Syllabus Analyse B1

door T. H. Koornwinder

Universiteit van Amsterdam, Faculteit Wiskunde en Informatica,

Vakgroep Wiskunde, cursus 1995/96

Ter inleiding

Nadat in het vak Analyse tijdens het eerste jaar van de studie uitvoerig analyse op $\mathbb{R}$ is behandeld, i.h.b. functies van één reële veranderlijke, zullen we ons nu in het tweede jaar van de studie 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. De behandeling van deze nieuwe stof zal uitgesmeerd worden over de colleges Analyse B1 en Analyse B2. Veel latere vakken zullen gebruik maken van het hier behandelde materiaal, bijv. 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 het goede begrip van deze syllabus dien je niet alleen vertrouwd te zijn met de stof van de eerstejaars-Analyse, maar ook met die van de vakken Lineaire Algebra en Topologie A. Aan Topologie A ontlenen we het hele begrippenapparaat rond metrische ruimtes, dat hier zal worden toegepast op $\mathbb{R}^n$ en deelverzamelingen van $\mathbb{R}^n$. 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 behandeld worden die betrekking hebben op de theorie van deze syllabus. Enige vertrouwdheid met Maple via het eerstejaars-vak Inleiding Computergebruik wordt hier verondersteld. In het bijzonder kan het je helpen als je Maple wat plaatjes laat tekenen van grafieken van functies in twee veranderlijken.

De opzet van deze syllabus is vergelijkbaar met die van de syllabus Analyse A3. We herhalen nog eens wat hierover in de inleiding van die syllabus werd geschreven.

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

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

Inhoudsopgave

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.

1 Genormeerde vectorruimtes en limieten

In dit hoofdstuk recapituleren we het materiaal uit de syllabus Topologie over genormeerde vectorruimtes en over operatornormen dat we 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.

1.1 Genormeerde vectorruimtes

In syll. Topologie, 4.1 werd de definitie gegeven van een genormeerde reële (of complexe) vectorruimte $(V, \|\,.\,\|)$. In syll. Topologie, 4.2 werd vervolgens bewezen dat een genormeerde vectrorruimte $(V, \|\,.\,\|)$ een metrische ruimte wordt t.o.v. de metriek

$$d(x,y) := \|x-y\| \quad (x,y \in V). \tag{1.1}$$

Lees deze paragrafen nog eens door.

In het vervolg zullen we een gegeven genormeerde vectorruimte $(V, \|\,.\,\|)$ ook automatisch als metrische ruimte opvatten met metriek (1.1). Convergentie van een rij (x_n) in V naar een punt $x \in V$ (cf. syll. Topologie, 1.7) kan dan op de volgende drie equivalente manieren beschreven worden:

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

Ook iedere deelverzameling van V wordt dan een metrische ruimte met de metriek (1.1).

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

$$\big|\, \|x\| - \|y\| \,\big| \le \|x-y\| \quad (x,y \in V). \tag{1.2}$$

1.2 We herhalen de definitie van inproductruimte, zie bijv. syll. Topologie, 4.5, Voorbeeld 1.

Definitie Een (reële) *inproductruimte* is een reële vectorruimte V met een afbeelding $(\,.\,,\,.\,)\colon V \times V \to \mathbb{R}$ (het *inproduct* op V) 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 $(v,v) > 0$.

In syll. Topologie, 4.5, Voorbeeld 1 werd ook bewezen dat een inproductruimte V met inproduct $(\,.\,,\,.\,)$ een genormeerde reële vectorruimte $(V, \|\,.\,\|)$ wordt t.o.v. de norm

$$\|v\| := \sqrt{(v,v)} \quad (v \in V). \tag{1.3}$$

In het vervolg zullen we een gegeven inproductruimte ook automatisch als genormeerde ruimte en als metrische ruimte opvatten overeenkomstig (1.3) en (1.1).

1.3 Voorbeeld (zie syll. Topologie, 4.5, Voorbeeld 1) 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

$$\langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n \quad (x, y \in \mathbb{R}^n). \tag{1.4}$$

(Merk op dat we dit inproduct met schuine haken schrijven.) Volgens (1.3) geeft dit aanleiding tot een norm op $\mathbb{R}^n$ die we altijd met enkele strepen zullen schrijven:

$$|x| := \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2} \quad (x \in \mathbb{R}^n). \tag{1.5}$$

Volgens (1.1) geeft dit weer aanleiding tot een metriek d op $\mathbb{R}^n$:

$$d(x, y) := |x - y| = \sqrt{(x_1 - y_1)^2 + \cdots (x_n - y_n)^2} \quad (x, y \in \mathbb{R}^n). \tag{1.6}$$

$\mathbb{R}^n$ voorzien van inproduct (1.4) en norm (1.5) heet *n-dimensionale Euclidische ruimte* en de metriek (1.6) heet *Euclidische metriek.*

1.4 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 $i = 1, \ldots, m$ (zie syll. Topologie, Stelling 3.9).

1.5 Er bestaan andere normen dan (1.5) op $\mathbb{R}^n$. Bijvoorbeeld maakt de norm (1.5) deel uit van de klasse van zogenaamde *ℓ^p-normen* $(1 \le p \le \infty)$ op $\mathbb{R}^n$ gedefinieerd door

$$\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{1/p} \quad (p \ge 1), \tag{1.7}$$

$$\|x\|_\infty := \max\{|x_1|, \ldots, |x_n|\}. \tag{1.8}$$

De norm in (1.8) is de sup-norm $\|\,.\|_s$ uit syll. Topologie, 4.5, Voorbeeld 2. Bij (1.7) benadrukken we i.h.b. de gevallen $p = 2$ en $p = 1$: $\|x\|_2 = |x|$ (cf. (1.5)) en $\|x\|_1 = |x_1| + \cdots + |x_n|$ (cf. syll. Topologie, Voorbeeld 3).

Opgave

(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

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\, \|x\|_2, \qquad \frac{1}{\sqrt{n}}\, \|x\|_2 \le \|x\|_\infty \le \|x\|_2. \tag{1.9}$$

1.6 In syll. Topologie, 4.9 werd gedefinieerd wanneer twee normen op een vectorruimte V *equivalent* zijn en in syll. Topologie, 4.10 werd vervolgens bewezen dat twee normen $\|\,.\|$ en $\|\,.\|^*$ equivalent zijn desda er $\alpha, \beta > 0$ bestaan zo dat

$$\beta\, \|x\|^* \le \|x\| \le \alpha\, \|x\|^* \quad \text{voor alle } x \in V. \tag{1.10}$$

Een zeer belangrijk resultaat werd gegeven in syll. Topologie, 6.17, Gevolg 1, nl. dat op een eindig-dimensionale reële vectorruimte iedere twee normen equivalent zijn. Gezien (1.10) levert dit:

Propositie Zij $\|\,.\,\|$ een willekeurige norm op $\mathbb{R}^n$ en zij $|\,.\,|$ de norm op $\mathbb{R}^n$ gegeven door (1.5). Dan zijn er $\alpha, \beta > 0$ zo dat

$$\alpha\,|x| \le \|x\| \le \beta\,|x| \quad \text{voor alle } x \in \mathbb{R}^n. \tag{1.11}$$

(Merk op dat dit klopt met (1.9).)

Laten d resp. D de metrieken zijn op $\mathbb{R}^n$ die corresponderen met de normen $|\,.\,|$ resp. $\|\,.\,\|$ uit bovenstaande Propositie. Beide metrieken leiden dan tot dezelfde topologische eigenschappen van $\mathbb{R}^n$ (of een deelverzameling van $\mathbb{R}^n$), zoals o.a. hieronder geformuleerd.

Gevolg
(a) Zij $V \subset \mathbb{R}^n$. Dan is V open in $(\mathbb{R}^n, D)$ desda V open is in $(\mathbb{R}^n, d)$.
(b) Zij (X, d_X) een metrische ruimte en laat f een afbeelding zijn van $\mathbb{R}^n$ naar X. Dan is $f\colon (\mathbb{R}^n, D) \to (X, d_X)$ continu desda $f\colon (\mathbb{R}^n, d) \to (X, d_X)$ continu is.
(c) Zij (X, d_X) een metrische ruimte, laat f een afbeelding zijn van X naar $\mathbb{R}^n$, en schrijf $f(x) = (f_1(x), \ldots, f_n(x))$. Dan zijn equivalent:
  (i) $f\colon (X, d_X) \to (\mathbb{R}^n, D)$ is continu.
  (ii) $f\colon (X, d_X) \to (\mathbb{R}^n, d)$ is continu.
  (iii) Voor $j = 1, \ldots, n$ is $f_j\colon (X, d_X) \to \mathbb{R}$ continu.

1.2 De operatornorm

In syll. Topologie, 4.11–4.14 werden definities en eigenschappen betreffende begrensde lineaire afbeeldingen gegeven. Lees dit nog eens na. Hieronder recapituleren we een paar zaken.

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

$$\begin{aligned} \|A\| &:= \inf\{\alpha \in \mathbb{R} \mid \|Av\| \le \alpha\,\|v\| \quad \forall v \in V\} && (1.12)\\ &= \min\{\alpha \in \mathbb{R} \mid \|Av\| \le \alpha\,\|v\| \quad \forall v \in V\} && (1.13)\\ &= \sup_{\substack{v \in V\\ \|v\| \le 1}} \|Av\|. && (1.14) \end{aligned}$$

We zullen het liefste werken met formule (1.14) voor de operatornorm $\|A\|$. Merk op dat (1.14) 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:

$$\|Av\| \le \|A\|\,\|v\| \quad (v \in V). \tag{1.15}$$

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

1.8 Propositie (zie syll. Topologie, 4.11) Laten V en W genormeerde reële vectorruimtes zijn, V bovendien eindig-dimensionaal. Dan zijn alle lineaire afbeeldingen van V naar W begrensd.

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.

Merk op dat de ruimte $L(\mathbb{R}^n, \mathbb{R}^m)$ een nm-dimensionale genormeerde vectorruimte is. Voor $n, m > 1$ is de door de operatornorm gegeven norm voor deze ruimte zeker niet gelijk aan enige ℓ^p-norm (1.7), (1.8) voor $\mathbb{R}^{mn}$, maar is er wel equivalent mee (cf. Propositie 1.6).

1.9 Opgave (cf. syll. Topologie, 4.15, Opgave 4) 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 $\|AB\| \le \|A\|\,\|B\|$.

1.10 Opgave 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 $\|A\| = |c|$.

1.11 **Opgave** 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 $\|A\| = \max_{j=1,\ldots,n} |\lambda_j|$ als A symmetrisch is met eigenwaarden $\lambda_1, \ldots, \lambda_n$.

1.3 Limieten van functies

In syll. Topologie, 3.1 werd gedefinieerd wanneer een afbeelding tussen twee metrische ruimtes continu is in een bepaald punt. Hiermee verband houdt de definitie van limiet:

1.12 Definitie Laten (X, d) en (Y, d') twee metrische ruimten zijn. Zij $a \in X$ en zij $f\colon X\backslash\{a\} \to Y$. 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:
$$0 < d(x, a) < \delta \quad \Longrightarrow \quad d'(f(x), b) < \varepsilon. \tag{1.16}$$

In bovenstaande Definitie is het toegestaan dat f op X inclusief het punt a gedefinieerd is, 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:
Een afbeelding $f\colon X \to Y$ is continu in a desda $\lim_{x\to a} f(x) = f(a)$.

1.13 **Opgave** Behoud de gegevens van Definitie 1.12. Bewijs dat $\lim_{x\to a} f(x) = b$ waar is voor hoogstens één $b \in Y$ als a een verdichtingspunt is van X, terwijl $\lim_{x\to a} f(x) = b$ waar is voor alle $b \in Y$ als a een geïsoleerd punt is van X.

1.14 Beschouw onder de aannamen van Definitie 1.12 de functie $x \mapsto d'(f(x), b)\colon X\backslash\{a\} \to \mathbb{R}$. Dan geldt:

$$\lim_{x\to a} f(x) = b \quad\Longleftrightarrow\quad \lim_{x\to a} d'(f(x), b) = 0.$$

Dus Definitie 1.12 kan worden teruggebracht tot het geval dat $Y = \mathbb{R}$.

1.15 We bespreken nog een paar speciale gevallen van Definitie 1.12.

1. $Y := \mathbb{R}^m$. Dan wordt (1.16) als volgt:

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

   Merk op dat de volgende drie uitspraken met elkaar equivalent zijn (cf. §1.4):
   (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 wordt (1.16) als volgt:

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

   Als X bovendien een omgeving van a is (m.a.w. a is een inwendig punt van X) dan is er volgens Opgave 1.13 hoogstens één $b \in Y$ zo dat $\lim_{x\to a} f(x) = b$.
3. $X \subset \mathbb{R}$ en $Y := \mathbb{R}$. Dan wordt (1.16) als volgt:

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

   Dit geval is eerder gedefinieerd in syll. Analyse A3, 3.17.

1.16 **Opgave** Zij (X, d) een metrische ruimte en zij $a \in X$. Laat $f, g\colon X\backslash\{a\} \to \mathbb{R}^m$ en $\alpha, \beta \in \mathbb{R}$. Neem aan dat $\lim_{x\to a} f(x) = p$ en $\lim_{x\to a} g(x) = q$ met $p, q \in \mathbb{R}^m$. Bewijs dat $\lim_{x\to a}(\alpha\, f(x) + \beta\, g(x)) = \alpha\, p + \beta\, q$.

1.17 **Opgave** 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:

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

(Hierbij hebben we in de linker limiet $h \in \mathbb{R}^n$ toegevoegd en in de rechter limiet $t \in \mathbb{R}$ om te benadrukken dat de linker limiet genomen wordt voor $h \to 0$ met h in een omgeving van 0 in $\mathbb{R}^n$ en de rechter limiet voor $t \to 0$ met t in een omgeving van 0 in $\mathbb{R}$.

1.18 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

$$f(rx) = r^\alpha f(x) \quad (x \in \mathbb{R}^n\backslash\{0\},\ r > 0). \tag{1.17}$$

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 $\mathbb{R}^n$. Merk op dat een homogene functie van graad α op $\mathbb{R}^n\backslash\{0\}$ geheel bepaald is door haar restrictie tot S^{n-1}. In het bijzonder is een homogene functie van graad 0 constant op elk van de halflijnen $\{rx \mid r > 0\}$ $(x \in S^{n-1})$.

Propositie Zij $f: \mathbb{R}^n\backslash\{0\} \to \mathbb{R}$ 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 $\alpha \leq 0$.

Bewijs S^{n-1} is een gesloten en begrensde deelverzameling van $\mathbb{R}^n$, dus compact (cf. syll. Topologie, Stelling 6.7). De functie f is continu op de compacte verzameling S^{n-1}, dus is daar begrensd (cf. syll. Topologie, Stelling 6.15). Laat $M := \sup_{x\in S^{n-1}} |f(x)|$. Dan $0 < M < \infty$ en $|f(x)| \leq M\,|x|^\alpha$. Veronderstel dat $\alpha > 0$ en zij $\varepsilon > 0$ gegeven. Dan geldt: als $|x| \leq (\varepsilon/M)^{1/\alpha}$ 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 $f(x) \neq 0$. 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 1.17). Veronderstel tenslotte dat $\alpha = 0$. Dan zijn er $x, y \in S^{n-1}$ zo dat $f(x) \neq f(y)$. Omdat $\lim_{r\downarrow 0} f(rz) = f(z)$ $(z = x, y)$, bestaat dus $\lim_{x\to 0} f(x)$ niet (cf. Opgave 1.17). □

1.19 Voorbeeld Laat

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

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

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

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

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

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 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 1.18?

Verdere vraagstukken

V1.1 Ga voor de volgende functies $f\colon \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ 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) $f(x,y) := \dfrac{x^4y}{(x^2+y^2)^2}$

V1.2 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 $f(x,x) := 2x$

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

2.1 Partiële en richtingsafgeleide

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

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \tag{2.1}$$

bestaat en we duiden de limiet (2.1) 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 (2.1) dan liever niet als $\frac{df}{dx}(a)$, wel eventueel als $\frac{df(x)}{dx}\big|_{x=a}$.

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

$$\lim_{h\to 0} \frac{f(a_1,\ldots,a_{j-1},a_j+h,a_{j+1},\ldots,a_n) - f(a_1,\ldots,a_{j-1},a_j,a_{j+1},\ldots,a_n)}{h} \tag{2.2}$$

bestaat en we duiden de limiet (2.2) dan aan met $(D_jf)(a)$. Als de limiet (2.2) bestaat voor alle $a \in E$ dan definieert dit zodoende een nieuwe functie $D_jf\colon x \mapsto (D_jf)(x)\colon E \to \mathbb{C}$, de *partiële afgeleide* van f *naar de $j$de coördinaat.*

We kunnen (2.2) wat compacter schrijven als volgt.

$$(D_jf)(a) := \lim_{h\to 0} \frac{f(a+he_j)-f(a)}{h}\,. \tag{2.3}$$

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

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

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

Als de partiële afgeleide D_jf bestaat op E dan schrijven we deze functie eventueel ook als $x \mapsto \frac{\partial f(x)}{\partial x_j}$, terwijl $(D_jf)(a)$ 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_jf en $\frac{\partial f}{\partial x_j}(a)$ voor $(D_jf)(a)$. (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.)

2.2 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.

Merk op dat in Voorbeeld 1.19 $(D_1 f)(0,0)$ en $(D_2 f)(0,0)$ beide bestaan en gelijk 0 zijn. 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.

2.3 Opgave Zij $f\colon \mathbb{R}^n \to \mathbb{R}$ 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.

2.4 Vervang in het rechterlid van (2.3) de vector e_j door een willekeurige vector $u \in \mathbb{R}^n$. Als de limiet dan bestaat, dan spreken we van de richtingsafgeleide:

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

$$(D_u f)(a) := \lim_{h \to 0} \frac{f(a+hu) - f(a)}{h}, \tag{2.4}$$

mits de limiet in het rechterlid bestaat.

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

$$g(t) := f(a+tu) \quad \text{dan geldt} \quad (D_u f)(a) = g'(0) \tag{2.5}$$

mits een van beide afgeleiden bestaat.

Merk op dat $(D_j f)(a) = (D_{e_j} f)(a)$, waarbij het linker lid gedefinieerd is door (2.3) en het rechter lid door (2.4). De notatie van D met subindex is dus dubbelzinnig, maar doorgaans is uit het verband wel duidelijk of de partiële afgeleide danwel de richtingsafgeleide bedoeld wordt.

Er volgt direct uit bovenstaande definitie dat, 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

$$(D_{\alpha u} f)(a) = \alpha\,(D_u f)(a). \tag{2.6}$$

De richtingsafgeleide in de richting 0 bestaat altijd:

$$(D_0 f)(a) = 0. \tag{2.7}$$

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.

2.5 Voorbeeld Neem f als in Voorbeeld 1.19. Dan geldt dat

$$\frac{f(rx,ry)-f(0,0)}{r}=\frac{xy^2}{x^2+r^2y^4}\xrightarrow{r\to 0}(D_{(x,y)}f)(0,0)=\begin{cases} y^2/x & \text{als } x\neq 0,\\ 0 & \text{als } x=0.\end{cases} \tag{2.8}$$

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.

2.6 Voorbeeld Neem f als in Voorbeeld 2.2. 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 coëfficiënt van t in de Taylorreeks rond 0 van de functie $t\mapsto f(x+tu,y+tv)$ 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 2.1, 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.

Merk op dat in Voorbeeld 2.5 $(D_{(u,v)}f)(0,0)$ 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.

2.2 *Totale afgeleide*

2.7 Zij $I\subset\mathbb{R}$ een open interval en zij $f\colon I\to\mathbb{R}$. 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) $\lim_{h\to 0} h^{-1}\left(f(a+h)-f(a)\right)=c$.

(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

$$f(a+h)=f(a)+c\,h+r(h) \tag{2.9}$$

voor h in een omgeving van 0, of equivalent

$$f(x) = f(a) + c\,(x-a) + r(x-a) \tag{2.10}$$

voor x in een omgeving van a. De eerste twee termen in het rechter lid van (2.10) 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 rechter lid van (2.9). 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 differentieerbaarhied 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 heeft A 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.

2.8 Zij nu E een open deelverzameling van $\mathbb{R}^n$, zij $f\colon E \to \mathbb{R}^m$ en zij $a \in E$. Zij $A\colon \mathbb{R}^n \to \mathbb{R}^m$ lineair. Naar analogie van (2.9) schrijven we

$$f(a+h) = f(a) + Ah + r(h) \quad (h \in \mathbb{R}^n \text{ zo dat } a+h \in E). \tag{2.11}$$

Het linker lid geeft een van h afhankelijke vector in $\mathbb{R}^m$. In het rechter lid 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 §2.7 zullen we f nu differentieerbaar noemen in het punt a als deze restvector van kleinere orde is in h voor $h \to 0$ dan $|h|$:

Definitie Zij E een open deelverzameling van $\mathbb{R}^n$. De afbeelding $f\colon E \to \mathbb{R}^m$ is *differentieerbaar* in zeker punt $a \in E$ met *afgeleide* $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ als $r(h)$ in (2.11) voldoet aan

$$\lim_{h \to 0} \frac{|r(h)|}{|h|} = 0. \tag{2.12}$$

In (2.12) 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 (2.12) is gedefinieerd op een omgeving van 0 in $\mathbb{R}^n$ met 0 er uit weggelaten.

We schrijven (2.11) samen met (2.12) ook wel compacter met behulp van het $o(|h|)$-symbool (spreek uit "kleine o van de norm van h") als volgt:

$$f(a+h) = f(a) + Ah + o(|h|) \quad \text{als } h \to 0 \text{ in } \mathbb{R}^n. \tag{2.13}$$

Per definitie bedoelen we hiermee hetzelfde als met (2.11) en (2.12) gecombineerd. De uitspraak (2.13) 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

$$\lim_{h\to 0} \frac{|f(a+h)-f(a)-Ah|}{|h|} = 0. \tag{2.14}$$

Hiermee is weer equivalent (cf. §1.15, onderdeel 1) dat

$$\lim_{h\to 0} |h|^{-1}\,(f(a+h)-f(a)-Ah) = 0, \tag{2.15}$$

en ook dat

$$\lim_{h\to 0} \frac{f_i(a+h)-f_i(a)-(Ah)_i}{|h|} = 0 \quad (i=1,\ldots,m). \tag{2.16}$$

2.9 Propositie Zij E een open deelverzameling van $\mathbb{R}^n$. Als de afbeelding $f\colon E\to\mathbb{R}^m$ 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 (2.15) 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 Opgave 1.16 toe. We verkrijgen dat $\lim\limits_{h\to 0}|h|^{-1}\,(B-A)h=0$. Dus er volgt met behulp van Opgave 1.17 dat voor elke $x\in\mathbb{R}^n\backslash\{0\}$ geldt:

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

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

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 $f\colon E\to\mathbb{R}^m$ 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 $f'\colon E\to L(\mathbb{R}^n,\mathbb{R}^m)$, die we de *(totale) afgeleide* van f op E noemen.

2.10 Opgave Zij $a\in E\subset\mathbb{R}^n$, E open en $f\colon E\to\mathbb{R}^m$. 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 $i=1,\ldots,m$, $f_i\colon E\to\mathbb{R}$ differentieerbaar is in a met afgeleide $f_i'(a)\colon\mathbb{R}^n\to\mathbb{R}$ gegeven door de $1\times n$ matrix $(A_{i1}\ldots A_{in})$.
Aanwijzing Gebruik de equivalentie van (2.15) met (2.16).

2.11 Opgave Zij $a\in E\subset\mathbb{R}^n$, E open en $f\colon E\to\mathbb{R}^m$. Zij f differentieerbaar in a met afgeleide $f'(a)$. Bewijs het volgende: Er bestaat een $\delta>0$ zo dat, als $|h|<\delta$ dan $a+h\in E$ en

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

Aanwijzing Gebruik (2.11) en (2.12).

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

Bewijs Dit volgt onmiddellijk uit Opgave 2.11. $\square$

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

Opgave 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.

2.13 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 (2.11): de restvector $r(h)$ is identiek 0.

2.3 Verband tussen verschillende types afgeleiden

2.14 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

$$f'(a) = \begin{pmatrix} (D_1 f_1)(a) & \dots & (D_n f_1)(a) \\ \vdots & & \vdots \\ (D_1 f_m)(a) & \dots & (D_n f_m)(a) \end{pmatrix} \tag{2.17}$$

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 (2.16) weten we dat

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

Dan levert Opgave 1.17 dat, voor elke $x \in \mathbb{R}^n \backslash \{0\}$, geldt dat

$$\lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{f_i(a+tx) - f_i(a) - (A(tx))_i}{|tx|} = 0.$$

Kies in het bijzonder $x := e_j$. Dan

$$\lim_{t \to 0} \frac{f_i(a+te_j) - f_i(a) - t\,(Ae_j)_i}{|t|} = 0. \tag{2.18}$$

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 (2.3) bestaat dus de partiële afgeleide $(D_j f_i)(a)$ en is deze gelijk aan A_{ij}. $\square$

Behoud 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:

$$\begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} = \begin{pmatrix} (D_1 f_1)(a) & \dots & (D_n f_1)(a) \\ \vdots & & \vdots \\ (D_1 f_m)(a) & \dots & (D_n f_m)(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}. \tag{2.19}$$

2.15 Als, in Stelling 2.14, $m = 1$, 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

$$f'(a) = ((D_1 f)(a) \dots (D_n f)(a)). \tag{2.20}$$

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:

$$f'(a)\,h = ((D_1 f)(a) \dots (D_n f)(a)) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{j=1}^{n} (D_j f)(a)\,h_j \quad (h \in \mathbb{R}^n). \tag{2.21}$$

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

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

$$(D_u f)(a) = f'(a)\,u = u_1\,(D_1 f)(a) + \dots + u_n\,(D_n f)(a). \tag{2.22}$$

Bewijs Het resultaat is evident als $u = 0$. Laat $u \neq 0$ en concludeer uit (2.14) 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 (2.4) en pas (2.21) toe. $\square$

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 (2.8) dat de functie f uit Voorbeeld 1.19 niet differentieerbaar is in 0.

2.16 Zij $a \in E \subset \mathbb{R}^n$, E open en $f\colon E \to \mathbb{R}$, differentieerbaar in a. In (2.20) schreven we $f'(a)$ als een $1 \times n$ matrix. Deze kunnen we uiteraard ook als (rij)vector in $\mathbb{R}^n$ opvatten. Deze vector heet de *gradiënt* van f in a en hij wordt genoteerd als

$$(\nabla f)(a) := ((D_1 f)(a), \ldots, (D_n f)(a)). \tag{2.23}$$

(Het symbool ∇ wordt uitgesproken als "nabla".) We zien dan dat de richtingsafgeleide (2.22) als inwendig product kan worden geschreven:

$$(D_u f)(a) = \langle u, (\nabla f)(a) \rangle. \tag{2.24}$$

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

2.18 Opgave Zij $f\colon \mathbb{R}^n \to \mathbb{R}$ 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 2.15 en Opgave 2.17).

2.19 Bekijk nu het geval $n = 1$ van Stelling 2.14, dus $a \in E \subset \mathbb{R}$, E open en $f\colon E \to \mathbb{R}^m$ 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 Zij $a \in E \subset \mathbb{R}$, E open en $f\colon E \to \mathbb{R}^m$ differentieerbaar in a. Bewijs dat $f'(a)$, opgevat als element van $\mathbb{R}^m$, ook verkregen kan worden uit

$$f'(a) = \lim_{\substack{h \to 0 \\ h \in \mathbb{R}}} \frac{f(a+h) - f(a)}{h}. \tag{2.25}$$

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 (2.1) voor de afgeleide van een scalaire functie afhangend van één reële variabele.

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

1) *totale afgeleide*: $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$, zie §2.8, 2.9, 2.14.
2) $(m = 1)$ *gradiënt*: $f'(a) = (\nabla f)(a) \in \mathbb{R}^n$ (rijvector), zie §2.15, 2.16.
3) $(n = 1)$ *afgeleide van vectorwaardige functie*: $f'(a) \in \mathbb{R}^m$ (kolomvector), zie §2.19.
4) $(m = 1)$ *richtingsafgeleide*: $(D_u f)(a) \in \mathbb{R}$ $(u \in \mathbb{R}^n)$, zie §2.4, 2.15, 2.16.
5) $(m = 1)$ *partiële afgeleide*: $(D_j f)(a) \in \mathbb{R}$ $(j = 1, \ldots, n)$, zie §2.1, 2.14, 2.15, 2.16.

2.21 Opgave Zij $a \in E \subset \mathbb{R}^n$ open. Zij $f, g\colon E \to \mathbb{R}^m$ 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

$$(fg)'(a) = g(a)\,f'(a) + f(a)\,g'(a), \quad (f/g)'(a) = \frac{g(a)\,f'(a) - f(a)\,g'(a)}{g(a)^2}\,. \tag{2.26}$$

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

2.22 Propositie Zij $E \subset \mathbb{R}^n$ open en convex (cf. syll. Topologie, 7.17). Zij $f\colon E \to \mathbb{R}^m$ 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 (2.22) 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 (2.5). Dus de functie $t \mapsto f(x + t(y - x))$ heeft afgeleide 0 op I, dus deze functie is constant op I. $\square$

Gevolg Zij $E \subset \mathbb{R}^n$ open en convex en zij $g\colon E \to L(\mathbb{R}^n, \mathbb{R}^m)$. Als er een differentieerbare afbeelding $f\colon E \to \mathbb{R}^m$ 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 Zij $f\colon \mathbb{R}^n \to \mathbb{R}^m$ 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 $f(x) = Ax + b$ $(x \in \mathbb{R}^n)$.

2.23 **Opgave** Zij $E \subset \mathbb{R}^n$ open.

(a) Bewijs dat E samenhangend is desda 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 (cf. syll. Topologie, 7.1 en 7.16).

(b) Zij E samenhangend. Zij $f\colon E \to \mathbb{R}^m$ 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 (a) en het bewijs van Propositie 2.22.

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

$$\sum_{j=1}^{n} x_j \,(D_j f)(x) = \alpha\, f(x) \quad (x \in \mathbb{R}^n \backslash \{0\}). \tag{2.27}$$

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

2.25 Definitie Zij $E \subset \mathbb{R}^n$ open. We zeggen dat een afbeelding $f\colon E \to \mathbb{R}^m$ *continu differentieerbaar* is op E als f differentieerbaar is op E en de afgeleide $f'\colon E \to L(\mathbb{R}^n, \mathbb{R}^m)$ 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 $f \in C^1(E)$.

De ruimte $L(\mathbb{R}^n, \mathbb{R}^m)$ kan geïdentificeerd 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 1.6(c) en (2.17) zien we dan het volgende in.

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

2.26 De volgende stelling geeft een beduidende en zeer bevredigende verscherping van Propositie 2.25.

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

Bewijs Gezien Propositie 2.25 en Opgave 2.10 is het voldoende om voor een afbeelding $f\colon E \to \mathbb{R}$ te bewijzen dat, als de partiële afgeleiden $D_j f$ bestaan op E en continu zijn, dan f differentieerbaar is op E.
Zij $a \in E$ en neem $r > 0$ zo dat $x \in E$ als $|x - a| < r$. Gezien (2.14) en Stelling 2.14 zal de differentieerbaarheid van f in a volgen als we kunnen bewijzen dat

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - (h_1\,(D_1 f)(a) + \cdots + h_n\,(D_n f)(a))|}{|h|} = 0. \tag{2.28}$$

We moeten dus $f(a+h) - f(a) - (h_1\,(D_1 f)(a) + \cdots + h_n\,(D_n f)(a))$ in absolute waarde afschatten als functie van h voor $|h| \to 0$. Neem $h \in \mathbb{R}^n$ met $|h| < r$ en schrijf $v_k := h_1 e_1 + \cdots + h_k e_k$ $(k = 0, 1, \ldots, n)$. Dan

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

Pas nu voor elke $k \in \{1, \ldots, n\}$ de middelwaardestelling (zie syll. Analyse A) toe op de functie $g(t) := f(a + v_{k-1} + t e_k)$. Er zijn dus getallen $\theta_k \in (0,1)$ $(k = 1, \ldots, n)$ zo dat

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

Dus

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

Dus

$$\frac{|f(a+h) - f(a) - (h_1\,(D_1 f)(a) + \cdots + h_n\,(D_n f)(a))|}{|h|} \\ \leq \sum_{k=1}^{n} \big|(D_k f)(a + v_{k-1} + \theta_k h_k e_k) - (D_k f)(a)\big|. \qquad (2.29)$$

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 $|(D_k f)(x) - (D_k f)(a)| < \varepsilon/n$ als $|x - a| < \delta$. Neem $|h| < \delta$. Dan $|v_{k-1} + \theta_k h_k e_k| < \delta$, dus het rechter lid van (2.29) is kleiner dan $\sum_{k=1}^{n}(\varepsilon/n) = \varepsilon$. Dit laat zien dat de limiet (2.28) geldt. $\square$

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 2.2.

2.27 Opgave Zij $f\colon \mathbb{R} \to \mathbb{R}$ 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 $F\colon \mathbb{R}^2 \to \mathbb{R}$ een C^1-functie is.
Aanwijzing Bewijs eerst dat $(D_1 F)(a, a) = \frac{1}{2}\, f''(a) = (D_2 F)(a, a)$.
Bereken nu $(D_1 F)(x, y)$ 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 $(D_1 F)(x, y) = \frac{1}{2}\, f''(x + \theta(y - x))$. Laat nu zien dat $D_1 F$ overal continu is. Een soortgelijk argument geldt voor $D_2 F$.

2.28 Opgave Zij $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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

V2.1 Bepaal de partiële afgeleiden van $f\colon E \to \mathbb{R}$ 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 $f(x, y, z) := x^y\, y^z\, z^x$

V2.2 Voor $f\colon \mathbb{R}^2 \to \mathbb{R}$ is gegeven:
$$f(0,1) = -\frac{\pi}{4}, \quad (D_1 f)(x,y) = y\,e^x \quad \text{en} \quad (D_2 f)(x,y) = e^x - \frac{1}{1+y^2}\,.$$
Bepaal f.

V2.3 Ga van de volgende functies $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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) $f(x,y) := \log\left(1 + \sqrt{|xy|}\right)$

V2.4 Zij $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $(0,0)$.

V2.5 Zij $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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) $f(x,y) := \dfrac{x^2y}{x^2+y^4}$
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?

V2.6 Toon aan dat de volgende functies $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $f(y^2,y) := 2y^2$.

V2.7 Zij $f\colon \mathbb{R}^n \to \mathbb{R}$.
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 $f'(0)$.

3 De kettingregel

3.1 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.

3.1 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

$$F'(a) = g'(f(a))\, f'(a). \tag{3.1}$$

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

$$\lim_{h \to 0} \frac{|F(a+h) - F(a) - g'(f(a))\, f'(a)\, h|}{|h|} = 0. \tag{3.2}$$

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:

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

Dus

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

Dus de limiet (3.2) zal volgen als we de twee volgende limieten kunnen bewijzen:

$$\lim_{h \to 0} \frac{|g(f(a+h)) - g(f(a)) - g'(f(a))\, (f(a+h) - f(a))|}{|h|} = 0, \tag{3.3}$$

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - f'(a)\, h|}{|h|} = 0. \tag{3.4}$$

De limiet (3.4) zegt juist dat f differentieerbaar is in a met afgeleide $f'(a)$, dus deze limiet is geldig. We zullen de limiet (3.3) bewijzen door te gebruiken dat g differentieerbaar is in $f(a)$ met afgeleide $g'(f(a))$, dus dat

$$|g(f(a) + k) - g(f(a)) - g'(f(a))\, k| = o(|k|) \quad \text{als } k \to 0 \text{ in } \mathbb{R}^m. \tag{3.5}$$

Laat nu $\varepsilon > 0$. Wegens (3.5) is er een $\delta > 0$ zo dat, als $|k| < \delta$ dan $f(a) + k \in U$ en

$$|g(f(a) + k) - g(f(a)) - g'(f(a))\,k| < \varepsilon\,|k|. \tag{3.6}$$

Wegens Opgave 2.11 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

$$|f(a+h) - f(a)| \le (\|f'(a)\| + 1)\,|h|. \tag{3.7}$$

Vanwege de continuïteit van f in a (cf. Propositie 2.12) kunnen we, door γ eventueel nog kleiner te nemen, bewerkstelligen dat bovendien $|f(a+h) - f(a)| < \delta$ als $|h| < \gamma$. Substitueer nu $k := f(a+h) - f(a)$ in (3.6) en combineer de resulterende ongelijkheid met ongelijkheid (3.7). 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 willkeurig genomen was, impliceert dit de limiet (3.3). $\square$

Formule (3.1) staat bekend als de *kettingregel* en ziet er formeel hetzelfde uit als de kettingregel voor differentieerbare functies van één variabele. Echter, het rechter lid van (3.1) moeten we lezen als de vermenigvuldiging (compositie) van twee lineaire afbeeldingen.

Formule (3.1) ziet er, in matrixvorm uitgeschreven, als volgt uit.

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

$$= \begin{pmatrix} (D_1g_1)(f(a)) & \dots & (D_mg_1)(f(a)) \\ \vdots & & \vdots \\ (D_1g_l)(f(a)) & \dots & (D_mg_l)(f(a)) \end{pmatrix} \begin{pmatrix} (D_1f_1)(a) & \dots & (D_nf_1)(a) \\ \vdots & & \vdots \\ (D_1f_m)(a) & \dots & (D_nf_m)(a) \end{pmatrix}. \tag{3.8}$$

Per matrixcomponent wordt dit

$$(D_jF_i)(a) = \sum_{k=1}^{m} (D_jf_k)(a)\,(D_kg_i)(f(a)), \quad i = 1, \dots, l, \quad j = 1, \dots, n. \tag{3.9}$$

Hier hebben we de factoren in de sommatie rechts in zo'n volgorde geschreven dat de formule gemakkelijker te onthouden is.

3.2 In het geval $l = 1$ van Stelling 3.1 zijn de functies g en $F := g \circ f$ reëelwaardig en we hebben de formules

$$F(x_1, \dots, x_n) = g(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) \quad (x \in E), \tag{3.10}$$

$$(D_jF)(a) = \sum_{k=1}^{m} (D_jf_k)(a)\,(D_kg)(f(a)), \quad j = 1, \dots, n. \tag{3.11}$$

Als f differentieerbaar is op E en g differentieerbaar op U, dan kunnen we (3.11) ook schrijven in de vorm

$$\frac{\partial}{\partial x_j} g(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) = \sum_{k=1}^{m} \frac{\partial f_k(x_1, \dots, x_n)}{\partial x_j}\, \frac{\partial g(y_1, \dots, y_m)}{\partial y_k}\Big|_{y=f(x)}. \tag{3.12}$$

3.3 In het geval $l = 1$ en $n = 1$ van Stelling 3.1 zijn de functies g en F reëelwaardig en hangen de functies f en F van slechts één reële variabele af. Dan nemen de formules (3.10), (3.11) en (3.12) de volgende vorm aan.

$$F(x) = g(f_1 x), \ldots, f_m(x)) \quad (x \in E), \tag{3.13}$$

$$F'(a) = \sum_{k=1}^{m} f_k'(a)\,(D_k g)(f(a)), \tag{3.14}$$

$$\frac{d}{dx} g(f_1(x), \ldots, f_m(x)) = \sum_{k=1}^{m} \frac{df_k(x)}{dx}\, \frac{\partial g(y_1, \ldots, y_m)}{\partial y_k}\bigg|_{y=f(x)}. \tag{3.15}$$

Formule (3.14) 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 (3.1). We formuleren Stelling 3.1 nog eens in het geval $l = 1$ en $n = 1$.

Stelling Zij $a \in E \subset \mathbb{R}$, E open, $f\colon E \to \mathbb{R}^m$, $f(E) \subset U \subset \mathbb{R}^m$, U open, $g\colon U \to \mathbb{R}$, $F = g \circ f\colon E \to \mathbb{R}$. Als f differentieerbaar is in a en g differentieerbaar in $f(a)$ dan is F differentieerbaar in a en geldt formule (3.14).

3.4 Opgave 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 (3.14), waarbij $m = 2$ en $g(y_1, y_2) := y_1\, y_2$. Voer een soortgelijke opdracht uit met betrekking tot de afgeleide van $f_1(x)/f_2(x)$.

3.5 Opgave Laten n, m, l, a, E, U, f en g zijn zoals in Stelling 3.1. 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 $f'(a)\colon \mathbb{R}^n \to \mathbb{R}^n$ inverteerbaar is en dat

$$g'(f(a)) = (f'(a))^{-1}. \tag{3.16}$$

3.2 *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.

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

$$\phi \circ \psi = \text{id op } E_2 \quad \text{en} \quad \psi \circ \phi = \text{id op } E_1. \tag{3.17}$$

(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 $F\colon E_1 \to \mathbb{R}$ en $f\colon E_2 \to \mathbb{R}$ C^1-afbeeldingen zijn zo dat

$$F(x) = f(\phi(x)) \quad (x \in E_1), \quad \text{of equivalent} \quad f(u) = F(\psi(u)) \quad (u \in E_2). \tag{3.18}$$

Herschrijving van (3.11) 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

$$(D_jF)(\psi(u)) = \sum_{k=1}^{n} (D_j\phi_k)(\psi(u))\,(D_kf)(u) \quad (u \in E_2), \tag{3.19}$$
$$(D_kf)(\phi(x)) = \sum_{j=1}^{n} (D_k\psi_j)(\phi(x))\,(D_jF)(x) \quad (x \in E_1). \tag{3.20}$$

Formule (3.19) 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 (3.20) doet dit in omgekeerde richting. Een andere manier van schrijven van (3.19), (3.20) is als volgt.

$$\left.\frac{\partial F(x)}{\partial x_j}\right|_{x=\psi(u)} = \sum_{k=1}^{n} \left.\frac{\partial \phi_k(x)}{\partial x_j}\right|_{x=\psi(u)} \frac{\partial f(u)}{\partial u_k}\,, \tag{3.21}$$
$$\left.\frac{\partial f(u)}{\partial u_k}\right|_{u=\phi(x)} = \sum_{j=1}^{n} \left.\frac{\partial \psi_j(u)}{\partial u_k}\right|_{u=\phi(x)} \frac{\partial F(x)}{\partial x_j}\,. \tag{3.22}$$

We laten nu in (3.21) en (3.22) 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 §2.1, is afgeraden. Soms is deze schrijfwijze in de praktijk echter toch wel handig.

$$\left.\frac{\partial}{\partial x_j}\right|_{x=\psi(u)} = \sum_{k=1}^{n} \left.\frac{\partial \phi_k(x)}{\partial x_j}\right|_{x=\psi(u)} \frac{\partial}{\partial u_k}\,, \tag{3.23}$$
$$\left.\frac{\partial}{\partial u_k}\right|_{u=\phi(x)} = \sum_{j=1}^{n} \left.\frac{\partial \psi_j(u)}{\partial u_k}\right|_{u=\phi(x)} \frac{\partial}{\partial x_j}\,. \tag{3.24}$$

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 (3.24) aanleiding tot de twee formules

$$\frac{\partial}{\partial u} = \left.\frac{\partial x(u,v)}{\partial u}\right|_{\substack{u=u(x,y)\\ v=v(x,y)}} \frac{\partial}{\partial x} + \left.\frac{\partial y(u,v)}{\partial u}\right|_{\substack{u=u(x,y)\\ v=v(x,y)}} \frac{\partial}{\partial y}\,, \tag{3.25}$$
$$\frac{\partial}{\partial v} = \left.\frac{\partial x(u,v)}{\partial v}\right|_{\substack{u=u(x,y)\\ v=v(x,y)}} \frac{\partial}{\partial x} + \left.\frac{\partial y(u,v)}{\partial v}\right|_{\substack{u=u(x,y)\\ v=v(x,y)}} \frac{\partial}{\partial y}\,. \tag{3.26}$$

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 linker lid 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

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

3.7 Voorbeeld Bekijk nu (3.25), (3.26) met (u,v) vervangen door (r,θ). Hiermee beschrijven we de overgang van poolcoördinaten (r,θ) naar Cartesische coördinaten (x,y):

$$\psi\colon (r,\theta) \mapsto (x,y) = (x(r,\theta), y(r,\theta)) := (r\cos\theta, r\sin\theta). \tag{3.27}$$

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

$$E_1 := \mathbb{R}^2 \backslash \{(x,0) \mid x \le 0\}, \quad E_2 := (0,\infty) \times (-\pi,\pi). \tag{3.28}$$

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 (3.27) 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 bijv. op het rechter halfvlak 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 (3.26) dat

$$\frac{\partial}{\partial r} = \frac{x}{\sqrt{x^2+y^2}}\,\frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2+y^2}}\,\frac{\partial}{\partial y}, \tag{3.29}$$

$$\frac{\partial}{\partial\theta} = -y\,\frac{\partial}{\partial x} + x\,\frac{\partial}{\partial y}. \tag{3.30}$$

Als we E_2 verticaal opschuiven en E_1 dienovereenkomstig veranderen dan blijven formules (3.29), (3.30) ongewijzigd (ga na). We kunnen de kunstmatige snede in E_1 elimineren door $\psi\colon E_2 \to E_1$ te beschouwen met

$$E_1 := \mathbb{R}^2 \backslash \{(0,0)\} \quad E_2 := (0,\infty) \times \mathbb{R}, \tag{3.31}$$

waarbij ψ nog steeds gedefinieerd is door (3.27). 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 bijv. een horizontale strook met wijdte 2π) die door ψ diffeomorf wordt afgebeeld.

De formule

$$f(r,\theta) = F(r\cos\theta, r\sin\theta) \tag{3.32}$$

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 (3.29), (3.30) blijven nog steeds geldig. De stilzwijgende veronderstelling is dat de linker leden van (3.29), (3.30) werken op $f(r,\theta)$ en de rechter leden op $F(x,y)$, waarbij f en F met elkaar verband houden volgens (3.32). 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 Laten F en f C^1-functies zijn op de door (3.31) gegeven verzamelingen E_1 resp. E_2 zo dat (3.32) geldt.

(a) Bewijs met behulp van (3.29) en (2.27) 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 (3.30) dat G eveneens homogeen van graad α is.

3.8 Laten (r,θ) en (x,y) met elkaar verbonden zijn door (3.27). 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:

$$\frac{\partial}{\partial x} = \cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial\theta}, \tag{3.33}$$

$$\frac{\partial}{\partial y} = \sin\theta\,\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\,\frac{\partial}{\partial\theta}. \tag{3.34}$$

Bij de eerste strategie gaan we uit van (3.25), (3.26) 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 (3.29), (3.30) 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 (3.33), (3.34).

Opgave Werk de details van beide geschetste methodes uit om (3.33), (3.34) af te leiden.

Verdere vraagstukken

V3.1 Definieer
$f\colon \mathbb{R} \to \mathbb{R}^2$ door $f(x) := (x^2, x-3)$ en
$g\colon \mathbb{R}^2 \to \mathbb{R}$ 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.

V3.2 Definieer
$f\colon \mathbb{R}^2 \to \mathbb{R}^3$ door $f(x,y) = (x+y, y^2, 3y+1)$ en
$g\colon \mathbb{R}^2 \to \mathbb{R}^2$ door $g(x,y) = (x^2+y^2, xy)$.
Berkeken $(f \circ g)'$ zowel rechtstreeks als m.b.v. de kettingregel.

V3.3 Laat $f\colon \mathbb{R}^2 \to \mathbb{R}$ een differentieerbare functie zijn. Zij $F\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $(D_2f)(x+y, x-y)$.

V3.4 Zij $f\colon \mathbb{R}^2 \to \mathbb{R}$ een differentieerbare functie. Definieer $F\colon (0,\infty) \times \mathbb{R} \to \mathbb{R}$ door $F(r,\theta) := f(r\cos\theta, r\sin\theta)$ (overgaan op poolcoördinaten). Bewijs dat

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

4 Meetkundige interpretaties van afgeleiden

4.1 Krommen

4.1 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 $\gamma\colon I \to \mathbb{R}^n$. De kromme γ heet *continu differentieerbaar* of C^1 als de afbeelding $\gamma\colon I \to \mathbb{R}^n$ een C^1-afbeelding is. (Als het interval I een begin- of eindpunt bevat, dan moeten we daar linker of rechter afgeleide 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 $\gamma\colon t \mapsto (\cos t, \sin t)\colon \mathbb{R} \to \mathbb{R}^2$ 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 $\gamma\colon I \to \mathbb{R}^n$ 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 $\gamma\colon I_1 \to \mathbb{R}^n$ en $\delta\colon I_2 \to \mathbb{R}^n$ als *equivalent* als er een bijectieve continue monotoon stijgende afbeelding $\phi\colon I_1 \to I_2$ is (dus ϕ heeft dan ook een continue inverse) zo dat

$$\gamma(t) = \delta(\phi(t)) \quad (t \in I_1). \tag{4.1}$$

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 $t \mapsto (\cos(t + t^3), \sin(t + t^3))\colon \mathbb{R} \to \mathbb{R}^2$. 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 $\gamma\colon I \to \mathbb{R}^n$ *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 $\gamma\colon I \to \mathbb{R}^n$ een C^1-kromme. Dan (cf. §2.8) geldt, voor $a \in I$, dat

$$\gamma(t) = \gamma(a) + (t - a)\,\gamma'(a) + \rho(t) \quad (t \in I), \tag{4.2}$$

waarbij $\rho\colon I \to \mathbb{R}^n$ voldoet aan $\lim_{t\to a} |t - a|^{-1}\,|\rho(t)| = 0$. Volgens §2.19 kunnen we de afgeleide $\gamma'(a)$ opvatten als vector in $\mathbb{R}^n$. Merk op uit (4.2) 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.

4.2 Opgave Laten γ en δ C^1-krommen in $\mathbb{R}^n$ zijn die in elkaar overgaan door een C^1-herparametrisering zoals in (4.1). 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.

4.3 Definitie Zij γ een C^1-kromme $\gamma\colon [a,b] \to \mathbb{R}^n$ 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. syll. Analyse A3, Definitie 7.1) een getal

$$\Lambda(P,\gamma) := \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|. \tag{4.3}$$

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

$$\Lambda(\gamma) := \sup_{P \text{ partitie van } [a,b]} \Lambda(P,\gamma). \tag{4.4}$$

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

Meetkundig gezien stelt $\Lambda(P,\gamma)$ de lengte voor van 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 (4.4) 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.

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

$$|v_1 + \cdots + v_k| \leq |v_1| + \cdots + |v_k|. \tag{4.5}$$

In zekere zin zijn Riemann-integralen op te vatten als continue sommen. Daarom zal het niet verbazen dat er een analogon van (4.5) geldt voor vectorwaardige functies:

Propositie Zij $f\colon [a,b] \to \mathbb{R}^n$ 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

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx. \tag{4.6}$$

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

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

Dus, met behulp van syll. Analyse A3, Gevolg 7.16 volgt er dat

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

Als $|y| \neq 0$ dan mogen we door y delen en verkrijgen we $|y| \leq \int_a^b |f(x)|\,dx$, i.e., (4.6). Als $|y| = 0$ dan is (4.6) triviaal. $\square$

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

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\,dt = \int_a^b \sqrt{\gamma_1'(t)^2 + \cdots + \gamma_n'(t)^2}\,dt. \tag{4.7}$$

Bewijs Zij $P = P(x_0, \ldots, x_n)$ een partitie van $[a,b]$. Dan geldt er met behulp van Propositie 4.4 dat

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

Dus, met behulp van (4.3),

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

Dan geeft (4.4) dat

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

waarmee we al hebben laten zien dat γ rectificeerbaar is. Formule (4.7) zal nu volgen als we ook de omgekeerde ongelijkheid kunnen bewijzen.

Zij $\varepsilon > 0$. De afbeelding $\gamma' \colon [a, b] \to \mathbb{R}^n$ is continu op een gesloten en begrensd (dus compact) interval en daarom uniform continu (specialiseer syll. Topologie, 4.6 en 6.19 of ga uit van het geval $n = 1$ dat bewezen werd in syll. Analyse A3, Stelling 3.38). 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

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

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

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

Opgave Zij γ als in Stelling 4.5, zij $\delta \colon [c, d] \to \mathbb{R}^n$ ook een C^1-kromme en laten γ en δ C^1-equivalent zijn d.m.v. $\phi \colon [a, b] \to [c, d]$ als in (4.1). Bewijs dat $\Lambda(\gamma) = \Lambda(\delta)$, m.a.w., C^1-equivalente krommen hebben dezelfde lengte.

4.6 We zullen nu het verband leggen tussen de analytisch ingevoerde functies cos en sin enerzijds, en de meetkundige invoering van deze functies anderzijds. (Dit was al beloofd in syll. Analyse A3, §6.23.) Het is hiertoe voldoende om aan te tonen dat de kromme (cirkelboog) $\gamma \colon t \mapsto (\cos t, \sin t) \colon [0, \theta] \to \mathbb{R}^2$ lengte θ heeft. Dit volgt nu echter onmiddellijk uit formule (4.7). 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, warbij π analytisch is ingevoerd als in syll. Analyse A3, Definitie 6.26.

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

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 krome γ_0 niet rectificeerbaar is als $\phi(t) := t \sin \dfrac{1}{t}$.

4.8 Er is een evidente fysische interpretatie van (4.7). De integrand $|\gamma'(t)|$ is de grootte van de snelheid ten tijde t. De totaal afgelegde weglengte is 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 (4.7) 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.

4.2 Meetkundige interpretatie van de gradiënt

4.9 We geven nu een meetkundige interpretatie van (2.24), 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 $f\colon E \to \mathbb{R}$, differentieerbaar in a. Nu zijn u en $(\nabla f)(a)$ beide vectoren in $\mathbb{R}^n$. Neem aan dat de gradiënt $(\nabla f)(a) \ne 0$ 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 (2.24) dat

$$(D_u f)(a) = |(\nabla f)(a)| \cos\theta \quad \text{en} \quad -|(\nabla f)(a)| \le (D_u f)(a) \le |(\nabla f)(a)| \quad (|u| = 1). \qquad (4.8)$$

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 formule (2.4) 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 (4.8) 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 Zij $a \in E \subset \mathbb{R}^2$, E open, $f\colon E \to \mathbb{R}$ differentieerbaar in a. Teken de grafiek van de functie $\theta \mapsto (D_{(\cos\theta, \sin\theta)} f)(a)\colon \mathbb{R} \to \mathbb{R}$. Hoe kunnen de richting en de grootte van de vector $(\nabla f)(a)$ uit deze grafiek worden afgeleid?

4.10 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 $f\colon E \to \mathbb{R}$. 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

$$l\colon x \mapsto f(a) + f'(a)\,(x - a) = f(a) + \sum_{j=1}^{n} (D_j f)(a)\,(x_j - a_j).$$

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*.

4.11 Zij $E \subset \mathbb{R}^n$ open en $f\colon E \to \mathbb{R}$. 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 $f\colon \mathbb{R}^2 \to \mathbb{R}$ gegeven door

$$f(x, y) := x^2 + y^2 \quad ((x, y) \in \mathbb{R}^2). \tag{4.9}$$

De niveauverzamelingen zijn

$$f^{-1}(c) = \{(x, y) \mid x^2 + y^2 = c\} \quad (c \geq 0), \tag{4.10}$$

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 $\gamma\colon I \to E$ een *niveaukromme* van een functie $f\colon E \to \mathbb{R}$ (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 $f\colon E \to \mathbb{R}$ 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 $n = 2$.

4.12 Zij nu $I \subset \mathbb{R}$ een open interval, $\gamma\colon I \to \mathbb{R}^n$ een C^1-kromme, $E \subset \mathbb{R}^n$ open zo dat $\gamma(I) \subset E$, en $f\colon E \to \mathbb{R}$ een C^1-functie. Dan volgt uit het speciale geval van de kettingregel (Stelling 3.3 samen met formules (3.13)–(3.15)) dat

$$(f \circ \gamma)'(t) = \sum_{k=1}^{n} \gamma_k'(t)\,(D_k f)(\gamma(t)) \quad (t \in I).$$

Gezien formule (2.23) voor de gradiënt kunnen we dit herschrijven als het inwendig product

$$(f \circ \gamma)'(t) = \langle \gamma'(t), (\nabla f)(\gamma(t)) \rangle \quad (t \in I). \tag{4.11}$$

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 en teken gekromde lijnen in E die overal loodrecht op dit vectorveld 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 "niveauverzamelingsdichtheid" bij a.

Voorbeeld Bekijk de functie f gegeven door (4.9). De niveauverzamelingen (4.10) ($c = r^2$) zijn precies de beeldverzamelingen van de krommen $\gamma_r\colon t \mapsto (r\cos t, r\sin t)\colon \mathbb{R} \to \mathbb{R}^2$ $(r \geq 0)$. 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.

Opgave 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.

4.13 We geven nog een fysisch geografische toepassing van het begrip gradiënt van 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 (4.11) en (2.24) kunnen we deze helling dus schrijven als

$$(f \circ \gamma)'(t) = \langle \gamma'(t), (\nabla f)(\gamma(t)) \rangle = (D_{\gamma'(t)} f)(\gamma(t)) \quad (t \in I), \tag{4.12}$$

waarbij in het rechterlid een richtingsafgeleide staat. We kunnen de meetkundige beschouwingen van §4.9 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 Zij f de hoogtefunctie als boven. Teken de hoogtelijnen en het gradiëntvectorveld in de omgeving van een bergpas.

Verdere vraagstukken

V4.1 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 $\gamma(t) = (a \cos t, a \sin t, bt)$.

V4.2 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

$$(x_2 - x_1) \sqrt{1 + b^{-2}}.$$

V4.3 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 $n = 3$.

V4.4 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 $f(x) := \langle x, x \rangle \, x$.

5 Middelwaardestelling; hogere partiële afgeleiden

5.1 Een generalisatie van de middelwaardestelling

5.1 Zij $f\colon I \to \mathbb{R}$ een differentieerbare functie op een open interval I en zij $a, b \in I$ met $a < b$. Dan zegt de middelwaardestelling (cf. syll. Analyse A) dat

$$f(b) - f(a) = f'(x)\,(b-a) \quad \text{voor zekere } x \in (a,b). \tag{5.1}$$

Een direct gevolg van (5.1) is de ongelijkheid

$$|f(b) - f(a)| \le |f'(x)|\,(b-a) \quad \text{voor zekere } x \in (a,b). \tag{5.2}$$

Dus als voor zekere $M \ge 0$ geldt dat $|f'(x)| \le M$ op I dan volgt er dat

$$|f(b) - f(a)| \le M\,|b-a| \quad (a, b \in I). \tag{5.3}$$

Dit laatste resultaat was al gegeven in syll. Analyse A3, Prop. 3.36.

Hoe staat het met analoga van deze resultaten voor differentieerbare afbeeldingen $f\colon E \to \mathbb{R}^m$, 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 (5.2), (5.3) met passende interpretatie geldig, maar gelijkheid (5.1) niet. Eigenschap (5.1) 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 (5.1) misgaan. Zie hiertoe het volgende tegenvoorbeeld.

5.2 Voorbeeld Definieer de differentieerbare functie $f\colon \mathbb{R} \to \mathbb{C}$ 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 $|f'(x)| = 1 \neq 0$ voor elke $x \in \mathbb{R}$. Dus (5.1) kan niet gelden voor $(a,b) := (0, 2\pi)$.

5.3 Propositie Zij $E \subset \mathbb{R}^n$ open. Zij $f\colon E \to \mathbb{R}$ 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

$$f(b) - f(a) = f'(x)\,(b-a) = \sum_{j=1}^{n} (D_j f)(x)\,(b_j - a_j) \quad \text{voor zekere } x \in (a,b). \tag{5.4}$$

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 (5.1) geeft dat $g(1) - g(0) = g'(t)$ voor zekere $t \in (0,1)$. □

5.4 Propositie Zij $E \subset \mathbb{R}^n$ open. Zij $f\colon E \to \mathbb{R}^m$ differentieerbaar op E. Zij $a, b \in E$ zo dat het verbindende lijnstuk $[a, b]$ geheel in E ligt. Dan

$$|f(b) - f(a)| \le \|f'(x)\|\,|b - a| \quad \text{voor zekere } x \in (a, b). \tag{5.5}$$

Alvorens het bewijs te geven maken we een paar opmerkingen. In (5.5) 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 (5.5) volstaan met de vectornorm $|f'(x)|$. Het speciale geval $n = 1$ van onderstaand bewijs is wellicht gemakkelijker. Onder de aannamen van Propositie 5.4 zullen we eerst bewijzen dat

$$|f(b) - f(a)|^2 = \langle f'(x)\,(b - a), f(b) - f(a)\rangle \quad \text{voor zekere } x \in (a, b). \tag{5.6}$$

Uitspraak (5.6) 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 (5.4) en (5.1).

Bewijs van de Propositie Eerst bewijzen we (5.6). Schrijf

$$\phi(t) := \langle f(a+t(b-a)), f(b)-f(a)\rangle = \sum_{i=1}^{m} (f_i(a+t(b-a))\,(f_i(b)-f_i(a)) \quad (0 \le t \le 1). \tag{5.7}$$

Dan is ϕ continu op $[0, 1]$ en er volgt met behulp van de kettingregel dat ϕ differentieerbaar is op $(0, 1)$ met

$$\begin{aligned}
\phi'(t) &= \sum_{i=1}^{m} \sum_{j=1}^{n} (D_j f_i)(a + t(b - a))\,(b_j - a_j)\,(f_i(b) - f_i(a)) \\
&= \sum_{i=1}^{m} \big((f'(a + t(b - a))\,(b - a)\big)_i \big(f(b) - f(a)\big)_i \\
&= \langle f'(a + t(b - a))\,(b - a), f(b) - f(a)\rangle.
\end{aligned} \tag{5.8}$$

(ga na). Toepassing van (5.1) geeft dat $\phi(1) - \phi(0) = \phi'(t)$ voor zekere $t \in (0, 1)$. Maar anderzijds volgt uit (5.7) dat $\phi(1) - \phi(0) = |f(b) - f(a)|^2$. Combinatie met (5.8) levert uitspraak (5.6).

We bewijzen nu uit (5.6) de uitspraak (5.5). Pas de ongelijkheid van Schwarz toe op het rechter lid van (5.6). 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 (5.5). Als $|f(b) - f(a)| = 0$ dan is (5.5) evident. $\square$

Als onmiddellijk gevolg van Propositie 5.4 verkrijgen we nu de generalisatie van (5.3).

5.5 Stelling Zij $E \subset \mathbb{R}^n$ open en convex. Zij $f: E \to \mathbb{R}^m$ differentieerbaar op E. Veronderstel dat voor zekere $M \geq 0$ geldt dat $||f'(x)|| \leq M$ voor alle $x \in E$. Dan

$$|f(b) - f(a)| \leq M\,|b - a| \quad (a, b \in E). \tag{5.9}$$

5.6 We generaliseren nu de definitie van Lipschitz-continuïteit, zoals die in syll. Analyse A3, §3.36 gegeven was voor reëelwaardige functies op een open interval. 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 $M \geq 0$ bestaat zo dat

$$d_Y(f(a), f(b)) \leq M\, d_X(a, b) \quad (a, b \in X). \tag{5.10}$$

I.h.b. is dus, voor $E \subset \mathbb{R}^n$, een afbeelding $f: E \to \mathbb{R}^m$ Lipschitz-continu als voor zekere $M \geq 0$ de eigenschap (5.9) geldt.

Opgave
(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.

5.7 **Opgave** 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 (5.9) 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 §3.8.)

5.8 **Opgave** 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 (5.9) 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 (3.33) en (3.34) bij (a). Gebruik Opgave 1.11(b) bij (c).

5.2 Hogere partiële afgeleiden

5.9 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

$$(D_{ij}f)(a) := (D_i(D_jf))(a), \tag{5.11}$$

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

$$(D_{j_1\ldots j_k}f)(a) := (D_{j_1}(D_{j_2\ldots j_k}f))(a), \tag{5.12}$$

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 2.26 is bovenstaande definitie van C^k-afbeelding voor het geval $k = 1$ compatibel met de eerder in §2.25 gegeven definitie van C^1-afbeelding.

Opmerking over 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

$$\frac{\partial^k f(x_1, \ldots, x_n)}{\partial x_{j_1} \ldots \partial x_{j_k}} := (D_{j_1\ldots j_k}f)(x).$$

5.10 Opgave 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.

5.11 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 Zij $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $(0,0)$.

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.

Propositie Zij $(a,b) \in E \subset \mathbb{R}^2$ met E open en zij $f\colon E \to \mathbb{R}$ 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

$$f(a+h,b+k)-f(a+h,b)-f(a,b+k)+f(a,b) = h\left((D_1f)(a+\theta_1h,b+k)-(D_1f)(a+\theta_1h,b)\right) \tag{5.13}$$

voor zekere $\theta_1 \in (0,1)$. Herschrijf nu het rechterlid van (5.13) door toepassing van de middelwaardestelling op de differentieerbare functie $y \mapsto (D_1f)(a+\theta_1h,y)$. Dit levert:

$$f(a+h,b+k)-f(a+h,b)-f(a,b+k)+f(a,b) = hk\,(D_{21}f)(a+\theta_1h,b+\theta_2k) \tag{5.14}$$

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

$$f(a+h,b+k)-f(a+h,b)-f(a,b+k)+f(a,b) = hk\,(D_{12}f)(a+\eta_1h,b+\eta_2k) \tag{5.15}$$

voor zekere $\eta_1,\eta_2 \in (0,1)$. Door combinatie van (5.13) en (5.14) zien we dat

$$(D_{21}f)(a+\theta_1h,b+\theta_2k) = (D_{12}f)(a+\eta_1h,b+\eta_2k). \tag{5.16}$$

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

$$\begin{aligned}|(D_{12}f)(x,y) - D_{12}f)(a,b)| < \varepsilon/2 \quad \text{als } |(x,y)-(a,b)| < \delta,\\ |(D_{21}f)(x,y) - D_{21}f)(a,b)| < \varepsilon/2 \quad \text{als } |(x,y)-(a,b)| < \delta.\end{aligned}$$

Neem $h, k > 0$ zo dat $|(h,k)| < \delta$. Dan zal $|(a+\theta_1 h, b+\theta_2 k)-(a,b)| < \delta$ en $|a+\eta_1 h, b+\eta_2 k)-(a,b)| < \delta$ als $\theta_1, \theta_2, \eta_1, \eta_2 \in (0,1)$. Dus er volgt uit (5.16) en de driehoeksongelijkheid dat $|(D_{21}f)(a,b) - (D_{12}f)(a,b)| < \varepsilon$. Omdat $\varepsilon > 0$ willekeurig genomen was, volgt de Propositie. $\square$

5.12 **Opmerking** De conclusie van bovenstaande Proposite 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 $f\colon E \to \mathbb{R}$ zo dat $D_1 f$, $D_2 f$ 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 $(D_{12}f)(a,b) = (D_{21}f)(a,b)$.

5.13 Stelling Zij $E \subset \mathbb{R}^n$ open en zij $k = 2, 3, \ldots$. Zij $f\colon E \to \mathbb{R}$ 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 $D_{i_{\sigma(1)} \ldots i_{\sigma(k)}} f = D_{i_1 \ldots i_k} f$. In het bijzonder geldt voor een C^2-afbeelding dat

$$(D_{ij}f)(x) = (D_{ji}f)(x) \quad (x \in E). \tag{5.17}$$

Bewijs Het geval van algemene k volgt uit het geval $k = 2$ (formule (5.17)) (waarom?). Het geval $k = 2$, op zijn beurt, volgt uit het geval $k = 2$, $n = 2$ (waarom?). Dit laatste geval is evident uit Propositie 5.11. $\square$

5.14 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.

Propositie Stel I_1 en I_2 zijn open intervallen in $\mathbb{R}$. Laat $f\colon I_1 \times I_2 \to \mathbb{R}$ een C^1-functie zijn. Neem een gesloten begrensd interval $[a,b] \subset I_2$ en definieer de functie $F\colon I_1 \to \mathbb{R}$ door

$$F(x) := \int_a^b f(x,y)\, dy \quad (x \in I_1). \tag{5.18}$$

Dan is F differentieerbaar op I_1 met afgeleide

$$F'(x) = \int_a^b (D_1 f)(x,y)\, dy \quad (x \in I_1). \tag{5.19}$$

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 (5.18) en (5.19) zijn goed gedefinieerd als Riemann-integralen. Merk nu op dat, voor x en $x+h$ in I_1,

$$\begin{aligned}\frac{F(x+h)-F(x)}{h} - \int_a^b (D_1 f)(x,y)\, dy &= \int_a^b \Big(\frac{f(x+h,y)-f(x,y)}{h} - (D_1 f)(x,y)\Big)\, dy\\ &= \int_a^b \big((D_1 f)(x+\theta_y h, y) - (D_1 f)(x,y)\big)\, dy \quad (5.20)\end{aligned}$$

voor zekere $\theta_y \in (0,1)$. Hierbij gebruikten we de middelwaardestelling (5.1) 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 (5.20) 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

$$\lim_{h\to 0}\Big(\frac{F(x+h)-F(x)}{h} - \int_a^b (D_1 f)(x,y)\,dy\Big) = 0. \qquad \square$$

De conclusie van bovenstaande stelling gaat door onder veel zwakkere aannamen. De geïnteresseerde lezer wordt verwezen naar Rudin, §9.42, 9.43.

5.15 Opgave 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)$\Rightarrow$(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.

We zien dat f door eigenschap (a) uniek bepaald is op een constante term na. De implicatie (a)$\Rightarrow$(b) van Opgave 5.15 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.

5.16 Opgave 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 5.15, maar dat (a) van Opgave 5.15 niet geldt.

5.3 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 §3.6 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.

5.17 Laten dus, net als in §3.6, E_1 en E_2 open deelverzamelingen zijn van $\mathbb{R}^n$ en laten $\phi\colon E_1 \to E_2$ en $\psi\colon E_2 \to E_1$ nu C^2-afbeeldingen zijn die voldoen aan (3.17), i.e., C^2-diffeomorfismen die elkaars inversen zijn. Laten $F\colon E_1 \to \mathbb{R}$ en $f\colon E_2 \to \mathbb{R}$ nu C^2-functies zijn die met elkaar verband houden volgens (3.18). Dus

$$\begin{array}{lll} \psi(\phi(x)) = x \quad (x \in E_1) & \text{en} & \phi(\psi(u)) = u \quad (u \in E_2), \\ F(x) = f(\phi(x)) \quad (x \in E_1) & \text{en} & f(u) = F(\psi(u)) \quad (u \in E_2). \end{array}$$

We gaan nu uit van de ongenummerde formule even onder (3.18), die we hier nog eens geven.

$$(D_l f)(u) = \sum_{j=1}^{n} (D_l \psi_j)(u)\, (D_j F)(\psi(u)) \quad (u \in E_2). \tag{5.21}$$

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 (5.21) en we passen de formule

$$(D_k(gh))(u) = g(u)\, (D_k h)(u) + (D_k g)(u)\, h(u) \tag{5.22}$$

voor de partiële afgeleide van het product van twee partieel differentieerbare functies toe (cf. §2.21). We verkrijgen dan:

$$(D_{kl} f)(u) = \sum_{j=1}^{n} (D_l \psi_j)(u)\, (D_k((D_j F) \circ \psi))(u) + \sum_{j=1}^{n} (D_{kl} \psi_j)(u)\, (D_j F)(\psi(u)) \quad (u \in E_2). \tag{5.23}$$

We willen vervolgens de laatste factor in de eerste sommatie van (5.23), dus de uitdrukking $(D_k((D_j F) \circ \psi))(u)$, gaan herschrijven met behulp van (5.21). We schrijven hiertoe (5.21) eerst als

$$(D_k(F \circ \psi))(u) = \sum_{i=1}^{n} (D_k \psi_i)(u)\, (D_i F)(\psi(u)) \quad (u \in E_2).$$

Dan vervangen we in deze formule F door $D_j F$. De formule blijft geldig omdat F een C^2-functie is, dus $D_j F$ een C^1-functie. We verkrijgen

$$(D_k((D_j F) \circ \psi))(u) = \sum_{i=1}^{n} (D_k \psi_i)(u)\, (D_{ij} F)(\psi(u)) \quad (u \in E_2).$$

Als we dit substitueren in (5.23) dan volgt er dat

$$(D_{kl}f)(u) = \sum_{i,j=1}^{n} (D_k\psi_i)(u)\,(D_l\psi_j)(u)\,(D_{ij}F)(\psi(u)) \\ + \sum_{j=1}^{n}(D_{kl}\psi_j)(u)\,(D_jF)(\psi(u)) \quad (u \in E_2).$$

Door substitutie van $u := \phi(x)$ verkrijgen we tenslotte het volgende resultaat.

$$\begin{aligned}(D_{kl}f)(\phi(x)) = \sum_{i,j=1}^{n} (D_k\psi_i)(\phi(x))\,(D_l\psi_j)(\phi(x))\,(D_{ij}F)(x) \\ + \sum_{j=1}^{n}(D_{kl}\psi_j)(\phi(x))\,(D_jF)(x) \quad (x \in E_1).\end{aligned} \tag{5.24}$$

Formule (5.24) 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 (5.24) kan worden geschreven als de volgende identiteit van partiële differentiaaloperatoren.

$$\frac{\partial^2}{\partial u_k\,\partial u_l} = \sum_{i,j=1}^{n} \frac{\partial x_i}{\partial u_k}\,\frac{\partial x_j}{\partial u_l}\,\frac{\partial^2}{\partial x_i\,\partial x_j} + \sum_{j=1}^{n} \frac{\partial^2 x_j}{\partial u_k\,\partial u_l}\,\frac{\partial}{\partial x_j}\,. \tag{5.25}$$

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 (5.25) 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 (5.25) 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 (3.22) worden geschreven als (3.24), welke we nu in nog meer uitgeklede vorm schrijven als

$$\frac{\partial}{\partial u_l} = \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l}\,\frac{\partial}{x_j}\,. \tag{5.26}$$

We leiden nu (5.25) formeel af uit (5.26).

$$\begin{aligned}\frac{\partial^2}{\partial u_k\,\partial u_l} &= \frac{\partial}{\partial u_k} \circ \Big(\sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l}\,\frac{\partial}{\partial x_j}\Big) = \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l}\,\frac{\partial}{\partial u_k} \circ \frac{\partial}{\partial x_j} + \sum_{j=1}^{n} \frac{\partial^2 x_j}{\partial u_k\,\partial u_l}\,\frac{\partial}{\partial x_j} \\ &= \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_l}\,\Big(\sum_{i=1}^{n} \frac{\partial x_i}{\partial u_k}\,\frac{\partial^2}{\partial x_i\,\partial x_j}\Big) + \sum_{j=1}^{n} \frac{\partial^2 x_j}{\partial u_k\,\partial u_l}\,\frac{\partial}{\partial x_j}\,.\end{aligned} \tag{5.27}$$

Hierbij gebruikten we in de eerste en derde identiteit formule (5.26) en in de tweede identiteit formule (5.22).

5.18 Voorbeeld We bekijken enerzijds Cartesische coördinaten (x, y) en anderzijds poolcoördinaten (r, θ), die met elkaar verbonden zijn volgens (3.27). In §3.8 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 (5.25) 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 (3.33) en (3.34) en dan te schrijven

$$\begin{aligned}\frac{\partial^2}{\partial x^2} &= \Big(\cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial\theta}\Big)\circ\Big(\cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial\theta}\Big)\\ &= \cos^2\theta\,\frac{\partial^2}{\partial r^2} - \frac{2\sin\theta\,\cos\theta}{r}\,\frac{\partial^2}{\partial r\partial\theta} + \frac{\sin^2\theta}{r^2}\,\frac{\partial^2}{\partial\theta^2} + \frac{\sin^2\theta}{r}\,\frac{\partial}{\partial r} + \frac{2\sin\theta\,\cos\theta}{r^2}\,\frac{\partial}{\partial\theta}\,.\end{aligned} \tag{5.28}$$

Hierbij gebruikten we in de tweede identiteit formule (5.22). 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 (5.27), echter met het volgende verschil. Het zou met de tweede identiteit van (5.27) gecorrespondeerd hebben om (5.22) 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 (3.33) en dan pas (5.22) 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 (3.29) substitueren en dan (5.22) toepassen. Veel handiger is nu om (3.29) te schrijven in de vorm

$$\frac{\partial}{\partial r} = \cos\theta\,\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial y}, \tag{5.29}$$

en dan af te leiden dat

$$\begin{aligned}\frac{\partial^2}{\partial r^2} &= \frac{\partial}{\partial r}\circ\Big(\cos\theta\,\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial y}\Big)\\ &= \cos\theta\,\frac{\partial}{\partial r}\circ\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial r}\circ\frac{\partial}{\partial y}\\ &= \cos^2\theta\,\frac{\partial^2}{\partial x^2} + 2\sin\theta\,\cos\theta\,\frac{\partial^2}{\partial x\partial y} + \sin^2\theta\frac{\partial}{\partial y^2}\\ &= \frac{x^2}{x^2+y^2}\,\frac{\partial^2}{\partial x^2} + \frac{2xy}{x^2+y^2}\,\frac{\partial^2}{\partial x\,\partial y} + \frac{y^2}{x^2+y^2}\,\frac{\partial^2}{\partial y^2}\,.\end{aligned} \tag{5.30}$$

Hierbij gebruikten we in de eerste en derde identiteit formule (5.29) en in de tweede identiteit formule (5.22).

Opgave 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 (5.28) af dat

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\,\frac{\partial^2}{\partial\theta^2} + \frac{1}{r}\,\frac{\partial}{\partial r}\,. \tag{5.31}$$

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 (5.30) en (3.29), formule (5.31) nogmaals af.

De partiële differentiaaloperator $\partial^2/\partial x^2 + \partial^2/\partial y^2$ wordt de *Laplace-operator* (in twee variabelen) genoemd. Het rechter lid van (5.31) geeft de Laplace-operator in poolcoördinaten.

5.19 Opgave Bepaal alle C^2-functies f op $(0,\infty)$ zo dat

$$\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big) f(\sqrt{x^2+y^2}) = 0 \quad ((x,y) \in \mathbb{R}^2\backslash\{(0,0)\}).$$

5.20 **Opgave**
(a) Zij $f(z) := z^n$ ($z \in \mathbb{C}$, $n = 0,1,2,\ldots$). Bewijs dat de limiet

$$f'(z) := \lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{f(z+h) - f(z)}{h} \quad (z \in \mathbb{C}) \tag{5.32}$$

bestaat en gelijk is aan $n\,z^{n-1}$.
(b) Zij f een complexwaardige functie op $\mathbb{C}$ waarvoor de limiet (5.32) bestaat. (Dan noemen we f *complex differentieerbaar*.) Bewijs dat

$$\frac{\partial f(x+iy)}{\partial x} = f'(x+iy), \quad \frac{\partial f(x+iy)}{\partial y} = i\,f'(x+iy) \qquad (x,y) \in \mathbb{R}^2). \tag{5.33}$$

5.21 **Opgave** Bewijs dat voor $n = 0,1,2,\ldots$ geldt dat

$$\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big)\Big\{(x \pm iy)^n\Big\} = 0. \tag{5.34}$$

Doe dit enerzijds door directe partiële differentiatie, anderzijds door over te gaan op poolcoördinaten en gebruik te maken van (5.31).

5.22 Opgave Zij $|x| := \sqrt{x_1^2 + \cdots + x_n^2}$ en zij Δ de partiële differentiaaloperator

$$\Delta := D_{11} + \cdots + D_{nn} = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\,. \tag{5.35}$$

(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

$$\Delta(|x|^\alpha) = \alpha\,(n + \alpha - 2)\,|x|^{\alpha-2}. \tag{5.36}$$

De operator Δ gegeven door (5.35) 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 (5.36) dat f gedefinieerd door $f(x) := |x|^{-n+2}$ harmonisch is op $\mathbb{R}^n \backslash \{0\}$.

5.23 Opgave Zij $|x| := \sqrt{x_1^2 + \cdots + x_n^2}$ en zij Δ gegeven door (5.35). Zij $f\colon (0, \infty) \to \mathbb{R}$ een C^2-functie. Bewijs dat

$$\Delta(f(|x|)) = f''(|x|) + (n-1)\,|x|^{-1}\,f'(|x|). \tag{5.37}$$

Verdere vraagstukken

V5.1 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

$$\frac{\partial^2 F}{\partial x^2} = \frac{1}{c^2}\,\frac{\partial^2 F}{\partial t^2} \tag{5.38}$$

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 (5.38) geschreven kan worden in de vorm

$$F(x, y) = f(x + ct) + g(x - ct)$$

waarin $f, g\colon \mathbb{R} \to \mathbb{R}$ twee tweemaal differentieerbare functies zijn.

6 Hogere totale afgeleide; Taylorreeks

Het belangrijkste deel van dit hoofdstuk is het laatste onderdeel, waar Taylorreeksen voor functies in meer veranderlijken worden behandeld. De twee eerste onderdelen zijn nodig ter voorbereiding van de Taylorreeksen, maar hebben ook belang op zich. We beginnen met algemeenheden over multilineaire afbeeldingen. Dan voeren we de totale afgeleide van hogere orde in, die blijkt te kunnen worden beschreven als een multilineaire afbeelding. Tenslotte blijkt de formule van Taylor dan compact te kunnen worden geschreven in termen van deze totale afgeleiden.

6.1 Multilineaire afbeeldingen

Multilineaire afbeeldingen zijn generalisaties van lineaire afbeeldingen. De hier beschreven theorie gaat niet erg diep, maar geeft toch nuttige kennis, niet alleen binnen het bestek van deze syllabus, maar ook in veel andere delen van de wiskunde waar multilineariteit opduikt.

6.1 Definitie Laten $V_1, \ldots, V_p$ en W reële vectorruimtes zijn. Een afbeelding $B\colon V_1 \times \cdots \times V_p \to W$ heet *multilineair* als voor elke $i \in \{1, \ldots, p\}$ geldt dat, bij vaste $v_1, \ldots, v_{i-1}$, $v_{i+1}, \ldots, v_p$, de afbeelding $v_i \mapsto B(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_p)\colon V_i \to W$ lineair is. De verzameling van al zulke multilineaire afbeeldingen wordt aangeduid met $\Lambda(V_1, \ldots, V_p; W)$. (Een analoge definitie geldt voor het geval dat de vectorruimtes complex zijn.)

De verzameling $\Lambda(V_1, \ldots, V_p; W)$ krijgt op een natuurlijke manier de structuur van een reële vectorruimte als volgt. Als $B_1, B_2 \in \Lambda(V_1, \ldots, V_p; W)$ en $\lambda, \mu \in \mathbb{R}$ dan definiëren we de multilineaire afbeelding $\lambda B_1 + \mu B_2$ door

$$(\lambda B_1 + \mu B_2)(v_1, \ldots, v_p) := \lambda\, B_1(v_1, \ldots, v_p) + \mu\, B_2(v_1, \ldots, v_p).$$

We kunnen een aantal speciale gevallen onderscheiden.

- $W = \mathbb{R}$. Dan spreken we van een *multilineaire vorm* en we schrijven $\Lambda(V_1, \ldots, V_p)$ i.p.v. $\Lambda(V_1, \ldots, V_p; \mathbb{R})$.
- $p = 2$. Dan spreken we van een *bilineaire afbeelding* of, als bovendien $W = \mathbb{R}$, van een *bilineaire vorm*.
- $p = 1$. Dan zijn we terug bij het geval van een *lineaire afbeelding* en $\Lambda(V; W) = L(V, W)$.
- $V_1 = \cdots = V_p$. Dan schrijven we $\Lambda_p(V, W)$ i.p.v. $\Lambda(V, \ldots, V; W)$. Als bovendien $W = \mathbb{R}$ dan schrijven we $\Lambda_p(V)$ i.p.v. $\Lambda(V, \ldots, V)$. We noemen een multilineaire afbeelding $B \in \Lambda_p(V, W)$ *symmetrisch* als $B(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) = B(v_1, \ldots, v_p)$ voor elke permutatie σ van $\{1, \ldots, p\}$. I.h.b. is een bilineaire afbeelding $B \in \Lambda_2(V, W)$ symmetrisch als $B(v_1, v_2) = B(v_2, v_1)$.
- $V_1 = \cdots = V_p$ met alle argumenten gelijk. Zij $B \in \Lambda_p(V, W)$. Dan noemen we de afbeelding $v \to B(v, v, \ldots, v)\colon V \to W$ *(homogeen) van graad p*. In het geval $p = 2$ noemen we de afbeelding *(homogeen) kwadratisch*. Als $W = \mathbb{R}$ dan spreken we van een *vorm van graad p op V*, resp. van een *kwadratische vorm op V*.

Opgave Bewijs dat de afbeelding die aan iedere matrix $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ de bilineaire vorm

$$B(x, y) := \sum_{i,j=1}^{n} A_{ij}\, x_i\, y_j = \langle x, Ay \rangle \quad (x, y \in \mathbb{R}^n) \tag{6.1}$$

toevoegt, een bijectieve lineaire afbeelding is van $L(\mathbb{R}^n, \mathbb{R}^n)$ op $\Lambda_2(\mathbb{R}^n)$. Bewijs dat de bilineaire vorm B symmetrisch is desda de matrix A symmetrisch is. Bewijs dat iedere kwadratische vorm f op $\mathbb{R}^n$ een unieke symmetrische $B \in \Lambda_2(\mathbb{R}^n)$ en symmetrische $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ bepaalt zo dat $f(x) = B(x, x) = \langle x, Ax \rangle$.

6.2 We zouden $(A_{ij})_{i,j=1,\ldots,n}$ in (6.1) de matrix van de bilineaire vorm $B \in \Lambda_2(\mathbb{R}^n)$ kunnen noemen. Algemener, als $B \in \Lambda_p(\mathbb{R}^n)$ een multilineaire vorm is dan kunnen we schrijven

$$B(x^{(1)}, \ldots, x^{(p)}) = \sum_{j_1,\ldots,j_p=1}^{n} B_{j_1\ldots j_p}\, x^{(1)}_{j_1} \ldots x^{(p)}_{j_p} \qquad (x^{(1)}, \ldots, x^{(p)} \in \mathbb{R}^n). \tag{6.2}$$

Dan noemen we $(B_{j_1\ldots j_p})_{j_1,\ldots,j_p=1,\ldots,n}$ de (gegeneraliseerde) *matrix* van B. Men kan met weinig moeite inzien dat er een bijectieve lineaire correspondentie is tussen multilineaire vormen (op $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ voor vaste p) en gegeneraliseerde matrices. I.h.b. bepaalt een multilineaire vorm een unieke gegeneraliseerde matrix en omgekeerd bepaalt een gegeneraliseerde matrix een unieke multilineaire vorm. Ook geldt dat de multilineaire vorm $B \in \Lambda_p(\mathbb{R}^n)$ symmetrisch is desda de bijbehorende gegeneraliseerde matrix $(B_{j_1\ldots j_p})_{j_1,\ldots,j_p=1,\ldots,n}$ symmetrisch is in haar indices $j_1, \ldots, j_p$.

6.3 **Opgave** Zij P een polynoom van graad $\le p$ in de n reële variabelen $x_1, \ldots, x_n$. Bewijs dat P unieke symmetrische multilineaire vormen $B^{(k)} \in \Lambda_k(\mathbb{R}^n)$ $(k = 1, \ldots, p)$ en $B^{(0)} \in \mathbb{R}$ bepaalt zo dat

$$P(x) = B^{(0)} + B^{(1)}(x) + B^{(2)}(x, x) + \cdots + B^{(p)}(x, \ldots, x) \quad (x \in \mathbb{R}^n). \tag{6.3}$$

Merk op dat omgekeerd het rechter lid van (6.3) voor gegeven symmetrische multilineaire vormen $B^{(k)} \in \Lambda_k(\mathbb{R}^n)$ $(k = 1, \ldots, p)$ en $B^{(0)} \in \mathbb{R}$ een polynoom $P(x)$ definieert van graad $\le p$ in de coördinaten $x_1, \ldots, x_n$ van x.

6.4 Nog wat algemener kunnen we, als $B \in \Lambda_p(\mathbb{R}^n, \mathbb{R}^m)$, schrijven:

$$\big(B(x^{(1)}, \ldots, x^{(p)})\big)_i = \sum_{j_1,\ldots,j_p=1}^{n} B_{i;j_1\ldots j_p}\, x^{(1)}_{j_1} \ldots x^{(p)}_{j_p} \qquad (x^{(1)}, \ldots, x^{(p)} \in \mathbb{R}^n,\ i = 1, \ldots, m). \tag{6.4}$$

Wederom noemen we $(B_{i;j_1\ldots j_p})_{i=1,\ldots,m;\ j_1,\ldots,j_p=1,\ldots,n}$ de *gegeneraliseerde matrix* van de multilineaire afbeelding B. We kunnen weer soortgelijke uitspraken doen als in de §6.2.

6.5 Definitie Laten $V_1, \ldots, V_p$ en W nu genormeerde reële vectorruimtes zijn. Een multilineaire afbeelding $B \in \Lambda(V_1, \ldots, V_p; W)$ heet *begrensd* als

$$||B|| := \sup_{\substack{v_1 \in V_1, \ldots, v_p \in V_p \\ ||v_1||, \ldots, ||v_p|| \leq 1}} ||B(v_1, \ldots, v_p)|| < \infty. \tag{6.5}$$

Opgave Laten $V_1, \ldots, V_p$ en W genormeerde reële vectorruimtes zijn en zij de multilineaire afbeelding $B \in \Lambda(V_1, \ldots, V_p; W)$ begrensd. Bewijs dat

$$||B(v_1, \ldots, v_p)|| \leq ||B|| \, ||v_1|| \ldots ||v_p|| \qquad (v_1, \ldots, v_p \in V). \tag{6.6}$$

6.6 **Opgave** Laten $V_1, \ldots, V_p$ en W genormeerde reële vectorruimtes zijn, $V_1, \ldots, V_p$ bovendien eindig-dimensionaal. Bewijs dat alle $B \in \Lambda(V_1, \ldots, V_p; W)$ begrensd zijn en dat, met $||B||$ gedefinieerd als in (6.5), $\Lambda(V_1, \ldots, V_p; W)$ een genormeerde reële vectorruimte wordt.

6.7 Opgave Laten V_1, V_2 en W reële vectorruimtes zijn. Bewijs het volgende. Er bestaat een bijectieve lineaire correspondentie

$$A \leftrightarrow B \colon L(V_1, L(V_2, W)) \leftrightarrow \Lambda(V_1, V_2; W) \quad \text{zo dat} \quad (Av_1)(v_2) = B(v_1, v_2) \tag{6.7}$$

als $v_1 \in V_1$ en $v_2 \in V_2$.

Deze uitspraak is een speciaal geval van de volgende propositie.

6.8 Propositie Laten $V_1, \ldots, V_p$ en W reële vectorruimtes zijn. Dan bestaat er een bijectieve lineaire correspondentie

$$A \leftrightarrow B \colon L(V_1, \Lambda(V_2, \ldots, V_p; W)) \leftrightarrow \Lambda(V_1, \ldots, V_p; W), \tag{6.8}$$

zo dat

$$(Av_1)(v_2, \ldots, v_p) = B(v_1, v_2, \ldots, v_p) \quad (v_1 \in V_1, v_2 \in V_2, \ldots, v_p \in V_p). \tag{6.9}$$

Het bewijs wordt als niet verplichte opgave aan de lezer overgelaten.

6.2 Hogere totale afgeleide

6.9 Zij $a \in E \subset \mathbb{R}^n$, E open en $f \colon E \to \mathbb{R}^m$ differentieerbaar in a met afgeleide $f'(a)$ in $L(\mathbb{R}^n, \mathbb{R}^m)$. We gaan weer even terug naar de definitie van $f'(a)$, zoals gegeven in (2.15). Dus $f'(a)$ is het unieke element van $L(\mathbb{R}^n, \mathbb{R}^m)$ zo dat

$$\lim_{h \to 0} |h|^{-1} \left(f(a+h) - f(a) - f'(a)\, h\right) = 0. \tag{6.10}$$

Hier ligt $|h|^{-1} \left(f(a+h - f(a) - f'(a)\, h\right)$ in $\mathbb{R}^m$, h ligt in $\mathbb{R}^n$, en de limiet wordt genomen t.o.v. de standaardnormen in $\mathbb{R}^n$ en $\mathbb{R}^m$. Omdat alle normen op een eindig-dimensionale vectorruimte equivalent zijn (cf. Propositie 1.6), verandert er niets aan de limiet in (6.10) wanneer we deze nemen t.o.v. een andere norm op $\mathbb{R}^m$ dan de standaardnorm. Dan hoeven we ook niet meer over de concrete ruimte $\mathbb{R}^m$ te spreken, maar kunnen we in plaats daarvan een eindig-dimensionale genormeerde reële vectorruimte $(V, ||\,.\,||)$ als beeldruimte voor f nemen.

We kunnen dus Definitie 2.8 (van differentieerbaarheid) als volgt herformuleren.

Definitie Zij $a \in E \subset \mathbb{R}^n$, E open en $f: E \to V$, waarbij $(V, ||\,.\,||)$ een eindig-dimensionale genormeerde reële vectorruimte is. We zeggen dat f *differentieerbaar* is in a met *afgeleide* $f'(a) \in L(\mathbb{R}^n, V)$ als (6.10) geldt.

We kunnen bovendien de uitspraak van Opgave 2.10 als volgt herformuleren.

Propositie Zij $a \in E \subset \mathbb{R}^n$, E open en $f: E \to V$, waarbij $(V, ||\,.\,||)$ een eindig-dimensionale genormeerde reële vectorruimte is. Kies een basis $e_1, \ldots, e_m$ voor V en schrijf $f(x) = (f_1(x), \ldots, f_m(x))$ t.o.v. deze basis. Dan is f differentieerbaar in a desda de functies $f_i: E \to \mathbb{R}$ differentieerbaar zijn in a. Voorts geldt er in het geval van differentieerbaarheid dat

$$(f'(a)\,h)_i = f_i'(a)\,h \quad (h \in \mathbb{R}^n,\ i = 1, \ldots, m). \tag{6.11}$$

Hier is $(f'(a)\,h)_i$ de coördiaat t.o.v. e_i van $f'(a)\,h \in V$.

6.10 We passen de formuleringen van §6.9 nu toe op de afbeelding $f': E \to L(\mathbb{R}^n, \mathbb{R}^m)$, waarbij $f: E \to \mathbb{R}^m$ een differentieerbare afbeelding is van een open deelverzameling E van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Dan is $L(\mathbb{R}^n, \mathbb{R}^m)$ een eindig-dimensionale reële vectorruimte (van dimensie nm) waarop we bijv. de operatornorm als norm kunnen kiezen. Voor V uit §6.9 substitueren we $L(\mathbb{R}^n, \mathbb{R}^m)$ en voor f uit §6.9 substitueren we f'. We kunnen dan dus zeggen dat f' differentieerbaar is in een punt a als voor zekere $A \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$ geldt dat

$$\lim_{h \to 0} |h|^{-1} \,(f'(a+h) - f'(a) - A\,h) = 0. \tag{6.12}$$

Dan is dus $(f')'(a)$, of kortweg $f''(a)$, gelijk aan A. We noemen $f''(a)$ de *tweede totale afgeleide* van f in a.

Laten we (6.12) aan een nadere beschouwing onderwerpen, waarbij we $f''(a)$ i.p.v. A schrijven. Van belang is de uitdrukking

$$r(h) := f'(a+h) - f'(a) - f''(a)\,h. \tag{6.13}$$

Hier liggen $f'(a+h)$ en $f'(a)$ in de ruimte $L(\mathbb{R}^n, \mathbb{R}^m)$. Omdat $f''(a) \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$ en $h \in \mathbb{R}^n$, zal $f''(a)\,h$ ook weer in $L(\mathbb{R}^n, \mathbb{R}^m)$ liggen. Dus $r(h) \in L(\mathbb{R}^n, \mathbb{R}^m)$. Het gaat er nu om dat $|h|^{-1}\,r(h) \to 0$ in $L(\mathbb{R}^n, \mathbb{R}^m)$ als $h \to 0$ in $\mathbb{R}^n$. Dit is equivalent met $\lim_{h \to 0} |h|^{-1}\,||r(h)|| = 0$, waarbij $||\,.\,||$ bijv. de operatornorm is op $L(\mathbb{R}^n, \mathbb{R}^m)$.

We willen nu kontakt maken met het speciale geval (6.7) van Propositie 6.8. Dat impliceert dat de afbeelding $(h, k) \mapsto (f''(a)\,h)\,k: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ bilineair is. We gebruiken voor deze bilineaire afbeelding dezelfde notatie $f''(a)$, dus

$$(f''(a)\,h)\,k = f''(a)\,(h, k) \quad (h, k \in \mathbb{R}^n). \tag{6.14}$$

Laat nu beide leden van (6.13) op $k \in \mathbb{R}^n$ werken. Dan

$$r(h)\,k = f'(a+h)\,k - f'(a)\,k - f''(a)\,(h, k) \quad (h \in E - a,\ k \in \mathbb{R}^n). \tag{6.15}$$

We kunnen dan een nieuwe definitie van $f''(a)$ geven, die echter equivalent is aan de zojuist gegeven definitie.

Definitie Zij $f\colon E \to \mathbb{R}^m$ een differentieerbare afbeelding van een open deelverzameling E van $\mathbb{R}^n$ naar $\mathbb{R}^m$. We zeggen dat f *twee maal differentieerbaar* is in een punt $a \in E$ met *tweede afgeleide* $f''(a) \in \Lambda_2(\mathbb{R}^n, \mathbb{R}^m)$ als $r(h)$, welke voor $h \in E - a$ als element van $L(\mathbb{R}^n, \mathbb{R}^m)$ gedefinieerd is door (6.15), voldoet aan

$$\lim_{h\to 0} |h|^{-1}\, r(h) = 0.$$

We kunnen in $L(\mathbb{R}^n, \mathbb{R}^m)$ een basis kiezen zo dat de coördinaten van een $(m \times n)$-matrix A t.o.v. deze basis juist de matrixelementen A_{ij} zijn. Gezien Stelling 2.14 wordt het (i,j)-de matrixelement van $f'(a)$ gegeven door $(D_j f_i)(a)$. Dus wegens Propositie 6.9 is f' differentieerbaar in a desda voor alle $i = 1, \ldots, m$ en alle $j = 1, \ldots, n$ de functie $D_j f_i\colon E \to \mathbb{R}$ differentieerbaar is in a. Dus dan bestaan de partiële afgeleiden van tweede orde van f_i in a, i.e. de uitdrukkingen $(D_{lj} f_i)(a)$. Voorts geldt er gezien (6.11) in het geval van differentieerbaarheid van f' in a dat

$$\big((f')'(a)\, h\big)_{ij} = (D_j f_i)'(a)\, h \quad (h \in \mathbb{R}^n).$$

Dus, met toepassing van (2.21) en met herschrijving van $(f')'$ als f'',

$$(f''(a)\, h)_{ij} = \sum_{l=1}^{n} (D_{lj} f_i)(a)\, h_l \quad (h \in \mathbb{R}^n).$$

Hier is $f''(a)$ opgevat als element van $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$. Laat nu $k = (k_1, \ldots, k_n) \in \mathbb{R}^n$, dan impliceert bovenstaande identiteit dat

$$\sum_{j=1}^{n} (f''(a)\, h)_{ij}\, k_j = \sum_{l,j=1}^{n} (D_{lj} f_i)(a)\, h_l\, k_j \quad (h, k \in \mathbb{R}^n).$$

Omdat

$$\sum_{j=1}^{n} (f''(a)\, h)_{ij}\, k_j = ((f''(a)\, h)\, k)_i = ((f''(a)\, (h, k))_i$$

(in de tweede identiteit gebruikten we (6.14)), concluderen we dat

$$(f''(a)\, (h, k))_i = \sum_{l,j=1}^{n} (D_{lj} f_i)(a)\, h_l\, k_j \quad (h, k \in \mathbb{R}^n). \tag{6.16}$$

Dus we hebben de volgende stelling bewezen.

Stelling Zij $f\colon E \to \mathbb{R}^m$ een differentieerbare afbeelding van een open deelverzameling E van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Zij $a \in E$. Dan is f twee keer differentieerbaar in a met tweede afgeleide $f''(a) \in \Lambda_2(\mathbb{R}^n, \mathbb{R}^m)$ desda als voor alle $j \in \{1, \ldots, n\}$ en alle $i \in \{1, \ldots, m\}$ geldt dat de functie $D_j f_i\colon E \to \mathbb{R}$ differentieerbaar is in a. Bovendien geldt (6.16) als $f''(a)$ bestaat.

6.11 Wanneer we (6.16) vergelijken met (6.4) dan zien we dat $\big((D_{lj}f_i)(a)\big)_{i=1,\ldots,m;\ l,j=1,\ldots,n}$ als de gegeneraliseerde matrix van de bilineaire afbeelding $f''(a)$ beschouwd kan worden. Er zal dus gelden dat $f''\colon E \to \Lambda_2(\mathbb{R}^n, \mathbb{R}^m)$ continu is desda alle functies $D_{lj}f_i$ continu zijn op E.

In §5.9 hebben we gedefinieerd wanneer $f\colon E \to \mathbb{R}^m$ een C^2-afbeelding heet. Deze definitie is in termen van existentie en continuïteit van de partiële afgeleiden van de f_i van eerste en tweede orde. Op grond van Stelling 6.10, Stelling 2.26 en de opmerking uit de vorige alinea zien we nu het volgende in.

Propositie Zij $E \subset \mathbb{R}^n$ open en $f\colon E \to \mathbb{R}^m$. Dan is f een C^2-afbeelding desda f twee keer differentieerbaar is op E en de afbeelding $f''\colon E \to \Lambda_2(\mathbb{R}^n, \mathbb{R}^m)$ continu is.

Gezien Stelling 5.13 zal voor een C^2-afbeelding $f\colon E \to \mathbb{R}^m$ (E open in $\mathbb{R}^n$) gelden dat de bilineaire afbeelding $f''(x) \in \Lambda_2(\mathbb{R}^n, \mathbb{R}^m)$ symmetrisch is voor elke $x \in E$.

De Definitie en Stelling uit §6.10 zijn speciaal van belang als $m = 1$, dus voor afbeeldingen $f\colon E \to \mathbb{R}$. Dan kunnen we de index i overal weglaten. Zo wordt (6.16) als volgt.

$$f''(a)\,(h,k) = \sum_{l,j=1}^{n} (D_{lj}f)(a)\,h_l\,k_j \quad (h,k \in \mathbb{R}^n). \tag{6.17}$$

Voor een C^2-functie $f\colon E \to \mathbb{R}$ geldt dan dat de bilineaire vorm $f''(x)$ symmetrisch is voor alle $x \in E$.

6.12 We kunnen bovenstaande redeneringen voortzetten voor totale afgeleiden van hogere orde. De definitie hieronder gebruikt volledige recurrentie naar de orde p van differentiatie, het bewijs van de stelling gebruikt volledige inductie naar p. We laten het bewijs echter achterwege. Wie het bovenstaande begrepen heeft en wie zich uitgedaagd voelt, moet zeker in staat zijn om het bewijs zelf te leveren.

Definitie Zij $f\colon E \to \mathbb{R}^m$ een afbeelding van een open deelverzameling E van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Zij $p \in \{2,3,4,\ldots\}$. Neem aan dat $f^{(p-1)}\colon E \to \Lambda_{p-1}(\mathbb{R}^n, \mathbb{R}^m)$ reeds gedefinieerd is en bestaat. Dan zeggen we dat f *p maal differentieerbaar* is in een punt $a \in E$ met *pde afgeleide* $f^{(p)}(a) \in \Lambda_p(\mathbb{R}^n, \mathbb{R}^m)$ als $r(h)$, dat we voor $h \in E - a$ als element van $\Lambda_{p-1}(\mathbb{R}^n, \mathbb{R}^m)$ definiëren door

$$\begin{aligned} r(h)\,(h^{(2)},\ldots,h^{(p)}) :=&f^{(p-1)}(a+h)\,(h^{(2)},\ldots,h^{(p)}) - f^{(p-1)}(a)\,(h^{(2)},\ldots,h^{(p)}) \\ &- f^{(p)}(a)\,(h,h^{(2)},\ldots,h^{(p)}) \qquad (h^{(2)},\ldots,h^{(p)} \in \mathbb{R}^n), \end{aligned} \tag{6.18}$$

voldoet aan

$$\lim_{h\to 0} |h|^{-1}\,r(h) = 0.$$

In bovenstaande Definitie wordt in feite de totale afgeleide van eerste orde van $f^{(p-1)}$ in a gedefinieerd. Hierbij wordt de identificatie van twee ruimtes gebruikt, zoals gegeven door (6.8) en (6.9).

Stelling Zij $f: E \to \mathbb{R}^m$ een afbeelding van een open deelverzameling E van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Zij $p \in \{2,3,4,\ldots\}$. Neem aan dat f $p-1$ keer differentieerbaar is op E. Zij $a \in E$. Dan is f p keer differentieerbaar in a met $p$de afgeleide $f^{(p)}(a) \in \Lambda_p(\mathbb{R}^n, \mathbb{R}^m)$ desda als voor alle $j_1, \ldots, j_{p-1} \in \{1, \ldots, n\}$ en voor alle $i \in \{1, \ldots, m\}$ geldt dat de functie $D_{j_1 \ldots j_{p-1}} f_i: E \to \mathbb{R}$ differentieerbaar is in a. Bovendien geldt, als $f^{(p)}(a)$ bestaat, dat

$$\big(f^{(p)}(a)\,(h^{(1)}, \ldots, h^{(p)})\big)_i = \sum_{j_1, \ldots, j_p = 1}^{n} (D_{j_1 \ldots j_p} f_i)(a)\, h^{(1)}_{j_1} \ldots h^{(p)}_{j_p} \qquad (h^{(1)}, \ldots, h^{(p)} \in \mathbb{R}^n). \tag{6.19}$$

Vergelijk formule (6.19) met formule (6.4). Dan zien we dat de gegeneraliseerde matrixelementen van de multilineaire afbeelding $f^{(p)}(a)\,(h^{(1)}, \ldots, h^{(p)})$ gegeven worden door

$$\big(f^{(p)}(a)\,(h^{(1)}, \ldots, h^{(p)})\big)_{i;\, j_1, \ldots, j_p} = (D_{j_1 \ldots j_p} f_i)(a).$$

Nu kunnen we een C^p-afbeelding, zoals gedefinieerd in 5.9, ook equivalent als volgt karakteriseren.

Propositie Zij $E \subset \mathbb{R}^n$ open en $f: E \to \mathbb{R}^m$. Dan is f een C^p-afbeelding desda f p keer differentieerbaar is op E en de afbeelding $f^{(p)}: E \to \Lambda_p(\mathbb{R}^n, \mathbb{R}^m)$ continu is. Voor een C^p-afbeelding f geldt bovendien dat de multilineaire afbeelding $f^{(p)}(x) \in \Lambda_p(\mathbb{R}^n, \mathbb{R}^m)$ symmetrisch is voor elke $x \in E$.

Bovenstaande Definitie en Stelling worden het meest gebruikt in het geval m=1. Dan is dus $f: E \to \mathbb{R}$ een reëelwaardige functie, $f^{(p)}(a)$ behoort tot $\Lambda_p(\mathbb{R}^n)$ en is dus een multilineaire vorm op $\mathbb{R}^n$, en in (6.19) kunnen we de subindices i weglaten. Formules met partiële afgeleiden van hogere orde kunnen vaak veel compacter geschreven worden als we met $f^{(p)}(a)$ werken. Daarom geven we nogmaals een definitie van dit object, zo dat die los van §85 gebruikt kan worden.

6.13 Definitie Zij $E \subset \mathbb{R}^n$ open en $f: E \to \mathbb{R}$ een C^p-functie. Zij $a \in E$. Definieer de multilineaire afbeelding $f^{(p)}(a) \in \Lambda_p(\mathbb{R}^n)$ door

$$f^{(p)}(a)\,(h^{(1)}, \ldots, h^{(p)}) := \sum_{j_1, \ldots, j_p = 1}^{n} (D_{j_1 \ldots j_p} f)(a)\, h^{(1)}_{j_1} \ldots h^{(p)}_{j_p} \qquad (h^{(1)}, \ldots, h^{(p)} \in \mathbb{R}^n). \tag{6.20}$$

Dan heet $f^{(p)}(a)$ de $p$de afgeleide van f in a.

Merk op dat wegens Stelling 5.13 de multilineaire afbeelding $f^{(p)}(a)$ symmetrisch is, d.w.z. dat het linker lid van (6.20) niet in waarde verandert als we de p variabelen $h^{(1)}, \ldots, h^{(p)}$ permuteren.

6.14 **Opmerking** Als f (voor m willekeurig) voldoet aan de aannamen van bovenstaande Stelling dan zien we door vergelijken van (6.19) met (6.20) dat

$$\big(f^{(p)}(a)\,(h^{(1)}, \ldots, h^{(p)})\big)_i = f_i^{(p)}(a)\,(h^{(1)}, \ldots, h^{(p)}) \quad (i = 1, \ldots, m). \tag{6.21}$$

6.3 Taylorreeksen

6.15 We recapituleren de *stelling van Taylor* voor reëelwaardige functies van één reële variabele (cf. syll. Analyse A).

Stelling Zij I een open interval in $\mathbb{R}$. Zij $f\colon I \to \mathbb{R}$ 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

$$f(a+h) = f(a) + f'(a)\,h + \frac{f''(a)\,h^2}{2!} + \cdots + \frac{f^{(p-1)}(a)\,h^{p-1}}{(p-1)!} + \frac{f^{(p)}(a+\theta h)\,h^p}{p!}. \tag{6.22}$$

Formule (6.22) 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 (5.1).

Neem a in (6.22) vast en herschrijf (6.22) 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

$$R_p(h) = \mathcal{O}(|h|^p) \quad \text{als } h \to 0$$

en

$$R_p(h) = \frac{f^{(p)}(a)\,h^p}{p!} + o(|h|^p) \quad \text{als } h \to 0.$$

Dus voor een C^p-functie f op een open interval I en $a \in I$ gelden ook de volgende varianten van de Taylorreeks (6.22).

$$f(a+h) = f(a) + \sum_{k=1}^{p-1} \frac{f^{(k)}(a)\,h^k}{k!} + \mathcal{O}(|h|^p) \quad \text{als } h \to 0 \tag{6.23}$$

en

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{f^{(k)}(a)\,h^k}{k!} + o(|h|^p) \quad \text{als } h \to 0. \tag{6.24}$$

6.16 We willen nu een analogon van (6.22) geven voor het geval f van meer dan één reële variabele afhangt, maar nog reëelwaardig is. Net als in het bewijs van Propositie 5.3 zullen we het resultaat bewijzen door het terug te brengen tot het geval van één variabele. Ter voorbereiding geven we een lemma.

Lemma Zij $a \in E \subset \mathbb{R}^n$ met E open, zij $f\colon E \to \mathbb{R}$ 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

$$F^{(k)}(t) = \sum_{j_1 \ldots j_k = 1}^{n} (D_{j_1 \ldots j_k} f)(a + t\,h)\, h_{j_1} \ldots h_{j_k} = f^{(k)}(a + t\,h)\,(h, \ldots, h). \tag{6.25}$$

Bewijs De tweede identiteit van (6.25) volgt uit (6.20). De eerste identiteit bewijzen we met volledige inductie naar k. We hebben al vaker gezien dat, als $g\colon E \to \mathbb{R}$ een C^1-functie is en $G(t) := g(a + t\,h)$, dan

$$G'(t) = \sum_{j=1}^{n} (D_j g)(a + t\,h)\, h_j. \tag{6.26}$$

Dit geeft het geval $k = 1$ van (6.25). Bovendien, als $p > k$, dan geeft (6.26) 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 (6.25) geeft dezelfde formule met k vervangen door $k + 1$. □

6.17 Stelling Zij $a \in E \subset \mathbb{R}^n$ met E open en zij $f\colon E \to \mathbb{R}$ 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

$$f(a + h) = f(a) + f'(a)\,h + \cdots + \frac{f^{(p-1)}(a)\,(h, \ldots, h)}{(p-1)!} + R_p(h), \tag{6.27}$$

waarbij

$$R_p(h) = \frac{f^{(p)}(a + \theta\,h)\,(h, \ldots, h)}{p!}\,. \tag{6.28}$$

Hierbij kunnen we schrijven

$$f^{(k)}(x)\,\underbrace{(h, \ldots, h)}_{k\ h\text{'s}} = \sum_{j_1, \ldots, j_k = 1}^{n} (D_{j_1 \ldots j_k} f)(x)\, h_{j_1} \ldots h_{j_k}. \tag{6.29}$$

Bovendien geldt er dat

$$R_p(h) = \mathcal{O}(|h|^p) \quad \text{als } h \to 0 \text{ in } \mathbb{R}^n \tag{6.30}$$

en

$$R_p(h) = \frac{f^{(p)}(a)\,(h, \ldots, h)}{p!} + o(|h|^p) \quad \text{als } h \to 0 \text{ in } \mathbb{R}^n. \tag{6.31}$$

Bewijs Schrijf $F(t) := f(a + t\,h)$ en pas Stelling 6.15 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 (6.25). Dit bewijst (6.27) samen met (6.28).

Voor het bewijs van (6.31) concluderen we uit (6.28) en (6.6) dat

$$|R_p(h) - f^{(p)}(a)\,(h,\ldots,h)\,/p!| \le ||f^{(p)}(a+\theta h) - f^{(p)}(a)||\,|h|^p.$$

Omdat f een C^p-functie is, is $f^{(p)}$ continu op E. Dit bewijst (6.31). Tenslotte volgt (6.30) uit (6.31), wederom met gebruik van (6.6). □

Laten we deze Stelling wat toelichten. In het geval van één variabele geeft formule (6.24) de benadering van de C^p-functie f door een polynoom (het zogenaamde *Taylor-polynoom*) $x \mapsto f(a) + \sum_{k=1}^{p} f^{(k)}(a)\,(x-a)^k/k!$ 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 we in (f) van §2.7 hebben bekeken en dat we gebruikten om een nieuwe karakerisering te geven van de afgeleide $f'(a)$. Goed beschouwd geeft (6.24) 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 (6.27) op grond van (6.31) en (6.29) 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 (wederom *Taylor-polynoom* genoemd) kunnen we in compacte vorm weergeven door de functie $x \mapsto f(a) + \sum_{k=1}^{p} f^{(k)}(a)\,(x-a,\ldots,x-a)/k!$. De coëfficiënten van het polynoom zijn nu niet langer getallen maar symmetrische multilineaire vormen (cf. Opgave 6.3). In principe zouden we (6.27) samen met (6.31) kunnen gebruiken voor een nieuwe karakterisering van de totale afgeleiden van f van hogere orde.

6.18 Behoud de aannamen van Stelling 6.17, behalve dat $f\colon E \to \mathbb{R}^m$ i.p.v. $f\colon E \to \mathbb{R}$. Dan is het nog steeds zinvol om (6.27) op te schrijven. We kunnen deze formule opvatten als de definitieformule voor $R_p(h) \in \mathbb{R}^m$. Formule (6.28) hoeft nu niet meer te gelden voor zekere $\theta \in (0,1)$. Maar (6.30) en (6.31) blijven gelden. We brengen het geval van algemene m simpelweg terug tot het geval $m = 1$ (i.e. Stelling 6.17) door gebruik van formule (6.21).

6.19 Zij $E \subset \mathbb{R}^n$ open en $f\colon E \to \mathbb{R}$ 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 $D_j\colon C^p(E) \to C^{p-1}(E)$. 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

$$D_{j_1}\ldots D_{j_k} f = D_{j_1\ldots j_k} f \quad \text{als } f \in C^p(E) \text{ en } k \le p.$$

Ook kunnen we sommen en scalaire producten van zulke lineaire afbeedlingen 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.

Het rechter lid van (6.29) 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.

Propositie Zij $E \subset \mathbb{R}^n$ open en zij $f\colon E \to \mathbb{R}$ een C^p-functie. Zij $k \leq p$. Dan geldt dat

$$\begin{aligned}\sum_{j_1,\ldots,j_k=1}^{n} h_{j_1}\ldots h_{j_k}\,D_{j_1\ldots j_k}f &= (h_1D_1+\cdots+h_nD_n)^k\,f\\ &= \sum_{k_1+\cdots+k_n=k}\frac{k!}{k_1!\ldots k_n!}\,h_1^{k_1}\ldots h_n^{k_n}\,D_1^{k_1}\ldots D_n^{k_n}\,f.\end{aligned}$$

Bewijs De eerste identiteit is triviaal. Merk nu op dat de operatoren $h_1D_1,\ldots,h_nD_n$ met elkaar commuteren. We kunnen dus de *multinomiaalformule*

$$(x_1+\cdots+x_n)^k = \sum_{k_1+\cdots+k_n=k}\frac{k!}{k_1!\ldots k_n!}\,x_1^{k_1}\ldots x_n^{k_n}$$

toepassen. (Deze formule volgt met volledige inductie naar n uit het speciale geval $n=2$, i.e. uit de binomiaalformule.) $\square$

Als een gevolg van deze Propositie kunnen we bijvoorbeeld de Taylorreeks (6.27) herschrijven als

$$f(a+h) = f(a) + \sum_{j=1}^{n} h_j\,(D_jf)(a) + \sum_{k=2}^{p-1}\;\sum_{k_1+\cdots k_n=k}\frac{h_1^{k_1}\ldots h_n^{k_n}\,(D_1^{k_1}\ldots D_n^{k_n}\,f)(a)}{k_1!\ldots k_n!} + R_p(h). \tag{6.32}$$

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:

$$\begin{aligned}f(x) &= f(a) + \sum_{j=1}^{n}(x_j-a_j)\,(D_jf)(a)\\ &+\sum_{k=2}^{p-1}\;\sum_{k_1+\cdots k_n=k}\frac{(x_1-a_1)^{k_1}\ldots(x_n-a_n)^{k_n}\,(D_1^{k_1}\ldots D_n^{k_n}\,f)(a)}{k_1!\ldots k_n!} + R_p(x-a).\end{aligned} \tag{6.33}$$

6.20 Opgave Zij $(a,b)\in\mathbb{R}^2$ en

$$f(x,y) := c_{00} + c_{10}\,(x-a) + c_{01}\,(y-b) + c_{20}\,(x-a)^2 + c_{11}\,(x-a)\,(y-b) + c_{02}\,(y-b)^2.$$

Druk de coëfficiënten c_{ij} uit in partiële afgeleiden van f in (a,b).

Verdere vraagstukken

V6.1 Geef de tweede en derde afgeleide (als multilineaire afbeelding) in het punt $(0,0,1)$ van de functie

$$f\colon \mathbb{R}^3 \to \mathbb{R}^2, \qquad f(x,y,z) = (x^2 \cos y - e^z \sin(x+y), xyz).$$

V6.2 Zij $f\colon \mathbb{R}^2 \to \mathbb{R}$ een functie en $a \in \mathbb{R}^2$. Bepaal bij f en a een polynoom $P\colon \mathbb{R}^2 \to \mathbb{R}$ van graad ≤ 2 zo dat

$$f(x) = P(x) + o(|x-a|^2) \quad \text{als } x \to a$$

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 $a := (0,0)$.

V6.3 Bepaal een polynoom $p_2(x,y)$ van graad ≤ 2 zo dat

$$\frac{1}{1-x+y} = p_2(x,y) + \mathcal{O}(|(x-1,y-1)|^3) \quad \text{als } (x,y) \to (1,1).$$

V6.4 Zij de functie $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $(x,y) \to (0,0)$.

7 Extremen

In syll. Analyse A3, §4.1 werden definities gegeven van verschillende soorten lokale maxima en minima die een reëelwaardige functie op een interval kan aannemen. Vervolgens werden in dat hoofdstuk criteria in termen van eerste of tweede afgeleide voor het aannemen van zulke extrema bewezen. Hieronder zullen we soortgelijk werk doen voor reëelwaardige functies op een deelverzameling van $\mathbb{R}^n$.

7.1 Definitie Zij (X, d) een metrische ruimte en $f \colon X \to \mathbb{R}$.

(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.

Zij (X, d) een metrische ruimte en $E \subset X$. In syll. Topologie, 2.2, 2.12 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.

7.2 Definitie Zij nu $E \subset \mathbb{R}^n$ en $f \colon E \to \mathbb{R}$. Dan is E als deelverzameling van $\mathbb{R}^n$ een metrische ruimte met metriek $d(x, y) = |x - y|$, en Definitie 7.1 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. $\mathbb{R}^n$.
Aanloog kunnen we een *lokaal inwendig minimum* en een *lokaal randminimum* definiëren.

7.3 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 reëele 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.

7.4 We gaan nu voor functies in meer veranderlijken de analoga bewijzen van Proposities 4.4 en 4.8 in syll. Analyse A3. Voor het gemak vatten we die resultaten nog eens samen.

Propositie Zij $I \subset \mathbb{R}$ een open interval, $a \in I$ en $f\colon I \to \mathbb{R}$.
(a) Zij f differentieerbaar in a. Als f een lokaal extremum aannneemt 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.

7.5 We onderzoeken nu voor functies van meer veranderlijken het verband tussen lokale extrema en eigenschappen van eerste en tweede totale afgeleide. Zoals we al een paar keer eerder gezien hebben, zullen resultaten vaak bewezen worden door herleiding tot het geval van één variabele.

Propositie Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f\colon E \to \mathbb{R}$. 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) volgt uit (2.5) en Propositie 7.4(a). Dan volgt (b) uit (a) en (2.22). □

Definitie Zij $E \subset \mathbb{R}^n$ en $f\colon E \to \mathbb{R}$. Een *stationair punt* van f is een inwendig punt a van E zo dat f differentieerbaar is in a en $f'(a) = 0$.

Opgave Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f\colon E \to \mathbb{R}$ 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.

7.6 Propositie Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f\colon E \to \mathbb{R}$ een C^2-functie. Als f een lokaal minimum (resp. maximum) aanneemt in a dan $f'(a) = 0$ en $f''(a)(h,h) \geq 0$ (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 7.5(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 $f''(a)(h,h) = F''(0) \geq 0$ op grond van formule (6.25) en Propositie 7.4(b). □

7.7 Stelling Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f\colon E \to \mathbb{R}$ een C^2-functie. Als $f'(a) = 0$ en $f''(a)(h,h) > 0$ (resp. < 0) voor alle $h \in \mathbb{R}^n \backslash \{0\}$ dan neemt f een sterk lokaal minimum (resp. maximum) aan in a.

Bewijs
$$\inf_{h \in \mathbb{R}^n \backslash \{0\}} \frac{f''(a)(h,h)}{|h|^2} = \inf_{h \in \mathbb{R}^n;\ |h|=1} f''(a)(h,h) = f''(a)(h^0,h^0) = \mu > 0$$

voor zekere $h^0 \in \mathbb{R}^n$ met $|h^0| = 1$ en voor zekere positieve μ, omdat de functie $h \mapsto f''(a)(h,h)$ 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 $f''(a)(h,h) \geq \mu |h|^2$ voor alle $h \in \mathbb{R}^n$. Toepassing van (6.27) en (6.31) geeft dat $f(a+h) - f(a) = \frac{1}{2} f''(a)(h,h) + o(|h|^2)$ als $h \to 0$ in $\mathbb{R}^n$. Dus er bestaat $\delta > 0$ zo dat voor $0 < |h| < \delta$ geldt dat

$$f(a+h) - f(a) \geq \tfrac{1}{2} f''(a)(h,h) - \tfrac{1}{4}\mu\,|h|^2 \geq \tfrac{1}{4}\mu\,|h|^2 > 0. \qquad \square$$

7.8 Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f\colon E \to \mathbb{R}$ een C^2-functie. We hebben zojuist gezien dat de eigenschappen van de kwadratische vorm $x \mapsto f''(a)(x,x)\colon \mathbb{R}^n \to \mathbb{R}$ een direct verband hebben met extremaaleigenschappen van f in a. We weten uit (6.16) dat

$$f''(a)(x,x) = \sum_{i,j=1}^{n} (D_{ij} f)(a)\, x_i\, x_j, \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n. \tag{7.1}$$

Deze kwadratische vorm wordt dus bepaald door de symmetrische matrix

$$H_f(a) := \begin{pmatrix} (D_{11}f)(a) & \ldots & (D_{1n}f)(a) \\ \vdots & & \vdots \\ (D_{n1}f)(a) & \ldots & (D_{nn}f)(a) \end{pmatrix}, \tag{7.2}$$

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 $\mathbb{R}^m$.) We geven nu een paar algemene definities en eigenschappen betreffende kwadratische vormen en passen die vervolgens toe op de kwadratische vorm uit (7.1).

7.9 Zij V een reële vectorruimte. In §6.1 definieerden we een kwadratische vorm op V als een afbeelding van de vorm $x \mapsto B(x,x)\colon V \to \mathbb{R}$, waarbij $B \in \Lambda_2(V)$, dus een bilineaire vorm op $V \times V$. We nemen aan dat de bilineaire vorm B symmetrisch is. (Er kan bewezen worden dat er bij een gegeven kwadratische vorm $x \mapsto B(x,x)$ een unieke symmetrische bilineaire vorm S is zo dat $B(x,x) = S(x,x)$, nl. $S(x,y) := \frac{1}{2}(B(x,y) + B(y,x))$.)

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 $B(y,y) < 0$.

Zij $A = (A_{ij})_{i,j=1,\dots,n}$ een symmetrische $n \times n$ matrix (dus $A_{ij} = A_{ji}$). Dan bepaalt A een symmetrische bilineaire vorm B op $\mathbb{R}^n \times \mathbb{R}^n$ (cf. (6.1)) en een bijbehorende kwadratische vorm op $\mathbb{R}^n$ gegeven door

$$B(x,x) = \langle Ax, x\rangle = \sum_{i,j=1}^{n} A_{ij}\, x_i\, x_j, \quad x \in \mathbb{R}^n. \tag{7.3}$$

We noemen de symmetrische matrix A *positief* of *negatief (semi)definiet* als de kwadratische vorm (7.3) de gelijknamige eigenschap heeft. Omdat de matrix A symmetrisch is, is er een orthonormale basis $f_1, \dots, f_n$ 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 $\widetilde{x}_1, \dots, \widetilde{x}_n$ t.o.v. deze basis hebben, dus $x = \sum_{j=1}^{n} \widetilde{x}_j\, f_j$. Dan volgt er uit (7.3) dat

$$B(x,x) = \sum_{j=1}^{n} \lambda_j \widetilde{x}_j^2.$$

Laat nu

$$\lambda_{\min} := \min_{j=1,\dots,n} \lambda_j, \quad \lambda_{\max} := \max_{j=1,\dots,n} \lambda_j.$$

Dan

$$\lambda_{\min} \sum_{j=1}^{n} \widetilde{x}_j^2 \le \sum_{j=1}^{n} \lambda_j\, \widetilde{x}_j^2 \le \lambda_{\max} \sum_{j=1}^{n} \widetilde{x}_j^2,$$

dus

$$\lambda_{\min}\, |x|^2 \le B(x,x) \le \lambda_{\max}\, |x|^2. \tag{7.4}$$

Bovendien zijn de ongelijkheden in (7.4) 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:

$$\begin{aligned} A \text{ is positief semidefiniet} &\iff \lambda_{\min} \ge 0, \\ A \text{ is positief definiet} &\iff \lambda_{\min} > 0, \end{aligned}$$

en analoog voor negatief (semi)definiet.

We kunnen nu Propositie 7.6 en Stelling 7.7 als volgt herformuleren in het geval van een minimum (het geval van een maximum is analoog).

7.10 Stelling Zij $E \subset \mathbb{R}^n$ open, $a \in E$ en $f\colon E \to \mathbb{R}$ 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.

Als het geval (c) van Stelling 7.10 zich voordoet en a bovendien een stationair punt is (dus als $f'(a)=0$ en de matrix $H_f(a)$ indefiniet is) dan noemen we a een *zadelpunt* van f. In zo'n geval bestaat er 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.

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 $(0,0)$.

7.11 Bij toepassingen van Stelling 7.10(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

$$\begin{array}{rcccc} \operatorname{tr} A & = & a+d & = & \lambda+\mu, \\ \det A & = & ad-b^2 & = & \lambda\mu. \end{array}$$

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 7.10 geeft:

Propositie Zij $E \subset \mathbb{R}^2$ open, $a \in E$ en $f\colon E \to \mathbb{R}$ 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.

7.12 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 $E \subset \mathbb{R}^2$. 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 7.10 of Propositie 7.11 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 7.4 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

V7.1 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) $f(x,y) := (y^2-1)\,(x^2-y^2)$.

V7.2 Zij $E \subset \mathbb{R}^2$. Bepaal voor elk van de volgende functies $f\colon E \to \mathbb{R}$ 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) $f(x,y) := 2x^4 - 3x^2y + y^2$, $E := \{(x,y) \mid 4x^2 + y \le 4\}$;
b) $f(x,y) := x^2 - x + 2y^2$, $E := \{(x,y) \mid x^2 + y^2 \le 1\}$;
c) $f(x,y) := (x^2 - y^2)\,e^{-x^2-y^2}$, $E := \{(x,y) \mid x^2 + y^2 \le 4\}$;
d) $f(x,y) := (x^2-1)\,(y^2-1)\,(x^2+y^2-1)$, $E := \{(x,y) \mid x^2 + y^2 \le 2\}$;
e) $f(x,y) := x^5 - (x^2 + x^3)\,y + y^2$, $E := \{(x,y) \mid -1 \le x \le 1,\ 0 \le y \le 1\}$;
f) $f(x,y) := x^4 + 9y^4$, $E := \{(x,y) \mid x^2 + 4y^2 \le 1\}$;
g) $f(x,y) := 2x^2 - 3y^2 - 2x + \frac{1}{2}$, $E := \{(x,y) \mid x^2 + y^2 \le 1\}$;
h) $f(x,y) := x^4 - x^2y^2 - 2x^2 + 2y^2$, $E := \{(x,y) \mid x^2 + y^2 \le 4\}$;
i) $f(x,y) := x^2y^2 + 2xy^2 + y^4$, $E := \{(x,y) \mid |x| \le 2,\ |y| \le 2\}$.

8 Maple-commando's

In dit hoofdstuk beschrijven we een aantal routines in Maple V, Release 3, die betrekking hebben op de voorgaande theorie. Voor basistechnieken in Maple verwijzen we naar het onderdeel Maple in het eerstejaars-vak Inleiding Computergebruik.

8.1 De commando's met grafische output kunnen het best uitgeprobeerd worden door op een Sun-werkstation in de Unix-shell het volgende commando te geven: `xmaple &`

Het plotten van 3-dimensionale grafieken kan gedaan worden met het commando plot3d (expr, x=a..b, y=c..d). Help-informatie hierover is te verkrijgen door `?plot3d` in te typen. (Op een soortgelijke manier kan men ook hulp vragen over andere commando's.) Je kunt bijvoorbeeld intypen

```
plot3d(x^2+y^2, x=-10..10, y=-10..10);
```

Er verschijnt dan een nieuw venster (3D window) waarin de grafiek van de functie $(x, y) \mapsto x^2 + y^2 \colon [-10, 10] \times [-10, 10] \to \mathbb{R}$ getekend wordt. Onder het menu *Axes* kun je *Framed* kiezen. Als je dan òf in het 3D window de letter `p` typt òf op de middelste muisknop drukt (om aan te geven dat de figuur opnieuw getekend moet worden), dan verschijnen er ook coördinaatassen. Meer informatie over het 3D window kun je krijgen via het menu *Help* rechtsboven. Je kunt het plaatje saven als postscript file door in het menu *File* onder *Print* voor *Postscript* te kiezen. Je kunt dan nog een naam opgeven voor de te vormen postscript-file (bijv. `plot1.ps`). Tenslotte kun je het plaatje dan printen door in de Unix-shell het commando `lpr plot1.ps` te geven. Wees echter terughoudend met printen! Bekijk de plaatjes zoveel mogelijk op het scherm.

Je zou nu ook de grafieken kunnen plotten van de functies uit Opgave 4.12. Experimenteer ook met de verschillende menu-opties in het 3D window.

8.2 De gradiënt van een functie kun je Maple laten berekenen met het commando grad(expr, variablelist). Eerst moet je echter het *linalg* package binnen Maple aanroepen door het commando `with(linalg):` in te typen. Vervolgens kun je bijv. het commando

```
grad(x^2-y^2, [x,y]);
```

intypen. Je kunt ook direct het gradiëntvectorveld laten tekenen. Roep daartoe eerst het *plots* package aan door `with(plots):` in te typen. Je zou dan bijv. kunnen intypen

```
gradplot(x^2-y^2, x=-10..10, y=-10..10);
```

en in een nieuw venster (Maple V 2D) verschijnt het gradiëntvectorveld.

8.3 Partiële differentiatie wordt met hetzelfde commando `diff` aangegeven als gewone differentiatie. Bijvoorbeeld:

$$
\begin{array}{rll}
\texttt{diff(sin(x),x);} & \text{bepaalt} & \dfrac{d \sin x}{dx} \\[2ex]
\texttt{diff(sin(x*y*z),x);} & \text{bepaalt} & \dfrac{\partial \sin(xyz)}{\partial x} \\[2ex]
\texttt{diff(sin(x*y*z),x,y);} & \text{bepaalt} & \dfrac{\partial^2 \sin(xyz)}{\partial x\, \partial y} \\[2ex]
\texttt{diff(sin(x*y*z),x\$2,z);} & \text{bepaalt} & \dfrac{\partial^3 \sin(xyz)}{\partial^2 x\, \partial z}
\end{array}
$$

Soms kan men output vereenvoudigen door direct erna het volgende commando te geven:

```
simplify(");
```

Terwijl `diff` op algebraïsche uitdrukkingen werkt, voert de operator `D` in Maple differentiatie van functies uit. Definieer bijvoorbeeld

```
f := x->sin(x);   g:=(x,y,z) -> sqrt(x^2+y^2+z^2);
```

Dan geldt bijv. het volgende:

`D(f);`	bepaalt	$x \mapsto f'(x)$
`D(f)(a);`	bepaalt	$f'(a)$
`D[1](g);`	bepaalt	$(x, y, z) \mapsto (D_1 g)(x, y, z)$.
`D[2,2](g)(u,v,w);`	bepaalt	$(D_{22} g)(u, v, w)$.
`D[1,2,3](g)(2,5,3);`	bepaalt	$(D_{123} g)(2, 5, 3)$

Probeer, met g als boven, nu eens het volgende:

```
sum(D[i,i](g)(x,y,z),i=1..3);
simplify(");
```

We hebben hier dus de Laplace-operator op g laten werken. Deze operator is ook kant en klaar beschikbaar in Maple, maar net zo als `diff`, werkend op algebraïsche uitdrukkingen en niet op functies. Bijvoorbeeld

```
with(linalg):
laplacian(sqrt(x^2+y^2+z^2),[x,y,z]);
```

bepaalt

$$\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\Big)\Big\{\sqrt{x^2 + y^2 + z^2}\Big\}$$

8.4 Zoals bekend kun je Maple V een Taylorreeks van een functie van één variabele laten uitrekenen met de routine `taylor` . Bijvoorbeeld `taylor(exp(x), x);` geeft $p_5(x) + \mathcal{O}(x^6)$, waarbij $p_5(x)$ een expliciet Taylorpolynoom van graad ≤ 5 is. We kunnen de orde van benadering zelf aangeven door bijv. `taylor(exp(x), x, 10);` Dit levert $p_9(x) + \mathcal{O}(x^{10})$, waarbij $p_9(x)$ een expliciet Taylorpolynoom is van graad ≤ 9. Taylorontwikkeling rond een ander punt kunnen we aangeven door bijv.

`taylor(exp(x), x=1);` of `taylor(exp(x), x=1, 10);`

wat $p_{n-1}(x-1) + \mathcal{O}((x-1)^n)$ levert met $n = 6$ resp. 10.

Taylorreeksen van functies in meer veranderlijken worden gegeven door de routine `mtaylor` . Hiertoe moet de procedure eerst ingelezen worden met het commando

```
readlib(mtaylor):
```

Nu levert bijv.

```
mtaylor(exp(x+y^2), [x,y], 10);
```

een Taylorpolynoom $p_9(x, y)$ zo dat $\exp(x + y^2) - p_9(x, y) = \mathcal{O}(|(x, y)|^{10})$. Deze resttermorde geeft `mtaylor` er echter niet bij in de output, dit in tegenstelling tot `taylor` . Als we het derde argument in `mtaylor` weglaten, dus geen gewenste orde van de restterm opgeven, dan wordt deze orde op 6 gesteld. We kunnen ook om andere punten dan $(0, 0)$ ontwikkelen door bijv.

`mtaylor(exp(x+y^2), [x=1,y]);` of `mtaylor(exp(x+y^2), [x=2,y=3]);` waarbij we ontwikkelen om $(1,0)$ resp. $(2,3)$.

Index