# **Kommentare zur Vektorgeometrie,** 12.5.25

Bis Seite 5 Hintergrund Information im Sinne von Felix Kleins *höherem Standpunkt.*

Lineare Gleichungen und lineare Gleichungssysteme gehen bis weit in die Antike zurück. Später, nachdem Fermat und Descartes Koordinatensysteme in die Geometrie eingeführt hatten, waren schon die Voraussetzungen für die Vektorrechnung vorhanden. Aber zu deren Definition kam es erst im 19. Jahrhundert, in Graßmanns "Lineale Ausdehnungslehre" von 1844. Er schreibt darin sehr beeindruckt, wie mühselige Rechnungen dadurch sehr verkürzt würden. Spätestens nach den Erfolgen von Banach und Fredholm in der Funktionalanalysis hatte sich der neue Begriff in der Mathematik durchgesetzt. Er erlaubte einen einheitlichen und eben auch vereinfachenden Blick auf eine Vielzahl von Situationen. Der Preis für diese Zusammenfassung war ein Abstraktionsgrad, der auch heute noch die *Lineare Algebra* unter Studienanfängern als schwierig gelten läßt. Um von einem *Vektorraum* und von *Linearkombinationen* seiner Elemente reden zu können, braucht man nur eine Menge von Objekten, die man addieren und mit reellen Zahlen multiplizieren kann. Das sind nicht nur die Koordinatentripel von Punkten aus der Schulmathematik. Viele den Schülerinnen und Schülern bekannte Situationen gehören dazu:

> Polynome (vom Grad $\leq n$ oder ohne eine solche Einschränkung) kann man mühelos addieren und mit reellen Zahlen multiplizieren.
> Oder Folgen von reellen Zahlen (die konvergent sind oder für die kein solcher Zusatz verlangt wird) lassen sich ohne Zusatzerklärungen addieren und mit reellen Zahlen multiplizieren.
> Auch Funktionen $f : [a, b] \mapsto \mathbb{R}$ (mit vielen Möglichkeiten für zusätzlich verlangte Eigenschaften: einmal oder mehrfach differenzierbar, oder $f(a) = 0$, oder auch $\lim_{x\to b} |f(x)| < \infty, \ldots$) sind nicht schwer zu addieren und mit reellen Zahlen zu multiplizieren.
> Sogar lineare Gleichungen in drei Unbekannten $x, y, z$ kann man addieren und mit reellen Zahlen multiplizieren.
> Und für die Matrizen von linearen Gleichungssystemen mit $m$ Gleichungen und $n$ Unbekannten gilt das auch.

All dies sind Beispiele für Vektorräume, die sogar einfacher sind als die zur Veranschaulichung von Vektoren verwendeten Pfeile. Bei den Pfeilen muss man ja hinzufügen, dass *verschiedene* Pfeile, die durch Parallelverschiebung ineinander überführbar sind, *denselben Vektor* repräsentieren. Daher wird auch von *Pfeilklassen* gesprochen. Und auch die *Parallelogrammaddition* der Pfeile erfordert eine etwas ausführlichere Erklärung als die Addition in den oben aufgeführten Beispielen. Damit ist keine Kritik an den Pfeilen beabsichtigt, denn in der *Analytischen Geometrie* geht es ja gerade darum, geometrische Objekte und geometrische Eigenschaften mit Formeln zu beschreiben und umgekehrt Formeln geometrisch zu interpretieren. Selbstverständlich gehört dazu auch die *Veranschaulichung* von Vektoren durch Pfeile. Nur darf man nicht so weit gehen zu sagen: "Vektoren **sind** Pfeile", weil das der vereinheitlichenden, der vereinfachenden Wirkung der Linearen Algebra ein erhebliches Hindernis in den Weg legt – denn wieso sind Polynome Pfeile oder gar Pfeilklassen?

Wie kann *"addieren und mit reellen Zahlen multiplizieren"* den Eindruck von Schwierigkeit hervorrufen? Meiner Meinung nach liegt das an der Erwartungshaltung. Anders als erwartet definiert die Mathematik *nicht*, was für ein Objekt ein Vektor **ist**, sondern sie sagt:

*Wir haben eine große Menge von Beispielen, in denen nach den Regeln der Vektorrechnung gerechnet werden kann und die daher Vektorräume genannt werden.* Natürlich wird man die Elemente von Vektorräumen als Vektoren bezeichnen – aber das bedeutet **nicht**, dass man so einen "Vektor" aus einem Vektorraum zu einem "Vektor" aus einem **anderen** Vektorraum addieren kann. Die Bezeichnung *ist ein Vektor* bedeutet nicht, dass man ein für immer definiertes Objekt vor sich hat, sondern es ist nur gemeint, dass man in der angesprochenen Situation die Regeln der Vektorrechnung benutzen kann.
Und diese Regeln sind wirklich nicht kompliziert. Addition der Vektoren und Multiplikation mit reellen Zahlen sind so definiert, *dass die einfachen Regeln, die wir vom Zahlenrechnen kennen, auch in der neuen Situation gelten.* Wenn wir den (erwarteten) Vektorraum mit $V$ und seine Elemente mit $u, v, w, \ldots$ bezeichnen,

dann muss für die **Addition** gelten:

**A1. Definiertheit(+):** Für alle $u, v \in V$ ist die Summe $u + v$ definiert und $u + v \in V$.
**A2. Es gibt einen Nullvektor:** $\vec{o} \in V$, sodass für alle $v \in V$ gilt $v + \vec{o} = v$.
**A3. Es gibt additive Inverse:** Zu jedem $v \in V$ gibt es ein $v^-$, sodass $v + v^- = \vec{o}$.
**A4. Assoziativgesetz:** Für alle $u, v, w \in V$ gilt $(u + v) + w = u + (v + w)$.
**A5. Kommutativgesetz:** Für alle $u, v \in V$ gilt $u + v = v + u$.

Tatsächlich ist noch etwas mehr, als hier formuliert ist, allgemein richtig und ich will das als Beispiel für die Abstraktheit der Linearen Algebra beweisen.
Mit Hilfe von A4 und A5 folgt, dass A2 auch ausführlicher lauten könnte: "Es gibt **genau einen** Nullvektor". Denn angenommen, man hätte zwei Vektoren $\vec{o}_1, \vec{o}_2$, für die A2 gilt, dann folgt aus A2 für $\vec{o}_2$ zunächst $\vec{o}_1 = \vec{o}_1 + \vec{o}_2$ und aus A2 für $\vec{o}_1$ folgt weiter $\vec{o}_1 + \vec{o}_2 = \vec{o}_2$, also $\vec{o}_1 = \vec{o}_2$. – Alle Nullvektoren werden (irreführend?) mit demselben Symbol bezeichnet.

Ebenso hätte A3 ausführlicher lauten können "Zu jedem $v \in V$ gibt es **genau ein** Inverses". Denn angenommen man hätte zu einem $v \in V$ zwei Inverse $v_1^-, v_2^-$, so könnte man zu $v + v_1^- = \vec{o}$ das andere Inverse addieren und mit A4 folgern

$$(v_2^- + v) + v_1^- = v_2^- + (v + v_1^-) = v_2^- + \vec{o}, \text{ also } v_1^- = \vec{o} + v_1^- = v_2^-.$$

Diese Folgerungen erhält man allein aus den formulierten Regeln, **ohne** dass man wissen muss, was für Objekte die Elemente von $V$ nun genauer sind.

Dasselbe wiederholt sich bei der Multiplikation mit reellen Faktoren, denn

für die **Multiplikation mit reellen Zahlen** $r, s \in \mathbb{R}$ muss gelten:

**M1. Definiertheit(·):** Für alle $r \in \mathbb{R}$ und alle $v \in V$ ist $r \cdot v$ definiert und $r \cdot v \in V$.
**M2. Wirkung von 1:** Für alle $v \in V$ gilt $1 \cdot v = v$ (Nur beinahe Folge von **M1, M3**).
**M3. Faktor-Assoziativgesetz:** Für alls $r, s \in \mathbb{R}$ und alle $v \in V$ gilt $(r \cdot s) \cdot v = r \cdot (s \cdot v)$.
**M4. Distributivgesetz 1:** Für alle $r \in \mathbb{R}$ und alle $u, v \in V$ gilt $r \cdot (u+v) = (r \cdot u) + (r \cdot v)$.
**M5. Distributivgesetz 2:** Für alle $r, s \in \mathbb{R}$ und alle $v \in V$ gilt $(r+s) \cdot v = (r \cdot v) + (s \cdot v)$.

Natürlich erwarten wir wegen $v = (1 + 0) \cdot v = v + 0 \cdot v$, dass für jedes $v \in V$ gilt $0 \cdot v = \vec{o}$. Um das mit A2 – A4 zu beweisen, muss man nur $v^-$ zu $v = v + 0 \cdot v$ addieren:

$$\vec{o} = v + v^- = v + v^- + 0 \cdot v = \vec{o} + 0 \cdot v = 0 \cdot v.$$

Ebenso erwarten wir für das Inverse, dass gilt: $v^- = (-1) \cdot v$ (und schreiben damit in Zukunft $u - v$ statt $u + v^-$). Auch dafür ist nicht die Rechnung sondern nur der Abstraktionsgrad schwierig, denn $(-1) \cdot v$ ist in der Tat additives Inverses von $v$:

$$\vec{o} = (1 - 1) \cdot v = (1 \cdot v) + ((-1) \cdot v).$$

Dass die uns wohlbekannten Regeln des Zahlenrechnens mit nur sehr geringen Zusätzen als *"Axiome eines Vektorraums"* gewählt werden können und damit in einer riesigen Menge von Situationen genutzt werden können, ist ein besonders klar erkennbarer Meilenstein eines **begrifflichen** Fortschritts in der Mathematik.
Zurück zur Schulmathematik. Wer sich an den Geometrieunterricht erinnert, dem wird offensichtlich etwas fehlen, denn dort wird von Anfang an von *senkrecht stehen*, von *Streckenlängen*, von *Abständen* und etwas später auch von *Winkeln* gesprochen – und diese Begriffe tauchen in der Vektorraumaxiomatik offenbar gar nicht auf. Aber ohne sie hat ein Vektor keine Länge und keine Richtung, d.h. ohne diese dem Vektorraum hinzuzufügenden Begriffe ist es nicht sehr suggestiv, Vektoren als Pfeile zu veranschaulichen.
Mir ist kein Beispiel bekannt, in dem die Definition eines mathematischen Begriffes zuerst formuliert wurde und erst nachher untersucht wurde, wo er nützlich sein könnte. So wurde von 1500 bis 1800 mit komplexen Zahlen gerechnet, ehe sie als zweidimensionale Zahlen interpretiert und damit der Vorstellungskraft näher gebracht wurden. Ebenso wurde nach Fermat und Descartes in Koordinatensystemen gerechnet, ohne dass das Fehlen des Vektorraumbegriffs auffiel, bis Graßmann beobachtete, wie viel übersichtlicher Rechnungen durch die Brille dieser Abstraktion werden. Dass Studierenden im ersten Semester die Begriffe Vektorraum und Skalarprodukt axiomatisch vorgesetzt werden können, hat als Voraussetzung, dass auf der Schule Polynome addiert und mit reellen Zahlen multipliziert wurden, dass lineare Gleichungen nach denselben Regeln manipuliert wurden (nämlich, um Gleichungssysteme zu vereinfachen), und eben auch, dass die Studienanfänger auf der Schule mit Koordinatengeometrie konfrontiert waren.
Beim Kennenlernen der Koordinatengeometrie steht der Vektorraumbegriff nicht an erster Stelle. Auf den Koordinatenachsen sind Einheitspunkte markiert und damit ist klar, wie lang Teilstrecken der Koordinatenachsen sind. Und natürlich sind Strecken, die durch Parallelverschiebung aus einander entstehen, gleich lang. Dann kommt der Auftritt des Pythagoras: Mit seiner Hilfe können die Längen der Diagonalen von Rechtecken und danach von Quadern berechnet werden. Geraden werden zunächst als *parametrisierte Geraden* beschrieben:

$$\mathit{Geradenpunkt}(t) = \mathit{Startpunkt} + t \cdot \mathit{Richtungsvektor},$$
$$\big(x(t), y(t), z(t)\big) = (p_1, p_2, p_3) + t \cdot (v_1, v_2, v_3).$$

Vektoren werden eingeführt als Differenzen der Koordinaten zweier Punkte:

$$(v_1, v_2, v_3) := (q_1, q_2, q_3) - (p_1, p_2, p_3).$$

Bei den in diesem Zusammenhang auftretenden Rechnungen, z.B. *Schnittpunkt einer parametrisierten Geraden mit einer Koordinatenebene*, würde ich hervorheben, dass nach denselben Regeln gerechnet wird wie beim Vereinfachen von linearen Gleichungssystemen oder bei Linearkombinationen von Polynomen $(r \cdot P(x) + s \cdot Q(x))$ – eben nach den oben formulierten Vektorraumaxiomen.
Die an Quadern geübte Längenberechnung führt jetzt zur Länge von Vektoren:

$$\big|(v_1, v_2, v_3)\big|^2 := v_1^2 + v_2^2 + v_3^2.$$

Ich habe in Schulbüchern wie in Programmiersprachen gesehen, dass diese Definition gemacht wird, während das dazu gehörige Skalarprodukt seltsamer Weise ignoriert wird. Der Weg dahin ist kurz. Betrachte zu einem Dreieck $ABC$ die Differenzvektoren

$$v := B - A,\ w := C - A,\ w - v = C - B.$$

Dies Dreieck ist bei $A$ rechtwinklig, wenn der Satz des Pythagoras in der Form
$|AB|^2 + |AC|^2 = |BC|^2$ gilt, oder mit den Vektorkomponenten

$$\begin{aligned}|v|^2 + |w|^2 &= v_1^2 + v_2^2 + v_3^2 + w_1^2 + w_2^2 + w_3^2 \\ &= |v - w|^2 = (v_1 - w_1)^2 + (v_2 - w_2)^2 + (v_3 - w_3)^2 \\ &= v_1^2 + v_2^2 + v_3^2 + w_1^2 + w_2^2 + w_3^2 - 2(v_1 \cdot w_1 + v_2 \cdot w_2 + v_3 \cdot w_3).\end{aligned}$$

Diese kurze Rechnung zeigt: Das Dreieck $ABC$ ist bei $A$ rechtwinklig genau dann wenn

$$v_1 \cdot w_1 + v_2 \cdot w_2 + v_3 \cdot w_3 = 0.$$

Allein das ist schon genug, um die *Definition* zu rechtfertigen

$$\boxed{Skalarprodukt(v, w) = v \bullet w := v_1 \cdot w_1 + v_2 \cdot w_2 + v_3 \cdot w_3,}$$

mit den einfachen Skalarprodukt Regeln:

1. positiv definit $\qquad 0 \leq |v|^2 = v \bullet v$
2. symmetrisch $\qquad v \bullet w = w \bullet v$
3. linear in jedem Argument $\qquad (r \cdot u + s \cdot v) \bullet w = r \cdot u \bullet w + s \cdot v \bullet w.$

Wer auf der Schule als Verallgemeinerung des Pythagoras den Kosinussatz

$$|CB|^2 = |AB|^2 + |AC|^2 - 2|AB| \cdot |AC| \cdot \cos(\angle(BAC)),$$

also auch $|w-v|^2 = |v|^2+|w|^2-2|v|\cdot|w|\cos(\angle(v,w))$, kennen gelernt hat, für den liefert obige kurze Rechnung eine zentrale Beziehung des Skalarprodukts zur Euklidischen Geometrie:

$$\boxed{|v| \cdot |w| \cdot \cos(\angle(v, w)) = v \bullet w.}$$

Diese Formel besagt: *Ist in einem Vektorraum ein Skalarprodukt gegeben, so sind dadurch auch Längen und Winkel gegeben – und umgekehrt.*

Die Umkehrung wird noch unterstrichen durch die Gleichung $|u + v|^2 - |u - v|^2 = 4\, u \bullet v.$

Beim Rechnen mit Zahlentripeln entsteht leicht folgender falscher Eindruck: Man hat sowohl die *Standard*-Basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ wie auch das oben hergeleitete *Standard*-Skalarprodukt und diese sehen so aus, als seien sie etwas Natürliches und besonders einfach. Tatsächlich gehört zu *jeder* Basis ein ganz einfach zu definierendes Skalarprodukt, sodass damit alle Rechnungen ebenso aussehen wie in der Standardform. Man muss nur verabreden, dass eine beliebige Basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ eine *Orthonormalbasis sein soll*, dass also zu definieren ist

$$1 \leq i \leq n \;\Rightarrow\; \vec{v}_i \bullet \vec{v}_i := 1 \text{ und } 1 \leq i \neq j \leq n \;\Rightarrow\; \vec{v}_i \bullet \vec{v}_j := 0.$$

Dann folgt für zwei beliebige Vektoren $\vec{x} = \sum_i x_i \cdot \vec{v}_i$, $\vec{y} = \sum_j y_j \cdot \vec{v}_j$ aus den Regeln für Skalarprodukte und den Verabredungen über $\vec{v}_i \bullet \vec{v}_j$:

$$\vec{x} \bullet \vec{y} = \sum_{i,j} x_i y_j \cdot \vec{v}_i \bullet \vec{v}_j = \sum_i x_i \cdot y_i.$$

Das so definierte Skalarprodukt wird also aus den Koeffizienten $x_i, y_j$ bezüglich der gewählten Basis nach derselben einfachen Formel berechnet, mit der wir zum ersten Mal (s.o.) ein

Skalarprodukt kennen gelernt haben. Und natürlich haben die Basisvektoren wie bei den Zahlentripeln die einfachen Koeffizienten $\{(1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$. Berechnungsformeln für Skalarprodukte werden erheblich komplizierter, wenn man sie aus den Koeffizienten bezüglich einer anderen, *nicht orthonormalen* Basis berechnen muss.

So weit mein Hintergrundkommentar. Von hier an ist mein Standpunkt näher an der Schule.

---

**Lineare Gleichungssysteme, Rechnen mit den Vektorraumaxiomen**

Das Wort *Vektor* taucht in der Schule erst in der Analytischen Geometrie der Oberstufe auf und auch lineare Gleichungssysteme werden erst in diesem Zusammenhang ausführlich behandelt – und natürlich mit der Interpretation versehen, dass ihre Lösungsmengen Ebenen, Geraden und Punkte beschreiben können.

Tatsächlich wird mit den auf Seite 2 zitierten Vektorraumaxiomen bereits früher gerechnet, ohne dass das explizit gesagt wird: Die Addition von Polynomen und deren Multiplikation mit reellen Zahlen erfolgt genau nach den Vektorraumaxiomen! Man hat dort schon deutlich vor der Analytischen Geometrie *Vektorräume* vor sich. Eine natürliche Aufgabe ist, die Koeffizienten eines höchstens quadratischen Polynoms $p(x) = a_0+a_1\cdot x+a_2\cdot x^2$ zu bestimmen, dessen Graph durch die drei Punkte $(x_1|w_1)$, $(x_2|w_2)$, $(x_3|w_3)$ mit $x_1 < x_2 < x_3$ geht. Das führt auf ein *lineares Gleichungssystem*, in dem die $x_j, w_j$ gegebene Zahlen sind und die Koeffizienten $a_0, a_1, a_2$ des Polynoms die gesuchten Unbekannten:

$$\begin{aligned} a_0 + x_1\cdot a_1 + x_1^2\cdot a_2 &= w_1\\ a_0 + x_2\cdot a_1 + x_2^2\cdot a_2 &= w_2\\ a_0 + x_3\cdot a_1 + x_3^2\cdot a_2 &= w_3 \end{aligned}$$

Die Form dieses Gleichungssystems suggeriert, die erste Zeile von den beiden anderen Gleichungen abzuziehen und die neuen Gleichungen als Gleichungssystem in nur noch zwei Unbekannten $a_1, a_2$ anzusehen. Das ist erstens eine Äquivalenzumformung, weil sie sich leicht *rückgängig machen* läßt und zweitens ist es ein guter erster Schritt zum Erlernen des Gauß Algorithmus, mit dem allgemeine lineare Gleichungssysteme effizient gelöst werden. Man erhält in diesem Beispiel:

$$\begin{aligned} a_0 \quad + \quad x_1\cdot a_1 \quad + \quad x_1^2\cdot a_2 &= w_1\\ (x_2-x_1)\cdot a_1 + (x_2^2-x_1^2)\cdot a_2 &= w_2-w_1\\ (x_3-x_1)\cdot a_1 + (x_3^2-x_1^2)\cdot a_2 &= w_3-w_1 \end{aligned}$$

Man kann schon an dieser Stelle sehen, dass man auch beim Vereinfachen linearer Gleichungen mit den Vektorraumaxiomen rechnet, dass man also einen weiteren Vektorraum vor sich hat! – Schließlich wird von der letzten Gleichung die mit $(x_3-x_1)/(x_2-x_1)$ multiplizierte zweite Gleichung abgezogen und damit in die nach $a_2$ auflösbare Gleichung

$$(x_3-x_1)(x_3-x_2)\cdot a_2 = w_3-(w_2-w_1)\cdot(x_3-x_1)/(x_2-x_1)$$

verwandelt. Danach kann man $a_1$ aus der zweiten, $a_0$ aus der ersten Gleichung berechnen. Falls sich $a_2 = 0$ ergibt, ist das Polynom linear oder konstant. – In diesem Beispiel hat man damit die Gaußsche Zeilen-Stufen-Form erreicht, also den Gauß Algorithmus vorbereitet.

Da Nullstellen von Polynomen (nicht nur auf der Schule) wichtig sind, möchte ich mit deren Hilfe und etwas linearer Argumentation die eben gelöste Aufgabe ohne Gleichsysteme neu

beantworten. Es ist nämlich sehr einfach, drei Polynome $P_j$ anzugeben, die an einer Stelle $x_j$ den Wert $P_j(x_j) = 1$ haben und an den beiden anderen Stellen den Wert 0:

$$P_1(x) := \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)},\ P_2(x) := \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)},\ P_3(x) := \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}.$$

Das gesuchte quadratische Polynom ist offenbar $p(x) = w_1 \cdot P_1(x) + w_2 \cdot P_2(x) + w_3 \cdot P_3(x)$. Da die Werte $w_j$ in diesem Ausdruck beliebig sind, kann man jedes höchstens quadratische Polynom in dieser Weise als sogenannte *Linearkombination* der Polynome $P_1, P_2, P_3$ darstellen. Dieser Trick funktioniert *ohne größeren Aufwand* auch bei Polynomen höheren Grades.
Man lernt daraus: Wenn lineare Gleichungssysteme nicht "allgemein" sind, sondern auf eine bestimmte Weise entstehen, dann kann es sehr viel einfachere Lösungsmöglichkeiten geben.

> Vor allem hoffe ich, dass das Vorhergehende überzeugend darlegt, dass man mit einer nur geringen Akzentverschiebung die Vektorrechnung früher auftreten lassen kann, sodass sie bereits bekannt ist, wenn in der Analytischen Geometrie dem Vektorraum der Ortsvektoren zusätzliche Eigenschaften hinzugefügt werden – nämlich die Länge von Vektoren und dass sie paarweise senkrecht sein können.

Ich werde ausführlich auf lineare Gleichungssysteme zurückkommen. Zunächst erweitere ich das Vokabular um wichtige Begriffe.
Die Vektorraumaxiome setzen voraus, dass Vektoren addiert und mit reellen Zahlen multipliziert werden können. Der Begriff *Linearkombination* drückt elegant aus, was man mit dieser Voraussetzung machen kann. Ich besitze ein Schulbuch (Jahrgang 2020), in dem dieser nützliche Begriff zwar in Aufgaben vorkommt, aber nicht definiert wird. Um vollständig zu sein:

> Gegeben seien *endlich* viele Vektoren $v_j \in V$ und ebenso viele reelle Zahlen $r_j \in \mathbb{R}$.
>
> **Definition.** Der Vektor $v := \sum_{j=1}^{n} r_j \cdot v_j$ heißt **Linearkombination** der $v_j$.
>
> Und eine dazu gehörige **Definition**: Eine Teilmenge $U$ eines Vektorraums $V$ heißt **Untervektorraum** von $V$, wenn $U$ die Vektorraumaxiome erfüllt.

Wie diese beiden Definitionen zusammenhängen, kann so ausgedrückt werden:

> *Alle Linearkombinationen von Elementen aus $U$ liegen wieder in $U$.*

Zu jeder Teilmenge $M \subset V$ kann man daher den von $M$ *erzeugten* Unterraum definieren, nämlich die Menge $U$ aller Linearkombinationen von Elementen aus $M$.

Der Begriff "Linearkombination" beschreibt nur, was man mit Vektoren machen kann. Um auch argumentieren zu können, braucht man einen weiteren Begriff, den ich für den wichtigsten der linearen Algebra halte.

> **Definition.** Die Vektoren der endlichen Menge $\{v_j;\ j = 1, \ldots, n\}$ heißen **linear abhängig**, falls es reelle Zahlen $r_j$ gibt, die nicht alle $= 0$ sind, sodass
> $$\sum_{j=1}^{n} r_j \cdot v_j = \vec{o}.$$
> Falls es keine solchen Zahlen $r_j$ gibt, heißen die $v_j$ **linear unabhängig**.

*Bemerkung 1.* Die *einzige* Linearkombination einer Menge linear unabhängiger Vektoren $\{v_j;\ j = 1, \ldots, n\}$, *die den Nullvektor ergibt*, ist daher $\sum_{j=1}^{n} 0 \cdot v_j = \vec{o}$.
*Bemerkung 2.* Eine Menge von Vektoren, die eine linear abhängige Teilmenge enthält, ist unmittelbar nach Definition ebenfalls linear abängig.
*Bemerkung 3.* Eine Menge von Vektoren, die den Nullvektor enthält, ist linear abhängig, denn $0 \cdot v + 1 \cdot \vec{o} = \vec{o}$ ist eine Linearkombination in der ein Koeffizient $\neq 0$ ist.

Um mit einem neuen Begriff arbeiten zu können, hilft es, zu wissen, was einfache Situationen sind. Dazu ein Beispiel. Wir können Zeilen-n-Tupel der Länge 5 addieren und mit reellen Zahlen multiplizieren. Sie bilden also einen Vektorraum. Um fünf dieser Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5$ als linear unabhängig nachzuweisen , muss von einem linearen Gleichungssystem für die fünf Koeffizienten $r_1, r_2, r_3, r_4, r_5$ der Linearkombination $\sum_{j=1}^{n} r_j \cdot \vec{v}_j = \vec{o}$ gezeigt werden, dass es **nur** die Lösung $r_1 = 0, r_2 = 0, r_3 = 0, r_4 = 0, r_5 = 0$ hat. Das kann mühsam sein, ist es aber nicht immer. Betrachte:

$$\begin{array}{rcl} \vec{v}_1 & = & (1,\ 2,\ 3,\ 4,\ 5) \\ \vec{v}_2 & = & (0,\ 3,\ 5,\ 7,\ 9) \\ \vec{v}_3 & = & (0,\ 0,\ 8,\ 8,\ 8) \\ \vec{v}_4 & = & (0,\ 0,\ 0,\ 1,\ 1) \\ \vec{v}_5 & = & (0,\ 0,\ 0,\ 0,\ 7) \end{array}$$

Dann ist

$$\sum_{j=1}^{n} r_j \cdot \vec{v}_j \quad = \quad \begin{array}{l} (1,\ 2,\ 3,\ 4,\ 5) \cdot r_1 \\ (0,\ 3,\ 5,\ 7,\ 9) \cdot r_2 \\ (0,\ 0,\ 8,\ 8,\ 8) \cdot r_3 \\ (0,\ 0,\ 0,\ 1,\ 1) \cdot r_4 \\ (0,\ 0,\ 0,\ 0,\ 7) \cdot r_5 \end{array} \quad = \ \vec{o} \ = \ (0,0,0,0,0)$$

Nun liefert die erste Spalte $r_1 = 0$, danach folgt aus der zweiten Spalte $r_2 = 0$. Danach liefert die dritte Spalte $r_3 = 0$. Mit $r_1 = r_2 = r_3 = 0$ liefert die vierte Spalte $r_4 = 0$ und aus der letzten Spalte folgt $r_5 = 0$. Gleichungssysteme dieser Form kann man also ohne Rechnung behandeln!

Die Frage: *Wie stellt man fest, ob eine gegebene Menge von Vektoren* $\{v_j;\ j = 1, \ldots, n\}$ *linear unabhängig ist oder nicht?* wird immer durch Lösen der Gleichung $\sum_{j=1}^{n} x_j \cdot v_j = \vec{o}$ beantwortet. Diese ist ein *lineares Gleichungssystem* und für Unabhängigkeit muss man daraus folgern können, dass $x_j = 0$, $j = 1, \ldots, n$ die *einzige* Lösung ist.

Auf den ersten Blick hängt das Aussehen dieses Gleichungssystems davon ab, was die Vektoren $v_j$ sind. Erwähnt habe ich bisher, dass sie gegebene Polynome oder lineare Gleichungen sein können. Wenige Leser werden überrascht sein, dass auch Zahlentripel oder Zahlen-n-tupel als die $v_j$ vorkommen können, wie in dem vorgeführten Beispiel. Die Methoden zum Lösen von Gleichungssystemen hängen nicht von diesem ersten Aussehen ab und weil sie ein so *wichtiges Werkzeug* sind, werden sie immer in folgender Form besprochen:

**Standard Form linearer Gleichungssysteme**

$$\begin{array}{c} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n = w_1 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + \ldots + a_{2n} \cdot x_n = w_2 \\ \vdots \\ a_{n1} \cdot x_1 + a_{n2} \cdot x_2 + \ldots + a_{nn} \cdot x_n = w_n \end{array}$$

Die wichtigste Umformung ist:

> *Verändere eine Gleichung durch Subtraktion einer Linearkombination der übrigen Gleichungen.*

Dies ist eine Äquivalenzumformung, weil man sie dadurch rückgängig machen kann, dass man *dieselbe* Linearkombination zu der veränderten Zeile wieder *addiert.* Auf dieser Umformung beruht ein wichtiges Lösungsverfahren, der **Gauß-Algorithmus**. (Beispiel S. 9)

*Ziel* des *Gauß-Algorithmus* ist, das Gleichungssystem in folgende Form zu bringen – wobei ich für die erste Diskussion annehme, dass die *ersten* Koeffizienten jeder Zeile $\neq 0$ sind. Denn dann hat man (wie in obigem Beispiel) ein ohne Rechnung lösbares Gleichungssystem erreicht, weil man, mit der letzten Zeile beginnend, die Werte der Variablen ablesen kann.

$$\begin{aligned} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n &= w_1 \\ b_{22} \cdot x_2 + \ldots + b_{2n} \cdot x_n &= \tilde{w}_2 \\ \vdots & \quad . \\ b_{nn} \cdot x_n &= \tilde{w}_n \end{aligned}$$

Dazu wird im ersten Schritt für $j = 2 \ldots n$ die erste Zeile mit $a_{j1}/a_{11}$ multipliziert und von der $j$-ten Zeile subtrahiert, um in der $j$-ten Zeile den Summanden mit $x_1$ zu eliminieren. Nun sind entweder auch alle $b_{2j} = 0$ und man geht zur dritten Spalte – oder man kann die unteren Zeilen so umsortieren, dass $b_{22} \neq 0$ ist. Danach kann der erste Schritt wiederholt werden: Für $j = 3 \ldots n$ wird die zweite Zeile mit $b_{j2}/b_{22}$ multipliziert und von der $j$-ten Zeile subtrahiert, um in jeder von diesen wieder den ersten Summanden zu 0 zu machen. Dies Verfahren wird bis zur letzten Spalte fortgesetzt. Dabei wird entweder die aufgeschriebene Zeilen-Stufen-Form erreicht (*Fall 1*) oder man kommt zur letzten Spalte in weniger als $n$ Schritten, weil bei der Durchführung das Umsortieren nicht immer möglich war, da alle weiteren Koeffizienten der bearbeiteten Spalte schon $= 0$ waren (*Fall 2*).

Was hat man mit dieser Umformumg erreicht?

*Fall 1: Die n-zeilige Endform wird erreicht.*

*Fall 1a:* $b_{nn} \neq 0$. Dann folgt $x_n = \tilde{w}_n/b_{nn}$ und jedes vorhergende $x_j$ kann aus der Zeile $j$ ausgerechnet werden, weil jedes $b_{jj} \neq 0$ ist – denn sonst wäre das Gauß Verfahren ja in weniger als $n$ Schritten zu Ende gegangen.

*Fall 1b:* $b_{nn} = 0, \tilde{w}_n \neq 0$. Dann ist das Gleichungssystem unlösbar. – Das hätte auch schon von Anfang an passieren können: Alle $a_{jk} = 0$, aber mindestens ein $w_j \neq 0$.

*Fall 1c:* $b_{nn} = 0, \tilde{w}_n = 0$. Dann kann für $x_n$ ein beliebiger Wert gewählt werden und die übrigen $x_j$ können für jede solche Wahl wie in (1a) berechnet werden. Das Gleichungssystem hat also unendlich viele Lösungen.

*Bemerkung.* Im Fall 1a sind die $n$ linken Seiten der Zeilen-Stufen-Form linear unabhängig. In (1b) sind die n Gleichungen linear unabhängig, aber nur die $(n-1)$ ersten linken Seiten. In (1c) sind die $n$ Gleichungen linear abhängig, die $n-1$ ersten sind linear unabhängig.

*Fall 2: Das Verfahren geht in weniger als n Schritten zu Ende.*

Sinngemäß passiert dasselbe wie in Fall 1: Die letzte Zeile hat genau eine oder gar keine oder unendlich viele Lösungen. Falls die letzte Zeile lösbar ist, können auch die vorhergehenden Zeilen gelöst werden, weil der erste Koeffizient jeder Zeile $\neq 0$ ist. In den Zeilen, nach denen der Stufensprung $> 1$ ist, können für weitere Variablen beliebige Werte gewählt werden.

Bei linearen Gleichungssystemen mit drei (vielleicht auch vier) Unbekannten können die verschiedenen Möglichkeiten, die bei der Durchführung des Gauß-Algorithmus auftreten, noch mit Hilfe von Fallunterscheidungen diskutiert werden. Ab vier Unbekannten ist es aber günstiger, wenn das Argumentieren mit linear abhängig und unabhängig weiter entwickelt ist als bis hier. Bevor ich dazu komme, behandele ich noch ein Beispiel. Mit zwei Ausnahmen

sind in folgendem Gleichungssystem die Koeffizienten feste Zahlen, an den Ausnahmen $b, c$ erkläre ich die Fallunterscheidungen.

$$\begin{aligned} 1x + 2y + 3z &= 4 \\ 2x + by + 7z &= 10 \\ 3x + 6y - cz &= 8 \end{aligned}$$

Als Schritt 1 subtrahieren wir zweimal Zeile 1 von Zeile 2 und dreimal Zeile 1 von Zeile 3, um in diesen beiden Zeilen die Variable $x$ zu eliminieren. Wir erhalten:

$$\begin{aligned} 1x \quad + \quad 2y \; + \; 3z &= 4 \\ (b-4)y \; + 1z &= 2 \\ 0y - (c+9)z &= -4 \end{aligned}$$

Falls $b = 4$ ist, ist in diesem Gleichungssystem auch die Variable $y$ so eliminiert, dass in den *beiden* letzten Gleichungen nur noch die Variable $z$ vorkommt. Falls $c \neq -7$ ist, ergeben sich aus den beiden Gleichungen *verschiedene* Werte für $z$. Das Gleichungssystem ist also *unlösbar*. Andernfalls, für $c = -7$, liefern die beiden letzten Gleichungen $z = 2$. Nun kann man in der ersten Gleichung $z = 2$ und für $y$ einen *beliebigen* Wert einsetzen und dann $x$ ausrechnen. Das System hat jetzt unendlich viele Lösungen! Damit sind die Möglichkeiten für $b = 4$ diskutiert.
Falls $b \neq 4$ ist, müßte ein weiterer Gauß-Schritt ausgeführt werden, um $y$ in der dritten Zeile zu eliminieren. Da dadurch keine weiteren Fälle auftreten, habe ich die dritte Gleichung so gewählt, dass $y$ schon im ersten Schritt eliminiert wurde. – In der letzten Zeile treten wieder zwei Fälle auf: Falls $c = -9$ ist, ist die letzte Zeile *unlösbar (*nämlich $0 \cdot z = -4$). Falls $c \neq -9$ ist, ergibt die letze Zeile einen eindeutigen Wert für $z$. Mit diesem Wert liefert die zweite Zeile $y = (2 - z)/(b - 4)$ und die erste Zeile gibt $x = 4 - 3z - 2y$.
Wir sehen also: Außer für spezielle Koeffizienten ist das Gleichungssystem *eindeutig lösbar*, für manche Koeffizienten ist es *unlösbar* und für andere spezielle Koeffizienten gibt es *unendlich viele Lösungen.* – Ich zeige später, dass es immer nur diese drei Möglichkeiten gibt.

Auf welche Weise man das Lösen linearer Gleichungssysteme im Einzelnen beschreibt, ist weniger eine mathematische Frage und eher Geschmacksache. Wichtig ist jedoch, dass dieses Lösen als wichtiges Handwerkszeug vermittelt wird. Da am Anfang immer drei Möglichkeiten bestehen: *Es kann eine einzige Lösung geben oder gar keine oder unendlich viele* – und da genauere Analysen zeigen, dass deshalb Rundungsfehler zu sehr falschen Ergebnissen führen können – scheint es mir wichtig, dass man sich nicht blind auf Rechnerausgaben verläßt, sondern dass man gelernt hat, was von solchen Ausgaben zu erwarten ist.
Um die Behandlung des Gauß-Algorithmus mit einer übersichtlichen Formulierung abzuschließen, führe ich als nächstes eine fundamentale Argumentation mit linear abhängig und unabhängig vor. Es geht darum, in einem Vektorraum eine *möglichst kleine* Menge von Vektoren zu finden, sodass deren Linearkombinationen *alle* anderen Vektoren ergeben. Ein schon bekanntes Beispiel dafür ist der Vektorraum aller Polynome vom Grad $\leq n$: Jedes Polynom $p(x) = \sum_{j=0}^{n} a_j \cdot x^j$ ist in *eindeutiger* Weise Linearkombination der Potenzen $p_0(x) = 1, p_1(x) = x, \ldots, p_n(x) = x^n$, denn $p$ ist nur dann dass Nullpolynom, wenn alle $a_j = 0$ sind.

### Minimale Erzeugendensysteme, maximal linear unabhängige Mengen

Um einen Vektorraum $V$ aus einer *möglichst kleinen* Teilmenge $M$ mit Linearkombinationen zu erzeugen, muss $M$ eine linear unabhängige Teilmenge sein, denn ein Vektor aus $M$, der Linearkombination anderer Vektoren aus $M$ ist, kann weggelassen werden, ohne dass die Menge der erzeugten Vektoren sich ändert. Und außerhalb von $M$ darf es keine Vektoren geben, die *nicht* Linearkombinationen von Vektoren aus $M$ sind – man sagt: *$M$ muss maximal linear unabhängig sein.*
Solche maximal linear unabhängigen Teilmengen wird man konstruieren, indem man nach einander immer mehr Vektoren hinzunimmt, die zusammen mit den bereits gewählten linear unabhängig sind. Ohne weitere Argumentation könnte die *Anzahl* der Elemente einer maximal linear unabhängigen Menge von der *Reihenfolge der Wahlen* weiterer Vektoren abhängen. Der **Austauschsatz von Steinitz** wird zeigen, dass das nicht der Fall ist.

Wir zeigen zuerst eine *Vorstufe des Austauschsatzes.*
**Voraussetzung:**
Es sei $M := \{v_1, v_2, \ldots, v_n\} \subset V$ maximal linear unabhängig und $\vec{o} \neq w \in (V \setminus M)$.
**Behauptung:**
Man kann $w$ gegen ein geeignetes $v_k \in M$ austauschen und so eine *andere* maximal linear unabhängige Teilmenge $M_*$ bekommen.
**Beweis:**
Weil $M$ maximal ist, ist $M \cup \{w\}$ linear abhängig. Das heißt, es gibt eine Linearkombination

$$s \cdot w + \sum_{j=1}^{n} r_j \cdot v_j = \vec{o},$$

deren Koeffizienten *nicht* alle $= 0$ sind. Genauer, erstens muss $s \neq 0$ sein, weil $M$ sonst linear abhängig wäre, und zweitens muss mindestens ein $r_k \neq 0$ sein, weil $w \neq \vec{o}$ gewählt wurde. Damit haben wir

$$\text{(L)} \qquad v_k = \big( - s \cdot w + \sum_{j=1,\, j\neq k}^{n} r_j \cdot v_j \big)/r_k, \quad w = -\big( \sum_{j=1}^{n} r_j \cdot v_j \big)/s.$$

In allen Linearkombinationen kann also $v_k$ durch diesen Ausdruck ersetzt werden, sodass sich Linearkombinationen ergeben, die statt $v_k$ den gewählten Vektor $w$ benutzen. Die Austauschmenge $M_* := (M \setminus \{v_k\}) \cup \{w\}$ erzeugt also weiterhin den Vektorraum $V$ und sie ist linear unabhängig, denn in einer Linearkombination des Nullvektors

$$\vec{o} = t \cdot w + \sum_{j=1,\, j\neq k}^{n} s_j \cdot v_j$$

wäre $t \neq 0$ wegen der linearen Unabhängigkeit der $v_j$ – aber damit wäre $w$ Linearkombination von $M \setminus \{v_k\}$ und nach Ersetzen von $w$ in Gleichung (L) wäre auch $v_k$ Linearkombination der übrigen $v_j$, im Widerspruch zur vorausgesetzten linearen Unabhängigkeit aller $v_j$.
*Ergebnis:* Die Austauschmenge $M_*$ ist wieder maximal linear unabhängig.

Nun werden zwei maximal linear unabhängige Teilmengen $M_v, M_w$ eines Vektorraums $V$ betrachtet. Der *Steinitzsche Austauschsatz* verallgemeinert die vorhergehende Argumentation, um die Elemente von $M_v$ gegen die von $M_w$ auszutauschen. Daraus folgt dann, dass $M_v$ und $M_w$ gleich viele Elemente enthalten. Das erlaubt die Definition:

> Die Dimension eines Vektorraums, $\dim(V)$, ist die Anzahl der Elemente einer maximal linear unabhängigen Teilmenge. Die Vektoren einer solchen Teilmenge heißen *Basis* von $V$.

### Austauschsatz von Steinitz

**Voraussetzung:**
$M_v = \{v_1, \ldots, v_m\}$ und $M_w = \{w_1, \ldots, w_n\}$ seien maximal linear unabhängige Teilmengen von $V$ und z.B. $m \geq n$.
**Behauptung:**
Man kann nach einander geeignete Elemente von $M_v$ durch die Elemente von $M_w$ ersetzen. Dabei ergibt sich $m = n$.
**Beweis:**
$w_1$ wird wie in dem vorhergehenden einfachen Austauschsatz gegen eines der $v_k$ getauscht. Die erhaltene maximal linear unabhängige Teilmenge heiße $M_1$.
Um $w_2$ auszutauschen, muss die vorherige Argumentation etwas ausgebaut werden. Die Menge $M_1 \cup \{w_2\}$ ist wie eben linear abhängig. Man kann also eine Linearkombination finden, die den Nullvektor ergibt und nicht alle Koeffizienten $= 0$ hat. Der Koeffizient von $w_2$ muss $\neq 0$ sein, weil $M_1$ linear unabhängig ist. Und es muss mindestens ein Koeffizient der verbliebenen $v_j$ auch $\neq 0$ sein, da $w_1$ und $w_2$ linear unabhängig sind. Deshalb kann ein $v_j$ (mit von 0 verschiedenem Koeffizienten) gegen $w_2$ ausgetauscht werden. Die erhaltene Menge heiße $M_2$, sie ist wie vorher maximal linear unabhängig.
Dies Argument kann wiederholt werden bis alle $w_j$ ausgetauscht sind. In jedem Schritt wird benutzt, dass das neu hinzugenommene $w_*$ in der Linearkombination des Nullvektors einen von 0 verschiedenen Koeffizienten hat und dass mindestens eines der verbliebenen $v_j$ einen Koeffizienten $\neq 0$ hat, weil die bereits ausgetauschten $w_j$ linear unabhängig sind.
Falls nach dem Austausch aller $w_j$ in der erhaltenen Menge $M_n$ noch Elemente aus $M_v$ übrig wären (nämlich, weil $m > n$ war), so hätte man den *Widerspruch*, dass die *maximal* linear unabhängige Menge $M_w$ in der *größeren* linear unabhängigen Menge $M_n$ enthalten wäre. Es ist also $M_w = M_n$. Damit ist der Satz von Steinitz bewiesen.

### Zusammenfassung zu linearen Gleichungssystemen

Um einen möglichst guten Überblick über die Lösungen eines linearen Gleichungssystems zu bekommen, ist es praktisch zu unterscheiden zwischen

> **Homogenen Systemen** – das sind solche, bei denen alle $w_j = 0$ sind
> und
> **Inhomogenen Systemen** – das sind solche, bei denen mindestens ein $w_j \neq 0$ ist.
> Zu jedem gegebenen inhomogenen System betrachtet man das *zugehörige* homogene System, indem man alle $w_j = 0$ setzt.

Man sieht leicht, dass die Differenz von zwei Lösungen eines inhomogenen Systems eine Lösung des dazu gehörigen homogenen Systems ist. Und ebenso ist die Summe aus einer Lösung des homogenen Systems und einer Lösung des inhomogenen Systems eine weitere Lösung des inhomogenen Systems. Das läßt sich so zusammenfassen:

*Man erhält* **alle** *Lösungen des inhomogenen Systems, indem man zu einer (möglichst leicht zu findenden) Lösung des inhomogenen Systems* **alle** *Lösungen des zugehörigen homogenen Systems addiert.*

Um die Größe der Lösungsmengen linearer Gleichungssysteme zu verstehen, genügt es daher, die Lösungsmengen homogener Systeme zu verstehen. Dabei haben die Linearkombinationen einen großen Auftritt, denn offensichtlich ist *jede Linearkombination* von Lösungen eines homogenen Systems wieder eine Lösung. Mit anderen Worten:

*Die Lösungsmenge eines homogenen Systems ist ein Vektorraum.*
Homogene Systeme haben immer mindestens die Null-Lösung.
Nur inhomogene Systeme können unlösbar sein.

Kann man schon im Voraus wissen, wie groß die Lösungsmenge ist? Also in der Terminologie der linearen Algebra:

*Wie groß ist die Dimension des Lösungsvektorraums des homogenen Systems?*

Dazu sehen wir uns den Gauß Algorithmus (G-A) noch einmal genauer an. Das Gleichungssystem wird gelöst in dem $n$-dimensionalen Vektorraum der Zahlen-n-Tupel, $\mathbb{R}^n$, wobei $n$ die Anzahl der Variablen $x_j$ ist. Die Umformungen des G-A verändern *nicht* die Maximalzahl linear unabhängiger Zeilen, weder des homogenen Systems, noch des inhomogenen Systems. *Nach Durchführung* des G-A kann man diese Maximalzahlen ablesen:

*Die Anzahl $m_h$ von Null verschiedener Zeilen des homogenen Systems und die Anzahl $m_i$ von Null verschiedener Zeilen des inhomogenen Systems ist für jedes System die Maximalzahl linear unabhängiger Zeilen.*

Und der G-A liefert:

*Das System ist lösbar, falls $m_h = m_i$ ist und unlösbar, falls $m_h < m_i$ ist.*
*Das System ist eindeutig lösbar, falls $n = m_h = m_i$ ist.*
*Falls $n > m_h = m_i$ ist, kann man an $n - m_h$ Stufen ebenso vielen Variablen beliebige Werte zuweisen. Der Lösungsraum hat die Dimension $n - m_h$.*

Der letzte Satz erfordert noch einen Beweis. Dazu werden folgende Werte den $(n - m_h)$ *frei wählbaren* Variablen zugewiesen ($n - m_h$ Zuweisungen von jeweils $n - m_h$ Werten):

$$w_1 = (1, 0, \ldots, 0),\ \ldots,\ w_{n-m_h} = (0, \ldots, 0, 1) \in \mathbb{R}^{n-m_h}$$

und dazu werden die Werte der *übrigen* Variablen aus dem Endzustand des G-A berechnet. Offensichtlich erhält man aus Linearkombinationen dieser speziellen Lösungen des homogenen Systems *alle Lösungen*. Außerdem sind diese Lösungen linear unabhängig, weil schon die Werte $w_1, \ldots, w_{n-m_h}$ der frei wählbaren Variablen linear unabhängig in $\mathbb{R}^{n-m_h}$ sind.

## Lineare Abbildungen

Man kann zur Geschichte der Mathematik vielleicht sagen, dass seit der zweiten Hälfte des 19. Jahrhunderts zusammen mit mathematischen Strukturen deren *strukturerhaltende Abbildungen* sehr betont worden sind. Was kann damit im Falle der Vektorräume gemeint sein? Bevor man anfängt zu argumentieren, kann man mit den Vektorraumaxiomen nichts anderes machen, als *Linearkombinationen von Vektoren* zu bilden. Deshalb sind für Vektorräume die strukturerhaltenden Abbildungen diejenigen, die Linearkombinationen respektieren. Sie werden *lineare Abbildungen* genannt und sind so definiert:

**Definition.** Gegeben seien zwei reelle Vektorräume $V, W$ und eine Abbildung $\ell : V \mapsto W$. Diese Abbildung heißt **linear** falls gilt:

$$\boxed{\forall v_1, v_2 \in V \text{ und } \forall r, s \in \mathbb{R} \text{ gilt } \ell(r \cdot v + s \cdot w) = r \cdot \ell(v) + s \cdot \ell(w)}$$

Diese Definition schließt ein, dass $V$ oder $W$ gleich dem Vektorraum $\mathbb{R}$ ist.
Unmittelbare Folgerung: Lineare Bilder linear abhängiger Vektoren sind linear abhängig, denn

$$\sum_{j=1}^n r_j \cdot v_j = \vec{o} \Rightarrow \sum_{j=1}^n r_j \cdot \ell(v_j) = \ell\left(\sum_{j=1}^n r_j \cdot v_j\right) = \ell(\vec{o}).$$

Kontraposition ergibt: Sind die Bilder $\ell(v_j)$ linear unabhängig, so auch die Urbilder $v_j$.

*Beispiele:*
Ein Vektorraum aus Polynomen (oder aus anderen Funktionen) wird *linear* in die reellen Zahlen abgebildet, wenn jedes Polynom auf seinen Wert an der Stelle $x = a$ abgebildet wird, $\ell : p(x) \mapsto p(a)$. Das wird schon bei der Polynomauswertung benutzt: Polynome sind Linearkombinationen von Potenzfunktionen und die Polynomwerte sind selbstverständlich einfach die Linearkombinationen der Werte der Potenzen.
Wenn wir in einer linearen Gleichung für eine oder mehrere der Variablen reelle Zahlen einsetzen, erhalten wir eine lineare Gleichung mit weniger Variablen. Auch dies Einsetzen liefert lineare Abbildungen von Vektorräumen aus linearen Gleichungen in andere solche Vektorräume.
Lineare Abbildungen können eine sehr große Hilfe sein, um lineare Unabhängigkeiten zu beweisen. Wir betrachten noch einmal das Problem, ein Polynom $n$-ten Grades, $q$, zu finden, das an $n$ Stellen $x_1 < x_2 < \ldots < x_n$ Werte $q(x_1) = w_1, \ldots, q(x_n) = w_n$ annehmen soll. Dies Problem wird ohne Rechnung gelöst mit Hilfe der folgenden Polynome $q_k$

$$q_k(x) := \prod_{j=1,\, j\neq k}^{n} \frac{x - x_j}{x_k - x_j} \quad \Rightarrow \quad q(x) = \sum_{k=1}^{n} w_k \cdot q_k(x).$$

Man kann fragen: *Sind die $q_k$ linear unabhängig?* Die linearen Abbildungen "Auswertung an der Stelle $x_\ell$" beantworten diese Frage ohne Rechnung mit "ja" (wegen $q_k(x_\ell) = 0$ für $k \neq \ell$):

$$\vec{o} = \sum_{k=1}^{n} r_k \cdot q_k(x) \;\Rightarrow\; 0 = \sum_{k=1}^{n} r_k \cdot q_k(x_\ell) = r_\ell \quad \text{für } \ell = 1, \ldots, n.$$

Lineare Abbildungen werden auch bei der Behandlung linearer Gleichungssysteme eingesetzt. Spaltenvektoren reeller Zahlen $(r_1|r_2|\ldots|r_n) \in \mathbb{R}^n$ können mit Hilfe der Koeffizienten $a_{jk}$ eines linearen Gleichungssystems *linear* in Spaltenvektoren $(w_1|w_2|\ldots|w_m) \in \mathbb{R}^m$ abgebildet werden (jede $j$-Zeile in (M) definiert eine lineare Abbildung $\mathbb{R}^n \mapsto \mathbb{R}$, alle zusammen eine lineare Abbildung $\mathbb{R}^n \mapsto \mathbb{R}^m$):

$$\text{(M)} \qquad w_j := a_{j1} \cdot r_1 + a_{j2} \cdot r_2 + \ldots + a_{jn} \cdot r_n = \sum_{k=1}^{k=n} a_{jk} \cdot r_k, \quad j = 1, \ldots, m.$$

Diese Zeilen werden zur Definition der *Matrixmultiplikation* benutzt, was die Übersichtlichkeit sehr verbessert:

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} := \begin{pmatrix} a_{11}\ a_{12} \dots a_{1n} \\ a_{21}\ a_{22} \dots a_{2n} \\ \vdots \\ a_{m1}\ a_{m2} \dots a_{mn} \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \quad \text{abgekürzt: } \vec{w} = A \cdot \vec{r}$$

Noch einmal: Die *Matrixmultiplikation* ist mit Gleichung (M) so definiert, dass hiermit dasselbe gemeint ist wie wie mit den $m$ Zeilen in (M).
Und Malpunkte $\cdot$ liefern immer in jedem Faktor *lineare* Abbildungen:

$$(xA + yB) \cdot \vec{r} = x(A \cdot \vec{r}) + y(B \cdot \vec{r}) \quad \text{und} \quad A \cdot (x\vec{r} + y\vec{s}) = x(A \cdot \vec{r}) + y(A \cdot \vec{s}).$$

Lineare Abbildungen sind also fast überall! Sie machen auch die Formulierung von linearen Gleichungssystemen übersichtlicher. Die Koeffizienten $a_{jk}$ des Gleichungssystems, in derselben Anordnung wie im Gleichungssystem, geben die Matrix einer linearen Abbildung. Die Variablen und die rechten Seiten werden als Spaltenvektoren geschrieben. Mit der Definition der Matrixmultiplikation liest sich das System dann so:

**Lineare Gleichungssysteme**
Trennung von Koeffizienten, Variablen und rechten Seiten mit Matrixmultiplikation

$$\begin{pmatrix} a_{11}\ a_{12} \dots a_{1n} \\ a_{21}\ a_{22} \dots a_{2n} \\ \vdots \\ a_{m1}\ a_{m2} \dots a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}.$$

Die durch die Matrix gegebene Abbildung bildet also die Lösungsvektoren auf die gegebenen rechten Seiten ab. Das Lösen dieses Gleichungssystems bedeutet also auch, zu gegebenen Punkten $\vec{w}$ deren Urbilder unter der linearen Abbildung zu finden. Deshalb wiederhole ich noch einmal: Solange man in einem Vektorraum nur die Vektorraumaxiome hat, sind die linearen Gleichungssysteme nicht nur wichtiges, sondern sogar einziges Handwerkszeug.
In der Analytischen Geometrie kommen neue Werkzeuge hinzu: Längen von Vektoren, paarweises senkrecht Stehen, der Satz des Pythagoras, das Skalarprodukt.

# Analytische Geometrie

**Noch einmal ganz von vorne.** Ein Stück karierten Papiers wird als Teil der Ebene betrachtet. Die Karolinien werden als *parallele* Geraden in Gedanken fortgesetzt. Dabei wird das Parallelenaxiom entweder explizit oder stillschweigend benutzt, denn wir können nicht experimentell feststellen, wie sich Geraden im Unendlichen verhalten. (Mehr dazu in meinem Text G1.) Der karierten Euklidischen Ebene wird noch ein Achsenkreuz mit Einheitspunkten hinzugefügt. Danach kann man zu jedem gezeichneten Punkt ein Paar reeller Zahlen, genannt *Koordinaten*, ablesen und zu jedem Paar von reellen Koordinaten kann man einen Punkt in die Ebene gezeichnet denken. Das sind die von Fermat und Descartes geschaffenen Voraussetzungen der analytischen Geometrie. Damit dies ein Anfang *ganz von vorne* ist, benutze ich zuächst Zahlenpaare $(v_1|v_2)$ und erst später Vektorvariable $\vec{v} = (v_1|v_2)$.
Da man die Koordinaten addieren und mit reellen Zahlen multiplizieren kann, bilden sie einen Vektorraum, den Vektorraum der *Ortsvektoren*. Geraden kommen zuerst als Graphen $\{(x|f(x));\, x \in \mathbb{R}\}$ sogenannter linearer Funktionen $f(x) = m \cdot x + b$ vor. (Wegen der Konstanten $b$ sind diese Funktionen nicht im Sinne der Linearen Algebra *lineare* Funktionen, aber der Name ist fest etabliert.) Mit Hilfe der Vektoren können Geraden jetzt auch beschrieben werden als "parametrisierte Geraden":

> *Geradenpunkt = Anfangspunkt plus Parameter mal Richtungsvektor*
> $g(t) = (p_1|p_2) + t \cdot (v_1|v_2),\ t \in \mathbb{R}$
> Oft wir $t$ als "Zeit" interpretiert, dann heißt der Richtungsvektor auch *Geschwindigkeitsvektor*.

Der Schnittpunkt von zwei Geraden $g(s) = (p_1|p_2) + s \cdot (v_1|v_2),\ h(t) = (q_1|q_2) + t \cdot (w_1|w_2)$ wird mit einem *linearen Gleichungssystem* in den zwei Unbekannten $s, t$ bestimmt:

$$\begin{array}{l} p_1 + s \cdot v_1 = q_1 + t \cdot w_1 \\ p_2 + s \cdot v_2 = q_2 + t \cdot w_2 \end{array} \text{ oder } \begin{array}{l} v_1 \cdot s - w_1 \cdot t = q_1 - p_1 \\ v_2 \cdot s - w_2 \cdot t = q_2 - p_2 \end{array} \text{ oder}$$
$$\begin{pmatrix} v_1 - w_1 \\ v_2 - w_2 \end{pmatrix} \cdot \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \end{pmatrix}.$$

Falls dies Gleichungssystem *unlösbar* ist, sind die Geraden parallel und verschieden.
Falls das Gleichungssystem *unendlich viele* Lösungen hat, stimmen die Geraden überein.
Falls das Gleichungssystem *eindeutig lösbar* ist, gibt die Lösung den Schnittpunkt.
Die geometrische Anwendung macht also die drei Möglichkeiten, die beim Lösen eines Gleichungssystems auftreten können, sehr anschaulich.

Lineare Gleichungen treten noch in anderer Weise zur Beschreibung von Geraden auf. Bezeichnet man die Punkte der Geraden mit $g(t) = (x|y)$, so kann man $t$ eliminieren und findet eine lineare Gleichung in $x, y$, die ebenfalls die Punkte der Geraden beschreibt:

$$\begin{array}{l} x = p_1 + t \cdot v_1 \mid \cdot v_2 \\ y = p_2 + t \cdot v_2 \mid \cdot v_1 \end{array} \text{ Subtraktion ergibt: } v_2 \cdot x - v_1 \cdot y = p_1 \cdot v_2 - p_2 \cdot v_1$$

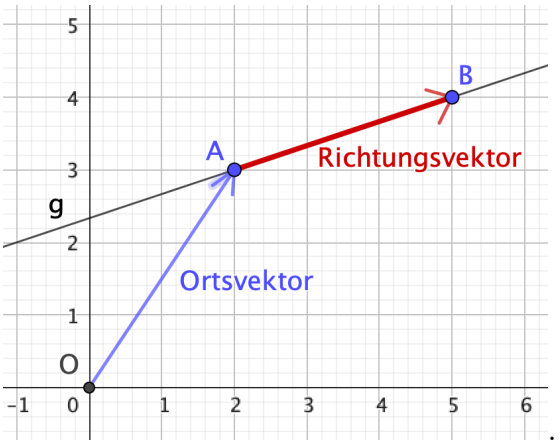
Man kommt von einer Gleichung $a \cdot x + b \cdot y = c$ zur parametrisierten Form zurück, indem man zwei Lösungen $(x_1|y_1), (x_2|y_2)$ der Gleichung bestimmt. Die *Differenz* der Ortsvektoren ist ein Richtungsvektor, also Lösung der *homogenen* Gleichung, z.B. $g(t) = (x_1|y_1) + t \cdot (-b|a)$.

Wenn zwei Geraden durch lineare Gleichungen $a_1 \cdot x + b_1 \cdot y = c_1,\ a_2 \cdot x + b_2 \cdot y = c_2$ gegeben sind, dann bestimmen beide Gleichungen zusammen den Schnittpunkt - und wie vorher: Bei Unlösbarkeit sind die Geraden parallel und verschieden, bei unendlich vielen Lösungen stimmen die Geraden überein.

Diese Fallunterscheidungen hat man auch schon bei der Geradengleichung $a \cdot x + b \cdot y = c$ selber: Wenn $a = b = c = 0$ ist, sind *alle* Punkte der Ebene Lösungen, wenn $a = b = 0, c \neq 0$ ist, dann gibt es keine Lösung und nur wenn $a \neq 0$ oder $b \neq 0$ (oder beides) gilt, beschreibt die Gleichung eine Gerade.

**Veranschaulichung von Vektoren.**

Solange man nur mit Punkten zu tun hat, bieten die Pfeildarstellungen noch keinen Gewinn. Sobald man aber eine parametrisierte Gerade mit Anfangspunkt und Richtungsvektor zeigen will, wird das Bild viel suggestiver, wenn man Ortsvektoren und Richtungsvektoren durch Pfeile vom Anfangspunkt zum Endpunkt visualisiert. Diese Bilder sind so überlegen, dass Vektoren immer durch Pfeile veranschaulicht werden.




**Längen und senkrecht Stehen.**

Von Anfang an gehören Abstände und senkrecht Stehen von Geraden – und jetzt auch Vektoren – zur Euklidischen Geometrie. Daher müssen diese Begriffe auch Teil der analytischen Geometrie sein. Dafür spielt der Satz des Pythagoras eine zentrale Rolle. Da die Koordinatenachsen als zueinander senkrecht gewählt wurden, sind die Längen der ersten und der zweiten Komponente eines Vektors Katheten in einem rechtwinkligen Dreieck, so dass mit Pythagoras die Länge eines Vektors und ebenso der Abstand zweier Punkte definiert werden muss durch

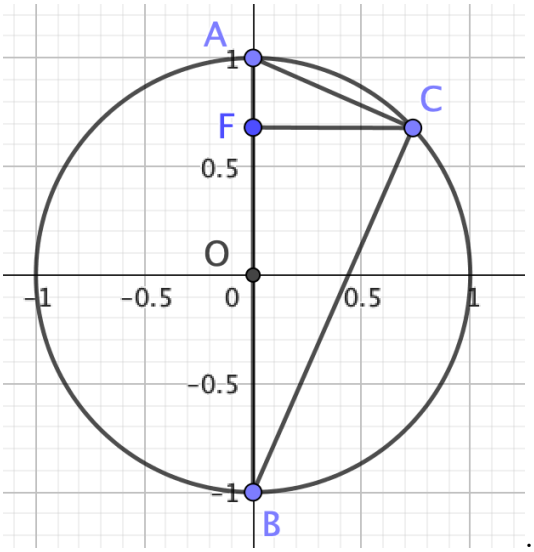
**Definition:** $|(v_1|v_2)|^2 := v_1^2 + v_2^2$ $\qquad$ $|(q_1 - p_1|q_2 - p_2)|^2 := (q_1 - p_1)^2 + (q_2 - p_2)^2$

**Kontrolle:**

Wir müssen überprüfen, ob mit dieser Definition der Pythagoras auch in rechtwinkligen Dreiecken gilt, deren Katheten *nicht* parallel zu den Koordinatenachsen sind. Dazu betrachte das nebenstehende Thalesdreieck $ABC$. Nach der eben gemachten Definition haben wir mit $C = (x|y)$ und $x^2 + y^2 = 1$:

$|AC|^2 = |FC|^2 + |AF|^2 = x^2 + (1 - y)^2$
$|BC|^2 = |FC|^2 + |BF|^2 = x^2 + (1 + y)^2$
und in der Tat folgt:
$|AC|^2 + |BC|^2 = 2x^2 + 2 + 2y^2 = 4 = |AB|^2$.

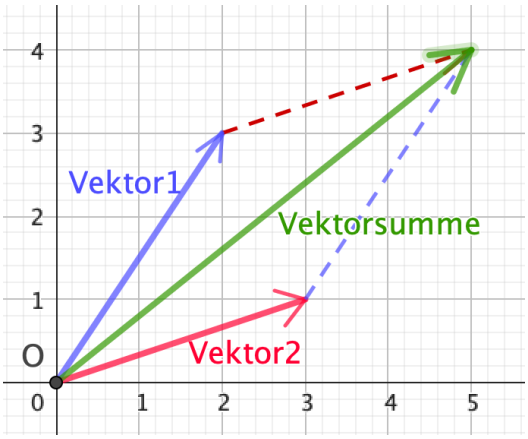
Die Abstandsformel gibt also korrekt wieder, was schon vor der analytischen Geometrie über Abstände in der Ebene bekannt war. Daher können Abstände jetzt direkt aus den Koordinaten der beteiligten Punkte berechnet werden.



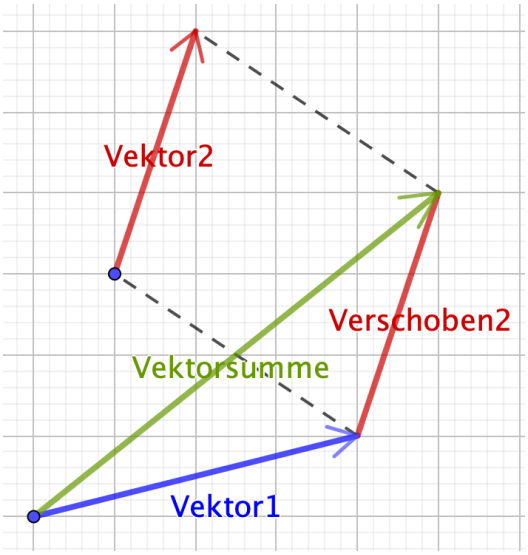

*Bemerkung.* Man kann auch den Höhensatz nachrechnen: $|AF| \cdot |FB| = |FC|^2$, $1 - y^2 = x^2$.

**Veranschaulichung der Vektoraddition.**

Während die Addition von Zahlenpaaren keine Schwierigkeiten macht:
$(v_1|v_2) + (w_1|w_2) = (v_1 + w_1|v_2 + w_2)$,
ist die Veranschaulichung etwas schwieriger, da die Zahlenpaare auf zwei verschiedene Weisen als Vektoren betrachtet werden.
Als *Ortsvektoren* werden sie immer vom Ursprung $O$ aus abgetragen, sodass zwei Vektoren als Ecke eines Parallelogramms betrachtet werden können. Die Summe dieser Vektoren ist dann die von $O$ ausgehende Parallelogrammdiagonale (grün).




Als *Richtungsvektoren* können sie an jeden Punkt der Ebene angehängt werden. Das bedeutet, dass alle Pfeile, die durch Parallelverschiebung aus einander hervorgehen, *denselben* Vektor repräsentieren. Deshalb spricht man bei diesen Vektoren auch von *Pfeilklassen*. Zum Addieren wird der zweite Vektor mit seinem Anfangspunkt an die Spitze des ersten Vektors gehängt. Die *Vektorsumme* ist dann der Pfeil (grün), der vom Anfangspunkt des ersten zum Endpunkt des angehängten zweiten Vektors zeigt. Das ist dann zwar *auch* die Diagonale des Parallelogramms, von dem die beiden Vektoren zwei Seiten sind. Aber das Addieren ist ausdrücklich als *an einander Hängen* gemeint.




Wenn man dies *an einander Hängen* mit einem parallelen Pfeil – statt des gezeichneten (blauen) Vektor1 – beginnt, so bekommt man einen Summenpfeil, der parallel zu dem gezeichneten grünen Pfeil ist. Oder: Man kann die Addition mit jedem Pfeil seiner Pfeilklasse beginnen.

**Erste Schritte mit der Vektordefinition.**

*Wie kann man feststellen, dass zwei Vektoren* $(v_1|v_2)$, $(w_1|w_2)$ *senkrecht zu einander sind?*
Trägt man diese Vektoren vom Ursprung $O = (0|0)$ ab, betrachtet sie also als Ortsvektoren zu den Endpunkten $V, W$, so sind sie genau dann senkrecht zu einander, wenn für das bei $O$ rechtwinklige Dreieck $OVW$ der Satz des Pythagoras gilt.
Der Differenzvektor von $V$ nach $W$ ist $(w_1 - v_1|w_2 - v_2)$. Der Satz des Pythagoras besagt

$$|OV|^2 + |OW|^2 = |VW|^2 \quad \text{oder}$$
$$(v_1^2 + v_2^2) + (w_1^2 + w_2^2) = (w_1 - v_1)^2 + (w_2 - v_2)^2$$
$$\text{ausgerechnet:} \qquad = w_1^2 + 2w_1v_1 + v_1^2 + w_2^2 + 2w_2v_2 + v_2^2$$
$$\text{Ergebnis: } (v_1|v_2) \perp (w_1|w_2) \;\Leftrightarrow\; w_1v_1 + w_2v_2 = 0$$

Man kann damit aus den Komponenten zweier Vektoren leicht ausrechnen, dass die beiden senkrecht zu einander sind. Ein *Beispiel* dazu:

Die Graphen zweier linearer Funktionen $f_1(x) = m_1 \cdot x + b_1$, $f_2(x) = m_2 \cdot x + b_2$ kann man als mit $x$ parametrisierte Geraden auffassen und eine bekannte Beziehung wiederfinden:

$g_1(x) = (x|f_1(x)) = (0|b_1) + x \cdot (1|m_1)$, $g_2(x) = (x|f_2(x)) = (0|b_2) + x \cdot (1|m_2)$.

Die beiden Graphen sind senkrecht zu einander, wenn ihre Richtungsvektoren es sind:

$(1|m_1) \perp (1|m_2) \Leftrightarrow 1 \cdot 1 + m_1 \cdot m_2 = 0 \Leftrightarrow \underline{m_1 \cdot m_2 = -1}$.

*Folgerung 1.* Ist eine Gerade durch die Gleichung $a \cdot x + b \cdot y = c$ gegeben, so gilt für ihren Richtungsvektor $(-b|a) \perp (a|b)$. Der Vektor $(a|b)$ ist also senkrecht zu der Geraden, er ist ein *Normalenvektor* der Geraden.

*Folgerung 2.* Der Abstand eines Punktes $P = (p_1|p_2)$ von einer gleichungsdefinierten Geraden $g$: $a \cdot x + b \cdot y = c$ ist gleich dem Radius des größten Kreises um $P$, der auf einer Seite von $g$ bleibt, der also $g$ als Tangente hat. Der Radius zum Berührpunkt ist *senkrecht* auf der Geraden. Um dessen Länge zu bestimmen, wird die *Normale* zu $g$ durch $P$, also die parametrisierte Gerade $n(t) = (p_1|p_2) + t \cdot (a|b)$, mit $g$ geschnitten (d.h. $t$ wird durch Einsetzen der parametrisierten Normale in die Gleichung für $g$ ausgerechnet):

$$a \cdot (p_1 + t \cdot a) + b \cdot (p_2 + t \cdot b) = c \text{ oder } t \cdot (a^2 + b^2) = c - (a \cdot p_1 + b \cdot p_2)$$

$$\text{Abstand}(P, g) = |t| \cdot \sqrt{a^2 + b^2} = \frac{|c - (a \cdot p_1 + b \cdot p_2)|}{\sqrt{a^2 + b^2}},$$

denn $\sqrt{a^2 + b^2}$ ist die Länge des Normalenvektors $(a|b)$, also $|t| \cdot \sqrt{a^2 + b^2}$ der Abstand$(P, g)$.

---

Nachdem man solche Rechnungen eine Weile gemacht hat, kann man vielleicht den Sprung schaffen, Zahlenpaare mit einer Vektorvariablen zu bezeichnen: $\vec{v} = (v_1|v_2)$. Man spart dadurch Schreibarbeit und vor allem werden die Formeln übersichtlicher. An eine ähnliche Vereinfachung bei Punkten $P = (p_1|p_2)$ wurden die Schülerinnen und Schüler außerdem schon früher gewöhnt. Parametrisierte Geraden sehen dann so aus: $g(t) = P + t \cdot \vec{v}$ und Längen von Vektoren lesen sich auch besser so $|\vec{v}|$ als so $\sqrt{v_1^2 + v_2^2}$.

**Beobachtungen, die zu Skalarprodukten führen.**

Mit dem Satz des Pythagoras wurde oben gezeigt: $\vec{v} \perp \vec{w} \Leftrightarrow v_1 \cdot w_1 + v_2 \cdot w_2 = 0$.
In Folgerung 1 haben wir den Normalenvektor $\vec{n} = (a|b)$ einer gleichungsdefinierten Geraden kennen gelernt. Bezeichnen wir Ortsvektoren $(x|y)$ oder $(x_1|x_2)$ mit $\vec{x}$, so tritt in der Geradengleichung wieder ein ebenso gebildeter Ausdruck auf: $n_1 \cdot x_1 + n_2 \cdot x_2 = c$.
In Folgerung 2 enhält die Abstandsformel wieder einen solchen Ausdruck: $a \cdot p_1 + b \cdot p_2$, besser geschrieben als $n_1 \cdot p_1 + n_2 \cdot p_2$.
Das sollte von der Wichtigkeit dieser Bildung überzeugen und folgende Definition rechtfertigen:

**Definition des Skalarproduktes zweier Vektoren** $\vec{v} = (v_1|v_2)$, $\vec{w} = (w_1|w_2) \in \mathbb{R}^2$

$$\vec{v} \bullet \vec{w} := v_1 \cdot w_1 + v_2 \cdot w_2,$$

mit den einfachen, für jedes Skalarprodukt geltenden Regeln:

1. positiv definit $\quad 0 \leq |\vec{v}|^2 = \vec{v} \bullet \vec{v}$
2. symmetrisch $\quad \vec{v} \bullet \vec{w} = \vec{w} \bullet \vec{v}$
3. linear in jedem Argument $\quad (r \cdot \vec{u} + s \cdot \vec{v}) \bullet \vec{w} = r \cdot (\vec{u} \bullet \vec{w}) + s \cdot (\vec{v} \bullet \vec{w})$.

**Beispiele für Vereinfachungen, die das Skalarprodukt bietet.**

(S1) Von $\vec{o}$ verschiedene, zu einander senkrechte Vektoren $\vec{v} \perp \vec{w}$ sind *linear unabhängig*: *Beweis:* Aus $r \cdot \vec{v} + s \cdot \vec{w} = \vec{o}$ folgt durch Skalarmultiplikation dieser Gleichung mit $\vec{v}$: $r \cdot |\vec{v}|^2 + s \cdot \vec{v} \bullet \vec{w} = r \cdot |\vec{v}|^2 + s \cdot 0 = \vec{o} \bullet \vec{v} = 0$, also $r = 0$. Ebenso folgt durch Skalarmultiplikation mit $\vec{w}$ auch $s = 0$.

(S2) Es seien $\vec{v} \perp \vec{w}$ zu einander senkrechte Einheitsvektoren, also $|\vec{v}| = 1, |\vec{w}| = 1$. Man nennt $\{\vec{v}, \vec{w}\}$ eine *Orthonormalbasis*. *Behauptung*: Für jeden Vektor $\vec{x} \in \mathbb{R}^2$ gilt:

$$\vec{x} = (\vec{x} \bullet \vec{v}) \cdot \vec{v} + (\vec{x} \bullet \vec{w}) \cdot \vec{w}$$

*Beweis:* Nach (S1)sind $\vec{v}, \vec{w}$ linear unabhängig, sodass es $x_v, x_w \in \mathbb{R}$ gibt mit $\vec{x} = x_v \cdot \vec{v} + x_w \cdot \vec{w}$. Skalarmultiplikation dieser Gleichung mit $\vec{v}$ gibt $x_v = \vec{x} \bullet \vec{v}$. Ebenso folgt $x_w = \vec{x} \bullet \vec{w}$.

Deshalb heißt das Skalarprodukt $\vec{x} \bullet \vec{v}$: *Komponente von $\vec{x}$ in Richtung $\vec{v}$.*

(S3) *Satz*: Die Summe der Quadrate der Längen der Diagonalen eines Parallelogramms ist gleich der Summe der Quadrate aller Seitenlängen. – Das liest sich in Formeln einfacher. Seien $\vec{v}, \vec{w}$ wie in Bild 1 Seite 17 die Seitenvektoren zweier benachbarter Parallelogrammseiten. Dann sind $\vec{w} - \vec{v}, \vec{w} + \vec{v}$ die Vektoren beider Diagonalen. Die Behauptung (mit Beweis) liest sich damit so:

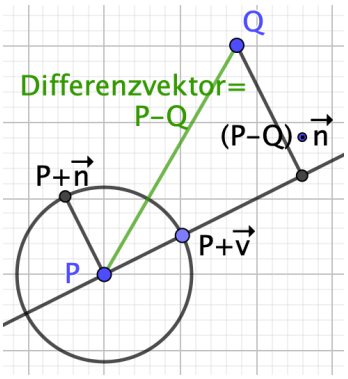
$$2|\vec{v}|^2 + 2|\vec{w}|^2 = |\vec{w} - \vec{v}|^2 + |\vec{w} + \vec{v}|^2 = |\vec{w}|^2 - 2\vec{w} \bullet \vec{v} + |\vec{v}|^2 \; + \; |\vec{w}|^2 + 2\vec{w} \bullet \vec{v} + |\vec{v}|^2 \quad \surd$$

(S4) Wie groß ist der Abstand eines Punktes $Q$ (=Ortsvektor) von einer parametrisierten Geraden $g(t) = P + t \cdot \vec{v}$ mit $|\vec{v}| = 1$ ?

Zu dem Richtungsvektor $\vec{v} = (v_1|v_2)$ gehört die Normale $\vec{n} = (-v_2|v_1)$ von $g$, ebenfalls mit $|\vec{n}| = 1$. Damit folgt:

$$\text{Abstand}(Q, g) = \big|(P - Q) \bullet \vec{n}\,\big|,$$

denn $(P-Q) \bullet \vec{n}$ ist die Komponente des Differenzvektors $P-Q$ in Richtung $\vec{n}$, also senkrecht zu $g$. Für gleichungsdefinierten Geraden $ax + by = c$ ist $\vec{n} = (a|b)/\sqrt{a^2 + b^2}$ ebenfalls bekannt.




Beachte: In allen diesen Fällen erhält man die Antwort mit Hilfe des Skalarprodukts (fast) ohne Rechnung. Man kann also schon bei der Behandlung weniger Aufgaben die Zeit wieder einsparen, die die Definition des Skalarproduktes gekostet hat. In der 3-dimensionalen Geometrie sind die Definitionen ungefähr dieselben und der Gewinn an Übersichtlichkeit und Zeitersparnis ist noch größer.

In der Elementargeometrie wird gern mit **Spiegelungen an Geraden** argumentiert. Diese Spiegelungen zu beschreiben, ist deshalb auch in der analytischen Geometrie wichtig. Die ersten Beispiele sind sehr einfach:

Spiegelung an der $x$-Achse: $(x|y) \mapsto (x|-y)$, oder $\text{Sp}_x((x|y)) = (x|-y)$.
Spiegelung an der $y$-Achse: $(x|y) \mapsto (-x|y)$, oder $\text{Sp}_y((x|y)) = (-x|y)$.
Spiegelung an der Winkelhalbierenden des 1. Quadranten: $(x|y) \mapsto (y|x) = \text{Wh}_1((x|y))$.
Spiegelung an der Winkelhalbierenden des 2. Quadranten: $(x|y) \mapsto (-y|-x)$.

Abbildungen kann man hinter einander ausführen. Z.B. wurde in der Elementargeometrie die 90°-Linksdrehung $\mathrm{D}^{90}$ durch Spiegelung an der $x$-Achse gefolgt von Spiegelung an der Winkelhalbierenden des ersten Quadranten erzeugt. Das geht jetzt auch analytisch:

$$\mathrm{D}^{90}((x|y)) = \mathrm{Wh}_1(\mathrm{Sp}_x((x|y)) = (-y|x).$$

**Abbildungsformeln für Spiegelungen an Ursprungsgeraden**

Nachdem die ersten Formeln so einfach waren,, sollen auch Abbildungsformeln für die *Spiegelung an einer gleichungsdefinierten Geraden g*: $a \cdot x_1 + b \cdot x_2 = 0$, die wie die Koordinatenachsen durch den Ursprung $O$ geht, hergeleitet werden. Wegen der guten Erfahrungen mit Einheitsvektoren nehmen wir an $a^2 + b^2 = 1$. Dann ist der Normalenvektor dieser Geraden $\vec{n} = (a|b)$, ein Richtungsvektor ist $\vec{v} = (-b|a)$, die Gleichung ist $\vec{n} \cdot \vec{x} = 0$.
**1.** Für die Spiegelung $\mathrm{Sp}_g$ an dieser Geraden gilt $\mathrm{Sp}_g \colon \vec{v} \mapsto \vec{v},\ \vec{n} \mapsto -\vec{n}$.
**2.** Ortsvektoren $\vec{x} = (x_1|x_2)$ können als Linearkombinationen von $\vec{v}, \vec{n}$ geschrieben werden.
**3.** Auch Spiegelungen sind lineare Abbildungen, respektierenn also Linearkombinationen.
Diese drei Punkte genügen, um ohne weitere Begriffe Abbildungsformeln für $\mathrm{Sp}_g$ anzugeben:

$$\begin{aligned}\mathrm{Sp}_g(\vec{x}) &= \mathrm{Sp}_g\left((\vec{x} \bullet \vec{v}) \cdot \vec{v} + (\vec{x} \bullet \vec{n}) \cdot \vec{n}\right) = (\vec{x} \bullet \vec{v}) \cdot \mathrm{Sp}_g(\vec{v}) + (\vec{x} \bullet \vec{n}) \cdot \mathrm{Sp}_g(\vec{n}) \\ &= (\vec{x} \bullet \vec{v}) \cdot \vec{v} - (\vec{x} \bullet \vec{n}) \cdot \vec{n} = \vec{x} - 2(\vec{x} \bullet \vec{n}) \cdot \vec{n}\end{aligned}$$

Diese übersichtliche **koordinatenfreie** Formel wird noch als **Matrixprodukt** geschrieben:

$$\begin{aligned}&= \begin{pmatrix} x_1 - 2(x_1 \cdot a + x_2 \cdot b) \cdot a \\ x_2 - 2(x_1 \cdot a + x_2 \cdot b) \cdot b \end{pmatrix} = \begin{pmatrix} (b^2 - a^2) \cdot x_1 - 2ab \cdot x_2 \\ -2ab \cdot x_1 + (a^2 - b^2) \cdot x_2 \end{pmatrix} \\ &= \begin{pmatrix} (b^2 - a^2) & -2ab \\ -2ab & (a^2 - b^2) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.\end{aligned}$$

Auch bei dieser Herleitung hat das Skalarprodukt hervorragende Dienste geleistet. Die Spalten der die Spiegelung beschreibenden Matrix sind zu einander senkrechte Einheitsvektoren. Üblicher ist, die Matrixkoeffizienten mit Hilfe von Winkeln hinzuschreiben. Dazu muss noch die Verbindung zwischen Skalarprodukten und sin und cos hergeleitet werden. Ehe wir dazu kommen, muss ein anderes Problem bearbeitet werden. In der Elementargeometrie sind die Spiegelsymmetrien von grundlegender Bedeutung. Deshalb muss nachgerechnet werden, dass in der analytischen Geometrie die Formeln für Spiegelungen an Geraden mit Skalarprodukten verträglich sind. Wir brauchen folgenden

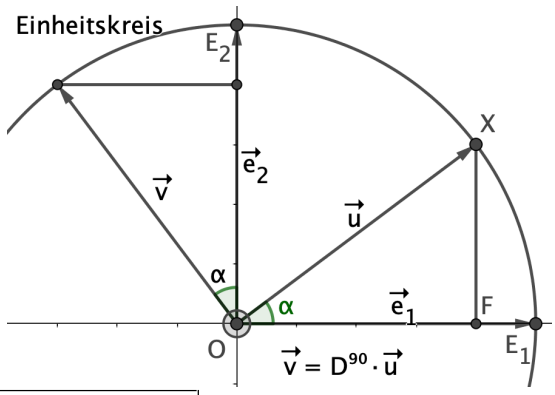
**Satz.** *Das Skalarprodukt zweier Vektoren stimmt überein mit dem Skalarprodukt ihrer Bilder unter Spiegelung an Geraden.* Es muss also gelten:

$$\vec{v} \bullet \vec{w} = \mathrm{Sp}_g(\vec{v}) \bullet \mathrm{Sp}_g(\vec{w}).$$

**Beweis.** Für die vier einfachen Spiegelungen am Anfang sieht man das ohne Rechnung. Danach müssen wir die Matrixformel für die Spiegelung benutzen, da das Skalarprodukt mit Hilfe der Vektorkomponenten in $x$- und $y$-Richtung definiert ist:

$$\begin{aligned}\mathrm{Sp}_g(\vec{v}) \bullet \mathrm{Sp}_g(\vec{v}) &= \begin{pmatrix} (b^2 - a^2) \cdot v_1 - 2ab \cdot v_2 \\ -2ab \cdot v_1 - (b^2 - a^2) \cdot v_2 \end{pmatrix} \bullet \begin{pmatrix} (b^2 - a^2) \cdot w_1 - 2ab \cdot w_2 \\ -2ab \cdot w_1 - (b^2 - a^2) \cdot w_2 \end{pmatrix} \\ &= (b^2 - a^2)^2 v_1 w_1 + 4a^2 b^2 v_2 w_2 + 4a^2 b^2 v_1 w_1 + (b^2 - a^2)^2 v_2 w_2 \\ &\quad - 2ab(b^2 - a^2) \cdot (v_1 w_2 + v_2 w_1) + 2ab(b^2 - a^2) \cdot (v_1 w_2 + v_2 w_1) \\ &= (b^2 + a^2)^2 \cdot (v_1 w_1 + v_2 w_2) = \vec{v} \bullet \vec{w}. \qquad \surd\end{aligned}$$

**Skalarprodukt und die Funktionen cos, sin.**




*Verabredungen:*
Ursprung $O = (0|0)$ des Einheitskreises.
Einheitspunkt $E_1 = (1|0) = \vec{e}_1$ der $x$-Achse.
Einheitspunkt $E_2 = (0|1) = \vec{e}_2$ der $y$-Achse.
Punkt $X = (x_1|x_2) = \vec{u}$ auf dem Einheitskreis.
Winkel $E_1OX = \alpha$ – Gradangabe nur mit Analysis aus $x_1, x_2$ ausrechenbar, ausführlich in Text A6.
$90°$ Grad Drehung $\mathrm{D}^{90}\colon \vec{u} \mapsto \vec{v},\ (x_1|x_2) \mapsto (-x_2|x_1)$.

**Definitionen und Beziehungen:**

$$\cos\alpha := x_1 = \vec{u}\bullet\vec{e}_1,\ \sin\alpha := x_2 = \vec{u}\bullet\vec{e}_2 = \vec{u}\bullet\mathrm{D}^{90}(\vec{e}_1)$$
$$-\sin\alpha = \vec{v}\bullet\vec{e}_1,\ \cos\alpha = \vec{v}\bullet\vec{e}_2 = \sin(\alpha+90°).$$

Beachte: Diese Formeln berechnen $\cos(\angle(\vec{e}_1,\vec{u}))$, $\sin(\angle(\vec{e}_1,\vec{u}))$, ohne dass der Winkel in Grad bekannt ist. Sinus und Kosinus sind Maßzahlen für die Größe von Winkeln. Wir können einige Vektoren angeben, deren Winkel mit $\vec{e}_1$ wir in Grad kennen, in den Fällen kann man sin und cos mit Skalarprodukten ausrechnen. Z.B.: Für $\vec{u} = (0{,}5|0{,}5\sqrt{3})$ ist das Dreieck $OE_1X$ gleichseitig und daher $\cos(60°) = \vec{u}\bullet\vec{e}_1 = 0{,}5$.

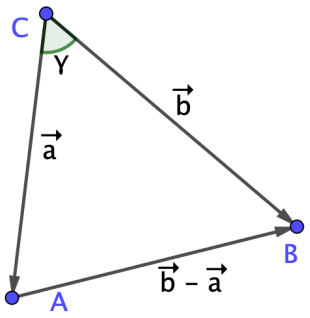
Für Vektoren $\vec{u}, \vec{v}$, die keine Einheitsvektoren sind, folgt natürlich: $\vec{u}\bullet\vec{v} = |\vec{u}|\cdot|\vec{v}|\cdot\cos(\angle(\vec{u},\vec{v}))$ und auch $\mathrm{D}^{90}(\vec{u})\bullet\vec{v} = |\vec{u}|\cdot|\vec{v}|\cdot\sin(\angle(\vec{u},\vec{v}))$.
Die Funktion $(\vec{u},\vec{v}) \mapsto \mathrm{D}^{90}(\vec{u})\bullet\vec{v} = u_1v_2 - u_2v_1 = -\mathrm{D}^{90}(\vec{v})\bullet\vec{u}$ ist wie das Skalarprodukt *linear* in jedem Vektorargument, aber wechselt das Vorzeichen bei Vertauschung der beiden Argumente. Diese Funktion heißt auch 2-dimensionale Determinante, $\det(\vec{u},\vec{v}) := \mathrm{D}^{90}(\vec{u})\bullet\vec{v}$.

**Der Kosinussatz für Dreiecke** ist eine unmittelbare Folge der Beziehung zwischen Skalarprodukt und cos. Er verallgemeinert den Satz des Pythagoras.

Der Kosinussatz berechnet die Länge $|AB| = c$ aus $|\vec{a}|$, $|\vec{b}|$ und aus $\gamma = \angle(ACB) = \angle(\vec{a},\vec{b})$:

$$\begin{aligned} c^2 &= (\vec{b}-\vec{a})\bullet(\vec{b}-\vec{a}) = |\vec{a}|^2 + |\vec{b}|^2 - 2\vec{a}\bullet\vec{b} \\ &= a^2 + b^2 - 2ab\cdot\cos\gamma. \end{aligned}$$

Zwei elementargeometrische Beweise stehen auf Seite 2 in Text G2.




Auch den **Flächeninhalt von Dreiecken** erhält man mit Hilfe des Skalarprodukts aus derselben Figur. Aus der Elementargeometrie ist bekannt *Fläche* $= \frac{1}{2}$ *Grundlinie mal Höhe*, also mit Skalarprodukten oder Koordinaten oder der Determinante ausgedrückt:

$$\mathrm{area}(ABC) = \frac{1}{2}ab\sin\gamma = \frac{1}{2}\mathrm{D}^{90}(\vec{a})\bullet\vec{b} = \frac{1}{2}(a_1b_2 - a_2b_1) = \frac{1}{2}\det(\vec{a},\vec{b}).$$

Läßt man die Faktoren $\frac{1}{2}$ weg, so erhält man den Flächeninhalt des von $\vec{a}$ und $\vec{b}$ *aufgespannten Parallelogramms.*

**Matrizen für Spiegelungen und Drehungen, beschrieben mit Winkeln.**
Wir hatten die Spiegelung $\mathrm{Sp}_g$ an einer Geraden $g$ mit Normale $\vec{n} = (a|b)$ und der Gleichung $\vec{n} \bullet \vec{x} = 0$ so beschrieben:

$$\mathrm{Sp}_g(\vec{x}) = \begin{pmatrix} (b^2 - a^2) & -2ab \\ -2ab & (a^2 - b^2) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Es ist üblich, die Gerade $g$ durch ihren Winkel $\alpha$ gegen die $x$-Achse zu beschreiben. Dann wird die Normale $\vec{n} = (-\sin\alpha|\cos\alpha)$ und die Gleichung $-\sin(\alpha) \cdot x_1 + \cos(\alpha) \cdot x_2 = 0$. Die Spiegelungsmatrix bekommt dann die Form

$$\begin{aligned} \mathrm{Sp}_g(\vec{x}) &= \begin{pmatrix} \cos^2(\alpha) - \sin^2(\alpha) & +2\sin(\alpha)\cos(\alpha) \\ +2\sin(\alpha)\cos(\alpha) & \sin^2(\alpha) - \cos^2(\alpha) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

Drehungen $\mathrm{Dr}^\alpha$ mit Drehwinkel $\alpha$ um den Ursprung werden erzeugt durch eine Spiegelung an der x-Achse ($(x_1|x_2) \mapsto (x_1|-x_2)$) gefolgt von einer Spiegelung an der Geraden mit dem *halben* Drehwinkel gegen die x-Achse. Das gibt:

$$\mathrm{Dr}^\alpha(\vec{x}) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Drehungen $\mathrm{Dr}^\alpha_P$ um einen Punkt $P = (p_1|p_2)$ bekommt man, wenn man zuerst die Parallelverschiebung von $P$ nach $O$ ausführt, dann um $O$ dreht und schließlich die Verschiebung rückgängig macht. Man bekommt:

$$\mathrm{Dr}^\alpha_P(\vec{x}) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x_1 - p_1 \\ x_2 - p_2 \end{pmatrix} + \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}.$$

Beim Zusammensetzen von Abbildungen gilt ein Assoziativgesetz:

$$B \cdot (A \cdot \vec{x}) = (B \cdot A) \cdot \vec{x}.$$

Man kann also die Matrix der Komposition linearer Abbildungen mit Matrixmultiplikation berechnen. Oder: *Maztrixmultiplikation repräsentiert die Komposition linearer Abbildungen!*
Damit beende ich den Überblick über die ebene analytische Geometrie und komme zur Raumgeometrie.

## Analytische Geometrie des 3-dimensionalen Euklidischen Raumes

Im Raum haben wir nicht die direkten Erfahrungen, die das Falten von Papier in der Ebene liefert. Wir können nicht eine Uhr in ihr Spiegelbild bewegen, wir können ihr Spiegelbild nur im Spiegel betrachten. Auch die 3-dimensionale Vorstellungskraft bleibt weit hinter der 2-dimensionalen zurück. Z.B.: Man verlängere die Seitenflächen eines nicht notwendig regulären Tetraeders zu Ebenen und frage: In wie viele Gebiete zerlegen diese vier Ebenen den Raum? Wie viele dieser Gebiete sind von Teilen aller vier Ebenen berandet? In welche dieser Gebiete kann man immer Kugeln legen, die alle vier Ebenen berühren? Wie kann man entscheiden, in welche der übrigen Gebiete man Kugeln legen kann, die alle vier Ebenen berühren? (Das Ergebnis hängt von dem Tetraeder ab.) Deshalb besteht ein guter Teil der 3-dimensionalen Geometrie darin, dass man die interessierende Figur mit geeigneten Ebenen schneidet und mit diesen Schnitten versucht, die Figur zu verstehen. Außerdem erwecken Schrägbilder und perspektive Bilder (Fotografien) ganz erfolgreich aus 2-dimensionalen Bildern 3-dimensionale Vorstellungen.

Im Gegensatz zu diesen Schwierigkeiten der anschaulichen Vorstellungen verallgemeinern sich wichtige Definitionen der ebenen analytischen Geometrie mühelos ins Dreidimensionale. Damit werde ich beginnen.

Ähnlich, wie wir die Ebene mit Quadraten gepflastert und ein Achsenkreuz hinzugefügt haben, können wir uns vorstellen, dass der Raum mit Würfeln gepflastert ist und das durch Verlängern der drei von einer Würfelecke ausgehenden Kanten drei Koordinatenachsen mit Einheitspunkten hinzugefügt sind. Damit bekommt jeder *Punkt* des Raumes drei reelle Zahlen als Koordinaten und dies Tripel betrachten wir von Anfang an als *Ortsvektor*:

$$P = (p_1|p_2|p_3) = \vec{p}.$$

*Geraden* können wir als *parametrisierte* Geraden wie im 2-dimensionalen Fall beschreiben:

$$\vec{g}(t) = P + t \cdot \vec{u}.$$

Für *parametrisierte Ebenen* werden an einen Punkt alle Linearkombinationen von zwei linear unabhängigen Vektoren angehängt:

$$\vec{x}(r,s) = P + r \cdot \vec{v} + s \cdot \vec{w}.$$

Raumpunkte haben den mit Hilfe des Pythagoras definierten Abstand und Vektoren eine Länge:

$$\text{Länge}(\vec{v})^2 = |\vec{v}|^2 = v_1^2 + v_2^2 + v_3^2,$$
$$\text{Abstand}(P,Q)^2 = |\vec{q} - \vec{p}|^2 := (q_1 - p_1)^2 + (q_2 - p_2)^2 + (q_3 - p_3)^2.$$

Wie im 2-dimensionalen Fall (Seite 18 unten) führt die Frage nach dem senkrecht Stehen von Vektoren noch einmal mit Hilfe des Pythagoras zur *Definition des Skalarproduktes*:

$$\vec{v} \bullet \vec{w} := v_1 w_1 + v_2 w_2 + v_3 w_3 \quad \text{mit} \quad \vec{v} \perp \vec{w} \Leftrightarrow \vec{v} \bullet \vec{w} = 0.$$

Wieder gelten die einfachen Regeln für Skalarprodukte:

| | |
|---|---|
| 1. positiv definit | $0 \leq \|\vec{v}\|^2 = \vec{v} \bullet \vec{v}$ |
| 2. symmetrisch | $\vec{v} \bullet \vec{w} = \vec{w} \bullet \vec{v}$ |
| 3. linear in jedem Argument | $(r \cdot \vec{u} + s \cdot \vec{v}) \bullet \vec{w} = r \cdot (\vec{u} \bullet \vec{w}) + s \cdot (\vec{v} \bullet \vec{w})$. |

Man hat dieselben unmittelbaren Folgerungen:

(S1) Paarweise auf einander senkrechte Vektoren $\vec{u}, \vec{v}, \vec{w}$ sind linear unabhängig, denn Skalarmultiplikation der Linearkombination $r \cdot \vec{u} + s \cdot \vec{v} + t \cdot \vec{w} = \vec{o}$ mit jedem der drei Vektoren liefert, dass der entsprechende Koeffizient $= 0$ ist.

(S2) Drei paarweise senkrechte Einheitsvektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ bilden eine sogenannte *Orthonormalbasis* und beliebige Vektoren $\vec{x} \in \mathbb{R}^3$ können mit Hilfe des Skalarproduktes *explizit* als Linearkombinationen geschrieben werden:

$$\vec{x} = \sum_{j=1}^{3} (\vec{x} \bullet \vec{e}_j) \cdot \vec{e}_j = |\vec{x}| \cdot \sum_{j=1}^{3} \cos(\angle(\vec{x}, \vec{e}_j)) \cdot \vec{e}_j$$

$$\text{Bekannte Folgerung: } \forall_{\vec{x} \in \mathbb{R}^3} \Rightarrow \sum_{j=1}^{3} \cos(\angle(\vec{x}, \vec{e}_j))^2 = 1.$$

Für die Grundaufgaben, Geraden und Ebenen bequem auch durch Gleichungen zu beschreiben und mit einander zu schneiden, fehlt uns nur noch ein Hilfsmittel:

*Wie findet man leicht zu zwei gegebenen linear unabhängigen Vektoren einen dritten, der auf beiden senkrecht steht?*


Diese Frage hat viele Leute so beschäftigt, dass Wikipedia ein besonders schönes Beispiel für dies Interesse dem Artikel "Rechte Hand Regel" hinzugefügt hat. Das Bild zeigt: Wenn man Daumen und Zeigefinger an die beiden gegebenen Vektoren hält, dann zeigt der Mittelfinger, wie man den zu ihnen senkrechten Vektor wählen sollte. Kann man das mathematisch nachmachen? Da wir Linearkombinationen beherrschen, müssen wir nur *aus dem Bild ablesen*, wie man zu Paaren von Basisvektoren den dritten findet. Aber das ist einfach:

$$(\vec{e}_1, \vec{e}_2) \mapsto +\vec{e}_3,\ (\vec{e}_2, \vec{e}_3) \mapsto +\vec{e}_1,\ (\vec{e}_3, \vec{e}_1) \mapsto +\vec{e}_2,$$
$$(\vec{e}_2, \vec{e}_1) \mapsto -\vec{e}_3,\ (\vec{e}_3, \vec{e}_2) \mapsto -\vec{e}_1,\ (\vec{e}_1, \vec{e}_3) \mapsto -\vec{e}_2,$$
$$(\vec{e}_j, \vec{e}_j) \mapsto \vec{o}$$

Wir erhalten eine Abbildung, die in der 3-dimensionalen Geometrie ähnlich hilfreich ist, wie die 90°-Drehung in der Ebene (beachte die Linearität im ersten und zweiten Argument):

$$\Big(\sum_{j=1}^{3} x_j \vec{e}_j, \sum_{k=1}^{3} y_k \vec{e}_k\Big) \mapsto (x_1 y_2 - x_2 y_1) \cdot \vec{e}_3 + (x_2 y_3 - x_3 y_2) \cdot \vec{e}_1 + (x_3 y_1 - x_1 y_3) \cdot \vec{e}_2$$

Und weil dieses **Kreuzprodukt** so wichtig ist, noch einmal in Matrixschreibweise:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}, \text{ wie gewünscht: } \vec{x} \times \vec{y} \perp \vec{x}, \vec{y}.$$

Um auch noch eine Beziehung zur 90°-Drehung zu betonen: Für $\vec{u} \perp \vec{e}$ ist $\vec{u} \mapsto \vec{e} \times \vec{u}$ die 90°-Drehung in der Ebene senkrecht zu $\vec{e}$. Die Richtung von $\vec{e}$ definiert die Oberseite dieser Ebene. Deshalb zeigt die Formel auch: Wenn man am Nordpol steht, dreht sich die Erde links herum, am Südpol dagegen rechts herum!

**Nun zu den Grundaufgaben.**

**Wie findet man Gleichungen zu Parametrisierungen?**
Die *parametrisierte Ebene* $\vec{x}(r,s) = P + r \cdot \vec{v} + s \cdot \vec{w}$ hat $\vec{n} := \vec{v} \times \vec{w}$ als Normalenvektor. Skalarmultiplikation mit $\vec{n}$ liefert die

**lineare Gleichung dieser Ebene** $\vec{x} \bullet \vec{n} = P \bullet \vec{n} = n_1x_1 + n_2x_2 + n_3x_3 = \text{const.}$

Umgekehrt kann man aus der Ebenengleichung eine Normale direkt ablesen und man kann eine Parametrisierung finden, indem man durch Lösen der Gleichung drei Punkte der Ebene bestimmt.
Der Schnitt von zwei Ebenen wird durch zwei lineare Gleichungen in drei Variablen beschrieben. Wie immer gibt es die drei Möglichkeiten: Dies System hat keine Lösung, falls die Ebenen parallel und verschieden sind. Das System hat eine 2-dimensionale Lösungsmenge, wenn die beiden Ebenen übereinstimmen. Und in den meisten Fällen ist die Lösungsmenge 1-dimensional, ist also eine Gerade.

Wie findet man zu einer *parametrisierten Gerade* $\vec{g}(t) = P + t \cdot \vec{u}$ zwei sie beschreibende lineare Gleichungen? Die Ortsvektoren $\vec{x}$ der Punkte dieser Geraden sind Lösungen des Gleichungssystems

$$\vec{x} \times \vec{u} = P \times \vec{u}.$$

Das sind *drei* lineare Gleichungen in den drei Unbekannten $x_1, x_2, x_3$. Wieso hat man eine 1-dimensionale Lösungsmenge? Die drei Gleichungen sind *linear abhängig*, denn Skalarmultiplikation der Vektorgleichung mit $\vec{u}$ liefert $\vec{o} = \vec{o}$.

**Wie schneidet man Geraden mit Ebenen?**
Entweder sind *beide parametrisiert*: $\vec{g}(t) = P + t \cdot \vec{u}$ und $\vec{x}(r,s) = P + r \cdot \vec{v} + s \cdot \vec{w}$. Dann ist die Vektorgleichung $\vec{g}(t) = \vec{x}(r,s)$ ein 3-dimensionales Gleichungssystem in den drei Unbekannten $r, s, t$. Wenn die Gerade in der Ebene liegt, ist die Lösungsmenge 1-dimensional. Wenn die Gerade parallel zu der Ebene aber nicht in ihr liegt, ist das homogene System 2-dimensional, das inhomogene aber 3-dimensional, also der unlösbare Fall. Und falls die Lösung eindeutig bestimmt ist, gibt sie den Schnittpunkt.

Oder *Gerade und Ebene sind gleichungsdefiniert.* Dann hat man drei lineare Gleichungen für den möglichen Schnittpunkt und es gibt wieder die drei eben aufgezählten Möglichkeiten.

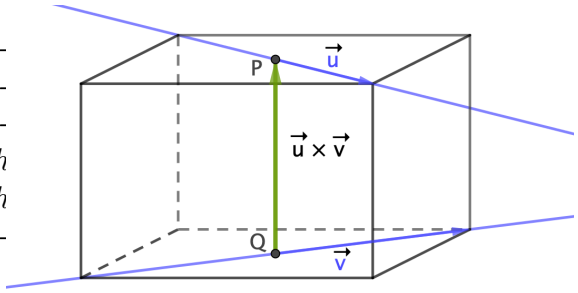
Oder die *Gerade ist parametrisiert, die Ebene gleichungsdefiniert.* Das ist der einfachste Fall: Man setzt die Gerade in die Ebenengleichung ein und bekommt *eine* lineare Gleichung für den Geradenparameter $t$ - wieder mit den drei Möglichkeiten.

Oder schließlich die *Gerade ist gleichungsdefiniert, die Ebene parametrisiert.* Dann muss man die Ebenenparametrisierung in die beiden Gleichungen für die Gerade einsetzen und bekommt zwei lineare Gleichungen für die beiden Ebenenparameter $r, s$, usw..

**Was kann mit zwei Geraden passieren?**
Zwei Geraden im Raum haben vier Möglichkeiten: Sie können zusammenfallen, sie können parallel sein, sie können sich schneiden oder sie können sich nicht treffen. In den ersten drei Fällen liegen die Geraden in einer Ebene und die Möglichkeiten sind schon in der ebenen Geometrie diskutiert. Ein einfaches Beispiel für den letzten Fall sind die *nicht* parallelen Diagonalen $\vec{d_1}(s) = P + s \cdot \vec{u}$, $\vec{d_2}(t) = Q + t \cdot \vec{v}$ in zwei gegenüber liegenden Quaderflächen, Bild nächste Seite. Solche Geraden heißen **windschief**.

**Behauptung.**

*Zu zwei windschiefen Geraden gibt es immer zwei parallele, verschiedene Ebenen, in denen die beiden Geraden liegen. Wenn man in Richtung der Ebenennormalen auf eine Ebene projiziert, so erhält man zwei sich schneidende Geraden. Und die Ebenennormale durch diesen Schnittpunkt ist die eindeutig bestimmte gemeinsame Normale der beiden windschiefen Geraden.*




**Beweis.**
Gegeben zwei windschiefe Geraden $\vec{g_1}(s) = P + s \cdot \vec{u}$, $\vec{g_2}(t) = Q + t \cdot \vec{v}$. Das bedeutet, dass $\vec{u}, \vec{v}$ linear unabhängig sind und die Geraden keinen gemeinsamen Punkt haben. Ein Vektor, der senkrecht zu beiden Geraden ist, muss proportional zum Kreuzprodukt der Richtungsvektoren sein, wir nehmen $\vec{n} := (\vec{u} \times \vec{v})/|\vec{u} \times \vec{v}|$. Dann liegt die Gerade $g_1$ in der Ebene $E_1 : \vec{x} \bullet \vec{n} = P \bullet \vec{n}$ und die Gerade $g_2$ liegt in der parallelen Ebene $E_2 : \vec{x} \bullet \vec{n} = Q \bullet \vec{n}$. Der *Abstand dieser Ebenen* ist der Betrag von $d$:

$$d := (Q - P) \bullet \vec{n}, \quad \text{Abstand}(E_1, E_2) = |d| \ .$$

Damit läßt sich die Gerade $g_1$ in Richtung der Ebenennormalen parallel verschieben, sodass sie als $g_1^*$ in die Ebene $E_2$ kommt und dort $g_2$ in einem Endpunkt $F := g_1^* \cap g_2$ des gemeinsamen Lotes von $g_1$ und $g_2$ schneidet.

$$\text{Parallel verschobene Gerade: } \vec{g_1^*}(s) = P + d \cdot \vec{n} + s \cdot \vec{u}.$$

Damit ist die Situation des "einfachen" Anfangsbeispiels praktisch erreicht (der Quader kann leicht hinzugefügt werden) und die allgemeine Situation windschiefer Geraden ist hergeleitet.

**Genauere Untersuchung des Kreuzproduktes.**
Wir wissen bisher nur $\vec{v} \times \vec{w} \perp \vec{v}, \vec{w}$, aber noch nichts über die Länge. Wir kennen nur die Proportionalität $r\vec{v} \times s\vec{w} = rs \cdot \vec{v} \times \vec{w}$. Als erstes möchte ich nachrechnen, dass das Kreuzprodukt von auf einander senkrechten Einheitsvektoren wieder ein Einheitsvektor ist. Das senkrecht Stehen wird so ausgenutzt:

$$\vec{v} \perp \vec{w} \ \Leftrightarrow \ |\vec{v} \bullet \vec{w}|^2 = 0 \ \Leftrightarrow \ (v_1w_1 + v_2w_2 + v_3w_3)^2 = 0$$
$$\Leftrightarrow (v_1^2w_1^2 + v_2^2w_2^2 + v_3^2w_3^2) = -2(v_1v_2w_1w_2 + v_2v_3w_2w_3 + v_3v_1w_3w_1).$$

Damit liefert die Formel von Seite 23 für das Kreuzprodukt, also

$$\vec{v} \times \vec{w} = (v_2w_3 - v_3w_2|v_3w_1 - v_1w_3|v_1w_2 - v_2w_1), \text{ die Behauptung:}$$
$$\vec{v} \perp \vec{w} \ \Rightarrow |\vec{v} \times \vec{w}|^2 = (v_1^2 + v_2^2 + v_3^2) \cdot (w_1^2 + w_2^2 + w_3^2) = |\vec{v}|^2 \cdot |\vec{w}|^2.$$

*Folgerung.* Für je zwei auf einander senkrechte Einheitsvektoren $\vec{v}, \vec{w}$ ist $\{\vec{v}, \vec{w}, \vec{v} \times \vec{w}\}$ eine (rechtshändige) Orthonormalbasis.
Als nächstes betrachte mit $\vec{z} = \vec{w} + a \cdot \vec{v}$ (und zuerst $|\vec{v}| = |\vec{w}| = 1$) das Kreuzprodukt

$$\vec{v} \times \vec{z} = \vec{v} \times \vec{w} \ \text{ mit } \ |\vec{z}|^2 = 1 + a^2, \ \ \sin(\angle(\vec{v}, \vec{z}))^2 = \frac{|\vec{w}|^2}{|\vec{z}|^2} = (1 + a^2)^{-1},$$

also $\quad |\vec{v} \times \vec{z}|^2 = |\vec{v}|^2 \cdot |\vec{z}|^2 \cdot \sin(\angle(\vec{v}, \vec{z}))^2 = \text{area}(\text{parallelogram}(\vec{v}, \vec{z}))^2.$

In Worten: *Die Länge des Kreuzproduktes zweier Vektoren ist gleich dem Flächeninhalt des von den Vektoren aufgespannten Parallelogramms.*
Diese Beziehung ist so wichtig, dass ich sie noch auf ganz andere Weise zeigen möchte.

In der Ebene hat das Argumentieren mit linearen Abbildungen (z.B. mit der Drehung $\mathrm{D}^{90}$) sehr geholfen. Im nächsten Beweis will ich Orthogonalprojektionen auf die Koordinatenebenen benutzen. Das sind Abbildungen mit denen Architekten und Ingenieure Grundrisse und Aufrisse ihrer Objekte zeichnen, damit man deren Maße ablesen und die Objekte herstellen kann. Gegeben sei die Orthonormalbasis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ eines Koordinatensystems.

**Definition der senkrechten Projektionen auf die Koordinatenebenen:**

$$\mathrm{OP}_3(\vec{x}) := \vec{x} - (\vec{x} \bullet \vec{e}_3)\cdot\vec{e}_3 = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix},\ \mathrm{OP}_2(\vec{x}) := \vec{x} - (\vec{x} \bullet \vec{e}_2)\cdot\vec{e}_2 = \begin{pmatrix} x_1 \\ 0 \\ x_3 \end{pmatrix},\ \mathrm{OP}_1(\vec{x}) := \begin{pmatrix} 0 \\ x_2 \\ x_3 \end{pmatrix}.$$

Für das Kreuzprodukt hatten wir diese Koordinatenbeschreibung:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Das von $\vec{x}$ und $\vec{y}$ *aufgespannte Parallelogramm* $P_{xy}$ wird in die drei Koordinatenebenen projiziert. Die projizierten Parallelogramme werden aufgespannt von $\mathrm{OP}_j(\vec{x})$, $\mathrm{OP}_j(\vec{y})$ für $j = 1, 2, 3$. Ihre Flächeninhalte seien $A_j := \mathrm{area}(\mathrm{OP}_j(P_{xy}))$.
Auf Seite 21 unten wurde hergeleitet:

$$A_1 = |x_2y_3 - x_3y_2|, \quad A_2 = |x_3y_1 - x_1y_3|, \quad A_3 = |x_1y_2 - x_2y_1|,$$

also $\quad A_1^2 + A_2^2 + A_3^2 = |\vec{x} \times \vec{y}|^2.$

Es bleibt noch, die geometrische Eigenschaft $A_1^2 + A_2^2 + A_3^2 = \mathrm{area}(P_{xy})^2$ zu zeigen, um erneut den *Betrag des Kreuzproduktes als gleich der Fläche des aufgespannten Parallelogramms* zu beweisen. Dazu rechnen wir aus, um welchen Faktor die Projektionen $\mathrm{OP}_j$ die Fläche von $P_{xy}$ verkleinern.
Projiziere also die Ebene $E_P$ von $P_{xy}$ mit Einheitsnormale $\vec{n}$ in Richtung eines Einheitsvektors $\vec{e}$ auf eine Ebene $E_K$ senkrecht zu $\vec{e}$. Die Schnitte von $E_P$ mit zu $E_K$ parallelen Ebenen heißen *Höhenlinien* von $E_P$. Sie werden durch die Projektion ohne Längenverkürzung auf $E_K$ projiziert. Die zu den Höhenlinien *senkrechten* Geraden in $E_K$ heißen *Falllinien* von $E_K$. Sie werden durch die Projektion verkürzt und der Verkürzungsfaktor ist gleich dem Verkleinerungsfaktor für Flächeninhalte bei der Projektion. Der Winkel $\alpha = \angle(\vec{e}, \vec{n})$ zwischen den beiden Ebenennormalen ist auch der *Schnittwinkel der beiden Ebenen.* Der Verkürzungsfaktor der Falllinien ist also gleich $\cos\alpha = \cos(\angle(\vec{e}, \vec{n}))$.
Das gilt jetzt für die drei Projektionen $\mathrm{OP}_j$ und wir haben

$$A_j = \mathrm{area}(P_{xy}) \cdot \cos(\angle(\vec{e}_j, \vec{n})), \quad j = 1, 2, 3$$

Auf Seite 24 oben wurde gezeigt $\sum_{j=1}^{3} \cos^2(\angle(\vec{e}_j, \vec{n})) = 1$. Damit ist die fehlende geometrische Eigenschaft $A_1^2 + A_2^2 + A_3^2 = \mathrm{area}(P_{xy})^2$ bewiesen.

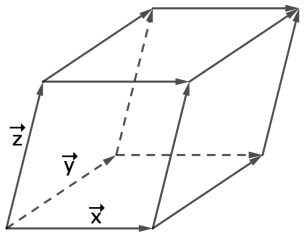
*Spezialfall.* Schneidet man von einem Würfel eine Ecke ab, so entsteht ein sogenanntes *rechtwinkliges Tetraeder.* Eine als "Pythagoras für Flächeninhalte" bezeichnete Aussage ist:

> *Die Quadratsumme der Flächen der drei rechtwinkligen Seitendreiecke eines rechtwinkligen Tetraeders ist gleich dem Quadrat der vierten Fläche.*

**Spatvolumina und Determinanten**

So wie zwei linear unabhängige Vektoren ein Parallelogramm aufspannen, so spannen drei linear unabhängige Vektoren $\vec{x}, \vec{y}, \vec{z}$ einen sogenannten Spat auf:

Definition des Spats: $\{r \cdot \vec{x} + s \cdot \vec{y} + t \cdot \vec{z};\ 0 \le r, s, t \le 1\}$




Falls die Vektoren paarweise senkrecht sind, ist der Spat ein Quader.
Das von $\vec{x}, \vec{y}$ aufgespannte Parallelogramm betrachten wir als *Grundfläche*, die Komponente von $\vec{z}$ senkrecht zur Grundfläche als *Höhe*. Das Volumen des Spats ist

*Grundfläche mal Höhe.*

Da wir den Flächeninhalt der Grundfläche mit Hilfe des Kreuzproduktes als einen senkrecht zur Grundfläche stehenden Vektor berechnen, brauchen wir diesen Vektor nur skalar mit $\vec{z}$ zu multiplizieren, um *Grundfläche mal Höhe* zu erhalten:

$$\text{Spatvolumen } = |(\vec{x} \times \vec{y}) \bullet \vec{z}\,|$$

$$\begin{aligned}(\vec{x} \times \vec{y}) \bullet \vec{z} &= (x_2y_3 - x_3y_2)z_1 + (x_3y_1 - x_1y_3)z_2 + (x_1y_2 - x_2y_1)z_3 \\ &= (x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_1z_2y_3 - x_2z_3y_1 - x_3z_1y_2\end{aligned}$$

Die trilineare Funktion $(\vec{x}, \vec{y}, \vec{z}) \mapsto (\vec{x} \times \vec{y}) \bullet \vec{z}$ ändert nach Definition des Kreuzproduktes ihr Vorzeichen, wenn man die Vektoren $\vec{x}$ und $\vec{y}$ vertauscht. *Aber sie hat mehr solche Symmetrien.* Die letzte Zeile des Ausdrucks in Koordinaten zeigt, dass die positiven Terme mit den negativen getauscht werden, wenn man die Vektoren $\vec{y}$ und $\vec{z}$ vertauscht. Man kann daraus folgern oder ebenfalls aus dem Koordinatenausdruck ablesen: Auch wenn man die Vektoren $\vec{x}$ und $\vec{z}$ vertauscht, ändert sich das Vorzeichen. Man kann also zu jedem Argument Linearkombinationen der *beiden anderen* Argumente addieren, ohne dass sich der Wert der Funktion $(\vec{x}, \vec{y}, \vec{z}) \mapsto (\vec{x} \times \vec{y}) \bullet \vec{z}$ ändert. Diese Funktion heißt "Determinante".

Damit haben wir das oft benutzte **Kriterium:**

`Die Determinante von drei Vektoren ist genau dann` $= 0$`,`
`wenn diese linear abhängig sind` – in Übereinstimmung mit der geometrischen Interpretation als Volumen.

**Die dreidimensionale Determinante:**

$$(\vec{x}, \vec{y}, \vec{z}) \ \mapsto\ (\vec{x} \times \vec{y}) \bullet \vec{z} = \det(\vec{x}, \vec{y}, \vec{z}) = \det(\vec{y}, \vec{z}, \vec{x}) =$$
$$x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2\ - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2$$

Die Determinante ist positiv für rechtshändige und negativ für linkshändige Tripel – das sind Tripel von Vektoren, die man mit der rechten bzw. der linken Hand veranschaulichen kann.

## Spiegelungen und Drehungen im Euklidischen Raum

Bisher war es nicht nötig, die Koeffizienten eines Vektors bezüglich einer Basis in die Koeffizienten bezüglich einer anderen Basis umzurechnen. Da diese sogenannten Basiswechsel aber im Zusammenhang mit linearen Abbildungen nützlich sind, brauche ich vorher noch eine **Abschweifung:**
So wie auf Seite 14 werden Lineare Abbildungen meistens mit Matrizen $A$ beschrieben, mit denen Vektoren so abgebildet werden, dass Spaltenmatrizen aus reellen Zahlen mit $A$ via Matrixmultiplikation in Spaltenmatrizen aus reellen Zahlen verwandelt werden. *Aber wo sieht man da die Vektoren?* Die Beschreibung einer linearen Abbildung durch eine Matrix *setzt voraus*, dass eine (oft orthonormale) Basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ gewählt wurde. Jeder Vektor $\vec{x}$ ist dann in eindeutiger Weise Linearkombination dieser Basisvektoren und die *Koeffizienten* dieser Linearkombination liefern die Spalten reeller Zahlen, die mit $A$ multipliziert werden:

$$\vec{x} = x_1 \cdot \vec{e}_1 + x_2 \cdot \vec{e}_2 + x_3 \cdot \vec{e}_3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 + a_{13} \cdot x_3 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + a_{23} \cdot x_3 \\ a_{31} \cdot x_1 + a_{32} \cdot x_2 + a_{33} \cdot x_3 \end{pmatrix}$$

$$\begin{aligned} \text{Bild}(\vec{x}) = & (a_{11} \cdot x_1 + a_{12} \cdot x_2 + a_{13} \cdot x_3) \cdot \vec{e}_1 \\ & + (a_{21} \cdot x_1 + a_{22} \cdot x_2 + a_{23} \cdot x_3) \cdot \vec{e}_2 \\ & + (a_{31} \cdot x_1 + a_{32} \cdot x_2 + a_{33} \cdot x_3) \cdot \vec{e}_3 \end{aligned}$$

Die Bildspalte ist dann noch nicht der Bildvektor von $\vec{x}$, sondern sie enthält die Koeffizienten der Linearkombination des Bildvektors bezüglich der gewählten Basis. Die Spalten der Matrix werden oft als Bilder der Basisvektoren bezeichnet, denn der erste Basisvektor hat die Koeffizienten $x_1 = 1, x_2 = 0, x_3 = 0$, sodass $\text{Bild}(\vec{x}_1)$ die Linearkombination

$$\text{Bild}(\vec{x}_1) = a_{11} \cdot \vec{e}_1 + a_{21} \cdot \vec{e}_2 + a_{31} \cdot \vec{e}_3$$

ist und ebenso für die anderen Basisvektoren. Die Spalten der Matrix sind also nicht wirklich die *Bilder* der Basisvektoren, aber sie sind deren Koeffizienten bezüglich der gewählten Basis. (Das schließt den Fall ein, dass die lineare Abbildung von einem Vektorraum in einen (eventuell) anderen geht und natürlich auch im Bildraum eine Basis gewählt sein muss.)

Falls der Vektorraum der $\mathbb{R}^n$ ist, ist dieser Unterschied zwischen Vektoren und Koeffizienten-Spalten nicht leicht zu verstehen, denn die Vektoren und die Spalten sehen einfach gleich aus. In einem Vektorraum aus Polynomen, z.B. vom Grad $\leq 3$, ist der Unterschied leicht zu sehen: Die (üblichen) Basisvektoren sind die Polynome $p_0(x) = 1, p_1(x) = x, p_2(x) = x^2, p_3(x) = x^3$ und die Koeffizientenspalte $(a_0|a_1|a_2|a_3)$ aus reellen Zahlen liefert das Polynom als Linearkombination:

$$p(x) = a_0 \cdot p_0(x) + a_1 \cdot p_1(x) + a_2 \cdot p_2(x) + a_3 \cdot p_3(x) \ .$$

Die lineare Abbildung "Differenzieren": $p \mapsto p'$ wird dann so beschrieben:

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Obwohl also Polynome differenziert werden, sieht man davon in der Beschreibung durch Matrixmultiplikation nichts mehr. Daher ist es wichtig, den Unterschied zwischen Vektoren und ihren Koeffizienten-Spalten bezüglich einer Basis zu verstehen. – Ende der Abschweifung.

Spiegelungen und Drehungen, die den Ursprung $O \in \mathbb{R}^3$ fest lassen, sind lineare Abbildungen der Ortsvektoren. Wenn ein anderer Punkt $P$ fest bleiben soll, wird das wie in $\mathbb{R}^2$ mit Hin- und Herschieben gemacht:

$$\vec{x} \mapsto A \cdot (\vec{x} - \vec{p}) + \vec{p}.$$

180° Drehungen um Geraden mit Richtungsvektor $\vec{v}$ der Länge 1, auch Spiegelungen an Geraden genannt, sind:

$$\begin{aligned} \mathrm{D}_{\vec{v}}^{180}(\vec{x}) = -\vec{x} + 2(\vec{v} \bullet \vec{x}) \cdot \vec{v} &= \begin{pmatrix} -x_1 + 2(v_1x_1 + v_2x_2 + v_3x_3) \cdot v_1 \\ -x_2 + 2(v_1x_1 + v_2x_2 + v_3x_3) \cdot v_2 \\ -x_3 + 2(v_1x_1 + v_2x_2 + v_3x_3) \cdot v_3 \end{pmatrix} \\ &= \begin{pmatrix} -1 + 2v_1^2 & 2v_1v_2 & 2v_1v_3 \\ 2v_1v_2 & -1 + 2v_2^2 & 2v_2v_3 \\ 2v_1v_3 & 2v_2v_3 & -1 + 2v_3^2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

Spiegelungen an Ebenen mit Einheitsnormalenvektor $\vec{n}$ (Ausrechnung der Matrix wie eben):

$$\mathrm{Sp}_{\vec{n}}(\vec{x}) = \vec{x} - 2(\vec{n} \bullet \vec{x}) \cdot \vec{n} = \begin{pmatrix} +1 - 2n_1^2 & -2n_1n_2 & -2n_1n_3 \\ -2n_1n_2 & +1 - 2n_2^2 & -2n_2n_3 \\ -2n_1n_3 & -2n_2n_3 & +1 - 2n_3^2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Falls der Vektoren $\vec{n} = (-\sin\alpha|\cos\alpha|0)$ in der $x_1$-$x_2$-Ebene liegt oder der Vektor $\vec{v}$ senkrecht zu dieser Ebene ist, kann man die Matrixbeschreibung von Seite 22 leicht ausdehnen, denn die dritte Koordinate von $\vec{x}$ bleibt ungeändert.
Spiegelung an der Ebene mit Normale $\vec{n} \perp \vec{e}_3$:

$$\vec{x} \mapsto \begin{pmatrix} \cos 2\alpha & \sin 2\alpha & 0 \\ \sin 2\alpha & -\cos 2\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Drehung um den Winkel $\alpha$ um die $z$-Achse (Richtungsvektor $= \vec{e}_3$):

$$\mathrm{Dr}^{\alpha}(\vec{x}) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Um die Matrix für eine Drehung mit dem Richtungsvektor $\vec{w}$ zu erhalten, braucht man die 180°-Drehung um die Gerade mit Richtungsvektor $\vec{v} := (\vec{v} + \vec{e}_3)/|\vec{v} + \vec{e}_3|$ (Winkelhalbierende von $\vec{v}, \vec{e}_3$), um die Drehachse zunächst in die $z$-Achse zu bewegen. Nach Anwendung der Drehung $\mathrm{Dr}^{\alpha}$ wird die Drehachse wieder in die Ausgangslage zurück bewegt, also:

$$\mathrm{Dr}_{\vec{w}}^{\alpha}(\vec{x}) = \mathrm{D}_{\vec{v}}^{180} \circ \mathrm{Dr}^{\alpha} \circ \mathrm{D}_{\vec{v}}^{180}(\vec{x}).$$

Damit stehen zwei Beschreibungen für Drehungen und Spiegelungen des $\mathbb{R}^3$ zur Verfügung.

---

Meiner Meinung nach besitzt man jetzt die Voraussetzungen, um in den Bibliotheken die Mathematikbücher zur Linearen Algebra und Analytischen Geometrie gut lesen zu können.