

A. BILDVEKTOREN UND SCHNITTOPERATOREN

Ausgehend von der Beobachtung, daß sich *Bewegungsträger* in Filmen wie starre Körper mit beschränkten Bewegungsmöglichkeiten verhalten, schreiben wir ihnen formal eine mathematische Funktion zu, die zu jedem Zeitpunkt ihren Ort in der jeweiligen Einstellung bezeichnet. Diese Funktion nennen wir *Bewegungsvariable* und benutzen diesen Begriff im Folgenden als Äquivalent des Bewegungsträgers, von dem sie erzeugt wird.

*

Das Konzept der Bewegungsvariablen ermöglicht uns, all das in einer Einstellung, was nicht zur Bewegung beiträgt, zur Beschreibung der Position des Bewegungsträgers zu benutzen. In diesem Sinne gehört sie einer gewissermaßen verallgemeinerten Ortskoordinate dieser Einstellung an. Wenn weiterhin t die eine Einstellung durchlaufende Zeitkoordinate darstellt, können wir eine Einstellung folgendermaßen parametrisieren:

$$s = (A_1(t), A_2(t), \dots , A_n(t), x, t)$$

wobei s Symbol der Einstellung ist, x Symbol der *verallgemeinerten Ortskoordinate*, t die Zeit und die $A_1(t)$, $A_2(t)$ etc. bis $A_n(t)$ die verschiedenen in der Einstellung enthaltenen zeitabhängigen Bewegungsvariablen der verschiedenen Bewegungsträger bezeichnen.

*

Eine solche Parametrisierung von s hat gewisse Ähnlichkeit mit dem, was in der Mathematik als *Vektor* bezeichnet wird. Im dreidimensionalen Raum etwa hat ein Vektor drei Komponenten, die wir als seine räumlichen Koordinaten bezeichnen. Benutzen wir für den Vektor etwa das Symbol v , für die x -, y - und z -Koordinaten die Symbole x , y und z , dann läßt sich schreiben:

$$v = (x, y, z)$$

In dieser Notation erkennt man die Ähnlichkeit mit unserer Parametrisierung der Einstellung s . Im Unterschied zu einem gewöhnlichen Vektor enthält s mehr als drei Koordinaten, von denen die Bewegungsvariablen selbst wieder abhängig von einer anderen Koordinate, nämlich der Zeit t

sind. Dies klingt kompliziert, aber der Mathematik ist es gelungen, solche Konstruktionen als mit dem üblichen Vektorbegriff im Wesentlichen identisch zu begreifen und zu behandeln. In diesem Sinne wollen wir eine Einstellung gelegentlich *Bildvektor* nennen. Das ist für das Folgende nicht unbedingt wichtig, erinnert aber daran, daß es sich bei Bildern nicht nur um ästhetische oder gar mythische Erscheinungen handelt, sondern daß in ihnen auch ein ganz prosaischer, geometrisch zu fassender Kern enthalten ist, mit denen man unter Umständen etwas Ähnliches anstellen kann wie mit einem gewöhnlichen Vektor.

*

Einen *Schnitt* begreifen wir in dieser Terminologie als einen *Operator*, der einem Zuschauer die raumzeitlichen Beziehungen zwischen zwei Bildvektoren vermittelt. Diesen Operator nennen wir *Schnittoperator*. Schnittoperatoren sind *Raumzeit-Operatoren*, die im Kopf des Zuschauers für die räumliche und zeitliche *Verschiebung* eines Bildvektors sorgen. Dies ist eine andere Formulierung dafür, daß das Geschehen des Film durch einen Schnitt von einer Ortskoordinate in eine andere verschoben wird. Bei einem Schnitt zwischen zwei Einstellungen s_1 und s_2

$$s_1 : s_2 = s_1 - C - s_2$$

sorgt der Schnittoperator C im Kopf des Zuschauers für eine Verschiebung der Raumzeit-Koordinate des aktuellen Geschehens aus dem Raum von s_1 in den Raum von s_2 . Das Symbol ":" ist hier wieder Symbol für einen Schnitt, wobei wir das Symbol "--" zunächst einmal als bloßes Trennungszeichen zwischen einer Einstellung und einem Operator verstehen.

*

Da *raum- und zeitverschiebende Operatoren* eine bekannte Klasse mathematischer Operatoren sind, lassen sich die Schnittoperatoren auf sie zurückführen. Die den Schnittoperatoren zugrunde liegenden Schnitte sind dann als künstliche Darstellungen dieser raumzeitverschiebenden Operatoren begreifbar, die vom Bewußtsein des Zuschauers für physikalisch real gehalten werden. Diese Betrachtungsweise ist nicht einfach, weil sie etwas sehr kompliziertes beschreibt. Dennoch vereinfacht sie die Analyse der Schnittformen insofern, als sie einen *deduktiven* Zugang zur Mannigfaltigkeit der möglichen Schnitte ermöglicht.

*

Es lohnt, sich die Mühe zu machen, diesen Prozeß einmal genau aufzuschlüsseln, um ihn genauer zu verstehen. Die Verhältnisse werden übersichtlicher, wenn wir uns die beiden Einstellungen s_1 und s_2 als kurz nacheinander von der gleichen Kamera aufgenommen denken. In einem Gedankenexperiment können wir uns dazu einen Raum mit einem auf seinem Boden aufgetragenen Liniengitter vorstellen, in dessen Zellen die geographischen Koordinaten jeder Zelle zehn-zentimetergenau eingeschrieben sind. Diese Koordinaten sollen automatisch mit abgebildet werden, wenn von diesem Raum ein Bild gemacht wird.

Schalten wir die Kamera ein, so beginnt sie am Ort x_1 ein Bild s_1 zu machen. Dabei verwandelt der Abbildungsprozeß den dreidimensionalen Raum, von dem die Schrift auf dem Boden ein Teil ist, durch eine Zentralprojektion in ein verkleinertes zweidimensionales Bild, welches sich schließlich nach Entwicklung, Fixierung und Kopierung als positives Kleinbild auf einem Filmstreifen wiederfindet.

Wird dieses in einem Kino projiziert, entsteht auf der Leinwand eine vergrößerte zweidimensionale Abbildung dieses kleinen Bildes. Auf ihm sieht der Betrachter ein Bild des aufgenommenen Raumes in perspektivischer Verkürzung, durch welche die der am Kamera am nächsten gewesene Zahl am Größten erscheint. Diese Zahl bezeichnen wir als die Ortskoordinaten x_1 der Einstellung s_1 . Sollte es sich dabei um eine repräsentative Einstellung handeln, verwandelt der Betrachter das von ihm wahrgenommene zweidimensionale Bild durch die Identifizierung von Bild und Abgebildeten in einen "wirklichen" dreidimensionalen Raum und bildet sich ein, er sehe diesen Raum mit der Koordinate x_1 dreidimensional vor sich.

*

Dadurch wird das Argument, das Projektionsbild ähnele nicht im mindestens dem Gesichtsfeld eines wirklichen Betrachters, vielleicht etwas klarer. Das tut es freilich nicht, denn ein solches ist verschwommen, unscharf, besteht aus einer Unmenge von Überlagerungen aufeinanderfolgender Saccaden, von denen manche sogar nur noch als Nachbild präsent sind, ein im mathematischen Sinne unglaubliches Durcheinander, auch zeitlich, das - ginge es in der Welt nach der Weltvorstellungen der Mathematik zu - eigentlich gar nicht als einheitlich Ganzes bezeichnenbar sein dürfte. Könnte man aber dennoch einmal ein solches Bild darstellen - und warum soll das nicht irgendwann gelingen - wäre dennoch gar nichts gewonnen: wenn ein Betrachter es nämlich auf der Leinwand wahrnimmt, würde er einer neuen Wirklichkeit begegnen und sich daraus ein

vermutlich noch wirreres Muster zusammenstellen, als das von Spielfilmen gewohnte, um sie zu begreifen. Es ist unsinnig vom Kamerabild als einem Objekt zu sprechen, das Ähnlichkeit mit dem "Bild" eines Sehenden hat. Richtig ist vielmehr, daß das Leinwandbild *ähnlich gesehen werden kann* wie die Wirklichkeit, deshalb kann man es auch mit ihr verwechseln. Diese Punkte sind fein, aber weittragend.

*

Wenn wir im Folgenden davon reden, daß ein Betrachter eine Einstellung s mit einem Bildraum x sehe, meinen wir immer diesen dreidimensionalen "wirklichen" Raum, den er sich am Ende der Produktionskette der Filmherstellung beim Sehen einbildet. Man muß sich aber darüber im klaren sein, daß damit eigentlich dieser komplexe mehrstufige Transformationsprozeß gemeint ist, wobei stillschweigend angenommen wird, daß er "ideal" abläuft und es keine größeren Komplikationen durch etwa optische Täuschungen, schlechte Bildqualität oder blanke Unaufmerksamkeit des Betrachters gibt.

*

Wenn wir dieses Experiment wiederholen und wieder s_1 an der Stelle x_1 aufnehmen, dann aber zusätzlich und noch während des Drehens die Kamera bewegen, ändern sich die abgebildeten Koordinaten gemäß dieser Bewegung. Den dabei entstehenden Teil der Einstellung nennen wir s_{bew} . Diese Bewegung kann beliebig sein, eine Fahrt etwa, ein Schwenk oder beides. Wenn wir nach einiger Zeit anhalten, werden nur noch die Koordinaten des Ortes abgebildet, bei dem wir angehalten haben. Diesen Ort nennen wir x_2 und den Teil der Einstellung in dem es keine Veränderung mehr gibt s_2 .

Wenn wir die dabei entstandenen Bilder später projizieren, sehen wir zunächst den Bildraum x_1 von s_1 mit seinen Koordinaten vor uns (ich möchte ein letztes mal daran erinnern, daß mit diesem so einfach klingenden Satz in Wirklichkeit die vorhin beschriebenen äußerst komplexen Verhältnisse gemeint sind), dann die wandernden Koordinaten aus dem sich verändernden Bildraum von s_{bew} und schließlich die Koordinaten von x_2 aus s_2 . Da wir den Prozeß in ganzer Länge sehen konnten, ist uns klar, was passiert ist: wir sagen daß der Raum von x_1 sich durch eine Kamerabewegung in den Raum von x_2 verwandelt hat oder daß das neue Bild s_2 aus dem ersten s_1 durch eine Kamerabewegung hervorgegangen ist. Diese Interpretation macht keine Schwierigkeiten, weil wir sie in ähnlicher Form bei jedem Spaziergang machen.

*

Schwieriger wird es, wenn wir während beim Betrachten dieses Vorgangs während der Fahrt s_{bew} die Augen schließen und sie erst wieder zum Beginn von s_2 öffnen. Das entspräche in etwa einem Spaziergang, bei dem man während des Gehens einige Zeit die Augen geschlossen (oder an etwas anderes gedacht) hat. Man weiß wieder, daß s_2 aus s_1 durch eine Kamerabewegung hervorgegangen ist, die in der Zwischenzeit stattgefunden hat, man weiß allerdings nicht mehr, durch welche Art von Bewegung. Da wir aber die Koordinaten von x_2 sehen und mit den Koordinaten von x_1 vergleichen können, wissen wir sofort, wie weit und in welcher Richtung die Kamera sich inzwischen bewegt haben muß. An der Änderung der Fluchtlinien des Gitters auf dem Boden können wir sogar erkennen, wie stark die Kamera sich dabei gedreht hat.

Dabei ist der tatsächliche Weg der Kamera von nebensächlicher Bedeutung. Man kann sich nämlich unendliche viele Wege vorstellen, auf denen die Kamera von x_1 kommend nach x_2 gefahren ist, für das Resultat spielt es keine Rolle. Es reicht, sich eine möglichst unkomplizierte, also eine gewissermaßen minimale Bewegung von x_1 nach x_2 vorzustellen, um sich über die räumlichen Verhältnisse Klarheit zu verschaffen.

*

Schwieriger wird es, wenn der Teil mit der Kamerabewegung ganz aus der Einstellung herausgenommen wird. Dann haben wir einen wirklichen Schnitt $s_1 : s_2$ vor uns, und auf einen solchen können uns unsere Spaziergänge, kann uns die Evolution, eigentlich unmöglich vorbereitet haben. Dennoch: wenn wir wissen, daß beide Filme von der gleichen Kamera aufgenommen sind, ist uns völlig klar, daß s_2 durch eine uns vorenthaltende Kamerabewegung aus s_1 hervorgegangen sein muß. Da wir die Koordinaten beider Einstellungen zu lesen und die Fluchtlinien des Gitters auf dem Boden miteinander zu vergleichen vermögen, sind wir weiterhin in der Lage, uns eine minimale Kamerabewegung vorstellen, die eine solche Verschiebung von x_1 nach x_2 erzeugen könnte. Damit verliert der Schnitt von s_1 nach s_2 für uns einen Großteil seines Geheimnisses.

Es gibt allerdings eine irritierende Stelle, den Moment genau zwischen s_1 und s_2 , an dem die Fahrt hätte stattfinden müssen. Wird sie weggelassen, müssen wir uns das Weggelassene Zeitintervall durch einen *Zeitsprung* Δt ergänzen. Leider haben wir zunächst einmal keine Ahnung, wie groß dieser Zeitsprung war. Es kann sein, daß die minimale Kamerabewegung, die

man sich vorstellt, ganz langsam vonstatten ging - dann wäre Δt relativ groß - oder ganz schnell, dann wäre der Zeitsprung unter Umständen sogar sehr kurz. Kopiert man während der Aufnahme fortwährend eine Uhr oder etwas Uhr-Ähnliches ein, könnte man den Zeitsprung direkt im Bild ablesen.

*

So weit so gut. Bis jetzt hatten wir es mit Vorgängen zu tun, die *real* abliefen, und deren Realität vom Zuschauer nur nachempfunden werden mußte. Wir kommen nun zu den für den Filmherstellungs- und Wahrnehmungsprozeß entscheidenden Komplikationen, durch die der Zuschauer *betrogen* wird und *betrogen werden will*. Zunächst einmal kann man sich vorstellen, daß die Kamera während der, sagen wir einmal zwanzig Sekunden, während welcher sie von x_1 nach x_2 fährt, abgeschaltet und erst dann wieder angeschaltet wird. Das spielt für den Zuschauer keine Rolle: da er s_{bew} ohnehin nicht sieht, braucht es auch gar nicht erst gedreht zu werden.

*

Eine größere Komplikation stellt dagegen dar, wenn die Kamera gar nicht fährt und stattdessen eine zweite Kamera die Einstellung s_2 an der Stelle x_2 aufnimmt. Da man das Bild nicht von dem von dem von ersten Kamera Gemachten unterscheiden kann, macht es keinen Unterschied, so lange behauptet und geglaubt wird, es wäre von einer Kamera aufgenommen. Tatsächlich muß es gar nicht erst behauptet werden. Werden die beiden Einstellungen hintereinander gezeigt, nimmt der Zuschauer weiterhin einfach an, daß s_2 durch eine Kamerabewegung aus s_1 hervorgegangen ist und verbindet die beiden Bildräume x_1 und x_2 weiterhin durch eine minimale Kamerafahrt.

*

Das ändert sich auch nicht, wenn wir den Raum x_2 samt dem Koordinatengitter komplett abbauen, ihn mit einem Lastwagen in eine andere Stadt transportieren, ihn dort wiederaufbauen und ihn dort mit einer zweiten Kamera als Einstellung s_2^* aufnehmen. Zwar stimmen die objektiven Koordinaten nicht mehr mit dem auf dem Bild sichtbaren überein, aber das ist für den Zuschauer auf dem Bild nicht zu erkennen. Für ihn ist s_2 nicht von s_2^* zu unterscheiden, es ist für ihn die gleiche Einstellung wie vorher. Wenn die beiden Einstellungen daher zusammengeschnitten werden, nimmt man weiterhin an, s_2 (bzw. s_2^*) wäre aus s_1 durch eine

Kamerafahrt hervorgegangen und versucht weiterhin, sich einen minimalen Weg vorzustellen, auf welchem dies geschehen konnte.

*

Stellen wir uns weiter vor, in einer anderen Stadt werde ein Raum aufgebaut, auf dessen Boden sich ein Koordinatensystem mit den zu x_2 gehörigen aufgemalten Zahlenwerten befindet: wird dort mit einer dritten Kamera eine Einstellung s_3 gemacht, die dem schon bekannten s_2 weitgehend ähnelt, so ändert sich für den Zuschauer auch nichts, wenn er s_1 und s_3 durch einen Schnitt $s_1 : s_3$ verbunden sieht. Wieder nimmt er an, s_3 wäre aus s_1 durch eine Kamerafahrt hervorgegangen, und wieder stellt er sich den gleichen minimalen Weg vor.

*

Sogar wenn s_3 objektiv vor s_1 gedreht worden sein sollte, kann ein Betrachter, wenn ihm die Einstellungen in der Reihenfolge $s_1 : s_3$ vorgeführt werden, keinen Unterschied zu einem der vorigen Schnitte entdecken. Er interpretiert sie weiterhin auf seine bewährte Weise, auch wenn dies mittlerweile den tatsächlichen Aufnahmeverhältnissen Hohn spricht. Mit anderen Worten: es kommt nicht auf den Aufnahmeprozeß an, sondern darauf, wie der Zuschauer das geschnittene Resultat interpretiert.

*

Das ändert sich auch nicht, wenn das Koordinatensystem im Bild weggelassen wird. Der Zuschauer wird sich weiterhin bemühen, eine Art Verbindung zwischen den beiden Einstellungen herzustellen. Dabei wird er versuchen, sich ein *eigenes Koordinatensystem* zurechtzubasteln. Dabei wird er aber nur selten von genauen Zahlen ausgehen, sondern eher die Koordinatensysteme unserer täglichen Orientierung benutzen, welche nach folgendem Muster zu funktionieren scheinen: "Hier ist mein Zimmer, rechts nebenan mein Schlafzimmer - vor meiner Tür gibt es ein Treppenhaus, gehe ich vier Stockwerke auf Stufen nach unten erreiche ich eine Straße, fünfzig Meter nach rechts bis zum Parkplatz, wenn ich vier Minuten mit dem Auto Richtung Stadt fahre, eine Kreuzung ... usw." Ein solches Ordnungssystem meinen wir, wenn wir in einer Einstellung

$$s = (A_1(t), A_2(t), \dots, A_n(t), x, t)$$

von verallgemeinerten Ortskoordinaten x sprechen. Beim Versuch einen beliebigen Schnitt $s_1 : s_2$ mit einem derartigen Koordinatensystem zu verbinden, wird ein Zuschauer versuchen, die beiden aufeinander folgenden Einstellungen zunächst einmal als durch eine Kamerabewegung auseinander hervorgegangen zu begreifen und sich dabei einen minimalen Kameraweg vorstellen, der diese Vermutung stützt. Ein solcher Weg ist letztlich nichts als eine Transformation der eigenen Wege, mit denen man den Zusammenhang von Örtlichkeiten erlebt und vor sich beschreibt. Da auch dieses subjektive Ordnungssystem von extrem unterschiedlichen Weltausschnitten, sowohl Details als auch Totalen, zusammengesetzt ist, erstaunt beim Betrachten eines Films auch nicht besonders, wenn bei der Aufnahme von s_1 und s_2 verschiedene Objektive mit ganz anderen Tiefenschärfeverhältnissen benutzt werden. Nimmt er derartiges überhaupt wahr (und den meisten Zuschauern ist das Näherbetrachten eines Objektes so selbstverständlich, daß er kaum bemerkt, wenn es im Film geschieht) denkt er sich eben während der Bewegung noch eine Heran- oder Wegfahrt dazu, oder wenn er technisches Vokabular liebt: einen Objektivwechsel.

*

Wenn wir also die den Schnittoperatoren zugrunde liegenden Schnitte als künstliche Darstellungen von raumzeitverschiebenden Operatoren begreifen, die vom Bewußtsein des Zuschauers für physikalisch real gehalten werden, meinen wir damit sehr schnell im Kopf ablaufende, gewissermaßen stenographische Verkürzungen dieser von ihm vermuteten Kamerabewegungen.
