

Wolfgang K. Seiler

Mathematische Visualisierung

Vorlesung an der Universität Mannheim
im Frühjahrssemester 2012

Kapitel 1

Visuelle Information

Wissenschaftliche Untersuchungen, Simulationen, Umfragen, Rechnungswesen und vieles mehr produzieren oft Daten in einem solchen Umfang, daß ein Mensch Wochen oder gar Jahre bräuchte, um sie auch nur zu lesen; um Zusammenhänge zu erkennen oder Hypothesen zu entwickeln und zu testen ist er praktisch immer auf Computerhilfe angewiesen. Computer können große Datenmengen sowohl speichern als auch nach festgelegten Kriterien verarbeiten; mit dem *Erkennen* unbekannter Zusammenhänge haben aber auch sie immer noch große Probleme.

Wir können solche Zusammenhänge anhand von Tabellen zwar auch nur in den seltensten Fällen erkennen, bei bildlichen Darstellungen allerdings springen sie uns oft direkt ins Auge. In diesem Punkt sind wir zumindest heute noch jedem Computer weit überlegen; ein einziger Blick – so zumindest unser Eindruck – verschafft uns eine Fülle von weitgehend sogar bereits aufbereiteten Informationen.

Die graphische Darstellung auch umfangreicher Datenmengen nach von uns festzulegenden Kriterien ist eine Aufgabe, die Computer problemlos meistern; in der mathematischen Visualisierung versuchen wir daher, vorhandene Datenmengen graphisch so aufzubereiten, daß wir unbekannte Zusammenhänge erkennen und bekannte Zusammenhänge einem von uns verschiedenen Betrachter vermitteln können.

Zur Vorbereitung darauf wollen wir uns in diesem ersten Kapitel zunächst mit den Möglichkeiten und Grenzen unseres visuellen Systems beschäftigen, danach mit den Techniken, die heute gebräuchliche Computergraphiksysteme zur Produktion von Bildern verwenden.

§ 1: Unser visuelles System

Unsere Augen zeigen uns *auf einen Blick* alles, was sich in unserem Gesichtsfeld befindet – zumindest haben wir diesen Eindruck. Tatsächlich aber können unsere Augen zu einem festen Zeitpunkt schon aus physiologischen Gründen nur einen kleinen Bruchteil unseres Gesichtsfelds abbilden, und selbst diesen Bruchteil nehmen wir oft nicht bewußt wahr.

Unsere Sehzellen sind nämlich nicht gleichmäßig über die Netzhaut verteilt, sondern konzentrieren sich massiv an einer Stelle. Nur durch (größtenteils unbewußte) Augenbewegungen schaffen wir es, das Licht von ständig wechselnden Teilen des Gesichtsfelds dorthin zu lenken, und erst unser Gehirn setzt die einzelnen Szenen nötigenfalls zusammen zu einem Gesamtbild. Wie Filmproduzenten schon seit langer Zeit ausnutzen, nehmen wir keinesfalls stets das gesamte Bild bewußt wahr; die Kontinuitätsfehler, die beim Zusammenschnitt einer Szene aus vielen, teilweise an verschiedenen Tagen gedrehten, Einstellungen praktisch immer entstehen, werden später von kaum einem Betrachter des Films wahrgenommen.

Die beiden Psychologen DANIEL SIMONS und DANIEL LEVIN testeten dies Mitte der Neunzigerjahre in einem kontrollierten Experiment, in dem sie von einem Schnitt zum nächsten sogar den Schauspieler durch einen anders aussehenden und anders gekleideten ersetzten; selbst dies bemerkte nur ein Drittel der Probanden. Kleinere Änderungen wie etwa das Verschwinden eines Schals blieben gar 90% aller Betrachter verborgen; siehe

DANIEL T. LEVIN, DANIEL J. SIMONS: Failure to detect changes to attended objects in motion pictures, *Psychonomic Bulletin & Review* 4 (4), 501–506, 1997.

In einem weiteren Experiment befragte ein Schauspieler mit einem Campusplan von Cornell in der Hand Passanten nach dem Weg. Mitten in der Diskussion gingen zwei weitere als Handwerker verkleidete Schauspieler, die eine Tür trugen, zwischen den beiden hindurch; währenddessen wurde der Frager durch einen anderen Schauspieler ersetzt. In mehr als

der Hälfte aller Fälle fiel dies den auskunftgebenden Passanten nicht auf; in einem zweiten Durchgang, bei dem die beiden Frager als völlig verschieden gekleidete Bauarbeiter auftraten, bemerkte gar nur ein Drittel den Wechsel; siehe

DANIEL J. SIMONS, DANIEL T. LEVIN: Failure to detect changes to people during a real-world interaction, *Psychonomic Bulletin & Review* 5 (4), 644–649, 1998.

Die Erklärung liegt natürlich darin, daß die auskunftgebenden Passanten sich auf den Campusplan und den Weg konzentrierten. Auch wenn sie sicherlich nicht den Eindruck hatten, irgendetwas in ihrer Umgebung nicht zu sehen, achteten sie schlichtweg nicht auf das Erscheinungsbild des Fragers und nahmen ihn deshalb auch nicht wahr. Hätte man sie später danach gefragt, hätten sie wahrscheinlich gesagt, daß sie sich nicht erinnern könnten. Bei den Bauarbeitern kam wohl dazu, daß sie einfach als solche klassifiziert wurden, genau wie ja auch viele Menschen das Verkaufspersonal eines Geschäfts auf der Straße nur schwer erkennen.

Als erste Konsequenz für die Visualisierung können wir daraus ableiten, daß ein Betrachter wohl auch bei der Aufnahme visualisierter Information nicht unbedingt jeden Aspekt unserer Graphiken automatisch zur Kenntnis nimmt. Der amerikanische Statistiker EDWARD ROLF TUFTE (*1942) leitet daraus ab, daß wir die *data-ink ratio*, also das Verhältnis von (darzustellenden) Daten und dazu verwendeter Tinte, maximieren sollten, daß wir also einen gegebenen Datensatz mit möglichst *wenig* Tinte visualisieren sollten. Das bedeutet vor allem den Verzicht auf überflüssige Effekte und Schnörkel, die zwar das Auge des Betrachters anziehen, aber keinen Beitrag zur Aufnahme der eigentlich relevanten Information leisten. Natürlich haben auch solche überfrachteten Graphiken ihren Nutzen: Wenn sie gut gemacht sind, können sie das Auge des Betrachters auf sich ziehen und ihn so überhaupt erst beispielsweise für den Artikel interessieren, zu dem die Graphik gehört. Solche journalistische Gesichtspunkte spielen aber in der Mathematischen Visualisierung keine Rolle: Hier wollen wir Graphiken gestalten, die einem interessierten Betrachter möglichst viel relevante Information bieten. Die Wahrnehmungspsychologie lehrt uns nicht nur, daß ein Betrachter – ohne dies zu merken – gelegentlich Informationen einfach übersieht;

sie lehrt uns umgekehrt auch, daß wir seine Aufmerksamkeit (innerhalb gewisser Grenzen) auch gezielt auf einen oder mehrere Phänomene richten können. Aus unserer Alltagserfahrung wissen wir, daß uns selbst Vorgänge am Rande unseres Gesichtsfelds, auf die wir gar nicht achten wollen, gelegentlich sofort „ins Auge fallen“, z.B. ein Blinklicht. Auch beim Lesen bleibt uns zwar, wenn wir nach einem bestimmten Wort auf dieser Seite suchen, meist nichts anderes übrig, als die gesamte Seite systematisch zu überfliegen. Wenn wir aber umblättern auf eine Seite, auf der ein einzelnes Wort **rot** statt schwarz gedruckt ist oder **fett** statt normal, springt uns dieses Wort sofort ins Auge, genauso auch, wenn ein Wort *schief* steht oder nicht auf derselben Linie wie der Rest des Textes.

Für solche sogenannte *präattentive Wahrnehmungen* ist das Areal unseres Gehirns verantwortlich, zu dem die Nervenimpulse von der Netzhaut als erstes geleitet werden, der visuelle Kortex 1 oder frontale Sehlappen. Welche Eindrücke dort verarbeitet werden, erforschte

ANNE TREISMANN: Preattentive Processing in Vision, *Computer Vision, Graphics, and Image Processing* **31**, 156–177, 1985

indem sie Probanden Bilder verschiedener Komplexität zeigte, auf denen diese gewisse Dinge finden sollten. Wenn die Zeit dafür kaum von der Komplexität abhängt, spricht dies für präattentive Wahrnehmung; wenn sie mit zunehmender Komplexität deutlich ansteigt, muß der Proband das Bild wohl absuchen.

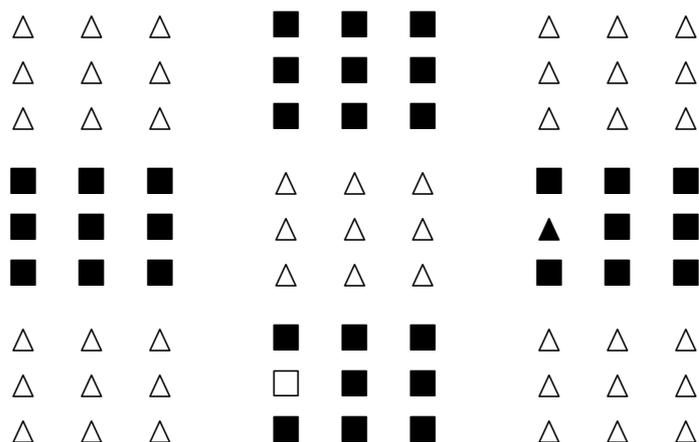
Wie ihre Experimente zeigten, gibt es eine ganze Reihe von Phänomenen, die wir präattentiv wahrnehmen können, wenn sie zumindest im lokalen Kontext isoliert auftreten. Dazu zählen beispielsweise

- Farbe
- Grauwerte
- Blinken
- Geometrische Gestalt
- Richtung von Linien
- Krümmung
- Bewegungsrichtung

- ausgefüllte oder nur als Umriß gezeichnete Figur
- zusammenhängende und getrennte Linien.

Wichtig ist ein hinreichend großer Unterschied zur Umgebung: Unter einer Vielzahl verschiedenfarbiger Punkte fällt der eine rote Punkt nicht besonders ins Auge, unter lauter blauen Punkten aber schon.

Wenn wir in einer Graphik solche Elemente verwenden, können wir das Auge des Betrachters sehr schnell auf einen bestimmten Bildteil lenken. Wir müssen diese Mittel allerdings sehr selektiv einsetzen, denn werden zu viele davon in einem einzigen Bild kombiniert, blockieren sie sich gegenseitig. Darüber hinaus können Elemente, die für sich allein sofort auffallen würden, durch ähnliche Elemente in der Nachbarschaft maskiert werden: In



sind das ausgefüllte Dreieck und das nur als Kontur gezeichnete Quadrat nicht so einfach zu finden wie in einem Diagramm, das nur die beiden relevanten 3×3 -Blöcke enthält:



Offensichtlich führen hier die Blöcke mit nur als Kontur gezeichneten Dreiecken zu einer Maskierung der beiden singulären Elemente.

Für die Visualisierung sollten wir festhalten, daß der Einsatz von Effekten, die zu präattentiver Wahrnehmung führen, das Auge des Betrachters zwar sehr schnell an eine bestimmte Stelle führen kann, daß dies aber nur dann gut funktioniert, wenn wir solche Effekte sparsam einsetzen:

Eine Graphik, in der es überall blinkt, hebt nichts hervor, sondern geht nur auf die Nerven.

§2: Farbe

Elektromagnetische Strahlung mit Wellenlängen zwischen 380 nm und 780 nm wird von uns als sichtbares Licht wahrgenommen; die kurzen Wellenlängen empfinden wir als blau, die langen als rot. Physikalisch korrekt wird Farbe daher beschrieben durch eine Funktion auf dem Intervall von etwa 380 bis etwa 780 nm beschrieben – wobei keineswegs nur stetige Funktionen in Betracht kommen.

Nun werden Farben aber nur selten mit dem Spektrometer betrachtet; für die Visualisierung interessiert nur der Eindruck auf das menschliche Auge. Dieses hat vier Arten von Photorezeptoren: Die sehr empfindlichen Stäbchenzellen, die nur bei schwachem Licht eine Rolle spielen, und die auch nur Helligkeitsinformationen liefern können, sowie drei Arten von weniger lichtempfindlichen Zäpfchenzellen, k , ℓ und m , die auf Grund von darin enthaltenen Sehfärbstoffen unterschiedlich auf Farben reagieren und somit bei gutem Licht gemeinsam ein farbiges Bild liefern können.

Die k -Zäpfchen (k wie kurzweilig) haben ihre maximale Empfindlichkeit im blauen Bereich und reagieren praktisch nur auf blaues Licht; die für mittlere Frequenzen zuständigen m -Zäpfchen überdecken, genau wie die Stäbchen, praktisch den gesamten Bereich des sichtbaren Lichts mit einer maximalen Empfindlichkeit im Grünen, und die ℓ -Zäpfchen für die langwelliges Licht haben zwar ein ausgeprägtes Maximum im roten Bereich, aber auch noch ein Nebenmaximum in blauen. Die entsprechenden Empfindlichkeitskurven findet man beispielsweise unter

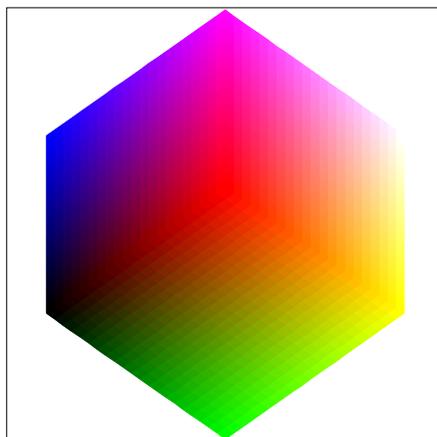
http://www.leifiphysik.de/web_ph09/grundwissen/13farbsehen/farbsehen.htm

Klassische wie auch digitale Photographie sowie Farbdarstellungen auf Fernseh- und Computermonitoren beruhen allesamt darauf, daß dem Auge etwas vorgesetzt wird, was die k , ℓ und m -Zäpfchen zur gleichen Reaktion veranlaßt wie das „echte“ Farbsignal.

Damit reicht es aus, Farben als Elemente eines dreidimensionalen Raums zu beschreiben. Da sich die Empfindlichkeitsbereiche der Zäpfchen stark überlappen, wäre es allerdings weder sinnvoll noch sonderlich praktikabel, eine Basis über die Ausgabewerte der drei Arten von Zäpfchen zu definieren. Stattdessen werden je nach Hauptanwendungsziel mehrere Farbmodelle betrachtet, die ausgehend von den unterschiedlichsten Ansätzen allesamt denselben dreidimensionalen Raum beschreiben. Wir wollen uns einige davon etwas genauer anschauen.

a) Das RGB-Modell

Dieses wohl bekannteste Farbmodell kombiniert sämtliche Farben aus den drei Grundfarben Rot, Grün und Blau. Der Farbraum wird also beschrieben als ein dreidimensionaler \mathbb{R} -Vektorraum mit Basisvektoren \vec{r} , \vec{g} und \vec{b} , die einem Rot, Grün bzw. Blau einer vorgegebenen Frequenz und Intensität entsprechen. Alle anderen Farben werden als Linearkombinationen $R\vec{r} + G\vec{g} + B\vec{b}$ dargestellt. Diese Darstellung ist insbesondere gut geeignet für die Farbdarstellung auf einem Monitor oder auch Fernsehschirm, wo Farben in genau dieser Weise erzeugt werden. Bei digitalen Bildern betrachtet man i.a. nur Koeffizienten R, G, B aus dem Intervall $[0, 1]$, so daß gewisse Grenzhelligkeiten nicht überschritten werden können. Bei der Darstellung mit einer Genauigkeit von acht Bit betrachtet man oft auch das 255-fache der entsprechenden Werte, gerundet zur nächsten ganzen Zahl. Die Abbildung zeigt den RGB-Würfel mit Blick auf seine „rote“ Ecke.



b) Grundlagen alternativer Modelle

Bei den meisten Alternativen zum RGB-Modell steht nicht die *Erzeugung* der Farben im Vordergrund, sondern deren Eindruck auf einen menschlichen Betrachter. Die Ausgabeimpulse der Zäpfchen werden noch im Sehapparat sofort weiterverarbeitet, und was wir bewußt zur Kenntnis nehmen, sind definitiv nicht die R-, G- und B-Anteile der Farben, sondern eher Helligkeiten, Farbsättigungen und ähnliche Eigenschaften.

Eine für die digitale Speicherung und Übermittlung wie auch für die Visualisierung interessante Tatsache ist, daß wir Helligkeiten viel feiner unterscheiden und viel höher auflösen können als Farbnuancen. Es liegt daher nahe, eine Basis zu wählen, in der auch die Gesamthelligkeit als Komponente auftritt, wobei diese Komponente entweder mit höherer Genauigkeit digitalisiert wird als die beiden anderen oder (häufiger) nur diese Komponente für *jedes* Pixel gespeichert wird, die beiden anderen aber nur für jedes zweite Pixel oder gar nur einmal pro Quadrat aus 2×2 Pixel. Die Wahl der Helligkeit als Basiskomponente hat auch den Vorteil, daß man dann in einfacher Weise diese Komponente für Schwarz-Weiß-Versionen der Bilder verwenden kann, was zum Beispiel beim Fernsehen eine wichtige Rolle spielt.

Wenn wir die Helligkeit als eine Basiskomponente wählen, stehen für die eigentliche chromatische Information nur noch zwei Basisvektoren zur Verfügung; diese können entweder aus zwei Farbanteilen des RGB-Systems abgeleitet werden oder (was für die Gestaltung photorealistischer Bilder gern angewandt wird) aus einer weiteren achromatischen Komponente wie etwa der Farbsättigung und nur *einer* chromatischen Komponente. Wir wollen zu jeder der beiden Alternativen ein Beispiel betrachten.

c) Das YCbCr-Modell

Bei den drei derzeit gebräuchlichen Fernsehstandards PAL, SECAM und NTSC sowie auch beim HDTV und beim JPEG-Format für digitale Bilder geht man den ersten Weg: Der erste Basisvektor \vec{y} beschreibt die Helligkeit oder *Luminanz*, die beiden weiteren Vektoren sind gewichtete

Differenzen zwischen dieser Helligkeit und dem Rot- oder Blauanteil des Farbvektors. Die Gewichte unterscheiden sich dabei in den einzelnen Fällen; da für die Hörer dieser Vorlesung das JPEG-Format interessanter sein dürfte als Fernsehstandards, betrachten wir das dort verwendete YCbCr-Modell.

Die Helligkeit ist, wie in allen diesen Systemem, als gewichtetes Mittel der drei Farbanteile definiert; die Gewichte L_R, L_G und $L_B \in (0, 1)$ bezeichnet man als *Lumared*, *Lumagreen* und *Lumablue*. Mit diesen Bezeichnungen ist die Helligkeit

$$Y = L_R R + L_G G + L_B B \quad \text{mit} \quad L_R + L_G + L_B = 1 ;$$

Standardwerte sind

$$L_R = \frac{299}{1000}, \quad L_G = \frac{587}{1000} \quad \text{und} \quad L_B = \frac{114}{1000} .$$

Die beiden Chrominanz sind festgelegt durch

$$C_b = \frac{B - Y}{2 - 2L_B} \quad \text{und} \quad C_r = \frac{R - Y}{2 - 2L_R} .$$

Damit haben wir eine neue Basis $(\vec{y}, \vec{c}_b, \vec{c}_r)$, mittels derer wir dieselben Farben beschreiben wie bezüglich der Basis $(\vec{r}, \vec{g}, \vec{b})$.

Ganz offensichtlich ist es wichtig, zwischen den beiden Basen hin- und herrechnen zu können, denn schließlich werden auch JPEG-Bilder auf RGB-Monitoren betrachtet, und CCD Chips in Digitalkameras messen zunächst einmal RGB-Komponenten, aus denen dann oft ein JPEG-Bild erzeugt wird.

Im vorliegenden Fall ist es einfach, konkrete Formeln zu finden:

$$\begin{aligned} Y &= L_R R + L_G G + L_B B , \\ C_b &= \frac{B - Y}{2 - 2L_B} = - \frac{L_R}{2 - 2L_B} \cdot R - \frac{L_G}{2 - 2L_B} \cdot G + \frac{1}{2} \cdot B , \\ C_r &= \frac{R - Y}{2 - 2L_R} = \frac{1}{2} \cdot R - \frac{L_G}{2 - 2L_R} \cdot G - \frac{L_B}{2 - 2L_R} \cdot B . \end{aligned}$$

Auch die Umkehrung lässt sich leicht ausrechnen:

$$\begin{aligned}
 R &= Y + (2 - 2L_R)C_r, \\
 B &= Y + (2 - 2L_B)C_b \quad \text{und} \\
 G &= \frac{Y - L_R R - L_B B}{L_G} \\
 &= \frac{(1 - L_R - L_B)Y - L_B(2 - 2L_B)C_b - L_R(2 - 2L_R)C_r}{L_G} \\
 &= Y - \frac{L_B(2 - 2L_B)}{L_G}C_b - \frac{L_R(2 - 2L_R)}{L_G}C_r,
 \end{aligned}$$

denn $1 - L_R - L_B = L_G$. Dies können wir auch mit Matrizen formulieren:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} L_R & L_G & L_B \\ \frac{-L_R}{2-2L_B} & \frac{-L_G}{2-2L_B} & \frac{1}{2} \\ \frac{1}{2} & \frac{-L_G}{2-2L_R} & \frac{-L_B}{2-2L_R} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

und

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 - 2L_R \\ 1 & -\frac{L_B(2-2L_B)}{L_G} & -\frac{L_R(2-2L_R)}{L_G} \\ 1 & 2 - 2L_B & 0 \end{pmatrix} \begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix}.$$

Der Übergang zwischen den beiden Basen kann also jeweils als Multiplikation mit einer Matrix interpretiert werden, und natürlich sind die beiden zugehörigen Matrizen invers zueinander.

Wenn wir die erste der beiden Matrizen mit A bezeichnen, so sind die Spalten von A die Bilder der Einheitsvektoren des RGB-Systems, umgerechnet ins YCbCr-System; wir erhalten also die Beziehungen

$$\begin{aligned}
 \vec{r} &= L_R \vec{y} - \frac{L_R}{2 - 2L_B} \vec{c}_b + \frac{1}{2} \vec{c}_r \\
 \vec{g} &= L_G \vec{y} - \frac{L_G}{2 - 2L_B} \vec{c}_b - \frac{L_G}{2 - 2L_R} \vec{c}_r \quad \text{und} \\
 \vec{b} &= L_B \vec{y} + \frac{1}{2} \vec{c}_b - \frac{L_B}{2 - 2L_R} \vec{c}_r.
 \end{aligned}$$

Entsprechend sind die Spalten von A^{-1} die Bilder der Einheitsvektoren des YCbCr-Systems, umgerechnet ins RGB-System, d.h.

$$\begin{aligned}\vec{y} &= \vec{r} + \vec{g} + \vec{b}, \\ \vec{c}_b &= -\frac{2-2L_B}{L_G}\vec{g} + (2-2L_B)\vec{b} \quad \text{und} \\ \vec{c}_r &= (2-2L_R)\vec{r} - \frac{2-2L_B}{L_G}\vec{g}.\end{aligned}$$

Zur Verdeutlichung seien die Gleichungen nochmals angegeben mit numerischen Koeffizienten, näherungsweise berechnet für die Standardwerte von L_R , L_G und L_B : Damit ist

$$A = \begin{pmatrix} 0,2990 & 0,5870 & 0,1140 \\ -0,1687 & -0,3313 & 0,5000 \\ 0,5000 & -0,4187 & -0,0813 \end{pmatrix}$$

und

$$A^{-1} = \begin{pmatrix} 1 & 0 & 1,4020 \\ 1 & -0,3441 & -0,7141 \\ 1 & 1,7720 & 0 \end{pmatrix},$$

also erhalten wir die Gleichungen

$$\begin{aligned}Y &= 0,2990 R + 0,5870 G + 0,1140 B \\ Cb &= -0,1687 R - 0,3313 G + 0,5000 B \\ Cr &= 0,5000 R - 0,4187 G - 0,0813 B\end{aligned}$$

und

$$\begin{aligned}\vec{r} &= 0,2990 \vec{y} - 0,1687 \vec{c}_b + 0,5000 \vec{c}_r \\ \vec{g} &= 0,5870 \vec{y} - 0,3313 \vec{c}_b - 0,4187 \vec{c}_r \\ \vec{b} &= 0,1140 \vec{y} + 0,5000 \vec{c}_b - 0,0813 \vec{c}_r.\end{aligned}$$

Entsprechend liefern Zeilen und Spalten von A^{-1} die Beziehungen

$$\begin{aligned}R &= Y + 1,402 Cr \\ G &= Y - 0,3441 Cb - 0,7141 Cr \\ B &= Y + 1,772 Cb\end{aligned}$$

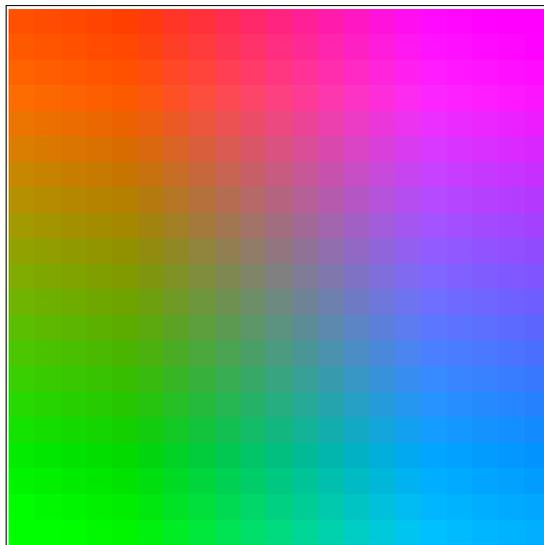
und

$$\vec{y} = \vec{r} + \vec{g} + \vec{b}$$

$$\vec{c}_b = -0,3441 \vec{g} + 10,772 \vec{b}$$

$$\vec{c}_r = 1,402 \vec{r} - 0,7141 \vec{g}.$$

Das Bild unten zeigt das CbCr-Rechteck für $Y = 0,5$. (Bei Schwarz-Weiß-Druck ist hier ein einfarbiges Rechteck mit Grauwert 0,5 zu sehen.)



d) Das HSV-Modell

Auch hier gehen wir wieder aus von den RGB-Werten einer Farbe, definieren nun aber ihre Helligkeit (den Value in der Terminologie des Modells) als den Maximalwert der drei Farbkomponenten:

$$V = \max(R, G, B).$$

Wenn dieses Maximum gleichzeitig auch Minimum ist, müssen alle drei Farbwerte gleich sein; in diesem Fall haben wir einen von V abhängigen Grauwert mit den Extremfällen weiß für $V = 0$ und schwarz für $V = 1$. Allgemeiner können wir die Differenz zwischen Maximum und Minimum als Maß für die Farbsättigung ansehen: Je größer sie ist, umso „bunter“ ist die Farbe.

Um einen Wert zwischen null und eins zu bekommen, dividieren wir diese Differenz noch durch das Maximum, was natürlich nur möglich ist,

wenn letzteres nicht verschwindet. Es verschwindet genau dann, wenn $R = G = B = 0$ ist, d.h. für schwarz; dessen „Farb“ättigung definieren wir formal als null. Wir definieren somit die Farbsättigung (**Saturation**) als

$$S = \begin{cases} 0 & \text{falls } R = G = B = 0 \\ \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} & \text{sonst} \end{cases} .$$

Bleibt noch eine Koordinate, der Farbton (**Hue**). Er muß offensichtlich davon abhängen, welcher der drei RGB-Werte maximal ist; andernfalls haben wir keine Chance auf eine Umkehrabbildung. Ansonsten soll die neue Koordinate die relative Position des mittleren RGB-Wert im Intervall zwischen Minimum und Maximum beschreiben, was natürlich nur dann sinnvoll ist, wenn sich diese beiden Werte unterscheiden; andernfalls setzen wir formal $H = 0$. Allgemein setzen wir

$$\Delta = \max(R, G, B) - \min(R, G, B)$$

und definieren zunächst eine Hilfsgröße h wie folgt:

$$h = \begin{cases} 0 & \text{falls } R = G = B \\ 0 + \frac{G-B}{\Delta} & \text{falls } \max(R, G, B) = R \\ 2 + \frac{B-R}{\Delta} & \text{falls } \max(R, G, B) = G \\ 4 + \frac{R-G}{\Delta} & \text{falls } \max(R, G, B) = B \end{cases} .$$

Die Brüche mit Nenner Δ können jeweils Werte zwischen -1 und 1 annehmen; wenn R das Maximum ist etwa entspricht der Wert -1 einer Farbe, bei der $R = G$ und $B = \min(R, G, B)$ ist; +1 erhalten wir für $R = B$ und $G = \min(R, G, B)$. Der Wertebereich von h geht also von -1 bis 5. Um Werte im Intervall $[0, 1]$ zu bekommen, setzen wir

$$H = \begin{cases} \frac{h}{6} & \text{falls } h > 0 \\ \frac{h}{6} + 1 & \text{falls } h < 0 \end{cases} .$$

Wenn sowohl R als auch G gleich dem Maximum sind, können wir h entweder nach der zweiten oder nach der dritten Zeile berechnen, erhalten aber in beiden Fällen die Werte $h = 1$ und $H = \frac{1}{6}$. Entsprechend liefern im Fall $G = B = \max(R, G, B)$ sowohl die dritte als auch die vierte Zeile die Werte $h = 3$ und $H = \frac{1}{2}$. Ist allerdings $R = B = \max(R, G, B)$, so liefert die Formel aus der zweiten Zeile den Wert

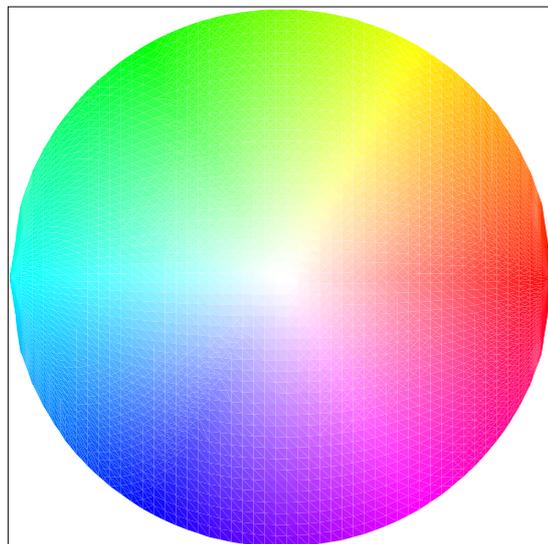
$h = -1$, die in der vierten dagegen $h = 5$. Der Wert von H ist jedoch in beiden Fällen $\frac{5}{6}$.

Die Primärfarben Rot, Grün und Blau haben $H = 0$, $\frac{1}{3}$ und $H = \frac{2}{3}$; zwischen 0 und $\frac{1}{3}$ folgt H im wesentlichen dem Verlauf des Spektrums von Rot bis Grün, dann geht es weiter zum blauen Ende des sichtbaren Spektrums bei $H = \frac{2}{3}$. Jenseits von $H = \frac{2}{3}$ sind zwar noch alle Werte $H < 1$ möglich, sie entsprechen aber keinen Spektralfarben, sondern sind nur als zusammengesetzte Farben realisierbar. Die entsprechenden Farben sind Mischungen aus Blau und Rot; man redet hier von der sogenannten *Purpurlinie*. Geht H gegen eins, so geht h gegen null; wir erreichen also wieder das Rot mit $H = 0$.

Dies legt es nahe, den Farbton nicht durch eine Zahl, sondern durch einen Winkel zu charakterisieren. Führen wir den *Farbwinkel*

$$\varphi = H \cdot 360^\circ$$

ein, so können wir S und φ als Polarkoordinaten in der Ebene betrachten und erhalten für konstantes $V = 1$ den folgenden HSV-Farbkreis:



§4: Abbildungen des Raums in die Ebene

In der Visualisierung haben wir oftmals Datensätze, die mathematisch in natürlicher Weise als Punkte eines Raums mit deutlich mehr als drei

Dimensionen aufgefaßt werden können. Unsere Vorstellungswelt ist nur dreidimensional, und unsere Computerbildschirme sind natürlich, genau wie das Papier auf das wir drucken, zweidimensional. Wie wir aus Erfahrung wissen, kann ein zweidimensionales Bild allerdings zumindest innerhalb gewisser Grenzen durchaus einen räumlichen Eindruck erwecken, und dieses Phänomen sollten wir in der Visualisierung auch soweit wie möglich ausnutzen. In diesem Paragraphen werden wir uns daher mit der Frage beschäftigen, wie wir eine dreidimensionale Szene auf einem zweidimensionalen Bildschirm so realistisch wie möglich darstellen können. Die meisten rechnerischen Details können wir dabei einem Computergraphikprogramm überlassen; die grundsätzlichen Strategien müssen aber natürlich wir festlegen und damit auch verstehen.

Die Zeichenfläche, auf die wir abbilden, ist beschränkt, typischerweise ein Rechteck; die Szene, die wir abbilden wollen, muß nicht unbedingt beschränkt sein; gelegentlich wollen wir auch Phänomene visualisieren, die sich auf dem gesamten \mathbb{R}^3 abspielen. Wie aus der Analysis bekannt, gibt es Abbildungen von \mathbb{R}^3 auf ein Rechteck, beispielsweise die Abbildung

$$\varphi: \begin{cases} \mathbb{R}^3 \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ (x, y, z) \mapsto \left(\arctan\left(-\frac{\sqrt{3}}{2}x + \frac{\sqrt{3}}{2}y\right), \arctan\left(z - \frac{x+y}{2}\right) \right) \end{cases}$$

Solche Abbildungen sind aber nicht Gegenstand der Computergraphik; bewußte Verzerrungen sind ein Darstellungsmittel der Visualisierung im engeren Sinne. Die Computergraphik betrachtet Abbildungen des gesamten \mathbb{R}^3 (oder zumindest einer möglichst großen Teilmenge davon) nach \mathbb{R}^2 , und läßt einfach alle Punkte außer Acht, deren Bild nicht im Zielrechteck liegt. Für die Konstruktion solcher Abbildungen von \mathbb{R}^3 nach \mathbb{R}^2 sind vor allem zwei Strategien gebräuchlich:

a) Axonometrien

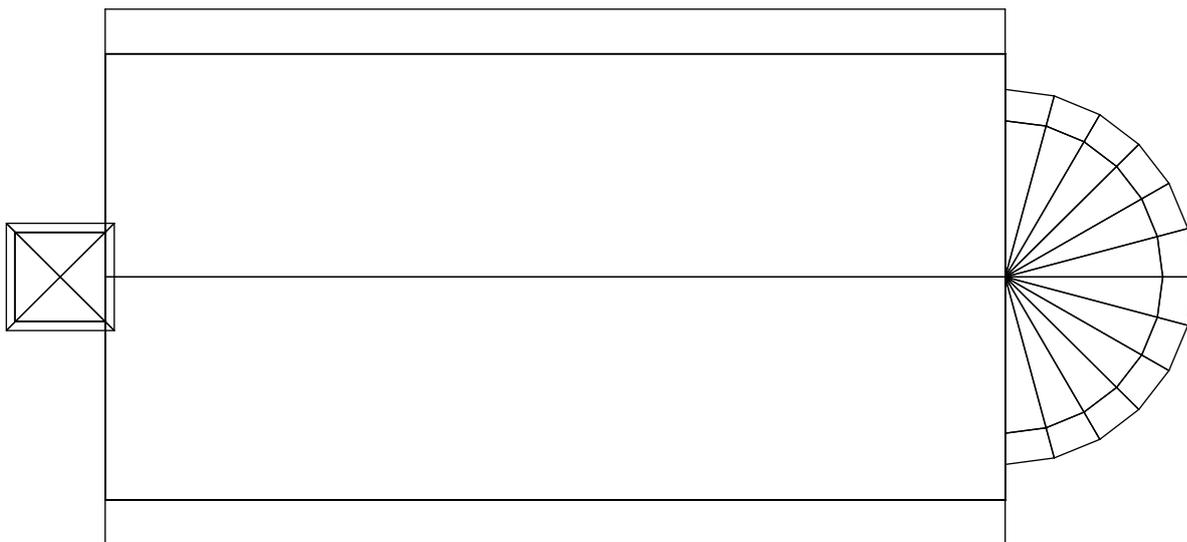
Hier geht es um lineare Abbildungen von \mathbb{R}^3 nach \mathbb{R}^2 . Eine solche Abbildung ist gegeben durch die Bilder $\vec{a}_1, \vec{a}_2, \vec{a}_3$ der drei Koordinatenein-

heitsvektoren des \mathbb{R}^3 ; sie hat die Form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x\vec{a}_1 + y\vec{a}_2 + z\vec{a}_3.$$

Diese Art von Abbildung wird vor allem für Konstruktionszeichnungen benutzt; die bekanntesten Beispiele sind Grundriß, Aufriß und Kreuzriß, bei denen jeweils \vec{a}_3, \vec{a}_1 bzw. \vec{a}_2 gleich dem Nullvektor sind und die beiden übrigen \vec{a}_i die Einheitsvektoren des \mathbb{R}^2 ; konkret handelt es sich also um die Abbildungen

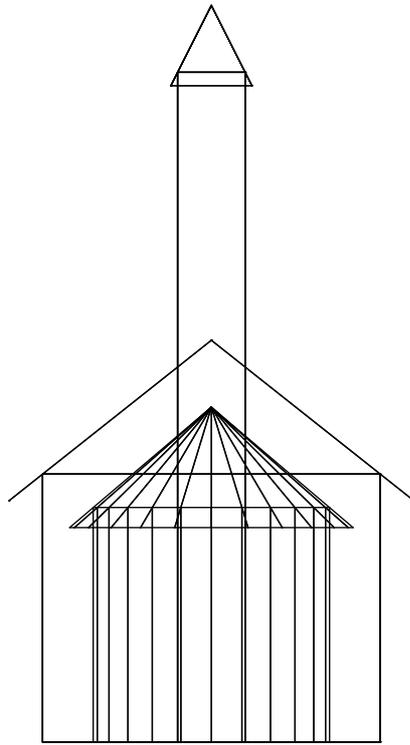
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ z \end{pmatrix}.$$



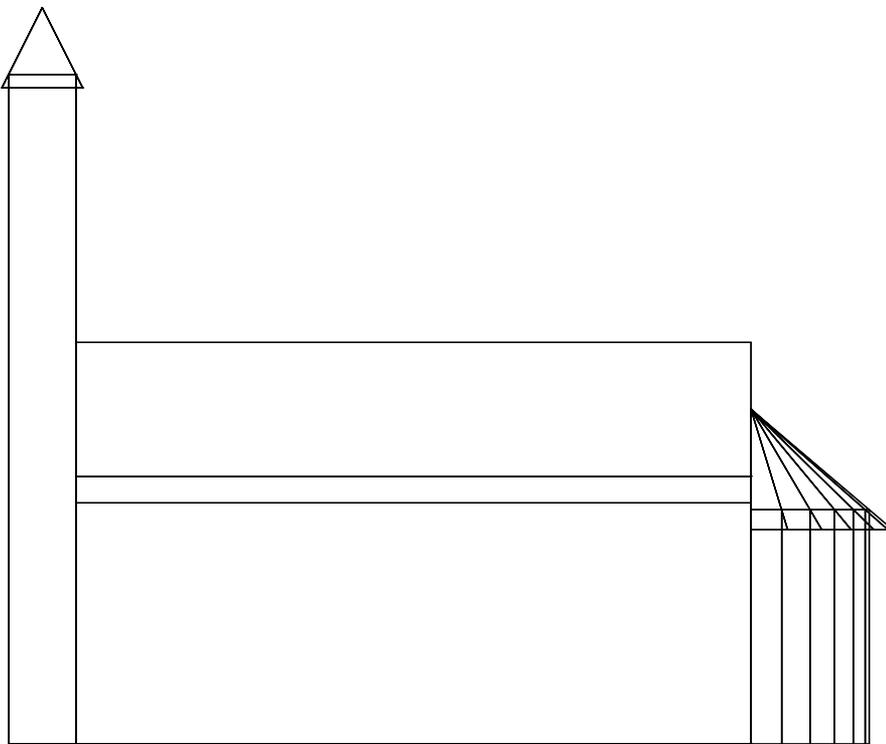
Grundriß

Einen „dreidimensionaleren“ Eindruck vermittelt die Isometrie nach DIN 5 (inzwischen aufgegangen in ISO 5456-3), bei der alle drei Vektoren Länge eins haben und zwei aufeinanderfolgende jeweils einen Winkel von 120° einschließen; konkret ist

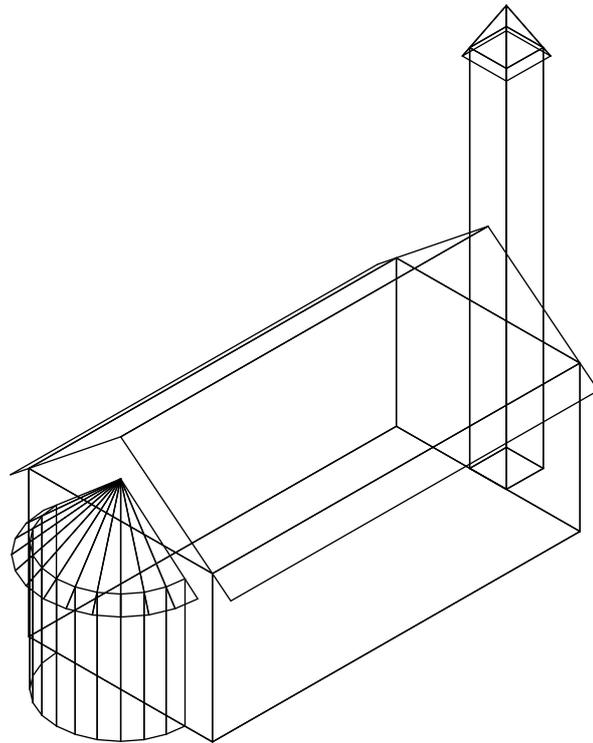
$$\vec{a}_1 = \begin{pmatrix} -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2} \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} \frac{1}{2}\sqrt{3} \\ -\frac{1}{2} \end{pmatrix} \quad \text{und} \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$



Aufriß



Kreuzriß



Isometrische Normalprojektion

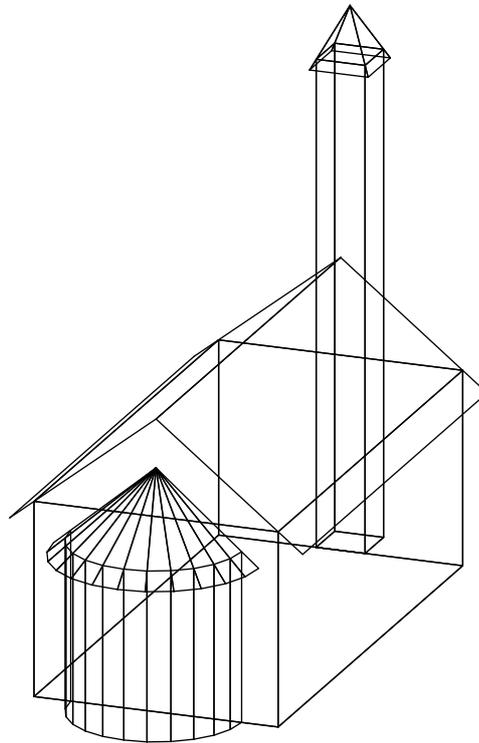
Eine Variante davon ist die *isometrische Normalprojektion*, bei der diese Vektoren jeweils mit $\sqrt{2/3}$ multipliziert werden.

Ein Kompromiss zwischen Isometrien und Aufriß ermöglichen die *Dimetrien*, bei denen zwei Vektoren die gleiche Länge haben, der dritte aber nur halb so groß ist. Auf diese Weise wird die Vorderansicht des Objekts betont, ohne daß der räumliche Eindruck verloren geht. Gebräuchlich sind die *Ingenieursaxonometrie* mit

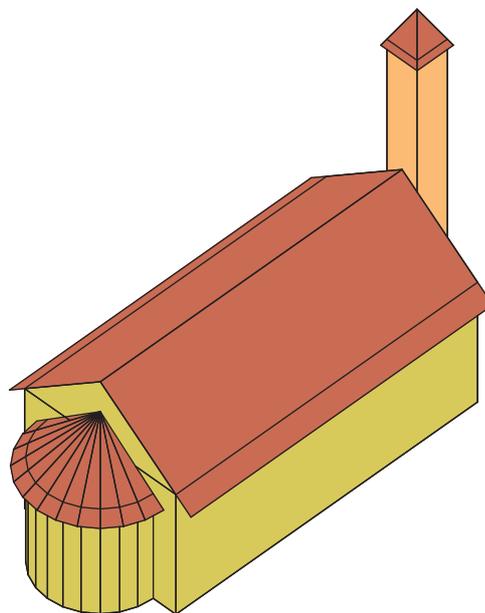
$$\vec{a}_1 = \begin{pmatrix} -\frac{3}{4} \\ -\frac{1}{4}\sqrt{7} \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} \frac{3}{4}\sqrt{7} \\ -\frac{1}{4} \end{pmatrix} \quad \text{und} \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

sowie die Dimetrie nach DIN 5, bei der die drei Vektoren genau halb so groß sind.

Anschaulicher, wenn auch in mancher Hinsicht weniger informativ, ist ein Bild, das an Stelle von Kanten Flächen projiziert und dabei auch auf verdeckte Linien achtet. Genau deshalb sieht ein solches Bild natürlicher aus, sagt dafür aber etwas weniger aus über die Geometrie. Ein Vorteil der Flächendarstellung ist natürlich die Möglichkeit Farben zu benutzen;



Ingenieursaxonometrie

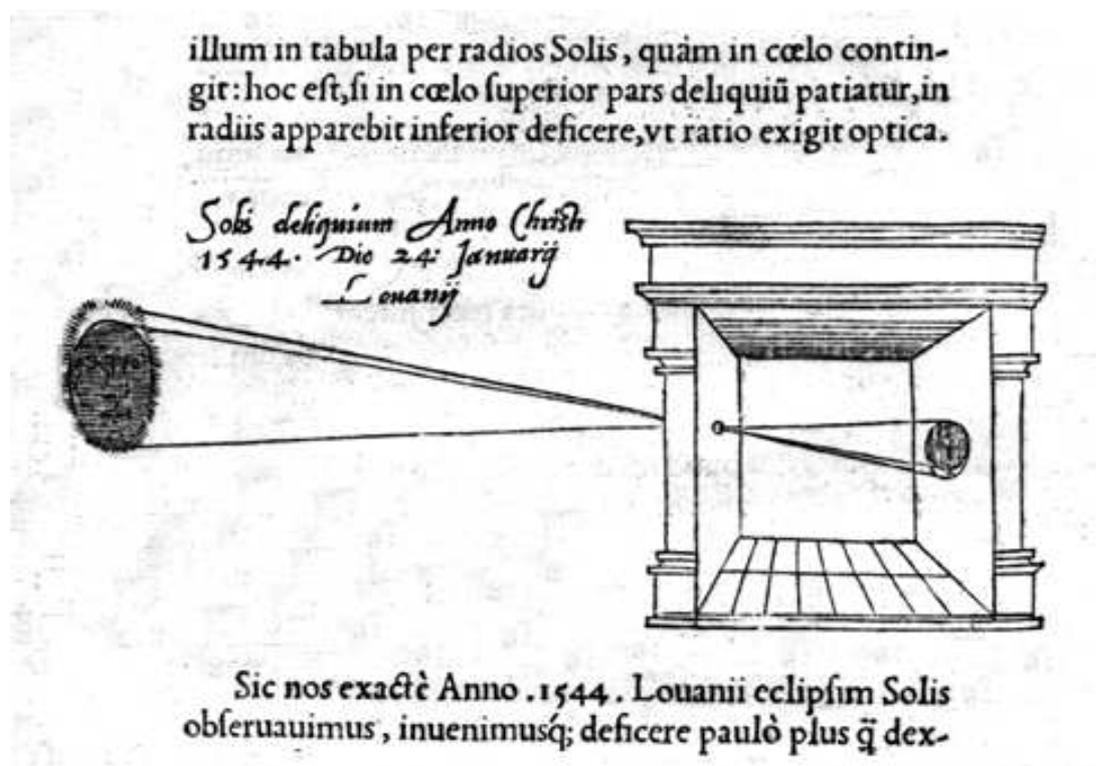


Isometrische Projektion mit Flächen

seinen vollen Nutzen entfaltet ein solches Bild dann, wenn man es am Bildschirm von allen Seiten betrachten kann.

b) Perspektivische Projektionen und Kameramodell

Die zweite Strategie, die eher für anschauliche bis photorealistische Darstellungen benutzt wird, orientiert sich an einem Kameramodell: Wir stellen eine Kamera in den \mathbb{R}^3 und identifizieren den Bildschirm (oder das betrachtete Fenster) mit dem Stück Film, das diese belichtet. Als Kamera nehmen wir ein altbewährtes Modell, das bereits im fünften vorchristlichen Jahrhundert von dem chinesischen Philosophen MO-TI oder MO TZU (470–391) beschrieben wurde und mit dem im darauffolgenden Jahrhundert auch ARISTOTELES (384–322) recht gute Erfahrungen machte: die *Camera obscura*.



Ihr großer Vorteil gegenüber moderneren Kameras ist, daß sie anstelle eines Objektivs nur eine Lochblende hat, so daß es keinerlei Probleme mit der Tiefenschärfe gibt (für die wir rechnerisch aufwendige Faltungsin-
tegrale bilden müßten). Natürlich kommt es immer wieder einmal vor, daß wir einen unscharfen Hintergrund brauchen, aber da können wir genauso gut mit rechnerisch erheblich billigerem Nebel arbeiten.

Die *Camera obscura* definiert allerdings keine Abbildung von \mathbb{R}^3 nach \mathbb{R}^2 : Zunächst einmal bildet sie nur Objekte ab, die sich *vor* der Kamera befinden. Diese Unzulänglichkeit der Realität ist mathematisch natürlich überhaupt kein Problem: Wir *definieren* einfach, daß auch bei Punkten hinter der Kamera der Bildpunkt gleich dem Schnittpunkt der Gerade zwischen Punkt und Lochblende mit der Filmebene sein soll. Ob diese Verbesserung der Natur wirklich wünschenswert ist, hängt von der Anwendung ab; statt darüber zu spekulieren, wollen wir lieber sehen, wie wir die Abbildung mathematisch beschreiben können. Sie ist offensichtlich keine lineare Abbildung von \mathbb{R}^3 nach \mathbb{R}^2 , denn jede solche Abbildung wäre in jedem Punkt des \mathbb{R}^3 definiert, während die *Camera obscura* zumindest die Punkte aus der zur Filmebene parallelen Ebene durch die Lochblende, also der Frontebene der Kamera, definitiv nicht in die Filmebene abbilden kann.

Erstaunlicherweise können wir die *Camera obscura* trotzdem durch eine lineare Abbildung beschreiben, indem wir letzteres Problem lösen und auch noch diesen Punkten (mit Ausnahme der Lochblende selbst) Bildpunkte zuordnen.

Diese Bildpunkte können, vernünftig interpretiert, nicht im \mathbb{R}^2 liegen: Betrachten wir einen von der Lochblende verschiedenen Punkt P auf der Frontebene der Kamera. Schieben wir diesen Punkt geringfügig nach vorne, also weiter weg von der Filmebene, so kann er von der Kamera abgebildet werden; sein Bildpunkt ist umso weiter vom „Mittelpunkt“ der Filmebene entfernt, je näher der Punkt an der Frontebene liegt. Die Punkte *auf* der Frontebene sollten daher abgebildet werden auch „unendlich ferne Punkte“ in Richtung des Verbindungsstrahls von der Lochblende zum abzubildenden Punkt.

c) Projektive Räume und Abbildungen

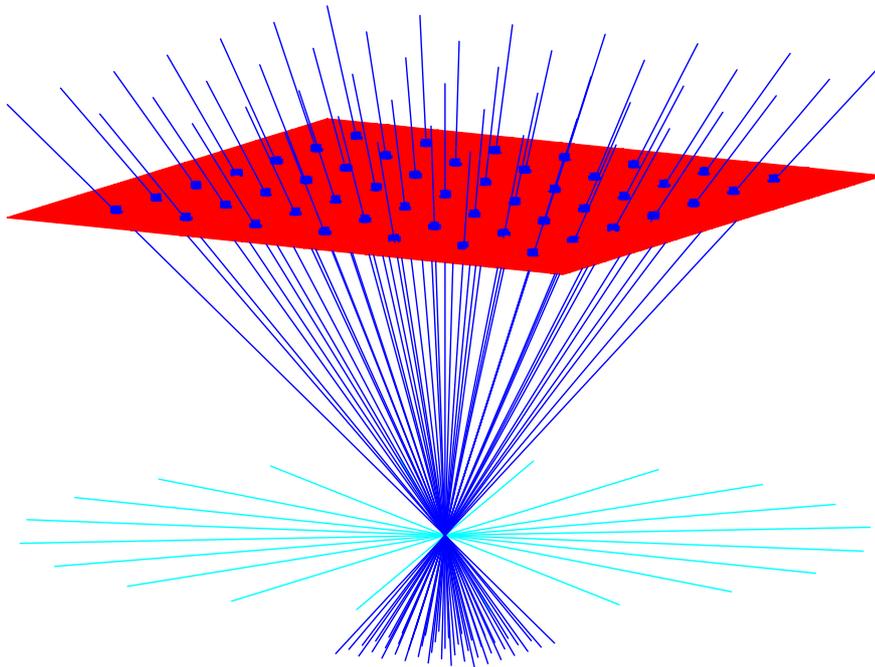
Dieses Konzept eines „unendlich fernen“ Punkts können wir mathematisch präzise definieren im Rahmen der projektiven Geometrie. Da es für die grundlegenden Definitionen gleichgültig ist, ob wir über den reellen Zahlen oder einem anderen Körper arbeiten, definieren wir die Begriffe zunächst allgemein:

Definition: $V \neq \{\vec{0}\}$ sei ein endlichdimensionaler Vektorraum über dem Körper k . Der *projektive Raum* $\mathbb{P}(V)$ ist die Menge aller eindimensionalen Untervektorräume von V ; diese Untervektorräume bezeichnen wir als die *Punkte* von $\mathbb{P}(V)$. Die Dimension von $\mathbb{P}(V)$ definieren wir als $\dim V - 1$; ein r -dimensionaler *projektiver Unterraum* von $\mathbb{P}(V)$ ist ein projektiver Raum der Form $\mathbb{P}(U)$ zu einem $(r + 1)$ -dimensionalen Untervektorraum $U \leq V$. Im Falle $r = 1$ sprechen wir von einer projektiven Geraden, bei $r = 2$ von einer projektiven Ebenen.

Um zu verstehen, daß dies unser Problem löst, betrachten wir zunächst die projektive Gerade $\mathbb{P}^1 = \mathbb{P}(\mathbb{R}^2)$. Ihre Punkte sind nach Definition die eindimensionalen Untervektorräume von \mathbb{R}^2 , d.h. Unterräume der Form $U = \mathbb{R}\begin{pmatrix} a \\ b \end{pmatrix}$ mit $\begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Nun betrachten wir in \mathbb{R}^2 die (klassische) Gerade $y = 1$. Sie enthält genau dann einen Vektor aus dem eindimensionalen Untervektorraum U , wenn $b \neq 0$ ist; in diesem Fall ist das der Punkt $(a/b, 1) \in g \subset \mathbb{R}^2$. Im Falle $b = 0$ ist $U = \mathbb{R}\begin{pmatrix} a \\ 0 \end{pmatrix} = \mathbb{R}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ die x -Achse, ist also parallel zur Geraden g und schneidet diese nicht.

Wir können daher die Gerade g als Teilmenge der projektiven Geraden $\mathbb{P}(\mathbb{R}^2)$ auffassen, indem wir den Punkt $(x, 1)$ mit dem Untervektorraum $\mathbb{R}\begin{pmatrix} x \\ 1 \end{pmatrix}$ identifizieren und erhalten so alle Punkte von $\mathbb{P}(\mathbb{R}^2)$ mit der einen Ausnahme $\mathbb{R}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Diesen Punkt bezeichnen wir als den „unendlich fernen Punkt“ von $\mathbb{P}(\mathbb{R}^2)$. Man beachte, daß dieser Punkt als Element von $\mathbb{P}(\mathbb{R}^2)$ keinerlei Sonderstellung einnimmt: Wenn wir an Stelle von g eine andere Gerade durch \mathbb{R}^2 betrachtet hätten, wäre bezüglich dieser Geraden ein anderer Punkt „unendlich fern“: Für die Gerade $bx - ay = 1$ wäre es der Punkt $\mathbb{R}\begin{pmatrix} a \\ b \end{pmatrix}$, d.h. wir können je nach Betrachtungsweise jeden Punkt von \mathbb{P}^1 als unendlich fern interpretieren.

Etwas komplizierter wird es für die projektive Ebene $\mathbb{P}^2 = \mathbb{P}(\mathbb{R}^3)$. Sie ist definiert als die Menge alle eindimensionalen Untervektorräume von \mathbb{R}^3 . Um sie mit einer klassischen Ebenen in Verbindung zu bringen, betrachten wir die Ebene $z = 1$ in \mathbb{R}^3 ; sie schneidet einen eindimensionalen Untervektorraum offensichtlich genau dann, wenn die dritte Koordinate ihres Basisvektors nicht verschwindet, denn dann können wir alle Komponenten durch die dritte teilen, und erhalten ein Element von E .



Falls die dritte Komponente verschwindet, können wir immer noch die erste und die zweite Komponente frei wählen; sie dürfen nur nicht beide gleichzeitig verschwinden. Die Punkte mit dieser Eigenschaft sind daher genau die Punkte der projektiven Geraden zum durch $z = 0$ definierten Untervektorraum des \mathbb{R}^3 .

Allgemein können wir im n -dimensionalen projektiven Raum $\mathbb{P}(\mathbb{R}^{n+1})$ stets eine „unendlich ferne“ Hyperebene auszeichnen und deren Komplement mit \mathbb{R}^n identifizieren.

Die Punkte eines projektiven Raums sind nach Definition die eindimensionalen Untervektorräume eines vorgegebenen Vektorraums V . Ein eindimensionaler Untervektorraum U ist eindeutig bestimmt durch jeden vom Nullvektor verschiedenen Vektor $\vec{u} \in U$; je zwei solche Vektoren unterscheiden sich durch eine von Null verschiedene reelle Konstante. Wir können einen Punkt $P \in \mathbb{P}(V)$ daher bestimmen durch die $n + 1$ Komponenten eines beliebigen Vektors $\vec{u} \neq \vec{0}$ des zugehörigen Untervektorraums U bezüglich einer gegebenen Basis von V . Diese Darstellung ist allerdings nicht eindeutig, da jedes vom Nulltupel verschiedene Vielfache denselben Punkt beschreibt. Wir sprechen daher von *homogenen Koordinaten* und schreiben die Tupel in der Form $(x_0 : x_1 : \dots : x_n)$. Wir werden im folgenden meist den Fall $V = \mathbb{R}^{n+1}$

betrachten; hier arbeiten wir natürlich mit der Standardbasis bestehend aus den $n + 1$ Koordinateneinheitsvektoren.

Da homogene Koordinaten nur bis auf eine multiplikative Konstante bestimmt sind, ist es nicht sinnvoll, die Koordinaten zweier Punkte zu addieren. Wir können sie aber mit einer Matrix multiplizieren und so einen neuen Punkt desselben oder eines anderen projektiven Raums definieren:

Lemma: $\varphi: V \rightarrow W$ sei eine von der Nullabbildung verschiedene lineare Abbildung zwischen zwei endlichdimensionalen Vektorräumen.

a) Ist φ injektiv, so definiert diese eine Abbildung $\mathbb{P}(V) \rightarrow \mathbb{P}(W)$, indem wir jedem eindimensionalen Teilraum $U \leq V$ denn dann ebenfalls eindimensionalen Untervektorraum $\varphi(U) \leq W$ zuordnen.

b) Falls $Z = \text{Kern } \varphi$ nicht der Nullraum ist, definiert diese Zuordnung eine Abbildung $\mathbb{P}(V) \setminus \mathbb{P}(Z) \rightarrow \mathbb{P}(W)$.

Der *Beweis* ist klar, denn nach dem Homomorphiesatz ist $\varphi(U)$ isomorph zu $U/(U \cap Z)$, ist also eindimensional, falls $U \cap Z = \{\vec{0}\}$ ist, und der Nullraum, falls $U \subseteq Z$. Im letzteren Fall erhalten wir keinen Punkt von $\mathbb{P}(W)$. ■

Wenn wir mit homogenen Koordinaten arbeiten, können wir φ durch seine Abbildungsmatrix beschreiben; haben V und W die Dimensionen $n + 1$ und $m + 1$, so ist das eine $(n + 1) \times (m + 1)$ -Matrix. Multiplizieren wir ein Tupel homogener Koordinaten mit dieser Matrix, so erhalten wir ein Koordinatentupel ein $(m + 1)$ -Tupel, das genau dann das Nulltupel ist, wenn das Ausgangstupel einen Punkt aus $\mathbb{P}(Z)$ beschreibt. Ersetzen wir das Ausgangstupel durch ein skalares Vielfaches, so wird das Bildtupel mit derselben Konstante multipliziert; die Abbildung ist also wohldefiniert auf dem Niveau der homogenen Koordinaten.

Definition: $\varphi: V \rightarrow W$ sei eine surjektive lineare Abbildung zwischen endlichdimensionalen Vektorräumen mit Kern Z . Falls $Z = \{\vec{0}\}$, bezeichnen wir die zugehörige Abbildung $\mathbb{P}(V) \rightarrow \mathbb{P}(W)$ als eine Projektivität; andernfalls bezeichnen wir die Abbildung $\mathbb{P}(V) \setminus \mathbb{P}(Z) \rightarrow \mathbb{P}(W)$ als eine Projektion von $\mathbb{P}(V)$ auf $\mathbb{P}(W)$ mit Zentrum $\mathbb{P}(Z)$.

d) Homogene Koordinaten

Die Definition von Punkten eines projektiven Raums als eindimensionale Untervektorräume eines Vektorraums ist zwar, wie wir gerade gesehen haben, theoretisch recht einsichtig; zum praktischen Rechnen können wir Punkte aber nicht durch unendliche Mengen darstellen. Ein eindimensionaler Vektorraum ist vollständig festgelegt durch seine Basis, also durch irgendeinen seiner vom Nullvektor verschiedenen Vektoren.

Falls wir eine Basis $\vec{b}_0, \dots, \vec{b}_n$ des $(n+1)$ -dimensionalen reellen Vektorraums V festlegen, können wir einen solchen Vektor \vec{v} bezüglich dieser Basis in der Form

$$\vec{v} = x_0 \vec{b}_0 + \dots + x_n \vec{b}_n$$

darstellen, wobei nicht alle x_i verschwinden dürfen, und das $(n+1)$ -tupel aus den Zahlen x_0, \dots, x_n legt den Punkt $\mathbb{R}\vec{v} \in \mathbb{P}(V)$ eindeutig fest.

Umgekehrt legt allerdings der Punkt das Tupel *nicht* eindeutig fest, denn an Stelle des Vektors \vec{v} könnten wir ja auch irgendeinen Vektor $\lambda\vec{v}$ mit $\lambda \neq 0$ als Basis von $\mathbb{R}\vec{v}$ wählen, wodurch auch alle x_i mit λ multipliziert würden. Wir bezeichnen die x_i daher als *homogene* Koordinaten des Punktes $\mathbb{R}\vec{v}$ und schreiben sie in der Form $(x_0 : x_1 : \dots : x_n)$ mit der Interpretation, daß für jedes $\lambda \neq 0$ gelten soll

$$(x_0 : x_1 : \dots : x_n) = (\lambda x_0 : \lambda x_1 : \dots : \lambda x_n).$$

Damit haben wir also für einen Punkt eines n -dimensionalen projektiven Raums $n+1$ Koordinaten, bei denen es allerdings nicht auf die absolute Größe, sondern nur auf die Verhältnisse ankommt. Wir werden gleich sehen, daß dies durchaus auch einige Vorteile mit sich bringt.

Betrachten wir der Einfachheit halber den dreidimensionalen Raum \mathbb{R}^3 und betten ihn über die Abbildung

$$\left\{ \begin{array}{l} \mathbb{R}^3 \rightarrow \mathbb{P}^3 \\ (x, y, z) \mapsto (x : y : z : u) \end{array} \right.$$

ein in den dreidimensionalen projektiven Raum. Im Bild liegen offenbar alle Punkte $(x : y : z : u) \in \mathbb{P}^3$ mit $u \neq 0$; Urbild dieses Punkts ist $(\frac{x}{u}, \frac{y}{u}, \frac{z}{u}) \in \mathbb{R}^3$.

Bewegungen im \mathbb{R}^3 werden bekanntlich beschrieben durch Abbildungen der Form

$$\vec{v} \mapsto A\vec{v} + \vec{b},$$

wobei A eine (orthogonale) 3×3 -Matrix ist und \vec{b} ein Vektor, der Verschiebungsvektor.

In homogenen Koordinaten können wir diese Abbildung interpretieren als Multiplikation von $\vec{v} \in \mathbb{R}^4$ mit einer Matrix A^* , bestehend aus der Matrix A , der zunächst der Vektor $\vec{v} \in \mathbb{R}^3$ als vierte Spalte angehängt wird, danach noch die vierte Zeile $(0 \ 0 \ 0 \ 1)$:

$$A^* = \begin{pmatrix} A & \vec{b} \\ \mathbf{0} & 1 \end{pmatrix},$$

wobei die fette Null für drei gewöhnliche Nullen steht. Mit dieser Bezeichnung ist dann, wie man leicht nachrechnet,

$$A^* \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \\ 1 \end{pmatrix} \quad \text{mit} \quad \begin{pmatrix} u \\ v \\ w \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \vec{b}.$$

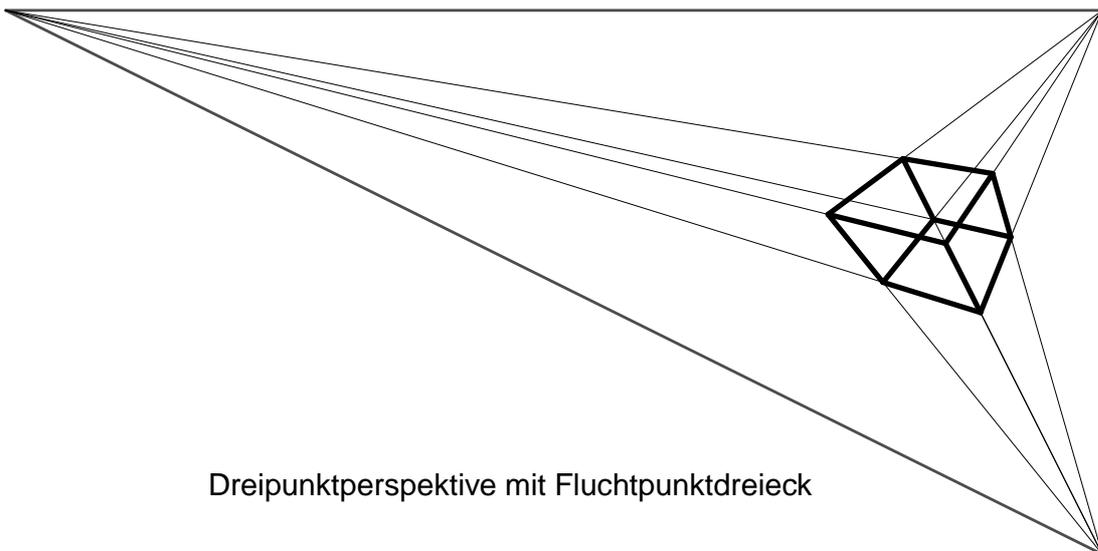
Die Darstellung links ist zwar rechnerisch etwas aufwendiger als die rechte, dafür aber etwas kompakter und hat vor allem den Vorteil, daß sich die Hintereinanderausführung zweier Abbildungen einfach durch ein Matrixprodukt beschreiben läßt.

Wirklich interessant wird es, wenn $u = 0$ wird: Der Punkt $(1 : 0 : 0 : u)$ etwa hat für $u \neq 0$ die kartesische Darstellung $(1/u, 0, 0)$, ist also der Punkt auf der x -Achse mit Koordinate $1/u$. Geht u gegen null, wandert der Punkt also (je nach Vorzeichen links oder rechts) ins Unendliche; $(1 : 0 : 0 : 0)$ kann somit als *unendlich ferner Punkt* der x -Achse aufgefaßt werden, und genauso ist allgemein $(x : y : z : 0)$ der unendlich ferne Punkt auf der vom Vektor (x, y, z) aufgespannten Geraden durch den Nullpunkt.

Solche unendlich ferne Punkte sind beispielsweise interessant für die Positionierung von Lichtquellen: Wenn wir etwa eine Szene mit Sonnenlicht beleuchten wollen, wäre es unsinnig, die genaue Position der Sonne in kartesischen Koordinaten anzugeben; es reicht völlig, wenn wir die Richtung angeben. Damit teilen wir dem System auch mit, daß nun für die Lichtintensität kein $1/r^2$ -Gesetz mehr gilt, sondern daß die Intensität unabhängig von der Entfernung ist.

Im Tutorium lightposition von NATE ROBINS kann man die Unterschiede zwischen den beiden Fällen demonstrieren.

Gelegentlich können wir unendlich ferne Punkte auch auf dem Bildschirm sehen: Falls wir an einem Würfel die Geraden durch zwei parallele Kanten betrachten, schneiden diese sich natürlich nirgends. Ihre Bilder unter einer perspektivischen Projektion dagegen können durchaus einen Schnittpunkt haben: Parallelen werden schließlich nicht auf Parallelen abgebildet. Diesem Schnittpunkt entspricht dann im Dreidimensionalen der unendlich ferne Punkt in Richtung der betrachteten Kantenrichtung. In manchen Fällen, wie bei der untenstehenden Abbildung, sind auch alle drei dieser Punkte zu sehen; sie bilden ein Dreieck, dessen Kantenurbilder den \mathbb{R}^3 im Unendlichen umrunden.



Dreipunktperspektive mit Fluchtpunktdreieck

Schließlich sind unendlich ferne Punkte auch nützlich, um den Definitionsbereich von Abbildungen zu erweitern und dadurch Fallunterscheidungen zu vermeiden: Beim Kameramodell war bislang aus gutem

Grund nur von Punkten vor oder hinter der Frontebene die Rede. Punkte die *auf* der Frontebenen liegen, aber vom Augpunkt verschieden sind, haben mit diesem eine Verbindungsgerade, die ganz in der Frontebene liegt und somit die Filmebene in keinem endlichen Punkt schneidet. Gemäß der üblichen Konvention kann man aber sagen, daß der unendlich ferne Punkt der Verbindungsgeraden gleichzeitig unendlich ferner Punkt der Filmebene ist, so daß sich dieser Punkt als Bildpunkt anbietet. Nachrechnen zeigt, daß dies in der Tat genau der Punkt ist, auf den auch die in homogenen Koordinaten ausgedrückte Abbildungsformel führt – sofern man das Ergebnis so erweitert, daß es keine Nenner mehr enthält. Somit kann man für alle Punkte des \mathbb{R}^3 mit Ausnahme des Augpunkts selbst einen Bildpunkt definieren, der in homogenen Koordinaten durch eine lineare Abbildung beschrieben werden kann. (Der Augpunkt würde unter dieser Formel auf den Punkt mit homogenen Koordinaten $(0 : 0 : 0 : 0)$ abgebildet, und so einen Punkt gibt es nicht.)

e) Verschiebungen, Drehungen, Streckungen

Um 4×4 -Matrizen in Aktion zu sehen, wollen uns einige gängige Transformationen anschauen. Am einfachsten sind die Translationen: Bei einer Verschiebung um den Vektor \vec{b} mit Komponenten x_0, y_0, z_0 ist die Matrix A gleich der Einheitsmatrix, die 4×4 -Matrix ist also

$$A^* = \begin{pmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Drehungen sind etwas komplizierter: Zwar ist der Verschiebungsanteil stets gleich dem Nullvektor, da die Drehachse punktweise fest bleibt, aber dafür kann der Matrixanteil ziemlich komplex werden. Beginnen wir daher mit dem einfachsten Fall einer Drehung um die z -Achse. Hier bleibt die z -Koordinate fest und für die (x, y) -Ebene haben wir das wohlbekanntes Sinus-Kosinus-Drehkästchen; insgesamt erhalten wir also die Matrix

$$A^* = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 & 0 \\ \sin \varphi & \cos \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ist die Drehachse eine beliebige Gerade durch den Nullpunkt, so können wir die Drehmatrix durch Basiswechsel aus der obigen erhalten, indem wir den Richtungsvektor der Drehachse als dritten Basisvektor nehmen, einen darauf senkrecht stehenden als ersten und das Kreuzprodukt dieser beiden als zweiten. Im neuen Koordinatensystem haben wir dann eine Drehung um die z -Achse; wir müssen also einfach, wie im ersten Studienjahr vielfach geübt, eine Abbildungsmatrix bezüglich einer neuen Basis darstellen.

Leider ist aber auch damit noch nicht die allgemeinste Rotation beschrieben: Wir müssen wir auch um Geraden rotieren können, die nicht durch den Nullpunkt gehen. Das läßt sich am einfachsten dadurch bewerkstelligen, daß wir zunächst um irgendeinen Vektor \vec{v} , der irgendeinen Punkt der Drehachse auf den Nullpunkt abbildet, *verschieben*, dann um die verschobene Drehachse rotieren, und schließlich noch um den Vektor $-\vec{v}$ verschieben.

Eine dritte, oft nützliche Transformation, ist die *Reskalierung*, die die Maßstäbe auf den drei Achsen unabhängig voneinander neu festlegt. Dies ist zwar keine Bewegung, aber doch oft nützlich: Beispielsweise kann man so Code zum Zeichnen einer Kugel wiederverwenden, um auch beliebige Ellipsoide zu zeichnen.

§5: Beleuchtung

Wenn wir einen einfarbigen Würfel in Richtung einer Raumdiagonalen projizieren, erhalten wir als Bild ein regelmäßiges Sechseck. Trotzdem muß eine Zeichnung schon sehr schlecht sein, wenn wir dort einen Würfel nicht von einem Sechseck unterscheiden können.

Der Grund dafür liegt darin, daß Licht praktisch nie völlig homogen aus allen Richtungen gleichzeitig kommt und daher von den verschiedenen Seitenflächen eines Würfels verschieden stark reflektiert wird; die Unstetigkeiten an den Würfelkanten lassen uns diese erkennen.

Die Richtung einer Polyederfläche läßt sich am einfachsten beschreiben durch ihren Normalenvektor, d.h. durch einen Vektor, der auf dieser

Fläche senkrecht steht. So ein Vektor läßt sich konstruieren als Kreuzprodukt zweier Kantenvektoren. Da dieses Kreuzprodukt nicht allgemein bekannt ist, sei es hier noch einmal kurz erklärt:

Das *Kreuzprodukt* oder *Vektorprodukt* ordnet zwei Vektoren aus $\vec{v}, \vec{w} \in \mathbb{R}^3$ einen dritten Vektor $\vec{v} \times \vec{w} \in \mathbb{R}^3$ zu; so ein Produkt existiert nur im Dreidimensionalen.

Der Vektor $\vec{v} \times \vec{w}$ ist festgelegt durch folgende Eigenschaften:

- $\vec{v} \times \vec{w}$ hat die Länge $|\vec{v} \times \vec{w}| = |\vec{v}| |\vec{w}| |\sin \angle(\vec{v}, \vec{w})|$.
Insbesondere ist also $\vec{v} \times \vec{w} = \vec{0}$, wenn \vec{v} und \vec{w} auf einer Geraden liegen, denn dann bilden sie einen Winkel von null oder 180 Grad, so daß der Sinus verschwindet.
- $\vec{v} \times \vec{w}$ steht senkrecht sowohl auf \vec{v} als auch auf \vec{w} .
Falls $\vec{v} \times \vec{w} \neq \vec{0}$ ist, spannen \vec{v} und \vec{w} eine Ebene auf, auf der (da wir im \mathbb{R}^3 sind) genau ein eindimensionaler Unterraum senkrecht steht. Darin gibt es allerdings für jede vorgegebene positive Länge zwei Vektoren, die sich durch ihr Vorzeichen unterscheiden. Um $\vec{v} \times \vec{w}$ eindeutig festzulegen, brauchen wir daher noch eine weitere Bedingung:
- Die drei Vektoren \vec{v}, \vec{w} und $\vec{v} \times \vec{w}$ bilden ein Rechtssystem, d.h. wenn sich die Finger der *rechten* Hand so ausrichten lassen, daß der Daumen in Richtung von \vec{v} zeigt, der Zeigefinger in Richtung von \vec{w} und der Mittelfinger in Richtung von $\vec{v} \times \vec{w}$.

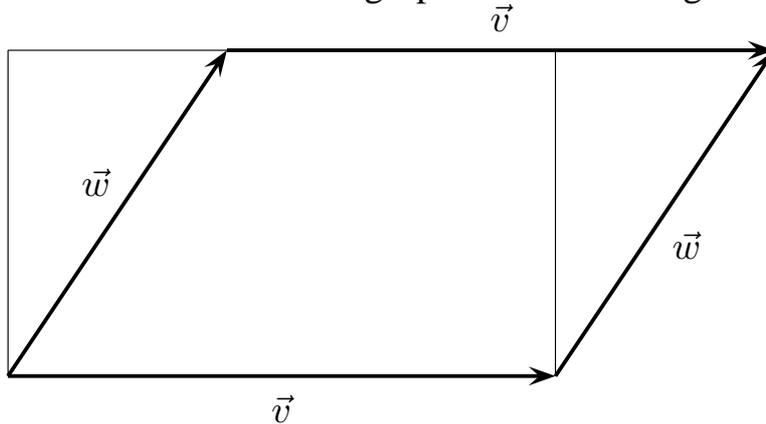
Alternativ kann man ein Rechtssystem auch so definieren, daß sich, ein von \vec{v} nach \vec{w} gedrehter Korkenzieher in Richtung $\vec{v} \times \vec{w}$ in den Kork bohrt. Ähnlich geht es auch mit Schrauben; da es allerdings neben den (üblichen) Rechtsschrauben auch die (seltenen) Linksschrauben gibt, ist diese Definition eventuell zirkulär: Alles hängt davon ab, wie man Rechtsschrauben definiert.

Aus jeder dieser Regeln folgt sofort die *Antikommutativität* des Vektorprodukts:

$$\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}.$$

Weitere Rechenregeln lassen sich leicht geometrisch ableiten: Da der Sinus eines Winkels gleich Gegenkathete durch Hypothenuse ist, ist in der

von \vec{v} und \vec{w} aufgespannten Ebenen $|\vec{w}| |\sin \angle(\vec{v}, \vec{w})|$ gleich der Länge des auf die senkrecht auf \vec{v} stehenden Geraden projizierten Vektors \vec{w} , das heißt also gleich der Höhe des in Abbildung 42 eingezeichneten Rechtecks. Die Länge des Vektors $\vec{v} \times \vec{w}$ ist daher gleich dem Flächeninhalt dieses Rechtecks und damit – wie eine Scherung zeigt – gleich der Fläche des von \vec{v} und \vec{w} aufgespannten Parallelogramms.



Geometrische Interpretation des Vektorprodukts

Dies zeigt das Distributivgesetz

$$\vec{v} \times (\vec{w} + \vec{u}) = \vec{v} \times \vec{w} + \vec{v} \times \vec{u},$$

für den zweiten Faktor, und wegen der Antikommutativität folgt daraus auch das für den ersten:

$$(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}.$$

Um das Vektorprodukt in Koordinaten ausrechnen zu können, müssen wir zunächst die Produkte der Koordinateneinheitsvektoren \vec{e}_i kennen. Da sie allesamt die Länge eins haben und paarweise aufeinander senkrecht stehen, ist klar, daß das Produkt zweier verschiedener dieser Vektoren bis aufs Vorzeichen gleich dem dritten ist; das Vorzeichen hängt ab von der Orientierung des Koordinatensystems. Das Produkt eines Vektors \vec{e}_i mit sich selbst ist natürlich, wie jedes Produkt eines Vektors mit sich selbst, gleich dem Nullvektor, denn der eingeschlossene Winkel ist null Grad.

Für die folgende Rechnung wollen wir annehmen, daß \vec{e}_1, \vec{e}_2 und \vec{e}_3 in dieser Reihenfolge ein Rechtssystem bilden; das ist beispielsweise dann

der Fall, wenn \vec{e}_1 nach rechts, \vec{e}_2 nach vorne und \vec{e}_3 nach oben zeigt. Dann folgt sofort, daß

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$$

ist, und nach einigen Fingerübungen auch findet man auch die Formeln

$$\vec{e}_2 \times \vec{e}_3 = \vec{e}_1 \quad \text{und} \quad \vec{e}_1 \times \vec{e}_3 = -\vec{e}_2 .$$

Die Produkte mit vertauschten Faktoren sind natürlich gerade das negative davon, und $\vec{e}_i \times \vec{e}_i = 0$. Für

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

ist also

$$\vec{v} \times \vec{w} = (v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3) \times (w_1\vec{e}_1 + w_2\vec{e}_2 + w_3\vec{e}_3),$$

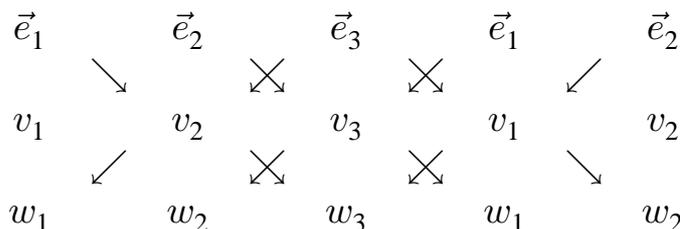
und nach obigen Rechenregeln ist dies gleich

$$\begin{aligned} & \sum_{i=1}^3 \sum_{j=1}^3 v_i w_j \vec{e}_i \times \vec{e}_j \\ &= (v_2 w_3 - v_3 w_2) \vec{e}_1 + (v_3 w_1 - v_1 w_3) \vec{e}_2 + (v_1 w_2 - v_2 w_1) \vec{e}_3, \end{aligned}$$

d.h.

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} .$$

Dies läßt sich dadurch merken, daß man im Schema



von \vec{e}_i ausgeht und als dessen Koeffizient das Zweierprodukt entlang der schrägen Linie nach rechts unten *positiv* und das entlang der schrägen Linie nach links unten *negativ* nimmt; man wendet also die SARRUSSche Regel an auf die „Determinante“

$$\begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} .$$

...