3.1 met behulp van **Theorem 21.3**

3.13. Een punt a heet een *verdichtingspunt* van een metrische ruimte (X, d) als voor iedere bal $B(a, r)$ geldt, dat $X \cap (B(a, r) \setminus \{a\}) \neq \emptyset$. Een punt $a \in X$ heet geïsoleerd als a niet een verdichtingspunt van X is.

Zij nu $f : X \setminus \{a\} \to Y$ een afbeelding van de metrische ruimte (X, d) naar (Y, d'). Bewijs: Als a een verdichtingspunt is van X dan is er ten hoogste één punt $b \in Y$ met $\lim_{x\to a} f(x) = b$. Bewijs vervolgens: Als a een geïsoleerd punt is van X, dan geldt $\lim_{x\to a} f(x) = b$ voor alle $b \in Y$.

4 Functierijen en –reeksen

4.1 Opgaven bij paragraaf 24

4.1. Zij
$$f_n(x) = \cos^n(x), \quad x \in \mathbb{R}, \quad n = 1, 2, \ldots.$$
Bepaal $\lim_{n\to\infty} f_n(x)$ en onderzoek of de rij (f_n) uniform op $\mathbb{R}$ convergeert.

4.2. Zij
$$f_n(x) = n \log(1 + x/n), \quad x \geq 0, \quad , n = 1, 2, \ldots.$$

1. Laat zien dat $\lim_{n\to\infty} f_n(x) = x$.
2. Laat zien dat de convergentie uniform is op ieder interval $[0, C]$, $C > 0$
3. Is de convergentie uniform op $[0, \infty)$?

4.3. Zelfde vragen als in opgave 4.2 voor de functies $g_n(x) = (1 + x/n)^n$.

4.2 Bij paragraaf 25

Eerst een resultaat voor alternerende functiereeksen, verwant aan de Stelling van Leibnitz over alternerende getallenrijen.

4.4 Propositie (Koornwinder, Analyse 2b)**.** Zij $(f_n)_{n=1}^{\infty}$ een functierij op $X \subset \mathbb{R}$ zo dat
$$f_n(x) \geq f_{n+1}(x) \geq 0 \quad (n \in \mathbb{N}, x \in X) \tag{4.1}$$
en $\lim_{n\to\infty} f_n = 0$ uniform op X. Dan convergeert de reeks $\sum_{k=1}^{\infty}(-1)^k f_k$ uniform op X.

Bewijs. Laat $n, m \in \mathbb{N}$ zo dat $m > n$. Met volledige inductie naar $m - n$ volgt er uit (4.1) dat
$$0 \leq (-1)^{n+1} \sum_{k=n+1}^{m} (-1)^k f_k(x) \leq f_{n+1}(x) \quad (x \in X).$$
Dus
$$\sup_{x\in X} \left| \sum_{k=n+1}^{m} (-1)^k f_k(x) \right| \leq \sup_{x\in X} |f_{n+1}|.$$
Met andere woorden, de reeks voldoet aan het Cauchy criterium **Theorem 25.6** voor uniforme convergentie. □

Vaak zul je moeten onderzoeken of een rij van functies puntsgewijs of uniform convergeert. Als men niet eenvoudig kan zien waar er sprake is van puntsgewijze convergentie, is onderzoeken van uniforme convergentie onbegonnen werk. Het volgende stappenplan voor onderzoek van uniforme convergentie van een rij $(f_n)_n$ op een deelverzameling X van $\mathbb{R}$, (of met het oog op de latere theorie $X \subset \mathbb{R}^n$ of X deel van een metrische ruimte) is ontleend aan de cursus analyse 2b van Prof T.H. Koornwinder.

1. Onderzoek of er puntsgewijze convergentie is. Als de reeks niet puntsgewijs convergeert, zal hij zeker niet uniform convergeren. Als er wel puntsgewijze convergentie is, ga naar 2.

2. Probeer de $\lim_{n\to\infty} f_n(x) =: f(x)$ te berekenen. Lukt dit naar 3 anders naar 8.

3. Bereken $g_n(x) := |f(x) - f_n(x)|$. Probeer $M_n := \sup_{x\in X} g_n(x)$ expliciet uit te rekenen. Ga naar 4 als dit lukt, ga anders naar 5 of 6 of 7.

4. Ga na of $\lim_{n\to\infty} M_n = 0$. Zoja, dan uniforme convergentie, zonee, dan geen uniforme convergentie.

5. Probeer getallen a_n te vinden zo dat $g_n(x) \leq a_n$ en $a_n \to 0$ voor $n \to \infty$. Als dit lukt dan is er uniforme convergentie

6. Ga over naar een deelverzameling X_1 van X, waarop M_n wel expliciet uitgerekend kan worden. Als daar niet $\lim_{n\to\infty} M_n = 0$ dan geldt er ook geen uniforme convergentie op X.

7. Probeer een rij $x_1, x_2, \ldots$ in X en $n_1 < n_2 < \ldots$ in $\mathbb{N}$ te vinden zo dat $g_{n_k}(x_k) \geq \varepsilon > 0$ voor $k = 1, 2, \ldots$. Dan geldt er geen uniforme convergentie op X.

8. Probeer met behulp van de uniforme Cauchy-voorwaarde na te gaan of er uniforme convergentie geldt.

Voor onderzoek van uniforme convergentie van een reeks van functies $\sum_{k=1}^{\infty} f_k$ kunnen analoge stappen worden doorlopen. Toch is er genoeg verschil met het rijgeval om de stappen voor het reeksgeval apart te vermelden:

1. Convergeert de functiereeks $\sum_{k=1}^{\infty} f_k$ puntsgewijs op X? Zoja, ga naar 2, zonee dan geen uniforme convergentie.

2. Is de reeks alternerend en is aan de voorwaarden van Propositie 4.4 voldaan? Zo ja, dan uniforme convergentie, zonee, ga naar 3.

3. Probeer $\mu_n \stackrel{\text{def}}{=} \sup_{x\in X} |f_n(x)|$ expliciet uit te rekenen. Lukt dit, ga dan naar 4, ga anders naar 5.

4. Ga na of $\sum_{n=1}^{\infty} \mu_n < \infty$. Zoja, dan uniforme convergentie op grond van de Weierstrasstoets, ga anders naar 5.

5. Probeer de puntsgewijze sommen $g_n(x) \stackrel{\text{def}}{=} \sum_{k=n+1}^{\infty} f_k(x)$ expliciet uit te rekenen. Ga naar 6 als dit lukt, ga anders naar 8 of 12.

6. Probeer $M_n \stackrel{\text{def}}{=} \sup_{x\in X} g_n(x)$ expliciet uit te rekenen. Ga naar 7 als dit lukt, ga anders naar 8 of 9 of 10 of 11 of 12.

7. Ga na of $\lim_{n\to\infty} M_n = 0$. Zoja, dan uniforme convergentie, zonee, dan geen uniforme convergentie.

8. Probeer getallen b_n te vinden zo dat $|f_n(x)| \leq b_n$ en $\sum_{n=1}^{\infty} b_n < \infty$. Als dit lukt dan is er uniforme convergentie op grond van de Weierstrass-toets.

9. Probeer getallen a_n te vinden zo dat $g_n(x) \leq a_n$ en $a_n \to 0$ voor $n \to \infty$. Als dit lukt dan is er uniforme convergentie

10. Ga over naar een deelverzameling X_1 van X, waarop M_n wel expliciet uitgerekend kan worden. Als daar niet $\lim_{n\to\infty} M_n = 0$ dan geldt er ook geen uniforme convergentie op X.

11. Probeer een rij $x_1, x_2, \cdots$ in X en $n_1 < n_2 < \ldots$ in $\mathbb{N}$ te vinden zo dat $g_{n_k}(x_k) \geq \varepsilon > 0$ voor $k = 1, 2, \ldots$. Dan geldt er geen uniforme convergentie op X

12. Probeer met behulp van de uniforme Cauchy-voorwaarde na te gaan of er uniforme convergentie geldt.

4.3 Opgaven bij paragraaf 25

4.5. Bepaal

$$\lim_{n\to\infty} \int_0^{10} (1 + x/n)^n dx$$

en motiveer je antwoord.

4.6. Bepaal

$$\lim_{n\to\infty} \int_0^{10} \frac{(3 + \sin^n(x))n}{4n + \cos^n(x)} dx$$

en motiveer je antwoord.

4.7. Voor welke $x \in \mathbb{R}$ convergeert de reeks

$$\sum_{n=1}^{\infty} \frac{x^n}{1 + x^{2n}}?$$

Op welke intervallen in $\mathbb{R}$ is de reeks uniform convergent?

4.8. Bewijs dat de reeks

$$\sum_{n=1}^{\infty} (-1)^n \frac{x^2 + n}{n^2}$$

op ieder begrensd interval uniform convergeert. Laat ook zien de de reeks nergens absoluut convergeert.

4.9. Beschouw de reeks

$$\sum_{n=0}^{\infty} (1 - x)x^n$$

op het interval $[0, 1]$. Bepaal de somfunctie. Is de convergentie uniform op $[0, 1]$?

4.10. Beschouw voor $\alpha \in \mathbb{R}$ de reeks

$$\sum_{n=1}^{\infty} n^{\alpha} x^{2n} (1 - x)^2.$$

a Onderzoek voor iedere α waar de reeks convergeert.

b Onderzoek voor iedere α of er sprake is van uniforme convergentie op het interval $[-1, 0]$.

c Onderzoek vervolgens voor iedere α of er sprake is van uniforme convergentie op het interval $[0, 1]$.

4.11. Beschouw de reeks

$$\sum_{n=1}^{\infty} \frac{nt^2}{n^3t^3+1}$$

a Bewijs dat deze reeks puntsgewijs convergeert op het interval $[0, \infty)$.

b Bewijs dat deze reeks uniform convergeert op ieder interval $[\delta, \infty)$, $(\delta > 0)$.

c Bewijs dat deze reeks niet uniform convergeert op $[0, 1]$ (Bekijk in $t = 1/N$ de "Cauchy-moot" $\sum_{N+1}^{2N}$.)

4.12. Beschouw voor $\alpha > 0$ de reeks

$$\sum_{n=1}^{\infty} t^{\alpha} e^{-n^2 t}$$

a Bewijs dat deze reeks puntsgewijs convergeert op het interval $[0, \infty)$.

b Bewijs dat deze reeks uniform convergeert op ieder interval $[\delta, \infty)$, $(\delta > 0)$.

c Bewijs dat voor iedere $\alpha > 1/2$ de reeks uniform convergeert op $[0, \infty)$.

d Laat zien dat voor $0 < \alpha \leq 1/2$ de reeks **niet** uniform convergeert op $[0, \infty)$.

4.4 Bij paragraaf 26

Het bewijs van **Theorem 26.6** kan wat vereenvoudigd worden. Merk op dat in het bewijs in het boek, en in het bewijs hieronder de techniek van het partieel sommeren wordt gebruikt die in paragraaf 2.9 is geïntroduceerd.

4.4.1 Stelling (Abel's Stelling)**.** Zij $f(x) = \sum_{n=0}^{\infty} a_n x^n$ een machtreeks met convergentiestraal $R > 0$. Als de reeks convergeert in $x = R$, respectievelijk $x = -R$ dan is f continu in $x = R$, respectievelijk $x = -R$.

Bewijs. Als in het boek gaat het erom de stelling te bewijzen voor $R = 1$ en $x = 1$. Schrijf $f_n(x) = \sum_{k=0}^{n} a_k x^k$ voor $x \in (-1, 1]$ en $s_n = \sum_{k=0}^{n} a_k$ $(n = 0, 1, 2, \ldots)$, $s = \sum_{k=0}^{\infty} a_k = f(0)$. We definiëren $s_{-1} = 0$, zodat voor alle $n \geq 0$ geldt $a_n = s_n - s_{n-1}$.

Zonder verlies van algemeenheid mogen we aannemen dat $s = 0$. Inderdaad, bekijk

$$g(x) = f(x) - s = (a_0 - s) + a_1 x + a_2 x^2 + \cdots .$$

Als we de stelling voor g kunnen bewijzen, dan hebben we ook een bewijs voor f. We gaan nu partieel sommeren (dit gaat net als pag 197, eerste display).

$$\begin{aligned} f_n(x) &= \sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} (s_k - s_{k-1}) x^k \\ &= \sum_{k=0}^{n} s_k x^k - x \sum_{k=0}^{n} s_{k-1} x^{k-1} = \sum_{k=0}^{n} s_k x^k - x \sum_{k=0}^{n-1} s_k x^k \\ &= (1-x) \sum_{k=0}^{n-1} s_k + s_n x^n \end{aligned} \tag{4.2}$$

Nu zal voor $|x| < 1$ gelden $\lim_{n\to\infty} s_n x^n = 0$ (want s_n is begrensd. Het geldt ook voor $x = 1$ omdat $s_n \to 0$, maar dat hebben we niet nodig.) Er volgt

$$f(x) = (1-x)\sum_{n=0}^{\infty} s_n x^n.$$

We moeten nu laten zien dat $\lim_{x\uparrow 1} f(x) = 0$. Zij $\varepsilon > 0$. Kies N zo groot dat $|s_n| < \varepsilon/2$, voor $n > N$. Dan is

$$\begin{aligned}
|(1-x)\sum_{n=0}^{\infty} s_n x^n| &\leq \left|(1-x)\sum_{n=0}^{N} s_n x^n\right| + \left|(1-x)\sum_{n=N+1}^{\infty} s_n x^n\right| \\
&\leq |H_N(x)| + (1-x)\sum_{n=N+1}^{\infty} |s_n| x^n \leq |H_N(x)| + (1-x)\frac{\varepsilon}{2}\sum_{n=N+1}^{\infty} x^n \qquad (4.3)\\
&= |(H_N(x)| + (1-x)\frac{\varepsilon}{2}\frac{x^{N_0+1}}{1-x} \leq |H_N(x)| + \frac{\varepsilon}{2}.
\end{aligned}$$

Hier is $H_N(x) = (1-x)\sum_{n=0}^{N} s_n x^n$. Dit is een continue functie in x met $H_N(1) = 0$. Er is dus een $\delta > 0$ zodanig dat $1-\delta < x < 1$ impliceert $|H_N(x)| < \varepsilon/2$. Voor deze δ geldt dan ook dat $1-\delta < x < 1$ impliceert $|f(x)| < \varepsilon$. $\square$

5 Differentiëren

5.1 Bij paragraaf 28

Zij f een reëelwaardige functie die gedefinieerd is op een open interval dat het punt a bevat. Er zijn drie gelijkwaardige manieren om uit te drukken dat f *differentieerbaar* is in a met *afgeleide* $A =: f'(a)$.

Definitie 5.1. 1.
$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=A;$$

2.
$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=A;$$

3. Er bestaat een functie g gedefinieerd op een open interval om 0 met de eigenschap dat $\lim_{h\to 0}\frac{g(h)}{h}=0$ en
$$f(a+h)=f(a)+Ah+g(h). \tag{5.1}$$

Merk op dat (1) en (2) de afgeleide uitdrukken als limiet van de richtingscoëfficiënten van de rechte lijnen door de punten $(a, f(a))$ en $(b, f(b))$, terwijl (3) uitdrukt dat er een goede affiene benadering voor $f(a+h)$ is in de buurt van a, nl $f(a)+Ah$. Voordeel van (3) is dat er niet gedeeld wordt door h, de uitdrukking (3) heeft ook zin voor functies op $\mathbb{R}^n$. Dan is h een vector en A een matrix, de *totale afgeleide*.

5.2 Lemma. In bovenstaande definitie zijn (2) en (3) equivalent.

Bewijs.
$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=A$$
is equivalent met
$\forall\varepsilon>0$ exists $\delta>0$ zodanig dat
$$z(h)=\left|\frac{f(a+h)-f(a)-Ah}{h}\right|<\varepsilon \quad \text{for } 0<|h|<\delta.$$

Neem nu $g(h)=|h|z(h)$, en (2) impliceert (3) is bewezen.

Omgekeerd impliceert (3) dat
$$\lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{h}=0$$
ofwel
$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=A.$$
Dit is (2). □

We geven een ander bewijs van de kettingregel gebaseerd op (3).

5.1.1 Stelling. Zij f_1 differentieerbaar in a en f_2 differentieerbaar in $f_1(a)$. Dan is $f_2 \circ f_1$ differentieerbaar in a en $(f_2 \circ f_1)'(a) = f_2'(f_1(a))f_1(a)$.

Bewijs. Volgens (3) moeten we bewijzen dat

$$f_2 \circ f_1(a+h) - f_2 \circ f_1(a) = f_2'(f_1(a))f_1(a)h + g(h), \tag{5.2}$$

waar $\lim_{h\to 0} g(h)/h = 0$. Het linkerlid is van de vorm $f_2(w+t) - f_2(w)$, met $w + t = f_1(a+h)$, $w = f_1(a)$ en dus $t = f_1(a+h) - f_1(a)$. We vinden dan, met gebruik van (5.1) voor f_2 en voor f_1 en bijbehorende g_1, g_2

$$\begin{aligned} f_2 \circ f_1(a+h) - f_2 \circ f_1(a) &= f_2'(f_1(a))(f_1(a+h) - f_1(a)) + g_2(f_1(a+h) - f_1(a)) \\ &= f_2'(f_1(a))(f_1'(a)h + g_1(h)) + g_2(f_1(a))(f_1'(a)h + g_1(h)) \\ &= f_2'(f_1(a))f_1'(a)h + [f_2'(f_1(a))g_1(h) + g_2(f_1'(a)h + g_1(h))] \end{aligned} \tag{5.3}$$

De uitdrukking $g_{21}(h) = [\ldots]$ zal de g-term zijn voor $f_2 \circ f_1$. We moeten laten zien dat voor iedere $\varepsilon > 0$ er een δ is zo dat $|g_{21}(h)| < \varepsilon|h|$ als $|h| < \delta$. Met de eigenschappen van g_1 zien we in dat $\exists \delta_1$ met

$$|f_2'(f_1(a))g_1(h)| < \varepsilon|h|/2, \tag{5.4}$$

als $|h| < \delta$. Vervolgens $\exists \delta_3$ en $C > 0$ zodanig dat $|h| < \delta_3$ impliceert

$$|f_1'(a)h + g_1(h)| < C|h|. \tag{5.5}$$

Tenslotte $\exists \delta_2$ met $|\tilde{h}| < \delta_2$ impliceert $|g_2(\tilde{h})| \leq \varepsilon|\tilde{h}|/(2C)$,

Als nu $|h| < \min\{\delta_3, \delta_2/C\}$ vinden we met (5.5)

$$|g_2(f_1'(a)h + g_1(h)) \leq \frac{\varepsilon}{2C}|f_1'(a)h + g_1(h)| \leq \frac{\varepsilon}{2}|h|. \tag{5.6}$$

Combinatie van (5.4) en (5.6) levert $|g_{12}(h)| < \varepsilon|h|$ voor $|h| < \min\{\delta_1, \delta_3, \delta_2/C\}$. □

5.2 *Opgaven bij paragraaf 29*

5.3. Zij f een continue functie op $(-1,1)$ die differentieerbaar is buiten 0. Veronderstel dat $\lim_{x\to 0} f'(x) = A$ bestaat. Bewijs dat f differentieerbaar is in 0 en dat $f'(0) = A$. [Pas de middelwaardestelling toe].

5.3 *Bij paragraaf 31*

De stelling van Taylor **31.3** en het gevolg **31.4**is vaak nuttiger voor het berekenen van limieten dan de stelling van l'Hospital.

5.4 Lemma. Veronderstel dat f oneindig vaak differentieerbaar is op het open interval $(-a, b)$, $a, b > 0$. Dan is er voor iedere $n \in \mathbb{N}$ en ieder deelinterval $I = [-a_1, b_1] \subset (-a, b)$ een constante $C = C_{I,f,n} > 0$ zodat voor de n-de restterm in de Taylorreeks van f geldt

$$R_n(x)| \leq C|x|^n, \quad x \in I$$

Bewijs. Het is voldoende op te merken dat men met stelling **31.3** voor C kan nemen het maximum van $|f^n(y)/n!|$ op het interval I. Dit bestaat vanwege Stelling 18.1. □

Voorbeeld 5.5. Bereken
$$\lim_{x\to 0}\frac{1-\cos(x^7)}{x^{14}}$$
Als je niet zo slim bent om $x^7 = y$ te substitueren, moet je wel erg vaak differentiëren wanneer je dit met l'Hospital aanpakt. Met Taylor gaat het zo. We weten dat $\cos x = 1 - x^2/2 + R_4(x)$, Volgens Lemma 5.4 is er een constante C zodat $|R_4(x)| \le C|x|^4$ op een interval I om 0. Dan geldt dus ook $\cos(x^7) = 1 - x^{14}/2 + R_4(x^7)$. Er volgt
$$\lim_{x\to 0}\frac{1-\cos(x^7)}{x^{14}} = \lim_{x\to 0}\frac{1-1+x^{14}/2+R_4(x^7)}{x^{14}} = 1/2 + \lim_{x\to 0}\frac{R_4(x^7)}{x^{14}} = \frac{1}{2}.$$
Want $|R_4(x^7)| \le C|x|^{28}$ op I.

Een erg ingewikkeld voorbeeld (hierin zit wel alles wat je over dit onderwerp kunt leren)

Voorbeeld 5.6. Bereken
$$\lim_{x\to 0}\frac{\cos x - 1 + x^2/2}{e^x - 2 + \cos x - x - \sin^3 x/6}$$

We schrijven $\cos x = 1 - x^2/2 + x^4/(4!) + R_6(x)$; dan $e^x = 1 + x + x^2/2 + x^3/(3!) + x^4/(4!) + R_5(x)$. Voor $\sin^3$ ontwikkelen we eerst $\sin(x) = x + R_3(x)$. Dus $\sin^3 x = x^3 + 3xR_3^2(x) + 3x^2R_3(x) + R_3^3(x)$. Dan is
$$\frac{\cos x - 1 + x^2/2}{e^x - \cos x - x - \sin^3 x/6} = \frac{x^4/(4!) + R_6(x)}{2x^4/(4!) + R_5(x) + R_6(x) - \frac{3x^2R_3(x)+3xR_3^2(x)+R_3^3(x)}{6}}.$$
Dit ziet er erger uit dan het is. Merk op dat iedere $|R_j(x)| \le C_j|x|^j$. Deel teller en noemer door x^4 en laat $x \to 0$. Er volgt dat de limiet $1/2$ is.

5.4 Opgaven bij paragraaf 31

Bereken voor de volgende uitdrukkingen de limiet als $x \to 0$.

5.7.
$$\frac{\sin x - x\cos x}{x^3}$$

5.8.
$$\frac{e^{x^2} - 2\cos x + 1}{\sin^2 x^2}$$

5.9.
$$\frac{e^{2x} - e^{2x^2} - \sin 2x}{\sin x - x}$$

6 Grote O en kleine o symbool

Landau heeft in het begin van de vorige eeuw het *grote O symbool* en het *kleine o symbool* ingevoerd. Dit is een handig gereedschap om met resten om te gaan die bijvoorbeeld optreden bij de definitie van afgeleide of Taylorreeks. Vaak spreekt men van het grote en kleine o-symbool van Landau.

6.1 Het grote O-symbool

We werken in $\mathbb{R}^n$

Definitie 6.1. Laat f een functie zijn, gedefinieerd op (een deel van) $\mathbb{R}^n$, naar $\mathbb{R}^m$ en g een positieve functie van een deel van $\mathbb{R}^n$. We zeggen

$$f(x) = O(g(x)) \quad \text{voor } x \to a$$

als $f(x)/g(x)$ begrensd is in de buurt van a. Meer precies

- Als $a \in \mathbb{R}^n$ (en f, g gedefinieerd op een omgeving van a): $\exists C, \delta > 0$ met

$$\left| \frac{f(x)}{g(x)} \right| < C \quad \text{als } 0 < |x - a| < \delta.$$

- Als $a = \infty$ (en f, g gedefinieerd op een verzameling van de vorm $\{|x| > C\}$): $\exists C, M > 0$ met

$$\left| \frac{f(x)}{g(x)} \right| < C \quad \text{als } |x| > M.$$

Soms schrijft men ook $f(x) \in O(g(x))$ voor $x \to a$.

6.2 Opgaven O

6.2. In deze opgaven werken we in $\mathbb{R}^2$. Laat zien dat $\sin(x^2 + y^4) - x^2 = O(|(x,y)|^4)$ voor $(x,y) \to 0$.

6.3. Als voor functies f, g, h op $\mathbb{R}^n$ geldt $f(x) = O(|g(x)|)$, $g(x) = O(|h(x)|)$, dan is $f(x) = O(h(x)|)$, $(x \to a)$.

6.3 Het kleine o-symbool

We werken in $\mathbb{R}^n$

Definitie 6.4. Laat f een functie zijn, gedefinieerd op (een deel van) $\mathbb{R}^n$, naar $\mathbb{R}^m$ en g een positieve functie op een deel van $\mathbb{R}^n$. We zeggen

$$f(x) = o(g(x)) \quad \text{voor } x \to a$$

als $f(x)/g(x)$ naar 0 gaat in de buurt van a. Meer precies Als $a \in \mathbb{R}^n$ (en f, g gedefinieerd op een omgeving van a) of als $a = \infty$ (en f, g gedefinieerd op een verzameling van de vorm $\{|x| > C\}$)

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0.$$

6.4 Opgaven o

6.5. In deze opgave werken we in $\mathbb{R}^2$. Laat zien dat $\sin(x^2 + y^4) - x^2 = o(|(x,y)|^\pi)$ voor $(x,y) \to 0$.

6.6. Als voor functies f, g, h op $\mathbb{R}^n$ geldt $f(x) = O(|g(x)|)$, $g(x) = o(|h(x)|)$, dan is $f(x) = o(h(x)|)$, $(x \to a)$.

6.7. Als $f(x) = o((|x|)$ en $g(x) = O(|x|)$, dan $(f+g)(x) = O(|x|)$, $(x \to 0)$.

Deel II

Functies van meer veranderlijken

Inleiding

Dit deel van het college is gebaseerd op de syllabus Analyse B1 van Tom Koornwinder.

`http://staff.science.uva.nl/~thk/edu/oldsyl/analb1.pdf`

6.5 Inleiding Analyse B1

Nadat in het vak Analyse op de Lijn en in het voorafgaande deel van dit college analyse op $\mathbb{R}$ is behandeld, i.h.b. functies van één reële veranderlijke, zullen we ons nu bezighouden met analyse op $\mathbb{R}^n$, i.h.b. met differentieerbare afbeeldingen van een open deelverzameling van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Wanneer we proberen om in deze situatie de analoga van de vertrouwde definities en stellingen uit de analyse op $\mathbb{R}$ te formuleren, dan zullen we allerlei nieuwe fenomenen ontmoeten. Veel latere vakken zullen gebruik maken van het hier behandelde materiaal, bijv. analyse op variëteiten, differentiaalmeetkunde, functietheorie, Liegroepen, partiële differentiaalvergelijkingen. Ook voor de toegepaste wiskunde, die zich vaak in hogere dimensies afspeelt, is de hier ontwikkelde theorie van groot belang.

Voor een goed begrip van de stof moet je vertrouwd zijn met tot nu toe behandelde analyse en lineaire algebra.

Het vak Lineaire Algebra zal van belang zijn omdat lineaire afbeeldingen de eenvoudigste voorbeelden zijn van de afbeeldingen die we gaan bekijken, en omdat veel eigenschappen van onze niet-lineaire afbeeldingen teruggebracht kunnen worden tot eigenschappen van hun linearisaties.

De analyse op $\mathbb{R}^n$ heeft een meer meetkundig karakter dan de analyse op $\mathbb{R}$. Het tekenen van plaatjes (bijv. voor $n = 2$) zal daarom veel steun kunnen bieden. Terwijl plaatjes in deze syllabus (nog) niet voorkomen, zullen ze des te meer op college en werkcollege getekend worden. Bij het zelf bestuderen van de stof is het ook raadzaam om veelvuldig plaatjes te tekenen.

In een laatste hoofdstuk zullen een aantal commando's uit Maple besproken worden, Mathematica zal gelijksoortige commando's hebben. Enige vertrouwdheid met Maple wordt hier verondersteld. Computeralgebra kan je helpen bij het tekenen van grafieken van functies in twee veranderlijken.

Er zijn twee soorten opgaven. Tussen de gewone tekst vind je geregeld opgaven, die je ertoe aansporen om met een net ingevoerd begrip of bewezen stelling direct zelf aan de gang te gaan door bijv. een voorbeeld of tegenvoorbeeld te bestuderen of door een eenvoudig aanvullend resultaat te bewijzen. Aan het eind van elk hoofstuk vind je wat concretere vraagstukken: echte sommen. Deze laatste zullen op het werkcollege behandeld worden. Vraagstukken van de eerste soort zullen soms op het werkcollege behandeld worden, maar het zal ook voorkomen dat de docent reeds tijdens het hoorcollege zo'n vraagstuk in dialoog met de studenten behandelt of dat je aangespoord wordt om het zelf als "huiswerk" te beantwoorden. Hoe dan ook, het maken van de vraagstukken van de eerste soort is een goede manier om bij te blijven met de behandelde stof.

Op het tentamen zullen vraagstukken van beide types voorkomen, echter meer vraagstukken van de tweede, concrete soort dan van de eerste, theoretische soort.

De organisatie van deze syllabus is als volgt. Hij is opgebouwd uit een aantal hoofdstukken, die meestal weer opgedeeld zijn in een paar deelhoofdstukken. Binnen een hoofdstuk is er een paragraafnummering van de vorm $a.b$, waarbij a het hoofdstuknummer is en b het volgnummer van de paragraaf binnen dat hoofdstuk. Als er bijv. ergens verwezen wordt naar Stelling $a.b$, dan wordt de Stelling in paragraaf $a.b$ bedoeld. Een zelfde conventie geldt voor Definitie $a.b$, Voorbeeld $a.b$, Opgave $a.b$, etc. Aan het slot van elk hoofdstuk volgen een aantal vraagstukken die genummerd zijn als V$a.b$, waarbij a weer het hoofdstuknummer is en b het volgnummer van het vraagstuk binnen dat hoofdstuk.

De paragraafnummers en vraagstuknummers zijn in de regel vet gedrukt. Soms zijn ze echter cursief gedrukt. Dit betekent dat die paragraaf of dat vraagstuk niet tot de verplichte tentamenstof behoort. Hetzelfde geldt voor bewijzen. Als het woord "Bewijs" vet is gedrukt, dan behoort het tot de vaste stof; als het cursief is gedrukt, dan is het facultatief. Soms zal de docent de onderdelen met cursieve aanduidingen niet behandelen. Voor studenten die wat dieper op de stof willen ingaan zijn deze gedeelten uiteraard aanbevolen materiaal.

Het wordt ten zeerste aangeraden om naast de syllabus ook wat boeken te raadplegen. Hier volgt een kleine selectie van aanvullende literatuur.

[1] J. H. J. Almering, *Analyse* (geheel herzien door H. Bavinck en R. W. Goldbach), Delftse Uitgeversmaatschappij, 6e druk, 1990.

[2] T. M. Apostol, *Calculus, Vol. II*, John Wiley & Sons, Second ed., 1969.

[3] J. J. Duistermaat & J. A. C. Kolk, *Syllabus Analyse C*, Mathematisch Instituut der Rijksuniversiteit te Utrecht, 1992.

[4] R. A. Kortram & A. van Rooij, *Analyse, functies van meer veranderlijken*, Epsilon Uitgaven 16, 1990.

[5] K.A. Ross *elementary analysis: The theory of calculus* Springer,1980

[6] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, Third ed., 1976. (zeer aanbevolen)

[7] W. Walter, *Analysis II* (in het Duits), Springer, 1990.

Dankwoord Bij het schrijven van deze syllabus heb ik veel ontleend aan de vroegere syllabus Analyse C van prof. dr. D. van Dulst uit 1984. Dit betreft vooral de vraagstukken, maar ook de algemene opzet en een aantal details. Ook heb ik, direct of indirect via Van Dulst's syllabus, inspiratie opgedaan bij Rudin [5]. Ook zeg ik dank aan dr. ir. A. B. J. Kuijlaars voor kritische opmerkingen en voor het leveren van enige vraagstukken.

Tom Koornwinder

6.6 Inleidende bewerking voor Analyse: van $\mathbb{R}$ naar $\mathbb{R}^n$

De huidige bewerking bestaat uit het aanpassen van de verwijzingen naar het boek K.A. Ross, Elementary Analysis, dat nu in het eerste jaar gebruikt wordt, en kleine aanpassingen in de tekst die deze syllabus hopelijk beter laten aansluiten op eerdere colleges. Een grote wijziging is het vervangen van het hoofdstuk "Hogere totale afgeleide; Taylorreeks" door het hoofdstuk "Taylorreeks in gebruik", omdat het hoofdstuk "Hogere totale afgeleide: Taylorreeks" voor veel eerste-jaars te ambitieus is. Ook is het hoofdstuk Maple-commando's geschrapt.

Graag wil ik Wieger Hinderks en Jim de Groot bedanken. Wieger heeft de syllabus grondig doorgelezen en voorzien van nuttig commentaar, dat ik verwerkt heb. Ik verwacht dat dit de bruikbaarheid van de syllabus zal bevorderen. Jim heeft het manuscript omgezet naar LaTeXen de nummering van secties en stellingen doorzichtiger gemaakt.

Jan Wiegerinck

7 Genormeerde vectorruimtes en limieten

In dit hoofdstuk behandelen we materiaal over genormeerde vectorruimtes en over operatornormen dat we later in deze syllabus nodig zullen hebben. Ook bespreken we nog een paar zaken over limieten van functies, i.h.b. van functies op $\mathbb{R}^n$. Over dit laatste zullen de vraagstukken bij dit hoofdstuk ook handelen.

7.1 Genormeerde vectorruimtes

In Lineaire Algebra werd de definitie gegeven van een reële (of complexe) vectorruimte met inproduct. We weten dat met behulp van een inproduct op natuurlijke manier de lengte of *norm* van een vector kan worden geïntroduceerd, evenals de afstand tussen twee vectoren. Dit afstandsbegrip hebben we al op zichzelf bestudeerd. Het geeft aanleiding tot *metrische ruimten*. In deze paragraaf bestuderen we ook de norm op zichzelf. Dit geeft aanleiding tot *genormeerde vectorruimten*.

We herhalen eerst de definitie van ruimte met inproduct.

7.1 Definitie. Een (reëel) *inproduct* op een reële vectorruimte V is een afbeelding $(\,.\,,\,.\,)\colon V \times V \to \mathbb{R}$ zo dat

(i) $(v, w) = (w, v)$,

(ii) $(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1(v_1, w) + \alpha_2(v_2, w) \quad (\alpha_1, \alpha_2 \in \mathbb{R})$,

(iii) als $v \neq 0$ dan $(v, v) > 0$.

Een (reële) *inproductruimte* is een paar $(V, (\,.\,,\,.\,))$ met V een reële vectorruimte en $(\,.\,,\,.\,)$ een inproduct op V.

7.2 Voorbeeld. Beschouw de reële vectorruimte $\mathbb{R}^n$ $(n \in \mathbb{N})$ en schrijf een element x van $\mathbb{R}^n$ als

$$x = x_1 e_1 + \cdots + x_n e_n = (x_1, \ldots, x_n),$$

waarbij de vectoren $e_1, \ldots, e_n$ de standaardbasis van $\mathbb{R}^n$ vormen en $x_1, \ldots, x_n$ de coördinaten van x zijn. Dan wordt $\mathbb{R}^n$ een inproductruimte met het standaard-inproduct

$$\langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n \quad (x, y \in \mathbb{R}^n). \tag{7.1}$$

(Merk op dat we het inproduct in $\mathbb{R}^n$ met schuine haken schrijven.)

7.3 Definitie. Zij V een reële vectorruimte. Een afbeelding $\|\,.\,\|:\ V \to \mathbb{R}_{\geq 0}$, $v \mapsto \|v\|$ noemen we een *norm* op V, indien deze de volgende eigenschappen bezit

(N1) $||v|| \geq 0$ en $\|v\| = 0 \Leftrightarrow v = 0$,

(N2) Voor $t \in \mathbb{R}$ geldt $\|tv\| = |t|\|v\|$,

(N3) $\|v + w\| \leq \|v\| + \|w\|$ (*Driehoeksongelijkheid*).

Een *genormeerde vectorruimte* is een paar $(V, \|\,.\,\|)$ met V een vectorruimte en $\|\,.\,\|$ een norm op V.

7.4. Op een vectorruimte V met inproduct $(\,.\,,\,.\,)$ kan men altijd een *norm* $\|\,.\,\|$ definiëren door

$$\|v\| := \sqrt{(v,v)} \quad (v \in V). \tag{7.2}$$

7.5. In een inproductruimte geldt de belangrijke *ongelijkheid van Cauchy-Schwarz Bunjakowski*. Deze wordt gebruikt om te laten zien dat (7.2) inderdaad een norm op V definieert.

7.6 Propositie (Cauchy, Schwarz, Bunjakowski)**.** Zij V een reële vectorruimte en $(\,.\,,\,.\,)$ een inproduct op V. Dan geldt

$$|(x,y)| \leq \|x\|\|y\|, \quad x, y \in V. \tag{7.3}$$

Bewijs. Als x of y de nulvector is hoeven we niets te bewijzen. Voor alle $t \in \mathbb{R}$ geldt

$$0 \leq (x+ty, x+ty) = (x,x) + 2t(x,y) + t^2(y,y).$$

Deze uitdrukking is als functie van t een niet negatief kwadratisch polynoom. Voor de discriminant vinden we dus

$$(2(x,y))^2 - 4(x,x)(y,y) \leq 0.$$

Herschrijven geeft (7.3). □

Opgave 7.1. Laat hiermee zien dat (7.2) inderdaad een norm op V definieert.

In het vervolg zullen we een gegeven inproductruimte ook automatisch als genormeerde ruimte opvatten overeenkomstig (7.2). De norm op $\mathbb{R}^n$ afkomstig van het standaard inwendig product zullen we altijd met enkele strepen schrijven:

$$|x| := \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \cdots + x_n^2} \quad (x \in \mathbb{R}^n). \tag{7.4}$$

Opgave 7.2. Zij $(V, \|\,.\,\|)$ een genormeerde vectorruimte. Bewijs dat

$$\big|\, \|x\| - \|y\| \,\big| \leq \|x - y\| \quad (x, y \in V). \tag{7.5}$$

7.7 Voorbeeld. Op $\mathbb{R}^n$ zijn verschillende normen mogelijk die niet alle afkomstig zijn van een inproduct:

(i) De ∞-*norm* $\|v\|_\infty = \max\{|v_1|, \ldots, |v_n|\}$,

(ii) De *1-norm* $\|v\|_1 = \sum_{j=1}^n |v_j|$,

(iii) En meer algemeen, voor $1 \leq p \leq \infty$, de *p-norm*

$$\|v\|_p = \left(\sum_{j=1}^n |v_j|^p\right)^{1/p}. \tag{7.6}$$

De 2-norm correspondeert met de norm afkomstig van het standaard inproduct. De andere normen zijn **niet** van een inproduct afkomstig.

Opgave 7.3.

(a) Bewijs dat $(\mathbb{R}^n, \|\,.\,\|_1)$ een genormeerde vectorruimte is.

(b) Teken het gebied $\{x \in \mathbb{R}^2 \mid \|x\|_p < 1\}$ in de gevallen $p = 1$, 2 en ∞.

(c) Bewijs, voor $x \in \mathbb{R}^n$, de ongelijkheden

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\,\|x\|_2, \qquad \frac{1}{\sqrt{n}}\,\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2. \tag{7.7}$$

Opgave 7.4.

(a) Bewijs dat in een inwendig productruimte V met door het inproduct geïnduceerde norm $\|\,.\,\|$ de *parallellogramidentiteit* geldt:

$$||v + w||^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

(In woorden: in een parallellogram is de som van de kwadraten van de lengten van de diagonalen gelijk aan de som van de kwadraten van de lengten van de zijden.)

(b) Laat zien dat ∞-norm en de 1-norm op $\mathbb{R}^n$ ($n \geq 2$) niet van een inproduct afkomstig zijn.

7.8. Op een vectorruimte V heten twee normen $\|\,.\|$ en $\|\,.\|^*$ *equivalent* dan en slechts dan als er $\alpha, \beta > 0$ bestaan zo dat

$$\beta\,\|x\|^* \leq \|x\| \leq \alpha\,\|x\|^* \quad \text{voor alle } x \in V. \tag{7.8}$$

7.9 Stelling. Iedere twee normen op een eindig-dimensionale vectorruimte V zijn equivalent.

Bewijs. We kunnen V identificeren met $\mathbb{R}^n$ voor een of andere $n \in \mathbb{N}$ en hoeven alleen de stelling voor $\mathbb{R}^n$ te bewijzen. We kiezen een basis $\{e_1, \ldots, e_n\}$ van $\mathbb{R}^n$. Zij $\|\,.\,\|$ een willekeurige norm op $\mathbb{R}^n$ en zij $|\,.\,|$ de standaard Euclidische norm op $\mathbb{R}^n$ gegeven door (7.4). Merk op dat equivalentie van normen een equivalentierelatie is. Het is dus voldoende om te laten zien dat er bestaan $\alpha, \beta > 0$ zo dat

$$\alpha\,|x| \leq \|x\| \leq \beta\,|x| \quad \text{voor alle } x \in \mathbb{R}^n. \tag{7.9}$$

Als we (7.9) bewijzen voor x met $|x| = 1$ zijn we klaar. Immers, voor $|x| = 0$ is er niets te bewijzen, en voor willekeurige $x \neq 0$ geldt $x = |x| \cdot x/|x|$, dus met N2

$$\alpha\,|x| = \alpha\,|x| \cdot |x/|x|| \leq |x| \cdot \|x/|x|\| \ \ (= \|x\|) \ \leq \beta\,|x| \cdot |x/|x|| = \beta\,|x|. \tag{7.10}$$

Eerst laten we zien dat de functie $\|\,.\,\|$ continu is op $\mathbb{R}^n$ met de norm $|\,.\,|$. Wel, zij $x = \sum_{i=1}^n x_i e_i, y \sum_{i=1}^n y_i e_i \in \mathbb{R}^n$ en zij $\varepsilon > 0$, Er geldt

$$\|x - y\| = \|\sum_{i-1}^n (x_i - y_i)e_i\| \leq \sum_{i=1}^n |x_i - y_i|\|e_i\|. \tag{7.11}$$

Laat $M = \max\{\|e_1\|, \ldots, \|e_n\|\}$. Dan geeft (7.11) in combinatie met (7.7), eerste formule,

$$\|x - y\| \leq M \sum_{i=1}^n |x_i - y_i| = M\|x - y\|_1 \leq M\sqrt{n}|x - y|.$$

Nemen we nu $\delta = \varepsilon/(M\sqrt{n})$, dan zien we $|x - y| < \delta \Rightarrow \|x - y\| < \varepsilon$, en $\|\,.\,\|$ is continu.

De verzameling $S = \{x : |x| = 1\}$ is de eenheidssfeer in $\mathbb{R}^n$. Deze is gesloten en begrensd, dus met Heine Borel, **Theorem 13.12** K.A. Ross, Elementary Analysis is S compact. Er volgt dat $\|\,.\,\|$ een maximum β en een minimum α aanneemt op S. Omdat $\|x\| = 0$ alleen als $x = 0$ moet $\alpha > 0$. Op S vinden we dus

$$\alpha|x| = \alpha \leq \|x\| \leq \beta = \beta|x|.$$

Daarmee is de stelling bewezen. □

(Merk op dat dit overeenkomt met (7.7).)

7.10. Een norm op een vectorruimte V geeft op natuurlijke manier aanleiding tot een afstandsbegrip op die vectorruimte: de *afstand* $d(v, w)$ tussen vectoren v en w zal de lengte van de verschilvector $v - w$ zijn, i.e.

$$d(v, w) = \|v - w\| \quad (v, w \in V). \tag{7.12}$$

De functie $d : V \times V \to \mathbb{R}_{\geq 0}$, $(v, w) \mapsto d(v, w)$ is inderdaad een metriek als in **Definition 13.1** in K.A. Ross, Elementary Analysis, en heet de *door de norm geïnduceerde* metriek op V.

Opgave 7.5. Laat met behulp van Definitie (7.3) zien dat de geïnduceerde metriek een metriek is.

Specialiseren naar $\mathbb{R}^n$ geeft de metriek afgeleid uit de norm die is afgeleid uit het inwendig product:

$$d(x, y) := |x - y| = \sqrt{(x_1 - y_1)^2 + \cdots (x_n - y_n)^2} \quad (x, y \in \mathbb{R}^n). \tag{7.13}$$

Vergelijk K.A. Ross, Elementary Analysis pag 80, Example 2. $\mathbb{R}^n$ voorzien van inproduct (7.1) en norm (7.4) heet *n-dimensionale Euclidische ruimte* en de metriek (7.13) heet *Euclidische metriek*.

In de Euclidische metriek d_E op $\mathbb{R}^3$ is de verzameling $B(a, r) = \{x : d_E(x, a) < r\}$ een bol met middelpunt a en straal r, In een algemene metrische ruimte (S, d) noemen we

$$B(a, r) = \{x \in S : d(x, a) < r\}$$

ook een *bol met middelpunt a en straal r*.

7.11 Voorbeeld. Zij X een willekeurige verzameling. De afbeelding d gedefinieerd door $d(x, x) = 0$, $d(x, y) = 1$ als $x \neq y$ is een metriek op X.

In het vervolg zullen we een gegeven genormeerde vectorruimte $(V, \|\,.\,\|)$ ook automatisch als metrische ruimte opvatten met metriek (7.12). Convergentie van een rij (x_n) in V naar een punt $x \in V$ kan nu worden gedefinieerd:

$$\lim_{n\to\infty} x_n = x \iff \lim_{n\to\infty} d(x_n, x) = 0 \iff \lim_{n\to\infty} \|x_n - x\| = 0.$$

De eerste equivalentie definieert $\lim_{n\to\infty} x_n = x$; de tweede en derde limiet zijn "gewone" limieten van rijen reëele getallen.

Ook iedere deelverzameling van V wordt dan een metrische ruimte met de metriek (7.12).

Opgave 7.6. Zij V een reële vectorruimte. Als een rij in zekere norm in V convergeert, dan convergeert hij ook in iedere norm die daarmee equivalent is.

7.2 Limieten van functies en continuïteit

7.12. Zij X een verzameling en $f\colon X \to \mathbb{R}^m$. Dan schrijven we

$$f(x) = f_1(x)\, e_1 + \cdots + f_m(x)\, e_m = (f_1(x), \ldots, f_m(x)) \quad (x \in X),$$

waarbij de functies f_i $(i = 1, \ldots, m)$ afbeeldingen zijn van X naar $\mathbb{R}$.

Als X bovendien een metrische ruimte is en $a \in X$ dan is f continu in a desda f_i continu is in a voor alle $i = 1, \ldots, m$. We zien dit in door $\mathbb{R}^m$ met de ∞-metriek te bekijken. Dan

$$\|f(x) - f(a)\|_\infty = \max_i \{|f_i(x) - f_i(a)|\}.$$

De uitspraak volgt nu ook voor de Euclidische norm met ongelijkheid (7.7).

7.13. We bespreken limieten en continuïteit in een paar speciale gevallen. Zij $f : X \to Y$ weer een afbeelding van een metrische ruimte (X, d) naar (Y, d').

1. $Y := \mathbb{R}^m$. Dan betekent $\lim_{x\to a} f(x) = b$ dat bij iedere $\varepsilon > 0$ er is een $\delta > 0$ met:

$$0 < d(x, a) < \delta \quad \Longrightarrow \quad |f(x) - b| < \varepsilon.$$

   Merk op dat de volgende drie uitspraken met elkaar equivalent zijn (Vergelijk de vorige opmerking §(7.12)):

   (a) $\lim_{x\to a} f(x) = b$.
   (b) $\lim_{x\to a} |f(x) - b| = 0$.
   (c) $\lim_{x\to a} f_i(x) = b_i \quad (i = 1, \ldots, m)$.

2. $X \subset \mathbb{R}^n$. Dan betekent $\lim_{x\to a} f(x) = b$ dat bij iedere $\varepsilon > 0$ er is een $\delta > 0$ met:

$$0 < |x - a| < \delta \quad \Longrightarrow \quad d'(f(x), b) < \varepsilon.$$

   Als X bovendien een omgeving van a is (m.a.w. a is een inwendig punt van X) dan is er hoogstens één $b \in Y$ zo dat $\lim_{x\to a} f(x) = b$, vgl Vraagstuk 3.4.3 (Analyse 1 aanvullingen Ross).

3. $X \subset \mathbb{R}$ en $Y := \mathbb{R}$. Dan betekent $\lim_{x\to a} f(x) = b$ gewoon $\forall\, \varepsilon > 0\, \exists\, \delta > 0$ met:

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |f(x) - b| < \varepsilon.$$

Opgave 7.7. Zij E een omgeving van 0 in $\mathbb{R}^n$ en zij (Y, d') een metrische ruimte. Laat $f\colon E\backslash\{0\} \to Y$. Bewijs:

$$\lim_{\substack{h\to 0\\ h\in\mathbb{R}^n}} f(h) = y \quad \Longrightarrow \quad \forall x \in \mathbb{R}^n\backslash\{0\} \quad \lim_{\substack{t\to 0\\ t\in\mathbb{R}}} f(t\,x) = y.$$

Vergelijk Stelling 3.3.3 (Analyse 1 aanvulling Ross) Hierbij hebben we in de linkerlimiet $h \in \mathbb{R}^n$ toegevoegd en in de rechterlimiet $t \in \mathbb{R}$ om te benadrukken dat de linkerlimiet genomen wordt voor $h \to 0$ met h in een omgeving van 0 in $\mathbb{R}^n$ en de rechterlimiet voor $t \to 0$ met t in een omgeving van 0 in $\mathbb{R}$.

7.14 Definitie. Laat f een reëelwaardige functie zijn op $\mathbb{R}^n\backslash\{0\}$ en laat $\alpha \in \mathbb{R}$. We noemen f *homogeen van graad* α als

$$f(rx) = r^\alpha f(x) \quad (x \in \mathbb{R}^n\backslash\{0\},\ r > 0). \tag{7.14}$$

Bijvoorbeeld de functies

$$f_1(x, y) := \frac{x^2 - y^2}{x^4 + y^4} \quad f_2(x, y) := \sin\left(\frac{x}{\sqrt{x^2 + y^2}}\right), \quad f_3(x, y) := \frac{xy^2}{x^2 + y^2}$$

zijn homogeen van graad -2, 0 en 1, respectievelijk.

Zij $S^{n-1} := \{x \in \mathbb{R}^n \mid |x| = 1\}$, dus de bol van straal 1 met middelpunt 0 in $\mathbb{R}^n$. Een homogene functie van graad α op $\mathbb{R}^n\backslash\{0\}$ is geheel bepaald door haar restrictie tot S^{n-1}. Inderdaad,

$$f(x) = f(|x|(x/|x|)) = |x|^\alpha f(x/|x|), \qquad \text{en } x/|x| \in S^{n-1}.$$

In het bijzonder is een homogene functie van graad 0 constant op elk van de halflijnen $\{rx \mid r > 0\}$ $(x \in S^{n-1})$.

7.15 Propositie. Zij $f\colon \mathbb{R}^n\backslash\{0\} \to \mathbb{R}$ continu, homogeen van graad α en niet identiek gelijk aan een constante. Dan geldt: $\lim_{x\to 0} f(x) = 0$ als $\alpha > 0$ en $\lim_{x\to 0} f(x)$ bestaat niet als $\alpha \le 0$.

Bewijs. S^{n-1} is een gesloten en begrensde deelverzameling van $\mathbb{R}^n$, dus compact (cf. K.A. Ross, Elementary Analysis, Stelling 13.12). De functie f is continu op de compacte verzameling S^{n-1}, dus is daar begrensd (cf. K.A. Ross, Elementary Analysis, Stelling 21.4 en Gevolg 21.5). Laat $M := \sup_{x\in S^{n-1}} |f(x)|$. Dan $0 < M < \infty$ en $|f(x)| \le M\,|x|^\alpha$ (waarom?). Veronderstel dat $\alpha > 0$ en zij $\varepsilon > 0$ gegeven. Dan geldt: als $|x| \le \delta := (\varepsilon/M)^{1/\alpha}$, dan $|f(x)| < \varepsilon$. Dus $\lim_{x\to 0} f(x) = 0$. Veronderstel nu dat $\alpha < 0$ en neem $x \in S^{n-1}$ zo dat $f(x) \ne 0$. Dan $f(rx) = r^\alpha f(x)$, dus $\lim_{r\downarrow 0} f(rx)$ bestaat niet, dus $\lim_{x\to 0} f(x)$ bestaat niet (cf. Opgave 7.7). Veronderstel tenslotte dat $\alpha = 0$. Dan zijn er $x, y \in S^{n-1}$ zo dat $f(x) \ne f(y)$. Omdat $\lim_{r\downarrow 0} f(rx) = f(x) \ne f(y) = \lim_{r\downarrow 0} f(ry)$ bestaat dus $\lim_{x\to 0} f(x)$ niet (cf. Opgave 7.7). □

7.16 Voorbeeld. Laat

$$f(x,y) := \frac{xy^2}{x^2+y^4} \quad \text{als} \quad (x,y) \in \mathbb{R}^2\backslash\{(0,0)\} \quad \text{en} \quad f(0,0) := 0.$$

Dan geldt voor alle $(x,y) \ne (0,0)$ dat

$$f(rx,ry) = r\,\frac{xy^2}{x^2+r^2y^4} \longrightarrow 0 = f(0,0) \quad \text{als} \quad r \downarrow 0.$$

Toch is f niet continu in $(0,0)$. Herschrijf f hiertoe als

$$f(x,y) = \frac{x/y^2}{1+(x/y^2)^2} \quad (y \ne 0), \qquad \text{dus} \qquad f(cy^2,y) = \frac{c}{1+c^2} \quad (y \ne 0).$$

Als $\lim_{(x,y)\to(0,0)} f(x,y)$ zou bestaan dan zou $\lim_{y\to 0} f(cy^2,y)$ onafhankelijk van c zijn, maar dit is niet het geval.

Opgave 7.8. Bewijs dat de functie f in bovenstaand voorbeeld begrensd is op $\mathbb{R}^2$. Teken de krommen in $\mathbb{R}^2$ waarop f constant is. Kan het niet bestaan van $\lim_{(x,y)\to(0,0)} f(x,y)$ ook ingezien worden met behulp van Propositie 7.15?

7.3 De operatornorm

We geven hieronder de definitie van begrensde lineaire afbeeldingen tussen genormeerde vectorruimten in algemene setting. In de latere hoofdstukken zullen we ons vooral bezighouden met de situatie dat de vectorruimten eindigdimensionaal zijn. De lineaire afbeeldingen zijn dan automatisch begrensd en zijn dan precies de lineaire afbeeldingen uit het college Lineaire Algebra.

7.17. Laten V en W genormeerde reële vectorruimtes zijn en zij $A\colon V \to W$ een lineaire afbeelding. Dan heet A *begrensd* als er een $\alpha > 0$ bestaat zo dat $\|Av\| \le \alpha\,\|v\|$ voor alle $v \in V$. Er geldt: A is begrensd desda A continu is.

Opgave 7.9. Bewijs deze uitspraak door te bewijzen: A is begrensd $\Longleftrightarrow$ A is continu in 0 $\Longleftrightarrow$ A is continu.

7.18 Definitie. Als A begrensd is dan wordt zijn *operatornorm* gedefinieerd door

$$\begin{aligned}
\|A\| &:= \inf\{\alpha \in \mathbb{R} \mid \|Av\| \le \alpha\,\|v\| \quad \forall v \in V\} && (7.15)\\
&= \min\{\alpha \in \mathbb{R} \mid \|Av\| \le \alpha\,\|v\| \quad \forall v \in V\} && (7.16)\\
&= \sup_{\substack{v\in V\\ \|v\|\le 1}} \|Av\|. && (7.17)
\end{aligned}$$

We zullen het liefste werken met formule (7.17) voor de operatornorm $\|A\|$. Merk op dat (7.17) voor willekeurige lineaire afbeeldingen $A\colon V \to W$ een waarde voor $\|A\|$ in $[0, \infty]$ geeft en dat A begrensd is desda $\|A\| < \infty$.

Als A begrensd is dan geldt:

$$\|Av\| \leq \|A\|\,\|v\| \quad (v \in V). \tag{7.18}$$

De verzameling van alle begrensde lineaire afbeeldingen van V naar W wordt genoteerd met $B(V, W)$. Deze verzameling vormt een lineaire deelruimte van de vectorruimte $L(V, W)$ van alle lineaire afbeeldingen van V naar W. Bovendien wordt $B(V, W)$ een genormeerde vectorruimte t.o.v. de operatornorm.

Opgave 7.10. Laat zien dat de operatornorm een norm is op $B(V, W)$.

Opgave 7.11. Beschouw $A := \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ als lineaire afbeelding van $(\mathbb{R}^2, |\,.\,|)$ naar zichzelf. Bewijs dat $\|A\| = \sup_{x^2+y^2=1}\big((\lambda x)^2 + (\mu y)^2\big)^{\frac{1}{2}} = \max\{|\lambda|, |\mu|\}$.

7.19 Propositie. Laten V en W genormeerde reële vectorruimtes zijn, V bovendien eindig-dimensionaal, met norm $\|.\|_V$, $\|.\|_W$ respectievelijk. Dan zijn alle lineaire afbeeldingen van V naar W begrensd.

Bewijs. Laat f een lineaire afbeelding van V naar W zijn, en laat $\{e_1, \ldots, e_n\}$ een basis van V zijn. Omdat V eindig dimensionaal is, zijn alle normen op V equivalent. We mogen dus de norm kiezen, en aannemen dat $\|x\|_V = \sum_{i=1}^n |x_j|$, waar $x = x_1e_1 + \cdots x_ne_n$. Laat $M = \max\{\|f(e_j)\|\}$. Er geldt voor alle $x \in V$ dat

$$\|f(x)\| = \|\sum_{i=1}^n x_i f(e_i)\| \leq \sum_{i=1}^n |x_i|\|f(e_i)\| \leq M \sum_{i=1}^n |x_i| = M\|x\|_V.$$

□

Als V en W genormeerde reële vectorruimtes zijn, maar V niet eindig-dimensionaal dan bestaan er onbegrensde lineaire afbeeldingen van V naar W.

In het vervolg van deze syllabus zullen we bijna altijd lineaire afbeeldingen $A\colon V \to W$ bekijken met V en W eindig-dimensionaal. Dan geldt, ongeacht de keuze van de normen op V en W, dat $B(V, W) = L(V, W)$. We zullen dan meestal de notatie $L(V, W)$ gebruiken.

De ruimte $L(\mathbb{R}^n, \mathbb{R}^m)$ bestaat dus uit de lineaire afbeeldingen van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Zo'n lineaire afbeelding wordt meestal geïdentificeerd een $m \times n$-matrix, waarvan de j-de kolom het beeld van de j-de eenheidsvector is. We kunnen $L(\mathbb{R}^n, \mathbb{R}^m)$ dus opvatten als $M(m, n, \mathbb{R})$ de nm-dimensionale ruimte van $m \times n$-matrices. De operator norm is een natuurlijke norm op $L(\mathbb{R}^n, \mathbb{R}^m)$ genormeerde vectorruimte is. Voor $n, m > 1$ is de operatornorm voor deze ruimte zeker niet gelijk aan een van de ℓ^p-norm (7.6), voor $\mathbb{R}^{mn}$, maar is er wel equivalent mee (cf. Propositie 7.8).

Opgave 7.12. Veronderstel dat U, V en W genormeerde reële vectorruimtes zijn en dat $B\colon U \to V$ en $A\colon V \to W$ begrensde lineaire afbeeldingen zijn. Bewijs dat $AB\colon U \to W$ begrensd is en dat $\|AB\| \leq \|A\|\,\|B\|$.

Opgave 7.13. Bewijs dat er met elke $A \in L(\mathbb{R}^n, \mathbb{R})$ een unieke $c \in \mathbb{R}^n$ correspondeerd zo dat $Ax = (x, c)$ $(x \in \mathbb{R}^n)$. Bewijs dat $\|A\| = |c|$.

Opgave 7.14. Beschouw $\mathbb{R}^n$ met standaard inwendig product en norm. Zij $A\colon \mathbb{R}^n \to \mathbb{R}^n$ lineair.

(a) Bewijs dat $\|A\| = 1$ als A orthogonaal is.

(b) Bewijs dat $\|A\| = \max_{j=1,\dots,n} |\lambda_j|$ als A symmetrisch is met eigenwaarden $\lambda_1, \dots, \lambda_n$.

(c) Laat zien dat $\max_{j=1,\dots,n} |\lambda_j|$ **geen** norm definieert op $L(\mathbb{R}^n, \mathbb{R}^m)$.

Verdere vraagstukken

Opgave 7.15. Ga voor de volgende functies $f\colon \mathbb{R}^2 \backslash \{(0,0)\} \to \mathbb{R}$ na of $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ bestaat en zo ja, bepaal deze limiet.

(a) $f(x,y) := \dfrac{xy}{x^2+y^2}$

(b) $f(x,y) := \dfrac{x^2\,y^2}{x^2+y^4}$

(c) $f(x,y) := \dfrac{xy}{x^2+y^4}$

(d) $f(x,y) := y\,\sin\dfrac{1}{x^2+y^2}$

(e) $f(x,y) := xy\,\log(x^2+y^2)$

(f) $f(x,y) := \dfrac{\log(1+x^2y^2)}{x^2+y^2}$

(g) $f(x,y) := \dfrac{\sin xy}{x^2+y^2}$

(h) $f(x,y) := \dfrac{x^4y}{(x^2+y^2)^2}$

Opgave 7.16. Toon aan dat de volgende functies $f\colon \mathbb{R}^2 \to \mathbb{R}$ continu zijn.

(a) $f(x,y) := \dfrac{\log(1+|xy|)}{|x|}$ als $x \neq 0$ en $f(0,y) := |y|$

(b) $f(x,y) := \dfrac{e^{x^2-y^2}-1}{x^2-y^2}$ als $x^2 \neq y^2$ en $f(x,y) := 1$ als $x^2 = y^2$

(c) $f(x,y) := \dfrac{\sin(x^2-y^2)}{x-y}$ als $x \neq y$ en $f(x,x) := 2x$

8 Differentiatie van functies van $\mathbb{R}^n$ naar $\mathbb{R}^m$

8.1 *Partiële en richtingsafgeleide*

Zij $I \subset \mathbb{R}$ een open interval en zij $f : I \to \mathbb{C}$. Zoals welbekend, noemen we de functie f *differentieerbaar* in een punt $a \in I$ als

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{8.1}$$

bestaat en we duiden de limiet (8.1) dan aan met $f'(a)$ of $(Df)(a)$. Als f differentieerbaar is in alle $a \in I$ dan definieert dit zodoende een nieuwe functie op I, de *afgeleide* van f, genoteerd als f' of Df of ook wel als de functie $x \mapsto \frac{df(x)}{dx}$. De notatie $\frac{df}{dx}$ voor de functie f' kan in theoretische beschouwingen beter vermeden worden, omdat dan al van te voren wordt aangenomen dat x de onafhankelijke variabele is. Evenzo schrijven we de limiet (8.1) dan liever niet als $\frac{df}{dx}(a)$, wel eventueel als $\frac{df(x)}{dx}\Big|_{x=a}$.

In dit hoofdstuk generaliseren we het begrip afgeleide naar afbeeldingen van open stukken in $\mathbb{R}^n$ naar $\mathbb{R}^m$. Wanneer we (8.1) voor functies van meer veranderlijken op proberen te schrijven, zien we dat we zouden delen door een vector h. Dat is een probleem dat op verschillende manier omzeild kan worden. Verschillende benaderingen leiden tot andere soorten afgeleide. Onderweg testen we in hoeverre de nieuwe concepten passen bij ons idee van afgeleide. De totale afgeleide uit Sectie 2.2. blijkt het meest van de eigenschappen van de afgeleide (8.1) te behouden, maar is misschien ook conceptueel het lastigst.

8.1 Definitie. Laat nu E een open deelverzameling van $\mathbb{R}^n$ zijn, zij $f : E \to \mathbb{C}$ en zij $j \in \{1, \dots, n\}$. We noemen de functie f *partieel differentieerbaar naar de j de coördinaat van het argument* in een punt $a = (a_1, \dots, a_n) \in E$ als

$$\lim_{h \to 0} \frac{f(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - f(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n)}{h} \tag{8.2}$$

bestaat en we duiden de limiet (8.2) dan aan met $(D_j f)(a)$. Als de limiet (8.2) bestaat voor alle $a \in E$ dan definieert dit zodoende een nieuwe functie $D_j f : x \mapsto (D_j f)(x) : E \to \mathbb{C}$, de *partiële afgeleide* van f *naar de $j$de coördinaat*.

We kunnen (8.2) wat compacter schrijven als volgt.

$$(D_j f)(a) := \lim_{h \to 0} \frac{f(a + he_j) - f(a)}{h} \,. \tag{8.3}$$

Partiële differentiatie komt er dus simpelweg op neer dat we alle onafhankelijke variabelen behalve de $j$de constant houden en dan de resulterende functie

$$g(x) := f(a_1, \dots, a_{j-1}, x, a_{j+1}, \dots, a_n)$$

van één variabele bekijken. Dan bestaat $(D_j f)(a)$ desda $g'(a_j)$ bestaat en beide uitdrukkingen zijn aan elkaar gelijk indien ze bestaan.

Als de partiële afgeleide $D_j f$ bestaat op E dan schrijven we deze functie eventueel ook als $x \mapsto \frac{\partial f(x)}{\partial x_j}$, terwijl $(D_j f)(a)$ geschreven kan worden als $\frac{\partial f(x)}{\partial x_j}\Big|_{x=a}$. In theoretische beschouwingen vermijden we echter liever weer de notaties $\frac{\partial f}{\partial x_j}$ voor $D_j f$ en $\frac{\partial f}{\partial x_j}(a)$ voor $(D_j f)(a)$. (Zulke notatie is bijvoorbeeld dubbelzinnig in een voorbeeld als $\frac{\partial f}{\partial x_1}(x_2, x_1)$, waarbij f een functie op $\mathbb{R}^2$ is.)

8.2 Voorbeeld. Laat de functie f gedefinieerd zijn op $\mathbb{R}^2$ door $f(x,y) := x + 3x^2y + 5x^4y^2$. Dan bestaan $D_1 f$ en $D_2 f$ overal op $\mathbb{R}^2$:

$$(D_1 f)(x,y) = 1 + 6xy + 20x^3y^2, \quad (D_2 f)(x,y) = 3x^2 + 10x^4y.$$

De partiële afgeleiden van de functies $D_1 f$ en $D_2 f$ bestaan ook weer overal. In het bijzonder vinden we:

$$(D_1(D_2 f))(x,y) = 6x + 40x^3y = (D_2(D_1 f))(x,y).$$

Het is niet algemeen waar dat $D_1(D_2 f) = D_2(D_1 f)$ zodra beide leden bestaan. Later zullen we een voorbeeld tegenkomen waarbij geen gelijkheid geldt.

We pakken Voorbeeld 7.16 weer even op, dus

$$f(x,y) := \frac{xy^2}{x^2 + y^4} \quad \text{als} \quad (x,y) \in \mathbb{R}^2 \backslash \{(0,0)\} \quad \text{en} \quad f(0,0) := 0. \tag{8.4}$$

Dan bestaan $(D_1 f)(0,0)$ en $(D_2 f)(0,0)$ en allebei zijn ze gelijk aan 0. Toch is f niet continu in $(0,0)$. Voor een functie in één variabele impliceert differentieerbaarheid in een punt echter continuïteit in dat punt. Partiële differentieerbaarheid is daarom een tamelijk zwakke eigenschap.

Opgave 8.1. Zij $f : \mathbb{R}^n \to \mathbb{R}$ zo dat voor alle $x \in \mathbb{R}^n$ en alle $j \in \{1, \ldots, n\}$ de partiële afgeleide $(D_j f)(x)$ bestaat. Bewijs het volgende. Als $(D_j f)(x) = 0$ voor alle x en j dan is $f(x) =$ constant.

8.3. Vervang in het rechterlid van vergelijking (8.3) de vector e_j door een willekeurige vector $u \in \mathbb{R}^n$. Als de limiet dan bestaat, dan spreken we van de richtingsafgeleide.

8.4 Definitie. Zij E een open deelverzameling van $\mathbb{R}^n$, zij $f : E \to \mathbb{C}$, $a \in E$ en $u \in \mathbb{R}^n$. De *richtingsafgeleide in de richting* u van de functie f in het punt a is gedefinieerd als

$$(D_u f)(a) := \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h}\,, \tag{8.5}$$

mits de limiet in het rechterlid bestaat.

8.5. Als we dus, met gegevens als boven, de functie g definiëren op een interval rond 0 door

$$g(t) := f(a + tu) \quad \text{dan geldt} \quad (D_u f)(a) = g'(0) \tag{8.6}$$

mits een van beide afgeleiden bestaat.

Merk op dat $(D_j f)(a) = (D_{e_j} f)(a)$, waarbij het linkerlid gedefinieerd is door (8.3) en het rechterlid door (8.5). We hebben dus twee notaties voor het zelfde object, maar doorgaans is uit het verband wel duidelijk of de partiële afgeleide danwel de richtingsafgeleide bedoeld wordt.

Als de richtingsafgeleide $(D_u f)(a)$ bestaat voor zekere $u \neq 0$, dan ook de richtingsafgeleide in de richting αu $(\alpha \in \mathbb{R})$ bestaat en

$$(D_{\alpha u} f)(a) = \lim_{h \to 0} \frac{f(a + h\alpha u) - f(a)}{h} = \lim_{h \to 0} \alpha \frac{f(a + h\alpha u) - f(a)}{h\alpha} = \alpha\, (D_u f)(a). \tag{8.7}$$

De richtingsafgeleide in de richting 0 bestaat altijd:

$$(D_0 f)(a) = 0. \tag{8.8}$$

Strikt genomen is de historisch gegroeide naam richtingsafgeleide misleidend, omdat twee vectoren u en v ongelijk 0 en met $v = \alpha u$ $(\alpha > 0)$ dezelfde richting vertegenwoordigen, terwijl $(D_u f)(a)$ en $(D_v f)(a)$ doorgaans niet aan elkaar gelijk zijn.

8.6 Voorbeeld. Neem f weer als in Voorbeeld (7.16), gegeven door (8.4). Dan geldt dat

$$\frac{f(rx,ry)-f(0,0)}{r} = \frac{xy^2}{x^2+r^2y^4} \overset{r\to 0}{\longrightarrow} (D_{(x,y)}f)(0,0) = \begin{cases} y^2/x & \text{als } x \neq 0, \\ 0 & \text{als } x = 0. \end{cases} \tag{8.9}$$

Dus, ook al bestaan in een zeker punt a de richtingsafgeleiden van een functie f in alle richtingen, dan hoeft f toch niet continu in a te zijn. Het bestaan van richtingsafgeleiden in alle richtingen is dus nog steeds niet een voldoende sterke eigenschap om als analogon in meer variabelen van differentieerbaarheid te dienen.

8.7 Voorbeeld. Neem f als in Voorbeeld 8.2, dus $f(x,y) := x+3x^2y+5x^4y^2$. Dan is $f(x+tu, y+tv)$ voor alle x,y,u,v een oneindig vaak differentieerbare functie in t, dus de richtingsafgeleide van f bestaat in alle punten (x,y) in alle richtingen (u,v). De (gewone) afgeleide $g'(0)$ van de functie $g : t \mapsto f(x+tu, y+tv)$ geeft dan de richtingsafgeleide:

$$(D_{(u,v)}f)(x,y) = u + 3(2xuy + x^2v) + 5(4x^3uy^2 + 2x^4yv).$$

Wanneer we dit vergelijken met de uitdrukkingen voor D_1f en D_2f in Voorbeeld 8.1, dan zien we dat

$$(D_{(u,v)}f)(x,y) = u\,(D_1f)(x,y) + v\,(D_2f)(x,y).$$

Dus $(D_{(u,v)}f)(x,y)$ is *lineair* in u en v, d.w.z. dat de afbeelding $(u,v) \mapsto (D_{(u,v)}f)(x,y) : \mathbb{R}^2 \to \mathbb{R}$ lineair is.

8.8. Merk op dat in Voorbeeld 8.6 $(D_{(u,v)}f)(0,0)$ verre van lineair is in u en v. We zullen spoedig zien dat lineariteit van de richtingsafgeleide in de richtingsvector geldt, zodra f “voldoende mooi” is.

8.2 *Totale afgeleide*

8.9. Zij $I \subset \mathbb{R}$ een open interval en zij $f : I \to \mathbb{R}$. Zij $a \in I$ en $c \in \mathbb{R}$. Dan zijn de volgende uitspraken alle equivalent:

(a) f is differentieerbaar in a met afgeleide gelijk aan c.

(b) $\lim_{h\to 0} h^{-1}\,(f(a+h) - f(a)) = c$.

(c) $\lim_{h\to 0} h^{-1}(f(a+h) - f(a) - c\,h) = 0$.

(d) $r(h) := f(a+h) - f(a) - c\,h$ voldoet aan $\lim_{h\to 0} h^{-1}\,r(h) = 0$.

(e) $r(h) := f(a+h) - f(a) - c\,h$ voldoet aan $\lim_{h\to 0} |h|^{-1}\,|r(h)| = 0$.

(f) $f(a+h) = f(a) + c\,h + o(|h|)$ als $h \to 0$.

Laten we (e) op een andere manier bekijken. Voor iedere $c \in \mathbb{R}$ is er een unieke functie r gedefinieerd op een omgeving van 0 in $\mathbb{R}$ zo dat

$$f(a+h) = f(a) + c\,h + r(h) \tag{8.10}$$

voor h in een omgeving van 0, of equivalent

$$f(x) = f(a) + c\,(x-a) + r(x-a) \tag{8.11}$$

voor x in een omgeving van a. De eerste twee termen in het rechterlid van (8.11) geven een functie $x \mapsto f(a) + c\,(x-a)$, waarvan de grafiek een rechte is door het punt $(a, f(a))$. De functie $x \mapsto r(x-a)$ geeft het verschil van de functies f en $x \mapsto f(a) + c\,(x-a)$. Differentieerbaarheid van f in a met afgeleide c wil dan zeggen dat de verschilfunctie $x \mapsto r(x-a)$ voor x nabij a in absolute waarde van kleinere orde is dan $|x-a|$. Meetkundig wil dat zeggen dat genoemde rechte raakt aan de grafiek van f in het punt $(a, f(a))$.

Bekijk nu het rechterlid van (8.10). Dat is opgebouwd uit een constante term $f(a)$, een term $c\,h$ die lineair is in h en een restterm $r(h)$ die, in het geval van differentieerbaarheid van f in a, van kleinere orde is in h dan een lineaire functie in h. Een lineaire functie $h \mapsto c\,h$ is een lineaire afbeelding van de vectorruimte $\mathbb{R}$ naar de vectorruimte $\mathbb{R}$. Noem deze lineaire afbeelding A, dus $Ah = c\,h$. Dan wordt de afbeelding A gerepresenteerd door de 1×1 matrix (c) en er is een lineaire 1-1 correspondentie tussen getallen c en afbeeldingen A.

We zouden dus in (a) van bovenstaande equivalente uitspraken evengoed kunnen zeggen dat f differentieerbaar is in a met als afgeleide de lineaire afbeelding A, en we zouden in uitspraken (c)–(f) de uitdrukking $c\,h$ kunnen vervangen door de uitdrukking Ah.

8.10. Zij nu E een open deelverzameling van $\mathbb{R}^n$, zij $f : E \to \mathbb{R}^m$ en zij $a \in E$. Zij $A : \mathbb{R}^n \to \mathbb{R}^m$ lineair. Naar analogie van (8.10) schrijven we

$$f(a+h) = f(a) + Ah + r(h) \qquad (h \in \mathbb{R}^n \text{ zo dat } a+h \in E). \tag{8.12}$$

Het linkerlid geeft een van h afhankelijke vector in $\mathbb{R}^m$. In het rechterlid is de eerste term een constante $f(a)$ in $\mathbb{R}^m$, de tweede term is een vector in $\mathbb{R}^m$ die lineair afhangt van $h \in \mathbb{R}^n$ en de derde term is de restvector $r(h)$ in $\mathbb{R}^m$ die nog afhangt van $h \in \mathbb{R}^n$. Analoog aan (e) in §8.9 zullen we f nu differentieerbaar noemen in het punt a als deze restvector van kleinere orde is, dwz sneller naar 0 gaat, in h voor $h \to 0$ dan $|h|$.

8.11 Definitie. Zij E een open deelverzameling van $\mathbb{R}^n$. De afbeelding $f : E \to \mathbb{R}^m$ is *differentieerbaar* in zeker punt $a \in E$ met *afgeleide* $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ als $r(h)$ in (8.12) voldoet aan

$$\lim_{h \to 0} \frac{|r(h)|}{|h|} = 0. \tag{8.13}$$

8.12. In (8.13) is $|r(h)|$ de (standaard)norm van $r(h)$ in $\mathbb{R}^m$ en is $|h|$ de (standaard)norm van h in $\mathbb{R}^n$. De functie $h \mapsto |h|^{-1}\,|r(h)|$ in (8.13) is gedefinieerd op een omgeving van 0 in $\mathbb{R}^n$ met 0 er uit weggelaten.

We schrijven (8.12) samen met (8.13) ook wel compacter met behulp van het $o(|h|)$-symbool (spreek uit “kleine o van de norm van h”) als volgt:

$$f(a+h) = f(a) + Ah + o(|h|) \quad \text{als} \quad h \to 0 \text{ in } \mathbb{R}^n. \tag{8.14}$$

Per definitie bedoelen we hiermee hetzelfde als met (8.12) en (8.13) gecombineerd. De uitspraak (8.14) zegt dus juist dat f differentieerbaar is in a met afgeleide A.

Nog een andere equivalente manier om te zeggen dat f differentieerbaar is in a met afgeleide A is dat

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - Ah|}{|h|} = 0. \tag{8.15}$$

Hiermee is weer equivalent (cf. §(7.13), onderdeel 1) dat

$$\lim_{h \to 0} |h|^{-1} \left(f(a+h) - f(a) - Ah\right) = 0, \tag{8.16}$$

en ook dat

$$\lim_{h \to 0} \frac{f_i(a+h) - f_i(a) - (Ah)_i}{|h|} = 0 \quad (i = 1, \ldots, m). \tag{8.17}$$

8.13 Propositie. Zij E een open deelverzameling van $\mathbb{R}^n$. Als de afbeelding $f : E \to \mathbb{R}^m$ in een punt $a \in E$ differentieerbaar is met afgeleide in $L(\mathbb{R}^n, \mathbb{R}^m)$ dan is deze afgeleide uniek.

Bewijs. Neem aan dat $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ allebei afgeleiden zijn van f in a. Dan geldt (8.16) en evenzo $\lim\limits_{h\to 0} |h|^{-1} (f(a+h) - f(a) - Bh) = 0$. Trek deze twee limietformules van elkaar af en pas de Opgave 3.4.1. uit (Analyse 1 aanvulling) toe. We verkrijgen dat $\lim\limits_{h\to 0} |h|^{-1} (B - A)h = 0$. Dus er volgt met behulp van Opgave 7.7 dat voor elke $x \in \mathbb{R}^n \backslash \{0\}$ geldt:

$$0 = \lim_{t\downarrow 0} \left(|t\,x|^{-1} (B - A)(t\,x)\right) = \lim_{t\downarrow 0} \left(|x|^{-1}\, t^{-1}\, t\,(B - A)x\right) = |x|^{-1} (B - A)x.$$

Dus $(B - A)x = 0$ als $0 \neq x \in \mathbb{R}^n$. Dus $B - A = 0$. □

8.14 Definitie. We noteren de afgeleide van f in a als $f'(a)$. Deze wordt ook wel *totale afgeleide* genoemd, om onderscheid te maken met het begrip partiële afgeleide. Als $f : E \to \mathbb{R}^m$ differentieerbaar is voor alle $a \in E$ dan is $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$ dus gedefinieerd voor alle $a \in E$ en we hebben een nieuwe afbeelding $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$, die we de *(totale) afgeleide* van f op E noemen.

Opgave 8.2. Zij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}^m$. Schrijf $f(x) = (f_1(x), \ldots, f_m(x))$. Zij $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, voorgesteld door de reële $m \times n$ matrix $(A_{ij})_{i=1,\ldots,m;\, j=1,\ldots,n}$. Bewijs dat f differentieerbaar is in a met afgeleide A desda, voor alle $i = 1, \ldots, m$, $f_i : E \to \mathbb{R}$ differentieerbaar is in a met afgeleide $f_i'(a) : \mathbb{R}^n \to \mathbb{R}$ gegeven door de $1 \times n$ matrix $(A_{i1} \ldots A_{in})$.
Aanwijzing Gebruik de equivalentie van (8.16) met (8.17).

8.15 Lemma. Zij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}^m$. Zij f differentieerbaar in a met afgeleide $f'(a)$.

Er bestaat een $\delta > 0$ zo dat, als $|h| < \delta$ dan $a + h \in E$ en

$$|f(a + h) - f(a)| \leq (\|f'(a)\| + 1)\,|h|.$$

Bewijs. Er geldt

$$|f(a + h) - f(a)| = |f'(a)h + o(|h|)| \leq (\|f'(a)\| + 1)\,|h|.$$

We gebruikten $|o(|h|)| \leq |h|$ als $|h| \to 0$ en (8.14). □

8.16 Propositie. Zij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}^m$. Als f differentieerbaar is in a dan is f ook continu in a.

Bewijs. Dit volgt onmiddellijk uit Lemma (8.15). □

Deze propositie geeft een indicatie dat we nu het juiste begrip van differentieerbaarheid voor functies in meer veranderlijken hebben.

Opgave 8.3. Zij $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ continu, homogeen van graad $\alpha \leq 0$ en niet identiek gelijk aan een constante. Breid f uit tot een functie $f : \mathbb{R}^n \to \mathbb{R}$ door een of andere keuze van $f(0)$. Bewijs dat f niet differentieerbaar is in 0.

8.17 Voorbeeld. Zij $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ en $f(x) := Ax$ ($x \in \mathbb{R}^n$). Dan is f overal differentieerbaar en $f'(x) = A$ (constant).

Inderdaad, er geldt voor $a, h \in \mathbb{R}^n$ dat

$$f(a + h) = A(a + h) = Aa + Ah = f(a) + Ah.$$

We zijn dus in de meest eenvoudige situatie van (8.12): de restvector $r(h)$ is identiek 0.

8.3 Verband tussen verschillende types afgeleiden

8.18 Stelling. Zij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}^m$, differentieerbaar in a. Dan bestaan alle partiële afgeleiden $(D_j f_i)(a)$ en $f'(a)$, als lineaire afbeelding van $\mathbb{R}^n$ naar $\mathbb{R}^m$, heeft matrix

$$f'(a) = \begin{pmatrix} (D_1 f_1)(a) & \dots & (D_n f_1)(a) \\ \vdots & & \vdots \\ (D_1 f_m)(a) & \dots & (D_n f_m)(a) \end{pmatrix} \tag{8.18}$$

Dus

$$f'(a)\, e_j = \sum_{i=1}^{m} (D_j f_i)(a)\, e_i \quad (j = 1, \dots, n).$$

Bewijs. Zij $A := f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$ en $A = (A_{ij})_{i=1,\dots,m;j=1,\dots,n}$. Neem $i \in \{1, \dots, m\}$ en $j \in \{1, \dots, n\}$ vast. Uit (8.17) weten we dat

$$\lim_{\substack{h\to 0 \\ h \in \mathbb{R}^n}} \frac{f_i(a+h) - f_i(a) - (Ah)_i}{|h|} = 0.$$

Voor iedere basisvector e_j substitueren we $h = h(t) = te_j$, $t \in \mathbb{R}$. Voor de samenstelling met de continue functie te_j geldt dan ook dat

$$\lim_{t\to 0} \frac{f_i(a + te_j) - f_i(a) - t\,(Ae_j)_i}{|t|} = 0. \tag{8.19}$$

Merk op dat $(Ae_j)_i = A_{ij}$. Dus

$$\lim_{t\to 0} \left| \frac{f_i(a + te_j) - f_i(a) - tA_{ij}}{t} \right| = 0.$$

Dus

$$\lim_{t\to 0} \frac{f_i(a + te_j) - f_i(a)}{t}$$

bestaat en is gelijk aan A_{ij}. Gezien (8.3) bestaat dus de partiële afgeleide $(D_j f_i)(a)$ en is deze gelijk aan A_{ij}. □

8.19. We houden de gegevens van bovenstaande stelling. Zij $h \in \mathbb{R}^n$ en $k := f'(a)\, h$, dus $k \in \mathbb{R}^m$. Als we deze identiteit in termen van coördinaten willen uitschrijven dan is het, zoals gebruikelijk bij een matrix werkend op een vector, handig om de de vectoren h en k als *kolomvectoren* te schrijven:

$$\begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} = \begin{pmatrix} (D_1 f_1)(a) & \dots & (D_n f_1)(a) \\ \vdots & & \vdots \\ (D_1 f_m)(a) & \dots & (D_n f_m)(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}. \tag{8.20}$$

8.20. Als we nu in Stelling 8.18 nemen $m = 1$, i.e., f is een reëelwaardige functie op E, dan is $f'(a)$ een lineaire afbeelding van $\mathbb{R}^n$ naar $\mathbb{R}$, die voorgesteld wordt door de $1 \times n$ matrix

$$f'(a) = ((D_1 f)(a) \dots (D_n f)(a)). \tag{8.21}$$

Zo'n $1 \times n$ matrix kan ook als een *rijvector* met n reële coördinaten worden opgevat. Voor $h \in \mathbb{R}^n$ zal nu $f'(a)\,h \in \mathbb{R}$. Meer expliciet wordt dit:

$$f'(a)\,h = ((D_1 f)(a) \dots (D_n f)(a)) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{j=1}^{n} (D_j f)(a)\,h_j \quad (h \in \mathbb{R}^n). \tag{8.22}$$

We kunnen nu ook een eenvoudig verband met de richtingsafgeleide geven.

8.21 Stelling. Zij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}$, differentieerbaar in a. Zij $u \in \mathbb{R}^n$. Dan bestaat de richtingsafgeleide $(D_u f)(a)$ en er geldt

$$(D_u f)(a) = f'(a)\,u = u_1\,(D_1 f)(a) + \cdots + u_n\,(D_n f)(a). \tag{8.23}$$

Bewijs. Het resultaat is evident als $u = 0$. Laat $u \neq 0$ en concludeer uit (8.15) dat dan ook de volgende limiet geldt.

$$\lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{|f(a + tu) - f(a) - f'(a)\,tu|}{|t|\,|u|} = 0.$$

Dus

$$\lim_{t \to 0} \frac{f(a + tu) - f(a) - t\,f'(a)\,u}{t} = 0.$$

Vergelijk dit met (8.5) en pas (8.22) toe. □

8.22 Opmerking. We zien uit bovenstaande Stelling dat, als f differentieerbaar is in a, de richtingsafgeleide $(D_u f)(a)$ lineair in u is. Dus ook: als de richtingsafgeleide $(D_u f)(a)$ niet lineair is in u, dan is f niet differentieerbaar in a.

Dus we zien uit formule (8.9) dat de functie f uit Voorbeeld 7.16 niet differentieerbaar is in 0.

8.23. Zij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}$, differentieerbaar in a. In (8.21) schreven we $f'(a)$ als een $1 \times n$ matrix, die we ook als (rij)vector in $\mathbb{R}^n$ opvatten. Deze vector heet de *gradiënt* van f in a en hij wordt genoteerd als

$$(\nabla f)(a) := ((D_1 f)(a), \dots, (D_n f)(a)). \tag{8.24}$$

(Het symbool ∇ wordt uitgesproken als "nabla".) We zien dan dat de richtingsafgeleide (8.23) als inwendig product kan worden geschreven:

$$(D_u f)(a) = \langle u, (\nabla f)(a) \rangle. \tag{8.25}$$

Opgave 8.4. Zij $c \in \mathbb{R}^n$ en definieer $f(x) := \langle x, c \rangle$ $(x \in \mathbb{R}^n)$. Merk op dat de functie $f : \mathbb{R}^n \to \mathbb{R}$ lineair is en homogeen van graad 1. Constateer dat $(\nabla f)(0)$ bestaat en bepaal deze gradiënt.

Opgave 8.5. Zij $f : \mathbb{R}^n \to \mathbb{R}$ continu en homogeen van graad $\alpha > 0$.

(a) Bewijs dat $f(0) = 0$.

(b) Bewijs dat f differentieerbaar is in 0 met afgeleide 0 als $\alpha > 1$.

(c) Zij f niet identiek gelijk aan 0. Bewijs dat, voor $0 < \alpha < 1$ en $0 \neq u \in \mathbb{R}^n$ de richtingsafgeleide $(D_u f)(0)$ niet bestaat. Concludeer hieruit dat f niet differentieerbaar is in 0 als $0 < \alpha < 1$.

(d) Zij $\alpha = 1$ en $0 \neq u \in \mathbb{R}^n$. Bewijs dat de richtingsafgeleide $(D_u f)(0)$ bestaat desda $f(-u) = -f(u)$ en dat dan $(D_u f)(0) = f(u)$. Concludeer hieruit dat, als $\alpha = 1$, f niet differentieerbaar is in 0 behalve als f lineair is (cf. Opmerking 8.20 en Opgave 8.4).

8.24. Bekijk nu het geval $n = 1$ van Stelling 8.18, dus $a \in E \subset \mathbb{R}$, E open en $f : E \to \mathbb{R}^m$ differentieerbaar in a. Dan is $f'(a)$ een lineaire afbeelding van $\mathbb{R}$ naar $\mathbb{R}^m$ met $m \times 1$ matrix

$$f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix}.$$

Zo'n $m \times 1$ matrix kan ook worden opgevat als een kolomvector met m coördinaten, dus als een element van $\mathbb{R}^m$. Beide interpretaties van $f'(a)$, als element van $L(\mathbb{R}, \mathbb{R}^m)$ en als element van $\mathbb{R}^m$ zullen we door elkaar gebruiken.

Opgave 8.6. Zij $a \in E \subset \mathbb{R}$, E open en $f : E \to \mathbb{R}^m$ differentieerbaar in a. Bewijs dat $f'(a)$, opgevat als element van $\mathbb{R}^m$, ook verkregen kan worden uit

$$f'(a) = \lim_{\substack{h \to 0 \\ h \in \mathbb{R}}} \frac{f(a+h) - f(a)}{h}. \tag{8.26}$$

Dus de afgeleide van de *vectorwaardige functie* f afhangend van één reële variabele is een vector en de formule voor deze afgeleide is exact dezelfde als formule (8.1) voor de afgeleide van een scalaire functie afhangend van één reële variabele.

8.25. We recapituleren nu de verschillende soorten afgeleiden die we in dit hoofdstuk ontmoet hebben. Zij $a \in E \subset \mathbb{R}^n$, E open, $f : E \to \mathbb{R}^m$. Dan onderscheiden we:

(1) *totale afgeleide*: $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$, zie §8.11, 8.14, 8.18.

(2) $(m = 1)$ *gradiënt*: $f'(a) = (\nabla f)(a) \in \mathbb{R}^n$ (rijvector), zie §8.20, 8.23.

(3) $(n = 1)$ *afgeleide van vectorwaardige functie*: $f'(a) \in \mathbb{R}^m$ (kolomvector), zie §8.24.

(4) $(m = 1)$ *richtingsafgeleide*: $(D_u f)(a) \in \mathbb{R}$ $(u \in \mathbb{R}^n)$, zie §8.4, 8.20, 8.23.

(5) $(m = 1)$ *partiële afgeleide*: $(D_j f)(a) \in \mathbb{R}$ $(j = 1, \ldots, n)$, zie §8.1, 8.18, 8.20, 8.23.

Opgave 8.7. Zij $a \in E \subset \mathbb{R}^n$ open. Zij $f, g : E \to \mathbb{R}^m$ differentieerbaar in a. Bewijs dat, voor $\lambda, \mu \in \mathbb{R}$, $(\lambda f + \mu g)'(a)$ bestaat en gelijk is aan $\lambda f'(a) + \mu g'(a)$. Neem nu aan dat $m = 1$. Bewijs dat ook fg en f/g differentieerbaar zijn in a (in het laatste geval mits $g(a) \neq 0$) en dat de afgeleiden gegeven worden door

$$(fg)'(a) = g(a)\, f'(a) + f(a)\, g'(a), \quad (f/g)'(a) = \frac{g(a)\, f'(a) - f(a)\, g'(a)}{g(a)^2}. \tag{8.27}$$

Geef analoge formules voor gradiënten $(\nabla(fg))(a)$, $(\nabla(f/g))(a)$ en voor de richtings- en partiële afgeleiden van fg en f/g.

8.26 Propositie. Zij $E \subset \mathbb{R}^n$ open en convex, i.e. voor iedere $x, y \in E$ is het lijnstuk $[x, y]$ dat x en y verbindt, bevat in E, (cf. K.A. Ross, Elementary Analysis, Stelling 22.5, Voorbeeld 3). Zij $f : E \to \mathbb{R}^m$ differentieerbaar op E en veronderstel dat $f'(x) = 0$ voor alle $x \in E$. Dan is er een $c \in \mathbb{R}^m$ zo dat $f(x) = c$ $(x \in E)$.

Bewijs. Het is voldoende de propositie te bewijzen voor $m = 1$. Het is ook voldoende te bewijzen dat f constant is op elk lijnstuk $[x, y]$ binnen E. Vanwege (8.23) zullen alle richtingsafgeleiden $D_u f$ gelijk 0 zijn op E. Veronderstel dat $x, y \in E$. Dan $x + t(y - x) \in E$ voor t in zeker open interval $I \supset [0, 1]$. Dus, voor $t \in I$,

$$0 = (D_{y-x}f)(x + t(y - x)) = \frac{d}{ds} f(x + t(y - x) + s(y - x))\Big|_{s=0} = \frac{d}{ds} f(x + s(y - x))\Big|_{s=t}.$$

Hierbij gebruikten we (8.6)(8.6). Blijkbaar heeft de functie $t \mapsto f(x + t(y - x))$ afgeleide 0 op I, dus deze functie is constant op I. □

8.27 Gevolg. Zij $E \subset \mathbb{R}^n$ open en convex en zij $g : E \to L(\mathbb{R}^n, \mathbb{R}^m)$. Als er een differentieerbare afbeelding $f : E \to \mathbb{R}^m$ bestaat zo dat $f'(x) = g(x)$ voor alle $x \in E$, dan is f uniek door g bepaald op een constante term na.

Opgave 8.8. Zij $f : \mathbb{R}^n \to \mathbb{R}^m$ differentieerbaar zo dat $f'(x) = A \in L(\mathbb{R}^n, \mathbb{R}^m)$ voor $x \in \mathbb{R}^n$ (dus constant). Bewijs dat er $b \in \mathbb{R}^m$ bestaat zo dat $f(x) = Ax + b$ ($x \in \mathbb{R}^n$).

Opgave 8.9. Zij $E \subset \mathbb{R}^n$ open. We noemen E lineair samenhangend als er voor elke x en y in E eindig veel punten $z_1, \ldots, z_k$ te vinden zijn zo dat alle lijnstukken $[x, z_1]$, $[z_1, z_2], \ldots, [z_{k-1}, z_k]$, $[z_k, y]$ ook in E liggen.

Zij E lineair samenhangend. Zij $f : E \to \mathbb{R}^m$ differentieerbaar op E en veronderstel dat $f'(x) = 0$ voor alle $x \in E$. Bewijs dat er een $c \in \mathbb{R}^m$ is zo dat $f(x) = c$ ($x \in E$).
Aanwijzing Gebruik het bewijs van Propositie 8.26.

Opgave 8.10 (Euler)**.** Zij $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ differentieerbaar Zij $\alpha \in \mathbb{R}$. Bewijs dat f homogeen is van graad α desda

$$\sum_{j=1}^{n} x_j \, (D_j f)(x) = \alpha \, f(x) \quad (x \in \mathbb{R}^n \backslash \{0\}). \tag{8.28}$$

Aanwijzing Bewijs eerst dat $\frac{d}{dr} f(rx)|_{r=1} = (D_x f)(x)$ (richtingsafgeleide) en vervolgens dat $\frac{d}{dr}(r^{-\alpha} f(rx))|_{r=1} = \sum_{j=1}^n x_j \, (D_j f)(x) = \alpha \, f(x)$.

8.28 Definitie. Zij $E \subset \mathbb{R}^n$ open. We zeggen dat een afbeelding $f : E \to \mathbb{R}^m$ *continu differentieerbaar* is op E als f differentieerbaar is op E en de afgeleide $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$ een continue afbeelding is. Hier wordt $L(\mathbb{R}^n, \mathbb{R}^m)$ beschouwd als metrische ruimte t.o.v. de operatornorm.

Continue differentieerbaarheid van f wordt ook aangeduid door te zeggen dat f een C^1-*afbeelding* is of dat $f \in C^1(E)$.

We zagen al in Sectie 1.3 dat de ruimte $L(\mathbb{R}^n, \mathbb{R}^m)$ geïdentificeerd kan worden met $\mathbb{R}^{mn}$, waarbij de coördinaten van $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ gegeven worden door de matrixelementen A_{ij} van A. Door toepassing van Gevolg 7.8(c) en (8.18) zien we dan het volgende in.

8.29 Propositie. Zij $E \subset \mathbb{R}^n$ open en zij $f : E \to \mathbb{R}^m$ differentieerbaar op E. Dan is f' continu op E desda elk van de partiële afgeleiden $D_j f_i$ continu is op E.

De volgende stelling geeft een beduidende en zeer bevredigende verscherping van de vorige Propositie.

8.30 Stelling. Zij $E \subset \mathbb{R}^n$ open en $f : E \to \mathbb{R}^m$. Dan is f continu differentieerbaar op E desda alle partiële afgeleiden $D_j f_i$ bestaan op E en continu op E zijn.

Bewijs. Opgave 8.2 geeft de directe implicatie ($\Rightarrow$). Voor de andere richting bewijzen we: "Als voor een afbeelding $f : E \to \mathbb{R}$ de partiële afgeleiden $D_j f$ op E bestaan en continu zijn, dan is f differentieerbaar op E." De vorige Propositie geeft dan dat f continu differentieerbaar is op E.

Zij $a \in E$ en neem $r > 0$ zodanig dat $x \in E$ als $|x - a| < r$. Gezien de definitie van differentieerbaar, (8.15), en Stelling 8.18 zal de differentieerbaarheid van f in a volgen als we kunnen bewijzen dat

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - (h_1\,(D_1 f)(a) + \cdots + h_n\,(D_n f)(a))|}{|h|} = 0. \tag{8.29}$$

We moeten dus $f(a+h)-f(a)-(h_1\,(D_1f)(a)+\cdots+h_n\,(D_nf)(a))$ in absolute waarde afschatten als functie van h voor $|h| \to 0$. Neem $h \in \mathbb{R}^n$ met $|h| < r$ en schrijf $v_k := h_1e_1 + \cdots + h_ke_k$ ($k = 0, 1, \ldots, n$). Dan

$$f(a+h) - f(a) = \sum_{k=1}^{n} (f(a+v_k) - f(a+v_{k-1})) = \sum_{k=1}^{n} (f(a+v_{k-1}+h_ke_k) - f(a+v_{k-1})).$$

Pas nu voor elke $k \in \{1, \ldots, n\}$ de middelwaardestelling (zie K.A. Ross, Elementary Analysis, Stelling 29.3) toe op de functie $g_k(t) := f(a + v_{k-1} + te_k)$. Er zijn dus getallen $\theta_k \in (0, 1)$ ($k = 1, \ldots, n$) zo dat

$$f(a+v_{k-1}+h_ke_k) - f(a+v_{k-1}) = g_k(h_k) - g_k(0) = h_k g_k'(\theta_k h_k) = h_k(D_kf)(a+v_{k-1}+\theta_k h_k e_k).$$

Dan

$$f(a+h) - f(a) = \sum_{k=1}^{n} h_k\,(D_kf)(a + v_{k-1} + \theta_k h_k e_k).$$

Dus

$$f(a+h) - f(a) - \sum_{k=1}^{n} h_k\,(D_kf)(a) = \sum_{k=1}^{n} h_k\,\big((D_kf)(a+v_{k-1}+\theta_kh_ke_k) - (D_kf)(a)\big).$$

Met de driehoeksongelijkheid en $|h_k| \le |h|$ volgt

$$\begin{aligned} &\frac{|f(a+h) - f(a) - (h_1\,(D_1f)(a) + \cdots + h_n\,(D_nf)(a))|}{|h|} \\ &\le \sum_{k=1}^{n} \big|(D_kf)(a+v_{k-1}+\theta_kh_ke_k) - (D_kf)(a)\big|. \end{aligned} \tag{8.30}$$

Zij $\varepsilon > 0$. Vanwege de continuïteit van de partiële afgeleiden is er een $\delta > 0$ zo dat, voor elke $k \in \{1, \ldots, n\}$, geldt dat $|(D_kf)(x) - (D_kf)(a)| < \varepsilon/n$ als $|x - a| < \delta$. Neem $|h| < \delta$. Dan

$$|v_{k-1} + \theta_k h_k e_k| = \left(\sum_{j=1}^{k-1} |h_j|^2 + (\theta_k h_k)^2\right)^{1/2} \le |h| < \delta,$$

dus het rechterlid van (8.30) is kleiner dan $\sum_{k=1}^{n}(\varepsilon/n) = \varepsilon$. Dit laat zien dat de limiet (8.29) geldt. □

8.31. Bovenstaande stelling geeft een gemakkelijk criterium voor differentieerbaarheid, dat o.a. toepasbaar is als f gegeven is door een mooie analytische uitdrukking in de variabelen $x_1, \ldots, x_n$. Bijvoorbeeld, als f een polynoom is in $x_1, \ldots, x_n$ dan zijn de partiële afgeleiden weer polynomen, dus continu, dus f is dan differentieerbaar, cf. Voorbeeld 8.2.

Opgave 8.11. Zij $f : \mathbb{R} \to \mathbb{R}$ een twee keer continu differentieerbare functie. Zij

$$F(x,y) := \frac{f(x) - f(y)}{x - y} \quad \text{als } x \neq y \quad \text{en zij } F(x,x) := f'(x).$$

Bewijs dat $F : \mathbb{R}^2 \to \mathbb{R}$ een C^1-functie is.
Aanwijzing Bewijs eerst dat $(D_1F)(a,a) = \frac{1}{2} f''(a) = (D_2F)(a,a)$.
Bereken nu $(D_1F)(x,y)$ voor $x \neq y$ en concludeer met behulp van de Taylorontwikkeling met restterm van de functie f rond het punt x dat er voor elke x, y een $\theta \in (0,1)$ is zo dat $(D_1F)(x,y) = \frac{1}{2} f''(x + \theta(y-x))$. Laat nu zien dat D_1F overal continu is. Een soortgelijk argument geldt voor D_2F.

Opgave 8.12. Zij $f : \mathbb{R}^2 \to \mathbb{R}$ gegeven door $f(x,y) = \dfrac{e^x - e^y}{x - y}$ als $x \neq y$ en $f(x,x) = e^x$ als $x \in \mathbb{R}$. Bewijs dat f een C^1-functie is en bereken f'.

Verdere vraagstukken

Opgave 8.13. Bepaal de partiële afgeleiden van $f : E \to \mathbb{R}$ als

(a) $E := \mathbb{R}^3$ en $f(x,y,z) := \sin(xyz)$

(b) $E := \mathbb{R}^3 \backslash \{0\}$ en $f(x,y,z) := \sqrt{x^2 + y^2 + z^2}$

(c) $E := \{(x,y,z) \in \mathbb{R}^3 \mid 0 < |(x,y,z)| < \pi/2\}$ en $f(x,y,z) := \tan\left(\sqrt{x^2 + y^2 + z^2}\right)$

(d) $E := \{(x,y,z) \in \mathbb{R}^3 \mid x > 0,\ y > 0,\ z > 0\}$ en $f(x,y,z) := x^y\, y^z\, z^x$

Opgave 8.14. Voor $f : \mathbb{R}^2 \to \mathbb{R}$ is gegeven:

$$f(0,1) = -\frac{\pi}{4}, \qquad (D_1f)(x,y) = y\, e^x \quad \text{en} \quad (D_2f)(x,y) = e^x - \frac{1}{1+y^2}.$$

Bepaal f.

Opgave 8.15. Ga van de volgende functies $f : \mathbb{R}^2 \to \mathbb{R}$ na of ze differentieerbaar zijn in $(0,0)$:

(a) $f(x,y) := \log(1 + x^2y^2)$

(b) $f(x,y) := \log(1 + |xy|)$

(c) $f(x,y) := \log\left(1 + \sqrt{|xy|}\right)$

Opgave 8.16. Zij $f : \mathbb{R}^2 \to \mathbb{R}$ met $f(0,0) = 0$ en $f(x,y)$ voor $(x,y) \neq (0,0)$ gegeven door:

(a) $f(x,y) := \dfrac{x^2y^2}{x^4 + y^4}$

(b) $f(x,y) := \dfrac{x^3 - y^3}{x^2 + y^2}$

(c) $f(x,y) := xy \sin \dfrac{1}{x^2 + y^2}$

(d) $f(x,y) := \dfrac{x^2 \sin y^2}{x^2 + y^4}$

(e) $f(x,y) := \dfrac{x^3}{x^2 + (x^2+y^2)^2}$

(f) $f(x,y) := x^2 \log(x^4 + y^2)$

Ga voor elk van deze functies na of (i) f continu is in $(0,0)$ en (ii) f differentieerbaar is in $(0,0)$.

Opgave 8.17. Zij $f : \mathbb{R}^2 \to \mathbb{R}$ met $f(0,0) = 0$ en $f(x,y)$ voor $(x,y) \neq (0,0)$ gegeven door:

(a) $f(x,y) := \dfrac{x^3}{x^2+y^2}$

(b) $f(x,y) := \dfrac{x^4}{x^2+y^2}$

(c) $f(x,y) := \dfrac{x^2 y}{x^2+y^4}$

(d) $f(x,y) := \dfrac{x^3}{x^2+y^4}$

Beantwoord voor elk van deze functies de volgende vragen:

(i) Is f continu in $(0,0)$?

(ii) Bestaan de partiële afgeleiden $D_1 f$ en $D_2 f$ in $(0,0)$?

(iii) Voor welke $u \in \mathbb{R}^2$ $(u \neq 0)$ bestaat de richtingsafgeleide $(D_u f)(0,0)$?

(iv) Is f differentieerbaar in $(0,0)$?

(v) Is f een C^1-functie?

Opgave 8.18. Toon aan dat de volgende functies $f : \mathbb{R}^2 \to \mathbb{R}$ C^1-functies zijn en bereken f'.

(a) $f(x,y) := xy \log(x^2+y^2)$ als $(x,y) \neq (0,0)$ en $f(0,0) := 0$

(b) $f(x,y) := \dfrac{\log(1 + x^2y^2)}{x^2}$ als $x \neq 0$ en $f(0,y) := y^2$

(c) $f(x,y) := \dfrac{e^{x^2-y^4} - 1}{x - y^2}$ als $x \neq y^2$ en $f(y^2, y) := 2y^2$.

Opgave 8.19. Zij $f : \mathbb{R}^n \to \mathbb{R}$.

(a) Laat $f(x) := |x|$. Toon aan dat f differentieerbaar is op $\mathbb{R}^n \backslash \{0\}$, dat f niet differentieerbaar is in 0, en dat $f'(a) = |a|^{-1}\, a$ als $a \neq 0$.

(b) Laat $f(x) := |x|^2$. Bereken $f'(a)$.

(c) Laat voor zekere $C > 0$ gelden dat $|f(x)| \leq C\,|x|^2$ $(x \in \mathbb{R}^n)$. Toon aan dat f differentieerbaar is in 0 en bereken $f'(0)$.

9 De kettingregel

9.1 *Formulering en bewijs van de kettingregel*

Als we bij functies in één variabele een samengestelde functie willen differentiëren, dan is de kettingregel het passende gereedschap. Hieronder geven we een analogon van de kettingregel voor functies in meer veranderlijken.

9.1 Stelling (kettingregel)**.** Zij $a \in E \subset \mathbb{R}^n$, E open, $f : E \to \mathbb{R}^m$, $f(E) \subset U \subset \mathbb{R}^m$, U open, $g : U \to \mathbb{R}^l$, $F := g \circ f$. Als f differentieerbaar is in a en g differentieerbaar in $f(a)$ dan is F differentieerbaar in a en

$$F'(a) = g'(f(a))\, f'(a). \tag{9.1}$$

Bewijs. Op grond van de gegevens beeldt F de open deelverzameling E van $\mathbb{R}^n$ af in $\mathbb{R}^l$. Ook hebben we de lineaire afbeeldingen $f'(a) : \mathbb{R}^n \to \mathbb{R}^m$ en $g'(f(a)) : \mathbb{R}^m \to \mathbb{R}^l$, dus $g'(f(a))\, f'(a)$ is lineair van $\mathbb{R}^n$ naar $\mathbb{R}^l$. Differentieerbaarheid van F in a met afgeleide $g'(f(a))\, f'(a)$ zal volgen als we kunnen bewijzen dat

$$F(a+h) - F(a) - g'(f(a))\, f'(a)\, h = o(|h|), \quad h \to 0. \tag{9.2}$$

We moeten dus $F(a+h) - F(a) - g'(f(a))\, f'(a)\, h$ in norm afschatten als functie van h voor $|h| \to 0$. We kunnen deze uitdrukking herschrijven als:

$$\begin{aligned} &F(a+h) - F(a) - g'(f(a))\, f'(a)\, h \\ &\qquad = g(f(a+h)) - g(f(a)) - g'(f(a))\,(f(a+h) - f(a)) \\ &\qquad\quad + g'(f(a))\,\bigl(f(a+h) - f(a) - f'(a)\, h\bigr). \end{aligned}$$

Dus

$$\begin{aligned} &|F(a+h) - F(a) - g'(f(a))\, f'(a)\, h| \\ &\qquad \le \bigl|g(f(a+h)) - g(f(a)) - g'(f(a))\,(f(a+h) - f(a))\bigr| \\ &\qquad\quad + \|g'(f(a))\|\, \bigl|f(a+h) - f(a) - f'(a)\, h\bigr|. \end{aligned}$$

Dus (9.2) zal volgen als we kunnen bewijzen:

$$g(f(a+h)) - g(f(a)) - g'(f(a))\,(f(a+h) - f(a)) = o(|h|), \quad h \to 0, \tag{9.3}$$
$$f(a+h) - f(a) - f'(a)\, h = o(|h|), \quad h \to 0. \tag{9.4}$$

Nu geldt (9.4) omdat f differentieerbaar is in a met afgeleide $f'(a)$. En (9.3) zullen we bewijzen door te gebruiken dat g differentieerbaar is in $f(a)$ met afgeleide $g'(f(a))$, dus dat

$$|g(f(a)+k) - g(f(a)) - g'(f(a))\, k| = o(|k|) \quad \text{als} \quad k \to 0 \quad \text{in} \quad \mathbb{R}^m. \tag{9.5}$$

Wegens Opgave 8.15 en de differentieerbaarheid van f in a bestaan er γ en $C > 0$ zo dat, als $h \in \mathbb{R}^n$ en $|h| < \gamma$, dan $a + h \in E$ en

$$|f(a+h) - f(a)| \le (\|f'(a)\| + 1)\, |h|. \tag{9.6}$$

Laat nu $\varepsilon > 0$. Wegens (9.5) is er een $\delta > 0$ zo dat, als $|k| < \delta$ dan $f(a) + k \in U$ en

$$|g(f(a)+k) - g(f(a)) - g'(f(a))\,k| < \varepsilon\,|k|. \tag{9.7}$$

Door γ nog kleiner te nemen, kunnen we ervoor zorgen dat $\gamma < \delta/(\|f'(a)\|+1)$. Dan is $|f(a+h) - f(a)| < \delta$ als $|h| < \gamma$. Substitueer nu $k := f(a+h) - f(a)$ in (9.7) en combineer de resulterende ongelijkheid met ongelijkheid (9.6). Dan volgt er dat, als $|h| < \gamma$, dan

$$|g(f(a+h)) - g(f(a)) - g'(f(a))\,(f(a+h) - f(a))| < \varepsilon\,(\|f'(a)\| + 1)\,|h|.$$

Omdat ε willekeurig genomen was, impliceert dit (9.3). □

9.2. Formule (9.1) staat bekend als de *kettingregel* en ziet er formeel hetzelfde uit als de kettingregel voor differentieerbare functies van één variabele. Echter, het rechterlid van (9.1) moeten we lezen als de vermenigvuldiging (compositie) van twee lineaire afbeeldingen.

Formule (9.1) ziet er, in matrixvorm uitgeschreven, als volgt uit.

$$\begin{pmatrix} (D_1F_1)(a) & \ldots & D_nF_1)(a) \\ \vdots & & \vdots \\ (D_1F_l)(a) & \ldots & (D_nF_l)(a) \end{pmatrix}$$

$$= \begin{pmatrix} (D_1g_1)(f(a)) & \ldots & (D_mg_1)(f(a)) \\ \vdots & & \vdots \\ (D_1g_l)(f(a)) & \ldots & (D_mg_l)(f(a)) \end{pmatrix} \begin{pmatrix} (D_1f_1)(a) & \ldots & (D_nf_1)(a) \\ \vdots & & \vdots \\ (D_1f_m)(a) & \ldots & (D_nf_m)(a) \end{pmatrix} \tag{9.8}$$

Per matrixcomponent wordt dit

$$(D_jF_i)(a) = \sum_{k=1}^{m} (D_jf_k)(a)\,(D_kg_i)(f(a)), \quad i = 1, \ldots, l, \quad j = 1, \ldots, n. \tag{9.9}$$

Hier hebben we de factoren in de sommatie rechts in de volgorde geschreven die in latere toepassingen handig is.

9.3. In het geval $l = 1$ van Stelling 9.1 zijn de functies g en $F := g \circ f$ reëelwaardig en we hebben de formules

$$F(x_1, \ldots, x_n) = g(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \quad (x \in E), \tag{9.10}$$

$$(D_jF)(a) = \sum_{k=1}^{m} (D_jf_k)(a)\,(D_kg)(f(a)), \quad (j = 1, \ldots, n). \tag{9.11}$$

Als f differentieerbaar is op E en g differentieerbaar op U, dan kunnen we (9.11) ook schrijven in de vorm

$$\frac{\partial}{\partial x_j} g(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) = \sum_{k=1}^{m} \frac{\partial f_k(x_1, \ldots, x_n)}{\partial x_j}\, \frac{\partial g(y_1, \ldots, y_m)}{\partial y_k}\Big|_{y=f(x)}. \tag{9.12}$$

9.4. In het geval $l = 1$ en $n = 1$ van Stelling 9.1 zijn de functies g en F reëelwaardig en hangen de functies f en F van slechts één reële variabele af. Dan nemen de formules (9.10), (9.11) en

(9.12) de volgende vorm aan.

$$F(x) = g(f_1 x), \ldots, f_m(x)) \quad (x \in E), \tag{9.13}$$

$$F'(a) = \sum_{k=1}^{m} f_k'(a)\,(D_k g)(f(a)), \tag{9.14}$$

$$\frac{d}{dx} g(f_1(x), \ldots, f_m(x)) = \sum_{k=1}^{m} \frac{df_k(x)}{dx}\, \frac{\partial g(y_1, \ldots, y_m)}{\partial y_k}\Big|_{y=f(x)}. \tag{9.15}$$

Formule (9.14) komt in de praktijk zo vaak voor dat het de moeite waard is om deze apart te onthouden en niet telkens weer af te leiden uit formule (9.1). We formuleren Stelling 9.1 nog eens in het geval $l = 1$ en $n = 1$.

9.5 Stelling. Zij $a \in E \subset \mathbb{R}$, E open, $f : E \to \mathbb{R}^m$, $f(E) \subset U \subset \mathbb{R}^m$, U open, $g : U \to \mathbb{R}$, $F = g \circ f : E \to \mathbb{R}$. Als f differentieerbaar is in a en g differentieerbaar in $f(a)$ dan is F differentieerbaar in a en geldt formule (9.14).

Opgave 9.1. Laten f_1 en f_2 differentieerbare functies zijn van $\mathbb{R}$ naar $\mathbb{R}$ en schrijf $F(x) := f_1(x)\, f_2(x)$. Het is welbekend dat dan $F'(x) = f_1'(x)\, f_2(x) + f_1(x)\, f_2'(x)$. Geef hiervan een nieuw bewijs door toepassing van (9.14), waarbij $m = 2$ en $g(y_1, y_2) := y_1\, y_2$. Voer een soortgelijke opdracht uit met betrekking tot de afgeleide van $f_1(x)/f_2(x)$.

Opgave 9.2. Laten n, m, l, a, E, U, f en g zijn zoals in Stelling 9.1. Neem aan dat $l = n$. Veronderstel dat $g(f(x)) = x$ $(x \in E)$.

(a) Bewijs dat $g'(f(a))\, f'(a) = I$ (identiteitsafbeelding op $\mathbb{R}^n$), dat $m \geq n$ en dat de $m \times n$ matrix $f'(a)$ rang n heeft.

(b) Neem bovendien aan dat $m = n$. Bewijs dat de lineaire afbeelding $f'(a) : \mathbb{R}^n \to \mathbb{R}^n$ inverteerbaar is en dat
$$g'(f(a)) = (f'(a))^{-1}. \tag{9.16}$$

9.2 Het effect van een coördinatentransformatie op de partiële afgeleiden

We bespreken nu met behulp van de kettingregel hoe partiële differentiaties transformeren bij een coördinatentransformatie. Eerst doen we dit in het algemene geval, dan voor het speciale geval van overgang van cartesische op poolcoördinaten of omgekeerd. In ieder geval moet je later de hieronder beschreven theorie in een concrete situatie kunnen uitvoeren.

9.6. Laten E_1 en E_2 open deelverzamelingen zijn van $\mathbb{R}^n$. Laten $\phi : E_1 \to E_2$ en $\psi : E_2 \to E_1$ C^1-afbeeldingen zijn zo dat

$$\phi \circ \psi = \text{id op } E_2 \quad \text{en} \quad \psi \circ \phi = \text{id op } E_1. \tag{9.17}$$

(Zo'n functie ϕ, bijectief en C^1 en met inverse die ook C^1 is, noemen we een *diffeomorfisme* of meer specifiek een C^1-*diffeomorfisme*.) Laten $F : E_1 \to \mathbb{R}$ en $f : E_2 \to \mathbb{R}$ C^1-afbeeldingen zijn zo dat

$$F(x) = f(\phi(x)) \quad (x \in E_1), \qquad \text{of equivalent} \qquad f(u) = F(\psi(u)) \quad (u \in E_2). \tag{9.18}$$

Herschrijving van (9.11) geeft het formulepaar

$$(D_jF)(x) = \sum_{k=1}^{n}(D_j\phi_k)(x)\,(D_kf)(\phi(x)) \quad (x \in E_1),$$

$$(D_kf)(u) = \sum_{j=1}^{n}(D_k\psi_j)(u)\,(D_jF)(\psi(u)) \quad (u \in E_2).$$

Door in de eerste formule $x := \psi(u)$ en in de tweede formule $u := \phi(x)$ te substitueren, verkrijgen we

$$(D_jF)(\psi(u)) = \sum_{k=1}^{n}(D_j\phi_k)(\psi(u))\,(D_kf)(u) \quad (u \in E_2), \tag{9.19}$$

$$(D_kf)(\phi(x)) = \sum_{j=1}^{n}(D_k\psi_j)(\phi(x))\,(D_jF)(x) \quad (x \in E_1). \tag{9.20}$$

Formule (9.19) drukt de partiële afgeleiden $(D_jF)(\psi(u))$ $(j = 1,\ldots,n)$ uit in termen van de partiële afgeleiden $(D_kf)(u)$ $(k = 1,\ldots,n)$ en (9.20) doet dit in omgekeerde richting. Een andere manier van schrijven van (9.19), 9.20 is als volgt.

$$\frac{\partial F(x)}{\partial x_j}\Big|_{x=\psi(u)} = \sum_{k=1}^{n}\frac{\partial \phi_k(x)}{\partial x_j}\Big|_{x=\psi(u)}\frac{\partial f(u)}{\partial u_k}\,, \tag{9.21}$$

$$\frac{\partial f(u)}{\partial u_k}\Big|_{u=\phi(x)} = \sum_{j=1}^{n}\frac{\partial \psi_j(u)}{\partial u_k}\Big|_{u=\phi(x)}\frac{\partial F(x)}{\partial x_j}\,. \tag{9.22}$$

We laten nu in (9.21) en (9.22) de functies f en F weg om de transformaties van de operatoren te benadrukken. We maken ons nu wel schuldig aan een schrijfwijze die eerder, in §8.1, is afgeraden. Soms is deze schrijfwijze in de praktijk echter toch wel handig.

$$\frac{\partial}{\partial x_j}\Big|_{x=\psi(u)} = \sum_{k=1}^{n}\frac{\partial \phi_k(x)}{\partial x_j}\Big|_{x=\psi(u)}\frac{\partial}{\partial u_k}\,, \tag{9.23}$$

$$\frac{\partial}{\partial u_k}\Big|_{u=\phi(x)} = \sum_{j=1}^{n}\frac{\partial \psi_j(u)}{\partial u_k}\Big|_{u=\phi(x)}\frac{\partial}{\partial x_j}\,. \tag{9.24}$$

Neem nu $n := 2$ en schrijf (x, y) i.p.v. (x_1, x_2) en (u, v) i.p.v. (u_1, u_2). Gebruik ook de notatie $(u(x, y), v(x, y))$ voor $\phi(x, y)$ en $(x(u, v), y(u, v))$ voor $\psi(u, v)$. Dan geeft (9.24) aanleiding tot de twee formules

$$\frac{\partial}{\partial u} = \frac{\partial x(u,v)}{\partial u}\Big|_{\substack{u=u(x,y)\\ v=v(x,y)}}\frac{\partial}{\partial x} + \frac{\partial y(u,v)}{\partial u}\Big|_{\substack{u=u(x,y)\\ v=v(x,y)}}\frac{\partial}{\partial y}\,, \tag{9.25}$$

$$\frac{\partial}{\partial v} = \frac{\partial x(u,v)}{\partial v}\Big|_{\substack{u=u(x,y)\\ v=v(x,y)}}\frac{\partial}{\partial x} + \frac{\partial y(u,v)}{\partial v}\Big|_{\substack{u=u(x,y)\\ v=v(x,y)}}\frac{\partial}{\partial y}\,. \tag{9.26}$$

Hierbij nemen we stilzwijgend aan dat de operatoren $\partial/\partial u$ en $\partial/\partial v$ in het linkerlid werken op $f(u, v)$ en dat het resultaat van de partiële differentiatie in het linkerlid geëvalueerd wordt voor $(u, v) = (u(x, y), v(x, y))$, terwijl de operatoren $\partial/\partial x$ en $\partial/\partial y$ in het rechterlid werken op $F(x, y)$. Hierbij is

$$f(u, v) = F(x(u, v), y(u, v)), \quad F(x, y) = f(u(x, y), v(x, y)).$$

9.7 Voorbeeld. Bekijk nu (9.25), (9.26) met (u, v) vervangen door (r, θ). Hiermee beschrijven we de overgang van poolcoördinaten (r, θ) naar Cartesische coördinaten (x, y):

$$\psi : (r, \theta) \mapsto (x, y) = (x(r, \theta), y(r, \theta)) := (r \cos \theta, r \sin \theta). \tag{9.27}$$

We willen dat $\psi : E_2 \to E_1$ een bijectieve C^1-afbeelding is voor zekere open deelverzamelingen E_1 en E_2 van $\mathbb{R}^2$ en dat de inverse $\phi : E_1 \to E_2$ van ψ eveneens C^1 is. We zouden bijvoorbeeld kunnen nemen

$$E_1 := \mathbb{R}^2 \backslash \{(x, 0) \mid x \leq 0\}, \quad E_2 := (0, \infty) \times (-\pi, \pi). \tag{9.28}$$

Dus E_1 is $\mathbb{R}^2$ met de negatieve x-as inclusief $(0, 0)$ er uit gesneden. Het is onvermijdelijk dat $(0, 0) \notin E_1$ (waarom?), maar desgewenst kan een andere halfrechte dan de negatieve x-as worden uitgesneden. Daartoe kunnen we E_2 verticaal verschuiven naar een willekeurige horizontale strook $(0, \infty) \times (-\pi + c, \pi + c)$ van wijdte 2π.

Om ϕ te bepalen moeten we (9.27) inverteren. Het is duidelijk dat $r = r(x, y) = \sqrt{x^2 + y^2}$. Dus voor $\theta = \theta(x, y)$ geldt dan dat

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}.$$

We kunnen geen gesloten uitdrukking voor θ zelf geven in termen van x en y welke geldig is op de hele E_1. Maar bijvoorbeeld op het rechterhalfvlak krijgen we

$$\theta = \theta(x, y) = \arcsin\Big(\frac{y}{\sqrt{x^2 + y^2}}\Big) \quad (x > 0).$$

Soortgelijke uitdrukkingen kunnen gegeven worden voor (x, y) in het bovenhalfvlak of in het benedenhalfvlak (ga na). Tezamen beschrijven deze uitdrukkingen $\theta(x, y)$ op de hele E_1. Hieruit zien we ook dat ϕ een C^1-afbeelding is (ga na). In de nu volgende berekeningen hebben we de expliciete uitdrukkingen voor $\theta(x, y)$ echter niet nodig.

Een eenvoudige berekening geeft dat

$$\begin{aligned}
\frac{\partial(r \cos \theta)}{\partial r} &= \cos \theta = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial(r \sin \theta)}{\partial r} = \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}, \\
\frac{\partial(r \cos \theta)}{\partial \theta} &= -r \sin \theta = -y, \quad \frac{\partial(r \sin \theta)}{\partial \theta} = r \cos \theta = x.
\end{aligned}$$

Dus er volgt uit (9.26)(9.26) dat

$$\frac{\partial}{\partial r} = \frac{x}{\sqrt{x^2 + y^2}} \frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2 + y^2}} \frac{\partial}{\partial y}, \tag{9.29}$$

$$\frac{\partial}{\partial \theta} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}. \tag{9.30}$$

Als we E_2 verticaal opschuiven en E_1 dienovereenkomstig veranderen dan blijven formules (9.29), (9.30) ongewijzigd (ga na). We kunnen de kunstmatige snede in E_1 elimineren door $\psi : E_2 \to E_1$ te beschouwen met

$$E_1 := \mathbb{R}^2 \backslash \{(0, 0)\} \quad E_2 := (0, \infty) \times \mathbb{R}, \tag{9.31}$$

waarbij ψ nog steeds gedefinieerd is door (9.27). Dan is ψ surjectief en C^1, maar niet langer injectief. Wel geldt dat $\psi(r, \theta) = \psi(r, \theta + 2\pi)$. Ook is ψ een *lokaal diffeomorfisme*, i.e., we kunnen bij elk punt van E_2 een open omgeving vinden (in dit geval bijvoorbeeld een horizontale strook met wijdte 2π) die door ψ diffeomorf wordt afgebeeld.

De formule
$$f(r,\theta)=F(r\cos\theta,r\sin\theta) \tag{9.32}$$
geeft nu een 1-1 correspondentie tussen C^1-functies F op E_1 en C^1-functies f op E_2 die 2π-periodiek zijn in hun tweede argument, i.e., die voldoen aan $f(r,\theta)=f(r,\theta+2\pi)$. Formules (9.29), (9.30) blijven nog steeds geldig. De stilzwijgende veronderstelling is dat de linkerleden van (9.29), (9.30) werken op $f(r,\theta)$ en de rechterleden op $F(x,y)$, waarbij f en F met elkaar verband houden volgens (9.32). We zien nu ook, vanwege de 2π-periodiciteit van $f(r,\theta)$ in θ, dat $\partial f(r,\theta)/\partial r$ en $\partial f(r,\theta)/\partial\theta$ eveneens 2π-periodiek zijn in θ.

Opgave 9.3. Laten F en f C^1-functies zijn op de door (9.31) gegeven verzamelingen E_1 resp. E_2 zo dat (9.32) geldt.

(a) Bewijs met behulp van (9.29) en (8.28) dat F homogeen is van graad α desda
$$r\,\frac{\partial f(r,\theta)}{\partial r}=\alpha\,f(r,\theta).$$

(b) Zij F homogeen van graad α. Definieer $G(x,y)$ door $G(r\cos\theta,r\sin\theta):=\partial f(r,\theta)/\partial\theta$. Bewijs met behulp van (9.30) dat G eveneens homogeen van graad α is.

9.8. Laten (r,θ) en (x,y) met elkaar verbonden zijn door (9.27). Om $\partial/\partial x$ en $\partial/\partial y$ uit te drukken in termen van $\partial/\partial r$ en $\partial/\partial\theta$ kunnen we twee strategieën volgen die beide zullen leiden tot:
$$\frac{\partial}{\partial x}=\cos\theta\,\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\,\frac{\partial}{\partial\theta}, \tag{9.33}$$
$$\frac{\partial}{\partial y}=\sin\theta\,\frac{\partial}{\partial r}+\frac{\cos\theta}{r}\,\frac{\partial}{\partial\theta}. \tag{9.34}$$

Bij de eerste strategie gaan we uit van (9.25), (9.26) en vervangen daarin (x,y) door (r,θ) en (u,v) door (x,y). We moeten dan de partiële afgeleiden van $r(x,y)$ en $\theta(x,y)$ naar x en y uitrekenen. Voor $r(x,y)$ is dit gemakkelijk, maar voor $\theta(x,y)$ wordt dit vervelend als we met arcsin-formules etc. zouden gaan werken. Voor $\partial\theta(x,y)/\partial x$ kunnen we bijv. als volgt redeneren. In plaats van $\partial\theta/\partial x$ beschouwen we $\partial(\tan\theta)/\partial x$. Enerzijds geldt
$$\frac{\partial(\tan\theta)}{\partial x}=\frac{\partial(y/x)}{\partial x}=\frac{-y}{x^2}=\frac{-\sin\theta}{r\cos^2\theta}$$
en anderzijds
$$\frac{\partial(\tan\theta)}{\partial x}=\frac{d(\tan\theta)}{d\theta}\,\frac{\partial\theta}{\partial x}=\frac{1}{\cos^2\theta}\,\frac{\partial\theta}{\partial x}\,.$$
Dus
$$\frac{1}{\cos^2\theta}\,\frac{\partial\theta}{\partial x}=\frac{-\sin\theta}{r\cos^2\theta}\,,\quad\text{dus}\quad\frac{\partial\theta}{\partial x}=\frac{-\sin\theta}{r}\,.$$
Strikt genomen is deze afleiding niet geldig voor $\theta=\pm\pi/2$, maar dit kunnen we ondervangen doordat we weten dat $\partial\theta(x,y)/\partial x$ continu is in (x,y) buiten $(0,0)$. (We zouden ook bovenstaande afleiding kunnen herhalen voor $\partial(\cot\theta)/\partial x$.)

Bij de tweede strategie herschrijven we (9.29), (9.30) in de volgende vorm.
$$\begin{pmatrix}\partial/\partial r\\ \partial/\partial\theta\end{pmatrix}=\begin{pmatrix}\frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}}\\ -y & x\end{pmatrix}\begin{pmatrix}\partial/\partial x\\ \partial/\partial y\end{pmatrix}.$$

Herschrijf nu de matrixelementen van bovenstaande matrix in termen van poolcoördinaten r, θ en inverteer de matrix. De resulterende vectorvergelijking is equivalent met (9.33), (9.34).

Opgave 9.4. Werk de details van beide geschetste methodes uit om (9.33) en (9.34) af te leiden.

Verdere vraagstukken

Opgave 9.5. Definieer
$f : \mathbb{R} \to \mathbb{R}^2$ door $f(x) := (x^2, x-3)$ en
$g : \mathbb{R}^2 \to \mathbb{R}$ door $g(x,y) := x^2 + xy + y^2$.
Bepaal $(g \circ f)'$ en $(f \circ g)'$ zowel rechtstreeks als m.b.v. de kettingregel.

Opgave 9.6. Definieer
$f : \mathbb{R}^2 \to \mathbb{R}^3$ door $f(x,y) = (x+y, y^2, 3y+1)$ en
$g : \mathbb{R}^2 \to \mathbb{R}^2$ door $g(x,y) = (x^2+y^2, xy)$.
Bereken $(f \circ g)'$ zowel rechtstreeks als m.b.v. de kettingregel.

Opgave 9.7. Laat $f : \mathbb{R}^2 \to \mathbb{R}$ een differentieerbare functie zijn. Zij $F : \mathbb{R}^2 \to \mathbb{R}$ gedefinieerd door $F(x,y) = f(x+y, x-y)$. Druk $(D_1F)(x,y) + (D_2F)(x,y)$ uit in termen van $(D_1f)(x+y, x-y)$ en $(D_2f)(x+y, x-y)$.

Opgave 9.8. Zij $f : \mathbb{R}^2 \to \mathbb{R}$ een differentieerbare functie. Definieer $F : (0,\infty) \times \mathbb{R} \to \mathbb{R}$ door $F(r,\theta) := f(r\cos\theta, r\sin\theta)$ (overgaan op poolcoördinaten). Bewijs dat

$$\left\{ \left(\frac{\partial f(x,y)}{\partial x} \right)^2 + \left(\frac{\partial f(x,y)}{\partial y} \right)^2 \right\} \Bigg|_{\substack{x = r\cos\theta \\ y = r\sin\theta}} = \left(\frac{\partial F(r,\theta)}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial F(r,\theta)}{\partial \theta} \right)^2 .$$

10 Meetkundige interpretaties van afgeleiden

10.1 Krommen

10.1. De term kromme wordt gebruikt als een meer meetkundige manier van beschrijven van afbeeldingen van een interval naar $\mathbb{R}^n$. Zij I een interval in $\mathbb{R}$. Een *(continue) kromme* (op I) in $\mathbb{R}^n$ is een continue afbeelding $\gamma : I \to \mathbb{R}^n$. De kromme γ heet *continu differentieerbaar* of C^1 als de afbeelding $\gamma : I \to \mathbb{R}^n$ een C^1-afbeelding is. (Als het interval I een begin- of eindpunt bevat, dan moeten we daar linker– of rechterafgeleide nemen.) Als we een kromme γ "tekenen" in $\mathbb{R}^n$ dan produceren we een deelverzameling $\gamma(I) = \{\gamma(t) \mid t \in I\}$ van $\mathbb{R}^n$: de beeldverzameling van γ. In het bijzonder zouden we zo krommen in het geval $n = 2$ kunnen tekenen, bijvoorbeeld de kromme $\gamma : t \mapsto (\cos t, \sin t) : \mathbb{R} \to \mathbb{R}^2$ levert zo de cirkel S^1 (of $\mathbb{T}$) met straal 1 en middelpunt $(0, 0)$ in $\mathbb{R}^2$. Ofschoon de beeldverzameling $\gamma(I)$ een belangrijk aspect van een kromme γ levert, is de *parametrisering* van de kromme, dus de manier waarop de punten van het interval I worden afgebeeld naar $\gamma(I)$, een onlosmakelijk onderdeel van het begrip kromme. Een tekening met volledige informatie over de kromme zou dus de beeldverzameling $\gamma(I)$ moeten geven, met bij elk punt x van $\gamma(I)$ de t-waarden uit de originelenverzameling $\gamma^{-1}(x)$ er bij geschreven.

Uiteraard kunnen we een kromme $\gamma : I \to \mathbb{R}^n$ ook grafisch volledig weergeven door haar *grafiek*, i.e. door de deelverzameling $\mathcal{G}(\gamma) := \{(t, \gamma(t)) \mid t \in I\}$ van $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$. Bijvoorbeeld bovenstaand voorbeeld van de kromme met de cirkel als beeldverzameling heeft als grafiek de spiraal $\{(t, \cos t, \sin t) \mid t \in \mathbb{R}\}$ in $\mathbb{R}^3$.

Soms beschouwt men twee krommen $\gamma : I_1 \to \mathbb{R}^n$ en $\delta : I_2 \to \mathbb{R}^n$ als *equivalent* als er een bijectieve continue monotoon stijgende afbeelding $\phi : I_1 \to I_2$ is (dus ϕ heeft dan ook een continue inverse) zo dat

$$\gamma(t) = \delta(\phi(t)) \quad (t \in I_1). \tag{10.1}$$

We spreken van een *herparametrisering* van de kromme. In het geval van C^1-krommen eisen we dan bovendien dat ϕ een C^1-afbeelding is met $\phi'(t) > 0$ (dus ook de inverse van ϕ is dan een C^1-afbeelding met strikt positieve afgeleide). Bijvoorbeeld bovenstaand voorbeeld van een kromme met $\gamma(I) = S^1$ is C^1-equivalent met de kromme $t \mapsto (\cos(t + t^3), \sin(t + t^3)) : \mathbb{R} \to \mathbb{R}^2$. Het gaat bij krommen dus vooral om de volgorde waarin de punten van $\gamma(I)$ successievelijk doorlopen worden.

Allerlei gedaantes zijn mogelijk voor een kromme. De afbeelding γ kan injectief zijn, maar de kromme kan zichzelf ook snijden, of zichzelf herhalen, zoals bij het voorbeeld van de cirkel. Vaak neemt men voor I een gesloten begrensd interval $[a, b]$ en zijn het begin- en eindpunt $\gamma(a)$ en $\gamma(b)$ van belang.

We noemen een kromme $\gamma : I \to \mathbb{R}^n$ *affien* als, voor zekere $v, w \in \mathbb{R}^n$ geldt dat $\gamma(t) = v + t\,w$ $(t \in I)$. Dan is $\gamma(I)$ dus een lijnstuk in $\mathbb{R}^n$.

Zij I een open interval in $\mathbb{R}$ en zij $\gamma : I \to \mathbb{R}^n$ een C^1-kromme. Dan (cf. §8.11) geldt, voor $a \in I$, dat

$$\gamma(t) = \gamma(a) + (t - a)\,\gamma'(a) + \rho(t) \quad (t \in I), \tag{10.2}$$

waarbij $\rho : I \to \mathbb{R}^n$ voldoet aan $\lim_{t\to a} |t - a|^{-1}\,|\rho(t)| = 0$. Volgens §8.24 kunnen we de afgeleide $\gamma'(a)$ opvatten als vector in $\mathbb{R}^n$. Merk op uit (10.2) dat de kromme γ nabij $t = a$ beschouwd kan worden als de som van de affiene kromme $t \mapsto \gamma(a) + (t - a)\,\gamma'(a)$ (die juist voor $t = a$ het punt $\gamma(a)$ bereikt) en de "verschilkromme" ρ waarbij $|\rho(t)|$ van orde kleiner dan $|t - a|$ is als $t \to a$. Dit rechtvaardigt om de lijn $\{\gamma(a) + (t - a)\,\gamma'(a) \mid t \in \mathbb{R}\}$ de *raaklijn* aan de kromme γ in het punt $\gamma(a)$ te noemen en de vector $\gamma'(a)$ de *raakvector* van de kromme γ in a (We houden deze terminologie zelfs aan als $\gamma'(a) = 0$.) De raakvector is dus een richtingsvector voor de raaklijn, maar ofschoon een richtingsvector van een rechte in $\mathbb{R}^n$ slechts bepaald is op vermenigvuldiging met een reële constante $\neq 0$ na, is de raakvector exact bepaald door de parametrisering van de kromme.

Een meer fysische interpretatie van een kromme $t \mapsto \gamma(t)$ krijgen we als we de parameter t als de tijd opvatten en $\gamma(I)$ als de baan in $\mathbb{R}^n$ van een puntdeeltje. Dan is de raakvector $\gamma'(a)$ de snelheidsvector van het deeltje ten tijde a.

Opgave 10.1. Laten γ en δ C^1-krommen in $\mathbb{R}^n$ zijn die in elkaar overgaan door een C^1-herparametrisering zoals in (10.1). Druk de raakvector van δ in $\phi(a)$ uit in termen van de de raakvector van γ in a. Concludeer dat voor C^1-equivalente C^1-krommen de raakvector bepaald is op vermenigvuldiging met een positieve reële constante na.

10.2 Definitie. Zij γ een C^1-kromme $\gamma : [a, b] \to \mathbb{R}^n$ gedefinieerd op een gesloten en begrensd interval. Associeer met elke partitie $P = P(x_0, x_1, \ldots, x_n)$ van $[a, b]$ (waarbij $a = x_0 < x_1 < \ldots < x_n = b$, cf. K.A. Ross, Elementary Analysis, Definitie 32.8) een getal

$$\Lambda(P, \gamma) := \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|. \tag{10.3}$$

Dan is de *lengte* van de kromme γ gedefinieerd door

$$\Lambda(\gamma) := \sup_{P \text{ partitie van } [a,b]} \Lambda(P, \gamma). \tag{10.4}$$

Deze lengte is dus mogelijk gelijk aan ∞. De kromme γ heet *rectificeerbaar* als $\Lambda(\gamma) < \infty$.

10.3. Meetkundig gezien stelt $\Lambda(P, \gamma)$ de lengte voor van een *ingeschreven polygoon*, het polygonale pad dat $\gamma(a)$ verbindt met $\gamma(b)$ via $\gamma(x_1), \gamma(x_2), \ldots, \gamma(x_{n-1})$, waarbij de punten $\gamma(x_i)$ in de goede volgorde op de kromme liggen. Volgens (10.4) is de lengte van de kromme dan het supremum van al deze benaderende lengtes.

Als de kromme γ bovendien C^1 is, dan zullen we zien dat γ rectificeerbaar is en dat we $\Lambda(\gamma)$ kunnen herschrijven als een Riemann-integraal. Eerst hebben we hiertoe nog een andere propositie nodig.

10.4. Voor eindige sommen van vectoren v_i in $\mathbb{R}^n$ geeft de driehoeksongelijkheid dat

$$|v_1 + \cdots + v_k| \leq |v_1| + \cdots + |v_k|. \tag{10.5}$$

In zekere zin zijn Riemann-integralen op te vatten als continue sommen. Daarom zal het niet verbazen dat er een analogon van (10.5) geldt voor vectorwaardige functies:

10.5 Propositie. Zij $f : [a, b] \to \mathbb{R}^n$ continu. Definieer $\int_a^b f(x)\, dx$ als het element van $\mathbb{R}^n$ met i-de coördinaat

$$\left(\int_a^b f(x)\, dx \right)_i := \int_a^b f_i(x)\, dx.$$

Dan geldt dat

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx. \tag{10.6}$$

Bewijs. Merk op dat alle Riemann-integralen goed gedefinieerd zijn omdat de functies f_i continu zijn en daardoor ook de functie $x \mapsto |f(x)| = \sqrt{(f_1(x))^2 + \cdots (f_n(x))^2}$. Schrijf $y_i := \int_a^b f_i(x)\,dx$ en $y := (y_1, \ldots, y_n)$. Dan

$$|y|^2 = \sum_{i=1}^n y_i^2 = \sum_{i=1}^n y_i \int_a^b f_i(x)\,dx = \int_a^b (y_1 f_1(x) + \cdots + y_n f_n(x))\,dx.$$

De ongelijkheid van Schwarz geeft dat

$$y_1 f_1(x) + \cdots + y_n f_n(x) \le |y|\,|f(x)|.$$

Dus, met behulp van K.A. Ross, Elementary Analysis, Stelling 33.4 en 33.3 volgt er dat

$$|y|^2 = \int_a^b (y_1 f_1(x) + \cdots + y_n f_n(x))\,dx \le \int_a^b |y|\,|f(x)|\,dx = |y| \int_a^b |f(x)|\,dx.$$

Als $|y| \neq 0$ dan mogen we door y delen en verkrijgen we $|y| \le \int_a^b |f(x)|\,dx$, i.e., (10.6). Als $|y| = 0$ dan is (10.6) triviaal. □

10.6 Stelling. Zij γ een C^1-kromme $\gamma : [a,b] \to \mathbb{R}^n$ gedefinieerd op een gesloten en begrensd interval. Dan is γ rectificeerbaar en er geldt dat

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\,dt = \int_a^b \sqrt{\gamma_1'(t)^2 + \cdots + \gamma_n'(t)^2}\,dt. \tag{10.7}$$

Bewijs. Zij $P = P(x_0, \ldots, x_n)$ een partitie van $[a,b]$. Dan geldt er met behulp van Propositie 10.5 dat

$$|\gamma(x_i) - \gamma(x_{i-1})| = \left| \int_{x_{i-1}}^{x_i} \gamma'(t)\,dt \right| \le \int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt \quad (i = 1, \ldots, n).$$

Dus, met behulp van (10.3),

$$\Lambda(P, \gamma) \le \int_a^b |\gamma'(t)|\,dt.$$

Dan geeft (10.4) dat

$$\Lambda(\gamma) \le \int_a^b |\gamma'(t)|\,dt,$$

waarmee we al hebben laten zien dat γ rectificeerbaar is. Formule (10.7) zal nu volgen als we ook de omgekeerde ongelijkheid kunnen bewijzen.

Zij $\varepsilon > 0$. De afbeelding $\gamma' : [a,b] \to \mathbb{R}^n$ is continu op een gesloten en begrensd (dus compact) interval en daarom uniform continu (Vergelijk K.A. Ross, Elementary Analysis, stelling 21.4). Dus er bestaat een $\delta > 0$ zo dat $|\gamma'(s) - \gamma'(t)| < \varepsilon$ als $|s - t| < \delta$. Neem nu een partitie $P(x_0, x_1, \ldots, x_n)$ van $[a,b]$ zo dat $x_i - x_{i-1} < \delta$ voor alle i. Dus

$$|\gamma'(t) - \gamma'(x_i)| < \varepsilon \quad \text{en} \quad |\gamma'(t)| \le |\gamma'(x_i)| + \varepsilon \quad \text{als } x_{i-1} \le t \le x_i.$$

Hieruit volgt dat

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt &\le (|\gamma'(x_i)| + \varepsilon)\,(x_i - x_{i-1}) = \left|\int_{x_{i-1}}^{x_i} \gamma'(x_i)\,dt\right| + \varepsilon(x_i - x_{i-1}) \\
&= \left|\int_{x_{i-1}}^{x_i} (\gamma'(t) + \gamma'(x_i) - \gamma'(t))\,dt\right| + \varepsilon(x_i - x_{i-1}) \\
&\le \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\,dt\right| + \left|\int_{x_{i-1}}^{x_i} (\gamma'(x_i) - \gamma'(t))\,dt\right| + \varepsilon\,(x_i - x_{i-1}) \\
&\le |\gamma(x_i) - \gamma(x_{i-1})| + 2\varepsilon\,(x_i - x_{i-1}).
\end{aligned}$$

Sommering over i van deze ongelijkheden levert

$$\int_a^b |\gamma'(t)|\,dt \le \Lambda(P,\gamma) + 2\varepsilon\,(b-a) \le \Lambda(\gamma) + 2\varepsilon\,(b-a).$$

Omdat ε willekeurig gekozen was, volgt er dat

$$\int_a^b |\gamma'(t)|\,dt \le \Lambda(\gamma).$$

□

Opgave 10.2. Zij γ als in Stelling 10.6, zij $\delta : [c,d] \to \mathbb{R}^n$ ook een C^1-kromme en laten γ en δ C^1-equivalent zijn d.m.v. $\phi : [a,b] \to [c,d]$ als in (10.1). Bewijs dat $\Lambda(\gamma) = \Lambda(\delta)$, m.a.w., C^1-equivalente krommen hebben dezelfde lengte.

10.7. We zullen nu het verband leggen tussen de analytisch (dwz. door de machtreeksdefinitie) ingevoerde functies cos en sin enerzijds, en de meetkundige invoering van deze functies anderzijds. De analytisch ingevoerde functies voldoen aan de bekende rekenregels over afgeleiden en hebben de eigenschap dat $\cos^2 + \sin^2 = 1$, dus het punt $(\cos t, \sin t)$ ligt op de eenheidscirkel. De meetkundige definitie gaat uit van de eenheidscirkel en zegt dat $(\cos t, \sin t)$ precies het punt P op de cirkel is, zodat de cirkelboog tussen $(1,0)$ en P lengte t heeft, (eventeel na een paar keer rond de cirkel gelopen te zijn). Het is dus voldoende om aan te tonen dat de kromme (cirkelboog) met de analytische definities van cos en sin $\gamma : t \mapsto (\cos t, \sin t) : [0,\theta] \to \mathbb{R}^2$ lengte θ heeft. Dit volgt nu echter onmiddellijk uit formule (10.7). De betreffende kromme is C^1 en

$$\Lambda(\gamma) = \int_0^\theta \sqrt{(\cos' t)^2 + (\sin' t)^2}\,dt = \int_0^\theta \sqrt{(\sin t)^2 + (\cos t)^2}\,dt = \int_0^\theta dt = \theta.$$

Voor $\theta := 2\pi$ vinden we terug dat de cirkel van straal 1 lengte 2π heeft.

Opgave 10.3. Zij $\phi : [0,1] \to \mathbb{R}$ continu, bovendien C^1 op elk interval $[\delta, 1]$ met $0 < \delta < 1$. Zij, voor $0 < \delta < 1$ de kromme γ_δ gegeven door $\gamma_\delta : t \mapsto (\cos(\phi(t)), \sin(\phi(t))) : [\delta, 1] \to \mathbb{R}^2$.

(a) Bewijs dat, voor $0 < \delta < 1$,

$$\Lambda(\gamma_\delta) = \int_\delta^1 |\phi'(t)|\,dt.$$

(b) Bewijs dat de kromme γ_0 niet rectificeerbaar is als $\lim_{\delta \downarrow 0} \Lambda(\gamma_\delta) = \infty$.

(c) Bewijs, met gebruik van a) en b), dat de kromme γ_0 niet rectificeerbaar is als $\phi(t) := t \sin \dfrac{1}{t}$.

10.8. Er is een evidente fysische interpretatie van formule (10.7). Als we $\gamma(t)$ weer opvatten als de positie van een puntdeeltje ten tijde t, dan is de integrand $|\gamma'(t)|$ de grootte van de snelheid van dat deeltje ten tijde t. De totale lengte van de weg die het deeltje aflegt tussen tijdstip a en b is (10.7), de integraal van de grootte van de snelheid als functie van de tijd.

Vermeldenswaard is het geval dat $|\gamma'(t)| = 1$ voor $a \le t \le b$. Als $a \le t_1 < t_2 \le b$ dan zien we uit (10.7) dat de kromme γ beperkt tot $[t_1, t_2]$ lengte $t_2 - t_1$ heeft. Als bovendien $a := 0$ dan is de kromme dus geparametriseerd door de afgelegde weglengte vanaf het beginpunt.

10.2 Meetkundige interpretatie van de gradiënt

10.9. We geven nu een meetkundige interpretatie van (8.25), i.e. van de formule $(D_u f)(a) = \langle u, (\nabla f)(a)\rangle$, waarbij $a \in E \subset \mathbb{R}^n$, E open en $f : E \to \mathbb{R}$, differentieerbaar in a. Nu zijn u en $(\nabla f)(a)$ beide vectoren in $\mathbb{R}^n$. Neem aan dat de gradiënt $(\nabla f)(a) \neq 0$ en neem $|u| = 1$, dus u doorloopt de eenheidssfeer S^{n-1} in $\mathbb{R}^n$. Zij $\theta \in [0, \pi]$ de hoek tussen de vectoren u en $(\nabla f)(a)$. Dan volgt uit (8.25) dat

$$(D_u f)(a) = |(\nabla f)(a)| \cos\theta \quad \text{en} \quad -|(\nabla f)(a)| \le (D_u f)(a) \le |(\nabla f)(a)| \quad (|u| = 1). \tag{10.8}$$

I.h.b. zijn er de gevallen

- $(D_u f)(a) = |(\nabla f)(a)| \iff \theta = 0 \iff u$ en $(\nabla f)(a))$ in zelfde richting;
- $(D_u f)(a) = -|(\nabla f)(a)| \iff \theta = \pi \iff u$ en $(\nabla f)(a)$ in tegengestelde richting;
- $(D_u f)(a) = 0 \iff \theta = \pi/2 \iff u$ en $(\nabla f)(a)$ onderling loodrecht.

Uit de definitie van richtingsafgeleide, formule (8.5), in het geval $|u| = 1$ zien we dat $(D_u f)(a)$ de helling (of mate van stijging) van $f(x)$ beschrijft wanneer we het argument x vanuit het punt a in de richting u volgen. Uit (10.8) volgt dat, wanneer f differentieerbaar is in a, de maximale helling bereikt wordt als u de zelfde richting heeft als de gradiëntvector $(\nabla f)(a)$. Omgekeerd kunnen we de vector $(\nabla f)(a)$ wat richting betreft karakteriseren als de richting vanaf a waarin f het hardste stijgt, en wat lengte betreft als de mate van stijging in de gevonden richting.

Opgave 10.4. Zij $a \in E \subset \mathbb{R}^2$, E open, $f : E \to \mathbb{R}$ differentieerbaar in a. Teken de grafiek van de functie $\theta \mapsto (D_{(\cos\theta, \sin\theta)} f)(a) : \mathbb{R} \to \mathbb{R}$. Hoe kunnen de richting en de grootte van de vector $(\nabla f)(a)$ uit deze grafiek worden afgeleid?

10.10. Als f een afbeelding is van een open deel van $\mathbb{R}^n$ naar $\mathbb{R}$ dan zijn er twee verschillende meetkundige beschrijvingen mogelijk van f en haar eventuele afgeleide. De eerste beschrijving werkt met de grafiek van f in de ruimte $\mathbb{R}^{n+1}$. De tweede beschrijving beperkt zich tot $\mathbb{R}^n$, maar geeft daar de niveauverzamelingen van f.

Eerst bespreken we de beschrijving met de grafiek. Zij $E \subset \mathbb{R}^n$ open en $f : E \to \mathbb{R}$. De *grafiek* van f is de deelverzameling $\mathcal{G}(f) := \{(x, f(x)) \mid x \in E\}$ van $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$. Als $n = 2$ dan kan men voor een gegeven f deze grafiek met de hand of met behulp van de computer proberen te tekenen. Differentieerbaarheid van f in een punt $a \in E$ wil zeggen dat

$$f(x) = f(a) + f'(a)\,(x - a) + o(|x - a|) \quad \text{als } x \to a \text{ in } \mathbb{R}^n.$$

M.a.w., een goede benadering voor de grafiek van f nabij het punt $(a, f(a))$ wordt gegeven door de grafiek $\mathcal{G}(l) := \{(x, f(a) + f'(a)\,(x - a)) \mid x \in \mathbb{R}^n\}$ van de functie

$$l : x \mapsto f(a) + f'(a)\,(x - a) = f(a) + \sum_{j=1}^{n} (D_j f)(a)\,(x_j - a_j).$$

Zo'n functie van $\mathbb{R}^n$ naar $\mathbb{R}$ van de vorm constante + lineaire functie noemen we *affiene functie*. De grafiek $\mathcal{G}(l)$ is een n-dimensionaal hypervlak in $\mathbb{R}^{n+1}$ dat gaat door het punt $(a, f(a))$; voor $n = 2$ dus een vlak in $\mathbb{R}^3$. Voor de grafiek $\{(x, r(x-a)) \mid x \in E\}$ van de verschilfunctie van f met l geldt dat $\lim_{x\to a} |x-a|^{-1} |r(x-a)| = 0$. Daarom is het gerechtvaardigd om $\mathcal{G}(l)$ het *raakhypervlak* te noemen aan de grafiek $\mathcal{G}(f)$ in het punt $(a, f(a))$. Voor $n = 2$ spreken we van *raakvlak*.

10.11. Zij $E \subset \mathbb{R}^n$ open en $f : E \to \mathbb{R}$. Zij $c \in \mathbb{R}$. De *niveauverzameling* van f bij niveau c is de deelverzameling $f^{-1}(c) = \{x \in E \mid f(x) = c\}$ van $\mathbb{R}^n$. De functie f wordt volledig beschreven door al haar niveauverzamelingen. Wanneer men zich beperkt tot een representatieve deelcollectie van niveaus c, dan geeft, althans voor continue f, het tekenen van de bijbehorende niveauverzamelingen (met het niveau er telkens bijgeschreven) al een redelijk beeld van de functie.

Bekijk als voorbeeld de C^1-functie $f : \mathbb{R}^2 \to \mathbb{R}$ gegeven door

$$f(x, y) := x^2 + y^2 \quad ((x, y) \in \mathbb{R}^2). \tag{10.9}$$

De niveauverzamelingen zijn

$$f^{-1}(c) = \{(x, y) \mid x^2 + y^2 = c\} \quad (c \geq 0), \tag{10.10}$$

dus concentrische cirkels rond $(0, 0)$. Elk van deze niveauverzamelingen is ook de beeldverzameling van een C^1-kromme in $\mathbb{R}^2$.

In het algemeen noemen we een kromme $\gamma : I \to E$ een *niveaukromme* van een functie $f : E \to \mathbb{R}$ (E open in $\mathbb{R}^n$) als $f(\gamma(t)) =$ constant voor $t \in I$. (Het is niet noodzakelijk dat de beeldverzameling van de niveaukromme een complete niveauverzameling is.)

Als $f : E \to \mathbb{R}$ een C^1-functie is op E dan kunnen we, voor elke $a \in E$, de afgeleide $f'(a)$ van f beschrijven door de gradiënt $(\nabla f)(a)$. Deze kunnen we tekenen door in elk punt a van E de vector $(\nabla f)(a)$ te tekenen d.m.v. een pijl die in a begint. Dit geeft een zogenaamd *vectorveld* op E: aan ieder punt van $E \subset \mathbb{R}^n$ wordt een vector uit $\mathbb{R}^n$ toegevoegd. Uiteraard tekent men vectorvelden het gemakkelijkst als $n = 2$.

10.12. Zij nu $I \subset \mathbb{R}$ een open interval, $\gamma : I \to \mathbb{R}^n$ een C^1-kromme, $E \subset \mathbb{R}^n$ open zo dat $\gamma(I) \subset E$, en $f : E \to \mathbb{R}$ een C^1-functie. Dan volgt uit het speciale geval van de kettingregel (Stelling 9.5 samen met formules (9.13) – (9.15)) dat

$$(f \circ \gamma)'(t) = \sum_{k=1}^{n} \gamma_k'(t)\,(D_k f)(\gamma(t)) \quad (t \in I).$$

Omdat $\nabla f = (D_1 f, \ldots, D_n f)$, (formule (8.24)) kunnen we dit herschrijven als het inwendig product

$$(f \circ \gamma)'(t) = \langle \gamma'(t), (\nabla f)(\gamma(t)) \rangle \quad (t \in I). \tag{10.11}$$

We concluderen dat de volgende drie beweringen equivalent zijn.

(a) $f(\gamma(t))$ is constant voor $t \in I$ (i.e., de functie f beperkt tot $\gamma(I)$ is constant; i.e., γ is een niveaukromme van f).

(b) $(f \circ \gamma)'(t) = 0$ voor alle $t \in I$.

(c) $(\nabla f)(\gamma(t)) \perp \gamma'(t)$ voor alle $t \in I$ (i.e., voor elke $t \in I$ staat de gradiënt van f in $\gamma(t)$ loodrecht op de raakvector van γ in t).

Onderdeel (c) geeft dus een nieuwe karakterisering van de niveaukrommen van f. Voor $n = 2$ in het bijzonder geeft dit vaak een heel aanschouwelijke indruk: Teken het gradiëntvectorveld ∇f van f. Krommen $\Gamma : I \to \mathbb{R}^n$ met de eigenschap dat in ieder punt ∇f gelijk is aan (een positief veelvoud van) Γ' heten *integraalkrommen van het vectorveld* ∇f. Teken nu de gekromde lijnen in E die overal loodrecht op dit vectorveld (dwz. op de integraalkrommen) staan (z.g. *orthogonale trajectoria* voor het vectorveld ∇f). Deze gekromde lijnen zijn bevat in niveauverzamelingen van f en worden, na parametrisering, niveaukrommen. Omgekeerd, nog steeds voor $n = 2$, als we niveaukrommen van f getekend hebben dan geldt voor elk punt a van E waardoor een C^1-niveaukromme gaat met raakvector ter plaatse ongelijk 0, dat de gradiënt van f in a loodrecht staat op deze raakvector en daardoor op een constante factor na bepaald is.

Laat $n = 2$. Laat $c_1, c_2 \in \mathbb{R}$, $c_2 > 0$ en laat c alle waarden $c_1 + kc_2$ ($k \in \mathbb{Z}$) doorlopen. Stel dat, voor al deze c-waarden de niveauverzamelingen $f^{-1}(c)$ getekend zijn. Voor dicht bij elkaar gelegen c-waarden kunnen we aannemen dat de bijbehorende niveauverzamelingen enigszins parallel lopen en dat het zinvol is om over hun onderlinge afstand te spreken. Nu kunnen we stellen dat de lengte van de vector $(\nabla f)(a)$ omgekeerd evenredig is met de afstand tussen twee opeenvolgende niveauverzamelingen nabij a (waarom?). We zouden ook kunnen zeggen dat de lengte van $(\nabla f)(a)$ evenredig is met de "dichtheid van de niveauverzamelingen" bij a.

10.13 Voorbeeld. Bekijk de functie f gegeven door (10.9). De niveauverzamelingen (10.10) ($c = r^2$) zijn precies de beeldverzamelingen van de krommen $\gamma_r : t \mapsto (r \cos t, r \sin t) : \mathbb{R} \to \mathbb{R}^2$ ($r \geq 0$). Nu geldt:

$$\gamma_r'(\theta) = (-r \sin\theta, r\cos\theta) \perp (2r\cos\theta, 2r\sin\theta) = (\nabla f)(r\cos\theta, r\sin\theta).$$

Merk op dat de functie f zijn kleinste waarde 0 aanneemt in $(0, 0)$ en verder nergens. Dit triviale feit is ook makkelijk af te lezen uit een tekening van gradiëntvectorveld en niveauverzamelingen.

We zien dat het gradiënt vectorveld bestaat uit

Opgave 10.5. Teken voor onderstaande functies $f : \mathbb{R}^2 \to \mathbb{R}$ de niveauverzamelingen. Bereken ∇f en teken het gradiëntvectorveld. Beschrijf de niveauverzamelingen door niveaukrommen, bereken hun raakvectoren en verifieer dat raakvector en gradiënt in elk punt loodrecht op elkaar staan.

(a) $f(x, y) := x^2 - y^2$.

(b) $f(x, y) := x^2 + y$.

Het punt $(0, 0)$ is uitzonderlijk voor de functie f uit (a). Probeer zoveel mogelijk bijzonderheden van het gedrag van f bij $(0, 0)$ te beschrijven.

10.14. We geven nog een fysisch geografische toepassing van het begrip gradiënt an een C^1-functie $f : E \to \mathbb{R}$ (E open in $\mathbb{R}^2$). Zij E een bergachtig terrein op aarde geprojecteerd op het horizontale vlak en f de functie die aan ieder punt van E de hoogte in het terrein toevoegt. De niveauverzamelingen van f zijn dan hoogtelijnen. Zij $\gamma : I \to E$ een C^1-kromme zo dat de deelverzameling $\gamma(I)$ van E een weg in het terrein voorstelt. De parametrisering bepaalt hoe de weg doorlopen moet worden. Neem aan dat $|\gamma'(t)| = 1$ voor alle t in I. Dus de weglengte vanaf parameterwaarde t_1 tot parameterwaarde $t_2 > t_1$ is precies $t_2 - t_1$. (Voor het gemak houden we geen rekening met de kromming van de aarde en nemen we aan dat de weglengte berekend wordt na projectie op het horizontale vlak.) Stel dat het interval I linker eindpunt 0 heeft, dan is de weglengte vanaf het beginpunt $\gamma(0)$ tot aan het punt $\gamma(t)$ dus t.

Met bovenstaande gegevens kunnen we $(f \circ \gamma)'(t)$ interpreteren als de helling van de weg ter plaatse t (dus bereken verticaal hoogteverschil $f(\gamma(t + \Delta t)) - f(\gamma(t))$ gedeeld door horizontaal afgelegde weg Δt, en laat dan $\Delta t \to 0$). Gezien (10.11) en (8.25) kunnen we deze helling dus schrijven als

$$(f \circ \gamma)'(t) = \langle \gamma'(t), (\nabla f)(\gamma(t)) \rangle = (D_{\gamma'(t)} f)(\gamma(t)) \quad (t \in I), \tag{10.12}$$

waarbij in het rechterlid een richtingsafgeleide staat. We kunnen de meetkundige beschouwingen van §10.9 nu toepassen, waarbij $u := \gamma'(t)$ en θ de hoek is tussen $\gamma'(t)$ en $(\nabla f)(\gamma(t))$. Dus als $a \in E$ en $\gamma(t) = a$ dan is op die plaats de helling van de weg, als functie van de wegrichting $\gamma'(t)$, maximaal als $\gamma'(t)$ in de zelfde richting wijst als $(\nabla f)(a)$, terwijl de helling 0 is als de wegrichting loodrecht op de gradiënt staat.

Opgave 10.6. Zij f de hoogtefunctie als boven. Teken de hoogtelijnen en het gradiëntvectorveld in de omgeving van een bergpas.

Verdere vraagstukken

Opgave 10.7. Bereken de lengte van het tussen $(a, 0, 0)$ en $(a, 0, 4\pi b)$ gelegen deel van de kromme (*schroeflijn*), die gegeven wordt door middel van de parametrisering $\gamma(t) = (a \cos t, a \sin t, bt)$.

Opgave 10.8. Zij γ de kromme in $\mathbb{R}^2$ die in poolcoördinaten gegeven is door $r = a\, e^{b\phi}$ (*logaritmische spiraal*) met $a > 0$, $b > 0$. Zij $(x_1, 0)$ en $(x_2, 0)$ twee opeenvolgende snijpunten van γ met de positieve x-as. Laat zien dat de lengte van het deel van γ dat tussen deze punten ligt gelijk is aan

$$(x_2 - x_1)\sqrt{1 + b^{-2}}.$$

Opgave 10.9. Zij $\gamma : \mathbb{R} \to \mathbb{R}^n$ een C^1-kromme met $|\gamma(t)| = 1$ voor alle $t \in \mathbb{R}$. Toon aan dat $\langle \gamma(t), \gamma'(t) \rangle = 0$ voor alle $t \in \mathbb{R}$. Geef hiervan een meetkundige interpretatie in het geval $n = 3$.

Opgave 10.10. Zij $a \in \mathbb{R}^n$ vast. Bepaal de afgeleiden van de volgende afbeeldingen:

(a) $f : \mathbb{R}^n \to \mathbb{R}$ met $f(x) := \langle a, x \rangle$

(b) $f : \mathbb{R}^n \to \mathbb{R}^n$ met $f(x) := \langle a, x \rangle\, x$

(c) $f : \mathbb{R}^n \to \mathbb{R}^n$ met $f(x) := \langle x, x \rangle\, x$.

11 Middelwaardestelling; hogere partiële afgeleiden

11.1 Een generalisatie van de middelwaardestelling

11.1. Zij $f : I \to \mathbb{R}$ een differentieerbare functie op een open interval I en zij $a, b \in I$ met $a < b$. Dan zegt de middelwaardestelling (cf. K.A. Ross, Elementary Analysis, Stelling 29.3) dat

$$f(b) - f(a) = f'(x)\,(b - a) \quad \text{voor zekere } x \in (a, b). \tag{11.1}$$

Een direct gevolg van (11.1) is de ongelijkheid

$$|f(b) - f(a)| \le |f'(x)|\,(b - a) \quad \text{voor zekere } x \in (a, b). \tag{11.2}$$

Dus als voor zekere $M \ge 0$ geldt dat $|f'(x)| \le M$ op I dan volgt er dat

$$|f(b) - f(a)| \le M\,|b - a| \quad (a, b \in I). \tag{11.3}$$

Hoe staat het met analoga van deze resultaten voor differentieerbare afbeeldingen $f : E \to \mathbb{R}^m$, waarbij E open in $\mathbb{R}^n$ is? Voor $m = 1$ zullen we zien dat alles (met passende interpretatie) geldig blijft. Voor algemene m blijven ongelijkheden (11.2), (11.3) met passende interpretatie geldig, maar gelijkheid (11.1) niet. Eigenschap (11.1) hoeft zelfs al niet meer te gelden als $n = 1$, $m > 1$, i.e., in het geval van vectorwaardige functies van één reële variabele. Ook voor complexwaardige functies van één reële variabele (dit is het geval $n = 1$, $m = 2$) kan (11.1) misgaan. Zie hiertoe het volgende tegenvoorbeeld.

11.2 Voorbeeld. Definieer de differentieerbare functie $f : \mathbb{R} \to \mathbb{C}$ door

$$f(x) := e^{ix} = \cos x + i \sin x.$$

Dan is enerzijds $f(2\pi) - f(0) = 1 - 1 = 0$, maar anderzijds geldt dat $f'(x) = i\,e^{ix}$. Dus $|f'(x)| = 1 \ne 0$ voor elke $x \in \mathbb{R}$. Dus (11.1) kan niet gelden voor $(a, b) := (0, 2\pi)$.

11.3 Propositie. Zij $E \subset \mathbb{R}^n$ open. Zij $f : E \to \mathbb{R}$ differentieerbaar op E. Zij $a, b \in E$ zo dat het verbindende lijnstuk $[a, b]$ geheel in E ligt. Schrijf $(a, b) := \{a + t(b - a) \mid 0 < t < 1\}$ voor het bijbehorende open lijnstuk. Dan

$$f(b) - f(a) = f'(x)\,(b - a) = \sum_{j=1}^{n} (D_j f)(x)\,(b_j - a_j) \quad \text{voor zekere } x \in (a, b). \tag{11.4}$$

Bewijs. Schrijf

$$g(t) := f(a + t(b - a)) \quad (0 \le t \le 1).$$

Dan is g continu op $[0, 1]$ en er volgt met behulp van de kettingregel dat g differentieerbaar is op $(0, 1)$ met

$$g'(t) = f'(a + t(b - a))\,(b - a)$$

(ga na). Toepassing van (11.1) geeft dat $g(1) - g(0) = g'(t)$ voor zekere $t \in (0, 1)$. □

11.4 Propositie. Zij $E \subset \mathbb{R}^n$ open. Zij $f : E \to \mathbb{R}^m$ differentieerbaar op E. Zij $a, b \in E$ zo dat het verbindende lijnstuk $[a, b]$ geheel in E ligt. Dan

$$|f(b) - f(a)| \le \|f'(x)\|\,|b - a| \quad \text{voor zekere } x \in (a, b). \tag{11.5}$$

11.5. Alvorens het bewijs te geven maken we een paar opmerkingen. In (11.5) komt de operatornorm van $f'(x)$ voor. Als $n = 1$ dan kunnen we $f'(x)$ opvatten als behorend tot $\mathbb{R}^m$ en kunnen we in (11.5) volstaan met de vectornorm $|f'(x)|$. Het speciale geval $n = 1$ van onderstaand bewijs is wellicht gemakkelijker. Onder de aannamen van Propositie 11.4 zullen we eerst bewijzen dat

$$|f(b) - f(a)|^2 = \langle f'(x)\,(b-a), f(b) - f(a)\rangle \quad \text{voor zekere } x \in (a,b). \tag{11.6}$$

Uitspraak 11.6 kan equivalent worden geformuleerd als

$$\langle f(b) - f(a) - f'(x)\,(b-a), f(b) - f(a)\rangle = 0 \quad \text{voor zekere } x \in (a,b).$$

en dit is weer equivalent met te zeggen dat

$$f(b) - f(a) = f'(x)\,(b-a) + v \quad \text{voor zekere } x \in (a,b) \text{ en voor zekere } v \perp f(b) - f(a).$$

Dit laatste kan opgevat worden als generalisatie van 11.4 en (11.1).

Bewijs van propositie 11.4. Eerst bewijzen we (11.6). Schrijf

$$\phi(t) := \langle f(a + t(b-a)), f(b) - f(a)\rangle = \sum_{i=1}^{m} (f_i(a + t(b-a))\,(f_i(b) - f_i(a)) \quad (0 \le t \le 1). \tag{11.7}$$

Dan is ϕ continu op $[0,1]$ en er volgt met behulp van de kettingregel dat ϕ differentieerbaar is op $(0,1)$ met

$$\phi'(t) = \sum_{i=1}^{m}\sum_{j=1}^{n} (D_j f_i)(a + t(b-a))\,(b_j - a_j)\,(f_i(b) - f_i(a)) \tag{11.8}$$

$$= \sum_{i=1}^{m} \big((f'(a + t(b-a))\,(b-a)\big)_i \big(f(b) - f(a)\big)_i \tag{11.9}$$

$$= \langle f'(a + t(b-a))\,(b-a), f(b) - f(a)\rangle. \tag{11.10}$$

(ga na). Toepassing van (11.1) geeft dat $\phi(1) - \phi(0) = \phi'(t)$ voor zekere $t \in (0,1)$. Maar anderzijds volgt uit (11.7) dat $\phi(1) - \phi(0) = |f(b) - f(a)|^2$. Combinatie met (11.10) levert uitspraak (11.6).

We bewijzen nu uit (11.6) de uitspraak (11.5). Pas de ongelijkheid van Schwarz toe op het rechterlid van (11.6). Dan volgt er dat

$$|f(b) - f(a)|^2 \le |f'(x)\,(b-a)|\,|f(b) - f(a)| \le ||f'(x)||\,|b-a|\,|f(b) - f(a)|.$$

Door beide zijden van deze ongelijkheid te delen door $|f(b) - f(a)|$ (mits $\neq 0$) volgt (11.5). Als $|f(b) - f(a)| = 0$ dan is (11.5) evident. $\square$

Als onmiddellijk gevolg van Propositie 11.4 verkrijgen we nu de generalisatie van (11.3).

11.6 Stelling. Zij $E \subset \mathbb{R}^n$ open en convex. Zij $f : E \to \mathbb{R}^m$ differentieerbaar op E. Veronderstel dat voor zekere $M \ge 0$ geldt dat $||f'(x)|| \le M$ voor alle $x \in E$. Dan

$$|f(b) - f(a)| \le M\,|b-a| \quad (a, b \in E). \tag{11.11}$$

11.7. In de één-dimensionale reële analyse noemen we een functie f op een open interval I *Lipschitz-continu* als er een M bestaat zodat $|f(x) - f(y)| \le M|x - y|$, $x, y \in I$. Dit kan men naar metrische ruimten generaliseren. Laten (X, d_X) en (Y, d_Y) metrische ruimtes zijn en zij $f : X \to Y$. We noemen de afbeelding f *Lipschitz-continu* als er $M \ge 0$ bestaat zo dat

$$d_Y(f(a), f(b)) \le M\, d_X(a, b) \quad (a, b \in X). \tag{11.12}$$

I.h.b. is dus, voor $E \subset \mathbb{R}^n$, een afbeelding $f : E \to \mathbb{R}^m$ Lipschitz-continu als voor zekere $M \ge 0$ de eigenschap (11.11) geldt.

Opgave 11.1.

(a) Bewijs dat Lipschitz-continuïteit uniforme continuïteit impliceert.

(b) Zij $E \subset \mathbb{R}^n$ open. Zij $f : E \to \mathbb{R}^m$ een C^1-afbeelding. Zij B een gesloten begrensde bal bevat in E. Bewijs dat de afbeelding $f : B \to \mathbb{R}^m$ Lipschitz-continu is.

Opgave 11.2. Laten r, θ poolcoördinaten zijn op $\mathbb{R}^2$. Vind een open, samenhangende, maar niet-convexe deelverzameling E van $\mathbb{R}^2\backslash\{(0, 0)\}$, waarop θ eenduidig als een C^1-functie gedefinieerd kan worden, waarop de totale afgeleide van θ in norm begrensd is door een constante M, maar waarop de Lipschitz-eigenschap (11.11) toch niet geldt. (Maak voor de mogelijke begrensdheid van de totale afgeleide van θ gebruik van de expliciete uitdrukking voor $\partial\theta/\partial x$ en $\partial\theta/\partial y$, zoals uitgerekend in de opgave bij §9.8.)

Opgave 11.3. Definieer een afbeelding $f : \mathbb{R}^2 \to \mathbb{R}^2$ door $f(x, y) := (x\sqrt{x^2 + y^2}, y\sqrt{x^2 + y^2})$.

(a) Bewijs dat f een C^1-functie is en dat

$$f'(r\cos\theta, r\sin\theta) = \begin{pmatrix} 2r\cos^2\theta + r\sin^2\theta & r\sin\theta\cos\theta \\ r\sin\theta\cos\theta & r\cos^2\theta + 2r\sin^2\theta \end{pmatrix}.$$

(b) Bewijs dat bovenstaande matrix eigenwaarden r en $2r$ heeft.

(c) Bewijs dat f voor iedere begrensde open convexe deelverzameling E van $\mathbb{R}^2$ aan (11.11) voldoet, waarbij $M = 2\sup_{(x,y)\in E}\sqrt{x^2 + y^2}$. Onderzoek in hoeverre deze keuze van M scherp is, bijv. als E de cirkelschijf van straal 1 rond $(0, 0)$ is.

Aanwijzing Gebruik (9.33) en (9.34 bij (a). Gebruik Opgave 7.14(b) bij (c).

11.2 Hogere partiële afgeleiden

11.8 Definitie. Zij $a \in E \subset \mathbb{R}^n$, E open, en zij $f : E \to \mathbb{C}$. Zij $i, j \in \{1, \ldots, n\}$. We zeggen dat de *partiële afgeleide van de tweede orde* $(D_{ij}f)(a)$ bestaat en gedefinieerd wordt door

$$(D_{ij}f)(a) := (D_i(D_jf))(a), \tag{11.13}$$

als voor zekere open omgeving U van a in E geldt dat de partiële afgeleide $(D_jf)(x)$ bestaat voor alle $x \in U$ en als de functie $D_jf : U \to \mathbb{C}$ in het punt a een partiële afgeleide $(D_i(D_jf))(a)$ heeft.

Algemener, als $j_1, \ldots, j_k \in \{1, \ldots, n\}$, dan zeggen we met recurrentie naar k dat de *partiële afgeleide van de k-de orde* $(D_{j_1\ldots j_k}f)(a)$ bestaat en gedefinieerd wordt door

$$(D_{j_1\ldots j_k}f)(a) := (D_{j_1}(D_{j_2\ldots j_k}f))(a), \tag{11.14}$$

als voor zekere open omgeving U van a in E geldt dat de partiële afgeleide $(D_{j_2\ldots j_k}f)(x)$ bestaat voor alle $x \in U$ en als de functie $D_{j_2\ldots j_k}f : U \to \mathbb{C}$ in het punt a een partiële afgeleide $(D_{j_1}(D_{j_2\ldots j_k}f))(a)$ heeft.

Wanneer $E \subset \mathbb{R}^n$ open en $k \in \{1, 2, \ldots\}$ dan noemen we een afbeelding $f : E \to \mathbb{R}^m$ een *C^k-afbeelding* (of *k keer continu differentieerbaar op E*) als voor alle $i = 1, \ldots, m$, alle $l = 1, \ldots, k$ en alle $j_1, \ldots, j_l \in \{1, \ldots, n\}$ geldt dat de partiële afgeleide $D_{j_1\ldots j_l}f_i$ overal op E bestaat en continu is op E. We noemen $f : E \to \mathbb{R}^m$ een *C^∞-afbeelding* als f een C^k-afbeelding is voor alle $k = 1, 2, \ldots$. In het bijzonder vermelden we hier de gevallen $m = 1$ (reëelwaardige C^k-functie) en $m = 2$ (te identificeren met complexwaardige C^k-functie).

Gezien Stelling 8.30 is bovenstaande definitie van C^k-afbeelding voor het geval $k = 1$ compatibel met de eerder in §8.28 gegeven definitie van C^1-afbeelding.

11.9 Notatie. Naar analogie van de notatie voor hogere afgeleiden van functies van één variabele schrijven we bijvoorbeeld

$$\frac{\partial^2 f(x,y)}{\partial x^2} := (D_{11}f)(x,y), \quad \frac{\partial^2 f(x,y)}{\partial x\,\partial y} := (D_{12}f)(x,y), \quad \frac{\partial^2 f(x,y)}{\partial y^2} := (D_{22}f)(x,y),$$

voor de partiële afgeleiden van de tweede orde van een C^2-functie f gedefinieerd op een open deelverzameling E van $\mathbb{R}^2$. Hierbij beschouwen we (x, y) als een willekeurig element van E. Je zou kunnen zeggen dat een partiële differentiaal operator als $\frac{\partial^2}{\partial x\,\partial y}$ voorafgaand aan een bepaalde uitdrukking die de variabelen x en y bevat, volgens een bepaald voorschrift een nieuwe uitdrukking geeft die de variabelen x en y bevat. Als we de partiële afgeleide willen evalueren in een concreet punt, zeg het punt $(5, 8)$, dan schrijven we voor $(D_{12}f)(5, 8)$

$$\text{niet} \quad \frac{\partial^2 f(5,8)}{\partial x\,\partial y} \qquad \text{maar} \quad \left.\frac{\partial^2 f(x,y)}{\partial x\,\partial y}\right|_{(x,y)=(5,8)}.$$

Algemener, als $f : E \to \mathbb{R}$ met $E \subset \mathbb{R}^n$ open een C^k-functie is dan schrijven we

$$\frac{\partial^k f(x_1, \ldots, x_n)}{\partial x_{j_1} \ldots \partial x_{j_k}} := (D_{j_1\ldots j_k}f)(x).$$

Opgave 11.4. Zij $E \subset \mathbb{R}^n$ open en $f : E \to \mathbb{R}^m$ zo dat alle partiële afgeleiden van orde $\leq k$ van $f_1, \ldots, f_m$ op E bestaan en de partiële afgeleiden van orde k van $f_1, \ldots, f_m$ bovendien continu zijn op E. Bewijs dat f een C^k-afbeelding is, i.e., dat de partiële afgeleiden van de f_i van orde $< k$ ook continu zijn op E.

Eerst gaan we verder in op de partiële afgeleiden van hogere orde. Het is in het algemeen niet waar dat $(D_{ij}f)(a)$ en $(D_{ji}f)(a)$, wanneer deze twee partiële afgeleiden van tweede orde beide bestaan, ook aan elkaar gelijk zijn. Zie een tegenvoorbeeld in de Opgave hieronder.

Opgave 11.5. Zij $f : \mathbb{R}^2 \to \mathbb{R}$ gegeven door

$$f(x,y) := \frac{xy^3}{x^2 + y^2} \quad \text{als } (x,y) \neq (0,0), \qquad f(0,0) := 0.$$

Bewijs dat f een C^1-functie is, dat $D_{12}f$ en $D_{21}f$ overal op $\mathbb{R}^2$ bestaan, maar dat $(D_{21}f)(0, 0) = 1 \neq 0 = (D_{12}f)(0, 0)$ en dat de functies $D_{21}f$ en $D_{12}f$ niet continu zijn in $(0, 0)$.

11.10. Als echter de afgeleiden van tweede orde bestaan en bovendien continu zijn, dan zal volgen dat $D_{ij}f = D_{ji}f$. Het zal blijken voldoende te zijn om dit te bewijzen voor C^2-functies op een open deelverzameling E van $\mathbb{R}^2$. Het bewijs maakt gebruik van een idee dat ook belangrijk is in de numerieke wiskunde: Voor een C^2-functie f in twee variabelen kunnen zowel $(D_{12}f)(a,b)$ als $(D_{21}f)(a,b)$ benaderd worden door het differentiequotiënt van tweede orde gegeven door

$$\frac{f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b)}{hk},$$

waarbij h,k reëel zijn en klein in absolute waarde.

11.11 Propositie. Zij $(a,b) \in E \subset \mathbb{R}^2$ met E open en zij $f : E \to \mathbb{R}$ een C^2-functie. Dan geldt dat $(D_{12}f)(a,b) = (D_{21}f)(a,b)$.

Bewijs. Zij $r > 0$ zo dat $(x,y) \in E$ als $|(x-a, y-b)| < r$. Neem $h,k > 0$ zo dat $|(h,k)| < r$. Dan ligt de gesloten rechthoek met hoekpunten (a,b), $(a+h,b)$, $(a,b+k)$ en $(a+h,b+k)$ geheel in E. Toepassing van de middelwaardestelling op de differentieerbare functie $x \mapsto f(x,b+k) - f(x,b)$ geeft dat

$$f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b) = h\big((D_1f)(a+\theta_1h,b+k) - (D_1f)(a+\theta_1h,b)\big) \tag{11.15}$$

voor zekere $\theta_1 \in (0,1)$. Herschrijf nu het rechterlid van (11.15) door toepassing van de middelwaardestelling op de differentieerbare functie $y \mapsto (D_1f)(a+\theta_1h,y)$. Dit levert:

$$f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b) = hk\,(D_{21}f)(a+\theta_1h,b+\theta_2k) \tag{11.16}$$

voor zekere $\theta_2 \in (0,1)$. Met een soortgelijke redenering kunnen we eerst de y-variabele en dan de x-variabele behandelen. We verkrijgen dan dat

$$f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b) = hk\,(D_{12}f)(a+\eta_1h,b+\eta_2k) \tag{11.17}$$

voor zekere $\eta_1,\eta_2 \in (0,1)$. Door combinatie van (11.15) en (11.16) zien we dat

$$(D_{21}f)(a+\theta_1h,b+\theta_2k) = (D_{12}f)(a+\eta_1h,b+\eta_2k). \tag{11.18}$$

Laat nu $\varepsilon > 0$. Vanwege de continuïteit van $D_{12}f$ en $D_{21}f$ is er een $\delta > 0$ zo dat $\delta \le r$ en

$$\begin{aligned} |(D_{12}f)(x,y) - (D_{12}f)(a,b)| < \varepsilon/2 \quad &\text{als} \quad |(x,y)-(a,b)| < \delta,\\ |(D_{21}f)(x,y) - (D_{21}f)(a,b)| < \varepsilon/2 \quad &\text{als} \quad |(x,y)-(a,b)| < \delta. \end{aligned}$$

Neem $h,k > 0$ zo dat $|(h,k)| < \delta$. Dan zal $|(a+\theta_1h,b+\theta_2k)-(a,b)| < \delta$ en $|a+\eta_1h,b+\eta_2k) - (a,b)| < \delta$ als $\theta_1,\theta_2,\eta_1,\eta_2 \in (0,1)$. Dus er volgt uit (11.18) en de driehoeksongelijkheid dat $|(D_{21}f)(a,b) - (D_{12}f)(a,b)| < \varepsilon$. Omdat $\varepsilon > 0$ willekeurig genomen was, volgt de Propositie. $\square$

11.12 Opmerking. De conclusie van bovenstaande Propositie volgt reeds uit zwakkere aannames. In Rudin, Theorem 9.41 wordt de volgende stelling bewezen.
Zij $(a,b) \in E \subset \mathbb{R}^2$ met E open, zij $f : E \to \mathbb{R}$ zo dat D_1f, D_2f en $D_{21}f$ overal op E bestaan, en zij $D_{21}f$ continu in (a,b). Dan bestaat $D_{12}f$ in (a,b) en $(D_{12}f)(a,b) = (D_{21}f)(a,b)$.

11.13 Stelling. Zij $E \subset \mathbb{R}^n$ open en zij $k = 2,3,\ldots$. Zij $f : E \to \mathbb{R}$ een C^k-afbeelding. Dan geldt voor alle $i_1,\ldots,i_k \in \{1,\ldots,n\}$ en voor alle permutaties σ van $\{1,\ldots,k\}$ dat $D_{i_{\sigma(1)}\cdots i_{\sigma(k)}}f = D_{i_1\ldots i_k}f$. In het bijzonder geldt voor een C^2-afbeelding dat

$$(D_{ij}f)(x) = (D_{ji}f)(x) \quad (x \in E). \tag{11.19}$$

Bewijs. Het geval van algemene k volgt uit het geval $k = 2$ (formule (11.19)) (waarom?). Het geval $k = 2$, op zijn beurt, volgt uit het geval $k = 2$, $n = 2$ (waarom?). Dit laatste geval is evident uit Propositie 11.11. $\square$

Ter voorbereiding van de volgende Opgave, maar ook vanwege het algemene belang bespreken we nu differentiatie van een integraal naar een parameter van de integrand.

11.14 Propositie. Stel I_1 en I_2 zijn open intervallen in $\mathbb{R}$. Laat $f : I_1 \times I_2 \to \mathbb{R}$ een C^1-functie zijn. Neem een gesloten begrensd interval $[a, b] \subset I_2$ en definieer de functie $F : I_1 \to \mathbb{R}$ door

$$F(x) := \int_a^b f(x, y)\, dy \quad (x \in I_1). \tag{11.20}$$

Dan is F differentieerbaar op I_1 met afgeleide

$$F'(x) = \int_a^b (D_1 f)(x, y)\, dy \quad (x \in I_1). \tag{11.21}$$

Bewijs. Omdat f een C^1-functie is, zijn, voor vaste x, de functies $y \mapsto f(x, y)$ en $y \mapsto (D_1 f)(x, y)$ continu op $[a, b]$, dus de integralen in (11.20) en (11.21) zijn goed gedefinieerd als Riemann-integralen. Merk nu op dat, voor x en $x + h$ in I_1,

$$\begin{aligned}\frac{F(x+h) - F(x)}{h} - \int_a^b (D_1 f)(x, y)\, dy &= \int_a^b \Big(\frac{f(x+h, y) - f(x, y)}{h} - (D_1 f)(x, y) \Big)\, dy \\ &= \int_a^b \big((D_1 f)(x + \theta_y h, y) - (D_1 f)(x, y) \big)\, dy \qquad (11.22)\end{aligned}$$

voor zekere $\theta_y \in (0, 1)$. Hierbij gebruikten we de middelwaardestelling (11.1) in het geval van één variabele. Neem nu x vast in I_1, neem $[c, d] \subset I_1$ zo dat $x \in (c, d)$. Dan is de verzameling $[c, d] \times [a, b]$ compact, dus $D_1 f$, die continu is op deze verzameling, is daar ook uniform continu. Neem $\varepsilon > 0$. Dan is er vanwege de uniforme continuïteit een $\delta > 0$ zo dat, als $|h| < \delta$ dan $x + h \in [c, d]$ en $|(D_1 f)(x + \theta h, y) - (D_1 f)(x, y)| < \varepsilon$ voor alle $y \in [a, b]$ en alle $\theta \in (0, 1)$. Dus, als $|h| < \delta$ dan kan de absolute waarde van de integrand in (11.22) naar boven afgeschat worden door ε, dus

$$\Big| \frac{F(x+h) - F(x)}{h} - \int_a^b (D_1 f)(x, y)\, dy \Big| \le \int_a^b \big| (D_1 f)(x + \theta_y h, y) - (D_1 f)(x, y) \big|\, dy \le \varepsilon (b - a).$$

Dus

$$\lim_{h \to 0} \Big(\frac{F(x+h) - F(x)}{h} - \int_a^b (D_1 f)(x, y)\, dy \Big) = 0.$$

$\square$

De conclusie van bovenstaande stelling gaat door onder veel zwakkere aannamen. De geïnteresseerde lezer verwijzen we naar Rudin, §9.42, 9.43.

Opgave 11.6. Laten u en v reëelwaardige C^1-functies zijn op $\mathbb{R}^2$. Bewijs dat de volgende twee uitspraken equivalent zijn.

(a) Er bestaat een reëelwaardige C^2-functie f op $\mathbb{R}^2$ zo dat $D_1 f = u$ en $D_2 f = v$.

(b) $D_2 u = D_1 v$.

Aanwijzing Merk voor het bewijs (b)⇒(a) op dat, als er een f bestaat met de eigenschappen in (a), dan, voor vaste $(a,b) \in \mathbb{R}^2$,

$$f(x,y) - f(a,b) = (f(x,y) - f(x,b)) + (f(x,b) - f(a,b)) = \int_b^y v(x,t)\,dt + \int_a^x u(t,b)\,dt.$$

Definieer nu $f(x,y)$ door het meest rechtse lid en laat zien dat $(D_1 f)(x,y)$ en $(D_2 f)(x,y)$ gelijk zijn aan $u(x,y)$ en $v(x,y)$, resp.

11.15. We zien dat f door eigenschap (a) uniek bepaald is op een constante term na. De implicatie (a)⇒(b) van Opgave 11.6 blijft doorgaan wanneer de functies gedefinieerd zijn op een willekeurige open deelverzameling van $\mathbb{R}^2$. Voor de implicatie in omgekeerde richting kunnen we volstaan met de eis dat de functies op een z.g. *enkelvoudig samenhangende* deelverzameling van $\mathbb{R}^2$ gedefinieerd zijn. Informeel gezegd bedoelen we hiermee een open samenhangende deelverzameling zonder gaten. Bijv. een schijf is enkelvoudig samenhangend maar een ring niet.

Opgave 11.7. Laten u en v gedefinieerd zijn als functies op $E := \mathbb{R}^2 \backslash \{(0,0)\}$ door

$$u(x,y) := \frac{-y}{x^2+y^2}\,, \quad v(x,y) := \frac{x}{x^2+y^2}\,.$$

Bewijs dat u en v C^1-functies zijn die voldoen aan (b) van Opgave 11.6, maar dat (a) van Opgave 11.6 niet geldt.

11.3 Coördinatentransformaties in partiële differentiaaloperatoren

We bespreken nu de transformatieformules voor de partiële differentiaaloperatoren van de tweede orde die corresponderen met een transformatie van coördinaten $x_1, \ldots, x_n$ naar coördinaten $u_1, \ldots, u_n$. Deze beschouwingen zijn analoog aan de formules in §9.6 en grijpen daar ook weer op terug. Na behandeling van het algemene geval zal het belangrijke speciale geval van overgang van cartesische op poolcoördianten of omgekeerd besproken worden. Ook hier geldt weer dat je later in staat moet zijn om deze transformaties in speciale situaties te kunnen uitrekenen.

11.16. Laten dus, net als in §9.6, E_1 en E_2 open deelverzamelingen zijn van $\mathbb{R}^n$ en laten $\phi : E_1 \to E_2$ en $\psi : E_2 \to E_1$ nu C^2-afbeeldingen zijn die voldoen aan (9.17), i.e., C^2-diffeomorfismen die elkaars inversen zijn. Laten $F : E_1 \to \mathbb{R}$ en $f : E_2 \to \mathbb{R}$ nu C^2-functies zijn die met elkaar verband houden volgens (9.18). Dus

$$\begin{array}{ccc} \psi(\phi(x)) = x \quad (x \in E_1) & \text{en} & \phi(\psi(u)) = u \quad (u \in E_2), \\ F(x) = f(\phi(x)) \quad (x \in E_1) & \text{en} & f(u) = F(\psi(u)) \quad (u \in E_2). \end{array}$$

We gaan nu uit van de ongenummerde formule even onder (9.18), die we hier nog eens geven.

$$(D_l f)(u) = \sum_{j=1}^{n} (D_l \psi_j)(u)\,(D_j F)(\psi(u)) \quad (u \in E_2). \tag{11.23}$$

Wanneer we het argument u weg zouden laten in deze formule dan hebben we links de C^1-functie $D_l f$ staan en rechts in de j-de term het product van de C^1-functies $D_l \psi_j$ en $(D_j F) \circ \psi$. Alle functies zijn gedefinieerd op E_2. We laten nu de operator D_k werken op de net beschreven functies aan beide zijden van (11.23) en we passen de formule

$$(D_k(gh))(u) = g(u)\,(D_k h)(u) + (D_k g)(u)\,h(u) \tag{11.24}$$

voor de partiële afgeleide van het product van twee partieel differentieerbare functies toe (cf. §8.7). We verkrijgen dan:

$$(D_{kl}f)(u) = \sum_{j=1}^{n}(D_l\psi_j)(u)\,(D_k((D_jF)\circ\psi))(u) + \sum_{j=1}^{n}(D_{kl}\psi_j)(u)\,(D_jF)(\psi(u)) \quad (u \in E_2). \quad (11.25)$$

We willen vervolgens de laatste factor in de eerste sommatie van (11.25), dus de uitdrukking $(D_k((D_jF)\circ\psi))(u)$, gaan herschrijven met behulp van (11.23). We schrijven hiertoe (11.23) eerst als

$$(D_k(F\circ\psi))(u) = \sum_{i=1}^{n}(D_k\psi_i)(u)\,(D_iF)(\psi(u)) \quad (u \in E_2).$$

Dan vervangen we in deze formule F door D_jF. De formule blijft geldig omdat F een C^2-functie is, dus D_jF een C^1-functie. We verkrijgen

$$(D_k((D_jF)\circ\psi))(u) = \sum_{i=1}^{n}(D_k\psi_i)(u)\,(D_{ij}F)(\psi(u)) \quad (u \in E_2).$$

Als we dit substitueren in (11.25) dan volgt er dat

$$\begin{aligned}(D_{kl}f)(u) = &\sum_{i,j=1}^{n}(D_k\psi_i)(u)\,(D_l\psi_j)(u)\,(D_{ij}F)(\psi(u))\\ &+\sum_{j=1}^{n}(D_{kl}\psi_j)(u)\,(D_jF)(\psi(u)) \quad (u \in E_2).\end{aligned}$$

Door substitutie van $u := \phi(x)$ verkrijgen we tenslotte het volgende resultaat.

$$\begin{aligned}(D_{kl}f)(\phi(x)) = &\sum_{i,j=1}^{n}(D_k\psi_i)(\phi(x))\,(D_l\psi_j)(\phi(x))\,(D_{ij}F)(x)\\ &+\sum_{j=1}^{n}(D_{kl}\psi_j)(\phi(x))\,(D_jF)(x) \quad (x \in E_1). \qquad (11.26)\end{aligned}$$

Formule (11.26) wordt in de praktijk bij concrete coördinatentransformaties geregeld toegepast. Als men de formule niet paraat heeft dan kan men hem opzoeken of opnieuw afleiden volgens bovenstaande methode, maar sneller gaat een afleiding op een meer formele manier die we nu zullen uitleggen.

Merk eerst op dat de essentie van (11.26) kan worden geschreven als de volgende identiteit van partiële differentiaaloperatoren.

$$\frac{\partial^2}{\partial u_k\,\partial u_l} = \sum_{i,j=1}^{n}\frac{\partial x_i}{\partial u_k}\,\frac{\partial x_j}{\partial u_l}\,\frac{\partial^2}{\partial x_i\,\partial x_j} + \sum_{j=1}^{n}\frac{\partial^2 x_j}{\partial u_k\,\partial u_l}\,\frac{\partial}{\partial x_j}\,. \qquad (11.27)$$

Hierbij vatten we de x_i op als functies $x_i(u) = \psi_i(u)$ van $u = (u_1, \ldots, u_n)$ en laten we de partiële differentiaaloperatoren van beide leden van (11.27) werken op $f(u)$ of $F(x)$, waarbij $F(x) = f(u(x))$ en $f(u) = F(x(u))$. De resulterende uitdrukkingen verkregen door beide leden van (11.27) op zulke functies te laten werken worden dan tenslotte zo geschreven dat òf alles van u afhangt òf alles van x afhangt, waarbij je weer op de bekende manier van x naar u kan switchen of omgekeerd. Evenzo kan de essentie van (9.22) worden geschreven als (9.24), welke we nu in nog meer uitgeklede vorm schrijven als

$$\frac{\partial}{\partial u_l} = \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l} \frac{\partial}{\partial x_j}\,. \tag{11.28}$$

We leiden nu (11.27) formeel af uit (11.28).

$$\begin{aligned}\frac{\partial^2}{\partial u_k\,\partial u_l} &= \frac{\partial}{\partial u_k} \circ \Big(\sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l} \frac{\partial}{\partial x_j}\Big) = \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l} \frac{\partial}{\partial u_k} \circ \frac{\partial}{\partial x_j} + \sum_{j=1}^{n} \frac{\partial^2 x_j}{\partial u_k\,\partial u_l} \frac{\partial}{\partial x_j} \\ &= \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l} \Big(\sum_{i=1}^{n} \frac{\partial x_i \partial u_k}{} \frac{\partial^2}{\partial x_i\,\partial x_j}\Big) + \sum_{j=1}^{n} \frac{\partial^2 x_j}{\partial u_k\,\partial u_l} \frac{\partial}{\partial x_j}\,.\end{aligned} \tag{11.29}$$

Hierbij gebruikten we in de eerste en derde identiteit formule (11.28) en in de tweede identiteit formule (11.24).

11.17 Voorbeeld. We bekijken enerzijds Cartesische coördinaten (x, y) en anderzijds poolcoördinaten (r, θ), die met elkaar verbonden zijn volgens (9.27). In §9.8 drukten we $\partial/\partial x$ en $\partial/\partial y$ uit in termen van partiële differentiaaloperatoren naar r en θ. Nu willen we hetzelfde doen voor $\partial^2/\partial x^2$ en $\partial^2/\partial y^2$. Als we hiertoe (11.27) willen toepassen dan moeten we dus de partiële afgeleiden van eerste en tweede orde van r en θ naar x en y berekenen. Liever vinden we de gewenste uitdrukkingen door uit te gaan van (9.33) en (9.34) en dan te schrijven

$$\begin{aligned}&\frac{\partial^2}{\partial x^2} = \Big(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r} \frac{\partial}{\partial \theta}\Big) \circ \Big(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r} \frac{\partial}{\partial \theta}\Big) \\ &= \cos^2\theta \frac{\partial^2}{\partial r^2} - \frac{2\sin\theta\,\cos\theta}{r} \frac{\partial^2}{\partial r \partial \theta} + \frac{\sin^2\theta}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\sin^2\theta}{r} \frac{\partial}{\partial r} + \frac{2\sin\theta\,\cos\theta}{r^2} \frac{\partial}{\partial \theta}\,.\end{aligned} \tag{11.30}$$

Hierbij gebruikten we in de tweede identiteit formule (11.24). Ook gebruikten we dat $(\partial/\partial r) \circ (\partial/\partial \theta) = (\partial/\partial \theta) \circ (\partial/\partial r)$ (wanneer werkend op een C^2-functie) en voegden we termen samen. Merk op dat deze afleiding in grote lijnen gelijk is aan de daarboven gegeven afleiding (11.29), echter met het volgende verschil. Het zou met de tweede identiteit van (11.29) gecorrespondeerd hebben om (11.24) toe te passen op de uitdrukking

$$\frac{\partial}{\partial x} \circ \Big(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r} \frac{\partial}{\partial \theta}\Big).$$

In dit voorbeeld is het echter handiger om $\partial/\partial x$ in de uitdrukking hierboven eerst ook te herschrijven met behulp van (9.33) en dan pas (11.24) toe te passen.

De verschillende handelwijzen zullen uiteindelijk allemaal tot hetzelfde resultaat leiden, maar er is feeling voor nodig om in te zien wat in een gegeven situatie het snelst en het soepelst gaat. Bijvoorbeeld, als we omgekeerd $\partial^2/\partial r^2$ willen uitdrukken in termen van partiële differentiaaloperatoren naar x en y dan geeft het omslachtige berekeningen wanneer we in $\frac{\partial}{\partial r} \circ \frac{\partial}{\partial r}$ twee keer (9.29) substitueren en dan (11.24) toepassen. Veel handiger is nu om (9.29) te schrijven in de vorm

$$\frac{\partial}{\partial r} = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}, \tag{11.31}$$

en dan af te leiden dat

$$
\begin{aligned}
\frac{\partial^2}{\partial r^2} &= \frac{\partial}{\partial r} \circ \Big(\cos\theta \,\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial y}\Big) \\
&= \cos\theta\,\frac{\partial}{\partial r} \circ \frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial r} \circ \frac{\partial}{\partial y} \\
&= \cos^2\theta\,\frac{\partial^2}{\partial x^2} + 2\sin\theta\,\cos\theta\,\frac{\partial^2}{\partial x \partial y} + \sin^2\theta\frac{\partial}{\partial y^2} \\
&= \frac{x^2}{x^2+y^2}\,\frac{\partial^2}{\partial x^2} + \frac{2xy}{x^2+y^2}\,\frac{\partial^2}{\partial x\,\partial y} + \frac{y^2}{x^2+y^2}\,\frac{\partial^2}{\partial y^2}\,. \qquad (11.32)
\end{aligned}
$$

Hierbij gebruikten we in de eerste en derde identiteit formule (11.31) en in de tweede identiteit formule (11.24).

Opgave 11.8. Bewijs dat

$$\frac{\partial^2}{\partial y^2} = \sin^2\theta\,\frac{\partial^2}{\partial r^2} + \frac{2\sin\theta\,\cos\theta}{r}\,\frac{\partial^2}{\partial r\,\partial\theta} + \frac{\cos^2\theta}{r^2}\,\frac{\partial^2}{\partial\theta^2} + \frac{\cos^2\theta}{r}\,\frac{\partial}{\partial r} - \frac{2\sin\theta\,\cos\theta}{r^2}\,\frac{\partial}{\partial\theta}\,.$$

Leid hieruit door combinatie met (11.30) af dat

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\,\frac{\partial^2}{\partial\theta^2} + \frac{1}{r}\,\frac{\partial}{\partial r}\,. \qquad (11.33)$$

Bewijs ook dat

$$\frac{\partial^2}{\partial\theta^2} = y^2\,\frac{\partial^2}{\partial x^2} - 2xy\,\frac{\partial^2}{\partial x\partial y} + x^2\,\frac{\partial^2}{\partial y^2} - x\,\frac{\partial}{\partial x} - y\,\frac{\partial}{\partial y}\,.$$

Leid hieruit, door combinatie met (11.32) en (9.29), formule (11.33) nogmaals af.

De partiële differentiaaloperator $\partial^2/\partial x^2 + \partial^2/\partial y^2$ wordt de *Laplace-operator* (in twee variabelen) genoemd. Het rechterlid van (11.33) geeft de Laplace-operator in poolcoördinaten.

Opgave 11.9. Bepaal alle C^2-functies f op $(0,\infty)$ zo dat

$$\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big) f(\sqrt{x^2+y^2}) = 0 \quad ((x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}).$$

Opgave 11.10.

(a) Zij $f(z) := z^n$ ($z \in \mathbb{C}$, $n = 0, 1, 2, \ldots$). Bewijs dat de limiet

$$f'(z) := \lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{f(z+h) - f(z)}{h} \quad (z \in \mathbb{C}) \qquad (11.34)$$

bestaat en gelijk is aan $n\,z^{n-1}$.

(b) Zij f een complexwaardige functie op $\mathbb{C}$ waarvoor de limiet (11.34) bestaat. (Dan noemen we f *complex differentieerbaar.*) Bewijs dat

$$\frac{\partial f(x+iy)}{\partial x} = f'(x+iy), \quad \frac{\partial f(x+iy)}{\partial y} = i\,f'(x+iy) \qquad (x,y) \in \mathbb{R}^2). \qquad (11.35)$$

Opgave 11.11. Bewijs dat voor $n = 0, 1, 2, \ldots$ geldt dat

$$\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big) \Big\{(x \pm iy)^n\Big\} = 0. \tag{11.36}$$

Doe dit enerzijds door directe partiële differentiatie, anderzijds door over te gaan op poolcoördinaten en gebruik te maken van (11.33).

Opgave 11.12. Zij $|x| := \sqrt{x_1^2 + \cdots + x_n^2}$ en zij Δ de partiële differentiaaloperator

$$\Delta := D_{11} + \cdots + D_{nn} = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\,. \tag{11.37}$$

(Bij de tweede schrijfwijze veronderstellen we dat Δ werkt op functies in de variabelen $x_1, \ldots, x_n$.) Zij $\alpha \in \mathbb{R}$. Bewijs dat

$$\Delta(|x|^\alpha) = \alpha\,(n + \alpha - 2)\,|x|^{\alpha - 2}. \tag{11.38}$$

11.18. De operator Δ gegeven door (11.37) heet *Laplace-operator* in n variabelen. Een C^2-functie f gedefinieerd op een open deelverzameling E van $\mathbb{R}^2$, die op E voldoet aan $\Delta f = 0$ heet *harmonisch* op E. Concludeer uit (11.38) dat f gedefinieerd door $f(x) := |x|^{-n+2}$ harmonisch is op $\mathbb{R}^n \backslash \{0\}$.

Opgave 11.13. Zij $|x| := \sqrt{x_1^2 + \cdots + x_n^2}$ en zij Δ gegeven door (11.37). Zij $f : (0, \infty) \to \mathbb{R}$ een C^2-functie. Bewijs dat

$$\Delta(f(|x|)) = f''(|x|) + (n - 1)\,|x|^{-1}\,f'(|x|). \tag{11.39}$$

Verdere vraagstukken

Opgave 11.14. We beschouwen de coördinatentransformatie

$$u = x + ct, \qquad v = x - ct$$

tussen het (x, t)-vlak en het (u, v)-vlak. Hierin is c een positieve constante.

(a) Laat zien dat de zogenaamde ééndimensionale golfvergelijking

$$\frac{\partial^2 F}{\partial x^2} = \frac{1}{c^2}\,\frac{\partial^2 F}{\partial t^2} \tag{11.40}$$

onder deze transformatie overgaat in

$$4\,\frac{\partial^2 F}{\partial u\,\partial v} = 0.$$

(b) Toon aan dat iedere C^2-functie F die voldoet aan de vergelijking (11.40) geschreven kan worden in de vorm

$$F(x, y) = f(x + ct) + g(x - ct)$$

waarin $f, g : \mathbb{R} \to \mathbb{R}$ twee tweemaal differentieerbare functies zijn.

12 Taylorreeks in gebruik

In dit hoofdstuk behandelen we Taylorreeksen voor functies in meer veranderlijken op zo'n manier dat deze goed toepasbaar zijn in het hoofdstuk over extremen, maar zonder een volkomen analogie met de situatie in één veranderlijke te bereiken. Het corresponderende hoofdstuk in Koornwinder's versie cf.

`http://staff.science.uva.nl/~thk/edu/oldsyl/analb1.pdf`

geeft meer inzicht hierin, maar vereist veel van het abstractievermogen van de lezer. Een aantal vraagstukken in dat hoofdstuk is ook met hetgeen we nu zullen behandelen oplosbaar.

In K.A. Ross, Elementary Analysis, Sectie 31.1 wordt voor één veranderlijke geraden hoe de coëfficiënten van de Taylorreeks er uit zien. We proberen dit ook in meer veranderlijken. Om van convergentievragen af te zijn beginnen we met een polynoom in n veranderlijken van de vorm.

$$P(x) = P(x_1, \ldots, x_n) = \sum_{j_1,\ldots,j_n=0}^{k} a_{j_1,\ldots,j_n} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}. \tag{12.1}$$

Ieder polynoom van graad k laat zich als (12.1) schrijven, waarbij natuurlijk $a_{j_1,\ldots,j_n} = 0$ als $j_1 + j_2 + \cdots + j_n > k$. We rekenen eens een hogere partiële afgeleide van P uit in 0 met $0 \le j_1, j_2, \ldots, j_n \le k$:

$$(D_1^{j_1} D_2^{j_2} \cdots D_n^{j_n} P)(0) = a_{j_1,\ldots,j_n} j_1! j_2! \cdots j_n!$$

Ofwel

$$a_{j_1,\ldots,j_n} = \frac{(D_1^{j_1} D_2^{j_2} \cdots D_n^{j_n} P)(0)}{j_1! j_2! \cdots j_n!}. \tag{12.2}$$

Deze formule laat zien hoe we de coëfficiënten van een polynoom uitdrukken in de partiële afgeleiden van dat polynoom in de oorsprong. Hij suggereert ook hoe je voor een functie f die op een omgeving van 0 C^∞ is, in $\mathbb{R}^n$ de Taylorreeks moet definiëren:

12.1 Definitie. De *Taylorreeks* van een functie f die C^∞ is op een omgeving van 0 in $\mathbb{R}^n$ is de reeks

$$\sum_{j_1,\ldots,j_n=0}^{\infty} \frac{(D_1^{j_1} D_2^{j_2} \cdots D_n^{j_n} f)(0)}{j_1! j_2! \cdots j_n!} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}. \tag{12.3}$$

Vaak nemen we de termen van gelijke graad bijeen. Dan krijgen we voor (12.3)

$$\sum_{k=0}^{\infty} P_k(x) \tag{12.4}$$

met

$$P_j(x) = \sum_{j_1+\cdots+j_n=k} \frac{(D_1^{j_1} D_2^{j_2} \cdots D_n^{j_n} f)(0)}{j_1! j_2! \cdots j_n!} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}. \tag{12.5}$$

De *k-de restterm* wordt gedefinieerd door

$$R_k(x) = f(x) - \sum_{j=0}^{k-1} P_j(x) = f(x) - \sum_{j_1+\cdots+j_n<k} \frac{(D_1^{j_1} D_2^{j_2} \cdots D_n^{j_n} f)(0)}{j_1! j_2! \cdots j_n!} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}. \tag{12.6}$$

De waarde van Taylors formule ligt hierin dat de reeks soms naar de functie convergeert of dat men kan inzien dat de restterm klein is. We merken nog op dat de uitdrukking (12.6) zin heeft wanneer $f \in C^{k-1}$ bij 0, en dat goede schattingen voor R_k mogelijk zijn als f tenminste in C^k bij 0. We zullen schattingen voor R_k bewijzen uit het één variabele geval.

Het volgende stuk is sectie 3 van het volgende hoofdstuk zo herschreven dat kennis en begrip van hogere totale afgeleiden en multilineaire afbeeldingen vermeden is. We recapituleren de *stelling van Taylor* voor reëelwaardige functies van één reële variabele (cf. K.A. Ross, Elementary Analysis, Sectie 31).

12.2 Stelling. Zij I een open interval in $\mathbb{R}$. Zij $f : I \to \mathbb{R}$ een C^p-functie voor zekere $p \geq 1$. Zij $a \in I$ en zij $h \in \mathbb{R}$ zo dat $a + h \in I$. Dan bestaat er een $\theta \in (0, 1)$ (afhangend van a en h) zo dat

$$f(a+h) = f(a) + f'(a)\,h + \frac{f''(a)\,h^2}{2!} + \cdots + \frac{f^{(p-1)}(a)\,h^{p-1}}{(p-1)!} + \frac{f^{(p)}(a+\theta h)\,h^p}{p!}\,. \tag{12.7}$$

12.3. Formule (12.7) staat bekend als de *Taylorreeks met restterm*. De Stelling geldt onder zwakkere aannamen (cf. Rudin, §5.15). Zo behoeft de $p$de afgeleide niet continu te zijn. Het geval $p = 1$ is precies de middelwaardestelling (11.1).

Neem a in (12.7) vast en herschrijf (12.7) voor variabele h als

$$f(a+h) = f(a) + \sum_{k=1}^{p-1} \frac{f^{(k)}(a)\,h^k}{k!} + R_p(h),$$

waarbij de restterm gegeven wordt door

$$R_p(h) := \frac{f^{(p)}(a+\theta\,h)\,h^p}{p!}\,,$$

met $\theta \in (0, 1)$ afhangend van h. Gezien de continuïteit van $f^{(p)}$ zal gelden dat

$$R_p(h) = \mathcal{O}(|h|^p) \quad \text{als} \quad h \to 0$$

en

$$R_p(h) = \frac{f^{(p)}(a)\,h^p}{p!} + o(|h|^p) \quad \text{als} \quad h \to 0.$$

Dus voor een C^p-functie f op een open interval I en $a \in I$ gelden ook de volgende varianten van de Taylorreeks (12.7).

$$f(a+h) = f(a) + \sum_{k=1}^{p-1} \frac{f^{(k)}(a)\,h^k}{k!} + \mathcal{O}(|h|^p) \quad \text{als} \quad h \to 0 \tag{12.8}$$

en

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{f^{(k)}(a)\,h^k}{k!} + o(|h|^p) \quad \text{als} \quad h \to 0. \tag{12.9}$$

12.4. We willen nu een analogon van (12.7) geven voor het geval f van meer dan één reële variabele afhangt, maar nog reëelwaardig is. We zullen we het resultaat bewijzen door het terug te brengen tot het geval van één variabele. Ter voorbereiding geven we een lemma.

12.5 Lemma. Zij $a \in E \subset \mathbb{R}^n$ met E open, zij $f : E \to \mathbb{R}$ een C^p-functie en zij $h \in \mathbb{R}^n$. Definieer de functie F op de open deelverzameling $\{t \in \mathbb{R} \mid a + t\,h \in E\}$ van $\mathbb{R}$ door $F(t) := f(a + t\,h)$. Dan is F een C^p-functie en er geldt voor $k = 1, \ldots, p$ dat

$$F^{(k)}(t) = \sum_{j_1 \ldots j_k = 1}^{n} (D_{j_1 \ldots j_k} f)(a + t\,h)\, h_{j_1} \ldots h_{j_k}. \tag{12.10}$$

Bewijs. Het bewijs gaat met volledige inductie naar k. We hebben al vaker gezien dat, als $g : E \to \mathbb{R}$ een C^1-functie is en $G(t) := g(a + t\,h)$, dan

$$G'(t) = \sum_{j=1}^{n} (D_j g)(a + t\,h)\, h_j. \tag{12.11}$$

Dit geeft het geval $k = 1$ van (12.10). Bovendien, als $p > k$, dan geeft (12.11) voor $g := D_{j_1 \ldots j_k} f$ dat

$$G'(t) = \sum_{j=1}^{n} (D_{j j_1 \ldots j_k} f)(a + t\,h)\, h_j.$$

Dus differentiatie van beide leden van (12.10) geeft dezelfde formule met k vervangen door $k + 1$. □

12.6 Stelling. Zij $a \in E \subset \mathbb{R}^n$ met E open en zij $f : E \to \mathbb{R}$ een C^p-functie. Zij $h \in \mathbb{R}^n$ zo dat $[a, a + h] \in E$. Dan is er $\theta \in (0, 1)$ zo dat

$$f(a + h) = \sum_{k=0}^{p-1} \frac{1}{k!} \left(\sum_{j_1, \ldots, j_k = 1}^{n} (D_{j_1 \ldots j_k} f)(a)\, h_{j_1} \ldots h_{j_k} \right) + R_p(h), \tag{12.12}$$

waarbij

$$R_p(h) = \frac{1p!}{\sum_{j_1, \ldots, j_p = 1}}^{n} (D_{j_1 \ldots j_p} f)(a + \theta\,h)\, h_{j_1} \ldots h_{j_p}. \tag{12.13}$$

Bovendien geldt er dat

$$R_p(h) = \mathcal{O}(|h|^p) \quad \text{als} \quad h \to 0 \quad \text{in} \quad \mathbb{R}^n \tag{12.14}$$

en

$$R_p(h) = \frac{1}{p!} \sum_{j_1, \ldots, j_p = 1}^{n} (D_{j_1 \ldots j_p} f)(a)\, h_{j_1} \ldots h_{j_p} + o(|h|^p) \quad \text{als} \quad h \to 0 \quad \text{in} \quad \mathbb{R}^n. \tag{12.15}$$

Bewijs. Schrijf $F(t) := f(a + t\,h)$ en pas Stelling 12.2 toe op F. Dan is er dus $\theta \in (0, 1)$ zo dat

$$F(1) = F(0) + F'(0) + \frac{F''(0)}{2!} + \cdots + \frac{F^{(p-1)}(0)}{(p-1)!} + \frac{F^{(p)}(\theta)}{p!}.$$

Substitueer nu (12.10). Dit bewijst (12.12) samen met (12.13).

Voor het bewijs van (12.15) merken we eerst op dat de partiële afgeleiden tot en met orde p continu zijn. Daarmee concluderen we uit (12.13) dat

$$\begin{aligned} &|R_p(h) - \frac{1}{p!} \sum_{j_1, \ldots, j_p = 1}^{n} (D_{j_1 \ldots j_p} f)(a)\, h_{j_1} \ldots h_{j_p}| \\ &\quad \le \frac{1}{p!} \sum_{j_1, \ldots, j_p = 1}^{n} \big|(D_{j_1 \ldots j_p} f)(a + \theta h) - (D_{j_1 \ldots j_p} f)(a)\big|\, |h_{j_1} \ldots h_{j_p}| = o(|h|^p). \end{aligned}$$

Dit is (12.15). Daarmee volgt (12.14) uit (12.15). □

12.7. Laten we deze Stelling wat toelichten. In het geval van één variabele geeft formule (12.9) de benadering van de C^p-functie f door een polynoom (het zogenaamde *Taylorpolynoom van orde* p) $x \mapsto f(a) + \sum_{k=1}^{p} f^{(k)}(a)\,(x-a)^k/k!$ van graad $\le p$ zo dat de benadering voor x nabij a zo goed mogelijk is, d.w.z. dat het verschil van $f(x)$ en het benaderende polynoom van orde kleiner dan $|x-a|^p$ is voor $x \mapsto a$. Dit geeft een generalisatie van het benaderende eerstegraads-polynoom $x \mapsto f(a) + f'(a)\,(x-a)$ dat de raaklijn in $(a, f(a))$ aan de grafiek van f bepaalt en dat we gebruikten om een nieuwe karakerisering te geven van de afgeleide $f'(a)$. Goed beschouwd geeft (12.9) een nieuwe karakterisering van de afgeleiden $f^{(k)}(a)$ tot/met orde p, nl. als de coëfficiënten (op een factor $1/k!$ na) van het best benaderende polynoom van graad $\le p$ nabij $x = a$.

Evenzo verkrijgen we uit (12.12) op grond van (12.15) een uitdrukking voor het polynoom van graad $\le p$ in de n variabelen $x_1, \ldots, x_n$ dat de functie f nabij $x = a$ zo goed mogelijk benadert, d.w.z. met een verschilfunctie van kleinere orde dan $|x-a|^p$ als $x \to a$. Dit polynoom heet weer het *Taylorpolynoom van orde* p. Een eenvoudig gevolg van deze stelling dat zeer belangrijk is, is het volgende resultaat.

12.8 Stelling. Zij $a \in E \subset \mathbb{R}^n$ met E open en zij $f : E \to \mathbb{R}$ een C^p-functie. Dan heeft f precies één Taylorpolynoom van orde p in a. Met andere woorden, Taylorontwikkelingen zijn uniek en als men op een of andere manier een polynoom P van graad $\le p$ vindt zodat $f(x) - P(x) = o(|x-a|^p)$, dan is P het Taylor polynoom.

Bewijs. Stel P_1 en P_2 zijn Taylorpolynomen van orde p voor f in a. Dan is $P_1 - P_2 = c_0 + c_1(x-a) + \cdots + c_p(x-a)^p$ een polynoom van graad $\le p$. Echter,

$$P_1 - P_2 = (P_1 - f) - (P_2 - f) = o(|x-a|^p).$$

Dit kan alleen als $c_j = 0$ voor iedere j, dus $P_1 = P_2$. $\square$

12.9. Behoud de aannamen van Stelling 12.6, behalve dat $f : E \to \mathbb{R}^m$ i.p.v. $f : E \to \mathbb{R}$. Dan is het nog steeds zinvol om (12.12) op te schrijven. We kunnen deze formule opvatten als de definitieformule voor $R_p(h) \in \mathbb{R}^m$. Formule (12.13) hoeft nu niet meer te gelden voor zekere $\theta \in (0,1)$, immers voor iedere component van f zal in het algemeen een eigen θ hebben. Maar (12.14) en (12.15) blijven gelden.

12.10. Zij $E \subset \mathbb{R}^n$ open en $f : E \to \mathbb{R}$ een C^p-functie. Zij $j \in \{1, \ldots, n\}$. Dan is $D_j f$ een C^{p-1}-functie en we kunnen D_j als lineaire afbeelding opvatten van de ruimte van C^p-functies op E naar de ruimte van C^{p-1}-functies op E. Meer compact schrijven we dit als $D_j : C^p(E) \to C^{p-1}(E)$. Hier wordt $C^p(E)$ een reële vectorruimte t.o.v. puntsgewijze optelling en scalaire vermenigvuldiging, i.e., $(f+g)(x) := f(x) + g(x)$ en $(\lambda f)(x) := \lambda f(x)$. Nu wordt het zinvol om producten van zulke lineaire afbeeldingen te bekijken, zoals $D_{j_1} \ldots D_{j_k}$. Dus

$$D_{j_1} \ldots D_{j_k} f = D_{j_1 \ldots j_k} f \quad \text{als} \quad f \in C^p(E) \quad \text{en} \quad k \le p.$$

Ook kunnen we sommen en scalaire producten van zulke lineaire afbeeldingen bekijken. Het zal nu bijv. duidelijk zijn wat we bedoelen met $(h_1 D_1 + \cdots + h_n D_n)\, f$ en $(h_1 D_1 + \cdots + h_n D_n)^k\, f$ als $f \in C^p(E)$, $k \le p$ en $h_1, \ldots, h_n \in \mathbb{R}$. Merk verder op dat de operatoren D_i en D_j commuteren (i.e. $D_i D_j = D_j D_i$) wanneer ze op C^2-functies werken.

De uitdrukking tussen haken in het rechter lid van (12.12) luidt

$$\sum_{j_1, \ldots, j_k = 1}^{n} h_{j_1} \ldots h_{j_k}\, (D_{j_1 \ldots j_k} f)(x).$$

Deze som bevat veel meer termen dan nodig is. We kunnen alle termen die uit elkaar verkregen worden door een permutatie van $\{1, \ldots, k\}$ (in de subindices van de tweede orde) samennemen.

12.11 Propositie. Zij $E \subset \mathbb{R}^n$ open en zij $f : E \to \mathbb{R}$ een C^p-functie. Zij $k \le p$. Dan geldt dat

$$\begin{aligned}\sum_{j_1,\dots,j_k=1}^{n} h_{j_1}\dots h_{j_k}\, D_{j_1\dots j_k} f &= (h_1 D_1 + \dots + h_n D_n)^k\, f \\ &= \sum_{k_1+\dots+k_n=k} \frac{k!}{k_1!\dots k_n!}\, h_1^{k_1}\dots h_n^{k_n}\, D_1^{k_1}\dots D_n^{k_n}\, f.\end{aligned}$$

Bewijs. De eerste identiteit is triviaal. Merk nu op dat de operatoren $h_1 D_1, \dots, h_n D_n$ met elkaar commuteren. We kunnen dus de *multinomiaalformule*

$$(x_1 + \dots + x_n)^k = \sum_{k_1+\dots+k_n=k} \frac{k!}{k_1!\dots k_n!}\, x_1^{k_1}\dots x_n^{k_n}$$

toepassen. (Deze formule volgt met volledige inductie naar n uit het speciale geval $n = 2$, i.e. uit de binomiaalformule.) □

12.12 Opmerking. Als een gevolg van deze Propositie kunnen we bijvoorbeeld de Taylorreeks (12.12) herschrijven als

$$f(a+h) = f(a) + \sum_{j=1}^{n} h_j\,(D_j f)(a) + \sum_{k=2}^{p-1} \sum_{k_1+\dots k_n=k} \frac{h_1^{k_1}\dots h_n^{k_n}\,(D_1^{k_1}\dots D_n^{k_n} f)(a)}{k_1!\dots k_n!} + R_p(h). \quad (12.16)$$

We hebben daarmee de uitdrukkingen (12.1) en (12.6) teruggevonden.

Uiteraard verandert er niets wezenlijks aan deze formule als we $h := x - a$ substitueren. Het is goed om ook aan die gedaante van de Taylorreeks gewend te raken:

$$\begin{aligned}f(x) &= f(a) + \sum_{j=1}^{n} (x_j - a_j)\,(D_j f)(a) \\ &+ \sum_{k=2}^{p-1} \sum_{k_1+\dots k_n=k} \frac{(x_1-a_1)^{k_1}\dots(x_n-a_n)^{k_n}\,(D_1^{k_1}\dots D_n^{k_n} f)(a)}{k_1!\dots k_n!} + R_p(x-a). \quad (12.17)\end{aligned}$$

Verdere vraagstukken

Opgave 12.1. Geef het Taylorpolynoom van orde 3 in het punt $(0, 0, 1)$ van de functies

$$f : \mathbb{R}^3 \to \mathbb{R}, \qquad f(x, y, z) = xyz.$$

en

$$f : \mathbb{R}^3 \to \mathbb{R}, \qquad f(x, y, z) = x^2 \cos y - e^z \sin(x + y).$$

Opgave 12.2. Zij $f : \mathbb{R}^2 \to \mathbb{R}$ een functie en $a \in \mathbb{R}^2$. Bepaal bij f en a een polynoom $P : \mathbb{R}^2 \to \mathbb{R}$ van graad ≤ 2 zo dat

$$f(x) = P(x) + o(|x - a|^2) \quad \text{als} \quad x \to a$$

als f en a als volgt gegeven zijn:

(a) $f(x, y) := x^2 y$ en $a := (1, -1)$;

(b) $f(x, y) := e^x \sin y$ en $a := (0, 0)$;

(c) $f(x,y) := \sqrt{1+x^2+y^2}$ en $a := (0,0)$.

Opgave 12.3. Bepaal een polynoom $p_2(x,y)$ van graad ≤ 2 zo dat

$$\frac{1}{1-x+y} = p_2(x,y) + \mathcal{O}(|(x-1,y-1)|^3) \quad \text{als} \quad (x,y) \to (1,1).$$

Opgave 12.4. Zij de functie $f : \mathbb{R}^2 \to \mathbb{R}$ gegeven door $f(x,y) := \exp(-(x^2+y^2)^{-1})$ als $(x,y) \neq (0,0)$ en $f(0,0) := 0$. Bewijs dat f een C^∞-functie is op $\mathbb{R}^2$. Bepaal voor iedere n een polynoom $p_n(x,y)$ van graad $\leq n$ zo dat $f(x,y) - p_n(x,y) = o(|(x,y)|^n)$ als $(x,y) \to (0,0)$.

13 Extremen

In het vervolg onderscheiden we verschillende soorten lokale maxima en minima die een reëelwaardige functie op een een deelverzameling van $\mathbb{R}^n$ kunnen aannemen. Vervolgens geven we criteria in termen van eerste of tweede afgeleide voor het aannemen van zulke extrema.

13.1 Definitie. Zij (X, d) een metrische ruimte en $f : X \to \mathbb{R}$.

(a) De functie f neemt een *lokaal maximum* aan in een punt $a \in X$ als er $\delta > 0$ bestaat zo dat $f(x) \leq f(a)$ voor alle $x \in X$ met $d(x, a) < \delta$, d.w.z. als er een omgeving U van a in X bestaat zo dat $\max_{x \in U} f(x) = f(a)$.

Analoog zeggen we dat de functie f een *lokaal minimum* aanneemt in een punt $a \in X$ als er $\delta > 0$ bestaat zo dat $f(x) \geq f(a)$ voor alle $x \in X$ met $d(x, a) < \delta$.

De functie f neemt een *lokaal extremum* aan in een punt $a \in X$ als f in dat punt een lokaal maximum of een lokaal minimum aanneemt.

(b) Bij lokale maxima maken we onderscheid tussen tussen *absolute* en *relatieve* maxima. De functie f neemt een *absoluut maximum* aan in een punt $a \in X$ als $f(x) \leq f(a)$ voor alle $x \in X$, d.w.z. als $\max_{x \in X} f(x) = f(a)$. De functie f neem een *relatief maximum* aan in a als f in a een lokaal maximum aanneemt dat niet absoluut is, d.w.z. dat de eigenschap heeft dat $f(b) > f(a)$ voor zekere $b \in X$.

Analoog kunnen we definiëren wanneer f een *absoluut* of *relatief minimum* aanneemt.

(c) Bij lokale maxima maken we voorts onderscheid tussen *sterke* en *zwakke* lokale maxima. De functie f neemt een *sterk lokaal maximum* aan in a als we $\delta > 0$ zelfs zo kunnen kiezen dat de strikte ongelijkheid $f(x) < f(a)$ geldt voor alle $x \in X$ met $x \neq a$ en $d(x, a) < \delta$. De functie f neemt een *zwak lokaal maximum* aan in a als f in a een lokaal maximum aanneemt dat niet sterk is, d.w.z. dat de eigenschap heeft dat er bij iedere $\delta > 0$ een $x \in X$ is met $x \neq a$ en $d(x, a) < \delta$ zo dat $f(x) = f(a)$.

Analoog kunnen we definiëren wanneer f een *sterk* of *zwak lokaal minimum* aanneemt.

13.2. Zij (X, d) een metrische ruimte en $E \subset X$. In K.A. Ross, Elementary Analysis Definition 13.6 en 13.8 werden de begrippen *inwendig punt* en *randpunt* (van E t.o.v. X) gedefinieerd. Een punt $a \in E$ zal dan óf inwendig punt óf randpunt van E zijn. Immers:

- Een punt $a \in E$ is *inwendig punt* van E als er $\delta > 0$ bestaat zo dat $x \in E$ voor alle $x \in X$ met $d(x, a) < \delta$, d.w.z. als er een omgeving U van a in X bestaat zo dat $U \subset E$.
- Een punt $a \in E$ is een *randpunt* van E als voor elke $\delta > 0$ er $x \in X$ bestaat met $d(x, a) < \delta$ en $x \notin E$, d.w.z. als elke omgeving van a in X niet-lege doorsnede met het complement van E heeft.

13.3 Definitie. Zij nu $E \subset \mathbb{R}^n$ en $f : E \to \mathbb{R}$. Dan is E als deelverzameling van $\mathbb{R}^n$ een metrische ruimte met metriek $d(x, y) = |x - y|$, en Definitie 13.1 is in deze situatie toepasbaar. Bij lokale maxima kunnen we nu ook onderscheid maken tussen *inwendige* en *rand*maxima. Als f een lokaal maximum aanneemt in $a \in E$ dan spreken we van een *lokaal inwendig maximum* als a een inwendig punt is van E t.o.v. $\mathbb{R}^n$, en we spreken van een *lokaal randmaximum* als a een randpunt is van E t.o.v. $\mathbb{R}^n$.

Analoog kunnen we een *lokaal inwendig minimum* en een *lokaal randminimum* definiëren.

13.4. Nog een opmerking over de terminologie. Als we spreken over (lokale) extrema van een functie, dan gaat het veel meer om de punten waar de functie een lokaal extremum aanneemt dan de extreme waarden van de functie in die punten. Dit is verwarrend als men dit vergelijkt met het begrip "maximum van een verzameling reële getallen". Het zou misschien zorgvuldiger zijn als men bij functies spreekt over extremaalpunten of maximaalpunten i.p.v. extrema of maxima, maar deze termen zijn niet zo ingeburgerd.

We gaan nu voor functies in meer veranderlijken de analoga bewijzen van resultaten die al vanaf het VWO bekend zijn. Voor het gemak vatten we die resultaten in onze terminologie nog eens samen. Vergelijk ook K.A. Ross, Elementary Analysis, Stelling 29.1 voor het eerste deel.

13.5 Propositie. Zij $I \subset \mathbb{R}$ een open interval, $a \in I$ en $f : I \to \mathbb{R}$.

(a) Zij f differentieerbaar in a. Als f een lokaal extremum aanneemt in a dan $f'(a) = 0$.

Zij f bovendien een C^2-functie.

(b) Als f een lokaal minimum (resp. maximum) aanneemt in a dan geldt dat $f''(a) \geq 0$ (resp. $f''(a) \leq 0$).

(c) Als $f'(a) = 0$ en $f''(a) > 0$ (resp. < 0) dan neemt f een sterk lokaal minimum (resp. maximum) aan in a.

Bewijs. Met betrekking tot het bewijs merken we op dat (b) een gevolg is van (c). Voor (c) gaan we als volgt te werk. Als $f'(a) = 0$, $f''(a) > 0$, dan is f' stijgend op een interval I om a, dus $f'(x) < 0$ als $x \in I$ en $x < a$, terwijl $f'(x) > 0$ als $x \in I$ en $x > a$. Dus f is dalend als $x \in I$ en $x < a$, en stijgend als $x \in I$ en $x > a$. In a heeft f dus een sterk locaal minimum. □

We onderzoeken nu voor functies van meer veranderlijken het verband tussen lokale extrema en eigenschappen van eerste totale afgeleide en de 2-de orde partiële afgeleiden. Zoals we al een paar keer eerder gezien hebben, zullen resultaten vaak bewezen worden door herleiding tot het geval van één variabele.

13.6 Propositie. Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f : E \to \mathbb{R}$. Laat f een lokaal extremum aannemen in a. Dan geldt:

(a) Zij $u \in \mathbb{R}^n$. Als de richtingsafgeleide $(D_u f)(a)$ bestaat dan is $(D_u f)(a) = 0$.

(b) Als de totale afgeleide $f'(a)$ bestaat dan is $f'(a) = 0$.

Bewijs. (a): (8.6) zegt dat $(D_u f)(a) = (\frac{d}{dt} f(a + ut))'(0) = 0$ volgens Propositie 13.5(a).

(b): Als uit $f'(a)$ bestaat, dan zeggen (8.23) en (a) dat $f'(a)u = (D_u f)(a) = 0$ voor iedere u. Dus $f'(a) = 0$. □

13.7 Definitie. Zij $E \subset \mathbb{R}^n$ en $f : E \to \mathbb{R}$. Een *stationair punt* van f is een inwendig punt a van E zo dat f differentieerbaar is in a en $f'(a) = 0$.

Opgave 13.1. Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f : E \to \mathbb{R}$ een C^1-afbeelding. Stel er zijn n C^1-krommen $\gamma_1, \ldots, \gamma_n$ in E zo dat $\gamma_j(0) = a$ $(j = 1, \ldots, n)$ en de raakvectoren $\gamma_1'(0), \ldots, \gamma_n'(0)$ lineair onafhankelijk zijn. Neem aan dat voor elke $j = 1, \ldots, n$ de functie $t \mapsto f(\gamma_j(t))$ een lokaal extremum aanneemt in 0. Bewijs dat a een stationair punt is van f.

13.8. Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f : E \to \mathbb{R}$ een C^2-functie. De rol die de tweede afgeleide speelt voor functies van één variabele, wordt overgenomen door het 2-de orde polynoom in de Taylorontwikkeling van een functie f van meer veranderlijken.

We weten dat dit polynoom gegeven wordt door

$$\sum_{i,j=1}^{n} (D_{ij}f)(a)\, x_i\, x_j, \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n. \tag{13.1}$$

De uitdrukking (13.1) heet een *kwadratische vorm*. Hij wordt bepaald door de symmetrische matrix

$$H_f(a) := \begin{pmatrix} (D_{11}f)(a) & \ldots & (D_{1n}f)(a) \\ \vdots & & \vdots \\ (D_{n1}f)(a) & \ldots & (D_{nn}f)(a) \end{pmatrix}, \tag{13.2}$$

de zogenaamde *Hesse-matrix* van f in a. (Verwar deze matrix niet met de $m \times n$ matrix $((D_j f_i)(a))_{i=1,\ldots,m;\, j=1,\ldots,n}$ van de lineaire afbeelding $f'(a)$, indien f een afbeelding is van een open deel van $\mathbb{R}^n$ naar $\mathbb{R}^m$.)

Uit lineaire algebra weten we dat

$$\sum_{i,j=1}^{n} (D_{ij}f)(a)\, x_i\, x_j = {}^{\tau}x H_f(a) x.$$

Hier is ${}^{\tau}x$ de getransponeerde van de vector $x \in \mathbb{R}^n$.

13.9 Propositie. Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f : E \to \mathbb{R}$ een C^2-functie. Als f een lokaal minimum (resp. maximum) aanneemt in a dan $f'(a) = 0$ en ${}^{\tau}h\, H_f(a)\, h \geq 0$ (resp. ≤ 0) voor alle $h \in \mathbb{R}^n$.

Bewijs. Stel f neemt een lokaal minimum aan in a. Dan is $f'(a) = 0$ wegens Propositie 13.6(b). Definieer voor vaste $h \in \mathbb{R}^n$ de C^2-functie F van één variabele door $F(t) := f(a + t\,h)$. Dan ${}^{\tau}h\, H_f(a)\, h = F''(0) \geq 0$ op grond van formule (12.10) en Propositie 13.5(b). □

13.10 Stelling. Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f : E \to \mathbb{R}$ een C^2-functie. Als $f'(a) = 0$ en ${}^{\tau}h\, H_f(a)\, h > 0$ (resp. < 0) voor alle $h \in \mathbb{R}^n \backslash \{0\}$ dan neemt f een sterk lokaal minimum (resp. maximum) aan in a.

Bewijs. Er geldt

$$\inf_{h \in \mathbb{R}^n \backslash \{0\}} \frac{{}^{\tau}h\, H_f(a)\, h}{|h|^2} = \inf_{h \in \mathbb{R}^n;\, |h|=1} {}^{\tau}h\, H_f(a)\, h = {}^{\tau}h^0\, H_f(a)\, h^0 = \mu > 0$$

voor zekere $h^0 \in \mathbb{R}^n$ met $|h^0| = 1$ en voor zekere positieve μ, omdat de functie $h \mapsto {}^{\tau}h\, H_f(a)\, h$ continu is op de compacte verzameling $\{h \in \mathbb{R}^n \mid |h| = 1\}$ en daarom een minimum aanneemt in zeker punt h^0. Dus ${}^{\tau}h\, H_f(a)\, h \geq \mu |h|^2$ voor alle $h \in \mathbb{R}^n$. Toepassing van (12.16) geeft dat

$$f(a + h) - f(a) = \frac{{}^{\tau}h\, H_f(a)\, h}{2} + o(|h|^2)$$

als $h \to 0$ in $\mathbb{R}^n$. Dus er bestaat $\delta > 0$ zo dat voor $0 < |h| < \delta$ geldt dat

$$f(a + h) - f(a) \geq \tfrac{1}{2}{}^{\tau}h\, H_f(a)\, h - \tfrac{1}{4}\mu\, |h|^2 \geq \tfrac{1}{4}\mu\, |h|^2 > 0.$$

□

We geven nu een paar algemene definities en eigenschappen betreffende kwadratische vormen en passen die vervolgens toe op de kwadratische vorm uit (13.1).

Schrijf V voor een eindigdimensionale vectorruimte $\mathbb{R}^n$, en zij B een symmetrische $n \times n$ matrix. Dan is $x \mapsto {}^\tau x\, B\, x$ een kwadratische vorm op V. We schrijven dit als $B(x,x)$ (naar analogie met het inproduct, waar $B = I$).

13.11 Definitie. De kwadratische vorm $x \mapsto B(x,x)$ noemen we *positief semidefiniet* als $B(x,x) \geq 0$ voor alle $x \in V$ en *positief definiet* als $B(x,x) > 0$ voor alle $x \in V\backslash\{0\}$. Analoog definiëren we *negatief (semi)definiet*. De kwadratische vorm heet *indefiniet* als er x en $y \in \mathbb{R}^n$ zijn zo dat $B(x,x) > 0$ en $B(y,y) < 0$.

13.12. Zij $A = (A_{ij})_{i,j=1,\ldots,n}$ een symmetrische $n \times n$ matrix (dus $A_{ij} = A_{ji}$). Dan weten we uit lineaire algebra dat A een symmetrische bilineaire vorm B op $\mathbb{R}^n \times \mathbb{R}^n$ bepaalt die gedefinieerd wordt door $B(x,y) = \langle Ay, x\rangle$ en een bijbehorende kwadratische vorm op $\mathbb{R}^n$ gegeven door

$$B(x,x) = \langle Ax, x\rangle = \sum_{i,j=1}^{n} A_{ij}\, x_i\, x_j, \quad x \in \mathbb{R}^n. \tag{13.3}$$

We noemen de symmetrische matrix A *positief* of *negatief (semi)definiet* als de kwadratische vorm (13.3) de gelijknamige eigenschap heeft. Omdat de matrix A symmetrisch is, is er een orthonormale basis $f_1, \ldots, f_n$ van $\mathbb{R}^n$ die bestaat uit eigenvectoren van A, dus $Af_j = \lambda_j\, f_j$ voor zekere $\lambda_j \in \mathbb{R}$. Laat nu $x \in \mathbb{R}^n$ coördinaten $\widetilde{x}_1, \ldots, \widetilde{x}_n$ t.o.v. deze basis hebben, dus $x = \sum_{j=1}^n \widetilde{x}_j\, f_j$. Dan volgt er uit (13.3) dat

$$B(x,x) = \sum_{j=1}^{n} \lambda_j \widetilde{x}_j^2.$$

Laat nu

$$\lambda_{\min} := \min_{j=1,\ldots,n} \lambda_j, \quad \lambda_{\max} := \max_{j=1,\ldots,n} \lambda_j.$$

Dan

$$\lambda_{\min} \sum_{j=1}^{n} \widetilde{x}_j^2 \leq \sum_{j=1}^{n} \lambda_j\, \widetilde{x}_j^2 \leq \lambda_{\max} \sum_{j=1}^{n} \widetilde{x}_j^2,$$

dus

$$\lambda_{\min}\, |x|^2 \leq B(x,x) \leq \lambda_{\max}\, |x|^2. \tag{13.4}$$

Bovendien zijn de ongelijkheden in (13.4) scherp. Want zij $f_{\min}$ eigenvector van A met eigenwaarde $\lambda_{\min}$ en $f_{\max}$ eigenvector van A met eigenwaarde $\lambda_{\max}$. Dan $\lambda_{\min}\, |f_{\min}|^2 = B(f_{\min}, f_{\min})$ en $\lambda_{\max}\, |f_{\max}|^2 = B(f_{\max}, f_{\max})$. We zien dus:

$$\begin{aligned} A \text{ is positief semidefiniet} &\Longleftrightarrow \lambda_{\min} \geq 0, \\ A \text{ is positief definiet} &\Longleftrightarrow \lambda_{\min} > 0, \end{aligned}$$

en analoog voor negatief (semi)definiet.

We kunnen nu Propositie 13.9 en Stelling 13.10 als volgt herformuleren in het geval van een minimum (het geval van een maximum is analoog).

13.13 Stelling. Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f : E \to \mathbb{R}$ een C^2-functie.

(a) Als f een lokaal minimum aanneemt in a dan is $f'(a) = 0$ en de symmetrische matrix $H_f(a)$ is positief semidefiniet (of, equivalent, alle eigenwaarden van $H_f(a)$ zijn ≥ 0).

(b) Als $f'(a) = 0$ en de symmetrische matrix $H_f(a)$ is positief definiet (of, equivalent, alle eigenwaarden van $H_f(a)$ zijn > 0) dan neemt f een sterk lokaal minimum aan in a.

(c) Als de symmetrische matrix $H_f(a)$ indefiniet is (of, equivalent, $H_f(a)$ heeft zowel positieve als negatieve eigenwaarden) dan neemt f geen lokaal extremum aan in a.

13.14 Definitie. We noemen a een *zadelpunt* van f als a een stationair punt van f is en f in a géén lokaal extremum aanneemt.

13.15. Als het geval (c) van Stelling 13.13 zich voordoet en a bovendien een stationair punt is (dus als $f'(a)$=0 en de matrix $H_f(a)$ indefiniet is) dan is a een *zadelpunt* van f. Want er bestaat dan een eigenvector v van $H_f(a)$ met eigenwaarde $\lambda > 0$ en een eigenvector w met eigenwaarde $\mu < 0$. Dan neemt de functie $t \mapsto f(a + tv)$ een sterk lokaal minimum aan in 0 en de functie $t \mapsto f(a+tw)$ een sterk lokaal maximum in 0. In sommige teksten wordt de term zadelpunt alleen gebruikt als de Hesse-matrix in het stationaire punt indefiniet is terwijl determinant ongelijk 0 is!

13.16 Voorbeeld. De drie functies $f(x, y) := x^2 + y^2$, $f(x, y) := -x^2 - y^2$, $f(x, y) := x^2 - y^2$ staan respectievelijk model voor de situaties dat f in $(0, 0)$ een sterk lokaal minimum danwel maximum aanneemt, of dat f een zadelpunt heeft in $(0, 0)$.

13.17. Bij toepassingen van Stelling 13.13(b) bepaalt men in elk stationair punt a van f de Hesse-matrix $H_f(a)$ en dient dan het teken van de eigenwaarden van deze matrix na te gaan. Men kan daartoe bijv. de eigenwaarden uitrekenen door de wortels λ van de algebraïsche vergelijking $\det(A - \lambda I) = 0$ te bepalen. Als $n = 2$ dan is dit slechts een tweedegraads vergelijking. Immers, als $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ dan

$$\det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ b & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - b^2.$$

In het geval $n = 2$ kan men eenvoudiger bepalen of een symmetrische $n \times n$ matrix positief of negatief definiet is, zonder de eigenwaarden echt uit te rekenen. We kunnen schrijven:

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} = O \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} O^{-1},$$

waarbij O een zekere orthogonale matrix is die de basistransformatie uitvoert waardoor A op diagonaalvorm komt. Dan zijn λ en μ de eigenwaarden van A. Het spoor van A (de som van de diagonaalelementen van A, genoteerd $\operatorname{tr} A$) en de determinant van A zijn invariant onder basistransformaties, dus

$$\begin{array}{rcccl} \operatorname{tr} A & = & a + d & = & \lambda + \mu, \\ \det A & = & ad - b^2 & = & \lambda\mu. \end{array}$$

We kunnen nu de volgende gevallen onderscheiden.

1. $\det A > 0$. Dan $ad > 0$ en $\lambda\mu > 0$. Dus $\lambda, \mu > 0$ of $\lambda, \mu < 0$.
   - 1.1. $a > 0$. Dan ook $d > 0$, dus $\lambda + \mu > 0$, dus $\lambda, \mu > 0$, dus A is positief definiet.
   - 1.2. $a < 0$. Dan ook $d < 0$, dus $\lambda + \mu < 0$, dus $\lambda, \mu < 0$, dus A is negatief definiet.
2. $\det A < 0$. Dan $\lambda < 0$, $\mu > 0$ of $\lambda > 0$, $\mu < 0$, dus A is indefiniet.
3. $\det A = 0$. Dan is één eigenwaarde 0 en de andere willekeurig, dus A is wel positief of negatief semidefiniet, maar niet positief of negatief definiet.

Combinatie met Stelling 13.13 geeft de volgende propositie.

13.18 Propositie. Zij $E \subset \mathbb{R}^2$ open, $a \in E$ en $f : E \to \mathbb{R}$ een C^2-functie zo dat $f'(a) = 0$. Als $\det H_f(a) < 0$ dan neemt f geen lokaal extremum aan in a. Als $\det H_f(a) > 0$ dan is $(D_{11})f(a) \neq 0$ en f neemt een sterk lokaal minimum of maximum aan in a al naar gelang $(D_{11}f)(a) > 0$ of < 0.

13.19. Hoe bepaalt men nu de punten waar een concreet gegeven C^2-functie $f : E \to \mathbb{R}$ lokale extrema aanneemt? Voor het gemak nemen we aan dat $E = \mathbb{R}^2$. Als f in een inwendig punt van E een lokaal extremum aanneemt dan zal dat een stationair punt zijn van f. Daarom bepalen we eerst de stationaire punten van f door oplossing van het stelsel van de twee (doorgaans niet-lineaire) vergelijkingen $(D_1 f)(x, y) = 0$, $(D_2 f)(x, y) = 0$. Laten we aannemen dat we de stationaire punten expliciet kunnen vinden. Vervolgens zou men Stelling 13.13 of Propositie 13.17 kunnen toepassen voor het nadere onderzoek van de stationaire punten. We rekenen dus voor ieder stationair punt a de Hesse-matrix $H_f(a)$ uit. Als we geluk hebben dan heeft $H_f(a)$ geen eigenwaarden 0, dus dan kunnen we besluiten tot een van de drie mogelijkheden sterk lokaal minimum, sterk lokaal maximum of zadelpunt. Als $H_f(a)$ een eigenwaarde 0 heeft dan moeten we echter andere methoden gebruiken om tot een beslissing te komen. Het boven beschreven rekenwerk kan vaak beperkt worden als zowel f als E invariant zijn onder een bepaalde symmetrie.

Vaak kan men al sneller beslissen wat de aard is van een stationair punt door de niveauverzameling van f bij een gegeven niveau (0 of misschien een andere waarde c) te tekenen en voor iedere samenhangscomponent van het complement van de niveauverzameling in te tekenen of $f(x) - c$ positief of negatief is. Dit is vooral aan te bevelen als $f(x, y)$ (of $f(x, y) - c$) expliciet te factoriseren is als een product van meer elementaire functies. Van snijpunten van niveaukrommen ziet men dan vaak onmiddellijk in dat het zadelpunten zijn, terwijl van een uniek stationair punt binnen een begrensde samenhangscomponent van het complement van de niveauverzameling onmiddellijk te zeggen is dat er een lokaal maximum of minimum wordt aangenomen. Maak hiertoe gebruik van de stelling dat een continue functie op een compacte verzameling een maximum en minimum moet aannemen.

Als $f : E \to \mathbb{R}$ met E niet open, dan moeten we ook randextrema onderzoeken. Zij E_1 de verzameling van randpunten van E. Als f in een punt a van E_1 een randextremum aanneemt dan zal de functie $f|_{E_1}$ (dus de beperking van f tot E_1) zeker een extremum aannemen in a, maar het omgekeerde hoeft niet te gelden. Maar om de randextrema van f te vinden is het een goede zaak om alle lokale extrema van $f|_{E_1}$ te bepalen en die dan aan nadere analyse te onderwerpen om te bepalen of ze ook lokale extrema voor f zijn. Vaak kan men de rand E_1 met behulp van een of meer C^2-krommen parametriseren. Als γ zo'n kromme is, dan kan men extrema voor $f|_{E_1}$ vinden door Propositie (13.5) toe te passen op de functie $t \mapsto f(\gamma(t))$. Een situatie waarbij men onmiddellijk kan beslissen dat een lokaal extremum (zeg een maximum) voor $f|_{E_1}$ in a ook een lokaal maximum voor f is, is wanneer $f|_{E_1}$ een absoluut maximum aanneemt in a en wanneer ook $f(a) \geq f(b)$ voor alle b waar f een inwendig lokaal maximum aanneemt. Dan zal f een absoluut maximum aannemen in a. Al deze zaken zal men tegenkomen bij de hieronder opgegeven vraagstukken.

Verdere vraagstukken

Opgave 13.2. Bepaal voor elk van de volgende functies $f : \mathbb{R}^2 \to \mathbb{R}$ de punten van $\mathbb{R}^2$ waar f een lokaal maximum of minimum aanneemt. Zeg van elk lokaal extremum of het: (i) sterk of zwak is; (ii) absoluut of relatief is.

(a) $f(x, y) := (3 - x)(3 - y)(3 - x - y)$;

(b) $f(x,y) := x^3 - 3xy^2$;

(c) $f(x,y) := (y-1)\,(x^2-y)^2$;

(d) $f(x,y) := 1 + x + y + x^2 - y^2$;

(e) $f(x,y) := x^4 + y^4 - 4xy$;

(f) $f(x,y) := (y^2-1)\,(x^2-y^2)$.

Opgave 13.3. Zij $E \subset \mathbb{R}^2$. Bepaal voor elk van de volgende functies $f : E \to \mathbb{R}$ de punten van E waar f een lokaal maximum of minimum aanneemt. Zeg van elk lokaal extremum of het: (i) sterk of zwak is; (ii) absoluut of relatief is; (iii) rand- of inwendig extremum is.

(a)	$f(x,y) := 2x^4 - 3x^2y + y^2$,	$E := \{(x,y) \mid 4x^2 + y \leq 4\}$;
(b)	$f(x,y) := x^2 - x + 2y^2$,	$E := \{(x,y) \mid x^2 + y^2 \leq 1\}$;
(c)	$f(x,y) := (x^2 - y^2)\, e^{-x^2-y^2}$,	$E := \{(x,y) \mid x^2 + y^2 \leq 4\}$;
(d)	$f(x,y) := (x^2-1)\,(y^2-1)\,(x^2+y^2-1)$,	$E := \{(x,y) \mid x^2 + y^2 \leq 2\}$;
(e)	$f(x,y) := x^5 - (x^2+x^3)\,y + y^2$,	$E := \{(x,y) \mid -1 \leq x \leq 1,\ 0 \leq y \leq 1\}$;
(f)	$f(x,y) := x^4 + 9y^4$,	$E := \{(x,y) \mid x^2 + 4y^2 \leq 1\}$;
(g)	$f(x,y) := 2x^2 - 3y^2 - 2x + \frac{1}{2}$,	$E := \{(x,y) \mid x^2 + y^2 \leq 1\}$;
(h)	$f(x,y) := x^4 - x^2y^2 - 2x^2 + 2y^2$,	$E := \{(x,y) \mid x^2 + y^2 \leq 4\}$;
(i)	$f(x,y) := x^2y^2 + 2xy^2 + y^4$,	$E := \{(x,y) \mid \lvert x\rvert \leq 2,\ \lvert y\rvert \leq 2\}$.

Index