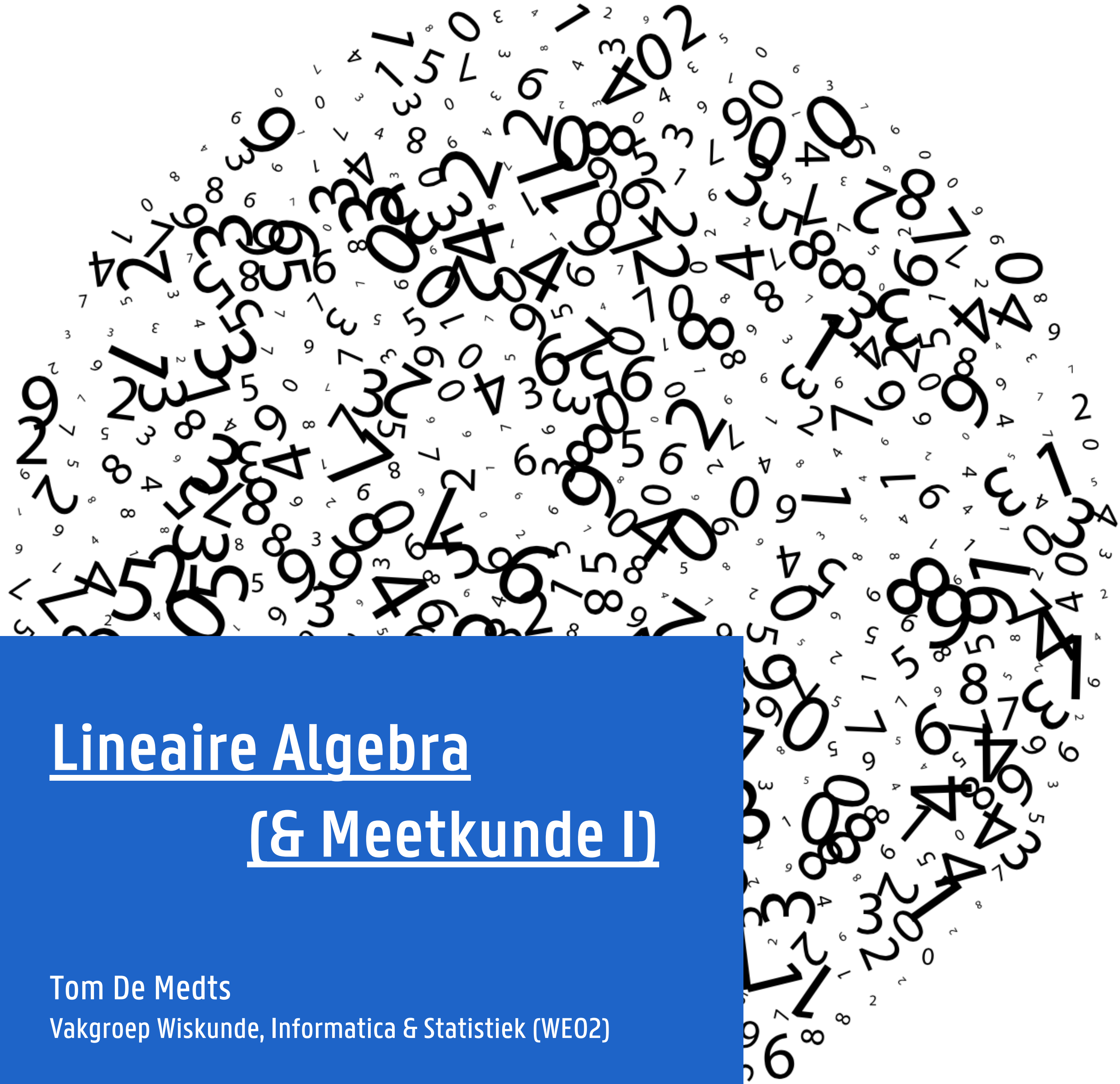
FACULTEIT WETENSCHAPPEN



Lineaire Algebra (& Meetkunde I)

Tom De Medts

Vakgroep Wiskunde, Informatica & Statistiek (WE02)

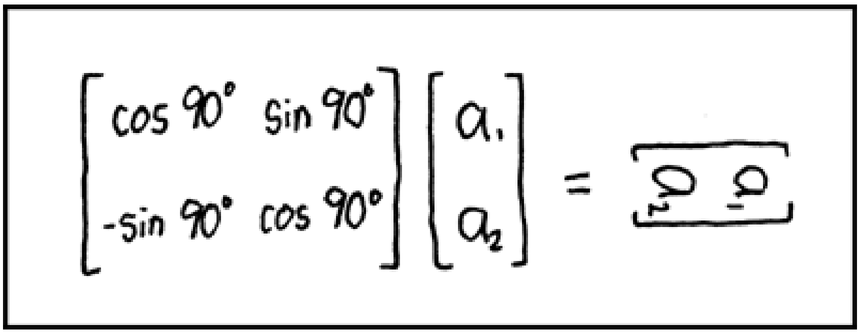
Opleidingen Wiskunde, Fysica & Sterrenkunde

Eerste bachelor, eerste semester

Academiejaar 2026–2027

UNIVERSITEIT GENT

Inhoudsopgave



http://xkcd.com/184/

Voorwoord

Deze cursus bevat het lesmateriaal van het vak “Lineaire Algebra en Meetkunde I” van de opleiding Wiskunde en van het vak “Lineaire Algebra” van de opleiding Fysica en Sterrenkunde. Het bevat de nodige basiskennis om op verder te bouwen later in de beide opleidingen. Het vak telt 6 studiepunten in de opleiding Wiskunde, en 5 in de opleiding Fysica en Sterrenkunde. Toch is er gekozen om met één cursus te werken, die alle informatie samenbundelt en handig kan zijn als je later nog even iets wil raadplegen. Dit betekent wel dat we niet alles zullen behandelen.

Hoewel de cursus vrij volledig is, wordt het toch warm aanbevolen om naar de lessen te komen, al is dit natuurlijk niet verplicht.

Vragen over de inhoud van de cursus, meldingen van typfouten of andere problemen? Steeds welkom voor of na de les, of stuur gerust een e-mail naar `tom.demedts@ugent.be`.

Hoofdstuk

0 Inleiding: De vectorruimte $\mathbb{R}^n$

Lineaire algebra is een deelgebied van de wiskunde, dat zich bezighoudt met de studie van vectoren, vectorruimten en lineaire afbeeldingen tussen vectorruimten.

De lineaire algebra staat centraal in de moderne wiskunde en haar toepassingen (waarbij we voornamelijk denken aan natuurwetenschappen, maar ook aan bijvoorbeeld de sociale wetenschappen). Een elementaire toepassing van de lineaire algebra is het oplossen van stelsels van lineaire vergelijkingen in meerdere onbekenden, maar minstens even belangrijk is het begrijpen van lineaire operatoren (bijvoorbeeld door middel van diagonalisatie). Aangezien de lineaire algebra zo centraal staat, is het niet verwonderlijk dat heel wat andere vakgebieden van de wiskunde hierop geënt zijn, zoals bijvoorbeeld abstracte algebra of functionaalanalyse. Ook worden in concrete toepassingen niet-lineaire wiskundige modellen vaak benaderd door lineaire modellen.

Veel van de basisinstrumenten van de lineaire algebra, in het bijzonder die met betrekking tot de oplossing van stelsels lineaire vergelijkingen, werden al in de Oudheid gebruikt. Maar de abstracte studie van vectoren en vectorruimten begint pas in de jaren 1600. De oorsprong van veel van deze ideeën is afkomstig van het gebruik van determinanten voor het oplossen van stelsels van vergelijkingen. Determinanten werden voor het eerst gebruikt door Leibniz in 1693, en de algemene *regel van Cramer* werd omstreeks 1750 ingevoerd door Gabriel Cramer. De kleinste-kwadratenmethode, die voor het eerst in de jaren 1790 door Carl Friedrich Gauss werd gebruikt, is een vroege en significante toepassing van de ideeën uit de lineaire algebra.

Het onderwerp begon haar moderne vorm aan te nemen in het midden van de 19e eeuw, toen veel ideeën en methoden uit vorige eeuwen werden veralgemeend in de abstracte algebra. Het was pas rond deze tijd dat Arthur Cayley terecht opmerkte:

> *There would be many things to say about this theory of matrices which should, it seems to me, precede the theory of determinants.*

In het begin van de 20e eeuw zorgde het gebruik van deze objecten in de speciale relativiteitstheorie, statistiek en kwantummechanica er voor dat de ideeën van de lineaire algebra zich buiten de zuivere wiskunde hebben verspreid.

De belangrijkste structuren van de lineaire algebra zijn vectorruimten en de lineaire afbeeldingen tussen deze vectorruimten. Een vectorruimte is een verzameling, waarvan de elementen bij elkaar op kunnen worden geteld en met *scalairen* of getallen kunnen worden vermenigvuldigd. In veel natuurlijke toepassingen zijn deze scalairen reële getallen, i.e. elementen van $\mathbb{R}$.

Vooraleer we de theorie van de vectorruimten algemeen —en dus eerder abstract— opbouwen, kijken we eerst even naar het specifieke voorbeeld van de vectorruimte $\mathbb{R}^n$. We gaan er van uit dat de lezer reeds vertrouwd is met matrices, en geen problemen heeft om met verzamelingen te werken.

Definitie. Zij n een geheel getal met $n \geq 1$. We definiëren de *vectorruimte* $\mathbb{R}^n$ als de verzameling V die bestaat uit alle $(n \times 1)$-matrices (dit zijn dus kolommatrices bestaande uit n rijen), voorzien van twee bewerkingen, met name de *optelling* en de *scalaire vermenigvuldiging*:

Optelling. We kunnen twee elementen van $\mathbb{R}^n$ optellen door hun componenten twee aan twee op te tellen:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix},$$

voor alle $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$.

Scalaire vermenigvuldiging. We kunnen een element van $\mathbb{R}^n$ vermenigvuldigen met een *scalair* (dit is een reëel getal) door elke component te vermenigvuldigen met die scalair:

$$\lambda \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix},$$

voor alle $\lambda, a_1, \ldots, a_n \in \mathbb{R}$.

Deze bewerkingen zijn niet willekeurig gekozen, en voelen heel natuurlijk aan. Ze voldoen aan een heel aantal interessante eigenschappen, waarvan we er een aantal oplijsten.

(1) Voor alle $v, w, u \in V$ is $(v + w) + u = v + (w + u)$.

(2) Er bestaat een $0_V \in V$ zodat $v + 0_V = 0_V + v = v$ voor alle $v \in V$.

(3) Voor alle $v \in V$ bestaat er een element $w \in V$ zodat $v+w = w+v = 0_V$.

(4) Voor alle $v, w \in V$ geldt dat $v + w = w + v$.

(5) Voor alle $v \in V$ en alle $\lambda, \mu \in \mathbb{R}$ geldt dat $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$.

(6) Voor alle $v \in V$ is $1 \cdot v = v$.

(7) Voor alle $v, w \in V$ en alle $\lambda, \mu \in \mathbb{R}$ is $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$ en $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

Merk op dat het element $0_V \in V$ waarvan sprake in (2) concreet kan voorgesteld worden door

$$0_V = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

De eigenschappen (1)–(7) zijn cruciaal om de structuur van $\mathbb{R}^n$ goed te begrijpen. We zullen dadelijk, in Hoofdstuk 2, precies deze 7 eigenschappen gebruiken om algemeen te *definiëren* wat we verstaan onder een vectorruimte. We zullen daarbij ook de structuur van de scalairen veralgemenen: waar we hier werken met de reële getallen, zullen we in onze algemene definitie werken met een willekeurig *veld*.

We kunnen vectorruimten eigenlijk pas goed begrijpen als we ook *lineaire afbeeldingen* tussen vectorruimten beschouwen.

Definitie. Beschouw vectorruimten $V = \mathbb{R}^n$ en $W = \mathbb{R}^m$. Een afbeelding $f\colon V \to W$ wordt een *lineaire afbeelding* genoemd, als ze de optelling en de scalaire vermenigvuldiging behoudt. Hiermee bedoelen we dat

$$\begin{aligned} f(u + v) &= f(u) + f(v) \quad \text{en} \\ f(\lambda \cdot v) &= \lambda \cdot f(v) \end{aligned}$$

voor alle $u, v \in V$ en alle $\lambda \in \mathbb{R}$.

Een typisch voorbeeld hiervan wordt gegeven door linkse vermenigvuldiging met een vast gekozen $m \times n$-matrix:

$$f \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

In heel wat toepassingen is het belangrijk om lineaire afbeeldingen van een vectorruimte V naar zichzelf te bestuderen, m.a.w. lineaire afbeeldingen $f\colon V \to V$; deze worden ook wel *lineaire operatoren* genoemd. Bijvoorbeeld, als we in de fysica bewegingen (van starre lichamen) bestuderen, dan worden deze bewegingen voorgesteld door middel van een lineaire operator op de vectorruimte $\mathbb{R}^3$ samengesteld met een translatie in $\mathbb{R}^3$. Een geavanceerder voorbeeld vinden we terug in de kwantummechanica, waar elke observabele wordt voorgesteld door middel van een zelf-toegevoegde lineaire operator op de toestandsruimte.

We zullen dan ook heel veel aandacht schenken aan de studie van lineaire afbeeldingen, en in het bijzonder van lineaire operatoren (zie Hoofdstuk 7).

Hoofdstuk

1 Inleidende begrippen

In dit eerste hoofdstuk definiëren we eerst *velden*. Deze structuren vormen een veralgemening van de rationale getallen $\mathbb{Q}$, de reële getallen $\mathbb{R}$ en de complexe getallen $\mathbb{C}$. Nadien bestuderen we veeltermen, matrices en stelsels van lineaire vergelijkingen over een willekeurig veld. We veronderstellen dat de lezer reeds kennis heeft gemaakt met veeltermen, matrices en stelsels lineaire vergelijkingen over $\mathbb{Q}$ of over $\mathbb{R}$.

We starten met een beknopte herhaling van het begrip verzameling; we veronderstellen dat de lezer hier reeds vertrouwd mee is. Onder een *verzameling* verstaan we een collectie van verschillende elementen; we gebruiken hiervoor de notatie

$$A = \{a, b, c, \dots\} .$$

Wanneer we willen benadrukken dat we een nieuwe verzameling invoeren, gebruiken we soms het symbool ":=" in plaats van "="; dit symbool kan gelezen worden als "is per definitie gelijk aan".

We zullen gebruik maken van de volgende courante begrippen: *deelverzameling*, *unie* (of vereniging), *doorsnede* (of intersectie), *verschil*.

Notatie 1.0.1. Beschouw twee verzamelingen A en B.

- Als A een *deelverzameling* van B is, noteren we dit met $A \subseteq B$. We gebruiken $A \subsetneq B$ als we willen benadrukken dat $A \subseteq B$ maar $A \neq B$.
- De *unie* van A en B noteren we met $A \cup B$.
- De *doorsnede* van A en B noteren we met $A \cap B$.
- Het *verschil* van A en B noteren we met $A \setminus B := \{a \mid a \in A, a \notin B\}$.
- Het *aantal elementen* van de verzameling A (dat niet noodzakelijk eindig is) noteren we met $|A|$; het wordt ook de *kardinaliteit* van A genoemd.
- De *ledige verzameling* $\{\,\}$ noteren we als $\emptyset$.

Onder een *functie* of *afbeelding* f van een verzameling A naar een verzameling B verstaan we een relatie tussen A en B met de eigenschap dat elk element van A met juist één element van B in relatie staat[1]. We noteren een functie of afbeelding van A naar B met $f\colon A \to B$, en we noteren het unieke element van B dat we aan een gegeven element a van A koppelen, als $f(a)$. Dit geheel noteren we ook kortweg als

$$f\colon A \to B\colon a \mapsto f(a).$$

1.1 Velden

We veronderstellen dat de lezer reeds vertrouwd is met de verzameling van natuurlijke getallen $\mathbb{N} = \{0, 1, 2, \dots\}$, de verzameling van gehele getallen $\mathbb{Z}$, de verzameling van rationale getallen $\mathbb{Q}$ en de verzameling van reële getallen $\mathbb{R}$. We zullen in deze cursus ook gebruik maken van de complexe getallen $\mathbb{C}$; deze voeren we hieronder beknopt in.

Definitie 1.1.1. Een *complex getal* is een uitdrukking van de vorm $a + bi$, waarin $a, b \in \mathbb{R}$ en i een nieuw 'getal' voorstelt met de eigenschap dat $i^2 = -1$.

De som en het product van twee complexe getallen $a + bi$ en $c + di$ met $a, b, c, d \in \mathbb{R}$ is als volgt gedefinieerd:

$$\begin{aligned} (a+bi) + (c+di) &= (a+c) + (b+d)i \\ (a+bi)(c+di) &= (ac-bd) + (ad+bc)i \end{aligned}$$

We noteren

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}.$$

De afbeelding

$$\iota\colon \mathbb{C} \to \mathbb{C}\colon a + bi \mapsto a - bi$$

is van groot belang; we noemen ι de *complexe toevoeging*, en noteren deze ook als $\iota(z) = \overline{z}$ voor alle $z \in \mathbb{C}$. We definiëren de *norm* $\mathsf{N}(z)$ van een complex getal als

$$\mathsf{N}(a+bi) = (a+bi)(\overline{a+bi}) = a^2 + b^2,$$

en de *modulus* $|z|$ van een complex getal z als de vierkantswortel van de norm, i.e.

$$|a+bi| = \sqrt{a^2+b^2}.$$

Lemma 1.1.2. (i) *Voor alle $z \in \mathbb{C}$ is $z + \overline{z} \in \mathbb{R}$ en $z\overline{z} \in \mathbb{R}_{\geq 0}$. Bovendien is $|z| = 0$ als en slechts als $z = 0$, en is $z = \overline{z}$ als en slechts als $z \in \mathbb{R}$.*

[1] Soms wordt het onderscheid tussen functie en afbeelding gemaakt, waarbij een functie dan een relatie is zodat elk element van A met *ten hoogste* één element van B in relatie staat, maar we zullen dit onderscheid meestal niet maken.

(ii) *Elk niet-nul element $a + bi \in \mathbb{C} \setminus \{0\}$ heeft een inverse voor de vermenigvuldiging, namelijk*

$$(a+bi)^{-1} = \frac{\overline{a+bi}}{\mathsf{N}(a+bi)} = \frac{a-bi}{a^2+b^2}\,.$$

Bewijs. (i) Zij $z = a + bi \in \mathbb{C}$. Per definitie is $z + \overline{z} = 2a \in \mathbb{R}$, en is $|z| = \sqrt{a^2+b^2} \in \mathbb{R}$. Merk op dat $a, b \in \mathbb{R}$, zodat $a^2 + b^2 \geq 0$, waarbij $a^2 + b^2 = 0$ als en slechts als $a = b = 0$.
Er geldt dat $z = a + bi = a - bi = \overline{z}$ als en slechts als $b = 0$, of dus als en slechts als $z \in \mathbb{R}$.

(ii) Merk op dat voor $a + bi \in \mathbb{C} \setminus \{0\}$ geldt dat $a^2 + b^2 \neq 0$, dus

$$\frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i \in \mathbb{C}.$$

Nu is

$$(a+bi)\frac{(a-bi)}{a^2+b^2} = \frac{a^2+b^2}{a^2+b^2} = 1.$$

□

Opmerking 1.1.3. Merk op dat we niet kunnen zeggen wanneer een element van $\mathbb{C}$ kleiner is dan een ander element van $\mathbb{C}$.

We komen nu tot de definitie van een veld; dit is een veralgemening van $\mathbb{Q}$, $\mathbb{R}$ en $\mathbb{C}$.

Definitie 1.1.4. Een *veld*[2] is een verzameling K, voorzien van twee bewerkingen, met name een *optelling* en een *vermenigvuldiging*, die we noteren als

$$\begin{aligned} &K \times K \to K : (a,b) \mapsto a + b\,,\\ &K \times K \to K : (a,b) \mapsto ab\,, \end{aligned}$$

met de volgende eigenschappen:

(K1) De optelling is *associatief*:
Voor alle $a, b, c \in K$ is $(a + b) + c = a + (b + c)$.

(K2) Er bestaat een *neutraal element* voor de optelling:
Er bestaat een $z \in K$ zodat voor alle $a \in K$ geldt dat $a+z = z+a = a$.
Dit element z is uniek (zie Lemma 1.1.8(i)); we noteren het als 0.

[2]In Nederland en soms ook in de buurt van Antwerpen wordt er gesproken over een *lichaam* i.p.v. een veld.

(K3) Elk element heeft een *invers element* voor de optelling:
Voor alle $a \in K$ bestaat er een element $b \in K$ zodat $a + b = b + a = 0$. Dit element b (corresponderend met a) is uniek (zie Lemma 1.1.8(iii)); we noteren het als $-a$.

(K4) De optelling is *commutatief*:
Voor alle $a, b \in K$ geldt dat $a + b = b + a$.

(K5) De vermenigvuldiging is *associatief*:
Voor alle $a, b, c \in K$ is $(ab)c = a(bc)$.

(K6) Er bestaat een *neutraal element* voor de vermenigvuldiging:
Er bestaat een $e \in K \setminus \{0\}$ zodat voor alle $a \in K$ geldt dat $ae = ea = a$. Dit element e is uniek (zie Lemma 1.1.8(ii)); we noteren het als 1.

(K7) Elk niet-nul element heeft een *invers element* voor de vermenigvuldiging:
Voor alle $a \in K \setminus \{0\}$ bestaat er een element $c \in K \setminus \{0\}$ zodat $ac = ca = 1$. Dit element c (corresponderend met a) is uniek (zie Lemma 1.1.8(iv)); we noteren het als a^{-1}.

(K8) De vermenigvuldiging is *commutatief*:
Voor alle $a, b \in K$ geldt dat $ab = ba$.

(K9) De vermenigvuldiging is *distributief* t.o.v. de optelling:
Voor alle $a, b, c \in K$ is $a(b+c) = ab+ac$ en dus ook $(b+c)a = ba+ca$.

De elementen van een veld noemen we vaak *scalairen*.

Opmerking 1.1.5. Om te verifiëren dat een gegeven structuur een veld is, mag je niet vergeten om de *inwendigheid* van de bewerkingen na te gaan: de optelling en vermenigvuldiging gaan van $K \times K$ naar K, dus het resultaat van de bewerkingen moet opnieuw een element van K zijn.

Opmerking 1.1.6. Ga zelf na dat voor $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ de eigenschappen (K1)–(K9) voldaan zijn; $\mathbb{Q}, \mathbb{R}$ en $\mathbb{C}$ zijn dus voorbeelden van velden. Bemerk dat $\mathbb{N}$ en $\mathbb{Z}$ echter geen velden zijn, aangezien niet elk niet-nul element inverteerbaar is voor de vermenigvuldiging.

Veelal worden in cursussen lineaire algebra hoofdzakelijk vectorruimten over de reële getallen beschouwd. Dit is een sterke beperking; er zijn heel wat toepassingen van de vectorruimtentheorie, zowel theoretische als praktische, waarbij het essentieel is dat willekeurige velden beschouwd worden. Daarenboven is er feitelijk geen verschil in de opbouw van de theorie van vectorruimten over $\mathbb{R}$ of vectorruimten over een veld K.

We zullen hier niet diep ingaan op algemene theorie van velden of op eigenschappen van andere velden dan $\mathbb{Q}, \mathbb{R}$ of $\mathbb{C}$; dit zou ons te ver leiden van het hoofdonderwerp van deze cursus, namelijk lineaire algebra en meetkunde.[3] We geven hieronder wel kort nog enkele voorbeelden van velden.

Voorbeelden 1.1.7. (1) Zij

$$\mathbb{R}[x] = \{a_0 + a_1x + \cdots + a_nx^n \mid n \in \mathbb{N}, a_i \in \mathbb{R}\}$$

de verzameling van de veeltermen over $\mathbb{R}$ in de variabele x met de klassieke optelling en vermenigvuldiging van veeltermen (zie ook paragraaf 1.2).
We kunnen ook "breuken" $\frac{f(x)}{g(x)}$, met $g(x) \in \mathbb{R}[x] \setminus \{0\}$ definiëren, waarbij we aannemen dat $\frac{f_1(x)}{g_1(x)} = \frac{f_2(x)}{g_2(x)}$ als en slechts als $f_1(x)g_2(x) = f_2(x)g_1(x)$. Zulke breuken noemt men *rationale functies over* $\mathbb{R}$ *in de variabele* x. We noteren de verzameling van rationale functies over $\mathbb{R}$ als

$$\mathbb{R}(x) = \left\{ \frac{f(x)}{g(x)} \mid f(x), g(x) \in \mathbb{R}[x], g(x) \neq 0 \right\},$$

en definiëren de bewerkingen

$$\frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)} = \frac{g_2(x)f_1(x) + g_1(x)f_2(x)}{g_1(x)g_2(x)}$$

en

$$\frac{f_1(x)}{g_1(x)} \cdot \frac{f_2(x)}{g_2(x)} = \frac{f_1(x)f_2(x)}{g_1(x)g_2(x)}.$$

Ga na dat $\mathbb{R}(x)$ een veld is.

(2) Voor de geïnteresseerde lezer geven we een voorbeeld van velden die slechts een eindig aantal elementen bevatten.
Zij p een priemgetal. Beschouw de verzameling van resten van gehele getallen na deling door p,

$$\mathbb{F}_p := \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{p-1}\}.$$

(We noteren $\overline{a}$ voor de rest van a na deling door p). Vermits de rest (na deling door p) van een optelling en een vermenigvuldiging van gehele getallen bekomen wordt door de resten (na deling door p) op te tellen, respectievelijk te vermenigvuldigen, en de rest van het resultaat (na deling door p) te nemen, zijn de volgende bewerkingen goed gedefinieerd:

$$\overline{n} + \overline{m} := \overline{n+m},$$

[3] De wiskundigen zullen veel meer leren over velden in de cursussen "Discrete Wiskunde I", "Algebra I" en "Algebra II".

$$\overline{n}\,\overline{m} := \overline{nm}.$$

Men kan verifiëren dat het feit dat p een priemgetal is impliceert dat $\mathbb{F}_p$ met deze bewerkingen een veld is; het is een *eindig veld*. Het veld $\mathbb{F}_p$ is (in zekere zin) het enige veld met p elementen.
Een voorbeeld voor zo een veld wordt gegeven door het veld $\mathbb{F}_2$ met 2 elementen. Dit veld speelt vooral in de informatica een grote rol bij de studie van de Booleaanse algebra.
Er bestaan ook eindige velden waarvan het aantal elementen een *priemmacht* p^h is, maar die zijn moeilijker te construeren. Zo speelt het veld met $2^8 = 256$ elementen een cruciale rol in de AES encryptie-standaard.

(3) Hoewel we hier niet ingaan op meer voorbeelden van velden is het belangrijk om in het achterhoofd te houden dat er zeer veel verschillende voorbeelden van velden bestaan! Het is zelfs onmogelijk om alle velden te classificeren.

De volgende elementaire eigenschappen van velden zullen we voortdurend gebruiken in het vervolg van de cursus. Ze tonen eigenlijk aan dat we in een willekeurig veld kunnen 'rekenen' zoals we in $\mathbb{Q}$, $\mathbb{R}$ en $\mathbb{C}$ doen.

Lemma 1.1.8. *Zij K een veld. Dan gelden volgende eigenschappen:*

(i) *Het neutraal element voor de optelling is uniek, m.a.w. stel dat er twee elementen $0, 0' \in K$ bestaan zodanig dat voor alle $a \in K$ geldt dat $0 + a = a + 0 = a$ en $0' + a = a + 0' = a$, dan is $0 = 0'$.*

(ii) *Het neutraal element voor de vermenigvuldiging is uniek, m.a.w. stel dat er twee elementen $1, 1' \in K$ bestaan zodanig dat voor alle $a \in K$ geldt dat $1a = a1 = a$ en $1'a = a1' = a$, dan is $1 = 1'$.*

(iii) *Het inverse van een element in K voor de optelling is uniek, m.a.w. zij $a \in K$ en stel dat er twee elementen $b, c \in K$ bestaan zodanig dat $a + b = b + a = 0$ en dat $a + c = c + a = 0$, dan is $b = c$. Het unieke inverse element voor de optelling noteren we als $-a$, en noemen we het* tegengestelde *van a. We spreken het uit als "min a".*

(iv) *Het inverse van een element in $K \setminus \{0\}$ voor de vermenigvuldiging is uniek, m.a.w. zij $a \in K \setminus \{0\}$ en stel dat er twee elementen $b, c \in K$ bestaan zodanig dat $ab = ba = 1$ en dat $ac = ca = 1$, dan is $b = c$. Het unieke inverse element voor de vermenigvuldiging noteren we als a^{-1} of als $\frac{1}{a}$, en noemen we het* inverse *van a. We spreken het uit als "a invers".*

(v) *Voor alle $a \in K$ geldt dat $a0 = 0$.*

(vi) *Voor alle* $a \in K$ *geldt dat* $-(-a) = a$.

(vii) *Voor alle* $a \in K$ *is* $-a = (-1)a = a(-1)$, *waarbij* -1 *het tegengestelde is van* 1. *Bovendien is* $-0 = 0$, *en geldt, voor alle* $a, b \in K$, *dat*

$$(-a)b = -(ab),\ -(a+b) = (-a) + (-b),\ (-a)(-b) = ab.$$

(viii) *Zij* $a, b \in K \setminus \{0\}$. *Dan is* $ab \neq 0$. *Anders gezegd, stel dat voor* $a, b \in K$ *geldt dat* $ab = 0$, *dan is ofwel*[4] $a = 0$, *ofwel* $b = 0$.

(ix) *Voor alle* $a, b \in K \setminus \{0\}$ *is* $(ab)^{-1} = a^{-1}b^{-1}$ *en is* $(a^{-1})^{-1} = a$.

Bewijs. We benadrukken dat men deze eigenschappen bewijst enkel gebruik makend van de eigenschappen (K1)–(K9) en eventueel reeds bewezen eigenschappen. Als voorbeeld bewijzen we (i); de andere eigenschappen zullen uitgewerkt worden in de oefeningenlessen.

Stel dus dat er twee elementen $0, 0' \in K$ bestaan zodanig dat voor alle $a \in K$ geldt dat $0 + a = a + 0 = a$ en $0' + a = a + 0' = a$. Door $a = 0'$ te kiezen in de uitspraak $0 + a = a$, komt er $0 + 0' = 0'$; anderzijds komt er, door $a = 0$ te kiezen in de uitspraak $a + 0' = a$, dat $0 + 0' = 0$. Door deze twee gelijkheden te vergelijken besluiten we dat $0 = 0'$. □

We noteren $a - b := a + (-b)$ voor alle $a, b \in K$ en $\frac{a}{b} := ab^{-1}$ voor alle $a \in K$ en $b \in K \setminus \{0\}$.

Naast velden zullen we in deze cursus ook gebruik maken van de begrippen *groep* en *ring*.

Definitie 1.1.9. Een *groep* is een verzameling G met daarop een bewerking

$$* : G \times G \to G : (a, b) \mapsto a * b$$

met de volgende eigenschappen:

(G1) Voor alle $a, b, c \in G$ is $(a * b) * c = a * (b * c)$.

(G2) Er bestaat een $e \in G$ zodat voor alle $a \in G$ geldt dat $a * e = e * a = a$.

(G3) Voor elke $a \in G$ bestaat er een element $b \in G$ zodat $a * b = b * a = e$.

Een groep is *abels* of *commutatief* als voor alle $a, b \in G$ geldt dat $a * b = b * a$.

Definitie 1.1.10. Een *ring* is een verzameling R met daarop een optelling en een vermenigvuldiging, die we noteren als

$$R \times R \to R : (a, b) \mapsto a + b,$$

[4] Het woord "ofwel" wordt steeds in wiskundige zin gebruikt, m.a.w. het betekent dat of het ene geldt, of het andere geldt, of beide gelden.

$$R \times R \to R : (a, b) \mapsto ab,$$

met de volgende eigenschappen:

(R1) Voor alle $a, b, c \in R$ is $(a + b) + c = a + (b + c)$.

(R2) Er bestaat een $0 \in R$ zodat voor alle $a \in R$ geldt dat $a+0 = 0+a = a$.

(R3) Voor alle $a \in R$ bestaat er een element $b \in R$ zodat $a + b = b + a = 0$.

(R4) Voor alle $a, b \in R$ geldt dat $a + b = b + a$.

(R5) Voor alle $a, b, c \in R$ is $(ab)c = a(bc)$.

(R6) Er bestaat een $1 \in R$ zodat voor alle $a \in R$ geldt dat $a1 = 1a = a$.

(R7) Voor alle $a, b, c \in R$ is $a(b + c) = ab + ac$ en $(b + c)a = ba + ca$.

Een ring is *commutatief* als voor alle $a, b \in R$ geldt dat $ab = ba$.

We geven enkele voorbeelden van groepen en ringen. We zullen in de loop van de cursus nog meer voorbeelden ontmoeten; natuurlijk bestaan er nog veel meer voorbeelden van groepen en van ringen.

Voorbeelden 1.1.11. (1) Zij K een veld. Ga na dat K dan ook een commutatieve ring is. Dus alle voorbeelden van velden zijn ook voorbeelden van ringen.

(2) Zij K een veld. Ga na dat K met als bewerking de optelling een abelse groep is (met $e = 0$). Ga na dat $K \setminus \{0\}$ met als bewerking de vermenigvuldiging een abelse groep is (met $e = 1$). Waarom is K met als bewerking de vermenigvuldiging geen groep?

(3) Zij R een ring. Ga na dat R met als bewerking de optelling een abelse groep is (met $e = 0$). Waarom is $R \setminus \{0\}$ met als bewerking de vermenigvuldiging niet noodzakelijk een groep?

(4) De gehele getallen $\mathbb{Z}$, met de gekende optelling en vermenigvuldiging, vormen een commutatieve ring, die geen veld is. De natuurlijke getallen $\mathbb{N}$ vormen geen ring; $\mathbb{N}$ met als bewerking de optelling vormt zelfs geen groep.

(5) De verzameling van de veeltermen $\mathbb{R}[x]$ over $\mathbb{R}$ in de variabele x met de optelling en vermenigvuldiging van veeltermen is een commutatieve ring (zie ook Lemma 1.2.2), die geen veld is.

(6) De verzameling van de $(n \times n)$-matrices over $\mathbb{R}$ met de optelling en vermenigvuldiging van matrices vormen een niet-commutatieve ring (zie ook Lemma 1.3.6).

Opmerking 1.1.12. (i) In iedere ring gelden ook de eigenschappen (i), (ii), (iii), (v), (vi) en (vii) van Lemma 1.1.8.

(ii) In een groep waar we de bewerking noteren als + gelden ook de eigenschappen (i), (iii) van Lemma 1.1.8.

1.2 Veeltermen

We veronderstellen dat de lezer vertrouwd is met veeltermen in één variabele over $\mathbb{R}$ (of over $\mathbb{Q}$). Het is echter mogelijk om veeltermen over een willekeurig veld te definiëren, zonder dat er iets essentieels verandert.

Definitie 1.2.1. Zij K een veld.

(i) Een *veelterm* (of *polynoom*) over K in één variabele x is een uitdrukking

$$a_m x^m + \cdots + a_1 x + a_0 \quad \text{met } m \in \mathbb{N} \text{ en } a_0, \ldots, a_m \in K.$$

We noteren een veelterm vaak met $f(x), g(x), \varphi(x), \psi(x), \ldots$ We zullen ook vaak gebruik maken van de compactere somnotatie,

$$a_0 + a_1 x + \cdots + a_n x^n = \sum_{i=0}^{n} a_i x^i.$$

(ii) Definieer de verzameling van alle veeltermen over K in één variabele als

$$K[x] := \{a_m x^m + \cdots + a_1 x + a_0 \mid m \in \mathbb{N}, a_0, \ldots, a_m \in K\}.$$

De optelling en vermenigvuldiging van twee veeltermen

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \quad \text{en} \quad g(x) = b_n x^n + \cdots + b_1 x + b_0$$

in $K[x]$ worden gegeven door[5]

$$f(x) + g(x) = (a_r + b_r)x^r + \cdots + (a_1 + b_1)x + (a_0 + b_0),$$

waarbij $r = \max\{m, n\}$, en

$$f(x)g(x) = \sum_{k=0}^{m+n} \Big(\sum_{i+j=k} a_i b_j \Big) x^k.$$

[5] Stel $a_i = 0$ voor alle $m < i \leq \max\{m, n\}$ en $b_j = 0$ voor alle $n < j \leq \max\{m, n\}$.

(iii) Zij $0 \neq f(x) \in K[x]$, dan is de *graad* van $f(x)$ de exponent van de hoogste macht van x waarvan de coëfficiënt in $f(x)$ niet nul is. We noteren de graad van $f(x)$ als $\deg f(x)$. Met andere woorden, als $\deg f(x) = m$, dan is[6]

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \quad \text{met } a_0, \ldots, a_m \in K, \; a_m \neq 0.$$

Voor elke $n \in \mathbb{N}$ noteren we de verzameling van veeltermen met graad ten hoogste n als P_n, m.a.w.

$$P_n = \{a_n x^n + \cdots + a_1 x + a_0 \mid a_0, \ldots, a_n \in K\} \subsetneq K[x].$$

(iv) Een veelterm $f(x) \in K[x]$ met $n := \deg f(x) \geq 0$ waarvan de coëfficiënt bij x^n gelijk is aan 1 noemen we een *monische veelterm.*
Een veelterm van graad nul noemen we ook een *constante veelterm.*

(v) Zij $f(x) = a_m x^m + \cdots + a_1 x + a_0 \in K[x]$ en $b \in K$, we noteren

$$f(b) := a_m b^m + \cdots + a_1 b + a_0 \in K.$$

We noemen $b \in K$ een *wortel* van $f(x)$ als en slechts als $f(b) = 0$.

Lemma 1.2.2. (i) *De verzameling $K[x]$ met de optelling en vermenigvuldiging van veeltermen is een commutatieve ring. We noemen $K[x]$ een* veeltermring *of* polynomenring *(over het veld K).*

(ii) *Voor alle $f(x), g(x) \in K[x]$ geldt:*

$$\begin{aligned} \deg(f(x) + g(x)) &\leq \max\{\deg f(x), \deg g(x)\} \\ \deg(f(x) \cdot g(x)) &= \deg f(x) + \deg g(x) \,. \end{aligned}$$

Bewijs. (i) Om aan te tonen dat $K[x]$ een ring is, moeten we gebruik maken van het feit dat K een veld is en dus voldoet aan de eigenschappen in Definitie 1.1.4 en Lemma 1.1.8. We geven wat toelichting bij enkele van de te bewijzen eigenschappen (R1)–(R7); werk zelf de details uit als oefening.

(R2) Het element $0 \in K$ is ook een element van $K[x]$ (een constante veelterm) en is het neutraal element voor de optelling in $K[x]$.

(R3) Ga na dat $(a_m x^m + \cdots + a_1 x + a_0) + (-a_m x^m - \cdots - a_1 x - a_0) = 0$.

(R5) Zij $f(x) = \sum_{i=0}^{m_1} a_i x^i \in K[x]$, zij $g(x) = \sum_{i=0}^{m_2} b_i x^i \in K[x]$, en zij $h(x) = \sum_{i=0}^{m_3} c_i x^i \in K[x]$. We tonen aan dat $f(x)\big(g(x)h(x)\big) =$

[6] We schrappen in $f(x)$ eerst alle termen van de vorm $0 \cdot x^i$ met $i > m$.

$\big(f(x)g(x)\big)h(x)$. Dit is niet moeilijk, maar het is een goede oefening op het gebruik van de somnotatie.

$$\begin{aligned}
f(x)\big(g(x)h(x)\big) &= f(x)\sum_{k=0}^{m_2+m_3}\Big(\sum_{i+j=k} b_i c_j\Big)x^k \\
&= \sum_{\ell=0}^{m_1+m_2+m_3}\Big(\sum_{n+k=\ell} a_n\big(\sum_{i+j=k} b_i c_j\big)\Big)x^\ell \\
&= \sum_{\ell=0}^{m_1+m_2+m_3}\Big(\sum_{n+i+j=\ell} a_n(b_i c_j)\Big)x^\ell \\
&= \sum_{\ell=0}^{m_1+m_2+m_3}\Big(\sum_{n+i+j=\ell} (a_n b_i)c_j\Big)x^\ell \quad \text{(wegens (K5))} \\
&= \sum_{\ell=0}^{m_1+m_2+m_3}\Big(\sum_{k+j=\ell}\big(\sum_{n+i=k} a_n b_i\big)c_j\Big)x^\ell \\
&= \Big(\sum_{k=0}^{m_1+m_2}\big(\sum_{n+i=k} a_n b_i\big)x^k\Big)h(x) \\
&= \big(f(x)g(x)\big)h(x).
\end{aligned}$$

(R6) Het element $1 \in K$ is ook een element van $K[x]$ (een constante veelterm) en is het neutraal element voor de vermenigvuldiging in $K[x]$.

(ii) Dit volgt direct uit de definitie van de optelling en vermenigvuldiging. □

We bespreken beknopt nog enkele eigenschappen en definities in verband met wortels van veeltermen en delers van veeltermen.

Stelling 1.2.3. *Zij $K[x]$ de polynomenring over een veld K. Zij $f(x), g(x) \in K[x]$ met $g(x) \neq 0$. Dan zijn er polynomen $q(x), r(x) \in K[x]$ zodat*

$$f(x) = g(x)q(x) + r(x) \text{ met } r(x) = 0 \text{ of } \deg r(x) < \deg g(x).$$

Zonder bewijs. (Het bewijs gebruikt de "staartdeling" voor veeltermen, analoog aan de staartdeling in $\mathbb{Z}$. Voor de wiskundigen verwijzen we naar de cursus "Discrete Wiskunde I".) □

Een belangrijk gevolg is dat een veelterm over K een wortel a heeft in K als en slechts als de veelterm deelbaar is door $x - a$.

Gevolg 1.2.4. *Zij* $0 \neq f(x) \in K[x]$*,* K *een veld, en zij* $a \in K$*. Dan is* $f(a) = 0$ *als en slechts als* $f(x) = (x - a)q(x)$ *voor zekere* $q(x) \in K[x]$*.*

Bewijs. Het is duidelijk dat uit $f(x) = (x - a)q(x)$ volgt dat $f(a) = 0$.

Stel omgekeerd dat $f(a) = 0$. Bovenstaande stelling geeft dat

$$f(x) = (x - a)q(x) + r(x),$$

met $\deg r(x) < \deg(x - a) = 1$ of $r(x) = 0$, met andere woorden, $r(x)$ is een constante veelterm. Uit $0 = f(a) = r(a)$ volgt dan $r(x) = 0$, bijgevolg is $f(x) = (x - a)q(x)$. □

Definitie 1.2.5. (i) Zij $f(x), g(x) \in K[x]$. Dan is $g(x)$ een *deler* van $f(x)$ als en slechts als er een $q(x) \in K[x]$ bestaat waarvoor $f(x) = g(x)q(x)$.

(ii) Zij $f(x) \in K[x]$, en zij a een wortel van $f(x)$. Dan is

$$f(x) = (x - a)^k q(x)$$

voor een bepaalde $k \in \mathbb{N}$ ($k \geq 1$) en $q(x) \in K[x]$ met $q(a) \neq 0$. We noemen k de *multipliciteit* van de wortel a.

(iii) Zij $f(x) \in K[x]$ een veelterm met $\deg f(x) > 0$. We noemen $f(x)$ *reducibel* over K als $f(x) = g(x)h(x)$ met $g(x), h(x) \in K[x]$, waarbij $\deg g(x) < \deg f(x)$ en $\deg h(x) < \deg f(x)$. We noemen f *irreducibel* als f niet reducibel is.

De volgende stelling staat gekend als de *grondstelling van de algebra*.

Stelling 1.2.6. *Zij* $f(x) \in \mathbb{C}[x]$ *met* $\deg f(x) \geq 1$*. Dan heeft* $f(x)$ *een wortel in* $\mathbb{C}$*.*

Zonder bewijs. (Voor de wiskundigen wordt dit bewezen in de cursussen "Complexe analyse" en "Algebra II".) □

Gevolg 1.2.7. *Zij* $f(x) \in \mathbb{C}[x]$ *een monische veelterm met* $\deg f(x) \geq 1$*. Dan is* $f(x) = \prod_{i=1}^{m}(x - a_i)^{n_i}$ *voor* $a_1, \ldots, a_m \in \mathbb{C}$ *en* $n_1, \ldots, n_m \in \mathbb{N}$*.*

Bewijs. Dit volgt uit Stelling 1.2.6 en Gevolg 1.2.4. □

1.3 Matrices

We veronderstellen dat de lezer reeds vertrouwd is met matrices over $\mathbb{Q}$ of over $\mathbb{R}$. We definiëren matrices over een willekeurig veld K.

Definitie 1.3.1. Zij K een veld.

(1) Zij $m, n \in \mathbb{N} \setminus \{0\}$. Een $m \times n$-*matrix* is een rechthoekig schema van elementen van K dat bestaat uit m *rijen* en n *kolommen*. Dit wordt op de volgende manier genoteerd:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

met $a_{ij} \in K$ voor alle $1 \leq i \leq m$ en $1 \leq j \leq n$. De scalairen a_{ij} noemen we de *componenten* of *elementen* van de matrix. Soms wordt ook de notatie met vierkante haken

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

gebruikt; er is geen enkel verschil in betekenis met de notatie met ronde haken die wij zullen hanteren.

(2) De verzameling van alle $m \times n$-matrices over K noteren we met $M_{m,n}(K)$ of $\mathrm{Mat}_{m,n}(K)$.

(3) Wanneer $n = m$ noteren we $M_{n,n}(K)$ als $M_n(K)$ of als $\mathrm{Mat}_n(K)$. We noemen deze matrices de *vierkante* matrices.

We zullen vaak gebruik maken van volgende notaties om matrices voor te stellen.

Notatie 1.3.2. De matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in M_{m,n}(K)$$

noteren we ook op de volgende manieren:

(1) $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ of $A = (a_{ij})_{i,j}$ of kortweg $A = (a_{ij})$, met $a_{ij} \in K$.

(2) $A = \begin{pmatrix} A_1 & \cdots & A_n \end{pmatrix}$ of $A = (A_1, \dots, A_n)$, met $A_i \in M_{m,1}(K)$ de kolommen van A voor $1 \leq i \leq n$.

(3) $A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}$ met $A_j \in M_{1,n}(K)$ de rijen van A voor $1 \leq j \leq m$. We zullen ook vaak de notatie $R_1, \ldots, R_m$ gebruik voor de rijen als het duidelijk is over welke matrix A het gaat.

We definiëren enkele bijzondere types matrices.

Definitie 1.3.3. Zij K een veld, $n, m \in \mathbb{N} \setminus \{0\}$.

(i) De *nulmatrix* is de matrix $(a_{ij}) \in M_{m,n}(K)$ met $a_{ij} = 0$ voor alle i, j. We noteren deze matrix als $0_{m,n}$ of gewoon 0; als $m = n$ gebruiken we ook 0_n.

(ii) Een matrix in $M_{m,1}(K)$ wordt een *kolommatrix* genoemd. Een matrix in $M_{1,n}(K)$ wordt een *rijmatrix* genoemd.
We noteren $K^m := M_{m,1}(K)$ voor de verzameling van de kolommatrices.

(iii) Een matrix $(a_{ij}) \in M_n(K)$ is een *diagonaalmatrix* als $a_{ij} = 0$ voor alle $i \neq j$. We noteren de diagonaalmatrix

$$\begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$

ook met $\mathrm{diag}(a_{11}, \ldots, a_{nn})$.

(iv) De *eenheidsmatrix* is de matrix $\mathrm{diag}(1, \ldots, 1) \in M_n(K)$. We noteren deze matrix als I_n.

(v) De matrix $(a_{ij}) \in M_n(K)$ is een *bovendriehoeksmatrix* als $a_{ij} = 0$ voor alle $i > j$.

(vi) De matrix $(a_{ij}) \in M_n(K)$ is een *onderdriehoeksmatrix* als $a_{ij} = 0$ voor alle $i < j$.

(vii) De matrix $(a_{ij}) \in M_n(K)$ is *symmetrisch* als $a_{ij} = a_{ji}$ voor alle i, j.

(viii) De matrix $(a_{ij}) \in M_n(K)$ is een *blokdiagonaalmatrix* met *blokken* $A_1, \ldots, A_k$, als A van de vorm

$$\begin{pmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & A_k \end{pmatrix}$$

is, met A_i een $n_i \times n_i$ matrix voor alle i.

We kunnen matrices optellen, vermenigvuldigen met een scalair of vermenigvuldigen met elkaar. De vermenigvuldiging van matrices is een bewerking die afgeleid is van wat men de *gewogen som* noemt:

$$\begin{pmatrix} a_1 & a_2 & \dots & a_k \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_k b_k \, .$$

Definitie 1.3.4. (1) De som van twee matrices $A = (a_{ij}) \in M_{m,n}(K)$ en $B = (b_{ij}) \in M_{m,n}(K)$ is de matrix $A + B := (a_{ij} + b_{ij}) \in M_{m,n}(K)$.

(2) Zij $\lambda \in K$, dan definiëren we het (scalair) product van de matrix $A = (a_{ij}) \in M_{m,n}(K)$ met de scalair λ als de matrix $\lambda A := (\lambda a_{ij}) \in M_{m,n}(K)$.

(3) Het product van twee matrices $A = (a_{ij}) \in M_{m,k}(K)$ en $B = (b_{ij}) \in M_{k,n}(K)$ is gedefinieerd als $(c_{ij}) := AB \in M_{m,n}(K)$ met

$$c_{ij} = \sum_{\ell=1}^{k} a_{i\ell} b_{\ell j} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ik} b_{kj}.$$

Merk op dat de som van matrices enkel gedefinieerd is voor matrices met evenveel rijen en evenveel kolommen. Het product AB van matrices is echter enkel gedefinieerd als A evenveel kolommen heeft als B rijen heeft. De ij-de component van AB is de gewogen som van de i-de rij van A met de j-de kolom van B. Het is een handige "vingertruc" om de ij-de component van het product te berekenen door met een vinger van de linkerhand over de i-de rij te gaan terwijl men met een vinger van de rechterhand over de j-de kolom gaat.

Opmerking 1.3.5. Een matrix $A = (A_1, \dots, A_n) \in M_{m,n}(K)$ rechts vermenigvuldigen met een kolommatrix $B \in M_{n,1}(K)$ geeft een kolommatrix die gelijk is aan een som van scalaire veelvouden van de kolommen van A:

$$(A_1, \dots, A_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = b_1 A_1 + \cdots + b_n A_n.$$

Een matrix $B = \begin{pmatrix} B_1 \\ \vdots \\ B_m \end{pmatrix} \in M_{m,n}(K)$ links vermenigvuldigen met een rijmatrix $A \in M_{1,m}$ geeft een rijmatrix die gelijk is aan een som van scalaire veelvouden van de rijen van B:

$$\begin{pmatrix} a_1 & \cdots & a_m \end{pmatrix} B = a_1 B_1 + \cdots + a_m B_m.$$

Merk op dat het product van twee $n \times n$-matrices opnieuw een $n \times n$-matrix is. In feite vormt, voor elke n, de verzameling van alle vierkante $n \times n$-matrices over K een ring. Meer bepaald hebben we het volgend resultaat.

Lemma 1.3.6. *Zij K een veld en $n, m \in \mathbb{N} \setminus \{0\}$.*

(i) *De verzameling van de matrices $M_{m,n}(K)$ met als operatie de matrixoptelling is een groep, met de nulmatrix $0_{m,n}$ als neutraal element.*

(ii) *De verzameling van de matrices $M_n(K)$ is een ring. Het neutraal element voor de optelling is de nulmatrix 0_n en het neutraal element voor de vermenigvuldiging is de eenheidsmatrix I_n.*

(iii) *Voor alle $A, B \in M_{m,n}(K)$, $\lambda, \mu \in K$ is*

$$(\lambda\mu)A = \lambda(\mu A), \quad (\lambda + \mu)A = \lambda A + \mu A, \quad \lambda(A + B) = \lambda A + \lambda B.$$

(iv) *Voor alle $A \in M_{m,k}(K)$, $B \in M_{k,n}(K)$ en $\lambda \in K$ is*

$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$

Bewijs. Deze eigenschappen volgen rechtstreeks uit de definities van optelling en vermenigvuldiging van matrices. We moeten natuurlijk gebruik maken van de eigenschappen van het veld K.

Eigenschappen (G3) en (R3) volgen uit het feit dat $(a_{ij}) + (-a_{ij}) = (a_{ij} + (-a_{ij})) = 0$. We werken de lastigste eigenschap (R5) volledig uit, en laten het bewijs van de andere eigenschappen als oefening.

Zij dus $A = (a_{ij}), B = (b_{ij}), C = (c_{ij}) \in M_n(K)$; we tonen aan dat $A(BC) = (AB)C$. We hebben

$$\begin{aligned} A(BC) &= \Big(\sum_{k=1}^{n} a_{ik}(BC)_{kj}\Big)_{1 \le i,j \le n} \\ &= \Big(\sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ik}\big(b_{k\ell}c_{\ell i}\big)\Big)_{1 \le i,j \le n} \end{aligned}$$

terwijl

$$\begin{aligned} (AB)C &= \Big(\sum_{\ell=1}^{n} (AB)_{i\ell}c_{\ell j}\Big)_{1 \le i,j \le n} \\ &= \Big(\sum_{k=1}^{n}\sum_{\ell=1}^{n} \big(a_{ik}b_{k\ell}\big)c_{\ell i}\Big)_{1 \le i,j \le n}. \end{aligned}$$

De gelijkheid van beide uitdrukkingen volgt nu uit (K5), de associativiteit van de vermenigvuldiging in het veld K. $\square$

Opmerking 1.3.7. Merk op dat de ring $M_n(K)$ niet commutatief is! We geven een voorbeeld van twee vierkante matrices waarvoor $AB \neq BA$:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \text{ maar } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Definitie 1.3.8. (i) Zij K een veld en $n \in \mathbb{N}\backslash\{0\}$. Een matrix $A \in M_n(K)$ is *inverteerbaar* als en slechts als er een matrix $B \in M_n(K)$ bestaat waarvoor $AB = BA = I_n$.

(ii) Als de matrix $A \in M_n(K)$ inverteerbaar is, dan is de matrix B met $AB = BA = I_n$ uniek bepaald. We noemen de matrix B de *inverse matrix* of kortweg de *inverse* van A en noteren deze met A^{-1}.

(iii) De verzameling van de inverteerbare matrices noteren we met $\mathsf{GL}_n(K)$.

Lemma 1.3.9. (i) *De verzameling* $\mathsf{GL}_n(K)$ *met als bewerking de matrix-vermenigvuldiging is een groep. Het neutrale element is* I_n.

(ii) *Zij* $A, B \in \mathsf{GL}_n(K)$. *Dan is* $(A^{-1})^{-1} = A$ *en* $(AB)^{-1} = B^{-1}A^{-1}$.

Bewijs. Oefening (gebruik Lemma 1.3.6). □

Niet iedere niet-nulmatrix is inverteerbaar.

Voorbeeld 1.3.10. Veronderstel dat de matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(K)$$

inverteerbaar zou zijn; dan bestaat er een $B = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(K)$ waarvoor $AB = BA = I_2$. Echter,

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

een tegenstrijdigheid. Hieruit volgt dat A niet inverteerbaar is.

Deze methode om na te gaan of een matrix al dan niet inverteerbaar is, is niet bijzonder praktisch. In Stelling 6.4.1 bewijzen we verscheidene meer bruikbare criteria om na te gaan of een matrix inverteerbaar is. We bekijken ook verscheidene methodes om het inverse van een inverteerbare matrix te bepalen.

Een ander zinvol begrip is het transponeren van een matrix.

Definitie 1.3.11. Zij $A = (a_{ij}) \in M_{m,n}(K)$. We definiëren de matrix $A^t = (b_{k\ell}) \in M_{n,m}(K)$ als $b_{ji} := a_{ij}$ voor alle $1 \leq i \leq m, 1 \leq j \leq n$. We noemen de matrix A^t de *getransponeerde matrix* van de matrix A.

De rijen van A worden dus de kolommen van A^t en de kolommen van A worden de rijen van A^t.

Voorbeeld 1.3.12.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}^t = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

We vermelden de volgende rekenregels voor de getransponeerde matrix.

Lemma 1.3.13. (i) *Voor alle* $A, B \in M_{m,n}(K)$ *is* $(A+B)^t = A^t + B^t$.

(ii) *Voor alle* $A \in M_{m,n}(K)$ *is* $(A^t)^t = A$.

(iii) *Voor alle* $A \in M_{m,k}(K), B \in M_{k,n}(K)$ *is* $(AB)^t = B^t A^t$.

(iv) *Voor alle* $A \in \mathsf{GL}_n(K)$ *is* $A^t \in \mathsf{GL}_n(K)$ *en* $(A^t)^{-1} = (A^{-1})^t$.

Bewijs. Oefening. $\square$

Notatie 1.3.14. We zullen het transponeren van matrices ook gebruiken als een notatie. We zullen een kolommatrix

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n = M_{n,1}(K)$$

dikwijls noteren als $(a_1 \cdots a_n)^t$ of $(a_1, \ldots, a_n)^t$.

De volgende definitie voelt misschien wat willekeurig aan, maar het nut ervan zal later nog duidelijk worden (zie Lemma 7.1.6).

Definitie 1.3.15. De som van de diagonaalelementen van een vierkante matrix A noemt men het *spoor* van A; we noteren $\mathrm{tr}(A) = a_{11} + \cdots + a_{nn}$.

Lemma 1.3.16. *Voor alle* $A, B \in M_n(K)$ *geldt* $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Bewijs. Dit is een eenvoudige oefening. $\square$

1.4 Stelsels van lineaire vergelijkingen

Een voorbeeld van een stelsel van 3 lineaire vergelijkingen in 3 onbekenden x, y, z over het veld $\mathbb{Q}$ is

$$\begin{cases} x + y + z = 3 \\ x - y + 2z = -1 \\ 3x - y + 5z = 1. \end{cases}$$

Door het elimineren van variabelen kunnen we de oplossingen van dit stelsel bepalen. Er blijken oneindig veel oplossingen te zijn: voor iedere $t \in \mathbb{Q}$ is $x = -\frac{3}{2}t + 1$, $y = \frac{1}{2}t + 2$, $z = t$ een oplossing van het stelsel.

In deze paragraaf ontwikkelen we een systematische methode om een willekeurig stelsel van m lineaire vergelijkingen in n onbekenden over een willekeurig veld K op te lossen.

Definitie 1.4.1. (i) Zij K een veld. Een stelsel over K van m lineaire vergelijkingen met n onbekenden[7] $x_1, x_2, \ldots, x_n$ is een collectie van m lineaire vergelijkingen over dezelfde n onbekenden, i.e. vergelijkingen van de vorm

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

waar $a_{ij} \in K$ voor alle $1 \leq i \leq m, 1 \leq j \leq n$ en $b_i \in K$ voor alle $1 \leq i \leq m$. (Het woord *stelsel* benadrukt het feit dat we deze vergelijkingen tegelijk bekijken, en niet individueel.)

(ii) Een *oplossing* van het bovenstaand stelsel is een n-tal bestaande uit n scalairen $c_1, \ldots, c_n \in K$ waarvoor geldt dat

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m. \end{cases}$$

Deze oplossing noteren we vaak als $x_1 = c_1, \ldots, x_n = c_n$ of als $(c_1, \ldots, c_n)^t$. De *oplossingsverzameling* van een stelsel lineaire vergelijkingen is de verzameling van alle oplossingen van het stelsel.

(iii) Een stelsel is *strijdig* als het geen oplossingen heeft, m.a.w. als de oplossingsverzameling gelijk is aan $\emptyset$.

(iv) Een stelsel is *homogeen* als het van de vorm

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

[7] Als er slechts een vast klein aantal onbekenden zijn worden deze ook wel genoteerd als $x, y, z, u, \ldots$.

is, voor $a_{ij} \in K$ met $1 \leq i \leq m, 1 \leq j \leq n$. Merk op dat een homogeen stelsel nooit strijdig is: $x_1 = x_2 = \cdots = x_n = 0$ is namelijk steeds een oplossing.

De lineaire vergelijkingen van een stelsel kunnen we vermenigvuldigen met een scalair en bij elkaar optellen. Om een stelsel op te lossen zullen we het eerst in een eenvoudigere vorm brengen. Hiervoor is het volgend lemma van cruciaal belang.

Lemma 1.4.2. *Beschouw een stelsel van m lineaire vergelijkingen met n onbekenden over het veld K. We noteren de vergelijkingen van dit stelsel als $R_1, \ldots, R_m$.*

We construeren een nieuw stelsel van m lineaire vergelijkingen met n onbekenden over het veld K door één van de volgende operaties toe te passen:

(I) *twee vergelijkingen van plaats verwisselen;*

(II) *een vergelijking R_i vervangen door λR_i voor een bepaalde $1 \leq i \leq m$ en $\lambda \in K \setminus \{0\}$;*

(III) *een vergelijking R_i vervangen door de vergelijking $R_i + \lambda R_j$ voor een bepaalde $1 \leq i \neq j \leq m$ en $\lambda \in K$.*

De oplossingsverzameling van het oorspronkelijke stelsel is gelijk aan de oplossingsverzameling van het nieuwe stelsel.

We noemen deze drie operaties ook wel *elementaire rij-operaties* of kortweg *EROs*. We verwijzen er naar als EROs van *type* I, II en III, respectievelijk[8].

Bewijs. We overlopen de drie verschillende operaties:

(I) Het is triviaal dat het verwisselen van twee vergelijkingen de oplossingen niet beïnvloedt.

(II) Het is evident dat $(c_1 \ \cdots \ c_n)^t$ een oplossing is van de vergelijking $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ als en slechts als het een oplossing is van de vergelijking $\lambda a_{i1}x_1 + \cdots + \lambda a_{in}x_n = \lambda b_i$.

(III) Om de notatie te vereenvoudigen beschouwen we een stelsel met slechts twee vergelijkingen, maar de redenering is ook geldig in een stelsel met m vergelijkingen.

[8] De volgorde van de drie types verschilt van bron tot bron. We gebruiken hier dezelfde conventie als die op de Wikipedia-pagina `https://nl.wikipedia.org/wiki/Elementaire_rijoperatie`; deze volgorde lijkt het vaakst voor te komen.

We beschouwen het stelsel

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2. \end{cases}$$

We passen operatie (1) toe door R_1 te vervangen door $R_1 + \lambda R_2$ met $\lambda \in K$; het nieuwe stelsel ziet er als volgt uit:

$$\begin{cases} (a_{11} + \lambda\, a_{21})x_1 + \cdots + (a_{1n} + \lambda\, a_{2n})x_n = b_1 + \lambda\, b_2 \\ \qquad\qquad\qquad\qquad a_{21}x_1 + \cdots + a_{2n}x_n = b_2. \end{cases}$$

Het is evident dat een oplossing van het oorspronkelijke stelsel ook een oplossing van het nieuwe stelsel is.
Stel nu omgekeerd dat $(c_1 \ \cdots \ c_n)^t$ een oplossing van het nieuwe stelsel is; dan is

$$\begin{aligned} (a_{11} + \lambda\, a_{21})c_1 + \cdots + (a_{1n} + \lambda\, a_{2n})c_n &= b_1 + \lambda\, b_2 \quad \text{en} \\ a_{21}c_1 + \cdots + a_{2n}c_n &= b_2. \end{aligned}$$

Door de eerste vergelijking min λ keer de tweede vergelijking te nemen, volgt dat ook $a_{11}c_1 + \cdots + a_{1n}c_n = b_1$. Bijgevolg is $(c_1 \ldots c_n)^t$ ook een oplossing van het oorspronkelijke stelsel. □

Om stelsels efficiënter te noteren, zullen we een stelsel vaak noteren als een matrixvergelijking. Het stelsel van m lineaire vergelijkingen met n onbekenden over K

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \qquad\qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

kunnen we immers veel compacter noteren als $AX = B$, waarbij A, X en B de volgende matrices zijn:

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Definitie 1.4.3. Beschouw een stelsel van m lineaire vergelijkingen met n onbekenden over K, genoteerd als $AX = B$, waarbij $A \in M_{m,n}(K)$, $B \in M_{m,1}(K)$ en X de kolommatrix is met de n onbekenden.

De $m \times (n+1)$-matrix $(A\ B)$ bekomen uit A door deze aan te vullen met de kolom B, noemen we de *uitgebreide matrix* van het stelsel. We noteren de uitgebreide matrix van dit stelsel ook als $(A|B)$; de streep geeft aan dat het de uitgebreide matrix is van een stelsel.

Voorbeeld 1.4.4. We hernemen het voorbeeldstelsel uit het begin van deze paragraaf:

$$\begin{cases} x + y + z = 3 \\ x - y + 2z = -1 \\ 3x - y + 5z = 1. \end{cases}$$

In matrixnotatie noteren we dit stelsel als $AX = B$, met

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 3 & -1 & 5 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad B = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}.$$

De uitgebreide matrix van dit stelsel is de 3×4-matrix

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 2 & -1 \\ 3 & -1 & 5 & 1 \end{pmatrix}, \text{ ook genoteerd als } \left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -1 & 2 & -1 \\ 3 & -1 & 5 & 1 \end{array}\right).$$

We 'vertalen' de operaties uit Lemma 1.4.2 naar operaties op de rijen van de uitgebreide matrix van een stelsel.

Definitie 1.4.5. Zij $A \in M_{m,n}(K)$ met rijen $R_1, \ldots, R_m$. We kunnen een nieuwe matrix construeren door één van de volgende operaties toe te passen op A:

(I) de rij R_i verwisselen met de rij R_j voor bepaalde $1 \leq i, j \leq m$;

(II) de rij R_i vervangen door λR_i voor een bepaalde $1 \leq i \leq m$ en $\lambda \in K \setminus \{0\}$;

(III) de rij R_i vervangen door $R_i + \lambda R_j$ voor bepaalde $1 \leq i, j \leq m$ en $\lambda \in K$.

Een dergelijke operatie noemen we een *elementaire rij-operatie* van type I, II of III respectievelijk.

Gevolg 1.4.6. *De oplossingsverzameling van een stelsel van lineaire vergelijkingen blijft dezelfde wanneer men op de uitgebreide matrix van het stelsel een eindig aantal elementaire rij-operaties na elkaar toepast.*

Bewijs. Dit volgt direct uit Lemma 1.4.2. □

Opmerking 1.4.7. De elementaire rij-operaties voor matrices in $M_{m,n}(K)$ kunnen alle verkregen worden door linkse vermenigvuldiging met een gepaste $m \times m$-matrix:

(I) de rij-operatie $R_i \leftrightsquigarrow R_j$ wordt verkregen door linkse vermenigvuldiging met de matrix

$$\begin{pmatrix} 1 & \cdots & & & & \cdots & 0 \\ \vdots & \ddots & & & & & \vdots \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & & & & \cdots & 1 \end{pmatrix},$$

i.e. de eenheidsmatrix I_m waarbij de 1'en op de (i,i)-de en de (j,j)-de positie vervangen werden door 0, en de 0'en op de (i,j)-de en de (j,i)-de positie vervangen werden door 1.

(II) de rij-operatie $R_i \rightsquigarrow \lambda R_i$ wordt verkregen door linkse vermenigvuldiging met de matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \lambda & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

i.e. de eenheidsmatrix I_m waarbij de 1 op de (i,i)-de positie vervangen werd door λ.

(III) de rij-operatie $R_i \rightsquigarrow R_i + \lambda R_j$ wordt verkregen door linkse vermenigvuldiging met de matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \lambda & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

i.e. de eenheidsmatrix I_m aangevuld met een λ op de (i,j)-de positie.

Als we een stelsel oplossen gaan we te werk door elementaire rij-operaties op de uitgebreide matrix toe te passen totdat de bekomen matrix de volgende specifieke gedaante heeft.

Definitie 1.4.8. Een matrix $A \in M_{m,n}(K)$ is een *rij-echelonmatrix* als de volgende twee voorwaarden voldaan zijn:

- Iedere rij is van de vorm $(0 \cdots 0)$, $(1 * \cdots *)$ of $(0 \cdots 0\ 1 * \cdots *)$, waar de sterretjes willekeurige scalairen voorstellen.
- Het eerste niet-nul element van de $(i+1)$-de rij ligt rechts van het eerste niet-nul element van de i-de rij; de nulrijen staan onderaan in de matrix.

We noemen de matrix A *rij-gereduceerd* als bovendien voldaan is aan de volgende derde voorwaarde:

- De elementen boven het eerste niet-nul element van iedere rij zijn allemaal gelijk aan nul.

Als een rij niet volledig gelijk is aan nul, wordt de plaats waar het eerste niet-nul element staat (dit is altijd een 1) een *spilplaats* of een *pivotplaats* genoemd. Een kolom waarin een spilplaats voorkomt noemen we een *spilkolom* of een *pivotkolom*.

Als A een rij-echelonmatrix is, zeggen we ook dat A in *rij-echelonvorm* staat. Aangezien we hier enkel rij-echelonmatrices beschouwen en geen kolom-echelonmatrices, hebben we het korter over *echelonmatrices* of matrices in *echelonvorm*.

Voorbeeld 1.4.9. We geven enkele voorbeelden van rij-echelonmatrices in $M_{3,4}(\mathbb{Q})$. De spilplaatsen zijn omcirkeld:

$$\begin{pmatrix} 1 & -3 & 4 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

De eerste twee van deze voorbeelden zijn niet rij-gereduceerd, het derde voorbeeld is dat wel. Een algemeen voorbeeld van een rij-echelonmatrix in $M_{4,n}(K)$ is

$$\begin{pmatrix} 0\cdots 0 & 1 & *\cdots * & * & *\cdots * & * & *\cdots * & * & *\cdots * \\ 0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots * & * & *\cdots * & * & *\cdots * \\ 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots * & * & *\cdots * \\ 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots * \end{pmatrix},$$

waar een sterretje een willekeurige scalair voorstelt. Als men in de bovenstaande matrix onderaan een nulrij plaatst blijft de matrix in echelonvorm.

Stelling 1.4.10. *Zij $A \in M_{m,n}(K)$. Men kan altijd een rij-gereduceerde matrix bekomen door een eindig aantal opeenvolgende elementaire rij-operaties op A toe te passen.*

Bewijs. Aangezien de nulmatrix reeds in rij-gereduceerde vorm staat, veronderstellen we dat $A \neq 0$. We gaan te werk in verschillende stappen:

Stap 1. Zoek de eerste kolom van A waarin een niet-nul element staat, noem deze kolom K_j.

Stap 2. Verwissel rijen, door rij-operaties van type I te gebruiken, zodat een niet-nul element in de bovenste rij R_i komt, noem dit $\lambda \in K \setminus \{0\}$.

Stap 3. Maak van die niet-nul component λ een 1 door een rij-operatie van type II te gebruiken: vermenigvuldig R_i met λ^{-1}.

Stap 4. Gebruik verschillende rij-operaties van type III om de elementen onder deze 1 nul te maken: alle rijen R_k met $k \leq i$ laten we staan, alle rijen R_k met $k > i$ vervangen we door de rij $R_k - a_{kj}R_i$.

Stap 5. De matrix A heeft nu de gedaante

$$\begin{pmatrix} 0\ldots 0 & 1 & *\cdots * \\ 0\cdots 0 & 0 & *\cdots * \\ \vdots & \vdots & \vdots \\ 0\cdots 0 & 0 & *\cdots * \end{pmatrix} = \begin{pmatrix} 0 & 1 & B \\ 0 & 0 & D \end{pmatrix}.$$

We passen nu stappen 1 tot 4 recursief toe op de matrix D, totdat er rechtsonder de laatst gevonden pivotplaats enkel nog nullen staan, of deze pivotplaats helemaal onderaan of helemaal rechts in de matrix staat.

Stap 6. De matrix staat nu reeds in echelonvorm. We maken ten slotte alle elementen boven elke pivotplaats nul door rij-operaties van type III toe te passen.

De matrix die we nu bekomen hebben is rij-gereduceerd. □

Opmerking 1.4.11. Het bewijs van voorgaande stelling is een constructief bewijs. Dit betekent dat we het bewijs kunnen toepassen om in de praktijk een matrix in een echelonmatrix om te zetten met behulp van elementaire rij-operaties. Hieronder werken we dit uit op een voorbeeld, in de oefeningenlessen worden nog heel wat meer voorbeelden besproken.

Deze techniek om een matrix naar rij-echelonvorm te brengen noemen we ook *rijreductie*.

Voorbeeld 1.4.12. We gaan verder met het voorbeeld 1.4.4. We passen de methode uit het bewijs van Stelling 1.4.10 toe op de matrix

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 2 & -1 \\ 3 & -1 & 5 & 1 \end{pmatrix}.$$

Aangezien het element linksboven reeds een 1 is, kunnen we direct overgaan naar Stap 4. Vervang R_2 door $R_2 - 1R_1$; dit geeft

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 1 & -4 \\ 3 & -1 & 5 & 1 \end{pmatrix}.$$

Vervang nu R_3 door $R_3 - 3R_1$; dit geeft

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 1 & -4 \\ 0 & -4 & 2 & -8 \end{pmatrix}.$$

Stap 4 is voltooid. We gaan verder met Stap 5 en werken dus verder met het omkaderde stuk van de matrix:

$$\left(\begin{array}{c|ccc|} 1 & 1 & 1 & 3 \\ \cline{2-4} 0 & -2 & 1 & -4 \\ 0 & -4 & 2 & -8 \\ \cline{2-4} \end{array}\right).$$

Stap 1 en Stap 2 zijn reeds in orde, het element linksboven is namelijk verschillend van nul. We gaan verder met Stap 3 en vervangen R_2 door $(-2)^{-1}R_2$:

$$\left(\begin{array}{c|ccc|} 1 & 1 & 1 & 3 \\ \cline{2-4} 0 & 1 & -\frac{1}{2} & 2 \\ 0 & -4 & 2 & -8 \\ \cline{2-4} \end{array}\right).$$

Nu gaan we verder met Stap 4 en vervangen R_3 door $R_3 + 4R_2$:

$$\left(\begin{array}{c|ccc|} 1 & 1 & 1 & 3 \\ \cline{2-4} 0 & 1 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 \\ \cline{2-4} \end{array}\right).$$

We moeten niet opnieuw Stap 5 doorlopen want er staat een nulmatrix rechtsonder de laatst verkregen 1. We gaan verder met Stap 6, we vervangen dus R_1 door $R_1 - R_2$ en krijgen de echelonmatrix

$$\begin{pmatrix} 1 & 0 & \frac{3}{2} & 1 \\ 0 & 1 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Wanneer de uitgebreide matrix van een stelsel een rij-echelonmatrix is, kunnen we de oplossing van het stelsel direct aflezen.

Stelling 1.4.13. *Zij $AX = B$ een stelsel van m lineaire vergelijkingen in n onbekenden over K, met $A \in M_{m,n}(K)$, $B \in M_{m,1}(K)$ en X een kolommatrix met de n onbekenden. Veronderstel dat de uitgebreide matrix $(A|B)$ een echelonmatrix is.*

(i) *Als de laatste kolom van $(A|B)$ een spilkolom is, is het stelsel strijdig.*

(ii) *Als de laatste kolom van $(A|B)$ geen spilkolom is, heeft het stelsel wel oplossingen. Zij $S \subseteq \{1, \ldots, n\}$ de verzameling van indices van alle spilkolommen van A. De oplossingen van het stelsel kunnen we als volgt beschrijven:*
De onbekenden x_i met $i \notin S$ worden vrij gekozen in K, stel $x_i = t_i \in K$. De onbekenden x_j met $j \in S$ zijn dan bepaald door

$$x_j = b_\ell - \sum_{k>j \text{ en } k \notin S}^{n} a_{\ell k} t_k, \tag{$*$}$$

waar ℓ het nummer van de rij van de spilplaats in de j^e-kolom is.

Bewijs. (i) Als de laatste kolom van $(A|B)$ een spilkolom is, betekent dit dat $(A|B)$ een rij bevat die gelijk is aan $(0 \cdots 0|1)$. Dit betekent dat het stelsel de vergelijking $0\,x_1 + \cdots + 0\,x_n = 1$ bevat; deze vergelijking heeft geen oplossingen. Het stelsel is dus strijdig.

(ii) Dit volgt uit de specifieke vorm van een echelonmatrix. In het bijzonder is $(*)$ precies de ℓ-de vergelijking van het stelsel. Merk hierbij op dat enerzijds $a_{\ell j} = 1$, want dit is precies het element op een spilplaats, en dat anderzijds $a_{\ell k} = 0$ voor alle $k < j$ en voor alle $k \in S \setminus \{j\}$. □

We kunnen dus aan de echelonvorm van de uitgebreide matrix aflezen 'hoeveel' oplossingen een stelsel heeft.

Gevolg 1.4.14. *Zij $AX = B$ een niet-strijdig stelsel met uitgebreide matrix $(A|B)$ in echelonvorm. Het aantal vrij te kiezen onbekenden is gelijk aan het aantal onbekenden min het aantal spilkolommen.*

Bewijs. Dit volgt rechtstreeks uit de voorgaande stelling. □

Opmerking 1.4.15. (i) Als de echelonvorm van de uitgebreide matrix een nulrij bevat, levert deze rij geen extra informatie op over de oplossingen van het stelsel. Een nulrij betekent namelijk dat $0\,x_1 + \cdots + 0\,x_n = 0$, dit is waar voor alle scalairen.

(ii) Uit het voorgaande gevolg volgt dat een stelsel exact één oplossing heeft als en slechts als iedere kolom van A een spilkolom is in $(A|B)$ en B geen spilkolom is. Enkel in dit geval is er een oplossing maar kunnen er geen onbekenden vrij gekozen worden.

(iii) De vorm van iedere oplossing die gegeven wordt in Stelling 1.4.13 is niet uniek. Je kan eventueel ook bepaalde x_j met $j \in S$ vrij kiezen en dan andere x_i met $i \notin S$ uitdrukken in functie van de gekozen onbekenden.

Voorbeeld 1.4.16. We gaan verder met het stelsel uit het begin van de paragraaf. In Voorbeeld 1.4.12 hebben we de echelonvorm van de uitgebreide matrix bepaald. We lezen nu de oplossingen van het stelsel af met behulp van de voorgaande stelling. De echelonvorm is gelijk aan (met de spilplaatsen omcirkeld)

$$\left(\begin{array}{ccc|c} 1 & 0 & \frac{3}{2} & 1 \\ 0 & 1 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

De laatste kolom is geen spilkolom dus het stelsel heeft oplossingen. De verzameling van indices van spilkolommen is $S = \{1, 2\}$. We stellen dus $x_3 = k \in K$, dan is $x_2 = 2 + \frac{1}{2}k$ en $x_1 = 1 - \frac{3}{2}k$. De oplossingsverzameling is gelijk aan $\{(1 - \frac{3}{2}k, 2 + \frac{1}{2}k, k)^t \mid k \in K\}$.

Tot slot vatten we samen hoe men de oplossingsverzameling van een willekeurig stelsel lineaire vergelijkingen kan bepalen:

(1) Bepaal de uitgebreide matrix $(A|B)$ van het stelsel.

(2) Breng de uitgebreide matrix in echelonvorm met behulp van het constructief bewijs van Stelling 1.4.10. Uit Gevolg 1.4.6 volgt dat de oplossingsverzameling van het stelsel hierdoor ongewijzigd blijft.

(3) Gebruik Stelling 1.4.13 om de oplossingsverzameling van het stelsel af te lezen.

Merk op dat deze methode neerkomt op het systematisch elimineren van zoveel mogelijk onbekenden.

Tot slot geven we een efficiënte methode om via rijreductie te bepalen of een matrix inverteerbaar is, en dan ook de inverse matrix te bepalen. We steunen hierbij op Opmerking 1.4.7 en (wat verrassender) op Gevolg 1.4.14.

Stelling 1.4.17. *Zij $A \in M_n(K)$ een vierkante matrix.*

(i) *De matrix A is inverteerbaar als en slechts als A door elementaire rijoperaties kan omgevormd worden tot de eenheidsmatrix I_n.*

(ii) *Als A inverteerbaar is, dan kunnen we A^{-1} bepalen door de "uitgebreide matrix" $\big(A \mid I_n\big)$ om te vormen naar een rij-gereduceerde matrix $\big(I_n \mid B\big)$; de resulterende matrix B is dan gelijk aan A^{-1}.*

Bewijs. (i) Merk op dat alle matrices uit Opmerking 1.4.7 inverteerbaar zijn. Als de matrix A door elementaire rij-operaties kan omgevormd worden tot de eenheidsmatrix I_n, dan volgt hieruit dat A het product is van inverteerbare matrices, en bijgevolg zelf inverteerbaar is.

Onderstel omgekeerd dat A inverteerbaar is. Dan heeft het stelsel $AX = 0$ een unieke oplossing (namelijk $X = 0$), want uit $AX = 0$ volgt $A^{-1}(AX) = 0$ en dus $X = 0$.

Zet nu de matrix A om via rijreductie naar een rij-gereduceerde matrix A' (wat altijd kan wegens Stelling 1.4.10). Merk op dat wegens Gevolg 1.4.6 het stelsel $A'X = 0$ ook de unieke oplossing $X = 0$ heeft. Uit Gevolg 1.4.14 volgt nu dat er juist n spilkolommen zijn, omdat er geen vrij te kiezen onbekenden zijn. Omdat A' echter een $n \times n$-matrix is, wil dit zeggen dat *elke* kolom van A' een spilkolom is. Dat kan enkel als $A' = I_n$.

(ii) Als A inverteerbaar is, weten we uit (i) dat we door middel van elementaire rij-operaties A kunnen omvormen tot de eenheidsmatrix I_n. Wegens Opmerking 1.4.7 wordt deze sequentie van elementaire rij-operaties gegeven door linksvermenigvuldiging met een zekere matrix R, dus $I_n = RA$. Als we nu deze rij-operaties loslaten op de grotere matrix $\big(A \mid I_n\big)$, dan zal de rechtse matrix I_n vermenigvuldigd worden met *dezelfde* matrix R, en dus omgevormd worden tot $\big(I_n \mid B\big) = \big(RA \mid RI_n\big)$. Dus $I_n = RA$ en $B = RI_n$, waaruit volgt dat $B = A^{-1}$. $\square$

Hoofdstuk 2

Vectorruimten

2.1 Definities en eerste eigenschappen

We gaan nu van start met een algemene studie van vectorruimten. Onze definitie is geïnspireerd op de eigenschappen die we in Hoofdstuk 0 hebben vastgesteld voor $\mathbb{R}^n$ uitgerust met de optelling en de scalaire vermenigvuldiging.

Definitie 2.1.1. Een *vectorruimte over een veld* K (of een K*-vectorruimte*) is een verzameling V met twee bewerkingen: de optelling

$$V \times V \to V\colon (v, w) \mapsto v + w,$$

en de vermenigvuldiging met scalairen

$$K \times V \to V\colon (\lambda, v) \mapsto \lambda v,$$

die aan de volgende eigenschappen voldoen:

(V1) Voor alle $v, w, u \in V$ is $(v + w) + u = v + (w + u)$.

(V2) Er bestaat een $0_V \in V$ zodat voor alle $v \in V$ geldt dat $v + 0_V = 0_V + v = v$.

(V3) Voor alle $v \in V$ bestaat er een element $w \in V$ zodat $v+w = w+v = 0_V$.

(V4) Voor alle $v, w \in V$ geldt dat $v + w = w + v$.

(V5) Voor alle $v \in V$ en alle $\lambda, \mu \in K$ geldt $(\lambda\mu)v = \lambda(\mu v)$.

(V6) Voor alle $v \in V$ is $1v = v$ (hier is $1 \in K$ het neutraal element voor de vermenigvuldiging in K).

(V7) Voor alle $v, w \in V$ en alle $\lambda, \mu \in K$ is $(\lambda+\mu)v = \lambda v+\mu v$ en $\lambda(v+w) = \lambda v + \lambda w$.

De eigenschappen (V1)–(V4) drukken uit dat V met als bewerking de optelling een abelse groep is, net 0_V als neutraal element.

Voorbeelden 2.1.2. (1) Zij K een willekeurig veld, en stel $V = \{0\}$. Dan vormt V, met als optelling $0 + 0 = 0$ en als scalaire vermenigvuldiging $\lambda 0 = 0$ voor alle $\lambda \in K$, een vectorruimte over K. We noemen dit de *nulruimte* (over K).

(2) Het standaardvoorbeeld van een vectorruimte over een veld K is de *kolommenruimte*

$$K^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_i \in K \right\}$$

met de componentsgewijze optelling en vermenigvuldiging met scalairen

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad \lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}.$$

Merk op dat dit een rechtstreekse veralgemening is van de vectorruimte $\mathbb{R}^n$ die we in Hoofdstuk 0 hebben ingevoerd. Het rekenen en redeneren in K^n verloopt dan ook geheel analoog als in $\mathbb{R}^n$.

(3) De verzameling van veeltermen in één variabele over een veld K,

$$K[x] := \left\{ \sum_{i=0}^{d} a_i x^i \mid d \in \mathbb{N}, a_i \in K \right\}$$

vormt een vectorruimte over K. (Ga dit zelf na als oefening.)

(4) Zij K een veld, en stel $V = M_{m,n}(K)$, de verzameling van $m \times n$-matrices, met de som en de scalaire vermenigvuldiging zoals gedefinieerd in Definitie 1.3.4. Dan is V een vectorruimte over K; dit volgt uit Lemma 1.3.6(i) en (iii).

(5) Het veld K zelf is een K-vectorruimte met als scalaire vermenigvuldiging de vermenigvuldiging van K.

(6) Het volgend voorbeeld laat zien dat de elementen van een vectorruimte zelf "ingewikkeldere" objecten kunnen zijn. (Dit idee zal later ook van belang zijn als we ruimten van homomorfismen en duale ruimten zullen bespreken.)
Zij K een veld, zij X een niet-lege verzameling, en beschouw de nieuwe verzameling

$$\mathbf{F} := \{f \colon X \to K\}$$

van alle mogelijke afbeeldingen van X naar K. We voorzien $\mathbf{F}$ met de bewerkingen "optelling" en "scalaire vermenigvuldiging" als volgt.

Veronderstel dat f en g twee willekeurige elementen van $\mathbf{F}$ zijn, m.a.w. twee willekeurige afbeeldingen van X naar K. Dan definiëren we een nieuwe afbeelding $f+g\colon X \to K$ (en dus $f+g \in \mathbf{F}$) door het voorschrift

$$(f+g)(x) := f(x) + g(x) \quad \text{voor alle } x \in X.$$

Veronderstel nu dat f een willekeurig element van $\mathbf{F}$ is en $\lambda \in K$ een willekeurige scalair; dan definiëren we een nieuwe afbeelding $\lambda f\colon X \to K$ (en dus $\lambda f \in \mathbf{F}$) door het voorschrift

$$(\lambda f)(x) := \lambda f(x) \quad \text{voor alle } x \in X.$$

Deze bewerkingen maken van $\mathbf{F}$ een K-vectorruimte. Ga zelf na dat (V1)–(V7) inderdaad voldaan zijn[1]. We merken op dat het element $0_{\mathbf{F}}$ dat nodig is in (V2) hier gegeven wordt door de nulafbeelding, i.e.

$$0_{\mathbf{F}}\colon X \to K \colon x \mapsto 0.$$

De volgende rekenregels in vectorruimten zullen we zeer vaak gebruiken:

Lemma 2.1.3. *Zij V een K-vectorruimte. We noteren het neutraal element voor de optelling in K met 0_K en het neutraal element voor de optelling in V met 0_V.*

(i) *Het neutraal element in V voor de optelling in V is uniek.*

(ii) *Elke $v \in V$ heeft een uniek tegengestelde voor de optelling; we noteren het als $-v$.*

(iii) *Voor alle $v \in V$ is $-(-v) = v$.*

(iv) *Voor alle $\lambda \in K$ is $\lambda 0_V = 0_V$.*

(v) *Voor alle $v \in V$ is $0_K v = 0_V$.*

(vi) *Voor alle $v \in V$ is $(-1)v = -v$. Bovendien geldt, voor alle $v \in V$ en $\lambda \in K$, dat*

$$(-\lambda)v = -(\lambda v) = \lambda(-v), \ \lambda v = (-\lambda)(-v), \ -(v+w) = -v + (-w).$$

(vii) *Als voor een $v \in V$ en $\lambda \in K$ geldt dat $\lambda v = 0$, dan is $\lambda = 0_K$ of $v = 0_V$.*

Bewijs. Oefening. (Dit wordt besproken in de oefeningenlessen.) □

[1] Gebruik hiervoor dat twee elementen $f, g \in \mathbf{F}$ gelijk zijn aan elkaar als en slechts als $f(x) = g(x)$ voor alle $x \in X$. Zie ook Opmerking 3.1.2 verderop.

Opmerking 2.1.4. (i) In het vorige lemma hebben we een verschillende notatie gebruikt voor $0_K \in K$ en $0_V \in V$. Vanaf nu zullen we deze twee nulelementen beide met 0 noteren; uit de context is het steeds duidelijk of er 0_K of 0_V bedoeld wordt.

(ii) We definiëren de aftrekking voor alle $v, w \in V$ als

$$v - w := v + (-w) = (-w) + v = -w + v.$$

2.2 Deelruimten, lineaire combinaties, span

Definitie 2.2.1. Een deelverzameling W van een K-vectorruimte V noemt men een *deelruimte* van V als W zelf een vectorruimte is voor de operaties van V; we noteren dit dan als $W \leq V$.

Opmerking 2.2.2. We benadrukken dat het symbool $\leq$ hier een notatie is, en niet zomaar mag gebruikt worden zoals we dit symbool zouden gebruiken voor getallen. In het bijzonder kunnen twee deelruimten $U \leq V$ en $W \leq V$ gerust tegelijk voldoen aan $U \not\leq W$ én $W \not\leq U$.

Het volgend criterium om na te gaan wanneer een deelverzameling een deelruimte is, is zeer eenvoudig maar zeer belangrijk.

Lemma 2.2.3. *Zij K een veld en V een K-vectorruimte, en zij $\emptyset \neq W \subseteq V$ een deelverzameling van V. Dan is W een deelruimte van V als en slechts als*

$$\lambda_1 w_1 + \lambda_2 w_2 \in W$$

voor alle $w_1, w_2 \in W$ en voor alle $\lambda_1, \lambda_2 \in K$.

Bewijs. Veronderstel eerst dat W een deelruimte is van V. Dit betekent precies dat W een vectorruimte is, met als optelling de (restrictie van) de optelling op V en als scalaire vermenigvuldiging de (restrictie van) de scalaire vermenigvuldiging van V. Hieruit volgt dat als $w_1, w_2 \in W$ en $\lambda_1, \lambda_2 \in K$, het element $\lambda_1 w_1 + \lambda_2 w_2$ opnieuw in W zit.

Veronderstel omgekeerd dat W een deelverzameling is van V die voldoet aan de eigenschap dat $\lambda_1 w_1 + \lambda_2 w_2 \in W$ voor alle $w_1, w_2 \in W$ en voor alle $\lambda_1, \lambda_2 \in K$.

We merken eerst op dat als we $\lambda_1 = \lambda_2 = 1$ kiezen, er volgt dat $w_1 + w_2 \in W$ voor alle $w_1, w_2 \in W$. Hieruit volgt dat de optelling een bewerking van $W \times W \to W$ is.

Als we nu $\lambda_2 = 0$ kiezen, volgt er dat $\lambda_1 w_1 \in W$ voor alle $\lambda_1 \in K$ en $w_1 \in W$. Hieruit volgt dat de scalaire vermenigvuldiging een bewerking van $K \times W \to W$ is.

We gaan na dat de eigenschappen van een vectorruimte in Definitie 2.1.1 voldaan zijn voor W, met als optelling de (restrictie van) de optelling op V en als scalaire vermenigvuldiging de (restrictie van) de scalaire vermenigvuldiging van V.

De eigenschappen (V1) en (V4)–(V7) gelden automatisch, want $W \subseteq V$. We verifiëren de twee resterende eigenschappen:

(V2) Als we $\lambda_1 = \lambda_2 = 0$ kiezen, volgt uit Lemma 2.1.3(v) dat $0w_1 + 0w_2 = 0 \in W$. Aangezien $W \subseteq V$ is ook $0 + w = w + 0 = w$ voor alle $w \in W$.

(V3) Uit Lemma 2.1.3(vi) volgt dat $v + (-1)v = 0$ voor alle $v \in V$. Stel nu dat $w \in W$, dan is $(-1)w \in W$ (stel $\lambda_1 = 0$ en $\lambda_2 = -1$). □

Ga zelf aan de hand van het criterium in het voorgaande lemma na dat de volgende voorbeelden van deelruimten inderdaad deelruimten zijn.

Voorbeelden 2.2.4. (1) Zij V een K-vectorruimte. Dan is de nulruimte $\{0\} \subseteq V$ een deelruimte.

(2) Zij $1 \leq m \leq n$. De verzameling

$$\{(a_1, \ldots, a_n)^t \mid a_i \in K \text{ en } a_{m+1} = \cdots = a_n = 0\} \subseteq K^n$$

vormt een deelruimte van K^n.

(3) De verzameling P_d van veeltermen in $K[x]$ van graad $\leq d$, vormt een deelruimte van $K[x]$.

(4) De verzameling van continue functies op een interval,

$$\mathbf{C} := \{f : [a, b] \to \mathbb{R} \mid f \text{ is een continue functie}\},$$

is een deelruimte van de $\mathbb{R}$-vectorruimte $\mathbf{F} = \{f : [a, b] \to \mathbb{R}\}$ van alle functies van $[a, b]$ naar $\mathbb{R}$.

Definitie 2.2.5. Zij V een K-vectorruimte en $S \subseteq V$ een deelverzameling. Een element $v \in V$ noemt men een *lineaire combinatie* van elementen van de verzameling S als v kan geschreven worden als

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

met $\lambda_i \in K$ en $v_i \in S$, voor $i = 1, \ldots, n$.

Merk op dat voor iedere $S \subseteq V$ het nulelement $0 \in V$ een lineaire combinatie van elementen van de verzameling S is, namelijk

$$0 = 0v_1 + 0v_2 + \cdots + 0v_n.$$

Deze lineaire combinatie noemt men de *triviale lineaire combinatie* (van de elementen $v_1, \ldots, v_n$).

Voorbeeld 2.2.6. Zij $V = \mathbb{Q}^3$, en beschouw de elementen $v_1 = (1,0,0)^t$ en $v_2 = (3,2,0)^t$ in V. Dan is $(0,1,0)^t \in V$ wel een lineaire combinatie van elementen van de verzameling $\{v_1, v_2\}$, want

$$(0,1,0)^t = (-\tfrac{3}{2})\, v_1 + \tfrac{1}{2}\, v_2;$$

anderzijds is $(0,0,1)^t \in V$ geen lineaire combinatie van elementen van de verzameling $\{v_1, v_2\}$, want

$$\lambda_1 v_1 + \lambda_2 v_2 = (\lambda_1 + 3\lambda_2, 2\lambda_2, 0)^t$$

heeft derde coördinaat gelijk aan 0 voor alle $\lambda_1, \lambda_2 \in \mathbb{Q}$, en kan dus nooit gelijk zijn aan $(0,0,1)^t$.

Opmerking 2.2.7. (i) We benadrukken nog eens dat een lineaire combinatie van elementen van S een *eindige* som is van elementen $\lambda_i v_i$, ook als S zelf oneindig is. We zullen geregeld een dergelijke lineaire combinatie schrijven als

$$\sum_{v \in S} \lambda_v v \tag{2.1}$$

waarbij slechts eindig veel van de λ_v verschillend van nul zijn; deze laatste voorwaarde is precies nodig om te garanderen dat er in (2.1) een eindige som staat. We omschrijven dit ook met de zegswijze dat *bijna alle* $\lambda_v = 0$; die “bijna alle” betekent dus “alle op een eindig aantal na”.

(ii) We spreken vaak kortweg over “een lineaire combinatie van $v_1, \ldots, v_n$” of ook over “een lineaire combinatie van de verzameling $\{v_1, \ldots, v_n\}$” in plaats van over “een lineaire combinatie van elementen van de verzameling $\{v_1, \ldots, v_n\}$”.

(iii) Er zit een lichte dubbelzinnigheid in het gebruik van de term “lineaire combinatie”. Volgens Definitie 2.2.5 is een lineaire combinatie het *resultaat* van de uitdrukking $\lambda_1 v_1 + \cdots + \lambda_n v_n$, maar anderzijds wordt het ook gebruikt om te verwijzen naar de *uitdrukking* $\lambda_1 v_1 + \cdots + \lambda_n v_n$ zelf. (Dat laatste is bijvoorbeeld duidelijk het geval als we het hebben over de triviale lineaire combinatie.) Deze dubbelzinnigheid is courant, maar kan initieel wat verwarrend zijn. Uit de context zal telkens wel blijken welk van beide betekenissen er bedoeld wordt.

Definitie 2.2.8. Zij V een K-vectorruimte, en beschouw een deelverzameling $S \subseteq V$. We definiëren de verzameling $\operatorname{span}(S)$ als de verzameling bestaande uit alle lineaire combinaties van elementen van S, met andere woorden,

$$\begin{aligned}\operatorname{span}(S) &:= \{\lambda_1 v_1 + \cdots + \lambda_k v_k \mid k \in \mathbb{N}, \lambda_1, \ldots, \lambda_k \in K, v_1, \ldots, v_k \in S\} \\ &= \Big\{\sum_{v \in S} \lambda_v v \mid \lambda_v \in K \text{ met bijna alle } \lambda_v = 0\Big\}.\end{aligned}$$

(Zie Opmerking 2.2.7(i) voor de tweede uitdrukking.) We zeggen dat $\operatorname{span}(S)$ de deelruimte *voortgebracht door S* is.[2]

Lemma 2.2.9. *Zij V een K-vectorruimte, en zij $S \subseteq V$.*

(i) *De verzameling* $\operatorname{span}(S)$ *is een deelruimte van V.*

(ii) *Als $S \subseteq T \subseteq V$, dan is* $\operatorname{span}(S) \leq \operatorname{span}(T)$.

Bewijs. (i) Beschouw twee willekeurige elementen w_1 en w_2 in $\operatorname{span}(S)$ en twee willekeurige scalairen α en β in K. Dan is

$$\begin{aligned}w_1 &= \lambda_1 v_1 + \cdots + \lambda_k v_k, \\ w_2 &= \mu_1 u_1 + \cdots + \mu_\ell u_\ell,\end{aligned}$$

voor zekere $\lambda_i, \mu_i \in K$ en $v_i, u_i \in S$, en dus

$$\alpha w_1 + \beta w_2 = \alpha\lambda_1 v_1 + \cdots + \alpha\lambda_k v_k + \beta\mu_1 u_1 + \cdots + \beta\mu_\ell u_\ell,$$

en dit heeft opnieuw de gedaante van een element in $\operatorname{span}(S)$. Uit Lemma 2.2.3 volgt dan dat $\operatorname{span}(S) \leq V$.

(ii) Het is evident dat elk element van $\operatorname{span}(S)$ ook in $\operatorname{span}(T)$ zit, dus $\operatorname{span}(S) \subseteq \operatorname{span}(T)$. Aangezien we reeds weten dat $\operatorname{span}(S)$ een deelruimte is van V, volgt hieruit dat $\operatorname{span}(S) \leq \operatorname{span}(T)$. □

Notatie 2.2.10. (i) We noteren de deelruimte $\operatorname{span}(S)$ ook met $\langle S \rangle$. Als $S = \{v_1, \ldots, v_k\}$, dan schrijven we ook $\operatorname{span}(v_1, \ldots, v_k)$ of $\langle v_1, \ldots, v_k \rangle$ in plaats van $\operatorname{span}(\{v_1, \ldots, v_k\})$ of $\langle \{v_1, \ldots, v_k\} \rangle$.

(ii) We noteren $\operatorname{span}(\{v\})$ ook als Kv, omdat elk element van $\operatorname{span}(\{v\})$ kan geschreven worden als λv met $\lambda \in K$.

[2] We spreken af dat $\operatorname{span}(\emptyset) = \{0\}$.

Voorbeeld 2.2.11. Beschouw de vectorruimte K^3. Er geldt dat

$$\text{span}(\{(1,0,0)^t, (0,1,0)^t, (0,0,1)^t\}) = K^3$$

en dat

$$\text{span}(\{(1,0,0)^t, (0,1,0)^t\}) = \{(\alpha, \beta, 0)^t \mid \alpha, \beta \in K\} \leq K^3.$$

2.3 Lineaire (on)afhankelijkheid en voortbrengendheid

Definitie 2.3.1. Zij V een K-vectorruimte. Een deelverzameling $S \subseteq V$ noemen we een *voortbrengende verzameling* voor V als $\text{span}(S) = V$. Met andere woorden, S is een voortbrengende verzameling voor V als elk element uit V een lineaire combinatie is van elementen uit S.

Voorbeeld 2.3.2. Beschouw de vectorruimte K^3. In het voorgaande voorbeeld hebben we aangetoond dat $\{(1,0,0)^t, (0,1,0)^t, (0,0,1)^t\}$ een voortbrengende verzameling voor K^3 is. Ook

$$\{(1,0,0)^t, (0,1,0)^t, (0,0,1)^t, (1,1,1)^t\}$$

is een voortbrengende verzameling voor K^3. De verzameling $\{(1,0,0)^t, (0,1,0)^t\}$ is echter geen voortbrengende verzameling voor K^3.

Definitie 2.3.3. Zij V een K-vectorruimte met $S \subseteq V$ een deelverzameling.

(i) We noemen de verzameling S *lineair afhankelijk* als er een eindige deelverzameling $\{v_1, \ldots, v_k\} \subseteq S$ bestaat[3] waarvoor

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$$

met $\lambda_1, \ldots, \lambda_k \in K$ waarbij $\lambda_i \neq 0$ voor minstens één $i \in \{1, \ldots, k\}$.

(ii) Als een verzameling S niet lineair afhankelijk is, noemen we S *lineair onafhankelijk*. Anders gezegd, S is lineair onafhankelijk als voor elke eindige deelverzameling $\{v_1, \ldots, v_k\} \subseteq S$ (met k verschillende elementen) het enkel mogelijk is dat

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$$

met $\lambda_1, \ldots, \lambda_k \in K$ als $\lambda_1 = \cdots = \lambda_k = 0$. Een eindige verzameling elementen is dus lineair onafhankelijk als en slechts als enkel de triviale lineaire combinatie van die elementen gelijk is aan 0.

[3] Indien niet anders vermeld veronderstellen wij door de schrijfwijze $\{v_1, \ldots, v_k\}$ dat de vectoren $v_1, \ldots, v_k$ twee aan twee verschillend zijn.

Opmerking 2.3.4. (i) Ook hier zullen we vaak zeggen dat de elementen $v_1, \dots, v_n$ lineair (on)afhankelijk zijn, in plaats van te zeggen dat de verzameling $\{v_1, \dots, v_n\}$ een lineair (on)afhankelijke verzameling is.

(ii) Om te bewijzen dat een deelverzameling $S \subseteq V$ lineair onafhankelijk is, zullen we vaak de volgende redenering maken:
We nemen aan dat er een lineaire combinatie $\sum_{i=1}^{k} \lambda_i v_i = 0$ is met $v_1, \dots, v_k$ willekeurige (verschillende) elementen van S en $\lambda_1, \dots, \lambda_k$ willekeurige elementen in K, en we tonen aan dat hieruit volgt dat $\lambda_1 = \cdots = \lambda_k = 0$.

Voorbeelden 2.3.5. (1) Beschouw de vectorruimte K^n, en stel

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

De deelverzameling

$$S = \{e_1, e_2, \dots, e_n\} \subset K^n$$

is een lineair onafhankelijke verzameling. Inderdaad, veronderstel dat $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$ voor $\lambda_1, \dots, \lambda_n \in K$, dan volgt dat

$$(\lambda_1, \lambda_2, \dots, \lambda_n)^t = (0, 0, \dots, 0)^t$$

en dus is $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Merk op dat S ook een voortbrengende verzameling voor K^n is.

(2) Beschouw de vectorruimte $\mathbb{Q}^3$, en stel

$$u = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad w = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}.$$

Dan is $S = \{u, v, w\} \subset \mathbb{Q}^3$ een lineair afhankelijke verzameling. Inderdaad, $-u + 2v - w = 0$, zodat we een niet-triviale lineaire combinatie vinden van $\{u, v, w\}$ die 0 is.

(3) Als $0 \in S \subseteq V$, dan is S lineair afhankelijk. We hebben immers dat $\lambda\, 0 = 0$ voor iedere $\lambda \in K \setminus \{0\}$.

(4) Beschouw de K-vectorruimte K. Elke deelverzameling van K bestaande uit meer dan 1 element is lineair afhankelijk. Inderdaad, stel dat $a, b \neq 0$; dan is $a^{-1}a + (-b^{-1})b = 0$.

Lemma 2.3.6. *Zij V een K-vectorruimte.*

(i) *Zij $T = \{v_1, \ldots, v_n\} \subseteq V$, en zij $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ een lineaire combinatie met $\lambda_i \neq 0$ voor een zekere i. Dan is v_i een lineaire combinatie van de overige $n-1$ elementen. Omgekeerd, als v_i een lineaire combinatie is van de overige $n-1$ elementen, dan is T lineair afhankelijk.*

(ii) *Twee elementen in V zijn lineair afhankelijk als en slechts als het ene een scalair veelvoud van het andere is.*

(iii) *Als $T \subseteq V$ lineair onafhankelijk is, en $v \in V$ behoort niet tot* $\operatorname{span}(T)$, *dan is $T \cup \{v\}$ nog steeds lineair onafhankelijk.*

Bewijs. (i) Als $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ met $\lambda_i \neq 0$, dan is $v_i = \sum_{j \neq i} -\frac{\lambda_j}{\lambda_i} v_j$. Omgekeerd, als $v_i = \sum_{j \neq i} \mu_j v_j$, stel dan $\mu_i := -1$; dan is $\sum_{j=1}^{n} \mu_j v_j = 0$ en dus is T lineair afhankelijk.

(ii) Dit volgt onmiddellijk uit (i) met $n = 2$, want dan is $n - 1 = 1$ en "een lineaire combinatie van de overige $n-1$ elementen" wordt dan eenvoudigweg "een scalair veelvoud van het andere element".

(iii) We bewijzen dit uit het ongerijmde. Onderstel dat $T \cup \{v\}$ wel lineair afhankelijk zou zijn. Er bestaat dan een lineaire combinatie

$$\lambda_1 v_1 + \cdots + \lambda_n v_n + \lambda_{n+1} v = 0$$

met $v_1, \ldots, v_n \in T$ zodat niet alle λ_i gelijk aan 0. Als hierin $\lambda_{n+1} = 0$ zou zijn, dan zou er een niet-triviale lineaire combinatie van elementen in T ontstaan die 0 is, in strijd met de lineaire onafhankelijkheid van T. Dus $\lambda_{n+1} \neq 0$, en uit (i) volgt nu dat v een lineaire combinatie is van $v_1, \ldots, v_n$. Maar dan zou $v \in \operatorname{span}(T)$, strijdig. Onze onderstelling was dus fout, en dus moet $T \cup \{v\}$ lineair onafhankelijk zijn. □

2.4 Basissen, dimensie

Nu we de noodzakelijke inleidende begrippen hebben ingevoerd, komen we tot het belangrijke concept van een basis van een vectorruimte.

Definitie 2.4.1. Zij V een K-vectorruimte. Een deelverzameling $\mathcal{B} \subseteq V$ is een *basis* voor V als aan de twee volgende voorwaarden voldaan is:

(1) $\mathcal{B}$ is een voortbrengende verzameling,

(2) $\mathcal{B}$ is een lineair onafhankelijke verzameling.

Voorbeelden 2.4.2. (1) De verzameling $e_1, \dots, e_n \in K^n$, zoals gedefinieerd in Voorbeeld 2.3.5(1), is een basis voor K^n; we noemen dit de *standaardbasis* voor de vectorruimte K^n.

(2) Beschouw de vectorruimte $\mathbb{R}^2$. Dan is de verzameling $\{(1, \sqrt{2})^t, (\pi, 3)^t\}$ een basis voor $\mathbb{R}^2$.

(3) De vectorruimte K^n heeft steeds meerdere basissen (behalve als $|K| = 2$ en $n = 1$). Zo is bijvoorbeeld voor alle $\lambda \in K \setminus \{0\}$ de verzameling $\{\lambda e_1, \dots, \lambda e_n\}$ een basis voor K^n. Als K een oneindig veld is zijn er dus zelfs oneindig veel verschillende basissen.

(4) Beschouw de vectorruimte $V = K[x]$ van veeltermen in één variabele over een veld K. Dan vormt de verzameling $\{1, x, x^2, x^3, \dots\}$ een basis voor V. Merk op dat deze basis oneindig veel elementen heeft, in tegenstelling tot de drie voorgaande voorbeelden.

Opmerking 2.4.3. Als T een lineair onafhankelijke verzameling is in een vectorruimte V, dan is T een basis voor de deelruimte $\operatorname{span}(T) \leq V$. Inderdaad, T is nog steeds lineair onafhankelijk in $\operatorname{span}(T)$, en T is per definitie voortbrengend voor $\operatorname{span}(T)$.

Stelling 2.4.4. *Een deelverzameling $\mathcal{B} \subseteq V$ is een basis voor V als en slechts als elk element $v \in V$ op unieke manier te schrijven is als een lineaire combinatie van de elementen uit $\mathcal{B}$.*

Bewijs. Veronderstel eerst dat $\mathcal{B}$ een basis is. Zij $v \in V$ willekeurig. Omdat $\mathcal{B}$ voortbrengend is, kunnen we v schrijven als een lineaire combinatie

$$v = \sum_{w \in \mathcal{B}} \lambda_w w \,,$$

waarbij slechts eindig veel van de $\lambda_w \in K$ verschillend van nul zijn. Veronderstel dat we v op twee manieren kunnen schrijven als een dergelijke lineaire combinatie:

$$v = \sum_{w \in \mathcal{B}} \lambda_w w = \sum_{w \in \mathcal{B}} \mu_w w \,.$$

Dan is de eindige som

$$\sum_{w \in \mathcal{B}} (\lambda_w - \mu_w) w = 0 \,,$$

en uit het feit dat $\mathcal{B}$ lineair onafhankelijk is, volgt dat $\lambda_w = \mu_w$ voor elke $w \in \mathcal{B}$. Dit bewijst de uniciteit.

Veronderstel nu omgekeerd dat elk element van V op unieke manier te schrijven is als een lineaire combinatie van de elementen van $\mathcal{B}$. Dan volgt reeds onmiddellijk dat $\mathcal{B}$ een voortbrengende verzameling is. We tonen nu aan dat $\mathcal{B}$ een lineair onafhankelijke verzameling is. Veronderstel dat er een lineaire combinatie van verschillende elementen in $\mathcal{B}$ bestaat die nul geeft, stel

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0\,.$$

Anderzijds is $0 \in V$ ook te schrijven als

$$0 \cdot v_1 + \cdots + 0 \cdot v_n = 0\,;$$

uit de veronderstelde uniciteit volgt nu dat $\lambda_i = 0$ voor elke i, wat precies aantoont dat $\mathcal{B}$ lineair onafhankelijk is. □

We zullen zien dat elke vectorruimte een basis heeft, en dat het aantal elementen van een basis uniek is (ook al is de basis zelf verre van uniek). Dit feit is van fundamenteel belang in de lineaire algebra, en heeft als gevolg dat we kunnen spreken van dimensies (zie Definitie 2.4.10 verderop).

Definitie 2.4.5. We noemen een K-vectorruimte V een *eindig-dimensionale vectorruimte* als V een eindige voortbrengende verzameling bevat. Een vectorruimte wordt *oneindig-dimensionaal* genoemd als ze niet eindig-dimensionaal is, i.e. als elke voortbrengende verzameling voor V oneindig is.

Het volgend lemma legt een verband tussen de grootte van voortbrengende verzamelingen en lineair onafhankelijke verzamelingen.

Lemma 2.4.6 (Lemma van Steinitz[4])**.** *Zij V een eindig-dimensionale K-vectorruimte met S een eindige voortbrengende verzameling, $|S| = n \in \mathbb{N}$. Dan heeft elke lineair onafhankelijke deelverzameling van V hoogstens n elementen.*

Bewijs. Zij $U \subseteq V$ met $|U| > n$, en beschouw $n+1$ verschillende elementen $w_1, \ldots, w_{n+1}$ in U. Zij $S = \{v_1, \ldots, v_n\}$. Vermits $\operatorname{span}(S) = V$ geldt

$$\begin{aligned} w_1 &= a_{11}v_1 + \cdots + a_{1n}v_n \\ &\ \vdots \\ w_{n+1} &= a_{n+1,1}v_1 + \cdots + a_{n+1,n}v_n. \end{aligned}$$

Beschouw het stelsel lineaire vergelijkingen

$$\begin{cases} a_{11}x_1 + \cdots + a_{n+1,1}x_{n+1} = 0 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{1n}x_1 + \cdots + a_{n+1,n}x_{n+1} = 0\,. \end{cases}$$

[4]Ernst Steinitz was een Duits wiskundige (1871–1928), die onder meer een zeer invloedrijk artikel "Algebraische Theorie der Körper" schreef.

Aangezien het stelsel homogeen is, is $x_1 = \cdots = x_{n+1} = 0$ een oplossing. We noteren de uitgebreide matrix van dit stelsel met $(A|0)$. Aangezien $A \in M_{n,n+1}(K)$ kunnen er maximaal n spilplaatsen zijn in de uitgebreide matrix. Uit Gevolg 1.4.14 volgt dat het aantal vrij te kiezen onbekenden steeds groter dan of gelijk aan $(n+1)-n = 1$ is. Bijgevolg heeft dit stelsel meer oplossingen dan enkel de oplossing $x_1 = \cdots = x_{n+1} = 0$.

Zij $x_1 = c_1, \ldots, x_{n+1} = c_{n+1}$ een oplossing met niet alle $c_i = 0$. Dan is

$$c_1 w_1 + \cdots + c_{n+1} w_{n+1} = \sum_{i=1}^{n+1} c_i \Big(\sum_{j=1}^{n} a_{ij} v_j \Big) = \sum_{j=1}^{n} \Big(\sum_{i=1}^{n+1} a_{ij} c_i \Big) v_j = 0.$$

Er is dus een niet-triviale lineaire combinatie van iedere verzameling met $n+1$ elementen, of anders gezegd, elke deelverzameling van V met ten minste $n+1$ elementen is lineair afhankelijk. $\square$

We kunnen nu het bestaan van basissen bewijzen, in de volgende sterke vorm.

Stelling 2.4.7. *Zij V een eindig-dimensionale K-vectorruimte. Zij T een lineair onafhankelijke deelverzameling van V en S een voortbrengende verzameling*[5] *die T bevat. Dan is er een basis $\mathcal{B}$ voor V zodat $T \subseteq \mathcal{B} \subseteq S$.*

Bewijs. Het bewijs is constructief: we zullen de verzameling T aanvullen met elementen van S tot we een basis verkrijgen.

Indien de elementen van T voortbrengend zijn, is T reeds een basis, en hoeven we niks meer te bewijzen. Veronderstel dus dat T niet voortbrengend is, met andere woorden, $\operatorname{span}(T) \neq V$. Merk nu op dat als $S \subseteq \operatorname{span}(T)$ zou zijn, dan zou ook $\operatorname{span}(S) \leq \operatorname{span}(T)$, maar $\operatorname{span}(S) = V$ (want S is voortbrengend), in strijd met $\operatorname{span}(T) \neq V$. Dus $S \not\subseteq \operatorname{span}(T)$, en dus kunnen we een $v \in S \setminus \operatorname{span}(T)$ kiezen. Uit Lemma 2.3.6(iii) volgt nu dat $T \cup \{v\}$ nog steeds lineair onafhankelijk is.

We herhalen nu deze procedure tot T voortbrengend is (en dus een basis); dit proces eindigt zeker, omdat wegens Lemma 2.4.6 een lineair onafhankelijke verzameling een begrensd aantal elementen bevat. $\square$

Een onmiddellijk gevolg van voorgaande stelling zegt dat elke lineair onafhankelijke verzameling kan "aangevuld worden tot een basis", en dat elke voortbrengende verzameling kan "beperkt worden tot een basis". Dit is Gevolg 2.4.8.

[5] We eisen *niet* dat S zelf eindig is.

Gevolg 2.4.8. (i) *Zij V een eindig-dimensionale K-vectorruimte, en zij $T \subseteq V$ een lineair onafhankelijke verzameling. Dan bestaat er een basis $\mathcal{B}$ voor V zodat $T \subseteq \mathcal{B}$.*

(ii) *Zij V een eindig-dimensionale K-vectorruimte, en zij $S \subseteq V$ een voortbrengende verzameling. Dan bestaat er een basis $\mathcal{B}$ voor V zodat $\mathcal{B} \subseteq S$.*

Bewijs. (i) Pas Stelling 2.4.7 toe met $S = V$ (V is immers een voortbrengende verzameling voor V).

(ii) Pas Stelling 2.4.7 toe met $T = \emptyset$ ($\emptyset$ is immers een lineair onafhankelijke verzameling van V). □

Gevolg 2.4.9. *Zij $V \neq \{0\}$ een eindig-dimensionale K-vectorruimte.*

(i) *Er is een basis in V.*

(ii) *Elke basis in V is eindig, en alle basissen hebben even veel elementen.*

Bewijs. (i) Dit volgt onmiddellijk uit Gevolg 2.4.8.

(ii) Als $\mathcal{B}$ een basis is volgt uit Lemma 2.4.6 dat $|\mathcal{B}|$ eindig is. Als $\mathcal{B}$ en $\mathcal{B}'$ basissen zijn voor V dan volgt uit Lemma 2.4.6 zowel $|\mathcal{B}| \leq |\mathcal{B}'|$ als $|\mathcal{B}'| \leq |\mathcal{B}|$, dus $|\mathcal{B}| = |\mathcal{B}'|$. □

Definitie 2.4.10. Zij V een eindig-dimensionale vectorruimte over een veld K. Het aantal elementen in een basis $\mathcal{B}$ voor V noemen we de *dimensie* van V; we noteren dit als $\dim V$, of ook als $\dim_K V$ als we het veld K expliciet willen vermelden. Als $\dim V = n$, dan noemen we V een *n-dimensionale* vectorruimte.

Opmerking 2.4.11. Per definitie nemen we aan dat de lege verzameling $\emptyset$ een basis is voor de nulruimte. Elke eindig-dimensionale vectorruimte heeft dan een basis, en een vectorruimte met dimensie 0 is dan de nulruimte.

Als we de dimensie van een vectorruimte V reeds kennen en we willen nagaan of een verzameling $\mathcal{B} \subseteq V$ een basis is, volstaat het om ofwel de lineair onafhankelijkheid, ofwel de voortbrengendheid, na te gaan, uiteraard op voorwaarde dat $\mathcal{B}$ het juiste aantal elementen bevat. Dit is Stelling 2.4.12.

Stelling 2.4.12. *Zij V een n-dimensionale K-vectorruimte, en zij $S \subseteq V$ een verzameling met precies n elementen.*

(i) *Als S lineair onafhankelijk is, dan is S een basis.*

(ii) *Als S voortbrengend is, dan is S een basis.*

Bewijs. (i) Veronderstel dat S lineair onafhankelijk is. We passen Gevolg 2.4.8(i) toe en vullen S aan tot een basis $\mathcal{B}$, dus $S \subseteq \mathcal{B}$. Omdat $\dim V = n$ is $|\mathcal{B}| = n = |S|$, maar dan moet $\mathcal{B} = S$.

(ii) Veronderstel dat S voortbrengend is. We passen Gevolg 2.4.8(ii) toe en beperken S tot een basis $\mathcal{B}$, dus $\mathcal{B} \subseteq S$. Omdat $\dim V = n$ is $|\mathcal{B}| = n = |S|$, maar dan moet $\mathcal{B} = S$. □

Stelling 2.4.7 en haar gevolgen blijven geldig voor oneindig-dimensionale vectorruimten, en hoewel de lineaire algebra die nodig is dezelfde is, wordt het bewijs ervan aanzienlijk moeilijker omwille van verzameling-theoretische complicaties. We verwijzen voor de wiskundigen naar de latere cursus "Logica".

Stelling 2.4.13. *Zij V een oneindig-dimensionale K-vectorruimte.*

(i) *Zij T een lineair onafhankelijke deelverzameling van V en S een voortbrengende verzameling die T bevat. Dan is er een basis $\mathcal{B}$ voor V zodat $T \subseteq \mathcal{B} \subseteq S$.*

(ii) *Er is een basis in V.*

(iii) *Elke basis in V is oneindig, en alle basissen hebben even veel*[6] *elementen.*

Voorbeelden 2.4.14. (1) De nulruimte heeft dimensie 0. Een vectorruimte V niet gelijk aan de nulruimte heeft dimensie ≥ 1.

(2) De vectorruimte K^n heeft dimensie n.

(3) De K-vectorruimte $M_{m,n}(K)$ heeft dimensie mn. Immers, de matrices U_{ij} waarin er een 1 staat op de (i,j)-de plaats en waarvan alle andere componenten gelijk zijn aan nul, vormen een basis vermits elke matrix een unieke lineaire combinatie is van de U_{ij}'s:

$$A = (a_{ij}) = a_{11}U_{11} + \cdots + a_{mn}U_{mn} = \sum_{i,j} a_{ij}U_{ij}\,.$$

(4) Beschouw de vectorruimte $K[x]$ van de veeltermen in één variabele over een veld K. Zoals we gezien hebben in Voorbeeld 2.4.2(4) is de verzameling $\{1, x, x^2, x^3, x^4, \dots\}$ van alle machten van de variabele x een basis voor deze ruimte. Bijgevolg is $K[x]$ een oneindig-dimensionale vectorruimte over K. De deelruimte van veeltermen van graad $< d$ heeft dimensie d; de verzameling $\{1, x, x^2, \dots, x^{d-1}\}$ vormt een basis voor deze deelruimte.

[6]Met "even veel" bedoelen we hier dat de kardinaliteiten dezelfde zijn, of nog, dat er een bijectie bestaat tussen elke twee basissen. Dit is sterker dan de uitspraak dat de basissen oneindig groot zijn.

2.5 Som en directe som van vectorruimten

Als $\mathcal{B}$ een basis is voor de n-dimensionale vectorruimte V, dan is, per definitie, elk element van V op een unieke manier te schrijven als een som van elementen uit de 1-dimensionale deelruimten Kv met $v \in \mathcal{B}$. We willen dit idee veralgemenen naar willekeurige deelruimten.

Definitie 2.5.1. Zij V een K-vectorruimte en zij $W_1, \dots, W_k \leq V$ deelruimten van V. De *som van de deelruimten* $W_1, \dots, W_k$ is gedefinieerd als

$$W_1 + \cdots + W_k := \left\{ \sum_{i=1}^{k} w_i \;\middle|\; w_i \in W_i \text{ voor alle } 1 \leq i \leq k \right\}.$$

Als ieder element v van $W := W_1 + \cdots + W_k$ op een unieke manier te schrijven is als $v = \sum_{i=1}^{k} w_i$ met $w_i \in W_i$, dan zeggen we dat W een *directe som van deelruimten* is en noteren dit als

$$W = W_1 \oplus \cdots \oplus W_k \,.$$

Merk op dat de (directe) som van een eindig aantal deelruimten van V opnieuw een deelruimte van V is. We noteren ook

$$\sum_{i=1}^{k} W_i := W_1 + \cdots + W_k \quad \text{en} \quad \bigoplus_{i=1}^{k} W_i := W_1 \oplus \cdots \oplus W_k.$$

Lemma 2.5.2. *Zij V een K-vectorruimte en zij W_1, W_2 en W drie deelruimten van V. Dan is $W = W_1 \oplus W_2$ als en slechts als $W = W_1 + W_2$ en $W_1 \cap W_2 = \{0\}$.*

Bewijs. Veronderstel eerst dat $W = W_1 \oplus W_2$; dan geldt per definitie reeds $W = W_1 + W_2$. Stel $w \in W_1 \cap W_2$; dan volgt uit $0 = 0 + 0 = w + (-w)$ dat $w = 0$, vermits de elementen van W maar op één manier te schrijven zijn als een som van een element uit W_1 en een element uit W_2.

Stel nu omgekeerd dat $W = W_1 + W_2$ en $W_1 \cap W_2 = \{0\}$. Er moet enkel nog aangetoond worden dat elk element van W niet op twee verschillende manieren te schrijven is als een som van elementen uit W_1 en W_2. Stel $w_1 + w_2 = w_1' + w_2'$ met $w_1, w_1' \in W_1$ en $w_2, w_2' \in W_2$. Dan is $w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{0\}$, dus $w_1 = w_1'$ en $w_2 = w_2'$. $\square$

Voorbeeld 2.5.3. Beschouw de vectorruimte $V = \mathbb{R}^3$, met deelruimten

$$W_1 = \langle (1,0,0)^t, (0,1,0)^t \rangle,$$

$$W_2 = \langle (1,2,3)^t, (2,3,4)^t \rangle,$$
$$W_3 = \langle (1,1,1)^t \rangle.$$

Dan is $V = W_1 + W_2$ en ook $V = W_1 + W_3$ maar $V \neq W_2 + W_3$. Bovendien is $V = W_1 \oplus W_3$, maar $V \neq W_1 \oplus W_2$ want $W_1 \cap W_2 \neq \{0\}$.

De volgende karakterisatie van directe sommen is erg nuttig. Het toont aan dat directe sommen inderdaad een natuurlijke veralgemening vormen van de notie van lineaire onafhankelijkheid.

Stelling 2.5.4. *Zij V een K-vectorruimte en zij $W_1, \dots, W_k \leq V$ deelruimten van V. Stel $W := W_1 + \cdots + W_k$. Dan is $W = W_1 \oplus \cdots \oplus W_k$ als en slechts als de enige manier om $0 \in W$ te schrijven als $0 = w_1 + \cdots + w_k$ met elke $w_i \in W_i$, de triviale manier is, d.w.z., $w_1 = \cdots = w_k = 0$.*

Bewijs. Als $W = W_1 \oplus \cdots \oplus W_k$, dan is ieder element $v \in W$ uniek te schrijven als een som $v = \sum_{i=1}^k w_i$ met $w_i \in W_i$, dus in het bijzonder geldt dit voor $v = 0$.

Veronderstel dus omgekeerd dat de enige manier om $0 \in W$ te schrijven als $0 = w_1 + \cdots + w_k$ met elke $w_i \in W_i$, de triviale manier is. Neem $v \in W$ willekeurig, dus $v = w_1 + \cdots + w_k$ met elke $w_i \in W_i$. Als we v ook nog zouden kunnen schrijven als $v = w'_1 + \cdots + w'_k$ met elke $w'_i \in W_i$, dan volgt hieruit dat
$$0 = v - v = (w_1 - w'_1) + \cdots + (w_i - w'_i),$$
waarbij elke $w_i - w'_i \in W_i$. Wegens onze veronderstelling kan dit enkel maar als elke $w_i - w'_i = 0$, en dus $w_i = w'_i$, voor elke i. Hieruit volgt dat v *uniek* te schrijven is als $v = w_1 + \cdots + w_k$ met elke $w_i \in W_i$, en dus is de som een directe som. □

Stelling 2.5.5. *Zij V een K-vectorruimte en zij $W_1, \dots, W_k$ deelruimten van V. Onderstel dat de som W van deze deelruimten een directe som $W = W_1 \oplus \cdots \oplus W_k$ is.*

(i) *Als $\mathcal{B}_i$ een basis is voor W_i, voor elke i, dan is $\mathcal{B} := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ een basis voor W.*

(ii) $\dim W = \dim W_1 + \cdots + \dim W_k$.

Bewijs. (i) Het is duidelijk dat $\mathcal{B}$ voortbrengend is voor W. Inderdaad, stel $v \in W$ willekeurig, dan is $v = w_1 + \cdots + w_k$ met elke $w_i \in W_i$. Aangezien elke w_i een lineaire combinatie is van $\mathcal{B}_i$, is $v = w_1 + \cdots + w_k$ een lineaire combinatie van $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k = \mathcal{B}$.

Vervolgens gaan we na dat $\mathcal{B}$ lineair onafhankelijk is. Stel dat er een lineaire combinatie

$$\sum_{v\in\mathcal{B}} \lambda_v v = \sum_{v\in\mathcal{B}_1} \lambda_v v + \cdots + \sum_{v\in\mathcal{B}_k} \lambda_v v = 0.$$

Merk op dat elke term $w_i := \sum_{v\in\mathcal{B}_i} \lambda_v v$ bevat is in W_i omdat $\mathcal{B}_i$ voortbrengend is voor W_i. Omdat $W = W_1 \oplus \cdots \oplus W_k$ volgt nu uit Stelling 2.5.4 dat elke $w_i = 0$. Omdat elke $\mathcal{B}_i$ lineair onafhankelijk is, volgt hieruit dat alle $\lambda_v = 0$. We besluiten dat $\mathcal{B}$ inderdaad lineair onafhankelijk is.

(ii) Dit volgt nu uit de definitie van dimensie omdat wegens (i) geldt dat $|\mathcal{B}| = |\mathcal{B}_1| + \cdots + |\mathcal{B}_k|$. □

We hebben ook een soort omgekeerde van Stelling 2.5.5, waarbij we een directe som verkrijgen door een basis "in stukken te splitsen":

Stelling 2.5.6. *Zij V een K-vectorruimte met basis $\mathcal{B}$, en veronderstel dat $\mathcal{B}$ de* disjuncte unie *is van $\mathcal{B}_1, \ldots, \mathcal{B}_k$, i.e. $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ en $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$ voor alle $i \neq j$. Dan is $V = \operatorname{span}(\mathcal{B}_1) \oplus \cdots \oplus \operatorname{span}(\mathcal{B}_k)$.*

Bewijs. Stel $W_i := \operatorname{span}(\mathcal{B}_i)$ voor elke i. We zullen aantonen dat enerzijds $V = W_1 + \cdots + W_k$, en dat anderzijds het criterium van Stelling 2.5.4 voldaan is.

Om aan te tonen dat $V = W_1 + \cdots + W_k$ beschouwen we een willekeurige $u \in V$. Aangezien $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ voortbrengend is voor V, kunnen we u schrijven als

$$u = \sum_{v\in\mathcal{B}_1} \lambda_v v + \cdots + \sum_{v\in\mathcal{B}_k} \lambda_v v.$$

Door $w_i := \sum_{v\in\mathcal{B}_i} \lambda_v v$ te stellen, voor elke i, zien we dat inderdaad $u = w_1 + \cdots + w_k$ met elke $w_i \in W_i$.

Onderstel nu dat $0 = w_1 + \cdots + w_k$ voor zekere $w_i \in W_i$; we moeten aantonen dat dan noodzakelijk $w_i = 0$ voor alle i. Aangezien $W_i = \operatorname{span}(\mathcal{B}_i)$ kunnen we elke w_i schrijven als een lineaire combinatie van $\mathcal{B}_i$, stel $w_i = \sum_{v\in\mathcal{B}_i} \lambda_v v$. Uit $0 = w_1 + \cdots + w_k$ volgt dan opnieuw

$$\sum_{v\in\mathcal{B}_1} \lambda_v v + \cdots + \sum_{v\in\mathcal{B}_k} \lambda_v v = 0,$$

en omdat $\mathcal{B}$ lineair onafhankelijk is, kan dit enkel als alle $\lambda_v = 0$, en dus zijn inderdaad alle $w_i = 0$. □

Ook indien $W_1 \cap W_2 \neq \{0\}$, is het mogelijk om een formule af te leiden voor $\dim(W_1 + W_2)$. Deze formule staat bekend als de *dimensiestelling voor deelruimten* of de *dimensiestelling van Grassmann*[7] .

Stelling 2.5.7 (Dimensiestelling voor deelruimten)**.** *Zij V een eindig-dimensionale K-vectorruimte en W_1, W_2 twee deelruimten van V. Dan is*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Bewijs. We bewijzen dit door een basis van $W_1 + W_2$ op te stellen waaruit we de te bewijzen dimensie-formule kunnen aflezen. We noteren $n := \dim(W_1)$, $m := \dim(W_2)$ en $k := \dim(W_1 \cap W_2)$.

Zij $\mathcal{C} = \{v_1, \dots, v_k\}$ een basis van $W_1 \cap W_2$. Aangezien $\mathcal{C} \subseteq W_1 \cap W_2 \subseteq W_1$, volgt uit Gevolg 2.4.8(i) dat men $\mathcal{C}$ kan uitbreiden tot een basis $\mathcal{B}_1$ van W_1. We noteren $\mathcal{B}_1 = \{v_1, \dots, v_k, w_1, \dots, w_{n-k}\}$. Op analoge manier kunnen we $\mathcal{C}$ ook uitbreiden tot een basis $\mathcal{B}_2$ van W_2, we noteren $\mathcal{B}_2 = \{v_1, \dots, v_k, z_1, \dots, z_{m-k}\}$.

In de rest van dit bewijs tonen we aan dat

$$\mathcal{B} := \{v_1, \dots, v_k, w_1, \dots, w_{n-k}, z_1, \dots, z_{m-k}\}$$

een basis is van $W_1 + W_2$.

- We gaan na dat $\mathcal{B}$ een voortbrengende verzameling voor $W_1 + W_2$ is. We nemen een willekeurig element $x \in W_1 + W_2$, dus $x = y_1 + y_2$ met $y_1 \in W_1$ en $y_2 \in W_2$. Aangezien $\mathcal{B}_1$ voortbrengend is voor W_1, is $y_1 \in \operatorname{span}(\mathcal{B}_1)$; analoog is $y_2 \in \operatorname{span}(\mathcal{B}_2)$. Hieruit volgt dat $x = y_1 + y_2 \in \operatorname{span}(\mathcal{B}_1 \cup \mathcal{B}_2) = \operatorname{span}(\mathcal{B})$; we besluiten dat $\mathcal{B}$ voortbrengend is voor $W_1 + W_2$.
- We gaan na dat $\mathcal{B}$ een lineair onafhankelijke verzameling is. We stellen

  $$U_1 := \operatorname{span}(w_1, \dots, w_{n-k}) \leq W_1 \quad \text{en} \quad U_2 := \operatorname{span}(z_1, \dots, z_{m-k}) \leq W_2.$$

  Merk vooreerst op dat, wegens Stelling 2.5.6, uit de constructie van $\mathcal{B}_1$ en $\mathcal{B}_2$ volgt dat

  $$\begin{aligned} (W_1 \cap W_2) \cap U_1 &= \{0\}, \\ (W_1 \cap W_2) \cap U_2 &= \{0\}. \end{aligned} \tag{2.2}$$

  Stel nu dat

[7] Hermann Günther Grassmann (1809–1877) was een Duitse polymath, die in zijn eigen tijd als taalkundige bekendstond, en nu vooral beroemd is omwille van zijn wiskundige bijdragen. Naast zijn werk als leraar middelbaar onderwijs was hij tevens natuurkundige, neohumanist en uitgever.

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 w_1 + \cdots + \beta_{n-k} w_{n-k} \\ + \gamma_1 z_1 + \cdots + \gamma_{m-k} z_{m-k} = 0 \quad (2.3)$$

voor scalairen $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-k}, \gamma_1, \ldots, \gamma_{m-k} \in K$, en stel

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_k v_k \in W_1 \cap W_2; \\ w &= \beta_1 w_1 + \cdots + \beta_{n-k} w_{n-k} \in U_1 \leq W_1; \\ z &= \gamma_1 z_1 + \cdots + \gamma_{n-k} z_{m-k} \in U_2 \leq W_2. \end{aligned}$$

dan is $v + w + z = 0$, en dus $v + w = -z \in W_1 \cap W_2$, waaruit volgt dat $z \in (W_1 \cap W_2) \cap U_2$, en uit (2.2) volgt dat $z = 0$. Analoog is $w \in (W_1 \cap W_2) \cap U_1$, en opnieuw uit (2.2) volgt dat $w = 0$. Dus $v = w = z = 0$, en omdat zowel $\mathcal{B}_1$ als $\mathcal{B}_2$ lineair onafhankelijke verzamelingen zijn, kan dit enkel als alle coëfficiënten α_i, β_i en γ_i gelijk zijn aan 0.

We hebben aangetoond dat $\mathcal{B}$ een basis is van $W_1 + W_2$. Dit bewijst het gestelde, aangezien $|\mathcal{B}| = k + (n - k) + (m - k) = n + m - k$. □

Opmerking 2.5.8. Opgelet: deze dimensiestelling kan *niet* worden uitgebreid tot meer dan twee deelruimten. Er geldt dus geen "inclusie-exlusie-principe" zoals bij de kardinaliteit van de unie van verzamelingen. De reden hiervoor is dat voor deelruimten U, W_1, W_2 in het algemeen *niet* geldt dat $U \cap (W_1 + W_2) = U \cap W_1 + U \cap W_2$. (Ga dit zelf na.)

Definitie 2.5.9. Zij V een K-vectorruimte en $W \leq V$ een deelruimte. Een deelruimte $W' \leq V$ is een *complement*[8] van W in V als $V = W \oplus W'$.

Merk op dat een complement van een deelruimte niet uniek bepaald is (zie Voorbeeld 2.5.11). Anders gezegd, uit $V = W \oplus W_1 = W \oplus W_2$ volgt *niet* dat $W_1 = W_2$.

Complementen bestaan wel altijd:

Lemma 2.5.10. *Zij W, U deelruimten van een K-vectorruimte V. Als $V = W + U$ dan bestaat er een deelruimte $Y \leq U$ zodat $V = W \oplus Y$. In het bijzonder heeft elke deelruimte $W \leq V$ een complement in V.*

Bewijs. Kies een basis $\mathcal{B}_W$ voor W en een basis $\mathcal{B}_U$ voor U. We passen Stelling 2.4.7 toe (of Stelling 2.4.13(i) indien V oneindig-dimensionaal is) met $T = \mathcal{B}_W$ en $S = \mathcal{B}_W \cup \mathcal{B}_U$, en vinden een basis voor V van de vorm

[8] Merk op dat dit *niet* het verzameling-theoretische complement $V \setminus W$ is; dit laatste is zelfs geen deelruimte.

$\mathcal{B}_W \cup \mathcal{C}$ met $\mathcal{C} \subseteq \mathcal{B}_U$. Neem $Y = \operatorname{span}(\mathcal{C})$, dan is $Y \leq U$ en elk element van V kan op een unieke manier geschreven worden als $w + y$ met $w \in W$ en $y \in Y$, dus $V = W \oplus Y$.

De laatste uitspraak volgt uit de eerste door $U = V$ te nemen. $\square$

Voorbeeld 2.5.11. Zij V de vectorruimte $V = \mathbb{R}^2$, en beschouw de volgende deelruimten van V:

$$\begin{aligned} W_1 &= \left\{ \left(\begin{smallmatrix} a \\ 0 \end{smallmatrix} \right) \mid a \in \mathbb{R} \right\}; \\ W_2 &= \left\{ \left(\begin{smallmatrix} 0 \\ b \end{smallmatrix} \right) \mid b \in \mathbb{R} \right\}; \\ W_3 &= \left\{ \left(\begin{smallmatrix} c \\ c \end{smallmatrix} \right) \mid c \in \mathbb{R} \right\}; \\ W_4 &= \left\{ \left(\begin{smallmatrix} d \\ 2d \end{smallmatrix} \right) \mid d \in \mathbb{R} \right\}; \end{aligned}$$

dan is elke W_i een complement van elke W_j met $j \neq i$. In het bijzonder zien we dat een complement van een deelruimte niet uniek bepaald is.

Tot nu toe hebben we het enkel gehad over de som en directe som van deelruimten van een vaste K-vectorruimte V. We definiëren nu de "uitwendige" directe som van een eindig aantal (verschillende) K-vectorruimten.

Definitie 2.5.12. Zij $W_1, \dots, W_m$ vectorruimten over K. Definieer de verzameling

$$\bigoplus_{i=1}^{m} W_i = W_1 \oplus \cdots \oplus W_m = \left\{ (a_1, \dots, a_m) \mid a_1 \in W_1, \dots, a_m \in W_m \right\}.$$

Definieer op $\bigoplus_{i=1}^m W_i$ de volgende optelling en vermenigvuldiging met scalairen,

$$(a_1, \dots, a_m) + (b_1, \dots, b_m) = (a_1 + b_1, \dots, a_m + b_m)$$

en

$$\lambda \cdot (a_1, \dots, a_m) = (\lambda a_1, \dots, \lambda a_m)$$

voor alle $a_i, b_i \in W_i$ en $\lambda \in K$.

We noemen $\bigoplus_{i=1}^m W_i$ de *directe som* van de vectorruimten $W_1, \dots, W_m$.

Lemma 2.5.13. *Zij $W_1, \dots, W_m$ vectorruimten over K. De verzameling $\bigoplus_{i=1}^m W_i$ met optelling en scalaire vermenigvuldiging zoals gedefinieerd in Definitie* 2.5.12*, is een K-vectorruimte.*

Bewijs. Oefening. $\square$

Opmerking 2.5.14. Wanneer de K-vectorruimten $W_1, \ldots, W_m$ allemaal deelruimten zijn van een bepaalde vectorruimte V dan noemen we de directe som $\bigoplus_{i=1}^m W_i$, zoals gedefinieerd in Definitie 2.5.1, ook wel de *inwendige directe som*.

De directe som $\bigoplus_{i=1}^m W_i$ in Definitie 2.5.12 voor willekeurige K-vectorruimten $W_1, \ldots, W_m$ noemen we ook wel de *uitwendige directe som*.

Het volgende lemma toont aan dat een uitwendige directe som ook een inwendige directe som is zoals in Definitie 2.5.1.

Lemma 2.5.15. *Zij $W_1, \ldots, W_m$ vectorruimten over K, en stel $V = \bigoplus_{i=1}^m W_i$ de uitwendige directe som van $W_1, \ldots, W_m$. Dan is V de inwendige directe som van de vectorruimten*

$$W_i' = \big\{(0, \ldots, a_i, \ldots, 0) \mid a_i \in W_i\big\},$$

i.e. $V = W_1' \oplus \cdots \oplus W_m'$.

Bewijs. Dit volgt uit het feit dat, voor alle $(a_1, \ldots, a_m) \in \bigoplus_{i=1}^m W_i$,

$$(a_1, a_2, \ldots, a_m) = (a_1, 0, \ldots, 0) + (0, a_2, \ldots, 0) + \cdots + (0, 0, \ldots, a_m)$$

de unieke manier is om $(a_1, \ldots, a_m)$ te schrijven als som van elementen uit $W_1', \ldots, W_m'$. $\square$

Opmerking 2.5.16. Voor de geïnteresseerde lezer vermelden we dat Definitie 2.5.12 als volgt kan veralgemeend worden naar een oneindig aantal vectorruimten. Zij I een verzameling die de rol speelt van *indexverzameling*, d.w.z. dat we de elementen gebruiken om objecten te nummeren. Typische voorbeelden zijn $I = \{1, \ldots, n\}$ of $I = \mathbb{N}$. Zij $\{W_i\}_{i \in I}$ een familie van K-vectorruimten, i.e. voor elke $i \in I$ is een K-vectorruimte W_i gegeven. Een element dat we verkrijgen door uit elke W_i een element a_i te kiezen[9], noteren we als $(a_i)_{i \in I}$. De verzameling

$$\prod_{i \in I} W_i = \big\{(a_i)_{i \in I} \mid a_i \in W_i\big\}$$

met daarop de componentsgewijze optelling en vermenigvuldiging met scalairen,

$$(a_i)_{i \in I} + (b_i)_{i \in I} = (a_i + b_i)_{i \in I}, \quad \lambda(a_i)_{i \in I} = (\lambda a_i)_{i \in I},$$

[9]Formeel betekent dit "kiezen" dat we een afbeelding $a \colon I \to \bigcup_{i \in I} W_i$ beschouwen zodat $a(i) \in W_i$ voor alle $i \in I$.

noemt men het *direct product* van de familie $\{W_i\}_{i\in I}$. Men verifieert dat $\prod_{i\in I} W_i$ een K-vectorruimte is.

De verzameling

$$\bigoplus_{i\in I} W_i = \Big\{(a_i)_{i\in I} \in \prod_{i\in I} W_i \mid \text{bijna alle } a_i = 0\Big\},$$

noemt men de *(uitwendige) directe som* van de familie $\{W_i\}_{i\in I}$. (Met *bijna alle* wordt bedoeld: alle op een eindig aantal na.) De verzameling $\bigoplus_{i\in I} W_i$ is een deelruimte van $\prod_{i\in I} W_i$.

Voor eindige families $\{W_i\}_{i\in I}$ geldt $\prod_{i\in I} W_i = \bigoplus_{i\in I} W_i$, en vinden we onze oorspronkelijke Definitie 2.5.12 terug.

Hoofdstuk

3 Lineaire afbeeldingen

Voor we de definitie geven van een lineaire afbeelding tussen twee vectorruimten, voeren we eerst nog enkele begrippen in in verband met afbeeldingen tussen twee verzamelingen.

3.1 Afbeeldingen en hun eigenschappen

Definitie 3.1.1. Zij A en B twee verzamelingen.

(i) Zij $f\colon A \to B$ een afbeelding, en zij $C \subseteq A$. We noteren

$$f(C) := \{f(c) \mid c \in C\} \subseteq B.$$

(ii) Zij $f\colon A \to B$ een afbeelding, en zij $D \subseteq B$. Het *inverse beeld* van D is de verzameling van alle elementen in A die op een element in D afgebeeld worden. We noteren[1]

$$f^{-1}(D) := \{a \in A \mid f(a) \in D\} \subseteq A.$$

(iii) Zij $f\colon A \to B$ een afbeelding, en zij $C \subseteq A$. Dan is $C \to B\colon c \mapsto f(c)$ ook een afbeelding; we noemen deze afbeelding de *restrictie* van f tot C. We noteren deze afbeelding met

$$f|_C\colon C \to B\colon c \mapsto f(c).$$

(iv) Een afbeelding $f\colon A \to B$ wordt *injectief* genoemd, indien elk element van B hoogstens één maal bereikt wordt door f, of nog, indien voor alle $a, a' \in A$ geldt dat

$$f(a) = f(a') \implies a = a';$$

we noemen f dan een *injectie* van A naar (of in) B.

[1] Opgelet, f^{-1} is géén afbeelding van B naar A, want een element van B wordt afgebeeld op een *deelverzameling* van A, die niet noodzakelijk uit 1 element bestaat. Zie echter anderzijds Definitie 3.1.4 verderop.

(v) Een afbeelding $f\colon A \to B$ wordt *surjectief* genoemd, indien elk element van B minstens één maal bereikt wordt door f, of nog, indien voor alle $b \in B$ geldt:

$$\text{er bestaat een } a \in A \text{ zodat } f(a) = b.$$

we noemen f dan een *surjectie* van A naar (of op) B.

(vi) Een afbeelding $f\colon A \to B$ wordt *bijectief* genoemd, indien elk element van B juist één maal bereikt wordt door f, of nog, indien voor alle $b \in B$ geldt:

$$\text{er bestaat juist één } a \in A \text{ zodat } f(a) = b.$$

we noemen f dan een *bijectie* van A naar B (of tussen A en B).

(vii) Voor elke verzameling A hebben we de *identieke afbeelding*

$$\mathbf{1}_A\colon A \to A\colon a \mapsto a.$$

Soms gebruiken we hiervoor ook de notatie id_A, of kortweg $\mathbf{1}$ of id als de verzameling duidelijk is uit de context.

Opmerking 3.1.2. Beschouw twee afbeeldingen $f\colon A \to B$ en $g\colon A \to B$. Dan geldt:

$$f = g \iff f(a) = g(a) \text{ voor alle } a \in A.$$

Lemma 3.1.3. (i) *Een afbeelding $f\colon A \to B$ is bijectief dan en slechts dan als ze injectief én surjectief is.*

(ii) *Als f een bijectie is van A naar B, dan is $|A| = |B|$, d.w.z. A en B hebben even veel elementen.*

Bewijs. Dit volgt rechtstreeks uit de definities. □

Definitie 3.1.4. Zij $f\colon A \to B$ een bijectie. Dan bestaat, voor iedere $b \in B$, de verzameling $f^{-1}(\{b\}) = \{a \in A \mid f(a) = b\}$ uit precies één element. We definiëren dan de afbeelding

$$f^{-1}\colon B \to A\colon b \mapsto \text{het unieke element in } f^{-1}(\{b\}).$$

We noemen f^{-1} de *inverse afbeelding* van f. Om die reden noemen we een bijectieve afbeelding ook wel een *inverteerbare* afbeelding.

De inverteerbare lineaire afbeeldingen zijn precies die afbeeldingen die inverteerbaar zijn met betrekking tot de bewerking “samenstelling”.

Definitie 3.1.5. Zij A, B, C drie verzamelingen, en zij g een afbeelding van A naar B en f een afbeelding van B naar C. We definiëren de afbeelding

$$f \circ g \colon A \to C \colon a \mapsto f(g(a)).$$

We noemen deze afbeelding de *samenstelling* van f en g, we spreken $f \circ g$ uit als f *na* g.

Lemma 3.1.6. *Zij $f \colon A \to B$ een afbeelding. Dan zijn de volgende uitspraken equivalent:*

(a) *Er bestaat een afbeelding $g \colon B \to A$ zodat $f \circ g = \mathbf{1}_B$ en $g \circ f = \mathbf{1}_A$.*

(b) *f is bijectief.*

Indien deze uitspraken voldaan zijn, dan is $g = f^{-1}$.

Bewijs. Veronderstel eerst dat (a) geldt. Dan is de afbeelding f injectief, want uit $f(a) = f(a')$ volgt dat $g(f(a)) = g(f(a'))$, en bijgevolg is $\mathbf{1}_A(a) = \mathbf{1}_A(a')$ en dus is $a = a'$. De afbeelding f is ook surjectief, want stel $b \in B$ willekeurig, dan is $b = \mathbf{1}_B(b) = f(g(b))$, en dus is b het beeld onder f van het element $g(b) \in A$. Dus f is bijectief, i.e. (b) geldt.

Veronderstel nu omgekeerd dat (b) geldt, en beschouw de inverse afbeelding $f^{-1} \colon B \to A$, die elk element $b \in B$ afbeeldt op het unieke element in A met $f(a) = b$. Dan is uiteraard $f(f^{-1}(b)) = b$ voor alle $b \in B$ en $f^{-1}(f(a)) = a$ voor alle $a \in A$, en dus geldt (a) met $g = f^{-1}$.

Ten slotte merken we op dat als f bijectief is en (a) geldt voor een zekere $g \colon B \to A$, dan is $f \circ g = \mathbf{1}_B = f \circ f^{-1}$, en uit de injectiviteit van f volgt dan $g = f^{-1}$. □

We komen nu tot de belangrijke definitie van *lineaire afbeeldingen*, met name de afbeeldingen tussen twee vectorruimten die structuur-bewarend zijn.

Definitie 3.1.7. Zij V en W twee vectorruimten over het veld K.

(i) Een afbeelding $f \colon V \to W$ is een *lineaire afbeelding* als

$$f(\lambda v_1 + \mu v_2) = \lambda f(v_1) + \mu f(v_2)$$

voor alle $\lambda, \mu \in K$ en alle $v_1, v_2 \in V$. Merk op dat in de bovenstaande identiteit de eerste $+$ de optelling is in de vectorruimte V en dat de tweede $+$ de optelling is in de vectorruimte W; een gelijkaardige opmerking geldt voor de scalaire vermenigvuldiging.
Een lineaire afbeelding wordt ook een *morfisme* van vectorruimten genoemd.

(ii) Een lineaire afbeelding $f\colon V \to V$ van een vectorruimte V naar zichzelf noemen we een *lineaire operator* op V, of ook soms een *endomorfisme* van V.

Opmerking 3.1.8. (i) Een lineaire afbeelding $f\colon V \to W$ zet lineaire combinaties om in lineaire combinaties. Inderdaad,

$$f\Big(\sum_{i=1}^{k} \lambda_i v_i\Big) = \sum_{i=1}^{k} \lambda_i f(v_i)$$

voor alle $v_1, \dots, v_k \in V$ en $\lambda_1, \dots, \lambda_k \in K$.

(ii) Zij $f\colon V \to W$ een lineaire afbeelding en zij $U \leq V$ een deelvectorruimte. Dan is de restrictie $f|_U\colon U \to W$ ook een lineaire afbeelding.

(iii) Zij V en W twee K-vectorruimten, en zij $f\colon V \to W$ een lineaire afbeelding. Dan is $f(0_V) = 0_W$. Inderdaad, beschouw de nulelementen $0_V \in V$, $0_W \in W$ en ook $0_K \in K$; uit de lineariteit van f volgt dan dat $f(0_V) = f(0_K 0_V) = 0_K f(0_V) = 0_W$.

Voorbeelden 3.1.9. (1) Zij $V = \mathbb{R}^2$ en $W = \mathbb{R}^3$. De afbeelding

$$f\colon V \to W\colon (x,y)^t \mapsto (x+y, 2x, -y)^t$$

is een lineaire afbeelding. Anderzijds is de afbeelding

$$g\colon V \to W\colon (x,y)^t \mapsto (x+y, 2+x, -y)^t$$

niet lineair.

(2) Zij V en W twee K-vectorruimten. De *nulafbeelding* $0\colon V \to W\colon v \mapsto 0_W$ is een lineaire afbeelding.

(3) Zij V een K-vectorruimte. De *identieke afbeelding* $\mathbf{1}_V\colon V \to V\colon v \mapsto v$ is een lineaire operator.

(4) Zij K een veld en $V = K^n$, $W = K^m$. De linkse vermenigvuldiging met een matrix $A \in M_{m,n}(K)$ is een lineaire afbeelding van V naar W. We noteren deze met

$$L_A\colon K^n \to K^m\colon v \mapsto Av.$$

We verifiëren dat dit een lineaire afbeelding is; we maken gebruik van Lemma 1.3.6:

$$L_A(\lambda v + \mu w) = A(\lambda v + \mu w) = \lambda Av + \mu Aw = \lambda L_A(v) + \mu L_A(w),$$

voor alle $v, w \in K^n$, $\lambda, \mu \in K$. We zullen verder zien dat elke lineaire afbeelding tussen eindig-dimensionale vectorruimten in zekere zin overeenkomt met de linkse vermenigvuldiging met een matrix (zie Stelling 4.1.8). Merk op dat we bijvoorbeeld de $f\colon V \to W$ uit (1) kunnen herschrijven als L_A met

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

(5) Zij $P_n \leq K[x]$, de vectorruimte van de veeltermen van graad $\leq n$. De *afleiding*

$$\frac{d}{dx}\colon P_n \to P_{n-1}\colon f(x) = \sum_{i=0}^{n} a_i x^i \mapsto \frac{df}{dx} = \sum_{i=1}^{n} i \cdot a_i x^{i-1},$$

is een lineaire afbeelding. De afbeelding $\frac{d}{dx}\colon K[x] \to K[x]$ is een lineaire operator op de vectorruimte $K[x]$.

(6) Zij $V = K^n$. De afbeelding

$$S\colon (\lambda_1, \ldots, \lambda_n)^t \mapsto (0, \lambda_1, \ldots, \lambda_{n-1})^t$$

is een lineaire operator op V. We noemen deze afbeelding de *shiftoperator.*

(7) Beschouw het Euclidisch vlak $V = \mathbb{R}^2$. De *rotatie* rond de oorsprong rond een willekeurige hoek θ is een lineaire afbeelding, namelijk

$$\rho_\theta\colon V \to V\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

Merk op dat $\rho_\theta = L_{A_\theta}$, waarbij $A_\theta = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$.

(8) Stel dat $V = \bigoplus_{i=1}^k W_i$ voor deelruimten $W_i \leq V$. Dan is de afbeelding

$$p_i\colon V \to W_i\colon v = w_1 + \cdots + w_k \mapsto w_i \quad \text{met } w_i \in W_i,$$

voor iedere $1 \leq i \leq k$ een lineaire afbeelding. Bemerk dat $v \in V$ op een unieke manier geschreven kan worden als $v = w_1 + \cdots + w_k$ met $w_i \in W_i$, aangezien V een directe som is.

Voorbeeld 3.1.9(8) is een voorbeeld van een projectie-operator:

Definitie 3.1.10. Een lineaire operator $p\colon V \to V$ op een K-vectorruimte V noemt men een *projectie-operator* of kortweg een *projectie* als $p(p(v)) = p(v)$ voor alle $v \in V$.

3.2 Kern en beeld van een lineaire afbeelding

We voeren nu de belangrijke concepten van het beeld en de kern van een lineaire afbeelding in.

Definitie 3.2.1. Zij $f\colon V \to W$ een lineaire afbeelding tussen twee willekeurige K-vectorruimten; dan is de *kern* van f de deelverzameling

$$\ker f := \{v \in V \mid f(v) = 0\}$$

van V. Het *beeld* van f definiëren we als de deelverzameling

$$\operatorname{im} f := f(V) := \{f(v) \mid v \in V\} = \{w \in W \mid \exists v \in V\colon f(v) = w\}$$

van W.

Lemma 3.2.2. *Zij $f\colon V \to W$ een lineaire afbeelding. Dan is $\ker f \leq V$ een deelruimte van V en $\operatorname{im} f \leq W$ een deelruimte van W.*

Bewijs. Dit volgt onmiddellijk uit Lemma 2.2.3. (Werk zelf de details uit.) $\square$

Voorbeeld 3.2.3. (1) Zij $V = K^n$, met standaardbasis $\{e_1, \dots, e_n\}$, en beschouw de shiftoperator $S\colon V \to V$ uit Voorbeeld 3.1.9(6). Dan is $\operatorname{im} S = \langle e_2, \dots, e_n \rangle$, terwijl $\ker S = \langle e_n \rangle$.

(2) Zij $V = \mathbb{Q}^2$ en $W = \mathbb{Q}^3$, en beschouw de lineaire afbeelding f die gegeven wordt door linkse vermenigvuldiging met de matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix}.$$

Dan is $\operatorname{im} f = \langle (1,2,3)^t \rangle \leq W$, terwijl $\ker f = \langle (2,-1)^t \rangle \leq V$.

Lemma 3.2.4. *Zij $f\colon V \to W$ een lineaire afbeelding. Dan geldt:*

(i) *f is injectief als en slechts als $\ker f = \{0\}$.*

(ii) *f is surjectief als en slechts als $\operatorname{im} f = W$.*

Bewijs. (i) Dat de kern van een injectieve afbeelding gelijk is aan de nulruimte volgt onmiddellijk uit de definities van injectiviteit en kern. Onderstel omgekeerd dat $\ker f = \{0\}$. Als $f(v) = f(v')$ voor $v, v' \in V$, dan volgt uit de lineariteit van f dat $f(v - v') = f(v) - f(v') = 0$, dus $v - v' \in \ker f$, maar dan is $v - v' = 0$ en dus $v = v'$.

(ii) Dit volgt onmiddellijk uit de definities van surjectiviteit en beeld. □

Lemma 3.2.5. *Zij $f\colon V \to W$ een lineaire afbeelding en zij $U \leq V$ een deelruimte. Beschouw de restrictie $f|_U$ van f tot U. Dan is*

$$\ker f|_U = \ker f \cap U \quad \text{en} \quad \operatorname{im} f|_U \leq \operatorname{im} f.$$

Bewijs. Oefening. □

3.3 Lineaire afbeeldingen en basissen

Om een lineaire afbeelding volledig te beschrijven, is het voldoende om het beeld van de basiselementen te geven. Dit zal een zeer nuttige manier blijken te zijn om over lineaire afbeeldingen na te denken.

We herinneren nog even aan de notatie uit Opmerking 2.2.7(i), die we in deze context vaak zullen gebruiken.

Lemma 3.3.1. *Zij V en W twee K-vectorruimten, en zij $\mathcal{B}$ een basis voor V.*

(i) *Zij $f\colon V \to W$ een lineaire afbeelding. Dan is f volledig bepaald door de beelden van de basiselementen $f(b)$ voor alle $b \in \mathcal{B}$. In het bijzonder is* $\operatorname{im}(f) = \operatorname{span}\big(\{f(b) \mid b \in \mathcal{B}\}\big)$.

(ii) *Omgekeerd, onderstel dat we voor elke $b \in \mathcal{B}$ een element $w_b \in W$ kiezen. Dan bestaat er een unieke lineaire afbeelding $f\colon V \to W$ zodat $f(b) = w_b$ voor alle $b \in \mathcal{B}$.*

Bewijs. (i) Elk element $v \in V$ is op unieke wijze te schrijven als een lineaire combinatie $\sum_{b \in \mathcal{B}} \lambda_b b$, $\lambda_b \in K$, met bijna alle $\lambda_b = 0$. Aangezien f lineair is, volgt dat

$$f(v) = \sum_{b \in \mathcal{B}} \lambda_b f(b)\,. \tag{3.1}$$

In het bijzonder is $\operatorname{im}(f)$ gelijk aan de verzameling van alle mogelijke lineaire combinaties van de elementen $f(b)$ voor $b \in \mathcal{B}$.

(ii) Veronderstel nu dat we voor elke $b \in \mathcal{B}$ een element $w_b \in W$ gekozen hebben. *Definieer* dan een afbeelding $f\colon V \to W$ als

$$f\Big(\sum_{b \in \mathcal{B}} \lambda_b b\Big) := \sum_{b \in \mathcal{B}} \lambda_b w_b.$$

Het is evident dat f dan een lineaire afbeelding is, en het is uiteraard ook de unieke lineaire afbeelding met $f(b) = w_b$ voor alle $b \in \mathcal{B}$. □

De volgende observatie is een eerste voorbeeld van het nut van basissen om lineaire afbeeldingen te beschrijven.

Stelling 3.3.2. *Zij V, W twee K-vectorruimten, zij $\mathcal{B}$ een basis voor V, en zij $f : V \to W$ een lineaire afbeelding. Dan geldt:*

(i) *f is surjectief als en slechts als $\{f(b) \mid b \in \mathcal{B}\}$ een voortbrengende verzameling is in W;*

(ii) *f is injectief als en slechts als $\{f(b) \mid b \in \mathcal{B}\}$ een lineair onafhankelijke verzameling is in W waarbij de elementen $f(b)$ twee aan twee verschillend zijn van elkaar[2];*

(iii) *f is bijectief als en slechts als $\{f(b) \mid b \in \mathcal{B}\}$ een basis is in W waarbij de elementen $f(b)$ twee aan twee verschillend zijn van elkaar.*

Bewijs. Om de injectiviteit en surjectiviteit van f uit te drukken, maken we gebruik van Lemmas 3.2.4 en 3.3.1(i).

(i) Dit volgt onmiddellijk uit Lemma 3.3.1(i), want f is surjectief als en slechts als $\operatorname{im}(f) = W$.

(ii) Onderstel eerst dat f een injectieve afbeelding is. Per definitie van injectiviteit zijn de elementen $f(b)$ dan twee aan twee verschillend van elkaar. Veronderstel nu dat $\sum_{b\in\mathcal{B}} \lambda_b f(b) = 0$ een lineaire combinatie is van $f(b)$'s (met bijna alle $\lambda_b = 0$) die gelijk is aan 0. Dan is $f\big(\sum \lambda_b b\big) = \sum \lambda_b f(b) = 0$, en dus is $\sum \lambda_b b = 0$ omdat $\ker(f) = \{0\}$. Aangezien $\mathcal{B}$ lineair onafhankelijk is, impliceert dit dat $\lambda_b = 0$ voor alle $b \in \mathcal{B}$. We besluiten dat de $f(b)$'s lineair onafhankelijk zijn.

    Omgekeerd, onderstel dat de $f(b)$'s twee aan twee verschillend en lineair onafhankelijk zijn; we zullen aantonen dat $\ker(f) = \{0\}$. Beschouw dus een element $v \in V$ met $f(v) = 0$, en schrijf $v = \sum \lambda_b b$ (met bijna alle $\lambda_b = 0$). Uit $f\big(\sum \lambda_b b\big) = 0$ volgt $\sum \lambda_b f(b) = 0$. Dit impliceert dat $\lambda_b = 0$ voor alle $b \in \mathcal{B}$, en dus is $v = 0$. Hieruit volgt dat $\ker(f) = \{0\}$; de afbeelding f is dus injectief.

(iii) Dit volgt nu uit (i) en (ii). □

Voorbeeld 3.3.3. Zij $K = \mathbb{R}$, en beschouw de afleiding $f = d/dx$ op de vectorruimte P_n uit Voorbeeld 3.1.9(5) hierboven. Als basis voor P_n kiezen we bijvoorbeeld $\{1, x, x^2, \ldots, x^n\}$, terwijl een basis voor P_{n-1} bijvoorbeeld

[2] We hadden dit ook kunnen formuleren door te zeggen dat $\{f(b) \mid b \in \mathcal{B}\}$ een lineair onafhankelijk *stel* is in W, maar we hebben het gebruik van stellen (i.e. verzamelingen waarin eenzelfde element meerdere malen kan optreden) bewust vermeden in deze cursus.

gegeven wordt door $\{1, x, x^2, \dots, x^{n-1}\}$. We bekijken nu de beelden van de basisvectoren, en we vinden

$$f(1) = 0,\ f(x) = 1,\ f(x^2) = 2x,\ \dots,\ f(x^n) = nx^{n-1}.$$

We stellen vast dat de verzameling $\{f(1), \dots, f(x^n)\}$ lineair afhankelijk is, want het bevat het element 0; hieruit volgt reeds dat f niet injectief is. Anderzijds kunnen we elk element van P_{n-1} schrijven als

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} = a_0 \cdot 1 + \frac{a_1}{2} \cdot 2x + \cdots + \frac{a_{n-1}}{n} \cdot nx^{n-1},$$

zodat de verzameling $\{f(1), \dots, f(x^n)\}$ wel voortbrengend is; hieruit volgt dat f surjectief is.

Definitie 3.3.4. (i) Een bijectieve lineaire afbeelding tussen twee K-vectorruimten noemen we een *isomorfisme*.

(ii) Als er tussen twee K-vectorruimten een bijectieve lineaire afbeelding bestaat dan zeggen we dat de vectorruimten *isomorf* zijn. Als twee K-vectorruimten V en W isomorf zijn, noteren we dit met $V \cong W$.

(iii) Een *automorfisme* van V is een isomorfisme van V naar zichzelf.

Voorbeeld 3.3.5. (1) Zij K een veld, en V een willekeurige K-vectorruimte. De identieke afbeelding $\mathbf{1}_V$ op V is een automorfisme van V.

(2) Zij K een veld, en $V = K^n$. Beschouw de afbeelding

$$f \colon V \to V \colon (a_1, a_2, \dots, a_{n-1}, a_n)^t \mapsto (a_n, a_{n-1}, \dots, a_2, a_1)^t.$$

Dan is f een automorfisme van V.

(3) De shiftoperator uit Voorbeeld 3.1.9(6) is geen automorfisme van V.

(4) Zij K een veld, en stel $V = K^n$ en $W = K^m$ voor zekere n, m. Dan is $V \cong W$ als en slechts als $n = m$. (Zie ook Gevolg 3.3.6.)

(5) Zij K een veld. Dan is $K^n \oplus K^m \cong K^{n+m}$ voor alle n, m.

Gevolg 3.3.6. *Twee K-vectorruimten zijn isomorf als en slechts als ze dezelfde dimensie hebben.*

Bewijs. Zij V en W twee K-vectorruimten, en zij $\mathcal{B}$ een basis van V.

Stel eerst dat $f \colon V \to W$ een isomorfisme is. Uit Stelling 3.3.2(iii) volgt dat $\{f(b) \mid b \in \mathcal{B}\}$ een basis van W is, waarbij de elementen $f(b)$ twee aan twee verschillend zijn van elkaar. Bijgevolg hebben de basissen van V en W even veel elementen.

Omgekeerd, stel dat V en W dezelfde dimensie hebben. Kies dan basissen $\mathcal{B}$ voor V en $\mathcal{C}$ voor W. Omdat V en W dezelfde dimensie hebben, bestaat er een bijectie $\beta\colon \mathcal{B} \to \mathcal{C}$. Definieer

$$f\colon V \to W\colon \sum_{b\in\mathcal{B}} \lambda_b b \mapsto \sum_{b\in\mathcal{B}} \lambda_b \beta(b);$$

dan is f een lineaire afbeelding, en $\{f(b) \mid b \in \mathcal{B}\} = \mathcal{C}$. Uit Stelling 3.3.2(iii) volgt nu dat f een bijectie is, en bijgevolg een isomorfisme is. $\square$

Gevolg 3.3.7. *Zij K een veld, en zij V een eindig-dimensionale K-vectorruimte. Stel $n = \dim_K(V)$. Dan is V isomorf met de K-vectorruimte K^n.*

Bewijs. Dit is een triviaal gevolg van Gevolg 3.3.6. $\square$

Opmerking 3.3.8. Het isomorfisme van V naar K^n in Gevolg 3.3.7 hangt dus af van de keuze van een basis voor V, en verschillende keuzes geven aanleiding tot verschillende isomorfismen. We gaan hier veel dieper op in in Hoofdstuk 4.

De volgende stelling geeft een verband tussen de dimensie van het beeld $\operatorname{im} f$ en de dimensie van de kern $\ker f$ van een willekeurige lineaire afbeelding f (vertrekkend uit een eindig-dimensionale vectorruimte).

Stelling 3.3.9 (Dimensiestelling voor lineaire afbeeldingen)**.** *Zij V een eindig-dimensionale K-vectorruimte, W een willekeurige K-vectorruimte, en $f\colon V \to W$ een lineaire afbeelding. Dan is*

$$\dim V = \dim\ker f + \dim\operatorname{im} f\,.$$

Bewijs. Kies een basis $\{b_1, \dots, b_k\}$ voor de deelruimte $\ker f$, dus $k = \dim\ker f$. Uit Gevolg 2.4.8(i) volgt dat we deze basis kunnen uitbreiden tot een basis $\{b_1, \dots, b_k, b_{k+1}, \dots, b_n\}$ voor V. Stel $U = \operatorname{span}(b_{k+1}, \dots, b_n)$. Uit Stelling 2.5.6 halen we dan dat $V = \ker f \oplus U$.

Uit Lemmas 3.2.4 en 3.2.5 volgt dat de restrictie $f|_U\colon U \to W$ een injectieve afbeelding is. Aangezien U een complement van $\ker f$ is, volgt er dat $\operatorname{im} f|_U = \operatorname{im} f$.

Dit impliceert dat $f|_U$ een isomorfisme is tussen U en $\operatorname{im} f$. Wegens Gevolg 3.3.6 is dan $\dim\operatorname{im} f = \dim U$. Vermits $\dim V = \dim\ker f + \dim U$ volgt het resultaat. $\square$

De volgende stelling is zeer belangrijk voor lineaire operatoren op eindig-dimensionale vectorruimten.

Gevolg 3.3.10 (Alternatief-stelling)**.** *Zij* V, W *twee* K*-vectorruimten en zij* $f\colon V \to W$ *een lineaire afbeelding. Onderstel dat* $\dim V = \dim W < \infty$*. De volgende eigenschappen zijn equivalent*[3]*:*

(a) f *is injectief;*

(b) f *is surjectief;*

(c) f *is bijectief.*

Bewijs. Merk op dat (c) equivalent is met "(a) en (b)", dus het volstaat te bewijzen dat (a) $\iff$ (b). Inderdaad, gebruik makend van de Dimensiestelling 3.3.9 vinden we:

$$\begin{aligned} f \text{ injectief} &\iff \ker f = \{0\} \\ &\iff \dim \ker f = 0 \\ &\iff \dim V = \dim \operatorname{im} f \\ &\iff \dim W = \dim \operatorname{im} f \\ &\iff \operatorname{im} f = W \\ &\iff f \text{ surjectief.} \end{aligned}$$

□

Opmerking 3.3.11. We benadrukken dat de alternatief-stelling gaat over lineaire afbeeldingen tussen vectorruimten van *gelijke* dimensie. In het bijzonder is ze van toepassing voor lineaire operatoren $f\colon V \to V$ met $\dim V$ eindig.

Voorbeeld 3.3.12. (1) Beschouw $f\colon V \to W$ zoals in Voorbeeld 3.2.3(2). Dan is $\dim \ker f = 1$ en $\dim \operatorname{im} f = 1$, zodat $\dim V = 2$ inderdaad gelijk is aan $\dim \ker f + \dim \operatorname{im} f$.

(2) Zij $K = \mathbb{R}$, en beschouw de afleiding $f = d/dx$ van P_n naar P_{n-1} zoals in Voorbeeld 3.1.9(5). Dan is $\dim \ker f = 1$ en $\dim \operatorname{im} f = n$ (want f is surjectief), zodat inderdaad $\dim P_n = n + 1 = \dim \ker f + \dim \operatorname{im} f$.

(3) Beschouw een projectie-operator $p\colon V \to V$ (zie Definitie 3.1.10). We beweren dat in dat geval $V = \ker p \oplus \operatorname{im} p$.
Inderdaad, enerzijds is $\ker p \cap \operatorname{im} p = 0$, want als $v \in \ker p \cap \operatorname{im} p$, dan kunnen we $v = p(u)$ schrijven voor een $u \in V$, maar omdat $v \in \ker p$ is dan $p(v) = 0$. Hieruit volgt dan $v = p(u) = p(p(u)) = p(v) = 0$.

[3] Het *equivalent* zijn van eigenschappen betekent dat *indien* één van de eigenschappen geldt, *dan* ook de andere gelden. In dit geval zegt de stelling dus dat, *indien* f injectief is, ze dan automatisch ook surjectief (en dus bijectief) is, en omgekeerd, *indien* f surjectief is, ze dan ook injectief (en dus bijectief) is. De stelling zegt dus zeker niet dat elke lineaire operator injectief, surjectief en/of bijectief zou zijn!

Anderzijds is $V = \ker p + \operatorname{im} p$, want elke $v \in V$ kunnen we schrijven als $v = u + p(v)$ met $u := v - p(v) \in \ker p$ en $p(v) \in \operatorname{im} p$.
Merk op dat de gelijkheid $V = \ker f \oplus \operatorname{im} f$ *niet geldig is* voor willekeurige operatoren $f\colon V \to V$! (Beschouw bijvoorbeeld de shift-operator uit Voorbeeld 3.1.9(6).)

3.4 Optelling en scalaire vermenigvuldiging van lineaire afbeeldingen

In deze sectie bekijken we de verzameling $\operatorname{Hom}_K(V, W)$ van lineaire afbeeldingen tussen twee K-vectorruimten V en W. We kunnen deze uitrusten met een optelling en scalaire vermenigvuldiging op een natuurlijke manier, en het blijkt dat we op deze manier een vectorruimte verkrijgen. Wat we hier doen, is analoog aan wat we kort hebben besproken in Voorbeeld 2.1.2(6).

Definitie 3.4.1. Zij K een veld, en V, W twee K-vectorruimten.

(i) De verzameling van alle lineaire afbeeldingen tussen V en W noteren we

$$\operatorname{Hom}_K(V, W) := \{f\colon V \to W \mid f \text{ een lineaire afbeelding}\},$$

en noemen we de *ruimte van homomorfismen van V naar W*, of ook wel de *ruimte van lineaire afbeeldingen van V naar W*, of kortweg de *homruimte van V naar W*. We gebruiken ook de notatie $\operatorname{Hom}(V, W)$ als het veld vastligt of als het duidelijk is over welk veld de vectorruimten beschouwd worden.

(ii) Zij $f, g \in \operatorname{Hom}_K(V, W)$ en $\lambda \in K$; dan definiëren we de afbeeldingen $f + g$ en λf als

$$\begin{cases} (f+g)(v) := f(v) + g(v) & \text{voor alle } v \in V, \text{ en} \\ (\lambda f)(v) := \lambda f(v) & \text{voor alle } v \in V. \end{cases}$$

De optelling en scalaire vermenigvuldiging in het rechterlid van de definitie zijn de optelling en de scalaire vermenigvuldiging in W.

Opmerking 3.4.2. Beschouw twee elementen $f, g \in \operatorname{Hom}_K(V, W)$. Dan geldt:

$$f = g \iff f(v) = g(v) \text{ voor alle } v \in V.$$

(Zie Opmerking 3.1.2.) Bovendien geldt, als $\mathcal{B}$ een willekeurige basis is voor V:

$$f = g \iff f(v) = g(v) \text{ voor alle } v \in \mathcal{B}.$$

(Ga dit zelf na.) We zullen dit in het vervolg vaak gebruiken.

Lemma 3.4.3. *Zij K een veld, en V, W twee K-vectorruimten. Zij $f, g \in \mathrm{Hom}_K(V, W)$ en $\lambda \in K$. Dan is $f + g \in \mathrm{Hom}_K(V, W)$ en $\lambda f \in \mathrm{Hom}_K(V, W)$. Met deze bewerkingen wordt de verzameling $\mathrm{Hom}_K(V, W)$ een K-vectorruimte. Het nulelement van deze vectorruimte is de nulafbeelding $0_{V,W}$.*

Bewijs. Aangezien W een vectorruimte is, beelden de afbeeldingen $f + g$ en λf elementen uit V af op elementen uit W. We gaan na dat deze afbeeldingen lineair zijn. Zij dus $v_1, v_2 \in V$ en $\mu_1, \mu_2 \in K$; dan is

$$\begin{aligned}(f+g)(\mu_1 v_1 + \mu_2 v_2) &= f(\mu_1 v_1 + \mu_2 v_2) + g(\mu_1 v_1 + \mu_2 v_2) \\ &= \mu_1 f(v_1) + \mu_2 f(v_2) + \mu_1 g(v_1) + \mu_2 g(v_2) \\ &= \mu_1(f(v_1) + g(v_1)) + \mu_2(f(v_2) + g(v_2)) \\ &= \mu_1 (f+g)(v_1) + \mu_2 (f+g)(v_2),\end{aligned}$$

en dus is $f + g \in \mathrm{Hom}_K(V, W)$;

$$\begin{aligned}(\lambda f)(\mu_1 v_1 + \mu_2 v_2) = \lambda f(\mu_1 v_1 + \mu_2 v_2) &= \lambda(\mu_1 f(v_1) + \mu_2 f(v_2)) \\ &= \mu_1 \lambda f(v_1) + \mu_2 \lambda f(v_2) = \mu_1 (\lambda f)(v_1) + \mu_2 (\lambda f)(v_2),\end{aligned}$$

en dus is $\lambda f \in \mathrm{Hom}_K(V, W)$.

Om aan te tonen dat $\mathrm{Hom}_K(V, W)$ een K-vectorruimte is, moeten we (V1)–(V7) nagaan. We werken uit dat de eigenschappen (V1), (V2) en (V3) gelden; de andere eigenschappen laten we als een oefening (analoog aan de verificatie van (V1)). We maken hierbij gebruik van Opmerking 3.4.2.

(V1) Beschouw $f, g, h \in \mathrm{Hom}(V, W)$, we moeten nagaan dat $(f + g) + h = f + (g + h)$. Kies $v \in V$ willekeurig, we gaan na dat $((f + g) + h)(v) = (f + (g + h))(v)$. Per definitie is

$$((f+g)+h)(v) = (f+g)(v) + h(v) = (f(v) + g(v)) + h(v),$$

en aangezien de optelling in W associatief is, is dit gelijk aan

$$f(v) + (g(v) + h(v)) = f(v) + (g+h)(v) = (f + (g+h))(v).$$

(V2) Voor alle $v \in V$ is $(f + 0)(v) = f(v) + 0(v) = f(v) + 0_W = f(v)$, dus is $f + 0 = f$ voor alle $f \in \mathrm{Hom}(V, W)$. Analoog is ook $0 + f = f$.

(V3) Zij $f \in \mathrm{Hom}(V, W)$; definieer de afbeelding $g \colon V \to W \colon v \mapsto -f(v)$. Het is duidelijk dat $g \in \mathrm{Hom}(V, W)$. Er geldt dat

$$(f+g)(v) = f(v) + g(v) = f(v) + (-f(v)) = 0_W = 0(v),$$

voor alle $v \in V$. Bijgevolg is $f + g = 0$; we noteren de afbeelding g ook vaak als $-f$. $\square$

Kort samengevat kunnen we zeggen dat $\mathrm{Hom}_K(V,W)$ een K-vectorruimte is, omdat W een K-vectorruimte is. In de volgende stelling gaan we na wat de dimensie van de vectorruimte $\mathrm{Hom}_K(V,W)$ is. We voeren eerst volgende notatie in.

Notatie 3.4.4 (Kronecker[4] delta)**.** Voor alle $i,j \in \mathbb{N}$ definiëren we

$$\delta_{ij} := \begin{cases} 1 & \text{als } i = j\,; \\ 0 & \text{als } i \neq j\,. \end{cases}$$

Stelling 3.4.5. *Zij K een veld, en V, W twee eindig-dimensionale K-vectorruimten. Dan is* $\dim \mathrm{Hom}(V,W) = \dim V \cdot \dim W$*.*

Bewijs. Zij $\mathcal{B} = \{v_1, \dots, v_n\}$ een basis voor V en $\mathcal{C} = \{w_1, \dots, w_m\}$ een basis voor W. Voor elke $1 \leq i \leq n$ en $1 \leq j \leq m$ stellen we $f_{ij} \in \mathrm{Hom}(V,W)$ gelijk aan de unieke lineaire afbeelding – zie Lemma 3.3.1(ii) – bepaald door

$$\begin{cases} f_{ij}(v_i) = w_j, \\ f_{ij}(v_k) = 0 \quad \text{voor alle } k \neq i. \end{cases} \tag{3.2}$$

Met andere woorden, er geldt dat $f_{ij}(v_k) = \delta_{ik} w_j$ voor alle i, j, k. We zullen aantonen dat de verzameling

$$S = \{f_{ij} \mid i = 1, \dots, n; j = 1, \dots, m\} \subseteq \mathrm{Hom}(V,W)$$

een basis van $\mathrm{Hom}(V,W)$ is.

- We tonen aan dat S lineair onafhankelijk is. Stel dat $\sum_{i,j} \lambda_{ij} f_{ij} = 0$ voor $\lambda_{ij} \in K$. Dan volgt, voor alle $k = 1, \dots, n$,

  $$\sum_{i,j} \lambda_{ij} f_{ij}(v_k) = \sum_{i,j} \lambda_{ij} \delta_{ik} w_j = \sum_{j} \lambda_{kj} w_j = 0\,;$$

  omdat $\mathcal{C}$ een basis is van W, volgt hieruit dat $\lambda_{kj} = 0$ voor alle k, j.
- We tonen aan dat S voortbrengend is. Zij $f \in \mathrm{Hom}(V,W)$ een willekeurige lineaire afbeelding van V naar W. Voor elke $k = 1, \dots, n$ schrijven we $f(v_k) = \sum_{j=1}^{m} \mu_{kj} w_j$, voor zekere $\mu_{kj} \in K$. Er volgt dat

  $$f(v_k) = \sum_{j=1}^{m} \mu_{kj} w_j = \sum_{i,j} \mu_{ij} \delta_{ik} w_j = \sum_{i,j} \mu_{ij} f_{ij}(v_k)$$

  voor elke k. Uit Opmerking 3.4.2 volgt nu dat $f = \sum_{i,j} \mu_{ij} f_{ij}$, dus is S voortbrengend.

[4] Genoemd naar de Duitse wiskundige Leopold Kronecker (1823–1891).

De verzameling S is dus een basis van $\mathrm{Hom}(V, W)$, en de stelling volgt aangezien $|S| = nm$. $\square$

Opmerking 3.4.6. Stelling 3.4.5 is niet langer waar indien V oneindig-dimensionaal is, zelfs in het eenvoudigste geval waarin $\dim W = 1$. (De stelling blijft wel waar als de dimensie van W oneindig is en V eindig-dimensionaal is.) We verwijzen hiervoor naar Opmerking 4.3.7 verderop.

3.5 De samenstelling van lineaire afbeeldingen

We zagen reeds dat $\mathrm{Hom}_K(V, W)$, voor twee K-vectorruimten V en W, uitgerust met de optelling en scalaire vermenigvuldiging, de structuur heeft van een K-vectorruimte (van dimensie $\dim(V) \cdot \dim(W)$. Als $V = W$ dan kunnen we nog verder gaan en $\mathrm{Hom}_K(V, V)$ uitrusten met een vermenigvuldiging, gevormd door de *samenstelling* van lineaire afbeeldingen. Samen met de optelling die we al hadden, krijgen we zo de structuur van een ring.

Lemma 3.5.1. *Zij V, W, U drie K-vectorruimten, zij $g \in \mathrm{Hom}(V, W)$ en $f \in \mathrm{Hom}(W, U)$. Dan is $f \circ g \in \mathrm{Hom}(V, U)$.*

Bewijs. Oefening. $\square$

Zij V een K-vectorruimte; dan is de verzameling $\mathrm{Hom}_K(V, V)$ van lineaire operatoren op V een K-vectorruimte. Deze ruimte heeft echter nog meer structuur.

Stelling 3.5.2. *Zij K een veld, en zij V een K-vectorruimte. Dan is de verzameling $\mathrm{Hom}_K(V, V)$ een ring, waarbij de vermenigvuldiging gegeven is door samenstelling van operatoren:*

$$(fg)(v) := (f \circ g)(v) = f(g(v)) \text{ voor alle } v \in V,$$

voor alle $f, g \in \mathrm{Hom}_K(V, V)$.

Bewijs. We weten reeds dat $\mathrm{Hom}_K(V, V)$ een abelse groep is voor de optelling, dus (R1)–(R4) zijn reeds voldaan.

(R5) We gaan na dat $f \circ (g \circ h) = (f \circ g) \circ h$ voor alle $f, g, h \in \mathrm{Hom}_K(V, V)$. Zij $v \in V$ willekeurig, dan is $(f \circ (g \circ h))(v) = f((g \circ h)(v)) = f(g(h(v)))$ en $((f \circ g) \circ h)(v) = (f \circ g)(h(v)) = f(g(h(v)))$.

(R6) Het neutraal element voor de samenstelling is de identieke afbeelding $\mathbf{1}_V$ op V. Immers, $(f \circ \mathbf{1}_V)(v) = f(\mathbf{1}_V(v)) = f(v)$ en $(\mathbf{1}_V \circ f)(v) = \mathbf{1}_V(f(v)) = f(v)$ voor alle $v \in V$.

(R7) We moeten nagaan dat voor alle $f, g, h \in \mathrm{Hom}_K(V, V)$ geldt dat

$$f(g+h) = fg + fh \quad \text{en} \quad (g+h)f = gf + hf\,.$$

We rekenen de linksdistributiviteit na. Voor alle $v \in V$ hebben we

$$\begin{aligned}\big(f(g+h)\big)(v) &= f \circ (g+h)(v) = f((g+h)(v)) \\ &= f(g(v) + h(v)) = f(g(v)) + f(h(v)) = (fg)(v) + (fh)(v)\,;\end{aligned}$$

de voorlaatste gelijkheid volgt uit de lineariteit van f. De verificatie van de rechtsdistributiviteit is volledig analoog. □

Definitie 3.5.3. De ruimte $\mathrm{Hom}_K(V, V)$ met als vermenigvuldiging de samenstelling vormt dus een ring; men noemt dit de *endomorfismenring* van V en noteert deze met $\mathrm{End}_K(V)$ of $\mathrm{End}(V)$.

Opmerking 3.5.4. (i) De ring $\mathrm{End}_K(V)$ is niet commutatief van zodra $\dim V > 1$. Als voorbeeld beschouwen we de vectorruimte K^2 en de lineaire operatoren

$$\begin{aligned}&f\colon V \to V\colon (a,b)^t \mapsto (b,0)^t \text{ en}\\ &g\colon V \to V\colon (a,b)^t \mapsto (0,a)^t.\end{aligned}$$

We hebben $(f \circ g)(a,b)^t = (a,0)^t$ en $(g \circ f)(a,b)^t = (0,b)^t$. Bijgevolg is $f \circ g \neq g \circ f$.

(ii) De ring $\mathrm{End}_K(V)$ heeft *nuldelers* als $\dim V > 1$, d.w.z. er bestaan elementen $f, g \in \mathrm{End}_K(V)$ met $f \neq 0$ en $g \neq 0$ en toch $fg = 0$. Als voorbeeld beschouwen we de vectorruimte K^2 en de lineaire operatoren

$$\begin{aligned}&f\colon V \to V\colon (a,b)^t \mapsto (a,0)^t \text{ en}\\ &g\colon V \to V\colon (a,b)^t \mapsto (0,b)^t.\end{aligned}$$

Dan is $(f \circ g)(a,b)^t = (0,0)^t$. Bijgevolg is $f \circ g = 0$.

3.6 Inverteerbare lineaire operatoren

We willen nu nog kort ingaan op inverteerbare (= bijectieve) lineaire operatoren van een vectorruimte V, de zogenaamde automorfismen van V (zie Definitie 3.3.4). We herinneren nog even aan Definitie 3.1.4 en Lemma 3.1.6.

Definitie 3.6.1. Zij V een K-vectorruimte. De verzameling van alle automorfismen van V noemt men de *algemene lineaire groep*, en noteert men als $\mathsf{GL}_K(V)$ of $\mathsf{GL}(V)$.

Lemma 3.6.2. (i) *Zij $f \in \operatorname{End}(V)$, en veronderstel dat f bijectief is. Dan is $f^{-1} \in \operatorname{End}(V)$ en f^{-1} is eveneens bijectief. Met andere woorden, als $f \in \mathsf{GL}(V)$, dan is $f^{-1} \in \mathsf{GL}(V)$.*

(ii) *De deelverzameling $\mathsf{GL}(V)$ van $\operatorname{End}(V)$ vormt een groep voor de samenstelling op $\operatorname{End}(V)$.*

Bewijs. (i) Het feit dat f^{-1} opnieuw een bijectieve afbeelding is, volgt direct uit de definitie. We tonen aan dat f^{-1} lineair is. Stel dus $v, w \in V$ en $\lambda, \mu \in K$ willekeurig; we willen aantonen dat

$$f^{-1}(\lambda v + \mu w) = \lambda f^{-1}(v) + \mu f^{-1}(w)\,.$$

Stel $f^{-1}(v) = x$ en $f^{-1}(w) = y$, m.a.w. $f(x) = v$ en $f(y) = w$. Dan is

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

omdat f lineair is; bijgevolg is

$$\begin{aligned} \lambda f^{-1}(v) + \mu f^{-1}(w) = \lambda x + \mu y = f^{-1}\big(\lambda f(x) + \mu f(y)\big) \\ = f^{-1}(\lambda v + \mu w)\,. \end{aligned}$$

(ii) Merk vooreerst op dat de samenstelling van twee automorfismen opnieuw een automorfisme is. Uit Stelling 3.5.2 volgt dat de samenstelling associatief is, dus (G1) is voldaan. Aangezien $\mathbf{1}_V$ een automorfisme is, is ook (G2) voldaan. In (i) hebben we aangetoond dat het inverse van een automorfisme opnieuw een automorfisme is, en aangezien $f \circ f^{-1} = f^{-1} \circ f = \mathbf{1}_V$ is (G3) ook voldaan. $\square$

Opmerking 3.6.3. Beschouw een eindig-dimensionale vectorruimte V met $\dim V = n$. In Stelling 4.1.8(iv) zullen we aantonen dat $\mathsf{GL}_K(V)$ en $\mathsf{GL}_n(K)$ (de groep van de inverteerbare $n \times n$-matrices over K) isomorfe groepen zijn. Dit verklaart waarom de notatie voor deze objecten gelijkaardig is.

Hoofdstuk

4 Coördinaten

In Gevolg 3.3.7 hebben we gezien dat elke n-dimensionale K-vectorruimte V isomorf is met de kolommenruimte K^n. Zo een isomorfisme hangt echter af van de keuze van een basis. In dit hoofdstuk zullen we zien hoe we aan elke vector een coördinaatvector kunnen hechten en hoe we op deze manier lineaire afbeeldingen kunnen voorstellen met behulp van matrixvermenigvuldiging. We zullen ook de invloed van verandering van basis op deze voorstellingen onderzoeken. Tot slot gaan we wat dieper in op de duale ruimte $\mathrm{Hom}(V, K)$ van een vectorruimte V.

4.1 Coördinaten en matrixvoorstellingen van lineaire afbeeldingen

We starten met enkele definities.

Definitie 4.1.1. Zij K een veld, zij V een n-dimensionale K-vectorruimte, en zij $\mathcal{B} = \{b_1, \dots, b_n\}$ een basis in V.[1]

(i) Als $v = \sum_{i=1}^n \lambda_i b_i$, dan noemen we $(\lambda_1, \dots, \lambda_n)^t$ de *coördinatenvector* van v ten opzichte van de basis $\mathcal{B}$. De elementen λ_i zijn de *coördinaten* van v ten opzichte van de basis $\mathcal{B}$.

(ii) Het isomorfisme

$$\mathrm{co}_{\mathcal{B}} \colon V \to K^n \colon \sum_{i=1}^n \lambda_i b_i \mapsto (\lambda_1, \dots, \lambda_n)^t,$$

noemen we het *coördinatenisomorfisme* ten opzichte van de basis $\mathcal{B}$.

Merk op dat het coördinatenisomorfisme ten opzichte van de basis $\{b_1, \dots, b_n\}$ iedere b_i afbeeldt op e_i, waarbij $\{e_1, \dots, e_n\}$ de standaardbasis is van K^n.

[1] In de context van coördinatenvectoren en matrixvoorstellingen nemen we steeds aan dat een basis $\mathcal{B} = \{b_1, \dots, b_n\}$ *geordend* is. Dat betekent dat we aan $\mathcal{B}$ een opsomming van haar elementen meegeven. Dan is bijvoorbeeld duidelijk wat de eerste, de tweede of de laatste basisvector is, etc.

We kunnen aan iedere lineaire afbeelding tussen twee eindig dimensionale vectorruimten een matrix associëren. Deze matrix hangt echter af van de keuze van de basissen.

Definitie 4.1.2. Zij V een n-dimensionale K-vectorruimte en W een m-dimensionale K-vectorruimte, en zij $f\colon V \to W$ een lineaire afbeelding. Beschouw een basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ van V en een basis $\mathcal{C} = \{c_1, \ldots, c_m\}$ van W. Voor iedere $j = 1, \ldots, n$ is

$$f(b_j) = \sum_{i=1}^{m} a_{ij} c_i \tag{4.1}$$

voor bepaalde scalairen $a_{ij} \in K$. De matrix $(a_{ij}) \in M_{m,n}(K)$ noemen we de *matrixvoorstelling* van f ten opzichte van de basissen $\mathcal{B}$ en $\mathcal{C}$.

We noteren de matrixvoorstelling van f ten opzichte van de basissen $\mathcal{B}$ en $\mathcal{C}$ als $A_{f,\mathcal{B},\mathcal{C}}$ of als A_f als het duidelijk is uit de context welke basissen we beschouwen.

Opmerking 4.1.3. Zij A_f de matrixvoorstelling van $f\colon V \to W$ ten opzichte van de basissen $\mathcal{B} = \{b_1, \ldots, b_n\}$ en $\mathcal{C} = \{c_1, \ldots, c_m\}$. Uit Definitie 4.1.2 volgt dat de j-de kolom van A_f de coördinatenvector van $f(b_j)$ ten opzichte van de basis $\mathcal{C}$ is.

In de praktijk stellen we de matrixvoorstelling A_f van een lineaire afbeelding f op door de coördinatenvectoren van $f(b_1), \ldots, f(b_n)$ t.o.v. $\mathcal{C}$ in de kolommen van A_f te schrijven.

Stelling 4.1.4. *Zij V een n-dimensionale K-vectorruimte en W een m-dimensionale K-vectorruimte. Beschouw een basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ van V en een basis $\mathcal{C} = \{c_1, \ldots, c_m\}$ van W.*

(i) *Voor iedere matrix $A \in M_{m,n}(K)$ bestaat er een unieke lineaire afbeelding $f\colon V \to W$ waarvoor $A_{f,\mathcal{B},\mathcal{C}} = A$.*

(ii) *Zij $f\colon V \to W$ een lineaire afbeelding en stel $A_f := A_{f,\mathcal{B},\mathcal{C}}$. Dan is*

$$\mathrm{co}_{\mathcal{C}} \circ f = L_{A_f} \circ \mathrm{co}_{\mathcal{B}},$$

met andere woorden, voor elke $v \in V$ hebben we

$$\mathrm{co}_{\mathcal{C}}(f(v)) = A_f \mathrm{co}_{\mathcal{B}}(v).$$

Bewijs. (i) Wegens Lemma 3.3.1(ii) bestaat er een unieke lineaire afbeelding $f\colon V \to W$ zodat de coördinatenvector van $f(b_i)$ ten opzichte van de basis $\mathcal{C}$ juist de i-de kolom van de matrix A is. Het gestelde volgt dan onmiddellijk.

(ii) Noteer $(a_{ij}) = A_f = A_{f,\mathcal{B},\mathcal{C}}$. Aangezien zowel $\mathrm{co}_\mathcal{C} \circ f$ als $L_{A_f} \circ \mathrm{co}_\mathcal{B}$ lineaire afbeeldingen zijn, volstaat het te bewijzen dat ze samenvallen op de basis $\mathcal{B}$, met andere woorden, dat $\mathrm{co}_\mathcal{C}(f(b_j)) = A_f \mathrm{co}_\mathcal{B}(b_j)$ voor elke j. Het linkerlid is gelijk aan $\mathrm{co}_\mathcal{C}(\sum_{i=1}^m a_{ij}c_i) = (a_{1j}, a_{2j}, \ldots, a_{nj})^t$ omwille van (4.1), en het rechterlid is gelijk aan $A_f e_j$, wat precies de j-de kolom is van de matrix A_f. Hieruit volgt het gestelde. □

Opmerking 4.1.5. (i) Stelling 4.1.4(i) toont dus aan dat een lineaire afbeelding volledig bepaald is door zijn matrixvoorstelling ten opzichte van de twee vast gekozen basissen.

(ii) We kunnen de gelijkheid $\mathrm{co}_\mathcal{C} \circ f = L_{A_f} \circ \mathrm{co}_\mathcal{B}$ uit Stelling 4.1.4(ii) schematisch weergeven door middel van een zogenaamd *commutatief diagram*:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\mathrm{co}_\mathcal{B}"] & W \arrow[d, "\mathrm{co}_\mathcal{C}"] \\ K^n \arrow[r, "L_{A_f}"'] & K^m \end{tikzcd}$$

We kunnen nu aantonen dat iedere lineaire afbeelding van K^n naar K^m gegeven is door linkse vermenigvuldiging met een matrix.

Gevolg 4.1.6. *Zij $f\colon K^n \to K^m$ een lineaire afbeelding. Dan bestaat er een matrix $A \in M_{m,n}(K)$ zodat $f = L_A$.*

Bewijs. We beschouwen de standaardbasis $\mathcal{B} = \{e_1, \ldots, e_n\}$ van $V := K^n$ en de standaardbasis $\mathcal{C} = \{e_1, \ldots, e_m\}$ van $W := K^m$. Als $v = (\lambda_1, \ldots, \lambda_n)^t \in K^n$, dan is de coördinatenvector van v t.o.v. $\mathcal{B}$ ook gegeven door $(\lambda_1, \ldots, \lambda_n)^t$, met andere woorden, $\mathrm{co}_\mathcal{B} = \mathbf{1}_V$; analoog is $\mathrm{co}_\mathcal{C} = \mathbf{1}_W$.

Definieer de matrix $A := A_{f,\mathcal{B},\mathcal{C}} \in M_{m,n}(K)$. Uit Stelling 4.1.4(ii) volgt nu dat $f = L_A$. □

Nu we weten dat er een een-op-een verband bestaat tussen lineaire afbeeldingen $f\colon V \to W$ en matrices $A \in M_{m,n}(K)$ (mits een vaste keuze van twee basissen) is het een natuurlijke vraag of dit aanleiding geeft tot een sterker "globaal verband" tussen *alle* lineaire afbeeldingen en *alle* $m \times n$-matrices. Dat is inderdaad het geval, zoals we in Stelling 4.1.8 zullen onderzoeken.

We tonen eerst volgend lemma aan:

Lemma 4.1.7. *Zij $A, B \in M_{m,n}(K)$. Dan is $A = B$ als en slechts als $L_A = L_B$ (als lineaire afbeeldingen van K^n naar K^m).*

Bewijs. Uiteraard volgt uit $A = B$ dat $L_A = L_B$. Onderstel dus omgekeerd dat $L_A = L_B$. Beschouw de standaardbasis $\mathcal{B} = \{e_1, \dots, e_n\}$ van K^n. Dan is $L_A(e_i)$ precies de i-de kolom van de matrix A, en analoog is $L_B(e_i)$ de i-de kolom van de matrix B. Uit $L_A = L_B$ volgt dus dat de i-de kolommen van A en B gelijk zijn, voor alle i, en dus is $A = B$. □

Stelling 4.1.8. *Zij V een n-dimensionale K-vectorruimte, W een m-dimensionale K-vectorruimte, en U een t-dimensionale K-vectorruimte. Kies een basis $\mathcal{B}$ voor V, $\mathcal{C}$ voor W en $\mathcal{D}$ voor U.*

(i) *De afbeelding*

$$\mathrm{Hom}_K(V, W) \to M_{m,n}(K) \colon f \mapsto A_f$$

(t.o.v. de basissen $\mathcal{B}$ en $\mathcal{C}$) is een isomorfisme van K-vectorruimten.

(ii) *Beschouw de afbeeldingen*

$$\begin{aligned} \mathrm{Hom}_K(V, W) &\to M_{m,n}(K) \colon\ f \mapsto A_f\,, \\ \mathrm{Hom}_K(W, U) &\to M_{t,m}(K) \colon\ g \mapsto A_g\,, \\ \mathrm{Hom}_K(V, U) &\to M_{t,n}(K) \colon\ h \mapsto A_h\,, \end{aligned}$$

die we in (i) *hebben ingevoerd. Dan geldt voor alle $f \in \mathrm{Hom}_K(V, W)$ en alle $g \in \mathrm{Hom}_K(W, U)$ dat $A_{g \circ f} = A_g A_f$.*

(iii) *De afbeelding*

$$\mathrm{End}_K(V) \to M_n(K) \colon f \mapsto A_f$$

(t.o.v. de basis $\mathcal{B}$) is een isomorfisme van K-vectorruimten, en voor alle $f, g \in \mathrm{Hom}_K(V)$ geldt dat $A_{g \circ f} = A_g A_f$.

(iv) *Zij $f \in \mathrm{End}_K(V)$, en zij A_f de corresponderende matrix t.o.v. de basis $\mathcal{B}$. Dan is f inverteerbaar als en slechts als A_f inverteerbaar is, en in dat geval is $(A_f)^{-1} = A_{f^{-1}}$, of explicieter neergeschreven,*

$$(A_{f,\mathcal{B},\mathcal{C}})^{-1} = A_{f^{-1},\mathcal{C},\mathcal{B}}.$$

We hebben dus een bijectieve afbeelding

$$\mathsf{GL}_K(V) \to \mathsf{GL}_n(K) \colon f \mapsto A_f\,,$$

en voor alle $f, g \in \mathsf{GL}_K(V)$ is $A_{g \circ f} = A_g A_f$.

Bewijs. (i) We verifiëren eerst dat de afbeelding $f \mapsto A_f$ lineair is, met andere woorden, dat

$$A_{\lambda f + \mu g} = \lambda A_f + \mu A_g$$

voor alle $f, g \in \mathrm{Hom}_K(V, W)$ en $\lambda, \mu \in K$.
Omdat $\mathrm{co}_\mathcal{C}$ een lineaire afbeelding is, hebben we

$$\mathrm{co}_\mathcal{C} \circ (\lambda f + \mu g) = \lambda \mathrm{co}_\mathcal{C} \circ f + \mu \mathrm{co}_\mathcal{C} \circ g.$$

Hierop passen we nu drie keer Stelling 4.1.4(ii) toe, en we krijgen

$$L_{A_{\lambda f + \mu g}} \circ \mathrm{co}_\mathcal{B} = \lambda L_{A_f} \circ \mathrm{co}_\mathcal{B} + \mu L_{A_g} \circ \mathrm{co}_\mathcal{B} = (\lambda L_{A_f} + \mu L_{A_g}) \circ \mathrm{co}_\mathcal{B}.$$

Aangezien $\mathrm{co}_\mathcal{B}$ een isomorfisme is en de matrixvermenigvuldiging distributief is, halen we hieruit dat

$$L_{A_{\lambda f + \mu g}} = \lambda L_{A_f} + \mu L_{A_g} = L_{\lambda A_f + \mu A_g}.$$

Uit Lemma 4.1.7 besluiten we nu dat $\lambda A_f + \mu A_g = A_{\lambda f + \mu g}$.

Uit Stelling 4.1.4(i) volgt dat $f \mapsto A_f$ een bijectie is. (Uit de existentie volgt surjectiviteit; uit de uniciteit volgt injectiviteit.) We besluiten dat deze afbeelding een isomorfisme van K-vectorruimten is.

(ii) Door drie keer Stelling 4.1.4(ii) toe te passen, vinden we

$$\begin{aligned} L_{A_{g \circ f}} \circ \mathrm{co}_\mathcal{B} &= \mathrm{co}_\mathcal{D} \circ (g \circ f) = (\mathrm{co}_\mathcal{D} \circ g) \circ f \\ &= (L_{A_g} \circ \mathrm{co}_\mathcal{C}) \circ f = L_{A_g} \circ (\mathrm{co}_\mathcal{C} \circ f) \\ &= L_{A_g} \circ (L_{A_f} \circ \mathrm{co}_\mathcal{B}) = (L_{A_g} \circ L_{A_f}) \circ \mathrm{co}_\mathcal{B}, \end{aligned}$$

en omdat $\mathrm{co}_\mathcal{B}$ een isomorfisme is, volgt hieruit dat $L_{A_{g \circ f}} = L_{A_g} \circ L_{A_f}$. Uit de associativiteit van de matrixvermenigvuldiging weten we dat het rechterlid ook nog gelijk is aan $L_{A_g A_f}$, en opnieuw volgt nu uit Lemma 4.1.7 dat $A_{g \circ f} = A_g A_f$.

(iii) Het resultaat volgt uit (i) met $V = W$ en (ii) met $V = W = U$.

(iv) Veronderstel eerst dat f inverteerbaar is, en beschouw de inverse afbeelding $f^{-1} \in \mathsf{GL}_K(V)$; dan is $f \circ f^{-1} = f^{-1} \circ f = \mathbf{1}_V$. Uit (ii) voor $V = W = U$ volgt dat

$$A_{f \circ f^{-1}} = A_f A_{f^{-1}} = A_{\mathbf{1}_V}, \quad A_{f^{-1} \circ f} = A_{f^{-1}} A_f = A_{\mathbf{1}_V}.$$

Aangezien $A_{\mathbf{1}_V} = I_n$, is $A_f A_{f^{-1}} = A_{f^{-1}} A_f = I_n$. Bijgevolg is de matrix A_f inverteerbaar; uit de uniciteit van het inverse van een matrix volgt dat $(A_f)^{-1} = A_{f^{-1}}$.

Veronderstel omgekeerd dat de matrix A_f inverteerbaar is, en stel $B := (A_f)^{-1}$. Uit Stelling 4.1.4(i) volgt dat $B = A_g$ voor een bepaalde $g \in \mathrm{End}_K(V)$. Uit $A_f B = B A_f = I_n$ volgt nu dat $A_{f \circ g} = A_{g \circ f} = A_{\mathbf{1}_V}$. Bijgevolg is $f \circ g = g \circ f = \mathbf{1}_V$, en uit Lemma 3.1.6 volgt dat f inverteerbaar is. $\square$

Opmerking 4.1.9. (i) De verzamelingen $\mathsf{GL}_K(V)$ en $\mathsf{GL}_n(K)$ zijn groepen. Stelling 4.1.8(iv) drukt uit dat de afbeelding $f \mapsto A_f$ een isomorfisme van groepen is.

(ii) De verzamelingen $\mathrm{End}_K(V)$ en $M_n(K)$ zijn beide voorbeelden van wat men K-algebra's noemt. Stelling 4.1.8(iii) drukt uit dat de afbeelding $f \mapsto A_f$ een isomorfisme van K-algebra's is.

4.2 Coördinatentransformaties

De beschrijving van elementen van een vectorruimte door hun coördinaten hangt af van de keuze van basissen. Ook de beschrijving van lineaire afbeeldingen door matrices hangt af van de keuzes van basissen. Men moet dan ook het verband onderzoeken tussen de coördinaten van een element in een vectorruimte ten opzichte van verschillende basissen, en het verband tussen matrixvoorstellingen van een lineaire afbeelding ten opzichte van verschillende basiskeuzes.

Definitie 4.2.1. Zij V een n-dimensionale K-vectorruimte. Beschouw twee basissen $\mathcal{B} = \{b_1, \dots, b_n\}$ en $\mathcal{B}' = \{b'_1, \dots, b'_n\}$ voor V. Dan is ieder element van $\mathcal{B}'$ een lineaire combinatie van elementen van $\mathcal{B}$:

$$b'_j = q_{1j}b_1 + \cdots + q_{nj}b_n = \textstyle\sum_{i=1}^n q_{ij}b_i \,,$$

voor bepaalde scalairen $q_{ij} \in K$.

De resulterende matrix $Q = (q_{ij}) \in M_n(K)$ noemt men de *matrix van basisverandering* of de *transitiematrix* van de (oude) basis $\mathcal{B}$ naar de (nieuwe) basis $\mathcal{B}'$.

Opmerking 4.2.2. (i) Zij Q de transitiematrix van de (oude) basis $\mathcal{B}$ naar de (nieuwe) basis $\mathcal{B}' = \{b'_1, \dots, b'_n\}$. De i-de kolom van Q is de coordinatenvector van b'_i ten opzichte van de (oude) basis $\mathcal{B}$. In de praktijk stellen we de transitiematrix Q dus op door de coördinatenvectoren van $b'_1, \dots, b'_n$ in de kolommen van Q te schrijven.

(ii) We beschouwen de vectorruimte K^n. Een vaak voorkomende situatie is deze waarbij we de transitiematrix willen bepalen van de standaardbasis naar een willekeurige basis $\mathcal{B} = \{b_1, \dots, b_n\}$. Iedere $b_i \in \mathcal{B}$ is gelijk aan zijn coördinatenvector t.o.v. de standaardbasis. Bijgevolg is $P = (b_1, \dots, b_n) \in M_n(K)$ de transitiematrix van de standaardbasis naar $\mathcal{B}$.

We herkennen de situatie uit Opmerking 4.2.2(i) in wat we gedaan hebben met matrixvoorstelling van lineaire afbeeldingen; zie Opmerking 4.1.3. Deze analogie is geen toeval:

Stelling 4.2.3. *Zij V een n-dimensionale K-vectorruimte. Beschouw twee basissen $\mathcal{B} = \{b_1, \ldots, b_n\}$ en $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ voor V, en zij Q de transitiematrix van de (oude) basis $\mathcal{B}$ naar de (nieuwe) basis $\mathcal{B}'$. Dan is*

$$Q = A_{\mathbf{1}_V, \mathcal{B}', \mathcal{B}}.$$

(Let hierbij op de volgorde van de basissen $\mathcal{B}'$ en $\mathcal{B}$!)

Bewijs. De transitiematrix $Q = (q_{ij})$ is gedefinieerd door elk basiselement b'_j te schrijven als $b'_j = \sum_{i=1}^n q_{ij} b_i$ (zie Definitie 4.2.1). Beschouw anderzijds de identieke afbeelding $\mathbf{1}_V \colon V \to V$. Uit Definitie 4.1.2 weten we dat de matrixvoorstelling $A = (a_{ij})$ van $\mathbf{1}_V$ ten opzichte van de basissen $\mathcal{B}'$ en $\mathcal{B}$ gegeven wordt door $\mathbf{1}_V(b'_j) = \sum_{i=1}^n a_{ij} b_i$ te schrijven. Aangezien $\mathbf{1}_V(b'_j) = b'_j$ zien we dat $q_{ij} = a_{ij}$ voor alle i, j, en dus is $Q = A = A_{\mathbf{1}_V, \mathcal{B}', \mathcal{B}}$. $\square$

Lemma 4.2.4. *Zij $\mathcal{B} = \{b_1, \ldots, b_n\}$ en $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ twee basissen voor V, en zij Q de transitiematrix van de (oude) basis $\mathcal{B}$ naar de (nieuwe) basis $\mathcal{B}'$.*

(i) *De matrix Q is inverteerbaar. De inverse Q^{-1} is de transitiematrix van de (nieuwe) basis $\mathcal{B}'$ naar de (oude) basis $\mathcal{B}$.*

(ii) *We hebben $L_Q \circ \mathrm{co}_{\mathcal{B}'} = \mathrm{co}_{\mathcal{B}}$, met andere woorden,*

$$Q\mathrm{co}_{\mathcal{B}'}(v) = \mathrm{co}_{\mathcal{B}}(v) \quad \textit{en} \quad \mathrm{co}_{\mathcal{B}'}(v) = Q^{-1}\mathrm{co}_{\mathcal{B}}(v)$$

voor alle $v \in V$.

$$\begin{tikzcd} & V \arrow[dl, "\mathrm{co}_{\mathcal{B}'}"'] \arrow[dr, "\mathrm{co}_{\mathcal{B}}"] & \\ K^n \arrow[rr, "L_Q"] & & K^n \end{tikzcd}$$

Bewijs. (i) Uit Stelling 4.2.3 weten we dat $Q = A_{\mathbf{1}_V, \mathcal{B}', \mathcal{B}}$. Uiteraard is $\mathbf{1}_V \colon V \to V$ inverteerbaar, met inverse $\mathbf{1}_V$, en uit Stelling 4.1.8(iv) volgt nu dat Q inverteerbaar is met inverse $Q^{-1} = A_{\mathbf{1}_V, \mathcal{B}, \mathcal{B}'}$. Opnieuw wegens Stelling 4.2.3 is dit de transitiematrix van de basis $\mathcal{B}'$ naar de basis $\mathcal{B}$.

(ii) Dit is een onmiddellijk gevolg van Stelling 4.1.4(ii) toegepast op $Q = A_{\mathbf{1}_V, \mathcal{B}', \mathcal{B}}$. $\square$

Vervolgens willen we bestuderen wat de invloed is op de matrixvoorstelling van een lineaire afbeelding $f \colon V \to W$ als we voor beide vectorruimten een nieuwe basis kiezen.

Stelling 4.2.5. *Zij V, W eindig-dimensionale K-vectorruimten en kies basissen $\mathcal{B}$, $\mathcal{B}'$ voor V en basissen $\mathcal{C}$, $\mathcal{C}'$ voor W. Beschouw de transitiematrix Q van $\mathcal{B}$ naar $\mathcal{B}'$ en de transitiematrix P van $\mathcal{C}$ naar $\mathcal{C}'$.*

Zij nu $f : V \to W$ een lineaire afbeelding en stel $A = A_{f,\mathcal{B},\mathcal{C}}$ en $A' = A_{f,\mathcal{B}',\mathcal{C}'}$. Dan is

$$A' = P^{-1}AQ.$$

Bewijs. Uit Lemma 4.2.4(ii) weten we dat $L_Q \circ \mathrm{co}_{\mathcal{B}'} = \mathrm{co}_{\mathcal{B}}$ en dat $L_P \circ \mathrm{co}_{\mathcal{C}'} = \mathrm{co}_{\mathcal{C}}$. Anderzijds halen we uit Stelling 4.1.4(ii) dat $L_A \circ \mathrm{co}_{\mathcal{B}} = \mathrm{co}_{\mathcal{C}} \circ f$ en $L_{A'} \circ \mathrm{co}_{\mathcal{B}'} = \mathrm{co}_{\mathcal{C}'} \circ f$.

Uit de associativiteit van de samenstelling volgt nu dat

$$L_{A'} \circ \mathrm{co}_{\mathcal{B}'} = L_P^{-1} \circ \mathrm{co}_{\mathcal{C}} \circ f = L_P^{-1} \circ L_A \circ \mathrm{co}_{\mathcal{B}} = L_P^{-1} \circ L_A \circ L_Q \circ \mathrm{co}_{\mathcal{B}'}.$$

$$\begin{tikzcd}
& V \arrow[rrr, "f"] \arrow[dl, "\mathrm{co}_{\mathcal{B}'}"'] \arrow[dr, "\mathrm{co}_{\mathcal{B}}"] & & & W \arrow[dl, "\mathrm{co}_{\mathcal{C}}"'] \arrow[dr, "\mathrm{co}_{\mathcal{C}'}"] & \\
K^n \arrow[rr, "L_Q"] & & K^n \arrow[r, "L_A"] & K^m \arrow[rr, "L_P^{-1}"] & & K^m
\end{tikzcd}$$

Uit Lemma 4.1.7 volgt nu dat $A' = P^{-1}AQ$. $\square$

Opmerking 4.2.6. Zij $f \in \mathrm{End}(V)$, en zij Q de transitiematrix van de basis $\mathcal{B}$ naar de basis $\mathcal{B}'$. Dan krijgt de formule in Stelling 4.2.5 de gedaante $A' = Q^{-1}AQ$.

Definitie 4.2.7. Zij $A \in M_n(K)$. De verzameling

$$\{Q^{-1}AQ \mid Q \in \mathsf{GL}_n(K)\}$$

noemt men de *conjugatieklasse* van A; de elementen van deze verzameling zijn de *geconjugeerden* of *toegevoegden* van A.

Als A een matrixvoorstelling is van een lineaire operator f op een n-dimensionale K-vectorruimte V, dan bestaat een conjugatieklasse van A juist uit alle mogelijke matrixvoorstellingen van f.

Men kan de vraag stellen of een lineaire afbeelding $f : V \to W$ "interessante" matrixvoorstellingen heeft. In het algemeen, wanneer we de basissen in het domein V en de beeldruimte W vrij kunnen kiezen, is er een eenvoudig antwoord:

Stelling 4.2.8. *Zij $f : V \to W$ met $n = \dim V$ en $m = \dim W$. Dan bestaat er een basis $\mathcal{B}$ van V en een basis $\mathcal{C}$ van W waarvoor*

$$A_{f,\mathcal{B},\mathcal{C}} = \left(\begin{array}{c|c} I_k & 0_{k,n-k} \\ \hline 0_{m-k,k} & 0_{m-k,n-k} \end{array}\right) \quad \textit{voor een bepaalde } k \leq n, m.$$

Bewijs. We schrijven $V = V' \oplus \ker f$, dan is V' een complement van $\ker f$. Zij $\{b_1, \ldots, b_k\}$ een basis voor V' en $\{b_{k+1}, \ldots, b_n\}$ een basis voor $\ker f$, dan is $\mathcal{B} := \{b_1, \ldots, b_n\}$ een basis voor V.

Merk op dat de restrictie $f_{V'} \colon V' \to W$ injectief is. Wegens Stelling 3.3.2 impliceert dit dat de verzameling $\{f(b_1), \ldots, f(b_k)\}$ lineair onafhankelijk is. Omwille van Gevolg 2.4.8(i) kunnen we deze verzameling uitbreiden tot een basis $\mathcal{C} := \{f(b_1), \ldots, f(b_k), c_{k+1}, \ldots, c_m\}$ voor W. Ten opzichte van deze basissen $\mathcal{B}$ en $\mathcal{C}$ is de matrixvoorstelling van f van de gezochte vorm. □

Het probleem is interessanter voor lineaire *operatoren* $f \colon V \to V$, waarbij het natuurlijk is om te eisen dat we dezelfde basis kiezen voor het domein en de beeldruimte. We zullen dit probleem uitgebreid behandelen in Hoofdstuk 7.

4.3 Lineaire vormen en duale vectorruimte

In deze paragraaf bestuderen we de *duale ruimte* $\mathrm{Hom}_K(V, K)$ van een vectorruimte V in meer detail. Dit concept speelt een belangrijke rol, niet alleen in vele deelgebieden van de wiskunde, maar ook in toepassingen in de fysica (kwantummechanica) en in de computerwetenschappen (machine learning).

Definitie 4.3.1. (i) Zij V een K-vectorruimte. De *duale ruimte* V^* van V is de vectorruimte

$$V^* := \mathrm{Hom}_K(V, K) = \{f \colon V \to K \mid f \text{ is een lineaire afbeelding}\},$$

waarbij we K beschouwen als 1-dimensionale K-vectorruimte.

(ii) De elementen van V^* noemen we ook *lineaire vormen op* V.

Lineaire vormen op $V = K^n$ hebben een eenvoudige expliciete gedaante.

Lemma 4.3.2. *Zij $V = K^n$ en $f \colon V \to K$ een lineaire vorm. Dan bestaan er scalairen $\mu_1, \ldots, \mu_n \in K$ zodat*

$$f((x_1, \ldots, x_n)^t) = \mu_1 x_1 + \cdots + \mu_n x_n$$

voor alle $x_1, \ldots, x_n \in K$.

Omgekeerd definieert deze uitdrukking voor elke keuze van $\mu_1, \ldots, \mu_n \in K$ een lineaire vorm f op V.

Bewijs. Dit volgt onmiddellijk uit Lemma 3.3.1, met $\mathcal{B}$ de standaardbasis van K^n. □

De volgende observatie kan handig van pas komen. Merk op dat we niet onderstellen dat $\dim V$ eindig zou zijn.

Lemma 4.3.3. *Zij* $f\colon V \to K$ *een lineaire vorm, en zij* $v \in V$ *met* $f(v) \neq 0$. *Dan is* $V = \ker(f) \oplus \langle v \rangle$.

Bewijs. We tonen aan dat $\ker(f) \cap \langle v \rangle = 0$ en dat $V = \ker(f) + \langle v \rangle$; het resultaat volgt dan uit Lemma 2.5.2.

Aangezien $f(v) \neq 0$, is het evident dat $\ker(f) \cap \langle v \rangle = 0$. Stel nu $w \in V$ willekeurig. Dan is

$$w = \bigl(w - f(w)f(v)^{-1}v\bigr) + f(w)f(v)^{-1}v.$$

We zien dat $w - f(w)f(v)^{-1}v \in \ker(f)$ en $f(w)f(v)^{-1}v \in \langle v \rangle$, waaruit het gestelde volgt. □

We willen de duale ruimte van een eindig-dimensionale K-vectorruimte V nu explicieter voorstellen door te vertrekken van een basis van V. We maken gebruik van Stelling 3.4.5 (met $W = K$), en meer bepaald van de basis die we in het bewijs van deze stelling hebben ingevoerd, gegeven door de afbeeldingen (3.2). Omdat $W = K$ is $m = 1$, zodat j in (3.2) enkel de waarde 1 kan aannemen. We kiezen voor $W = K$ de basis $\mathcal{C} = \{1\}$. Dan is

$$\begin{cases} f_{i1}(e_i) = 1, \\ f_{i1}(e_k) = 0 \quad \text{voor alle } k \neq i, \end{cases}$$

en de verzameling van alle f_{i1} vormt een basis voor $V^* = \operatorname{Hom}(V, K)$. Uiteraard hoeven we de index 1 niet langer op te nemen in de notatie. Dit geeft aanleiding tot de volgende definitie, waarin we f_{i1} noteren als ε_i.

Definitie 4.3.4. Zij V een eindig-dimensionale K-vectorruimte en $\mathcal{B} = \{e_1, \dots, e_n\}$ een basis voor V. Definieer de lineaire vormen

$$\varepsilon_i\colon V \to K\colon \sum_{j=1}^{n} \lambda_j e_j \mapsto \lambda_i$$

voor alle $i = 1, \dots, n$. Stel nu $\mathcal{B}^* := \{\varepsilon_1, \dots, \varepsilon_n\}$. We noemen $\mathcal{B}^*$ de *duale basis* van $\mathcal{B}$.

Merk op dat $\varepsilon_i(e_j) = \delta_{ij}$ voor alle i, j, waarbij δ de Kronecker delta is die we ingevoerd hebben in Notatie 3.4.4.

Voorbeeld 4.3.5. Beschouw $V = K^n$ voorzien van de standaardbasis $\mathcal{B} = \{e_1, \dots, e_n\}$ zoals in Voorbeeld 2.4.2(1). Dan is ε_i de lineaire vorm gegeven door

$$\varepsilon_i \colon (\lambda_1, \dots, \lambda_n)^t \mapsto \lambda_i,$$

voor elke $i \in \{1, \dots, n\}$. We noemen dit ook wel de *i-de coördinaatafbeelding* op K^n. De lineaire vorm in Lemma 4.3.2 is dan precies $\mu_1\varepsilon_1 + \cdots + \mu_n\varepsilon_n$.

Opmerking 4.3.6. Als V eindig-dimensionaal is, dan volgt uit de voorgaande discussie (of rechtstreeks uit Stelling 3.4.5) dat $\dim V^* = \dim V$. Bijgevolg is $V \cong V^*$. Een expliciet isomorfisme wordt gegeven door een basis $\mathcal{B} = \{e_1, \dots, e_n\}$ voor V te kiezen, vervolgens de duale basis $\mathcal{B}^* = \{\varepsilon_1, \dots, \varepsilon_n\}$ op te stellen, en dan het isomorfisme van V naar V^* te beschouwen dat elke e_i afbeeldt op ε_i. Dit isomorfisme is echter niet "canoniek": het hangt af van de keuze van de basis.

Het is interessant om op te merken dat er wél een canoniek isomorfisme bestaat van V naar V^{**}, waarbij dus $V^{**} = \operatorname{Hom}(V^*, K)$. Inderdaad, we kunnen dit isomorfisme opschrijven zonder te refereren naar een basis – we spreken van een *coördinaatvrije* constructie – door elke $v \in V$ af te beelden op de afbeelding $g_v \colon V^* \to K \colon f \mapsto f(v)$. Het is een goede oefening om aan te tonen dat de afbeelding $v \mapsto g_v$ een injectieve lineaire afbeelding is van V naar V^{**}; aangezien V eindig-dimensionaal is, volgt hieruit dat het een isomorfisme is.

Opmerking 4.3.7. Zij V een *oneindig-dimensionale* K-vectorruimte en $\mathcal{B}$ een basis van V. Net zoals in Definitie 4.3.4 definiëren we, voor iedere $b \in \mathcal{B}$,

$$\varepsilon_b \colon V \to K \colon \sum_{v \in \mathcal{B}} \lambda_v v \to \lambda_b.$$

Er geldt dat $\varepsilon_b \in V^*$. In dit geval is de verzameling $\{\varepsilon_b \mid b \in \mathcal{B}\}$ echter geen basis voor V^*! (Ze is wel lineair onafhankelijk, maar niet voortbrengend.) Sterker nog, voor oneindig-dimensionale vectorruimten is zelfs steeds $\dim V^* > \dim V$, en dus zijn V^* en V niet isomorf aan elkaar.

Voorbeeld 4.3.8. Voor de geïnteresseerde lezer geven we een voorbeeld ter toelichting van de vorige opmerking. Beschouw de oneindig-dimensionale K-vectorruimte $V = K[x]$ van veeltermen over een veld K, met basis $\mathcal{B} = \{1, x, x^2, \dots\}$. Als $b = x^k \in \mathcal{B}$, dan is de duale afbeelding ε_b de afbeelding $\varepsilon_b \colon K[x] \to K$ die een veelterm afbeeldt op de coëfficiënt horend bij de term x^k.

Beschouw nu de lineaire vorm $\varphi \colon K[x] \to K \colon f(x) \mapsto f(1)$. Dan is φ niet te schrijven als lineaire combinatie van de ε_b's. Inderdaad, stel dat

$\varphi = \sum_{b\in\mathcal{B}} \lambda_b\varepsilon_b$, waarbij de som een eindige som is. Als het basiselement met hoogste graad dat voorkomt in deze som, gelijk is aan x^k, dan is $\varphi(x^{k+1}) = 0$, terwijl per definitie van φ we zouden moeten hebben dat $\varphi(x^{k+1}) = 1$; dit is een contradictie. We besluiten dat $\varphi \in V^*$ terwijl $\varphi \notin \operatorname{span}(\{\varepsilon_b \mid b \in \mathcal{B}\})$.

Als $f\colon V \to W$ een lineaire afbeelding is, kunnen we met f een afbeelding $f^*\colon W^* \to V^*$ associëren.

Lemma 4.3.9. *Zij V en W twee K-vectorruimten, en $f\colon V \to W$ een lineaire afbeelding.*

(i) *Voor elke $\varphi \in W^*$ is $\varphi \circ f \in V^*$.*

(ii) *De afbeelding $f^*\colon W^* \to V^*\colon \varphi \mapsto \varphi \circ f$ is een lineaire afbeelding.*

Bewijs. (i) Aangezien $f \in \operatorname{Hom}_K(V, W)$ en $\varphi \in \operatorname{Hom}_K(W, K)$, volgt uit Lemma 3.5.1 dat $\varphi \circ f \in \operatorname{Hom}_K(V, K) = V^*$.

(ii) Om na te gaan dat f^* een lineaire afbeelding is, gaan we na dat voor alle $\varphi, \psi \in W^*$ en alle $\lambda, \mu \in K$ geldt dat

$$f^*(\lambda\varphi + \mu\psi) = \lambda f^*(\varphi) + \mu f^*(\psi)\,.$$

Dit is equivalent met de gelijkheid

$$(\lambda\varphi + \mu\psi) \circ f = \lambda(\varphi \circ f) + \mu(\psi \circ f)\,,$$

die volgt door beide leden te laten inwerken op elk element $v \in V$. $\square$

Definitie 4.3.10. De lineaire afbeelding $f^*\colon W^* \to V^*\colon \varphi \mapsto \varphi \circ f$ noemen we de *duale afbeelding* van f.

Opmerking 4.3.11. Zij V een n-dimensionale vectorruimte met basis $\mathcal{B} = \{b_1, \ldots, b_n\}$. We beschouwen de duale ruimte $V^* = \operatorname{Hom}_K(V, K)$. We kiezen in de K-vectorruimte K de basis $\{1\}$. Als we Stelling 4.1.8(i) toepassen vinden we dat $f \mapsto A_f$ een isomorfisme bepaalt van V^* naar $M_{1,n}(K)$, de vectorruimte van de rijvectoren. Concreet, zij $f \in V^*$, dan is

$$A_f = (f(b_1), \ldots, f(b_n)) \in M_{1,n}(K).$$

Voor $V = K^n$ met de standaardbasis krijgen we dat een $f \in V^*$ met de notatie uit Lemma 4.3.2 als matrixvoorstelling precies $A_f = (\mu_1, \ldots, \mu_n) \in M_{1,n}(K)$ heeft. We zien dus dat het erg natuurlijk is om de elementen van V voor te stellen als *kolommatrices* en de elementen van V^* als *rijmatrices*.

Stelling 4.3.12. *Zij V een n-dimensionale K-vectorruimte en W een m-dimensionale K-vectorruimte. Zij $f\colon V \to W$ een lineaire afbeelding, en zij $f^*\colon W^* \to V^*$ de corresponderende duale afbeelding tussen de duale ruimten. Als A_f de matrixvoorstelling is ten opzichte van een basis $\mathcal{B}$ in V en een basis $\mathcal{C}$ in W, dan is de getransponeerde matrix A_f^t de matrixvoorstelling van f^* ten opzichte van de duale basis van $\mathcal{C}$ en de duale basis van $\mathcal{B}$.*

Bewijs. Oefening. (Dit wordt besproken in de oefeningenlessen.) $\square$

Hoofdstuk

5 De rang van een matrix

In dit kortere hoofdstuk voeren we de *rang* van een matrix in. We gebruiken dit begrip om de oplosbaarheid van een stelsel van lineaire vergelijkingen te bepalen, evenals de ‘grootte’ van de oplossingsverzameling. We maken van de gelegenheid gebruik om affiene deelruimten in te voeren, die niet alleen nodig zijn in de context van stelsels, maar ook erg nuttig zijn vanuit meetkundig oogpunt, zoals we later in Hoofdstuk 9 meer in detail zullen zien.

5.1 De rang van een matrix

Definitie 5.1.1. Zij $A \in M_{m,n}(K)$ een matrix met kolommen $A_1, \dots, A_n$ en rijen $R_1, \dots, R_m$.

(i) We definiëren de *kolomruimte* van A als de deelruimte van K^m opgespannen door de kolommen van A, dus

$$C(A) := \operatorname{span}(A_1, \dots, A_n) \leq K^m.$$

Analoog definiëren we de *rijruimte* van A als de deelruimte van K^n opgespannen door de rijen van A, dus

$$R(A) := \operatorname{span}(R_1, \dots, R_m) \leq K^n.$$

Merk op dat we in principe de rijen zouden moeten transponeren tot kolommen om in K^n terecht te komen, maar dat zullen we meestal niet expliciet doen.

(ii) We definiëren de *kolomrang* van A als de dimensie van $C(A)$ en de *rijrang* van A als de dimensie van $R(A)$.

Stelling 5.1.2. *Zij $A \in M_{m,n}(K)$. Dan is de kolomrang van A gelijk aan de rijrang van A.*

Bewijs. Noteer de kolommen van A als $A_1, \dots, A_n$ en de rijen als $R_1, \dots, R_m$. Stel $s := \dim C(A)$ en $r := \dim R(A)$.

Beschouw een willekeurige basis $\mathcal{B} = \{B_1, \ldots, B_s\}$ van de deelruimte $C(A) \leq K^m$, en stel B gelijk aan de matrix gevormd door deze s kolommen, dus $B = \begin{pmatrix} B_1 & \cdots & B_s \end{pmatrix} \in M_{m,s}(K)$. Aangezien $\mathcal{B}$ een basis is, is iedere kolom A_i een lineaire combinatie van $\mathcal{B}$. We noteren voor alle $1 \leq k \leq n$

$$A_k = \lambda_{1k} B_1 + \cdots + \lambda_{sk} B_s,$$

met $\lambda_{ij} \in K$. Definieer de matrix $D = (\lambda_{ij}) \in M_{s,n}(K)$. Per definitie is dan $A = BD$.

Noteer met $\{D_1, \ldots, D_s\}$ de rijen van D. Uit $A = BD$ volgt dat iedere rij R_i van A een lineaire combinatie is van $\{D_1, \ldots, D_s\}$; bijgevolg is

$$r = \dim \operatorname{span}(R_1, \ldots, R_m) \leq \dim \operatorname{span}(D_1, \ldots, D_s) \leq s.$$

We hebben dus aangetoond dat $r \leq s$.

Wanneer we de redenering opnieuw doen voor de matrix A^t bekomen we dat $s \leq r$. We concluderen dat $r = s$. □

Definitie 5.1.3. Aangezien de kolomrang en de rijrang van een matrix A dus altijd gelijk zijn aan elkaar, noemen we dit eenvoudigweg de *rang* van de matrix A. We noteren dit als $\operatorname{rk}(A)$, dus

$$\operatorname{rk}(A) := \dim C(A) = \dim R(A).$$

Opmerking 5.1.4. De rang $\operatorname{rk}(A)$ is dus gekarakteriseerd als het (unieke) getal r zodat er een verzameling bestaat van r kolommen (resp. rijen) van A die lineair onafhankelijk is, en iedere verzameling van meer dan r kolommen (resp. rijen) van A lineair afhankelijk is.

Voorbeelden 5.1.5. (1) Er geldt dat $\operatorname{rk}(0_{m,n}) = 0$ en dat $\operatorname{rk}(I_n) = n$.

(2) Zij $A = (a_{ij}) \in M_{m,n}(K) \setminus \{0\}$ met $a_{ij} = a_{k\ell}$ voor alle $1 \leq i, k \leq m$ en $1 \leq j, \ell \leq n$. Dan is $\operatorname{rk}(A) = 1$.

(3) Zij A een echelonmatrix. Dan is $\operatorname{rk}(A)$ gelijk aan het aantal spilplaatsen. Inderdaad, de verzameling van de spilkolommen is immers een basis voor de ruimte voortgebracht door de kolommen van A.

De volgende eigenschappen van de rang van een matrix zijn eenvoudig maar belangrijk.

Lemma 5.1.6. *Zij $A \in M_{m,n}(K)$ met $\operatorname{rk}(A)$ de rang van A.*

(i) *Beschouw de lineaire afbeelding $L_A \colon K^n \to K^m \colon v \mapsto Av$. Dan is*

$$\operatorname{im}(L_A) = C(A) = \operatorname{span}(A_1, \ldots, A_n);$$

bijgevolg is $\dim(\operatorname{im} L_A) = \operatorname{rk}(A)$.

(ii) *Elementaire rij-operaties laten de rang van een matrix invariant.*

(iii) *Zij* $A \in M_{m,n}(K)$ *en* $B \in M_{n,r}(K)$. *Dan is* $\mathrm{rk}(AB) \le \mathrm{rk}(A)$ *en* $\mathrm{rk}(AB) \le \mathrm{rk}(B)$.

(iv) *Zij* $A \in \mathsf{GL}_n(K)$ *een inverteerbare matrix en zij* $B \in \mathrm{Mat}_{n,r}(K)$. *Dan is* $\mathrm{rk}(AB) = \mathrm{rk}(B)$.

(v) *Zij* $A \in \mathsf{GL}_n(K)$ *een inverteerbare matrix en zij* $B \in \mathrm{Mat}_{r,n}(K)$. *Dan is* $\mathrm{rk}(BA) = \mathrm{rk}(B)$.

Bewijs. (i) Zij $\{e_1, \dots, e_n\}$ de standaardbasis in K^n. Het beeld van L_A wordt voortgebracht door $\{Ae_1, \dots, Ae_n\}$ (zie Lemma 3.3.1(i)). Nu is $Ae_i \in K^m$ gelijk aan de i-de kolom van A.

(ii) Na het toepassen van een elementaire rij-operatie blijft de ruimte opgespannen door de rijen van A onveranderd.

(iii) Merk op dat $\mathrm{im}(L_{AB})$ een deelruimte is van $\mathrm{im}\, L_A$, want $(AB)v = A(Bv) \in \mathrm{im}\, L_A$ voor elke $v \in K^r$. In het bijzonder is $\dim(\mathrm{im}(L_{AB})) \le \dim(\mathrm{im}\, L_A)$, en uit (i) volgt dat $\mathrm{rk}(AB) \le \mathrm{rk}(A)$.
De andere ongelijkheid volgt nu door transponeren:

$$\mathrm{rk}(AB) = \mathrm{rk}((AB)^t) = \mathrm{rk}(B^t A^t) \le \mathrm{rk}(B^t) = \mathrm{rk}(B).$$

(iv) Uit (iii) volgt reeds dat $\mathrm{rk}(AB) \le \mathrm{rk}(B)$. Omdat A inverteerbaar is, kunnen we echter B nog herschrijven als $B = A^{-1} \cdot AB$, zodat opnieuw uit (iii) volgt dat $\mathrm{rk}(B) = \mathrm{rk}(A^{-1} \cdot AB) \le \mathrm{rk}(AB)$. Deze twee ongelijkheden samen geven ons $\mathrm{rk}(B) = \mathrm{rk}(AB)$.

(v) Analoog aan (iv). □

Men kan dus de rang van een matrix berekenen door de matrix eerst in echelonvorm te brengen met behulp van elementaire rij-operaties en dan de rang af te lezen (zie Voorbeeld 5.1.5(3)).

5.2 Affiene deelruimten

We zullen zo dadelijk het concept van rang van een matrix gebruiken om de oplosbaarheid en het aantal oplossingen van een stelsel lineaire vergelijkingen weer te geven. In Lemma 5.3.1 tonen we aan dat de oplossingsverzameling van een niet-strijdig stelsel een zogenaamde *affiene deelruimte* is. We voeren dit begrip nu in.

Definitie 5.2.1. Zij V een K-vectorruimte, beschouw een element $v \in V$ en een deelruimte $W \leq V$. Definieer de deelverzameling

$$v + W := \{v + w \mid w \in W\} \subseteq V\,;$$

$v+W$ is dus de verzameling van alle sommen $v+w$ met $w \in W$. We noemen $v + W$ een *affiene deelruimte* of een *lineaire variëteit*.

Opmerking 5.2.2. Als $v \notin W$, dan is de affiene deelruimte $v + W$ geen deelruimte van V! Immers, als $v \notin W$, dan is ook $-v \notin W$ waardoor $0 = v + (-v) \notin v + W$. We kunnen $v + W$ zien als een *verschuiving* van de deelruimte W. (Als $v \in W$, dan is $v + W = W$ wel een deelruimte van V.)

Om het verschil tussen deelruimten en affiene deelruimten te benadrukken, zullen we soms spreken over *vectordeelruimten* versus *affiene deelruimten*.

Voorbeeld 5.2.3. (1) Zij $W \leq V$ een deelruimte. Dan is W ook steeds een affiene deelruimte. We kunnen W immers schrijven als $0 + W$.

(2) Beschouw de vectorruimte $\mathbb{R}^3$. De deelverzameling

$$S = \{(1 + 2r, -1 - s, r + s)^t \mid r, s \in \mathbb{R}\}$$

is een affiene deelruimte van $\mathbb{R}^3$, we kunnen S immers ook schrijven als

$$S = (1, -1, 0)^t + \operatorname{span}(\{(2, 0, 1)^t, (0, -1, 1)^t\}).$$

Deze affiene deelruimte is geen deelruimte van $\mathbb{R}^3$.

Zij $v \in V$ en $w \in W \leq V$, dan zijn de verzamelingen $v+W$ en $(v+w)+W$ gelijk. Het volgend lemma toont aan dat elke affiene deelruimte een unieke deelruimte bepaalt waarvan ze de verschuiving is.

Lemma 5.2.4. *Zij V een vectorruimte, $v_1, v_2 \in V$ en W_1, W_2 deelruimten van V. Dan is $v_1+W_1 = v_2+W_2$ als en slechts als $W_1 = W_2$ en $v_1-v_2 \in W_1$.*

Bewijs. Als $W_1 = W_2$ en $v_1 - v_2 = w_0 \in W_1$, dan is

$$v_1 + W_1 = \{v_1 + w \mid w \in W_1\} \text{ en } v_2 + W_2 = \{v_1 - w_0 + w \mid w \in W_1\}.$$

Als w door alle elementen van W_1 loopt dan loopt $w - w_0$ ook door alle elementen van W_1. De twee verzamelingen zijn dus gelijk.

Omgekeerd, zij $v_1 + W_1 = v_2 + W_2$. Stel $w_0 = v_1 - v_2$, dan volgt dat $w_0 + W_1 = W_2$. Vermits $0 \in W_2$ moet $w_0 \in W_1$. Maar dan is $w_0 + W_1 = W_1$, en bijgevolg $W_1 = W_2$. $\square$

Definitie 5.2.5. Zij $v+W$ een affiene deelruimte van een K-vectorruimte V. Dan definiëren we de *dimensie* van $v+W$ als $\dim(v+W) := \dim W$.

Opmerking 5.2.6. We moeten nagaan dat de definitie van dimensie van een affiene deelruimte *goed gedefinieerd* is. Hiermee bedoelen we dat de neergeschreven formule niet mag afhangen van de keuze van de elementen die we gebruiken om de formule neer te schrijven. Meer bepaald moeten we hier nagaan dat, als $v_1+W_1 = v_2+W_2$ twee schrijfwijzen zijn van *dezelfde* affiene deelruimte, dat dan $\dim(W_1) = \dim(W_2)$. Dat dit inderdaad het geval is, volgt onmiddellijk uit Lemma 5.2.4.

5.3 De oplossingsverzameling van een stelsel

We keren nu terug naar de stelsels, en we tonen aan dat de oplossingsverzameling van een niet-strijdig stelsel een affiene deelruimte is.

Lemma 5.3.1. *Beschouw het stelsel $AX = w$ met $A \in M_{m,n}(K)$, $w \in K^m$ en X een kolomvector met n onbekenden. Beschouw de lineaire afbeelding*

$$L_A \colon K^n \to K^m \colon v \mapsto Av.$$

(i) *Het stelsel $AX = w$ is strijdig als en slechts als $w \notin \operatorname{im} L_A$.*

(ii) *Als het stelsel $AX = w$ niet strijdig is, dan is de oplossingsverzameling een affiene deelruimte. Als $v_0 \in K^n$ een oplossing is, dan is de oplossingsverzameling gegeven door $L_A^{-1}(w) = v_0 + \ker L_A$.*

(iii) *Als het stelsel homogeen is, m.a.w. van de vorm $AX = 0$, dan is de oplossingsverzameling gelijk aan de deelruimte* $\ker L_A$.

Bewijs. Merk vooreerst op dat de oplossingsverzameling van $AX = w$ gelijk is aan

$$\{v \in K^n \mid Av = w\} = \{v \in K^n \mid L_A(v) = w\} = L_A^{-1}(w).$$

(i) De oplossingsverzameling $L_A^{-1}(w)$ is ledig als en slechts als $w \notin \operatorname{im} L_A$.

(ii) Stel dat er een oplossing $L_A(v_0) = w$ is. Zij nu $u \in K^n$ willekeurig. Dan is u een oplossing van het stelsel als en slechts als $L_A(u) = w = L_A(v_0)$, als en slechts als $L_A(u - v_0) = 0$, als en slechts als $u - v_0 \in \ker L_A$, als en slechts als $u \in v_0 + \ker L_A$.

(iii) Dit is een bijzonder geval van (ii), rekening houdend met het feit dat een homogeen stelsel nooit strijdig is, en steeds de nuloplossing $v_0 = 0$ heeft. $\square$

Stelling 5.3.2. *Beschouw het stelsel $AX = w$ met $A \in M_{m,n}(K)$, $w \in K^m$ en X een kolomvector met n onbekenden.*

(i) *Het stelsel heeft een oplossing als en slechts als* $\mathrm{rk}(A) = \mathrm{rk}(A|w)$.

(ii) *Stel dat het stelsel niet strijdig is. Dan is de dimensie van de oplossingsverzameling gelijk aan* $n - \mathrm{rk}(A)$.

Bewijs. (i) Zij $\{A_1, \dots, A_n\}$ de kolommen van de matrix A. Er geldt dat

$$\begin{aligned} &\mathrm{rk}(A) = \mathrm{rk}(A|w) \\ &\qquad \iff \dim \mathrm{span}(A_1, \dots, A_n) = \dim \mathrm{span}(A_1, \dots, A_n, w) \\ &\qquad \iff \mathrm{span}(A_1, \dots, A_n) = \mathrm{span}(A_1, \dots, A_n, w) \\ &\qquad \iff w \in \mathrm{span}(A_1, \dots, A_n). \end{aligned}$$

Uit Lemma 5.3.1(i) weten we dat het stelsel een oplossing heeft als en slechts als $w \in \mathrm{im}(L_A)$. De bewering volgt nu wegens Lemma 5.1.6(i).

(ii) Uit het bovenstaande lemma volgt dat de oplossingsverzameling de affiene deelruimte $v_0 + \ker L_A$ is, met v_0 een oplossing. We hebben dat

$$\begin{aligned} \dim(v_0 + \ker L_A) &= \dim(\ker L_A) \\ &= \dim(K^n) - \dim(\mathrm{im}\, L_A) = n - \mathrm{rk}(A), \end{aligned}$$

waar we gebruik gemaakt hebben van achtereenvolgens Definitie 5.2.5, Stelling 3.3.9 en Lemma 5.1.6(i). □

Wanneer we een stelsel met even veel onbekenden als vergelijkingen beschouwen, hebben we de volgende sterkere eigenschap.

Gevolg 5.3.3. *Beschouw het stelsel $AX = w$, $A \in M_n(K)$. Als* $\mathrm{rk}(A) = n$, *dan heeft het stelsel een unieke oplossing.*

Bewijs. Aangezien $\mathrm{rk}(A) = n$ moet ook $\mathrm{rk}(A|w) = n$, en dus volgt uit Stelling 5.3.2 dat het stelsel een oplossing heeft. Aangezien $n - \mathrm{rk}(A) = 0$ volgt dat de oplossingsverzameling dimensie 0 heeft, en dus is er een unieke oplossing. □

We zullen zien in Stelling 6.4.1 verderop dat een matrix $A \in M_n(K)$ met $\mathrm{rk}(A) = n$ steeds inverteerbaar is. De unieke oplossing van het stelsel $AX = w$ met $A \in M_n(K)$ en $\mathrm{rk}(A) = n$, wordt dan gegeven door $x = A^{-1}w$. (Merk op dat inderdaad $A(A^{-1}w) = w$.)

Opmerking 5.3.4. We hebben nu ogenschijnlijk twee verschillende stellingen gezien in verband met de oplossingsverzameling van een lineair stelsel, met name Stelling 1.4.13 en Stelling 5.3.2. We gaan na dat deze twee stellingen in feite op hetzelfde neerkomen.

Zij dus $AX = w$ een stelsel in m vergelijkingen en n onbekenden. Omwille van Lemma 5.1.6(ii) mogen we, om Stelling 5.3.2 toe te passen, aannemen dat de uitgebreide matrix $(A|w)$ in echelonvorm staat.

Volgens Stelling 1.4.13(i) is het stelsel strijdig als en slechts als de laatste kolom van $(A|w)$ een spilkolom is, terwijl volgens Stelling 5.3.2(i) het stelsel strijdig is als en slechts als $w \notin \operatorname{im} L_A$. Dit komt inderdaad op hetzelfde neer, want de laatste kolom van $(A|w)$ is een spilkolom als en slechts als $\operatorname{rk}(A|w) = \operatorname{rk}(A) + 1$, wat op zijn beurt equivalent is met het feit dat $w \notin \operatorname{span}(A_1, \dots, A_n)$ waarbij $A_1, \dots, A_n$ de kolommen van A zijn.

Veronderstel nu dat het stelsel $AX = w$ niet strijdig is. Volgens Gevolg 1.4.14 is de dimensie van de oplossingsverzameling (daar uitgedrukt als "het aantal vrij te kiezen onbekenden") gelijk aan het aantal onbekenden min het aantal spilkolommen van de echelonmatrix, terwijl volgens Stelling 5.3.2(ii) deze dimensie gelijk is aan $n - \operatorname{rk}(A)$. Dat dit op hetzelfde neerkomt, volgt nu onmiddellijk uit Voorbeeld 5.1.5(3).

Voorbeeld 5.3.5. Beschouw het stelsel $AX = w$ met 3 vergelijkingen en 6 onbekenden over $\mathbb{R}$, met uitgebreide matrix

$$(A|w) = \left(\begin{array}{cccccc|c} 1 & 2 & 0 & 3 & 4 & 0 & 5 \\ 0 & 0 & 1 & 6 & 7 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right).$$

De verzameling S van de indices van de spilkolommen is $\{1, 3, 6\}$. Volgens de formule in Stelling 1.4.13(ii) is de oplossingsverzameling

$$\bigl\{(5 - 2t_1 - 3t_2 - 4t_3,\ t_1,\ 8 - 6t_2 - 7t_3,\ t_2,\ t_3,\ 9)^t \mid t_1, t_2, t_3 \in K\bigr\}.$$

Dit is gelijk aan de verzameling

$$\begin{pmatrix} 5 \\ 0 \\ 8 \\ 0 \\ 0 \\ 9 \end{pmatrix} + \left\{ t_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -3 \\ 0 \\ -6 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} -4 \\ 0 \\ -7 \\ 0 \\ 1 \\ 0 \end{pmatrix} \;\middle|\; t_1, t_2, t_3 \in K \right\}.$$

We zien dat dit inderdaad een affiene deelruimte is, en dat de dimensie van deze affiene deelruimte ook exact het aantal vrij gekozen variabelen is.

Het is eenvoudig na te gaan dat de 3 basisvectoren van de deelruimte behorende bij de affiene ruimte in $\ker L_A$ zitten. Aangezien

$$\dim(\ker L_A) = n - \dim(\operatorname{im} L_A) = 6 - 3 = 3,$$

is dit ook de volledige kern.

Hoofdstuk

6 Determinanten

In dit hoofdstuk introduceren we een belangrijke afbeelding op $M_n(K)$ met waarden in K, namelijk de *determinantafbeelding* (of kortweg de *determinant*). We zullen deze eerst definiëren en een aantal eigenschappen oplijsten.

Voor de studenten van de opleiding Fysica en Sterrenkunde volstaat het om de definities en deze eigenschappen te kennen te kunnen toepassen (zonder bewijs). Voor de studenten van de opleiding Wiskunde gaan we een stap verder, door de determinant (opnieuw) in te voeren als een multilineaire afbeelding op de kolommen van een matrix. Dit blijkt bijzonder nuttig te zijn om voorgaande eigenschappen te bewijzen op een erg efficiënte manier.

Nadien zullen we de belangrijkste toepassing van de determinant behandelen, met name het verband met inverteerbaarheid van matrices.

6.1 Permutaties

Alvorens we de determinant kunnen definiëren, moeten we een aantal begrippen en notaties invoeren in verband met permutaties van $\{1, \ldots, n\}$.

Definitie 6.1.1. Zij $n \in \mathbb{N} \setminus \{0\}$.

(i) Een *permutatie* van $\{1, \ldots, n\}$ is een bijectie $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$. De verzameling van alle permutaties van $\{1, \ldots, n\}$ noteren we als $\mathrm{Sym}(n)$ of S_n.

(ii) We stellen een permutatie σ vaak voor door de opeenvolgende beelden $\sigma(1), \ldots, \sigma(n)$ tussen haakjes te plaatsen; we noteren dus $\sigma = (i_1, \ldots, i_n)$ waarbij $i_k = \sigma(k)$.

(iii) Zij $\sigma = (i_1, \ldots, i_n)$ een willekeurige permutatie. We kunnen steeds $(i_1, \ldots, i_n)$ omvormen in $(1, \ldots, n)$ door een aantal keer na elkaar twee elementen van plaats te verwisselen. Als dit mogelijk is in k stappen, dan definiëren we het *teken* van de permutatie σ als $\mathrm{sgn}(\sigma) = (-1)^k$; het teken is dus steeds 1 of -1. We spreken ook van *even* permutaties (als het teken 1 is) en *oneven* permutaties (als het teken -1 is), en we noemen dit de *pariteit* van de permutatie.

Opmerking 6.1.2. (i) Het is mogelijk om op een heel aantal verschillende manieren $(i_1, \dots, i_n)$ om te vormen in $(1, \dots, n)$ door een aantal keer twee elementen van plaats te verwisselen. Men kan aantonen dat het echter niet mogelijk is om dit enerzijds te doen in een even aantal stappen en anderzijds in een oneven aantal stappen. Hieruit volgt dat het teken van een permutatie goed gedefinieerd is.

(ii) Als σ een permutatie is van $\{1, \dots, n\}$, dan is ook de inverse afbeelding σ^{-1} een permutatie van diezelfde verzameling. Bovendien is $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$, want als σ bekomen wordt door k keer na elkaar twee elementen van plaats te verwisselen, dan wordt σ^{-1} bekomen door diezelfde k verwisselingen uit te voeren in de omgekeerde volgorde.

We bekijken de verzameling S_n voor kleine waarden van n.

Voorbeeld 6.1.3. (1) Zij $n = 2$. Dan zijn er juist twee permutaties van de verzameling $A = \{1, 2\}$, met name

$$\sigma_1 \colon A \to A \colon \begin{cases} 1 & \mapsto & 1 \\ 2 & \mapsto & 2\,, \end{cases}$$

en

$$\sigma_2 \colon A \to A \colon \begin{cases} 1 & \mapsto & 2 \\ 2 & \mapsto & 1\,. \end{cases}$$

We noteren $\sigma_1 = (1, 2)$ en $\sigma_2 = (2, 1)$, het is duidelijk dat $\operatorname{sgn}(\sigma_1) = (-1)^0 = 1$ en $\operatorname{sgn}(\sigma_1) = (-1)^1 = -1$.

(2) Zij $n = 3$. Dan zijn er juist zes permutaties van de verzameling $A = \{1, 2, 3\}$, met name

$$(1, 2, 3),\ (1, 3, 2),\ (2, 1, 3),\ (2, 3, 1),\ (3, 1, 2),\ (3, 2, 1)\,.$$

Het is duidelijk dat $\operatorname{sgn}(1, 2, 3) = 1$ en dat $\operatorname{sgn}(1, 3, 2) = \operatorname{sgn}(2, 1, 3) = \operatorname{sgn}(3, 2, 1) = (-1)^1 = -1$. We bepalen $\operatorname{sgn}(2, 3, 1)$:

$$(2, 3, 1) \to (2, 1, 3) \to (1, 2, 3),$$

er geldt dus dat $\operatorname{sgn}(2, 3, 1) = (-1)^2 = 1$. Analoog is $\operatorname{sgn}(3, 1, 2) = 1$.

Merk op dat $|S_n| = n! = 1 \cdot 2 \cdots n$ voor alle $n \in \mathbb{N} \setminus \{0\}$. Men kan aantonen dat S_n een groep is, met als bewerking de samenstelling van permutaties; de groep S_n is niet-abels zodra $n \geq 3$.

6.2 Determinanten: definitie, eigenschappen

We definiëren nu de determinant van een willekeurige vierkante matrix.

Definitie 6.2.1. Zij $A \in M_n(K)$. De *determinant* van A is de scalair

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n} \in K\,.$$

Voor elke waarde van $n \in \mathbb{N} \setminus \{0\}$ krijgen we dus een afbeelding

$$\det\colon M_n(K) \to K \colon A \mapsto \det(A)$$

die we dan de *determinantafbeelding* of kortweg de *determinant* noemen.

Voorbeeld 6.2.2. Als voorbeeld berekenen we wat de determinant van een willekeurige 2×2-matrix en 3×3-matrix is.

(1) Zij $A = (a_{ij}) \in M_2(K)$. Uit Voorbeeld 6.1.3 weten we dat $S_2 = \{(1,2),(2,1)\}$ met $\operatorname{sgn}(1,2) = 1$ en $\operatorname{sgn}(2,1) = -1$. Bijgevolg is

$$\det A = \sum_{\sigma \in \{(12),(21)\}} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} = a_{11}a_{22} - a_{21}a_{12}.$$

(2) Zij $A = (a_{ij}) \in M_3(K)$. Uit Voorbeeld 6.1.3 weten we dat

$$S_3 = \{(1,2,3),(2,3,1),(3,1,2),(1,3,2),(3,2,1),(2,1,3)\}$$

met

$$\begin{aligned} &\operatorname{sgn}(1,2,3) = \operatorname{sgn}(2,3,1) = \operatorname{sgn}(3,1,2) = 1 \quad \text{en} \\ &\operatorname{sgn}(1,3,2) = \operatorname{sgn}(3,2,1) = \operatorname{sgn}(2,1,3) = -1. \end{aligned}$$

Er volgt dat

$$\begin{aligned} \det A = \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma)\, a_{\sigma(1)1}a_{\sigma(2)2}a_{\sigma(3)3} \\ = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33}. \end{aligned}$$

Deze formule voor de determinant van een 3×3-matrix wordt ook wel de *regel van Sarrus* genoemd.

$$\left| \begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array} \right.$$

In de praktijk zullen we niet vaak gebruik maken van de formule in Definitie 6.2.1 om determinanten uit te rekenen, maar de formule is wel van groot theoretisch belang en zullen we dus nog vaker ontmoeten in deze cursus.

Een gebruiksvriendelijkere methode om de determinant van een matrix te bepalen, is om de matrix te *ontwikkelen* naar een rij of naar een kolom:

Constructie 6.2.3. We construeren de afbeelding $\det\colon M_n(K) \to K$ recursief. We definiëren de determinant dus eerst voor $M_1(K)$. Vervolgens nemen we aan dat de determinantafbeelding geconstrueerd is voor $M_{n-1}(K)$ en geven een formule om de determinant op $M_n(K)$ te bepalen in functie van de determinant voor matrices in $M_{n-1}(K)$.

- Voor 1×1-matrices stellen we $\det(a) := a$.
- We definiëren de $(n-1) \times (n-1)$-matrix A_{ij} als de matrix bekomen door in A de i-de rij en de j-de kolom te schrappen:

$$A_{ij} := \left(\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ \hline a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{array}\right)$$

- We nemen aan dat "det" gedefinieerd is voor $(n-1)\times(n-1)$ matrices; dan definiëren we, voor een vast gekozen $k \in \{1, \ldots, n\}$,

$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} a_{ki} \det(A_{ki}). \tag{6.1}$$

  We zeggen dat we *ontwikkelen* naar de k-de rij.

We zullen in Stelling 6.3.4 aantonen dat de determinantafbeelding gegeven door formule (6.1) samenvalt met deze die we hebben ingevoerd in Definitie 6.2.1.

De determinantafbeelding heeft een hele hoop interessante eigenschappen, waarvan we hieronder de voornaamste oplijsten. In de oefeningen kunnen deze het rekenwerk sterk vereenvoudigen.

Lemma 6.2.4. *Zij $A, B \in M_n(K)$ twee willekeurige $n \times n$-matrices. Noteer de kolommen van A als $A_1, \ldots, A_n$. Dan geldt:*

(i) $\det A^t = \det A$.

(ii) $\det(AB) = \det A \cdot \det B = \det(BA)$.

(iii) $$\det(A_1, \ldots, \lambda A_i + \mu A'_i, \ldots, A_n) = \lambda \det(A_1, \ldots, A_i, \ldots, A_n) + \mu \det(A_1, \ldots, A'_i, \ldots, A_n)$$

voor alle $i \in \{1, \ldots, n\}$*, alle* $\lambda, \mu \in K$*, en alle* $A_i, A'_i \in M_{n,1}(K)$.

(iv) $\det(A_{i_1}, \ldots, A_{i_n}) = \operatorname{sgn}(i_1, \ldots, i_n) \det(A_1, \ldots, A_n)$.

(v) *Als twee willekeurige kolommen van* A *gelijk zijn, dan is* $\det A = 0$.

(vi) $\det(A_1, \ldots, A_n) = \det(A_1, \ldots, A_i + \lambda A_k, \ldots, A_n)$ *voor alle* $\lambda \in K$ *en alle* $k \neq i$.

(vii) *We kunnen de determinant van een matrix* A *bepalen door te ontwikkelen naar de* k*-de kolom, i.e.*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik}).$$

De eigenschappen (iii)–(vi) *zijn ook geldig voor rijen in plaats van kolommen.*

Dit lemma wordt bewezen op p. 109 verderop.

6.3 Determinanten als multilineaire afbeeldingen

Deze hele paragraaf 6.3 maakt geen deel uit van de leerstof voor de studenten van de opleiding Fysica en Sterrenkunde.

Om determinanten te doorgronden, beginnen we even van vooraf aan. We zullen aantonen dat er een unieke afbeelding

$$D \colon M_n(K) \to K$$

bestaat die voldoet aan de volgende drie eigenschappen, waarbij we voor elke matrix $A \in M_n(K)$ de notatie $A = (A_1, \ldots, A_n)$ gebruiken, met $A_1, \ldots, A_n$ de kolommen van A.

($\mathbf{D_1}$) D is *multilineair* in de kolommen van de matrices, i.e.

$$D(A_1, \ldots, \lambda A_i + \mu A'_i, \ldots, A_n) = \lambda D(A_1, \ldots, A_i, \ldots, A_n) + \mu D(A_1, \ldots, A'_i, \ldots, A_n)$$

voor alle $i \in \{1, \ldots, n\}$, alle $\lambda, \mu \in K$, en alle $A_i, A'_i \in M_{n,1}(K)$;

($\mathbf{D_2}$) als de matrix $A \in M_n(K)$ twee opeenvolgende gelijke kolommen heeft, dan is $D(A) = 0$;

($\mathbf{D_3}$) $D(I_n) = 1$.

Opmerking 6.3.1. Hoewel zo dadelijk zal blijken dat D noodzakelijk de determinantafbeelding is die we in paragraaf 6.2 als "det" hebben genoteerd, gebruiken we heel bewust een andere notatie D, omdat we dat op dit ogenblik nog niet weten en een hele redenering willen opbouwen rond dergelijke afbeeldingen D.

We tonen aan dat er ten hoogste één afbeelding kan bestaan die aan de eigenschappen ($\mathbf{D_1}$)–($\mathbf{D_3}$) voldoet.

Stelling 6.3.2. *Zij $D\colon M_n(K) \to K$ een afbeelding die voldoet aan de eigenschappen* ($\mathbf{D_1}$)–($\mathbf{D_3}$).

(i) *Als $A \in M_n(K)$ twee gelijke kolommen heeft, dan is $D(A) = 0$.*

(ii) *Stel $A = (A_1, \ldots, A_n) \in M_n(K)$, en zij $\sigma = (i_1, \ldots, i_n)$ een permutatie van $\{1, \ldots, n\}$. Dan is*

$$D(A_{i_1}, \ldots, A_{i_n}) = \operatorname{sgn}(\sigma) D(A).$$

(iii) *Voor elke matrix $A \in M_n(K)$ geldt*

$$D(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}\,, \tag{6.2}$$

met andere woorden,

$$D(A) = \det(A)$$

waarbij $\det(A)$ *de determinant is die we hebben ingevoerd in Definitie 6.2.1.*

(iv) *Voor alle $A, B \in M_n(K)$ geldt dat $D(AB) = D(A)D(B)$.*

In het bijzonder toont de formule (6.2) *aan dat er ten hoogste één afbeelding D kan bestaan die aan de eigenschappen* ($\mathbf{D_1}$)–($\mathbf{D_3}$) *voldoet.*

Bewijs. We leiden stap voor stap een aantal eigenschappen af uit de eigenschappen ($\mathbf{D_1}$)–($\mathbf{D_3}$). Zij $A = (A_1, \ldots, A_n) = (a_{ij}) \in M_n(K)$.

(1) Uit ($\mathbf{D_2}$) volgt dat $D(A_1, \ldots, A_k + A_{k+1}, A_k + A_{k+1}, \ldots, A_n) = 0$; door nu ($\mathbf{D_1}$) toe te passen op het linkerlid van deze gelijkheid krijgen we

$$D(A_1, \ldots, A_k, A_{k+1}, \ldots, A_n) = -D(A_1, \ldots, A_{k+1}, A_k, \ldots, A_n).$$

Dus D verandert van teken als twee opeenvolgende kolommen verwisseld worden.

(2) Als $A_i = A_j$ voor een $i \neq j$, dan is $D(A) = 0$: dit volgt uit het voorgaande door kolommen te verwisselen tot de twee gelijke kolommen naast elkaar staan. Dit toont reeds (i) aan.

(3) Uit het voorgaande punt volgt met een analoog argument als in punt (1) dat $D(A)$ van teken verandert als we twee willekeurige kolommen verwisselen.

(4) Zij $\sigma = (i_1, \ldots, i_n)$ een permutatie van $(1, \ldots, n)$. Dan is

$$D(A_{i_1}, \ldots, A_{i_n}) = \pm D(A_1, \ldots, A_n)\,.$$

Door opeenvolgende verwisselingen van kolommen kan men $(A_{i_1}, \ldots, A_{i_n})$ omvormen in $(A_1, \ldots, A_n)$. Bij elke verwisseling verandert het teken. Deze verwisselingen zijn precies dezelfde verwisselingen als degene die we nodig hebben om $(i_1, \ldots, i_n)$ om te vormen in $(1, \ldots, n)$. Uit de definitie van $\operatorname{sgn}(\sigma)$ volgt dat dit teken gelijk is aan $\operatorname{sgn}(i_1, \ldots, i_n)$. We vinden dus dat

$$D(A_{i_1}, \ldots, A_{i_n}) = \operatorname{sgn}(\sigma) D(A) \quad \text{met } \sigma = (i_1, \ldots, i_n);$$

dit bewijst (ii).

(5) Stel $B = (b_{ij})$. Dan is $C := AB \in M_n(K)$ een matrix waarbij $C = (C_1, \ldots, C_n)$ is, met $C_k = b_{1k}A_1 + \cdots + b_{nk}A_n$. Eigenschap $(\mathbf{D_1})$ geeft dat

$$D(C_1, \ldots, C_n) = \sum_{i_1, i_2, \ldots, i_n} b_{i_1,1} b_{i_2,2} \cdots b_{i_n,n} D(A_{i_1}, \ldots, A_{i_n})\,,$$

waarbij de i_j onafhankelijk van elkaar lopen van 1 tot n. Als $i_k = i_\ell$, dan bevat de matrix $(A_{i_1}, \ldots, A_{i_n})$ twee gelijke kolommen en is dus $D(A_{i_1}, \ldots, A_{i_n}) = 0$. In de som komen dus enkel de termen voor met $\sigma = (i_1, \ldots, i_n)$ een permutatie van $\{1, \ldots, n\}$. Wanneer we (ii) toepassen, bekomen we

$$D(C_1, \ldots, C_n) = D(A_1, \ldots, A_n) \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n}\,,$$

waarbij σ loopt door de verzameling S_n van alle permutaties van $\{1, \ldots, n\}$.

(6) Nemen we in (5) nu $A = I_n$, dan volgt uit $(\mathbf{D}_3)$ dat $D(A) = 1$ en vinden we met $C = AB = B$ dat

$$D(B) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n}\,,$$

en dit bewijst de expliciete formule in (iii).

(7) Uit (5) en (6) volgt dat voor $C = AB$ geldt dat $D(C) = D(A)D(B)$, en dit bewijst (iv). $\square$

Opmerking 6.3.3. We hebben op dit moment nog niet bewezen dat de afbeelding "det" voldoet aan de eigenschappen ($\mathbf{D_1}$)–($\mathbf{D_3}$). We zouden dit nu rechtstreeks kunnen aantonen (en het is een interessante oefening om dit te doen), maar in plaats daarvan zullen we aantonen dat de afbeelding uit Constructie 6.2.3 voldoet aan ($\mathbf{D_1}$)–($\mathbf{D_3}$). Bijgevolg valt deze afbeelding samen met de afbeelding "det", waaruit dan meteen ook volgt dat "det" zelf voldoet aan ($\mathbf{D_1}$)–($\mathbf{D_3}$).

Stelling 6.3.4. *Voor iedere $k \in \{1, \dots, n\}$ is de afbeelding $D\colon M_n(K) \to K$ gedefinieerd in Constructie 6.2.3 gelijk aan de determinantafbeelding. We kunnen dus de determinant van een matrix A bepalen door te ontwikkelen naar de k-de rij, i.e.*

$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} a_{ki} \det(A_{ki}).$$

Bewijs. We bewijzen per inductie op n dat D voldoet aan ($\mathbf{D_1}$), ($\mathbf{D_2}$) en ($\mathbf{D_3}$). Het is duidelijk dat dit geldt voor $n = 1$. We veronderstellen dus dat ($\mathbf{D_1}$), ($\mathbf{D_2}$) en ($\mathbf{D_3}$) gelden voor iedere matrix in $M_{n-1}(K)$ en tonen aan dat dit ook zo is voor $M_n(K)$.

(1) We tonen aan dat $D(A)$ lineair afhangt van de ℓ-de kolom, voor elke ℓ, i.e.

$$\begin{aligned} &D(\overbrace{A_1, \dots, \lambda A_\ell + \mu A'_\ell, \dots, A_n}^{C}) \\ &\qquad = \lambda D(\underbrace{A_1, \dots, A_\ell, \dots, A_n}_{B}) + \mu D(\underbrace{A_1, \dots, A'_\ell, \dots, A_n}_{B'}). \end{aligned}$$

Stel $C = (c_{ij})$, $B = (b_{ij})$ en $B' = (b'_{ij})$. Zij nu $j \in \{1, \dots, n\}$. Als $j \neq \ell$, dan is wegens de inductiehypothese

$$D(C_{kj}) = \lambda D(B_{kj}) + \mu D(B'_{kj}),$$

terwijl $c_{kj} = b_{kj} = b'_{kj}$. Voor $j = \ell$ is $C_{k\ell} = B_{k\ell} = B'_{k\ell}$, terwijl de coëfficiënt $c_{k\ell}$ gelijk is aan

$$c_{k\ell} = \lambda b_{k\ell} + \mu b'_{k\ell}.$$

Bijgevolg is

$$c_{kj} D(C_{kj}) = \lambda b_{kj} D(B_{kj}) + \mu b'_{kj} D(B'_{kj})$$

voor elke $j \in \{1, \dots, n\}$; uit de formule (6.1) volgt nu dat de afbeelding D lineair is in de kolommen.

(2) Onderstel dat twee opeenvolgende kolommen van A gelijk zijn, $A_\ell = A_{\ell+1}$. Voor $j \neq \ell, \ell+1$ hebben de matrices A_{kj} twee gelijke kolommen, dus is, wegens de inductiehypothese, $D(A_{kj})$ gelijk aan 0. De matrices $A_{k,\ell}$ en $A_{k,\ell+1}$ zijn gelijk, dus zijn ook $D(A_{k,\ell}) = D(A_{k,\ell+1})$. In de ontwikkeling naar de k-de rij komen $D(A_{k,\ell})$ en $D(A_{k,\ell+1})$ voor met tegengesteld teken. Deze feiten samen tonen aan dat $D(A) = 0$.

(3) Als $A = I_n$ dan is $a_{kj} = 0$ voor alle $j \neq k$, dus $D(A) = a_{kk}D(A_{kk})$. We hebben $a_{kk} = 1$, en wegens de inductiehypothese is ook $D(A_{kk}) = D(I_{n-1}) = 1$, dus $D(A) = 1$.

Dit toont aan dat D voldoet aan $(\mathbf{D_1})$, $(\mathbf{D_2})$ en $(\mathbf{D_3})$. Uit Stelling 6.3.2(iii) volgt nu dat $D(A) = \det(A)$ voor alle $A \in M_n(K)$. □

De combinatie van Stellingen 6.3.2 en 6.3.4 is bijzonder krachtig. Samen bewijzen ze het bestaan én de uniciteit van een afbeelding die voldoet aan $(\mathbf{D_1})$–$(\mathbf{D_3})$, maar bovendien hebben we in dit proces heel wat eigenschappen afgeleid. Dat zijn precies de eigenschappen uit Lemma 6.2.4. We geven hier het bewijs mee van deze eigenschappen, waarvan de meeste zonder enige moeite volgen uit wat we al hebben bewezen. Enkel het eerste deeltje, met name de uitspraak $\det A = \det A^t$, vereist nog een kort argument.

Bewijs van Lemma 6.2.4. Stel $C = (c_{ij}) = A^t$; dan is

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) c_{1,\sigma(1)} \cdots c_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) c_{\sigma^{-1}(1),1} \cdots c_{\sigma^{-1}(n),n} = \sum_{\tau \in S_n} \operatorname{sgn}(\tau^{-1}) c_{\tau(1),1} \cdots c_{\tau(n),n} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) c_{\tau(1),1} \cdots c_{\tau(n),n} = \det A^t\,,
\end{aligned}$$

waarbij we gebruik gemaakt hebben van Opmerking 6.1.2(ii). Dit bewijst (i). Eigenschap (iii) is precies eigenschap $(\mathbf{D_1})$. Eigenschappen (ii), (iv) en (v) zijn reeds vermeld in Stelling 6.3.2. Eigenschap (vi) volgt uit eigenschappen (iii) en (v). Uit het feit dat $\det A = \det A^t$ volgt ten slotte dat deze eigenschappen ook geldig zijn wanneer men rijen in plaats van kolommen beschouwt. Op deze wijze volgt eigenschap (vii) ook uit Stelling 6.3.4. □

6.4 Inverteerbaarheid van matrices

Voor de studenten van de opleiding Fysica en Sterrenkunde zijn de bewijzen *in deze paragraaf 6.3 niet te kennen. Het is wel belangrijk om de definities* *en de stellingen zelf te beheersen.*

Met behulp van de determinant kunnen we enkele karakterisaties van inverteerbare matrices bewijzen.

Stelling 6.4.1. *Zij $A \in M_n(K)$; dan zijn de volgende eigenschappen equivalent:*

(a) $\det(A) \neq 0$;

(b) $\mathrm{rk}(A) = n$;

(c) $A \in \mathsf{GL}_n(K)$.

Bewijs. We zullen aantonen dat (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).

Zij $A = (A_1, \ldots, A_n)$, we tonen (a) $\Rightarrow$ (b) aan uit het ongerijmde. Als $\mathrm{rk}(A) \neq n$, dan zijn de kolommen van A lineair afhankelijk. Er is dus een eerste kolom, stel A_k, die een lineaire combinatie is van de voorgaande kolommen. Dus $A_k = \lambda_1 A_1 + \cdots + \lambda_{k-1} A_{k-1}$. Vervangen we A_k door deze lineaire combinatie, dan vinden we

$$\begin{aligned} \det A = \lambda_1 \det(A_1, \ldots, A_{k-1}, A_1, A_{k+1}, \ldots, A_n) + \cdots \\ + \lambda_{k-1} \det(A_1, \ldots, A_{k-1}, A_{k-1}, A_{k+1}, \ldots, A_n) = 0\,; \end{aligned}$$

de matrices in elke term hebben immers twee gelijke kolommen. Uit (a) volgt dus (b).

Als $\mathrm{rk}(A) = n$, dan zijn de kolommen van A lineair onafhankelijk en vormen ze een basis voor de kolommenruimte K^n. De standaardbasisvectoren $e_1, \ldots e_n$ zijn dus lineaire combinaties van de kolommen van A. Dus

$$I_n = \left(\sum b_{i1} A_i, \ldots, \sum b_{in} A_i\right) = AB\,.$$

Ook is $\mathrm{rk}(A^t) = n$, zodat om dezelfde reden een matrix C bestaat zodat $I_n = A^t C$, of dus $I_n = C^t A$. Hieruit volgt dan dat

$$C^t = C^t I_n = C^t(AB) = (C^t A)B = I_n B = B,$$

en dus is zowel $AB = I_n$ als $BA = I_n$. We besluiten dat A inverteerbaar is (met inverse B); uit (b) volgt dus (c).

Veronderstel nu dat (c) geldt; dan is

$$1 = \det I_n = \det(AA^{-1}) = \det(A) \cdot \det(A^{-1})\,,$$

waaruit volgt dat $\det A \neq 0$. Dus (c) impliceert (a). □

We vinden nu ook een expliciete formule om het inverse van een matrix te berekenen. We voeren daartoe de volgende terminologie in.

Definitie 6.4.2. (i) De (i, j)-de *minor* van A is gedefinieerd als $\det(A_{ij})$, waar $A_{ij} \in M_{n-1}(K)$ de matrix is bekomen door in A de i-de rij en de j-de kolom te schrappen.

(ii) De (i, j)-de minor met zijn teken $\gamma_{ij} = (-1)^{i+j} \det A_{ij}$ noemt men de *cofactor* van het element a_{ij} uit A.

(iii) We definiëren de *adjunct* van een matrix $A \in M_n(K)$ als

$$\operatorname{adj}(A) := \begin{pmatrix} \gamma_{11} & \gamma_{21} & \cdots & \gamma_{n1} \\ \gamma_{12} & \gamma_{22} & \cdots & \gamma_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{1n} & \gamma_{2n} & \cdots & \gamma_{nn} \end{pmatrix},$$

of dus $\operatorname{adj}(A) = (\gamma_{ij})^t$. (Bemerk het transponeren.)

Stelling 6.4.3. (i) *Voor alle* $A \in M_n(K)$ *is* $A \operatorname{adj}(A) = \operatorname{adj}(A) A = (\det A) I_n$.

(ii) *Voor alle* $A \in \mathsf{GL}_n(K)$ *is* $A^{-1} = (\det A)^{-1} \operatorname{adj}(A)$.

Bewijs. (i) Het i-de diagonaalelement van $A \operatorname{adj}(A)$ is gelijk aan

$$\begin{aligned} \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} \gamma_{i1} \\ \gamma_{i2} \\ \vdots \\ \gamma_{in} \end{pmatrix} \\ &= a_{i1}\gamma_{i1} + \cdots + a_{in}\gamma_{in} \\ &= (-1)^{i+1} a_{i1} \det A_{i1} + \cdots + (-1)^{i+n} a_{in} \det A_{in}\,; \end{aligned}$$

dit is de ontwikkeling van A naar de i-de rij en is dus gelijk aan $\det A$. Analoog vinden we dat het i-de diagonaal element van $\operatorname{adj}(A)A$ gelijk is aan de ontwikkeling van A naar de i-de kolom, dus ook gelijk is aan $\det A$.

We bekijken vervolgens de niet-diagonaal elementen van $A \operatorname{adj}(A)$ en $\operatorname{adj}(A)A$. Stel dus $j \neq k$, en beschouw het (j, k)-de element van $A \operatorname{adj}(A)$. Definieer de matrix $\tilde{A}$ bekomen uit A door de k-de rij van A te vervangen door de j-de rij van A. De matrix $\tilde{A}$ heeft twee gelijke rijen en $\det \tilde{A}$ is dus gelijk aan nul. Dit impliceert dat de ontwikkeling naar de k-de rij nul oplevert, dus

$$\tilde{a}_{k1}\tilde{\gamma}_{k1} + \cdots + \tilde{a}_{kn}\tilde{\gamma}_{kn} = 0\,.$$

Nu is $\tilde{a}_{ki} = a_{ji}$ en $\tilde{\gamma}_{ki} = \gamma_{ki}$ voor alle $i = 1, \dots, n$, en dus

$$(A \operatorname{adj}(A))_{jk} = a_{j1}\gamma_{k1} + \cdots + a_{jn}\gamma_{kn} = 0\,.$$

Analoog toont men aan dat $(\operatorname{adj}(A)A)_{jk} = 0$ als $k \neq j$.

(ii) Als $A \in \mathsf{GL}_n(K)$, dan impliceert Stelling 6.4.1 dat $\det A \neq 0$, en uit (i) vinden we dat

$$A^{-1} = (\det A)^{-1} \operatorname{adj}(A)\,.$$

□

We geven nog een ander gevolg mee van Stelling 6.4.1.

Definitie 6.4.4. Zij $A \in \mathrm{Mat}_{m,n}(K)$, en zij $k \leq \min\{m, n\}$. Een *$k \times k$-minor* van A is de determinant van een $k \times k$-deelmatrix die we verkrijgen door het schrappen van $m - k$ rijen en $n - k$ kolommen uit A.

Gevolg 6.4.5. *Zij $A \in \mathrm{Mat}_{m,n}(K)$, en $k \leq \min\{m, n\}$. Dan is $\mathrm{rk}(A) \geq k$ als en slechts als er een $k \times k$-minor van A is die verschillend is van 0.*

Bewijs. Veronderstel eerst dat $\mathrm{rk}(A) < k$. Dan vormen elke k rijen van A een lineair afhankelijke verzameling. In het bijzonder heeft elke $k \times k$-deelmatrix k afhankelijke rijen, en dus heeft een dergelijke deelmatrix rang kleiner dan k. Uit Stelling 6.4.1 volgt nu onmiddellijk dat elke $k \times k$-minor gelijk is aan 0.

Veronderstel nu dat $\mathrm{rk}(A) \geq k$. Kies k rijen $R_1, \dots, R_k$ uit A zodat de $k \times n$-matrix B gevormd door de rijen $R_1, \dots, R_k$ rang k heeft. Omdat B nu ook kolomrang k heeft (zie Stelling 5.1.2), kunnen we in B nu ook k kolommen $C_1, \dots, C_k$ vinden zodat de matrix C gevormd door de kolommen $C_1, \dots, C_k$ rang k heeft. Nu is C een $k \times k$-matrix met rang k, en opnieuw uit Stelling 6.4.1 besluiten we dat $\det C$ een $k \times k$-minor is verschillend van 0. □

Gevolg 6.4.5 wordt vaak toegepast voor $k = 2$, omdat 2×2-minoren natuurlijk heel eenvoudig te berekenen zijn.

Hoofdstuk

7 Lineaire operatoren

De doelstelling van dit hoofdstuk is om voor een gegeven lineaire operator een "goede" matrixvoorstelling te vinden die ons in staat stelt om deze operator beter te begrijpen. Dit is zowel voor theoretische als voor praktische doeleinden zeer belangrijk.

Voor de studenten van de opleiding Wiskunde gaan we nadien dieper in op de deelring $K[f]$ van $\operatorname{End}_K(V)$ voor een vaste lineaire operator $f \in \operatorname{End}_K(V)$ en bewijzen we de beroemde stelling van Cayley–Hamilton.

7.1 De karakteristieke veelterm van een lineaire operator

We willen onze voorgaande studie van determinanten van matrices nu toepassen op de studie van lineaire operatoren. Vooreerst zullen we aan elke lineaire operator een bijzondere veelterm kunnen associëren, die we de *karakteristieke veelterm* noemen.

Een belangrijke vaststelling vooraf is dat het zinvol is om te spreken van de determinant van een lineaire operator, zoals we nu nagaan.

Lemma 7.1.1. *Zij V een eindig-dimensionale K-vectorruimte en $f \in \operatorname{End}(V)$. Beschouw twee matrixvoorstellingen A en A' van f ten opzichte van verschillende basissen van V. Dan is $\det A = \det A'$.*

Bewijs. Uit Stelling 4.2.5 weten we dat A' toegevoegd is aan A, m.a.w. er bestaat een $P \in \mathsf{GL}_n(K)$ zodat $A' = P^{-1}AP$. Uit Lemma 6.2.4(ii) volgt dan dat

$$\begin{aligned}\det A' &= \det(P^{-1}AP) = \det P^{-1} \cdot \det A \cdot \det P \\ &= \det A \cdot \det P^{-1} \cdot \det P = \det A \cdot \det(P^{-1}P) = \det A\,. \qquad \square\end{aligned}$$

Bijgevolg kunnen we de *determinant van een lineaire operator* als volgt definiëren:

Definitie 7.1.2. Zij V een eindig-dimensionale K-vectorruimte, en zij $f \in \mathrm{End}(V)$. Dan definiëren we $\det f := \det A$ waarbij A een matrixvoorstelling van f t.o.v. een willekeurige basis is. Omwille van Lemma 7.1.1 is dit *goed gedefinieerd*.

Opmerking 7.1.3. In wat volgt, zullen we matrices nodig hebben die gedefinieerd zijn over $K[x]$ in plaats van over een veld K. De elementen in een dergelijke matrix zijn dus zelf veeltermen. We kunnen deze matrices echter op dezelfde manier manipuleren als we gewoon zijn, zolang we maar geen inversen nodig hebben.

Zo is de determinant van een matrix $A \in M_n(K[x])$ nog steeds gegeven door Definitie 6.2.1, en is de formule $A \cdot \mathrm{adj}(A) = (\det A) I_n$ nog steeds geldig. Daarentegen is het *niet* langer waar dat $A \in M_n(K[x])$ inverteerbaar is als en slechts als $\det A \neq 0$.

We komen nu tot het belangrijke begrip van de karakteristieke veelterm van een lineaire operator.

Lemma 7.1.4. *Zij V een eindig-dimensionale vectorruimte over een veld K, en zij $f \in \mathrm{End}_K(V)$ een lineaire operator, met matrixvoorstelling A (ten opzichte van een willekeurige basis voor V). Dan is*

$$\chi_f(x) := \det(xI_n - A) \in K[x]$$

een monische veelterm van graad n in de variabele x, en deze veelterm hangt niet af van de keuze van de basis voor V.

Bewijs. Zij A een matrixvoorstelling van f; dan is $xI_n - A$ een $n \times n$-matrix over $K[x]$. Merk op dat de variabele x enkel voorkomt op de diagonaal van de matrix $xI_n - A$. Stellen we nu $B = xI_n - A$ in de formule (6.2), dan zien we dat in elke term van deze formule één element van elke rij en elke kolom van B voorkomt, waarbij elke term van deze formule een element is van $K[x]$. Vermits x enkel voorkomt in de elementen op de diagonaal van B, is er maar één term van graad n, namelijk

$$b_{11} \cdots b_{nn} = (x - a_{11}) \cdots (x - a_{nn}) .$$

Hieruit volgt dat $\chi_f(x)$ een veelterm van graad n is en dat deze veelterm monisch is.

Zij nu A' een andere matrixvoorstelling voor f; dan is $A' = P^{-1}AP$ met $P \in \mathsf{GL}_n(K)$, en dus

$$\begin{aligned} \det(xI_n - A) &= \det(P^{-1}(xI_n - A)P) \\ &= \det(xI_n - P^{-1}AP) = \det(xI_n - A') . \end{aligned}$$ □

Definitie 7.1.5. (i) Zij $f\colon V \to V$ een lineaire operator op een eindig-dimensionale vectorruimte V over K. De veelterm $\chi_f(x)$ gedefinieerd in Lemma 7.1.4 noemt men de *karakteristieke veelterm* van f.

(ii) Zij $A \in M_n(K)$. De veelterm $\chi_A(x) := \det(xI_n - A) \in K[x]$ noemt men de *karakteristieke veelterm* van de matrix A. De karakteristieke veelterm van de matrix A is dus gelijk aan de karakteristieke veelterm van de lineaire operator $L_A\colon K^n \to K^n$.

(iii) Als $\chi(x)$ de karakteristieke veelterm is van een lineaire operator (resp. van een matrix), dan noemen we de vergelijking $\chi(x) = 0$ de *karakteristieke vergelijking* van de lineaire operator (resp. van de matrix).

Lemma 7.1.6. *Zij $A = (a_{ij}) \in M_n(K)$ met karakteristieke veelterm*

$$\chi_A(x) = x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1x + c_0.$$

Dan is $c_0 = (-1)^n \det A$ en $c_{n-1} = -\operatorname{tr} A$.

Bewijs. Door de gelijkheid

$$\chi_A(x) = \det(xI_n - A) = x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1x + c_0$$

te beschouwen voor $x = 0$ bekomen we $c_0 = \det(-A) = (-1)^n \det A$.

Om in te zien dat $c_{n-1} = -\sum_{i=1}^n a_{ii}$ stellen we opnieuw $B = xI_n - A$ in de formule (6.2). Vermits x enkel voorkomt in de elementen op de diagonaal van B, is er maar één term van graad $\geq n-1$, namelijk

$$b_{11} \cdots b_{nn} = (x - a_{11}) \cdots (x - a_{nn})\,;$$

hieruit volgt dat c_{n-1} gelijk is aan de coëfficiënt bij x^{n-1} in deze uitdrukking, en dus $c_{n-1} = -(a_{11} + \cdots + a_{nn}) = -\operatorname{tr} A$. □

Opmerking 7.1.7. Uit Lemmas 7.1.4 en 7.1.6 volgt dat het spoor van geconjugeerde matrices gelijk is. Dit volgde in feite ook reeds uit Lemma 1.3.16, want $\operatorname{tr}(P^{-1}AP) = \operatorname{tr}(P^{-1} \cdot AP) = \operatorname{tr}(AP \cdot P^{-1}) = \operatorname{tr}(A)$.

7.2 Eigenwaarden en eigenvectoren

Een essentieel begrip in de theorie van de lineaire operatoren is het begrip van een eigenvector.

Definitie 7.2.1. Zij V een K-vectorruimte en $f\colon V \to V$ een lineaire operator.

(i) Een niet-nul element $0 \neq v \in V$ is een *eigenvector* van f als $f(v) = \lambda v$ voor een $\lambda \in K$.

(ii) Zij $0 \neq v \in V$ een eigenvector van f met $f(v) = \lambda v$. Dan is λ een *eigenwaarde* van f.

(iii) Zij λ een eigenwaarde van f. Dan is $\ker(\lambda \mathbf{1}_V - f)$ de *eigenruimte* van f bij de eigenwaarde λ.

We beschouwen voor elke $A \in M_n(K)$ de operator $L_A\colon K^n \to K^n$. Als we bovenstaande definitie toepassen op deze operator, bekomen we de volgende definitie voor de eigenwaarden en eigenvectoren van een matrix.

Definitie 7.2.2. Zij $A \in M_n(K)$ een $n \times n$-matrix over K.

(i) Een niet-nul element $0 \neq v \in K^n$ is een (rechtse[1]) *eigenvector* van A als $Av = \lambda v$ voor een $\lambda \in K$.

(ii) Zij $0 \neq v \in K^n$ een eigenvector van A met $Av = \lambda v$. Dan is λ een *eigenwaarde* van A.

(iii) Zij λ een eigenwaarde van A. Dan is

$$E_\lambda := \ker L_{\lambda I_n - A} = \{w \in K^n \mid (\lambda I_n - A)w = 0\}$$

de *eigenruimte* van A bij de eigenwaarde λ.

Opmerking 7.2.3. Stel dat λ een eigenwaarde van f is. Dan is de eigenruimte van f bij λ de unie van de verzameling van alle eigenvectoren bijhorende bij λ en de nulvector $0 \in V$ (die per definitie geen eigenvector is). We gaan dit gemakkelijk na:

$$\begin{aligned} E_\lambda = \ker(\lambda \mathbf{1}_V - f) &= \{v \in V \mid (\lambda \mathbf{1}_V - f)v = 0\} \\ &= \{v \in V \mid f(v) = \lambda v\} \\ &= \{0\} \cup \{v \in V \setminus \{0\} \mid f(v) = \lambda v\}. \end{aligned}$$

Voorbeelden 7.2.4. (1) Zij $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in M_n(K)$. De elementen van de standaardbasis $e_1, \ldots, e_n$ van K^n zijn eigenvectoren van A met respectieve eigenwaarden $\lambda_1, \ldots, \lambda_n$. (Zie ook Stelling 7.3.2 verderop.)

(2) Op de eindig-dimensionale ruimte K^n voor $n \in \mathbb{N}$ heeft de shiftoperator

$$S\colon K^n \to K^n\colon (a_1, \ldots, a_n)^t \mapsto (0, a_1, \ldots, a_{n-1})^t$$

eigenvectoren. De vectoren $(0, \ldots, 0, a_n)^t$ zijn eigenvectoren met bijhorende eigenwaarde 0 als $a_n \neq 0$.

[1]We kunnen ook *linkse* eigenvectoren definiëren door A te beschouwen als een lineaire operator op de *rijruimte* K^n via rechtse vermenigvuldiging; de linkse eigenvectoren van A zijn dus de (getransponeerde van) de rechtse eigenvectoren van A^t.

(3) We beschouwen nu de shiftoperator op de oneindig-dimensionale vectorruimte $K^{\mathbb{N}} = \prod_{i\in\mathbb{N}} K$ van oneindige rijen van elementen van K. De *rechtse shiftoperator*

$$S\colon K^{\mathbb{N}} \to K^{\mathbb{N}}\colon (a_0,\ldots,a_n,\ldots)^t \mapsto (0,a_0,a_1,\ldots,a_n,\ldots)^t$$

heeft geen eigenvectoren. Immers, als voor een $v \in K^{\mathbb{N}}$

$$S(v) = (0,a_0,\ldots,a_n,\ldots)^t = \lambda(a_0,\ldots,a_n,\ldots)^t$$

voor een zekere $\lambda \in K$, dan is $\lambda a_0 = 0$. Als $\lambda = 0$ volgt onmiddellijk dat $a_i = 0$ voor alle i; als $\lambda \neq 0$ is $a_0 = 0$, maar dan volgt uit $\lambda a_i = a_{i-1}$ voor alle $i \geq 1$ opnieuw dat alle $a_i = 0$. We concluderen dat uit $S(v) = \lambda v$ voor een $\lambda \in K$ volgt dat $v = 0$. Bijgevolg heeft S geen eigenvectoren. Beschouwen we daarentegen de *linkse shiftoperator*

$$T\colon K^{\mathbb{N}} \to K^{\mathbb{N}}\colon (a_0,\ldots,a_n,\ldots)^t \mapsto (a_1,\ldots,a_n,\ldots)^t,$$

dan is elk element van K een eigenwaarde voor T! Inderdaad, voor elke $\lambda \in K$ en elke $a \in K$ is de oneindige rij

$$v = (a, \lambda a, \lambda^2 a, \lambda^3 a, \ldots)$$

een eigenvector voor T met eigenwaarde λ. Als K een oneindig veld is, vinden we dus oneindig veel verschillende eigenwaarden voor deze lineaire operator T.

We zullen ons vanaf nu beperken tot eindig-dimensionale vectorruimten. We bespreken een methode om voor een lineaire operator f op een eindig-dimensionale vectorruimte V (of matrix $A \in M_n(K)$) alle eigenwaarden en eigenvectoren te bepalen.

In Definitie 7.1.5 hebben we de karakteristieke veelterm van een lineaire operator f en van een matrix ingevoerd. We bewijzen nu een zeer handig criterium om te bepalen welke elementen van K eigenwaarden bijhorend bij f zijn.

Stelling 7.2.5. (i) *Zij $A \in M_n(K)$. Een scalair $\lambda \in K$ is een eigenwaarde van A als en slechts als $\chi_A(\lambda) = \det(\lambda I_n - A) = 0$.*

(ii) *Zij V een n-dimensionale K-vectorruimte en $f\colon V \to V$ een lineaire operator. Een scalair $\lambda \in K$ is een eigenwaarde van f als en slechts als $\chi_f(\lambda) = \det(\lambda\,\mathbf{1} - f) = 0$.*

Bewijs. (i) Een element $\lambda \in K$ is een eigenwaarde van A als en slechts als er een $v \in K^n \setminus \{0\}$ bestaat waarvoor $(\lambda I_n - A)v = 0 \in K^n$, of nog, als en slechts als het homogeen $n \times n$-stelsel $(\lambda I_n - A)X = 0$ een niet-nul oplossing heeft, waarbij X een kolomvector is met n onbekenden.
Uit Stelling 5.3.2(ii) volgt dat $(\lambda I_n - A)X = 0$ een niet-nul oplossing heeft als en slechts als $\mathrm{rk}(\lambda I_n - A) < n$. Wegens Stelling 6.4.1 is dit op zijn beurt equivalent met $\det(\lambda I_n - A) = 0$.

(ii) Zij $\mathcal{B}$ een willekeurige basis van V en zij A de matrixvoorstelling van f t.o.v. $\mathcal{B}$. Uit Stelling 4.1.4(ii) volgt dat $v \in V$ een eigenvector is voor f met eigenwaarde λ als en slechts als $\mathrm{co}_{\mathcal{B}}(v) \in K^n$ een eigenvector is voor A met eigenwaarde λ. □

Gevolg 7.2.6. *Een lineaire operator f op een n-dimensionale K-vectorruimte V (of een matrix $A \in M_n(K)$) heeft hoogstens n eigenwaarden.*

Bewijs. De veelterm $\chi_f(x)$ (resp. $\chi_A(x)$), die van graad n is, kan hoogstens n wortels hebben in K. □

Constructie 7.2.7. Met behulp van de vorige stelling kunnen we nu een methode opstellen om alle eigenwaarden met bijhorende eigenvectoren van een matrix $A \in M_n(K)$ te bepalen:

(1) Bepaal de karakteristieke veelterm $\chi_A(x)$.

(2) Los de vergelijking $\chi_A(x) = 0$ op met $x \in K$. Er zijn maar eindig veel oplossingen $\lambda_1, \ldots, \lambda_s$ van deze vergelijking in K, dit zijn de eigenwaarden van A in K.

(3) Bepaal voor elke λ_i, $i = 1, \ldots, s$, de eigenruimte

$$E_{\lambda_i} := \{w \in K^n \mid (\lambda_i I_n - A)w = 0 \in K^n\}.$$

Opmerking 7.2.8. De eerste stap van Constructie 7.2.7 kan men uitvoeren door de determinant van een matrix te berekenen; dit kan men steeds (met de nodige rekentijd) doen door bijvoorbeeld te ontwikkelen naar een geschikte rij of kolom. De derde stap komt neer op het oplossen van een lineair stelsel; dit kan men steeds (met de nodige rekentijd) doen door bijvoorbeeld rijreductie naar de echelonvorm uit te voeren.

De enige stap waarvoor we geen methode besproken hebben is de tweede stap, namelijk het bepalen van de wortels van een veelterm in één variabele van graad n. Zolang $n \leq 4$ is het steeds mogelijk om alle oplossingen van de karakteristieke veelterm te berekenen. Voor hogere graad zijn er niet altijd algemene oplossingsmethoden bekend; het zoeken naar oplossingsmethoden voor veeltermen van hogere graad heeft tot de ontwikkeling van heel wat algebraïsche theorieën geleid. Zo kan men bijvoorbeeld *bewijzen* dat er (in zekere zin) geen algemene oplossingsmethode kan bestaan voor het bepalen van de oplossingen van veeltermvergelijkingen van graad 5 of hoger. De bespreking van dit probleem (over een willekeurig veld) is voor de wiskundigen onderwerp van de cursus “Algebra II”.

7.3 Diagonaliseren van operatoren

We passen nu de theorie van de eigenwaarden en eigenvectoren toe om een gegeven lineaire operator f in een zo eenvoudig mogelijke gedaante te brengen. In het ideale geval kunnen we een basis vinden zodat de matrix van f een diagonaalmatrix is; we noemen de operator in dat geval *diagonaliseerbaar*.

Het diagonaliseren van matrices is in concrete toepassingen van uitermate groot belang. Een operator wordt vaak gegeven als een welbepaalde matrix, bijvoorbeeld bekomen door het uitvoeren van experimenten, en het diagonaliseren ervan (indien dit mogelijk is) zet de operator om in een zeer begrijpelijke en interpreteerbare gedaante: het geeft voor elke basisvector een “expansie- of contractiefactor” weer, zodat het gemakkelijk te visualiseren of interpreteren valt wat het effect is van het toepassen van de operator.

Definitie 7.3.1. (i) Een operator f op een eindig-dimensionale vectorruimte V is *diagonaliseerbaar* als er een matrixvoorstelling van f is die een diagonaalmatrix is.

(ii) Een matrix $A \in M_n(K)$ is *diagonaliseerbaar* als de lineaire operator $L_A \colon K^n \to K^n$ diagonaliseerbaar is.

De volgende eenvoudige observatie legt een cruciaal verband tussen diagonaliseerbaarheid en eigenvectoren.

Stelling 7.3.2. *Een lineaire operator f op een eindig-dimensionale K-vectorruimte is diagonaliseerbaar als en slechts als V een basis heeft bestaande uit eigenvectoren voor f.*

Bewijs. Stel dat f ten opzichte van de basis $\mathcal{B} = \{b_1, \dots, b_n\}$ als matrixvoorstelling een diagonaalmatrix $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ heeft; dan is, voor alle $1 \leq i \leq n$,

$$f(b_i) = 0b_1 + \cdots + 0b_{i-1} + \lambda_i b_i + 0b_{i+1} + \cdots + 0b_n = \lambda_i b_i.$$

Dus $\mathcal{B}$ is een basis van V bestaande uit eigenvectoren van f.

Zij omgekeerd $\mathcal{B} = \{b_1, \dots, b_n\}$ een basis van eigenvectoren voor de operator $f \in \operatorname{End}(V)$. Uit $f(b_i) = \lambda_i b_i$ volgt dat de matrixvoorstelling van f ten opzichte van $\mathcal{B}$ een diagonaalmatrix is. $\square$

Omwille van Definitie 7.3.1 en Opmerking 4.2.6 is een matrix $A \in M_n(K)$ diagonaliseerbaar als en slechts als er een matrix $P \in \mathsf{GL}_n(K)$ bestaat zodat $P^{-1}AP$ een diagonaalmatrix is. Het is interessant om op te merken dat we deze berekening niet expliciet hoeven uit te voeren als we al beschikken over een basis van eigenvectoren:

Lemma 7.3.3. *Onderstel dat $A \in M_n(K)$ diagonaliseerbaar is, en zij $\mathcal{B} = \{b_1, \dots, b_n\}$ een bijhorende basis van K^n van eigenvectoren voor A, dus $Ab_i = \lambda_i b_i$ voor elke i. Stel dan $P := \big(b_1, \dots, b_n\big) \in M_n(K)$ en $D := \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in M_n(K)$.*

Dan is $D = P^{-1}AP$.

Bewijs. Uit Opmerking 4.2.2(ii) weten we dat P precies de transitiematrix is van de standaardbasis van K^n naar de basis $\mathcal{B}$. De matrixvoorstelling van L_A ten opzichte van de standaardbasis is uiteraard de matrix A zelf. Anderzijds is de matrixvoorstelling van L_A ten opzichte van $\mathcal{B}$ precies gelijk aan de matrix D, want $Ab_i = \lambda_i b_i$ voor elke i. Het resultaat volgt nu uit Opmerking 4.2.6. □

Stelling 7.3.4. *Zij V een n-dimensionale vectorruimte, en zij $f \in \operatorname{End}(V)$.*

- (i) *De eigenvectoren van f die horen bij verschillende eigenwaarden, zijn lineair onafhankelijk van elkaar.*
- (ii) *Als f precies n verschillende eigenwaarden heeft, dan is f diagonaliseerbaar.*
- (iii) *Als f diagonaliseerbaar is, dan is V de directe som van de eigenruimten van f horende bij de verschillende eigenwaarden van f.*

Bewijs. (i) Zij $\{v_1, \dots, v_k\}$ een verzameling van eigenvectoren van een operator f met $f(v_i) = \lambda_i v_i$ waarbij alle λ_i verschillend zijn. We redeneren per inductie op k. Voor $k = 1$ is de uitspraak evident, want de eigenvector v_1 is niet gelijk aan 0 en dus is $\{v_1\}$ een lineair onafhankelijke verzameling.

Onderstel nu dat we al weten dat $\{v_1, \dots, v_{k-1}\}$ lineair onafhankelijk is, en onderstel dat $a_1, \dots, a_k \in K$ scalairen zijn zodat

$$a_1 v_1 + \cdots + a_k v_k = 0\,. \tag{7.1}$$

We nemen het beeld van het linkerlid onder f, en we bekomen

$$f(a_1 v_1 + \cdots + a_k v_k) = a_1 \lambda_1 v_1 + \cdots + a_k \lambda_k v_k = 0\,. \tag{7.2}$$

Vermenigvuldig vergelijking (7.1) met λ_k en trek deze af van vergelijking (7.2); dan is

$$a_1(\lambda_1 - \lambda_k)v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0\,.$$

Aangezien $\{v_1, \ldots, v_{k-1}\}$ lineair onafhankelijk is, kan dit enkel maar als $a_i(\lambda_i - \lambda_k) = 0$ voor elke $i < k$. Omdat $\lambda_i \neq \lambda_k$ voor elke i per onderstelling, volgt hieruit dat $a_i = 0$ voor elke $i < k$, en dan is ook de overblijvende a_k gelijk aan 0. We besluiten dat ook $\{v_1, \ldots, v_k\}$ lineair onafhankelijk is.

(ii) Aangezien er n verschillende eigenwaarden zijn voor f, volgt er uit (i) dat er n lineair onafhankelijke eigenvectoren zijn voor f. Een lineair onafhankelijke verzameling met n elementen is een basis voor V. Uit Stelling 7.3.2 volgt nu het gestelde.

(iii) Veronderstel dat $\lambda_1, \ldots, \lambda_\ell$ de verschillende eigenwaarden van f zijn. Aangezien f diagonaliseerbaar is, is er een basis $\mathcal{B}$ van V bestaande uit eigenvectoren. Stel nu, voor elke $i \in \{1, \ldots, \ell\}$,

$$\mathcal{B}_i := \{v \in \mathcal{B} \mid f(v) = \lambda_i v\}.$$

Uiteraard is $\mathcal{B}$ dan de disjuncte unie van $\mathcal{B}_1, \ldots, \mathcal{B}_\ell$. Stel nu $V_i :=$ span$(\mathcal{B}_i)$ voor elke i. Uit Stelling 2.5.6 volgt dan dat

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_\ell\,. \tag{7.3}$$

Per definitie is elke V_i bevat in de λ_i-eigenruimte van f, en omgekeerd volgt uit (i) en (7.3) dat elke eigenvector horende bij de eigenwaarde λ_i bevat is in V_i. Dus elke V_i is precies gelijk aan de eigenruimte horende bij de eigenwaarde λ_i, en uit (7.3) volgt nu het gestelde. □

Niet alle lineaire operatoren zijn diagonaliseerbaar. We werken twee voorbeelden uit van matrices over $\mathbb{R}$ die niet diagonaliseerbaar zijn, en een voorbeeld van een matrix die wel diagonaliseerbaar is:

Voorbeelden 7.3.5. (1) Beschouw de matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{R}).$$

De karakteristieke vergelijking van A is

$$\chi_A(x) = (x-1)(x^2-1) = (x-1)^2(x+1);$$

deze heeft twee oplossingen, namelijk 1 en -1, waarbij de oplossing 1 multipliciteit 2 heeft. We hebben dus twee verschillende eigenwaarden $\lambda_1 = 1$ en $\lambda_2 = -1$. De eigenruimte bij $\lambda_1 = 1$ is gelijk aan

$$E_1 = \{(r, s, s)^t \mid r, s \in \mathbb{R}\},$$

de eigenruimte bij $\lambda_2 = -1$ is gelijk aan

$$E_{-1} = \{(0, r, -r)^t \mid r \in \mathbb{R}\}.$$

We vinden dus dat

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

een basis van $\mathbb{R}^3$ is die bestaat uit eigenvectoren. Bijgevolg is A diagonaliseerbaar; de matrixvoorstelling van $L_A \colon \mathbb{R}^3 \to \mathbb{R}^3 \colon v \mapsto Av$ t.o.v. $\mathcal{B}$ is de diagonaalmatrix $D = \operatorname{diag}(1, 1, -1)$.

(2) Beschouw de matrix

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \in M_3(\mathbb{R}).$$

De karakteristieke vergelijking van B is

$$\chi_B(x) = (x-1)(x+1)^2;$$

deze heeft twee oplossingen, namelijk 1 en -1, waarbij de oplossing -1 multipliciteit 2 heeft. We hebben dus twee verschillende eigenwaarden $\lambda_1 = 1$ en $\lambda_2 = -1$. De eigenruimte bij $\lambda_1 = 1$ is gelijk aan

$$E_1 = \{(r, 0, 0)^t \mid r \in \mathbb{R}\},$$

de eigenruimte bij $\lambda_2 = -1$ is gelijk aan

$$E_{-1} = \{(0, 0, r)^t \mid r \in \mathbb{R}\}.$$

Er zijn te weinig lineair onafhankelijke eigenvectoren om een basis van $\mathbb{R}^3$ te construeren; bijgevolg is B niet diagonaliseerbaar.

(3) Beschouw de matrix

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \in M_3(\mathbb{R}).$$

De karakteristieke vergelijking van C is

$$\chi_C(x) = (x-1)(x^2+1);$$

deze vergelijking heeft slechts één oplossing over $\mathbb{R}$, namelijk 1. We hebben dus één eigenwaarde $\lambda = 1$ over $\mathbb{R}$. De eigenruimte bij $\lambda = 1$ is gelijk aan

$$E_1 = \{(r,0,0)^t \mid r \in \mathbb{R}\}.$$

We hebben dus opnieuw te weinig lineair onafhankelijke eigenvectoren om een basis van $\mathbb{R}^3$ te construeren; bijgevolg is C niet diagonaliseerbaar over $\mathbb{R}$.

In het voorgaande voorbeeld hebben we twee matrices B en C bekeken die niet diagonaliseerbaar zijn over $\mathbb{R}$. Deze twee voorbeelden van niet-diagonaliseerbare operatoren zijn echter verschillend van aard.

In voorbeeld (3) heeft de karakteristieke vergelijking van C slechts één oplossing over $\mathbb{R}$ (multipliciteiten van de oplossingen meegerekend). De operator is niet diagonaliseerbaar over $\mathbb{R}$, omdat als we de wortels tellen met hun multipliciteit, de karakteristieke vergelijking *niet genoeg wortels* heeft.

In voorbeeld (2) heeft de karakteristieke vergelijking van B wel genoeg wortels, namelijk één van multipliciteit 1 en één van multipliciteit 2. Maar in dit geval zijn de dimensies van de eigenruimten te klein om een basis van eigenvectoren te hebben.

Voorbeeld 7.3.6. We bekijken nu wat er gebeurt als we de matrix C uit Voorbeeld 7.3.5(3) beschouwen over het veld $\mathbb{C}$ in plaats van over het veld $\mathbb{R}$. Stel dus

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \in M_3(\mathbb{C}).$$

De karakteristieke vergelijking van C is (met $i \in \mathbb{C}$ met $i^2 = -1$)

$$\chi_C(x) = (x-1)(x^2+1) = (x-1)(x+i)(x-i);$$

deze vergelijking heeft nu drie verschillende oplossingen over $\mathbb{C}$, namelijk $1, -i, i$. De matrix C heeft dus drie verschillende eigenwaarden, namelijk

$\lambda_1 = 1, \lambda_2 = -i, \lambda_3 = i$, en uit Stelling 7.3.4(ii) volgt onmiddellijk dat C diagonaliseerbaar is.

Voor de volledigheid bepalen we de eigenruimten van C over $\mathbb{C}$. De eigenruimte bij $\lambda_1 = 1$ is gelijk aan $E_1 = \{(r, 0, 0)^t \mid r \in \mathbb{C}\}$. De eigenruimte bij $\lambda_2 = -i$ is gelijk aan $E_{-i} = \{(0, r, -ir)^t \mid r \in \mathbb{C}\}$. De eigenruimte bij $\lambda_3 = i$ is gelijk aan $E_i = \{(0, r, ir)^t \mid r \in \mathbb{C}\}$. We vinden dus dat

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -i \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix} \right\}$$

een basis van $\mathbb{C}^3$ is die bestaat uit eigenvectoren over $\mathbb{C}$. Hieruit volgt dus nogmaals dat A diagonaliseerbaar is over $\mathbb{C}$; de matrixvoorstelling van $L_A \colon \mathbb{C}^3 \to \mathbb{C}^3 \colon v \mapsto Av$ t.o.v. $\mathcal{B}$ is de diagonaalmatrix $D = \operatorname{diag}(1, i, -i)$.

In het vorig voorbeeld hebben we aangetoond dat de matrix C over $\mathbb{C}$ wel diagonaliseerbaar is. Maar natuurlijk is ook niet iedere matrix over $\mathbb{C}$ diagonaliseerbaar: als we de matrix B over het veld $\mathbb{C}$ gaan beschouwen, dan blijft deze niet diagonaliseerbaar.

Het probleem dat de karakteristieke veelterm niet genoeg wortels heeft, komt vaak voor over willekeurige velden K. De volgende definitie is erg relevant in deze context.

Definitie 7.3.7. Zij $\varphi(x) \in K[x]$ een veelterm van graad $n \geq 1$. We zeggen dat $\varphi(x)$ *splijt over* K als $\varphi(x)$ precies n wortels heeft (multipliciteiten meegerekend), met andere woorden, als we $\varphi(x)$ kunnen schrijven als een product van n lineaire factoren.

Als een veelterm $\varphi(x)$ splijt, dan kunnen we deze schrijven als

$$\varphi(x) = \prod_{i=1}^{t} (x - \lambda_i)^{n_i}, \text{ met } \lambda_i \in K, \tag{7.4}$$

waarbij $\lambda_1, \ldots, \lambda_t$ de *verschillende* wortels van $\varphi(x)$ zijn.

Als we bovenstaande notatie gebruiken, zullen we altijd stilzwijgend onderstellen dat $\lambda_1, \ldots, \lambda_t$ *verschillend zijn.*

Merk dus op dat het splijten van de karakteristieke veelterm $\chi_f(x)$ over K een *nodige voorwaarde* (maar geen voldoende voorwaarde) is opdat f diagonaliseerbaar zou zijn over K. Inderdaad, als f diagonaliseerbaar is, dan is $\chi_f(x) = \chi_D(x)$ voor een diagonaalmatrix $D = \operatorname{diag}(d_1, \ldots, d_n)$. Het is duidelijk dat dan $\chi_D(x) = (x - d_1) \cdots (x - d_n)$, dus $\chi_f(x)$ splijt over K (en de n wortels zijn precies $d_1, \ldots, d_n$).

Opmerking 7.3.8. (i) Uit Gevolg 1.2.7 weten we dat elke veelterm van graad n over de complexe getallen $\mathbb{C}$ splijt. Dat is niet het geval voor veeltermen over $\mathbb{R}$, zoals we hebben gezien in Voorbeeld 7.3.5(3), omdat de veelterm $x^2 + 1$ niet splijt over $\mathbb{R}$.

(ii) Voor de wiskundigen wordt in de cursus “Algebra II” aangetoond dat als een veelterm $\varphi(x)$ *niet* splijt over K, er dan steeds een groter veld bestaat dat K bevat waarover $\varphi(x)$ wel splijt, net zoals we in Voorbeeld 7.3.6 het veld $\mathbb{R}$ hebben vergroot tot het veld $\mathbb{C}$. Een dergelijk veld wordt dan een *splijtveld* van $\varphi(x)$ genoemd.

We onderzoeken nu wanneer een lineaire operator $f\colon V \to V$ diagonaliseerbaar is. We zullen vanaf nu steeds veronderstellen dat de karakteristieke veelterm $\chi_f(x)$ splijt over K. De volgende definities zijn van cruciaal belang.

Definitie 7.3.9. Zij V een K-vectorruimte en $f\colon V \to V$ een lineaire operator met als karakteristieke veelterm

$$\chi_f(x) = \prod_{i=1}^{t} (x - \lambda_i)^{n_i} \quad \text{met } \lambda_1, \ldots, \lambda_t \in K.$$

(We benadrukken nogmaals dat dit een assumptie is.)

(i) De *algebraïsche multipliciteit* van een eigenwaarde λ_i van f is gelijk aan n_i, i.e. de multipliciteit van λ_i als wortel van de karakteristieke veelterm $\chi_f(x)$.

(ii) De *meetkundige multipliciteit* van een eigenwaarde λ_i van f is gelijk aan

$$d_i := \dim E_{\lambda_i} = \dim \ker(\lambda_i \mathbf{1} - f).$$

(De letter “d” in de notatie verwijst naar “dimensie” van de eigenruimte.)

Merk op dat zowel de algebraïsche als de meetkundige multipliciteit van een eigenwaarde steeds groter dan of gelijk aan 1 is.

Lemma 7.3.10. *Zij V een K-vectorruimte en $f\colon V \to V$ een lineaire operator met karakteristieke veelterm*

$$\chi_f(x) = \prod_{i=1}^{t} (x - \lambda_i)^{n_i} \quad \textit{met } \lambda_1, \ldots, \lambda_t \in K.$$

Dan is $1 \le d_i \le n_i$ voor elke i.

Bewijs. We zullen aantonen dat $(x - \lambda_i)^{d_i}$ een deler is van het karakteristiek polynoom van f; dit impliceert dan dat $d_i \leq n_i$.

Beschouw nu een vaste i, en stel $d = d_i$ en $\lambda = \lambda_i$. Zij $\{v_1, \ldots, v_d\}$ een basis van de eigenruimte E_{λ_i}. We breiden deze basis uit tot een basis van V, stel $\mathcal{B} = \{v_1, \ldots, v_d, w_1, \ldots, w_s\}$, met $s = n - d$. We stellen de matrixvoorstelling op van f ten opzichte van deze basis. Deze wordt bekomen door de matrix te beschouwen waarvan de kolommen

$$f(v_1), \ldots, f(v_d), f(w_1), \ldots, f(w_s)$$

zijn. We hebben $f(v_i) = \lambda v_i$ voor alle $i \in \{1, \ldots, d\}$; over de beelden $f(w_1), \ldots, f(w_s)$ hebben we geen bijkomende informatie. Dan is de matrixvoorstelling van f t.o.v. $\mathcal{B}$ gelijk aan

$$A = \left(\begin{array}{c|c} \lambda I_d & B \\ \hline 0 & C \end{array}\right)$$

voor zekere (onbekende) matrices $B \in M_{d,n-d}(K)$ en $C \in M_{n-d}(K)$.

Er geldt dat $\chi_f(x) = \det(xI_n - A)$. Als we deze determinant bepalen door telkens (d opeenvolgende malen) te ontwikkelen naar de eerste kolom, volgt er dat $(x - \lambda)^d$ een deler is van $\chi_f(x) = \det(xI_n - A)$. $\square$

Uit het voorgaand lemma volgt dat, als de algebraïsche multipliciteit n_i van een eigenwaarde gelijk is aan 1, de meetkundige multipliciteit d_i dan ook gelijk is aan 1.

We kunnen nu een criterium bewijzen voor het al dan niet diagonaliseerbaar zijn van een operator.

Stelling 7.3.11. *Zij V een K-vectorruimte en $f\colon V \to V$ een lineaire operator met karakteristieke veelterm*

$$\chi_f(x) = \prod_{i=1}^{t} (x - \lambda_i)^{n_i} \quad \text{met } \lambda_1, \ldots, \lambda_t \in K.$$

Dan is f diagonaliseerbaar als en slechts als $d_i = n_i$ voor alle i, met andere woorden, als en slechts als voor iedere eigenwaarde de meetkundige multipliciteit gelijk is aan de algebraïsche multipliciteit.

Bewijs. Aangezien de graad van $\chi_f(x)$ gelijk is aan n, is $\dim(V) = n = n_1 + \cdots + n_t$. Voor iedere eigenruimte E_{λ_i} kiezen we een basis $\mathcal{B}_i$, die dus d_i elementen bevat. Uit Stelling 7.3.4(i) volgt dat de verzameling $T := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_t$ lineair onafhankelijk is.

Nu splitsen we op in twee gevallen:

(i) Stel dat $d_i = n_i$ voor alle i. Dan is

$$|T| = d_1 + \cdots + d_t = n_1 + \cdots + n_t = n = \dim(V).$$

Dus T is een basis van V bestaande uit eigenvectoren van f. Uit Stelling 7.3.2 volgt dat f diagonaliseerbaar is.

(ii) Stel dat $d_i < n_i$ voor een bepaalde i, en merk op dat steeds $d_j \leq n_j$ voor alle j wegens Lemma 7.3.10. Dan is

$$|T| = d_1 + \cdots + d_t < n_1 + \cdots + n_t = n = \dim(V).$$

Hieruit volgt dat T geen basis is van V. Aangezien iedere eigenvector van f in $\mathrm{span}(T)$ zit, kunnen we nooit $\dim(V)$ lineair onafhankelijke eigenvectoren vinden. Dus f is niet diagonaliseerbaar. $\square$

We illustreren dit nieuwe criterium aan de hand van enkele voorbeelden.

Voorbeelden 7.3.12. (1) We beschouwen de matrix $A \in M_3(\mathbb{R})$ uit Voorbeeld 7.3.5. Aangezien $\chi_A(x) = (x-1)^2(x+1)$, kunnen we Stelling 7.3.11 toepassen. De eigenwaarde λ_1 heeft algebraïsche multipliciteit 2, de eigenwaarde λ_2 heeft algebraïsche multipliciteit 1. Uit de bepaling van de eigenruimten volgt dat de eigenwaarde λ_1 meetkundige multipliciteit 2 heeft; we kunnen dus onmiddellijk concluderen dat A diagonaliseerbaar is.

(2) Beschouw de matrix $B \in M_3(\mathbb{R})$ uit Voorbeeld 7.3.5. Aangezien $\chi_B(x) = (x-1)(x+1)^2$ kunnen we Stelling 7.3.11 toepassen. De eigenwaarde λ_1 heeft algebraïsche multipliciteit 1, de eigenwaarde λ_2 heeft algebraische multipliciteit 2. Uit de bepaling van de eigenruimten volgt dat de eigenwaarde λ_2 meetkundige multipliciteit 1 heeft; we kunnen dus onmiddellijk concluderen dat B niet diagonaliseerbaar is.

(3) Beschouw de shiftoperator S op K^n uit Voorbeeld 7.2.4(2), en beschouw de standaardbasis $\mathcal{B} = \{e_1, \ldots, e_n\}$ van K^n. Ten opzichte van deze basis $\mathcal{B}$ heeft S de matrixvoorstelling

$$A := \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Herhaaldelijk ontwikkelen naar de laatste kolom geeft $\chi_A(x) = x^n$; we kunnen dus Stelling 7.3.11 toepassen. De shiftoperator S heeft dus 1 eigenwaarde $\lambda = 0$ met algebraïsche multipliciteit n. De eigenruimte bij λ is gelijk aan $\{(0, \ldots, 0, k) \mid k \in K\}$; bijgevolg is de meetkundige multipliciteit van λ gelijk aan 1. We besluiten dat S niet diagonaliseerbaar is.

Opmerking 7.3.13. Zij V een K-vectorruimte en $f\colon V \to V$ een lineaire operator met karakteristieke veelterm

$$\chi_f(x) = \prod_{i=1}^{t} (x - \lambda_i)^{n_i} \quad \text{met } \lambda_1, \ldots, \lambda_t \in K.$$

Als f niet diagonaliseerbaar is, is het toch steeds mogelijk om een "mooie" matrixvoorstelling voor f te geven.

Er bestaat namelijk steeds een basis voor V zodat de matrixvoorstelling van f ten opzichte van deze basis een zeer eenvoudige vorm heeft, namelijk een blokdiagonaalmatrix

$$\left(\begin{array}{c|c|c|c} J_1 & & & \\ \hline & J_2 & & \\ \hline & & \ddots & \\ \hline & & & J_m \end{array}\right),$$

waarbij de blokken matrices zijn met een unieke eigenwaarde λ, met 1-en juist onder de diagonaal, en met alle andere componenten gelijk aan nul, i.e. blokken van de vorm

$$\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}.$$

Matrices met deze bijzondere vorm noemen we *Jordan*[2] *matrices*. De blokdiagonale matrixvoorstelling van f noemt men de *Jordan normaalvorm* van f.

Voor een bewijs van dit feit verwijzen we voor de wiskundigen naar de cursus "Algebra I".

[2]Genoemd naar de Franse wiskundige Marie Ennemond Camille Jordan (1838–1922).

Tot slot geven we nog een interessant verband tussen de determinant en het spoor van een lineaire operator, en diens eigenwaarden.

Stelling 7.3.14. *Zij $f\colon V \to V$ een lineaire operator op een n-dimensionale vectorruimte V. Onderstel dat de karakteristieke veelterm $\chi_f(x)$ splijt over K, en schrijf $\chi_f(x)$ zoals in* (7.4)*. Dan is*

$$\det(f) = \prod_{i=1}^{t} \lambda_i^{n_i} \quad \textit{en} \quad \operatorname{tr}(f) = \sum_{i=1}^{t} n_i \lambda_i.$$

Anders gezegd, de determinant en het spoor zijn respectievelijk het product en de som van alle eigenwaarden, waarbij we elke eigenwaarde even vaak meerekenen als zijn algebraïsche multipliciteit.

Bewijs. Uit Lemma 7.1.6 weten we dat we de karakteristieke veelterm kunnen uitschrijven als

$$\chi_f(x) = x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1 x + c_0$$

met $c_0 = (-1)^n \det A$ en $c_{n-1} = -\operatorname{tr} A$. Anderzijds weten we uit (7.4) dat

$$\chi_f(x) = \prod_{i=1}^{t} (x - \lambda_i)^{n_i},$$

waarbij de λ_i de verschillende eigenwaarden van f zijn omwille van Stelling 7.2.5. Als we in deze laatste uitdrukking de coëfficiënten van x^0 en x^{n-1} berekenen, vinden we het gevraagde resultaat. $\square$

7.4 De minimaalveelterm van een lineaire operator

Deze hele paragraaf 7.4 maakt geen deel uit van de leerstof voor de studenten van de opleiding Fysica en Sterrenkunde.

Het zal in het vervolg belangrijk blijken om uit een gegeven lineaire operator nieuwe lineaire operatoren te construeren door middel van veeltermen.

Definitie 7.4.1. Zij $f \in \operatorname{End}(V)$ een lineaire operator.

(i) Voor elk natuurlijk getal ℓ stellen we

$$f^\ell := \underbrace{f \circ \cdots \circ f}_{\ell \text{ keer}},$$

waarbij we $f^0 := \mathbf{1}_V$ definiëren. (Merk op dat de notatie f^ℓ in feite niet nieuw is, en precies overeenkomt met de ringstructuur die we hebben ingevoerd in Stelling 3.5.2.)

(ii) Elke uitdrukking $\sum_{i=0}^d a_i f^i$ met $a_i \in K$ definieert een lineaire operator op V, gegeven door

$$\Big(\textstyle\sum_{i=0}^d a_i f^i\Big)(v) = \sum_{i=0}^d a_i f^i(v)$$

voor alle $v \in V$.

(iii) Zij nu $\varphi(x) = \sum_{i=0}^d a_i x^i \in K[x]$ een veelterm. Dan definiëren we $\varphi(f) := \sum_{i=0}^d a_i f^i \in \operatorname{End}(V)$.

(iv) We definiëren de deelverzameling $K[f]$ van $\operatorname{End}(V)$ als

$$K[f] := \Big\{\sum_{i=0}^d a_i f^i \mid a_i \in K\Big\} = \big\{\varphi(f) \mid \varphi(x) \in K[x]\big\}.$$

Opmerking 7.4.2. Onderstel dat V eindig-dimensionaal is. Zij A_f een matrixvoorstelling van f ten opzichte van een willekeurige basis $\mathcal{B}$ voor V. Stel nu $\varphi(x) \in K[x]$ en $g := \varphi(f)$ zoals in Definitie 7.4.1(iii). Uit Stelling 4.1.8(iii) volgt dan dat de matrixvoorstelling van g ten opzichte van $\mathcal{B}$ gegeven wordt door de matrix

$$A_g = A_{\varphi(f)} = \varphi(A_f) \in M_n(K).$$

Stelling 7.4.3. *Zij V een eindig-dimensionale K-vectorruimte, en beschouw $f \in \operatorname{End}(V)$.*

(i) *De deelverzameling $K[f]$ is een commutatieve deelalgebra (d.w.z. een K-deelruimte en een commutatieve deelring) van* $\operatorname{End}(V)$.

(ii) $\dim_K K[f] \leq n^2$.

(iii) *Er bestaat een $d \in \mathbb{N}$ zodat de verzameling $\{1, f, f^2, \ldots, f^{d-1}\}$ een basis vormt voor $K[f]$, en $\dim K[f] = d$.*

(iv) *Indien $f^d = -\sum_{i=0}^{d-1} a_i f^i$, met d zoals in* (iii), *dan is*

$$\mu(x) := x^d + \sum_{i=0}^{d-1} a_i x^i \in K[x]$$

de monische veelterm van kleinst mogelijke graad zodat $\mu(f)$ de nuloperator is.

(v) *Zij $\varphi(x) \in K[x]$ een veelterm waarvoor geldt dat $\varphi(f) = 0$. Dan is $\mu(x)$ een deler van $\varphi(x)$.*

Bewijs. (i) Het is duidelijk dat $K[f]$ een deelruimte vormt van $\operatorname{End}(V)$. We bewijzen nu dat de samenstelling van twee operatoren $g, h \in K[f]$ opnieuw tot $K[f]$ behoort. Schrijf dus $g = \varphi(f)$ en $h = \psi(f)$ voor zekere $\varphi(x), \psi(x) \in K[x]$; dan is

$$g \circ h = \varphi(f) \circ \psi(f) = (\varphi\psi)(f)\,,$$

waarbij $(\varphi\psi)(x)$ het product voorstelt van de veeltermen $\varphi(x)$ en $\psi(x)$. Dit toont aan dat $g \circ h \in K[f]$, en dus is $K[f]$ een deelring. Het is ook duidelijk dat $g \circ h = h \circ g$ omdat $\varphi\psi = \psi\varphi$, en dus is $K[f]$ een commutatieve deelring.

(ii) Uit Stelling 3.4.5 volgt dat $\dim \operatorname{End}(V) = n^2$. Omdat $K[f]$ een deelruimte is van $\operatorname{End}(V)$ besluiten we dat $\dim K[f] \leq n^2$.

(iii) De verzameling $\{1, f, f^2, f^3, \dots\} \subset K[f]$ brengt de ruimte $K[f]$ voort. Omdat $\dim K[f] \leq n^2$ eindig is, kan een lineair onafhankelijke verzameling in $K[f]$ hoogstens n^2 elementen bevatten. Lemma 2.3.6(iii) impliceert nu dat er een kleinste natuurlijk getal d bestaat zodanig dat $f^d \in \langle 1, f, f^2, \dots, f^{d-1} \rangle$. We hebben dus een lineaire relatie van de vorm

$$f^d + \sum_{i=0}^{d-1} a_i f^i = 0 \tag{7.5}$$

met $a_i \in K$, en omdat d minimaal gekozen is, is de verzameling $T := \{1, f, f^2, \dots, f^{d-1}\}$ lineair onafhankelijk. Per inductie volgt, door (7.5) te vermenigvuldigen met machten van f, dat voor alle $k > d$ eveneens $f^k \in \operatorname{span}(T)$. De verzameling T brengt dus $K[f]$ voort. Omdat we reeds weten dat T ook lineair onafhankelijk is, is het een basis voor $K[f]$.

(iv) Merk op dat $\mu(x)$ inderdaad een monische veelterm is zodat $\mu(f) = 0$. Veronderstel nu dat er een veelterm $\psi(x)$ zou zijn van graad kleiner dan d zodat $\psi(f) = 0$; dan zou hieruit volgen dat er een lineaire relatie bestaat tussen de elementen $1, f, f^2, \dots, f^{d-1}$, in strijd met (iii).

(v) Zij $\varphi(x) \in K[x]$ zodat $\varphi(f) = 0$. Uit Stelling 1.2.3 volgt dat er veeltermen $q(x), r(x) \in K[x]$ zijn met $\deg r(x) < \deg \mu(x)$ of $r(x) = 0$, zodat

$$\varphi(x) = \mu(x)q(x) + r(x)\,.$$

Hieruit volgt dat $0 = \varphi(f) = \mu(f)q(f) + r(f) = r(f)$; uit (iv) volgt dat $r(x) = 0$. We besluiten dat $\varphi(x) = \mu(x)q(x)$. □

Definitie 7.4.4. Zij $f \in \operatorname{End}(V)$, V een K-vectorruimte. De monische veelterm $\mu_f(x)$ van kleinste graad waarvoor $\mu_f(f) = 0$ noemt men de *minimaalveelterm* van de operator f.

Voorbeelden 7.4.5. (1) Beschouw de matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \in M_2(\mathbb{R})$. Dan is $A^2 = \left(\begin{smallmatrix} 7 & 10 \\ 15 & 22 \end{smallmatrix}\right)$ en bijgevolg is $A^2 = 5A + 2I_2$. De minimaalveelterm van A is dus $\mu_A(x) = x^2 - 5x - 2$.

(2) Beschouw de matrix $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \in M_2(\mathbb{R})$. Dan is $A^2 = 0$, en dus is de minimaalveelterm van A gelijk aan $\mu_A(x) = x^2$.

(3) Beschouw een matrix van de vorm $A = \lambda I_n \in M_n(\mathbb{R})$. (We noemen een dergelijke matrix soms een *scalaire matrix*, omdat het een scalair veelvoud van de eenheidsmatrix is.) Dan is de minimaalveelterm van A gelijk aan $\mu_A(x) = x - \lambda$. Ook het omgekeerde is waar: als voor een matrix $A \in M_n(K)$ geldt dat $\deg \mu_A(x) = 1$, dan is A een scalaire matrix.

Opmerking 7.4.6. Als $\mu_f(x)$ de minimaalveelterm is van een lineaire operator $f \in \operatorname{End}(V)$, met V een n-dimensionale vectorruimte, dan is $\deg \mu_f(x) = \dim K[f] \leq \dim \operatorname{End}(V) = n^2$. We zullen later zien dat $\dim K[f] \leq n$; zie Gevolg 7.5.2.

Opmerking 7.4.7. (i) Het is belangrijk om het verschil in te zien tussen de veeltermring $K[x]$ en de ring $K[f]$. Zo is $K[x]$ oneindig-dimensionaal, terwijl $K[f]$ eindig-dimensionaal is. Een ander verschil is dat $K[x]$ geen nuldelers bevat terwijl $K[f]$ nuldelers kan bevatten (namelijk als $\mu_f(x)$ een reducibele veelterm is). Zie ook Opmerking 3.5.4(ii).

(ii) Wat we eigenlijk hebben aangetoond in het bewijs van Stelling 7.4.3(i), is dat de afbeelding

$$\gamma \colon K[x] \to K[f] \colon \varphi(x) \mapsto \varphi(f)$$

een *ringmorfisme* is, dit wil zeggen, het is een afbeelding tussen twee ringen die de optelling en de vermenigvuldiging respecteert. (De vermenigvuldiging in $K[x]$ is de gewone vermenigvuldiging van veeltermen, de vermenigvuldiging in $K[f]$ is de samenstelling van lineaire operatoren.) Het ringmorfisme γ is surjectief, maar niet injectief. Stelling 7.4.3(v) zegt precies dat de kern van γ bestaat uit de verzameling van alle veeltermen in $K[x]$ die deelbaar zijn door $\mu_f(x)$.

7.5 De stelling van Cayley–Hamilton

Deze hele paragraaf 7.5 maakt geen deel uit van de leerstof voor de studenten van de opleiding Fysica en Sterrenkunde.

De karakteristieke veelterm is een uitermate belangrijke invariant van een lineaire operator. De stelling van Cayley–Hamilton[3] drukt uit dat een lineaire operator steeds voldoet aan zijn eigen karakteristieke vergelijking.

Stelling 7.5.1 (Stelling van Cayley–Hamilton)**.** (i) *Zij* $A \in M_n(K)$ *met karakteristieke veelterm* $\chi_A(x)$*. Dan is* $\chi_A(A) = 0 \in M_n(K)$*.*

(ii) *Zij* V *een* n*-dimensionale vectorruimte over* K*, en zij* $f \in \mathrm{End}(V)$ *met karakteristieke veelterm* $\chi_f(x)$*. Dan is* $\chi_f(f)$ *de nuloperator op* V*.*

Bewijs. (i) We noteren in dit bewijs $I := I_n$. Schrijf

$$\chi_A(x) = x^n + \sum_{i=0}^{n-1} c_i x^i = \det(xI - A) \in K[x]. \tag{7.6}$$

Beschouw de matrix $xI - A$ in $M_n(K[x])$, en zijn adjunctmatrix

$$\mathrm{adj}(xI - A) = \big(p_{ij}(x)\big) \in M_n(K[x]).$$

Iedere $p_{ij}(x) \in K[x]$ is een veelterm met $\deg p_{ij}(x) \leq n-1$, want de elementen van de adjunctmatrix zijn determinanten van $(n-1)\times(n-1)$-deelmatrices van $xI - A$. Bijgevolg is

$$\mathrm{adj}(xI - A) = \big(p_{ij}(x)\big) = B_{n-1}x^{n-1} + B_{n-2}x^{n-2} + \cdots + B_1 x + B_0$$

waarbij elke $B_i \in M_n(K)$. Wegens (7.6) geldt nu de volgende gelijkheid van veeltermen[4] in de variabele x met coëfficiënten in $M_n(K)$:

$$\begin{aligned} I \cdot x^n + c_{n-1} I \cdot x^{n-1} + \cdots + c_1 I \cdot x + c_0 I &= \chi_A(x) I \\ &= \det(xI - A) I = (xI - A)\,\mathrm{adj}(xI - A) \\ &= (I \cdot x - A)(B_{n-1} \cdot x^{n-1} + B_{n-2} \cdot x^{n-2} + \cdots + B_1 \cdot x + B_0). \end{aligned}$$

Wanneer we nu voor elke k de coëfficiënt van x^k in linker- en rechterlid vergelijken, bekomen we

$$I = B_{n-1},$$

[3]Genoemd naar de Britse wiskundige Arthur Cayley (1821–1895) en de Ierse wiskundige Sir William Rowan Hamilton (1805–1865). Deze laatste is ook beroemd omwille van zijn ontdekking van de quaternionen, en de bijhorende vandalistische daad op de Broom Bridge in Dublin.

[4]Opgelet, dit zijn veeltermen waarvan de coëfficiënten behoren tot een *niet-commutatieve* ring, en de vermenigvuldiging van dergelijke veeltermen is dus ook niet commutatief. Anderzijds commuteert de variabele x wel met alles. Dit heeft als gevolg dat het niet toegestaan is om dergelijke veeltermen te "evalueren" in een matrix. Bijvoorbeeld, de veelterm $\varphi(x) = Ax - I$ kunnen we ook schrijven als $xA - I$, maar $AB - I$ en $BA - I$ zijn in het algemeen verschillend, dus $\varphi(B)$ heeft geen betekenis.

$$
\begin{aligned}
c_i I &= B_{i-1} - AB_i \quad \text{voor alle } i = 1, \dots, n-1, \\
c_0 I &= -AB_0.
\end{aligned}
$$

We links-vermenigvuldigen nu elk van deze vergelijkingen met de gepaste A^i, en we bekomen

$$
\begin{aligned}
A^n &= A^n B_{n-1}, \\
c_i A^i &= A^i B_{i-1} - A^{i+1} B_i \quad \text{voor alle } i = 1, \dots, n-1, \\
c_0 I &= -AB_0.
\end{aligned}
$$

Wanneer we deze $n+1$ vergelijkingen optellen bekomen we

$$
A^n + \sum_{i=0}^{n-1} c_i A^i = A^n B_{n-1} + \sum_{i=1}^{n-1} (A^i B_{i-1} - A^{i+1} B_i) - AB_0 = 0.
$$

(ii) Dit volgt onmiddellijk uit (i) en Opmerking 7.4.2. □

Gevolg 7.5.2. *Zij $f \in \mathrm{End}(V)$ een lineaire operator op een eindig-dimensionale vectorruimte V. De karakteristieke veelterm $\chi_f(x)$ van f is een veelvoud van de minimaalveelterm $\mu_f(x)$.*

In het bijzonder is $\dim K[f] = \deg \mu_f(x) \leq \dim V$.

Bewijs. Uit de stelling van Cayley–Hamilton volgt dat voor $\chi_f(x) \in K[x]$ geldt dat $\chi_f(f) = 0$. Uit Stelling 7.4.3(v) volgt dat de minimaalveelterm $\mu_f(x)$ een deler is van $\chi_f(x)$.

Aangezien $\deg \chi_f(x) = \dim V$, volgt hieruit ook nog $\deg \mu_f(x) \leq \dim V$. Anderzijds weten we uit Opmerking 7.4.6 dat $\dim K[f] = \deg \mu_f(x)$. □

We hebben nog een sterker verband tussen de minimaalveelterm en de karakteristieke veelterm.

Stelling 7.5.3. *Zij $f \in \mathrm{End}(V)$ een lineaire operator op een eindig-dimensionale vectorruimte V. Dan is elke eigenwaarde λ van f een wortel van de minimaalveelterm $\mu_f(x)$. In het bijzonder hebben $\mu_f(x)$ en $\chi_f(x)$ dezelfde wortels (maar met een mogelijks verschillende multipliciteit).*

Bewijs. Zij $v \in V$ een eigenvector voor f met eigenwaarde λ. We passen de gelijkheid $\mu_f(f) = 0$ toe op de vector v. Aangezien $f^i(v) = \lambda^i v$ voor elke $i \geq 0$, zien we dat $\mu_f(f)(v) = \mu_f(\lambda)v$, en uit $\mu_f(f) = 0$ volgt dan dat de scalair $\mu_f(\lambda)$ gelijk aan 0 moet zijn. Dus λ is een wortel van de veelterm $\mu_f(x)$. □

Opmerking 7.5.4. Als $\chi_f(x)$ splijt over K, dan zal ook $\mu_f(x)$ splijten over K, want $\mu_f(x) \mid \chi_f(x)$. Concreet wil dit zeggen dat als we $\chi_f(x)$ schrijven zoals in (7.4), dan

$$\chi_f(x) = \prod_{i=1}^{t}(x - \lambda_i)^{n_i} \quad \text{en} \quad \mu_f(x) = \prod_{i=1}^{t}(x - \lambda_i)^{\ell_i}$$

waarbij $1 \leq \ell_i \leq n_i$ voor elke i wegens Gevolg 7.5.2 en Stelling 7.5.3.

Opmerking 7.5.5. Het is misschien erg verleidelijk om de stelling van Cayley–Hamilton te "bewijzen" door in de definitie $\chi_A(x) = \det(xI_n - A)$ gewoon $x = A$ in te vullen en hieruit te besluiten dat $\chi_A(A) = \det(AI_n - A) = \det(0) = 0$. Dit houdt om vele redenen geen steek. Zo is $\chi_A(A)$ een *matrix*, terwijl de determinant van de nulmatrix een *scalair* is, dus zelfs de "types" van linkerlid en rechterlid zijn niet compatibel. We verwijzen naar

`https://en.wikipedia.org/wiki/Cayley-Hamilton_theorem`

voor verdere uitleg en andere juiste bewijzen en foutieve bewijspogingen van deze belangrijke stelling.

Hoofdstuk

8 Inproduct-ruimten

Een inproduct-ruimte is een vectorruimte V over $\mathbb{R}$ of $\mathbb{C}$ die bijkomend is uitgerust met een nieuwe operatie, het inproduct, die met elk koppel van vectoren $v, w \in V$ een scalair $\langle v, w\rangle$ associeert. Het inproduct zal ons in staat stellen om meetkundige concepten zoals *lengtes* en *loodrechte stand* (orthogonaliteit) in te voeren. Als we voor het inproduct op een eindig-dimensionale vectorruimte het "standaard inproduct" kiezen, dan zal dit overeenkomen met de notie van lengte en orthogonaliteit zoals we die al kennen in een Euclidische ruimte.

8.1 Definitie en eerste eigenschappen

In dit hoofdstuk werken we met vectorruimten over $\mathbb{R}$ of $\mathbb{C}$. We zullen vaak de definities en eigenschappen omtrent het veld $\mathbb{C}$ in Definitie 1.1.1 en Lemma 1.1.2 gebruiken; we doen dit vaak zonder expliciete referenties te geven. Zo zullen we in dit hoofdstuk vaak gebruik maken van de *complexe toevoeging* $z \mapsto \overline{z}$ ingevoerd in Definitie 1.1.1.

We beginnen met de algemene definitie van een inproduct-ruimte.

Definitie 8.1.1. Zij V een vectorruimte over het veld $K = \mathbb{R}$ of $K = \mathbb{C}$. Een afbeelding $\langle\cdot,\cdot\rangle : V \times V \to K$ wordt een *inproduct* op V genoemd, als aan de volgende voorwaarden is voldaan:

- [*toegevoegd-symmetrisch*] $\langle v, w\rangle = \overline{\langle w, v\rangle}$ voor alle $v, w \in V$; in het bijzonder is $\langle v, v\rangle \in \mathbb{R}$ voor alle $v \in V$;
- [*lineair in het eerste argument*] $\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$ voor alle $u, v, w \in V$ en alle $a, b \in K$;
- [*positief-definiet*] $\langle v, v\rangle > 0$ voor alle $v \in V \setminus \{0\}$.

Het paar $(V, \langle\cdot,\cdot\rangle)$ wordt dan een *inproduct-ruimte* genoemd; het inproduct zelf wordt vaak weggelaten indien het duidelijk is uit de context, en men zegt dan eenvoudigweg dat V een inproduct-ruimte is.

Als V een inproduct-ruimte is, dan definiëren we de *norm* van een element $v \in V$ als
$$\|v\| := \sqrt{\langle v, v \rangle}.$$

Merk op dat indien $(V, \langle \cdot, \cdot \rangle)$ een inproduct-ruimte is, en $W \leq V$ een deelruimte is van V, dan is ook $(W, \langle \cdot, \cdot \rangle)$ een inproduct-ruimte.

Voor we verder gaan, vermelden we een aantal eenvoudige gevolgen van de definiërende eigenschappen van inproduct-ruimten.

Lemma 8.1.2. *Zij* $(V, \langle \cdot, \cdot \rangle)$ *een inproduct-ruimte over* $K = \mathbb{R}$ *of* $\mathbb{C}$. *Dan gelden volgende eigenschappen, voor alle* $u, v, w \in V$ *en alle* $\lambda, \mu \in K$.

(i) [*bi-additiviteit*] $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ *en* $\langle u, v+w \rangle = \langle u, v \rangle + \langle u, w \rangle$;

(ii) [*sesqui-lineariteit*] $\langle \lambda v, \mu w \rangle = \lambda \overline{\mu} \cdot \langle v, w \rangle$;

(iii) [*homogeniteit van de norm*] $\|\lambda v\| = |\lambda| \cdot \|v\|$.

Bovendien is $\langle \cdot, \cdot \rangle$ niet-ontaard, *i.e. als* $u \in V$ *zodanig is dat* $\langle u, v \rangle = 0$ *voor alle* $v \in V$, *dan is noodzakelijk* $u = 0$.

Bewijs. Oefening. □

De norm geeft aanleiding tot een notie van afstand, en op deze wijze wordt V een zogenaamde *metrische ruimte*.

Definitie 8.1.3. Zij V een inproduct-ruimte. Dan definiëren we, voor elke twee elementen $v, w \in V$, de *afstand* tussen v en w als
$$\operatorname{dist}(v, w) := \|w - v\|.$$

Uit Lemma 8.1.2(iii) zien we dat dist symmetrisch is, i.e. $\operatorname{dist}(v, w) = \operatorname{dist}(w, v)$ voor alle $v, w \in V$.

We geven een aantal voorbeelden van inproduct-ruimten. In het vervolg van de cursus zullen Voorbeelden 8.1.4(1) en (3) regelmatig gebruikt worden.

Voorbeelden 8.1.4. (1) Zij $K = \mathbb{R}$, en $V = \mathbb{R}^n$. Dan definiëren we het *standaard inproduct* op V als de afbeelding
$$\langle \cdot, \cdot \rangle \colon V \times V \to K \colon (v, w) \mapsto v^t \cdot w,$$
of expliciet,
$$\langle (x_1, \dots, x_n)^t, (y_1, \dots, y_n)^t \rangle = x_1 y_1 + \dots + x_n y_n.$$

Dit inproduct wordt ook het *Euclidische inproduct* genoemd, en deze inproduct-ruimte wordt de *Euclidische n-dimensionale ruimte*[1] genoemd. Soms wordt hiervoor de notatie $\mathbb{E}^n$ gebruikt.
We verkrijgen de volgende formule voor de afstand tussen twee vectoren:

$$\operatorname{dist}((x_1,\ldots,x_n)^t,(y_1,\ldots,y_n)^t) = \sqrt{(y_1-x_1)^2+\cdots+(y_n-x_n)^2}\,.$$

(2) Het standaard inproduct is zeker niet het enige mogelijke inproduct op $\mathbb{R}^n$. Zo kunnen we bijvoorbeeld op $\mathbb{R}^3$ ook de afbeelding

$$\langle (x_1,x_2,x_3)^t,(y_1,y_2,y_3)^t\rangle = x_1y_1+2x_2y_2+3x_3y_3$$

beschouwen; dit is evenzeer een inproduct. Ten opzichte van dit inproduct heeft ook het begrip "afstand" een nieuwe betekenis; we hebben nu namelijk

$$\operatorname{dist}((x_1,x_2,x_3)^t,(y_1,y_2,y_3)^t) = \sqrt{(y_1-x_1)^2+2(y_2-x_2)^2+3(y_3-x_3)^2}\,.$$

Bij het gebruik van deze begrippen is het dus steeds belangrijk om in het achterhoofd te houden wat het onderliggende inproduct is.

(3) Zij $K=\mathbb{C}$, en $V=\mathbb{C}^n$. Dan definiëren we het *standaard inproduct* op V als de afbeelding

$$\langle\cdot,\cdot\rangle\colon V\times V\to K\colon (v,w)\mapsto v^t\cdot\overline{w}\,,$$

waarbij $\overline{w}$ de vector is die bekomen wordt door de complexe toevoeging te nemen van de coördinaten van w; expliciet is

$$\langle (x_1,\ldots,x_n)^t,(y_1,\ldots,y_n)^t\rangle = x_1\overline{y_1}+\cdots+x_n\overline{y_n}\,.$$

(4) Zij V een 4-dimensionale vectorruimte over $\mathbb{R}$, en zij $(e_1,\ldots,e_4)$ een basis voor V. We definiëren nu

$$\langle (x_1,\ldots,x_4),(y_1,\ldots,y_4)\rangle := x_1y_1+x_2y_2+x_3y_3-x_4y_4\,.$$

Dan is $(V,\langle\cdot,\cdot\rangle)$ een zogenaamde *Minkowski*[2] *ruimte*, en $\langle\cdot,\cdot\rangle$ wordt het *Minkowski inproduct* genoemd. Nochtans is $(V,\langle\cdot,\cdot\rangle)$ *géén* inproduct-ruimte! Inderdaad, deze afbeelding is symmetrisch en lineair, maar niet positief-definiet. De Minkowski ruimte speelt een belangrijke rol in de fysica, voornamelijk in Einstein's speciale relativiteitstheorie; de eerste drie dimensies stellen de 3-dimensionale ruimte voor, terwijl de vierde dimensie de tijd voorstelt. Een vector in de Minkowski ruimte wordt dan ook een *event* (gebeurtenis) genoemd.

[1] Genoemd naar Euclides van Alexandrië, een Hellenistisch wiskundige, die rond het jaar 300 v.Chr. werkzaam was in de bibliotheek van Alexandrië. Hij wordt soms de "vader van de meetkunde" genoemd.

[2] Genoemd naar de Duitse wiskundige Hermann Minkowski (1864–1909).

(5) Zij $a, b \in \mathbb{R}$ met $a < b$, en zij $C[a, b]$ de verzameling van continue complexwaardige functies op het interval $[a, b]$. Stel

$$\langle f, g \rangle := \int_a^b f(t)\overline{g(t)}dt\,.$$

Dan is $(C[a, b], \langle \cdot, \cdot \rangle)$ een (oneindig-dimensionale) inproduct-ruimte.

(6) Zij B een mogelijks oneindige verzameling. We definiëren de *reeksruimte* $\ell^2(B)$ (lees: kleine el-twee van B) als

$$\ell^2(B) := \left\{ f \colon B \to \mathbb{C} \;\middle|\; \sum_{b \in B} |f(b)|^2 < \infty \right\}.$$

Dan definieert

$$\langle f, g \rangle := \sum_{b \in B} f(b)\overline{g(b)}$$

een inproduct op $\ell^2(B)$.

Opmerking 8.1.5. Voor de geïnteresseerde lezer vermelden we dat een inproduct-ruimte $(V, \langle \cdot, \cdot \rangle)$ een *Hilbertruimte* [3] genoemd wordt als V *compleet* is als metrische ruimte, d.w.z. dat elke Cauchy-rij in V convergeert naar een element van V. In het bijzonder is elke eindig-dimensionale inproduct-ruimte een Hilbertruimte. (We verwijzen de wiskundigen naar een toekomstige cursus Analyse voor een nauwkeurige behandeling van deze begrippen in algemene metrische ruimten.) De inproduct-ruimte in Voorbeeld 8.1.4(5) is géén Hilbertruimte; de inproduct-ruimte in Voorbeeld 8.1.4(6) is dat wél.

Hilbertruimten over $\mathbb{C}$ spelen een cruciale rol in de kwantummechanica. Hierbij wordt de status van een fysisch systeem voorgesteld door een vector (of preciezer nog, door een straal van vectoren) in een complexe Hilbertruimte, en wordt voorgesteld als een zogenaamde *ket* die als $|\psi\rangle$ wordt genoteerd. De elementen van de *duale* Hilbertruimte (zie hoofdstuk **??** voor de definitie van duale vectorruimte) worden *bra*s genoemd en als $\langle\phi|$ genoteerd. Deze bras beelden dus kets af op complexe getallen, en dit geeft aanleiding tot Paul Dirac's *bra-ket* notatie $\langle\phi|\psi\rangle := \langle\phi|\,(|\psi\rangle) \in \mathbb{C}$. Deze bra-kets komen precies overeen met het inproduct van de oorspronkelijke Hilbertruimte.

Opmerking 8.1.6. De geïnteresseerde lezer kan zich afvragen waarom we ons bij het bestuderen van inproduct-ruimten beperken tot $\mathbb{R}$ en $\mathbb{C}$, en niet

[3] Genoemd naar de Duitse wiskundige David Hilbert (1862–1943), één van de meest invloedrijke en universele wiskundigen van de 19e en vroege 20e eeuw.

werken over willekeurige velden. Een eerste beperking die zich uiteraard opdringt, is dat we het begrip "positief" moeten kunnen definiëren (het inproduct moet positief-definiet zijn), en daarom moet ons veld in elk geval een *geordend deelveld* bevatten en bijgevolg karakteristiek[4] 0 hebben. Ten tweede moet elk positief element het kwadraat zijn van een ander element van het veld; dit is immers nodig om de norm te kunnen definiëren als de vierkantswortel van een zeker element. Ten derde willen we (om diepgaandere redenen) dat in elk geval de eindig-dimensionale inproduct-ruimten steeds compleet zijn als metrische ruimten, en dat blijkt bijvoorbeeld nooit het geval te zijn als het onderliggend veld een eigenlijk deelveld van $\mathbb{R}$ of $\mathbb{C}$ zou zijn.

We bewijzen een belangrijke eigenschap in verband met de norm van willekeurige inproduct-ruimten.

Stelling 8.1.7 (Ongelijkheid van Cauchy–Schwarz[5]). *Zij V een inproduct-ruimte over $\mathbb{R}$ of $\mathbb{C}$, en $v, w \in V$ willekeurig. Dan geldt steeds*

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|\,,$$

en de gelijkheid geldt dan en slechts dan als v en w lineair afhankelijk zijn.

Bewijs. Indien $w = 0$ is de bewering triviaal voldaan; stel dus $w \neq 0$. Als $v = \lambda w$ voor zekere $\lambda \in \mathbb{C}$, dan is

$$|\langle v, w\rangle| = |\langle \lambda w, w\rangle| = |\lambda| \cdot |\langle w, w\rangle| = |\lambda| \cdot \|w\|^2 = \|\lambda w\| \cdot \|w\| = \|v\| \cdot \|w\|.$$

Veronderstel dus $v \neq \lambda w$ voor alle $\lambda \in \mathbb{C}$, en definieer nu $\lambda = \langle w, w\rangle^{-1}\langle v, w\rangle$. Merk op dat $\overline{\lambda} = \langle w, w\rangle^{-1}\langle w, v\rangle$, en bijgevolg

$$\lambda\langle w, v\rangle = \overline{\lambda}\langle v, w\rangle = \lambda\overline{\lambda}\langle w, w\rangle = \langle w, w\rangle^{-1}\langle v, w\rangle\langle w, v\rangle\,.$$

Omdat het inproduct positief-definiet is, is

$$\begin{aligned}
0 &< \langle v - \lambda w, v - \lambda w\rangle \\
&= \langle v, v\rangle - \lambda\langle w, v\rangle - \overline{\lambda}\langle v, w\rangle + \lambda\overline{\lambda}\langle w, w\rangle \\
&= \langle v, v\rangle - \langle w, w\rangle^{-1}\langle v, w\rangle\langle w, v\rangle \\
&= \|v\|^2 - \|w\|^{-2}|\langle v, w\rangle|^2
\end{aligned}$$

en omdat zowel $\|v\|$, $\|w\|$ als $|\langle v, w\rangle|$ niet-negatieve reële getallen zijn, volgt het resultaat. □

[4]De *karakteristiek* van een veld is het kleinste geheel getal $n > 0$ zodat $\underbrace{1 + \cdots + 1}_{n \text{ keer}} = 0$, indien een dergelijke n bestaat; in het andere geval stellen we de karakteristiek gelijk aan 0.

[5]Genoemd naar de Franse wiskundige Augustin-Louis Cauchy (1789–1857) en de Duitse wiskundige Karl Hermann Amandus Schwarz (1843–1921).

Een gevolg van de ongelijkheid van Cauchy–Schwarz is dat de norm aan de driehoeksongelijkheid voldoet.

Gevolg 8.1.8. *Zij* $(V, \langle \cdot, \cdot \rangle)$ *een inproduct-ruimte over* $\mathbb{R}$ *of* $\mathbb{C}$.

(i) [Driehoeksongelijkheid] *Voor alle* $v, w \in V$ *hebben we*

$$\|v + w\| \leq \|v\| + \|w\| .$$

(ii) [Stelling van Pythagoras] *Als* $v, w \in V$ *met* $\langle v, w \rangle = 0$*, dan geldt*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Bewijs. We hebben

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \|v\|^2 + \langle v, w \rangle + \langle w, v \rangle + \|w\|^2 \\ &= \|v\|^2 + \langle v, w \rangle + \overline{\langle v, w \rangle} + \|w\|^2 \\ &\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 \quad \text{(Cauchy–Schwarz)} \\ &= \big(\|v\| + |w\|\big)^2 . \end{aligned}$$

□

8.2 Orthogonaliteit

We hebben reeds gezien dat we in een inproduct-ruimte een zinvolle notie van afstand hebben. Ook het begrip “orthogonaliteit”, of anders gezegd, het “loodrecht op elkaar staan” van vectoren, houdt steek in een inproduct-ruimte. We kunnen niet zomaar een goed begrip van hoek definiëren op een eenvoudige manier, maar we zullen in Hoofdstuk 9 zien dat dit wel kan bij reële inproduct-ruimten.

Definitie 8.2.1. Beschouw een inproduct-ruimte $(V, \langle \cdot, \cdot \rangle)$.

(i) Een vector $v \in V$ staat *loodrecht* op een vector $w \in V$ als

$$\langle v, w \rangle = 0 .$$

Merk op dat de relatie “loodrecht op” symmetrisch is. We zeggen ook dat de vectoren v en w *orthogonaal* zijn (ten opzichte van elkaar). We noteren dit met $v \perp w$.

(ii) Een basis $\mathcal{B}$ voor V wordt *orthogonaal* genoemd als geldt dat $\langle b, b' \rangle = 0$ voor alle $b \neq b' \in \mathcal{B}$. Dit betekent dus dat de basiselementen twee aan twee loodrecht op elkaar staan.

(iii) Een basis $\mathcal{B}$ voor V wordt *orthonormaal* genoemd als geldt dat $\langle b, b' \rangle = 0$ voor alle $b \neq b' \in \mathcal{B}$ en $\langle b, b \rangle = 1$ voor alle $b \in \mathcal{B}$. Dit betekent dus dat de basiselementen twee aan twee loodrecht op elkaar staan, en elk norm 1 hebben.

(iv) Als $W \leq V$ een deelruimte is, noemen we

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ voor alle } w \in W\}$$

het *orthogonaal complement* van W.

In Lemma 8.2.5 tonen we aan dat $W^\perp$ een complement (zie Definitie 2.5.9) is van W, met andere woorden dat $W \oplus W^\perp = V$.

Lemma 8.2.2. *Het orthogonaal complement $W^\perp$ van een deelruimte $W \leq V$ is een deelruimte van V. Er geldt dat $W \cap W^\perp = \{0\}$.*

Bewijs. Stel dat $v_1, v_2 \in W^\perp$ en $\lambda_1, \lambda_2 \in K$. Dan is

$$\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \lambda_1 \langle v_1, w \rangle + \lambda_2 \langle v_2, w \rangle = 0$$

voor iedere $w \in W$, en dus $\lambda_1 v_1 + \lambda_2 v_2 \in W^\perp$; bijgevolg is $W^\perp$ een deelruimte.

Stel nu dat $v \in W \cap W^\perp$; dan is $\langle v, v \rangle = 0$, en uit het positief-definiet zijn volgt dat $v = 0$. □

Voorbeeld 8.2.3. (1) Zij $K = \mathbb{R}$ of $\mathbb{C}$, en beschouw de vectorruimte K^n met het standaard inproduct. Het is duidelijk dat de standaardbasis $\{e_1, \ldots, e_n\}$ een orthonormale basis is.

(2) Zij $(V, \langle \cdot, \cdot \rangle)$ een inproduct-ruimte. Als $\{v_1, \ldots, v_n\}$ een orthogonale basis is van V, dan is

$$\left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|} \right\}$$

een orthonormale basis van V.

In de volgende stelling wordt aangetoond dat iedere eindig-dimensionale deelruimte een orthonormale basis bezit. In Stelling 8.3.1 verderop zullen we een efficiënte expliciete methode geven om een dergelijke orthonormale basis te construeren.

Stelling 8.2.4. *Zij $K = \mathbb{R}$ of $\mathbb{C}$, en zij $(V, \langle \cdot, \cdot \rangle)$ een eindig-dimensionale inproduct-ruimte over K, stel van dimensie d. Dan bezit V een orthonormale basis $\{b_1, \ldots, b_d\}$.*

Bewijs. We bewijzen de uitspraak met inductie naar d.

Veronderstel dus eerst dat $d = 1$, en zij $0 \neq w \in V$, dan is $\{b_1\}$ met

$$b_1 = \frac{1}{\|w\|} w$$

de gezochte basis.

Veronderstel nu dat $d > 1$; de inductiehypothese stelt dat de uitspraak waar is voor deelruimten van dimensie $d-1$, en we bewijzen de uitspraak nu voor een willekeurige deelruimte V van dimensie d.

Zij $0 \neq w \in V$ en definieer de 1-dimensionale vectorruimte $W = Kw$. Beschouw de lineaire vorm

$$f \colon V \to K \colon v \mapsto \langle v, w \rangle\,;$$

merk op dat $f(w) = \langle w, w \rangle \neq 0$. De kern van deze lineaire vorm is precies het orthogonaal complement $W^\perp$ van W in V. Uit Lemma 4.3.3 volgt nu dat $V = W^\perp \oplus W$; in het bijzonder is $\dim W^\perp = d - 1$.

Door de inductiehypothese weten we dat $W^\perp$ een orthonormale basis $\{b_2, \dots, b_d\}$ bezit. Neem $b_1 = \frac{1}{\|w\|} w$; dan is $\{b_1\}$ een basis voor W. Omdat $V = W \oplus W^\perp$, volgt hieruit dat $\{b_1, \dots, b_d\}$ een basis is voor V. Men verifieert nu eenvoudig dat deze basis orthonormaal is. □

Lemma 8.2.5. *Zij $K = \mathbb{R}$ of $\mathbb{C}$, en zij $(V, \langle \cdot, \cdot \rangle)$ een eindig-dimensionale inproduct-ruimte over K. Voor elke deelruimte $W \leq V$ is het orthogonaal complement $W^\perp$ in V een complementaire ruimte, i.e. $W \oplus W^\perp = V$.*

In het bijzonder is $\dim(W^\perp) = \dim V - \dim W$.

Bewijs. Vermits $W \cap W^\perp = \{0\}$ wegens Lemma 8.2.2 moeten we enkel nog aantonen dat $W + W^\perp = V$.

Wegens Stelling 8.2.4 bestaat er een orthonormale basis $\{b_1, \dots, b_d\}$ voor W. Zij $v \in V$; dan is

$$v = \sum_{i=1}^{d} \langle v, b_i \rangle b_i + \Big(v - \sum_{i=1}^{d} \langle v, b_i \rangle b_i \Big).$$

De eerste term is een element van W. We tonen aan dat de tweede term $z := v - \sum_{i=1}^{d} \langle v, b_i \rangle b_i$ bevat is in $W^\perp$. Het volstaat om aan te tonen dat ieder basiselement van W loodrecht staat op z. Voor elke $j \in \{1, \dots, d\}$ hebben we, wegens de lineariteit van het inproduct in het eerste argument,

$$\langle z, b_j \rangle = \langle v, b_j \rangle - \sum_{i=1}^{d} \langle v, b_i \rangle \langle b_i, b_j \rangle = \langle v, b_j \rangle - \langle v, b_j \rangle = 0. \qquad \square$$

Opmerking 8.2.6. De uitdrukking $\sum_{i=1}^{d} \langle v, b_i \rangle b_i$ uit bovenstaand bewijs zal precies de *orthogonale projectie* van v op W zijn, die we zo dadelijk zullen invoeren. In principe hadden we deze uitdrukking al kunnen gebruiken als definitie, maar dan is het niet evident dat dit onafhankelijk is van de keuze van de orthonormale basis $\{b_1, \dots, b_d\}$ voor W. Zoals we dadelijk zullen zien, zullen we Lemma 8.2.5 net *gebruiken* om de orthogonale projectie te definiëren. Nadien volgt gemakkelijk dat dit overeenkomt met bovenstaande uitdrukking; zie Lemma 8.2.9(ii).

Gevolg 8.2.7. *Zij $K = \mathbb{R}$ of $\mathbb{C}$, en zij $(V, \langle \cdot, \cdot \rangle)$ een eindig-dimensionale inproduct-ruimte over K. Voor elke deelruimte $W \leq V$ is $(W^\perp)^\perp = W$.*

Bewijs. De inclusie $W \leq (W^\perp)^\perp$ is evident (denk goed na over de betekenis van $(W^\perp)^\perp$). Omdat $\dim(W^\perp)^\perp = \dim V - (\dim V - \dim W) = \dim W$, volgt nu dat de inclusie een gelijkheid is. $\square$

We bespreken nu de orthogonale projectie op een deelruimte, die we hebben vermeld in Opmerking 8.2.6.

Definitie 8.2.8. Zij $K = \mathbb{R}$ of $\mathbb{C}$, en zij $(V, \langle \cdot, \cdot \rangle)$ een eindig-dimensionale inproduct-ruimte over K.

(i) Beschouw een deelruimte $W \leq V$ en een element $v \in V$. Aangezien $V = W \oplus W^\perp$, is elke $v \in V$ op een unieke manier te schrijven als $v = w + u$ met $w \in W$ en $u \in W^\perp$. We definiëren de *orthogonale projectie* van v op W als het element $w \in W$, en we noteren dit als $\text{proj}_W(v) := w$.

(ii) Als W een 1-dimensionale deelruimte is, stel $W = \langle w \rangle$, dan noteren we $\text{proj}_W(v)$ ook als $\text{proj}_w(v)$.

Lemma 8.2.9. *Zij $K = \mathbb{R}$ of $\mathbb{C}$, en zij $(V, \langle \cdot, \cdot \rangle)$ een eindig-dimensionale inproduct-ruimte over K. Beschouw een deelruimte $W \leq V$.*

(i) *Voor alle $w \in W$ is $\text{proj}_W(w) = w$; voor alle $u \in W^\perp$ is $\text{proj}_W(u) = 0$. In het bijzonder is de orthogonale projectie een projectie-operator in de betekenis van Definitie* 3.1.10.

(ii) *Zij $\{b_1, \dots b_d\}$ een orthonormale basis van W. Dan is*

$$\text{proj}_W(v) = \sum_{i=1}^{d} \langle v, b_i \rangle b_i$$

voor elke $v \in V$.

(iii) *Zij* $w \in W$ *met* $w \neq 0$. *Dan is*

$$\operatorname{proj}_w(v) = \frac{\langle v, w\rangle}{\langle w, w\rangle} w$$

voor elke $v \in V$.

Bewijs. (i) Dit volgt uit $w = w + 0 \in W \oplus W^\perp$ en $u = 0 + u \in W \oplus W^\perp$.

(ii) Net als in het bewijs van Gevolg 8.2.5 is

$$v = \sum_{i=1}^{d} \langle v, b_i\rangle b_i + \Big(v - \sum_{i=1}^{d} \langle v, b_i\rangle b_i\Big) \in W \oplus W^\perp.$$

(iii) Stel $W = \langle w\rangle$; dan is $\{w/\|w\|\}$ een orthonormale basis voor W. Uit (ii) volgt dan dat

$$\operatorname{proj}_w(v) = \operatorname{proj}_W(v) = \left\langle v, \frac{w}{\|w\|}\right\rangle \frac{w}{\|w\|} = \frac{\langle v, w\rangle}{\|w\|^2} w. \qquad \square$$

We tonen nu aan dat de kortste afstand van een element tot een gegeven deelruimte gegeven wordt door de orthogonale projectie.

Stelling 8.2.10. *Zij* $K = \mathbb{R}$ *of* $\mathbb{C}$, *en zij* $(V, \langle\cdot,\cdot\rangle)$ *een eindig-dimensionale inproduct-ruimte over* K. *Beschouw een deelruimte* $W \leq V$ *en* $v \in V$. *Dan geldt voor elke* $w \in W \setminus \{\operatorname{proj}_W(v)\}$ *dat* $\operatorname{dist}(v, w) > \operatorname{dist}(v, \operatorname{proj}_W(v))$.

Bewijs. Ontbind $v - w$ volgens de decompositie $V = W \oplus W^\perp$ als $v - w = y + z$ met $y \in W$ en $z \in W^\perp$. Dan is per definitie $y = \operatorname{proj}_W(v - w) = \operatorname{proj}_W(v) - w$, en dus is $z = v - w - y = v - \operatorname{proj}_W(v)$. Aangezien $y \perp z$ volgt uit de Stelling van Pythagoras (Gevolg 8.1.8(ii)) dat

$$\|v - w\|^2 = \|y + z\|^2 = \|y\|^2 + \|z\|^2 = \|y\|^2 + \|v - \operatorname{proj}_W(v)\|^2.$$

Merk op dat $y \neq 0$ omdat $w \neq \operatorname{proj}_W(v)$ per assumptie, en dus is

$$\|v - w\|^2 > \|v - \operatorname{proj}_W(v)\|^2. \qquad \square$$

8.3 Gram–Schmidt orthonormalisatieproces

Tot slot geven we de beloofde methode om een orthonormale basis te construeren: het zogenaamde Gram–Schmidt orthonormalisatieproces.

Stelling 8.3.1 (Gram–Schmidt[6] orthonormalisatieproces)**.** *Zij $K = \mathbb{R}$ of $\mathbb{C}$, en zij $(V, \langle\cdot,\cdot\rangle)$ een eindig-dimensionale inproduct-ruimte over K. Zij W een deelruimte van V met basis $\{v_1, \dots, v_d\}$. We definiëren nu, op recursieve wijze,*

$$\begin{aligned}
b_1 &= v_1, \\
b_2 &= v_2 - \operatorname{proj}_{b_1}(v_2), \\
b_3 &= v_3 - \operatorname{proj}_{b_1}(v_3) - \operatorname{proj}_{b_2}(v_3), \\
b_4 &= v_4 - \operatorname{proj}_{b_1}(v_4) - \operatorname{proj}_{b_2}(v_4) - \operatorname{proj}_{b_3}(v_4), \\
&\ \vdots \\
b_d &= v_d - \sum_{i=1}^{d-1} \operatorname{proj}_{b_i}(v_d).
\end{aligned}$$

Dan is $\{b_1, \dots, b_d\}$ een orthogonale basis voor W, en dus is

$$\left\{ \frac{b_1}{\|b_1\|}, \dots, \frac{b_d}{\|b_d\|} \right\}$$

een orthonormale basis voor W.

Bewijs. We bewijzen eerst dat $\{b_1, \dots, b_d\}$ een basis is voor W. Per definitie van de elementen b_i geldt dat $v_i \in \operatorname{span}(b_1, \dots, b_i)$ voor elke i, zodat $W = \operatorname{span}(v_1, \dots, v_d) \leq \operatorname{span}(b_1, \dots, b_d)$. Hieruit volgt dat de verzameling $\{b_1, \dots, b_d\}$ voortbrengend is voor W, en aangezien ze precies $d = \dim W$ elementen bevat, kunnen we uit Stelling 2.4.12(ii) besluiten dat ze een basis is.

We bewijzen nu de orthogonaliteit. Door Lemma 8.2.9(iii) te gebruiken zien we dat de algemene formule om b_j te bepalen gegeven wordt door

$$b_j = v_j - \sum_{i=1}^{j-1} \frac{\langle v_j, b_i\rangle}{\langle b_i, b_i\rangle} b_i. \tag{8.1}$$

We bewijzen nu per inductie op j dat $b_1, \dots, b_j$ paarsgewijze orthogonaal zijn. Voor $j = 1$ valt er niks te bewijzen. Stel dus $j > 1$, en veronderstel dat we reeds weten (door de inductiehypothese) dat $b_1, \dots, b_{j-1}$ paarsgewijze orthogonaal zijn. We maken nu gebruik van (8.1), en we bekomen, voor alle $k < j$, dat

$$\langle b_j, b_k\rangle = \langle v_j, b_k\rangle - \sum_{i=1}^{j-1} \frac{\langle v_j, b_i\rangle}{\langle b_i, b_i\rangle} \langle b_i, b_k\rangle$$

[6] Genoemd naar de Deense actuaris en wiskundige Jørgen Pedersen Gram (1850–1916) en de Duitse wiskundige Erhard Schmidt (1876–1959), hoewel dit reeds eerder verschenen was in werk van Laplace en Cauchy.

$$
\begin{aligned}
&= \langle v_j, b_k\rangle - \frac{\langle v_j, b_k\rangle}{\langle b_k, b_k\rangle}\langle b_k, b_k\rangle \\
&= 0.
\end{aligned}
$$

Dit toont aan dat $b_j \perp b_k$ voor alle $k < j$ en bewijst dus dat $b_1, \dots, b_j$ paarsgewijze orthogonaal zijn. □

We sluiten deze paragraaf af met een observatie over de transitiematrix tussen twee orthonormale basissen in een inproduct-ruimte.

Stelling 8.3.2. *Beschouw een inproduct-ruimte $(V, \langle ., .\rangle)$ over $K = \mathbb{R}$ of $\mathbb{C}$. Zij $\mathcal{B} = \{b_1, \dots, b_n\}$ en $\mathcal{B}' = \{b'_1, \dots, b'_n\}$ twee orthonormale basissen van V, en zij Q de transitiematrix van $\mathcal{B}$ naar $\mathcal{B}'$. Dan is $Q^t\overline{Q} = I_n$.*[7]

Bewijs. We moeten aantonen dat $Q^t\overline{Q} = I_n$, of dus dat $\sum_k q_{ki}\overline{q}_{kj} = \delta_{ij}$ voor alle $i, j = 1 \dots, n$. Er geldt dat $b'_i = \sum_k q_{ki} b_k$, en aangezien de basissen orthonormaal zijn is $\langle b_i, b_j\rangle = \delta_{ij}$ en $\langle b'_i, b'_j\rangle = \delta_{ij}$, voor alle i, j. We vinden dat

$$
\begin{aligned}
\delta_{ij} = \langle b'_i, b'_j\rangle &= \Big\langle \sum_k q_{ki} b_k, \sum_\ell q_{\ell j} b_\ell \Big\rangle \\
&= \sum_k \sum_\ell q_{ki}\overline{q}_{\ell j}\langle b_k, b_\ell\rangle \\
&= \sum_k q_{ki}\overline{q}_{kj}.
\end{aligned}
$$

□

Voor inproduct-ruimten over $\mathbb{R}$ hebben we dus dat de transitiematrix tussen twee orthonormale basissen voldoet aan $Q^tQ = I_n$. Een dergelijke matrix noemt men een *orthogonale matrix*. Voor inproduct-ruimten over $\mathbb{C}$ zijn deze transitiematrices precies de *unitaire matrices*. Zie ook Definitie 8.5.6 verderop.

Voorbeeld 8.3.3. Beschouw het Euclidisch vlak $V = \mathbb{R}^2$ met het standaard inproduct, en beschouw de rotatie ρ_θ rond de oorsprong over een hoek θ uit Voorbeeld 3.1.9(7). De matrix $A_\theta = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$ zal dan elke orthonormale basis afbeelden op een orthonormale basis, en is dus een orthogonale matrix. We rekenen inderdaad eenvoudig na dat

$$
A_\theta^t A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
$$

[7] Hierbij is $\overline{Q} := (\overline{q}_{ij})$.

8.4 Hermitische en symmetrische operatoren

We bestuderen het diagonaliseerbaar zijn van bijzondere klassen van operatoren op inproduct-ruimten over $\mathbb{R}$ en $\mathbb{C}$.

Definitie 8.4.1. (i) Zij V een willekeurige inproduct-ruimte over $\mathbb{C}$ en zij $f \in \mathrm{End}_{\mathbb{C}}(V)$ een lineaire operator. We zeggen dat een lineaire operator $g \in \mathrm{End}_{\mathbb{C}}(V)$ *hermitisch toegevoegd*[8] is aan f als voor alle $v, w \in W$ geldt dat

$$\langle f(v), w\rangle = \langle v, g(w)\rangle\,. \tag{8.2}$$

De operator g is uniek bepaald door f omdat het inproduct niet-ontaard is; we noteren[9] deze ook wel als $g = f^\dagger$. We noemen de operator f *hermitisch* of *zelf-toegevoegd* als $f^\dagger = f$, met andere woorden als voor alle $v, w \in W$ geldt dat

$$\langle f(v), w\rangle = \langle v, f(w)\rangle\,. \tag{8.3}$$

(ii) Beschouw nu de n-dimensionale inproduct-ruimte $V = \mathbb{C}^n$ met het standaard inproduct $\langle v, w\rangle = v^t\overline{w}$. Zij $A, B \in M_n(\mathbb{C})$. We noemen B *hermitisch toegevoegd* aan A als de operator L_B hermitisch toegevoegd is aan L_A, en we noteren $B = A^\dagger$. We noemen A een *hermitische matrix* of een *zelf-toegevoegde matrix* als $A^\dagger = A$. Concreet, als $B = A^\dagger$, dan geldt voor alle $v, w \in \mathbb{C}^n$ dat

$$v^t A^t \overline{w} = (Av)^t\overline{w} = v^t(\overline{\overline{B}w}) = v^t\overline{B}\,\overline{w}\,,$$

waaruit volgt dat

$$A^\dagger = \overline{A^t}\,.$$

De matrix A is dus hermitisch als en slechts als $A^t = \overline{A}$.

(iii) We kunnen precies dezelfde definitie beschouwen voor $\mathbb{R}^n$ in plaats van $\mathbb{C}^n$. In dat geval vervangen we overal "hermitisch" door "symmetrisch". We spreken dus van de *symmetrisch-toegevoegde operator*, *symmetrische operatoren* en *symmetrische matrices.* Een reële matrix is dus symmetrisch als $A^t = A$, in overeenstemming met Definitie 1.3.3(vii).

Stelling 8.4.2. *Zij f een hermitische operator op $\mathbb{C}^n$ of een symmetrische operator op $\mathbb{R}^n$. Dan geldt:*

[8] Genoemd naar de Franse wiskundige Charles Hermite (1822–1901).

[9] Vaak wordt ook de notatie f^* gebruikt, maar we gebruiken deze notatie al voor de duale afbeelding; zie Definitie 4.3.10. Onze notatie $f^\dagger$ is zeer gebruikelijk in de fysica.

(i) *Alle eigenwaarden van f zijn reëel.*

(ii) *Eigenvectoren behorende bij twee verschillende eigenwaarden zijn orthogonaal.*

(iii) *De operator f is diagonaliseerbaar, en er bestaat steeds een orthonormale basis van eigenvectoren voor f.*

Bewijs. We nemen aan dat f een hermitische operator is op $\mathbb{C}^n$; het geval van een symmetrische operator f op $\mathbb{R}^n$ is analoog[10].

(i) Stel dat λ een eigenwaarde is van f met eigenvector v, m.a.w. $f(v) = \lambda v$. Nu is

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle.$$

Omdat $v \neq 0 \in \mathbb{C}^n$, is $\langle v, v\rangle \neq 0$, zodat $\lambda = \overline{\lambda}$ en dus $\lambda \in \mathbb{R}$.

(ii) Stel dat λ een eigenwaarde is van f met eigenvector v en μ een eigenwaarde is van f met eigenvector w met $\lambda \neq \mu$. Aangezien $\lambda, \mu \in \mathbb{R}$ is

$$\lambda\langle v, w\rangle = \langle f(v), w\rangle = \langle v, f(w)\rangle = \mu\langle v, w\rangle.$$

Er volgt dat $(\lambda - \mu)\langle v, w\rangle = 0$, dus is $\langle v, w\rangle = 0$.

(iii) We tonen aan dat er een orthonormale basis van eigenvectoren van f bestaat door gebruik te maken van inductie naar de dimensie. Voor 1-dimensionale ruimten is de uitspraak triviaal.
Als inductiehypothese nemen we aan dat er voor iedere hermitische operator op $\mathbb{C}^{n-1}$ een orthonormale basis van eigenvectoren is. Stel dat f een eigenvector v met eigenwaarde λ heeft; deze bestaat zeker want de karakteristieke vergelijking van f heeft steeds een wortel in $\mathbb{C}$. Noem W het orthogonaal complement van $\mathbb{C}v$, dus $\mathbb{C}^n = \mathbb{C}v \oplus W$ en $\dim(W) = n - 1$ (zie Gevolg 8.2.5). Voor $w \in W$ geldt dus dat $\langle v, w\rangle = 0$. We tonen aan dat ook $\langle v, f(w)\rangle = 0$:

$$\langle v, f(w)\rangle = \langle f(v), w\rangle = \lambda\langle v, w\rangle = 0.$$

Hieruit volgt dat $f(W) \subseteq W$, wat betekent dat de restrictie van f tot W een hermitische operator is op W. Omdat $\dim W = n - 1$ volgt nu uit de inductiehypothese dat W een orthonormale basis van

[10] In principe kan dit ook afgeleid worden uit het eerste geval door "uitbreiding van scalairen" van $\mathbb{R}$ naar $\mathbb{C}$, maar we hebben de juiste achtergrondkennis nog niet opgebouwd om hier een formele betekenis aan te geven.

eigenvectoren heeft voor de restrictie van f op W; deze basis noteren we met $\{w_1, \ldots, w_{n-1}\}$. Aangezien $\mathbb{C}^n = \mathbb{C}v \oplus W$, is

$$\left\{\tfrac{1}{\|v\|}v, w_1, \ldots, w_{n-1}\right\}$$

een orthonormale basis van V bestaande uit eigenvectoren van f. $\square$

Opmerking 8.4.3. In de praktijk gaan we vaak als volgt te werk om een orthonormale basis van eigenvectoren van een hermitische of symmetrische operator te vinden. We bepalen eerst de eigenwaarden en eigenruimten voor f; aangezien f diagonaliseerbaar is, is V de directe som van de eigenruimten. Voor iedere eigenruimte apart bepalen we nu een orthonormale basis bestaande uit eigenvectoren van f. Wanneer we al deze basisvectoren samenvoegen krijgen we een basis van V. Uit Stelling 8.4.2(ii) volgt dan dat al deze basisvectoren loodrecht op elkaar staan.

8.5 Lineaire groepen

De groep $\mathsf{GL}_n(K)$ van inverteerbare $n \times n$-matrices over een veld K noemt men de *algemene lineaire groep*. *Lineaire groepen* zijn deelgroepen van de algemene lineaire groep.

Definitie 8.5.1. Een *elementaire $n \times n$-matrix* is een matrix van de vorm

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \lambda & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \lambda & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

i.e. gelijk aan de eenheidsmatrix plus een veelvoud van een matrixeenheid U_{ij}, $i \neq j$. We noteren $E_{ij}(\lambda)$ voor de elementaire matrix met het element λ op de ij-de plaats.

Links vermenigvuldigen met een elementaire matrix $E_{ij}(\lambda)$ komt overeen met het vervangen van de i-de rij in een matrix door de i-de rij plus λ keer de j-de rij, dus met een elementaire rij-operatie van type III (zie Definitie 1.4.5 en Opmerking 1.4.7).

Opmerking 8.5.2. (1) Rechtsvermenigvuldiging met een elementaire matrix $E_{ij}(\lambda)$ komt overeen met het vervangen van de j-de kolom door de j-de kolom plus λ keer de i-de kolom.

(2) Elementaire matrices hebben determinant gelijk aan 1 en behoren dus tot $\mathsf{GL}_n(K)$.

(3) Sommige bronnen, waaronder ook de Wikipedia-pagina

`https://nl.wikipedia.org/wiki/Elementaire_matrix`,

definiëren drie types elementaire matrices, overeenkomstig met de drie types van elementaire rij-operaties die we hebben ingevoerd.

Stelling 8.5.3. *Elke matrix $A \in \mathsf{GL}_n(K)$ is van de vorm*

$$A = SD,$$

met S een product van elementaire matrices en $D = \mathrm{diag}(1, \ldots, 1, d)$ met $d = \det A$.

Bewijs. Het is voldoende aan te tonen dat door elementaire rij-operaties van type III een matrix A in $\mathsf{GL}_n(K)$ kan gereduceerd worden tot een matrix van de vorm $\mathrm{diag}(1, \ldots, 1, d)$.

Vermits de matrix rang n heeft is er een element in de eerste kolom dat niet nul is. Door de eerste rij te vervangen door de rij bekomen door bij de eerste rij een geschikt veelvoud van de rij waarin dit element staat op te tellen bekomen we een matrix met op de $(1, 1)$-de plaats een 1. (Indien $a_{11} \neq 0$ maar alle andere elementen van de eerste kolom wel nul zijn, tellen we eerst de eerste rij op bij de tweede rij, om de voorgaande operatie mogelijk te maken.)

Door nu voor $i = 2, \ldots, n$, de i-de rij te vervangen door een rij bekomen door bij de i-de rij een geschikt veelvoud van de eerste rij op te tellen, bekomen we voor elke $i = 2, \ldots, n$ op de $(i, 1)$-de plaats een 0. We hebben nu reeds een matrix van de vorm

$$A' = \begin{pmatrix} 1 & * \\ 0 & B \end{pmatrix},$$

waarbij B een $(n-1) \times (n-1)$-matrix is met $\det B \neq 0$, vermits $0 \neq \det A = \det A' = \det B$. Als de orde van B strikt groter is dan 1 kunnen we dus hetzelfde herhalen voor de matrix B, zonder de eerste rij van de matrix A' te veranderen. Inductief bekomen een bovendriehoeksmatrix met als diagonaal $(1, \ldots, 1, d)$, met $d = \det A$.

Met elementaire rij-operaties van type III kunnen we nu ook de elementen boven de diagonaal nul maken. Vermits de determinant van elementaire matrices gelijk is aan 1, is de determinant d van de gereduceerde matrix gelijk aan de determinant van de oorspronkelijke matrix A. $\square$

Definitie 8.5.4. De deelgroep $\mathsf{SL}_n(K)$ van $\mathsf{GL}_n(K)$ die bestaat uit de matrices met determinant gelijk aan 1 noemt men de *speciale lineaire groep.*

Stelling 8.5.5. *Elke matrix in* $\mathsf{SL}_n(K)$ *is het product van elementaire matrices.*

Bewijs. Dit volgt onmiddellijk uit Stelling 8.5.3. □

We definiëren enkele deelgroepen van $\mathsf{GL}_n(K)$ en $\mathsf{SL}_n(K)$ die een rol spelen in verschillende delen van de wiskunde.

Definitie 8.5.6. De *orthogonale groep* en de *speciale orthogonale groep* zijn de deelgroepen van $\mathsf{GL}_n(K)$ gegeven door respectievelijk

$$\begin{aligned} \mathsf{O}_n(K) &:= \{A \in \mathsf{GL}_n(K) \mid AA^t = A^tA = I_n\}\,, \\ \mathsf{SO}_n(K) &:= \mathsf{O}_n(K) \cap \mathsf{SL}_n(K)\,. \end{aligned}$$

Merk op dat de elementen van $\mathsf{O}_n(K)$ steeds determinant ± 1 hebben. Indien $K = \mathbb{R}$, worden deze groepen vaak genoteerd als respectievelijk $\mathsf{O}(n)$ en $\mathsf{SO}(n)$.

Veronderstel nu dat K een veld is, voorzien van een *involutie* σ, i.e. een veldautomorfisme van de orde 2 (dus $\sigma^2 = 1_K$ en $\sigma \neq 1_K$); we noteren $a^\sigma = \overline{a}$ voor alle $a \in K$. We definiëren nu voor alle $A = (a_{ij}) \in M_n(K)$ de *toegevoegde*[11] *matrix* $\overline{A} := (\overline{a}_{ij})$. De *unitaire groep* en de *speciale unitaire groep* zijn dan de deelgroepen van $\mathsf{GL}_n(K)$ gegeven door respectievelijk

$$\begin{aligned} \mathsf{U}_n(K,\sigma) &:= \{A \in \mathsf{GL}_n(K) \mid A\overline{A}^t = \overline{A}^tA = I_n\}\,, \\ \mathsf{SU}_n(K,\sigma) &:= \mathsf{U}_n(K,\sigma) \cap \mathsf{SL}_n(K)\,. \end{aligned}$$

Indien $K = \mathbb{C}$, waarbij de toevoeging de complexe toevoeging is, worden deze groepen vaak genoteerd als respectievelijk $\mathsf{U}(n)$ en $\mathsf{SU}(n)$.

De elementen van de orthogonale groep noemen we *orthogonale matrices*, de elementen van de unitaire groep noemen we *unitaire matrices.*

Opmerking 8.5.7. De reële groepen $\mathsf{O}(n)$ en $\mathsf{SO}(n)$, en de complexe groepen $\mathsf{U}(n)$ en $\mathsf{SU}(n)$, zijn voorbeelden van *Lie*[12] *groepen*; behalve de groepsstructuur hebben ze ook de bijkomende structuur van een differentieerbare variëteit, die bovendien compatibel is met de groepsstructuur.

[11] Dit dient niet verward te worden met het toegevoegd zijn van twee matrices (aan elkaar), zoals ingevoerd in Definitie 4.2.7.

[12] Genoemd naar de Noorse wiskundige Marius Sophus Lie (1842–1899), die in feite één van de grondleggers is van wat we nu kennen als groepentheorie.

8.6 De Euclidische groep $\mathsf{E}(n)$

In deze laatste paragraaf van dit hoofdstuk besteden we aandacht aan de isometrieën van de Euclidische ruimte $\mathbb{R}^n$.

Definitie 8.6.1. (i) Beschouw de Euclidische ruimte $\mathbb{R}^n$. Een *isometrie* van $\mathbb{R}^n$ is een bijectieve afbeelding φ van $\mathbb{R}^n$ naar zichzelf die de afstand bewaart, i.e. $\operatorname{dist}(\varphi(v), \varphi(w)) = \operatorname{dist}(v, w)$ voor alle $v, w \in \mathbb{R}^n$.

(ii) De verzameling van alle isometrieën van $\mathbb{R}^n$ vormt een groep (met de samenstelling als bewerking), die we de *isometriegroep* van $\mathbb{R}^n$ noemen; dit wordt ook de *Euclidische n-dimensionale groep* genoemd, en we noteren deze groep als $\mathsf{E}(n)$ of $\mathsf{ISO}(n)$.

(iii) Voor elke $v \in \mathbb{R}^n$ beschouwen we de afbeelding

$$T_v \colon \mathbb{R}^n \to \mathbb{R}^n : w \mapsto w + v\,.$$

Het is duidelijk dat $T_v \in \mathsf{E}(n)$, en dat de verzameling

$$\mathsf{T}(n) := \{T_v \mid v \in \mathbb{R}^n\}$$

een deelgroep vormt van $\mathsf{E}(n)$. We noemen de elementen T_v *translaties* van $\mathbb{R}^n$, en de deelgroep $\mathsf{T}(n)$ de *translatiegroep* van $\mathbb{R}^n$.

Merk op dat isometrieën niet noodzakelijk lineair zijn.

Lemma 8.6.2. *Zij $Q \in \mathsf{O}(n)$. Dan is $\varphi := L_Q \in \mathsf{E}(n)$.*

Bewijs. Voor alle $v, w \in \mathbb{R}^n$ hebben we

$$\varphi(v) \cdot \varphi(w) = Qv \cdot Qw = (Qv)^t(Qw) = v^tQ^tQw = v^tw = v \cdot w\,,$$

en dus bewaart φ het inproduct. Hieruit volgt dat φ ook de afstand bewaart. Aangezien Q inverteerbaar is, is φ bijectief. □

Stelling 8.6.3. *Zij $\mathsf{E}_0(n)$ de verzameling van de elementen van $\mathsf{E}(n)$ die de oorsprong 0 vasthouden. Dan is*

(i) $\mathsf{E}_0(n) = \mathsf{O}(n)$ *(waarbij we L_Q identificeren met Q),*

(ii) $\mathsf{E}(n) = \mathsf{T}(n)\mathsf{O}(n)$.

(iii) *Elk element van $\mathsf{E}(n)$ is op unieke wijze te schrijven als de samenstelling van een element van $\mathsf{T}(n)$ en een element van $\mathsf{O}(n)$.*

Bewijs. (i) Uit het voorgaande lemma volgt dat $\mathsf{O}(n) \subseteq \mathsf{E}_0(n)$. We tonen aan dat ook $\mathsf{E}_0(n) \subseteq \mathsf{O}(n)$.

Zij $\varphi \in \mathsf{E}_0(n)$. Aangezien φ de afstand bewaart en 0 fixeert, bewaart het ook de norm. Uit de relatie $v \cdot w = \frac{1}{2}\big(\|v+w\|^2 - \|v\|^2 - \|w\|^2\big)$ volgt dan dat φ ook het inproduct bewaart. Hieruit volgt dan op zijn beurt dat φ ook orthogonaliteit bewaart.

We tonen vervolgens aan dat φ een lineaire afbeelding is. Stel dus $v, w \in \mathbb{R}^n$ en $\lambda, \mu \in \mathbb{R}$, en beschouw $u = \varphi(\lambda v + \mu w) - \lambda\varphi(v) - \mu\varphi(w)$. We moeten aantonen dat $u = 0$, en omdat φ surjectief is en het inproduct niet-ontaard is (zie Lemma 8.1.2), volstaat het hiertoe om aan te tonen dat $u \cdot \varphi(z) = 0$ voor alle $z \in \mathbb{R}^n$. Door gebruik te maken van de bilineariteit van het inproduct, evenals het feit dat φ het inproduct bewaart, vinden we

$$\begin{aligned} u \cdot \varphi(z) &= \varphi(\lambda v + \mu w) \cdot \varphi(z) - \lambda\varphi(v) \cdot \varphi(z) - \mu\varphi(w) \cdot \varphi(z) \\ &= (\lambda v + \mu w) \cdot z - \lambda v \cdot z - \mu w \cdot z \\ &= 0\,. \end{aligned}$$

Dus $\varphi \in \mathsf{GL}_n(\mathbb{R})$. Aangezien φ zowel afstand als orthogonaliteit bewaart, zal φ elke orthonormale basis afbeelden op een orthonormale basis. Uit Stelling 8.3.2 volgt nu dat $\varphi \in \mathsf{O}(n)$.

(ii) Zij $\varphi \in \mathsf{E}(n)$, en stel $\varphi(0) = v$. Beschouw dan de translatie $T_{-v} = T_v^{-1}$; deze zal v afbeelden op 0, zodat de samenstelling $T_v^{-1} \circ \varphi$ de oorsprong 0 zal vasthouden. Dus $T_v^{-1} \circ \varphi \in \mathsf{E}_0(n) = \mathsf{O}(n)$, en bijgevolg

$$\varphi \in T_v \cdot \mathsf{O}(n) \subset \mathsf{T}(n)\mathsf{O}(n)\,. \qquad \square$$

(iii) Dit volgt nu uit het voorgaande, omdat $\mathsf{T}(n) \cap \mathsf{O}(n) = 1$.

Definitie 8.6.4. (i) Zij $\varphi \in \mathsf{E}(n)$, en schrijf $\varphi = T_v\,\varphi_0$, waarbij $\varphi_0 \in \mathsf{E}_0(n) = \mathsf{O}(n)$ uniek bepaald is door φ. Aangezien $\varphi_0 \in \mathsf{O}(n)$, is $\det \varphi_0 \in \{1, -1\}$. Als $\det \varphi_0 = 1$, noemen we φ een *directe isometrie* of een *verplaatsing* (soms ook *starre verplaatsing* of *starre beweging* genoemd, vooral in de mechanica van starre lichamen); als $\det \varphi_0 = -1$, noemen we φ een *indirecte isometrie.*

(ii) De verzameling van alle directe isometrieën van $\mathbb{R}^n$ vormt een deelgroep van $\mathsf{E}(n)$ die we als $\mathsf{E}^+(n)$ noteren, en de *verplaatsingsgroep* van $\mathbb{R}^n$ noemen.

Opmerking 8.6.5. Uit Stelling 8.6.3 volgt dat $\mathsf{E}^+(n) = \mathsf{T}(n)\mathsf{SO}(n)$, en meer bepaald dat elk element van $\mathsf{E}^+(n)$ op unieke wijze te schrijven is als de samenstelling van een element van $\mathsf{T}(n)$ en een element van $\mathsf{SO}(n)$.

Voorbeeld 8.6.6. Beschouw het Euclidisch vlak $V = \mathbb{R}^2$. We beweren dat de elementen van $\mathsf{SO}(2)$ precies gegeven worden door de rotaties rond de oorsprong:

$$\mathsf{SO}(2) = \{A_\theta \mid \theta \in \mathbb{R}\} = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}.$$

Inderdaad, als we uitdrukken dat een willekeurige matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ moet voldoen aan $A^t A = I_2$, dan krijgen we

$$\begin{cases} a^2 + c^2 = 1, \\ b^2 + d^2 = 1, \\ ab + cd = 0. \end{cases}$$

Dit zegt precies dat de kolommen $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ en $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ een orthonormale basis vormen! Inderdaad, de eerste vergelijking zegt dat $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ norm 1 heeft, de tweede vergelijking dat $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ norm 1 heeft, en de derde vergelijking dat $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ en $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ inproduct 0 hebben.

Nu is een vector in $\mathbb{R}^2$ met norm 1 altijd te schrijven als $\left(\begin{smallmatrix} \cos\theta \\ \sin\theta \end{smallmatrix}\right)$ voor een zekere hoek θ. Als we dus deze keuze maken voor $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$, dan is de tweede kolom $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ gelijk aan $\left(\begin{smallmatrix} -\sin\theta \\ \cos\theta \end{smallmatrix}\right)$ of $\left(\begin{smallmatrix} \sin\theta \\ -\cos\theta \end{smallmatrix}\right)$, en dus is

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{of} \quad A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

In het eerste geval is $\det A = 1$, in het tweede geval is $\det A = -1$. Aangezien $A \in \mathsf{SO}(2)$ moet $\det A = 1$, en we vinden $A = A_{-\theta}$.

Uit Opmerking 8.6.5 volgt nu dat elke directe isometrie van $\mathbb{R}^2$ de gedaante

$$\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2 \colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos\theta - y\sin\theta + s \\ x\sin\theta + y\cos\theta + t \end{pmatrix}$$

heeft, voor zekere $s, t \in \mathbb{R}$ en een zekere hoek $\theta \in \mathbb{R}$, en analoog heeft elke indirecte isometrie de gedaante

$$\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2 \colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos\theta + y\sin\theta + s \\ x\sin\theta - y\cos\theta + t \end{pmatrix}$$

voor zekere $s, t \in \mathbb{R}$ en een zekere hoek $\theta \in \mathbb{R}$.

Hoofdstuk

9 De Euclidische ruimte $\mathbb{R}^n$

Dit hele hoofdstuk 9 maakt geen deel uit van de leerstof voor de studenten van de opleiding Fysica en Sterrenkunde.

In Hoofdstuk 8 hebben we inproduct-ruimten over $\mathbb{R}$ en $\mathbb{C}$ bestudeerd. In dit hoofdstuk bestuderen we de Euclidische ruimte $\mathbb{R}^n$ in meer detail, en in het bijzonder behandelen we een aantal meetkundige aspecten.

We krijgen bovendien nog meer bijkomende structuur als we de Euclidische ruimte $\mathbb{R}^3$ in drie dimensies beschouwen, omdat we op een dergelijke ruimte ook nog eens een *vectorieel product* zullen kunnen definiëren.

9.1 Hoeken in een reële inproduct-ruimte

In Definitie 8.2.1(i) hebben we voor algemene inproduct-ruimten over $\mathbb{R}$ en $\mathbb{C}$ gedefinieerd wanneer twee vectoren v, w orthogonaal of loodrecht zijn; dit is het geval als $\langle v, w\rangle = 0$. Vanuit het klassieke standpunt zouden we het begrip orthogonaliteit graag koppelen aan een notie van een hoek, en in het bijzonder zodanig dat orthogonaliteit overeenkomt met een hoek van 90° oftewel $\pi/2$ (radiaal). Algemenere hoeken definiëren blijkt echter, in onze context, enkel zinvol te zijn in inproduct-ruimten over $\mathbb{R}$ en niet over $\mathbb{C}$.

Veronderstel vanaf nu dat $(V, \langle ., .\rangle)$ een inproduct-ruimte is over $\mathbb{R}$. We kunnen de ongelijkheid van Cauchy–Schwarz (zie Stelling 8.1.7) voor alle $v, w \in V \setminus \{0\}$ herschrijven in de vorm

$$-1 \leq \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \leq 1\,.$$

Bijgevolg is de volgende definitie zinvol.

Definitie 9.1.1. Zij $(V, \langle ., .\rangle)$ een reële inproduct-ruimte. Voor elke twee niet-nul vectoren $v, w \in V$ definiëren we de *hoek* tussen v en w als

$$\angle(v, w) := \arccos \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}\,,$$

waarbij we aannemen dat arccos waarden aanneemt in het interval $[0, \pi]$.

Aangezien het inproduct symmetrisch is, is het duidelijk dat ook $\angle$ symmetrisch is, i.e. $\angle(v, w) = \angle(w, v)$ voor alle $v, w \in V$.

Opmerking 9.1.2. (i) Merk op dat $\angle(v, w) = 0$ als en slechts als v en w een positief veelvoud zijn van elkaar, en dat $\angle(v, w) = \pi$ als en slechts als v en w een negatief veelvoud zijn van elkaar, in overeenstemming met onze intuïtie over het hoek-begrip.

(ii) Merk anderzijds op dat $\angle(v, w) = \pi/2$ als en slechts als $\langle v, w\rangle = 0$. De bovenstaande definitie van hoek is dus compatibel met Definitie 8.2.1(i).

We bekijken nu in detail hoe de formules er uitzien als we ons beperken tot de Euclidische n-dimensionale ruimte. Beschouw dus de inproduct-ruimte $\mathbb{E}^n = (\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ waarbij het inproduct het standaard Euclidische inproduct is zoals in Voorbeeld 8.1.4(1). Stel $v = (x_1, \ldots, x_n)^t$ en $w = (y_1, \ldots, y_n)^t$, dan is

$$\langle v, w\rangle = x_1y_1 + \cdots + x_ny_n\,,$$

en bijgevolg krijgen we voor de hoek

$$\angle(v, w) = \arccos \frac{x_1y_1 + \cdots + x_ny_n}{\sqrt{(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2)}}\,.$$

Notatie 9.1.3. Wanneer we werken in de Euclidische ruimte, is het zeer gebruikelijk om het inproduct van twee vectoren $u, v \in \mathbb{R}^n$ niet als $\langle u, v\rangle$ te noteren, maar simpelweg als $u \cdot v$ of uv. Overeenkomstig zullen we ook soms $\|v\|^2 = \langle v, v\rangle$ schrijven als vv of zelfs v^2. We zullen ook zelf in het vervolg deze conventies aanhouden. Dit zou geen verwarring mogen veroorzaken, aangezien er geen andere natuurlijke vermenigvuldiging is gedefinieerd op een vectorruimte. Zie echter wel de volgende paragraaf 9.2, waar we in $\mathbb{R}^3$ wel een vermenigvuldiging zullen invoeren die een paar van vectoren op een vector afbeeldt; dit vectorieel product zullen we echter steeds als $u \times v$ noteren.

9.2 Het vectorieel product in $\mathbb{R}^3$

In deze paragraaf bestuderen we de Euclidische ruimte $\mathbb{E}^3$ verder. Meer bepaald zullen we het *vectorieel product* invoeren (ook wel *kruisproduct* of *vectorproduct* genoemd). Dit product zal een paar van vectoren in $\mathbb{R}^3$ afbeelden op een nieuwe vector in $\mathbb{R}^3$ (in tegenstelling tot het inproduct dat een paar van vectoren afbeeldt op een element van het grondveld $\mathbb{R}$). Echter, om tot een intrinsieke definitie te komen (d.w.z. een definitie die niet afhangt van de keuze van een coördinatenstelsel, of dus van een orthonormale basis) zal het nodig zijn om de vectorruimte te *oriënteren*.

Definitie 9.2.1. Beschouw de standaardbasis $\mathcal{B} = \{e_1, \dots, e_n\}$ van $\mathbb{R}^n$ en beschouw een willekeurige geordende basis $\mathcal{B}' = \{f_1, \dots, f_n\}$. Zij Q de transitiematrix van $\mathcal{B}$ naar $\mathcal{B}'$. (Uit opmerking 4.2.2(ii) weten we dat Q precies de matrix is met kolommen $f_1, \dots, f_n$.) We noemen de basis $\mathcal{B}'$ *positief georiënteerd* of *direct georiënteerd* als $\det Q > 0$, en we noemen $\mathcal{B}'$ *negatief georiënteerd* of *indirect georiënteerd* als $\det Q < 0$. (Merk op dat $\det Q \neq 0$ omdat een transitiematrix altijd inverteerbaar is.)

Beschouw de groepen $\mathsf{O}(n)$ en $\mathsf{SO}(n)$ die we hebben ingevoerd in Definitie 8.5.6.

Lemma 9.2.2. *Beschouw de Euclidische ruimte $\mathbb{R}^n$ met standaardbasis $\mathcal{B} = \{e_1, \dots, e_n\}$. Zij $\mathcal{B}' = \{f_1, \dots, f_n\}$ een willekeurige orthonormale basis van $\mathbb{R}^n$, en zij Q de transitiematrix van $\mathcal{B}$ naar $\mathcal{B}'$. Als $\mathcal{B}'$ positief georiënteerd is, dan is $Q \in \mathsf{SO}(n)$, i.e. $QQ^t = Q^tQ = I_n$ en $\det Q = 1$.*

Bewijs. Uit Stelling 8.3.2 weten we dat $Q^tQ = I_n$ en dus is $Q \in \mathsf{O}(n)$ (zie Definitie 8.5.6); in het bijzonder is $\det(Q) \in \{1, -1\}$. Omdat $\mathcal{B}'$ positief georiënteerd is, volgt er nu dat $\det(Q) = 1$, en dus $Q \in \mathsf{SO}(n)$. □

Vanaf nu werken we in de 3-dimensionale Euclidische ruimte $V = \mathbb{R}^3$ met standaardbasis $\mathcal{B} = \{e_1, e_2, e_3\}$. We definiëren nu een afbeelding van $V \times V$ naar V die gebruik maakt van deze basis $\mathcal{B}$. We tonen verderop aan dat deze definitie niet afhangt van de keuze van de positief georiënteerde orthonormale basis.

Definitie 9.2.3. Zij $v = (x_1, x_2, x_3)^t \in \mathbb{R}^3$ en $w = (y_1, y_2, y_3)^t \in \mathbb{R}^3$ twee willekeurige elementen. Dan definiëren we het *vectorieel product* van v en w als

$$v \times w := (x_2y_3 - x_3y_2,\ x_3y_1 - x_1y_3,\ x_1y_2 - x_2y_1)^t\,. \tag{9.1}$$

We kunnen dit symbolisch schrijven als

$$v \times w = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix},$$

waarbij we benadrukken dat dit slechts een symbolische schrijfwijze is, aangezien de e_i's vectoren zijn, terwijl de x_i's en y_i's reële getallen zijn. We kunnen dit exact neerschrijven als

$$v \times w = \sum_{i=1}^{3} \begin{vmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} e_i\,. \tag{9.2}$$

Lemma 9.2.4. *De definitie $v \times w$ is onafhankelijk van de keuze van de positief georiënteerde orthonormale basis $\mathcal{B} = \{e_1, e_2, e_3\}$.*

Bewijs. Stel dat $\mathcal{B}' = \{f_1, f_2, f_3\}$ een andere positief georiënteerde orthonormale basis is, met Q de transitiematrix; uit Lemma 9.2.2 volgt dat $Q \in \mathsf{SO}(3)$.

Veronderstel nu dat v en w ten opzichte van de nieuwe basis $\mathcal{B}'$ coördinaten $(x_1', x_2', x_3')^t$, respectievelijk $(y_1', y_2', y_3')^t$ hebben. Uit Lemma 4.2.4(iii) volgt (na transponeren) dat

$$\begin{aligned} \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} &= \begin{pmatrix} x_1' & x_2' & x_3' \end{pmatrix} Q^t \,, \\ \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} &= \begin{pmatrix} y_1' & y_2' & y_3' \end{pmatrix} Q^t \,. \end{aligned}$$

Merk verder ook op dat uit $QQ^t = I_3$ volgt dat

$$\begin{pmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \end{pmatrix} = \begin{pmatrix} q_{i1} & q_{i2} & q_{i3} \end{pmatrix} Q^t$$

voor elke $i \in \{1, 2, 3\}$. Uiteindelijk verkrijgen we dus

$$\begin{aligned} \sum_{i=1}^{3} \begin{vmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} e_i &= \sum_{i=1}^{3} \det \left(\begin{pmatrix} q_{i1} & q_{i2} & q_{i3} \\ x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \end{pmatrix} \cdot Q^t \right) e_i \\ &= \det(Q) \cdot \sum_{i=1}^{3} \begin{vmatrix} q_{i1} & q_{i2} & q_{i3} \\ x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \end{vmatrix} e_i \,, \end{aligned}$$

en door deze laatste determinant te ontwikkelen naar de eerste rij, krijgen we

$$\begin{aligned} \sum_{i=1}^{3} \begin{vmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} e_i &= \det(Q) \cdot \sum_{i=1}^{3} \sum_{j=1}^{3} \begin{vmatrix} \delta_{j1} & \delta_{j2} & \delta_{j3} \\ x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \end{vmatrix} q_{ij} e_i \\ &= \det(Q) \cdot \sum_{j=1}^{3} \begin{vmatrix} \delta_{j1} & \delta_{j2} & \delta_{j3} \\ x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \end{vmatrix} f_j \,. \end{aligned}$$

Aangezien $\det(Q) = 1$ zien we dat de definitie

$$v \times w = \sum_{i=1}^{3} \begin{vmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} e_i$$

inderdaad onafhankelijk is van de keuze van de positief georiënteerde orthonormale basis. $\square$

We illustreren nu het nut van het vectorieel product door een aantal interessante eigenschappen aan te tonen. Zoals we kunnen zien, levert ook het combineren van het inproduct en het vectorproduct nuttige informatie op.

Stelling 9.2.5. *Onderstel dat $u, v, w \in \mathbb{R}^3$ en $a \in \mathbb{R}$, dan geldt:*

(i) $v \times w = -w \times v$;

(ii) $a(v \times w) = (av) \times w = v \times (aw)$;

(iii) $(u + v) \times w = u \times w + v \times w$;

(iv) $v \times w = 0 \iff v$ *en* w *zijn lineair afhankelijk;*

(v) [tripel-product] $u(v \times w) = (u \times v)w$;

(vi) $v \times w \perp v$ *en* $v \times w \perp w$*, met andere woorden,* $v(v \times w) = w(v \times w) = 0$;

(vii) [formule van Lagrange[1]] $u \times (v \times w) = (u \cdot w)v - (u \cdot v)w$;

(viii) [identiteit van Lagrange] $(v \times w)^2 = v^2w^2 - (vw)^2$;

(ix) *als* $v \neq 0$ *en* $w \neq 0$*, dan geldt:* $\|v \times w\| = \|v\| \cdot \|w\| \cdot \sin\angle(v, w)$;

(x) *als* $v \times w \neq 0$*, dan is* $\{v, w, v \times w\}$*, in deze volgorde, een positief georiënteerde basis voor* $\mathbb{R}^3$;

(xi) [Jacobi[2] identiteit] $u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0$;

(xii) $(u \times v) \times (u \times w) = u(v \times w) \cdot u$;

(xiii) *als* $u \cdot w = v \cdot w$ *én* $u \times w = v \times w$ *met* $w \neq 0$*, dan is* $u = v$*.*

Bewijs. Beschouw de standaardbasis $\mathcal{B}$ voor $\mathbb{R}^3$, en stel $u = (u_1, u_2, u_3)^t$, $v = (v_1, v_2, v_3)^t$ en $w = (w_1, w_2, w_3)^t$.

(i),(ii),(iii) Onmiddellijk uit (9.2) en basiseigenschappen van de determinant.

(iv) Uit (9.2) zien we dat $v \times w$ gelijk is aan 0 als en slechts als de drie (2×2)-minoren van de matrix $A = \begin{pmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$ nul zijn. Wegens Gevolg 6.4.5 is dit equivalent met het feit dat $\mathrm{rk}(A) < 2$, of nog, met het feit dat de twee rijen v en w lineair afhankelijk zijn.

[1]Joseph-Louis Lagrange (Giuseppe Lodovico Lagrangia) (1736–1813) was een wiskundige en astronoom van Italiaanse afkomst, die later in Frankrijk en Pruisen werkte. Samen met Leonhard Euler is hij één van de grootste wiskundigen van de 18e eeuw.

[2]Genoemd naar Carl Gustav Jacob Jacobi (1804–1851).

(v), (vi) Uit (9.2) volgt dat

$$u \cdot (v \times w) = \sum_{i=1}^{3} \begin{vmatrix} \delta_{1i} & \delta_{2i} & \delta_{3i} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} u_i = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix},$$

en analoog is

$$(u \times v) \cdot w = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

De gelijkheid (v) volgt nu omdat deze twee determinanten aan elkaar gelijk zijn via rijverwisselingen. Anderzijds zien we dat

$$v \cdot (v \times w) = \begin{vmatrix} v_1 & v_2 & v_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = 0\,,$$

en hieruit volgt (vi).

(vii) Dit is een beetje een vervelende berekening gebruik makend van (9.1). Stel $y = v \times w$, dan is $y = (y_1, y_2, y_3)^t$ met

$$\begin{aligned} y_1 &= v_2 w_3 - v_3 w_2\,, \\ y_2 &= v_3 w_1 - v_1 w_3\,, \\ y_3 &= v_1 w_2 - v_2 w_1\,. \end{aligned}$$

Stel nu $z = u \times y$, dan is $z = (z_1, z_2, z_3)^t$ met

$$\begin{aligned} z_1 &= u_2 y_3 - u_3 y_2 = u_2 v_1 w_2 - u_2 v_2 w_1 - u_3 v_3 w_1 + u_3 v_1 w_3\,, \\ z_2 &= u_3 y_1 - u_1 y_3 = u_3 v_2 w_3 - u_3 v_3 w_2 - u_1 v_1 w_2 + u_1 v_2 w_1\,, \\ z_3 &= u_1 y_2 - u_2 y_1 = u_1 v_3 w_1 - u_1 v_1 w_3 - u_2 v_2 w_3 + u_2 v_3 w_2\,. \end{aligned}$$

Stel ten slotte $x = (u \cdot w)v - (u \cdot v)w = (x_1, x_2, x_3)^t$, dan vinden we

$$\begin{aligned} x_1 &= (u_1 w_1 + u_2 w_2 + u_3 w_3) v_1 - (u_1 v_1 + u_2 v_2 + u_3 v_3) w_1\,, \\ x_2 &= (u_1 w_1 + u_2 w_2 + u_3 w_3) v_2 - (u_1 v_1 + u_2 v_2 + u_3 v_3) w_2\,, \\ x_3 &= (u_1 w_1 + u_2 w_2 + u_3 w_3) v_3 - (u_1 v_1 + u_2 v_2 + u_3 v_3) w_3\,; \end{aligned}$$

we stellen vast dat de vectoren z en x inderdaad aan elkaar gelijk zijn.

(viii) Hoewel de identiteit van Lagrange ook gemakkelijk rechtstreeks kan aangetoond worden met behulp van (9.1), zullen we ze trachten af te leiden uit reeds eerder bewezen eigenschappen. We vinden

$$(v \times w)^2 = (v \times w)(v \times w)$$

$$
\begin{aligned}
&= ((v \times w) \times v) \cdot w && \text{(wegens (v) met } u = v \times w) \\
&= -(v \times (v \times w)) \cdot w && \text{(wegens (i))} \\
&= \big((v \cdot v)w - (v \cdot w)v\big) \cdot w && \text{(wegens (vii))} \\
&= (v \cdot v)(w \cdot w) - (v \cdot w)(v \cdot w) \\
&= v^2 w^2 - (v \cdot w)^2 \,.
\end{aligned}
$$

(ix) Dit volgt onmiddellijk uit de identiteit van Lagrange, aangezien per definitie van de hoek $vw = \|v\| \cdot \|w\| \cdot \cos \angle(v, w)$, en bijgevolg

$$
\begin{aligned}
\|v \times w\|^2 &= \|v\|^2 \, \|w\|^2 - \|v\|^2 \, \|w\|^2 \cos^2 \angle(v, w) \\
&= \|v\|^2 \, \|w\|^2 \sin^2 \angle(v, w) \,.
\end{aligned}
$$

(x) Aangezien $v \times w \neq 0$ zijn v en w lineair onafhankelijk, en omdat $v \times w$ orthogonaal staat op zowel v als w is het drietal $\{v, w, v \times w\}$ lineair onafhankelijk en vormt bijgevolg een basis voor $\mathbb{R}^3$. Stel $v \times w = (y_1, y_2, y_3)^t$; dan wordt de transitiematrix van de standaardbasis naar deze basis gegeven door

$$
Q = \begin{pmatrix} v_1 & w_1 & y_1 \\ v_2 & w_2 & y_2 \\ v_3 & w_3 & y_3 \end{pmatrix},
$$

en de determinant van deze matrix is precies het tripel-product van de drie vectoren v, w en $v \times w$. We bekomen

$$
\det(Q) = (v \times w)(v \times w) = (v \times w)^2 > 0 \,,
$$

en bijgevolg is de basis $\{v, w, v \times w\}$ positief georiënteerd.

(xi) Om de Jacobi identiteit aan te tonen, volstaat het om drie keer de formule van Lagrange toe te passen en de vergelijkingen op te tellen:

$$
\begin{aligned}
u \times (v \times w) &= (u \cdot w)v - (u \cdot v)w \,, \\
v \times (w \times u) &= (v \cdot u)w - (v \cdot w)u \,, \\
w \times (u \times v) &= (w \cdot v)u - (w \cdot u)v \,.
\end{aligned}
$$

We zien dat de som van de drie rechterleden inderdaad 0 is.

(xii) We vertrekken opnieuw van de formule van Lagrange, en we maken gebruik van (v) en (vi):

$$
\begin{aligned}
(u \times v) \times (u \times w) &= \big((u \times v) \cdot w\big)u - \big((u \times v) \cdot u\big)w \\
&= \big(u \cdot (v \times w)\big)u \,.
\end{aligned}
$$

(xiii) Veronderstel dat $u \cdot w = v \cdot w$ en $u \times w = v \times w$. Dan is $(u - v) \cdot w = 0$ en $(u - v) \times w = 0$, en bijgevolg staat de vector $u - v$ enerzijds orthogonaal op w, terwijl anderzijds $u - v$ en w lineair afhankelijk zijn. Dit kan enkel als $u - v = 0$, en dus $u = v$. $\square$

Opmerkingen 9.2.6. (i) Het tripel-product van drie vectoren $u, v, w \in \mathbb{R}^3$ wordt ook soms genoteerd als $\{u\ v\ w\}$. Zoals we hebben aangetoond in het bewijs van Stelling 9.2.5(v), is het tripel-product van drie vectoren gelijk aan de determinant van de matrix waarvan de rijen (of kolommen) bestaan uit de coördinaten van de respectieve vectoren ten opzichte van een positief georiënteerde orthonormale basis. Uit de eigenschappen van determinanten kan men bijgevolg bijkomende eigenschappen afleiden over het tripel-product.

(ii) Uit het voorgaande volgt dat we het vectorieel product van twee vectoren v en w ook nog intrinsiek (dus zonder gebruik te maken van coördinaten) als volgt kunnen definiëren: Als de vectoren lineair afhankelijk zijn, is hun vectorieel product gelijk aan nul. In het andere geval stellen we $v \times w$ gelijk aan de unieke vector die voldoet aan:

(a) hij staat loodrecht op v en op w;

(b) zijn lengte is gelijk aan $\|v\| \cdot \|w\| \cdot \sin \angle(v, w)$;

(c) $\{v, w, v \times w\}$ vormt een positief georiënteerde basis van $\mathbb{R}^3$.

Ga zelf na dat dit de vector $v \times w$ uniek bepaalt, en dat dit overeenkomt met de eerdere definitie van het vectorieel product.

(iii) Noch het inproduct, noch het vectorieel product, voldoen aan een schrappingswet, m.a.w. noch uit $u \cdot w = v \cdot w$, noch uit $u \times w = v \times w$, kan men besluiten dat $u = v$ indien $w \neq 0$. Zie echter Stelling 9.2.5(xiii).

(iv) Eigenschap (i) samen met de Jacobi identiteit (xi) tonen aan dat het vectorieel product aan de vectorruimte $\mathbb{R}^3$ de structuur geven van een zogenaamde *Lie algebra*. In feite valt deze Lie algebra samen met de Lie algebra afkomstig van de Lie groep $\mathsf{SO}(3)$.

(v) De natuurlijke vraag dringt zich op of er gelijkaardige vectoriële producten kunnen gedefinieerd worden in $\mathbb{R}^n$ voor $n \neq 3$. In elk geval willen we dat een dergelijk vectorproduct, behalve aan de zeer natuurlijke eigenschappen (i)–(iii), ook nog voldoet aan (vi) en aan eigenschap (ix), of equivalent hiermee, aan de identiteit van Lagrange. Verrassend genoeg kan men aantonen dat een dergelijk vectorieel product enkel kan bestaan in dimensies 3 en 7! De diepere reden hiervoor is het feit dat zogenaamde genormeerde delingsalgebra's over $\mathbb{R}$ enkel bestaan in dimensies 1 ($\mathbb{R}$), 2 ($\mathbb{C}$), 4 ($\mathbb{H}$, de *quaternionen*) en 8 ($\mathbb{O}$, de *octonionen*). Binnen het raam van deze cursus kunnen we hier niet verder op ingaan.

(vi) Het is wel mogelijk om een willekeurige n-dimensionale vectorruimte V uit te breiden tot zijn zogenaamde *uitwendige algebra* $\Lambda(V)$, en in deze algebra kunnen we dus wel vectoren vermenigvuldigen met elkaar. Het product van twee vectoren v en w van V is nu geen vector, maar een *bi-vector* $v \wedge w$. Als $n = 3$ kunnen we met deze bi-vector een gewone vector laten overeenkomen door de *Hodge dualiteit* toe te passen, en het resultaat is precies het vectorieel product: $v \times w = *(v \wedge w)$.

9.3 Affiene deelruimten in $\mathbb{R}^n$

We beginnen nu onze studie van de zogenaamde affiene meetkunde; we herinneren de lezer aan Definitie 5.2.1 waar we affiene deelruimten hebben ingevoerd (in willekeurige K-vectorruimten). In het vervolg beperken we ons tot reële vectorruimten, waar de begrippen een vertrouwde betekenis hebben.

We zullen vooreerst de affiene deelruimten van $\mathbb{R}^n$ nader bestuderen en meetkundig interpreteren. Affiene deelruimten van enkele specifieke dimensies geven we meetkundige benamingen.

Definitie 9.3.1. Beschouw de vectorruimte $\mathbb{R}^n$ met $n \geq 2$.

(i) Een affiene deelruimte van dimensie 0 noemen we een *punt*.

(ii) Een affiene deelruimte van dimensie 1 noemen we een *rechte*.

(iii) Een affiene deelruimte van dimensie 2 noemen we een *vlak*.

(iv) Een affiene deelruimte van dimensie $n-1$ noemen we een *hypervlak*.

In $\mathbb{R}^3$ zijn vlakken en hypervlakken dezelfde objecten. We zeggen ook dat een hypervlak *codimensie* gelijk aan 1 heeft.

Definitie 9.3.2. Beschouw de vectorruimte $\mathbb{R}^n$. Een affiene deelruimte van $\mathbb{R}^n$ wordt ook wel een *Euclidische deelruimte* van $\mathbb{R}^n$ genoemd.

(i) Zij D een affiene deelruimte van $\mathbb{R}^n$. Wegens Lemma 5.2.4 bestaat er een unieke deelruimte $W \leq \mathbb{R}^n$ waarvoor $D = v + W$ voor een $v \in \mathbb{R}^n$. We noemen W de *geassocieerde vectordeelruimte* of de *onderliggende vectordeelruimte* van D en noteren $W =: D_0$. We zeggen dat D een *getranslateerde* is van W.

(ii) Zij $D = v + D_0$ met $D_0 \leq \mathbb{R}^n$. Elke vector in D noemen we een *plaatsvector* van D. Elke vector in D_0 noemen we een *richtingsvector* van D.

Aangezien iedere 1-dimensionale deelruimte van de vorm $\{\lambda w \mid \lambda \in \mathbb{R}\}$ is, kunnen we een rechte voorstellen als

$$D = \{v + \lambda w \mid \lambda \in \mathbb{R}\}$$

voor zekere $v \in \mathbb{R}^n$ en $w \in \mathbb{R}^n \setminus \{0\}$. Hierbij is v een plaatsvector van D, en is w een richtingsvector, zodat dus $D_0 = \{\lambda w \mid \lambda \in \mathbb{R}\}$.

Het is duidelijk dat een affiene deelruimte D van dimensie k uniek bepaald is door 1 plaatsvector en k lineair onafhankelijke richtingsvectoren.

Opmerking 9.3.3. (Voor de geïnteresseerde lezer) Hoewel we hier spreken over affiene deelruimten, hebben we in feite nog niet vermeld wat we verstaan onder de *affiene ruimte*. We willen hier niet al te diep ingaan op de algemene axiomatische definitie van deze structuren, maar wegens hun belang in ondermeer de fysica (bijvoorbeeld mechanica en kinematica) willen we hier toch even een aantal aspecten nader toelichten.

Een affiene ruimte is een verzameling E van elementen die we punten noemen. In tegenstelling tot vectorruimten geven we aan geen enkel punt een bijzondere rol; een affiene ruimte heeft dus geen "oorsprong". Wel veronderstellen we dat de *onderlinge ligging* van de punten gemodelleerd is op een vectorruimte V: elk koppel (a, b) van elementen van E bepaalt een element van V, dat we als $\overrightarrow{ab}$ noterenzodanig dat $\overrightarrow{ab} + \overrightarrow{bc} = \overrightarrow{ac}$. De elementen van V worden in deze context *vrije vectoren* van E genoemd. Meer bepaald is elke vrije vector een equivalentieklasse van equipollente[3] koppels van punten van E, en voor elke $a \in E$ bestaat er een unieke[4] $b \in E$ zodat $\overrightarrow{ab}$ gelijk is aan de gegeven (vrije) vector.

Wanneer we nu een specifiek punt $o \in E$ uitkiezen, bepaalt elk ander punt een vector, en omgekeerd bepaalt elke (vrije) vector $v \in V$ een uniek punt a van E zodat $v = \overrightarrow{oa}$. De *gepunte ruimte* E_o, i.e. de verzameling E samen met het uitverkoren punt o, krijgt op die manier de structuur van een vectorruimte, en die vectorruimte is isomorf met de onderliggende vectorruimte V.

In deze terminologie is een plaatsvector van een affiene deelruimte dus een *gebonden vector*, i.e. de vector hangt af van de keuze van de oorsprong o; anderzijds zijn de richtingsvectoren vrije vectoren, die niet afhangen van de keuze van o.

[3]Twee koppels van punten (a, b) en (c, d) worden *equipollent* genoemd als hun beeld onder θ gelijk is, m.a.w. als $\overrightarrow{ab} = \overrightarrow{cd}$.

[4]Dit volgt precies uit de hierboven vermelde bijectiviteit.

We bewijzen een meetkundige karakterisatie van affiene deelruimten.

Lemma 9.3.4. *Beschouw de vectorruimte $\mathbb{R}^n$. Een deelverzameling $D \subseteq \mathbb{R}^n$ is een affiene deelruimte van $\mathbb{R}^n$ als en slechts als*

$$y + a(z - y) \in D \quad \textit{voor alle } y, z \in D \textit{ en alle } a \in \mathbb{R}\,. \tag{9.3}$$

Meetkundig zegt dit dus precies dat een deelverzameling $D \subseteq \mathbb{R}^n$ een affiene deelruimte is, dan en slechts als elke rechte door 2 punten van D volledig in D ligt.

Bewijs. Veronderstel eerst dat $D \subseteq \mathbb{R}^n$ een affiene deelruimte is van $\mathbb{R}^n$; per definitie is dan $D = W + v$ voor een zekere deelruimte W en een $v \in \mathbb{R}^n$. Neem $y, z \in D$ en $a \in \mathbb{R}$ willekeurig. Stel $y = w_y + v$ en $z = w_z + v$ met $w_y, w_z \in W$; dan is

$$\begin{aligned} y + a(z - y) &= (1 - a)(w_y + v) + a(w_z + v) \\ &= (1 - a)w_y + aw_z + v \in W + v = D\,, \end{aligned}$$

en dus is (9.3) inderdaad voldaan.

Veronderstel omgekeerd dat (9.3) geldt. Neem $y \in D$ willekeurig; we zullen aantonen dat de verzameling $D - y := \{d - y \mid d \in D\}$ een (vector)deelruimte is van $\mathbb{R}^n$, en hieruit volgt dan dat D een affiene deelruimte is. Neem dus $v, w \in D - y$ en $a \in \mathbb{R}$ willekeurig; we moeten enerzijds aantonen dat $v + w \in D - y$, en anderzijds dat $av \in D - y$. Stel dus $v = d_v - y$ en $w = d_w - y$ met $d_v, d_w \in D$; dan is

$$av = \big(y + a(d_v - y)\big) - y \in D - y\,.$$

Anderzijds is

$$(v + w)/2 = (d_v - y + d_w - y)/2 = \big(d_v + \tfrac{1}{2}(d_w - d_v)\big) - y \in D - y\,,$$

en omdat we reeds hebben aangetoond dat $D - y$ gesloten is onder scalaire vermenigvuldiging, volgt hieruit dat ook $v + w \in D - y$. □

In Lemma 5.3.1 toonden we aan dat de oplossing van een niet-strijdig stelsel vergelijkingen steeds een affiene deelruimte is. In Stelling 9.3.6 zullen we aantonen dat het omgekeerde ook geldig is. Het volgend lemma bevat de essentie van het argument. We herinneren aan de definitie van de rijruimte $R(A)$ van een matrix A, ingevoerd in Definitie 5.1.1(i), en aan de definitie van het orthogonaal complement van een deelruimte, ingevoerd in Definitie 8.2.1(iv).

Lemma 9.3.5. *Zij* $A \in M_{m,n}(\mathbb{R})$. *Dan is* $\ker(L_A) = R(A)^\perp$.

Bewijs. Noteer de rijen van A als $R_1, \ldots, R_m$. Voor elke $v \in \mathbb{R}^n$ is Av gelijk aan de matrix $(R_1 v, \ldots, R_m v)^t$, en dus is $Av = 0$ als en slechts als $R_i v = 0$ voor alle $i \in \{1, \ldots, m\}$. Omdat per definitie $R(A) = \text{span}(R_1, \ldots, R_m)$, is dit ook equivalent met $Rv = 0$ voor alle $R \in R(A)$, wat volgens[5] het standaard inproduct op $\mathbb{R}^n$ equivalent is met $v \in R(A)^\perp$. □

Stelling 9.3.6. *Zij* D *een affiene deelruimte in de Euclidische ruimte* $\mathbb{R}^n$*, en stel* $\dim D = d$. *Dan geldt:*

(i) *D is gelijk aan de oplossingsverzameling van een stelsel van $n-d$ lineaire vergelijkingen in n onbekenden over $\mathbb{R}$.*

(ii) *Stel dat $D = \{X \in \mathbb{R}^n \mid AX = w\}$, dan is $D_0 = \{X \in \mathbb{R}^n \mid AX = 0\}$.*

Bewijs. Zij $D = v + D_0$ met $D_0 \leq \mathbb{R}^n$ de geassocieerde vectordeelruimte en $v \in \mathbb{R}^n$. Wegens Lemma 8.2.5 is $\dim D_0^\perp = n - d$. Kies een basis $\mathcal{B} = \{b_1, \ldots, b_{n-d}\}$ voor $D_0^\perp \leq \mathbb{R}^n$ en beschouw de matrix $A \in M_{n-d,n}(\mathbb{R})$ waarvan de *rijen* precies de basisvectoren $b_1, \ldots, b_{n-d} \in \mathbb{R}^n$ zijn. Wegens Lemma 9.3.5 en Gevolg 8.2.7 is $D_0 = \ker(L_A) = \{X \in \mathbb{R}^n \mid AX = 0\}$.

Stel nu $w := Av \in \mathbb{R}^m$. Aangezien $D = v + D_0$ geldt nu, voor elke $X \in \mathbb{R}^n$,

$$X \in D \iff X - v \in D_0 \iff A(X - v) = 0 \iff Ax = w,$$

en dus is $D = \{X \in \mathbb{R}^n \mid AX = w\}$ de oplossingsverzameling van het stelsel $AX = w$ met $n - d$ vergelijkingen en n onbekenden. □

Opmerking 9.3.7. Het bewijs van voorgaande stelling geeft ons ook een methode om, gegeven een affiene deelruimte, een stelsel te bepalen waarvan deze affiene ruimte de oplossingsverzameling is. Bemerk dat dit stelsel zeker niet uniek is.

Als bijzonder geval kunnen we enerzijds de hypervlakken een eenvoudige beschrijving geven, en anderzijds een interessante meetkundige interpretatie bekomen van de voorgaande stelling.

Gevolg 9.3.8. (i) *Zij H een hypervlak in $\mathbb{R}^n$. Dan bestaan er reële getallen $a_1, \ldots, a_n, b \in \mathbb{R}$ zodat*

$$H = \{(x_1, \ldots, x_n)^t \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n + b = 0\}.$$

[5]Merk op dat de rijruimte $R(A)$ gezien wordt als deelruimte van $\mathbb{R}^n$ *na transponeren*, zodat het standaard inproduct $\langle v, w \rangle = v^t w$ hier de gedaante $\langle R, v \rangle = Rv$ krijgt.

(ii) *Zij D een affiene deelruimte van dimensie d. Dan bestaan er $n-d$ hypervlakken $H_1, \dots, H_{n-d}$ zodanig dat $D = H_1 \cap \dots \cap H_{n-d}$.*

Bewijs. (i) Uit Stelling 9.3.6 volgt dat een $(n-1)$-dimensionale affiene deelruimte de oplossingsverzameling is van $n-(n-1) = 1$ lineaire vergelijking.

(ii) Aangezien een d-dimensionale affiene deelruimte de oplossingsverzameling van $n-d$ lineaire vergelijkingen is, volgt het gestelde uit (i). □

Het is duidelijk dat er door twee verschillende punten in $\mathbb{R}^3$ een rechte gaat. We veralgemenen dit feit naar k punten in $\mathbb{R}^n$.

Lemma 9.3.9. *Beschouw k punten $p_1, \dots, p_k$ in $\mathbb{R}^n$.*

(i) *Voor alle $i, j = 1, \dots, k$ geldt*

$$p_i + \operatorname{span}(p_1 - p_i, \dots, p_k - p_i) = p_j + \operatorname{span}(p_1 - p_j, \dots, p_k - p_j).$$

(ii) *Stel dat D een affiene deelruimte is met $p_1, \dots, p_k \in D$, dan geldt er dat $p_1 + \operatorname{span}(p_2 - p_1, \dots, p_k - p_1) \subseteq D$.*

Bewijs. (i) We hebben

$$\operatorname{span}(p_1 - p_i, \dots, p_k - p_i) = \operatorname{span}(p_1 - p_j, \dots, p_k - p_j)$$

aangezien elke vector $p_\ell - p_i$ ook in het rechterlid bevat is, en elke vector $p_\ell - p_j$ ook in het linkerlid bevat is. Aangezien $p_i - p_j$ ook tot deze deelruimte W behoort, volgt nu uit Lemma 5.2.4 dat $p_i + W = p_j + W$.

(ii) Als $p_1, \dots, p_k \in D$, dan is $D = p_1 + D_0$. In het bijzonder hebben we $p_2 - p_1, \dots, p_k - p_1 \in D_0$, dus is ook $\operatorname{span}(p_2 - p_1, \dots, p_k - p_1) \subseteq D_0$. Hieruit volgt het gestelde. □

Definitie 9.3.10. Beschouw k punten $p_1, \dots, p_k$ in $\mathbb{R}^n$. Dan noemen we de affiene deelruimte

$$p_1 + \operatorname{span}(p_2 - p_1, \dots, p_k - p_1)$$

de *affiene deelruimte bepaald door* of *opgespannen door* de punten $p_1, \dots, p_k$. Uit Lemma 9.3.9 volgt dat dit de kleinste affiene deelruimte is die de punten $p_1, \dots, p_k$ bevat.

Opmerking 9.3.11. We hebben $\dim(p_1 + \operatorname{span}(p_2 - p_1, \dots, p_k - p_1)) \leq k - 1$; de gelijkheid geldt als en slechts als de verzameling van richtingsvectoren $\{p_2 - p_1, \dots, p_k - p_1\}$ lineair onafhankelijk is.

We definiëren wanneer twee affiene deelruimten parallel zijn:

Definitie 9.3.12. (i) Zij $D = v + D_0$ een affiene deelruimte van V. Een vector $w \in V$ noemen we *parallel* aan D, als en slechts als $w \in D_0$; we noteren dit als $w \parallel D$.

(ii) Beschouw twee affiene deelruimten $D = v + D_0$ en $D' = v' + D'_0$ van V. Dan noemen we D en D' parallel aan elkaar als en slechts als $D_0 \leq D'_0$ of $D'_0 \leq D_0$. We noteren dit met $D \parallel D'$.
In het bijzonder zijn twee affiene deelruimten D en D' van dezelfde dimensie parallel aan elkaar als en slechts dan als $D_0 = D'_0$.

Anderzijds veralgemenen we ook het begrip van orthogonale vectoren, dat we hebben ingevoerd in Definitie 8.2.1, tot willekeurige affiene deelruimten van de Euclidische deelruimte $\mathbb{R}^n$.

Definitie 9.3.13. (i) Zij $D = v + D_0$ een affiene deelruimte van $\mathbb{R}^n$. We zeggen dat een vector $w \in \mathbb{R}^n$ *orthogonaal op* D staat als en slechts als w orthogonaal staat op elke vector van D_0; we noteren dit als $w \perp D$.

(ii) Beschouw twee affiene deelruimten $D = v + D_0$ en $D' = v' + D'_0$ van V. Dan staan D en D' orthogonaal op elkaar als en slechts als elke vector van D_0 loodrecht staat op elke vector van D'_0. We noteren dit met $D \perp D'$.

Stel dat $D \perp D'$, dan is noodzakelijk $D_0 \cap D'_0 = \{0\}$. Immers, in de Euclidische ruimte is de nulvector het enige element dat orthogonaal op zichzelf staat.

9.4 Hypervlakken in $\mathbb{R}^n$

Een hypervlak in $\mathbb{R}^n$ is een affiene deelruimte van dimensie $n - 1$.

Definitie 9.4.1. Zij H een hypervlak in $\mathbb{R}^n$. Uit Gevolg 9.3.8 volgt dat H van de vorm

$$H = \{(x_1, \dots, x_n)^t \in \mathbb{R}^n \mid a_1x_1 + \cdots + a_nx_n + b = 0\}$$

is, voor zekere vaste reële getallen $a_1, \dots, a_n, b \in \mathbb{R}$. Wanneer we de vector $(a_1, \dots, a_n)^t \in \mathbb{R}^n$ noteren als $\mathfrak{n}$, kunnen we de vergelijking van H herschrijven als

$$H = \{v \in \mathbb{R}^n \mid \mathfrak{n} \cdot v + b = 0\}.$$

Elke niet-nul vector $\mathfrak{m}$ met $\mathfrak{m} \perp H$ (of equivalent, $\mathfrak{m} \perp H_0$) noemen we een *normaalvector* voor H.

Zij H gegeven door de vergelijking $\mathfrak{n} \cdot v + b = 0$. Er volgt dat $\mathfrak{n} \perp H_0$, en aangezien $\dim(H_0)^\perp = n - (n-1) = 1$, volgt dat iedere normaalvector van H een scalair veelvoud van $\mathfrak{n}$ is.

Lemma 9.4.2. *Beschouw in $\mathbb{R}^n$ een niet-nul vector $\mathfrak{n} = (a_1, \ldots, a_n)^t$, en een willekeurig punt $p = (p_1, \ldots, p_n)^t$. Dan is er een uniek hypervlak met normaalvector $\mathfrak{n}$ dat het punt p bevat.*

Bewijs. Alle hypervlakken met normaalvector $\mathfrak{n}$ hebben de gedaante

$$H = \{(x_1, \ldots, x_n)^t \in \mathbb{R}^n \mid a_1x_1 + \cdots + a_nx_n + b = 0\}$$

voor een zekere $b \in \mathbb{R}$. Het is nu duidelijk dat de voorwaarde $p \in H$ het getal b uniek bepaalt; we krijgen $b = -(a_1p_1 + \cdots + a_np_n)$, en bijgevolg kunnen we de vergelijking van H herschrijven als

$$H = \{(x_1, \ldots, x_n)^t \in \mathbb{R}^n \mid a_1(x_1 - p_1) + \cdots + a_n(x_n - p_n) = 0\},$$

of nog, in vectoriële notatie,

$$H = \{v \in \mathbb{R}^n \mid \mathfrak{n} \cdot (v - p) = 0\}. \qquad \square$$

Lemma 9.4.3. *Beschouw in $\mathbb{R}^n$ twee hypervlakken*

$$\begin{aligned} H &= \{v \in \mathbb{R}^n \mid \mathfrak{n} \cdot v + b = 0\} \text{ en} \\ H' &= \{v \in \mathbb{R}^n \mid \mathfrak{n}' \cdot v + b' = 0\}. \end{aligned}$$

Dan is $H \parallel H'$ als en slechts als $\mathfrak{n}$ en $\mathfrak{n}'$ lineair afhankelijk zijn.

Bewijs. Omdat $\dim(H) = \dim(H')$, is $H \parallel H'$ als en slechts als $H_0 = H'_0$, waarbij $H_0 = \{v \in \mathbb{R}^n \mid \mathfrak{n} \cdot v = 0\}$ en $H'_0 = \{v \in \mathbb{R}^n \mid \mathfrak{n}' \cdot v = 0\}$. Als $\mathfrak{n}$ en $\mathfrak{n}'$ evenredig zijn, dan is uiteraard $H_0 = H'_0$. Omgekeerd, als $H_0 = H'_0$ dan zijn $\mathfrak{n}$ en $\mathfrak{n}'$ normaalvectoren van hetzelfde hypervlak, en bijgevolg zijn ze evenredig aan elkaar. $\square$

Als $n > 2$, dan kunnen twee hypervlakken nooit orthogonaal op elkaar staan. Stel immers dat $H \perp H'$; dan volgt uit de dimensiestelling dat

$$\dim(H_0 + H'_0) = (n-1) + (n-1) - 0 = 2n - 2 > n,$$

een tegenstrijdigheid. Nochtans definiëren we hieronder orthogonale hypervlakken; we benadrukken dat deze definitie *niet* conform is met Definitie 9.3.13, maar dat er precies door deze beschouwingen geen verwarring mogelijk is. (In het geval $n = 2$ blijken beide begrippen met elkaar samen te vallen.)

Definitie 9.4.4. We zeggen dat twee hypervlakken H en H' van $\mathbb{R}^n$ *orthogonaal op elkaar* staan, als en slechts als hun normaalvectoren orthogonaal op elkaar staan.

Aangezien normaalvectoren op een evenredigheidsfactor na uniek bepaald zijn, is deze definitie onafhankelijk van de gekozen normaalvectoren.

We veralgemenen de definitie van orthogonale hypervlakken en definiëren de hoek tussen twee hypervlakken.

Definitie 9.4.5. We definiëren de *hoek* tussen twee hypervlakken als de *kleinste* hoek ingesloten door hun normaalvectoren.

Merk op dat de hoek $\alpha = \angle(\mathfrak{n}, \mathfrak{n}')$ onveranderd blijft als we $\mathfrak{n}$ of $\mathfrak{n}'$ vervangen door een positief veelvoud, maar verandert in $\pi - \alpha$ als we $\mathfrak{n}$ of $\mathfrak{n}'$ vervangen door een negatief veelvoud. (Indien we beide vervangen door een negatief veelvoud bekomen we uiteraard opnieuw α.) De kleinste hoek is dus ofwel α, ofwel $\pi - \alpha$, en is bijgevolg steeds bevat in het interval $[0, \pi/2]$. Indien de hoek 0 is, zijn de hypervlakken parallel aan elkaar; indien de hoek $\pi/2$ is, zijn ze orthogonaal aan elkaar.

Lemma 9.4.6. *Beschouw in $\mathbb{R}^n$ twee hypervlakken H en H', met respectieve normaalvectoren $\mathfrak{n}$ en $\mathfrak{n}'$. Dan is*

$$\angle(H, H') = \arccos \frac{|\mathfrak{n} \cdot \mathfrak{n}'|}{\|\mathfrak{n}\| \cdot \|\mathfrak{n}'\|} .$$

Bewijs. Aangezien $\arccos(-t) = \pi - \arccos(t)$ voor alle $t \in [-1, 1]$, en arccos de waarden $[0, \pi/2]$ aanneemt op het interval $[0, 1]$, volgt deze formule uit de definities. □

9.5 Toepassingen

In deze sectie bekijken we verscheidene toepassingen van de voorgaande secties. We bestuderen een aantal resultaten in verband met de onderlinge ligging van rechten, vlakken en hypervlakken in $\mathbb{R}^n$.

Lemma 9.5.1. *Zij H een hypervlak met normaalvector $\mathfrak{n}$, en L een rechte met richtingsvector r. Dan is $L \perp H$ als en slechts als $\mathfrak{n}$ en r lineair afhankelijk zijn.*

Bewijs. Per definitie is $L \perp H$ als en slechts als elke vector van L_0 orthogonaal is met elke vector van H_0. De vectoren van L_0 zijn precies de scalaire veelvouden van r, en een vector staat loodrecht op $H_0 = \{x \in \mathbb{R}^n \mid x \cdot \mathfrak{n} = 0\}$ als en slechts als hij evenredig is met een (willekeurige) normaalvector van H. Dit bewijst het gestelde. □

Merk op dat in $\mathbb{R}^3$ geldt dat twee vectoren lineair afhankelijk zijn als en slechts als hun vectorieel product 0 is. In $\mathbb{R}^3$ bekomen we dus dat $L \perp H$ als en slechts als $\mathfrak{n} \times r = 0$.

In Definitie 8.1.1 definieerden we de afstand tussen twee vectoren in een inproduct-ruimte. We willen deze definitie nu uitbreiden tot de afstand tussen twee willekeurige affiene deelruimten.

Definitie 9.5.2. (i) Zij D en D' twee affiene deelruimten in $\mathbb{R}^n$. We definieren de afstand tussen D en D' als de kortst mogelijke afstand tussen een punt van D en een punt van D'. We noteren dit als $\operatorname{dist}(D, D')$.

(ii) Zij D een affiene deelruimte en zij $p \in \mathbb{R}^n \setminus D$ een punt niet op D gelegen. Een *loodlijn* uit p op D is een rechte L door p die loodrecht staat op D en die D snijdt. We tonen zo dadelijk aan dat er voor elke p en D een unieke dergelijke rechte bestaat, zodat we kunnen spreken van *de* loodlijn uit p op D.

In de volgende stelling maken we gebruik van de orthogonale projectie op een deelruimte; zie Definitie 8.2.8.

Stelling 9.5.3. *Zij $D = q + D_0$ een affiene deelruimte en zij $p \in \mathbb{R}^n \setminus D$ een punt niet op D gelegen. Stel $r := \operatorname{proj}_{D_0^\perp}(p - q)$ de orthogonale projectie van $p - q$ op $D_0^\perp$, en stel $L := \{p + \lambda r \mid \lambda \in \mathbb{R}\}$. Dan geldt:*

(i) *De rechte L is de unieke loodlijn van p op D, en $L \cap D = \{p - r\}$.*

(ii) $\operatorname{dist}(p, D) = \operatorname{dist}(p, L \cap D) = \|r\|$.

Bewijs. (i) Aangezien L een richtingsvector heeft in $D_0^\perp$, staat L loodrecht op D. We bepalen nu $L \cap D$. Merk op dat

$$p - q = \operatorname{proj}_{D_0}(p - q) + r.$$

We hebben nu enerzijds dat $q + \operatorname{proj}_{D_0}(p - q) \in q + D_0 = D$, terwijl anderzijds $q + \operatorname{proj}_{D_0}(p - q) = p - r \in L$. Aangezien $p \notin D$ is $L \not\subseteq D$, en we besluiten dat $L \cap D = \{p - r\}$. Dus L is een loodlijn van p op D.

Om de uniciteit te bewijzen, nemen we aan dat M een willekeurige loodlijn is van p op D. Stel $M \cap D = \{z\}$. Dan is $M \subseteq p + D_0^\perp$, zodat $z \in (p + D_0^\perp) \cap D$. Maar dan is $D = z + D_0$ en $p + D_0^\perp = z + D_0^\perp$, waaruit volgt dat

$$(p + D_0^\perp) \cap D = (z + D_0^\perp) \cap (z + D_0) = z + (D_0^\perp \cap D_0) = \{z\}.$$

Hieruit volgt dat z niet afhangt van de keuze van M, en bijgevolg is M de unieke rechte door p en dit punt z.

(ii) Zij $w \in D \setminus \{q + \text{proj}_{D_0}(p-q)\}$; dan is $w - q \in D_0 \setminus \{\text{proj}_{D_0}(p-q)\}$. We passen Stelling 8.2.10 toe en bekomen

$$\text{dist}(p-q, w-q) > \text{dist}(p-q, \text{proj}_{D_0}(p-q)).$$

Dit is equivalent met

$$\text{dist}(p, w) > \text{dist}(p, q + \text{proj}_{D_0}(p-q)) = \text{dist}(p, L \cap D),$$

waaruit volgt dat $\text{dist}(p, D) = \text{dist}(p, L \cap D)$. Merk ten slotte op dat $\text{dist}(p, L \cap D) = \text{dist}(p, p - r) = \|r\|$. □

We bestuderen nu het specifieke geval van de loodlijn uit een punt op een hypervlak en de afstand van een punt tot een hypervlak in $\mathbb{R}^n$.

Stelling 9.5.4. *Zij H een hypervlak met vergelijking $\mathfrak{n} \cdot v + b = 0$, en p een punt in $\mathbb{R}^n \setminus H$. Dan geldt:*

(i) *De loodlijn vanuit p op H wordt gegeven door*

$$L = \{p + \lambda\mathfrak{n} \mid \lambda \in \mathbb{R}\}.$$

(ii) *De afstand van p tot H is gelijk aan*

$$\text{dist}(p, H) = \frac{|\mathfrak{n} \cdot p + b|}{\|\mathfrak{n}\|}.$$

Bewijs. (i) Zij L de loodlijn vanuit p op H. Aangezien $L \perp H$, volgt uit Lemma 9.5.1 dat $\mathfrak{n}$ een richtingsvector is voor L, en aangezien uiteraard p een plaatsvector is, wordt een willekeurige vector van L gegeven door $v = p + \lambda\mathfrak{n}$ met $\lambda \in \mathbb{R}$.

(ii) We passen Stelling 9.5.4(ii) toe. Zij $q \in H$ willekeurig; dan is $H = q + H_0$, en $q \cdot \mathfrak{n} + b = 0$. Merk op dat $H_0^\perp = \mathbb{R}\mathfrak{n}$. We kunnen de projectie op $H_0^\perp$ nu bepalen met behulp van Lemma 8.2.9(iii), en we vinden

$$\begin{aligned}\text{dist}(p, H) &= \|\text{proj}_{\mathfrak{n}}(p-q)\| \\ &= \left\| \frac{(p-q) \cdot \mathfrak{n}}{\mathfrak{n} \cdot \mathfrak{n}} \mathfrak{n} \right\| = \frac{|p \cdot \mathfrak{n} - q \cdot \mathfrak{n}|}{\|\mathfrak{n}\|} = \frac{|p \cdot \mathfrak{n} + b|}{\|\mathfrak{n}\|}.\end{aligned}$$ □

We bestuderen nu de loodlijn uit een punt op een rechte en de afstand van een punt tot een rechte in $\mathbb{R}^n$. We bekijken eerst even het geval $n = 3$ apart, en geven nadien het resultaat voor algemene n.

Stelling 9.5.5. *Zij p een punt in $\mathbb{R}^3$, en zij M een rechte met plaatsvector q en richtingsvector r. Veronderstel dat $p \notin M$. Dan is*

$$\operatorname{dist}(p, M) = \frac{\|r \times (p-q)\|}{\|r\|}.$$

Bewijs. Zij y de projectie van p op de rechte M. Dan is $r \times (p-q) = r \times ((p-y)+(y-q)) = r \times (p-y)$, en uit Stelling 9.2.5(ix) halen we dat

$$\|r \times (p-y)\| = \|r\| \cdot \|p-y\| \cdot \sin(\pi/2) = \|r\| \cdot \|p-y\|,$$

zodat

$$\operatorname{dist}(p, M) = \operatorname{dist}(p, y) = \|p-y\| = \frac{\|r \times (p-q)\|}{\|r\|}. \qquad \square$$

Stelling 9.5.6. *Zij p een punt in $\mathbb{R}^n$ en M een rechte met plaatsvector q en richtingsvector r. Veronderstel dat $p \notin M$. Dan is de loodlijn L vanuit p op M de rechte door de punten p en $q + \frac{r\cdot(p-q)}{\|r\|^2}r$. De afstand van p tot M is gegeven door*

$$\operatorname{dist}(p, M) = \frac{\sqrt{\|r\|^2\,\|p-q\|^2 - \big(r\cdot(p-q)\big)^2}}{\|r\|}.$$

Bewijs. We passen opnieuw Stelling 9.5.4 toe. We hebben $M = q + M_0$, met $M_0 = \mathbb{R}r$, en we gebruiken opnieuw Lemma 8.2.9(ii). Hieruit halen we

$$y := q + \operatorname{proj}_{M_0}(p-q) = q + \left((p-q)\cdot\frac{r}{\|r\|}\right)\frac{r}{\|r\|} = q + \frac{r\cdot(p-q)}{\|r\|^2}r.$$

Ten slotte vinden we

$$\begin{aligned}
\operatorname{dist}(p, M) &= \operatorname{dist}(p, y) \\
&= \left\| q - p + \frac{r\cdot(p-q)}{\|r\|^2} r \right\| \\
&= \sqrt{\|q-p\|^2 + \left(\frac{r\cdot(p-q)}{\|r\|^2}\right)^2 \|r\|^2 + 2\frac{r\cdot(p-q)}{\|r\|^2}\,(q-p)\cdot r} \\
&= \frac{\sqrt{\|r\|^2\,\|p-q\|^2 - \big(r\cdot(p-q)\big)^2}}{\|r\|}. \qquad \square
\end{aligned}$$

Opmerking 9.5.7. In het geval $n = 3$ herleidt de formule uit Stelling 9.5.6 zich tot deze uit Stelling 9.5.5 door middel van de identiteit van Lagrange; zie Stelling 9.2.5(viii).

In feite kunnen we deze observatie gebruiken om een alternatief bewijs te geven voor Stelling 9.5.6. Wegens Stelling 9.5.5 weten we immers dat het resultaat geldig is in $\mathbb{R}^3$. Maar een punt en een willekeurige rechte in $\mathbb{R}^n$ (met $n \geq 3$) kunnen steeds ingebed worden in een gemeenschappelijk vlak, zodat we de restrictie kunnen nemen tot een 3-dimensionale Euclidische deelruimte van $\mathbb{R}^n$ en daarin het gekende resultaat kunnen toepassen. Voor $\mathbb{R}^2$ moeten we net omgekeerd eerst een extra dimensie toevoegen om het resultaat in $\mathbb{R}^3$ te kunnen toepassen.

In de drie-dimensionale Euclidische ruimte $\mathbb{R}^3$ hebben twee niet-parallelle (hyper)vlakken de bijzondere eigenschap dat ze snijden in een rechte, waarvan de richting dus bepaald is door één vector. We vinden die vector als volgt terug:

Lemma 9.5.8. *Zij H en H' twee niet-parallelle vlakken in $\mathbb{R}^3$, met respectieve normaalvectoren $\mathfrak{n}$ en $\mathfrak{n}'$. Dan is $H \cap H'$ een rechte met richtingsvector $\mathfrak{n} \times \mathfrak{n}'$.*

Bewijs. Aangezien H en H' niet parallel zijn aan elkaar, snijden ze noodzakelijk in een rechte $L = H \cap H'$. Omdat $\mathfrak{n} \perp H$, is in het bijzonder $\mathfrak{n} \perp L$, en analoog $\mathfrak{n}' \perp L$. Aangezien ook $\mathfrak{n} \times \mathfrak{n}'$ orthogonaal staat op zowel $\mathfrak{n}$ als $\mathfrak{n}'$, en de richting van $\mathfrak{n} \times \mathfrak{n}'$ uniek bepaald is door deze eigenschap (zie Opmerking 9.2.6(ii)), volgt hieruit dat $\mathfrak{n} \times \mathfrak{n}' \parallel L$. $\square$

Ten slotte beschouwen we de situatie waarin twee rechten in $\mathbb{R}^3$ elkaar kruisen, d.w.z. ze zijn noch parallel, noch snijdend. Een *gemeenschappelijke loodlijn* is een rechte die deze twee rechten loodrecht snijdt. We tonen aan dat een dergelijke gemeenschappelijke loodlijn bestaat en uniek is, en we bepalen de vergelijking ervan. Dit zal ons ook in staat stellen om de afstand tussen twee kruisende rechten te bepalen.

Stelling 9.5.9. *Beschouw twee kruisende rechten L en M in $\mathbb{R}^3$. Dan is er een unieke rechte N die zowel L als M orthogonaal snijdt. Als L en M gegeven zijn door*

$$\begin{aligned} L &= \{p + \lambda r \mid \lambda \in \mathbb{R}\}, \\ M &= \{q + \lambda s \mid \lambda \in \mathbb{R}\}, \end{aligned}$$

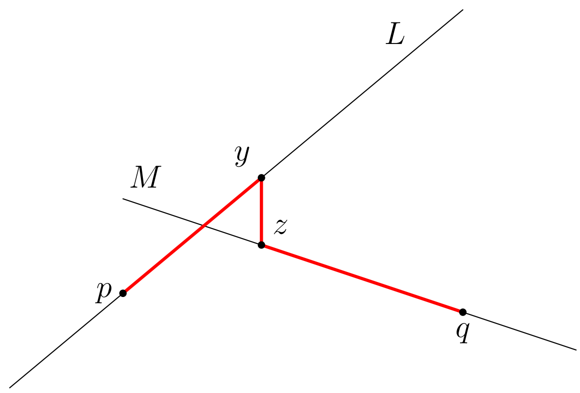
dan heeft N richtingsvector $r \times s$, en de snijpunten van N met L en M zijn respectievelijk gegeven door

$$y = p + \frac{(q-p) \cdot \big(s \times (r \times s)\big)}{\|r \times s\|^2}\, r \qquad \text{en}$$

$$z = q + \frac{(q-p)\cdot\big(r\times(r\times s)\big)}{\|r\times s\|^2}\, s\,.$$

Bovendien is de afstand van L tot M gelijk aan

$$\mathrm{dist}(L,M) = \mathrm{dist}(y,z) = \frac{|(q-p)\cdot(r\times s)|}{\|r\times s\|}\,.$$



Bewijs. Aangezien L en M respectieve richtingsvectoren r en s hebben, waarbij r en s lineair onafhankelijk zijn omdat L en M niet parallel zijn, moet een gemeenschappelijke loodlijn in elk geval richtingsvector $u = r\times s$ $(\neq 0)$ hebben.

Beschouw nu het vlak α door L en parallel met u, en het vlak β door M en parallel met u. Merk op dat α richtingsvectoren r en u heeft, en dat β richtingsvectoren s en u heeft. Omdat $\{r,s,u\}$ een basis is, volgt hieruit dat $\alpha \nparallel \beta$, en dus is $\alpha\cap\beta$ een rechte, die uiteraard u als richtingsvector heeft, en die zowel L als M snijdt. Omgekeerd moet elke gemeenschappelijke loodlijn van L en M bevat zijn in zowel α als β, en dit toont dus aan dat $\alpha\cap\beta$ de unieke gemeenschappelijke loodlijn N van L en M is.

We zoeken nu een expliciete vergelijking voor α. Omdat α richtingsvectoren r en $r\times s$ heeft, is een normaalvector van α gegeven door $n_\alpha = r\times(r\times s)$. We kunnen dus α voorstellen door de vectorvergelijking

$$\alpha = \big\{v\in\mathbb{R}^3 \mid (v-p)\cdot\big(r\times(r\times s)\big) = 0\big\}\,. \tag{9.4}$$

Het snijpunt z van N met M is nu precies gelijk aan $\alpha\cap M$. We beschouwen dus een punt $z = q+\lambda s$ van M, en we drukken uit dat $z\in\alpha$. We bekomen

$$\lambda = \frac{(p-q)\cdot\big(r\times(r\times s)\big)}{s\cdot\big(r\times(r\times s)\big)}\,.$$

Stelling 9.2.5(i) en (v) leert ons dat

$$s \cdot \big(r \times (r \times s)\big) = (s \times r) \cdot (r \times s) = -\|r \times s\|^2 ,$$

en we bekomen de gezochte uitdrukking voor het punt z. Op analoge wijze bekomen we de uitdrukking voor het snijpunt y van N met L.

Ten slotte berekenen we de afstand $\mathrm{dist}(L, M)$. De kortste afstand tussen een punt van L en een punt van M is duidelijkerwijze gegeven door de afstand langs de gemeenschappelijke loodlijn, dus $\mathrm{dist}(L, M) = \mathrm{dist}(y, z)$. In principe kan men via de gevonden uitdrukkingen voor y en z nu $\mathrm{dist}(y, z)$ berekenen. Een efficiëntere methode gaat als volgt. We kunnen de vector $q-p$ opsplitsen in een stuk langs L, een stuk langs N en een stuk langs M. Meer bepaald schrijven we $q-p = (q-z)+(z-y)+(y-p)$, en merk op dat $(q-z)\cdot(r \times s)$ en $(y-p) \cdot (r \times s)$ beide 0 zijn. Bijgevolg is

$$(q - p) \cdot (r \times s) = (z - y) \cdot (r \times s) .$$

Schrijf nu $z - y = \lambda(r \times s)$, en gebruik de gelijkheid in de Ongelijkheid van Cauchy–Schwarz (Stelling 8.1.7); dan vinden we

$$\mathrm{dist}(y, z) = \|z - y\| = \frac{|(z - y) \cdot (r \times s)|}{\|r \times s\|} = \frac{|(q - p) \cdot (r \times s)|}{\|r \times s\|} ,$$

en we vinden de gezochte formule terug. □

Opmerking 9.5.10. Indien we geïnteresseerd zijn in een expliciete vergelijking voor N, kunnen we zeer eenvoudig te werk gaan door N te beschrijven als de doorsnede van de vlakken α en β die we in het bewijs hebben gebruikt. Stel zoals gewoonlijk $p = (p_1, p_2, p_3)^t$, en analoog voor q, r en s. Uit vergelijking (9.4) volgt dat α bestaat uit de vectoren $(x, y, z)^t \in \mathbb{R}^3$ die voldoen aan de vergelijking

$$\begin{vmatrix} x - p_1 & y - p_2 & z - p_3 \\ r_1 & r_2 & r_3 \\ r_2 s_3 - r_3 s_2 & r_3 s_1 - r_1 s_3 & r_1 s_2 - r_2 s_1 \end{vmatrix} = 0 ,$$

en analoog bestaat β uit de vectoren $(x, y, z)^t \in \mathbb{R}^3$ die voldoen aan

$$\begin{vmatrix} x - q_1 & y - q_2 & z - q_3 \\ s_1 & s_2 & s_3 \\ r_2 s_3 - r_3 s_2 & r_3 s_1 - r_1 s_3 & r_1 s_2 - r_2 s_1 \end{vmatrix} = 0 .$$

We kunnen nu de gemeenschappelijke loodlijn N omschrijven als de verzameling van vectoren $(x, y, z)^t \in \mathbb{R}^3$ die aan beide vergelijkingen voldoen.