Hoofdstuk 2

Impliciete functies

2.1 Jacobiaan van een afbeelding

Een $\mathbb{R}^n$-$\mathbb{R}^m$ vectorwaardige functie $\mathbf{f} = (f_1, f_2, \ldots, f_m)$ is afleidbaar in een inwendig punt $\mathbf{a}$ van haar domein juist als elke component afleidbaar is in $\mathbf{a}$, nl. als voor elke $i = 1, 2, \ldots, m$ en voor elke $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$ met $\|\mathbf{h}\|$ voldoende klein

$$f_i(\mathbf{a}+\mathbf{h}) = f_i(\mathbf{a}) + \frac{\partial f_i}{\partial x_1}(\mathbf{a}) \cdot h_1 + \cdots + \frac{\partial f_i}{\partial x_n}(\mathbf{a}) \cdot h_n + \|\mathbf{h}\| r_i(\mathbf{h})$$

met $\lim_{\mathbf{h}\to\mathbf{0}} r_i(\mathbf{h}) = 0$. Met behulp van het matrixproduct kunnen we dit herschrijven als

$$\begin{pmatrix} f_1(\mathbf{a}+\mathbf{h}) \\ \vdots \\ f_m(\mathbf{a}+\mathbf{h}) \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{a}) \\ \vdots \\ f_m(\mathbf{a}) \end{pmatrix} + \underbrace{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}}_{=:\, D\mathbf{f}(\mathbf{a})} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \|\mathbf{h}\| \begin{pmatrix} r_1(\mathbf{h}) \\ \vdots \\ r_m(\mathbf{h}) \end{pmatrix}$$

met $r_i\colon \mathbb{R}^n \to \mathbb{R}$, $\lim_{\mathbf{h}\to\mathbf{0}} r_i(\mathbf{h}) = 0$ $(i = 1, \ldots, m)$.

2.1.1 Definitie. Als de functie $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ afleidbaar is in $\mathbf{a}$, dan wordt de matrix $D\mathbf{f}(\mathbf{a})$ de **Jacobiaanse matrix** (of **totale afgeleide**) van $\mathbf{f}$ in $\mathbf{a}$ genoemd. Als $m = n$, dan wordt de determinant $\det(D\mathbf{f}(\mathbf{a}))$ de **Jacobiaanse determinant** van $\mathbf{f}$ genoemd.

2.1.2 Opmerkingen.

1. We zullen ook $\det(D\mathbf{f}(\mathbf{a}))$ als $\dfrac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\mathbf{a})$ noteren.[1]

2. Als $f\colon \mathbb{R}^n \to \mathbb{R}$ afleidbaar is in $\mathbf{a}$, dan is haar Jacobiaanse matrix een rijvector, gegeven door haar gradiënt in $\mathbf{a}$, nl. $Df(\mathbf{a}) = (\nabla f)(\mathbf{a}) = (\frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a}))$ (zie 1.8.3).

[1] zie §4.2 en 6.3.1

We kunnen de afleidbaarheid van $\mathbf{f}$ in $\mathbf{a}$ hiermee beknopt vectorieel noteren als[2]

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a}) \cdot \mathbf{h} + \|\mathbf{h}\|\, \mathbf{r}(\mathbf{h}),$$

met $\mathbf{r}\colon \mathbb{R}^n \to \mathbb{R}^m$, $\lim_{\mathbf{h}\to\mathbf{0}} r(\mathbf{h}) = 0$.

2.1.3 Opmerking. $D\mathbf{f}(\mathbf{a}) \in \mathbb{R}^{m\times n}$ is de matrix-voorstelling van de lineaire benadering van $\mathbf{f}$: $\mathbb{R}^n \to \mathbb{R}^m$ in de omgeving van $\mathbf{a}$, als we zoals vroeger de oorsprong in $\mathbb{R}^n$ verplaatsen naar het punt $\mathbf{a}$ en de oorsprong in $\mathbb{R}^m$ naar het punt $\mathbf{f}(\mathbf{a})$. We noemen deze lineaire afbeelding de **Jacobiaan-afbeelding** van $\mathbf{f}$ in $\mathbf{a}$ en blijven deze met $D\mathbf{f}(\mathbf{a})$ noteren. Meetkundig gezien betekent dit opnieuw dat de raakruimte aan de grafiek van $\mathbf{f}$ in het punt $(\mathbf{a}, \mathbf{f}(\mathbf{a})) \in \mathbb{R}^{n+m}$ bestaat: die raakruimte is juist de grafiek van de lineaire afbeelding $D\mathbf{f}(\mathbf{a})$.

Dat dit een conceptuele manier is om afleidbaarheid te benaderen, blijkt ondermeer uit de kettingregel: als $\mathbf{f} = (f_1, \ldots, f_m)\colon \mathbb{R}^n \to \mathbb{R}^m$ afleidbaar is in $\mathbf{a}$ en $\mathbf{g} = (g_1, \ldots, g_k)\colon \mathbb{R}^m \to \mathbb{R}^k$ afleidbaar in $\mathbf{f}(\mathbf{a})$, dan is volgens de kettingregel in meerdere veranderlijken

$$\frac{\partial((\mathbf{g}\circ\mathbf{f})_i)}{\partial x_j}(\mathbf{a}) = \frac{\partial(g_i\circ\mathbf{f})}{\partial x_j}(\mathbf{a}) = \sum_{q=1}^{m} \frac{\partial g_i}{\partial y_q}(\mathbf{f}(\mathbf{a})) \cdot \frac{\partial f_q}{\partial x_j}(\mathbf{a}), \quad (i = 1, \ldots, k)$$

of, in matrixnotatie,

$$\begin{pmatrix} \dfrac{\partial((\mathbf{g}\circ\mathbf{f})_1)}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial((\mathbf{g}\circ\mathbf{f})_1)}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial((\mathbf{g}\circ\mathbf{f})_k)}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial((\mathbf{g}\circ\mathbf{f})_k)}{\partial x_n}(\mathbf{a}) \end{pmatrix} =$$

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial y_1}(\mathbf{f}(\mathbf{a})) & \cdots & \dfrac{\partial g_1}{\partial y_m}(\mathbf{f}(\mathbf{a})) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_k}{\partial y_1}(\mathbf{f}(\mathbf{a})) & \cdots & \dfrac{\partial g_k}{\partial y_m}(\mathbf{f}(\mathbf{a})) \end{pmatrix} \cdot \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}$$

m.a.w.,

$$\boxed{D(\mathbf{g}\circ\mathbf{f})(a) = D\mathbf{g}(\mathbf{f}(\mathbf{a})) \cdot D\mathbf{f}(\mathbf{a})}$$

zodat 'de beste lineaire benadering van de samenstelling gelijk is aan de samenstelling[3] van de beste lineaire benaderingen', wat we ook zonder formeel bewijs hadden durven vermoeden.

2.1.4 Oefening. *Als* $\mathbf{f}$ *zelf een lineaire afbeelding is* $\mathbb{R}^n$-$\mathbb{R}^m$ *met matrixvoorstelling* A*, dan is* $D\mathbf{f}(\mathbf{a}) = A$ *in elk punt* $\mathbf{a} \in \mathbb{R}^n$*.*

Het resultaat uit de vorige oefening hoeft niet te verbazen: het drukt uit dat de beste lineaire benadering van een lineaire afbeelding juist die afbeelding zelf is.

[2] het product in het rechterlid is het matrixproduct van $D\mathbf{f}(\mathbf{a})$ met de kolomvector $\mathbf{h}$

[3] het matrixproduct komt immers overeen met de samenstelling van de corresponderende lineaire afbeeldingen

2.2 Impliciete functies

Als een gegeven vlakke kromme Γ bepaald is door een vergelijking Γ: $f(x,y)=0$, dan kan het gebeuren dat we deze vergelijking *expliciet* kunnen oplossen naar y, en zo een functie $y(x)$ kunnen definiëren waarvan de grafiek juist samenvalt met de kromme Γ (m.a.w., waarvoor $f(x,y(x))=0$). We zeggen dan dat de functie $y(x)$ *impliciet* bepaald is door de vergelijking $f(x,y)=0$.

2.2.1 Voorbeeld. Γ: $x^2+y^2=1$ bepaalt een cirkel met middelpunt 0 en straal 1. We kunnen hier de vergelijking expliciet oplossen naar y, en vinden $y(x)=\pm\sqrt{1-x^2}$.

Zoals uit het voorbeeld blijkt, zullen we er rekening moeten mee houden dat zo'n impliciete vergelijking niet steeds een eenduidige expliciete functie bepaalt.
Voor minder eenvoudige vergelijkingen $f(x,y)=0$ is het meestal moeilijk (of zelfs onmogelijk) om $y(x)$ expliciet op te lossen in termen van elementaire functies (zelfs als f een elementaire functie is die continu afleidbaar is), zoals bijv. Γ: $\sin y = xy$. Het is daardoor ook moeilijker te zien in welk(e) interval(len) we de vergelijking kunnen oplossen naar y. Toch kunnen we zonder $y(x)$ expliciet te bepalen al informatie krijgen over de afgeleide $y'(x)$:

2.2.2 Voorbeeld. Als $x^2+(y(x))^2=1$ op een open interval $I\subseteq\mathbb{R}$, dan kunnen we beide leden van de gelijkheid afleiden op I. We vinden dat $2x+2y(x)y'(x)=0$ op I, zodat $y'(x)=-x/y(x)$ als $y(x)\neq 0$.

In het algemeen vinden we voor een kromme Γ: $f(x,y)=0$ door impliciet afleiden dat een expliciete functie $y(x)$ voldoet aan (kettingregel in meerdere veranderlijken)

$$\frac{\partial f}{\partial x}(x,y(x))+\frac{\partial f}{\partial y}(x,y(x))y'(x)=0. \tag{2.1}$$

We hebben hierbij wel verondersteld:
(i) een expliciete functie $y(x)$ bestaat op een gepast open interval $I\subseteq\mathbb{R}$
(ii) $y(x)$ is afleidbaar op I.
We zullen een voldoende voorwaarde geven waaronder een (eenduidige) expliciete functie $y(x)$ bestaat in een voldoende kleine omgeving van een gegeven punt $(a,b)\in\Gamma$.

2.2.3 Voorbeeld. De kromme Γ: $x^2+y^2=1$ bepaalt een eenduidige oplossing $y(x)$ in een omgeving van elk punt (a,b) op Γ waarvoor $-1<a<1$, maar geen eenduidige oplossing in een omgeving (hoe klein ook) van $(\pm 1,0)$.

Meetkundig gezien verwachten we een eenduidige oplossing $y(x)$ in een omgeving van een punt $(x,y(x))\in\Gamma$ op voorwaarde dat de raaklijn aan Γ in $(x,y(x))$ niet verticaal is. In de vergelijking (2.1) is $y'(x)$ eenduidig bepaald zodra $\frac{\partial f}{\partial y}(x,y(x))\neq 0$.[4] Dit zal inderdaad een voldoende voorwaarde blijken:

2.2.4 Stelling (Stelling van de impliciete functies in het vlak). *Zij f een $\mathbb{R}^2$-$\mathbb{R}$ functie die van de klasse C^1 is in een omgeving van $(a,b)\in\mathbb{R}^2$. Als $f(a,b)=0$ en $\frac{\partial f}{\partial y}(a,b)\neq 0$, dan geldt:*

[4] terwijl het geval $\frac{\partial f}{\partial y}(x,y(x))=0$ intuïtief overeenkomt met $y'(x)=\pm\infty$, d.w.z., een verticale raaklijn (de raaklijn aan Γ in $(x_0,y(x_0))$ heeft idd. als vergelijking $\frac{\partial f}{\partial x}(x_0,y(x_0))(x-x_0)+\frac{\partial f}{\partial y}(x_0,y(x_0))(y-y(x_0))=0$)

1. *Er bestaan $\delta > 0$ en $\varepsilon > 0$ zo dat voor elke $x \in]a-\delta, a+\delta[$ een unieke $y(x) \in]b-\varepsilon, b+\varepsilon[$ bestaat waarvoor*
$$f(x, y(x)) = 0.$$
2. *De afbeelding y: $]a-\delta, a+\delta[\to \mathbb{R}$ is van de klasse C^1.*
3. *Als $f(x,y)$ bovendien van de klasse C^k ($k > 1$) is in een omgeving van (a,b), dan is ook $y(x)$ van de klasse C^k in een omgeving van a.*

Bewijs. We mogen aannemen dat $\frac{\partial f}{\partial y}(a,b) > 0$ (het geval $\frac{\partial f}{\partial y}(a,b) < 0$ volgt analoog).

1. Omdat $\frac{\partial f}{\partial y}$ continu is, volgt door behoud van teken dat ook $\frac{\partial f}{\partial y}(x,y) > 0$ in een zekere rechthoek $[a-\varepsilon, a+\varepsilon] \times [b-\varepsilon, b+\varepsilon]$. Daardoor is de functie $y \mapsto f(a,y)$ strikt stijgend in $[b-\varepsilon, b+\varepsilon]$. Omdat $f(a,b) = 0$ is dus $f(a, b-\varepsilon) < 0$ en $f(a, b+\varepsilon) > 0$. Omdat f continu is, volgt (opnieuw door behoud van teken) dat $f(x, b-\varepsilon) < 0$ en $f(x, b+\varepsilon) > 0$ voor elke x in een zeker interval $]a-\delta, a+\delta[$.
Kies nu willekeurig $x \in]a-\delta, a+\delta[$. Dan is ook de functie $y \mapsto f(x,y)$ strikt stijgend op $[b-\varepsilon, b+\varepsilon]$ (verklein desnoods δ zo dat $\delta \leqslant \varepsilon$). Door de tussenwaardestelling heeft deze functie een (uniek, omdat ze strikt stijgend is) nulpunt $y(x) \in]b-\varepsilon, b+\varepsilon[$.

2. I.h.b. is $y(a) = b$. In feite kunnen we in deel 1. $\varepsilon > 0$ zo veel verkleinen als we willen. We vinden dus voor elke (voldoend kleine) $\varepsilon > 0$ een $\delta > 0$ zo dat voor elke $x \in]a-\delta, a+\delta[$ een unieke $y(x)$ bestaat met $|y(x) - b| < \varepsilon$ waarvoor $f(x, y(x)) = 0$. De functie $y(x)$ is dus continu in a. Wat we juist bewezen hebben, kunnen we ook toepassen in een willekeurig punt $(x, y(x))$ in de plaats van het punt (a,b): f voldoet ook in dit punt aan de voorwaarden van de stelling. We besluiten dat y continu is in elk punt $x \in]a-\delta, a+\delta[$.[5]
Door de middelwaardestelling is
$$f(x+h, y+k) = f(x,y) + \frac{\partial f}{\partial x}(x+\theta h, y+\theta k) \cdot h + \frac{\partial f}{\partial y}(x+\theta h, y+\theta k) \cdot k$$
voor zekere $\theta \in [0,1]$ (afhankelijk van h, k). Kiezen we i.h.b. (x,y) en $(x+h, y+k)$ op de kromme met vergelijking $f(x,y) = 0$, m.a.w. kiezen we $y := y(x)$ en $k := y(x+h) - y(x)$, dan vinden we
$$\frac{y(x+h) - y(x)}{h} = \frac{k}{h} = -\frac{\dfrac{\partial f}{\partial x}(x+\theta h, y+\theta k)}{\dfrac{\partial f}{\partial y}(x+\theta h, y+\theta k)}. \tag{2.2}$$
Omdat $y(x)$ continu is, is dus ook $|\theta k| \leqslant |k| = |y(x+h) - y(x)| \to 0$ als $h \to 0$. Uit (2.2) volgt dus dat
$$y'(x) = \lim_{h \to 0} \frac{y(x+h) - y(x)}{h} = -\frac{\dfrac{\partial f}{\partial x}(x, y(x))}{\dfrac{\partial f}{\partial y}(x, y(x))} \tag{2.3}$$
omdat ook $\frac{\partial f}{\partial x}$ en $\frac{\partial f}{\partial y}$ continu zijn. Uit de gelijkheid (2.3) volgt dat $y(x)$ van de klasse C^1 is op een omgeving van a.

[5] Tot hier in het bewijs hoeven we niet te vragen dat f van de klasse C^1 is: het volstaat dat f en $\frac{\partial f}{\partial y}$ bestaan en continu zijn op een omgeving van (a,b).

3. Per inductie: als f van klasse C^2 is, dan zijn $\frac{\partial f}{\partial x}$ en $\frac{\partial f}{\partial y}$ van klasse C^1, en bewezen we al dat $y(x)$ van klasse C^1 is. Door (2.3) is dan ook $y'(x)$ van klasse C^1, m.a.w. $y(x)$ is van klasse C^2, enz. $\square$

Meer algemeen vragen we ons af wanneer we in een stelsel vergelijkingen, bijv.

$$\begin{cases} f_1(x,y,z) = 0 \\ f_2(x,y,z) = 0 \end{cases}$$

twee van de veranderlijken (lokaal) kunnen oplossen als functie van de derde veranderlijke, laat ons zeggen y en z als functie van x. (Meetkundig gezien bepaalt het stelsel nu een kromme in de drie-dimensionale ruimte, en we vragen ons af wanneer de veranderlijke x de parameter in een parametervoorstelling van de kromme kan zijn.)
Laat ons voor het gemak aannemen dat de oorsprong $\mathbf{0} = (0,0,0)$ op de kromme ligt. Omdat we aannemen dat f_1 en f_2 continu afleidbaar zijn, en afleidbaarheid betekent dat f_1 en f_2 goed benaderd kunnen worden door lineaire functies, bekijken we eerst het stelsel van de lineaire benaderingen in $\mathbf{0}$

$$\begin{cases} \dfrac{\partial f_1}{\partial x}(\mathbf{0}) \cdot x + \dfrac{\partial f_1}{\partial y}(\mathbf{0}) \cdot y + \dfrac{\partial f_1}{\partial z}(\mathbf{0}) \cdot z = 0 \\ \\ \dfrac{\partial f_2}{\partial x}(\mathbf{0}) \cdot x + \dfrac{\partial f_2}{\partial y}(\mathbf{0}) \cdot y + \dfrac{\partial f_2}{\partial z}(\mathbf{0}) \cdot z = 0 \end{cases}$$

Bij een gegeven vaste $x \in \mathbb{R}$ moeten we dus een lineair stelsel in y, z oplossen. Een voldoende voorwaarde dat dit stelsel een unieke oplossing heeft is dat

$$\det \begin{pmatrix} \dfrac{\partial f_1}{\partial y}(\mathbf{0}) & \dfrac{\partial f_1}{\partial z}(\mathbf{0}) \\ \\ \dfrac{\partial f_2}{\partial y}(\mathbf{0}) & \dfrac{\partial f_2}{\partial z}(\mathbf{0}) \end{pmatrix} \neq 0.$$

De algemene vorm van de stelling van de impliciete functies zegt dat dit ook een voldoende voorwaarde is voor het (uniek) oplosbaar zijn van het oorspronkelijk (niet-gelineariseerd) stelsel in functie van x in een omgeving van $\mathbf{0}$.
In het algemeen, voor een stelsel van m vergelijkingen in $n+m$ veranderlijken (dat we willen oplossen in functie van de eerste n veranderlijken) noteren we de optredende functies $f_1, \ldots, f_m$ vectorieel als $\mathbf{f} = (f_1, \ldots, f_m)$ en noteren we de veranderlijken als $(\mathbf{x}, \mathbf{y}) = (x_1, \ldots, x_n, y_1, \ldots, y_m)$. Het stelsel

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n, y_1, y_2 \ldots, y_m) = 0 \\ \\ f_2(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0 \\ \qquad\qquad \vdots \\ f_m(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0 \end{cases}$$

is dan

$$\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{0}.$$

We noteren d.m.v. $D_1\mathbf{f}(\mathbf{x},\mathbf{y}) \in \mathbb{R}^{m\times n}$, resp. $D_2\mathbf{f}(\mathbf{x},\mathbf{y}) \in \mathbb{R}^{m\times m}$, de Jacobiaanse matrix van de functie $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x},\mathbf{y})$ (bij vaste $\mathbf{y}$), resp. $\mathbf{y} \mapsto \mathbf{f}(\mathbf{x},\mathbf{y})$ (bij vaste $\mathbf{x}$), m.a.w.

$$D_1\mathbf{f}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{x},\mathbf{y}) & \dots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{x},\mathbf{y}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\mathbf{x},\mathbf{y}) & \dots & \dfrac{\partial f_m}{\partial x_n}(\mathbf{x},\mathbf{y}) \end{pmatrix}$$

en

$$D_2\mathbf{f}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1}(\mathbf{x},\mathbf{y}) & \dots & \dfrac{\partial f_1}{\partial y_m}(\mathbf{x},\mathbf{y}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial y_1}(\mathbf{x},\mathbf{y}) & \dots & \dfrac{\partial f_m}{\partial y_m}(\mathbf{x},\mathbf{y}) \end{pmatrix}.$$

2.2.5 Stelling (Stelling van de impliciete functies). *Zij* $\mathbf{f}$ *een functie* $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ *van de klasse* C^k *(*$k \geqslant 1$*) in een omgeving van* $(\mathbf{a},\mathbf{b}) \in \mathbb{R}^n \times \mathbb{R}^m$*. Als* $\mathbf{f}(\mathbf{a},\mathbf{b}) = \mathbf{0}$ *en* $\det(D_2\mathbf{f}(\mathbf{a},\mathbf{b})) \neq 0$*, dan geldt:*

1. *Er bestaan* $\delta > 0$ *en* $\varepsilon > 0$ *zo dat voor elke* $\mathbf{x} \in B(\mathbf{a},\delta)$ *een unieke* $\mathbf{y}(\mathbf{x}) \in B(\mathbf{b},\varepsilon)$ *bestaat waarvoor*
$$\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x})) = \mathbf{0}.$$
2. *De afbeelding* $\mathbf{y}\colon B(\mathbf{a},\delta) \to \mathbb{R}^m$ *is van de klasse* C^k*.*

De stelling van de impliciete functies speelt een belangrijke rol bij het ontwikkelen van de theorie van deelvariëteiten van $\mathbb{R}^n$ (cursus *Differentiaalmeetkunde*).
De partiële afgeleiden van $\mathbf{y}(\mathbf{x})$ vinden we opnieuw door de identiteit $\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x})) = \mathbf{0}$ impliciet af te leiden. Stellen we $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^{n+m}\colon \mathbf{g}(\mathbf{x}) = (\mathbf{x},\mathbf{y}(\mathbf{x}))$, dan is

$$D\mathbf{f}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} D_1\mathbf{f}(\mathbf{x},\mathbf{y}) & D_2\mathbf{f}(\mathbf{x},\mathbf{y}) \end{pmatrix} \in \mathbb{R}^{m\times(n+m)} \quad \text{en} \quad D\mathbf{g}(\mathbf{x}) = \begin{pmatrix} I_{n\times n} \\ D\mathbf{y}(\mathbf{x}) \end{pmatrix} \in \mathbb{R}^{(n+m)\times n}$$

(met $I_{n\times n} \in \mathbb{R}^{n\times n}$ de eenheidsmatrix) zodat (kettingregel)

$$D_1\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x})) + D_2\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x})) \cdot D\mathbf{y}(\mathbf{x}) = \mathbf{0}.$$

Omdat $\det(D_2\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x}))) \neq 0$ in een omgeving van $(\mathbf{a},\mathbf{b})$, besluiten we dat

$$\boxed{D\mathbf{y}(\mathbf{x}) = -(D_2\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x})))^{-1} \cdot D_1\mathbf{f}(\mathbf{x},\mathbf{y}(\mathbf{x}))}.$$

Opnieuw geeft de notatie met afgeleiden als matrices/lineaire afbeeldingen een beknopte schrijfwijze die de analogie met het geval $m = n = 1$ duidelijk weergeeft.

2.3 Stelling van de inverse functies

Als toepassing van de stelling van de impliciete functies geven we een meerdimensionale veralgemening van de stelling i.v.m. afleidbaarheid van de inverse van een (inverteerbare) afleidbare functie $f\colon \mathbb{R} \to \mathbb{R}$ uit *Analyse I*:[6]

$$\text{Als } f'(a) \neq 0, \text{ dan is de inverse } f^{-1} \text{ afleidbaar in } f(a) \text{ en is } (f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

2.3.1 Definitie. Een C^k**-diffeomorfisme** van $G_1 \subseteq \mathbb{R}^n$ op $G_2 \subseteq \mathbb{R}^n$ (G_1 en G_2 open) is een bijectie $\mathbf{f}\colon G_1 \to G_2$ waarvoor zowel $\mathbf{f}$ als $\mathbf{f}^{-1}\colon G_2 \to G_1$ van de klasse C^k zijn.
Een $\mathbb{R}^n$-$\mathbb{R}^n$ functie $\mathbf{f}$ is een **lokaal** C^k**-diffeomorfisme** in $\mathbf{a} \in \mathbb{R}^n$ als

1. $\mathbf{f}$ een bijectie is van een omgeving W_1 van $\mathbf{a}$ naar een omgeving W_2 van $\mathbf{f}(\mathbf{a})$[7]
2. $\mathbf{f}$ van de klasse C^k is op een (open)[8] omgeving U_1 van $\mathbf{a}$
3. $\mathbf{f}^{-1}$ van de klasse C^k is op een (open) omgeving U_2 van $\mathbf{f}(\mathbf{a})$.

2.3.2 Opmerking. Men kan aantonen dat men steeds $U_1 = W_1$ en $U_2 = W_2$ kan kiezen,[9] zodat $\mathbf{f}$ een lokaal C^k-diffeomorfisme is als en slechts als een open omgeving U van $\mathbf{a}$ en een open omgeving V van $\mathbf{f}(\mathbf{a})$ bestaan zo dat $\mathbf{f}_{|U}\colon U \to V$ een C^k-diffeomorfisme is (stelling 2.3.B).

2.3.3 Stelling (Stelling van de inverse functies). *Zij* $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ *een functie van klasse* C^k *(*$k \geqslant 1$*) in een omgeving van* $\mathbf{a} \in \mathbb{R}^n$.
Als $\det(D\mathbf{f}(\mathbf{a})) \neq 0$, *dan geldt:*

1. $\mathbf{f}$ *is een lokaal* C^k*-diffeomorfisme in* $\mathbf{a}$.

2. $D\mathbf{f}^{-1}(\mathbf{f}(\mathbf{a})) = (D\mathbf{f}(\mathbf{a}))^{-1}$.

Bewijs. 1. We kunnen de inverse functie $\mathbf{f}^{-1}$ beschouwen als de functie die impliciet gedefinieerd wordt door de vergelijking $\mathbf{x} = \mathbf{f}(\mathbf{y})$. Stel dus

$$\mathbf{F} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n : \mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{f}(\mathbf{y}) - \mathbf{x}.$$

We gaan de voorwaarden van stelling 2.2.5 na: $D_2\mathbf{F}(\mathbf{f}(\mathbf{a}), \mathbf{a}) = D\mathbf{f}(\mathbf{a})$ is inverteerbaar, en $\mathbf{F}$ is van de klasse C^k in een omgeving van $(\mathbf{f}(\mathbf{a}), \mathbf{a})$, omdat $\frac{\partial \mathbf{F}}{\partial x_j}(\mathbf{x}, \mathbf{y})$ constant zijn en

$$\frac{\partial \mathbf{F}}{\partial y_j}(\mathbf{x}, \mathbf{y}) = \frac{\partial \mathbf{f}}{\partial y_j}(\mathbf{y}).$$

Wegens stelling 2.2.5 bestaan dus $\delta > 0$ en $\varepsilon > 0$ zo dat voor elke $\mathbf{x} \in U := B(\mathbf{f}(\mathbf{a}), \delta)$ een unieke $\mathbf{y}(\mathbf{x}) \in B(\mathbf{a}, \varepsilon)$ bestaat waarvoor $\mathbf{f}(\mathbf{y}(\mathbf{x})) = \mathbf{x}$ en is $\mathbf{y}(\mathbf{x})$ een C^k-afbeelding $U \to \mathbb{R}^n$. Dit toont aan dat $\mathbf{y}(\mathbf{x})$ de inverse afbeelding is van de bijectie $\mathbf{f}\colon \mathbf{y}(U) \to U$.[10] (De afbeelding $\mathbf{y}(\mathbf{x})$ is niet noodzakelijk surjectief op $B(\mathbf{a}, \varepsilon)$.) We gaan na dat $\mathbf{y}(U)$ een omgeving is van $\mathbf{a}$. Omdat $\mathbf{f}$ continu is, vinden we een $r \in]0, \varepsilon[$ met de eigenschap dat $\mathbf{f}(\boldsymbol{\xi}) \in U$ voor elke $\boldsymbol{\xi} \in B(\mathbf{a}, r)$. Dan is $B(\mathbf{a}, r) \subseteq \mathbf{y}(U)$, want als $\boldsymbol{\xi} \in B(\mathbf{a}, r)$, dan is $\boldsymbol{\xi}$ het unieke element $\mathbf{z} \in B(a, \varepsilon)$ waarvoor $\mathbf{f}(\mathbf{z}) = \mathbf{f}(\boldsymbol{\xi}) \in U$, m.a.w. $\boldsymbol{\xi} = \mathbf{y}(\mathbf{f}(\boldsymbol{\xi})) \in y(U)$.

[6] we zullen nu wel veronderstellen dat f minstens continu afleidbaar is in een omgeving van a

[7] W_1, W_2 zijn niet noodzakelijk open

[8] noodzakelijk open, want we hebben klasse C^k enkel gedefinieerd op open verzamelingen

[9] vergewis je ervan dat dat niet vanzelfsprekend kan!

[10] omdat $\mathbf{f}(\mathbf{y}(U)) = U$, is $\mathbf{f}\colon \mathbf{y}(U) \to U$ surjectief; ook is $\mathbf{f}$ injectief op $\mathbf{y}(U)$: als $\mathbf{f}(\mathbf{y}(\mathbf{x}_1)) = \mathbf{f}(\mathbf{y}(\mathbf{x}_2))$ voor zekere $\mathbf{x}_1, \mathbf{x}_2 \in U$, dan is $\mathbf{x}_1 = \mathbf{x}_2$, en dus ook $\mathbf{y}(\mathbf{x}_1) = \mathbf{y}(\mathbf{x}_2)$

2. Deze uitdrukking volgt door de gelijkheid $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$ af te leiden in $\mathbf{a}$ en de kettingregel toe te passen. $\square$

2.3.A Eigenschap. *Zij $G \subseteq \mathbb{R}^n$ een open verzameling en $\mathbf{f}$ een continue afbeelding $G \to \mathbb{R}^m$. Als $U \subseteq \mathbb{R}^m$ open is, dan is ook $\mathbf{f}^{-1}(U) := \{\mathbf{x} \in G : \mathbf{f}(\mathbf{x}) \in U\} \subseteq \mathbb{R}^n$ open.*

Bewijs. Kies willekeurig $\mathbf{a} \in \mathbf{f}^{-1}(U)$. Het volstaat aan te tonen dat $B(\mathbf{a}, \delta) \subseteq \mathbf{f}^{-1}(U)$ voor zekere $\delta > 0$. Bij definitie is $\mathbf{f}(\mathbf{a}) \in U$. Omdat U open is, bestaat $\varepsilon > 0$ zo dat $B(\mathbf{f}(\mathbf{a}), \varepsilon) \subseteq U$. Omdat $\mathbf{f}$ continu is in $\mathbf{a}$, bestaat $\delta > 0$ zo dat $\mathbf{f}(\mathbf{x}) \in B(\mathbf{f}(\mathbf{a}), \varepsilon) \subseteq U$ voor elke $\mathbf{x} \in B(\mathbf{a}, \delta)$ (omdat G open is, behoort zulke $\mathbf{x}$ automatisch tot het domein van $\mathbf{f}$ als we $\delta > 0$ klein genoeg kiezen), d.w.z. dat $B(\mathbf{a}, \delta) \subseteq \mathbf{f}^{-1}(U)$. $\square$

We kunnen nu de equivalente definitie 2.3.2 van lokaal C^k-diffeomorfisme aantonen:

2.3.B Stelling. *Als een functie $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ een lokaal C^k-diffeomorfisme in $\mathbf{a} \in \mathbb{R}^n$, dan bestaan er een open omgeving U van $\mathbf{a}$ en een open omgeving V van $\mathbf{f}(\mathbf{a})$ zo dat $\mathbf{f}_{|U}\colon U \to V$ een C^k-diffeomorfisme is.*

Bewijs. Kies U_1, U_2, W_1, W_2 zoals in definitie 2.3.1. Kies $\varepsilon > 0$ zo dat $B(\mathbf{f}(\mathbf{a}), \varepsilon) \subseteq U_2 \cap W_2$. Omdat $\mathbf{f}$ continu is in $\mathbf{a}$, vinden we $\delta > 0$ zo dat $\mathbf{f}(\mathbf{x}) \in B(\mathbf{f}(\mathbf{a}), \varepsilon)$ voor elke $\mathbf{x} \in B(\mathbf{a}, \delta) =: U$. Door $\delta > 0$ klein genoeg te kiezen, kunnen we er ook voor zorgen dat $B(\mathbf{a}, \delta) \subseteq U_1 \cap W_1$. Noem $\mathbf{g} := \mathbf{f}^{-1}{}_{|B(\mathbf{f}(\mathbf{a}),\varepsilon)}$. Omdat $\mathbf{f}(U)$ een deelverzameling van het domein van $\mathbf{g}$ is, is $\mathbf{f}(U) = \mathbf{g}^{-1}(U)$. Omdat $\mathbf{g}$ continu is, is $\mathbf{g}^{-1}(U)$ open (eigenschap 2.3.A). Er volgt dat $\mathbf{f}_{|U}\colon U \to \mathbf{f}(U)$ een C^k-diffeomorfisme is. $\square$

Tot slot geven we nog een aanvulling over vectorwaardige functies op compacte verzamelingen. We weten uit *Analyse I* dat de inverse van een continue strikt stijgende functie over een interval eveneens continu is. Hier volgt de versie voor een vectorfunctie in verschillende veranderlijken. Ons bewijs gebruikt daarbij *compactheid,* wat voor $n = 1$ niet het geval was. Ook is er natuurlijk geen sprake van 'stijgen' of 'dalen', want dat heeft geen betekenis voor functies van meerdere veranderlijken.

2.3.C Stelling. *Is $\mathbf{f} : K \to \mathbb{R}^m$ continu en injectief over de compacte verzameling $K \subset \mathbb{R}^n$, dan is haar inverse $\mathbf{f}^{-1} : \mathbf{f}(K) \to K$ eveneens continu en injectief over een compacte verzameling, nl. $\mathbf{f}(K)$.*

Bewijs. Zij $\mathbf{y}_0$ een willekeurig punt van $\mathbf{f}(K)$. We zullen het rijenkenmerk voor continuïteit van $\mathbf{f}^{-1}{}_{|\mathbf{f}(K)}$ in $\mathbf{y}_0$ gebruiken: *is $(\mathbf{y}_n)_n$ een willekeurige rij uit $\mathbf{f}(K)$ die naar $\mathbf{y}_0$ convergeert, dan $\mathbf{f}^{-1}(\mathbf{y}_n) \to \mathbf{f}^{-1}(\mathbf{y}_0)$.* Wegens de injectiviteit is $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ en $\mathbf{y}_n = \mathbf{f}(\mathbf{x}_n)$ voor unieke $\mathbf{x}_0, \mathbf{x}_1, \ldots$ uit K. Het gegeven is dus, dat $\mathbf{f}(\mathbf{x}_n) \to \mathbf{f}(\mathbf{x}_0)$, en we moeten aantonen dat $\mathbf{x}_n \to \mathbf{x}_0$, m.a.w.

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n \in \mathbb{N})(n > N \implies \|\mathbf{x}_0 - \mathbf{x}_n\| < \varepsilon).$$

Als dat niet zo was, dan zou er een $\varepsilon_0 > 0$ bestaan met de eigenschap dat

$$(\forall N \in \mathbb{N})(\exists n \in \mathbb{N})(n > N \,\&\, \|\mathbf{x}_0 - \mathbf{x}_n\| \geqslant \varepsilon_0).$$

Bij $N = 1$ vinden we dan een $n_1 > 1$ met $\|\mathbf{x}_0 - \mathbf{x}_{n_1}\| \geqslant \varepsilon_0$, bij $N = n_1$ een $n_2 > n_1$ met $\|\mathbf{x}_0 - \mathbf{x}_{n_2}\| \geqslant \varepsilon_0$ enzovoort. Er ontstaat een deelrij $(\mathbf{x}_{n_k})_k$ van $(\mathbf{x}_n)_n$. Wegens K compact heeft die een deelrij $\mathbf{x}_{n_{k_l}} \to \boldsymbol{\xi} \in K$. Die limiet kan niets anders zijn dan $\mathbf{x}_0$. Inderdaad, door continuïteit volgt uit $\mathbf{x}_{n_{k_l}} \to \boldsymbol{\xi}$ dat $\mathbf{f}(\mathbf{x}_{n_{k_l}}) \to \mathbf{f}(\boldsymbol{\xi})$. Maar uit het gegeven $\mathbf{f}(\mathbf{x}_n) \to \mathbf{f}(\mathbf{x}_0)$ volgt dat de deelrij $(\mathbf{f}(\mathbf{x}_{n_{k_l}}))_l$ dezelfde limiet $\mathbf{f}(\mathbf{x}_0)$ heeft. Vandaar is $\mathbf{f}(\boldsymbol{\xi}) = \mathbf{f}(\mathbf{x}_0)$ en, door injectiviteit, $\boldsymbol{\xi} = \mathbf{x}_0$. We hebben dus $\mathbf{x}_{n_{k_l}} \to \mathbf{x}_0$, maar dat is in tegenspraak met het feit dat $\|\mathbf{x}_0 - \mathbf{x}_{n_{k_l}}\| \geqslant \varepsilon_0$ voor alle l. □

2.4 Oefeningen

1. In welke punten van $\mathbb{R}^{n+m}$ hebben de volgende vergelijkingen lokaal een unieke oplossing van de opgegeven m veranderlijken als functie van de andere n veranderlijken?

2. Bepaal een gelijkheid voor de partiële afgeleiden van deze functie van n veranderlijken door impliciet afleiden.

3. Bepaal de vergelijking van de raakruimte in een gegeven punt van $\mathbb{R}^{n+m}$ aan het oppervlak bepaald door de gegeven vergelijking(en).

   (a) $xy - \sin y = 0 \quad$ (y in functie van x).

   (b) $xy - \sin(y/x) = 0 \quad$ (x in functie van y).

   (c) $z^m + yz + x = 0$ ($m \in \mathbb{N}$, $m \geqslant 2$) $\quad$ (z in functie van (x, y)).

   (d) $ze^z = x - 2y \quad$ (z in functie van (x, y)).

   (e) $\begin{cases} y^z = x \\ x + y + z = 3 \end{cases} \quad$ ((y, z) in functie van x).

   (f) $\begin{cases} u^2 + v^2 + 2x = 0 \\ uv - y = 0 \end{cases} \quad$ ((u, v) in functie van (x, y)).