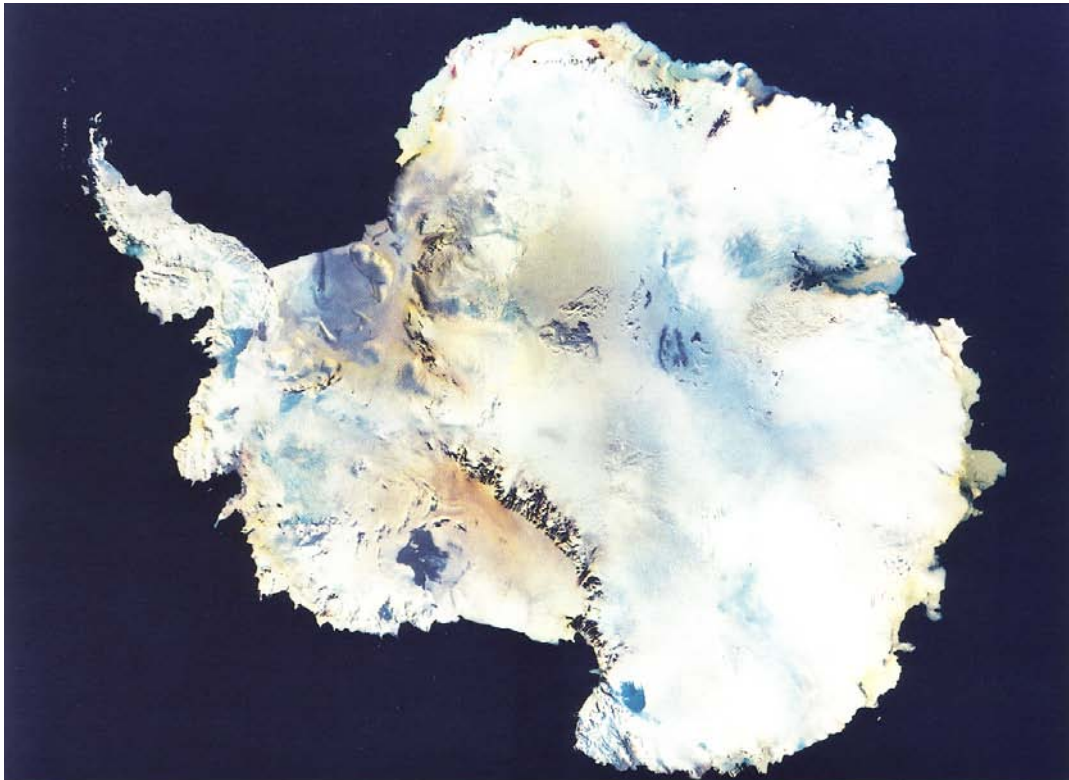


Erdfernerkundung



Inhaltsverzeichnis

1	Einführung	1
1.1	Definition	1
1.2	Historisches	4
1.3	Ziele der Erdfernerkundung	6
1.4	Ein Beispiel – Definition der Anforderungen	8
1.5	Aufbau der Vorlesung	10
2	Satellitenbahnen	12
2.1	Historisches	12
2.2	Physikalische Grundlagen	14
2.2.1	Eingeschränktes Zweikörperproblem	14
2.2.2	Satellitenbahnen und Erhaltungsgrößen	15
2.2.3	Ellipsenbahnen	18
2.3	Typische Bahnen von erdgebundenen Satelliten	22
2.3.1	Typen von Erdsatelliten	22
2.3.2	Erderkundungs- und Aufklärungssatelliten	22
2.3.3	Wetter- und Kommunikationssatelliten	24
2.3.4	Navigationsatelliten	25
2.3.5	Forschungssatelliten	26
2.3.6	Zusammenfassung	27
2.4	Stabilisierung des Satelliten	28
2.4.1	Stabilisierungsverfahren	28
2.4.2	Sonnensynchrone Bahnen	29
2.5	Bahnstörungen	31
2.5.1	Übersicht	31
2.5.2	Reibung in der Hochatmosphäre	33
2.5.3	Bahnstörungen in Folge der Erdabplattung	37
2.5.4	Zusammenfassung	39
2.6	Einbringen des Satelliten in seine Bahn	40
2.6.1	Allgemeine Probleme	40
2.6.2	Raketengleichung	40
2.6.3	Mehrstufenprinzip	42
2.6.4	Transferbahnen (Hohmann-Bahnen)	45
2.7	*Spezielle Bahnen	49
2.7.1	Lagrange-Punkte	49
2.7.2	Planetares Billard	54
3	Satelliteninstrumente	61
3.1	Grundlagen	61
3.1.1	Atmosphärische Transmission	61
3.1.2	Technische Randbedingungen	64
3.1.3	Grobklassifizierung	65
3.1.4	Kenngrößen für Satelliteninstrumente	66

3.2	Passive Instrumente im sichtbaren Spektrum	66
3.2.1	Kurzer Rückblick	67
3.2.2	Kameratypen	67
3.2.3	Abbildung durch Objektive	69
3.2.4	Aufzeichnung der Abbildung: photographische Emulsion	69
3.2.5	Warum Falschfarben?	73
3.2.6	Charakterisierung eines photographischen Systems	75
3.2.7	Elektronische Aufzeichnung statt photographischer Emulsion	81
3.2.8	Scanner: Spiegel statt Objektiv	86
3.2.9	Instrumente der nächsten Generation	90
3.2.10	Zusammenfassung	92
3.3	Passive Instrumente im nicht-sichtbaren	92
3.3.1	Thermisches Infrarot	93
3.3.2	Mikrowelleninstrumente	99
3.4	Aktive Instrumente	101
3.4.1	Altimeter und Scatterometer	101
3.4.2	SAR Synthetic Aperture Radar	104
3.5	Sounder	106
3.5.1	Prinzip	106
3.6	Anwendungsbeispiele	113
3.6.1	Archäologie und Anthropologie	113
3.6.2	Vegetation	114
3.6.3	Hydrosphäre	119
4	Missionen	130
4.1	Der klassische Geograph: LandSat	130
4.2	Ein Meteorologe: MeteoSat	131
4.3	Ein tiefer gelegter Meteorologe: TIROS bis NOAA	132
4.4	Der Generalist für Umweltfragen: EnviSat	134
4.5	Small is beautiful: CHAMP und Diamant	136
4.6	Lost and found: GPS-NAVSTAR	137
4.7	Mission to Planet Earth	138
4.7.1	Zielsetzung	139
5	Kommunikation	143
5.1	Historisches	143
5.1.1	Rückführung des Filmmaterials	143
5.1.2	Onboard-Scans	145
5.1.3	Technische Anmerkungen	145
5.2	Information	145
5.2.1	Kommunikation	146
5.2.2	Was ist Information?	148
5.2.3	Informationsgehalt	149
5.2.4	Shannon-Funktion und Entropie	152
5.2.5	Redundanz	152
5.2.6	Abtasttheorem (Signal als kontinuierliche Funktion der Zeit)	159
5.2.7	Wahrnehmungspsychologie: Verwendung von Hyperzeichen im Erkennen	160
5.2.8	Zusammenfassung	165
5.3	Codierung	165
5.3.1	Quellencodierung	166
5.3.2	Kanalcodierung	173
5.3.3	Zusammenfassung	181
5.4	Technische Anmerkungen	181
5.4.1	Überblick	181
5.4.2	Lösungsansätze	182

5.4.3	Kommunikation über Satelliten	185
6	Datenaufbereitung	187
6.1	Korrekturverfahren	187
6.1.1	Fehlende Pixel	188
6.1.2	Periodisch fehlende Streifen	189
6.1.3	Periodische Streifen	191
6.1.4	Random Noise	191
6.1.5	Dunstkorrektur	193
6.2	Bildverbesserung	195
6.2.1	Kontrastanreicherung	195
6.2.2	Kantenanreicherung	199
6.3	Automatische Klassifikation	202
6.3.1	Distanzverfahren	202
6.3.2	Mehrdimensionale Klassifikationsverfahren	205
6.4	Managing Gigabytes	206
6.4.1	Verlustfreie Bildformate	207
6.4.2	Und JPEG?	207
6.4.3	Fraktale Bildkompression	210
A	Anhang	216
A.1	Liste der verwendeten Symbole	216
A.2	Nützliche Konstanten	218
A.3	Abkürzungen	218
A.4	Spickzettel	220
B	Hinweise zur Lösung einiger Aufgaben	221
B.1	Kapitel 2	221
B.1.1	Fragen	221
B.1.2	Aufgaben	227
B.2	Kapitel 3	230
B.2.1	Fragen	230
B.2.2	Aufgaben	235
B.3	Kapitel 6	236
B.3.1	Fragen	236
B.3.2	Aufgaben	238

Kapitel 1

Einführung

Erdfernerkundung hat sicherlich einen ästhetischen Aspekt, vgl. z.B. Abb. 1.1 oder Titelbild (AVHRR-Aufnahme der Antarktis). Davon sprechen auch die zahlreichen Bücher [9, 45, 56], in denen meist photographische Aufnahmen der Erde vom Weltraum aus gezeigt werden – häufig mehr unter ästhetischen als unter wissenschaftlichen Aspekten. Viele dieser Aufnahmen werden heutzutage kommerziell vertrieben, z.B. für Ikonos unter <http://www.spaceimaging.com/> oder <http://www.euspaceimaging.com/>, auf deren Seiten auch eine Vielzahl von Beispielen (inkl. Enduring Freedom oder Olympische Spiele) zu finden sind, oder für verschiedene aktuelle Satelliten wie ASTER, EROS, SPOT 5, IRS, QuickBird, Ikonos und LandSat 7 unter <http://www.satimagingcorp.com/index.html>.

Die typischen Beispiele der Erdfernerkundung, wie in Abb. 1.2 Osnabrück und das Osnabrücker Schloss in einer Aufnahme des russischen Ikonos-Satelliten, erinnern an deren frühe Ziele: (militärische) Aufklärung und Kartographie.

Moderne Erdfernerkundung ‘photographiert’ jedoch nicht nur die Erdoberfläche mit ihren offensichtlichen Merkmalen sondern dringt, z.B. zur Erkundung von Bodenschätzen oder für archäologische Fragestellungen, in den Erdboden ein oder stellt im optischen nicht sichtbare Phänomene wie das Ozonloch dar, vgl. Abb. 1.3.

Abbildung 1.3 gibt gleichzeitig einen Hinweis auf die Vorteile der Satelliten gestützten Erdfernerkundung gegenüber Messungen vom Boden aus: Erdfernerkundung ermöglicht globale Betrachtungen und kann damit die dynamischen Aspekte eines Phänomens besser beschreiben. Und Erdfernerkundung ermöglicht Langzeitbeobachtungen (Monitoring) und ist damit zur Untersuchung von langfristigen Veränderungen (Stichwort ‘global change’ oder ‘climate change’) unerlässlich.

Tabelle 1.1 gibt einen Überblick über einige universell einsetzbare Satelliteninstrumente mit ihren wichtigsten Eigenschaften. Viele von ihnen werden wir im Laufe der Vorlesung genauer diskutieren, ebenso wie Instrumente, die speziellen Fragestellungen dienen.

1.1 Definition

Das einfachste (und wohl auch älteste) Fernerkundungsinstrument dürften Kameras an Drachen (!), Ballonen und später auch Flugzeugen gewesen sein. Erdfernerkundung, oder zumindest die Satelliten gestützte Erdfernerkundung, ist auch ein Teilgebiet der Extraterrestrik, d.h. der Erforschung des Weltraums und der Forschung vom Weltraum aus. Da ich im Bereich der Extraterrestrischen Physik ausgebildet bin, habe ich diesen Aspekt natürlich auch im Auge – das wird Ihnen im Laufe der Vorlesung manchmal auffallen, z.B. mit Hinweisen auf Fernerkundung an anderen Planeten, was im Rahmen der vergleichenden Atmosphärenforschung auch zum Verständnis der Erdatmosphäre wichtig ist.

Ballone und Raketen mit komplizierteren, auch auf die Messung von elektromagnetischen Feldern und geladenen Teilchen ausgelegten Instrumenten, bildeten bis in die siebziger Jahre das Rückgrat der extraterrestrischen Forschung. Heute treten beide zunehmend hinter den

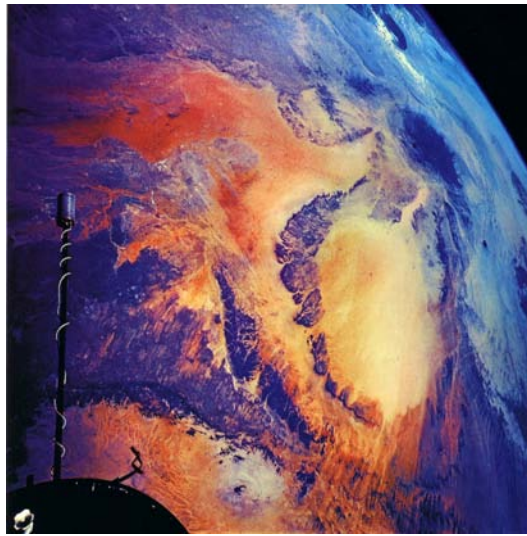


Abbildung 1.1: Wüstenlandschaft des Fessan (Lybien), aufgenommen 1966 von Gemini XI mit einer Hasselblad mit 38 mm Objektiv aus einer Höhe von 324 km [9]

Satelliten zurück, die eine für langfristige Anwendungen geeignete Plattform für komplexe Systeme einander ergänzender Instrumente bilden. Insgesamt ergibt sich damit die folgende Definition:

Definition 1 *Unter Extraterrestrischer Physik versteht man Forschung in allen Disziplinen der Physik, einschließlich der Astro- und der Geophysik, die mit Hilfe von Satelliten, Raketen oder hochfliegenden Ballonen durchgeführt wird. Dazu gehören selbstverständlich auch die Interpretation der Beobachtungen und Vergleiche mit theoretischen Untersuchungen.*

Aus dieser Definition wird einerseits eine teilweise etwas willkürliche Abgrenzung zu einigen Gebieten der Geo- und Astrophysik deutlich, andererseits zeigt sich jedoch auch der häufig interdisziplinäre Charakter der extraterrestrischen Physik. Letzterer ergibt sich u.a. dadurch, dass auf einzelnen Raumfahrzeugen eine Vielzahl von Instrumenten aus unterschiedlichen Bereichen geflogen wird, so dass die Planung und Vorbereitung der Mission ebenso wie die anschließende Datenauswertung dazu führt, sich nicht nur mit dem eigenen Instrument, sondern auch mit den Techniken und Messungen anderer Experimente zu beschäftigen. In den letzten Jahren ist dieser interdisziplinäre Charakter noch deutlicher geworden, da eine immer größer werdende Zahl der mit den Methoden der Extraterrestrischen Physik untersuchten

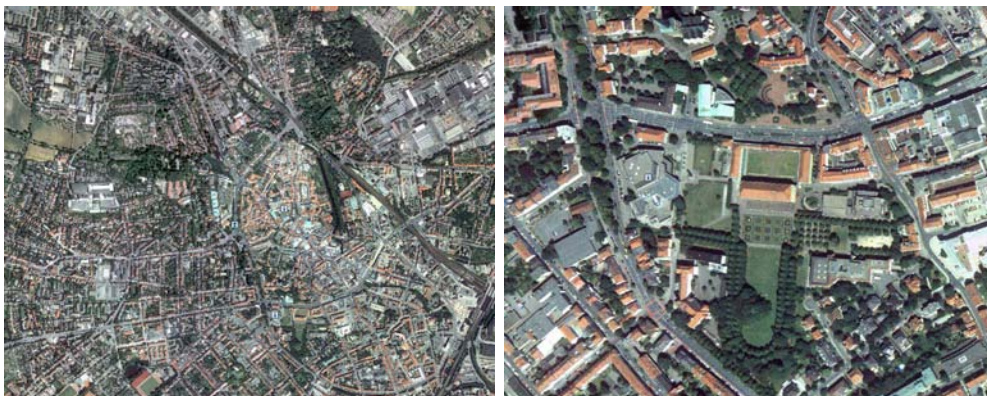


Abbildung 1.2: Osnabrück (links) und Osnabrücker Schloss (rechts) vom Satelliten Ikonos

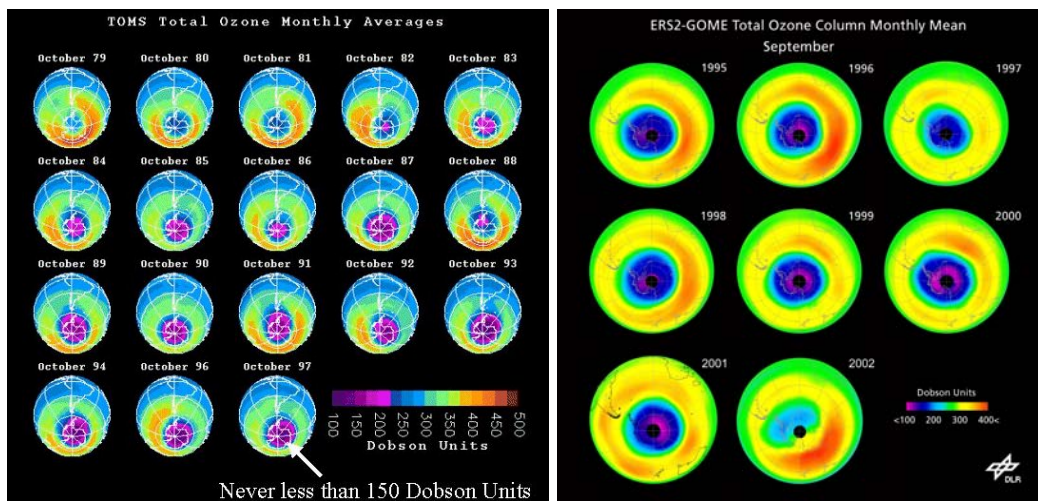


Abbildung 1.3: Ozonloch über der Antarktis: links TOMS-Messungen von 1979 bis 1997 (Quelle <http://www.geo.arizona.edu/Antevs/nats104/00lect28ozonspol.gif>), rechts GOME-Messungen zwischen 1995 und 2002 (Quelle: http://www.caf.dlr.de/caf/aktuelles/news-archiv/bilderarchiv/gome/ozonloch2002/x_images/gome_to3_momean-antarctica.jpg)

Fragestellungen aus Bereichen der Biologie, Ökologie, Meteorologie oder Geophysik kommt – insbesondere natürlich auch die uns im Rahmen der Erdfernerkundung interessierenden Fragen. Selbst primär zur Erdfernerkundung gebaute Satelliten wie die POES/NOAA-Serie verfügen über Kombinationen von Instrumenten zur Wettervorhersage, einem Universalinstrument AVHRR zur Erdfernerkundung und dem Space Environment Monitor SEM zur Untersuchung energiereicher Teilchen und Plasmen in der Erdmagnetosphäre und über den Polkappen.

Wenn wir den häufig vorgebrachten Einwand der hohen Kosten der Weltraumfahrt berücksichtigen, stellt sich die Frage: Was können wir durch Beobachtungen von Satelliten aus lernen, was wir nicht auch vom Erdboden aus untersuchen können? Auf diese Frage gibt es eine Vielzahl möglicher Antworten.; einige Beispiele seien die:

- die störenden Einflüsse der Erdatmosphäre auf die elektromagnetische Strahlung einerseits und der Erdmagnetosphäre auf energiereiche Teilchen andererseits lassen sich teilweise ausschalten. Von Satelliten aus ist der Blick nach draußen daher nicht nur auf wenige Wellenlängenbereiche beschränkt, sondern es lassen sich auch Röntgen- und Gammaastronomie betreiben. Auch können energiereiche Teilchen, Plasmen und elektromagnetische Felder im Sonnensystem in situ nachgewiesen und untersucht werden – diese sind aber für die Erdfernerkundung nicht unerheblich, da sie insbesondere die höhere Atmosphäre beeinflussen und Einfluss auf die Kommunikation zwischen Satellit und Bodenstation haben.
- andere Planeten können untersucht werden, wodurch wir die Erde besser verstehen und in einen größeren Zusammenhang stellen können. Für die Erdfernerkundung ist dies unter dem Gesichtspunkt der vergleichenden Planetenforschung von Bedeutung. Insbesondere ein Vergleich der Planetenatmosphären ist für die Untersuchung des 'climate change' von Interesse, da der anthropogene Treibhauseffekt eng mit einem Anstieg der CO_2 -Konzentration in der Atmosphäre verbunden ist und Mars und Venus zwei Planeten sind, deren Atmosphären im wesentlichen aus CO_2 bestehen aber dennoch sehr unterschiedliche Eigenschaften haben.
- im Hinblick auf Erdfernerkundung ermöglichen der große Abstand des Beobachters (Satelliten) zum Erdboden und seine Relativbewegung einen schnellen Überblick über weite Bereiche der Erde (remote sensing). Dies ist wichtig für viele umweltrelevante Fragen und unser Verständnis globaler Zusammenhänge.

Satellitentyp	Bahnhöhe (km) Umlaufart	Repetition (Tage)	Streifen- breite (km)	räumliche Auflösung im Nadir (m x m)	Aufnahme- systeme	Wellenlänge der Kanäle in μm
Landsat 1 Landsat 2 Landsat 3	930; fast polar	18	185	79 x 56 nominelle, 79 x 79 effektive Pixelgröße	MSS und RBV * nur Landsat 3	RBV MSS 0,475-0,575 0,6- 0,7 0,580-0,680 0,7- 0,8 0,690-0,830 0,8- 0,9 0,510-0,750 * 0,8- 1,1 10,4-12,6 *
Landsat 4 Landsat 5 Landsat 6 *	705; fast polar	16	185	30 x 30 120 x 120 m im TIR-Kanal-6 13 x 15 (pan*)	TM und MSS * nur Landsat 6	TM MSS 0,45-0,52 0,6-0,7 0,52-0,60 0,7-0,8 0,63-0,69 0,8-0,9 0,76-0,90 0,9-1,1 1,55-1,75 10,4-12,5 2,08-2,35 0,50-0,90 (pan*)
SPOT 1 SPOT 2 SPOT 3	830; fast polar	26 aber variabel 1-5	60	20 x 20	HRV	multispekt. pan 0,50-0,59 0,51-0,73 0,61-0,68 0,79-0,89
HCMM	620; fast polar	16	716	500 x 500 600 x 600	HCMR	0,55-1,1 10,5-12,5
SEASAT	800; fast polar	14	4 000	25 x 25	SAR und VIRR	SAR VIRR L-Band 0,49-0,94 250 mm 10,5-12,5
NOAA-9 NOAA-10 NOAA-11 NOAA-12	850; fast polar	1 x tägl. 2 x tägl.	3 000	1 100 x 1 100	AVHRR * nicht auf NOAA-10	0,58-0,68 0,725-1,1 3,55-3,93 10,3-11,3 11,4-12,4 *
METEOSAT SMS/GOES GMS, INSAT	36 000; geostat.	30 Minuten	„full disk“	2 500 x 2 500 5 000 x 5 000	VISSR (MSR)	0,4-1,1 5,7-7,1 10,5-12,5
ERS-1	785; fast polar	35	100	12,5 x 12,5	AMI ATSR	AMI ATSR C-Band 1,6; 3,7 56,5 mm 10,8; 12,0
Spaceshuttle STS-7; -11;	230 – 330; geneigt	Einzelmission 1983, 1984	140	20 x 20	MOMS-01	0,575-0,625 0,825-0,975
Spacelab Mission D-2;	230 – 330; geneigt	Einzelmission 1993	78 37 für Band 5	13,5 x 13,5 4,5 x 4,5 für Band 5	MOMS-02	multispekt. pan 0,440-0,505 0,55-0,77 0,530-0,575 0,55-0,77 0,645-0,680 0,55-0,77 0,810-0,760
Spacelab-1 STS-9	240 – 257 geneigt	Einzelmission 1983	189	20 x 20	Metric Camera	Schwarzweiß- und Farbinfrarotfilm
KOSMOS;	270; geneigt	Einzelmission 3 – 4 /Jr.	80	5 – 10	KFA-1 000	Schwarzweiß- und Spektronzonalfilm 0,57-0,67; 0,67-0,81
Spaceshuttle STS-2; STS-13	225 – 352; geneigt	Einzelmission 1981; 1984	180 – 285		LFC	Schwarzweißfilm 0,40-0,90

Tabelle 1.1: Wichtige Satelliteninstrumente und ihre Parameter [71]

1.2 Historisches

Erdfernerkundung oder remote sensing ist keine Erfindung des Satellitenzeitalters (vgl. z.B. Diskussionen und Abbildungen in [16, 117]). Zwar ist der Begriff ‘remote sensing’ erst in den Sechzigern im US Office of Naval Research eingeführt worden, zu einer Zeit, als sich die Verwendung ‘spionierender’ Satelliten aus dem militärischen Bereich heraus in den zivilen Bereich ausdehnte. Im wörtlichen Sinne bedeutet remote sensing, dass etwas aus der Ferne wahrgenommen wird, d.h. man möchte etwas über ein Objekt erfahren, ohne in direktem physischem Kontakt mit ihm zu stehen. In diesem Sinne ist Photographie ein Beispiel für remote sensing, ebenso basiert die gesamte Astrophysik auf remote sensing.

Photographie ist das entscheidende Stichwort zur Geschichte des remote sensing. Die ersten Ansätze zum remote sensing finden sich bereits in der Lochkamera und in der Kamera Obscura. Für die Verwendung photographischer Systeme in einem gewissen Abstand von der Erdoberfläche ist das erste überlieferte Photo 1858 von einem Ballon in der Nähe von Paris aufgenommen worden.

Außer Ballonen als Kameraträger wurden bereits sehr früh automatisierte unbemannte Fernerkundungssysteme entwickelt. Dazu gehörten Kameras, die an Drachen befestigt waren

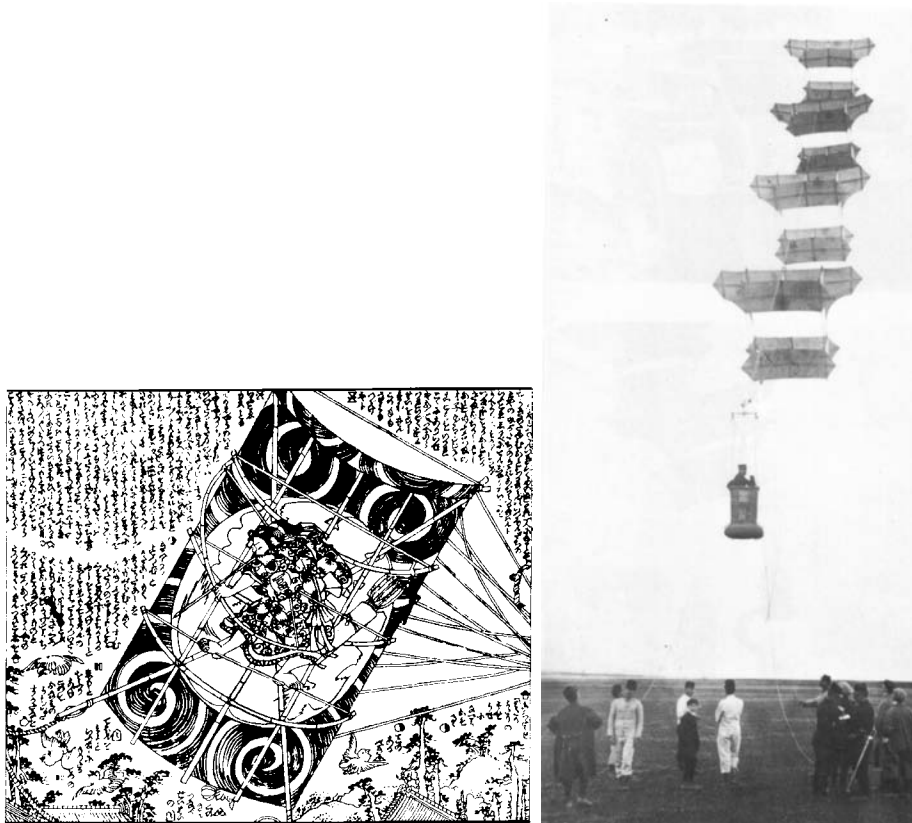


Abbildung 1.4: Erdfernerkundung ohne photographische Aufzeichnung in Japan (links) und Europa (rechts) [85]

(Strain und Engle [117] zeigen dazu eine aus 600 m Höhe gemachte Drachenaufnahme von San Franzisko kurz nach dem Erdbeben im Jahre 1906). Diese Kameras mussten automatisiert sein in dem Sinne, dass sie selbsttätig zu einer bestimmten Zeit den Verschluss öffneten bzw. bei besseren Versionen nach einer Belichtung auch den Film transportierten, den Verschluss erneut spannten und in einer bestimmten zeitlichen Sequenz die Belichtungen wiederholten. Nach der Entwicklung besonders leichter automatisierter Kameras wurden auch Tauben (!) als Kameraträger eingesetzt.

Mit dem Aufkommen des Flugzeugs boten sich neue Möglichkeiten der Erdfernerkundung. Auch wenn Bewegung und Vibrationen der Flugzeuge eine Schwierigkeit bei den Aufnahmen darstellten, so eröffneten sie doch neue Möglichkeiten: Kameras an Drachen und Tauben waren automatisiert und ließen sich daher nicht auf ein Zielgebiet ausrichten. Mit einem Flugzeug dagegen bot sich die Möglichkeit, sehr gezielt Luftaufnahmen eines bestimmten Gebietes unter vorgegebenen Bedingungen zu machen. Diese Möglichkeiten wurden während des ersten Weltkrieges zur militärischen Nutzung intensiviert und weiterentwickelt. Nach dem ersten Weltkrieg fanden sich die ersten systematischen zivilen Anwendungen in der Erdfernerkundung, z.B. für forstwirtschaftliche Kontrollflüge, geologische Untersuchungen und zur Sammlung von Luftbildern als Grundlage für die Kartographie.

Während des zweiten Weltkrieges wurden Luftbilder intensiv zur Aufklärung verwendet, darunter auch zur Erfolgskontrolle und Zielplanung für Luftangriffe auf Städte, vgl. z.B. [54]. Seit dieser Zeit wurden auch Versuche durchgeführt, Erdfernerkundung nicht nur im sichtbaren Bereich des elektromagnetischen Spektrums zu betreiben, sondern auch im Infraroten und durch die Verwendung von Radar. Ohne diese beiden Sensortypen würde die Fernerkundung heutzutage wesentlich weniger Informationen liefern.

Der nächste Sprung in der Erdfernerkundung erfolgte mit dem Beginn des Satellitenzeit-

Archaeology and anthropology	Oceanography
Cartography	surface temperature
Geology	geoid
surveys	bottom topography
mineral resources	winds, waves and currents
Land use	circulation
urban land use	mapping of sea ice
agricultural land use	oil pollution monitoring
soil survey	Meteorology
health of crops	weather systems tracking
soil moisture and evapotranspiration	weather forecasting
yield predictions	sounding for atmospheric profiles
rangelands and wildlife	cloud classification
forestry — inventory	Climatology
forestry, deforestation, acid rain, disease	atmospheric minority constituents
Civil Engineering	surface albedo
site studies	desertification
water resources	Natural disasters
transport facilities	floods, earthquakes,
Water resources	volcanoes, forest fires,
surface water, supply, pollution	subsurface coal fires,
underground water	landslides
snow and ice mapping	Planetary Studies
Coastal studies	
erosion, accretion, bathymetry	
sewage, thermal and chemical pollution monitoring	

Tabelle 1.2: Verwendung von remote sensing Instrumenten [16]

alters, markiert durch den Start von Sputnik 1 im Oktober 1957. Der erste zivile Erdfernerkundungssatellit wurde 1960 gestartet, es handelte sich um einen Meteorologiesatelliten der TIROS-Serie (Television and Infra-Red Observation Satellite). Wie der Name dieses Satelliten bereits andeutet, diente er vielfältigen Zwecken: der Meteorologie durch Erzeugung von Bildern im sichtbaren wie Infraroten Bereich ebenso wie der Kommunikation. Die gleichzeitige Verwendung als Kommunikationssatellit ist übrigens ein bei geostationären Meteorologiesatelliten häufig verwendetes Verfahren, auch die uns aus dem Wetterbericht geläufigen MeteoSats haben eine, wenn auch begrenzte Aufgabe im Bereich der Kommunikation. Teilweise handelt es sich dabei um das Aufsammeln wissenschaftlicher Daten von unbemannten Stationen, z.B. im Rahmen des ARGOS-Netzes durch die NOAA-Satelliten, oder um die Weiterverteilung bereits bearbeiteter meteorologischer Daten vom globalen Netz der geostationären Satelliten [16].

Dieser kurze Rückblick ist allerdings lückenhaft insofern, als dass er erst mit der Verwendung von photographischen Systemen, d.h. in der Mitte des 19. Jahrhunderts beginnt: erst das Aufkommen der Photographie erlaubt eine Aufzeichnung der beobachteten Szene. Die Beobachtung erfolgte jedoch schon früher durch den Menschen. Als Träger wurden Drachen verwendet, eine Variante, die in Japan anscheinend schon seit dem 11. Jahrhundert bekannt ist (eine Darstellung aus dem 18. Jahrhundert findet sich im linken Teil von Abb. 1.4). Diese Verfahren wurden in der zweiten Hälfte des 19. Jahrhunderts auch in Europa angewandt: zum einen war die aufkommende Photographie noch nicht so automatisiert, dass man Instrumente hätte alleine fliegen lassen können, zum anderen diente der Manlifter für Studien zur Entwicklung von Trägersystemen für Menschen, sprich Flugzeugen.

1.3 Ziele der Erdfernerkundung

Nachdem der Mensch nun über Jahrzehnte versucht hat, seine Kameras in immer größeren Höhen zu betreiben, welche Ziele hat er sich davon versprochen und was hat er damit erreicht? Die ersten Ideen hatten sicherlich etwas mit Entdeckergeist zu tun: man wollte sich einen Überblick verschaffen, einen anderen Blickwinkel bekannten Dingen gegenüber finden. Heutzutage ist die Erdfernerkundung wesentlich zielgerichteter, insbesondere sind die Fragen, die an die Satellitendaten gestellt werden bzw. die zur Entwicklung neuer Instrumententypen

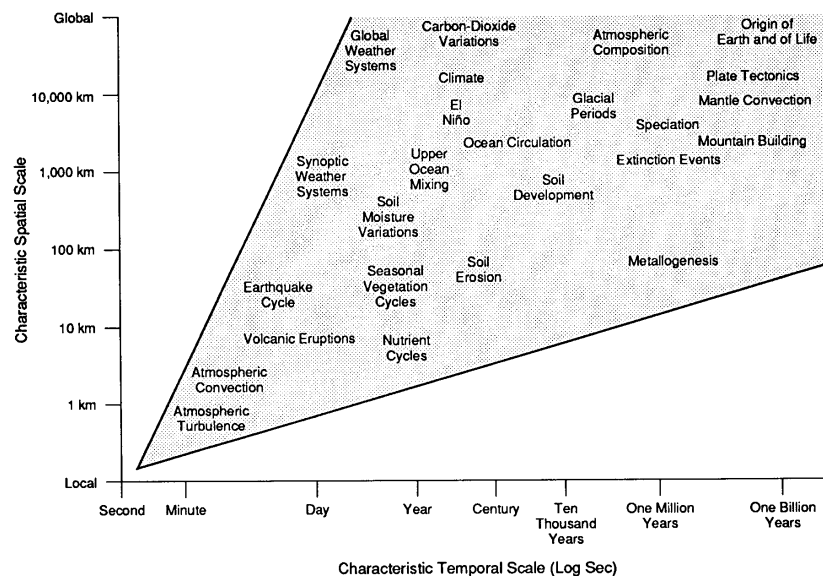


Abbildung 1.5: Charakteristische räumliche und zeitliche Skalen für einige der durch remote sensing beobachteten Schlüsselprozesse auf der Erde [24]

führen, wesentlich spezieller. Tabelle 1.2 gibt dazu einen Überblick über die Bereiche der Erderkundung, in denen die Verwendung von remote sensing Instrumenten üblich ist.

Für einige dieser Bereiche ist der Nutzen der Erdfernerkundung sofort offensichtlich: in der Kartographie bedarf es keiner näheren Erläuterung; in der Meteorologie haben wir uns an die Verwendung von Satellitendaten für die tägliche Wettervorhersage gewöhnt; Instrumente, die in der Ozeanographie zur Bestimmung des Geoids, der Meeresbodentopographie sowie der Windgeschwindigkeiten, Wellen und Strömungen dienen, werden wir im Zusammenhang mit den aktiven Mikrowelleninstrumenten in Abschn. 3.4.1 kennen lernen.

Andere Anwendungen von remote sensing dagegen sind weniger offensichtlich. Wie und mit welchen Instrumenten soll man z.B. Archäologie von einem Satelliten aus betreiben? Wie kann man Wasserreservoirs unter der Erdoberfläche erkennen? Einige dieser Anwendungsbeispiele werden im Laufe der Vorlesung (und im Skript) näher erläutert werden.

Aus diesen unterschiedlichen Fragestellungen ergeben sich unterschiedliche Anforderungen an die Satelliten und Satellitenexperimente. Ein Blick auf Tabelle 1.2 zeigt, dass die verschiedenen Fragestellungen auch mit unterschiedlichen zeitlichen und räumlichen Skalen verbunden sind. Für Kartographie z.B. benötigt man eine gute räumliche Auflösung. Die zeitliche Auflösung dagegen kann gering sein, dafür sind im Hinblick auf natürliche wie auf anthropogene Änderungen Zeitskalen in der Ordnung von einem Jahr ausreichend, für geologische Veränderungen (Kontinentaldrift, Gebirgsbildung und -zerstörung) sogar Zeitskalen von Millionen Jahren. Wettersysteme dagegen werden auf ganz anderen Skalen untersucht. Hier sind räumliche Skalen im Bereich von einigen zehn oder hundert Kilometern völlig ausreichend. Allerdings sind Wettersysteme schnell-veränderlich, d.h. sie müssen kontinuierlich oder zumindest auf Zeitskalen deutlich kleiner als ein Tag beobachtet werden. Abbildung 1.5 gibt eine Übersicht über die zur Betrachtung verschiedener Phänomene notwendigen räumlichen und zeitlichen Skalen.

Das räumliche Auflösungsvermögen ist durch die Eigenschaften des Instruments und die Flughöhe des Satelliten festgelegt, das zeitliche Auflösungsvermögen dagegen ergibt sich aus der Wiederholfrequenz der Satellitenbahn. Insbesondere bei den Wettersatelliten zeigen sich dabei deutliche Unterschiede: Satelliten der Meteosat-Serie befinden sich in einer geostationären Umlaufbahn, d.h. sie haben mit relativ schlechter räumlicher Auflösung kontinuierlich den gleichen Bereich der Erde im Blickfeld. Hier ist die Wiederholfrequenz durch die Eigenschaften des Instruments gegeben, so erzeugen die Meteosats alle halbe Stunde

eine komplette Aufnahme ihres Blickfeldes. Die Wettersatelliten der NOAA Klasse (National Oceanographic and Atmospheric Administration) dagegen fliegen in niedrigen polaren Orbits. Damit ergibt sich ein wesentlich besseres räumliches Auflösungsvermögen und damit die Möglichkeit zum Studium kleinerer Details (wichtig insbesondere bei der Prognose und Diagnose extremer Wetterlagen), allerdings um den Preis, dass jeder einzelne Satellit einen bestimmten Bereich nur ein oder zweimal pro Tag überfliegt. Daher lassen sich mit einem einzelnen Satelliten keine Studien zur Dynamik von Wettersystemen betreiben, hier ist eine größere Zahl von Satelliten notwendig (planmäßig sollten sich stets drei bis vier der NOAA-Satelliten im Orbit befinden).

1.4 Ein Beispiel – Definition der Anforderungen

Die Themenbereiche, die die physikalischen Grundlagen der Erdfernerkundung umfassen, können wir uns mit Hilfe eines einfachen Planspiels klar machen.

Ihre Aufgabe ist der Bau eines Satelliten mit geeigneter Instrumentation zur Detektion von Waldbränden in ausgedehnten und dünn besiedelten Gebieten (z.B. Nationalparks in den USA). Diese Waldbrände sollen nicht erst dann entdeckt werden, wenn die Rauchsäule so hoch ist, dass der halbe Bundeststaat in Flammen steht und die Rauchsäule von allen anderen Bundesstaaten aus auch vom Boden gesehen werden kann, sondern bereits dann, wenn man das Feuer noch kontrolliert (z.B. durch Brandschneisen) abbrennen kann.

Die Lösung dieses Problems ist nicht eindeutig, sehr unterschiedliche Konfigurationen von Satellit und Instrument(en) sind denkbar. Eine Suche nach der Lösung des Problems kann folgende Strategie verfolgen:

1. was sind die physikalischen (und mit einem remote sensing Instrument nachweisbaren) Eigenschaften des zu untersuchenden Objekts?
2. mit welcher Art von Instrument(en) können diese nachgewiesen werden?
3. mit welcher Genauigkeit müssen diese Größen gemessen werden?
4. wo müssen sie gemessen werden?
5. wie oft müssen sie gemessen werden?
6. welche Information über die Eigenschaften des Objekts soll an den Beobachter (d.h. die Bodenstation) übertragen werden?
7. wann soll übertragen werden?

Die erste Frage ist die entscheidende, hier ist etwas Brainstorming hilfreich. Wenn wir Feuer nachweisen wollen, brauchen wir uns nur unsere Vorstellung von Feuer zu vergegenwärtigen: der offenbare Aspekt von Feuer sind Flammen, d.h. veränderliche, rötlich-gelbe Leuchterscheinungen und Wärme. Oder in physikalischer Terminologie: ein Feuer emittiert Licht im sichtbaren Bereich des elektromagnetischen Spektrums und Wärme, d.h. elektromagnetische Strahlung im thermischen Infrarot.

Und die Instrumente (Pkt. 2)? Zum Nachweis der Emission im sichtbaren Bereich reicht es doch, eine einfache Kamera zu verwenden, die im sichtbaren Bereich des Spektrums beobachtet und damit die zum Feuer gehörenden Flammen nachweist. Allerdings: Feuer ist auch mit Rauchentwicklung verbunden, d.h. unsere Kamera kann das Feuer wahrscheinlich gar nicht sehen, da es unter einer Rauchwolke den Blicken entzogen ist. Also müssten wir beim Nachweis eines Feuers im sichtbaren Bereich des Spektrums die Abbildungen auf Flammen und/oder Rauch untersuchen.

Wo wir gerade bei Rauchwolken sind: was macht ein derartiges Instrument eigentlich bei einem Wolken bedeckten Himmel? Dann sieht ein Instrument im sichtbaren Bereich des elektromagnetischen Spektrums nur die Oberkante der Wolken, alles darunter entzieht sich seinen Blicken – und der Wald könnte großflächig in Flammen stehen bevor eine Wolkenlücke einmal einen Blick auf das Feuer gewährt.

Unser Waldbrandnachweiser muss also um etwas ergänzt (oder durch etwas ersetzt) werden, dass auch bei Wolkenbedeckung Informationen vom Erdboden aufnehmen und auswerten

kann. Wolkenbedeckung kann von Mikrowellenstrahlung durchdrungen werden, auch im thermischen Infrarot kann man, zumindest teilweise, durch die Wolken sehen. Und ein Sensor im thermischen Infrarot kann auch einen Aspekt eines Waldbrandes nachweisen, nämlich die Wärmeentwicklung – auch dann, wenn das eigentliche Feuer bereits unter einer Rauchwolke verschwindet.

Unsere Wunschinstrumentierung könnte daher ein Sensor im thermischen Infrarot sein, eventuell ergänzt (für Schönwetter) um einen Sensor im sichtbaren Licht, da im sichtbaren die Informationen für einen menschlichen Dateninterpretier natürlich einfacher auszuwerten sind.

Bevor wir uns (gemäß Pkt. 3) an eine Spezifikation der Instrumente machen, müssen wir unser Konzept noch überprüfen: zwar würde diese Instrumentkombination die zu untersuchenden Effekte nachweisen – aber wie oft würde sie einen Fehlalarm auslösen? Unser Infrarotsensor reagiert auch thermische Emission – das ist nicht zwingend ein Waldbrand, auch ein Hochofenabstich, eine startende Rakete, ein Nutz- oder Schadfeuer, das Abfackeln von Erdgas an einer Ölbohrstelle, ein Leck in einer Ölpipeline u.v.a.m. sind mit Wärme und damit einem Signal im thermischen Infrarot verbunden. Damit der Satellit nicht immer zu Waldbrandalarm gibt, müssen wir ein Verfahren entwickeln, Waldbrand von den anderen Wärmequellen zu unterscheiden. Bei gutem Wetter kann dabei sicherlich die Kamera im sichtbaren Bereich helfen: das spricht dafür, beide Instrumente zu fliegen und nicht nur den Sensor im thermischen IR alleine. Zusätzlich können wir dem Satelliten eine Karte mitgeben, auf der Hochöfen und andere stationäre Wärmeproduzenten eingetragen sind, so dass an diesen Stellen ein Infrarotsignal nicht als Waldbrand identifiziert wird.

Bei der Instrumentspezifikation (Pkt. 3) sind wichtige Gesichtspunkte das räumliche Auflösungsvermögen und sowie das spektrale/thermische Auflösungsvermögen. Auf Grund der verwendeten großen Wellenlängen haben Sensoren im thermischen Infrarot ein schlechteres Auflösungsvermögen als solche im sichtbaren: größenordnungsmäßig 100 m im thermischen IR vergleichen mit größenordnungsmäßig 10 m im sichtbaren Bereich des Spektrums. Eine Kamera erkennt den Ort also genauer als ein IR-Sensor.¹ Sie muss ihn auch genauer erkennen, da die Flammen im sichtbaren Bereich (insbesondere wenn unter der Rauchwolke verborgen) einen geringeren Kontrast mit der Umgebung bilden als es die Wärmequelle im IR tut. Daher können Flammen im sichtbaren erst entdeckt werden, wenn sie einen großen Teil des durch die Bodenauflösung gegebenen Bildpunktes (Pixels) ausfüllen, im thermischen IR dagegen muss nur ein kleiner Teil des Pixels in Flammen stehen um dessen Helligkeitswert so weit anzuheben, dass er sich gegen die Umgebung deutlich abhebt. Auf dieser Basis können wir uns jetzt mit einigen ESA oder NASA-Technikern zusammen setzen und die Detailspezifikationen der Instrumente erarbeiten – oder sie arbeiten das Skript durch und kehren dann zu diesem Beispiel zurück, um die Spezifikationen selbst zu geben.

Die Punkte 4 und 5 unserer Liste betreffen nicht mehr die Instrumentierung sondern den Satelliten bzw. genauer dessen Bahn. Selbst wenn die Forstbehörden des Bundesstaats Washington die Auftraggeber für den Satelliten und die Hauptabnehmer der Daten sind, wird es uns nicht gelingen, den Satelliten so zu positionieren, dass er immer über Seattle steht, da sich die Erde unter der Satellitenbahn weg dreht. Daher ist es am besten, eine globale Waldbrandüberwachung anzubieten und die Daten an andere Interessenten zu verkaufen. Allerdings brauchen wir keine vollständige globale Überwachung: in der Arktis und Antarktis fehlen die für einen Waldbrand erforderlichen Wälder. Der Satellit kann also auf den Überflug über hohe Breiten verzichten. Dies lässt sich durch die Wahl der Inklination seiner Bahn erreichen. Die Flughöhe des Satelliten in seiner Bahn hat Konsequenzen für das Bodenauflösungsvermögen und den vom Satelliten während eines Überfluges beobachteten Bereich (und damit auch für die Wiederholfrequenz, d.h. die Frequenz, mit der ein bestimmter Punkt auf der Erde von dem Satelliten eingesehen werden kann – für Waldbrandüberwachung ist es nicht ganz optimal, wenn der Satellit nur alle 2 Wochen vorbeischaud). Beide Größen lassen

¹Bei ausgedehnten Wäldern in unbewohnten Gebieten reden wir allerdings nicht über 1 km mal 1 km große Parzellen sondern eher über Gebiete mit Kantenlängen von einigen 10 km, d.h. selbst ein Auflösungsvermögen von 1 km wäre noch hilfreich.

sich durch geschickte Kombination von Flughöhe und Spezifikation des Instruments über einen relativ weiten Bereich einstellen.

Bleiben noch die beiden letzten Punkte. Diese betreffen die Behandlung der von den Instrumenten gewonnenen Daten. Für die Waldbrandentdeckung ist es sicherlich nicht sinnvoll, wenn die Daten stets erst an Bord gespeichert werden und dann einmal im Monat an eine Bodenstation gegeben werden. Andererseits ist es bei direkter Datenübertragung für einen Überwacher am Boden vielleicht nicht zwingend notwendig, zweimal am Tag ein neues Bild von den Wäldern um Seattle zu erhalten und zu interpretieren. Daher könnte man für einen Waldbrandsatelliten ein anderes Verfahren der Datenverarbeitung vorsehen: die Datenverarbeitung an Bord des Satelliten soll ein Feuer automatisch erkennen (z.B. durch Vergleich der Aufnahmen mit einer gespeicherten Karte und/oder den Aufnahmen der Vortage; Auslösen von Feueralarm wenn in einem oder mehreren Bildpunkten ein vorher fest gelegter Helligkeitswert überschritten wurde). Die Datenübertragung würde dann wie folgt aussehen: kein Anzeichen für ein Feuer, so wird ein kurzes ok übertragen (sonst weis die Bodenstation nicht, ob der Satellit kein Feuer gesehen hat oder ob es technische Probleme gibt) und die Bodenstation fordert das Bild nur bei Bedarf an. Findet sich ein Anzeichen für Feuer, so wird ein Alarm ausgelöst und das Bild übertragen. Auf Grund des Alarms weis die Bodenstation auch, dass dieses Bild mit hoher Priorität genauer zu untersuchen ist und nicht nur eine Routineaufnahme darstellt. Durch eine derartige automatisierte Vorverarbeitung wird also nicht nur die zu übertragende Datenmenge relativ klein gehalten sondern auch die Aufmerksamkeit der Beobachter am Boden auf den entscheidenden Datensatz gelenkt.

1.5 Aufbau der Vorlesung

Mit Hilfe des voran gegangenen Beispiels haben wir die wesentlichen Anforderungen an das Design einer Satellitenmission definiert:

- die Bahn, die (mit den Eigenschaften des Instruments) die Wiederholfrequenz und das Bodenaufklärungsvermögen bestimmt,
- die Auswahl und Spezifizierung des primären Satelliteninstruments,
- Auswahl und Spezifizierung ergänzender Instrumenten,
- die Datentransmission, sowie
- die automatische Datenverarbeitung (an Bord oder am Boden) zur Identifikation.

Das Skript folgt in seinem Aufbau dieser Anforderungsliste.

In Kapitel 2 werden wir Satellitenbahnen diskutieren. Dazu gehören die aus den Kepler'schen Gesetzen bestimmten elementaren Bahnen, Kreisbahn oder Ellipse, ebenso wie kompliziertere Bahnen wie das scheinbare Ruhen eines Satelliten im Lagrange-Punkt auf der Verbindungsachse Sonne–Erde oder das für Missionen zu anderen Planeten wichtige planetare Billard. Beispiele für typische Satellitenbahnen werden ebenso vorgestellt wie Bahnstörungen und ihre Ursachen. Dabei werden wir auch lernen, dass einige der Bahnstörungen nicht nur negative Auswirkungen haben sondern auch zur Erzeugung bestimmter Bahnen verwendet werden können. Wir werden kurz auch das Einbringen des Satelliten in seine Bahn diskutieren und seine Stabilisierung in der Bahn.

In Kapitel 3 werden Beispiele für typische Instrumente auf Erdfernerkundungssatelliten vorgestellt. Dabei werden auch ältere Instrumente genauer diskutiert:

- sie sind einfach, so dass sich die physikalischen Grundlagen gut verstehen lassen,
- sie bilden das Rückgrat für alle Untersuchungen zum global change und werden zur Erzeugung einer kontinuierlichen Datenbasis teilweise heute noch nachgebaut bzw. am Leben erhalten,
- sie sind Universalinstrumente und nicht auf eine sehr spezielle Fragestellung optimiert,
- die meisten modernen Instrumente sind eher eine Weiterentwicklung dieser klassischen Instrumente als Neuentwicklungen, d.h. die Grundlagen haben sich nicht geändert, lediglich einige technische Details.

Das Kapitel enthält auch Hinweise auf Internet-Ressourcen. ...

Kapitel 4 stellt sehr kurz einige Satellitenfamilien bzw. Missionen vor: auf unseren Wettersatelliten beziehen wir uns immer als MeteoSat – z.Z. sind wir allerdings bei bereits bei MeteoSat-8 und dieser unterscheidet sich von seinen Vorgängern nicht unerheblich.

Bisher haben wir Erdfernerkundung eher unter dem messtechnischen Aspekt betrachtet. Wie bei einer Labormessung müssen die Daten aber auch ausgewertet werden. Bevor wir uns mit einigen Aspekten der Datenaufbereitung, insbesondere der Bildverarbeitung, in Kapitel 6 beschäftigen können, müssen wir einen Aspekt des remote Sensing betrachten, der bei einer Labormessung nicht in dieser Form auftritt: die Kommunikation über hunderte oder Tausende von Kilometern entlang gestörter Kommunikationskanäle. Dazu werden wir uns in Kapitel 5 mit den Grundlagen der Kommunikation beschäftigen, insbesondere mit der Quellencodierung zur Verminderung der Redundanz und mit der Kanalcodierung, d.h. dem Zufügen von Redundanz um eine Fehlererkennung und -korrektur zu ermöglichen. In Kapitel 6 werden wir einen kurzen Ausflug in die Bildbearbeitung vornehmen. In beiden Kapiteln werden uns Bekannte aus der täglichen Arbeit mit dem Computer begegnen, wie z.B. der Lempel-Ziv-Algorithmus als Grundlage von ZIP und der Bildkompression in GIF und PNG oder JPEG als verlustbehaftetes Bildformat.

Fragen und Aufgaben finden sich in allen Kapiteln (außer 4), Musterlösungen für einen Teil der Aufgaben finden Sie im Anhang.

Literatur ist ebenfalls am Ende jedes Kapitels angegeben. Wirklich empfehlenswert als Begleitung für die gesamte Vorlesung ist der Sabins [101] als ausführliches, mit vielen Beispielen versehenes Werk, und Cracknell und Hayes [16] als kompaktes Bändchen. Die NASA stellt ferner ein Online-Tutorial zur Erdfernerkundung zur Verfügung, das viele (auch militärische oder von Flugzeugen gemachte) Beispielaufnahmen enthält: <http://rst.gsfc.nasa.gov/start.html>. Das Tutorial verweist auch auf weitere online-Ressourcen.

Kapitel 2

Satellitenbahnen

In diesem Kapitel werden die Grundlagen der Bahnmechanik besprochen. Dazu gehört die Bewegung eines Testkörpers im Schwerfeld eines wesentlich massiveren Körpers und die sich daraus ergebenden grundlegenden Bahnen, die durch Kegelschnitte beschrieben werden (Abschn. 2.2). Für die Erdfernerkundung sind von diesen Bahnen nur die geschlossenen Bahnen, d.h. Kreis- und Ellipsenbahnen, interessant. Beispiele für typische Bahnen von Erdsatelliten, insbesondere im Hinblick auf Flughöhe und Inklination, werden diskutiert (Abschn. 2.3). In Abschn. 2.7 werden neben diesen typischen Bahnen auch spezielle Bahnen diskutiert, die weniger für die Erdfernerkundung als vielmehr für die Erforschung des Planetensystems von Bedeutung sind: das scheinbare Verharren eines Satelliten im Lagrange-Punkt auf der Verbindungslinie Erde–Sonne und das u.a. von den Voyager und Pioneer Raumsonden durchgeführte interplanetare Billard, bei dem die Sonden im Gravitationsfeld eines Planeten in Richtung auf den nächsten Planeten abgelenkt werden.

2.1 Historisches

Die Flugbahn eines Satelliten um einen Planeten ist, ebenso wie die Bahn eines Planeten um ein Zentralgestirn, durch die Anziehungskraft zwischen den beiden Körpern sowie die Bahngeschwindigkeit des Satelliten/Planetens bestimmt.

Die Beschreibung antriebsloser Satellitenbahnen kann daher mit Hilfe der Kepler'schen Gesetze erfolgen:¹

1. Die Planeten bewegen sich auf elliptischen Bahnen in deren einem Brennpunkt die Sonne steht.²
2. Die Radiusvektoren von der Sonne zu den Planeten überstreichen in gleichen Zeiten gleiche Flächen (*Flächensatz*).
3. Die dritten Potenzen der großen Halbachsen a sind proportional zu den Quadraten der Umlaufzeiten, d.h.

$$\frac{T^2}{a^3} = c_k \tag{2.1}$$

mit c_k als einer für alle Planeten gleichen Konstante.

Zur Anwendung der Kepler'schen Gesetze auf Satelliten ist jeweils das Wort Planet durch Satellit zu ersetzen und Sonne bzw. Zentralgestirn durch Planet.

Anschaulich sind die Kepler'schen Gesetze in Abb. 2.1 dargestellt.

¹Kepler veröffentlichte seine ersten beiden Gesetze 1609 in *Astronomica nova de motibus stellae Martis*, das dritte nach weiteren Beobachtungen und einer verfeinerten Datenanalyse zehn Jahre später in *Harmonices mundi libri V*.

²Das erste Kepler'sche Gesetz beinhaltet, dass die Bahnen der Planeten geschlossene Bahnen sind. Das ist nicht trivial, eine alternative Form der Bewegung mit nicht geschlossenen Bahnen wären Präzessionsorbits.

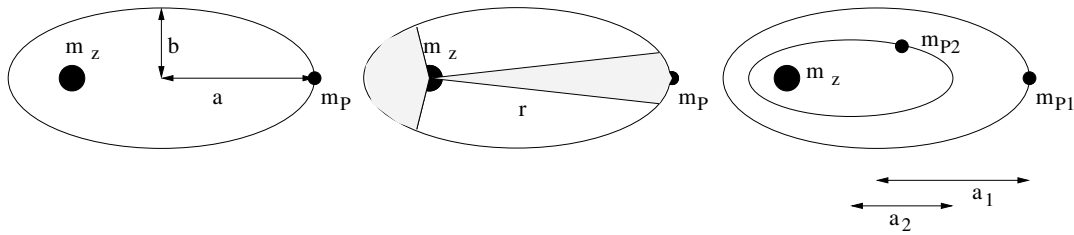


Abbildung 2.1: Kepler'sche Gesetze

Die Kepler'schen Gesetze sind empirisch, d.h. sie beruhen auf Beobachtungen und sind nicht aus grundlegenden Prinzipien abgeleitet. Erst Newton gelang das Verständnis dieser Gesetze in dem er erkannte, dass das von ihm gefundene Gravitationsgesetz nicht nur den Fall eines Apfels auf der Erde beschreibt sondern ebenso auch die Bewegung eines Planeten um die Sonne.³ Da ein Planet, so Newton's Argumentation, sich nicht auf einer gradlinigen Bahn bewegt, muss gemäß Trägheitsgesetz eine Kraft auf ihn wirken. Diese ist die Gravitationskraft des Zentralkörpers, gegeben als

$$\vec{F} = -\gamma \frac{M_Z m_p}{r^2} \vec{e}_r \quad (2.2)$$

mit M_Z und m_p als den Massen von Zentralkörper und Planet, r als deren Abstand, \vec{e}_r als dem Einheitsvektor entlang der Verbindungsachse zwischen den beiden (mit dem Zentralkörper als Ausgangspunkt) und $\gamma = 6.67 \cdot 10^{-11} \text{ Nm}^2/\text{kg}^2$ als der universellen Gravitationskonstante.

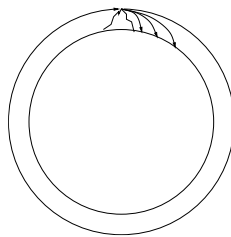
Unter der Annahme einer Kreisbewegung lässt sich die Bahngeschwindigkeit des Planeten anschaulich einfach herleiten, da die Zentralkraft die Gravitationskraft ist, d.h. es

$$\frac{m_p v_K^2}{r} = \gamma \frac{M_Z m_p}{r^2} \quad \rightarrow \quad v_K = \frac{\gamma M_Z}{r} \quad (2.3)$$

Diese Kreisbahngeschwindigkeit, auch als erste kosmische Geschwindigkeit bezeichnet, ist eine ausgezeichnete Geschwindigkeit, da sie die niedrigste Geschwindigkeit ist, mit der sich ein Körper auf einer Umlaufbahn um den Zentralkörper halten kann. Für einen Körper auf der Erde beträgt die erste kosmische Geschwindigkeit 7.9 km/s oder 28 440 km/h.

Auch eine Ellipsenbahn lässt sich mit Newtons Apfelanalogie zumindest plausibel machen. Stellen wir uns dazu vor, dass der Körper mit einer Horizontalgeschwindigkeit größer als der ersten kosmischen Geschwindigkeit geworfen wird. Dann ist die Fallstrecke während einer Zeiteinheit die gleiche wie bei der Kreisbahn, die dabei zurück gelegte horizontale Strecke jedoch größer, d.h. der Körper entfernt sich vom Zentralkörper, r wird also größer. Mit zunehmendem r wird auch die potentielle Energie des Körpers größer. Da die Bewegung nach dem Abwurf antriebslos erfolgt, muss daher die kinetische Energie des Körpers abnehmen und damit auch seine Geschwindigkeit. Ist die Gesamtenergie W_{ges} des Körpers beim Start

³Newton veröffentlichte diese Betrachtungen zuerst 1687 in den *Principia*. Dort gibt er auch eine anschauliche Erklärung: beim waagerechten Wurf hängt die Flugweite von der horizontalen Geschwindigkeit des Körpers ab; mit zunehmender Abwurfgeschwindigkeit wächst die Flugweite. Wird der Körper sehr schnell geworfen, so wird seine Flugweite so groß, dass die Erdkrümmung nicht mehr vernachlässigt werden kann. Eine Kreisbahn entsteht dann, wenn sich die Erde während einer Zeiteinheit des Fluges um genau das Stückchen unter dem Körper krümmt, um dass er in dieser Zeiteinheit in Richtung auf das Zentrum gefallen ist:



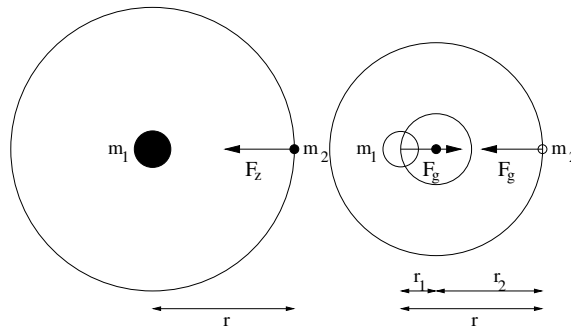


Abbildung 2.2: Bewegung um den gemeinsamen Schwerpunkt

größer als seine potentielle Energie $W_{p,\infty}$ im Unendlichen, so entweicht der Körper aus dem Schwerfeld des Zentralkörpers. Für den Fall $W_{\text{ges}} < W_{p,\text{infy}}$ wird so lange kinetische Energie in potentielle Energie des Körpers umgewandelt, wie seine Geschwindigkeit die Kreisbahngeschwindigkeit bei seinem momentanen Abstand übersteigt. Wird dieser Punkt überschritten, so fällt der Körper pro Zeiteinheit eine größere Strecke als sich die lokale Kreisbahn unter ihm wegkrümmt, d.h. der Körper bewegt sich unter Umwandlung von potentieller in kinetische Energie auf den Zentralkörper zu.

Mit dieser anschaulichen Betrachtung können wir zwar keine Ellipsenbahnen begründen (die Betrachtung zeigt nicht, dass die Bahnen geschlossen sind), aber der Flächensatz wird plausibel: große Abstände sind mit geringen Geschwindigkeiten verbunden, kleine Abstände mit großen.

Das dritte Kepler'sche Gesetz lässt sich ebenfalls elementar begründen. Auf einer Kreisbahn (als Spezialfall der Ellipse) gilt in jedem Punkt der Bahn (2.3). Ersetzt man die Geschwindigkeit v_K durch den Quotienten aus Umfang $2\pi r$ der Kreisbahn und Umlaufzeit T , so ergibt (2.3) eine Konstante:

$$\frac{r^3}{T^2} = \frac{\gamma M_Z}{(2\pi)^2}. \quad (2.4)$$

2.2 Physikalische Grundlagen

Planeten- und Satellitenbahnen sind durch die gegenseitige Anziehung zwischen den beteiligten Körpern bestimmt. Dabei wird z.B. eine Planetenbahn um ein Zentralgestirn durch die anderen Planeten gestört, eine Satellitenbahn kann durch den Einfluss des Mondes gestört werden.

2.2.1 Eingeschränktes Zweikörperproblem

Aber selbst bei Beschränkung auf zwei Körper müssen wir die Beschreibung genauer hinterfragen. Betrachten wir dazu zwei Körper gleicher Masse m , die sich im Schwerfeld des jeweils anderen Körpers bewegen. Dann ist es nicht möglich zu sagen, dass der eine Körper ruht und der andere sich im Schwerfeld des ersten auf einer Kreisbahn um diesen bewegt. Stattdessen bewegen sich beide Körper um den gemeinsamen Massenmittelpunkt.

Da bei Planeten- und Satellitenbewegungen der eine Körper wesentlich massereicher ist als der andere, liegt der gemeinsame Schwerpunkt sehr dicht an bzw. im massenreicheren Körper und die anschauliche Darstellung der Bewegung eines Körpers um den anderen ist gerechtfertigt.

Betrachten wir dazu zwei Massen m_1 und m_2 an den Orten \vec{r}_1 und \vec{r}_2 . Ihr Massenmittelpunkt befindet sich am Ort \vec{r}_{MM} , d.h. da kein Drehmoment wirkt muss gelten $m_1(\vec{r}_1 - \vec{r}_{\text{MM}}) + m_2(\vec{r}_2 - \vec{r}_{\text{MM}}) = 0$. Mit $\vec{r} = \vec{r}_1 - \vec{r}_2$ als dem die beiden Massen verbindenden Vektor

ergibt sich für die Anstände der Massen vom gemeinsamen Schwerpunkt

$$\vec{r}_1 - \vec{r}_{\text{MM}} = \frac{m_2}{m_1 + m_2} \vec{r} \quad \text{und} \quad \vec{r}_2 - \vec{r}_{\text{MM}} = -\frac{m_1}{m_1 + m_2} \vec{r}. \quad (2.5)$$

Mit \vec{F}_1 als der von m_2 auf m_1 ausgeübten Kraft und entsprechend \vec{F}_2 als der auf m_2 wirkenden Kraft ergibt sich

$$\vec{F}_1 = m_1 \ddot{\vec{r}}_1 = m_1 \ddot{\vec{r}}_{\text{MM}} + \frac{m_1 m_2}{m_1 + m_2} \ddot{\vec{r}} \quad \text{und} \quad \vec{F}_2 = m_2 \ddot{\vec{r}}_2 = m_2 \ddot{\vec{r}}_{\text{MM}} - \frac{m_1 m_2}{m_1 + m_2} \ddot{\vec{r}}. \quad (2.6)$$

Da die Kräfte entgegengesetzt gleich sind, $\vec{F}_1 = -\vec{F}_2$, ergibt sich daraus $m_1 \ddot{\vec{r}}_{\text{MM}} = -m_2 \ddot{\vec{r}}_{\text{MM}}$ und damit für allgemeine Massen

$$\ddot{\vec{r}}_{\text{MM}} = 0, \quad (2.7)$$

d.h. der Massenmittelpunkt des Zweikörpersystems wird nicht beschleunigt sondern befindet sich im Zustand der Ruhe oder der gleichförmigen Bewegung.

Damit reduziert sich (2.6) auf

$$\vec{F}_1 = -\vec{F}_2 = m_2 \ddot{\vec{r}}_2 = \frac{m_1 m_2}{m_1 + m_2} \ddot{\vec{r}}. \quad (2.8)$$

Einsetzen der Gravitationskraft (2.2) liefert

$$\ddot{\vec{r}} = -\frac{\gamma(m_1 + m_2)}{r^3} \vec{r} = -\frac{\mu}{r^2} \vec{e}_r. \quad (2.9)$$

Formal ist diese Bewegungsgleichung äquivalent zu der sich bei Verwendung der Gravitationskraft (2.2) ergebenden, allerdings wird die Masse des Zentralkörpers ersetzt durch die Summe μ der Massen der beiden Körper. Für einen Planeten und sein Zentralgestirn bzw. einen Satelliten und seinen Planeten ist die Gesamtmasse μ jedoch praktisch mit der des Zentralkörpers identisch,⁴ so dass die vereinfachende Vorstellung, dass der Satellit vom Planeten angezogen und auf seine Umlaufbahn gezwungen wird, für praktische Anwendungen ausreichend ist.

2.2.2 Satellitenbahnen und Erhaltungsgrößen

Die formale Herleitung von Planeten- bzw. Satellitenbahnen finden Sie in Lehrbüchern zur theoretischen Physik, z.B. [27, 31], oder in Büchern zur Bahn- bzw. Himmelsmechanik, z.B. [8, 10, 102, 104]. An dieser Stelle wird nur der generelle Weg der Herleitung kurz skizziert um den Zusammenhang zwischen Eigenschaften der Bahn und den Erhaltungsgrößen noch einmal zu rekapitulieren.

Energieerhaltung (1. Integral der Bewegung)

Aus der Bewegungsgleichung (2.9) lässt sich durch Multiplikation mit $2\vec{r}$ das Energieintegral herleiten. Aus $2\vec{r}\ddot{\vec{r}} = -2\mu/r^3 \vec{r}\ddot{\vec{r}}$ lässt sich durch Zusammenfassen des Produkts aus einer Größe und ihrer Ableitung zur Ableitung der quadrierten Größe⁵ bilden, d.h. wir erhalten $\dot{\vec{v}}^2 = -\mu/r^3 \dot{\vec{r}}^2$. Integration liefert das *Energieintegral*

$$v^2 - \frac{2\mu}{r} = h_c \quad (2.10)$$

mit h_c als einer Integrationskonstante, die mit den Anfangswerten v_0 und r_0 verknüpft ist gemäß $h_c = v_0^2 - 2\mu/r_0$ und der Gesamtenergie E_0 proportional ist: $E_0 = \frac{1}{2} m h_c$.

⁴Die Masse der Sonne ist um mehr als 5 Größenordnungen größer als die der Erde, vgl. Abschn. A.2, die Masse der Erde ist um ca. 21 Größenordnungen größer als die eines großen Erdfernerkundungssatelliten.

⁵Dies können Sie durch Anwendung der Kettenregel auf das Ergebnis leicht einsehen. Das Produkt $2\vec{r}\ddot{\vec{r}}$ lässt sich schreiben als $\dot{\vec{v}}^2$, da gilt

$$\frac{d\dot{\vec{v}}^2}{dt} = 2\vec{v}\dot{\vec{v}} = 2\vec{r}\ddot{\vec{r}}.$$

Momentintegral (Drehimpulserhaltung)

Das Momentintegral, und damit der Flächensatz, lässt sich durch vektoriell Multiplikation der Bewegungsgleichung (2.9) mit \vec{r} herleiten:

$$\vec{r} \times \ddot{\vec{r}} = -\frac{\mu}{r^3} \vec{r} \times \vec{r}. \quad (2.11)$$

Die rechte Seite der Gleichung verschwindet. Dann ist auch $\vec{r} \times \ddot{\vec{r}} = 0$, d.h. die Bewegung erfolgt in einer Ebene. Gleichung (2.11) lässt sich unter Berücksichtigung der Kettenregel schreiben als $d(\vec{r} \times \dot{\vec{v}})/dt = 0$. Integration liefert das *Momentintegral* (Drehimpulserhaltung)

$$\vec{r} \times \dot{\vec{v}} = \vec{\sigma}_c \quad (2.12)$$

mit $\vec{\sigma}_c$ als einer Integrationskonstante. Skalare Multiplikation mit \vec{r} liefert die *Gleichung der Bahnebene*

$$\vec{r} \vec{\sigma}_c = 0, \quad (2.13)$$

d.h. σ_c steht senkrecht auf der Bahnebene.

Flächensatz

Aus der Gleichung der Bahnebene (2.13) lässt sich der Flächensatz durch Übergang auf Polarkoordinaten herleiten. Fällt in kartesischen Koordinaten die Bahnebene in die $z = 0$ -Ebene, so ist $\vec{\sigma}_c = (0, 0, \sigma_c)$ und das Momentintegral (2.12) reduziert sich auf $x\dot{y} - y\dot{x} = 0$. Einsetzen von $(x, y) = r(\cos \varphi, \sin \varphi)$ liefert den *Flächensatz* $r^2 \dot{\varphi} = \sigma_c = \text{const.}$ In Polarkoordinaten ist das Flächenelement $dA = \frac{1}{2} r^2 d\varphi$ und wir erhalten

$$A = \frac{\sigma_c}{2} (t_2 - t_1), \quad (2.14)$$

d.h. der Ortsvektor überstreicht in gleichen Zeitintervallen $t_2 - t_1$ gleiche Flächen A .

Laplace-Integral

Vektorielle Multiplikation der Bewegungsgleichung (2.9) mit dem Momentintegral (2.12) liefert

$$\vec{\sigma}_c \times \ddot{\vec{r}} = -\frac{\mu}{r^3} (\vec{r} \times \dot{\vec{v}}) \times \vec{r} = -\frac{\mu}{r^3} [\dot{\vec{v}} (\vec{r} \cdot \vec{r}) - \vec{r} (\dot{\vec{v}} \cdot \vec{r})]. \quad (2.15)$$

Mit $\vec{r} \cdot \dot{\vec{r}} = r^2$ und $\vec{r} \cdot \dot{\vec{v}} + \dot{\vec{v}} \cdot \vec{r} = 2vr$ ergibt sich $\dot{\vec{v}} \cdot \vec{r} = rv$ und damit

$$\vec{\sigma}_c \times \ddot{\vec{r}} = -\frac{\mu}{r^3} (r^2 \dot{\vec{v}} - rv\vec{r}) = -\mu \frac{r\dot{\vec{v}} - v\vec{r}}{r^2} = -\mu \frac{d}{dt} \left(\frac{\vec{r}}{r} \right) \quad (2.16)$$

bzw.

$$\frac{d}{dt} (\vec{\sigma}_c \times \vec{r}) - \mu \frac{d}{dt} \left(\frac{\vec{r}}{r} \right) = 0. \quad (2.17)$$

Integration liefert das *Laplace-Integral*

$$\vec{\sigma}_c + \dot{\vec{v}} + \mu \frac{\vec{r}}{r} = \vec{\lambda}_c \quad (2.18)$$

mit $\vec{\lambda}_c$ als Integrationskonstante. Skalare Multiplikation von (2.18) mit $\vec{\sigma}_c$ liefert $\vec{\sigma}_c \cdot \vec{\lambda}_c = 0$, d.h. der *Laplace-Vektor* $\vec{\lambda}_c$ liegt in der Bahnebene.

Da die Bewegungsgleichung als vektorielle DGL 2^{ter} Ordnung sechs Integrationskonstanten benötigt, wir aber mit den bisher bestimmten h_c , $\vec{\sigma}_c$ und $\vec{\lambda}_c$ sieben haben, können letztere nicht linear unabhängig sein. Die Beziehung zwischen ihnen lässt sich durch Quadrieren des Laplace-Integral bestimmen zu

$$\lambda_c^2 = \mu^2 + h_c \sigma_c^2. \quad (2.19)$$

	Kreis	Ellipse	Parabel	Hyperbel
Exzentrizität ε	0	$0 < \varepsilon < 1$	1	> 1
Bahnparameter p	a	$a(1 - \varepsilon^2)$	p	$a(\varepsilon^2 - 1)$
Große Halbachse a	a	a	–	a
Kleine Halbachse b	a	$a\sqrt{1 - \varepsilon^2}$	–	$a\sqrt{\varepsilon^2 - 1}$
Perigäum r_P	a	$a(1 - \varepsilon)$	$\frac{1}{2}p = q$	$a(\varepsilon - 1)$
Apogäum r_A	a	$a(1 + \varepsilon)$	∞	∞

Tabelle 2.1: Kegelschnitte

Bahngleichung

Die Bahngleichung lässt sich durch skalare Multiplikation des Laplace-Integrals (2.18) mit \vec{r} bestimmen:

$$\vec{r} \cdot (\vec{\sigma}_c \times \vec{v}) + \mu \frac{\vec{r} \cdot \vec{r}}{r} = -\vec{\lambda}_c \cdot \vec{r}. \quad (2.20)$$

Anwendung des Momentintegrals (2.12) und zyklisches Vertauschen der Vektoren im Spatprodukt liefert $-\sigma_c^2 + \mu r = -\lambda_c r \cos \tau$ mit τ als dem Winkel zwischen dem Ortsvektor \vec{r} und dem Laplace-Vektor $\vec{\lambda}_c$. Auflösen nach r gibt die *Gleichung der Satellitenbahn*

$$r = \frac{p}{1 + \varepsilon \cos \tau} \quad (2.21)$$

mit dem *Bahnparameter*

$$p = \frac{\sigma_c^2}{\mu} > 0 \quad (2.22)$$

und der *Exzentrizität*

$$\varepsilon = \frac{\lambda}{\mu} > 0. \quad (2.23)$$

Kegelschnitte

Die Bahngleichung (2.21) beschreibt *Kegelschnitte*, wobei die Exzentrizität (2.23) die Art des Kegelschnitts gibt (vgl. auch Tabelle 2.1):

$0 \leq \varepsilon < 1$ der Nenner in der Bahngleichung (2.21) ist immer von Null verschieden, so dass sich eine geschlossene, im endlichen verlaufende Ellipsenbahn ergibt.

$\varepsilon = 1$ der Nenner der Bahngleichung (2.21) verschwindet für $\tau = \pi$, d.h. die Bahn hat einen Punkt im Unendlichen (Parabel).

$\varepsilon > 1$ der Nenner der Bahngleichung (2.21) verschwindet für zwei Werte, d.h. die Bahn hat zwei Punkte im Unendlichen (Hyperbel).

Die realisierte Bahn hängt von den Anfangsbedingungen ab. Einsetzen der Beziehung (2.19) in die Definition (2.23) liefert

$$\varepsilon = \sqrt{1 + h_c \frac{\sigma_c^2}{\mu^2}} \quad (2.24)$$

mit der Energiekonstanten h_c gemäß (2.10) bzw. bei Verwendung der Anfangsbedingungen $h_c = v_0^2 - 2\mu/r_0$.

Bei der Ellipsenbahn ist $\varepsilon < 1$ und damit $h_c < 0$, d.h. die Bahn ist eine gebundene. Die *elliptische Anfangsgeschwindigkeit* wird damit zu $v_0^2 < 2\mu/r_0$. Entsprechend ergibt sich mit $\varepsilon = 1$ und damit $h_c = 0$ die *parabolische Anfangsgeschwindigkeit* zu $v_0^2 = 2\mu/r_0$. Eine *hyperbolische Anfangsgeschwindigkeit* ergibt sich für $\varepsilon > 1$ und damit $h_c = -\mu^2/\sigma^2$ zu $v_0^2 = 2\mu/r_0 - \mu^2/\sigma_c^2$.

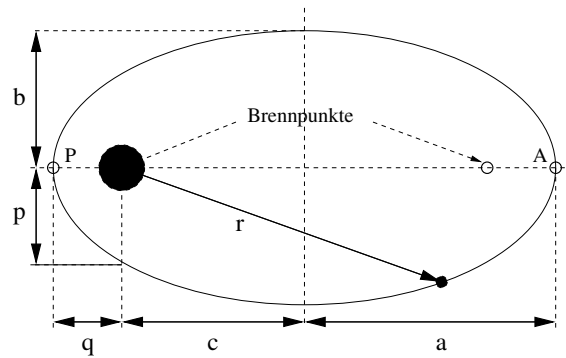


Abbildung 2.3: Kenngrößen einer elliptischen Bahn

Ein Spezialfall der Ellipsenbahn ist die Kreisbahn mit $\varepsilon = 0$ und damit $\sigma_c^2 = p\mu$ mit $p = r_0$ und $v_0 = \text{const.}$ Die *Kreisbahngeschwindigkeit* oder *erste kosmische Geschwindigkeit* ist damit

$$v_k^2 = \frac{\mu}{r}. \quad (2.25)$$

Sie ist gerade ausreichend damit ein Satellit um einen sphärischen Himmelskörper entlang eines Großkreises fliegen kann.⁶ Für die Erde ist $v_k = 7.9$ km/s.

2.2.3 Ellipsenbahnen

Planeten und Satelliten bewegen sich in der Regel auf Ellipsenbahnen oder deren Spezialfall, den Kreisbahnen. Eine Ellipse wird durch ihre große und kleine Halbachse, a und b , beschrieben, vgl. Abb. 2.3. Im Spezialfall des Kreises ist $a = b = r$. Die Punkte A und P markieren das Apo- und das Perizentrum, d.h. den Punkt der Bahn am weitesten vom Zentralkörper entfernt bzw. diesem am nächsten. Bei der Erde werden diese als *Apogäum* und *Perigäum* bezeichnet, bei der Sonne als *Aphel* und *Perihel*. Ihre Abstände vom Zentralkörper sind durch die Exzentrizität ε und den Bahnparameter p bestimmt:

$$r_P = \frac{p}{1 + \varepsilon} \quad \text{und} \quad r_A = \frac{p}{1 - \varepsilon}. \quad (2.26)$$

Die Achse zwischen Apo- und Perizentrum wird als Apsidenlinie bezeichnet.

Energetik

Die Geschwindigkeit in verschiedenen Punkten der Bahn lässt sich aus dem Energiesatz (2.10) bestimmen. Die Integrationskonstante h_c kann in irgend einem Punkt der Bahn bestimmt werden, also auch im Perigäum. Dort stehen Orts- und Geschwindigkeitsvektor senkrecht auf einander und es gilt

$$|\vec{v}_P| = r_P \left(\frac{d\eta}{dt} \right)_P \quad (2.27)$$

mit der wahren Anomalie η als dem Winkel zwischen dem Ortsvektor und der Apsidenlinie. Gemäß Flächensatz (2.14) ist $|\sigma_c| = r^2\dot{\eta}$ und wir erhalten für die Energiekonstante im Perigäum

$$h_c = \frac{1}{2} \frac{\sigma_c^2}{r_P^2} - \frac{\mu}{r_P} = \frac{\mu}{2r_P^2} (p - 2r_P), \quad (2.28)$$

⁶Ein Großkreis ist ein Kreis auf einer Kugel, bei dem der Kugelmittelpunkt in der Kreisebene liegt. Die Äquatorebene der Erde wäre ein Beispiel, ebenso alle Kreisbahnen entlang eines festen Längengrades. Entsprechend ist die Bahn entlang des 50^{ten} Breitengrades keine Großkreisbahn, da die Bahnebene nicht durch den Erdmittelpunkt geht – als antriebslose Satellitenbahn wäre eine derartige Bahn nicht realisierbar.

v [km/s]	v/v_k	v/v_p	a [km]	ε	D [km]
2	0.25	0.18	3 290	0.94	200
4	0.51	0.36	3 650	0.74	920
6	0.76	0.54	4 480	0.42	2 580
7.91	1.00	0.707	6 378	0.00	6 378
8	1.01	0.71	6 500	0.02	6 620
9	1.14	0.80	9 010	0.34	11 640
10	1.26	0.89	15 740	0.60	25 100
11.0	1.391	0.984	99 500	0.94	192 600
11.10	1.403	0.9924	221 000	0.97	435 000
11.185	1.414	1.000	∞	1.00	∞

Tabelle 2.2: Flugbahnen bei verschiedenen Abschussgeschwindigkeiten im Perigäum

wobei im letzten Schritt (2.22) verwendet wurde.

Setzt man die entsprechenden Werte aus Tabelle 2.1 ein, so erhalten wir für die Gesamtenergie der Ellipsenbahn

$$h = -\frac{\mu}{2a}, \quad (2.29)$$

für die Parabel $h = 0$ und für die Hyperbel $h = \mu/(2a)$, d.h. die Gesamtenergie hängt nur von der großen Halbachse a und der Masse μ des Zentralgestirns (plus Satellit/Planet). Die Exzentrizität der Bahn hat keinen Einfluss auf die Energie, d.h. ein Satellit auf einer sehr exzentrischen Bahn hat die gleiche Gesamtenergie wie einer auf einer die Ellipse umschreibenden Kreisbahn.

Für die Geschwindigkeit entlang der Bahn ergibt sich nach Einsetzen der Gesamtenergie in die Energiegleichung (2.10) die *Binet'sche Gleichung*

$$v = \sqrt{\mu \left(\frac{2}{r} - \frac{1}{a} \right)}. \quad (2.30)$$

Beim Kreis ist $a = r$ und es ergibt sich wieder die erste kosmische Geschwindigkeit

$$v_1 = \sqrt{\frac{\mu}{r}} = v_k. \quad (2.31)$$

Für eine Parabel ($a = \infty$) ergibt sich die *zweite kosmische Geschwindigkeit* zu

$$v_2 = \sqrt{\frac{2\mu}{r}} = \sqrt{2} v_k. \quad (2.32)$$

Diese Geschwindigkeit ist notwendig, um aus dem Schwerfeld des Zentralkörpers zu entweichen. Für die Erde beträgt diese *Fluchtgeschwindigkeit* 11.2 km/s, die Fluchtgeschwindigkeit aus dem Sonnensystem beträgt 16.7 km/s.

Mit Hilfe der Binet'schen Gleichung (2.30) ergibt sich für die Geschwindigkeiten in Apo- und Perizentrum der Bahn

$$v_A = \sqrt{\frac{\mu}{a} \frac{1-\varepsilon}{1+\varepsilon}} \quad \text{und} \quad v_P = \sqrt{\frac{\mu}{a} \frac{1+\varepsilon}{1-\varepsilon}}. \quad (2.33)$$

Die Geschwindigkeiten im Peri- und Apozentrum hängen also nicht nur von der großen Halbachse der Ellipse sondern auch von der Exzentrizität ab: je größer die Exzentrizität, um so größer die Unterschiede in den Geschwindigkeiten. Ein Molniya-Satellit z.B. fliegt auf einer elliptischen Bahn mit $\varepsilon = 0.75$ und hat damit im Perigäum eine um einen Faktor 7 höhere Geschwindigkeit als im Apogäum.

Tabelle 2.2 gibt einen Überblick über die sich in Abhängigkeit von der Abschussgeschwindigkeit im Perigäum ergebenden Bahnen (das entspricht der Abwurfgeschwindigkeit auf Newton's Berg). Die Geschwindigkeiten sind auch in Einheiten der Kreisbahngeschwindigkeit v_k

a [km]	a/R	H [km]	T [min]	n [°/min]	v_k [km/s]
6 378	1.0000	7	84.49	4.261	7.91
6 400	1.0034	29	84.92	4.239	7.89
6 500	1.0191	129	86.92	4.142	7.83
6 600	1.0348	229	88.94	4.048	7.77
6 700	1.0505	329	90.97	3.958	7.71
6 800	1.0661	429	93.01	3.871	7.66
6 900	1.0818	529	95.07	3.787	7.60
7 000	1.0975	629	97.15	3.706	7.55
7 200	1.1288	829	101.34	3.552	7.44
7 400	1.1602	1 029	105.58	3.410	7.34
7 600	1.1916	1 229	109.89	3.276	7.24
7 800	1.2229	1 429	114.26	3.151	7.15
8 000	1.2542	1 629	118.69	3.033	7.06
9 000	1.4111	2 629	141.62	2.542	6.65
10 000	1.5678	3 629	165.64	2.174	6.31
15 000	2.3518	8 630	5 ^h 4 ^m 43 ^s	1.814	5.15
20 000	3.1357	13 630	7 ^h 49 ^m 8 ^s	0.7674	4.47
42 160	6.611	35 790	23 ^h 56 ^m 4 ^s	0.2507	3.07
384 400	60.266	–	27 ^d 7 ^h 43 ^m	0.00915	1.02

Tabelle 2.3: Bahnparameter als Funktion der großen Halbachse

und der parabolischen Geschwindigkeit v_p gegeben. Außerdem sind angegeben die Bahnparameter große Halbachse a und Exzentrizität ε sowie die maximale Entfernung D vom Startpunkt. Die horizontalen Linien markieren jeweils den Übergang zur Kreisbahn bzw. unten die Annäherung an die parabolische Bahn.

Tabelle 2.3 gibt die Bahnparameter als Funktion der großen Halbachse. Die unteren beiden Bahnen entsprechen der geostationären Bahn und der Mondbahn, die Bahnen zwischen den beiden horizontalen Linien sind typisch für Erdfernerkundungssatelliten. Angegeben sind neben der großen Halbachse a die große Halbachse in Einheiten des Erdradius, a/R , die mittlere Flughöhe H über dem Erdboden, die Umlaufzeit T , die mittlere Winkelgeschwindigkeit sowie die Kreisbahngeschwindigkeit. Auf diese Tabelle werden wir im Laufe der weiteren Diskussion noch häufiger zurück greifen.

Lage der Bahn im Raum

Bisher haben wir die Ellipsenbahn mit ihren Parametern große und kleine Halbachse, a und b , Exzentrizität ε und Bahnparameter p zur Beschreibung der Bahn eingeführt. Um die Bewegung eines Satelliten zu beschreiben, benötigen wir zusätzliche Parameter, insbesondere die Lage der Bahn im Raum.

Der wichtigste Bahnparameter ist die *Inklination* i , d.h. die Neigung der Bahnebene gegenüber der Äquatorebene. Damit können z.B. wir zwischen äquatorialen Bahnen ($i = 0$) und polaren Bahnen ($i = 90^\circ$) unterscheiden – während erstere für Wettersatelliten oder zum Studium der tropischen Ozeane sicherlich gut geeignet sind, eignen sich letztere eher zum Studium des (ant)arktischen Ozonlochs. Die Untergrenze des Werts der Inklination ist, falls der Satellit nicht nach Einbringen in seine Bahn manövriert wird, durch die geographische Breite seines Startorts bestimmt: von Kiruna lässt sich kein Satellit in eine Bahn mit einer Inklination von weniger als 68° einschieseln, der Einschuss in eine polare Bahn dagegen ist möglich. Daher versucht man, Raketenstartplätze in möglichst niedriger geographischer Breite zu betreiben (Kourou, der Startplatz der ESA, liegt bei N5 W52).

Der zweite Parameter zur Beschreibung der Lage der Bahnebene ist die *Rektaszension* Ω . Sie gibt die Lage der Knotenlinie der beiden Ebene im Bezug auf die Richtung des Frühjahrspunkts an, wobei der aufsteigende Punkt dieser Linie in der Richtung zeigt, in

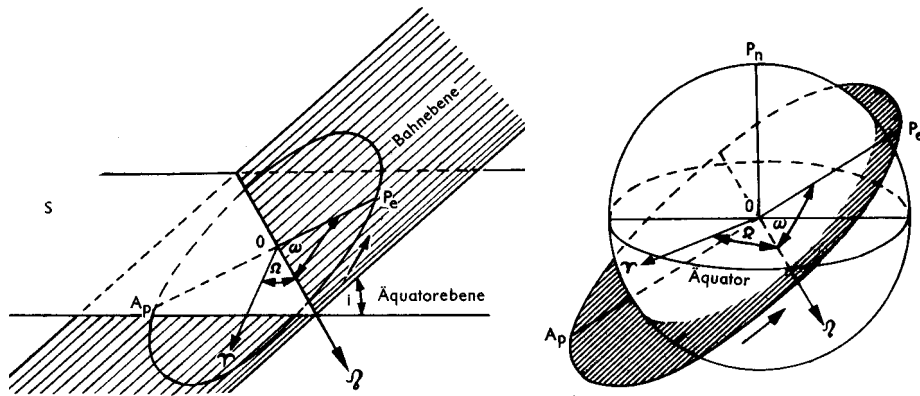


Abbildung 2.4: Festlegung der Bahnebene und -ellipse eines Satelliten gegenüber der Äquatorebene und dem Frühlingspunkt [10]

der der Satellit den Äquator von Nord nach Süd überquert. Oder pragmatischer: die Rektaszension legt fest, zu welcher Tageszeit ein Satellit den Äquator bei einer bestimmten geographischen Länge südwärts fliegend überquert.

Jetzt sind die Bahn und die Lage der Bahnebene festgelegt, zusätzlich muss noch die Lage der Bahn in der Bahnebene bestimmt werden durch den Winkel ω zwischen der Knotenlinie und der Apsidenlinie. Zur Festlegung der genauen Bewegung benötigen wir einen sechsten Parameter, der angibt, wo auf der Bahn sich der Satellit zu einem bestimmten Zeitpunkt (Epoche) befunden hat. Dafür eignen sich die Zeiten von Perigäums- oder Knotendurchgängen. Die Parameter sind in Abb. 2.4 zusammen gefasst.

Einige Spezialfälle für Bahnen sind:

- *äquatoriale Bahnen*, d.h. die Inklination verschwindet, $i = 0$,
- *polare Bahnen* mit einer Inklination nahe 90° ,
- *stark exzentrische Bahnen* mit $\varepsilon > 0.1$ werden für einige Kommunikationssatelliten verwendet,
- *geostationäre Bahnen* mit $\varepsilon = 0$, $i = 0$ und $T = 24$ h,
- *sonnensynchrone Bahnen* in denen die Bahnebene des Satelliten so ausgerichtet ist, dass sie stets in einem festen Winkel zur Erde-Sonne-Achse steht.

Abbildung 2.5 fasst schematisch einige spezielle Bahnen zusammen: (A) gibt eine polare Kreisbahn in niedriger Höhe, wie sie zur Erderkundung und -überwachung und zur militärischen Aufklärung verwendet wird. (B) zeigt eine äquatoriale Kreisbahn in großer Höhe, z.B. eine geostationäre Bahn für Wetter- und Kommunikationssatelliten. (C) gibt eine stark exzentrische Bahn mit relativ großer Inklination wie sie z.B. von den Molniya-Kommunikationssatelliten genutzt wird.

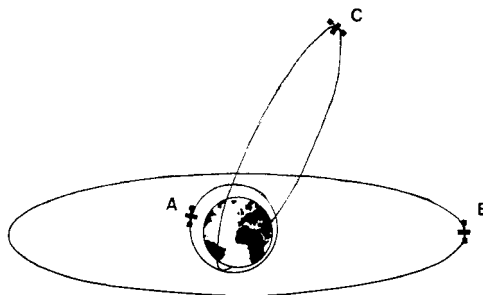


Abbildung 2.5: Beispiele für Satellitenbahnen [132]

2.3 Typische Bahnen von erdgebundenen Satelliten

Erdsatelliten bewegen sich auf Kreis- oder Ellipsenbahnen. Welcher der in Tabelle 2.3 gegebenen Bahnen die Bahn eines bestimmten Satelliten ähnelt, hängt von den an den Satelliten gestellten Anforderungen ab. Eine grobe Unterteilung können wir nach Satellitentypen vornehmen.

2.3.1 Typen von Erdsatelliten

Erdsatelliten lassen sich nach ihrer (hauptsächlichen) Aufgabe klassifizieren. Die wichtigsten Typen sind:

- Erderkundungssatelliten, in ihrer militärischen Variante als Aufklärungssatelliten bezeichnet;
- Wettersatelliten, die nicht nur, wie durch ihren Namen nahe gelegt, zur Wetterbeobachtung eingesetzt werden, sondern auch Instrumente für allgemeine atmosphärische oder umweltbezogene Untersuchungen an Bord haben, wie z.B. das *Advanced Very High Resolution Radiometer* AVHRR auf den verschiedenen NOAA Satelliten, mit denen Schnebedeckungen ebenso wie Ölfilme auf Gewässern vermessen werden können oder der Space Environment Monitor SEM zur Messung von Plasmen und energiereichen Teilchen.
- Kommunikationssatelliten, die als Relaisstationen zwischen verschiedenen Bodenstationen, zwischen einem Raumflugkörper und einer Bodenstation oder zwischen verschiedenen Raumflugkörpern eingesetzt werden können;
- Navigationssatelliten, die einem Beobachter auf der Erde oder im erdnahen Weltraum eine sehr genaue Positionsbestimmung gestatten;
- Forschungssatelliten, die vielfältige Aufgaben haben können und, da wir Erderkundungssatelliten bereits separat aufgelistet haben, nicht auf die Erde sondern nach draußen gerichtet sind.

2.3.2 Erderkundungs- und Aufklärungssatelliten

Erderkundungssatelliten werden auf vielfältige Weise eingesetzt. Waren sie ursprünglich eher experimentelle Einrichtungen, die hauptsächlich zur Kartographie verwendet wurden, so haben sie sich heute zu hochgradig komplexen Systemen von Instrumenten entwickelt, die zur Untersuchung von so verschiedenen Phänomenen wie Schadstoffeinträge in Gewässer, Abholzung des tropischen Regenwaldes, Änderungen im Vegetationsmuster (Desertifikation), Beurteilung von Waldschäden, Meeresströmungen, Planktongehalt im Meerwasser, Ausdehnung des Ozonlochs und vielem mehr eingesetzt werden. Der militärische Zwilling Bruder des Erderkundungssatelliten ist der Aufklärungssatellit.

Die meisten Erderkundungs- und Aufklärungssatelliten fliegen auf nahezu kreisförmigen Bahnen mit hoher Inklination. Letztere ist nicht nur durch den Startort festgelegt, sondern gibt gleichzeitig die höchste geographische Breite bis zu der der Satellit gelangt. Eine hohe Inklination ist daher notwendig, um die Beobachtungen nicht nur auf einen schmalen Breitenstreifen um den Äquator zu beschränken. Die Bahnen sind kreisförmig, um auf allen Abschnitten des Orbits möglichst nahe an der Erdoberfläche zu sein und damit ein gutes Bodenaufklärungsvermögen zu erreichen. Militärische Späher (Big Bird, KH-11) fliegen in Höhen zwischen 200 und 300 km,⁷ zivile Erderkundungssatelliten (z.B. SPOT, LandSat, SeaSat, Nimbus) zwischen 600 und 1000 km, vgl. Tabelle. 1.1. Auf Grund der geringen Flughöhen liegen die Umlaufzeiten typischerweise im Bereich von 90 bis 100 min, vgl. Tabelle 2.3.

Die Bedeutung der Bahnparameter für einen speziellen Satelliten kann man auch aus der Projektion der Satellitenbahn auf die Erdoberfläche, der *Bodenspur*, erschließen. Abbildung 2.6 zeigt als Beispiel die Bodenspur eines LandSat-Satelliten. Die Bahn ist nahezu polar,

⁷Größere Flughöhen verringern das Bodenaufklärungsvermögen, geringere Flughöhen sind nicht sinnvoll, da dann die Reibung in der Restatmosphäre die Lebensdauer des Satelliten drastisch verringert bzw. eine sehr häufige Bahnkorrektur und damit das Mitführen großer Treibstoffmengen erforderlich macht.

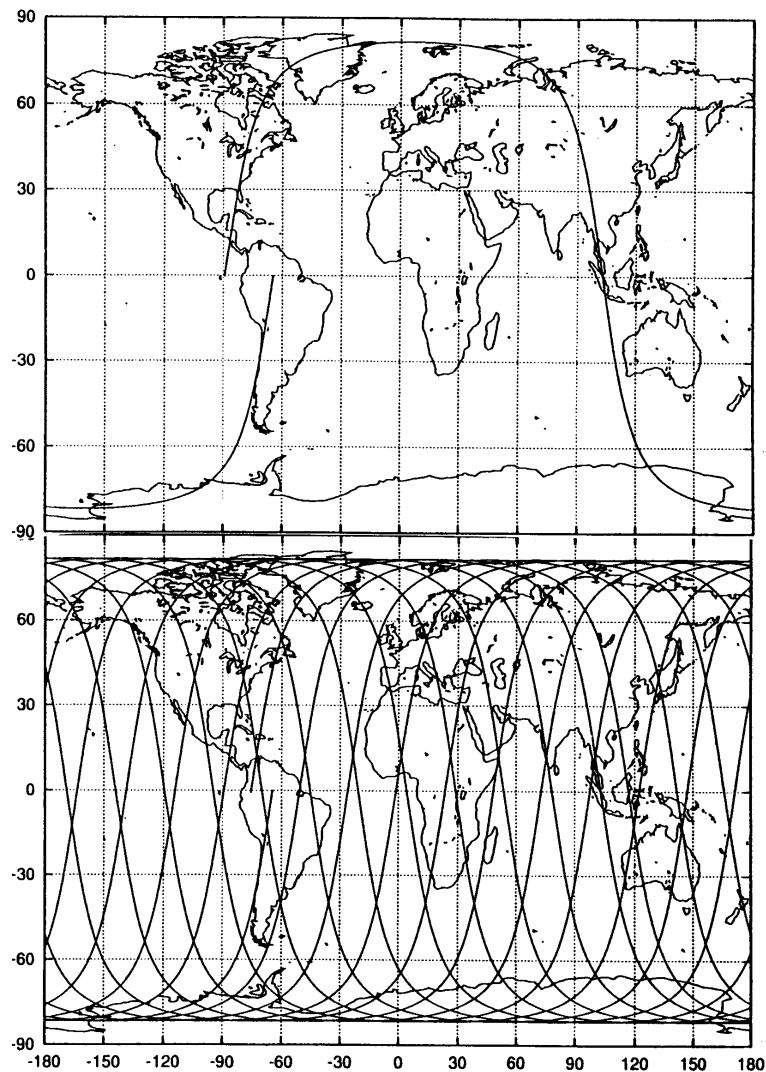


Abbildung 2.6: Bodenspur von LandSat für einen Umlauf (oben) und alle Umläufe eines Tages (unten) [19]

so dass ein großer Breitenbereich überdeckt wird (oberes Teilbild). Da sich die Erde unter der Satellitenbahn weg dreht, werden im Laufe des Tages verschiedene Bereiche der Erde überflogen, so dass sich bereits innerhalb eines Tages eine Überdeckung weiter Bereiche der Erde ergibt (unteres Teilbild). Da der Blickwinkel der Instrumente auf diesem Satelliten zu gering ist, um den gesamten Bereich zwischen zwei benachbarten Bodenspuren beobachten zu können, ist das Überdeckungsmuster der folgenden Tage jeweils versetzt – daher schließt sich die Bodenspur im unteren Teilbild über Südamerika nicht. Für eine komplette Überdeckung sind in diesem Beispiel 248 Umläufe erforderlich, entsprechend einer Dauer von 17 Tagen.

Die Verwendung einer sonnensynchronen Bahn liefert zusätzliche Informationen, da auf Grund der konstanten Beleuchtungsverhältnisse aus der Länge des Schattens eines Objekts auf seine Höhe zurück geschlossen werden kann, ebenso wie auf seine Struktur (siehe auch Abb. 2.12). Die Verlängerung des Schattens zwischen aufeinander folgenden Umläufen kann damit ein Indiz für bauliche Veränderungen geben. Sonnensynchrone Orbits waren aus diesen Gründen in der Frühzeit der Erdsatelliten im wesentlichen für militärische Aufklärung von Interesse. Im Zuge der immer detaillierteren Untersuchungen des spektralen Reflektionsvermögens z.B. zur Bestimmung von Pflanzenwachstum und -gesundheit sind sonnensyn-

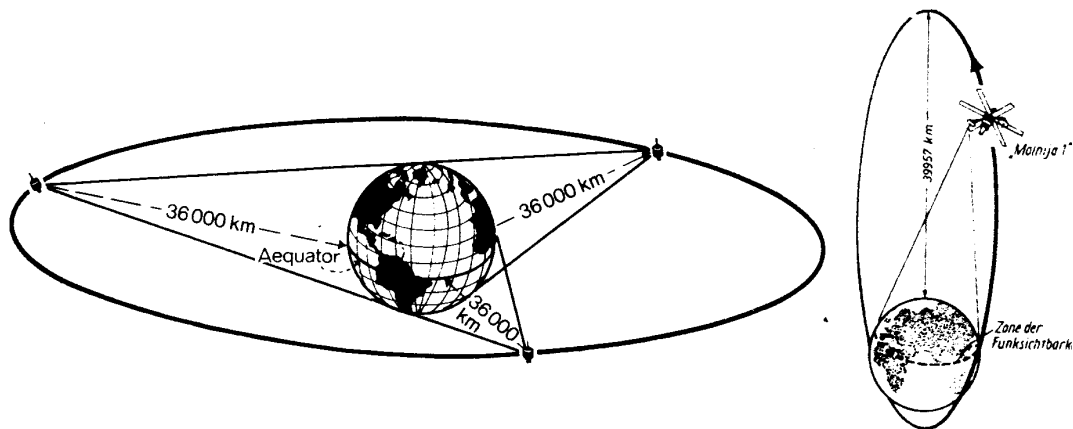


Abbildung 2.7: Kommunikationssatelliten: Überdeckung der Erdoberfläche mit Hilfe von drei geostationären Satelliten (links) und Molniya-Orbit (rechts)

chrone Bahnen ebenfalls von Interesse, da dann das einfallende Spektrum nur durch die atmosphärischen Parameter variiert wird und nicht nur zusätzlich durch Sonnenstand und damit Auftreffwinkel des Lichts auf die Pflanze.

2.3.3 Wetter- und Kommunikationssatelliten

Wetter- und Kommunikationssatelliten haben sehr ähnliche Orbits und können daher gemeinsam behandelt werden. Beide befinden sich häufig auf geostationären Orbits in einer Höhen von ca. 36 000 km über dem Erdboden (vgl. Tabelle 2.3). Ihre Umlaufzeit beträgt 24 Stunden, d.h. der Satellit scheint über einem Punkt fest zu stehen. Das ist für meteorologische Anwendungen interessant, da der Satellit auf diese Weise die lokalen Veränderungen beobachtet. Und für die Kommunikation ist es günstig, da der Satellit vom Boden gesehen stets am gleichen Punkt am Himmel steht und daher Antennen nicht nachgeführt werden müssen. Der Satellit ist für die Kommunikation daher genauso wie eine Bodenstation eine ortsfeste Relais-Station. Für die Kommunikation zwischen Kontinenten befinden sich die Satelliten über den Ozeanen, für die Kommunikation innerhalb eines Kontinents über diesem. Drei äquidistante Satelliten sind ausreichend, um eine nahezu vollständige Überdeckung der Erdoberfläche zu gewährleisten, vgl. linkes Teilbild in Abb. 2.7.

Das Problem geostationärer Satelliten ist die Überfüllung des Orbits: allein 1991 wurden 26 Satelliten in geostationäre Orbits gestartet, die meisten davon zivile Kommunikationssatelliten. Bei der Überfüllung ist der mechanische Gesichtspunkt unerheblich, vom Standpunkt der Kommunikation dagegen ist ein Mindestabstand von 2° notwendig, d.h. es sind insgesamt 180 Plätze zu vergeben, verglichen mit 26 Starts alleine in einem Jahr. Lösungsvorschläge dieser technischen Variante eines Überbevölkerungsproblems werden in [95] diskutiert. Eine Lösungsmöglichkeit besteht in der Verwendung von Orbits ähnlich denen der Molniya-Satelliten oder die Nutzung von pseudo-geostationären Orbits mit einer Inklination von einigen Grad, die aus Sicht eines Beobachters auf der Erde eine Oszillation des Satelliten entlang eines Längengrades über den durch die Inklination definierten Bereich bewirken würde.

Geostationäre Bahnen sind per Definition äquatoriale Bahnen, da der Erdmittelpunkt als Massenmittelpunkt in der Bahnebene liegen muss. Mit von Null verschiedenen Inklinationen lassen sich zwar auch Orbits erreichen, bei denen die Umlaufzeit 24 Stunden beträgt, jedoch oszilliert der Satellit dann entlang eines Längengrades um den Äquator, wie oben diskutiert. Andererseits hat aber jeder Satellit, der von einem Startplatz in einem gewissen Abstand vom Äquator gestartet wird, mindestens die Inklination, die der geographischen Breite seines Startorts entspricht. In höherer Breite gestartete Satelliten können daher nur über energieaufwendige Manöver in äquatoriale Bahnen gebracht werden – was das Mitführen

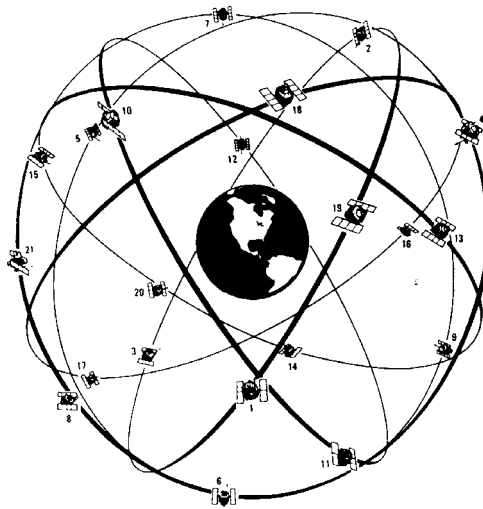


Abbildung 2.8: Konfiguration der GPS-Satelliten [63]

von Treibstoff und entsprechend leistungsfähigen Antriebsaggregaten erfordert und damit zu Lasten der transportierbaren Nutzlast geht.

Auf Grund ihrer geographischen Lage konnte die damalige Sowjetunion (FSU, Former Soviet Union) mit Startplätzen oberhalb von 30° N kein System geostationärer Kommunikationssatelliten aufbauen. Stattdessen wurde ein System von Kommunikationssatelliten in hochgradig elliptischen Bahnen betrieben. Diese Molniya-Satelliten (Molniya = Blitz) haben ein Perigäum bei ca. 400 km, ein Apogäum bei ca. 40 000 km und eine Inklination von 65°. Damit ergibt sich eine Umlaufzeit von ca. 12 Stunden. Da das Apogäum über dem Nordpol gewählt wurde, vgl. rechter Teil von Abb. 2.7, halten sich die Satelliten während ca. 2/3 ihrer Umlaufzeit im Empfangsbereich einer Bodenstation der FSU auf und bewegen sich zugleich nur relativ langsam entlang ihrer Bahn. Eine vollständige Überdeckung wird durch die Verwendung jeweils eines Satellitenpaares erreicht, wobei der zweite Satellit in einem um 90° versetzten Orbit fliegt.

2.3.4 Navigationssatelliten

Navigationssatelliten sind primär ebenfalls militärische Systeme, die es Einheiten bzw. einzelnen Soldaten ermöglichen sollen, ihren Standort genau zu bestimmen, und die natürlich auch für die Lenkung automatischer Waffensysteme verwendet werden. Das System hat aber auch zivile Anwendungen, insbesondere die Funkortung für die Luft- und Seefahrt. Auch Erdfernerkundungssatelliten benutzen GPS zur genauen Positionsbestimmung, die bodengebundenen Landvermesser ebenfalls.

Gerade im Bereich der Navigation liegt daher eine starke Kopplung zwischen ziviler und militärischer Nutzung vor. Das bekannteste System, NAVSTAR bzw. GPS (Global Positioning System), wurde ursprünglich in unterschiedlichen Modi für die militärische und die zivile Nutzung betrieben: die militärische Variante mit einer Genauigkeit der Ortsbestimmung von ca. 1 m, die zivile mit einer Genauigkeit im Bereich von einigen 10 Metern. Dieses künstliche Verrauschen der zivil genutzten Signale wurde Mitte der Neunziger Jahre abgeschaltet, so dass jetzt auch zivilen Nutzern eine Genauigkeit im Bereich von Metern zur Verfügung steht.⁸

⁸Bei kleinen kommerziellen GPS-Empfängern beträgt die Genauigkeit in den horizontalen Koordinaten wenige Meter, kann jedoch bei schlechter Satellitenüberdeckung geringer sein. Ein Problem ist die Genauigkeit der vertikalen Koordinate – diese kann auch mal um 100 m versetzt sein. Das bedeutet z.B. auf einem Leuchtturm stehend eine Höhenkoordinate von -32 m, d.h. 32 m unter NormalNull, zu erhalten. Einfache GPS-Systeme sind nicht in der Lage, schnelle Änderungen in Kurs und Geschwindigkeit korrekt zu erkennen, da ein Teil der Auswertung auf Dead Reckoning basiert.

Das Verfahren der Ortsbestimmung ist ähnlich der zur Bestimmung der Höhe eines Gebäudes oder Berges verwendeten Triangulation. Um einen Ort im Raum zu bestimmen, ist es ausreichend, drei Orte mit genau bekannter Lage anzupeilen und die dazu gehörigen Richtungswinkel zu bestimmen.⁹ Auf die GPS-Satelliten übertragen bedeutet dies, dass mindestens drei von ihnen im Blickfeld eines jeden Punktes auf der Erdoberfläche sein müssen. Praktisch wird das GPS-System durch 21 Satelliten auf sechs nahezu kreisförmigen Orbits mit einer Umlaufhöhe von ca. 20 000 km und einer Periode von 12 Stunden realisiert, vgl. Abb. 2.8. Die Idee des Systems basiert darauf, dass von jedem Punkt der Erde mindestens vier dieser Satelliten angepeilt werden können.¹⁰ Der vierte Satellit ist deshalb wichtig, weil das System im Gegensatz zu bodengebundenen Peilstationen eine zeitliche Komponente hat: die Satelliten sind nicht ortsfest sondern bewegen sich relativ zum Erdboden. Vereinfacht gesagt soll der vierte Satellit eine zeitliche Synchronisation ermöglichen: aus seinem Signal wird die genaue Zeit und damit der genaue Ort der drei für die Ortsbestimmung verwendeten Satelliten ermittelt.¹¹ Eine ausführliche Darstellung des GPS gibt [46].

2.3.5 Forschungssatelliten

Forschungssatelliten bilden eine inhomogene Gruppe mit ihrer wissenschaftlichen Fragestellung angepassten Bahnen. Einige Beispiele für erdgebundene Forschungssatelliten sind:

- für die Erforschung des interplanetaren Raums von erdgebundenen Satelliten aus werden stark exzentrische Orbits verwendet, z.B. IMP (Interplanetary Monitoring Platform). Dadurch befindet sich der Satellit lange Zeit außerhalb der Erdmagnetosphäre; das Bahnkonzept ist ähnlich dem der Molniya-Satelliten. Eine Alternative ist eine Raumsonde im Lagrange Punkt. Letzteres ist der Punkt auf der Achse Sonne–Erde, in dem sich die Gravitationskräfte der beiden genau aufheben. Beispiele sind ISEE-3 (International Sun Earth Explorer), SOHO (Solar and Heliospheric Observatory) und WIND. Letzteres sind allerdings keine Satelliten mehr in dem Sinne, dass sie einen Zentralkörper umkreisen, sondern Raumsonden bzw. Raumflugkörper.
- Satelliten zur Erforschung der Erdmagnetosphäre befinden sich ebenfalls auf stark exzentrischen Orbits, allerdings mit geringerer Flughöhe. Die starke Exzentrizität erlaubt es, die radiale Variation der Magnetosphäre zu untersuchen; Beispiele sind ISEE-1,2 und Cluster-II.
- Für astrophysikalische Fragestellungen und solare Beobachtungen werden meist nahezu kreisförmige Bahnen in Höhen von einigen hundert Kilometern verwendet. Dies ergibt sich vielfach aus der Größe der Satelliten, ein VW-Bully ist klein und leicht gegen einige von ihnen. Diese Satelliten wurden vom Space-Shuttle ausgesetzt, führen aber auf Grund ihrer Größe und Masse kaum Treibstoff für Bahnmanöver mit sich, müssen also im wesentlichen dem Shuttle-Orbit folgen. Dies hat auch einen Vorteil: die Satelliten können mit Hilfe des Shuttle gewartet oder repariert werden, so wurde z.B. dem Hubble Space Telescope HST eine Brille verpasst, um seine Abbildungsqualität zu verbessern. Beispiele sind:
 - Solwind P78-1 ($H = 530$ km, $i = 98^\circ$) zur Erforschung der Ionosphäre und der Magnetosphäre sowie zur Beobachtung der Sonne (Koronagraph). Dieser militärische Forschungssatellit wurde nach 6 1/2 Jahren erfolgreicher Arbeit, obwohl noch voll funktionsfähig, von der US Air Force als Zielsatellit für Anti-Satellitenwaffen zweckentfremdet;
 - Solar Maximum Mission SMM ($H = 490$ km, $i = 29^\circ$) als Sonneobservatorium. SolarMax war der erste Satellit, der von einem Shuttle zu Reparaturzwecken wieder eingefangen

⁹Die Genauigkeit des Verfahrens hängt natürlich von der Lage der drei Orte ab – im Extremfall, dass alle drei auf einer Linie liegen, kann man den eigenen Ort nicht bestimmen sondern nur eine Linie angeben, auf der er liegen muss. Entsprechend ungenau wird das Verfahren, wenn alle drei angepeilten Punkte dicht bei einander liegen.

¹⁰Ein einfacher GPS-Empfänger kann die Informationen von bis zu 12 GPS-Satelliten verarbeiten – sofern er so viele im Blickfeld hat.

¹¹Für die sehr genaue Ortsbestimmung (Landvermessung, einige Rettungssysteme) verwendet man ein *differentielles GPS (DGPS)*: zusätzlich zu den Satellitensignalen werden die Signale einer ortsfesten Bodenstation als Referenz verwendet. Damit lassen sich die Satelliteninformationen korrigieren und Genauigkeiten im Bereich von einigen Zentimetern erreichen.

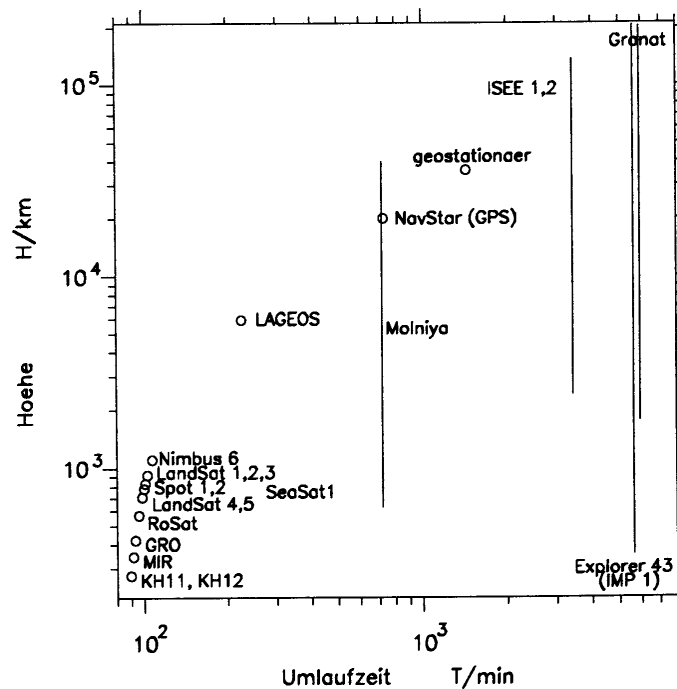


Abbildung 2.9: Höhen und Umlaufzeiten verschiedener Satelliten. Die Kreise markieren Satelliten auf nahezu kreisförmigen Bahnen, die Linien Satelliten in sehr exzentrischen Orbits.

wurde. Er stürzte 1989 nach 10 Jahren Betriebszeit in der Folge hoher solarer Aktivität ab – und wurde damit ein Opfer seines Beobachtungsobjektes.

- Solar A (Yokoh) ($H = 650 \text{ km}$, $i = 31^\circ$) dient der Röntgenbeobachtung der Sonne und wurde, wie seine hohe Flugbahn andeutet, nicht vom Shuttle sondern von einer Rakete gestartet.
- Gamma Ray Observatory GRO, später umbenannt in Compton Gamma Ray Observatory CGRO ($H = 420 \text{ km}$, $i = 29^\circ$) zur Gammabeobachtung und zur Suche nach neuen Röntgenquellen im Universum.
- Hubble Space Telescope HST ($H = 615 \text{ km}$, $i = 29^\circ$) als nicht von der Erdatmosphäre beeinträchtigtes astronomisches Teleskop.

Die sehr ähnlichen Inklinationen der Bahnen dieser Forschungssatelliten ergeben sich durch die geographische Breite des Startorts, im Falle von SMM, GRO und HST wurde vom Kennedy Space Center (Cape Canaveral) auf einer Breite von N29 gestartet, Yokoh wurde von Tanegashima bei N31 gestartet.

2.3.6 Zusammenfassung

Die Bahn eines Satelliten wird der Zielsetzung der Mission angepasst. Militärische Satelliten (KH-11, KH-12 in Abb. 2.9) fliegen in geringer Höhe auf Kreisbahnen. Erdkundungssatelliten (LandSat, SeaSat, SPOT, Nimbus) bewegen sich ebenfalls auf Kreisbahnen, jedoch in größerer Höhe um die Einflüsse der Reibung geringer zu halten und bei guter Bodenauflösung trotzdem große Bereiche überblicken zu können. Im Höhenbereich zwischen diesen beiden Arten von Erdfernerkundungssatelliten befinden sich die Raumstationen (MIR) und astrophysikalische Forschungssatelliten (RoSat, GRO). Hier sind die Bahndetails stark durch die Erreichbarkeit des Orbits mit Raumfähren oder -kapseln bestimmt. In Höhen oberhalb 2 000 km finden sich auf Kreisbahnen nur noch die NAVSTAR-Satelliten des Navigationssystems GPS, die geostationären Kommunikations- und Wettersatelliten sowie vereinzelt Forschungssatelliten (LAGEOS). Die meisten Orbits hoch fliegender Satelliten sind stark exzen-

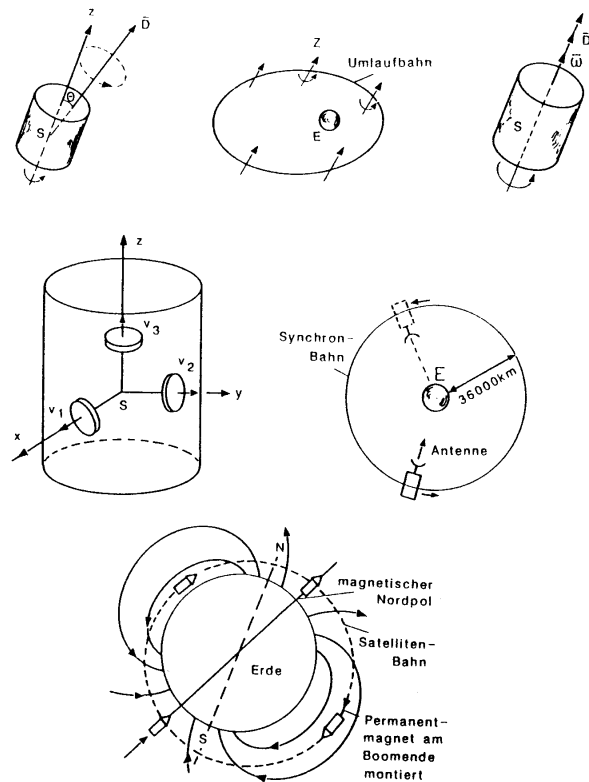


Abbildung 2.10: Stabilisierungssysteme: Spinstabilisierung (oben), Drei-Achsen-Stabilisierung (Mitte) und passive Stabilisierung (unten) [36]

trisch: einerseits auf Grund der Zielsetzung (bei Molniya lang dauernder Aufenthalt über der FSU, in der Magnetosphärenforschung Überdeckung eines weiten radialen Bereichs), andererseits aber auch, weil es energetisch wesentlich weniger aufwendig ist, einen Satelliten in eine Ellipsenbahn mit weit entferntem Apogäum zu bringen als in eine Kreisbahn mit entsprechendem Radius.

2.4 Stabilisierung des Satelliten

Die Stabilisierung eines Satelliten auf seiner Bahn betrifft nicht mehr die Bahn sondern die Lage des Satelliten in dieser Bahn und damit die Ausrichtung der auf dem Satelliten befindlichen Instrumente relativ zur Bahn und zum Target. Neben den Instrumenten muss auch die Antenne ausgerichtet werden, so dass diese stets auf eine Empfangs- oder Relaisstation gerichtet ist.

Formal muss die Bewegung des Satelliten in zwei Teile zerlegt werden: (a) die Bewegung des Massenmittelpunkts auf seiner Bahn, wie bisher in diesem Kapitel betrachtet, und (b) die Bewegung des Satelliten um seinen Massenmittelpunkt. Um die Ausrichtung der Instrumente und Antennen zu gewährleisten, bedient man sich verschiedener Stabilisierungssysteme, für Zusammenfassungen siehe z.B. [25, 36].

2.4.1 Stabilisierungsverfahren

Abbildung 2.10 gibt einen Überblick über verschiedene Lageregelungs- oder Stabilisierungsverfahren. Diese lassen sich in drei Klassen einteilen: Spinstabilisierung (oben), Drei-Achsen-Stabilisierung (Mitte) und passive Stabilisierung (unten).

Spinstabilisierung

Das einfachste Verfahren zur Stabilisierung eines Satelliten oder einer Raumsonde ist die Spinstabilisierung: der Satellit rotiert um eine Achse mit großem Trägheitsmoment, meist um die Figurenachse. Die Rotationsachse $\vec{\omega}$ ist parallel zum raumfesten Drehimpulsvektor \vec{L} . Auf Grund der Drehimpulserhaltung widersetzt sich der Satellit wie ein Kreisel jeder Änderung seiner Drehachse. Für praktische Anwendungen reichen zur Stabilisierung Spingeschwindigkeiten in der Größenordnung von einer Umdrehung pro Sekunde (Beispiele: TIROS I: 1/6 U/s; Syncom 2.7 U/s; Helios 1 U/s).

Spinstabilisierung ist für die Erdfernerkundung unpraktisch: sofern die Instrumente nicht parallel zur Drehachse ausgerichtet sind, ergibt sich ein ständig wechselndes Blickfeld. Für Aufnahmen der Erde ist Spinstabilisierung daher keine geeignete Form der Lagereglung. Für andere Typen von Instrumenten, z.B. Instrumente zur Messung energiereicher Teilchen, dagegen sind spinnende Plattformen von Vorteil: durch die Änderung der Blickrichtung wird der gesamte Raum abgescannt und man kann auf diese Weise die Richtungsverteilung der einfallenden Teilchen bestimmen und nicht nur die Teilchenintensitäten aus einer bestimmten Richtung. Ausnahme sind auch viele meteorologische Satelliten: hier wird die Spinstabilisierung gleichzeitig zum Scannen genutzt mit dem Vorteil eines einfachen Instruments ohne bewegliche mechanische Bestandteile. Die Konfiguration, bei der der Satellit gleichsam seine Bahn entlang spinnert, wird als Cartwheel bezeichnet.

Da die Drehachse bei spinnenden Raumsonden in der Regel senkrecht auf der Bahnebene steht, gibt es keine Möglichkeit, eine Richtantenne mit fester Ausrichtung auf eine Bodenstation auf dem Raumfahrzeug zu montieren. Daher muss zur Gewährleistung der Funkverbindung die Antenne relativ zum Satelliten rotieren, um eine Ausrichtung auf die Erde zu gewährleisten. Die Genauigkeit, mit der diese Gegenrotation erfolgt, ist so groß, dass für Helios 1 erst 12 Jahre nach dem Start des Satelliten eine Abweichung auftrat, die so groß war, dass der relativ schmale Strahl der Antenne die Erde nicht mehr traf. Bei einer Rotationsperiode von 1/s heißt das, dass der Antennenstrahl erst nach ca. 360 Mio Rotationen um wenige Grad von seiner ursprünglichen Richtung abgewichen ist.

Drei-Achsen-Stabilisierung

Bei der Drei-Achsen-Stabilisierung wird die momentane Lage des Satelliten durch Momentenaustausch mit Stellgliedern, z.B. Schwungrädern, gehalten. Dadurch lässt sich eine feste Ausrichtung von Antennen bzw. Instrumenten erreichen. Drei-Achsen-Stabilisierung erfolgt bei vielen geostationären Satelliten (Kommunikation), Erdfernerkundungssatelliten und bei vielen interplanetaren Raumsonden (z.B. Pioneer, Voyager, ISEE-3).

Passive Stabilisierungssysteme

Passive Stabilisierungssysteme nutzen die Umgebungsbedingungen zur Orientierung, z.B. Strahlungsdruck, Gravitationsgradienten oder terrestrisches Magnetfeld. Im unteren Teil von Abb. 2.10 ist als Beispiel die Lagestabilisierung des ersten deutschen Forschungssatelliten, Azur, durch das Magnetfeld dargestellt. Im Satelliten ist ein starker Permanentmagnet eingebaut. Da sich der Satellit auf einer polaren Umlaufbahn befindet, dreht er sich während eines Umlaufs um die Erde auch einmal um seine Achse. Bei Abweichungen aus seiner Lage parallel zu den Magnetfeldlinien bewirkt der Permanentmagnet ein rücktreibendes Moment, so dass die Satellitenachse stets parallel zu den Feldlinien bleibt und stets das gleiche Ende des Satelliten in Richtung auf den magnetischen Nordpol weist.

2.4.2 Sonnensynchrone Bahnen

Die sonnensynchrone Bahn ist ein Beispiel für die Stabilisierung der Bahnebene durch geschickte Ausnutzung einer Bahnstörung. Bei der sonnensynchronen Bahn ist die Bahnebene des Satelliten so ausgerichtet, dass sie stets in einem festen Winkel, meist senkrecht, zur

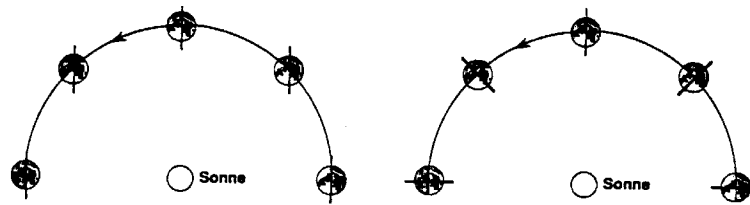


Abbildung 2.11: Polare und sonnensynchrone Bahn [19]

Erde–Sonne-Achse steht. Dazu muss eine feste Beziehung zwischen großer Halbachse a , Exzentrizität ε und Inklination i erfüllt sein, vgl. (2.45)

Abbildung 2.11 zeigt ein Beispiel für eine sonnensynchrone Bahn. Im linken Teil der Abbildung ist eine normale polare Bahn dargestellt: da der Bahndrehimpulsvektor und damit die Bahnebene raumfest sind, ändert sich ihre Orientierung im Bezug auf die Achse Sonne–Erde während der Rotation der Erde um die Sonne. Bei einer sonnensynchronen Bahn dagegen driftet der Bahndrehimpulsvektor und damit die Bahnebene derart, dass sie stets in einem festen Winkel zur Erde–Sonne-Achse stehen. Der Vorteil der sonnensynchronen Bahn liegt bei Erdfernerkundungssatelliten in den gleichbleibenden Beleuchtungsverhältnissen. Dies ist nicht nur für die Klassifikation von Bildelementen (vgl. Abschn. 6.3) wichtig, da ein Blatt dem Auge bzw. der Kamera anders erscheint, je nachdem, ob es von schräg oben oder senkrecht angestrahlt wird, sondern auch für die Entdeckung von baulicher Tätigkeit: diese erfolgt häufig über Veränderungen im Schattenwurf. Auch lassen sich in der Draufsicht keine Details der vertikalen Struktur erkennen – im Schatten dagegen schon, vgl. Abb. 2.12.

Für Beobachtungen der Sonne ist ein sonnensynchrones Orbit mit einer Bahnebene ungefähr senkrecht zur Erde–Sonne-Achse ebenfalls gut geeignet, da der Satellit die Sonne während des gesamten Umlaufs im Blickfeld hat und die Beobachtungspause auf der Nachtseite der Erde entfällt. Auf einer derartigen sonnensynchronen Bahn wird der Satellit während des gesamten Orbits von der Sonne angestrahlt – bei der Energieversorgung mit Solarzellen ein großer Vorteil.

Die Neigung der Bahnebene zur Erde–Sonne-Achse kann beliebig sein, je nach Anforderungen. Bahnebenen nahezu senkrecht zu dieser Achse haben den Vorteil, dass die Solarzellen stets in der Sonne sind, allerdings werden alle Punkte der Erde um 6 bzw. 18 Uhr lokaler Zeit überflogen – und damit im Winter in beiden Abschnitten des Orbits in der Dämmerung bzw. im Dunkeln. Umgekehrt ist der Schattenwurf zu diesen Zeiten jedoch recht ausgeprägt. Im anderen Extrem, Bahnebene parallel zur Erde–Sonne-Achse, passiert überfliegt der Satellit in



Abbildung 2.12: Ein Gegenstand kann manchmal über seinen Schatten wesentlich einfacher zu identifizieren sein als direkt, wie in dieser Luftaufnahme am Mast der Hochspannungsleitung deutlich wird [71]

Accelerations of spacecraft in earth orbit. The estimates and the relative uncertainties are given for a typical geosynchronous satellite and LAGEOS. M , R and $J_{\ell m}$ are the mass, the radius and the geopotential coefficients of the earth; r is the average satellite–earth distance; the subscripts m, \odot and v refer to the moon, the sun and Venus; S/m , C_D and V are the satellite’s area-to-mass ratio, drag coefficient and velocity relative to the atmosphere (of density ρ); Φ is the solar constant; A the earth’s albedo; α and $\Delta T/T$ are the satellite’s absorption coefficient and the fractional temperature difference between parts of its surface.

Origin of acceleration	Formula	Parameters (in CGS units)	Relative uncertainty in the parameters	Geosynchronous satellite $a = 42,160 \text{ km}$, $S/m = 0.1 \text{ cm}^2/\text{g}$ in cm/s^2		LAGEOS $a = 12,270 \text{ km}$, $S/m = 0.007 \text{ cm}^2/\text{g}$ in cm/s^2	
				Acceleration	Uncertainty	Acceleration	Uncertainty
1. Earth’s monopole	GM/r^2	$GM = 3.986 \cdot 10^{20}$	$5 \cdot 10^{-9}$	22	$2 \cdot 10^{-7}$	280	$3 \cdot 10^{-6}$
2. Earth’s oblateness	$\frac{3GM R^2}{r^4} J_{20}$	$J_{20} = 4.84 \cdot 10^{-4}$ $R = 6.378 \cdot 10^8$	$6 \cdot 10^{-7}$	$7.4 \cdot 10^{-4}$	$4.5 \cdot 10^{-10}$	0.1	$6 \cdot 10^{-8}$
2. Low-order geopotential harmonics ($\ell = m = 2$)	$\frac{3GM R^2}{r^4} J_{22}$	$J_{22} = 2.81 \cdot 10^{-6}$	$4 \cdot 10^{-4}$	$4.3 \cdot 10^{-6}$	$2 \cdot 10^{-9}$	$6.0 \cdot 10^{-4}$	$2 \cdot 10^{-7}$
2. <i>Ibidem</i> , ($\ell = m = 6$)	$\frac{7GM R^6}{r^7} J_{66}$	$J_{66} = 2.36 \cdot 10^{-7}$	$7 \cdot 10^{-3}$	$4.4 \cdot 10^{-10}$	$3 \cdot 10^{-12}$	$8.6 \cdot 10^{-6}$	$6 \cdot 10^{-8}$
2. High-order geopotential harmonics ($\ell = m = 18$)	$\frac{19GM R^{18}}{r^{19}} J_{1818}$	$J_{1818} = 7 \cdot 10^{-9}$	0.3	negligible	negligible	$2.7 \cdot 10^{-10}$	$1 \cdot 10^{-10}$
3. Perturbations due to the Moon	$\frac{2GM_m r}{r_m^3}$	$GM_m = GM/81.3$ $r_m \approx 3.8 \cdot 10^{10}$	10^{-5} 10^{-9}	$7.3 \cdot 10^{-4}$	$7 \cdot 10^{-9}$	$2.1 \cdot 10^{-4}$	$2 \cdot 10^{-9}$
3. Perturbations due to the Sun	$\frac{2GM_\odot r}{r_\odot^3}$	$GM_\odot = 3.33 \cdot 10^8 GM$ $r_\odot \approx 1.5 \cdot 10^{13}$	$3 \cdot 10^{-8}$ 10^{-8}	$3.3 \cdot 10^{-4}$	negligible	$9.6 \cdot 10^{-5}$	negligible
3. Perturbations due to other planets (e.g., Venus)	$\frac{2GM_v r}{r_v^3}$	$GM_v = 0.82 \cdot GM$ $r_v \geq 4 \cdot 10^{12}$	10^{-4} 10^{-7}	$4.3 \cdot 10^{-8}$	negligible	$1.3 \cdot 10^{-8}$	negligible
4. General relativistic corrections	$\frac{GM}{c^2 r} \frac{GM}{r^2}$	$\frac{GM}{c^2} = 0.44$	$5 \cdot 10^{-9}$	$2.3 \cdot 10^{-9}$	negligible	$9.5 \cdot 10^{-8}$	negligible
5. Atmospheric drag $V = \text{volume}$, $S = \text{surface}$	$C_D \frac{S}{2m} \rho V^2$	$C_D = 2-4$ $\rho = \rho(r)$	1 1	negligible	negligible	$3 \cdot 10^{-10}$	$3 \cdot 10^{-10}$
6. Solar radiation pressure	$\frac{S}{m} \frac{\Phi}{c}$	$\Phi = 1.38 \cdot 10^6$	$S: 0.01-0.3$ $\Phi: 0.001$	$4.6 \cdot 10^{-6}$	$9 \cdot 10^{-7}$	$3.2 \cdot 10^{-7}$	$6 \cdot 10^{-9}$
7. Radiation pressure due to the albedo of the Earth A	$\frac{S}{m} \frac{\Phi}{c} \left(\frac{R}{r}\right)^2 A$	$A \approx 0.4$	Opt. coeff.: 0.01–0.1 $A: 0.1-1$ $S: 0.01-0.3$	$4.2 \cdot 10^{-8}$	10^{-8}	$3.4 \cdot 10^{-8}$	$8 \cdot 10^{-9}$
8. Recoil from thermal emission	$\frac{4}{3} \frac{S}{m} \frac{\Phi}{c} \alpha \frac{\Delta T}{T}$	$\alpha \approx 0.1-0.7$ $T \approx 300 \text{ K}$ $\Delta T \sim 1-20 \text{ K}$	0.01–0.1 0.01 0.1	$3 \cdot 10^{-8}$	$6 \cdot 10^{-9}$	$5 \cdot 10^{-9}$	$1 \cdot 10^{-9}$

Tabelle 2.4: Auf einen Satelliten wirkende Kräfte [8]

einem Teil des Orbits um 12 Uhr lokaler Zeit, im anderen um Mitternacht. Auf der Tagseite ist daher ein senkrechter Einfall der Sonne ohne bzw. mit minimalem Schattenwurf gegeben, auf der Nachtseite ist mit Instrumenten im sichtbaren Bereich des elektromagnetischen Spektrums keine Messung möglich. Dafür wird aber (außer den hohen Breiten der Winterhemisphäre) jeder Punkt der Erde mindestens einmal bei Tageslicht überflogen. Welches dieser Orbits (bzw. welches dazwischen liegende) sinnvoll ist, hängt von der Fragestellung ab. Bei sonnensynchronen Satelliten wird meistens zur Beschreibung der Bahn der Zeitpunkt (in lokaler Zeit) angegeben, an dem sie auf der Tagseite den Äquator kreuzen.

2.5 Bahnstörungen

Bisher haben wir Satellitenbahnen als antriebslose Bahnen ausschließlich unter dem Einfluss der Gravitationskraft betrachtet. In diesem Abschnitt werden wir auf einige Störungen von Satellitenbahnen eingehen und insbesondere auf die Umstände, unter denen diese auftreten. Wir werden auch lernen, dass Bahnstörungen nicht immer etwas Negatives sein müssen sondern auch positiv angewendet werden können.

2.5.1 Übersicht

Bahnstörungen werden in der Bewegungsgleichung derart berücksichtigt, dass zusätzlich zur Gravitationskraft eine Störkraft \vec{F}_S eingeführt wird und die erweiterte Bewegungsgleichung gelöst wird. Einige dieser Kräfte sind in Tabelle 2.4 zusammen gefasst. Zusätzlich enthält die Tabelle die sich aus diesen Kräften ergebenden Störbeschleunigungen für einen Satelliten in einem geostationären Orbit, also einen typischen Kommunikations- oder Wettersatelliten, und für LAGEOS, einen Geodäsie-Satelliten auf einer Flughöhe von ca. 12 000 km.

Die Störbeschleunigungen lassen sich klassifizieren in (a) gravitative Effekte (1–3 in Tabelle 2.4), (b) Reibung in der Atmosphäre (5), (c) Strahlungsdruck (6,7), (d) Sonnenwind, (e) thermische Emission (8), (f) zufällige Einflüsse, sowie (g) atmosphärische Gezeiten.

Gravitative Einflüsse

Gravitative Einflüsse haben eine Vielzahl von Ursachen. Als erste ist die Ungenauigkeit in der Bestimmung der Gravitationskonstanten γ zu nennen. Diese Störung ist allerdings von allen gravitativen Einflüssen die Unbedeutendste.

Einen größeren Einfluss hat die Abplattung der Erde, in Tab. 2.4 nach verschiedenen Ordnungen aufgeschlüsselt. Die Erdaplattung hat einen größeren Einfluss auf niedrig fliegende Satelliten als auf solche in höheren Orbits – in letzterem Fall kommt die Erde der idealisierten Punktmasse bereits recht nahe, in ersterem dagegen erfährt der Satellit die Abweichungen von der Idealgestalt.

Ein weiterer gravitativer Effekt sind die Einflüsse anderer Himmelskörper. Diese sind, abgesehen von der Anziehungskraft des Mondes auf hoch fliegende Satelliten, zu vernachlässigen. Die Gravitation dritter Körper ist jedoch für spezielle Bahnen, wie die Aufhängung von Satelliten im Lagrange-Punkt von Bedeutung und bei Bahnen eingefangener Körper (z.B. der Planet Levy-Shoemaker 9 im Schwerefeld des Jupiter; teilweise auch beim planetaren Billard).

Reibung in der Hochatmosphäre

Die durch die Reibung in der Hochatmosphäre bewirkte Kraft ist um so größer, je niedriger das Orbit des Satelliten, da die Dichte der Atmosphäre mit zunehmender Höhe ungefähr exponentiell abnimmt. Eine Abschätzung des Effekts ist insbesondere für sehr niedrig fliegende Satelliten schwierig, da sich die Eigenschaften der Hochatmosphäre in Abhängigkeit von der solaren Aktivität erheblich ändern können.

Strahlungsdruck

Hier muss zwischen dem direkten Strahlungsdruck durch die auf den Satelliten auftreffende solare Strahlung unterschieden werden und dem indirekten Strahlungsdruck, der durch von der Erde bzw. der Erdatmosphäre reflektierte solare Strahlung entsteht. Beide Effekte sind für normale Erdsatelliten unerheblich, sie spielen jedoch bei leichten und großen Satelliten (z.B. die Echo-Ballon-Satelliten) eine nicht zu vernachlässigende Rolle.

Sonnenwind

Der Sonnenwind ist ein ca. 1 Mio K heißes, kontinuierlich von der Sonne abströmendes Plasma mit einer Dichte von einigen Teilchen pro cm^3 . Da ein Plasma aus geladenen Teilchen besteht, wird es vom Erdmagnetfeld abgelenkt, d.h. der Sonnenwind gelangt nicht in die Erdmagnetosphäre. Letztere hat eine Ausdehnung von mindestens 10 Erdradien, d.h. der Sonnenwind beeinflusst nur Satelliten, die in extrem hohen und/oder sehr exzentrischen Orbits fliegen und meist auch die Aufgabe haben, eben diesen Sonnenwind zu vermessen. Erdfernerkundungssatelliten kommen normalerweise nicht in den Einflussbereich des Sonnenwindes.

Thermische Emission

Die thermische Emission geschieht in Folge der Erwärmung des Satelliten durch die absorbierte solare Einstrahlung. Diese thermische Emission bewirkt auf Grund der Impulserhaltung einen Rückstoß auf den Satelliten. Diese Bahnstörung kann daher als die Umkehrung des Strahlungsdrucks angesehen werden. Der Einfluss dieser Störgröße hängt ab von der Geometrie des Satelliten, dem verwendeten Lagestabilisierungsverfahren und dem Auftreten und der Geometrie herausragender Merkmale wie Antennen und Sonnensegel. Als Faustregel gilt auch hier: der Satellit muss groß und leicht sein, damit diese Störgröße zu merkbaren Effekten führen kann.

Zufällige Einflüsse

Zufällige Einflüsse umfassen z.B. Meteoriten als direkt auf den Satelliten wirkende Kräfte und solare Flares, die über eine Aufheizung und Ausdehnung der Hochatmosphäre zu einer

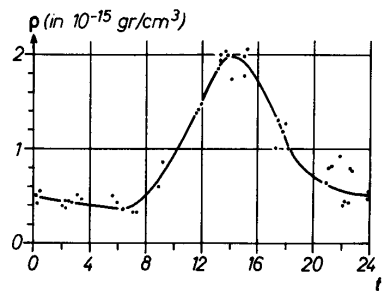


Abbildung 2.13: Tägliche Dichteveriation in einer Höhe von 560 km, geschlossen aus der Bahn von Vanguard 2 [58]

verstärkten Satellitenabbremung durch Reibung führen. Auf Grund ihrer zufälligen Natur lassen sich diese Störgrößen weder im Zeitpunkt oder der Rate des Auftretens noch in ihrer Größe klassifizieren.

Atmosphärische Gezeiten

Atmosphärische Gezeiten sind ein Teilaspekt der Reibung in der Hochatmosphäre, d.h. sie wirken auf einen Satelliten in niedrigem Orbit stärker als auf einen hoch fliegenden. Sie werden als separater Effekt aufgeführt, da die atmosphärischen Gezeiten eine periodisch wirkende Störgröße bilden, so dass es im Gegensatz zur normalen Reibung unter ungünstigen Umständen zu einer Resonanz zwischen den Gezeiten und bestimmten Eigenschaften des Orbits kommen kann.

Formale Behandlung

Formal werden die Einflüsse aller Störgrößen im Rahmen der Störungsrechnung behandelt. Ziel ist die Beschreibung der durch die Störgrößen bewirkten Änderung der Bahnparameter a , ε , ω , i und Ω in Abhängigkeit von der Zeit. In der Störungsrechnung werden die auf den Satelliten wirkenden Kräfte zerlegt in eine systematische Kraft auf Grund des Schwerefeldes der Erde und eine Störbeschleunigung \vec{a}_S . Damit ergibt sich für die Bewegungsgleichung

$$\ddot{\vec{r}} = -\mu \frac{\vec{r}}{r^3} + \vec{a}_S . \quad (2.34)$$

Die Störbeschleunigung wird in drei Komponenten zerlegt, eine radiale Störbeschleunigung a_r , eine senkrecht zum Ortsvektor \vec{r} wirkende Beschleunigung a_s in der Bahnebene, und eine Störbeschleunigung a_ω senkrecht zur Bahnebene. Die Komponenten a_r und a_s wirken auf die Bahnparameter a , ε und ω ; sie werden z.B. durch Reibung und Erdabplattung verursacht. Die Störbeschleunigung a_ω dagegen wirkt auf i und Ω , Ursache ist z.B. die Erdabplattung.

Bei allen Störungen wird zusätzlich zwischen periodischen Störungen und sekulären unterschieden. Letztere erfolgen zeitlich stets gleichsinnig.

2.5.2 Reibung in der Hochatmosphäre

Reibung in der Hochatmosphäre bewirkt eine Abbremung des Satelliten und damit ein Absinken auf niedrigere Bahnen bis hin zum Absturz. Reibung ist aber nicht nur ein Störeffekt – die beobachtete Abbremung eines Satelliten wird umgekehrt auch zur Messung der Temperatur¹² bzw. Dichte der Hochatmosphäre verwendet. Aus Veränderungen der Satellitenbe-

¹²Temperaturmessung in der Hochatmosphäre ist ein nicht-triviales Problem. Zwar sind die Temperaturen sehr hoch, etliche hundert bis zu 2000 K, jedoch ist die Dichte der Atmosphäre so gering, dass keine messbare Wärmemenge durch Stöße an einen Temperaturfühler übertragen werden kann. Daher entfallen konventionelle Temperaturmessverfahren. Auf Grund der geringen Dichte und der starken thermischen Emission des Erdbodens ist es auch nicht möglich, eine Effektivtemperatur über den emittierten Strahlungsstrom

wegung hat man bereits früh die starke Variabilität der Hochatmosphäre erkannt.

Abbildung 2.13 zeigt als Beispiel die täglichen Dichteveriationen in einer Höhe von 560 km, wie sie aus den Bahnvariationen von Vanguard 2 hergeleitet wurden. In Übereinstimmung mit dem starken Tagesgang der Temperatur in der Hochatmosphäre, zwischen 600 und 1800 K je nach Tageszeit und solarer Aktivität, treten in der Dichte im Verlauf des Tages Variationen um bis zu einem Faktor 5 auf. Das Dichtemaximum tritt kurz nach 14 Uhr lokaler Zeit auf, d.h. kurz nach dem Maximum der solaren Einstrahlung. Diese Verzögerung erklärt sich durch Wärmeleitungsvorgänge in der Atmosphäre [5], eine ähnliche Verzögerung des Temperaturgangs ist uns auch aus der Troposphäre bekannt. Die Dichteveriation entlang eines Breitenkreises lässt sich auch als die Wanderung einer ‘Gasglocke’ (diurnal bulge) beschreiben, die die Erde dem Sonnenstand folgend umkreist. Oder weniger geozentrisch formuliert: die Dichteglocke befindet sich ortsfest annähernd in solarer Richtung und die Erde dreht sich im Laufe des Tages einmal unter dieser Glocke hinweg.

Zur Abschätzung des Einfluss der Reibung setzen wir die Gravitationskraft als Zentralkraft ein und erhalten mit m_S als der Masse des Satelliten und M_E als der Masse der Erde

$$\frac{m_S v^2}{r} = \gamma \frac{M_E m_S}{r^2} \quad \text{und damit} \quad \gamma M_E = v^2 r . \quad (2.35)$$

Damit lässt sich die kinetische Energie schreiben als

$$E_{\text{kin}} = \frac{m_S v^2}{2} = \frac{\gamma M_E m_S}{2r} . \quad (2.36)$$

Für die potentielle Energie gilt

$$E_{\text{pot}} = mg(h) h = \frac{\gamma M_E m_S}{r} \quad \text{und damit} \quad E_{\text{pot}} = 2E_{\text{kin}} . \quad (2.37)$$

Für die Reibungskraft machen wir einen Newton’schen Ansatz, d.h. die Reibungskraft ist proportional der Geschwindigkeit: $F_{\text{Reib}} = -Dv$. Die Proportionalitätskonstante D berücksichtigt:

- die Querschnittsfläche A des Satelliten: je größer A , um so mehr Teilchen können Impuls auf den Satelliten übertragen;
- die Dichte ρ der Atmosphäre: je größer ρ , um so mehr Teilchen können Impuls auf den Satelliten übertragen;
- die Geschwindigkeit v des Satelliten als ein Maß für die Größe des Impulsübertrags pro Stoß; und den
- Widerstandsbeiwert c_D , der die Form des Satelliten und damit die Effizienz des Impulsübertrags berücksichtigt.

Mit Av als dem Volumen, das pro Zeiteinheit vom Satelliten durchquert wird ist $Av\rho$ die Luftmasse, die den Satelliten pro Zeiteinheit trifft. Diese überträgt pro Zeiteinheit einen Impuls $(Av\rho)v$. Berücksichtigt man den Widerstandsbeiwert c_D , so ergibt sich für die verzögernde Kraft

$$F_{\text{Reib}} = -\frac{1}{2} c_D \rho A v^2 . \quad (2.38)$$

Der Widerstandsbeiwert beträgt ca. 2 für inelastische und ca. 4 für elastische Stöße [10].

Multiplikation von (2.38) mit v und Umschreiben der linken Seite ergibt für die Änderung der kinetischen Energie mit der Zeit

$$\dot{E}_{\text{kin}} = \frac{d}{dt} \left(\frac{m_S v^2}{2} \right) = -\frac{\rho A c_D v^3}{2} . \quad (2.39)$$

zu bestimmen. In der Hochatmosphäre lässt sich die Temperatur daher nur noch über die kinetische Energie der Teilchen definieren gemäß Boltzmann-Verteilung, d.h. mit der mittleren Teilchengeschwindigkeit $\langle v \rangle$ ergibt sich für die Temperatur T : $m\langle v \rangle^2/2 = f k_B T/2$ mit f als der Zahl der Freiheitsgrade und k_B als der Boltzmann-Konstanten.

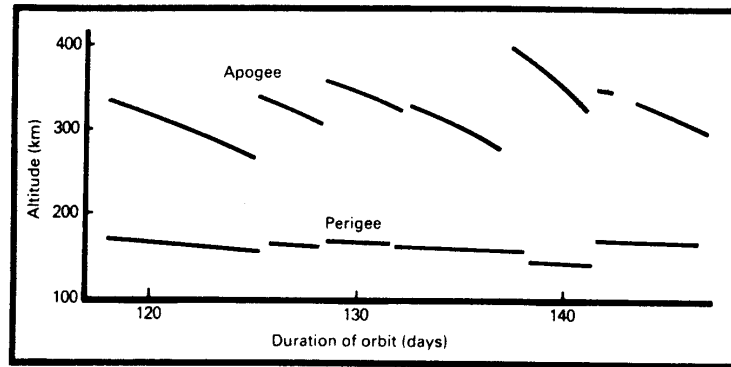


Abbildung 2.14: Einfluss des Luftwiderstands auf eine elliptische Satellitenbahn [63]. Die Diskontinuitäten sind Folge von Bahnmanövern

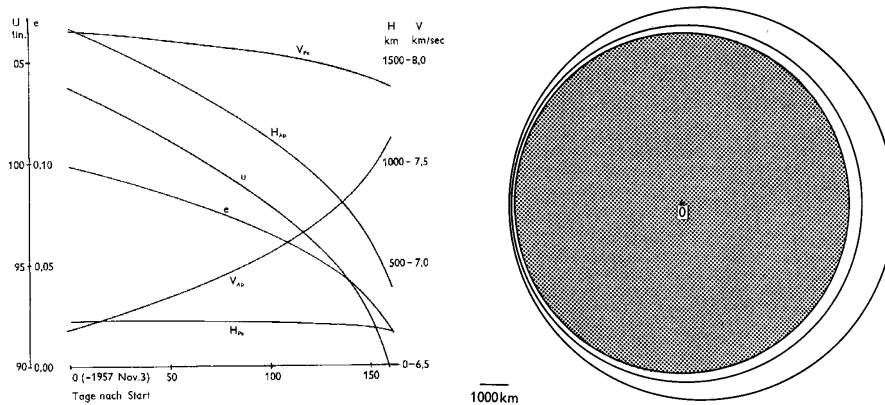


Abbildung 2.15: Änderung der Bahnparameter von Sputnik 2 in Folge des Luftwiderstands (links) und Anfangs- und Endbahn von Sputnik 2 (rechts) [10]

Die Rate der Änderung der Flughöhe eines Satelliten lässt sich aus der Rate der Änderung der kinetischen Energie abschätzen:

$$\frac{d}{dt} E_{\text{kin}} = \frac{d}{dr} E_{\text{kin}} \frac{dr}{dt} \quad (2.40)$$

bzw. nach Auflösen und Einsetzen von (2.38)

$$-\frac{dr}{dt} = \frac{\rho v A C_D r}{m_S} \quad (2.41)$$

Da sich Umlaufzeiten oder -frequenzen genauer bestimmen lassen als Flughöhen, ist es sinnvoller, die Änderung der Umlaufzeit $T = 2\pi r/v$ zu betrachten. Unter Berücksichtigung von (2.37) ergibt sich für die Abhängigkeit der Umlaufzeit von der potentialen Energie $T^2 = 4\pi^2 r^3 / (\gamma M_E)$ und damit nach Ableiten für die zeitliche Änderung der Umlaufzeit

$$2T \frac{dT}{dt} = \frac{4\pi^2 3r^2}{\gamma M_E} \frac{dr}{dt} \quad \Rightarrow \quad \frac{dT}{dt} = -\frac{3\pi A C_D r}{m_S} \rho \quad (2.42)$$

Die Änderung der Umlaufzeit eines Satelliten in einem kreisförmigen Orbit ist also direkt proportional der Dichte der umgebenden Atmosphäre. Die Bahnschrumpfung, d.h. die Verringerung der Umlaufhöhe bzw. Umlaufzeit, ist um so größer, je größer das Verhältnis aus Fläche A und Masse m_S des Satelliten ist. Große Satelliten geringer Masse werden daher stärker abgebremst als kleine kompakte Satelliten.

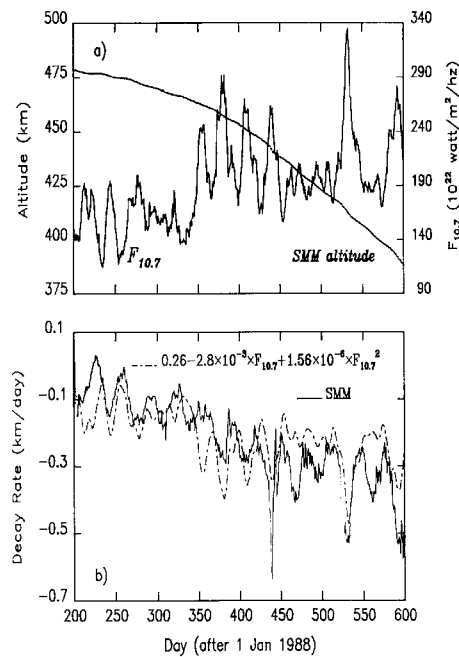


Abbildung 2.16: Abbremsung des SMM-Satelliten durch Zunahme solarer Aktivität [67]

Die Herleitung von (2.42) geht von einer Kreisbahn aus. Das ist sinnvoll, da gerade die niedrig fliegenden Späher sich auf nahezu kreisförmigen Orbits befinden. Bei einer elliptischen Bahn wirkt sich die Reibung im Detail etwas anders aus: die verzögernde Kraft ist im Perigäum größer als im Apogäum. Als Konsequenz verringert sich die kinetische Energie des Satelliten im Perigäum stärker als im Apogäum. Dadurch wird beim folgenden Apogäum eine deutlich geringere Höhe erreicht während die Perigäumshöhe nur geringfügig verringert wird. Als Konsequenz nimmt die Apogäumshöhe langsamer ab als die Perigäumshöhe. Dies ist am Beispiel des Satelliten Cosmos 1097 in Abb. 2.14 gezeigt. Die sprunghaften Anstiege in den Höhen ergeben sich durch Feuern der Booster, um den Satelliten auf eine höhere Umlaufbahn zu bringen und damit vor dem sofortigen Absturz zu bewahren.

Die schnellere Abnahme der Apogäums- im Vergleich zur Perigäumshöhe bewirkt, dass die elliptische Bahn sich im Laufe der Zeit immer stärker einer Kreisbahn annähert. Kurz vor dem Absturz eines Satelliten geht die Exzentrizität sehr schnell gegen Null, so dass man $\dot{\varepsilon}$ verwenden kann, um die Lebensdauer eines Satelliten abzuschätzen. Die Veränderung der Exzentrizität ε , der Umlaufzeit U , der Apogäums- und Perigäumshöhen H_{Ap} und H_{Pe} sowie der Geschwindigkeiten v_{Ap} und v_{Pe} in diesen Punkten sind in Abb. 2.15 für Sputnik 2 gezeigt.

Die Lebenserwartung eines Satelliten, der nur noch eine Umlaufzeit von 88–89 min hat, ist gering. Die untere Grenze für T beträgt 87 min, entsprechend einer Höhe von ca. 150 km.

Absturz durch solare Aktivität: SMM

Besonders stark ist die Satellitenabbremmung zu Zeiten hoher solarer Aktivität. Dann wird die Hochatmosphäre durch die harte elektromagnetische Strahlung stärker aufgeheizt und dehnt sich insgesamt aus. Dadurch werden dichtere Luftschichten in größere Höhen gehoben – oder aus der Sicht des Satelliten: die Dichte der Atmosphäre auf seiner Bahn nimmt zu und damit auch die Abbremsung. So sind des öfteren Satelliten durch erhöhte solare Aktivität abgestürzt bzw. schneller abgestürzt als geplant. Prominenteste Beispiele sind das Raum-Laboratorium SkyLab und SolarMax.

Abbildung 2.16 zeigt im oberen Teil die Flughöhe von SMM (Solar Maximum Mission, SolarMax, 1980–1989) zusammen mit der 10.7 GHz Radio-Strahlung der Sonne als einem Maß

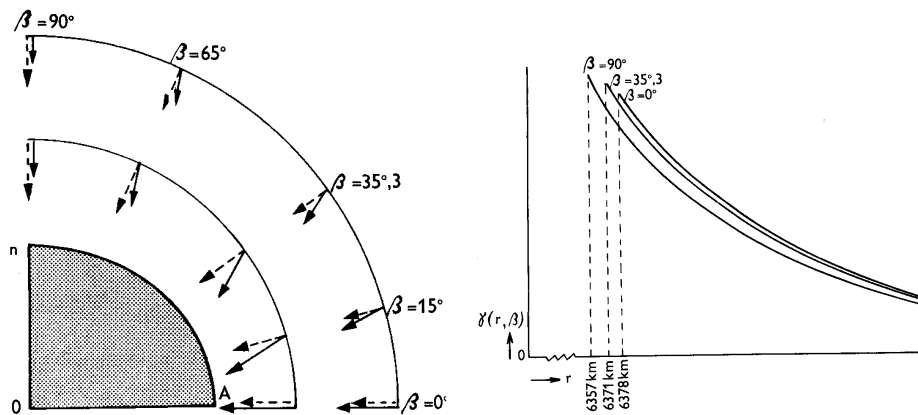


Abbildung 2.17: Abplattung der Erde: Richtung der Schwerkraftbeschleunigung (links) und Wert der Schwerkraftbeschleunigung für verschiedene Breiten in Abhängigkeit vom radialen Abstand (rechts) [10]

für die solare Aktivität und damit die Aufheizung der Hochatmosphäre. Letzterer variiert mit der Rotation der Sonne, d.h. auf einer Zeitskala von ca. 27 Tagen. Im Laufe der Zeit steigt das allgemeine Niveau der Radiostrahlung an, gleichzeitig nimmt die Rate der Abnahme der Flughöhe von SMM zu. Dies allein ist noch kein Zeichen für eine Kopplung von Flughöhe und solarer Aktivität, da selbst bei konstanten Umgebungsbedingungen die Flughöhe mit zunehmender Zeit schneller abnimmt.

Die direkte Kopplung wird in den Details des unteren Teilbildes von Abb. 2.16 deutlicher. Hier sind die Schwankungen im Radiofluss reziprok dargestellt und auf ein Äquivalent zur Änderungsrate der Flughöhe normiert (gestrichelte Kurve). Zusätzlich ist die Änderungsrate der Flughöhe aufgetragen (durchgezogene Linie). Man erkennt, dass die Änderung der Flughöhe dann besonders klein ist, wenn die reziproke Radioemission ein Maximum hat, d.h. wenn die solare Aktivität gering ist. Umgekehrt ist bei hoher solarer Aktivität die Änderungsrate der Satellitenhöhe groß, d.h. der Satellit verliert schnell an Höhe und stürzt letztendlich ab.

2.5.3 Bahnstörungen in Folge der Erdabplattung

Eine weitere Bahnstörung in niedrigen Orbits ist die Abplattung der Erde, in schematischer Form im linken Teil von Abb. 2.17 dargestellt. Die Abplattung bewirkt eine Abweichung der Richtung der Schwerkraftbeschleunigung von der radialen Richtung, in übertriebener Form dargestellt durch die ausgezogenen Vektoren. Auf einen Satelliten im Orbit wirkt die Schwerkraftbeschleunigung daher nicht nur in radialer Richtung sondern hat auch eine tangential Komponente. Letztere führt zu einer Änderung der Geschwindigkeit entlang des Orbits: wird der Satellit verzögert, so verringert sich seine Flughöhe, im umgekehrten Fall erhöht sie sich. Die Abplattung der Erde beträgt mit $R_E = 6278.2$ km als dem äquatorialen Radius und P_E als dem Polarradius $(R_E - P_E)/R_E \approx 1/298.2$.

Im rechten Teil von Abb. 2.17 ist die Gravitationsbeschleunigung als Funktion des Abstands für verschiedene Breiten β dargestellt. Zwei Effekte werden deutlich: für einen Satelliten in äquatorialer Bahn ist die Erdabplattung irrelevant, da er die Erde immer an der breitesten Stelle umkreist und somit die Schwerkraftbeschleunigung immer radial nach innen gerichtet ist, die Abplattung also keine Auswirkungen hat. Ferner sehen wir, dass mit zunehmendem radialen Abstand die durch die Abplattung bewirkte Änderung in Richtung und Betrag der Gravitationsbeschleunigung immer kleiner wird, da die Erde aus der Sicht eines Satelliten mit zunehmendem Abstand immer mehr zum Massenpunkt zusammen schmilzt.

Für eine quantitative Abschätzung entwickeln wir das Gravitationspotential nach Multi-

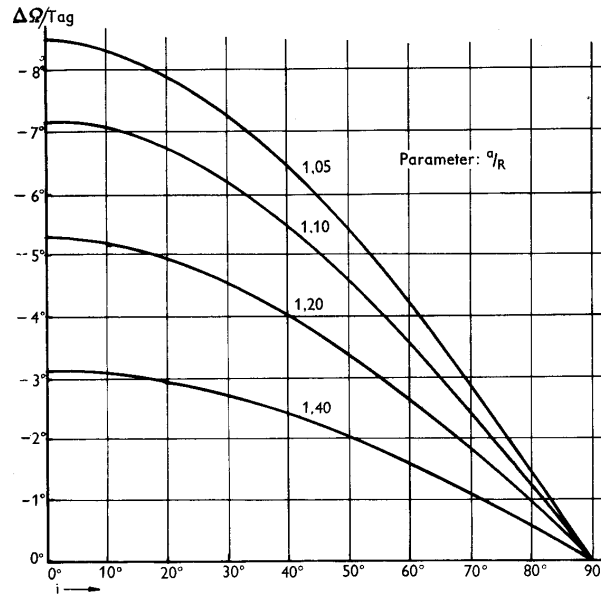


Abbildung 2.18: Präzession der Bahnebene in Folge der Erdabplattung als Funktion der Inklination i und des mittleren Bahnabstandes a/R_E [10]

polen [58]:

$$U = \frac{\mu}{r} \left[1 + \frac{1}{2r^2} \frac{C - A}{M_E} (1 - 3 \sin^2 \phi) \right] + \dots = \frac{\mu}{r} \left[1 + \frac{J}{3} \frac{R_E^2}{r^2} (1 - 3 \sin^2 \phi) \right] + \dots \quad (2.43)$$

mit ϕ als Breite, C als Trägheitsmoment um die Rotationsachse (z -Achse) und A als Trägheitsmoment um eine senkrecht dazu stehende x -Achse:

$$A = \int (z^2 + y^2) dm, \quad C = \int (x^2 + y^2) dm \quad \text{und} \quad J = \frac{3}{2} \frac{C - A}{R_E^2 M_E} \approx 1.6234 \cdot 10^{-3}. \quad (2.44)$$

Präzession

Für die Satellitenbahn ergeben sich aus der Abplattung der Erde zwei Konsequenzen: (a) die Drehung der Bahnebene (Präzession) und (b) die Drehung der Bahnellipse (Apsidendrehung). Die Herleitungen zu beiden Prozessen finden Sie in [8, 92]. Die wesentlichen Ergebnisse für die Präzession der Bahn sind

$$\frac{\Delta\Omega}{\Delta T} = -2\pi J \left(\frac{R_E}{p} \right)^2 \cos^2 i \quad \text{bzw.} \quad \frac{\Delta\Omega}{\text{Tag}} = \frac{10^\circ}{\left(\frac{a}{R_E} \right)^{7/2} (1 - \varepsilon^2)^2} \cos i. \quad (2.45)$$

Die Präzession hängt also von der Inklination i ab und verschwindet, wie erwartet, für äquatoriale Bahnen, vgl. Abb. 2.18. Sie hängt ferner von der Exzentrizität ε der Bahn und der Flughöhe ab, hier gegeben durch die große Halbachse a/R_E in Einheiten des Erdradius. Als Beispiel ergibt sich für Sputnik 2 mit $a/R_E = 1.13$, $\varepsilon = 0.09$ und $i = 65.3^\circ$ eine Drehung der Bahnebene von $2.77^\circ/\text{Tag}$; für Explorer 7 mit $a/R_E = 1.128$, $\varepsilon = 0.0377$ und $i = 50.3^\circ$ ergibt sich $\Delta\Omega = -4.27^\circ/\text{Tag}$. Das Vorzeichen der Bahndrehung sagt aus, ob die Bahn nach Ost oder West ausweicht – das Minuszeichen gilt für den Normalfall, dass die Bewegung des Satelliten in Ost–West-Richtung erfolgt.¹³ Für beide Satelliten waren die beobachteten Werte etwas geringer als die berechneten (Sputnik 2.72° , Explorer 4.21°). Diese Abweichungen

¹³Das ist auch die am häufigsten vorkommende Richtung, da in diesem Fall die Erddrehung beim Abschuss ausgenutzt wird. Bei Bahnen in entgegengesetzter Richtung erfolgt die Präzession mit umgekehrtem Vorzeichen (retrograd).

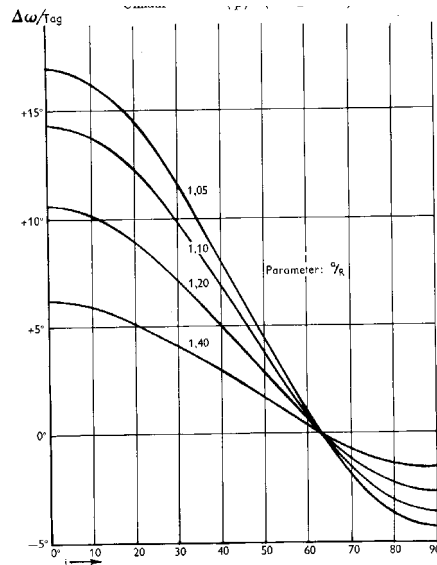


Abbildung 2.19: Apsidendrehung als Funktion der Inklination und des mittleren Abstands [10]

zwischen beobachteter und berechneter Bahndrehung erlauben eine genauere Bestimmung von J .

Eine Nutzanwendung der Drehung der Bahnebene ist die sonnensynchrone Bahn. Hier muss die Drehung der Bahnebene genau 360° pro Jahr betragen, so dass die Bahnebene immer in einem festen Winkel zur Sonne–Erde-Achse steht. Mit einer für einen Erderkundungssatelliten typischen kreisförmigen Bahn ($\varepsilon = 0$) mit großer Inklination (z.B. $i = 80^\circ$) erhalten wir aus (2.45) als Bahnhöhe für die gewünschte sonnensynchrone Bahn $(a/R_E)^{7/2} = 1.76$ und damit $H = a - R_E = 1100$ km.

Apsidendrehung

Die mit der Abplattung der Erde verbundene Verzerrung des Gravitationsfelds bewirkt neben der Drehung der Bahnebene auch eine Drehung der Bahnellipse in dieser Ebene (Apsidendrehung) mit

$$\frac{\Delta\omega}{\text{Tag}} = \frac{5^\circ(5 \cos^2 i - 1)}{\left(\frac{a}{R_E}\right)^{7/2}} (1 - \varepsilon^2). \quad (2.46)$$

Die Apsidendrehung verschwindet, wenn der Zähler Null wird, d.h. für Bahnen mit einer Inklination i von 63.4° , vgl. Abb. 2.19. Daher haben die Bahnen der Molniya-Satelliten genau diese Inklination. Ansonsten würde sich innerhalb von Tagen bis Wochen das Apogäum von der Nord- auf die Südhemisphäre drehen und die Satelliten würden nicht mehr die meiste Zeit eines Umlaufs als Relaisstationen über der FSU stehen.

Für Inklinationen größer oder kleiner als 63.4° erfolgt die Apsidendrehung in unterschiedliche Richtungen, mit zunehmendem Abstand wird sie immer geringer.

2.5.4 Zusammenfassung

Die wichtigsten Bahnstörungen für Satelliten sind die Abbremsung durch Reibung, die für niedrig fliegende Satelliten zum Absturz führen kann, und die durch die Abplattung der Erde bedingte Abweichung des Gravitationsfeldes von der Kugelsymmetrie, die zu einer Drehung der Bahnebene (Präzession) und zu einer Drehung der Bahn in der Ebene (Apsidendrehung) führt. Bahnstörungen können messtechnisch eingesetzt werden, so liefert die Abbremsung

von Satelliten Informationen über die anderweitig nicht messbare Dichte bzw. Temperatur der Hochatmosphäre. Die durch die Abplattung der Erde bewirkten Effekte lassen sich ebenfalls nutzen: die Präzession kann zur Erzeugung sonnensynchroner Orbits verwendet werden, die Apsidendrehung zur Stabilisierung einer sehr exzentrischen Flugbahn wie der der Molniya-Satelliten. Die Auswertung von Bahnstörungen im Rahmen der Satellitengeodäsie wird ausführlich diskutiert in Schneider [104].

2.6 Einbringen des Satelliten in seine Bahn

Das Einbringen eines Satelliten in seine Umlaufbahn setzt sich meistens aus mehreren Prozessen zusammen: dem Einbringen des Satelliten in eine niedrige kreisförmige Umlaufbahn mit Hilfe einer Rakete oder des Space Shuttle und dem anschließenden, auf Grund der Begrenzung der mitgeführten Treibstoffvorräte energetisch möglichst günstig zu gestaltenden Transfer in sein Ziel-Orbit (Hohmann-Bahnen). Transferbahnen werden nicht nur für die Positionierung von Satelliten in einer Erdumlaufbahn von Interesse sondern ebenso für Missionen zu anderen Planeten bzw., nahe liegender, zum Mond.

2.6.1 Allgemeine Probleme

Der Start eines Satelliten ist, insbesondere wenn die zur Verfügung stehende Energie optimal ausgenutzt werden soll, kein trivialer Vorgang. Die drei wichtigsten Probleme sind:

- die Eigenbewegung der Erde: die Erdrotation sowie die Rotation der Erde um die Sonne sind mit Eigengeschwindigkeiten von ca. 0.5 km/s bzw. 30 km/s verbunden. Diese relativ hohen Geschwindigkeiten können unterstützend verwendet werden – allerdings nur bei Start in Richtung dieser Bewegungen. Bei Start in die entgegengesetzte Richtung müssen diese Geschwindigkeiten zusätzlich überwunden werden.
- der große Luftwiderstand in niedrigen Höhen verhindert große Startgeschwindigkeiten. Aber selbst bei vernachlässigbarem Luftwiderstand könnte man keine zu großen Anfangsgeschwindigkeiten verwenden, da dann die initiale Beschleunigung zu groß wäre und damit starke Belastungen für die mechanische Struktur von Satelliten und Instrumenten mit sich bringen würde, von der Belastung der Astronauten auf bemannten Missionen ganz zu schweigen.
- der zum Antrieb erforderliche Brennstoff muss mit angehoben werden, d.h. also auch mit beschleunigt werden.

2.6.2 Raketengleichung

Die Notwendigkeit, den Treibstoff mit anzuheben, führt auf die Raketengleichung. Darin wird die Beschleunigung einer veränderlichen Masse (die Rakete wird durch das Verbrennen der Treibgase leichter) unter Einwirkung einer konstanten Kraft (die Verbrennungsgase strömen mit konstanter Rate und konstantem Impuls aus) betrachtet.

Gravitationsfrei und luftleer

Der einfachste Ansatz ist eine Rakete im gravitationsfreien und luftleeren Raum. Formal ist nur die Impulserhaltung zu betrachten mit den Impulsen $m(t)v(t)$ der Rakete und dem Impuls $(v(t) - v_0)\beta dt$ der ausgestoßenen Treibgase mit v_0 als der Ausströmgeschwindigkeit und β als der Ausströmrates: $dm = \beta dt$. Die Rakete erfährt in einem Zeitintervall dt beim Ausströmen des Treibgases eine Änderung dp des Impulses gegenüber ihrem Impuls p am Anfang des Zeitintervalls:

$$p + dp = (m - \beta dt)(v + dv) \quad \text{bzw.} \quad p = (m - \beta dt)(v + dv) + \beta dt(v - v_0). \quad (2.47)$$

Ausmultiplizieren und Vernachlässigung des Gliedes $dm dv$ liefert

$$dv = -v_0\beta dt/m \quad (2.48)$$

Geschwindigkeits- verhältnis v_B/v_0	Massen- verhältnis m_0/m_B	Wirkungs- grad η_A
0.001	1.001	0.001
0.01	1.010	0.010
0.1	1.105	0.095
0.5	1.65	0.385
1.	2.72	0.582
1.594	4.93	0.647
2.	7.4	0.625
3.	20.1	0.524
4.	55.	0.299
5.	149.	0.169
6.	403.	0.089
8.	3000	0.021
10.	22 000	0.0045
57.	$6 \cdot 10^{24}$	$0.5 \cdot 10^{-21}$

Tabelle 2.5: Äußerer Wirkungsgrad einer Rakete

und damit für die *Schubkraft* F , die einer Rakete der momentanen Masse m die Beschleunigung dv/dt erteilt,

$$F = v_0 \beta . \quad (2.49)$$

Ist v_0 während der gesamten Brenndauer t_B konstant, so lässt sich (2.48) einfach integrieren und wir erhalten als *Raketenformel*

$$\Delta v = v_0 \ln \frac{m_0}{m_B} \quad (2.50)$$

mit m_0 als der Gesamtmasse der Rakete zum Zeitpunkt $t = 0$ und m_B als der Raketenmasse bei Brennschluss. Δv ist der Geschwindigkeitszuwachs der Rakete bzw. bei Beschleunigung aus der Ruhe die Geschwindigkeit am Ende der Beschleunigungsphase.

Der Geschwindigkeitszuwachs Δv kann formal nahezu beliebig gesteigert werden, wenn das Verhältnis m_0/m_B groß genug wird. Allerdings ist dies zum einen aus technischen Gründen (den Treibstoff umschließende Struktur) nicht einfach, zum anderen wird der Wirkungsgrad für große Massenverhältnisse sehr klein, vgl. Tabelle 2.5.

Als Wirkungsgrad einer Rakete wird das Verhältnis ihrer kinetischen Energie bei Brennschluss zur insgesamt im Treibstoff enthaltenen Energie definiert:

$$\eta = \frac{E_{\text{kin},B}}{E_{\text{kin},\infty}} = \frac{m_B v_B^2}{m_T v_\infty^2} \quad (2.51)$$

mit v_∞ als der Geschwindigkeit, die die ausströmenden Treibgase haben würden, wenn ihr gesamter Wärmeinhalt in die kinetische Energie des ausgestoßenen Gases verwandelt wäre. Würde man eine derartige ideale Rakete auf dem Prüfstand abbrennen, so würde keine Energie an die fest verankerte Rakete übertragen sondern alle Energie in die Ausströmung gehen. Bei der realen Rakete ist die Ausströmgeschwindigkeit v_0 geringer als v_∞ , so dass die kinetische Energie der Ausströmung gegeben ist als $E_{\text{kin},T} = m_T v_0^2/2$. Der Wirkungsgrad (2.51) wird daher aufgespalten gemäß

$$\eta = \eta_I \eta_A \quad \text{mit} \quad \eta_I = \frac{v_0^2}{v_\infty^2} \quad \text{und} \quad \eta_A = \frac{m_B v_B^2}{m_T v_\infty^2} \quad (2.52)$$

als dem inneren und äußeren Wirkungsgrad. Der innere Wirkungsgrad gibt ein Maß dafür, welcher Teil der in den heißen Brennkammerngasen enthaltenen Energie in gerichtete kinetische Energie der austretenden Gase umgewandelt wird. Er liegt zwischen 0.36 und 0.64 [25]. Der äußere Wirkungsgrad gibt den Bruchteil der in den ausströmenden Raketengasen steckenden

Energie, der in kinetische Energie der Rakete umgewandelt werden kann. Der Wirkungsgrad ist Null, wenn die Rakete auf dem Prüfstand festgehalten wird, er würde eins, wenn in jeder Flugphase die Ausströmgeschwindigkeit ihrem Betrage nach gleich der momentanen Flugeschwindigkeit v der Rakete wäre. In diesem Fall würde sich der Gasstrahl bezüglich eines raumfesten Systems in Ruhe befinden. Im realen Fall ist der Wirkungsgrad geringer, ein Teil der Energie steckt dann in der Geschwindigkeit, mit der die Rückstoss-gase durch den Raum fliegen. Für konstante Ausströmgeschwindigkeit können wir die Raketengleichung (2.50) einsetzen und erhalten für den äußeren Wirkungsgrad

$$\eta_A = \frac{(v_B/v_0)^2}{\exp(v_B/v_0) - 1} = \frac{\left(\ln \frac{m_0}{m_B}\right)^2}{\frac{m_0}{m_B} - 1}. \quad (2.53)$$

Tabelle 2.5 fasst die äußeren Wirkungsgrade für verschiedene Geschwindigkeits- bzw. Massenverhältnisse zusammen. Der optimale Wirkungsgrad von 0.647 ergibt sich für ein Massenverhältnis von 4.93 (vernünftige Annahme) und ein Geschwindigkeitsverhältnis von 1.594.

Luftleer, aber Gravitation

Die Rakete im luftleeren und gravitationsfreien Raum kann vielleicht als Annäherung für Booster bei der Durchführung von Bahnmanövern betrachtet werden, die eigentliche Anwendung einer Rakete, den Start vom Erdboden aus, beschreibt sie jedoch nicht. Um diesen zu beschreiben müssen wir in der Bewegungsgleichung zusätzlich zur Schubkraft die Gravitation berücksichtigen und erhalten mit (2.49)

$$F = v_0\beta - g(h) \quad (2.54)$$

für den senkrechten Start, beim schiefen Schuss muss die Betrachtung vektoriell durchgeführt werden. Integration liefert (für konstantes v_0 und unter Vernachlässigung der Höhenabhängigkeit von g)

$$\Delta v = v_0 \ln \frac{m_0}{m_B} - gt_B \quad (2.55)$$

wobei sich die Zeit t_B aus der Masse des Treibstoffs und der Ausströmrates bestimmen lässt.

2.6.3 Mehrstufenprinzip

Bei einer einfachen Rakete ist das Massenverhältnis der begrenzende Faktor für die erreichbaren Geschwindigkeiten, eine optimale Ausnutzung des Wirkungsgrads würde ein Geschwindigkeitsverhältnis von ca. 1.6 liefern (vgl. Tabelle 2.5). Mit einer Ausströmgeschwindigkeit von 3 km/s ergibt sich dann eine Endgeschwindigkeit von 4.8 km/s, noch deutlich unter der ersten kosmischen Geschwindigkeit und damit viel zu gering, um einen Satelliten in ein Orbit zu bringen.

Mit einer mehrstufigen Rakete kann man dieses Problem leicht umgehen. Anschaulich ist das einsichtig: eine mehrstufige Rakete zündet eine erste Raketstufe. Ist diese ausgebrannt, so wird ihre Hülle abgeworfen. Damit reduziert sich die von der zweiten Stufe zu beschleunigende Masse. Dieser Vorgang wiederholt sich mit der zweiten Stufe.

Tabelle 2.6 gibt eine Illustration des Mehrstufenprinzips. Im oberen Teil ist eine Einstufenrakete betrachtet mit einer Startmasse m_o von 50.68 t. Von der Gesamtmasse entfallen 42.2 t auf Treibstoff, 8.44 t auf die Rakete selbst und 40 kg auf die Nutzlast. Die Ausströmgeschwindigkeit beträgt 2.7 km/s. Damit ergibt sich nach der Raketengleichung (2.50) eine Geschwindigkeit von 4.8 km/s bei Brennschluss der Rakete, d.h. die Rakete erreicht nicht die für eine Kreisbahn notwendige Geschwindigkeit. Im unteren Teil der Tabelle stehen für Brennstoff, Nutzlast und Raketensstruktur die gleichen Massen zur Verfügung. Diese Massen werden jedoch auf drei Raketestufen verteilt, wobei die erste Stufe alleine 95% der Masse der vorher betrachteten Einstufenrakete hat. Da in dieser ersten Stufe aber etwas weniger Treibstoffmasse vorhanden ist und die Restmasse der Rakete nach Brennschluss der

<i>Einstufenrakete:</i>		
Nutzlast		$m_S = 0.04 \text{ t}$
Treibstoff		$m_T = 42 - 20 \text{ t}$
Rakete		$m_R = 8.44 \text{ t}$
→ Masse beim Start		$m_0 = 50.68 \text{ t}$
Masse bei Brennschluss		$m_B = 8.48 \text{ t}$
→ Massenverhältnis		$m_0/m_B = 5.98$
Ausströmgeschwindigkeit		$v_0 = 2.7 \text{ km/s}$
→ Endgeschwindigkeit (2.50)		$v_R = 4.8 \text{ km/s}$
<i>Mehrstufenrakete:</i>		
Nutzlast		$m_S = 0.04 \text{ t}$
3. Stufe Rakete		$m_{R3} = 0.04 \text{ t}$
Treibstoff		$m_{T3} = 0.20 \text{ t}$
2. Stufe Rakete		$m_{R2} = 0.40 \text{ t}$
Treibstoff		$m_{T2} = 2.00 \text{ t}$
1. Stufe Rakete		$m_{R1} = 8.00 \text{ t}$
Treibstoff		$m_{T1} = 40 \text{ t}$
→ Gesamtmasse beim Start		$m_0 = 50.68 \text{ t}$
davon Gesamtmasse Rakete		$m_R = 8.44 \text{ t}$
davon Gesamtmasse Treibstoff		$m_T = 42.20 \text{ t}$
1. Stufe Masse beim Start = Gesamtmasse		$m_{01} = m_0 = 42.20 \text{ t}$
Masse bei Brennschluss		$m_{B1} = 10.68 \text{ t}$
Geschwindigkeitszuwachs		$\Delta v_1 = 4.21 \text{ km/s}$
Brennschlussgeschwindigkeit der 1. Stufe		$v_1 = 4.21 \text{ km/s}$
2. Stufe Masse bei Zündung der zweiten Stufe		$m_{02} = 2.68 \text{ t}$
Masse bei Brennschluss der zweiten Stufe		$m_{B2} = 0.68 \text{ t}$
Geschwindigkeitszuwachs		$\Delta v_2 = 3.71 \text{ km/s}$
Brennschlussgeschwindigkeit der zweiten Stufe $v_1 + \Delta v_2$		$v_2 = 7.92 \text{ km/s}$
3. Stufe Masse bei Zündung der dritten Stufe		$m_{03} = 0.28 \text{ t}$
Masse bei Brennschluss der dritten Stufe		$m_{03} = 0.28 \text{ t}$
Geschwindigkeitszuwachs		$\Delta v_3 = 3.39 \text{ km/s}$
Brennschlussgeschwindigkeit der dritten Stufe $v_1 + \Delta v_2 + \Delta v_3$..		$v_3 = v_B = 11.31 \text{ km/s}$

Tabelle 2.6: Vergleich Ein- und Dreistufenrakete, nach [25]

ersten Stufe auch noch die Massen der anderen beiden Stufen umfasst, ergibt sich eine Geschwindigkeitszunahme von nur 4.21 km/s gegenüber den 4.8 km/s bei der Einstufenrakete. Die zweite Stufe, nun der Restmasse der ausgebrannten ersten Stufe entledigt, wird mit dieser Anfangsgeschwindigkeit gezündet und erreicht durch den Schub der zweiten Stufe eine Geschwindigkeitszunahme um weitere 3.71 km/s, entsprechend einer Gesamtgeschwindigkeit von 7.92 km/s, d.h. die beiden unteren Stufen alleine könnten die Nutzlast zusammen mit der dritten Stufe bereits in eine Erdumlaufbahn bringen. Nach Absprengen der ausgebrannten zweiten Stufe wird die dritte Stufe gezündet. Nach deren Brennschluss ist eine Endgeschwindigkeit von 11.31 km/s erreicht. Während die Einstufenrakete bei gleicher Masse nicht einmal in der Lage gewesen wäre, einen Satelliten in ein Erdorbit zu befördern, kann mit der Dreistufenrakete eine Geschwindigkeit oberhalb der Fluchtgeschwindigkeit von 11.2 km/s erreicht werden, d.h. es können sogar planetare Missionen gestartet werden.

Mehrstufenrakete formal

Um zu einer formalen Beschreibung der Endgeschwindigkeit bei einer Mehrstufenrakete zu gelangen, betrachten wir im Folgenden eine Zweistufen-Rakete. Die erste Stufe hat die Masse $m_1 = m_{T1} + m_{R1}$ mit m_{T1} als der Masse des Treibstoffes und m_{R1} als der Restmasse der Struktur der ersten Stufe. Analog bezeichnet der Index 2 die Massen der zweiten Stufe

(wobei bei einer vorhandenen Nutzlast deren Masse in m_{R2} mit enthalten ist). Nach der Raketengleichung (2.50) ergibt sich die Geschwindigkeit Δv_1 nach Brennschluss der ersten Stufe zu:

$$\Delta v_1 = v_{01} \ln \frac{m_{T1} + m_{R1} + m_{T2} + m_{R2}}{m_{R1} + m_{T2} + m_{R2}}. \quad (2.56)$$

Die zweite Stufe m_2 wird nun von der leergebrannten ersten Stufe abgetrennt (Verlust der Masse m_{R1}). Nach Brennschluss der zweiten Stufe ergibt sich ein Geschwindigkeitszuwachs von

$$\Delta v_2 = v_{02} \ln \frac{m_{T2} + m_{R2}}{m_{R2}}. \quad (2.57)$$

Da die zweite Stufe bei ihrer Zündung bereits die durch (2.56) gegebene Geschwindigkeit Δv_1 erreicht hatte, ergibt sich für die Gesamtgeschwindigkeit nach Brennschluss der zweiten Stufe $v_{B2} = \Delta v_1 + \Delta v_2$. Vereinfachend nehmen wir an, dass bei beiden Stufen die Ausströmgeschwindigkeiten gleich sind: $v_{01} = v_{02}$. Dann ergibt sich

$$v_{B2} = v_0 \ln \frac{(m_{T1} + m_{R1} + m_{T2} + m_{R2})(m_{T2} + m_{R2})}{(m_{R1} + m_{T2} + m_{R2})m_{R2}}. \quad (2.58)$$

Erweitert man den Bruch im Logarithmus mit $(m_{R1} + m_{R2})$, so ergibt sich nach Ausklammern, Ausmultiplizieren und Umformen

$$v_{B2} = v_0 \ln \frac{m_1 + m_2}{(m_{R1} + m_{R2})} + v_0 \ln \left(1 + \frac{m_{R1} m_{R2}}{m_{R2} (m_{R1} + m_{R2} + m_{T2})} \right). \quad (2.59)$$

Dabei stellt der erste Term auf der rechten Seite die Endgeschwindigkeit dar, die man nach der Raketenformel (2.50) für eine einstufige Rakete mit derselben Startmasse $m_1 + m_2$ und derselben Treibstoffmasse $m_{T1} + m_{T2}$ wie das gesamte Zweistufenaggregat erhalten hätte. Der zweite Term stellt den Geschwindigkeitsgewinn durch die Verwendung einer zweiten Stufe dar, dieser Term ist stets positiv.

Setzt man insgesamt n Stufen aufeinander und trennt jeweils die unterste Stufe nach Brennschluss ab, so ergibt sich für die Endgeschwindigkeit der letzten Stufe einschließlich der Nutzlast: $v_B = \sum \Delta v_i$. Für alle Raketenstufen gilt die Raketenformel (2.50). Unter der Voraussetzung, dass die Ausströmgeschwindigkeit v_0 für alle Stufen gleich ist, ergibt sich

$$v_B = v_0 \ln \left(\frac{m_{01}}{m_{B1}} \frac{m_{02}}{m_{B2}} \cdots \frac{m_{0n}}{m_{Bn}} \right). \quad (2.60)$$

mit den Größen m_{0i} und m_{Bi} als den Gesamtmassen vor Zünden bzw. nach Brennschluss der i -ten Stufe, gegeben durch

$$\begin{aligned} m_{01} &= m_N + \sum_{i=2}^n (m_{Ri} + m_{Ti}) + (m_{R1} + m_{T1}), \\ m_{B1} &= m_N + \sum_{i=2}^n (m_{Ri} + m_{Ti}) + m_{R1}, \\ m_{02} &= m_N + \sum_{i=3}^n (m_{Ri} + m_{Ti}) + (m_{R2} + m_{T2}), \\ &\vdots \\ m_{Bn} &= m_N + m_{Rn}. \end{aligned} \quad (2.61)$$

Dabei ist m_N die Masse der Nutzlast. Die aus diesen Massen gebildete, als Argument im Logarithmus von (2.59) auftretende Produktkette

$$\frac{M_0}{M_B} = \prod_{i=1}^n \left(\frac{m_{0i}}{m_{Bi}} \right) \quad (2.62)$$

wird als das totale Massenverhältnis bezeichnet. Dann lässt sich (2.59) analog zur Raketenformel (2.50) auch in der Form

$$v_B = v_0 \ln \frac{M_0}{M_B} \quad (2.63)$$

schreiben.

Die Raketengleichung (2.50) erlaubt nur eine ungefähre Abschätzung der Geschwindigkeit bei Brennschluss der Rakete. In der Realität müssen die Reibung in der Erdatmosphäre und die (mit der Höhe abnehmende) Gravitationskraft berücksichtigt werden. Allerdings gibt es auch eine ‘Starthilfe’: die Bewegung des Startplatzes verschafft der Rakete in einem raumfesten System bereits eine Anfangsgeschwindigkeit. Für planetare Missionen ist die Starthilfe die Rotation der Erde um die Sonne. Die Umlaufgeschwindigkeit beträgt $v_U=29.8$ km/s, die am Ort der Erde benötigte Entweichgeschwindigkeit aus dem Sonnensystem v_e^s 46.4 km/s. Nutzt man die bereits durch die Rotation der Erde um die Sonne vorgegebene Geschwindigkeit, so reduziert sich die dem Satelliten bzw. der Rakete zusätzlich zu erteilende Geschwindigkeit auf $\Delta v = v_e^s - v_U = 16.6$ km/s. Dies gilt natürlich wenn der Start in der Bewegungsrichtung der Erde erfolgt. Ein Start in die Gegenrichtung ist praktisch unmöglich, da dann die Eigengeschwindigkeit der Erde zusätzlich aufgehoben werden müsste, der Rakete also eine Gesamtgeschwindigkeit von 76.2 km/s zu erteilen wäre. Für einen Start von Erdsatelliten ist der Effekt der Eigenrotation der Erde geringer, da die Kreisbahngeschwindigkeit 7.8 km/s beträgt, die Umlaufgeschwindigkeit am Erdboden aber nur 0.46 km/s.

2.6.4 Transferbahnen (Hohmann-Bahnen)

Eine normale Rakete besitzt keine ausreichende Schubkraft, um z.B. einen Kommunikationssatelliten in ein geostationäres Orbit zu bringen. Wird mit dem Space-Shuttle gestartet, so kann bei sehr leichten Satelliten nur eine Maximalhöhe von ca. 1 000 km erreicht werden, realistischere Höhen liegen eher zwischen 400 und 600 km. Um den Satelliten dann in ein höheres Orbit zu überführen, benötigt man zusätzlichen Schub. Wie dieses Verfahren vom energetischen Standpunkt optimiert werden kann, wird durch die Hohmann-Bahnen beschrieben. Andere Transferbahnen existieren ebenfalls, diese sind jedoch wesentlich energieaufwendiger.

Hohmann hat die nach ihm benannten Bahnen bereits vor dem 1. Weltkrieg unter Anwendung der Gesetze der klassischen Himmelmechanik berechnet. Sein Ziel war zu ermitteln, wie groß, wie schwer und wie leistungsfähig ein raketentriebenes Raumschiff sein müsste, um auf Sonnen umrundenden Ellipsenbahnen bei geringstem Energieaufwand zu anderen Planeten zu gelangen. Hohmann veröffentlichte diese Untersuchungen 1925 unter dem Titel ‘Die Erreichbarkeit der Himmelskörper’ (Oldenburg, München) mit den Kapiteln: Loslösung von der Erde; Rückkehr zur Erde; Freie Fahrt im Raum; Umfahrung anderer Himmelskörper; und: Landung auf anderen Himmelskörpern. Wernher von Braun kommentierte dazu: ‘Seine Untersuchungen schlossen nicht nur das primäre Problem des Verlassens des Schwerefeldes der Erde ein, sondern auch die delikate Aufgabe des Wiedereintritts in die Erdatmosphäre mit super-orbitalen Rückkehrgeschwindigkeiten’. Die wichtigsten theoretischen Grundlagen der Manövrierung von Raumsonden wurden also bereits gelegt lange bevor sich die ersten Möglichkeiten zur technischen Realisierung solcher ‘hochfliegenden’ Pläne abzeichneten.

Hohmann-Bahnen

Hohmann-Bahnen bestimmen den energetisch günstigsten Übergang zwischen zwei Orbits. Hierbei kann es sich um den Übergang von einer niedrigen Erdumlaufbahn in ein geostationäres Orbit handeln oder um den Übergang von einer Bahn um die Erde auf eine um einen anderen Planeten. Als Näherung wollen wir das Problem soweit vereinfachen, dass die Anziehungskraft auf den Satelliten nur von einem Zentralkörper ausgeht (also z.B. zwei Satellitenbahnen um die Erde oder der Übergang von der Erd- auf die Marsbahn ohne Einschwenken in eine Umlaufbahn um diesen Planeten, so dass dessen Anziehungskraft vernachlässigt werden kann). Zusätzlich wird angenommen, dass es sich um kreisförmige Bahnen handelt,

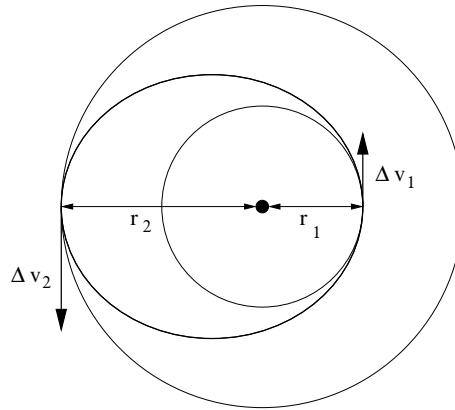


Abbildung 2.20: Hohmann-Bahn als Übergangselipse zwischen zwei kreisförmigen Orbits

die alle in einer Ebene liegen, was z.B. für Bahnen zwischen den Planeten keine bedeutende Einschränkung darstellt.

Eine Möglichkeit des Überganges zwischen zwei kreisförmigen Bahnen mit den Radien r_1 und r_2 ist eine elliptische Trajektorie, vgl. Abb. 2.20. Diese Übergangselipse wird als Hohmann-Ellipse oder Hohmann-Bahn bezeichnet. Ihre große Halbachse a ergibt sich zu $a = \frac{1}{2}(r_1 + r_2)$, da die Bahn in ihrem Perihel bzw. Perigäum die innere, in ihrem Aphel bzw. Apogäum die äußere Kreisbahn berührt. Nach der Binet'schen Gleichung (2.30) ergibt sich dann für jeden Punkt der Bahn die Geschwindigkeit zu

$$v^2 = 2\mu \cdot \left(\frac{1}{r} - \frac{1}{r_1 + r_2} \right). \quad (2.64)$$

Damit lassen sich die Geschwindigkeiten im Perihel ($r = r_1$) und Aphel ($r = r_2$) bestimmen zu

$$v_P = \sqrt{\frac{\mu}{r_1}} \sqrt{\frac{2}{r_1/r_2 + 1}} \quad \text{und} \quad v_A = \sqrt{\frac{\mu}{r_2}} \sqrt{\frac{2}{1 + r_1/r_2}}. \quad (2.65)$$

Erfolgt der Start nun von der inneren Kreisbahn aus in Richtung der Umlaufbewegung des Raumfahrzeuges, so hat der Körper bereits vor dem Start eine Geschwindigkeit, die der Kreisbahngeschwindigkeit $v_{K1} = \sqrt{\mu/r_1}$ auf diesem Orbit entspricht. Zum Übergang auf die Hohmann-Bahn ist dann eine Geschwindigkeitsänderung $\Delta v_1 = v_1 - v_{K1}$ erforderlich, wobei gemäß der Definition der Bahn Δv_1 tangential zur Kreisbahn ist. Die Geschwindigkeitsänderung durch das Feuern der Schubdüsen erfolge instantan, d.h. in einer Zeit, die klein gegenüber der Umlaufzeit ist. Dann erhält man die Geschwindigkeitsänderung zu

$$\Delta v_1 = \sqrt{\frac{\mu}{r_1}} \left(\sqrt{\frac{2}{r_1/r_2 + 1}} - 1 \right). \quad (2.66)$$

Mit diesem Geschwindigkeitszuwachs kann das Raumfahrzeug die elliptische Bahn bis zur Kreisbahn mit r_2 durchlaufen. Da seine Geschwindigkeit an diesem Punkt kleiner ist als die Kreisbahngeschwindigkeit beim Radius r_2 , würde das Raumfahrzeug dann auf der zweiten Hälfte der Ellipse wieder nach innen laufen. Daher muss in diesem Punkt nochmals ein Geschwindigkeitszuwachs erreicht werden, der gegeben ist durch

$$\Delta v_2 = \sqrt{\frac{\mu}{r_2}} \left(1 - \sqrt{\frac{2}{1 + r_1/r_2}} \right). \quad (2.67)$$

Der gesamte Geschwindigkeitszuwachs ergibt sich als die Summe der beiden zu

$$\Delta v = v_{K1} \left(1 - \sqrt{\frac{r_1}{r_2}} \right) \left[\sqrt{\frac{2}{r_1/r_2 + 1}} \left(1 + \sqrt{\frac{r_1}{r_2}} \right) - 1 \right]. \quad (2.68)$$

Planet	Flugzeit/Jahre
Merkur	0.29
Venus	0.40
Mars	0.71
Jupiter	2.73
Saturn	6.05
Uranus	16.1
Neptun	30.6
Pluto	45.5

Tabelle 2.7: Flugzeiten für Hohmann-Übergänge zu anderen Planeten

Diese Gleichung ist plausibel: Betrachtet man den Fall $r_1 \rightarrow r_2$, so geht Δv gegen Null, d.h. um auf derselben Kreisbahn zu bleiben, ist keine Geschwindigkeitsänderung nötig. Umgekehrt ergibt sich für $r_2 \rightarrow \infty$ ein $\Delta v \rightarrow v_{K1}(\sqrt{2}-1)$. Das ist aber gerade die Fluchtgeschwindigkeit eines Körpers, der sich in einer Höhe r_1 befindet.

Da wir wieder annehmen, dass die Geschwindigkeitsänderungen innerhalb vernachlässigbar kleiner Zeitintervalle stattfinden, können wir für diese kurze Dauer die Gravitation vernachlässigen und die benötigte Treibstoffmenge nach der Raketenformel (2.50) abschätzen.

Will man von der äußeren Kreisbahn auf die innere gelangen, so muss der Raumflugkörper jeweils um eine Geschwindigkeit Δv_2 und Δv_1 verzögert werden. Dabei wird exakt die gleiche Energie benötigt, wie beim umgekehrten Übergang von der inneren Bahn in die äußere, obwohl der Übergang hier von einer Bahn größerer zu einer Bahn geringerer Gesamtenergie erfolgt. Daher ist es sinnvoll, zur Kennzeichnung des Gesamtenergieverbrauchs eine charakteristische Geschwindigkeitsänderung Δv_{char} einzuführen mit

$$\Delta v_{\text{char}} = |\Delta v_1| + |\Delta v_2|. \quad (2.69)$$

Die Flugdauer zwischen zwei Orbits ist gleich der halben Umlaufzeit auf der Hohmann-Ellipse. Durch Einsetzen der Halbachse a in das dritte Kepler'sche Gesetz ergibt sich

$$\Delta t = \frac{P_1}{2} \cdot \sqrt{\left(\frac{1+r_2/r_1}{2}\right)^3}. \quad (2.70)$$

Darin ist $P_1 = 2\pi\sqrt{r_1^3/\mu}$ die Umlaufperiode auf der inneren Bahn, für die Erde ist diese $P_1 = 1$ Jahr. Für Hohmann-Übergänge zu den anderen Planeten ergeben sich die in Tabelle 2.7 gegebenen Zeiten. Die Flugzeit zum Neptun beträgt z.B. 30 Jahre. Das ist wesentlich länger als die 11 Jahre, die Voyager für diese Strecke benötigt hat, da Hohmann-Bahnen nur für ein Einschwenken auf das andere Orbit entscheidend sind, um einfach an einem Planeten vorbeizufliegen, muss nur die Geschwindigkeit ausreichend groß sein (bei Voyager Entweichgeschwindigkeit aus dem Sonnensystem). Hier deutet sich bereits eine Möglichkeit an, einen schnelleren Übergang zu erreichen, der allerdings energetisch aufwendiger ist.

Andere Transferbahnen

So lange das Verhältnis r_1/r_2 von Ziel- zu Ausgangsradius kleiner als 12 ist, haben Hohmann-Bahnen den Vorteil, energetisch am günstigsten zu sein. Dieses Verhältnis mag groß erscheinen, bei planetaren Übergängen ließe sich damit z.B. ein Orbit hinter der Saturnbahn erreichen, aber der Übergang benötigt sehr viel Zeit. So würde selbst zu unserem nächsten Nachbarn, Mars, der Flug noch mehr als 8 Monate dauern, zum Saturn wären es gar 6 Jahre. Der zeitintensivste Teil der Hohmann-Bahn ist die (nach dem zweiten Kepler'schen Gesetz) langsame Annäherung an das Apogäum der Übergangsellipse, das gleichzeitig der Ausgangspunkt für die äußere Kreisbahn ist. Ein schnellerer Übergang sollte daher gerade an dieser Stelle größere Geschwindigkeiten haben. Dazu müsste man dem Raumfahrzeug am Startpunkt der Übergangsbahn auf der inneren Kreisbahn einen stärkeren Schub erteilen (und damit mehr Treibstoff verbrauchen), als er für eine Hohmann-Bahn erforderlich ist. Dann

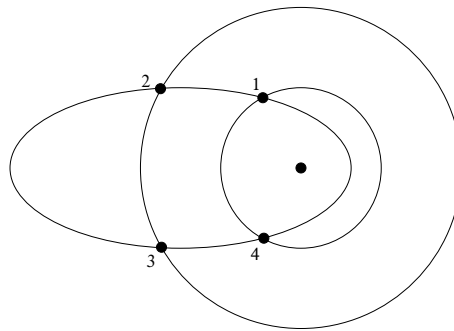


Abbildung 2.21: Übergangsbahn aus Ellipsenbögen

liegt das Apogäum der Übergangsellipse hinter der angestrebten äußeren Bahn. Also muss am Ort des äußeren Orbits relativ stark abgebremst werden, um auf die gewünschte Umlaufbahn einzuschwenken. Die Geschwindigkeitsänderung in diesem Punkt ist dabei nicht, wie beim Hohmann-Übergang, tangential zur Bahn. Die Zeit zum Erreichen des äußeren Orbits wird bei dieser Art von Übergängen um so geringer, je exzentrischer die Übergangsellipse ist. Allerdings steigt die benötigte Treibstoffmenge an, lediglich in bestimmten Abstandsverhältnissen gibt es Einzelfälle, in denen dieser Übergang energetisch günstiger ist als eine Hohmann-Bahn.

So wie beim Übergang vom Transfer-Orbit zur äußeren Bahn eine nicht-tangentiale Geschwindigkeitsänderung erfolgt, kann man auch beim Übergang von der inneren auf die Übergangsbahn eine nicht-tangentiale Geschwindigkeitsänderung vornehmen. Dadurch verkürzt sich die Länge der Übergangsbahn nochmals beträchtlich, vgl. Abb. 2.21.

Für große Verhältnisse von r_1/r_2 oberhalb 11.94 ist die Hohmann-Bahn nicht mehr verbrauchsoptimal, dann erfolgt ein verbrauchsoptimaler Übergang mit Hilfe von drei Impulsen [102]. Dabei erhält der Raumflugkörper anfangs einen zu starken Impuls, der ihn in eine extrem exzentrische Übergangsellipse mit einem Apogäum hinter der gewünschten Zielbahn befördert. Im Apogäum dieser Übergangsbahn wird nochmals ein Vorwärtsschub erteilt, der die Ellipse aufweitet und zu einem Perigäum auf der Bahn des Zielplaneten führt. In diesem Perigäum wird Gegenschub gegeben, damit sich die Ellipsenbahn zu einer Kreisbahn verkleinert.

Bisher haben wir immer kurze Schubimpulse betrachtet. Ein völlig anderes Konzept ergibt sich jedoch, wenn man eine kontinuierliche Beschleunigung annimmt. Ionenraketeten [25, 75, 100] z.B. sollten während der gesamten Reise mit kontinuierlichem Schub arbeiten. Dann ergibt sich eine nach außen immer weiter werdende Spiralbahn um das Gravitationszentrum. Diese Bahn und für planetare Missionen erforderliche Manöver werden z.B. in [25] diskutiert.

Reise zu anderen Planeten

Im Vorangegangenen sind die Transferbahnen zwischen zwei Satellitenorbits dargestellt worden. Diese Betrachtung ist wesentlich vereinfacht gegenüber der Reise von der Erde zu einem anderen Planeten oder auch nur zum Mond. Beim Übergang von einem Satellitenorbit zu einem anderen müssen nur zwei (bzw. beim Drei-Impulsübergang drei) Geschwindigkeitsänderungen vorgenommen werden, ansonsten ‘fällt’ der Satellit seinen Weg alleine.

Für die Reise von der Erde zu einem anderen Planeten sind jedoch noch eine Anzahl zusätzlicher Punkte zu beachten:

1. Bei Reisen zwischen den Planeten ist der Zentralkörper die Sonne. Bevor wir unsere Raumsonde auf eine Übergangsellipse im Schwerfeld der Sonne schicken können, müssen wir sie erst aus dem Schwerfeld der Erde in ein Orbit heben, das ähnlich dem der Erde um die Sonne läuft. Die erste Stufe zum Erreichen eines anderen Planeten besteht also darin, der Raumsonde die Fluchtgeschwindigkeit zum Verlassen des Anziehungsbereiches

der Erde zu erteilen.

2. Von dieser Bahn aus lässt sich ein einfacher Hohmann-Übergang auf die Bahn des anderen Planeten vornehmen. Hierbei muss jedoch beachtet werden, dass der Planet auch zu der Zeit, an der unsere Raumsonde seine Bahn erreicht, am entsprechenden Ort sein muss. Aus dieser Randbedingung erklärt sich, warum es bei Missionen zu anderen Planeten stets so genannte Startfenster gibt, d.h. Zeitbereiche, während derer ein Start insofern sinnvoll ist, als dass die Raumsonde nicht nur die Bahn des zu besuchenden Planeten, sondern auch den Planeten selbst erreicht.
3. Die Sonde befindet sich jetzt auf der Bahn des Planeten, nicht aber auf einer Bahn um den Planeten herum. Jetzt muss eine Umkehrung des ersten Schritts erfolgen und die Sonde muss auf eine Bahn um den Zielplaneten einschwenken, d.h. sie muss sich aus dem Schwerefeld der Sonne in das des Zielplaneten begeben.
4. Nachdem sich die Sonde nun in einer Bahn um den Zielplaneten befindet, können wir diesen vermessen (z.B. die Radarvermessung der Venus durch Magellan) oder wir können die Sonde bzw. eine Landekapsel der Sonde auf die Planetenoberfläche senden. Für eine weiche Landung stellt die Abbremsung das Hauptproblem dar, hier muss gleichsam ein Startvorgang mit umgekehrter Zeitrichtung durchgeführt werden. Wichtig ist dabei auch der korrekte Eintritt in die Atmosphäre des Planeten, um ein Verglühen zu verhindern (gute Diskussion z.B. in [75]).
5. Die Sonde/Landekapsel hat dem Zielplaneten Proben entnommen, die zur Erde zurück gebracht werden sollen. Damit muss der gesamte Vorgang von (1) bis (4) erneut durchlaufen werden, jetzt jedoch haben Erde und Zielplanet getauschte Rollen. Bei einer Rückführung erweist es sich als ausgesprochen hilfreich, auf das bereits von Hohmann vorgeschlagene und in den Apollo-Mondlandungen verwirklichte Konzept des 'Beibootes' zurück zu greifen. Die Idee dabei ist, nicht die komplette Raumsonde zu landen, sondern nur einen möglichst kleinen Teil der Sonde (Lander). Der Vorteil ist offensichtlich: der Lander benötigt für seinen Start vom Zielplaneten wesentlich weniger Treibstoff als die ganze Raumsonde benötigen würde.

Zusammenfassung

Bei einer Rakete handelt es sich um einen Körper veränderlicher Masse, der durch eine konstante Kraft beschleunigt wird. Die bei dieser Beschleunigung erreichbare Endgeschwindigkeit bestimmt sich aus der Ausströmgeschwindigkeit der Verbrennungsgase und dem Verhältnis der Massen bei Start und Brennschluss (Raketengleichung). Das Mehrstufenprinzip erlaubt es, durch geschickte Verteilung der Massen von Struktur und Treibstoff auf wesentlich höhere Geschwindigkeiten zu beschleunigen, als es mit einer einfachen Rakete möglich wäre, da die folgenden Stufen jeweils nicht mehr die leere Hülle der vorhergehenden Stufen beschleunigen müssen. Energetisch günstige Übergänge zwischen zwei Satelliten- oder Planetenbahnen lassen sich durch Übergangsellipsen (Hohmann-Bahnen) realisieren.

2.7 *Spezielle Bahnen

Dieser Abschnitt ist für die Erdfernerkundung nicht relevant, mag aber vielleicht einige von Ihnen aus technischen Gründen interessieren, insbesondere, da die Bahnen einerseits physikalisch anschaulich sind (Lagrange-Punkt), andererseits mit bekannten Missionen (planetares Billard bei Voyager) verbunden sind.

2.7.1 Lagrange-Punkte

Die Bahn von ISEE-3 ist ein klassisches Beispiel für ein Dreikörperproblem. Die Raumsonde befindet sich am Librationspunkt zwischen Sonne und Erde, d.h. am Ort von ISEE-3 herrscht ein Gleichgewicht zwischen der Anziehung durch die Sonne und der durch die Erde. Praktisch wird diese Bahn durch ein Halo-Orbit um den Lagrange-Punkt herum realisiert: würde

sich die Sonde direkt im Lagrange-Punkt befinden, so wäre aufgrund der starken solaren Radiostrahlung eine Kommunikation nicht möglich. Der Orbit um den Lagrange-Punkt herum dient auch gleichzeitig der Stabilisierung der Sonde.

Formal lässt sich der Lagrange-Punkt scheinbar aus dem Gleichgewicht zwischen der Anziehung der Erde und der der Sonne berechnen. Mit R als dem Abstand zwischen Sonne und Erde ergibt sich der Abstand r dieses Punktes von der Erde wegen

$$F_E = F_\odot \quad \rightarrow \quad \frac{m_s M_E \gamma}{r^2} = \frac{m_s M_\odot \gamma}{(R-r)^2} . \quad (2.71)$$

zu

$$r = \frac{R \sqrt{\frac{M_E}{M_\odot}}}{1 + \sqrt{\frac{M_E}{M_\odot}}} \approx r = 266\,084 \text{ km} . \quad (2.72)$$

Stabil ist dieser Punkt in der so diskutierten Form nicht, da eine kleine Auslenkung den Satelliten stets in Richtung des Körpers stürzen lassen würde, in dessen Richtung er ausgelenkt wurde.

Allerdings ist die obige Betrachtung zum Lagrange-Punkt physikalisch auch nicht korrekt, da wir es mit einem rotierenden Bezugssystem zu tun haben: die Verbindungsachse Erde–Sonne rotiert um den gemeinsamen Massenmittelpunkt, der allerdings nur leicht gegen den Massenmittelpunkt der Sonne versetzt ist. Außerdem müssen wir das Problem als echtes Dreikörperproblem lösen, um auch ein Gefühl für die Stabilität des Lagrange-Punktes zu entwickeln.

Nehmen wir als Bezugssystem ein kartesisches Koordinatensystem mit dem Ursprung im Massenmittelpunkt des Systems Sonne–Erde und der x -Achse entlang der gemeinsamen Verbindungslinie.¹⁴ Zur Herleitung der Bewegungsgleichung seien jetzt die Einheiten von Masse, Entfernung und Zeit so gewählt, dass $\gamma(M+m)$, der konstante Abstand zwischen den beiden schweren Körpern und die Gravitationskonstante γ gleich 1 gesetzt werden. Aus Keplers drittem Gesetz ergibt sich dann, dass die konstante Winkelgeschwindigkeit n der beiden schweren Körper in einem Inertialsystem ebenfalls gleich 1 ist. In diesen Einheiten ist m das Verhältnis der kleineren Masse zur Summe der beiden Massen. Die Abstände der Massen M und m vom Massenschwerpunkt sind dann m und $(1-m)$. Das Bezugssystem sei so gewählt, dass es mit der Winkelgeschwindigkeit $n=1$ um den Massenschwerpunkt in der Bahnebene der beiden schweren Körper rotiert. Dann befinden sich die beiden schweren Massen m und M stets an festen Punkten auf der x -Achse. Die Bewegung einer Testmasse am Punkt (x, y) ist dann gegeben durch die Bewegungsgleichung

$$\begin{aligned} \frac{d^2 x}{dt^2} - 2 \frac{dy}{dt} &= x - \frac{(1-m)(x-m)}{r_1^3} - \frac{m(x-1+m)}{r_2^3} \quad \text{und} \\ \frac{d^2 y}{dt^2} + 2 \frac{dx}{dt} &= y - \frac{(1-m)y}{r_1^3} - \frac{my}{r_2^3} \end{aligned} \quad (2.73)$$

mit $r_1 = \sqrt{(x+m)^2 + y^2}$ und $r_2 = \sqrt{(x-1+m)^2 + y^2}$ als den Abständen der Testmasse von den Massen M und m . Die Bewegungsgleichung ist ähnlich der Bewegungsgleichung eines geladenen Teilchens in einem konstanten homogenen Magnetfeld und einem elektrischen Feld. Der zweite Term auf der linken Seite gibt die Coriolisbeschleunigung, auf der rechten Seite stehen die Zentrifugalbeschleunigung und die Gravitationskraft der beiden schweren Massen. Diese Terme sind Funktionen nur von (x, y) und können aus einer Potentialfunktion W abgeleitet werden mit

$$W(x, y) = -\frac{1}{2}(x^2 + y^2) - \frac{1-m}{r_1} - \frac{m}{r_2} . \quad (2.74)$$

¹⁴Eingeschränktes Dreikörperproblem: die Testmasse wird als klein gegen die beiden anderen Massen betrachtet, so dass der Massenmittelpunkt aller drei Massen mit dem Massenmittelpunkt der beiden großen Massen zusammen fällt.

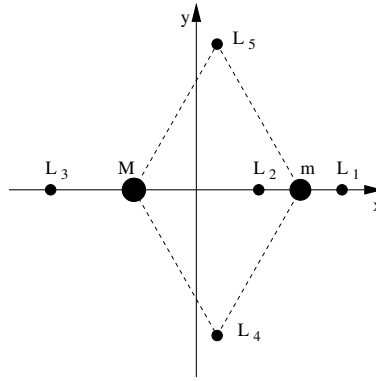


Abbildung 2.22: Lage der fünf Lagrange-Punkte im rotierenden Bezugssystem

Damit werden die Bewegungsgleichungen zu

$$\frac{d^2x}{dt^2} - 2\frac{dy}{dt} = -\frac{\partial W}{\partial x} \quad \text{und} \quad \frac{d^2y}{dt^2} + 2\frac{dx}{dt} = -\frac{\partial W}{\partial y} . \quad (2.75)$$

Multiplikation mit $2dx/dt$ bzw. $2dy/dt$ und anschließende Integration über die Zeit liefert $\frac{1}{2}(x^2 + y^2) = -W(x, y) + C$, wobei sich die Integrationskonstante C aus den Anfangsbedingungen ergibt. Diese Herleitung ist analog zu der Herleitung der Energieerhaltung für einen Massenpunkt in einem wirbelfreien Kraftfeld. In diesem Fall ist die Jacobi-Konstante C jedoch nicht die Energie, da wir uns in einem rotierenden Bezugssystem befinden. Zu beachten ist, dass die Corioliskraft keine Arbeit leistet und damit die Erhaltungsgleichung nicht beeinflusst. Dennoch ist die Corioliskraft für die Bewegung des Testkörpers von entscheidender Bedeutung.

Da die kinetische Energie niemals negativ werden kann, ergibt sich aus der Erhaltungsgleichung die wichtige Konsequenz, dass die Bewegung nur in den Bereichen der (x, y) -Ebene erfolgen kann, in denen gilt $W(x, y) \leq C$. Diese Einschränkungen hängen von der Jacobi-Konstante und damit von den Anfangsbedingungen ab. Für $W = C$ wird die Grenze eines erlaubten Bereiches definiert, die sogenannten Hill's Zero-Velocity Curves. Diese Bezeichnung ergibt sich daraus, dass der Testkörper diese Kurven nur dann erreichen kann, wenn seine Geschwindigkeit in dem rotierenden System verschwindet. Um die Topologie dieser erlaubten Bereiche zu verstehen, sollten wir uns zuerst klar machen, dass $-E$ sehr groß wird, wenn entweder $(x^2 + y^2)$, d.h. das Quadrat des Abstandes der Testmasse vom Ursprung, sehr groß wird, oder wenn entweder r_1 oder r_2 sehr klein werden. Wenn daher $-C$ groß ist, muss sich die Testmasse entweder außerhalb eines großen Kreises mit Radius $x^2 + y^2 = -2C$ aufhalten oder aber innerhalb eines sehr schmalen Kreises um einen der beiden schweren Körper. Die Radien dieser Kreise sind jeweils gegeben durch $r_1 = (m - 1)/C$ und $r_2 = -m/C$. Im ersten Fall empfindet der Testkörper die beiden Massen ungefähr als ein Gravitationszentrum, im anderen Fall sieht der Testkörper nur die eine Masse, die er umkreist.

An dieser Stelle ist es sinnvoll, die Punkte zu bestimmen, an denen die Gesamtkraft auf den Testkörper Null ist, d.h. die Punkte, an denen er im rotierenden Bezugssystem in Ruhe bleibt. Dazu ist gefordert

$$-\nabla W = \vec{r} - (1 - m)\frac{\vec{r}_1}{r_1^3} - m\frac{\vec{r}_2}{r_2^3} = 0 , \quad (2.76)$$

mit $\vec{r} = (x, y)$, $\vec{r}_1 = (x + m, y)$ und $\vec{r}_2 = (x - 1 + m, y)$. Da das Massenzentrum im Ursprung liegt, ist dann $\vec{r} = (1 - m)\vec{r}_1 + m\vec{r}_2$. Einsetzen liefert

$$(1 - m)\left[\frac{1}{r_1^3} - 1\right]\vec{r}_1 + m\left[\frac{1}{r_2^3} - 1\right]\vec{r}_2 = 0 . \quad (2.77)$$

Wenn sich diese Gleichgewichtspunkte nicht auf der x -Achse befinden, so kann die Gleichung nur dann erfüllt werden, wenn $r_1 = r_2 = 1$, d.h. zwei der Gleichgewichtspunkte sind die

Spitzen eines gleichseitigen Dreiecks, dessen Basis durch die Strecke zwischen den beiden schweren Massen gegeben ist (L_4 und L_5 in Abb. 2.22).

Um die Lösungen von (2.77) zu bestimmen, die auf der x -Achse liegen, ist es hilfreich, eine dimensionslose Größe $\alpha = x + m$ einzuführen. Damit lässt sich die Gleichung umwandeln zu

$$(1 - m) \frac{\alpha}{|\alpha|^3} + m \frac{(\alpha - 1)}{|\alpha - 1|^3} = \alpha - m. \quad (2.78)$$

Die Fälle $\alpha=0$ und $\alpha=1$ entsprechen dabei den Situationen, dass sich die Testmasse am Ort von M bzw. m befindet. Die Gleichung hat stets drei reelle Lösungen, wir werden sie hier für den eingeschränkten Fall lösen, dass $m \ll M$ ist, d.h. $m \ll 1$. Diese Situation ist für die meisten Anwendungen im Sonnensystem, also z.B. für das Erde-Mond-System oder das Sonne-Jupiter-System gegeben.

Eine offensichtliche Lösung ist der L_3 -Punkt, an dem sich der Testkörper genau entgegengesetzt zur kleineren Masse m befindet. Diese Lösung können wir finden, wenn wir in (2.78) $\alpha = -1 + O(m)$ setzen, so dass sich $\alpha = -1 + \frac{7}{12}m + O(m^3)$ ergibt. Betrachtet man nur einen anziehenden Körper, so befindet sich der Testkörper an diesem Punkt in einem Gleichgewicht aus der Anziehung des Hauptkörpers und der Zentrifugalkraft. Die Anwesenheit der zweiten anziehenden Masse bewirkt eine zusätzliche Anziehungskraft, die zu einer Verstärkung der Anziehung führt, wobei sich der Gleichgewichtspunkt geringfügig nach links verschiebt. Diese neue Kraft wird durch einen Zuwachs der Zentrifugalkraft ausgeglichen.

Zwei weitere Gleichgewichtspunkte befinden sich an Punkten dicht an der kleinen Masse, deren Störungen den Radius des Keplerorbits um die Masse M beeinflussen. In diesem Falle müssen wir die Ordnung in m der kleinen Größe $\alpha \ll 1$ finden. Dazu setzen wir versuchsweise $\alpha = 1 + \lambda m q$. Darin ist $q > 0$ und λ in der Größenordnung von 1. Damit ergibt sich für (2.78)

$$\frac{m^{(1-2q)} \lambda}{|\lambda|^3} = 3\lambda m q + O(m^2 q), \quad (2.79)$$

wobei die beiden größeren Terme für $q = 1/3$ von der gleichen Ordnung sind. Dann erhalten wir $|\lambda|^3 = \frac{1}{3}$ und damit für die beiden Gleichgewichtspunkte $\alpha = 1 \pm (m/3)^{1/3} + O(m^{2/3})$. Dabei nimmt der Abstand von der kleineren Masse nur mit der dritten Wurzel ab. Damit ergibt sich für die fünf Lagrange-Punkte:

$$\begin{aligned} L_1 &= [1 - m + (m/3)^{1/3}, 0], \\ L_2 &= [1 - m - (m/3)^{1/3}, 0], \\ L_3 &= [-1 - 5m/12, 0], \\ L_4 &= [1/2 - m, -\sqrt{3}/2], \\ L_5 &= [1/2 - m, \sqrt{3}/2]. \end{aligned} \quad (2.80)$$

Diese fünf Punkte sind in Abb. 2.22 im rotierenden Bezugssystem dargestellt.

Um die Form der Zero-Velocity-Kurven mit $W(x, y) = \text{const}$ zu bestimmen, ist es nützlich, die Werte von W an den fünf stationären Punkten L_i zu bestimmen zusammen mit dem topologischen Charakter dieser Punkte. Das kann durch die Bestimmung der zweiten Ableitung von W und die Bestimmung des Vorzeichens der Jacobi-Determinante $J = (\partial_{xx} W)(\partial_{yy} W) - (\partial_{xy} W)^2$ an den Lagrange-Punkten geschehen. Dabei ergibt sich $W(L_5) = W(L_4) > W(L_3) > W(L_1) > W(L_2)$. Dabei sind L_4 und L_5 Maxima, die anderen Lagrange-Punkte dagegen sind Sattelpunkte. Damit lassen sich die Zero-Velocity-Kurven eindeutig definieren, wenn wir die Jacobi-Konstante von sehr großen negativen Werten (z.B. die schon weiter oben diskutierten Grenzen) durch die kritischen Werte ansteigen lassen. Qualitativ ergeben sich dabei fünf Möglichkeiten:

I $C < W(L_2)$: Die Bewegung ist nur innerhalb zweier kleiner Kreise um die beiden massiven Körper erlaubt, oder außerhalb eines großen Kreises, der beide massiven Körper umfasst. In diesem Falle kann die Testmasse als ein Satellit der beiden massiven Körper (die von der Testmasse nur als ein Gravitationszentrum gesehen werden) betrachtet werden.

- II $W(L_1) > C > W(L_2)$: Zwischen den beiden erlaubten Regionen um die beiden massiven Körper öffnet sich ein Kanal. Der Testkörper kann sich, wenn er sich einmal in der Nähe der beiden massiven Körper befindet, nicht aus deren Nähe entfernen. In diesem Bild können Satelliten zwischen den beiden Körpern ausgetauscht werden, der Testkörper kann aber das System nicht verlassen. Eine entsprechende Beschreibung wäre z.B. auch für die Materie in einem Doppelsternsystem angemessen. Diese Topologie erklärt auch, warum die meisten Trajektorien zwischen Erde und Mond die Form einer 8 annehmen, wobei der Schnittpunkt in der Nähe von L_2 liegt.
- III $W(L_3) > C > W(L_1)$: Der erlaubte Bereich öffnet sich hinter der kleineren Masse und der Testkörper kann durch dieses Loch entweder von der schwereren Masse entweichen oder, wenn er sich ursprünglich im äußeren Bereich befunden hat, in die Nähe der schwereren Masse vordringen. Der verbotene Bereich hat ungefähr die Form eines Hufeisens, die L_4 und L_5 liegen in der Nähe der Spitzen, L_3 liegt innerhalb des Hufeisens.
- IV $W(L_5) = W(L_4) > C > W(L_3)$: Auch hinter der schwereren Masse öffnet sich ein Kanal, dadurch sind die Zero-Velocity-Kurven wieder voneinander getrennt. Die Bewegung ist dann nur noch in zwei relativ schmalen Bereichen um L_4 und L_5 verboten.
- V $C > W(L_4) = W(L_5)$: Die Bewegung kann in der gesamten (x, y) -Ebene erfolgen.

Diese Klassifikation bzw. Topologie bleibt selbst dann noch erhalten, wenn einige der bei der Herleitung gemachten Einschränkungen nicht mehr gelten. Ist die kleinere Masse m nicht infinitesimal, so verschieben sich die drei Lagrange-Punkte lediglich entlang der x -Achse des Systems. Selbst wenn die Orbits der beiden Massen eine gewisse, allerdings nicht zu große Exzentrizität aufweisen, ist eine Klassifikation mit Hilfe der Jacobi-Konstanten immer noch sinnvoll, insbesondere im Hinblick auf Stabilität oder Entweichen der Testmasse.

Eine der entscheidendsten und auch interessantesten Fragen betrifft die Stabilität der Lagrange-Punkte, d.h.: Wenn die Testmasse anfangs in einen geringen Abstand von einem der Lagrange-Punkte gesetzt wurde, bleibt sie dann in der Nähe dieses Punktes oder entfernt sie sich immer weiter vom Lagrange-Punkt? Betrachten wir nur die zweiten Ableitungen des Potentials, so wäre keiner der Lagrange-Punkte stabil, da keiner von ihnen einem Minimum im Potential entspricht. Allerdings wirkt in unserem System auch noch die geschwindigkeitsabhängige Corioliskraft. Damit ergibt sich, dass die beiden Lagrange-Punkte L_4 und L_5 stabil sind [8]. Die Stabilität dieser beiden Lagrange-Punkte ist auch praktisch bewiesen durch die Trojaner, zwei Familien von Asteroiden, die im System Sonne–Jupiter stabil um diese Punkte oszillieren. Die Punkte befinden sich dabei auf dem Jupiter–Orbit und laufen jeweils um 60° versetzt vor bzw. hinter dem Planeten her. Auch scheinen viele der Saturnmonde kleine, Trojaner-ähnliche Begleiter zu haben.

Die Punkte L_1 bis L_3 sind instabil, jedoch gibt es Bahnen um diese Punkte herum, die stabil sind. Im Falle des oben erwähnten ISEE-3 Orbits handelt es sich um eine elliptische Bahn um den L_3 -Lagrange-Punkt. Die Stabilisierung durch die Ellipsenbahn folgt dem Prinzip des Kreisels. Die Ellipsenbahn ist mit einem Drehimpuls verbunden, wobei, da die Bahnebene senkrecht auf der Achse Sonne–Erde steht, der Drehimpulsvektor entlang eben dieser Achse ausgerichtet ist. Wird jetzt der Satellit etwas aus der Bahn ausgelenkt, so ist das Gleichgewicht der Anziehungskräfte von Sonne und Erde verletzt und es wirkt eine Nettokraft in Richtung auf einen der Körper. Das bedeutet aber nichts anderes, als würde man versuchen, einen Kreisel zu kippen: auf einen Drehimpuls wirkt ein Drehmoment (im Falle des normalen Kreisels die Schwerkraft, im Falle des Satelliten die Netto-Anziehungskraft durch den jetzt dominierenden Körper). Der Kreisel antwortet darauf, indem er versucht, seinen Drehimpulsvektor in Richtung des Drehmomentvektors zu ändern. Dadurch weicht er dem Umkippen aus und führt eine Präzessionsbewegung aus. Für Satelliten im Lagrange-Punkt zwischen Sonne und Erde ermöglichen diese Stabilisierungsorts auch eine ungestörte Kommunikation mit dem Satelliten, da sich dieser dann, von der Erde aus gesehen, um die Radioquelle Sonne bewegt aber nie direkt vor ihr steht.

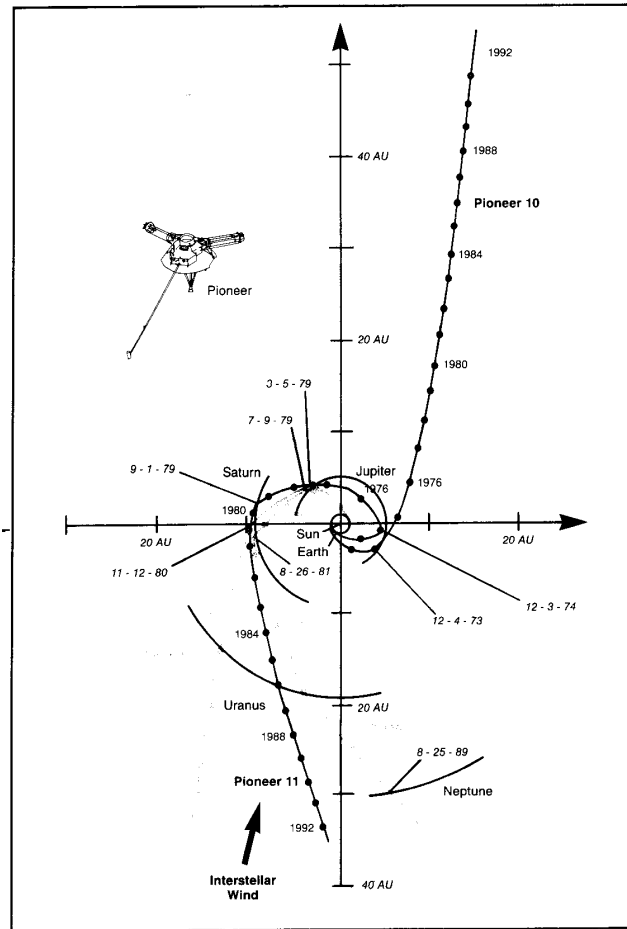


Abbildung 2.23: Bahnen der Voyager- und Pioneer-Sonden [6]

2.7.2 Planetares Billard

Die bisher betrachteten Bewegungen waren alle periodische Bewegungen mit der Erde oder Sonne als Zentralkörper bzw. um den Lagrange-Punkt. Bei Missionen zu einem anderen Planeten werden vielfach die bereits in Abschn. 2.6.4 betrachteten Hohmann-Bahnen verwendet, bei Missionen zu mehreren Planeten ergeben sich unregelmäßige Bahnen. Abbildung 2.23 zeigt als Beispiele die Bahnen der Pioneer- und Voyager-Sonden. Auffallend sind die starken Änderungen der Bahn in der Nähe anderer Himmelskörper. Diese Bahnänderungsmanöver werden als 'gravity assisted navigation', Gravitationsumlenkung oder Swingby bezeichnet.

Das Prinzip des Swingby beruht darauf, dass bei genügend nahem Vorbeiflug an einem Himmelskörper (Planeten oder Monde) das Gravitationspotential und die Eigenbewegung dieses Himmelskörpers ausgenutzt werden, um eine Impulsänderung zu erzielen [13, 26, 73]. Diese Impulsänderungen werden ausgenutzt, um mehrere Planeten mit einer Sonde besuchen zu können bzw. mehrere Monde auf einer Mission anzufliegen. Die erforderliche Genauigkeit dieses Billardspiels lässt sich am Beispiel der Mariner-Mission illustrieren: Der Fehler beim Venus-Vorbeiflug durfte maximal 300 km betragen bei einer Flugstrecke von 250 Millionen km. Der echte Fehler betrug nur 20 km. Ein Scharfschütze müsste bei entsprechender Präzision ein 10-Pfennigstück über eine Entfernung von 250 km treffen [26]. Aber nicht nur die Präzision ist wichtig für eine derartige Schleudertour, im Falle von sehr ausgedehnten Missionen wie der Voyager-Mission ist auch die relative Stellung der Planeten zueinander von Bedeutung: für den Voyager-Flug lagen alle diese Planeten auf einer Kurve, deren Krümmungsrichtung sich nicht änderte. Es entstand also kein Zick-Zack-Kurs, der einen

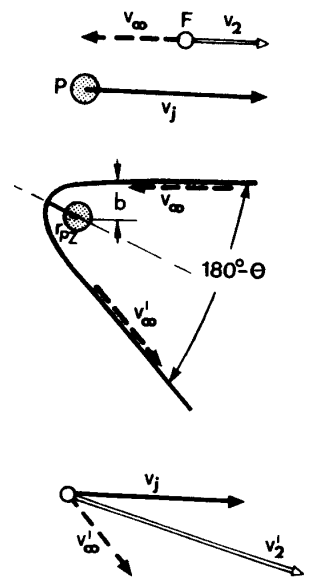


Abbildung 2.24: Swingby an einem Planeten [26]

extrem hohen Energiebedarf bedeutet hätte. Diese Voraussetzung ist jedoch aufgrund der langen Umlaufzeiten der äußeren Planeten nur alle 177 Jahre erfüllt, vor dem Voyager-Start im Jahre 1977 war das zur Zeit der napoleonischen Kriege zuletzt der Fall.

Um das Prinzip des Schwungholens an einem Planeten zu verstehen, ist es wichtig, die Eigenbewegung des Planeten zu berücksichtigen. Würde die Raumsonde in die Nähe eines ruhenden Körpers gelangen, so würde sie, vorausgesetzt ihre Anfangsgeschwindigkeit ist ausreichend groß, beim Anflug durch die in Bewegungsrichtung wirkende Gravitation beschleunigt werden, dann aber beim Weiterflug hinter dem Planeten wieder abgebremst werden. Dabei würde sich eine Hyperbelbahn ergeben, wobei die Geschwindigkeit auf beiden Ästen die gleiche ist, d.h. es findet zwar eine Impulsänderung statt aber keine Energieänderung. In der Relativität bewegt sich der Planet jedoch und die Sonde kann einen verschwindend kleinen Teil der kinetischen Energie des Planeten abzweigen und zur Erhöhung der eigenen Geschwindigkeit verwenden.

Anschaulich beschreibt Giese [26] diese Technik folgendermaßen: Im Prinzip ähnelt der Vorgang einem irdischen Experiment, bei dem ein Tennisball senkrecht an einem Haus mit einer Anfangsgeschwindigkeit v_1 hochgeworfen wird und mit verringerter Geschwindigkeit v_2 an einem Balkon vorbeikommt, wo ihn ein Tennisspieler mit seinem Schläger (Geschwindigkeit v_j) weiter nach oben schlägt. Auf diese Weise erreicht der Ball wieder eine höhere Geschwindigkeit $v'_2 > v_2$ und steigt weiter auf als ihn die Person vom Boden aus eigener Kraft hätte schleudern können.

Übertragen auf ein Raumfahrzeug F , das durch Wechselwirkung mit einem Planeten P von einer heliozentrischen Geschwindigkeit v_2 auf eine höhere Geschwindigkeit v'_2 beschleunigt werden soll, ergibt sich das Prinzip der Swingby-Technik wie in Abb. 2.24 dargestellt. Dieser Vorgang lässt sich in drei Abschnitte unterteilen:

(1) Zunächst holt der Planet in dem vereinfacht dargestellten Fall mit seiner Eigengeschwindigkeit v_j die sich mit einer Geschwindigkeit v_2 bewegendes Raumfahrzeug ein. Ein Beobachter auf dem Planeten würde daher die Sonde mit einer Geschwindigkeit $v_\infty = v_2 - v_j$ auf sich zukommen sehen.

(2) Der auf dem Planeten befindliche Beobachter würde dann feststellen, dass sich das Raumfahrzeug auf einer Hyperbelbahn um den Planeten herumschwingt, wobei der minimale Abstand im Perizentrum mit r_{Pz} bezeichnet ist. Schließlich verlässt das Raumfahrzeug den Wirkungsbereich des Planeten wieder mit einer Geschwindigkeit v'_∞ . Dabei hat sich aber nicht

der Betrag der Geschwindigkeit geändert, sondern lediglich die Richtung (d.h. im System des Planeten ist die Energie der Raumsonde nach der Wechselwirkung unverändert).

(3) Die heliozentrische Geschwindigkeit der Raumsonde nach der Wechselwirkung ist dagegen gegeben durch $\vec{v}'_2 = \vec{v}_j + \vec{v}'_\infty$. Die Raumsonde hat also nach der Wechselwirkung mit dem Planeten eine höhere Geschwindigkeit.

Formal kann man den Vorgang als einen elastischen Stoß zwischen dem Raumfahrzeug und dem Planeten beschreiben oder als einen Streuprozess zwischen Sonde und Planet, wobei das Gravitationsfeld des Planeten die Ursache der Streuung ist. Dieser Streuvorgang ist durch die Einschussgeschwindigkeit v_∞ und den Stoßparameter b gekennzeichnet. Dadurch wird der Ablenkwinkel θ gegenüber der ursprünglichen Richtung \vec{v}_∞ festgelegt. Der Energiesatz für die Hyperbel war gegeben durch die Binet'sche Gleichung (2.30) zu

$$v^2 = \mu \left(\frac{2}{r} + \frac{1}{a} \right). \quad (2.81)$$

Die beiden Extremgeschwindigkeiten im Unendlichen bzw. im Perizentrum sind damit gegeben durch

$$v_\infty^2 = v_\infty'^2 = \frac{\mu}{a} \quad \text{bzw.} \quad v_{\text{Pz}}^2 = \mu \left(\frac{2}{r_{\text{Pz}}} + \frac{1}{a} \right). \quad (2.82)$$

Aus der Drehimpulserhaltung ergibt sich mit v_{Pz} als der Geschwindigkeit im Perizentrum $\sigma = \sqrt{\mu a(\epsilon^2 - 1)} = r_{\text{Pz}} v_{\text{Pz}}$. Aus der Hyperbelgeometrie lässt sich über

$$\cos \frac{(180^\circ - \theta)}{2} = \frac{a}{\sqrt{a^2 + a^2(\epsilon^2 - 1)}} \quad (2.83)$$

der Ablenkwinkel bestimmen zu

$$\sin \frac{\theta}{2} = \frac{1}{\epsilon}. \quad (2.84)$$

Durch Einsetzen in den Drehimpulssatz wird

$$r_{\text{Pz}}^2 v_{\text{Pz}}^2 = \mu a(\epsilon^2 - 1) = r_{\text{Pz}}^2 \mu \left(\frac{2}{r_{\text{Pz}}} + \frac{v_\infty^2}{\mu} \right) \quad (2.85)$$

und es ist

$$\frac{b}{a\epsilon} = \sin \left(90^\circ - \frac{\theta}{2} \right) = \sqrt{1 - 1/\epsilon^2}. \quad (2.86)$$

Damit lassen sich der Ablenkwinkel und der Stoßparameter bestimmen zu

$$\sin \frac{\theta}{2} = \frac{1}{1 + \frac{r_{\text{Pz}} v_\infty^2}{\mu}} \quad \text{und} \quad b = r_{\text{Pz}} \sqrt{1 + \frac{2\mu}{r_{\text{Pz}} v_\infty^2}}. \quad (2.87)$$

Bei sehr weitem (großes r_{Pz}) und/oder sehr schnellem Vorbeiflug erfolgt also kaum eine Ablenkung. Andererseits wird die Ablenkung um so größer, je dichter die Raumsonde an den Planeten kommt, d.h. je kleiner r_{Pz} wird. Damit lässt sich eine maximale Ablenkung bestimmen, da die geringste Annäherung durch den Planetenradius bestimmt ist (es kann also in keinem Falle eine Ablenkung um 180° erfolgen). Je massereicher der Planet ist, umso näher kann man diesem Wert kommen. Der Stoßparameter b gibt an, wohin die Raumsonde gelenkt werden muss, um bei vorgegebenem r_{Pz} und v_∞ die gewünschte Ablenkung θ zu erreichen.

Außer bei den Voyager- und Pioneer-Sonden wurde diese Technik des Swingby unter anderem auch bei Mariner 10 zum Besuch von Venus und Merkur angewendet; bei Ulysses, um aus der Ebene der Ekliptik zu gelangen (Ausnutzung des Gravitationspotentials vom Jupiter – Jupiter ist für solche Manöver sehr gut geeignet, da er nach der Sonne das Objekt mit der größten Masse im Sonnensystem ist) und bei Galileo zum Einschwenken in eine Jupiter-Umlaufbahn unter Ausnutzung der Gravitation der Planetenmonde. Zusätzlich wird

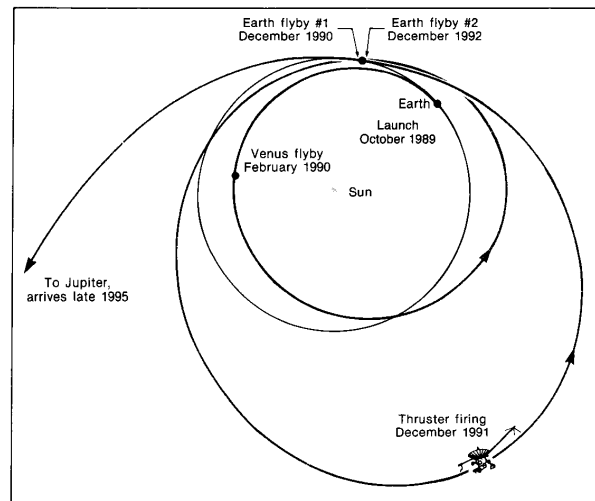


Abbildung 2.25: VEEGA (Venus-Earth-Earth gravity assist) Manöver der Raumsonde Galileo: mehrfaches Schwungholen, um die für den Flug zum Jupiter benötigte Geschwindigkeit zu erreichen [6]

das Schwungholen an Planeten und Monden bei Galileo nicht nur zur Bahnänderung, sondern auch zur Geschwindigkeitserhöhung eingesetzt. Ursprünglich sollte Galileo zusammen mit einer Flüssigtreibstoff-Rakete vom Space-Shuttle in einer niedrigen (ca. 400 km) Erdumlaufbahn ausgesetzt werden und dann mit Hilfe dieser Rakete in Richtung auf den Jupiter geschossen werden. Nach der Explosion der Challenger im Januar 1986 wurde jedoch der Transport von Flüssigtreibstoffen im Space-Shuttle verboten, Galileo konnte daher nur mit einer weniger explosiven aber auch weniger schubstarken Feststoffrakete ausgesetzt werden. Um dennoch die notwendige Geschwindigkeit zu erreichen, wurde Galileo auf einen Kurs gebracht, der es ihm ermöglichte, einmal an der Venus und zweimal an der Erde Schwung zu holen (VEEGA: Venus-Earth-Earth gravity assist, vgl. Abb. 2.25), um dann seine drei Jahre dauernde Reise zum Jupiter zu beginnen.

Literatur

Allgemeine Einführungen in Satellitenbahnen und ihre Störungen geben die Bändchen von Sagirov [102] und Bohrmann [10]. Dort werden auch Hohmann- und andere Transfer-Bahnen behandelt. Spezielle Bahnen, insbesondere die Stabilität von Satelliten in Lagrange-Punkten und interplanetares Billard, werden diskutiert in Bertotti und Farinella [8]. Für die formalen Betrachtungen (Zweikörper- und eingeschränktes Dreikörperproblem) kann auch ein Lehrbuch zur Theoretischen Physik zu Rate gezogen werden.

Fragen

1. Die Bahnen von nicht manövrierfähigen Satelliten werden durch die Kepler'schen Gesetze beschrieben. Geben Sie diese an (verbal, wo möglich auch formal).
2. Die Kepler'schen Gesetze sind empirisch bestimmt. Welche Grundkonzepte und -gleichungen werden zu ihrer formalen Herleitung benötigt?
3. Eine elliptische Bahn wird, ebenso wie eine Kreisbahn, durch ein Kräftegleichgewicht bestimmt. Für die Kreisbahn muss die Gravitationskraft gleich der Zentripetalkraft sein. Welche Kräfte wirken auf den Satelliten auf einer elliptischen Bahn (bitte mit Skizze)?
4. Bei der Herleitung der Kepler'schen Gesetze macht man den Übergang vom allgemeinen

- zum eingeschränkten Zweikörperproblem. Welche Annahmen gehen dabei ein? Erläutern Sie, unter welchen Bedingungen diese Vereinfachung sinnvoll ist.
5. Die Herleitung der Kepler'schen Gesetze beruht auf fundamentalen Erhaltungssätzen. Geben Sie die Erhaltungssätze an und ihren Zusammenhang mit der Bewegungsgleichung.
 6. Satellitenbahnen sind Kegelschnitte. Kreuzen Sie die wahren Aussagen an:
 - Ist die Anfangsgeschwindigkeit eines Satelliten kleiner der ersten kosmischen Geschwindigkeit, so ergibt sich eine Wurfparabel.
 - Wird die Geschwindigkeit eines Satelliten durch Reibung auf einen Wert kleiner der Kreisbahngeschwindigkeit verringert, so stürzt der Satellit auf einer Spiralbahn ab.
 - Ein Satellit mit einer Geschwindigkeit größer der Kreisbahngeschwindigkeit entweicht aus dem Schwerefeld der Erde.
 - Eine hinreichende Bedingung für eine Ellipsenbahn ist eine Bahngeschwindigkeit kleiner der zweiten kosmischen Geschwindigkeit.
 - Die Geschwindigkeit eines Satelliten im Perigäum darf nicht kleiner sein als die Kreisbahngeschwindigkeit.
 - Die Geschwindigkeit eines Satelliten im Apogäum darf nicht kleiner sein als die Kreisbahngeschwindigkeit.
 7. Bestimmen Sie aus der Binet'schen Gleichung den Zusammenhang zwischen Kreisbahngeschwindigkeit und erster kosmischer Geschwindigkeit sowie zwischen Kreisbahngeschwindigkeit und zweiter kosmischer Geschwindigkeit.
 8. Durch welche Parameter ist die Lage einer Satellitenbahn im Raum bestimmt? Welche zusätzlichen Parameter sind zur Angabe der Satellitenposition erforderlich? Benennen und erklären Sie die Parameter.
 9. Erläutern Sie den Begriff Inklination.
 10. Bestimmen Sie die Bahngeschwindigkeiten eines Satelliten im Perigäum und im Apogäum.
 11. Auch ohne Kenntnis der Gravitationskonstante kann man angeben, wieviel mal masse-reicher die Sonne ist als die Erde. Man braucht dazu außer allbekannten Daten über die Jahres- und Monatslänge nur das Verhältnis des Abstände von Sonne bzw. Mond von der Erde (400:1), nicht aber die absoluten Abstände. Wie funktioniert das?
 12. Für ein studentisches Kleinsatellitenprojekt hat die Uni Osnabrück auf dem Westerberg eine Raletenstartrampe errichtet. Der Satellit soll auf einer antriebslosen Bahn fliegen; zur Auswahl stehen ein Instrument zur Untersuchung des Wasserdampftransports über tropischen Ozeanen oder ein Instrument zur Untersuchung des Anteils flüssigen Wassers in Eis. Welches Instrument ist für den Satelliten besser geeignet?
 13. Bahnmanöver: welche der folgenden Aussagen ist wahr?
 - Eine Erhöhung der Bahngeschwindigkeit führt auf ein höheres Orbit.
 - Eine Änderung der Bahngeschwindigkeit senkrecht zur Bahnebene führt stets auf eine Drehung der Bahnebene unter Beibehaltung der Bahn in der Bahnebene.
 - Ein dauerhaft höheres Orbit lässt sich nur mit mindestens zwei Geschwindigkeitsänderungen erreichen.
 - Um in ein ansonsten identisches Orbit senkrecht zur Ausgangsbahnebene zu gelangen, muss die Kraft in einem Winkel von 135° zur Flugrichtung stehen.
 - Um in ein ansonsten identisches Orbit senkrecht zur Ausgangsbahnebene zu gelangen, muss die Kraft in einem Winkel von 45° zur Flugrichtung stehen.
 14. Eine Sojus-Kapsel nähert sich der Internationalen Raumstation ISS auf einem kreisförmigen Orbit wenige Kilometer unterhalb der Umlaufbahn der Raumstation (ebenfalls kreisförmig, beide Bahnen liegen bereits in einer Ebene). Skizzieren Sie, welche Manöver der Kommandant der Sojus-Kapsel ausführen muss, um an der Raumstation andocken zu können (Achtung, beim Andocken sollte die Relativgeschwindigkeit Null sein).
 15. Zum Nachdenken: Um, insbesondere im Hinblick auf Langzeitmissionen, etwas mehr Bewegung zu haben, darf die Mannschaft eines Space-Shuttle jetzt Tischtennis spielen. Aus Platzgründen geht das allerdings erst nach dem Aussetzen der Nutzlast in der offenen

Ladebucht (Sport an frischer Luft ist ohnehin gesünder). Diskutieren Sie, ob sich aus dieser Situation veränderte Spielgewohnheiten ergeben, z.B. unter den folgenden Gesichtspunkten: kann man einen normalen Aufschlag machen (den Ball hochwerfen und dann schlagen)? Oder sollte man einen modifizierten Aufschlag entwickeln? Was passiert mit dem Ball, wenn der Partner nicht trifft? Gibt es Situationen, in denen Sie den Ball in einem späteren Orbit wieder einfangen können? Macht es einen Unterschied, ob das Shuttle (wie normal) mit der geöffneten Ladebucht nach unten zur Erde blickend fliegt oder mit einer nach außen weisenden Ladebucht? Wenn Ihnen die Situation zu unübersichtlich ist, betrachten Sie einfache Grenzfälle: was passiert mit Bällen, die (a) senkrecht nach oben, (b) senkrecht nach unten, (c) direkt in Flugrichtung, (d) entgegen der Flugrichtung und (e) senkrecht zur Bahnebene (also nach links oder rechts) geworfen werden.

16. Veranschaulichen Sie sich, warum die Gesamtenergie der Ellipsenbahn nur von der großen Halbachse abhängt, nicht jedoch von der Exzentrizität. Kann man diesen Zusammenhang gezielt nutzen, um zumindest für kurze Zeiten möglichst große Abstände von der Erdoberfläche zu erreichen?
17. Erdfernerkundungssatelliten werden häufig in sonnensynchrone Bahnen gebracht. Was versteht man unter einer sonnensynchronen Bahn, welche Vorteile hat sie?
18. Geben Sie typische Bahnparameter für Erdfernerkundungssatelliten an und begründen Sie, warum diese so gewählt werden.
19. Erläutern Sie Verfahren zur Lagestabilisierung eines Satelliten. Welche Verfahren sind für Erdfernerkundungssatelliten geeignet, welche nicht?
20. Bahnstörungen müssen nicht unbedingt Störungen sein sondern können auch gezielt zur Beeinflussung von Satellitenbahnen eingesetzt werden. Nennen Sie Beispiele.
21. Benennen Sie mindestens drei Störprozesse, die eine Satellitenbahn beeinflussen können und charakterisieren Sie diese kurz.
22. Reibung in der Hochatmosphäre kann auch zur Bestimmung der Dichte der Atmosphäre verwendet werden. Skizzieren Sie die Herleitung. Welche Größe wird als beobachtbare Größe verwendet?
23. Ein auf einer elliptischen Bahn befindlicher Satellit wird durch Reibung abgebremst. Wie verändert sich die Bahn? Erklären Sie, warum die Änderungen so erfolgen.
24. Geben Sie die Mindestflughöhe und Mindestumlaufdauer für einen Satelliten an. Was begrenzt diese Werte?
25. Die Mindestflughöhe eines Erdsatelliten ist durch Reibung in der Atmosphäre bestimmt. Gibt es auch eine maximale Flughöhe? Wenn ja, wodurch ist diese bestimmt und wie können Sie sie abschätzen?
26. Beschreiben Sie die Bahnstörungen, die sich aus der Erdabplattung ergeben. Wofür werden diese Störungen gezielt ausgenutzt?
27. Warum muss man bei einem Übergang von einem Orbit großer Höhe auf ein Orbit niedrigerer Höhe die gleiche Energie verwenden wie bei einem Übergang in Gegenrichtung? Bestimmen Sie die Gesamtenergie beider Bahnen. Wo bleibt die beim Übergang aufgebrauchte Energie beim Übergang von der höheren zur niedrigeren Bahn?
28. Warum kann eine Rakete bzw. ein Satellit die Atmosphäre beim Start problemlos durchdringen während abstürzende Satelliten verglühen?

Aufgaben

1. Ein Molnyia-Satellit hat ein Perigäum von 400 km und ein Apogäum von 40 000 km. Bestimmen Sie die jeweiligen Geschwindigkeiten.
2. Ein Satellit der Masse $m = 1000$ kg befindet sich in einer polaren Umlaufbahn in 400 km Höhe. Durch einen einmaligen gleichförmigen Kraftstoß von 10 Sekunden Dauer wird der Satellit auf eine äquatoriale Bahn gebracht. Berechnen Sie die dazu erforderliche Kraft.
3. Die Jupitermonde befolgen ebenfalls Kepler's drittes Gesetz: die Kuben der großen Halbachsen dividiert durch die Quadrate der Umlaufzeiten ergeben eine Konstante. Ist es die gleiche Konstante wie für Jupiter und die anderen Planeten beim Umlauf um die Sonne?

- Spezielle Frage im Bezug auf den Jupiter-Mond Io: sein Abstand vom Planeten beträgt $4.5 \cdot 10^5$ km, seine Umlaufzeit T 1 d 18 h 28 min. Wie groß ist die Masse des Jupiter?
4. Berechnen Sie die Umlaufzeit eines Satelliten um den Mond, wenn dieser sich in einem Orbit mit einer Höhe von 100 km über der Mondoberfläche befindet. Wie groß wäre die Umlaufzeit einer entsprechenden Bahn um die Erde? (Mondparameter: Masse $7.3 \cdot 10^{23}$ kg, Radius 1738 km; Erdparameter: Masse $6 \cdot 10^{24}$ kg, Radius 6378 km)
 5. Der kleinste Abstand von Halleys Kometen zur Sonne beträgt $8.9 \cdot 10^{10}$ m, seine Umlaufzeit 76 Jahre. Welcher Art ist die Bahn? Berechnen sie (a) die Länge der großen Halbachse, (b) die Exzentrizität der Bahn, und (c) den Abstand des Aphels von der Sonne. Bestimmen Sie die Verhältnisse der Geschwindigkeiten (Bahn- und Winkelgeschwindigkeit) in Perihel und Aphel.
 6. Zwischen Erde und Sonne gibt es einen Punkt, an dem sich die Gravitationskräfte der beiden auf einen Satelliten aufheben. Wo liegt dieser neutrale Punkt? Wie groß sind die Bahn- und Winkelgeschwindigkeit eines Satelliten in diesem Punkt? Vergleichen Sie mit der Erde. (Abstand Sonne-Erde 1 AU = $149 \cdot 10^6$ km).
 7. Sie sind im Untersuchungsausschuss zum Coloumbia-Absturz mit der Frage konfrontiert, ob eine Rettung der Astronauten auf die ISS möglich gewesen wäre. Die Coloumbia hätte zwar nicht andocken können, aber da auf ISS Raumanzüge vorhanden sind, hätte die Crew im Prinzip umsteigen können. Sie sollen sich mit der Frage auseinandersetzen, ob Coloumbia (Flughöhe 278 km, Inklination 39°) die ISS (385 km, Inklination 59.6°) überhaupt hätte erreichen können. Coloumbia hat keinen Treibstoff für großartige Bahnmanöver, der Treibstoffvorrat reicht nur für einen Abstieg auf ein 120 km hohes Orbit (von dort an bremst die Erdatmosphäre ausreichend, um das Shuttle in einen geregelten Abstieg zu zwingen) sowie eine Sicherheitsreserve von 20%. Wäre mit diesem Treibstoff ein Aufstieg in das ISS-Orbit möglich gewesen?
 8. Der französische Erdfernerkundungssatellit SPOT fliegt in einem nahezu kreisförmigen Orbit in einer Höhe von 832 km. Bestimmen Sie die Bahngeschwindigkeit des Satelliten und seine Geschwindigkeit entlang der Bodenspur.
 9. Ein Satellit der Masse $m = 1000$ kg soll von einem Orbit in 400 km Höhe auf eine geostationäre Bahn in 36 000 km Höhe gebracht werden. Erläutern Sie das Verfahren und berechnen Sie die erforderlichen Geschwindigkeitsänderungen sowie die dazu erforderlichen Kräfte (Annahme: die Impulsänderung erfolgt gleichförmig über einen Zeitraum von jeweils 100 s).
 10. Eine Rakete bringt einen LandSat-Satelliten in ein kreisförmiges Orbit von 200 km Höhe. Der Satellit soll in ein Orbit mit einer Flughöhe von 850 km transferiert werden. Beschreiben Sie die Bahn. Welche Geschwindigkeitsänderungen müssen erfolgen? Wieviel Energie müssen Sie bei einem Satelliten der Masse 4 t dafür aufwenden? Wie lange dauert der Übergang? Mit welchen Geschwindigkeiten bewegt sich der Satellit in den beiden kreisförmigen Orbits?
 11. Eine Rakete bringt einen Satelliten ($m = 1000$ kg) in eine Umlaufbahn um die Sonne (Näherung: Satellit befindet sich im Erdorbit, ist jedoch gravitativ nicht mehr an die Erde sondern bereits an die Sonne gebunden). Der Satellit soll auf eine Merkurbahn/Jupiterbahn transferiert werden. Welche Energie ist aufzuwenden? (Masse Sonne $1.9 \cdot 10^{30}$ kg, Erdbahnradius 1 AU=149 Mio km, Merkurbahn 0.387 AU, Jupiterbahn 5.2 AU, Gravitationskonstante $6.67 \cdot 10^{-11}$ Nm²/kg²)

Kapitel 3

Satelliteninstrumente

Fernerkundungsinstrumente auf Satelliten sammeln Informationen über die Erdoberfläche, die Ozeane bis hin zu den Ozeanböden und die Atmosphäre. Entsprechend der vielfältigen Fragestellungen und beobachteten Objekte hat sich eine große Vielfalt von unterschiedlichen Sensoren entwickelt, eine sehr umfassende Zusammenstellung gibt [62]. In diesem Kapitel werde ich mich auf die Vorstellung von Messprinzipien und Universalinstrumenten beschränken, Weiterentwicklungen werden nur skizziert. Zu den verschiedenen Typen von Instrumenten werden einige kurze Anwendungsbeispiele vorgestellt, weitere Anwendungsbeispiele, die teilweise auch auf der Kombination verschiedener Instrumententypen basieren, werden am Ende des Kapitels vorgestellt.

3.1 Grundlagen

Satelliteninstrumente wurden und werden für eine Vielzahl unterschiedlicher Anwendungen und Fragestellungen entwickelt. Einige Satelliteninstrumente und Missionen sind eher experimentell, d.h. sie dienen im wesentlichen zur Entwicklung und zum Test von Instrumenten. Andere Missionen verfolgen eine gezielte Fragestellung, z.B. eine genaue Kartographie des Meeresbodens. Diese Missionen sind in der Regel nur über einen gewissen Zeitraum aktiv. Eine weitere Klasse von Instrumenten und Missionen dient dem Monitoring, d.h. der kontinuierlichen Beobachtung. Wettersatelliten oder Instrumente zur Beobachtung von Ozon oder atmosphärischer Zusammensetzung sind typische Beispiele.

Trotz der Vielfalt dieser Instrumente haben alle mit ähnlichen Randbedingungen zu kämpfen: (a) der Transmission der beobachteten elektromagnetischen Strahlung durch die Atmosphäre und (b) den eingeschränkten Ressourcen an Bord eines Satelliten. Wir werden beide Punkte kurz ansprechen und dann mit einer Klassifikation der Instrumente beginnen.

3.1.1 Atmosphärische Transmission

Die Atmosphäre ist für die Erdfernerkundung einerseits Forschungsobjekt, andererseits auch Störobjekt: bei klassischen kartographischen Fragestellungen wird an der Erdoberfläche reflektiertes Sonnenlicht betrachtet, d.h. das Licht muss zweimal durch die Atmosphäre bevor es detektiert wird.

Die Atmosphäre

Da die Atmosphäre auch Untersuchungsgegenstand der Erdfernerkundung ist, lohnt es sich, ihre wesentlichen Merkmale hier kurz zusammen zu fassen.

Abbildung 3.1 gibt einen Überblick über die vertikale Struktur der Atmosphäre. Diese wird an Hand des Temperaturprofils in verschiedene Schichten, vulgo Stockwerke, eingeteilt:

(1) Die *Troposphäre* ist die Wetterschicht des Planeten. Sie ist die einzige Schicht, in der sich in nennenswertem Maße Wasserdampf findet – und sie ist die Schicht, die vom Menschen

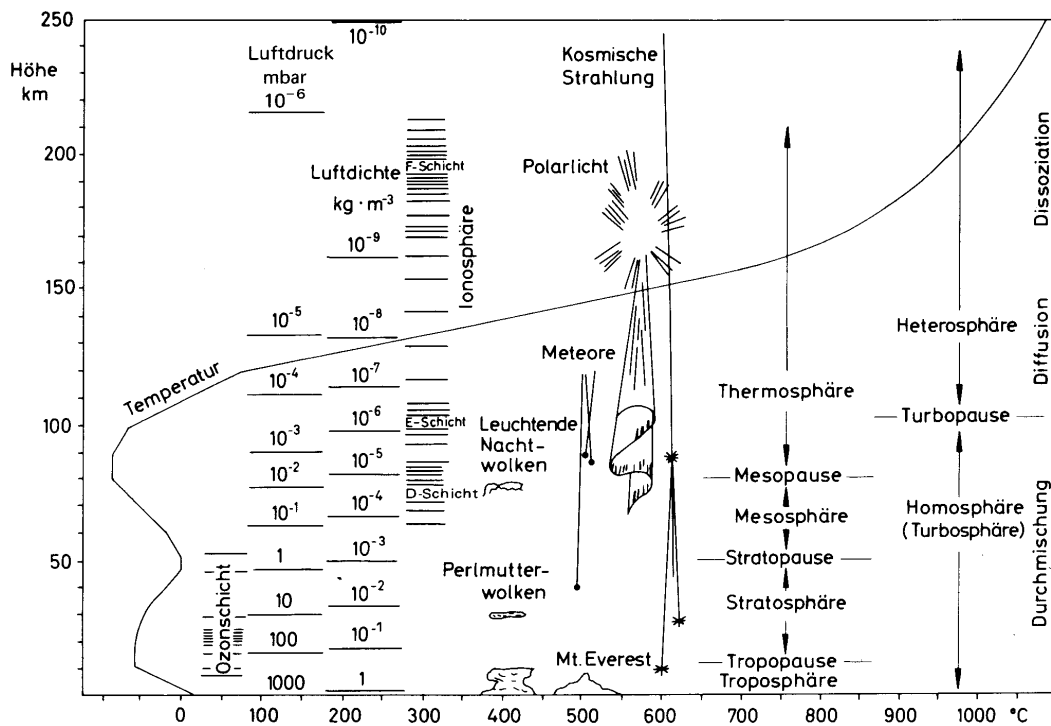


Abbildung 3.1: Aufbau der Atmosphäre [69]

am direktesten beeinflusst wird durch den Eintrag von (Spuren-)Gasen und Schadstoffen ebenso wie durch thermische Emission. Die Troposphäre hat einen negativen Temperaturgradienten von ca. 6.5 K/km. Die Oberkante der Troposphäre befindet sich in einer Höhe zwischen 16 km (Äquator) und 8 km (Pole) und ist als Inversion im Temperaturverlauf kenntlich. Daher liegt diese *Tropopause* gleichsam wie ein Deckel auf der Troposphäre und behindert den Stoffaustausch mit der darüber liegenden Stratosphäre.

(2) Die *Stratosphäre* ist durch einen positiven Temperaturgradienten gekennzeichnet. Diese Temperaturzunahme erfolgt im wesentlichen durch die Absorption solarer UV-Strahlung durch das Ozon – die Ozonschicht ist ein wesentlicher Bestandteil der Stratosphäre, auch wenn Ozon nur ein Spurengas ist. Die Stratosphäre erstreckt sich bis in eine Höhe von ca. 50 km und wird begrenzt von der *Stratopause*. Da hier der Temperaturgradient wieder negativ wird bildet die Stratopause keine Inversion, ein Transport über die Stratopause ist also möglich.

(3) Die *Mesosphäre*, manchmal auch als Ignorosphäre bezeichnet, erstreckt sich in eine Höhe bis ca. 90 km. Dies ist die Atmosphärenschicht über die am wenigsten bekannt ist, da sie für eine direkte Messung mit Flugzeugen oder Ballonen nicht zugänglich ist und eine indirekte Messung mangels besonderer Spurengase oder sonstiger Eigenarten ebenfalls nicht möglich ist. Ein ästhetischer, vielleicht auch umweltrelevanter Aspekt der Mesosphäre sind die nachtleuchtenden Wolken, dünne Eisschleier in Höhen von ca. 85 km. In der Mesosphäre herrscht wieder ein negativer Temperaturgradient, d.h. die Temperatur nimmt mit zunehmender Höhe ab.

(4) Ab einer Höhe von ca. 90 km beginnt die *Thermosphäre*. Hier nimmt die Temperatur wieder zu bis auf Werte von bis zu 2000 K. Die Erwärmung entsteht durch die Absorption der harten elektromagnetischen Strahlung der Sonne sowie über den Polkappen auch durch einfallende energiereiche Teilchen (z.B. Polarlicht). Die Thermosphäre ist stark durch die solare Einstrahlung beeinflusst, ihre Eigenschaften zeigen daher einen deutlichen Tag-Nacht-Gang, eine Variation mit den Jahreszeiten und natürlich Veränderungen im Laufe des 11-Jahre dauernden Solarzyklus. Die Thermosphäre hat keine Oberkante, ab 500 km Höhe beginnt jedoch

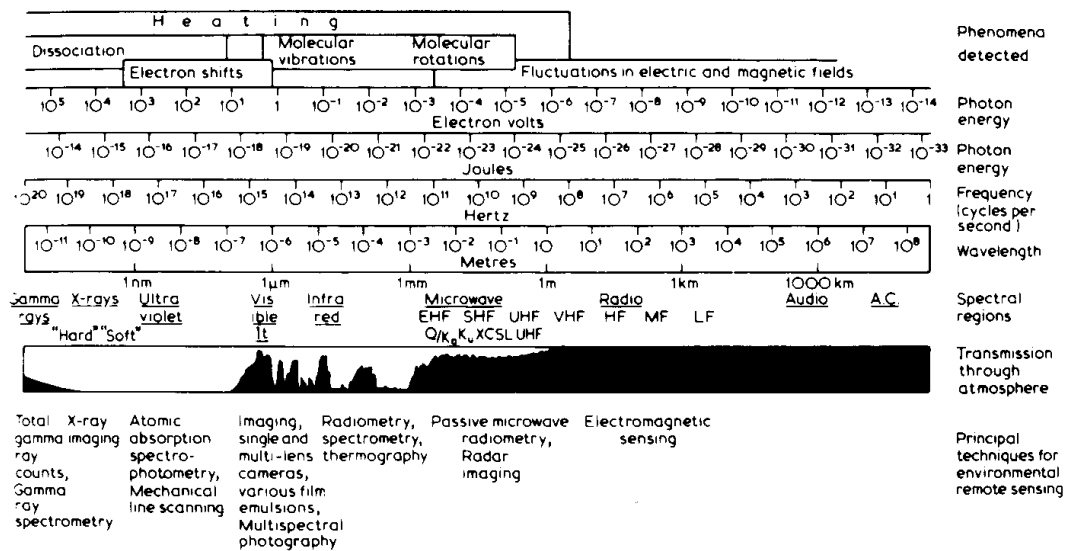


Abbildung 3.2: Elektromagnetische Spektrum, nachweisbare Phänomene und atmosphärische Transmission, [16] auf der Basis von [4]

die *Exosphäre*: hier haben Wasserstoffatome eine hinreichend große thermische Energie und auf Grund der geringen Dichte auch sehr große mittlere freie Weglängen und können daher in den Weltraum entweichen. Dies ist als *Geokorona* in Satellitenaufnahmen zu erkennen.

Die Erwärmung der Thermosphäre beruht auf der Emission harter elektromagnetischer Strahlung. Der grundlegende physikalische Prozess ist dabei die Ionisation (Photoeffekt und mit zunehmender Energie auch Compton-Effekt). Daher ist ein Teil der Thermosphäre ionisiert. Diese Schicht, in der neben Neutralgas auch Ladungsträger auftreten können, wird als *Ionosphäre* bezeichnet. Sie beginnt in einer Höhe von ca. 75 km, wobei dort der relative Anteil der Ladungsträger nur sehr gering ist. Der Ionisationsgrad steigt mit zunehmender Höhe, ab ca. 250 km ist die Atmosphäre vollständig ionisiert. Die Ionosphäre beeinflusst die Ausbreitung von Radiowellen und ist daher für die Kommunikation zwischen Satellit und Bodenstation wichtig.

Als Konsequenz ändert sich auch die Zusammensetzung der Atmosphäre. Bis in eine Höhe von 100 km ist die Ionisation gering und die Dichte groß genug, um für eine Durchmischung aller Gasbestandteile zu sorgen. In dieser Homosphäre besteht die Luft zu 78% aus N_2 , 21% O_2 und ca. 1% Ar sowie in der Troposphäre bis zu 4% Wasserdampf. Außerdem enthält die Atmosphäre Spurengase wie Ozon, Kohlendioxid und Methan. Oberhalb dieser Höhe beginnt die Heterosphäre – die geringere Dichte und damit reduzierte Kollisionsrate zwischen den Molekülen verhindert eine effiziente Durchmischung der verschiedenen Molekül- und Atomsorten, so dass diese Entmischen und sich mit eigener Skalenhöhe anordnen: die schweren Moleküle weiter unten, in größeren Höhen überwiegen leichte Moleküle und Atome (Sedimentation). Außerdem bewirkt die zunehmende Ionisierung eine Dissoziation der Moleküle, so dass mit zunehmender Höhe der Anteil der Atome gegenüber dem der Moleküle zunimmt.

Atmosphärische Transmission

Um die Möglichkeiten und Einschränkungen der Erdfernerkundung zu verstehen, müssen wir uns mit der atmosphärischen Transmission auseinandersetzen. Abbildung 3.2 zeigt das elektromagnetische Spektrum, die in verschiedenen Bereichen nachweisbaren Phänomene und die atmosphärische Transmission.

Im sichtbaren Bereich ist die atmosphärische Transmission nahezu eins, eine Einschränkung entsteht lediglich durch die Streuung des Lichtes an den Luftmolekülen. Diese ist ins-

besondere bei großen Frequenzen, d.h. im blauen Bereich des sichtbaren Spektrums, von Bedeutung. Geht man weiter zu kürzeren Wellenlängen, so geht die Transmission schnell gegen Null: die UV-Strahlung wird im wesentlichen von stratosphärischem Ozon absorbiert. Härtere Strahlung (EUV, Röntgen, γ) wird bereits in der höheren Atmosphäre absorbiert und sorgt dort für die Bildung der Ionosphäre. Auf Grund dieser Absorption ist es weder möglich, Röntgen- oder Gammaastronomie vom Erdboden aus zu betreiben noch den Erdboden mit Röntgen- oder γ -Instrumenten zu untersuchen.¹ γ s mit sehr hoher Energie können die Atmosphäre wieder passieren, sie erzeugen in der unteren Atmosphäre durch Paarbildung Elektronen und Positronen, die ihrerseits wieder γ s erzeugen, die ihrerseits wieder Paarbildung erfahren, die usw. Der Prozess wird als Luftschauer bezeichnet, ist jedoch auf Grund des seltenen Einfalls von γ s mit hinreichender Energie für die Erdfernerkundung nicht von Bedeutung.

Auch zum langwelligen Ende des Spektrums wird die Transmission schnell gering. Im Infrarot finden sich starke Absorptionsbanden. Diese Strahlung regt in der Atmosphäre mehratomige Moleküle zum Schwingen an und wird daher absorbiert.² In einigen Bereichen im Infrarot, in sogenannten Fenstern, ist die atmosphärische Transmission groß. Diese Bereiche werden zur Detektion der vom Erdboden bzw. Objekten am Erdboden emittierten Wärmestrahlung genutzt.

Mit weiter zunehmender Wellenlänge wird die Atmosphäre wieder nahezu transparent, in diesem Bereich werden aktive und passive Radiowelleninstrumente eingesetzt.

Tabelle 3.1 fasst typischerweise in der Erdfernerkundung verwendete Instrumente, die von ihnen betrachteten Wellenlängenbereiche sowie einige Anwendungen zusammen.

3.1.2 Technische Randbedingungen

Satelliteninstrumente verwenden Messprinzipien, die auch in Laborexperimenten bzw. teilweise im täglichen Leben Anwendung finden. Allerdings unterscheiden sich die Satelliteninstrumente von ihren erdgebundenen Verwandten in einigen Eigenschaften:

- Satelliteninstrumente müssen leicht sein – selbst ein moderner Satellit wie EnviSat mit einer Masse von 8140 kg hat nur eine Gesamtnutzlast von 2150 kg für die wissenschaftlichen Instrumente, Interfaces und Kommunikation. Ein typisches Teilchenteleskop mit weitem Energiebereich auf einer interplanetaren Raumsonde dagegen muss mit einer Masse von ca. 1 kg auskommen.
- Satelliteninstrumente müssen einen geringen Stromverbrauch haben. Selbst bei einer Solarzellenfläche von 70 m² erzeugt EnviSat nur eine Leistung von maximal 6.5 kW mit einer mittleren Leistung von 1.9 kW.
- Satelliteninstrumente müssen mechanisch und thermisch sehr stabil sein: beim Start mit einer Rakete oder einem Shuttle wirkt das Mehrfache der Erdbeschleunigung als systematische Kraft. Zusätzlich treten starke stochastische Kräfte durch Vibrationen und Rütteln auf. Die thermische Belastung ist beim Start gering (der Satellit befindet sich zu der Zeit noch in einer schützenden Umhüllung), erweist sich dagegen im Orbit als zu berücksichtigendes Problem: auf Grund der Ausrichtung der Instrumente auf die Erde befindet sich für längere Zeiträume eine Seite des Satelliten im Schatten während die andere der solaren Einstrahlung ausgesetzt ist. Entsprechend groß werden die Temperaturgradienten und damit auch mechanische Spannungen.

¹Es gibt eine kleine Einschränkung zu dieser Aussage. Beim Durchgang durch Materie wird elektromagnetische Strahlung nach dem Bouguer–Lambert–Beer’schen Gesetz abgeschwächt: $I = I_0 e^{-\mu x}$ mit μ als Absorptionskoeffizient und x als der Schichtdicke der durchsetzten Materie. Sehr starke Röntgen- oder Gammaquanten können daher auch nach Durchdringen der Atmosphäre in einem Detektor noch ein messbares Signal erzeugen. So wurden z.B. zur Zeit des Kalten Krieges auf militärischen Satelliten auch γ -Detektoren geflogen, um die bei Kernwaffenexplosionen emittierte Gammastrahlung nachzuweisen, z.B. Nachweis der südafrikanischen Bemühungen um Kernwaffen. Das Verfahren ist jedoch nur sinnvoll für atmosphärische Kernwaffentests, unterirdische Tests lassen sich durch die von ihnen ausgelösten Wellen besser im Seismographennetzwerk nachweisen.

²Die wesentlichen Absorber sind Wasserdampf, Kohlendioxid, Methan und Stickoxide, d.h. die Spurengase, die wir in der Klimadiskussion als treibhauswirksame Spurengase bezeichnen.

Remote Sensor Types	Part of Spectrum utilized	Remarks/Use Note: 1 Å = 10 ⁻¹⁰ m; 1 µm = 10 ⁻⁶ m
Scintillation Counters Gamma-ray spectrometers Geiger counters	<0.03 - 0.100 Å	Measurement of emitted natural radiation by gamma-ray detectors or Geiger counters
Scanners with photomultipliers Image orthicons and cameras with filtered IR film >2900 Å	100 Å - 0.4 µm	Records incident natural radiation. Imaged ultra-violet spectroscopy available.
Cameras Using conventional B&W and color film Using infrared film (B&W and IR color)	4,000 - 7,000 Å (0.4 - 0.7 µm) 6,000 - 9,000 Å (0.6 - 0.9 µm)	B&W film for high spatial detail Improved spatial detail through contrast. Greater reflectance gradients useful for vegetation surveys
Multispectral units	3,000 - 9,000 Å (0.3 - 0.9 µm)	Individual narrow band scenes available with multi-camera systems
Lidar Laser radar	4,000 - 11,000 Å (0.4 - 1.1 µm)	Monochromatic active systems for measuring backscattered radiation from the atmosphere (particles)
Radiometers	Usually IR and microwave bands	Generally measures total radiation in a wide band in the infrared or microwave regions. Imagery obtained by scanning techniques.
Photometers	4,000 - 7,000 Å (0.4 - 0.7 µm)	Measures luminous flux in various bands of the optical region for distribution, color, etc.
Spectrometers	In any spectral region	Narrow-band data available sequentially - Electromagnetic radiation amplitude vs frequency
Solid State Detectors Single detectors, Line arrays Matrices	1 µm - 1 mm	Single detecting element used in scanners, radiometers. 1-D and 2-D arrays for sequential data gathering.
Radars	1 mm - 0.8 m	Narrow band active systems. Both analog and imagery available.
Radiometers (microwave)	1 mm - 0.8 m	Passive systems. Both analog and imagery available.

Tabelle 3.1: Übersicht über Sensoren der Fernerkundung, typische Wellenlängen- und Anwendungsbereiche [62]

- Satelliteninstrumente (und insbesondere ihre Elektronik) müssen unempfindlich gegen Strahlenschäden sein. Diese entstehen durch die Ionisation durch energiereiche geladene Teilchen der kosmischen Strahlung, aus solaren Ereignissen oder aus den Strahlungsgürteln der Magnetosphäre. Am stärksten gefährdet sind Satelliten, die die Strahlungsgürtel der Erde durchfliegen – letztere geben damit gleichsam eine Obergrenze für den Orbit vor.
- Die Datenübertragung von Satelliteninstrumenten zu Bodenstationen ist eingeschränkt. So verfügt EnviSat über eine Übertragungsrate von 2 kBit/s für den Uplink, d.h. von der Bodenstation an den Satelliten, und 4096 kBit/s für den Downlink, d.h. die eigentliche Datenübertragung. Das sind sehr große Raten im Vergleich zu den wenigen Bit/s aus der Frühzeit der Raumfahrt oder bei den Voyager Missionen - im Vergleich zu den anfallenden Daten alleine im optischen Bereich sind es jedoch immer noch kleine Raten.³

3.1.3 Grobklassifizierung

Die Grobklassifikation von Instrumenten zur Erdfernerkundung erfolgt nach den Kriterien aktives oder passives Instrument und bildgebendes oder nicht-bildgebendes Verfahren.

Da Erdfernerkundung, wie der Name besagt, remote sensing ist und nicht in-situ Messung, werden alle relevanten Größen indirekt, nämlich über die ausgesandte, reflektierte oder absorbierte elektromagnetische Strahlung untersucht.

Passive Instrumente erzeugen kein Signal sondern nehmen die vom Untersuchungsgegenstand emittierte (z.B. thermisches Infrarot), partiell absorbierte (UV und Ozon, Sounding-Verfahren) oder reflektierte Strahlung (sichtbares Licht, nahes Infrarot) auf. Passive Instrumente bestehen daher aus einem Sensor und der anschließenden Auswerte-Elektronik.

³Eine simple 5 MPixel Digitalkamera erzeugt ein Bild von 2560 mal 1920 Pixel mit einer Farbtiefe von 24 Bit bzw. 8 Bit je RGB-Kanal. Als Bitmap abgespeichert hat dieses Bild eine Größe von 117 MBit oder knapp 15 MB. Selbst bei einem Downlink von 4096 kBit/s würde die Übertragung des Bildes noch fast 29 Sekunden dauern – eine Zeit, in der sich der Satellit um 1800 km entlang seiner Bodenspur bewegt hat.

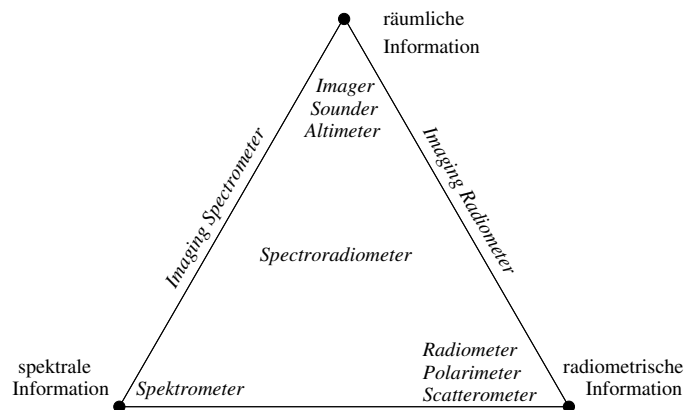


Abbildung 3.3: Beziehung zwischen verschiedenen Klassen von Sensoren

Aktive Instrumente dagegen erzeugen ein Signal (z.B. Radiowellen beim Radar) und untersuchen das reflektierte Signal bezüglich Intensität, Laufzeit, Polarisation usw. Aktive Instrumente benötigen daher neben dem Sensor auch einen Signalgeber – und natürlich für beide die notwendige Elektronik. Aktive Satelliteninstrumente stellen höhere Anforderungen an die Stromversorgung – militärische Radarsatelliten haben früher kleine Kernreaktoren an Bord gehabt, was im Falle eines Startunfalls oder Absturz⁴ zu beträchtlichen ökologischen Folgen führen kann.

3.1.4 Kenngrößen für Satelliteninstrumente

Die wichtigsten Kenngrößen von Satelliteninstrumenten sind die Auflösungsvermögen. Je nach Anforderung an das Instrument unterscheidet man verschiedene Auflösungsvermögen:

- Bodenauflösungsvermögen:
- radiometrisches Auflösungsvermögen:
- thermisches Auflösungsvermögen:
- zeitliches Auflösungsvermögen:

Schon alleine in Anbetracht der großen Datenfülle und der begrenzten zur Aufnahme zur Verfügung stehenden Zeit kann kein Instrument in allen Auflösungsvermögen optimiert sein: je mehr Pixel und Kanäle zur Verfügung stehen, um so weniger Photonen können während der Belichtungszeit in einen Kanal einfallen, so dass dieser dann durch ein hohes Rauschen und geringes radiometrisches Auflösungsvermögen wenig Informationen liefert.

Abbildung 3.3 gibt einen Überblick über die Beziehung zwischen den verschiedenen Klassen von Instrumenten und ihren jeweiligen Auflösungsvermögen: im Zentrum das Spectroradiometer als Universalinstrument, zu den Eckpunkten jeweils an spezielle Fragestellungen angepasste und im entsprechenden Auflösungsvermögen optimierte Instrumente.

3.2 Passive Instrumente im sichtbaren Bereich des elektromagnetischen Spektrums

Das klassische Instrument der Erdfernerkundung ist die Kamera mit einem konventionellen Film. Hierbei handelt es sich um ein passives Instrument, da die Kamera das vom Untersuchungsgegenstand reflektierte Sonnenlicht aufzeichnet, und um ein bildgebendes Instrument, d.h. es wird ein zwei-dimensionales Abbild des Gegenstandes erzeugt.

⁴Kosmos 954 stürzte 1978 über Kanada ab, Kosmos 1042 stürzte 1983 in den Indischen Ozean – in beiden Fällen führte die günstige Absturzposition dazu, dass keine Beeinträchtigungen der Anthroposphäre auftraten.

3.2.1 Kurzer Rückblick

Die ersten auf Raumfahrzeugen eingesetzten Erdfernerkundungsinstrumente waren ebenfalls Kameras. Das erste Bild aus dem Orbit wurde, mit einer sehr schlechten Auflösung, 1959 vom US-Satelliten Explorer-6 mit einer TV-Kamera aus einem Hohlspiegel und einem Phototransistor aufgenommen: die Rotation des spin-stabilisierten Explorer lieferte einen Scan senkrecht zur Flugrichtung (entsprechend einer Zeile eines Fernsehbildes), die Bewegung entlang seiner Bahn lieferte die einzelnen Zeilen. Die Übertragung dieses ersten Fernsehbildes dauerte nahezu 40 min. Das erste Farbbild aus dem Orbit war ebenfalls eine Video-Aufnahme, 1967 vom sowjetischen Molniya-1 Satelliten gemacht.

Der Großteil der frühen Erdaufnahmen ist wurde jedoch konventionell mit Kamera und Film gemacht: im August 1960 gelang es dem Satelliten Discoverer-14 (CORONA – militärische Aufklärungssatelliten der USA) erstmals einen im Orbit belichteten Film erfolgreich zur Erde zurück zu senden – der Film wurde in der Luft abgefangen. Während der folgenden 12 Jahre, bis März 1972, wurden 95 weitere CORONA Missionen geflogen. Dieses Konzept der Filmrückführung wurde erste 1976 mit den KH-Satelliten (Keyhole) zu Gunsten opto-elektronischer Aufnahmetechniken aufgegeben, die alten militärischen Aufnahmen wurden Mitte der Neunziger Jahre deklassifiziert,⁵ so dass sie jetzt auch zivilen Nutzern zur Verfügung stehen und damit den Zeitraum kontinuierlicher Erdbeobachtung um mehr als eine Dekade rückwärts verlegen.

Der erste Photosatellit der UdSSR, Zenit, wurde 1962 gestartet und trug ein bei VEB Carl-Zeis-Jena gebautes Instrument, die MKF-6 (Multi-Kanal-Fotografie-6), das das Grundprinzip der modernen Erdfernerkundung prägen wird: eine Multispektralkamera aus sechs, in einem 2 mal 3 Raster montierten Einzelkameras, die mit unterschiedlichen Filtern im Spektralbereich von 0.45–0.9 μm , d.h. bis ins nahe IR hinein, ausgestattet jeweils die gleiche Szene aufnehmen. Diese Kameras wurden auch auf bemannten Missionen und von Flugzeugen eingesetzt.

Konventionelle Filmaufnahmen haben den Nachteil der erforderlichen Rückführung des Filmmaterials. Daher wurde, trotz deutlich reduzierter Qualität der Aufnahmen, bereits Mitte der Sechziger Jahre mit der Verwendung von Video-Kameras experimentiert, z.B. aus den meteorologischen Satelliten der TIROS-Serie (erster Start 1960), auf den ersten Landsats (an 1972) sowie auf ESSA (1966), DMSP (1965) und Meteor (1969).

Den Durchbruch in der Datenaufnahme brachten die CCDs, die die Entwicklung fortschrittlicher Scanner ebenso wie der Multispektralverfahren erlaubten. Die ersten auf der Basis dieser Technologie eingesetzten Instrumente wurden auf Flugzeugen eingesetzt, ihre Weiterentwicklung, der MSS, wurde mit Beginn der LandSat-Serie 1972 auch auf Satelliten geflogen. Aus diesem ‘Urmodell’ sind viele Weiterentwicklungen entstanden, von denen wir einige weiter unten genauer betrachten werden. Der LandSat Multispektralscanner MSS ist für Langzeituntersuchungen im Rahmen des global change bzw. climate change von so großer Bedeutung, dass LandSat 5 als der letzte einen MSS tragende LandSat nach über 20 Dienstjahren immer noch aktiv ist und selbst sein etwas älterer Bruder LandSat 4 im Standby-Betrieb gehalten wird, um gegebenenfalls LandSat 5 bei einem Ausfall des Satelliten bzw. des MSS zu ersetzen.

3.2.2 Kameratypen

Ein konventionelles photographisches System besteht aus einer Kamera mit einem Objektiv und einer photographischen Emulsion. Die Eigenschaften dieses Systems werden durch die Art der Kamera, die Abbildungseigenschaften des Objektivs und die Eigenschaften der photographischen Emulsion beschrieben.

Kameras lassen sich einteilen in Einzelobjektivkameras, mehrlinsige Kameras (Multispektralkameras), Streifenkameras und Panoramakameras. Bei der *Einzelobjektivkamera* handelt

⁵Nähere Informationen und Beispielbilder unter <http://edc.usgs.gov/guides/displ.html> und <http://edc.usgs.gov/products/satellite/declass1.html>, andere Beispiele und eine Diskussion militärischer Aufklärung, inkl. U-2, unter <http://rst.gsfc.nasa.gov/Intro/Part2.26e.html>

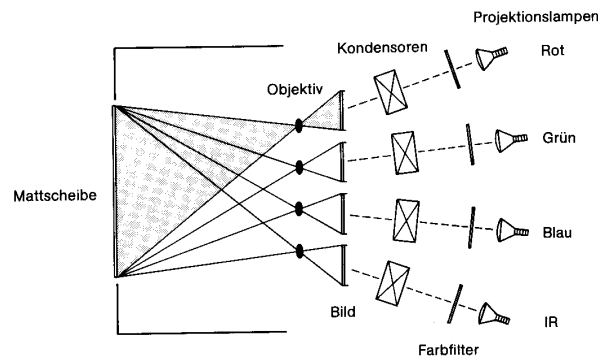


Abbildung 3.4: Farbmischprojektor: Rekonstruktion eines Farb- oder Falschfarbenbildes-Falschfarbenbild aus mehreren mit unterschiedlichen Filtern belichteten S/W-Filmen [71]

es sich um einen ganz normalen Photoapparat. Um ein gutes Bodenaufklärungsvermögen zu erreichen, werden meistens größere Negativformate als in der konventionellen Photographie verwendet mit Filmbreiten bis zu 19 cm [61, 128], verglichen mit den 2.4 cm eines konventionellen Rollfilms. Derartige Kameras werden vielfach als Handkameras bei bemannten Missionen (Gemini, SkyLab, Spacelab, Space-Shuttle) verwendet, auf dem Space-Shuttle hat sich eine modifizierte Hasselblad-Mittelformatkamera mit 6 cm Negativformat durchgesetzt.

Die *mehrlinsige Kamera* wird auch als *Multispektralkamera* bezeichnet. Sie besteht aus mehreren, miteinander gekoppelten identischen Kameras, die gleichzeitig ausgelöst werden. Jedes der Objektive hat den gleichen Blickwinkel, die gleiche Brennweite und beobachtet das gleiche Gebiet. Jedoch ist jedes Objektiv mit einem anderen Filter versehen, so dass gleichzeitig Aufnahmen eines Gebietes in verschiedenen Bereichen des elektromagnetischen Spektrums gemacht werden. Je nach Filter können in den einzelnen Kameras der Multispektralkamera auch unterschiedliche Filmmaterialien verwendet werden. Anstelle eines einzigen Farb- oder Schwarzweißfilms erhält man auf diese Weise Schwarzweißfilme, die sich in Grauton und Dichte je nach Intensität der einzelnen Spektralbereiche unterscheiden. Der entscheidende Vorteil der Multispektralkamera gegenüber einer einlinsigen Kamera mit einem Farbfilm liegt in der photographischen Emulsion: Schwarzweißfilme sind empfindlicher und feiner als Farbfilme, was zu einem besseren Auflösungsvermögen und Kontrast führt. Die Farbinformation wird durch die verschiedenen Filter vor den einzelnen Kameras erzeugt. Zusätzlich haben diese Kameras den Vorteil, dass sie bei entsprechenden Filtern auch in Wellenlängenbereichen sehen können, in denen das menschliche Auge nicht empfindlich ist (nahes Infrarot), und dass Wellenlängen, die stark dunstempfindlich sind (blau), beim späteren Zusammenmischen weggelassen werden können. Dadurch wirkt sich die atmosphärische Streuung nicht so negativ auf das Auflösungsvermögen aus, für Beispiele siehe z.B. [101].

Das Original lässt sich aus diesen Bildern durch die Verwendung eines Farbmischprojektor rekonstruieren (vgl. Abb. 3.4). Die auf diese Weise zusammengemischten Bilder können in den 'echten' Farben erzeugt werden oder je nach Kombination von Filmen und Filtern auch als *Falschfarbenbilder*. Letztere Methode kann zur besseren Darstellung relativ geringer Kontraste oder zur Darstellung des Infrarotanteils verwendet werden.

Die Streifenkamera wird nur in der militärischen Aufklärung verwendet. Hier bleibt der Verschluss der Kamera offen und der Film wird mit hoher Geschwindigkeit bewegt. Bei der Panoramakamera ist das Objektiv beweglich oder einem starren Objektiv ist ein bewegliches Prisma vorgesetzt, wodurch das Blickfeld des Satelliten in Streifen, die senkrecht zur Flugrichtung liegen, abgetastet wird. Der Nachteil dieses Verfahrens ist die Verzerrung zu den Seiten dieser Streifen, der Vorteil liegt in der schnellen Abtastung großer Flächen (normalerweise wird nur ein relativ schmaler Streifen entlang der Bodenspur von der Kamera eingesehen).

3.2.3 Abbildung durch Objektive

Photographische Objektive sind aus mehreren Linsen aufgebaut, da (a) eine einzelne Linse zu dick wäre und (b) auf diese Weise Linsenfehler und chromatische Effekte verringert werden können. Die Abbildung lässt sich durch die Linsengleichung beschreiben. Mit f als der Brennweite der Linse, g als dem Abstand zwischen Objekt und Linse (Gegenstandsweite, entspricht der Flughöhe h) und b als dem Abstand zwischen Linse und Filmebene (Bildweite) gilt dann

$$\frac{1}{f} = \frac{1}{g} + \frac{1}{b}. \quad (3.1)$$

Bei Satellitenbeobachtungen betragen die typischen Objektweiten g einige hundert Kilometer, während die Brennweiten und der Abstand Linse – Filmebene nur im Bereich von Zentimetern oder Dezimetern liegen. Dann gilt $f \approx b$, d.h. der Abstand Linse – Filmebene ist gleich der Brennweite. Für das Verhältnis aus Gegenstandsgröße G und Bildgröße B (Abbildungsmaßstab) ergibt sich

$$\frac{B}{G} = \frac{b}{g} = \frac{b}{h}, \quad (3.2)$$

bzw. unter Berücksichtigung von $f \approx b$

$$\frac{B}{G} = \frac{f}{h}. \quad (3.3)$$

Für den Abbildungsmaßstab, und damit auch für das Bodenaufhebungsvermögen, ist daher bei vorgegebener Flughöhe die Brennweite von entscheidender Bedeutung. Für zivile Satellitenkameras liegen die Brennweiten im Bereich von 44 mm (Multispektralkamera UdSSR [128]) über 80 mm (NASA-Hasselblad, z.B. Apollo) und 305 mm (Zeiss-Kamera, Spacelab [71]) bis hin zu 450 mm (Earth Terrain Camera auf SkyLab [101]). Insbesondere im militärischen Bereich werden auch Objektive mit Brennweiten im Bereich von Metern verwendet (Spiegelobjektive). In Abhängigkeit von der Brennweite verändert sich der Öffnungswinkel der Objektive und damit die Fläche, die eingesehen werden kann. Für die oben gegebenen Beispiele liegen die Blickwinkel zwischen ca. 135° und 45° .

Das Blickfeld F_B einer Satellitenkamera lässt sich aus (3.3) durch Ersetzen der Bildgröße B durch die Kantenlänge k des Films bestimmen

$$F_B = \frac{g \cdot k}{f} = \frac{h \cdot k}{f}. \quad (3.4)$$

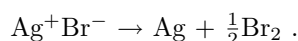
Für typische zivile Anwendungen liegen die Blickfelder in Bereichen zwischen 80 und 200 km.

3.2.4 Aufzeichnung der Abbildung: photographische Emulsion

Die Aufzeichnung mit photographischen Emulsionen wird in der modernen Erdfernerkundung von bemannten Stationen aus betrieben (Filmrückfuhr lässt sich einfach lösen), auf einigen Aufklärungssatelliten soll es immer noch Filmmaterialien geben.

Schwarz-Weiß-Film

Die wichtigste photographische Emulsion ist der Schwarzweißfilm. Dieser besteht aus einer Gelatineschicht mit einer Dicke von ca. $100 \mu\text{m}$, in die Silberhalogenidkörnchen (AgBr, auch AgJ oder AgCl) von einer Größe von 1 bis $10 \mu\text{m}$ eingebettet sind. Diese Schicht befindet sich auf einer Trägerschicht, die die mechanische Stabilität gewährleisten soll. Auftreffende elektromagnetische Strahlung bewirkt durch Ionisation des Br^- (Photoeffekt) eine Aufspaltung in Silber- und Bromatome:



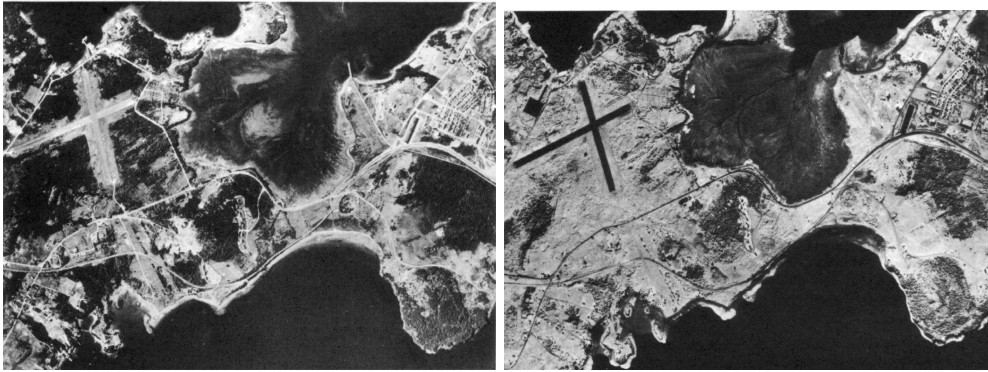


Abbildung 3.5: Panchromatische Aufnahme (links) im Vergleich zur Aufnahme mit einem Infrarotfilm [101]

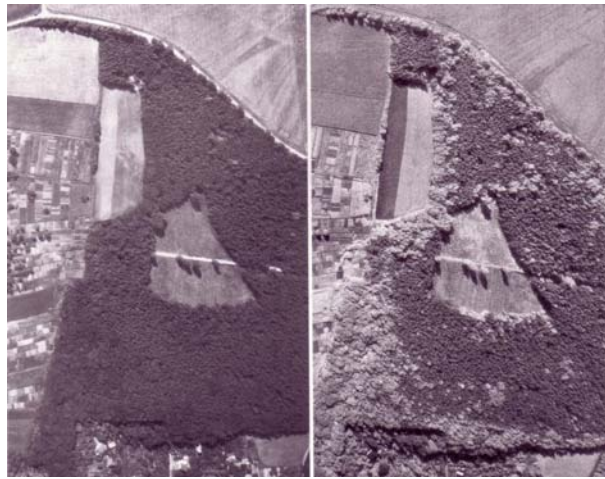


Abbildung 3.6: Vegetation panchromatisch und im Infrarotfilm [71]

Normalerweise kehrt sich diese Reaktion von selbst wieder um. Wenn jedoch innerhalb eines Körnchens eine ausreichende Zahl von Silberatomen (typischerweise 4) geformt wurde, ist dieser Vorgang irreversibel und der Film enthält ein latentes Bild. Dieses latente Bild wird durch die Entwicklung fixiert (Negativ). Da zur Auslösung der Reaktion das Photon eine Mindestenergie haben muss, gibt es auch eine maximale Wellenlänge, bis zu der ein Schwarzweißfilm empfindlich ist. Diese liegt im Bereich von 0.4 bis $0.5 \mu\text{m}$. Schwarzweißfilme sprechen daher auf den blauen, violetten und den UV-Teil des Spektrums an. Panchromatische Schwarzweißfilme werden besonders empfindlich gemacht, um auf den ganzen Bereich des sichtbaren Spektrums anzusprechen. Da der UV- und der Blauanteil stark dunstempfindlich sind, werden diese Wellenlängen häufig durch den Vorsatz eines Filters ausgeschaltet. Allerdings gibt es auch Filme, die gezielt für die UV-Photographie gedacht sind. Diese können nur bei sehr klarem Wetter und bei geringen Flughöhen eingesetzt werden. Ein Anwendungsbeispiel ist die Feststellung von Ölflecken auf Wasser.

Beim *Infrarotfilm* wird die Filmempfindlichkeit bis in den infraroten Bereich ausgedehnt, allerdings bleibt der Film auch für den gesamten sichtbaren und den UV-Bereich empfindlich. Häufig wird daher der für atmosphärische Streuung besonders empfindliche kurzwellige Bereich durch Filter ausgeschaltet. Der Film zeigt dann in verschiedenen Graustufen den Bereich von grün bis zum nahen Infrarot, die maximale Wellenlänge liegt im nahen Infrarot bei $0.9 \mu\text{m}$. Die Verwendung eines Infrarotfilms eignet sich insbesondere für Vegetationsuntersuchungen aber auch zur Kontrastanreicherung zwischen Vegetation und anderen Strukturen:

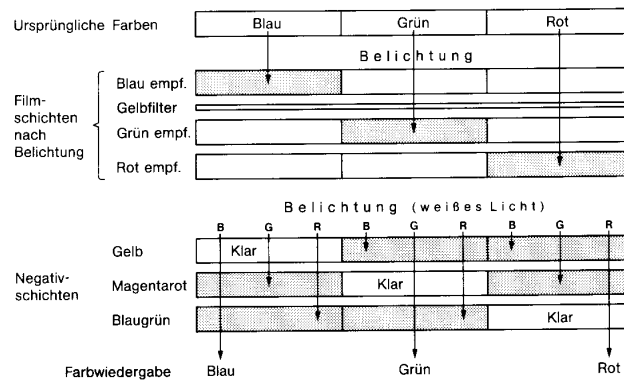


Abbildung 3.7: Schichtstruktur eines Farbfilms [71]

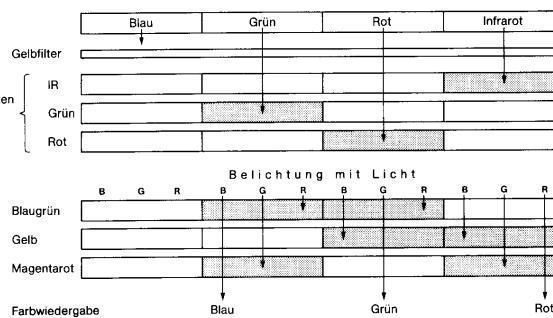


Abbildung 3.8: Schichtstruktur eines IR-Farbfilms [71]

so ist im der panchromatischen Aufnahme im linken Teil von Abb. 3.5 der Kontrast zwischen Landebahn und umgebenden Grasfläche nur gering während im Infrarotfilm auf der rechten Seite die Landebahn deutlich gegen das Gras abgehoben ist.

Aber nicht nur der Unterschied zwischen Vegetation und Landebahn wird im Infrarotfilm deutlicher, auch die Details der Vegetation. Während in der panchromatischen Aufnahme im linken Teil von Abb. 3.6 die Bäume alle gleich aussehen, zeigen sich im Infrarotfilm (rechtes Teilbild) deutliche Unterschiede zwischen den Pflanzen am Rand des Wäldchens und denen in dessen Zentrum, Zeichen für Unterschiede in der Pflanzengesundheit und/oder -ernährung.

Farbfilm

Farbfilme dagegen bestehen aus drei oder mehr photosensitiven Schichten, die in unterschiedlichen Spektral- bzw. Farbbereichen empfindlich sind, vgl. Abb. 3.7. Die Schichten sind, von oben nach unten, blauempfindlich ($0.4\text{--}0.5\ \mu\text{m}$), grünempfindlich ($0.5\text{--}0.6\ \mu\text{m}$) und rot empfindlich ($0.6\text{--}0.7\ \mu\text{m}$). Da die rot- und grünempfindlichen Schichten auch für blaues Licht empfindlich sind, befindet sich hinter der blauempfindlichen Schicht ein Gelbfilter. Nach der Entwicklung entstehen aus diesen Schichten die Komplementfarben. Wie beim Schwarzweißfilm ist auch hier wieder die Schwärzungsdichte der einzelnen Negativschichten proportional der Intensität der einfallenden Strahlung. Durch die Belichtung mit weißem Licht auf photosensitives Papier lassen sich Positive erstellen.

Nahes Infrarot: Falschfarbenaufnahme

Ein Spezialfall des Farbfilms ist der Infrarotfarbfilm. Bei ihm ist die obere photosensitive Schicht gegen Infrarot empfindlich und nicht, wie beim normalen Farbfilm, gegen Blau (vgl. Abb. 3.8). Die Schichten decken daher den Bereich vom nahen Infrarot bis Grün ab, der

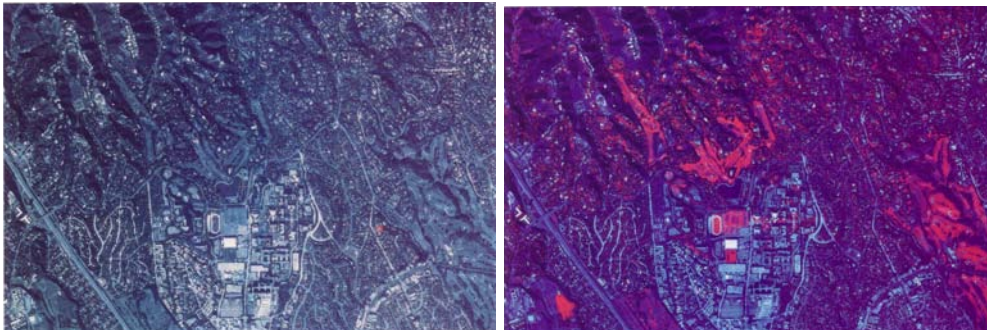


Abbildung 3.9: Vergleich einer Szene in konventioneller Farbaufnahme (links) und Falschfarbaufnahme (rechts) mit einem IR-Farbfilm [101]

Blauanteil geht, wie auch beim Infrarotfilm, vollständig verloren. Während beim normalen Farbfilm eine Wiedergabe in den normalen Farben erfolgen kann, ist dies beim Infrarotfilm nicht sinnvoll: zum einen fehlt der blaue Anteil, zum anderen ist die Information für den nicht sichtbaren Infrarotanteil vorhanden. Beim infrarotempfindlichen Schwarzweißfilm entsteht dadurch kein Problem, da die Abbildung nur in Grautönen erfolgt. Beim Farbfilm bedient man sich einer willkürlichen Zuordnung der einzelnen Schichten zu Farben, so dass für jede der Schichten eine der drei Primärfarben Blau, Grün und Rot verwendet wird. Die Zuordnung zwischen diesen Farben und den Schichten ist willkürlich. Die gängige Zuordnung bildet den infraroten Teil rot ab. Damit muss für das sichtbare Rot eine andere Farbe gewählt werden, in der Regel Grün. Für den verbleibenden grünen Bereich wird das nicht mit aufgenommene Blau verwendet. Auf diese Weise wird ein *Falschfarbenbild* erzeugt.

Bei der konventionellen Zuordnung der Schichten in der Falschfarbendarstellung wird die Vegetation auf Grund ihres hohen spektralen Reflektionsvermögens im nahen IR rot abgebildet. Ein Beispiel für eine städtische Szene ist in Abb. 3.9 gezeigt. Dort zeichnet sich z.B. das Stadium in der Bildmitte im Falschfarbenbild wesentlich deutlicher ab als im normalen Farbbild.

Anwendung einfacher photographischer Systeme

Die bisher betrachteten relativ einfachen photographischen Systeme waren in der Frühphase der Erderkundung die einzigen vorhandenen Systeme und wurden in Aufklärungssatelliten, in der Kartographie und in der Umweltforschung eingesetzt. Heute wird die einfache Photographie (wobei das Aufnahmemedium nicht mehr zwingend ein Film sein muss, sondern durch CCDs ersetzt sein kann, s.u.) im wesentlichen in der Kartographie eingesetzt. Für die meisten umweltbezogenen Fragen werden kompliziertere Systeme verwendet, allerdings können die einfachen photographischen Systeme auch in Bereichen Informationen liefern, die man vielleicht nicht unbedingt mit Erdfernerkundung in Verbindung bringt. Als ein Beispiel sei die ‘Volkszählung von oben’ erwähnt, bei der in dünn besiedelten oder mit nomadisierenden Völkern besiedelten Gebieten die Größe der Bevölkerung abgeschätzt werden kann. Bei Tagaufnahmen sind Dörfer als dunkle Flecken auf dem Bild zu erkennen. Aus der Größe dieser Flecken lässt sich eine grobe Schätzung über die Bevölkerung in den Dörfern vornehmen, aus der Gesamtzahl der Flecken eine Schätzung der Bevölkerung in einer bestimmten Region. Ein Beispiel aus dem Sudan wird in Calder [14] vorgestellt. Bei Nachtaufnahmen kann man Lichtquellen erkennen. In den Industrienationen heben sich daher die Städte bzw. Ballungsgebiete auf Nachtaufnahmen deutlich ab. Am Rand der Sahelzone dagegen gibt es keine nennenswerte Besiedlung, die hier beobachteten hellen Punkte stammen von den Feuern nomadisierender Hirtengruppen, auf deren Verteilung und Anzahl auf diese Weise zurückgeschlossen werden kann [117]. Auch Gebiete, in denen Brandrodung getrieben wird, lassen sich so erkennen.

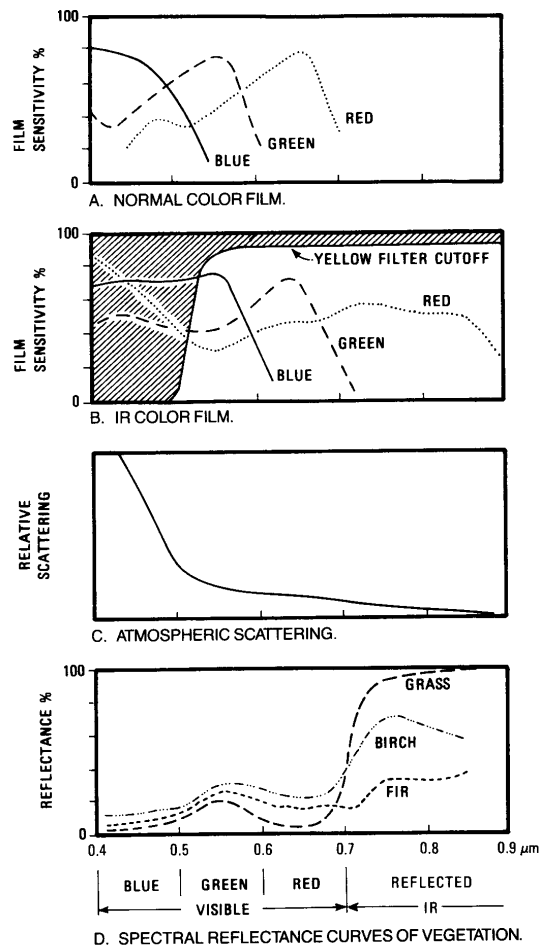


Abbildung 3.10: Spektrale Empfindlichkeit eines Farb- und eines Infrarotfarbfilmes zusammen mit der spektralen Abhängigkeit der atmosphärischen Streuung und des Reflektionsvermögens von Pflanzen [101]

3.2.5 Warum Falschfarben?

Sowohl bei den Schwarzweiß- als auch bei den Farbfilmen gibt es jeweils auch Varianten, die im nahen Infrarot empfindlich sind. Was ist nun das Interessante an diesem Bereich, der für das menschliche Auge nicht mehr wahrnehmbar ist, bzw. durch Falschfarbenbilder sichtbar gemacht werden muss? Bevor wir diesen Punkt diskutieren, sollten wir uns noch einmal klar machen, dass wir hier Sensoren besprechen, die die reflektierte Infrarotstrahlung der Sonne aufzeichnen, nicht aber um Sensoren, die das thermische Infrarot (d.h. die von terrestrischen Gegenständen abgegebene Wärmestrahlung) erkennen können.

Abbildung 3.10 zeigt dazu die spektrale Empfindlichkeit eines Farbfilms (A) bzw. eines Infrarotfarbfilms (B). Kurve (C) zeigt die Streuung in der Atmosphäre in Abhängigkeit von der Wellenlänge: Die Streuung ist am stärksten im blauen Licht (Dunststreuung), mit zunehmender Wellenlänge wirkt sich die atmosphärische Streuung immer weniger aus, so dass sich schärfere Aufnahmen erzeugen lassen.

Der zweite Vorteil der Verwendung des nahen Infrarot wird aus dem unteren Teil von Abb. 3.10 deutlich: Gerade beim Übergang vom sichtbaren zum infraroten Bereich ändern sich die Reflektionseigenschaften vieler Materialien sehr stark. Im Infraroten erzeugen Pflanzen ein relativ starkes Signal (wesentlich deutlicher als im sichtbaren Bereich, damit lässt sich Vegetation leicht von Sand, Wasser oder Gegenständen unterscheiden). Da der infrarote An-

Band	1 (0.5–0.6 μm)	2 (0.6–0.7 μm)	3 (0.7–0.8 μm)	4 (0.8–1.1 μm)
	Reflectance (%)			
<i>Rock and soil materials and covers</i>				
Sand	5.19	4.32	3.46	6.71
Loam 1% H ₂ O	6.70	6.79	6.10	14.01
Loam 20% H ₂ O	4.21	4.02	3.38	7.57
Ice	18.30	16.10	12.20	11.00
Snow	19.10	15.00	10.90	9.20
Cultivated land	3.27	2.39	1.58	(not given)
Clay	14.34	14.40	11.99	(not given)
Gneiss	7.02	6.54	5.37	10.70
Loose soil	7.40	6.91	5.68	(not given)
<i>Vegetation</i>				
Wheat (low fertilizers)	3.44	2.27	3.56	8.95
Wheat (high fertilizers)	3.69	2.58	3.67	9.29
Water	3.75	2.24	1.20	1.89
Barley (healthy)	3.96	4.07	4.47	9.29
Barley (mildewed)	4.42	4.07	5.16	11.60
Oats	4.02	2.25	3.50	9.64
Oats	3.21	2.20	3.27	9.46
Soybean (high H ₂ O)	3.29	2.78	4.11	8.67
Soybean (low H ₂ O)	3.35	2.60	3.92	11.01

Tabelle 3.2: Spektralabhängigkeit des Reflektionsvermögens für verschiedene Böden, Bodenbedeckungen und Pflanzen für die vier Bänder des Multi-Spectral-Sensors MSS [4]

teil rot abgebildet wird, erscheint Vegetation auf einem Falschfarbenbild rot. Das Verhältnis von reflektierter Strahlung im sichtbaren Bereich zur reflektierten Strahlung im nahen Infrarot kann sogar verwendet werden, um verschiedene Vegetationsformen zu unterscheiden: so reflektiert ein Tannenwald im Infraroten nur etwas mehr als im roten Bereich, eine Birke bereits deutlich mehr und Gras reflektiert im roten Bereich nur sehr wenig Strahlung, im Infraroten dagegen sehr viel. Tabelle 3.2 gibt eine Übersicht über die Reflektivität verschiedener Böden und Pflanzen in vier verschiedenen Frequenzbereichen. Die Angabe der Frequenzbereiche erfolgt in Übereinstimmung mit den vier Spektralbändern des Multi-Spectral-Sensor MSS auf ERTS-1 (Earth Resources Technology Satellite - später auf LandSat 1 umgetauft).

Verschiedene Pflanzen und insbesondere der Gesundheitszustand dieser Pflanzen lassen sich aufgrund der Abhängigkeit des Reflektionsvermögens im Infraroten von der typischen Blattstruktur der Pflanzen erkennen. Die Hauptrolle spielt dabei das Mesophyll: Die Grenzschichten zwischen kleinen Luft- und Wasserbläschen in diesem Gewebe reflektieren das Infrarot besonders gut (vgl. Abschn. 3.6.2, insbesondere Abb. 3.45 und Abb. 3.46). Aus dem Infrarotbild lässt sich eine Abschätzung über Wachstums- und Reifegrad bzw. über die Pflanzengesundheit ableiten: 'gestresste' Pflanzen reflektieren weniger Infrarot, da sich der Wassergehalt in ihnen verringert hat und damit die internen Blattstrukturen verändert sind, vgl. z.B. [101]. Angewendet wird dieses Verfahren in der Waldschadensforschung ebenso wie in der Ernteüberwachung. Auch lassen sich Informationen über die in einem bestimmten Gebiet angebauten Pflanzen gewinnen. Diese Interpretationen sind jedoch relativ schwierig und erfordern hohe Genauigkeit. Satelliteninstrumente werden daher so geeicht, dass Aufnahmen von großflächigen Monokulturen angefertigt werden und als Eichkurven dienen [61], vgl. Abschn. 3.6.2 und Abschn. 6.3.

So attraktiv dieses Verfahren für die Erforschung der Biosphäre auch ist, es ist immer noch ein großer Aufwand in der Forschung und Modellierung notwendig, um die Erkennung

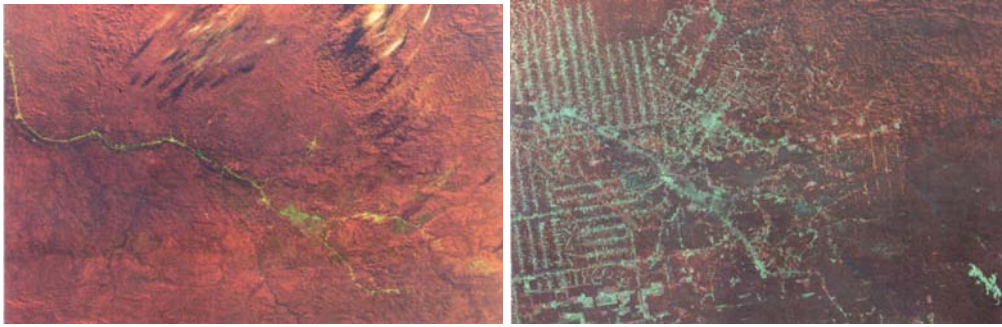


Abbildung 3.11: Abholzung des Amazonas: rot sind Waldflächen, grün in den Wald geschlagene Schneisen. Die linke Aufnahme stammt von 1973, die rechte von 1987 [14]

verschiedener Vegetationsformen und des Zustandes dieser Pflanzen auf große Datensätze anzuwenden. Insbesondere hängen die Reflektionseigenschaften der Pflanzen von den Beleuchtungsbedingungen (Sonnenstand, Wolken, spektrale Verteilung der Einstrahlung), dem Beobachtungsort (Hanglage, meteorologische und klimatische Bedingungen, Aerosole, Wasserdampfgehalt der Atmosphäre) und natürlich den Eigenschaften der Pflanzen ab (Form der Pflanzen, Dichte und Verteilung der Pflanzen, Untergrund, Temperatur, tageszeitliche Variationen). Probleme entstehen insbesondere bei stark durchmischter Vegetation und in Laubwäldern z.B. durch Wind, da Blätter in Abhängigkeit von ihrer Stellung völlig unterschiedliche Reflektionseigenschaften haben können (Extremfall der Pappel mit relativ dunkler grüner Oberfläche und fast silberner Blattunterseite).

Ein wichtiger Anstoß zur Entwicklung der Infrarotphotographie stammt aus dem militärischen Bereich. Bei Luftaufnahmen im sichtbaren Bereich ist es schwierig, einen Panzer unter einem Tarnnetz von natürlicher Vegetation zu unterscheiden. Im nahen Infrarot dagegen gibt sich der Panzer dadurch zu erkennen, dass sein Tarnnetz nicht wie für Vegetation typisch, ein hohes Reflektionsvermögen besitzt.

Anwendungsbeispiele

Abbildung 3.11 zeigt ein Beispiel für die Anwendung von Falschfarbenaufnahmen im Umweltmonitoring: die Abholzung des Amazonas. Rot sind die Waldflächen dargestellt, grünlich die Einschlagschneisen. Die obere Aufnahme stammt aus dem Jahre 1973 und zeigt im wesentlichen eine Schneise, wahrscheinlich einem Fluss oder einer anderen natürlichen Struktur folgend. Die rechte Aufnahme stammt aus dem Jahre 1987. Die Schneise von 1973 ist noch als Referenz zu identifizieren, die rechtwinklig abzweigenden Schneisen sind eindeutig menschlichen Ursprungs und markieren Abholzung und Zugang zu Einschlag- und Rodungsbereichen.

3.2.6 Charakterisierung eines photographischen Systems

Die Eigenschaften eines Photoapparates beschreibt man durch die Brennweite und Lichtstärke des Objektivs, die Verschlussgeschwindigkeit und die Geschwindigkeit des Films. Für jede Aufnahmesituation kann man dann in Abhängigkeit von der Beleuchtung und dem Abstand zum Objektiv einige Eigenschaften der Aufnahme, wie Blickwinkel, Auflösungsvermögen und Kontrast, abschätzen. Im Prinzip könnte man die gleichen Kriterien auch bei einem Aufnahmesystem auf einem Satelliten verwenden. Da sich bei diesem aber der Abstand zum Objekt nicht ändert (Erderkundungssatelliten haben kreisförmige Bahnen) und die Lichtverhältnisse ebenfalls nahezu konstant sind (moderne Erderkundungssysteme überfliegen den Äquator immer zur gleichen Zeit), ist es einfacher, die Systeme nach der Qualität der von ihnen erzeugten Bilder zu beurteilen. Dazu werden als Kenngrößen verwendet:

- das räumliche Auflösungsvermögen (Bodenauflösungsvermögen),

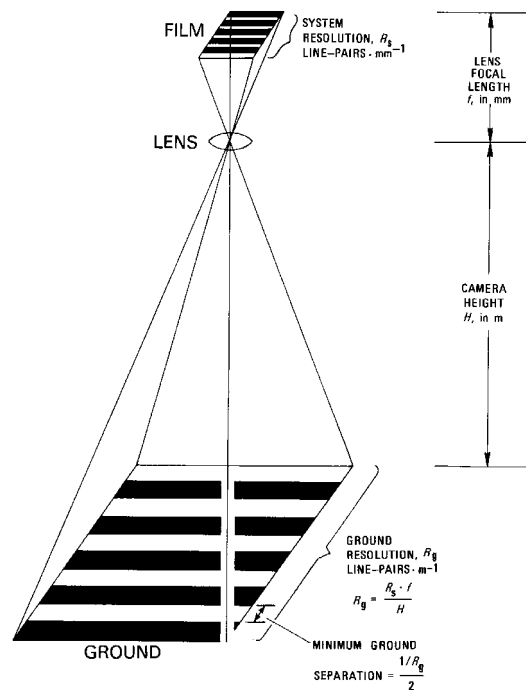


Abbildung 3.12: Zur Definition des Bodenaufklärungsvermögens [101]

- das spektrale Auflösungsvermögen,
- das radiometrische Auflösungsvermögen,
- das thermale Auflösungsvermögen und
- das zeitliche Auflösungsvermögen.

Bodenaufklärungsvermögen

Das Bodenaufklärungsvermögen bezeichnet die Fähigkeit eines Satelliteninstruments, einen auf der Erdoberfläche befindlichen Linientest aufzulösen (vgl. Abb. 3.12). Das Bodenaufklärungsvermögen R_b ist nach (3.3) gegeben zu

$$R_b = \frac{g}{f \cdot R_k} = \frac{h}{f \cdot R_k}, \quad (3.5)$$

worin R_k das Auflösungsvermögen des Systems Kamera-Film bezeichnet. Dieses setzt sich zusammen aus dem Auflösungsvermögen des Objektivs und dem der photographischen Emulsion. Letzteres wird in der Regel in Linienpaare/mm angegeben. Nach (3.5) wird das Bodenaufklärungsvermögen auch durch die Flughöhe des Satelliten bestimmt. Ein Instrument von einem frühen Landsat (900 km Flughöhe) würde daher auf einen militärischen Späher (Flughöhe von 180 km) ein um einen Faktor 5 besseres Bodenaufklärungsvermögen haben. Das Bodenaufklärungsvermögen wird auch als IFOV (Instantaneous Field of View) bezeichnet. Dies entspricht bei Scannern der Kantenlänge der Fläche, die auf genau ein Pixel (Bildelement) des Bildes abgebildet wird.

Das Bodenaufklärungsvermögen ist in Abb. 3.13 für einen in einer Höhe von 150 km fliegenden Satelliten in Abhängigkeit vom Auflösungsvermögen des Systems Kamera - Film und von der Brennweite des Objektivs angegeben. Die Abschätzung für das Bodenaufklärungsvermögen ist insofern optimistisch, als dass (a) die Flughöhe des Satelliten militärisch zu niedrig ist und (b) bei den sehr langen Brennweiten die Beugung das Auflösungsvermögen begrenzt.

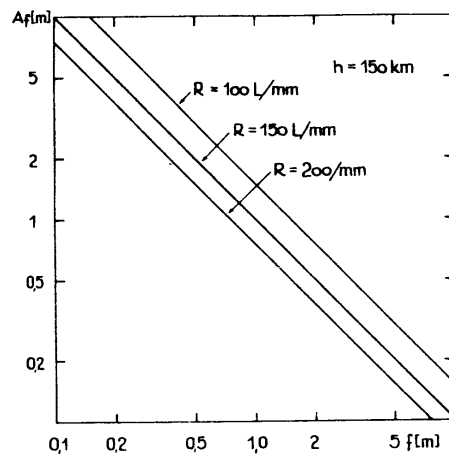


Abbildung 3.13: Bodenauflösungsvermögen in Abhängigkeit von der Brennweite des Objektivs und dem Linienauflösungsvermögen des Systems Kamera - Film (Bormann, 1980)

Diese Beugung bestimmt auch das *Auflösungsvermögen des Objektivs*, welches in das R_k in (3.5) eingeht. Das Winkelaufklärungsvermögen θ des Objektivs ist:

$$\theta = \frac{a}{f} = \frac{\lambda}{D} \quad (3.6)$$

mit a als der in der Brennweite f aufgelösten Distanz, λ der Wellenlänge des Lichts und D dem Öffnungsdurchmesser des Objektivs. Dieser Abstand a auf der Filmebene kann nicht beliebig klein gemacht werden, da durch Beugung aus Punkten Beugungsscheibchen werden und der Mindestabstand zur Auflösung zweier benachbarter Punkte größer sein muss als der Durchmesser dieser Beugungsscheibchen (vgl. Abb. 3.14). λ ist vorgegeben, da wir im sichtbaren Bereich arbeiten. Dann bestimmt die Blendenöffnung D das Winkelaufklärungsvermögen. Im Falle von Satelliten kann D nicht beliebig groß gemacht werden, da einerseits die Baugröße durch den Satelliten eingeschränkt wird und andererseits die einfallende Lichtmenge reduziert werden muss. Bei einem mittleren λ von $5 \cdot 10^{-5} \text{ cm}$ ergibt sich für eine Blendenöffnung von 1 cm (5 cm, 20 cm, 1 m) und eine Brennweite von 100 mm nach (3.6) in der Filmebene ein Durchmesser der Beugungsscheibchen von $5 \cdot 10^{-4} \text{ cm}$ ($1 \cdot 10^{-4}$; $2,5 \cdot 10^{-5}$; $5 \cdot 10^{-6} \text{ cm}$). Die letzten Werte liegen ungefähr in der Größe der Körnchen in einer photographischen Emulsion. Ist die Blendenöffnung sehr eng oder die Brennweite sehr groß, so kann der Durchmesser des Beugungsscheibchens den Durchmesser der Körnchen bei weitem übertreffen, d.h. das Objektiv und nicht die Körnigkeit des Films begrenzen das Auflösungsvermögen. Bei geringen Brennweiten oder großen Blendenöffnungen dagegen kann der Durchmesser des Beugungsscheibchens in der Größenordnung der Körnchengröße des Films oder noch dar-

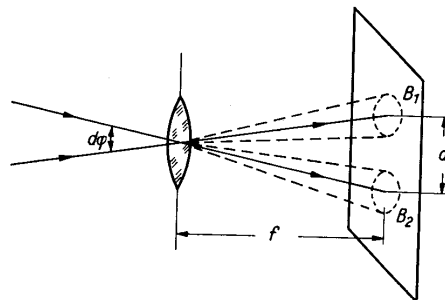


Abbildung 3.14: Winkelaufklärungsvermögen [7]

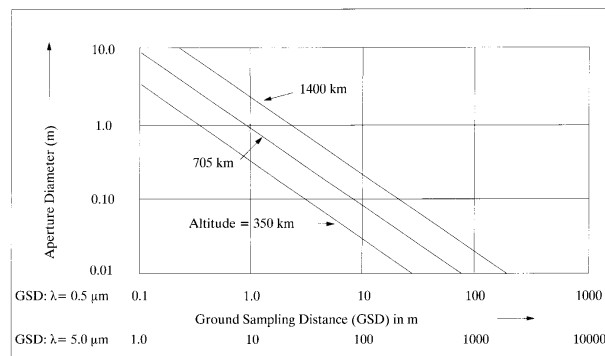


Abbildung 3.15: Einfluss des Winkelauflösungsvermögens [62]

unter liegen, so dass in diesem Falle die Eigenschaften der photographischen Emulsion den größeren Einfluss auf das Bodenaufklärungsvermögen haben. Abbildung 3.15 gibt das durch das Winkelaufklärungsvermögen entstehende Bodenaufklärungsvermögen in Abhängigkeit von der Apertur und der Flughöhe des Satelliten.

Objektive zur photographischen Aufklärung sollten also einerseits eine große Brennweite haben, um ein möglichst gutes Bodenaufklärungsvermögen zu erreichen. Andererseits ist die Brennweite aber auch begrenzt durch die Beugung und die große Länge und Masse von aus Linsen zusammengesetzten langbrennweitigen Objektiven. Lange Brennweiten werden daher durch die Verwendung von Spiegelobjektiven realisiert, wobei die Baulänge bis auf $1/5$ der Brennweite verringert werden kann [128]. Spiegelobjektive haben zusätzlich den Vorteil, dass sie weniger lichtstark sind. Daher muss ihr Öffnungsdurchmesser gegenüber dem eines Linsenobjektivs vergrößert werden, was die Einflüsse der Beugung verringert.

Auch die *photographische Emulsion* trägt zum Auflösungsvermögen des Systems Kamera-Film bei. Verschiedene Faktoren definieren deren Auflösungsvermögen: zum einen das räumliche Auflösungsvermögen der Emulsion, d.h. die Größe der Silberkörnchen in einem Schwarzweißfilm, zum anderen die Geschwindigkeit und der Kontrast des Filmes.

Unter der Geschwindigkeit eines Filmes versteht man die Zeit, die der Film einer vorgegebenen Strahlung ausgesetzt sein muß, um eine ausreichende Bildung von latenten Bildern zu erreichen. Abbildung 3.16 zeigt die charakteristische Kurve, die den Zusammenhang zwischen der Belichtung und der Schwärzung des Filmes angibt: ist die Belichtung zu gering, so gibt es kein Signal. In einem linearen Bereich ist die Schwärzung der Belichtung proportional, für stärkere Belichtung dagegen kommt es zu einer Sättigung. Die charakteristischen Kurven hängen von der Geschwindigkeit des Filmes ab, sowohl die Lage der Kurven, d.h. der Einsatzpunkt eines Signals und der Beginn der Sättigung, als auch der Bereich, in dem die Kurve linear ansteigt und die Steigung in diesem Bereich. Ist die Belichtung so stark, dass die Sättigung erreicht ist, so verringert sich das Auflösungsvermögen der Emulsion, da die einzelnen Körnchen auswaschen und zusammenwachsen. Außerdem werden Objekte unterschiedlicher Helligkeit, die aber alle hell genug sind, um eine Schwärzung innerhalb der Sättigung zu bewirken, nicht mehr voneinander getrennt.

Der *Kontrast* beschreibt die Reaktion des Filmes auf unterschiedliche Belichtungen. Er kann ebenfalls mit Hilfe der charakteristischen Kurve verstanden werden. Wenn die Belichtung außerhalb des linearen Bereiches liegt, so bewirken signifikante Unterschiede in der Belichtung kaum Unterschiede in der Schwärzung, d.h. für den Kontrast ist es wichtig, dass ein Film geeigneter Geschwindigkeit gewählt wird. Je nach Steigung der charakteristischen Kurve im linearen Bereich ist der Film mehr oder weniger gegen leichte Unterschiede in der Helligkeit der betrachteten Objekte empfindlich. Andererseits reagiert ein empfindlicher Film sehr empfindlich gegen zu kurze oder lange Belichtungszeiten, d.h. der Spielraum bei der Belichtung wird klein.

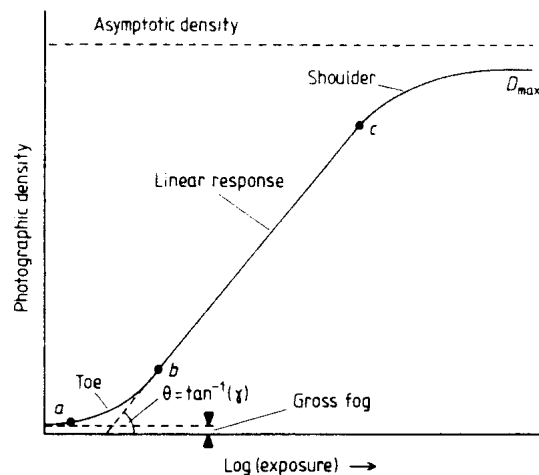


Abbildung 3.16: Charakteristische Kurve eines Films [59]

Das räumliche Auflösungsvermögen der Emulsion wird durch die bereits oben erwähnte Korngröße eines Films gegeben, diese liegt zwischen 1 und 10 μm bzw. 40 Linien/mm bis 100 Linien/mm [93], bei feineren Filmen bis 200 Linien/mm [12]. Dieses räumliche Auflösungsvermögen ist mit der Geschwindigkeit des Films gekoppelt: Sind die Körner groß, so ist die Auflösungsvermögen klein und es werden nur wenige Photonen benötigt, um eine Schwärzung zu bewirken, der Film ist also schnell. Je feiner die Körner sind, umso langsamer wird der Film, wobei aber das räumliche Auflösungsvermögen besser wird. Zum Vergleich: für ein Standardobjektiv (Brennweite 105 mm, Öffnung 5 cm) beträgt das durch die Beugung bedingte maximale Auflösungsvermögen 300 Linien/mm, d.h. bei einem derartigen Objektiv begrenzt der Film die Auflösung [93].

Beispiele für das Bodenauflösungsvermögen von photographischen Satellitensysteme gab es lange Zeit nur aus dem zivilen Bereich, für den militärischen Bereich gibt es eine Vielzahl von Gerüchten und Vermutungen – einiges hat sich heute durch die Deklassifizierung militärischer Aufnahmen überprüfen lassen (vgl. Abschn. 3.2.1). Sehr konventionelle, von Hand bediente Kameras wurden z.B. auf SkyLab verwendet, einem zeitweise bemannten Observatorium in einer Höhe von 435 km. Verwendet wurde eine Multispektralkamera aus 6 Einzelkameras mit unterschiedlichen Filmen und Filtern für Wellenlängen zwischen 0.4 und 0.9 μm . Mit einer Objektivbrennweite von 152 mm und einer Kantenlänge der Filme von 70 mm ergab sich ein Blickfeld von 163 x 163 km und eine Bodenauflösung zwischen 60 und 145 m [101]. Eine weitere Kamera (Earth Terrain Camera) mit einer Brennweite von 457 mm und einem Blickfeld von 109 x 109 km wurde mit verschiedenen Filmen (Farbe, Schwarzweiß, IR - Schwarzweiß) betrieben und erreichte je nach Film ein Bodenauflösungsvermögen von 15 m bis 30 m. Mit den Scannern auf den neueren Erdbeobachtungssatelliten (z.B. SPOT und LandSat) lassen sich aus einer Höhe von 900 km Bodenauflösungsvermögen zwischen 10 m und 80 m erreichen. Kommerzielle Systeme wie Ikonos oder SPOT 5 liefern Bodenauflösungsvermögen von ca. 1 m, wobei das hohe Bodenauflösungsvermögen im panchromatischen Modus erreicht wird, die Farbbilder meist ein um einen Faktor 3 bis 4 geringeres Auflösungsvermögen haben. Die auf CDs vertriebenen Ikonos-Bilder sind häufig Kombinationen aus panchromatischen und mehrfarbigen Bildern: erstere werden verwendet, um die letzteren in ihrer Auflösung zu verbessern.

Die neuesten Entwicklungen (in der Regel Scanner und nicht unbedingt photographische Emulsionen) erlauben deutlich bessere Auflösungsvermögen, vgl. Abb. 3.17. Im militärischen Bereich scheint man damit unter die 10 cm Grenze zu kommen; das Lesen der Schlagzeile der Bildzeitung aus dem Orbit wird damit vielleicht bald möglich werden – ob es sinnvoll ist, möchte ich lieber nicht kommentieren.

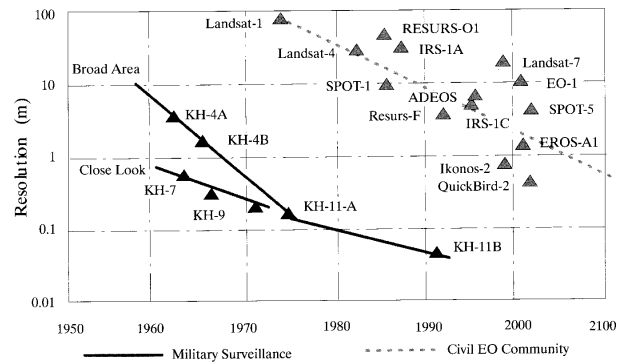


Abbildung 3.17: Mit verschiedenen Systemen erreichbares Bodenaufklärungsvermögen [62]

Im zivilen Bereich werden häufig Aufnahmen von SPOT 5 oder Ikonos verwendet; letzterer wurde z.B. in den Medien nach dem 11. September 2001 verwendet (http://www.space.com/news/ikonos_wtc_010912.html). Beispiele zu diesen und anderen Satelliten sind im Web unter den folgenden Adressen zu finden: <http://www.spaceimaging.com/>, <http://www.euspaceimaging.com/> und <http://www.satimagingcorp.com/index.html>. Einige große hochauflösende Bilder von Ikonos sind für private Zwecke beim DLR unter <http://www.caf.dlr.de/caf/satellitendaten/missionen/ikonos/> zu finden.

Spektrales Auflösungsvermögen

Das emphysemale Auflösungsvermögen bezeichnet die Fähigkeit des Systems, Wellenlängenbereiche zu trennen. Bei einer Multispektralkamera z.B. werden drei bis sechs relativ breite Wellenlängenbereiche getrennt. Die spektrale Auflösung ist für die ersten, im wesentlichen auf Kartographie und militärische Aufklärung ausgelegten Systeme, von untergeordneter, aber, wie die frühe Entwicklung der Multispektralkameras zeigt, nicht zu vernachlässigender Bedeutung gewesen. Heutige moderne Systeme versuchen auch die spektrale Auflösung zu verfeinern. Beim High-Resolution Imaging Spectrometer HIRIS z.B. wird eine spektrale Auflösung des Wellenlängenbereiches 0.4 bis $2.5\mu\text{m}$ (sichtbares Licht bis mittleres Infrarot) in 200 Kanäle angestrebt [3].

Die Motivation für ein sehr gutes spektrales Auflösungsvermögen ist im Prinzip die gleiche, wie die zur Verwendung des nahen Infrarot: verschiedene Materialien haben eine wellenlängenabhängige Reflektivität (vgl. Abb. 3.10 und Tabelle 3.2). Objekte, die in einer Farb- oder Schwarzweißaufnahme sehr ähnlich aussehen, können bei stark verfeinerter Spektralanalyse aber deutlich unterschieden werden. Dadurch lassen sich unter anderem verschiedene Pflanzenspezies besser unterscheiden und es lassen sich sogar deren physiologische Parameter wie Alter und Feuchtigkeitsgehalt bestimmen (vgl. einige der Artikel in [24]).

Thermales Auflösungsvermögen

Das thermale Auflösungsvermögen ist für die Infrarotsensoren von Bedeutung, die im Bereich des thermalen Infrarot arbeiten, d.h. die von einem Körper emittierte Wärmestrahlung betrachten. Hierbei handelt es sich eigentlich nur um eine andere Formulierung für ein spektrales Auflösungsvermögen, siehe Abschn. 3.3.1.

Radiometrisches Auflösungsvermögen

Das radiometrische Auflösungsvermögen bezeichnet die Fähigkeit, Helligkeitswerte zu unterscheiden. Bei schlechter radiometrischer Auflösung verringert sich auch das Bodenaufklärungsvermögen, insbesondere wenn die beobachteten Objekte nicht sehr kontrastreich sind, vgl.

Abb. 6.4. Kontrastanreicherung bei der Weiterverarbeitung kann zwar einen Teil dieser Ungenauigkeiten wieder ausgleichen, ist aber zur vollständigen Kompensation nicht ausreichend, vgl. Abschn. 6.2.1. Man kann sich die Bedeutung des Kontrastes für das Auflösungsvermögen einfach veranschaulichen, Sabins [101] gibt Testtafeln, die den Einfluss des Kontrastes auf das Auflösungsvermögen des Auges demonstrieren. Das radiometrische Auflösungsvermögen (bei Detektoren statt Filmen) wird in Bits angegeben, gängige Werte sind 7 Bit (128 Grauwerte, LandSat MSS) oder 8 Bit (256 Grauwerte, LandSat TM oder SPOT-HRV).

Zeitliches Auflösungsvermögen

Das zeitliche Auflösungsvermögen bezeichnet die Wiederholfrequenz, d.h. die Zeit die zwischen zwei aufeinanderfolgenden Aufnahmen eines bestimmten Gebietes vergeht. Auf Grund der großen Eigengeschwindigkeit von Satelliten ist das zeitliche Auflösungsvermögen, außer bei geostationären Satelliten, nicht durch die Kamera-Verschlusszeiten oder vorgegebene Wiederholzeiten gegeben, sondern durch die durch den Satellitenorbit bestimmte Zeit zwischen zwei aufeinanderfolgenden Flügen über ein bestimmtes Gebiet.

Verschlechterung des Auflösungsvermögens

Bisher haben wir das (Boden-)Auflösungsvermögen einer Satellitenkamera unter optimalen Bedingungen betrachtet. Eine Verschlechterung des Auflösungsvermögens entsteht einerseits durch die Eigenschaften des Instruments, andererseits durch Streuung und Absorption in der Atmosphäre.

Die instrumentell bedingten Verschlechterungen entstehen durch den verwendeten Film (Filmempfindlichkeit, Kontrastempfindlichkeit), das Objektiv (Kontrastverschlechterung, Öffnungswinkel, Abbildungsfehler können das Auflösungsvermögen zu den Rändern hin um bis zu einen Faktor 2 verringern [61]) und die Bewegung des Satelliten. Letztere führt zu einer Bildverschiebung d in der Filmebene, gegeben durch

$$d[\mu\text{m}] = 1000 \cdot \delta \cdot v \left[\frac{\text{km}}{\text{s}} \right] \cdot \frac{f}{g} \left[\frac{\text{mm}}{\text{km}} \right] \cdot t [\text{s}] \quad (3.7)$$

mit t als der Belichtungszeit, δ als dem Faktor der Vorwärtskompensation und v als der Subsattelitengeschwindigkeit [12, 128]. Die Vorwärtsbewegung führt zu einer Verschmierung der Bilddetails in Bewegungsrichtung.

Die zweite Einflussgröße lässt sich quantitativ weniger genau beschreiben, da die Ausprägung dieser Einflüsse mit der atmosphärischen Variabilität variiert. Die Streuung (Dunst, zur genauen Beschreibungen vgl. z.B. [66]) in der Atmosphäre bewirkt eine Kontrastverschlechterung. Als Gegenmaßnahme werden Anti-Dunstfilter verwendet, die das blaue atmosphärische Streulicht unterdrücken. Bei Multispektralkameras werden die Wellenlängenbereiche bevorzugt, in denen die Streuung geringer ist (vgl. auch Abb. 3.10) bzw. verwenden diese zu einer nachträglichen Dunstkorrektur, vgl. Abschn. 6.1.5. Zusätzlich bewirken Dichteschwankungen in der Atmosphäre zufällige Winkelverschiebungen der betrachteten Objekte (Szintillation oder Pulsation). Die damit verbundene Kontrastverschlechterung ist um so größer, je näher die Kamera den Turbulenzonen ist und lässt sich auf Grund der Relativbewegung zwischen Aufnahmesystem und Objekt nicht wie bei einem astronomischen Teleskop durch Despeckeln korrigieren.

3.2.7 Elektronische Aufzeichnung statt photographischer Emulsion

Die bisher diskutierten photographischen Systeme sind gleichsam der Urahn der Satelliteninstrumente zur Erderkundung. Die Technik einer Kamera ist relativ einfach und gut verstanden, allerdings hat die Verwendung eines Films als Aufzeichnungsmedium verschiedene Nachteile: (a) die 'Lebensdauer' des Systems ist durch den mitgeführten Filmvorrat begrenzt und (b) der belichtete Film muss entweder zur Erde zurückgeführt werden oder an Bord des Satelliten entwickelt, elektronisch abgetastet und dann als Digitalsignal zum Boden gesandt

werden. Photographische Kameras werden daher heutzutage im wesentlichen nur noch auf bemannten Missionen (Space-Shuttle) verwendet.

Das Problem der Filmrückführung wurde bei militärischen Aufklärungssatelliten aufwendig gelöst, vgl. Abb. 5.1. Diese fliegen meist in niedrigen Orbits, so dass ihre Lebensdauer kurz ist. Die FSU hat nach jeweils 1 bis 2 Wochen Flugzeit ihre Satelliten mittels eines ‘geordneten Absturzes’ über eigenem Gebiet zu Boden gebracht und so die Filmkapseln geborgen. Um eine kontinuierliche Überdeckung zu erreichen, benötigt man entsprechend viele Satelliten. Amerikanische Satelliten dagegen packen die belichteten Filme in Kapseln, die abgeworfen werden. Die Kapsel sendet während ihres Falls Funksignale aus, so dass ihre Flugbahn von Schiffen und Flugzeugen aus verfolgt werden kann. In einer Höhe von 16 km öffnet die Kapsel einen Fallschirm und wird dann entweder von einem Flugzeug aufgefangen oder beim Absturz ins Meer von einem Schiff geborgen (vgl. z.B. Wolf et al., 1983).

Videokameras

Dem konventionellen photographischen Verfahren am ähnlichsten ist das RBV (Return Beam Vidicon). Vorteil dieser Systeme gegenüber konventionellen photographischen Systemen ist ihre hohe Quantenausbeute, d.h. sie arbeiten auch unter ungünstigen Beleuchtungsverhältnissen. Ihr Hauptnachteil sind ein schlechtes räumliches und radiometrisches Auflösungsvermögen.

Der wichtigste Bestandteil einer Videokamera ist ein Vidicon. Dabei handelt es sich um einen photosensitiven Schirm, dessen elektrische Leitfähigkeit je nach einfallender Lichtintensität variiert. Das Blickfeld wird wie bei einer konventionellen Kamera über ein Objektiv auf diesen Schirm projiziert. Dieser hält das Bild für eine gewisse Zeit fest, während derer der Schirm elektronisch abgetastet wird (zum Prinzip vgl. z.B. [59]).

Eine verbesserte Version des Vidicons (SEC-Vidicon) durch Sekundärelektronenvervielfachung erreicht eine erhöhte Quantenausbeute und ein etwas besseres radiometrisches Auflösungsvermögen. Dabei wird das zu scannende Bild nicht direkt von den Photonen produziert, sondern die Photonen fallen auf einen Schirm, aus dem sie Elektronen auslösen. Diese Elektronen werden in der Röhre auf einen zweiten Schirm hin beschleunigt, aus dem sie Sekundärelektronen herausschlagen. Diese werden von einer positiven Elektrode abgezogen und erzeugen dort ein Potentialmuster, das das ursprünglich einfallende Muster reproduziert und als Ausgangssignal abgetastet wird. Ein Photoemitter am Eintrittsfenster kann so gewählt werden, dass das Vidicon für bestimmte Wellenlängenbereiche empfindlich ist. Damit besteht die Möglichkeit zu Multispektralaufnahmen. Letztere kann man auch durch den Vorsatz geeigneter Filter vor ein einfaches Vidicon erzeugen.

Während ein normaler Fernsehschirm, der ja nach der Umkehrung dieses Prinzips arbeitet, normalerweise in 625 Zeilen abgetastet wird, werden in den in Satelliten installierten Vidicons rund 4500 Rasterzeilen abgetastet. Vidicons waren eine Standardausstattung der ersten zivilen Erderkundungssatelliten (LandSat 1-3), sind aber heutzutage größtenteils von Scannern verdrängt. Auf LandSat 1 und 2 wurden drei parallele RBV Kameras eingesetzt für Multispektralaufnahmen in den Wellenlängenbereichen $0.475 - 0.575 \mu\text{m}$ (Blaugrün), $0.58 - 0.68 \mu\text{m}$ (Gelbrot) und $0.68 - 0.73 \mu\text{m}$ (Rot und Infrarot). Diese Kameras konnten alle 25 Sekunden ein Gebiet von $185 \times 185 \text{ km}$ aufnehmen. Mit der Zeilenzahl des Vidicons ergibt sich daraus eine Auflösung von 80 m [17, 71]. Die Bilder der einzelnen Kameras entsprechen jeweils Aufnahmen, wie sie ein Schwarzweißfilm liefern würde, farbige Aufnahmen werden daraus durch Überlagerung erzeugt.

Auf LandSat 3 wurden nur zwei RBV-Kameras eingesetzt, die parallel geschaltet waren (vgl. Abb. 3.18). Beide Kameras hatten nur einen Breitbandkanal von $0.51 - 0.75 \mu\text{m}$ (Gelb bis nahes Infrarot), d.h. auf die spektrale Auflösung wurde verzichtet. Dafür hatten sie aber ungefähr doppelt so große Brennweiten wie die LandSat 1 und 2 Instrumente. Dadurch ergibt sich ein nur halb so großes Blickfeld ($99 \times 99 \text{ km}$), so dass zwei Kameras notwendig sind, um das gleiche Gesamtblickfeld auszuleuchten. Die größere Brennweite führte zu einem räumlichen Auflösungsvermögen von 30 m [17]. Die LandSat-RBV-Bilder wurden verwendet, um die gleichzeitig aufgenommenen MSS-Bilder in ihrer räumlichen Auflösung zu verbessern.

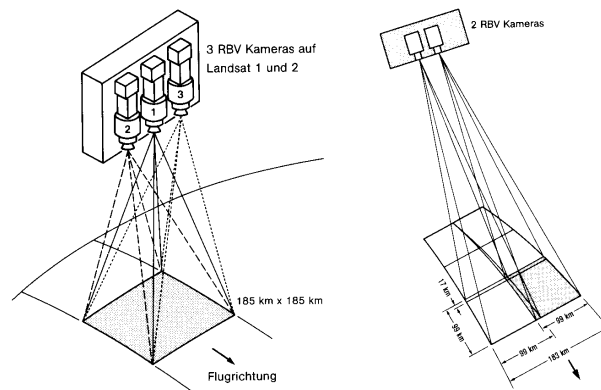


Abbildung 3.18: LandSat Videokameras [71]

Die RBV-Kameras haben sich nur begrenzt bewährt. Ihr Hauptproblem ist die geringe radiometrische Auflösung, so dass weniger als die Hälfte der LandSat RBV Aufnahmen eine gute Qualität aufwies ([17] und Zitate darin). Für die Missionen zu anderen Planeten wurden RBV-Kameras für die Beobachtungen im optischen und nahen infraroten Bereich eingesetzt, z.B. auf Viking und Mariner (Erforschung des Mars), auf Voyager (äußere Planeten und ihre Monde) und Mariner 10 (Venus).

CCDs: Charged Coupled Devices

Bei einer Videokamera erfolgt eine elektronische Aufzeichnung auf einen Schirm, d.h. abgesehen von der elektronischen Aufzeichnung ist das Verfahren dem der konventionellen Photographie sehr ähnlich. Bei den Charged Coupled Devices (CCDs) wird der Film ersetzt durch eine Reihe lichtempfindlicher Sensoren, d.h. es wird im Prinzip nur ein sehr schmaler Streifen senkrecht zur Flugrichtung des Satelliten abgetastet. Die Eigenbewegung des Satelliten bewirkt jedoch, dass dieser Streifen immer weiter verschoben wird, so dass die ganze Fläche abgetastet wird.

Die Wirkungsweise des CCD beruht auf dem photoelektrischen Effekt: auf einen Halbleiter fallendes Licht erzeugt ein Elektron-Loch Paar. Diese Elektronen werden in Potentialsenken gefangen, die aus vielen kleinen Elektroden bestehen. Dort sammeln sich die Elektronen bis ihre Anzahl durch 'charge coupling' der Sensorelektrode zu einer Ausleselektrode übergeben wird.

Der Aufbau eines einzelnen CCD ist im linken Teil von Abb. 3.19 gezeigt. Die Elektrode ist gegenüber dem Halbleiter durch eine dünne Siliziumoxidschicht getrennt (Verwandschaft mit dem Metalloxid-Silizium-Transistor MOS). Die Elektrode wird auf einer geringen positiven Spannung gehalten, die gerade gross genug ist, um die positiven Löcher in das p-dotierte Silizium zu treiben und die Elektronen in eine dünne Schicht unterhalb der Elektrode zu ziehen. Die von der einfallenden Strahlung in dieser Verarmungsschicht erzeugten Elektronen-Lochpaare werden genauso getrennt, die von der Strahlung erzeugten Elektronen sammeln sich ebenfalls unterhalb der Elektrode. In diesem Speicherbereich hat sich dann eine Elektronenladung geformt, die eine Funktion der Intensität der einfallenden Strahlung ist. Ein CCD verhält sich daher wie ein strahlungsgetriebener Kondensator.

Ordnet man viele dieser Elektroden auf einem einzigen Siliziumchip an und isoliert die Verarmungsregionen gegeneinander, so entwickelt jede von ihnen eine Ladung, die der auf sie fallenden Strahlung proportional ist. Ein 'Beleuchtungsmuster' wird also in ein räumlich digitalisiertes Ladungsmuster verwandelt. Für die technische Lösung des Problems des Auslesens dieses Ladungsmusters sei z.B. auf [59] verwiesen.

Die Verwendung von CCDs in der Erdfernerkundung hat, entsprechend der technischen Entwicklung, drei Stufen durchlaufen:

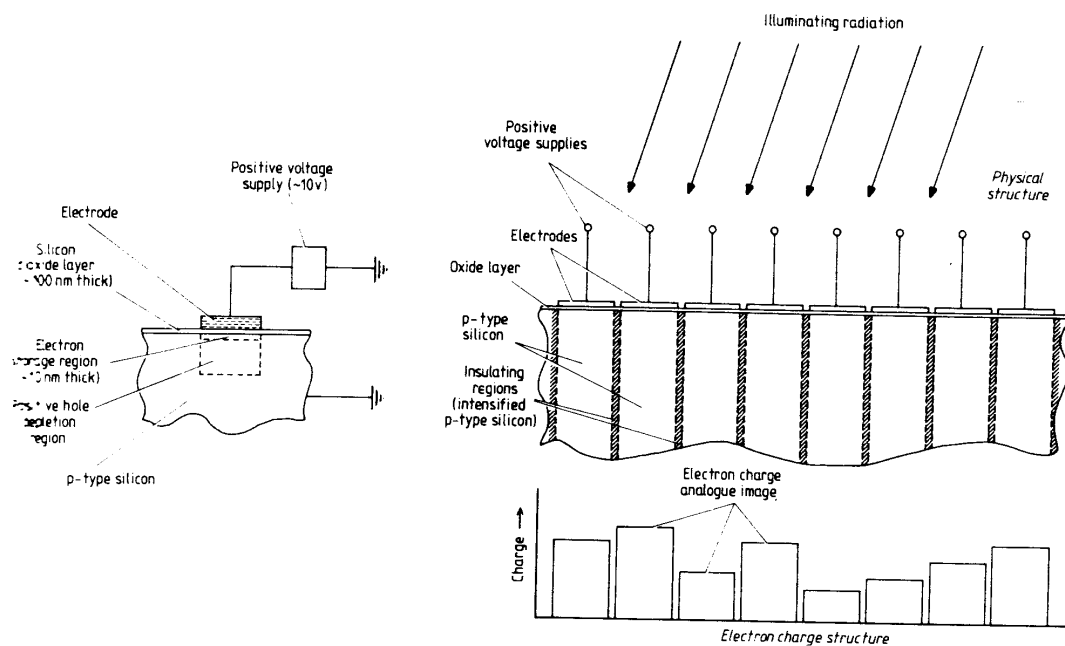


Abbildung 3.19: Aufbau eines einzelnen CCD und eine Zeile von CCDs [59]

1. die Verwendung einer Zeile von CCDs in der Kehrbesentechnik bei der die Bewegung des Satelliten zum Scannen ausgenutzt wird;
2. die Verwendung von Einzel-CCDs (bzw. kleiner Zeilen mit einer CCD pro Spektralkanal) beim Scanner, der nur jeweils ein Pixel zur Zeit aufnimmt und durch eine Kombination aus Satellitenbewegung und Bewegung eines Spiegels senkrecht zur Satellitenbahn scant;
3. und in den neuesten Satelliten (insbesondere kommerziellen wie Diamant) die Verwendung von CCD-Arrays, d.h. instrumentell kehrt die Erdfernerkundung jetzt zu ihren Ursprüngen zurück, der Aufzeichnung der Information auf einem zwei-dimensionalen Medium.

Kehrbesentechnik

Der prinzipielle Aufbau eines optoelektronischen Abtasters ist im linken Teil von Abb. 3.20 gezeigt. Eine CCD-Sensorzeile kann aus bis zu 2000 CCDs bestehen [71]. Diese Sensorzeile

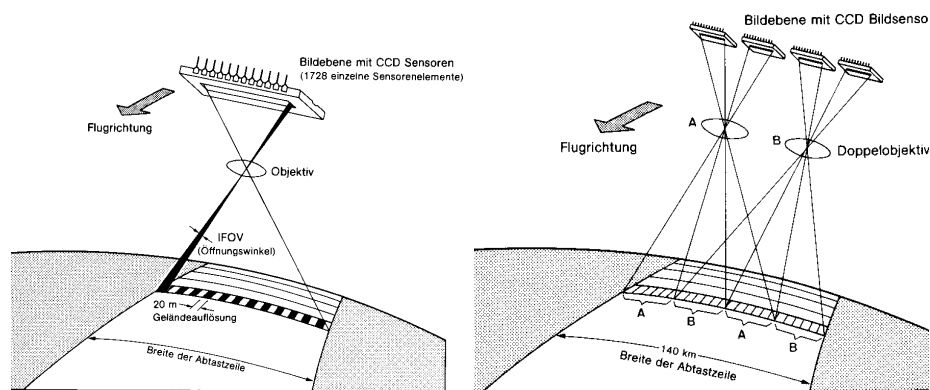


Abbildung 3.20: Optoelektronischer Abtaster (links) und Verbesserung des Auflösungsvermögens durch die MOMS-Doppeloptik [71]

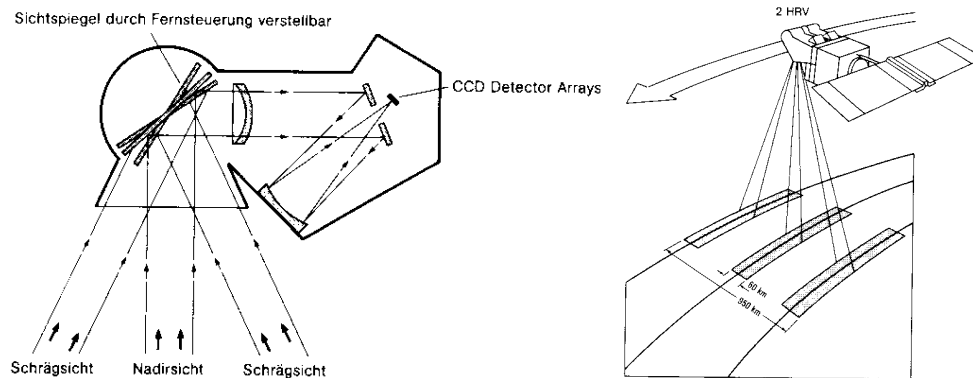


Abbildung 3.21: Prinzip der High Resolution Visible (HRV) Kamera auf SPOT und Veränderung des Aufnahmegebiets durch Verstellung des Spiegels [71]

wird wie ein Kehrbesen über das abzutastende Gelände 'geschoben' ('push broom scanning', Kehrbesenabtasten). Das gesamte System besteht also aus einem Objektiv und einer Zeile von CCDs, die den Film in einer konventionellen Kamera ersetzt. Das Verfahren entspricht dem einer Streifenkamera. Mechanische Bauteile (z.B. Objektivverschluss) sind nicht notwendig, wodurch sich eine hohe Zuverlässigkeit ergibt. Die Instrumente arbeiten einfach und energiesparend, so dass sie für Anwendungen auf Satelliten optimal sind.

Die Begrenzung des Bodenaufösungsvermögens ist durch die Zahl der einzelnen Sensorelemente in einer Bildzeile gegeben. Um eine höhere Auflösung zu erreichen, werden mehrere Bildzeilen miteinander gekoppelt. Im rechten Teil der Abbildung ist dazu der Modulare Optoelektronische MultispektralScanner (MOMS) dargestellt, bei dem durch eine Doppeloptik aus zwei genau aufeinander abgestimmten Objektiven der abzutastende Streifen auf insgesamt vier CCD-Zeilen gegeben wird. Damit ist die Auflösung um einen Faktor 4 verbessert, da bei gleicher Breite der Abtastzeile am Boden eine vierfache Anzahl von Bildpunkten erzeugt wird. Im Multispektralbetrieb (ein Modul misst im sichtbaren, eins im infraroten Spektralbereich) ergibt sich bei Aufnahmen vom Shuttle aus eine Bodenaufösung von 20 m.

CCDs werden auch auf dem zur Zeit wahrscheinlich effektivsten Universal-System zur Erdfernerkundung SPOT eingesetzt. Abbildung 3.21 zeigt das Funktionsprinzip der High Resolution Visible (HRV) Kamera auf SPOT. Das photosensitive Element besteht aus einer Reihe von ca. 6000 CCDs [71]. Bei einem Blickfeld von 60 km Breite ergibt sich eine Auflösung von 10 m (panchromatischer Modus). Da zwei Instrumente parallel geschaltet sind, ergibt sich ein Gesamtblickfeld von 117 km (3 km Überlappung). Die Optik besteht aus einem Spiegelteleskop und einem davor angebrachten Drehspiegel. Dieser Drehspiegel steht während einer Aufnahme fest (er fungiert nicht als scannendes Element!), kann jedoch so bewegt werden, dass die Kamera etwas zur Seite blickt (rechter Teil in Abb. 3.21). Dadurch können innerhalb eines 950 km breiten Streifens beliebige 60 km breite Streifen herausgegriffen werden. Dadurch kann dann ein Gebiet beobachtet werden, ohne dass es direkt überflogen wird. Angewendet wird dies z.B. bei schnell veränderlichen Phänomenen (z.B. die Beobachtung der Ölbrände während des ersten Golfkrieges) oder um Aufnahmen, die bei wolkenbedecktem Himmel erfolgten, bereits kurze Zeit später wiederholen zu können und nicht erst nach 26 Tagen (Wiederholfrequenz durch die Bahn). Durch das Schwenken des Spiegels können auch echte stereoskopische Aufnahmen eines Gebiets erstellt werden.

Das HRV-Instrument auf SPOT kann in zwei Modi betrieben werden, als Multispektralkamera (drei Spektralbereiche, vgl. Tabelle 3.3) oder im panchromatischen Mode. In letzterem erreicht das Instrument ein Bodenaufösungsvermögen von 10 m. Die Multispektralanalyse erfolgt durch eine aus verschiedenen Filtern bestehende Scheibe vor dem CCD, bei der immer das Filter des gewünschten Frequenzbereiches in den Strahlengang gedreht wird.

Trotz der guten räumlichen Auflösung von SPOT sollte man berücksichtigen, dass der Thematic Mapper von Landsat mit seiner besseren spektralen Auflösung und dem größeren

	Multispectral mode	Panchromatic mode
Spectral bands:		
green	0.50 to 0.59 μm	0.51 to 0.73 μm
red	0.61 to 0.68 μm	—
reflected IR	0.79 to 0.89 μm	—
Angular field of view	4.13°	4.13°
Ground resolution cell (nadir viewing)	20 by 20 m	10 by 10 m
Detectors per spectral band	3000	6000
Ground-swath width (nadir viewing)	60 km	60 km

Tabelle 3.3: Charakteristika des SPOT High Resolution Visible HRV Instruments [101]

überdeckten Spektralbereich für viele Anwendungen Vorteile hat (z.B. in geologischen Fragestellungen). Die Suche nach Mineralien- oder Ölvorkommen erfordert keine so gute räumliche Auflösung, dafür ist dort die spektrale Auflösung, insbesondere der Vergleich von Band 5 und 7 des TM von entscheidender Bedeutung. Ein Mixen der TM und SPOT Aufnahmen zur Verbesserung des Auflösungsvermögens, so wie früher mit den MSS und RBV Aufnahmen von LandSat und heute mit den panchromatischen und mehrfarbigen Aufnahmen von Ikonos gemacht, ist nicht möglich, da die Aufnahmen weder zeitlich noch räumlich synchronisiert werden können.

Einen nahezu identischen Aufbau hat die Multispektralkamera auf Giotto, die 1986 die ersten Nahaufnahmen eines Kometen (Halley) zur Erde zurücksandte [100] und damit das Bild eines Kometen als einer riesigen, durch das All fliegenden Kartoffel schuf.

3.2.8 Scanner: Spiegel statt Objektiv

Bisher haben wir uns von der konventionellen photographischen Aufzeichnung nur durch Verwendung eines anderen aufzeichnenden Mediums entfernt, jedoch die Abbildung mit Hilfe eines Linsenobjektivs beibehalten. In diesem Abschnitt werden wir nun Systeme betrachten, in denen das Objektiv durch einen oszillierenden Spiegel ersetzt wurde. Bekannteste Beispiele sind der Multispektralscanner (MSS, für Details und Datenprodukte siehe <http://edc.usgs.gov/products/satellite/mss.html>), der auf allen LandSat-Flügen bis einschließlich LandSat 5 verwendet wird, und der Thematic Mapper (TM, für Details und Datenprodukte siehe <http://edc.usgs.gov/products/satellite/tm.html>), der nur auf den neueren LandSat-Satelliten (LandSat 4 und 5) als Ersatz für die RBV-Kamera verwendet wird. Der TM unterscheidet sich vom MSS durch ein besseres räumliches und radiometrisches Auflösungsvermögen, arbeitet aber nach dem gleichen Prinzip. Er wurde für den aktuellen LandSat 7 zum Enhanced Thematic Mapper+ (ETM+, siehe <http://edc.usgs.gov/products/satellite/landsat7.html>) weiter entwickelt.

Abbildung 3.22 zeigt im linken Teil die Arbeitsweise eines Multispektralabtasters. Das Funktionsprinzip dieses Scanners beruht auf einem oszillierenden Spiegel, der Geländestreifen senkrecht zur Flugrichtung abtastet. Dieser Spiegel projiziert die einfallende Strahlung über eine Spiegeloptik auf ein Spektrometer, das in den vier Wellenlängenbereichen Grün (0.5 – 0.6 μm), Rot (0.6 – 0.7 μm) und nahes Infrarot (0.7 – 0.8 μm und 0.8 – 1.1 μm) die Intensität der Strahlung registriert. Diese Wellenlängenbereiche sind nicht scharf getrennt, sondern haben Ansprechfunktion, die einen Schwerpunkt in der Intervallmitte haben und zu den Rändern auch in die Messbereiche benachbarter Bänder reichen (vgl. Abb. 3.23 links).

Die bei einer Spiegeldrehung vom MSS abgetasteten Bereiche sind 476 m breit. Dies ergibt sich aus der fürs Abtasten benötigten Zeit und der Geschwindigkeit des Satelliten über Grund. Damit die räumliche Auflösung nicht auf diese 476 m beschränkt ist, werden 6 Detektoren pro Wellenlängenband betrieben, so dass jeder einzelne abgetastete Streifen 79 m breit ist. Die von den Detektoren gemessenen Intensitätswerte werden in 64 Kanäle eingeteilt, d.h. die

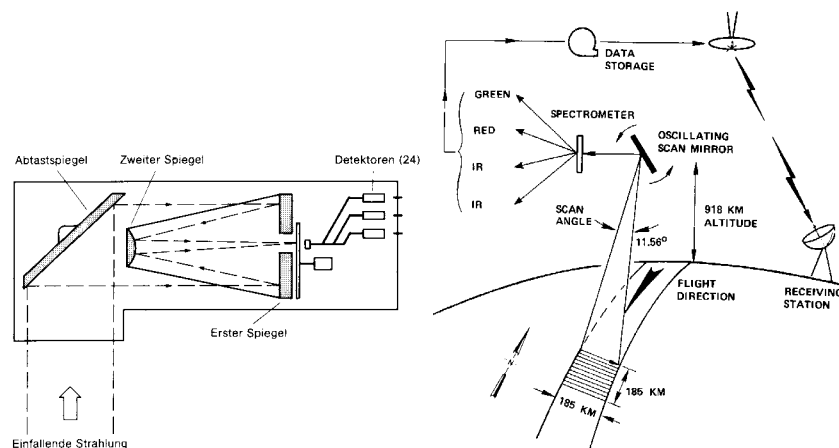


Abbildung 3.22: Prinzip des Multispektralscanners MSS auf LandSat (links, [71]) und Blickfeld und Arbeitsweise für die Scanner auf LandSat 1 – 3 (rechts [101]). LandSat 4 und 5 fliegen in einer Höhe von 705 km und haben Scanwinkel von 14.9°

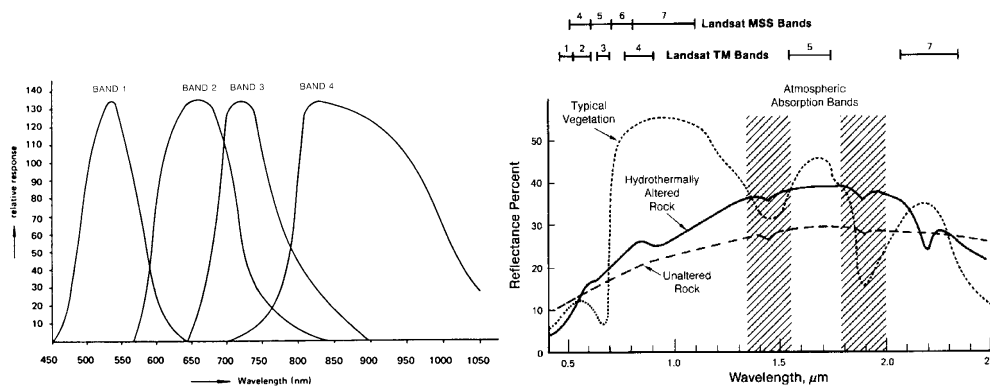


Abbildung 3.23: Links: Landsat TM Wellenlängenbereiche [16]; rechts: Spektralbereiche für die beiden Landsatinstrumente TM und MSS, zum Vergleich ist das spektrale Reflektionsvermögen für Vegetation und Gestein gezeigt [101]

photometrische Auflösung ist auf 6 Bit begrenzt.

Diese 476 m breiten Streifen werden kontinuierlich abgescannt, d.h. man könnte im Prinzip ein einziges großes Bild der Erdoberfläche entlang der Satellitenspur erstellen. Um später mit den Aufnahmen der RBV-Kameras vergleichen zu können (Kombination zur Verbesserung des Bodenaufhebungsvermögens), wird der Datenstrom nach jeweils 185 km abgebrochen. Das entspricht den 25 Sekunden, die die RBV-Kamera zum Abtasten eines Bildes benötigt. Obwohl die Auflösung des MSS geringer ist als die der RBV-Kamera, hat sich das System gegenüber der RBV Kamera bewährt, insbesondere dadurch, dass es unempfindlicher gegen Belichtungsprobleme ist und eine bessere spektrale Auflösung bietet.

Auf den LandSat-Satelliten der zweiten Generation ist die RBV-Kamera durch den Thematic Mapper TM ersetzt. Der TM basiert auf dem gleichen Prinzip wie der MSS, ist diesem gegenüber aber in den folgenden Punkten verbessert:

- das räumliche Auflösungsvermögen beträgt 30 m (79 m beim MSS): die gesamte Zeilenbreite ist durch die Satellitenbewegung zu 480 m gegeben, jedoch erfolgt, außer im thermischen Infrarot, eine Aufteilung auf 16 Detektoren (6 im MSS).
- der TM hat sieben Spektralbänder (4 beim MSS, für die Spektralbereiche vergleiche Abbildung 11.17). Der Spektralbereich des TM erstreckt sich dabei über einen breiteren Bereich vom Blauen bis ins thermische Infrarot (Kanal 6 von $10.4 - 12.5 \mu\text{m}$ ist in Abb. 3.23 nicht

Band	Wavelength, μm	Characteristics
1	0.45 to 0.52	Blue-green—no MSS equivalent. Maximum penetration of water, which is useful for bathymetric mapping in shallow water. Useful for distinguishing soil from vegetation and deciduous from coniferous plants.
2	0.52 to 0.60	Green—coincident with MSS band 4. Matches green reflectance peak of vegetation, which is useful for assessing plant vigor.
3	0.63 to 0.69	Red—coincident with MSS band 5. Matches a chlorophyll absorption band that is important for discriminating vegetation types.
4	0.76 to 0.90	Reflected IR—coincident with portions of MSS bands 6 and 7. Useful for determining biomass content and for mapping shorelines.
5	1.55 to 1.75	Reflected IR. Indicates moisture content of soil and vegetation. Penetrates thin clouds. Good contrast between vegetation types.
6	10.40 to 12.50	Thermal IR. Nighttime images are useful for thermal mapping and for estimating soil moisture.
7	2.08 to 2.35	Reflected IR. Coincides with an absorption band caused by hydroxyl ions in minerals. Ratios of bands 5 and 7 are potentially useful for mapping hydrothermally altered rocks associated with mineral deposits.

Tabelle 3.4: Spektralbänder des Thematic Mapper und ihre Beziehung zur Interpretation der Bilder bzw. zu konkreten Fragestellungen an einzelne Kanäle [101]

6 gezeigt).

- das radiometrische Auflösungsvermögen ist besser (auf 256 Kanäle erweitert).

Die Bezeichnung Thematic Mapper wird verständlich, wenn man sich die Auswahl der Wellenlängenbereiche einmal genauer ansieht. Diese sind so gewählt, dass nicht nur die Überlagerung dieser Bilder wieder ein Gesamtbild erzeugt, sondern dass jeder Spektralbereich für sich schon Auskünfte geben kann, z.B. über Eigenschaften der Vegetation, die Bodenfeuchtigkeit oder über Öl- oder Mineralienvorkommen (vgl. Tab. 3.4). Sabins [101] gibt eine Fülle von Beispielen, in denen Einzelaufnahmen aller Kanäle gezeigt und diskutiert werden, teilweise auch mit vergleichenden Karten und Vergleichsaufnahmen im sichtbaren Bereich.

Als Beispiel zeigt Abb. 3.24 zeigt Aufnahmen der einzelnen Kanäle des TM für eine Szene des St. Francis River im Nordosten von Arkansas zusammen mit der Karte des betreffenden Gebiets. Hinzuweisen ist auf die dünnen und hohen Wolken im oberen Teil der Aufnahme, die insbesondere im Kanal 1 (blau) zu erkennen sind, mit zunehmender Wellenlänge aber immer weniger deutlich sind und im nahen IR (Band 4) verschwinden. Der eigentliche Flusslauf zeichnet sich gegenüber der Marsch in den sichtbaren Kanälen ab, vermischt jedoch mit dem Feuchtgebiet im nahen Infrarot (ein Hinweis auf die hohe Bioproduktivität im Wasser) und im thermischen IR (Flusswasser und Marsch haben gleiche Temperatur). Außerdem wird das schlechte Bodenaufklärungsvermögen der IR-Aufnahme deutlich sowie die Sichtbarkeit der hohen Wolken im Band 6: diese haben eine geringe Temperatur und zeichnen sich daher deutlich gegen die Erdoberfläche ab.

Wir hatten in Abschn. 3.2.6 verschiedene Auflösungsvermögen als charakteristische Merkmale eines Erderkundungsinstrumentes eingeführt. Die Bedeutung des Auflösungsvermögens (räumliches, spektrales und radiometrisches) eines Scanners für die Identifikation von Objekten soll anhand von Abb. 3.25 illustriert werden. Innerhalb eines Bildelements (Pixels) befinden sich aufgrund des begrenzten räumlichen Auflösungsvermögens mehrere verschiedene Objekte, z.B. Vegetation auf rotem Sandstein. Sowohl Vegetation als auch roter Sandstein haben eine typische Spektralcharakteristik. Der Scanner nimmt jedoch nur die kombinierte Spektralcharakteristik (unteres Teilbild) wahr, da ja beide Objekte innerhalb eines Pixels liegen. Zusätzlich wird diese kombinierte Kurve nicht so detailliert wahrgenommen wie im unteren Teil der Abbildung gezeigt, sondern sie ist gerastert, z.B. in die vier in der Abbildung angedeuteten Spektralbereiche des MSS und in eine gewisse Zahl von Helligkeitswerten (radiometrische Auflösung). Eine eindeutige Identifikation des aufgenommenen Objektes (oder Objektclusters) erfordert daher: (a) eine gute räumliche Auflösung, um möglichst we-

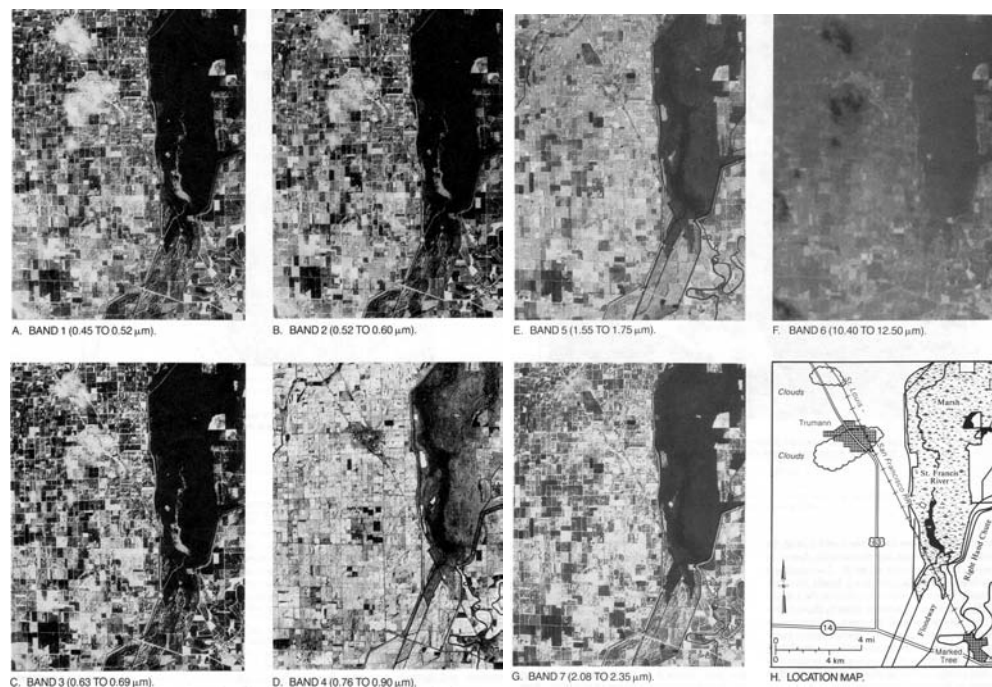


Abbildung 3.24: Einzelbilder der verschiedenen TM Kanäle während einer Aufnahme [101]

nig verschiedene Objekte innerhalb eines Pixels zu haben, und (b) eine gute spektrale und radiometrische Auflösung, um die kombinierte Spektralcharakteristik möglichst gut reproduzieren zu können. Dann sollte es möglich sein, zu erkennen, dass verschiedene Objekte in dem einen Pixel zusammengefasst wurden und wie groß ihr relativer Anteil ist. Mischpixel (Mixed Pixels) bilden den Hauptteil aller übertragenen Pixel: beim MSS sind für mitteleuropäische Verhältnisse nur etwa 1% der Pixel rein, also von nur einem Material erzeugt. Aufgrund des verbesserten räumlichen Auflösungsvermögens sind beim TM schon 18% und bei SPOT sogar 38% der Pixel rein [77].

Interessanterweise lassen sich unter bestimmten Umständen auch Objekte identifizieren, die eigentlich zu kein sind, um bei vorgegebenem Bodenaufklärungsvermögen noch erkannt werden zu können. Ein oft zu beobachtendes Beispiel hierfür ist die Abbildung von Straßen und Wegen, die oft nur wenige Meter breit sind und trotzdem auf den LandSat-Bildern deutlich zu erkennen sind. Diese Situation ist in Abb. 3.26 dargestellt: ein heller Weg (Hellig-

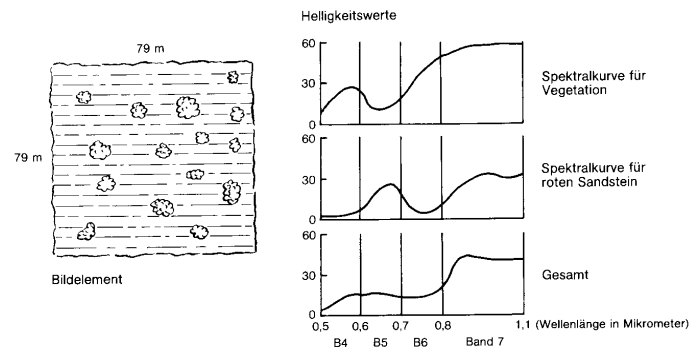


Abbildung 3.25: Integration des spektralen Sensors innerhalb des unmittelbaren Blickfeldes (d.h. innerhalb eines Pixels) eines multispektralen Scanners [71]. Die Bänder entsprechen denen des LandSat MSS

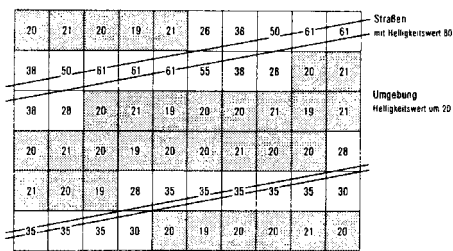


Abbildung 3.26: Erscheinungen, die kleiner sind als ein Bildelement, werden sichtbar, wenn sie sich deutlich vom Hintergrund abheben, wie an diesem Beispiel zweier Wege gezeigt wird [71]

keitswert z.B. 80) durchquert eine Landschaft mit relativ geringem Helligkeitswert (ca. 20). Die Helligkeitswerte aller Bildelemente, die von der Straße beeinflusst werden, liegen deutlich über der des Hintergrundes. Der Weg ist daher sichtbar, nimmt aber eine Breite von ein bis zwei Bildelementen ein, d.h. er erscheint als 79 bzw. 158 m breiter Strich und ist daher stark überzeichnet.⁶ Die Identifikation als Weg ist nur deshalb möglich, weil die Ausdehnung in einer Richtung zwar unter Pixelgröße war, in der anderen Richtung aber weit über Pixelgröße. Einzelne sehr kleine, aber sehr helle Objekte würden zwar das entsprechende Pixel aufhellen, wären aber nicht zu identifizieren.

Ein dem LandSat-Spektralabtaster ähnliches Instrument ist das Advanced Very High Resolution Radiometer AVHRR auf vielen Satelliten der NOAA- und TIROS-N Serie. Die räumliche Auflösung dieses Scanners ist mit ungefähr 1 km wesentlich schlechter als die der LandSat-Instrumente. Dafür wird aber ein 3 000 km breiter Streifen abgescannt (statt der 185 km beim LandSat MSS). Dadurch ergibt sich eine leichte Überlappung der bei zwei aufeinanderfolgenden Orbits gescannten Fläche, so dass innerhalb eines Tages die ganze Erdoberfläche gescannt werden kann. Diese relativ geringe Auflösung ist für verschiedene meteorologische und ozeanographische Anwendungen vollkommen ausreichend, insbesondere da sie mit einer besseren Zeitauflösung verbunden ist. So werden die AVHRR-Aufnahmen in den USA häufig in der Wettervorhersage verwendet. Die Oberflächentemperaturen der Ozeane können bestimmt werden, da es zwei im thermischen Infrarot arbeitende Kanäle gibt. Schneebedeckungen, Wasserressourcen, Landformen und großflächige Vegetationsuntersuchungen werden ebenfalls mit den Daten dieses Instruments durchgeführt (vgl. [16] und Beispiele in [24]).

Auf dem Prinzip des Multispektralabtasters arbeitende Kameras werden jetzt auch für planetare Missionen vorgesehen, z.B. Mars 94/96 [81].

3.2.9 Instrumente der nächsten Generation

In den 90er Jahren ist eine internationale Kampagne 'Mission to Planet Earth' [3] angelaufen,⁷ die eine vollständige Untersuchung der Erde inklusive der Atmosphäre zum Ziel hat. Die Fragestellungen umfassen dabei u.a.: den hydrologischen Zyklus (Verdunstung, Niederschlag, Transport von Wasserdampf, Wolken), die ozeanische und atmosphärische Zirkulation (mit besonderer Berücksichtigung der Oberflächenwinde), die atmosphärische Zusammensetzung (z.B. Wasserdampf, Wolken und Aerosole), eine globale Beobachtung der troposphärischen Chemie, die Untersuchung der Stratosphäre (Ozon, Chemie und Wind), die Strahlungsbilanz und Oberflächentemperaturen sowie die Temperatur der Meeresoberfläche und die damit

⁶Das Phänomen entspricht dem aus der Kartographie bekannten Problem der Darstellung von Straßen in Karten mit kleinem Maßstab durch den Prozess der Verdrängung: eine Autobahn soll auf einer Karte erkennbar sein, egal ob im Maßstab 1:25 000 einer Wanderkarte oder im Maßstab 1:500 000 einer Autokarte. In beiden Fällen hat die Autobahn, in der Realität mit Mittelstreifen und Böschung vielleicht 25 mm breit, eine Darstellungsbreite von 2 mm. Beim großen Maßstab entspricht dies in der Natur einer noch nicht realitätsfernen Breite von 50 m, beim kleinen Maßstab dagegen einer Breite von 1 km.

⁷und partiell auch wieder im Sande versickert, vgl. Abschn. 4.7.

verbundenen Windsysteme. Die verschiedenen Instrumente sollen aber nicht nur zur Untersuchung der Atmosphäre verwendet werden. Andere Fragestellungen beschäftigen sich mit biologischer Aktivität und Ökosystemen, dem Vorkommen von Mineralien und Pflanzenarten für Geologie, Land- und Forstwirtschaft, der Oberflächenmineralogie, der Glaziologie und der Entstehung und möglichen Vorhersage von Erdbeben. Diese verschiedenen Messziele sollen durch eine Vielzahl von auf verschiedene Satelliten verteilte Instrumente untersucht werden.

Viele der für diese Mission geplanten oder teilweise auch schon gebauten Instrumente sind in ihren Möglichkeiten den bisher betrachteten Instrumenten bei weitem überlegen. Einige von ihnen seien hier kurz vorgestellt, um einen Einblick in ihre Möglichkeiten (und teilweise auch Zielsetzungen) zu geben.

Das *High-Resolution Imaging Spectrometer* (HIRIS) soll die detaillierte und genau lokalisierte Messung geologischer, biologischer und physikalischer Prozesse vornehmen. Dazu sind gefordert eine spektrale Auflösung von 200 Kanälen im Spektralbereich zwischen 0.4 und 2.5 μm (sichtbares Licht bis nahes Infrarot) und ein Bodenaufklärungsvermögen von 30 m bei einem Blickfeld von 30 km

Das *High-Resolution Imaging Spectrometer* (HRIS) ist HIRIS ähnlich und dient dem remote sensing von Land- und Küstenbereichen. Dazu gehört die Kartierung der geologischen Strukturen und der Vegetation. Anwendungsziel ist ein Überblick über landwirtschaftliche, geologische und Wald-Ressourcen. Der Spektralbereich umfasst wie bei HIRIS den Bereich von 0.4 bis 2.5 μm . Die spektrale Auflösung ist besser als 20 nm, jedoch können nur 10 Spektralbänder für jedes Bild ausgelesen werden. Insofern verschafft HIRIS einen Überblick (z.B. Vegetation oder Gestein), HRIS gestattet dann durch Verwendung der für die jeweilige Gattung interessanten Spektralbänder eine genauere Bestimmung. Die Bodenaufklärung liegt zwischen 20 und 50 m, das Blickfeld beträgt zwischen 20 und 60 km.

Das *Medium Resolution Imaging Spectrometer* (MERIS) hat die Hauptaufgabe, die Farbe des Ozeans zu bestimmen, woraus sich u.a. auf den Planktongehalt zurückschließen lässt (wichtig für den CO_2 -Kreislauf). Dazu können 9 von 60 Spektralbändern im Bereich von 0.4 bis 1.04 μm ausgewählt werden. Das Bodenaufklärungsvermögen beträgt 500 m (Ozeane sind groß und durchmischen sich relativ schnell), die Breite des Blickfelds beträgt zwischen 1000 und 1500 km.

Das *Moderate-Resolution Imaging Spectrometer: Tilt* (MODIS-T, <http://modis.gsfc.nasa.gov/>) und *Nadir* (MODIS-N) dient der Messung der Landvegetation und Geologie, der Ozeanoberflächentemperatur und Farbe, und von atmosphärischen Aerosolen und Wolken. Der Spektralbereich von 0.4 bis 1.04 μm ist in 64 Bänder eingeteilt.

MODIS und HIRIS dienen im wesentlichen dem Verständnis bio-geochemischer Zyklen, des hydrologischen Zyklus, geologischer Prozesse und langzeitlicher Klimavariationen. Aufgrund der feinen Spektralaufklärung können die Instrumente Gruppen von Organismen identifizieren und damit die komplexen Muster von Ökosystemen und deren Veränderungen auf einer globalen Skala aufzeichnen. Aufgrund seiner besseren räumlichen Auflösung ist HIRIS gleichsam das Zoom-Objektiv von MODIS.

Das *Intermediate and Thermal Infrared Radiometer* (ITIR) soll hochaufgelöste Bilder im nahen und thermischen Infrarot im Spektralbereich von 0.85 bis 11.7 μm erzeugen. Das räumliche Auflösungsvermögen beträgt je nach Wellenlänge zwischen 15 und 60 m ('heutige' Instrumente wie TIR auf HCMM haben eine Auflösung von 600 m). ITIR ist die Ergänzung zu HIRIS im thermalen Infrarot, die Aufgaben umfassen Oberflächenbiologie, Mineralogie, Bodenbeschaffenheit, Bodentemperatur und geothermische Ressourcen.

Außer mit den Imaging-Instrumenten sind diese Satelliten auch mit verschiedenen Typen von Soundern ausgestattet, die teilweise im sichtbaren, teilweise im Mikrowellenbereich arbeiten.

Der *Atmospheric Infrared Sounder* (AIRS, <http://www-airs.jpl.nasa.gov/>) dient zur Messung der atmosphärischen Temperatur und der Spurengase. Die räumliche Auflösung beträgt dabei zwischen 15 und 50 km. AIRS kann bis zu 49° zu beiden Seiten seiner Bodenspur scannen. Damit ist bei zwar geringer räumlicher Auflösung ein schneller Überblick über weite Bereiche möglich. Die spektrale Auflösung ist mit 115 Kanälen hoch genug, um die Identi-

fikation einzelner Spurengase zu ermöglichen. Bei der Messung des Temperaturprofils wird eine Genauigkeit von 1 K erwartet bei einer vertikalen Auflösung von 1 km.

Das *Advanced Microwave Scanning Radiometer* (AMSR, <http://nsidc.org/daac/amr/>) ist ein passives Mikrowelleninstrument, das im Bereich zwischen 6.6 und 31.6 GHz arbeitet. Aufgabe ist die Messung atmosphärischen Wasserdampfes, der Oberflächentemperatur und des Oberflächenwindes über dem Meer. Die Temperaturauflösung liegt im Bereich von 1 K, die räumliche Auflösung variiert mit der Frequenz zwischen 9 und 50 km. Abgetastet werden Blickfelder von 1200 km Seitenlänge.

Die *Advanced Microwave Sounding Unit* (AMSU, <http://aqua.nasa.gov/AMSU3.html>) ist ein Radiometer zur Messung der atmosphärischen Temperatur- und Feuchteprofile. Die Temperaturauflösung beträgt 0.25 bis 1.3 K für Höhen zwischen der Oberfläche und 40 km. Die Temperaturprofile werden innerhalb von 15 Kanälen zwischen 23.8 und 89 GHz gemessen. Wasserdampf wird in fünf Kanälen zwischen 89 und 138 GHz gemessen. Das IFOV beträgt 50 km für die Temperaturmessung und 15 km für den Wasserdampf. Das Blickfeld des Instruments erstreckt sich 50° zu beiden Seiten der Bodenspur.

Das *Altimeter* (ALT) ist ein Radaraltimeter zur Untersuchung der Topographie von Ozeanen und Oberflächen. Verwendet werden zwei Frequenzbänder (5.3 GHz und 13.6 GHz), die Genauigkeit in der Bestimmung der Höhe soll 2 cm betragen.

Das *Atmospheric Lidar* (ATLID, <http://esapub.esrin.esa.it/bulletin/bullet95/MARINI.pdf>) ist ein Lidar, d.h. es arbeitet wie Radar, jedoch werden hier keine Mikrowellen eingesetzt, sondern es wird ein Laser verwendet. Aufgabe des Instruments ist die Bestimmung der Höhe von Wolkenoberkanten, atmosphärischen Inversionsschichten und der Verteilung von Aerosolschichten. Verwendet wird Strahlung im Bereich von 1.06 oder 1.53 μm . Die räumliche Auflösung beträgt 100 bis 500 m in vertikaler und 10 bis 50 km in horizontaler Richtung.

Der *Laser Atmospheric Wind Sounder* (LAWS, <http://www.lidar.com/windswg/windswg2.htm>) ist ein Doppler-Lidarsystem zur Messung der troposphärischen Luftbewegungen (Prinzip: Laser-Doppler-Anemometrie). Die vertikale Auflösung beträgt 1 km, das Blickfeld 100 km². Die Geschwindigkeit kann auf 1 m/s genau bestimmt werden.

Das *Geoscience Laser Ranging System* (GLRS, <http://icesat.gsfc.nasa.gov/12538-r10.pdf>) ist ein Laser-Altimeter zur Bestimmung der Bewegung der Erdplatten. Dazu sind verschiedene Punkte auf der Oberfläche mit reflektierenden Würfeln ausgestattet, so dass ihre Wanderung verfolgt werden kann. Ferner dient das GLRS zur topographischen Kartierung von Eis- und Landflächen sowie Wolkenoberkanten. Im Gegensatz zum Radaraltimeter ist der Durchmesser des abtastenden Strahls am Boden mit 80 bis 290 m klein.

3.2.10 Zusammenfassung

Wir haben in diesem Abschnitt verschiedene Instrumente kennengelernt, die das von der Erde reflektierte Licht im sichtbaren und nahen infraroten Bereich aufzeichnen. Die wichtigsten waren lange Zeit konventionelle photographische Systeme, die außer bei bemannten und vielleicht auch militärischen Missionen immer mehr an Bedeutung verlieren, und durch Scanner und opto-elektronische Detektoren ersetzt werden – heute allerdings mit der Verwendung von CCD-Arrays wieder an ihre Ursprünge zurück kehren. Wichtigste Kenngrößen dieser Instrumente sind das räumliche und das spektrale Auflösungsvermögen. Insbesondere die Kombination verschiedener Spektralkanäle und die Verwendung des nahen Infrarot sind bei der Interpretation der Bilder (z.B. Identifikation von Vegetation) wichtig.

3.3 Passive Instrumente im nicht-sichtbaren

Passive Instrumente im nicht-sichtbaren Bereich detektieren die von den Untersuchungsobjekten emittierte Strahlung. Für terrestrische Objekte ist dies im wesentlichen die thermische Infrarot-Strahlung. Bei der Sonne und anderen astrophysikalischen Objekten wird vom Satelliten auch die harte elektromagnetische Strahlung vom Ultraviolett über Röntgen- bis hin

zur γ -Strahlung untersucht, da diese auf Grund der atmosphärischen Absorption vom Boden aus nicht gemessen werden kann – und umgekehrt natürlich mögliche terrestrische Strahlungsquellen in diesem Bereich vom Satelliten nicht detektiert werden können.

Weitere Instrumente detektieren die von terrestrischen Objekten emittierte Mikrowellenstrahlung (passive Mikrowelleninstrumente) oder den über das gesamte Spektrum integrierten Strahlungsstrom (Solarkonstante bzw. deren reflektierter Anteil).

3.3.1 Thermisches Infrarot

Instrumente, die im thermischen Infrarot arbeiten, messen die Temperaturstrahlung von Objekten. Diese Objekte können die Oberflächen der Ozeane, Vegetation, Landoberflächen aber auch Wolken sein. Will man die Daten dieser Instrumente auswerten, so benötigt man eine Umrechnung der Intensität der elektromagnetischen Strahlung über einen gewissen Wellenlängenbereich im Infraroten auf eine Temperatur. Die physikalische Grundlage hierfür sind die Strahlungsgesetze, insbesondere das Stefan-Boltzmann-Gesetz und das Wien'sche Verschiebungsgesetz.

Die Strahlungsgesetze setzen voraus, dass sich ein System im thermodynamischen Gleichgewicht befindet und erlauben dann eine Charakterisierung des von einem Körper ausgesandten Strahlungsfeldes in Abhängigkeit von der Temperatur. Betrachten wir dazu einen Hohlraum, dessen Wände auf einer konstanten Temperatur T gehalten werden (z.B. in einem Wärmebad dieser Temperatur). Von diesen Wänden werde Strahlung in den Hohlraum emittiert und die aus diesem auf die Wände fallende Strahlung wird von letzterem ganz oder zum Teil absorbiert. Je größer die Strahlungsmenge im Hohlraum ist umso größer wird auch die auf die Wände fallende Menge sein, und um so größer auch die von ihnen absorbierte Menge. Die Menge der Strahlung im Hohlraum wird daher zu- oder abnehmen, bis der Verlust durch Absorption der von den Wänden ununterbrochen emittierten Strahlung gerade das Gleichgewicht hält. Die Gleichgewichtsverteilung der Strahlung, sowohl im Bezug auf ihre Dichte als auch im Bezug auf ihre Frequenz, wird dann durch die Temperatur der Wände vollkommen bestimmt.

Das *Kirchhoff-Planck'sche Gesetz* beschreibt das Verhältnis von spektralem Absorptionsvermögen a_λ zu spektralem Emissionsvermögen ϵ_λ für einen beliebigen Körper im thermodynamischen Gleichgewicht:

$$\epsilon_\lambda = a_\lambda \cdot S_\lambda(T). \quad (12.4)$$

Darin ist die Ergiebigkeit $S_\lambda(T)$ eine universelle Funktion der Temperatur und der Frequenz. Sie wird auch als Quellfunktion bezeichnet. Für $a_\lambda = 1$ ergibt sich maximale Absorption und damit auch die maximale Ausstrahlung für den Körper. Ist $a_\lambda = 1$ für alle λ , so handelt es sich um einen *schwarzen Körper* und die universelle Funktion der Temperatur wird durch das *Planck'sche Strahlungsgesetz* beschrieben. Dieses gibt an, wie viel Strahlungsenergie pro Wellenlängeneinheit von einer Flächeneinheit eines schwarzen Körpers in einer Zeiteinheit in einen Einheitsraumwinkel abgestrahlt wird:

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{\lambda T}} - 1} \quad \text{oder} \quad B_\nu(T) = \frac{2h\nu^3}{c^2} \cdot \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \quad (3.8)$$

Die gesamte Ausstrahlung eines schwarzen Körpers ergibt sich durch Integration über die Planck'sche Strahlungskurve. Der Zusammenhang zwischen Gesamtstrahlungsstrom und Temperatur wird durch das *Stefan-Boltzmann Gesetz* beschrieben:

$$q = \pi \int_0^\infty B_\lambda(T) d\lambda = \sigma T^4, \quad (3.9)$$

Die gesamte Ausstrahlung eines schwarzen Körpers ist also einzig eine Funktion seiner Temperatur und der universellen Konstante $\sigma = 5.67051 \cdot 10^{-8} \text{ W}/(\text{m}^2 \text{ K}^4)$. Für beliebige andere (d.h. nicht-schwarze) Körper lässt sich bei der Definition eines mittleren Emissionsvermögens $\bar{\epsilon}$ die von der Fläche in den Halbraum abgegebene Gesamtenergie bestimmen zu

$$q = \sigma \bar{\epsilon} T^4. \quad (3.10)$$

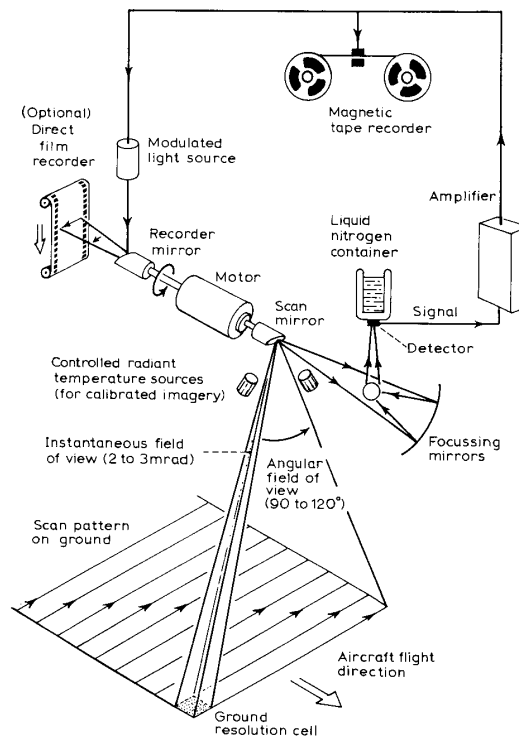


Abbildung 3.27: Prinzip eines Infrarotscanners [16]

Das Stefan-Boltzmann Gesetz erlaubt eine einfache Abschätzung der Temperatur eines Körpers. Es wird auch auf Körper angewendet, die keine schwarzen Strahler sind. In diesem Falle definiert man dann die Effektivtemperatur eines Körpers als die Temperatur, die ein schwarzer Strahler haben müsste, um den gleichen Gesamtstrahlungsstrom zu emittieren:

$$q = \sigma T_{\text{eff}}^4 . \quad (3.11)$$

Das *Wien'sche Verschiebungsgesetz* ergibt sich aus dem Planckschen Strahlungsgesetz durch Ableitung nach λ und Nullsetzen. Es beschreibt die Wellenlänge des Maximums der Planckschen Kurve in Abhängigkeit von der Temperatur des Strahlers:

$$\lambda_{\text{max}} \cdot T = \text{const} = 2884 \mu\text{m K} . \quad (3.12)$$

Mit Hilfe des Wien'schen Verschiebungsgesetzes lässt sich bei bekannter Wellenlänge des Maximums unter der Annahme, dass der Körper ein schwarzer Strahler sei, dessen Temperatur abschätzen.

Nachdem wir die Interpretation der Daten kurz beschrieben haben, wollen wir uns jetzt den Grundlagen des Gerätes zuwenden. Abbildung 3.27 zeigt das Funktionsprinzip eines Scanners, der im thermischen Infrarot arbeitet. Dieser Scanner unterscheidet sich von den opto-mechanischen Scannern wie z.B. dem MSS auf LandSat nur geringfügig. Auch hier wird mit Hilfe eines Drehspiegels ein Streifen senkrecht zur Flugrichtung abgetastet. Die Signale werden mit Hilfe einer Spiegeloptik auf einen Detektor projiziert. Der entscheidende Unterschied zu einem im sichtbaren Licht arbeitenden Scanner liegt in der Betriebsweise des Detektors: dieser wird auf 73°K heruntergekühlt, um sein thermisches Rauschen zu verringern. Kühlt man nicht, so hat man eine Mischung aus Detektorrauschen und Signal, wobei das Rauschen in der Größenordnung des Signals liegt. Die Kühlung erfolgt mit flüssigem Stickstoff. Die in Abbildung 12.10 angedeutete Speicherung der Signale auf Film erfolgte nur bei den ersten Infrarotinstrumenten, heutzutage wird nur Digital verarbeitet.

Scanner im Infrarotbereich arbeiten innerhalb der atmosphärischen Fenster zwischen 3.5 – 5.5 μm und zwischen 8.0 – 14.0 μm (vgl. auch Abb. 3.4). Thermisches Infrarot wird häufig

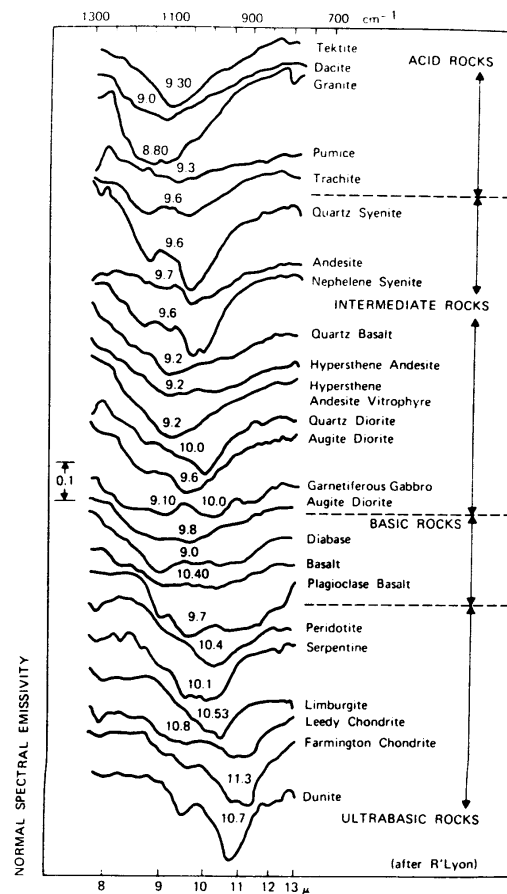


Abbildung 3.28: Spektrale Signaturen von Gestein. Viele Gesteinsarten unterscheiden sich durch ihre spektrale Emission im thermischen Infrarot voneinander [4]

auch von Flugzeugen aus beobachtet, Beispiele für Satelliteninstrumente sind der Kanal 6 des Thematic Mapper auf LandSat und das TIR-Instrument auf der Heat Capacity Mapping Mission HCMM. Letzteres lieferte die ersten relativ hochauflösenden Daten im thermischen Infrarot mit einem Bodenaufklärungsvermögen von 600 m und einem radiometrischen Auflösungsvermögen von 256 Kanälen innerhalb einer Temperaturskala von 260 K bis 340 K (entsprechend einer Temperaturauflösung von 0.4 K). Mit dieser Auflösung lassen sich Wälder, landwirtschaftlich genutzte Flächen und bebauten Gebiete unterscheiden, ja sogar der innerstädtische Temperaturgradient (im Zentrum wärmer als zu den Rändern) ist deutlich erkennbar. Selbst Autobahnen zeigen sich im thermischen Infrarot als deutliche Wärmebänder.

Eine ausführliche Beschreibung von Daten und Datenverarbeitung sowie Dateninterpretation für HCMM findet sich in Großmann [32], der auch eine große Zahl von Beispielen für bearbeitete und unbearbeitete Aufnahmen im thermischen Infrarot zeigt. Man sollte beachten, dass die Teilinformationen über den Zustand der Vegetation nicht aus dem thermischen Infrarot kommen (das wäre die Temperatur der Pflanze), sondern aus dem nahen Infrarot (reflektiertes Infrarot, sehr stark von der internen Blattstruktur und damit dem Stress, dem die Pflanze ausgesetzt ist, abhängig). Fragen an die thermischen Infrarotaufnahmen sind eher von der Art, ob sich anthropogene Einflüsse auf den Wärmehaushalt von Landflächen erkennen lassen [32]. Hat man ein ausreichend feines spektrales Auflösungsvermögen, so lassen sich z.B. auch einzelne Gesteinsarten unterscheiden. Wie in Abb. 3.28 gezeigt, haben verschiedene Gesteinsarten durchaus unterschiedliche Charakteristika im spektralen Emissionsvermögen im infraroten Bereich. Ein anderes Beispiel für die Anwendung von Infrarotsensoren ist das Temperature Humidity Infrared Radiometer (THIR) auf Nimbus-7. Dieses nutzt

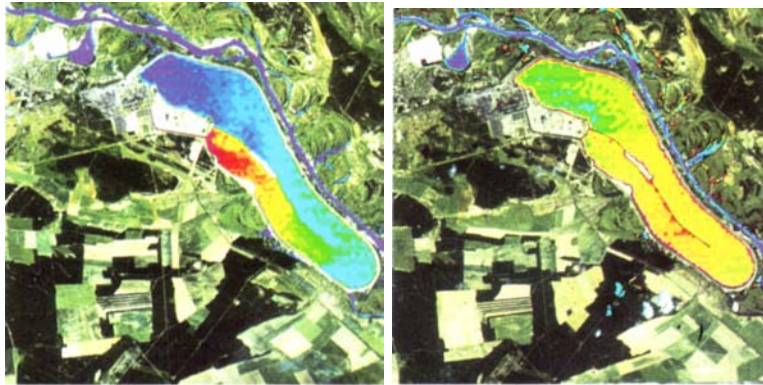


Abbildung 3.29: Tschernobyl 1985 (links) und 1986 (rechts), Aufnahme im thermischen Infrarot, LandSat [14]

die Infrarotstrahlung der Wolken aus, um Karten der Wolkenbedeckung zu erzeugen, gemessen wird in einem relativ schmalen Band um $11.5 \mu\text{m}$. Hier ist z.B. die Kopplung zwischen Veränderungen in der Wolkenbedeckung und im Zirkulationssystem der Ozeane von Interesse (z.B. El Niño und die Southern Oscillation, ein Wechsel zwischen warmen und kalten Oberflächenströmungen im Südpazifik, die sowohl für die regionale Ökologie und Ökonomie als auch für das großräumige Klima eine starke Bedeutung haben, vgl. Abbildung 10-4 in [16]).

Anwendung Tschernobyl

Abbildung 3.29 zeigt, wie Abb. 3.11, einen Vorteil der Erdfernerkundung mit Hilfe von Satelliten: das Monitoring und damit die Entdeckung von Veränderungen. Hier sind Aufnahmen im thermischen Infrarot (Wasserflächen) mit einer Aufnahme im sichtbaren (Landflächen) kombiniert für Tschernobyl. Die linke Aufnahme wurde von LandSat 1985 gemacht: die Einleitung des Kühlwassers aus dem Reaktorblock in den See ist als deutliches thermisches Signal zu erkennen. Entsprechend ist der See am anderen Ende auf Grund des aus den Flüssen zugeführten Wassers relativ kalt. Die rechte Aufnahme aus dem Jahre 1986 zeigt den still gelegten Reaktor, erkennbar an der nun relativ gleichmäßigen Temperatur des Sees, dem Fehlen einer ausgeprägten Temperaturdifferenz zwischen Ein- und Auslass des Reaktorblocks sowie der thermischen Isolation vom benachbarten kühleren Flusssystem.

Die thermische Infrarot-Aufnahme erlaubt damit, im Gegensatz zur Aufnahme im sichtbaren Bereich, nicht nur die Identifikation von strukturellen Veränderungen (Baumaßnahmen bzw. in diesem Fall Zerstörung der Reaktorgebäude) sondern würde auch bei intakter Struktur eine Identifikation veränderter Betriebsmodi erlauben.

Abbildung 3.30 zeigt ein Beispiel für eine weiter bearbeitete Aufnahme im thermischen Infrarot: in dieser Aufnahme sind Informationen von verschiedenen (meist geostationären) Wettersatelliten zu einer innerhalb von insgesamt 3 h entstandenen globalen Aufnahme kombiniert. Mit Hilfe eines Wolkenidentifikationsalgorithmus wurde bestimmt, ob eine bestimmte Stelle bewölkt war oder nicht, entsprechend wurde der Bildpunkt einer der Temperaturskalen für Wolken, Eis und Schnee, Wasser oder Land zugeordnet. Das Bodenaufklärungsvermögen beträgt ungefähr 40 km.

Anwendung: Unterirdischer Wasserlauf

Dass Satelliten zur Untersuchung der Hydrosphäre, d.h. der Wasserschicht des Planeten verwendet werden, ist einleuchtend, insbesondere wenn man berücksichtigt, dass 70% der Erdoberfläche durch Wasser bedeckt sind. Der größte Teil dieses Wassers befindet sich in den Ozeanen, süßwasserhaltige Gewässer machen nur einen kleinen Teil der Oberflächengewässer

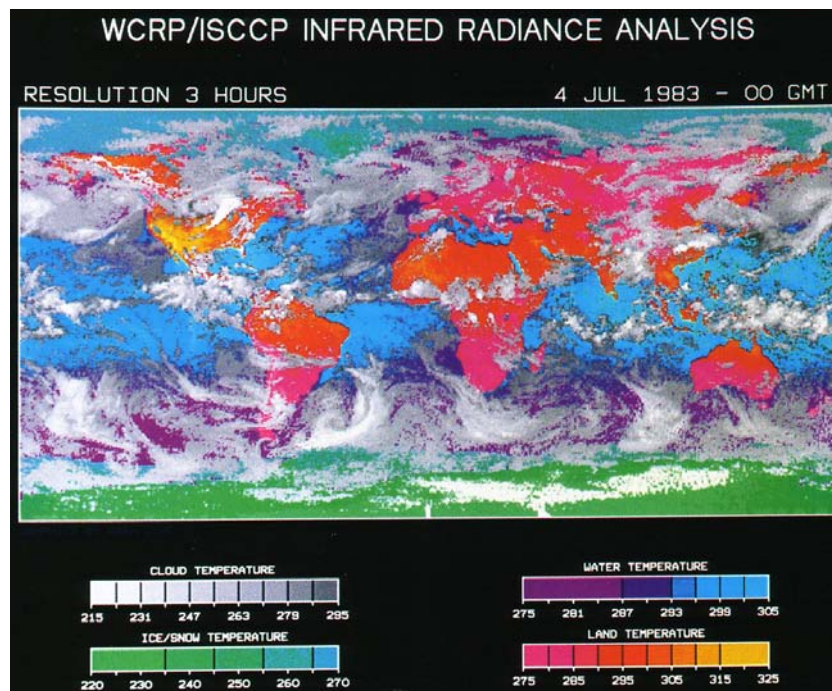


Abbildung 3.30: Zusammengesetzte Aufnahme von Wolken-, Land- und Meeresoberflächentemperaturen [91]

aus. Weitere Süßwasservorkommen befinden sich unterhalb der Erdoberfläche, so findet sich z.B. ein sehr großes Reservoir fossilen Wassers unter der Sahara [51]. Die Entdeckung und Beobachtung unterirdischer Wasserreservoirs und -ströme durch Satelliten mag vielleicht auf den ersten Blick eine etwas befremdliche Aufgabe der Erdfernerkundung sein. Warum sollte man durch den Blick von außen auf die Erde in der Lage sein, Wasservorkommen unterhalb der Erdoberfläche zu erkennen? Was sind die spezifischen Merkmale solcher Wasservorkommen?

Wenn man sich Aufnahmen im thermischen Infrarot, d.h. im Bereich der von einem Körper emittierten Wärmestrahlung ansieht, so fallen Wasserflächen in der Regel sofort auf, da sie aufgrund ihrer großen Wärmekapazität tagsüber oft kühler, nachts aber wärmer sind als ihre Umgebung – und da sie kaum einen Tagesgang in der Temperatur zeigen (Beispiele z.B. in [101]). Insbesondere das Signal eines Wasserkörpers als warmes Element auf einer nächtlichen Aufnahme kann man sich bei der Suche nach bzw. Überwachung von unterirdischen Was-

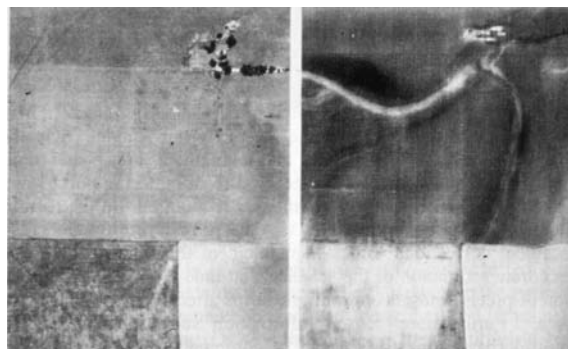


Abbildung 3.31: Unterirdische Wasserläufe werden im thermischen Infrarot sichtbar [16]

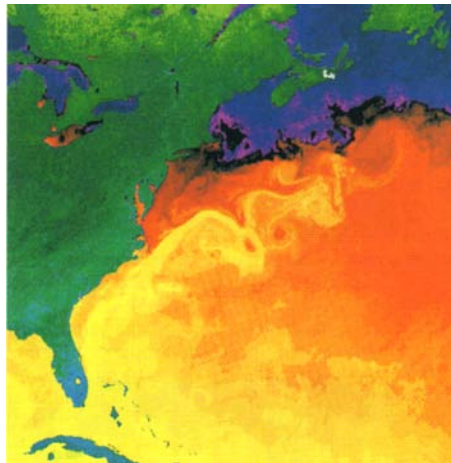


Abbildung 3.32: Der Golfstrom im thermischen Infrarot [119]

serflächen zu Nutzen machen. Abbildung 3.31 zeigen dies sehr anschaulich in einem Vergleich einer im sichtbaren Licht gemachten Aufnahme mit einer Aufnahme im thermischen Infrarot. Im sichtbaren Licht lassen sich unter landwirtschaftlich genutzten Flächen keine Hinweise auf Wasser finden. Im thermischen Infrarot zeichnen sich nachts dagegen deutlich zwei Wasseradern ab. Deren Temperaturunterschied zur Umgebung lässt sogar einige Rückschlüsse auf die Eigenschaften dieser Wasseradern zu: die eine Ader zeichnet sich sehr hell ab, d.h. sie ist deutlich wärmer als der sie umgebende Boden. Das lässt auf eine relativ schnelle Strömung des Wassers schließen, da bei langsamer Strömung ein Wärmeaustausch mit der Umgebung stattgefunden hätte. Diese wasserführende Ader besteht daher aus ziemlich grobem Material, das eine relativ hohe Durchflussgeschwindigkeit erlaubt. Eine andere Wasserader dagegen zeichnet sich nur wesentlich schwächer gegenüber der Umgebung ab, das Material muss also relativ schlecht wasserdurchlässig sein, so dass Wärmeaustausch mit der Umgebung stattfinden kann.

So attraktiv dieses Verfahren auch erscheinen mag, unterirdische Wasseradern mit Hilfe von Satellitenaufnahmen im thermischen Infrarot aufzuspüren, so deutlich muss man auch die Einschränkung sehen: liegen die Wasseradern oder Reservoirs in einer größeren Tiefe als ca. 5 m, so ist (a) der Tagesgang der Temperatur zu gering und die Wasservolumina haben ungefähr die Umgebungstemperatur angenommen und (b) selbst bei schnell fließenden Gewässern ist das bis an die Oberfläche dringende Signal zu schwach, um im Rauschen erkannt zu werden.

Anwendung Meere: Golfstrom

Die Untersuchung von Oberflächentemperaturen von Gewässern ist eine der Hauptanwendungen für Sensoren im thermischen Infrarot – auch das Beispiel in Abb. 3.29 basiert ja auch auf der Untersuchung von Oberflächentemperaturen.

Abbildung 3.32 zeigt als Beispiel den Golfstrom. Deutlich zu erkennen ist der Kern des Golfstroms (gelb), wie er sich von Florida relativ dicht entlang der Ostküste in nordöstlicher Richtung ausbreitet. Durch Instabilitäten kommt es zur Ausbildung von Wirbeln und damit zu einer Auflösung der eigentlichen Stromstruktur.

Die relativ deutliche Grenze zwischen kaltem (blau) und wärmerem (rot, gelb) Wasser im oberen Teil der Abbildung weist allerdings auch auf ein Problem bei der Verwendung von Falschfarbenaufnahmen hin: die Grenze wird durch die Veränderung der Farbe von Rot nach Blau wesentlich deutlicher hervor gehoben als ein Bereich, in dem sich die Farbe von Gelb nach Rot ändert – obwohl der Temperatursprung der gleiche ist. Daher stets Vorsicht vor Rot in Falschfarbenaufnahmen: rot hat eine hohe Signalwirkung, ist bei Falschfarben aber

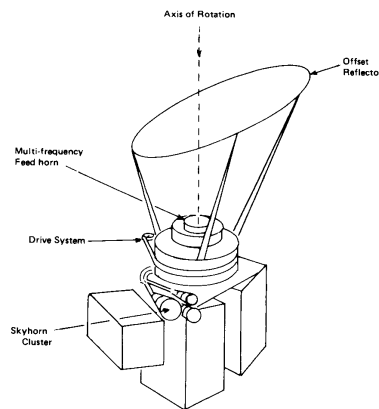


Abbildung 3.33: Scannendes Multifrequenz Mikrowellenradiometer [16]

häufig einfach nur eine Abstufung wie jede andere Farbe auch. Daher werden trotz moderner Aufnahmetechniken in der Medizin CT-Scans und ähnliches immer noch in Graustufen dargestellt und nicht in Farben.

3.3.2 Mikrowelleninstrumente

Detektoren im Mikrowellenbereich haben gegenüber anderen Detektoren zwei Vorteile: (a) sie arbeiten auch bei Bewölkung und (b) sie arbeiten auch nachts, da sie nicht die reflektierte solare Strahlung messen, sondern die von den Körpern emittierte Mikrowellenstrahlung. Dieses Prinzip haben sie mit den Detektoren im thermischen Infrarot gemeinsam. Dem steht jedoch ein entscheidender Nachteil der Mikrowellendetektoren gegenüber: auf Grund der großen Wellenlängen ist das räumliche Auflösungsvermögen dieser Detektoren gering (vgl. (3.6)). Auch ist der emittierte Strahlungsstrom im Mikrowellenbereich nur gering, d.h. die radiometrische Auflösung ist schlecht.

Passive Mikrowellensensoren arbeiten bei Frequenzen zwischen 200 GHz und 1 GHz bzw. Wellenlängen zwischen 0.15 cm und 30 cm. Das räumliche Auflösungsvermögen verringert sich daher um drei bis vier Größenordnungen, wenn wir vom infraroten Bereich in den Mikrowellenbereich gehen. Für die kürzeste Wellenlänge (Frequenz von 37 GHz) des Scanning Multichannel Microwave Radiometer SMMR auf Nimbus-7 ergibt sich eine Pixelgröße von 18 km x 27 km, für seine längste Wellenlänge (entsprechend 6.6 GHz) eine von 95 km x 148 km. Die Hauptanwendungsgebiete von passiven Mikrowelleninstrumenten liegen daher in den Bereichen, in denen ein sehr gutes räumliches Auflösungsvermögen nicht unbedingt erforderlich ist. Dazu gehören die Meteorologie (Messung von Temperaturprofilen in der Atmosphäre) und die Ozeanographie (z.B. Meereisbedeckung, insbesondere die Veränderungen der Meereisbedeckung in der Arktis und Antarktis als eine Art Frühwarnsystem für globale Temperaturänderungen).

Abbildung 3.33 zeigt den schematischen Aufbau eines scannenden Mikrowellenradiometers. Gescannt wird durch zwei Bewegungen, die Rotation des Instruments (entspricht dem Abtasten senkrecht zur Bewegungsrichtung) und die Vorwärtsbewegung des Satelliten. Da das einfallende Signal nur sehr gering ist, ist eine Reduktion des Rauschens des Empfängers ein wesentlicher Bestandteil. Der Empfänger besteht im wesentlichen aus einer Hochgewinnantenne, anschließender Signalverarbeitung und einer Quelle für ein Referenzsignal. Zum Arbeitsprinzip und zum Verständnis technischer Details sei auf Rees [93] verwiesen.

Abbildung 3.34 zeigt als Beispiel für Daten von einem passiven Mikrowelleninstrument die über den Zeitraum von Januar 1979 bis Dezember 1986 gemittelten Polarisationsunterschiede (horizontal vs. vertikal) bei 37 GHz, beobachtet vom Scanning Multi-Channel Radiometer SMMR auf Nimbus-7. Über unbewachsenen Oberflächen ist diese Differenz relativ groß (25 bis 35 K), mit zunehmender Vegetationsdichte verringert sich diese. Daher kann ein

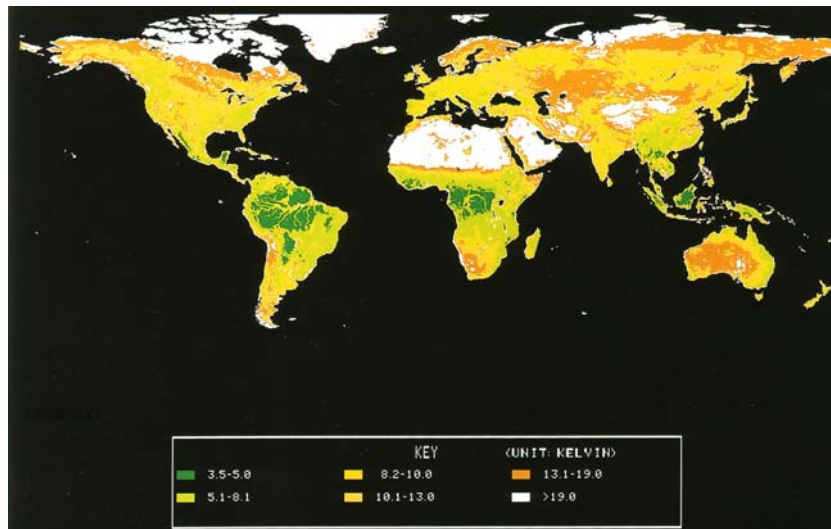


Abbildung 3.34: Polarisationsunterschiede bei 37 GHz beobachtet mit dem Scanning Multi-Channel Radiometer SMMR auf Nimbus-7 [15]

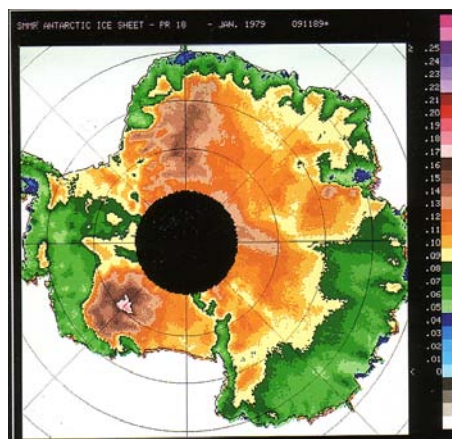


Abbildung 3.35: Polarisationsunterschiede bei 18 GHz beobachtet mit dem Scanning Multi-Channel Radiometer SMMR auf Nimbus-7 über der Antarktis im Januar 1979 [121]

Mikrowelleninstrument auch zur großräumigen Überwachung von Vegetation und zur Untersuchung von Desertifikation verwendet werden. In der Abbildung erkennt man die höchsten Werte über den großen Wüsten (Sahara, Namib, Rub Al Khali usw.), die geringsten über den Regenwäldern Brasiliens, Kameruns und Indonesiens. In der Sahara sind die Bergregionen von Ahaggar, Adrar und Air deutlich an ihrem Bewuchs zu erkennen. Die Unterschiede in der Polarisation entstehen im wesentlichen durch den Wassergehalt, so dass Abb. 3.34 im wesentlichen eine Kombination aus Pflanzenbewuchs und Bodenfeuchte darstellt – für die Detaillklassifikation der Vegetation sicherlich weniger geeignet als Kombinationen aus sichtbarem Licht und nahem IR; unter dem Aspekt der Wüstenbildung auf Grund des betrachtene allgemeineren Kriterium Feuchte bzw. Wassergehalt dagegen eher geeignet.

Polarisationsunterschiede bei 18 GHz basieren ebenfalls auf der Existenz von flüssigem Wasser im Bildbereich. Ein Anwendungsbeispiel ist in Abb. 3.35 gezeigt: Polarisationsunterschiede über dem arktischen Eisschild. Blau sind die Bereiche mit einem signifikanten Gehalt an flüssigem Wasser, d.h. auch die Bereiche, in denen im Sommer der Eisschild schmilzt.

Eis lässt sich zwar auch im sichtbaren und im thermischen IR gut nachweisen, jedoch

ist die Dynamik des Eises dort nur durch sein saisonales Verschwinden bzw. Auftauchen zu erkennen. Diese Verdahren weisen also nur ein vollständiges Verschwinden der Eisdecke nach. Für die Bedeckung mit Meereis ist dies sinnvoll. Ein Eisschild verändert sich jedoch lange Zeit in seiner Mächtigkeit bis er an den Rändern so weit abgeschmolzen ist, das dort auch sein zurückweichen zu erkennen ist. Schmelzende Eis ist jedoch gleichbedeutend mit flüssigem Wasser, so dass die Daten aus Abb. 3.35 einen Hinweis auf die Bereiche geben, in denen der arktische Eisschild taut.

3.4 Aktive Instrumente

Das Paradebeispiel für ein aktives Instrument ist das Radar. Es ist, wie die anderen aktiven Mikrowelleninstrumente, ein vergleichsweise selten auf Satelliten eingesetztes Instrument, da es als aktives Instrument selber Strahlung emittieren muss und damit einen sehr hohen Energiebedarf hat. Radarsatelliten sind häufig mit kleinen Kernreaktoren ausgestattet und können so aufgrund der von ihnen emittierten γ -Strahlung leicht identifiziert werden [87]. Die Reaktoren sollen, wenn der Satellit abzusinken beginnt, vom Satelliten getrennt und in eine höhere Umlaufbahn geschossen werden. Diese Trennung funktioniert nicht immer, so dass einige Reaktoren abgestürzt sind (z.B. Kosmos 954 im Januar 1978 über Kanada, Kosmos 1402 im Februar 1983 über dem Indischen Ozean). Die wichtigsten zivilen Radarinstrumente auf Satelliten sind die SeaSat-1 Instrumente, die Instrumente auf ERS-1 (Earth Resource Satellite) sowie die im Shuttle verwendeten SIR-A und SIR-B Instrumente. Die Raumsonde Magellan ist mit einem Radar zur Kartographie der Venus ausgestattet.

Radar basiert auf der Verwendung von Mikrowellen, hat jedoch gegenüber dem oben beschriebenen passiven Instrument den Vorteil einer höheren räumlichen Auflösung. Mit dem aktiven Instrument kann man außerdem gegenüber einem passiven Instrument zusätzliche Informationen gewinnen. Die zusätzlichen Messgrößen sind dabei die Laufzeit des Signals vom Satelliten zum Boden und zurück, die Dopplerverschiebung der Frequenz aufgrund der relativen Bewegung des Satelliten gegenüber dem Boden und die Polarisation der Strahlung. Mit diesen Informationen lassen sich drei unterschiedliche Typen aktiver Mikrowelleninstrumente konstruieren: Altimeter, Scatterometer und das Synthetic Aperture Radar SAR.

3.4.1 Altimeter und Scatterometer

Das Altimeter ist ein Instrument, das aus der Reflektion eines senkrecht nach unten gestrahlten Mikrowellenpulses den Abstand des Satelliten gegenüber dem Boden bestimmt. Derartige Instrumente wurden verwendet auf SkyLab, GEOS-3, SeaSat und ERS-1. Die Idee hierbei ist eine genaue Bestimmung des Erdgeoids (das ist eine fiktive Äquipotentialfläche der Erde in ungefähr der Höhe der Erdoberfläche). Besonders interessant sind hier die Ozeane, da deren Oberfläche bzw. die Abweichungen der Meereshöhe von der eines Referenzellipsoids die direktesten Signaturen für die Abweichung des Erdgeoids von diesem Ellipsoid geben. Dazu muss die über einen langen Zeitraum gemittelte Höhe der Meeresoberfläche bestimmt werden (der lange Zeitraum ist erforderlich, da man sonst die Wellenhöhe bestimmen würde, nicht aber die mittlere Höhe der Meeresoberfläche). Dazu müssen zwei Größen genau bestimmt werden, (1) die Höhe des Satelliten über der senkrecht darunter liegenden Meeresoberfläche und (2) die genaue Position des Satelliten.

Zur Bestimmung der Höhe des Satelliten über der Meeres- oder Landoberfläche sendet das SeaSat-Altimeter kurze Impulse ($3.2 \mu\text{s}$) bei einer Frequenz von 13.5 GHz aus, die Wiederholfrequenz dieser Pulse beträgt 1020 Hz. Das Prinzip dieser Messungen ist in Abb. 3.36 angedeutet. Das Altimeter hat eine Genauigkeit von ± 10 cm in der Bestimmung der Höhe des Geoids. Diese Genauigkeit ergibt sich nicht einfach aus der Laufzeit des Signals, sondern erfordert genaue Korrekturen für (a) den Abstand der Altimeterantenne vom Massenmittelpunkt des Raumfahrzeuges, (b) interne Instrumentungenauigkeiten z.B. durch Signallaufzeiten, (c) den Einfluss der Atmosphäre, (d) Veränderungen im Reflektionsverhalten in Abhängigkeit von der Rauigkeit oder Bewegung der Oberfläche, (e) die Topographie der Ozeanoberfläche

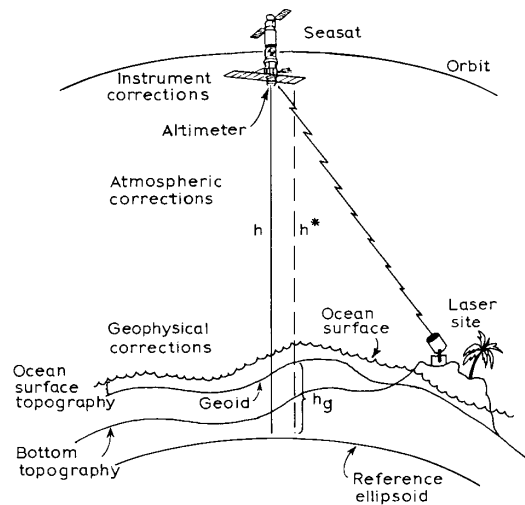


Abbildung 3.36: Funktionsweise des SeaSat-Altimeters [16]

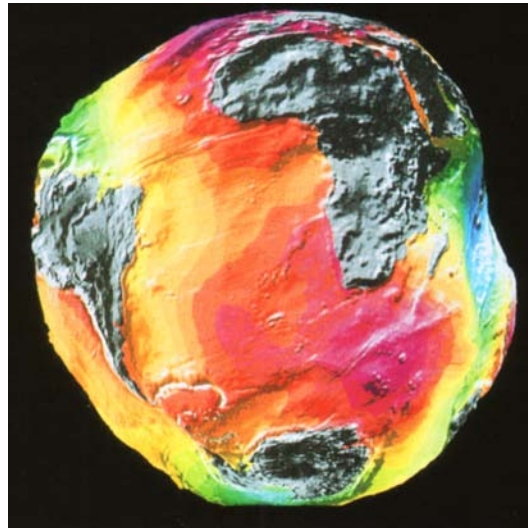


Abbildung 3.37: Geoid: 'wahre' Erdform [34]

durch Zirkulationen, barometrische Effekte und Luftbewegung, (f) Ungenauigkeiten im Referenzgeoid, (g) Gezeiten und (h) zufällige Fehler. Die sich aus der Reflektion des Signals ergebende Höhe wird mit einer Genauigkeit von ± 5 cm vermessen, die Ungenauigkeit aufgrund der oben genannten Abweichungen führt auf den Gesamtfehler von ± 10 cm. Zum Vergleich: vor SeaSat war der Geoid lediglich mit einer Genauigkeit von ± 1 m bekannt.

Die zur Zeit genaueste Vermessung des Geoids ist durch den im Juni 1991 gestarteten ERS-1 erfolgt. Die Messgenauigkeit in der Höhe liegt bei wenigen Zentimetern. Für die Schwankungen in der Meereshöhe, die wir ja sonst gerne als den absoluten Bezugspunkt für Höhenmessungen (Höhe über NN) angeben, ergeben sich beachtliche Größen. So ragt die Wasserfläche nordöstlich von Australien bis zu 85 m über den Durchschnittswert empor und liegt im nördlichen indischen Ozean bis zu 105 m darunter. Die Ursachen für solche langwelligeren Ondulationen in der Form des Geoids sind wahrscheinlich Temperaturschwankungen im flüssigen Erdmantel. Kühler ist er insbesondere in jenen Regionen, wo kontinentale Platten während ihrer Subduktion noch nicht vollständig erwärmt sind. Je kühler das Magma ist, umso größer ist aber auch seine Dichte und damit die Gravitation an dieser Stelle. Alles,

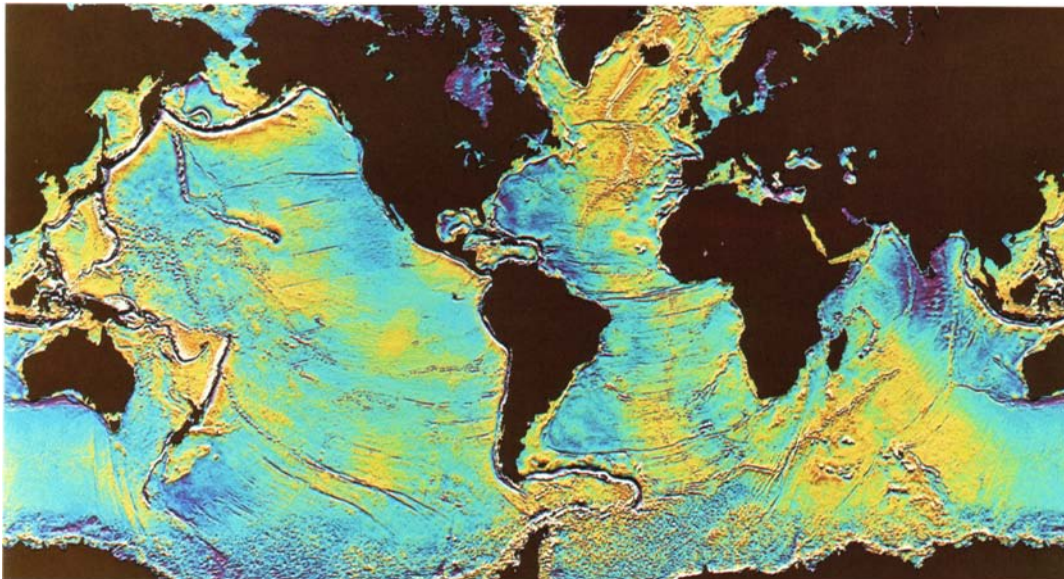


Abbildung 3.38: Sea Floor Mapping: die Form des Ozeanbodens mit Gräben und Kontinenten [14]

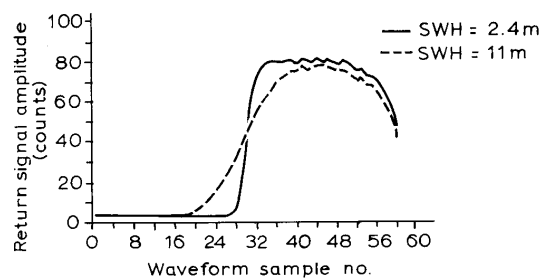


Abbildung 3.39: Form des reflektierten Signals des SeaSat-Altimeters für verschiedene signifikante Wellenhöhen [16]

was darüber liegt, wird entsprechend stärker angezogen, es bildet sich also eine Delle aus. Umgekehrt bildet sich an den Stellen, wo das Magma sehr heiß ist, eine Beule aus. Das Geoid (Bildmittelpunkt auf den südlichen Atlantik gerichtet) ist in Abb. 3.37 dargestellt.

Ein häufig gezeigtes Ergebnis des SeaSat-Instruments ist eine indirekt bestimmte Größe. Die Abweichungen der Meereshöhe von einem Referenzgeoid lassen sich verwenden, um daraus Schlüsse über die Topographie des Meeresbodens zu ziehen, d.h. die Höhe der Wasseroberfläche wird in eine Höhenskala für den Meeresboden umgerechnet (sicherlich ein nicht triviales Verfahren, das auf verschiedenen Annahmen beruht). Damit ist eine Darstellung des Reliefs des Ozeanbodens (Sea Floor Mapping) und damit auch der gesamten Erdoberfläche (im Sinne von Erdkruste) möglich., vgl. Abb. 3.38. Diese Darstellung zeigte erstmals direkt die einzelnen Kontinentalplatten und die sich auf Grund der Kontinentaldrift auftuenden Gräben bzw. sich aufwerfenden Gebirgszüge.

Während die Laufzeit des Signals eine Information über die Höhe des Satelliten über Seeoberfläche gibt, gibt die Form des reflektierten Signal eine Information über die signifikante Wellenhöhe: das ursprünglich emittierte Signal entspricht einem Rechteckimpuls. Wird dieser von einer waagerechten, vollständig glatten Fläche reflektiert, so werden alle Bestandteile gleichzeitig reflektiert und das Signal kommt als scharfer Sprung an. Ist die Meeresoberfläche dagegen 'uneben', so trifft ein Teil des Signals auf niedriger liegende Teile als der Rest (der Durchmesser oder die Punktgröße des SeaSat-Mikrowellensignals auf der Wasserober-

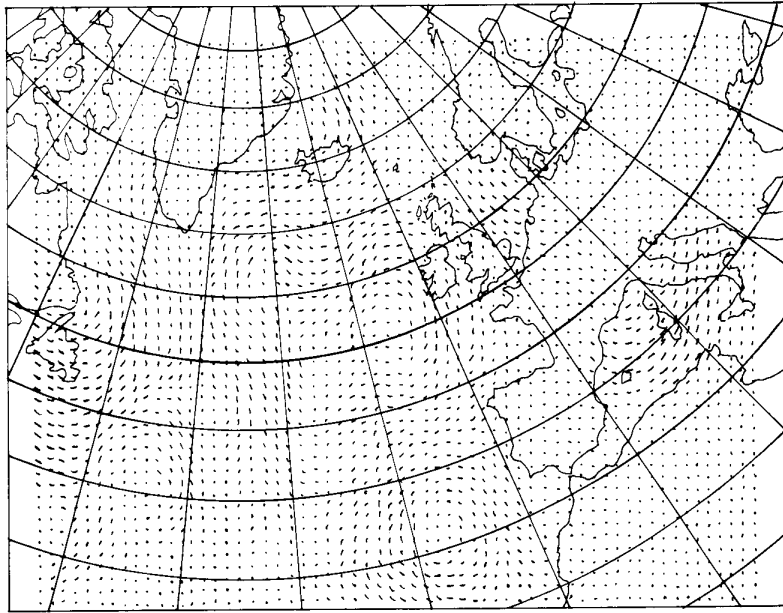


Abbildung 3.40: Windgeschwindigkeiten bestimmt mit dem Scatterometer von SeaSat [16]

fläche beträgt zwischen 1.6 km und 12 km), damit ist aber auch seine Laufzeit länger und der Rechteckimpuls wird zu einem kontinuierlichen Anstieg verwischt. Abbildung 3.39 zeigt die Form eines zurückgekehrten Impulses in Abhängigkeit von der signifikanten Wellenhöhe SWH. Zusätzlich zur Bestimmung der Meereshöhe ist auch eine Bestimmung der Windgeschwindigkeit (nur Betrag, nicht die Richtung) über der Wasseroberfläche möglich.

Das Scatterometer wird ähnlich dem Altimeter zur Untersuchung des Zustandes der See verwendet. Statt eines einzelnen Radarpulses wird ein Ensemble von vier Pulsen benutzt. Dadurch ist eine Bestimmung nicht nur des Betrages der Windgeschwindigkeit, sondern auch ihrer Richtung möglich. Abbildung 3.40 zeigt ein Beispiel für die mit Hilfe des Scatterometers bestimmten Windprofile. Details zum Instrument und zum Verfahren finden sich in [16].

3.4.2 SAR Synthetic Aperture Radar

Im Gegensatz zu Alti- und Scatterometer erzeugt das eigentliche *Radar* räumlich aufgelöste Bilder. Der Output ist damit ähnlich dem von optischen Kameras oder Scannern, jedoch hat Radar den Vorteil, Wolken zu durchdringen und auch nachts zu arbeiten. Das räumliche Auflösungsvermögen ist gegenüber dem passiver Mikrowelleninstrumente deutlich verbessert. Konventionelles Radar erfordert die Verwendung einer großen Antenne. Da diese auf einem Satelliten nicht mitgeführt werden kann, verwendet man ein Synthetic Aperture Radar (SAR): die große Antenne wird dabei aus der Bewegung einer kleinen Antenne durch die Eigenbewegung des Satelliten zusammengesetzt. Synthetic Aperture Radar wird nicht nur von Satelliten sondern ebenso von Flugzeugen aus verwendet. Abbildung 3.41 zeigt dazu am Beispiel eines Flugzeuges, wie durch die Bewegung der Antenne eine größere Array- oder Antennenlänge erzeugt wird.

SAR verwendet relativ kleine Antennen mit einem weiten Richtstrahl. Wie in Abb. 3.41 angedeutet, ist dieser Strahl zur Seite gerichtet. Dadurch werden vom bewegten Sensor vom gleichen Bodenobjekt eine Vielzahl von Radarechos aus verschiedenen Positionen empfangen. Sowohl die Amplitude als auch die Phase des zurückkommenden Signals werden für einen Zeitraum T gespeichert. Dieser Zeitraum T wird definiert als die Zeit, in der die beobachteten Objekte innerhalb eines bestimmten Abstandes R von der Antenne liegen. Die während dieser Zeit zurückgelegte Flugstrecke (in Abb. 3.41 AB) entspricht dann der Länge L des

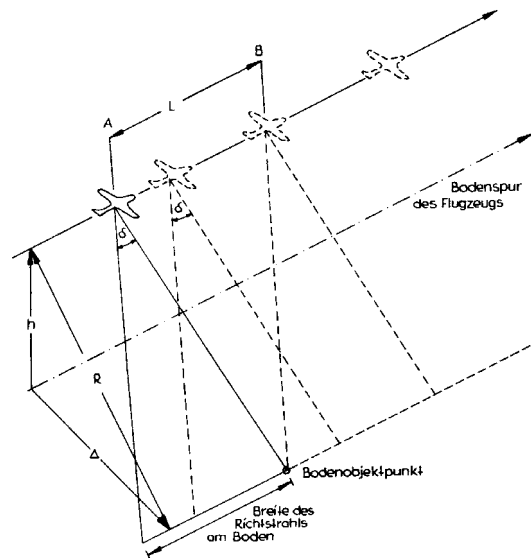


Abbildung 3.41: Veranschaulichung der Formierung der synthetischen Arraylänge eines SAR [12]

synthetischen Arrays. Mit l als der Länge der Antenne ist diese dann gegeben zu [12]

$$L \approx R \cdot \frac{\lambda}{l} . \quad (3.13)$$

Die aufgrund der Relativbewegung zwischen Antenne und Objekt verursachte Dopplerverschiebung des zurückkommenden Signals wird mit der Frequenz und Phasenlage eines stabilen Oszillators verglichen. Das Interferenzsignal kann entweder in Form eines Hologramms auf Film oder auf Magnetband aufgezeichnet werden, die Wiederherstellung des Bildes erfolgt mit Hilfe kohärenter optischer Prozessoren oder elektronischer Rechanlagen. Im so bearbeiteten Radarbild ergibt sich somit ein Objektpunkt aus der kohärenten Summierung einer Vielzahl individueller Radarechos, wobei sich positive und negative Dopplerverschiebungen aufheben. Beispiele für Radaraufnahmen finden sich z.B. in [101].

Abbildung 3.42 fasst einige Eigenschaften bei der Abbildung mit Radar zusammen: da das Aufnahmeverfahren auf einem zur Seite geneigten Strahl basiert, entstehen hinter erhöhten Geländemerkmale Schattenbereiche über die keine Information gewonnen werden können. Der Helligkeitswert des reflektierten Radarpulses hängt von der Geländestruktur ab: eckige geometrische Strukturen, also auch viele vom Menschen errichtete Strukturen, reflektieren die Puls stark, erzeugen also einen großen Helligkeitswert. An diffusen Oberflächen (z.B. Vegetation) wird ein Teil des Strahls auf den Sender zurück gestreut, es entstehen mittlere Helligkeitswerte. Glatte Oberflächen, z.B. Gewässer) reflektieren den Radarstrahl nicht in Richtung des Senders, diese Flächen erscheinen dunkel.

Die Charakteristika verschiedener auf Raumfahrzeugen eingesetzter SARs sind in Tabelle 3.5 zusammengefasst. Das Bodenaufklärungsvermögen liegt im Bereich von 25 bis 38 m, ist also mit dem der Multispektralkameras und -scanner ohne weiteres zu vergleichen. Allerdings fehlen die Spektralinformationen, die gerade das LandSat TM Instrument so interessant und wichtig machen. Damit ist zwar beim Radar der Vorteil der Unempfindlichkeit gegenüber Bewölkung gegeben, dieser muss aber gegen die Nachteile des fehlenden spektralen Auflösungsvermögens und des hohen Energiebedarfs abgewogen werden. Haupteinsatzgebiet von Radarsatelliten ist die Ozeanüberwachung. Diese erfolgt zum einem im militärischen Bereich (Aufenthalt und Bewegungsrichtung von Schiffen), zum anderen unter verschiedenen geophysikalischen oder meteorologischen Fragestellungen. Für die Fernerkundung der landbedeckten Erdoberfläche eignen sich die im optischen Bereich arbeitenden Instrumente eher.

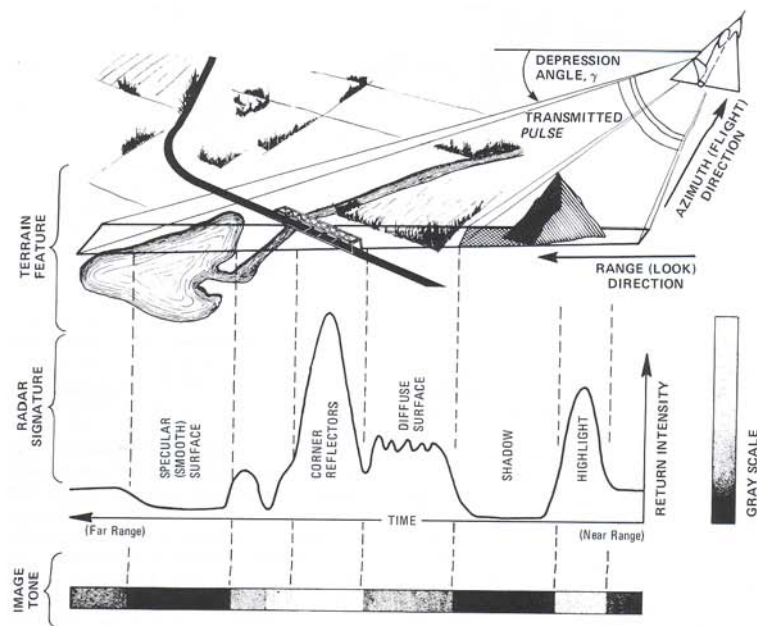


Abbildung 3.42: SAR: reflektiertes Signal in Abhängigkeit von den Geländeeigenschaften [101]

Aktive Mikrowelleninstrumente werden nicht nur zur Erkundung der Erde verwendet. Auf der Raumsonde Magellan z.B. sind mehrere aktive Mikrowelleninstrumente zur Kartierung der Venus eingesetzt. Ähnlich SeaSat ist Magellan mit verschiedenen Instrumenten ausgestattet: Hauptbestandteil ist das SAR, das eine globale Karte der Venus mit einer Bodenauflösung von 150 m erstellt hat. Dazu kommt ein Altimeter mit einer Auflösung von 30 m zur Bestimmung des Gravitationsfeldes und ein Radiometer mit einer Auflösung von 2 K. Diese Instrumente erlauben ein Verständnis der auf der Venus wirkenden geologischen Prozesse und damit auch eine Überprüfung unserer Vorstellungen über die Entwicklung eines Planeten. Insbesondere die Vergleiche zwischen verschiedenen erdähnlichen Planeten (z.B. Venus und Erde) sind hier von Bedeutung. Die Magellan-Ergebnisse sind in zwei Sonderbänden des JGR veröffentlicht (Journal of Geophysical Research 97, E8 und 97, E10) sowie in einer Folge von drei Artikeln in Sterne und Weltraum [53].

3.5 Sounder

3.5.1 Prinzip

Wir haben bisher im sichtbaren Bereich nur Imaging-Instrumente betrachtet, d.h. Instrumente, die uns das Äquivalent zu einer Photographie erzeugen - die zweidimensionale Darstellung eines Gebietes. Für einige Anwendungen, insbesondere zur Untersuchung der Atmosphäre, sind diese Bilder wenig hilfreich. Zur Messung von Spurengasen, Aerosol oder der Ozonkonzentration bedient man sich eines Sounders. Das generelle Merkmal eines Sounders ist, dass die Information über die Verteilung einer bestimmten Eigenschaft (z.B. Konzentration von CO_2) in Abhängigkeit von der Höhe bestimmt wird. Eine gute Übersicht über verschiedene Typen von Soundern geben Houghton et al. [47].

Sounder arbeiten nach drei verschiedenen Prinzipien: Okkultation (Limb-Viewing), Streuung und Emission. Okkultationsinstrumente sind dabei am einfachsten zu verstehen, ihr Prinzip ist in Abb. 3.43 dargestellt. Die Absorption des Lichts von der Sonne oder einem Stern lässt Rückschlüsse auf die Zusammensetzung der Atmosphäre in Abhängigkeit von der Höhe zu, in der das Licht die Atmosphäre durchsetzt hat. Dabei muss auf die Länge des in der Atmosphäre zurückgelegten Weges korrigiert werden. Vorteil der Methode sind die

Characteristics	Seasat (1978)	SIR-A (1981)	SIR-B (1984)
Orbit inclination	108°	38°	57°
Wavelength	L-band (23.5 cm)	L-band (23.5 cm)	L-band (23.5 cm)
Spatial resolution	25 m	38 m	25 m
Latitude coverage	72°N to 72°S	50°N to 35°S	58°N to 58°S
Altitude	790 km	250 km	225 km
Image-swath width	100 km	50 km	40 km
Depression angle, γ	67° to 73°	37° to 43°	30° to 75° (variable)
At average γ	70°	40°	52°
Smooth criterion	$h < 1.0$ cm	$h < 1.5$ cm	$h < 1.2$ cm
Rough criterion	$h > 5.7$ cm	$h > 8.3$ cm	$h > 6.8$ cm
Polarization	HH	HH	HH

Tabelle 3.5: Charakteristika der SeaSat und Shuttle SARs [101]

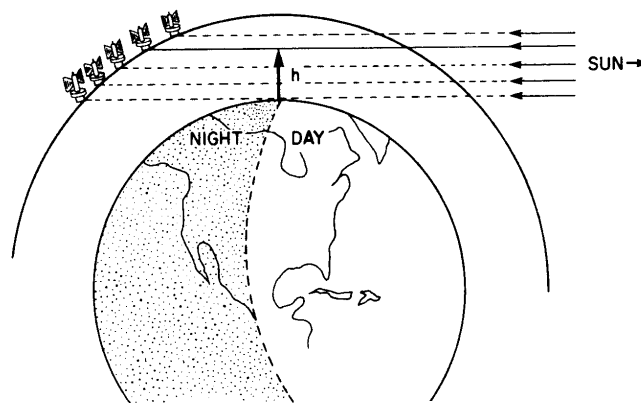


Abbildung 3.43: Sonnenuntergang aus der Sicht des Sounders SAM (Stratospheric Aerosol Measurement) auf Nimbus-7. Das Instrument beginnt, die Sonnenscheibe auf einer tangentialen Höhe h von ungefähr 350 km zu scannen und folgt der Sonne, bis sie untergegangen ist. Dadurch werden unterschiedliche Schichten der Atmosphäre nacheinander abgetastet [47]

langen Wege des Lichts durch die Atmosphäre, so dass sich selbst schwache Emissions- und Absorptionssignale nachweisen lassen. Außerdem ist die Höhenmessung im Gegensatz zu den vertikal arbeitenden Geräten direkt. Der Nachteil des Verfahrens ist die Erfordernis einer externen Lichtquelle (Sonne oder Stern), wodurch das Verfahren auf bestimmte Zeiten und Raumbereiche festgelegt ist.

Die Spurenstoffe oder Aerosole, die sich mit Hilfe eines Sounders untersuchen lassen, hängen von den verwendeten Wellenlängen ab. Das bereits erwähnte SAM (Stratospheric Aerosol Measurement) auf Nimbus-7 z.B. betrachtet das Sonnenlicht in einem Frequenzband, das um $1 \mu\text{m}$ zentriert ist. In diesem Bereich ist die Absorption durch atmosphärische Bestandteile vernachlässigbar. Jede Abschwächung des Sonnenlichtes muss daher durch Streuung an Aerosolen oder Rayleigh-Streuung an Molekülen entstanden sein. Die Messung von Ozonkonzentrationen erfolgt ebenfalls mit Limb-Viewing Soundern. Hier werden die Absorptionseigenschaften des Ozons im UV-Bereich ausgenutzt. Die Instrumente arbeiten meistens in mehreren engen Frequenzbereichen innerhalb des Wellenlängenbereiches von 0.25 bis $0.34 \mu\text{m}$.

	CO ₂ [ppm]	CH ₄ [ppm]	CFC-11 [ppt]	CFC-12 [ppt]	N ₂ O [ppb]
1750–1800	280	0.8	0	0	288
1990	353	1.72	280	484	310
jährl. Zunahme	1.8	0.015	9.5	17	0.8
Lebensdauer [a]	50–200	10	65	130	150

Tabelle 3.6: Klimawirksame Spurengase, deren Konzentration stark von menschlicher Aktivität beeinflusst wird

Sounder arbeiten aber nicht exklusiv im sichtbaren Licht. Einige Instrumente messen z.B. Spurengase durch deren Emission im Infrarotbereich. Hier ergibt sich die Information über die Höhe aus der Abhängigkeit der Emissionsbanden vom Druck, d.h. für diese Instrumente ist ein exzellentes spektrales Auflösungsvermögen erforderlich.

Eine andere Methode des Sounding beruht auf der Streuung. Die Hauptanwendung dieser Methode ist z.B. der Nachweis von Regen. Wenn ein Radarpuls durch ein Regengebiet geht, so werden die Radarechos proportional zur Rückstreuung der Impulse an den Regentropfen reflektiert, wobei tiefer in die Wolke eindringende Pulse später reflektiert werden. Auf diese Weise lässt sich ein vertikales Profil der Intensität des Niederschlages erzeugen. Eine weitere Anwendung von Streuinstrumenten ist die Messung der Ozonkonzentration mit Hilfe der Rückstreuung der solaren UV-Strahlung (TOMS BUV = Backscattered UV).

Treibhauswirksame Spurengase und Ozon

Eng mit dem Thema des Treibhauseffektes und einer möglichen globalen Erwärmung verknüpft sind die Spurengase, hier natürlich insbesondere die treibhauswirksamen Spurengase. Messnetze für Spurengase müssten im Prinzip ähnlich dem Netz der Wetterstationen installiert werden, allerdings stellt sich dort, genauso wie auch bei den Wetterstationen, das Problem, wie diese Messungen auch über den Ozeanen und in anderen schwer zugänglichen Bereichen erfolgen können. Die Verteilungen, Konzentrationen, Quellen und Senken sowie die Lebensdauer dieser Spurengase in der Atmosphäre sind einige der zu untersuchenden Fragen. Alle diese Parameter werden mit möglichst großer Genauigkeit benötigt, wenn man ein Klimamodell auf seine Realitätsnähe überprüfen will und auf der Basis dieses Modells Voraussagen über mögliche Klimaänderungen machen will.

Der stärkste Teil der anthropogenen Verschmutzung spielt sich in der Troposphäre ab. Die Troposphäre ist die unterste Atmosphärenschicht, die sich vom Boden bis in Höhen von, je nach geographischer Breite, 8 bis 15 km erstreckt. In dieser Schicht finden die Wettervorgänge der Atmosphäre statt, hier ist der vertikale und horizontale Austausch kräftig, auch findet ein Austausch mit Meer-, Land- und Eisflächen statt. Die Tropopause bildet die obere Grenze der Troposphäre. Sie ist durch eine deutliche Temperaturinversion markiert. Diese verhindert den Austausch von Gasen mit der darüberliegenden Stratosphäre nahezu vollständig. Lediglich im Bereich der Jetströme kann ein jahreszeitenabhängiger Austausch zwischen diesen Schichten stattfinden (zur Struktur, Chemie und Physik der Atmosphäre siehe z.B. [28, 48, 49, 98]). Allerdings kann z.B. bei Vulkanexplosionen die aufsteigende Asche- und Gaswolke so mächtig und heiß sein, dass sie bis in die Stratosphäre hinein ragt und dort einen Eintrag an Aerosolen und Schwefeloxid bewirkt. Diese Wolke kann sich global verteilen (z.B. Mt. Pinatobu) und baut sich nur langsam wieder ab.

Die wichtigsten der klimawirksamen Spurengase sind Kohlendioxid, Methan, verschiedene Flourchlorkohlenwasserstoffe und Stickoxide. Die Konzentrationen, Wachstumsraten und atmosphärischen Lebensdauern dieser Gase sind in Tabelle 3.6 zusammengefasst. Aber wie bestimmt man solche Daten? Wie weist man Spurengase nach? Die meisten diese Gase sind nicht sichtbar, d.h. eine Satellitenaufnahme im sichtbaren Bereich hilft bei ihrer Identifikation nicht. Spurengas-, ebenso wie Aerosolkonzentrationen werden mit Sounding-Instrumenten gemessen. In Abschn. 3.5.1 hatten wir das Verfahren kurz am Beispiel des SAM II (Stratospheric Aerosol Measurement) erläutert. SAM II ist ein Okkultationsinstrument, bei dem die

Atmosphäre in dünne horizontale Streifen zerlegt wird. Nachteil eines Okkultationsinstrumentes ist, dass es nur oberhalb einer Höhe von einigen Kilometern arbeiten kann, da bei geringerer Höhe z.B. Gebirge in den Lichtweg treten. Diese Instrumente eignen sich also eher für das Studium der Stratosphäre und der obersten Schichten der Troposphäre, weniger aber für den im Zusammenhang mit dem Spurengaseintrag interessanteren Bereich der unteren und mittleren Troposphäre. Für Messungen in der Troposphäre benötigt man daher Instrumente, die senkrecht oder schräg nach unten blicken. Die Messungen werden dadurch verkompliziert, dass die Stratosphäre vom Instrument ebenfalls gesehen wird und dass Wolken die Messungen stark beeinflussen können. Die Möglichkeit, troposphärische Aerosole und Spurengase von Satelliten aus zu messen, ist erst in den letzten Jahren erfolgreich demonstriert worden [39]. Eine Übersicht darüber, welche Instrumente sich für die Messung troposphärischer Aerosole und Spurengase eignen, gibt Levine [68]. Die meisten der Instrumente befinden sich noch in der Erprobungs- und Entwicklungsphase, so dass sie heutzutage nur auf dem Space-Shuttle Verwendung finden, aber noch nicht routinemäßig auf Satelliten eingesetzt sind. In der nächsten Satellitengeneration (Mission to Planet Earth) sollen diese Instrumente jedoch Standard werden, SCIAMACHY auf dem 2002 gestarteten EnviSat ist ein Beispiel.

MAPS (Measurement of Air Pollution from Satellites) ist ein Gasfilter-Korrelationsradiometer, das 1981 und 1984 auf Shuttle-Flügen eingesetzt wurde. MAPS misst die vom System Erde-Atmosphäre emittierte Strahlung. Diese Strahlung wird mit der Referenzstrahlung einer geschwärzten Aluminiumplatte verglichen, die nicht durch atmosphärische Einflüsse verändert ist. Drei Detektoren messen die Strahlung des Erde-Atmosphäre-Systems und der Referenzplatte, wobei die Strahlung vor jedem Detektor einen gasgefüllten Hohlraum durchsetzen muss, der es erlaubt, die Messungen jeweils auf einen bestimmten Höhenbereich in der Atmosphäre zu beziehen [82, 94]. Korrigiert man diese Daten gegen die Beeinflussung durch Kohlendioxid, Stickoxide, Ozon, Wasserdampf und Wolken, so ist dieses Instrument am empfindlichsten gegenüber Veränderungen in der Kohlenmonoxid-Konzentration im Höhenbereich zwischen 3 und 8 km, d.h. in der mittleren Troposphäre. Die Weiterentwicklung von MAPS ist TRACER (Tropospheric Radiometer for Atmospheric Chemistry and Environmental Research). TRACER wird in der Lage sein, sowohl Kohlenmonoxid als auch Methan zu messen. Messungen sollen möglich sein für die atmosphärische Grenzschicht in 0 bis 3 km Höhe und für die darüberliegende freie Troposphäre. Das Interesse am Kohlenmonoxid erklärt sich dadurch, dass sich CO als ein gutes Näherungsmaß für die troposphärische Ozonkonzentration verwenden lässt (d.h. das durch den Menschen zusätzlich verursachte Ozon in Bodennähe, das zu den sommerlichen Ozonwarnungen führt). Harris [39] zeigt Beispiele für die Kohlenmonoxid- und Ozonmessungen mit MAPS.

Eine direkte Messung der troposphärischen Ozonkonzentration ist kaum möglich, da der größte Teil des Gesamt ozons in der Stratosphäre liegt, d.h. jeder Satellit würde im wesentlichen das stratosphärische Ozon sehen. Allerdings gibt es auch eine indirekte Möglichkeit, das troposphärische Ozon zu bestimmen. TOMS (Total Ozone Mapping Spectrometer) misst seit 1979 täglich und global den Gesamt ozongehalt der Atmosphäre (d.h. die Summe aus dem troposphärischen und dem stratosphärischen Ozon). TOMS arbeitet nach dem Prinzip der zurückgestreuten UV-Strahlung (BUV, backscattered UV) und ist das Instrument, auf dessen Daten die Informationen über das antarktische Ozonloch beruhen, vgl. Abb. 1.3. Während in hohen Breiten die Ozonkonzentration sehr variabel ist (was sich ja auch im saisonalen Auftreten des Ozonloches zeigt), ist in den Tropen bis hin zu mittleren Breiten die stratosphärische Ozonkonzentration praktisch konstant, d.h. in mittleren und niedrigen Breiten geben die Schwankungen in der Gesamt ozonkonzentration gleichzeitig die Schwankungen im troposphärischen Ozon wieder.

Die Verwendung von Satellitendaten hat hier, wie auch in vielen anderen Bereichen der Erdfernerkundung wieder zwei Vorteile, (a) die kontinuierliche Messung und (b) die Messung auch über entlegenen Gegenden. Ob wir ohne Satellitenmessungen bereits soviel über das antarktische Ozonloch und insbesondere dessen saisonale Variabilität wüssten, ist zweifelhaft. Ohne kontinuierliche Messungen wäre es wahrscheinlich kaum möglich gewesen, diese saisonale Abhängigkeit zu entdecken. Eine Beschreibung der TOMS-Beobachtungen findet

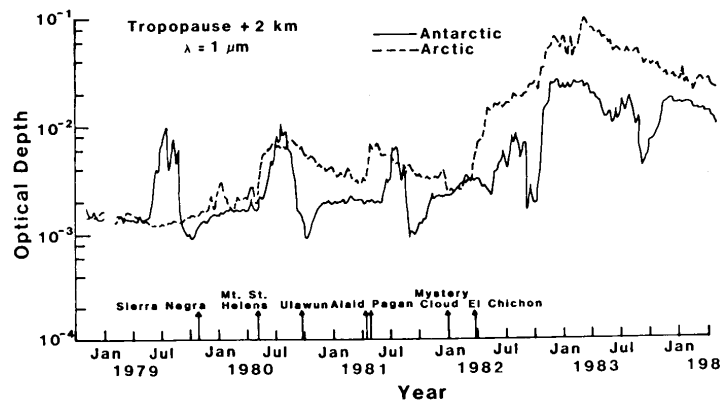


Abbildung 3.44: Stratosphärische Aerosol Messungen (SAM II) der optischen Tiefe als einem Maß für die Abschwächung der einfallenden solaren Strahlung in der Arktis und Antarktis [16]. Die Pfeile deuten größere Vulkanausbrüche an

sich in Schoeberl [106].

Auch zeigen die Beobachtungen der verschiedenen Spurengase, dass Luftverschmutzung von städtischen/industriellen ebenso wie aus landwirtschaftlichen Aktivitäten nicht nur ein lokales Problem ist, sondern dass sich die eingetragenen Schadstoffe bis in die entlegensten Bereiche der Erde ausgedehnt haben [28, 29, 99].

Vulkane

Die Warnung vor Vulkanausbrüchen ist z.B. auch für den Luftverkehr wichtig (aktuell hat Anfang Oktober 2004 gerade mal wieder der Mt. St. Helens ein Überflugverbot), da schon häufiger Flugzeuge, die in die Staubwolke eines Vulkans gerieten, durch den Ausfall eines oder mehrerer Triebwerke zur Notlandung gezwungen waren. Interessant sind diese Staubwolken auch aufgrund ihrer Fähigkeit, bis in die Stratosphäre vorzudringen. Hier bewirken sie eine Abschwächung der einfallenden solaren Strahlung. Außer den direkten Informationen über die Wolke aus den Imaging-Daten werden hier zum genaueren Studium, insbesondere auch zur Untersuchung der Abschwächung der einfallenden solaren Strahlung, Sounding-Daten verwendet. Abbildung 3.44 zeigt ein Beispiel für die Messungen des SAM II. Die optische Tiefe ist ein Maß für die Schwächung der einfallenden solaren Strahlung verwenden.⁸ Nach den größeren Vulkaneruptionen kommt es in der Arktis zu Erhöhungen der optischen Tiefe. Die jährlichen Peaks in der Antarktis dagegen sind nicht durch Vulkaneruptionen bedingt sondern entstehen im Zusammenhang mit den im antarktischen Winter geschlossenen Zirkulationssystemen. Während der Ausbruch des Mount St. Helens bei seinem Ausbruch 1980 nur ein Signal in der Arktis erzeugt (relativ hohe nördliche Breite des Vulkans!), hat sich die Staubwolke des El Chichon global über beide Hemisphären verteilt.

Klima und Wolken

Im Zusammenhang mit der Anwendung von Satelliten scheinen Wolken ein triviales Thema zu sein. Selbstverständlich kann man Satelliten zur Untersuchung von Wolken verwenden, das wird uns ja in jeder Wettervorhersage meist recht anschaulich anhand der MeteoSat-Bilder demonstriert. Und wo eine Wolke ist, da kann es auch regnen. Insgesamt klingt das sehr einfach. Bei einer genaueren Untersuchung geht es jedoch um das Verständnis der Einflüsse von Wolken auf unsere Umwelt, d.h. auf Menge und Eigenschaften des Niederschlages einerseits und auf die Bedeutung der Wolken für das Klima andererseits.

⁸Eine optische Tiefe von 1 bedeutet, dass die einfallende Strahlung in dieser Tiefe um einen Faktor $1/e$ abgeschwächt ist.

Die Bedeutung von Wolken als Lieferant für Niederschläge ist offensichtlich. Die wichtigsten Stichworte umfassen dabei die Bedeutung von Aerosolen als Kondensationskeime für Hydrometeore (Wassertröpfchen) und die Reinigung der Atmosphäre durch das Auswaschen von Aerosolen und eines Teiles der Spurengase. Dieses Auswaschen der Spurengase bewirkt gleichzeitig eine Veränderung der Chemie des Niederschlagswassers (Stichwort: saurer Regen). Wird bei der Auflösung der Wolke kein bis auf den Boden reichender Niederschlag erzeugt, so entfällt der Auswascheffekt der Hydrometeore, stattdessen werden Aerosole und Schadstoffe in der Atmosphäre vertikal ebenso wie horizontal umgelagert.

Wolken erzeugen aber nicht nur Niederschlag, sie tragen auch zu einer Modifikation des Strahlungstransports in der Erdatmosphäre bei und sind damit ein wichtiger Bestandteil in Klimamodellierungen. Ihr Einfluss auf den Strahlungshaushalt ist insbesondere im Zusammenhang mit dem anthropogen verstärkten Treibhauseffekt teilweise kontrovers diskutiert worden: eine Erwärmung des Erdbodens, z.B. als Folge eines Anstiegs des CO_2 -Gehalts, würde zu einer zusätzlichen Verdunstung und Wolkenbildung führen. Dabei würden sowohl der Wasserdampfgehalt der Atmosphäre als auch der Wolkenwassergehalt zunehmen. Welche Auswirkungen sich daraus für den Strahlungstransport ergeben, hängt davon ab, welcher Anteil des Wasserdampfes in den Flüssigwassergehalt von Wolken überführt wurde, von der Tröpfchengröße in diesen Wolken und von der Höhe der Wolken in der Atmosphäre. Nach momentanem Stand geht man davon aus, dass tiefliegende Wolken einen abkühlenden Effekt haben, da sie die globale Albedo erhöhen (d.h. mehr einfallende solare Strahlung in den Weltraum zurück reflektieren), dagegen nur eine relativ geringe zusätzliche Absorption der von der Erde emittierten Infrarotstrahlung bewirken. Hohe Wolken dagegen lassen die einfallende solare Strahlung ungestört passieren, absorbieren aber zusätzlich von der Erde emittierte Infrarotstrahlung und führen damit zu einer Erwärmung. Die Bedeutung von Wolken für Niederschlag und Strahlungstransport kann sicherlich nicht getrennt betrachtet werden. Zwischen beiden Phänomenen gibt es Verbindungen, z.B. greifen die Aerosole einerseits selbst in den Strahlungstransport ein, sind andererseits für die Bildung von Wolken unerlässlich, werden aber auch durch Niederschlag ausgewaschen.

Wollen wir die Bedeutung und Wechselwirkung von Wolken mit der Atmosphäre im Hinblick auf Niederschlagsbildung ebenso wie auf die Beeinflussung des Strahlungstransports in einem Modell beschreiben, so benötigen wir (a) eine gewisse Zahl an Messgrößen, die uns den Zustand der Atmosphäre in Abhängigkeit von der Zeit und von den dreidimensionalen Raumkoordinaten beschreiben, und (b) physikalische und chemische Gesetzmäßigkeiten, die die Wechselwirkungen der Größen untereinander beschreiben. Ein Teil der verknüpfenden Gesetze wird durch mikrophysikalischen Untersuchungen geliefert (Bildung von Hydrometeoren und deren Aufnahme von Spurengasen und Aerosolen). Einige der anderen Beobachtungsgrößen lassen sich durch Satellitenbeobachtungen bestimmen.

Wollen wir eine Wolke verstehen, so benötigen wir als Minimaldatensatz u.a. die folgenden Messgrößen:

- die Evapotranspiration des betrachteten Gebietes, d.h. ein Maß für die Fähigkeit der Erdoberfläche, Wasserdampf an die Atmosphäre abzugeben.
- den Feuchtigkeitsgehalt des betrachteten Gebietes als ein rückwirkendes Maß für über dem Gebiet niedergegangene Niederschläge (vgl. auch die Diskussion im Zusammenhang mit der Schneebedeckung in Abschn. 3.6.3) ebenso wie als ein Maß für die mögliche Verdunstung aus diesem Gebiet.
- den H_2O -Gehalt der Atmosphäre, der sich zusammensetzt aus dem Wasserdampfgehalt der Atmosphäre ebenso wie dem Flüssigwassergehalt der Wolken. Die Unterscheidung zwischen Wasserdampfgehalt und Flüssigwassergehalt ist wichtig, da beide unterschiedliche Einflüsse auf den Strahlungstransport haben.
- die mittlere Tröpfchengröße in den Wolken (oder besser eine Verteilung der Tröpfchengrößen), da die Tröpfchengröße Auswirkungen auf den Strahlungstransport hat. Entsprechende Informationen müssen auch für die feste Phase vorliegen, d.h. Schnee, Eiskristalle und Hagelkörner.

- allgemeine Informationen über die Atmosphäre wie Schichtung, vertikale Luftbewegungen, Druck, Temperatur, Dichte, Frontensysteme.
- der Gehalt an Aerosolen (inkl. Größenverteilung und Zusammensetzung) in der Atmosphäre, da diese als Kondensationskeime zur Tröpfchenbildung beitragen.
- der Gehalt an Spurengasen in der Atmosphäre, da diese selbst den Strahlungstransport beeinflussen können und in die Chemie der Hydrometeore eingreifen.
- die räumliche (horizontal und vertikal!) und (jahres)zeitliche Verteilung dieser Parameter.

Die Gesetze bzw. empirischen Regeln, die eine Verknüpfung dieser Meßgrößen im Wolkenmodell beschreiben sollen, umfassen u.a.

- die Kondensation von Hydrometeoren in Abhängigkeit von den Umgebungsbedingungen wie Übersättigung, Temperatur sowie Vorhandensein, Größenspektrum und Art von Aerosolen.
- das Wachstum von Hydrometeoren in Abhängigkeit von den Umgebungsbedingungen.
- die Umlagerung bzw. das Auswaschen von Spurengasen und Aerosolen durch diese Hydrometeore.
- die Chemie der gasförmigen und in Hydrometeoren gelösten bzw. an diesen angelagerten atmosphärischen Bestandteile sowie deren mögliche Rückwirkung auf die Wechselwirkung der Hydrometeore mit ihrer Umgebung, z.B. durch Veränderung des pH-Wertes eines Hydrometeors.
- die Dynamik innerhalb der Wolke, die sowohl die Bildung bzw. Auflösung von Hydrometeoren als auch deren Wechselwirkung mit der Umgebung bestimmt.
- die Dynamik der Atmosphäre insgesamt (vertikale Luftbewegung, Schichtung, erzwungene und freie Konvektion) und die sich daraus ergebende Wechselwirkung zwischen der Luftmasse in der Wolke und der umgebenden Atmosphäre mit ungesättigter Luft.
- die spektralen Absorptions-, Emissions- und Reflektionseigenschaften von Wasserdampf, Wassertröpfchen, Schnee, Eis und Hagel in Abhängigkeit von den Eigenschaften (insbesondere der Größe) der Hydrometeore.
- die spektralen Absorptionseigenschaften der anderen atmosphärischen Bestandteile, insbesondere der Aerosole und Spurengase.

Ein Teil dieser Gesetzmäßigkeiten (insbesondere die oberen vier) lassen sich durch mikrophysikalische Untersuchungen bestimmen, die anderen durch Beobachtungen in der Atmosphäre direkt bzw. indirekt durch Satelliten.

Im Zusammenhang mit der Modellierung von Wolken haben wir bereits einige wichtige Input-Parameter kennengelernt. Als letzter dieser Input-Parameter wurde die räumliche und zeitliche Verteilung der Größen erwähnt. Darin zeigt sich, dass eine in-situ Bestimmung der Parameter nicht mit vertretbarem Aufwand durchgeführt werden kann. Stattdessen erscheint es sinnvoll, auf Satellitenbeobachtungen zurückzugreifen. Zu nennen sind hier die üblichen Imaging Instrumente, die im sichtbaren Bereich des elektromagnetischen Spektrums arbeiten und im thermischen Infrarot.

Mit diesen Instrumenten lassen sich Wolkenbedeckungen ebenso wie Wolkenalbedos bestimmen. Durch die Verwendung des Infrarotkanals (bzw. eines im thermischen Infrarot arbeitenden Instruments) lässt sich die Temperatur der Wolkenobergrenze bestimmen und damit durch Vergleich mit der Standardatmosphäre oder besser noch durch Bezug auf ein Sounding Instrument oder einen geeigneten Radiosondenaufstieg die Höhe der Wolkenoberkante.

Damit erhalten wir Informationen über die Bereiche, in denen sich Wolken bilden, die Bewegung dieser Wolken in der Atmosphäre (dient auch der indirekten Messung von Windgeschwindigkeiten in größeren Höhen) und die Auflösung dieser Wolken durch Niederschlag oder Verdunstung. Diese Beobachtungsgrößen entsprechen also dem Output (bzw. einem Teil-Output) eines Wolkenmodells: sie zeigen uns Wolkenbildung und -auflösung und den Transport von Hydrometeoren und darin enthaltenen atmosphärischen Bestandteilen an.

Ein Output-Parameter, der in diesen Beobachtungen noch nicht enthalten ist, ist der Niederschlag. Will man diesen nicht auf der Basis des Netzes von Beobachtungsstationen

bestimmen (z.B. um eine feinere räumliche Auflösung zu erhalten, um auch sehr lokale Niederschlagsgebiete zu erfassen, oder mit dem Ziel der besseren Erfassung auch ausgedehnter (Wasser-)Flächen ohne dichtes Netz an Beobachtungsstationen), so bieten sich auch hier Möglichkeiten der Fernerkundung an: zum einen die Verwendung von Wetterradar, zum anderen die Verwendung von Satellitenbeobachtungen im Mikrowellenbereich.

Auch die Input-Parameter lassen sich größtenteils mit Hilfe von Satelliteninstrumenten bestimmen. Dazu werden Kombinationen einer Vielzahl von Imaging und Sounding Instrumenten benötigt, die in verschiedenen Bereichen des elektromagnetischen Spektrums messen. Als Beispiel für die Komplexität der Fragestellungen aber auch für die Möglichkeiten der Untersuchung, die sich durch eine geschickte Kombination von Instrumenten ergeben können, sei hier nur der Wasserdampfgehalt der Atmosphäre erwähnt. Die Strahlung im $6.7 \mu\text{m}$ Band ist stark durch die Absorption durch Wasserdampf und die Temperaturverteilung in der Atmosphäre bestimmt. Um daraus die Wasserdampfverteilung zu bestimmen, muss also gleichzeitig eine Temperaturverteilung bestimmt werden, z.B. durch einen Radiosondenaufstieg oder ein entsprechendes Sounding Instrument.

Weiter oben hatten wir Aerosole und Spurengase als weitere Input-Parameter kennen gelernt. Deren Vertikalprofile lassen sich ebenfalls mit Sounding Instrumenten bestimmen. Schwieriger dagegen wird es wieder bei der Bestimmung von mittleren Tröpfchengrößen in den Wolken. Hier ist noch zu überprüfen, ob geschickte Kombinationen vieler Beobachtungen (insbesondere die Verwendung der spektralen Reflektions- und Emissionsvermögen) es erlauben, zumindest grobe Mittelwerte für den oberen Bereich der Wolke anzugeben. Für die Bestimmung des gesamten Flüssigwassergehaltes in Wolken (also eine Summation über die einzelnen Tröpfchen) deutet sich die Möglichkeit einer Kombination von Daten im sichtbaren, infraroten und Mikrowellenbereich an. Das Verfahren ist bisher durch Kombination von Radiosonden- und Satellitendaten überprüft worden, sollte aber aufgrund der verbesserten Instrumentierung der neueren Satellitentypen auch mit Hilfe der Satellitendaten alleine durchgeführt werden können.

Ein derartiges Wolkenmodell sollte übrigens nicht alleine als Selbstzweck zur Beschreibung einer Wolke gesehen werden, sondern ist als ein Bestandteil eines wesentlich größeren und komplexeren Klimamodells zu sehen.

3.6 Anwendungsbeispiele

Bei den einzelnen Instrumenten wurden bereits Anwendungsbeispiele vorgestellt, in diesem Abschnitt finden Sie eine lose Sammlung von vielleicht nicht ganz so intuitiven Beispielen.

3.6.1 Archäologie und Anthropologie

Archäologie und Anthropologie sind beides eher bodenständige Disziplinen, die man nicht gerade mit der Forschung vom Weltraum aus in Verbindung bringen würde. Dieser Eindruck ist auch insofern völlig korrekt, als dass beide Disziplinen nicht mit Erdfernerkundung in dem Sinne betraut sind, dass ein wissenschaftliches Ziel definiert wird und dann ein Instrument zur Erreichung dieses Zieles entworfen und gebaut wird. Beide Disziplinen bedienen sich jedoch gelegentlich der Erdfernerkundung als eines willkommenen Hilfsmittels.

Die Archäologie bedient sich der Erdfernerkundung allerdings nicht erst seitdem diese durch Satelliten erfolgt. Auch normale Luftbilder wurden bereits im Dienste der Archäologie verwendet [17, 118]. Adams et al. [1] z.B. verwenden Radar um die Landnutzung durch die Mayas zu untersuchen. Häufig kann man auch am Wachstum von Pflanzen erkennen, ob sich unter dem Boden Ruinenreste befinden oder nicht: ist der Boden über einer Ruine (z.B. einem Mauerrest) nur relativ dünn, so unterscheidet er sich sowohl im Feuchtigkeits- als auch im Nährstoffgehalt von seiner Umgebung. Als Konsequenz wachsen hier die Pflanzen, insbesondere Nutzpflanzen, anders und unterscheiden sich in Größe und häufig auch Farbe von ihren benachbarten Artgenossen. Barret und Curtis [4] illustrieren dieses Phänomen unter

Verwendung von Luftbildern recht eindrucksvoll. Sabins [101] diskutiert in diesem Zusammenhang die Verwendung von Infrarotfarbfilmern (bzw. entsprechenden Daten von Scannern) zur Identifikation von Siedlungs- oder Infrastrukturen römischen Ursprunges. Auch die Muster der Landnutzung prähistorischer Indianer in Arizona sind auf diese Weise untersucht worden.

Strain und Engele [117] beschreiben als ein Beispiel für die Verwendung einer Kombination verschiedener Satelliteninstrumente in der Archäologie die Entdeckung von Ubar, einer verschwundenen Stadt, die T.E. Lawrence (Lawrence von Arabien) als das ‘Atlantis der Wüste’ beschrieb. Gradlinige Strukturen, die auf den LandSat- und SPOT-Aufnahmen schwach erkennbar waren, konnten in der freien Natur nicht mit irgendwelchen besonderen Strukturen am Boden in Verbindung gebracht werden. Die Kombination mit Space-Shuttle Radaraufnahmen machte jedoch deutlich, dass es sich bei diesen Linien um alte Karawanenpfade handeln musste: zwar zeigt sich keine Delle im Wüstensand, da dieser alle Rinnen gleichmäßig zuweht, aber die Radaraufnahmen zeigen ein Signal, da hier der Boden verfestigt ist und die Sandschicht etwas dicker ist als in der Umgebung. Die Karawanenpfade konvergierten in einem engen Bereich, in dem man bei Ausgrabungen die Ruinen der Stadt fand.

Radar wird auch verwendet, um trocken gefallene Flusstäler (Wadis) zu identifizieren. Diese sind in Nordafrika weit verbreitet und können als gute Hinweise bei der Suche nach alten Siedlungsplätzen dienen, da Ansiedlungen stets an Wasser gebunden sind. Da der Wüstensand extrem trocken ist, kann Radar die oberste Sandschicht durchdringen und darunter liegende Strukturen erfassen (das gleiche Prinzip hat auch im vorangegangenen Beispiel die Karawanenpfade deutlich hervortreten lassen). Damit lassen sich die ehemaligen Flussläufe, die weder in Aufnahmen im sichtbaren Bereich noch durch Untersuchungen vor Ort erkennbar sind, identifizieren. Strain und Engele [117] diskutieren in diesem Zusammenhang die Sanddünen im Tifernine Dünenfeld in Ostalgerien. In diesem Wüstengebiet lassen sich an einigen Stellen Hinweise auf ein prähistorisches Flusssystem finden. Dass das Klima an diesem Ort früher anders gewesen sein muss, belegen auch die etwa 6 000 Jahre alten Felszeichnungen im Tassili-N-Ajjer am Rande dieses Gebietes, die mit Jagszenen, Giraffen, Wildschafen, Flusspferden und Antilopen ebenfalls auf ein gemäßigteres Klima hindeuten.

3.6.2 Vegetation

Zur Untersuchung der Vegetation sind neben den Daten im sichtbaren Bereich des elektromagnetischen Spektrums insbesondere die Spektralkanäle im nahen Infrarot wichtig. Einige Beispiele für derartige Untersuchungen sind hier zusammen gestellt, die Grundlagen für eine Klassifikation werden im Zusammenhang mit der Datenverarbeitung in Abschn. 6.3 vorgestellt.

Grundlagen der Pflanzenbeurteilung, Vegetationsindex

Wir hatten im Zusammenhang mit der Diskussion der verschiedenen Satelliteninstrumente in Abschn. 3.2.5 bereits erwähnt, dass Beobachtungen im nahen Infrarot zur Bestimmung verschiedener Vegetationstypen und ihres Gesundheitszustandes verwendet werden können.

Abbildung 3.45 zeigt die Struktur eines Pflanzenblattes. Das vollständige Spektrum der solaren Strahlung fällt auf dieses Blatt. Die Bestandteile eines Blattes sind relativ transparent und absorbieren, mit Ausnahme der Pigmente im Chloroplast, keine nennenswerte Energie im sichtbaren Licht und nahen Infrarot. Das Blatt enthält viele Luft-Wasser-Übergänge zwischen den Zellwänden und dem interzellulären Raum. Der Unterschied im Brechungsindex des Wassers in den Zellwänden und der Luft in den Bereichen dazwischen bewirkt eine effiziente interne Reflektion des einfallenden Lichtes. Absorption findet nur in den Wellenlängen statt, die für die Photosynthese benötigt werden. Die wichtigsten Blattpigmente in diesem Bereich sind Chlorophyll a und b, Karotin und Xantophyll, die alle in der Nähe von $0.445 \mu\text{m}$ (blau) absorbieren. Zusätzlich absorbiert Chlorophyll auch in der Nähe von $0.645 \mu\text{m}$ (rot). Ein geringer Teil der absorbierten Energie wird als Fluoreszenzlicht emittiert, der größte Teil wird durch Photosynthese umgewandelt.

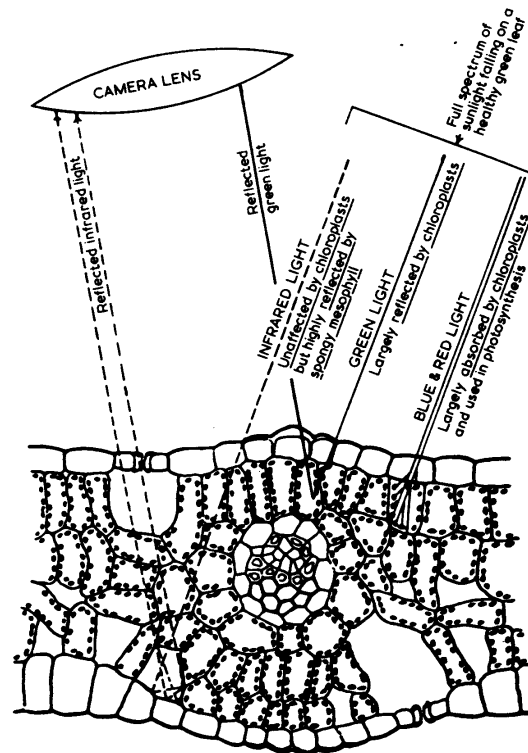


Abbildung 3.45: Struktur eines Pflanzenblattes und seine Reflektionsmerkmale [4]

Für eine gesunde Pflanze ergibt sich das in Abb. 3.46 durch die durchgezogene Kurve dargestellte spektrale Reflektionsvermögen. Bei einer kranken Pflanze ist die Photosyntheserate geringer, so dass die Absorption im roten und blauen Bereich des Spektrums geringer ist und dort mehr Strahlung reflektiert wird. Der Hauptunterschied zwischen einer gesunden und einer kranken Pflanze zeigt sich jedoch am besten im nahen Infrarot: während ein gesundes Blatt durch das Mesophyll an der Blattunterseite den Großteil der Strahlung reflektiert, wird bei einer kranken Pflanze das Mesophyll porös und die Reflektivität im nahen Infrarot nimmt ab.

Zur Beschreibung der Bioproduktivität einer Pflanze lässt sich ein 'normalized difference vegetation index' NDVI definieren. Dieser ist gegeben als das normierte Verhältnis $(IR - Rot)/(IR + Rot)$ [2, 109]. Dieser NDVI ist einer der wichtigsten Indizes zur Beschreibung der Vegetation geworden. Der NDVI wird vereinfachend als Vegetationsindex bezeichnet.

Waldschäden

Diese Eigenschaft von Vegetation, im nahen Infrarot ein helles Signal zu erzeugen, lässt sich auf verschiedene Weise einsetzen. Das einfachste Beispiel ist die Forstwirtschaft, bei der auf diese Weise vom Satelliten aus eine großflächige Bestandserfassung auch in nur schwer zugänglichen Gebieten erfolgen kann. Da diese Untersuchungen schnell wiederholt werden können, lässt sich ein Waldschadensbericht erstellen, der auch die Veränderungen berücksichtigt, d.h. Fragen untersucht wie: (a) wie schnell breitet sich ein Schaden aus, (b) wie breitet sich der Schaden aus (da gibt es sicherlich Unterschiede, ob ein Schaden in der Folge des sauren Regens im wahrsten Sinne des Wortes von außen über die Pflanzen kommt oder ob es eine Pflanzenkrankheit bzw. ein Schädlingsbefall ist, die/der sich von einem oder mehreren Herden her ausbreitet), (c) gibt es Bereiche, die relativ resistent gegen diesen Schaden erscheinen und wodurch zeichnen sich diese Bereiche aus (z.B. Bodenqualität, die einen Einfluss auf die Widerstandsfähigkeit der Pflanzen allgemein haben kann; geringerer Stress

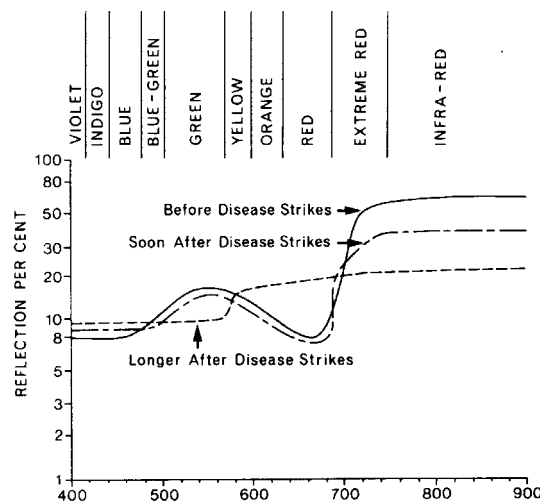


Abbildung 3.46: Charakteristisches spektrales Reflektionsvermögen gesunder und kranker Blätter [4]

auf die Pflanzen durch weniger exponierte Lage; bessere Nährstoffversorgung)? Ziel dabei kann es z.B. auch sein, bei einer Wiederaufforstung Areale mit unterschiedlichen Eigenschaften zu definieren und die jeweils für diese Umgebungsbedingungen am besten geeigneten Pflanzen dort anzupflanzen.

Globale Landbedeckung

Eine ganz andere Fragestellung verfolgt die mit den AVHRRs (Advanced Very High Resolution Radiometer) der NOAA-Satelliten angefertigte 'Inventur' der Vegetation unseres Planeten [18, 79, 122]. Insbesondere Änderungen in der Landbedeckung sind von Bedeutung; diese beeinflussen das Klima u.a. durch Veränderung der Albedo, Evapotranspiration und durch die Emission von treibhauswirksamen Spurengasen. Für diese Aufgabe werden die AVHRR-Instrumente verwendet, da diese mit einer relativ groben räumlichen Auflösung von 1.1 km gegenüber LandSat und SPOT Beobachtungen noch handhabbare Datenmengen erzeugen. Auch ist die Wiederholrfrequenz wesentlich größer, so dass sich saisonale Variationen in der Vegetation besser studieren lassen. LandSat und SPOT haben zwar eine Wiederholrfrequenz von ca. 2 Wochen, aber da kann natürlich leicht mal bei dem einen oder anderen Überflug das Zielgebiet unter einer Wolkendecke verborgen sein, so dass es erst nach mehreren Monaten wieder beobachtet werden kann.

Allerdings ist die Klassifikation der Vegetation nicht ganz so einfach wie man es sich vielleicht aus der obigen Faustformel für den Vegetationsindex wünschen würde [18]. Zum ersten ist der NDVI jeder einzelnen Pflanze abhängig u.a. von der Beleuchtung (Intensität, Einfallswinkel), der Nährstoff- und Wasserversorgung, der Schadstoffbelastung und der Qualität des Bodens. Zusätzlich muss zur Bewertung der reflektierten Strahlung außer dem Reflektionsvermögen des einzelnen Blattes bekannt sein die gesamte vom Satelliten einsehbare Blattoberfläche und die Neigung der einzelnen Blätter (z.B. durch Wind), da Blattober- und -unterseite unterschiedliche Reflektionsvermögen haben können. Dadurch wird die Trennung zwischen verschiedenen Pflanzenspezies weniger scharf, entsprechendes gilt auch für die Beurteilung ihres Gesundheitszustandes. Zusätzlich finden sich innerhalb eines Pixels verschiedene Pflanzenarten, es wird ein Teil des Bodens und, bei Vorhandensein, auch des Gesteins im Pixel erfasst (vgl. die Diskussion über Mischpixel im Zusammenhang mit Abb. 3.25). Dadurch ergeben sich Unterschiede in der Klassifikation der Vegetation, die von den Bearbeitungsverfahren der einzelnen Beobachter und ihren Modellannahmen abhängen. Zur Illustration, welche unterschiedlichen Resultate die aus den AVHRR-Aufnahmen nach verschiedenen Ver-

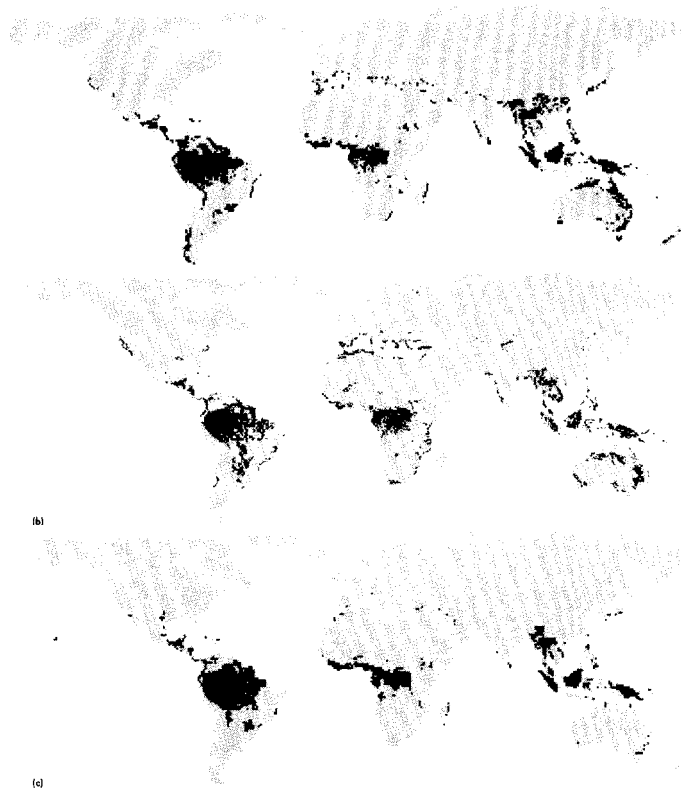


Abbildung 3.47: Globale Verteilung von 'breitblättrigen immergrünen Bäumen' bestimmt nach verschiedenen Verfahren aus AVHRR-Daten [18]

fahren bestimmten Verteilungen der Vegetation liefern können, zeigt Abb. 3.47 die globale Verteilung einer bestimmten Vegetationsart, in diesem Falle der 'breitblättrigen immergrünen Bäume', wie sie von drei verschiedenen Arbeitsgruppen bestimmt wurde. In allen drei Interpretationen gibt es bestimmte Kernbereiche, in denen diese Vegetation vorkommt, insbesondere global in den niedrigen Breiten (Regenwälder). Die Ränder dieser Bereiche und die Vorkommen dieser Pflanzengruppe außerhalb der Kerngebiete unterscheiden sich in den drei Bildern jedoch teilweise beträchtlich. Die starke Abweichung zu den Rändern hin erklärt sich zumindest teilweise aus der Tatsache, dass dort die Wahrscheinlichkeit für Mischpixel immer größer wird und dementsprechend die Identifikation der Vegetation ungenauer wird bzw. im Rahmen einer eher regionalen Studie mit Daten mit besserer räumlich Auflösung durchgeführt werden sollte.

Waldbrände

Ebenfalls einfach zu erkennen sind Wald- und Flächenbrände. Im sichtbaren Bereich sind sie in der Regel an ihrer Rauchwolke zu erkennen, die sich gut gegenüber der Umgebung abzeichnet. Noch besser zu identifizieren sind sie im thermischen Infrarot als deutliche Wärmeinseln. Hier ist einmal die Information über die Vernichtung von Biomasse und damit den Kohlendioxideintrag in die Atmosphäre interessant (insbesondere natürlich bei weit ausgedehnten Bränden in schwer zugänglichen Bereichen). Forstwirtschaftler und Biologen interessieren sich aber auch für die Zeit nach dem Waldbrand und studieren dazu das Verheilen der Brandnarben. Untersucht werden soll dabei die Selbstheilungskraft der Natur sowie die Charakteristika der Pionierpflanzen, die als erste ein zerstörtes Gebiet wiederbesiedeln. Wissen dieser Art wird z.B. im Rahmen des Fire-Managements in den nordamerikanischen Nationalparks benötigt. Fire-Management bedeutet dabei, dass man heutzutage bestrebt ist, natürlich ausbrechende

Feuer (Blitzschlag) nicht mehr zu löschen sondern brennen zu lassen. Die dahinterliegende Idee orientiert sich an der Tatsache, dass es auch früher Waldbrände gegeben hat, die zwar individuelle Pflanzen zerstört haben, die die Vegetation insgesamt aber überlebt hat, und dass diese Brände für die Bestandspflege wichtig sind: zu dichtes Unterholz wird ausgedünnt, die Alterspyramide der Bäume wird modifiziert (es gibt weniger alte und damit gegenüber Feuer und Schädlingen sehr empfindliche Bäume) und einige Pflanzen benötigen Feuer sogar, um sich zu vermehren: bei vielen Eukalyptusarten z.B. ist die Samenkapsel sehr hart und benötigt Feuer um aufzuplatzen – hier vernichtet ein Feuer zwar einen alten Bestand, legt aber gleichzeitig auch die Grundlage zur selbsttätigen Wiederaufforstung und damit auch zur Bestandserneuerung.

Versalzung

Ebenfalls recht einfach durch Betrachtung eines Kanals bzw. eines Falschfarbenbildes zu erkennen sind Gebiete starker Versalzung. Dabei gibt es natürliche Gebiete wie z.B. Salztäler in den Dünenfeldern der großen arabischen Wüste (Rub el Khali) oder in der Dasht-e-Kavir (große Salz- und Geröllwüste im Iran). Letztere ist besonders interessant, da sich hier durch zeitweise aus dem Grund hervortretendes Wasser neue Salzablagerungen bilden, die sich durch die Fernerkundung gut dokumentieren lassen. Auch die Parceval-Seen in Australien sind ein gutes Beispiel, an dem sich Versalzung studieren lässt.

Versalzung ist aber nicht nur ein natürlicher Vorgang sondern an vielen Stellen, gerade in den Trockengebieten der Erde, ein durch menschliche Eingriffe hervorgerufener oder beschleunigter Vorgang. Diese ‘anthropogene Versalzung’ entsteht durch Überbewirtschaftung landwirtschaftlich genutzter Flächen in Gebieten mit hoher Verdunstungsrate. Das Prinzip ist einfach: aufgrund der hohen Verdunstungsrate werden große Mengen an Wasser auf die Felder gepumpt, um überhaupt ein gewisses Maß an Bewässerung zu erreichen. Von diesem Wasser verdunstet aber der größte Teil, wobei sich die Mineralsalze auf der Oberfläche ablagern und so eine Salzsicht bilden. Diese Salzsicht ist der weiteren landwirtschaftlichen Nutzung der Fläche abträglich, da es kaum Pflanzen (und schon gar keine Nutzpflanzen) gibt, die gegen Salz resistent sind. Versalzung wird insbesondere dann zu einem Problem, wenn die landwirtschaftlich genutzten Flächen sich immer weiter ausdehnen bzw. nach einsetzender Versalzung verlagert werden. Jacobs [51] diskutiert als ein Beispiel die Versalzung in Folge künstlicher Bewässerung in den Randgebieten der Sahara

Eng mit dem Problem der künstlichen Bewässerung verbunden ist aber auch die Verkleinerung natürlicher Wasserreservoirs durch Wasserentnahme aus diesen bzw. aus ihren Zuflüssen. Das prominenteste Beispiel dürfte die Verkleinerung des Aralsee sein.

Bodenschätze

Ein Anwendungsbeispiel für Erdfernerkundung in den Geowissenschaften ist die Suche nach Bodenschätzen mit Hilfe von Pflanzen als Indikatoren. Hierfür werden die Informationen über die Geländeform und die Gesteinsarten verwendet. Diese Verfahren versagen jedoch, wenn man ein mit Vegetation bedecktes Gelände betrachtet: dann gibt es keine direkten Informationen mehr über das Oberflächengestein. Hier bedient man sich indirekter Verfahren, wie in Abb. 3.48 illustriert: befindet sich eine Pflanze über einer Erzader, im Beispiel Kupfer, so nimmt die Pflanze durch ihre Wurzeln einen Teil dieses Erzes auf (linkes Teilbild). Dadurch verändert sich aber das spektrale Reflektionsvermögen der Pflanzen (Beispiele im rechten Teil von Abb. 3.48). Diese Pflanzen können dann nach den gleichen Prinzipien, wie man es auch bei der Identifikation von Pflanzen- bzw. Bodenbedeckung allgemein macht, identifiziert werden. Die wichtigste Veränderung im spektralen Reflektionsvermögen ist die Blauverschiebung des steilen Anstieges beim Übergang zum nahen Infrarot. Diese Verschiebung scheint bei vielen Mineralien und verschiedenen Pflanzensorten aufzutreten und könnte daher als ein generelles Merkmal für verdeckte Erzlagerstätten dienen [101]. Allerdings bedarf es hier noch genauerer Untersuchungen.

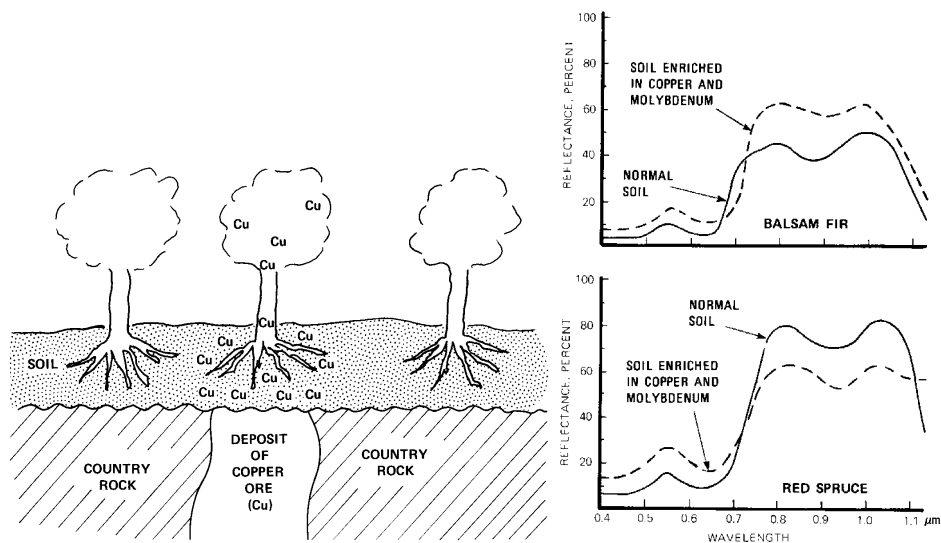


Abbildung 3.48: Übergang von unterliegenden Metallen in den Boden und die Pflanzen bei einer verdeckten Kupferader (oben) und Veränderung des spektralen Reflektionsvermögens dieser dotierten Pflanzen [101]

3.6.3 Hydrosphäre

Die Beobachtung der Ozeane erfolgt klassisch im thermischen Infrarot, vgl. Abb. 3.32. Beobachtungen in anderen Spektralbereichen geben jedoch auch über andere als nur die physikalischen Eigenschaften Auskunft, vgl. Abb. 3.52.

Schnee und Meereis

Daten über die Schneebedeckung werden in der Meteorologie benötigt zur Bestimmung der Art und Menge des Niederschlages, wobei insbesondere die Verfahren zur Bestimmung der Dicke der Schneedecke bzw. ihres Wasseräquivalents noch sehr in den Kinderschuhen stecken. Andere Interessenten an Daten über die Schneebedeckung sind die Klimatologen, die Straßenmeistereien (Räumdienste) und die Betreiber von hydro-elektrischen Anlagen. Insbesondere letztere interessiert natürlich das Wasseräquivalent der Schneedecke, da dieses nach der Schneeschmelze zur Stromerzeugung zur Verfügung steht.

Die Schneebedeckung wird übrigens nicht nur als Selbstzweck für alle Fragen im Zusammenhang mit Schnee benötigt. Zusätzlich wird Schneebedeckung auch benötigt, um den Vegetationsindex genauer zu bestimmen bzw. Verfälschungen dieser Größe durch Schneebedeckung zu erkennen [72].

Die Bestimmung der Meereisbedeckung ist ebenfalls für die Meteorologie von Interesse, da die globalen Wetterformationen (Lage der Polar- und Subtropenfronten und damit die Wanderung einzelner Tief- und Hochdruckgebiete) auch von der Temperatur des Untergrundes abhängen. Zusätzlich sind Informationen über Meereisbedeckungen natürlich auch für die Schifffahrt von Bedeutung. Die Bestimmung der Meereisbedeckung wird durch die verschiedenen Formen des Eises verkompliziert: so erzeugen 'schnelles Eis', Packeis, Eisberge, Eisschollen, sowie frisches und altes (in der korrekten Terminologie einjähriges und mehrjähriges) Eis unterschiedliche Signaturen. Zur Bestimmung der Eisbedeckung werden Aufnahmen im nahen Infrarot (z.B. LandSat MSS Band 7), im thermischen Infrarot sowie Radaraufnahmen verwendet. Alle diese Verfahren können im Prinzip als gleichwertig betrachtet werden, allerdings haben thermisches Infrarot und Radar den Vorteil, dass sie auch während der (Polar)Nacht eingesetzt werden können. Aufnahmen im thermischen Infrarot haben zusätzlich den Vorteil, dass sie eine Abschätzung der Dicke der Eisschicht erlauben: Wasser ist wärmer als Eis, erscheint also im Infrarotbild heller. Andererseits wirkt die Eisschicht aber auch als

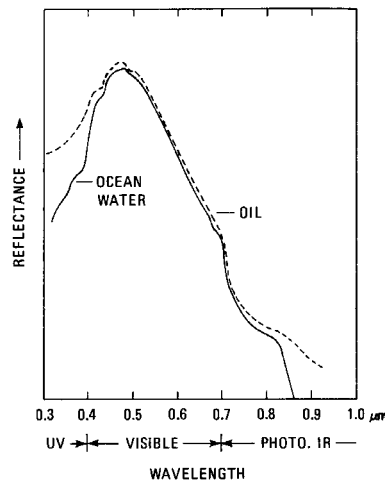


Abbildung 3.49: Spektrales Reflektionsvermögen von Seewasser und einer dünnen Ölschicht [101]

ein Isolator für das darunter befindliche Wasser, so dass zwischen Eis und Wasser meist ein deutlicher Temperaturunterschied besteht. Ist die Eisschicht nur sehr dünn, so lässt sie mehr Strahlungsenergie des wärmeren Wassers durch als eine dicke Eisschicht und erscheint damit heller.

Bei relativ hoher Wiederholfrequenz (z.B. tägliche Aufnahmen) lässt sich aus der Verlagerung charakteristischer Merkmale im Eis (Schollenform, eingefrorene Eisberge) auf die Driftrichtung und -geschwindigkeit der Eismasse zurückschließen. Dadurch lassen sich Eiswarnungen für naheliegende Schifffahrtsrouten geben.

Ölfilme auf Wasser

Hierbei geht es um die Identifikation von Ölverschmutzern und natürlichen Ölaustrittsgebieten. Natürliche Ölverschmutzer tragen ungefähr das fünffache der Menge zur Ölverschmutzung der Meere bei, die durch Ölbohrungen vor der Küste eingetragen wird, sind aber dennoch klein gegenüber der Ölverschmutzung, die durch defekte Pipelines entsteht [101]. Zur Identifikation eines Ölfilmes können UV, sichtbares Licht und nahes Infrarot ebenso wie thermisches Infrarot und Radar verwendet werden.

Abbildung 3.49 zeigt das spektrale Reflektionsvermögen von Seewasser und von einem dünnen Ölfilm für den Bereich vom Ultravioletten bis hin zum nahen Infrarot. Öl auf Wasser kommt in verschiedenen dicken Schichten vor: was wir von einer Pfütze oder einer nassen Straße als schillernden Ölfilm kennen, ist ein sehr dünner Film. Dieser kann sich als geschlossener Film nur auf relativ ruhigen Wasserflächen ausbilden, auf bewegttem Wasser zerreisst der Film sofort und bildet Inseln. Diese Filme sind auch relativ leicht im sichtbaren Licht zu erkennen. Im Falle eines Ölunfalls, d.h. des Austritts größerer Mengen Öls, kommt es nicht zur Bildung dünner Filme, vielmehr ist der der größte Teil der betroffenen Fläche von einer dicken, breiigen Ölschicht bedeckt, die bräunlich bis schwarz aussieht. In diesem Bereich ist die größte Menge des Öls konzentriert. Insbesondere diese dicken Schichten gilt es zu entdecken.

Relativ dünne Schichten lassen sich insbesondere im UV-Licht als helle Bereiche erkennen, da das Reflektionsvermögen einer dünnen Ölschicht für UV-Strahlung wesentlich höher ist als für Wasser (vgl. Abb. 3.49). In diesem Bereich ist der Unterschied des spektralen Reflektionsvermögens von Öl und Wasser am größten, d.h. im UV-Bereich sollten die Instrumente am leichtesten in der Lage sein, einen Ölfilm nachzuweisen. Allerdings wird die UV-Strahlung in der Atmosphäre stark gestreut und absorbiert, so dass eine Einsatz derartiger Instrumente auf Satelliten nicht sinnvoll ist. Mit einem in einer Höhe unterhalb 1 000 m fliegenden Flug-

zeug dagegen lassen sich im UV-Licht noch Ölfilme von einer Dicke von 0.15 mm erkennen. Allerdings haben auch Gischt und einige Algen sehr helle Signaturen im UV-Bereich, so dass hier andere Spektralbereiche ebenfalls überprüft werden sollten, um eine Verwechslungsgefahr zu vermeiden.

Im sichtbaren Bereich lassen sich zwei Effekte zur Entdeckung eines Ölfilms verwenden, zum einen das spektrale Reflektionsvermögen wie in Abb. 3.49 dargestellt, zum anderen die Eigenschaft des Ölfilms, kleine Wellen zu dämpfen, wodurch die Wasseroberfläche ruhiger wird. Diese ruhigen Wasserflächen reflektieren jedoch weniger Licht, so dass sie sich gegenüber dem sauberen umgebenden Wasser dunkel abzeichnen. Auf diese Weise lässt sich der Ölfilm vom Satelliten aus dann weniger direkt (d.h. über das veränderte spektrale Reflektionsvermögen) als vielmehr indirekt, nämlich über seine Auswirkungen auf die Feinstruktur der Wasseroberfläche, beschreiben. Mit Radar lassen sich derartige Bereiche ruhigeren Wassers übrigens auch nachweisen. Dieser indirekte Nachweis hat den Vorteil, dass er selbst bei dicken Ölschichten noch funktioniert.

Auch im thermischen Infrarot lassen sich Ölfilme erkennen. Das erscheint auf den ersten Blick eher überraschend, da ja das Öl die gleiche Temperatur haben sollte wie das Wasser, auf dem es schwimmt. Diese Betrachtung ist auch korrekt, zumindest dann, wenn man den Ölfilm durch Temperaturmessung mit einem Thermometer nachweisen wollte. Allerdings hängt die abgegebene Strahlung nicht nur von der Temperatur des Körpers ab, sondern auch von seinem Emissionsvermögen (vergleiche (3.10)). Das Emissionsvermögen von reinem Wasser beträgt 0.993. Befindet sich ein dünner Ölfilm auf dem Wasser, so beträgt das Emissionsvermögen nur noch 0.972, d.h. es wird insgesamt weniger Strahlung emittiert. Bestimmt man die Strahlungstemperatur (oder Effektivtemperatur) dann nach (3.11), so erhält man für das reine Wasser eine höhere Strahlungstemperatur als für den Ölfilm. Betrachten wir ein Beispiel: bei einer Wassertemperatur von 18°C würde sich für reines Wasser eine Strahlungstemperatur von 17.5°C ergeben, für den Ölfilm aber eine von 15.9°C, d.h. ein Unterschied von 1.6 K, d.h. selbst mit einem einfachen IR-Scanner zuverlässig nachgewiesen werden kann.

El Niño

Zur Untersuchung der Meere ist die Erdfernerkundung, ebenso wie bei der Untersuchung der Atmosphäre, schon alleine deshalb das Mittel der Wahl, weil man nicht, wie über Land, ein dichtes Netz von Beobachtungsstationen aufbauen kann. Beispiele für die Anwendung von Satelliten bei der Untersuchung der Ozeane hatten wir im Zusammenhang mit Ölverschmutzung und Meereisbedeckung bereits kennen gelernt.

Zu den elementarerer Fragen, die mit Hilfe von Satellitenmessungen beantwortet werden sollen, gehört die Bestimmung der Oberflächentemperatur. Abbildung 3.50 zeigt einen Überblick über die Oberflächentemperaturen der Ozeane für die Monate Januar (oben) und Juli (unten). Es bildet sich ein deutliches zonales System aus mit den höchsten Temperaturen um den Äquator und niedrigeren Temperaturen an den Polen. Diese extremen Bereiche zeigen kaum eine jahreszeitliche Variation, während in den mittleren Breiten eine deutliche jahreszeitliche Abhängigkeit der Oberflächentemperatur zu erkennen ist, die sich einerseits aus der veränderten Einstrahlung, andererseits aber auch durch infolge der veränderten Einstrahlung geänderte atmosphärische und ozeanische Zirkulationsmuster ergibt.

Während die in Abb. 3.50 gezeigte Temperaturverteilung fast trivial ist, sind es die auf mittleren räumlichen Skalen stattfindenden Phänomene, die ein größeres Interesse hervorrufen. Dazu gehören z.B. die Struktur des Golfstroms und El Niño, das Christkind. Der Golfstrom (vgl. Abb. 3.32) ist eine warme Meereströmung, die sich von der Saragossa-See über den Atlantik bis vor die Küste Norwegens erstreckt und für das Klima in Mittel- und Nordeuropa von regulierender Bedeutung ist. Die Existenz des Golfstromes als warme Strömung (und damit Heizung für unseren Bereich) ist lange bekannt, aber erst Satellitenaufnahmen haben einige Eigenarten des Golfstromes gezeigt. So fließt der Golfstrom nicht wie ein ruhiges Band dahin, sondern er bildet Verwirbelungen und Mäander. Manchmal können sich solche Wirbel ablösen und als warme Wasserringe in Gebiete nördlich oder südlich des Golfstromes vorstoßen. Umgekehrt kann durch solche Wirbel auch kaltes Wasser aus den umgebenden Be-

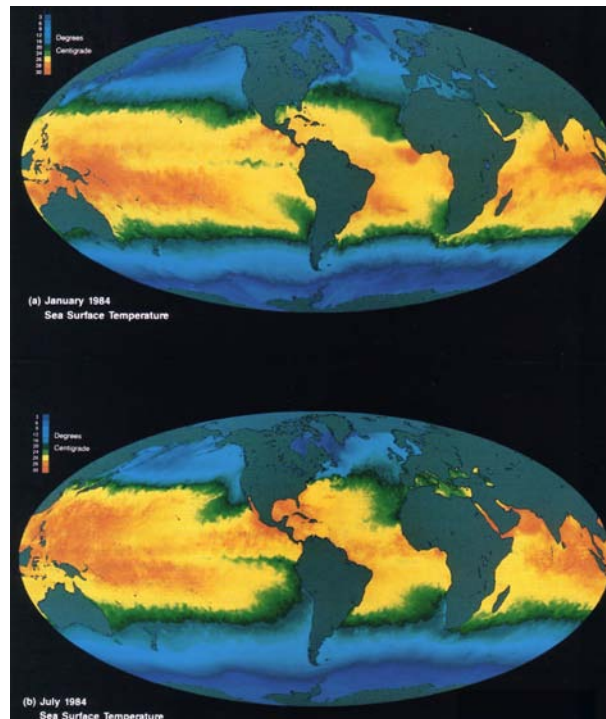


Abbildung 3.50: Mittlere Oberflächentemperatur der Ozeane für Januar (oben) und Juli (unten). Während sich in den Tropen und den Polregionen kein Jahresgang der Temperatur findet, ist in den mittleren Breiten eine deutliche Jahreszeitenabhängigkeit zu erkennen [83]

reichen in den Golfstrom eingezogen werden und dort einen Kaltwasserring bilden. Inwieweit diese Wirbel Rückwirkungen auf Biologie und Chemie der Ozeane sowie auf atmosphärische Vorgänge haben, bleibt noch zu klären [108].

Während der Golfstrom ein dauerhaftes Merkmal der Zirkulation im Atlantik ist, gibt es auch Zirkulationsanomalien, die nach heutigen Kenntnissen nicht vorhersagbar sind. Ein Beispiel ist El Niño: vor der peruanischen Küste verschwinden gelegentlich zu Weihnachten die Fische, da sich um diese Zeit das Meerwasser plötzlich und stark erwärmt und damit für die Fische keine ausreichende Menge an Nährstoffen zur Verfügung steht. Aufgrund seines Auftretens um die Weihnachtszeit herum wird dieses Phänomen als El Niño bezeichnet - das Christkind. Diese Periode warmen Oberflächenwassers dauert in der Regel bis in den März an, danach normalisiert sich die Temperaturverteilung wieder. In einigen Fällen dauert El Niño jedoch länger und kann sich sogar über Jahre erstrecken. Damit ergeben sich entsprechend katastrophale Folgen für die lokale Ökonomie. Diese langandauernden El Niños treten nur alle paar Jahre auf, bisher gibt es keine Vorhersagemöglichkeiten dafür. Die Temperaturanomalie des warmen Oberflächenwassers und vor allen Dingen ihre räumliche Ausdehnung und zeitliche Entwicklung lassen sich mit den Infrarotsensoren von Satelliten sehr gut erfassen.

Den Schlüssel zum Verständnis dieses Phänomens bildet die Southern Oscillation (Walker-Zirkulation): diese beschreibt die bisher nicht vollständig verstandene Kopplung des Luftdrucksystems im Ostpazifik mit dem des Westpazifiks. Die Passatwinde als Bestandteil der Hadley-Zirkulation werden durch die beiden beständigen Hochdruckgebiete im Pazifik - über den Osterinseln im Süden und vor Kalifornien im Norden - umgelenkt, so dass sie vor der amerikanischen Pazifikküste in Richtung auf den Äquator und damit auf die innertropische Konvergenzzone zu wehen. Der Südostpassat treibt den Humboldtstrom und schiebt damit Oberflächenwasser von der Küste weg. Dadurch kann kälteres, nährstoffreicheres Wasser aus der Tiefe aufquellen. Dieses zeigt sich in den Satellitenaufnahmen als ein kühler Bereich vor der südamerikanischen Pazifikküste (vgl. unteres Teilbild in Abb. 3.51). Das weggetriebe-

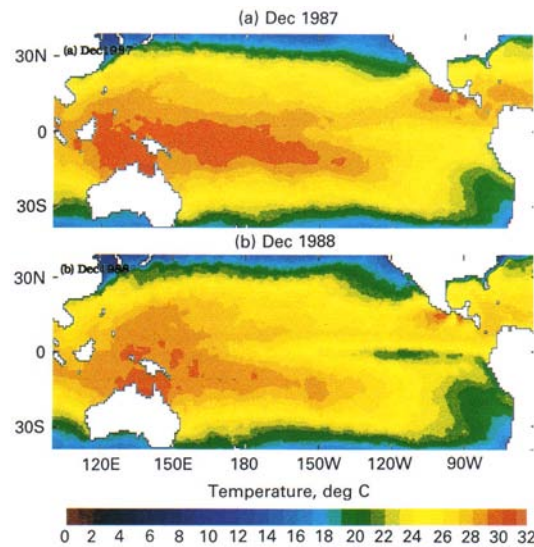


Abbildung 3.51: Oberflächentemperaturen im äquatorialen Pazifik während (Dezember 1987) und nach (Dezember 1988) einem El Niño [83]

ne warme Wasser dagegen staut sich im westlichen Pazifik. Die hohen Verdunstungsraten aus dieser Warmwassermasse führen zu starken Regenfällen in Indonesien und Südostasien. Zusammengefasst bedeutet also ein kalter Ostpazifik gleichzeitig starke Regenfälle in Indonesien. Diese Kopplung zwischen zwei Ereignissen in weit auseinanderliegenden Bereichen der Erde wird als Telekonnektion bezeichnet.

In der Atmosphäre besteht entsprechend der Wassertemperaturen eine negative Druckanomalie über Indonesien, die mit einer positiven Druckanomalie im pazifischen Hochdruckgebiet gekoppelt ist. Dabei entsteht eine Zirkulationszelle, die entlang eines Breitenkreises, und nicht, wie bei der großskaligen Zirkulation sonst üblich, entlang eines Längengrades ausgerichtet ist. Diese Zelle wird von der normalen Hadley-Zirkulation überlagert. Im Nettoeffekt bewirkt dann der Auftrieb kalten Äquatorwassers vor der amerikanischen Pazifikküste eine Abkühlung und damit eine Schwächung der innertropischen Konvergenzzone. Damit geht aber auch dem Südost-Passat ein Teil seines Antriebs verloren, er wird schwächer. Ein schwächerer Südost-Passat kann aber nicht mehr soviel warmes Wasser von der peruanischen Küste wegtreiben, die innertropische Konvergenzzone wird wieder gestärkt, damit auch der Südost-Passat und der Kreislauf kann erneut beginnen.

Das El Niño Phänomen ist mit dem Zusammenbruch des Passatwindsystems im Westpazifik und einem Umspringen dieser Passate in bodennahe Westwinde gekoppelt. Dann strömt das in den Westpazifik getriebene warme Wasser wieder in Richtung auf die amerikanische Ostküste zurück und blockiert dort den Auftrieb kalten Tiefenwassers, der eigentliche El Niño entsteht, vgl. Abb. 3.51 oben. Durch die Westwinde kommt es zu einem Aufsteigen feuchter Luft an der amerikanischen Ostküste, die Feuchtigkeit kondensiert und es kommt in normalerweise trockenen Gebieten zu starken Regenfällen. Die dann trockene Luft strömt in großen Höhen wieder Richtung Westen zurück, so dass es über Indonesien zur Ausbildung einer ungewöhnlichen Trockenheit kommt.

Die Auswirkungen von El Niño betreffen jedoch anscheinend nicht nur die unmittelbaren Bereiche dieser Walker-Zirkulation. Die Warmwasseranomalie in den Tropen kann die Zirkulation bis in die mittleren Breiten beeinflussen. In Zusammenhang mit dem Auftreten von El Niños könnte das Auftreten von Überschwemmungen in Kalifornien, Dürren in Afrika und Indien sowie starker Trockenheit in der ex-UdSSR und anderen Teilen der Erde stehen (vgl. z.B. Keppler, 1988). Bei einem weiteren Studium dieses Phänomens wird sich die Verwendung von Satelliten als hilfreich erweisen, da dadurch viele der zu beobachtenden Parameter

(Oberflächentemperaturen, Windgeschwindigkeiten, Wolken, Luftfeuchtigkeit, Niederschlag) an allen Orten standardisiert bestimmt werden können, so dass sich die Beeinflussung entfernter Regionen durch El Niño besser nachweisen lässt.

Die Farbe des Meeres

Das Schlagwort 'die Farbe des Meeres' bezieht sich eigentlich weniger auf die bisher genannten Untersuchungen, auch wenn die Aufnahmen im thermischen Infrarot normalerweise als eindrucksvolle farbige Aufnahmen dargestellt werden. Bei 'die Farbe des Meeres' handelt es sich um Untersuchungen des Planktons, d.h. der elementarsten Form von Leben im Meer, deren Verteilung gleichzeitig auch die Möglichkeiten der höheren Meeresbewohner, von den kleinsten Garnelen bis zu den größten Walen, bestimmt, da Plankton den Anfang der Nahrungskette bildet. Wichtigstes Instrument ist hier der Coastal Zone Color Scanner CZCS auf Nimbus-7. Der Name des Instrumentes ist irreführend, da sich die Beobachtungen keinesfalls auf die Küstenregionen beschränkt, sondern die Meere insgesamt umfasst. Und die Bezeichnung 'Color', die zu dem Schlagwort von der 'Farbe des Meeres' geführt hat? Beim CZCS handelt es sich um einen Multispektralscanner, der dem MSS auf LandSat und den AVHRRs auf den NOAA-Satelliten ähnlich ist. Im Zusammenhang mit der Identifikation von Vegetation hatten wir bereits auf die Verwendung verschiedener Kanäle zur Identifikation des spektralen Reflektionsvermögens hingewiesen. Dabei zeigte sich, dass die Verwendung breiter Kanäle die Unterscheidung erschwert, da über verschiedenen Spektralbereiche mit unterschiedlichem Reflektionsvermögen hinweggemittelt wird. Daher ist für Studien der Vegetation ein besseres spektrales Auflösungsvermögen anzustreben. Da beim CZCS die Zahl der Spektralkanäle ebenso wie bei den anderen Instrumenten begrenzt ist, hat man sich dadurch beholfen, dass man für die Identifikation von Plankton wichtige enge Spektralbereiche ausgeschnitten hat. Damit lässt sich dann zwar kein Bild in den realen Farben mehr rekonstruieren, wie es z.B. bei den LandSat Aufnahmen möglich ist, aber hieraus ergibt sich keine Einschränkung für das Instrument, da es nicht wie die LandSat-Instrumente als Universalerdkundungsinstrument konzipiert ist, sondern bereits als zweckgebundenes Instrument. Die typischen Spektralkanäle von LandSat- und NOAA-Instrumenten sind in Tabelle 3.1 aufgelistet, der CZCS dagegen misst in 6 Bändern von 0.433 - 0.453 μm , 0.51 - 0.53 μm , 0.54 - 0.56 μm , 0.66 - 0.68 μm , 0.6 - 0.8 μm und 10.5 - 12.5 μm . Das Bodenaufklärungsvermögen beträgt 825 m, was aber in Anbetracht der zu untersuchenden Flächen schon fast zu gut ist (Problem der Datenflut im Vergleich zu den Skalen, auf denen sich räumliche Veränderungen ergeben). Mit diesen Kanälen lassen sich die Chlorophyll-Pigmente des Planktons sehr gut und eindeutig nachweisen.

Die CZCS-Ergebnisse haben im wesentlichen die Erkenntnisse bestätigt, die man bereits seit einem Jahrhundert Meereskunde kannte: Plankton ist in den flachen Küstenbereichen wesentlich stärker vorhanden als im freien Ozean. Die neuen Ergebnisse betrafen dagegen zum einen die prinzipielle Möglichkeit, die Gewässerbereiche an Arktis und Antarktis zu studieren und die jahreszeitlichen Veränderungen zu untersuchen. So zeigte sich z.B. dass in diesen beiden sehr kalten Bereichen die Algenblüte im Frühjahr in ihrer Intensität und damit auch Biomasseproduktivität ohne weiteres mit der in den küstennahen Gewässern gemäßiger Breiten vergleichbar ist, vgl. Abb. 3.52. Ein deutlicher Unterschied zwischen Land und Meer zeigt sich in den Tropen: während hier die Bioproduktivität an Land sehr groß ist (zumindest, wenn der Mensch ihr freien Lauf ließe und nicht durch Brandrodung zur Zeit eine Abnahme der Biomasse bewirken würde), sind die tropischen Meere die Bereiche geringster Bioproduktivität, sie werden daher auch als die Wüsten der Meere bezeichnet (daher auch die wirtschaftlichen Einbrüche durch El Niño).

Geowissenschaften

Die Geowissenschaften bedienen sich des vollständigen Spektrums der Erdfernerkundungsinstrumente. Im sichtbaren Bereich sind Satellitenaufnahmen schon alleine deshalb von Bedeutung, will sie einen guten Überblick über weite Bereiche erlauben. Dadurch lassen sich die

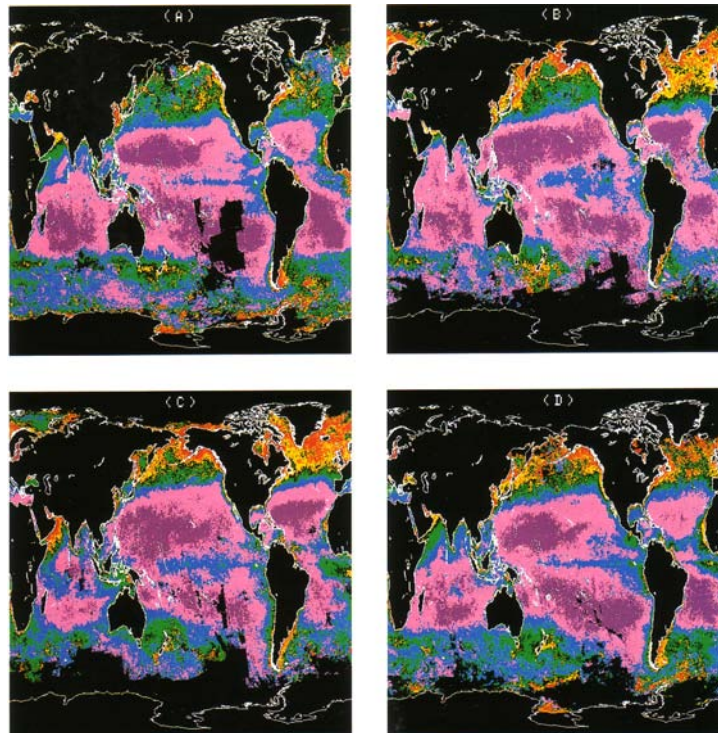


Abbildung 3.52: Die Farbe des Meeres: Pigmentkonzentration (Farbskala: violett – blau – grün – gelb – orange – rot) im Winter (A), Frühjahr (B), Sommer (C) und Herbst (D) [78]

großen Strukturen wie die Beziehungen zwischen verschiedenen Landformen, Deformationsmuster und Muster von Entwässerungsnetzen erkennen. Besonders wertvoll haben sich die LandSat-Aufnahmen bei der Identifikation von Falten und Bruchlinien erwiesen: Aufgrund lokaler Unterschiede (Vegetation, Beleuchtung) lassen sich diese Strukturen auf Luftbildaufnahmen (auch auf Mosaiken aus Luftbildaufnahmen) häufig nicht eindeutig identifizieren. Auf den meisten der Satellitenbilder zeichnen sich jedoch solche langen Linienstrukturen deutlich ab. Die Identifikation dieser Linien ist auch von allgemeinem Interesse, da sie häufig die großen Spalten- und Faltensysteme repräsentieren, die für die Ausbildung von Erdbeben von Bedeutung sind.

Neben der großräumigen Struktur ist aber auch die Beschaffenheit des Gesteins von Interesse. Hier werden neben Aufnahmen im sichtbaren Bereich insbesondere solche im thermischen Infrarot verwendet, da hier verschiedene Gesteine unterschiedliche Signaturen in ihrem spektralen Emissionsvermögen aufweisen (vgl. Abb. 3.28). Instrumente im thermischen Infrarot erlauben auch den Nachweis unterirdischer Kohlenfeuer, die durch Selbstentzündung entstehen und, wenn sie sich in der Nähe eines Abbaugebietes befinden, zum Eindringen giftiger Gase in die Grube führen können.

Radar wird ebenfalls in den Geowissenschaften verwendet. Zum einen erlaubt Radar, im Gegensatz zu anderen Instrumenten, die Bestimmung von Höhen, d.h. es lässt sich eine wirklich drei-dimensionale Topographie erstellen (vergleiche das SeaSat Altimeter, wie in Abschn. 3.4.1 beschrieben). Radar-Imaging wird ebenfalls verwendet, da Radar in trockenen Gebieten die Sandschicht durchdringen kann und somit Informationen über den darunterliegenden festen Grund liefert.

Eine weitere wichtige Anwendung von Satelliteninstrumenten in den Geowissenschaften ist die Beobachtung von Vulkanen. Verwendet werden hier Instrumente im sichtbaren Bereich und nahen Infrarot zur Beobachtung der Lavaströme und der Rauchwolken, sowie Instrumente im thermischen Infrarot zur Beobachtung des Vulkanmassivs direkt und zur Abschätzung des Austrittszeitpunktes der Lava (je jünger die Lavaschicht ist, um so wärmer ist sie noch,

d.h. die Temperatur erlaubt Rückschlüsse auf das Alter). Vulkanbeobachtungen werden u.a. durchgeführt zum Zwecke der Warnung der Anlieger vor einem zu erwartenden Ausbruch. Da die lavahaltige Erde an Vulkanhängen sehr fruchtbar ist, sind gerade Gebiete an Vulkanen häufig relativ dicht besiedelt. Dass die Erde dort fruchtbar ist, kann man auch in Falschfarbenaufnahmen erkennen, wo die vulkanischen Berge ein wesentlich stärkeres Signal im nahen Infrarot erzeugen als benachbarte, nicht-aktive Berge.

Literatur

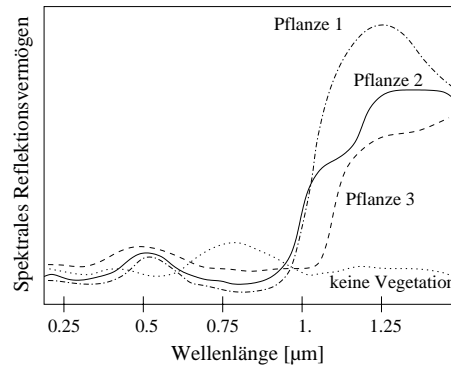
Einen nahezu vollständigen Überblick über Instrumente zur Erdfernerkundung bietet [62].

Die wesentlichen Grundlagen zur Anwendung und Funktionsweise von Sensoren auf Satelliten finden sich in Rees [93], Sabins [101], Cracknell und Hayes [16], Curran [17] und Löffler [71]. Kitchin [59] beschreibt Sensoren allgemein, dort ist insbesondere der Abschnitt über die photographische Emulsion gut. Die Instrumente der nächsten Generation sind kurz in Baker [3] beschrieben, ansonsten in verschiedenen NASA-Publikationen unter dem Obertitel 'Earth System Science'. Einige Instrumente oder Instrumenttypen sind auch in Rycroft [100] beschrieben.

Fragen

1. Welche der folgenden Satelliteninstrumente können Sie zur Begutachtung von Vegetation verwenden:
 - konventioneller Schwarz-Weiß-Film
 - konventioneller Farbfilm
 - Multispektralscanner im sichtbaren Bereich
 - IR-Farbfilm
 - Thermischen IR-Scanner
 - Multispektralkamera (konventionelle Filme)
 - Multispektralscanner mit Kanälen im nahen IR
 - IR-Schwarz-Weiß-Film
 - Mikrowellenradiometer
2. In der Frühzeit der (bemannten) Erdfernerkundung wurde mit Multispektralkameras gearbeitet. Erklären Sie das Prinzip. Was sind Vor- oder Nachteile gegenüber einer Aufnahme mit einem Farbfilm?
3. Erläutern Sie den Unterschied zwischen einem konventionellen und einem IR-Farbfilm. Wozu werden IR-Farbfilme verwendet? Wie werden die Informationen dargestellt (IR ist nicht sichtbar!)?
4. Erläutern Sie, warum Multispektralinformationen für die Erdfernerkundung interessant sind.
5. Skizzieren Sie den Aufbau eines photographischen Systems, erläutern Sie den Begriff des Bodenaufklärungsvermögens und geben sie eine Beziehung zwischen Bodenaufklärungsvermögen, Brennweite, Flughöhe und dem Auflösungsvermögen des Films.
6. Beim Farb-IR-Film verzichtet man auf eine der konventionellen Farbschichten und verwendet stattdessen eine Schicht, die IR-empfindlich ist. Auf welche Schicht verzichtet man und warum?
7. Welche der folgenden Parameter beeinflussen das Bodenaufklärungsvermögen eines optischen Systems:
 - Flughöhe
 - Filmkantenlänge
 - Filmauflösungsvermögen
 - Brennweite des Objektivs
 - Öffnung des Objektivs
 - verwendeter Film (S/W oder Farbe)

- Wellenlänge
 - Blickwinkel des optischen Systems
 - Abstand Objektiv–Filmebene
8. Verwendet man ein LandSat Instrument (Flughöhe 900 km) auf einem militärischen Späher (Flughöhe 180 km), so verändern sich das Blickfeld und das Bodenaufklärungsvermögen. Wie?
 9. Das spektrale Reflektionsvermögen verschiedener Pflanzen und Nicht-Vegetation ist in der folgenden Abbildung dargestellt. Wie würden Sie vier Spektralkanäle wählen (bitte in Abbildung markieren), um Vegetation möglichst sicher identifizieren und zumindest ansatzweise auch klassifizieren zu können.



10. Welche Faktoren verschlechtern das Bodenaufklärungsvermögen?
11. Geben Sie die charakteristischen Größen (Auflösungsvermögen) eines optischen Systems an und charakterisieren Sie diese kurz
12. Der Wunsch nach einem hohen Bodenaufklärungsvermögen ist bei einem Erdfernerkundungsinstrument zumindest aus Sicht der Kartographen verständlich. Warum interessieren sich viele Anwender auch für ein gutes spektrales Auflösungsvermögen?
13. Welche der folgenden Probleme können unter Verwendung eines im thermischen Infrarot arbeitenden Instruments untersucht werden:
 - Identifikation von 'Wärmeverschmutzern' (Kraftwerke, Autobahnen etc.)
 - Unterscheidungen verschiedener Vegetationsformen und Pflanzengesundheit
 - Unterscheidung verschiedener Gesteinsarten
 - Bestimmung der Wolkenbedeckung
 - Identifikation unterirdischer Wasserläufe oder -lager
 - Bestimmung der Wolkenhöhe
 - Beobachtung von Atmosphäre–Ozean–Wechselwirkung wie El Niño
14. Erläutern Sie die Kehrbesentechnik. Was sind die Vorteile gegenüber einem konventionellen photographischen System?
15. Skizzieren und beschreiben Sie die wesentlichen Elemente eines Multi-Spektralscanners (Hinweis: es gibt verschiedene Realisierungsformen, bitte eine relativ genau beschreiben, die anderen nur kurz).
16. Auf Thematic Mapper Aufnahmen von Landsat sind trotz eines Bodenaufklärungsvermögens von 30 m Strukturen mit einer Breite von nur wenigen Metern (Eisenbahnlinien, Molen u.ä.) deutlich zu erkennen. Erläutern Sie, möglichst mit Hilfe einer Skizze, warum.
17. Was ist ein Mixed Pixel? Wodurch wird sein Auftreten begünstigt? Haben Sie Vorschläge für Instrumententwickler, mit deren Hilfe das Auftreten von Mixed Pixels reduziert oder zumindest in der anschließenden Datenauswertung korrigiert werden kann?
18. Welche Messverfahren werden bei der Untersuchung der Atmosphäre (insbesondere Zusammensetzung und Spurengase) verwendet?
19. Erläutern Sie das Verfahren des Limb-Sounding.
20. Erläutern Sie das Messprinzip eines Ozonsensors, der vom Boden oder im Limb-Viewing eingesetzt werden kann. Welche physikalischen Grundprinzipien und -gleichungen gehen ein?

21. Welche Wellenlängenbereiche können sinnvoll zum Nachweis von Kohlendioxid in der Atmosphäre verwendet werden?
 - UV
 - weiches Röntgenlicht
 - sichtbares Licht
 - nahes Infrarot
 - thermisches Infrarot
 - Mikrowellen
22. In welchem Wellenlängenbereich sind Aerosol-Messungen sinnvoll und warum?
23. Welches Bodenaufklärungsvermögen haben Sensoren im thermischen Infrarot? Kann man das Bodenaufklärungsvermögen in den Bereich optischer Sensoren bringen? Begründen Sie.
24. Skizzieren Sie den Aufbau eines Infrarotscanners (thermisches IR) und beschreiben Sie das Funktionsprinzip.
25. Welche Größe wird von einem Infrarotscanner gemessen und wie wird sie anschließend interpretiert?
26. Erläutern Sie den Begriff der Effektivtemperatur. Geben Sie die relevante(n) Gleichung(en).
27. Welche Vor- und Nachteile haben passive Mikrowellendetektoren gegenüber optischen Instrumenten und Instrumenten, die im thermischen Infrarot arbeiten? Nennen Sie Anwendungsbeispiele.
28. Erläutern Sie, wie mit Hilfe eines aktiven Mikrowelleninstrumentes vom Satelliten aus die Wellenhöhe bestimmt werden kann.
29. Aus dem gemessenen Gesamtstrahlungsstrom lässt sich die Effektivtemperatur eines Körpers bestimmen. Steht diese in einer systematischen oder zufälligen Beziehung zu dessen realer Temperatur?

Aufgaben

1. Auf dem Shuttle wurde für 'hübsche Bilder' die Earth Terrain Camera mit einer Brennweite von 450 mm und einem Mittelformatfilm mit 6 cm Kantenlänge verwendet (quadratisches Bild). Bestimmen Sie für eine Flughöhe von 400 km das Blickfeld auf dem Boden sowie das Bodenaufklärungsvermögen bei einer Filmempfindlichkeit von 300 Linien/mm.
2. Ein leichtes Teleobjektiv einer Standard-Spiegelreflexkamera hat eine Brennweite von 100 mm und eine Blendenöffnung von 5 cm. Für ein mittleres λ von $5 \cdot 10^{-5}$ cm ist der Durchmesser des Beugungsscheibchens in der Filmebene zu bestimmen. Ist es sinnvoll, in dieser Kamera einen Film mit einem Auflösungsvermögen von 500 Linien/mm einzusetzen?
3. Sie haben ein Teleobjektiv mit einer Brennweite von 500 mm und einen Film mit einem Auflösungsvermögen von 500 Linien/mm. Welche Öffnung sollte das Objektiv haben, damit Sie auch im nahen IR ($\lambda = 1 \mu\text{m}$) ein maximales Bodenaufklärungsvermögen erhalten?
4. Auflösungsvermögen (Hubble verkehrt): Aufgrund der äußeren Abmessungen amerikanischer Spionagesatelliten wird manchmal spekuliert, dass diese Satelliten mit einem Instrument ähnlich dem Hubble-Space-Teleskop ausgestattet sind – nur mit dem Unterschied, dass das Instrument zur Erde blickt. Das Teleskop hat eine Brennweite von 6 m. (a) Bestimmen Sie das Bodenaufklärungsvermögen für den Fall, dass das Instrument (i) mit einem CCD aus einer Digitalkamera (ca. 2500 mal 2000 Pixel) betrieben wird oder (ii) mit einem konventionellen Film mit einer Auflösung von 500 Linien/mm (ebenfalls als Kleinbild-Film mit Maßen 36 mm mal 24 mm). Diskutieren Sie jeweils für eine militärische Flughöhe von 250 km und eine zivile Flughöhe von 900 km. (b) Welche Blendenöffnung muss das Teleskop für die verschiedenen Konfigurationen mindestens haben? (c) Wie groß ist der auf einem Bild abgebildete Bodenbereich jeweils? (d) Wieviele Orbits vergehen, bis das Gebiet erneut überflogen wird? (e) Welche Zeit steht maximal für eine Aufnahme zur Verfügung? (f) Wie groß ist die Verringerung des Auflösungsvermögens

- durch die Eigenbewegung des Satelliten bei einer Belichtungszeit von $1/500$ s? (g) Lässt sich dieser Störeinfluss verringern?
5. Der Thematic Mapper auf LandSat scannt einen 185 km breiten Streifen auf dem Boden mit einem Bodenaufklärungsvermögen von 30 m ab. Wieviel Zeit steht für einen Streifen zur Verfügung und wie schnell muss der Spiegel (beidseitig beschichtet) rotieren? Wie groß ist die Belichtungszeit für ein einzelnes Pixel? Eine Szene hat eine Länge von 170 km (bei der Streifenbreite von 185 km). Wieviele Pixel hat eine Szene? Wie lange dauert die Aufnahme einer vollständigen Szene?
 6. Der französische Fernerkundungssatellit SPOT fliegt in einem nahezu kreisförmigen Orbit in einer Höhe von 832 km. Bestimmen Sie die Bahngeschwindigkeit des Satelliten und die Geschwindigkeit, mit der er seine Bodenspur abscannt. Welche Auswirkungen hat die Bodengeschwindigkeit auf Aufnahmen.
 7. Sie haben einen Satelliten, auf dem eine Videokamera mit einem Blickwinkel von 1° montiert ist. Ihr Ziel ist es, die gesamte Erdoberfläche zu filmen. Welche Inklination müssen Sie wählen? Bestimmen Sie die Zahl der benötigten Umläufe in Abhängigkeit von der Flughöhe. Diskutieren Sie die Vor- und Nachteile besonders niedriger oder hoher Orbits.
 8. Der Multispectralscanner MSS auf LandSat tastet bei einer Spiegeldrehung einen 476 m breiten Bereich quer zur Flugrichtung ab. Um ein Bodenaufklärungsvermögen von 79 m zu kommen, wird dieser Streifen in jedem Wellenlängenband auf 6 CCD-Elemente abgebildet. Die Breite des abgetasteten Bereichs senkrecht zur Flugrichtung beträgt 185 km. Bestimmen Sie die Rotationsfrequenz des Spiegels (beidseitig reflektierend). Welche Zeit steht für die Aufnahme eines einzelnen Pixels zur Verfügung (die Pixel sind quadratisch)? Vergleichen Sie diesen Wert mit dem Wert, der sich ergäbe würde, wenn pro Spektralkanal nur ein CCD-Element zur Verfügung stehen würde.
 9. Berechnen Sie den Gesamtstrahlungsstrom, den ein Körper mit einer Temperatur von 35°C und einer Emissivität von 0.95 emittiert ($\sigma = 5.67 \cdot 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$). Welche Effektivtemperatur hat dieser Körper?
 10. Die Erde hat eine Effektivtemperatur von 255 K und emittiert Infrarotstrahlung im Bereich zwischen 3 und $100 \mu\text{m}$. Bei welcher Wellenlänge liegt das Maximum der terrestrischen Infrarotstrahlung? Scanner im thermischen Infrarot nutzen nur einen Bereich zwischen 3.5 und $5.5 \mu\text{m}$ sowie zwischen 8 und $14 \mu\text{m}$. Warum?
 11. Auf der Nordsee schwimmt ein dünner Ölfilm. Wasser und Ölfilm haben eine Temperatur von 20°C ; Wasser hat eine Emissivität von 0.993, ein Ölfilm auf Wasser eine von 0.972. Wie gut sollte das thermische Auflösungsvermögen eines Infrarotscanners mindestens sein, um den mit Ölfilm verschmutzten Bereich der Nordsee von der sauberen Nordsee zu unterscheiden?
 12. Beurteilen Sie, ob man von einem militärischen Aufklärungssatelliten aus die Schlagzeile der Bild oder Prawda lesen kann (Hinweis: die größten annehmbaren Brennweiten sind 6 m, das Filmauflösungsvermögen kann maximal 500 Linien/mm betragen). Welchen Öffnungsdurchmesser würden Sie benötigen? Die Flughöhe kann mit 200 km abgeschätzt werden.
 13. Bestimmen Sie das Bodenaufklärungsvermögen Ihres Auges, wenn Sie sich als Beobachter auf einem der LandSat-Satelliten befinden würden (Flughöhen in Tab. 1.1) und vergleichen Sie mit dem Auflösungsvermögen des TM oder MSS. Hinweis zum Auge: Die Rezeptoren haben einen Durchmesser von $4 \mu\text{m}$, die 'Bildweite' des Auges beträgt 20 mm, die effektive Weite einer Rezeptorzelle ist $4/20000$ oder $1/5000$ der Bildweite [101]. Welche Rezeptorgröße würden Sie benötigen, um ein dem TM vergleichbares Bodenaufklärungsvermögen zu erreichen? Ist dieser Wert sinnvoll? In welchem Bereich des elektromagnetischen Spektrums müssten Sie dann 'sehen'?
 14. Der oszillierende Spiegel des TM vollendet 14 Scans während einer Sekunde. Welche Zeit wird dabei für ein einzelnes Pixel aufgewendet? Vergleichen Sie mit den Verschlusszeiten normaler Kameras.

Kapitel 4

Missionen

In diesem Kapitel werden einige ausgewählte Missionen zur Erdfernerkundung vorgestellt. Dabei werden Klassiker wie die LandSat und NOAA-Satelliten ebenso berücksichtigt wie Europas modernster Erdfernerkundungssatellit EnviSat und einige eher unbekanntere kleinere Missionen wie CHAMP.

4.1 Der klassische Geograph: LandSat

LandSat ist ein US-amerikanisches Satellitenprogramm. Das LandSat-Programm hat ursprünglich aus einer Mischung aus technologischen Zielen und kontinuierlicher Erdbeobachtung bestanden. 1983 ging das System von der NASA and NOAA über, seit dem hat der technologische Aspekt nur geringe Bedeutung und das Schwergewicht liegt auf der kontinuierlichen Erdbeobachtung. Die wesentlichen Merkmale der LandSat-Satelliten sind in Tab. 4.1 zusammen gefasst. Eine Übersicht über alle LandSat-Satelliten inklusive ihrer Instrumentierung findet sich auf der LandSat Homepage der NASA unter <http://geo.arc.nasa.gov/sge/landsat/landsat.html>.

Der erste LandSat-Satellit wurde 1972 unter dem Namen ERTS-1 (Earth Resources Technology Satellite) gestartet, später in LandSat 1 umbenannt. Die experimentellen technologischen Aspekte dieses Satelliten waren seine Verwendung einer Videokamera (RBV, später nur für planetare Missionen verwendet, nicht jedoch in der Erdfernerkundung) und des Prototyps des Multispektralscanners MSS.

LandSat 2 und 3 folgten 1975 bzw. 1978 mit ähnlicher Ausstattung, ihre Bilder begegnen uns in verschiedenen älteren Atlanten. Auf LandSat 4 und 5, gestartet 1982 bzw. 1984, wurde die Videokamera durch den Thematic Mapper TM ersetzt, außerdem verfügten beide über

S/C	S/C Launch	Sensor Completion	Data Resolution (m)	Data Communications	Orbital Altitude	S/C Operator(s)	End of Service/ Status
LS-1 (ERTS)	Jul. 23, 1972	RBV MSS, DCS	80 80	DD (Direct Downlink) 2 WBVTR	907 km	NASA	Jan. 6, 1978
LS-2	Jan. 22, 1975	RBV MSS,DCS	80 80	DD with 2 WBVTR	908 km	NASA	1983
LS-3	Mar. 5, 1978	RBV MSS,DCS	30 80	DD with 2 WBVTR	915 km	NASA	Jan. 7, 1983
LS-4	Jul. 16, 1982	MSS TM, GPS	80 30	DD TDRSS	705 km	NOAA ('83)/ Eosat ('85)	standby Dec. 1993
LS-5	Mar. 1, 1984	MSS TM, GPS	80 30	DD TDRSS	705 km	NOAA ('84)/ Eosat ('85)	operational as of 2001
LS-6	Oct. 5, 1993	ETM	15 (PAN) 30 (MS)	DD with re- corders	launch failure (contact lost during launch)		
LS-7	Apr. 15, 1999	ETM+	15 (PAN) 30 (MS)	DD with re- corders	705 km	NOAA	

Tabelle 4.1: Übersicht über die LandSat Satelliten [62]

Landsat sensor	MSS (LS-1-5)	TM (LS-4/5)	ETM (on LS-6)	ETM+ (on LS-7)
Spectral bands (all bands in μm)	1) 0.5 - 0.6 2) 0.6 - 0.7 3) 0.7 - 0.8 4) 0.8 - 1.1	1) 0.45 - 0.52 VNIR 2) 0.52 - 0.60 VNIR 3) 0.63 - 0.69 VNIR 4) 0.76 - 0.90 VNIR 5) 1.55 - 1.75 SWIR 7) 2.08 - 2.35 SWIR 6) 10.4 - 12.5 TIR	P) 0.52 - 0.90 VNIR 1) 0.45 - 0.52 VNIR 2) 0.52 - 0.60 VNIR 3) 0.63 - 0.69 VNIR 4) 0.76 - 0.90 VNIR 5) 1.55 - 1.75 SWIR 7) 2.08 - 2.35 SWIR 6) 10.4 - 12.5 TIR	P) 0.52 - 0.90 VNIR 1) 0.45 - 0.52 VNIR 2) 0.53 - 0.61 VNIR 3) 0.63 - 0.69 VNIR 4) 0.78 - 0.90 VNIR 5) 1.55 - 1.75 SWIR 7) 2.09 - 2.35 SWIR 6) 10.4 - 12.5 TIR
Swath width	185 km	185 km	185 km	185 km
Spatial resolution	80 m	30 m VNIR/SWIR 120 m TIR	15 m PAN, 30 m VNIR/SWIR, 120 m TIR	15 m PAN 30 m VNIR/SWIR 60 m TIR
Radiometric resolution	6 bit	8 bit	9 bit (8 bit transmitted)	9 bit (8 bit transmitted)
Band-to-band registration		0.2 pixel (90%)	0.2 pixel (90%)	0.2 pixel (90%)
Geodetic accuracy without ground control		500 m (90%)	1000 m (90%)	400 m (90%)
Data rate	15 Mbit/s	85 Mbit/s	2 x 85 Mbit/s	2 x 75 Mbit/s
Instrument mass	64 kg	258 kg	288 kg scanner, plus 81 kg AEM	318 kg scanner, plus 103 kg AEM, plus 20 kg cable harness
Average power	50 W	332 W	490 W	590 W
Telescope aperture	23 cm	40.6 cm	40.6 cm	40.6 cm

Tabelle 4.2: Charakteristika der LandSat-Instrumente MSS, TM und ETM+ [62]

ein, allerdings dem DoD Programm unterstellte, GPS-System. Auch fliegen beide in einem niedrigeren Orbit: durch die größere Streifenbreite und die geringere Umlaufzeit verringert sich die Wiederholfrequenz des Überflugs von 18 auf 16 Tage. Die Orbits sind sonnensynchron.

1995 ging LandSat 6 während des Starts verloren: zwar wurde er von der Rakete in ein niedriges Orbit eingetragen, nach dem Zünden der Booster zum Übergang auf ein Transfer-Orbit ging jedoch der Kontakt verloren. Da auch keine Radarechos des Satelliten oder seiner Trümmer gefunden werden konnten, vermutet man, dass die Lagekontrolle nicht funktioniert hat und LandSat von seinen Booster in einen Ozean katapultiert wurde.

1999 wurde als Ersatz LandSat 7 gestartet. Dieser trägt nur noch ein Instrument, den Enhanced Thematic Mapper+ als eine Kombination aus altem TM und MSS mit deutlich verbessertem Bodenaufklärungsvermögen. Die sieben Spektralkanäle sind denen des TM ähnlich, ein zusätzlicher panchromatischer Kanal ermöglicht das hohe Bodenaufklärungsvermögen (der Ansatz ist von SPOT bekannt). Die Instrumentdaten sind mit denen von MSS und TM in Tabelle 4.2 zusammen gefasst. Die wichtigsten Informationen über LandSat 7 finden sich auf der LandSat 7 Homepage der NASA unter <http://landsat.gsfc.nasa.gov/LandSat7>, Informationen über den Enhanced Thematic Mapper ETM+ unter <http://edc.usgs.gov/products/satellite/landsat7.htmlETM+>.

Während für die ersten LandSat Satelliten neben der Instrumenterprobung Kartographie und Geländeklassifikation wichtigste Aspekte der Missionen waren, ist für die jüngsten LandSat-Vertreter das Monitoring, d.h. die kontinuierliche Beobachtung, ein wichtigste Aspekt. Gerade für alle den 'global change' betreffende Fragen ist eine langandauernde, kontinuierliche Datenbasis erforderlich. Daher wurde LandSat 5 nicht abgeschaltet sondern wird parallel zu LandSat 7 betrieben, da LandSat 5 der letzte Satellit ist, der einen MSS trägt. Sollte LandSat 5 bzw. der auf ihm befindliche MSS ausfallen, so wird der auf StandBy stehende LandSat 4 als Ersatz zur Fortsetzung der MSS Aufnahmen aktiviert werden.

4.2 Ein Meteorologe: MeteoSat

MeteoSat ist das Wettersatellitenprogramm der ESA. Es wurde 1972 ins Leben gerufen, der erste Satellit MeteoSat-1 wurde 1977 in einen geostationären Orbit gebracht und arbeitete bis 1979. Sein Nachfolger, MeteoSat-2 war von 1981 bis 1988 aktiv und wurde 1991 in eine andere Position manöviert und von einem anderen Träger weiter betrieben. Der letzte Satellit der ersten Serie, MeteoSat-7, wurde 1997 gestartet. MeteoSat hat eine Gesamtmasse von 322 kg. Der erste Satellit der zweiten Serie, MeteoSat 8, wurde im Januar 2004 in Betrieb genommen.

Die MeteoSat Satelliten sind spinstabilisiert mit 100 Umdrehungen pro Minute, die Spinachse steht senkrecht zur Bahnebene (Cartwheel, häufig bei Wettersatelliten verwendet, da keine Verschleiß-empfindlichen mechanisch beweglichen Teile vorhanden sind, zu deren Betrieb auch noch Energie benötigt wird). Die Rotation dient gleichzeitig zum Scannen. Da die Fragestellungen der Satelliten einfach sind, nämlich Wettervorhersage, verfügen sie einheitlich über ein relativ einfaches Instrument VISSR (Visible and Infrared Spin Scan Radiometer), auch bezeichnet als MSR (Multispectral Radiometer) oder MVIR (Meteosat Visible and Infrared Radiometer). VISSR hat drei Kanäle: VIS zwischen 0.5 und 0.9 μm für die Aufnahmen im sichtbaren Bereich, wie sie uns bei der Wettervorhersage gezeigt werden, WV zwischen 5.7 und 7.1 μm detektiert die Emission von Wasserdampf, und TIR im Bereich von 10.5–12.5 μm misst im thermischen Infrarot. Das Bodenaufklärungsvermögen beträgt im VIS 2.5 km, in den beiden anderen Kanälen 5 km.

Die Entwicklung der MSG (MeteoSat Second Generation) Satelliten führt zu keiner Änderung des Grundkonzepts, nämlich das Scannen mit der Stabilisierung des Spacecraft zu verbinden. VISSR wird ersetzt durch seine Erweiterung SEVIRI0 (Spinning Enhanced Visible and Infrared Imager) mit insgesamt 12 Spektralbändern (5 im sichtbaren bzw. nahen IR, 2 für Wasserdampf und 5 im thermischen Infrarot). Zusätzlich werden die MSG Satelliten mit GERB (Geostationary Earth Radiation Budget) um ein Instrument erweitert: die Messung der globalen Strahlungsbilanz ist sowohl zum Verständnis der Variationen der Solar-konstante als auch zum Verständnis von climate change erforderlich. Die Masse der Satelliten wird sich dadurch auf 2040 kg erhöhen, die Leistungsaufnahme auf 600 W; MeteoSat-8 als erster Vertreter dieser neuen Generation wurde im Januar 2004 in Betrieb genommen.

Eine vollständige Übersicht über MeteoSat-8 mit Informationen auch zu seinen Vorgängern, findet sich auf der Seite seines Betreibers EumetSat unter <http://www.eumetsat.de/>.

4.3 Ein tiefer gelegter Meteorologe: TIROS bis NOAA

Einen anderen Ansatz verfolgt das US-amerikanische meteorologische Programm TIROS (Television Infrared Observation Satellite): statt geostationärer Orbits werden hier Satelliten in niedrigen polaren Orbits verwendet. Durch ein gegenüber Kartographie-Satelliten deutlich reduziertes Bodenaufklärungsvermögen haben die Instrumente ein großes Blickfeld und damit eine für die Wettervorhersage ausreichende Wiederholrate.

Dieses Programm besteht eigentlich aus mehreren Serien von Satelliten, die jeweils aus ihren Vorgängern entwickelt wurden. Diese Satelliten tragen neben den klassischen Instrumenten zur Meteorologie auch weitere Instrumente für umweltbezogene Fragestellungen. Am bekanntesten ist vielleicht das AVHRR (Advanced Very High Resolution Radiometer) auf den NOAA-Satelliten, das ähnlich den MSS- und TM-Instrumenten ein sehr gutes Universalinstrument ist. Die verschiedenen Satelliten und Serien sind in Tabelle 4.3 zusammen gefasst.

Die erste Serie dieser Satelliten umfasst TIROS-1 bis TIROS-10. Die wichtigsten Instrumente sind Videokameras mit Weitwinkel (TV-WA) bzw. Tele (TV-NA). Diese wurden auf allen Satelliten der ersten TIROS-Serie geflogen, die Instrumente haben ein Blickfeld von $1200 \times 1200 \text{ km}^2$ bei einer Bodenaufklärung von 2.5–3 km bzw. $120 \times 120 \text{ km}^2$ bei einer Auflösung von 0.3–0.8 km. Beide Videokameras arbeiten im panchromatischen Modus, Aufnahmen im thermischen Infrarot wurden erst mit dem Scanning Radiometer (SR, Masse 2 kg, Leistungsaufnahme 1 W) ab TIROS-2 möglich. Auf allen weiteren Satelliten der ersten Serie wurden neben den Videokameras weitere Instrumente geflogen, viele davon eher experimenteller Art.

TIROS-9 ist der erste ‘Cartwheel’ Meteorologiesatellit: bei ihm steht, wie bereits bei den MeteoSats erwähnt, die Spinachse senkrecht auf der Bahnebene, so dass die Spinbewegung gleichzeitig ein Scannen der Erdoberfläche ermöglicht. TIROS-9 und 10 waren gleichzeitig auch die Testsatelliten für die zweite Generation von Meteorologie-Satelliten: TOS (TIROS Operational Satellite) oder ESSA (Environmental Science and Services Administration). Auch hier ist der Schwerpunkt die Beobachtung der Wolkenbedeckung und der Wettersysteme. Die Instrumentierung besteht aus dem AVCS (Advanced Vidicon Camera System) mit

Satellite Name	Launch Date	Period (min)	Perigee (km)	Apo-gee (km)	Inclin. (°)	Remarks/Sensor Complement
TIROS-1	01 Apr 60	99.2	796	867	48.3	1 TV-WA and 1 TV-NA
TIROS-2	23 Nov 60	98.3	717	837	48.5	1 TV-WA, 1 TV-NA, passive & active IR scan
TIROS-3	12 Jul 61	100.4	854	937	47.8	2 TV-WA, HB, IR, IRP
TIROS-4	08 Feb 62	100.4	817	972	48.3	1 TV-WA, IR, IRP, HB
TIROS-5	19 Jun 62	100.5	680	1119	58.1	1 TV-WA, 1 TV-MA
TIROS-6	18 Sep 62	98.7	783	822	58.2	1 TV-WA, 1 TV-MA
TIROS-7	19 Jun 63	97.4	713	743	58.2	2 TV-WA, IR, ion probe, HB
TIROS-8	21 Dec 63	99.3	796	878	58.5	1st APT TV direct readout & 1 TV-WA
Nimbus-1	28 Aug 64	98.3	487	1106	98.6	3 AVCS, 1 APT, HRIR, three-axis stabilization
TIROS-9	22 Jan 65	119.2	806	2967	96.4	First "wheel," 2 TV-WA, global coverage
TIROS-10	02 Jul 65	100.6	848	957	98.6	Sun synchronous, 2 TV-WA
ESSA-1	03 Feb 66	100.2	800	965	97.9	1st operational system, 2 TV-WA, FPR
ESSA-2	28 Feb 66	113.3	1561	1639	101.0	2 APT, global operational APT
Nimbus-2	15 May 66	108.1	1248	1354	100.3	3 AVCS, HRIR, MRIR, APT
ESSA-3	02 Oct 66	114.5	1593	1709	101.0	2 AVCS, FPR
ATS-1	06 Dec 66	24 hr	41,257	42,447	0.2	SSCC (Spin Scan Cloudcover Camera)
ESSA-4	26 Jan 67	113.4	1522	1656	102.0	2 APT
ESSA-5	20 Apr 67	113.5	1556	1635	101.9	2 AVCS, FPR
ATS-2	05 Nov 67	24 hr	41,166	41,222	0.4	MSSCC (Multicolor Spin scan Cloud Camera)
ESSA-6	10 Nov 67	114.8	1622	1713	102.1	2 APT TV
ESSA-7	16 Aug 68	114.9	1646	1691	101.7	2 AVCS, FPR, S-band
ESSA-8	15 Dec 68	114.7	1622	1682	101.8	2 APT TV
ESSA-9	26 Feb 69	115.3	1637	1730	101.9	2 AVCS, FPR, S-band
Nimbus-3	14 Apr 69	107.3	1232	1302	101.1	HRIR, IRIS, MRIR, MUSE, SIRS, IDICS, IRLS, 2 APT, 2 AVCS, 2 SR, FPR, three-axis stabilization
TIROS-1	23 Jan 70	115.1	1648	1700	102.0	SIRS B, IRIS, SCR, THIR, , FWS, IDCS,
Nimbus-4	15 Apr 70	107.1	1200	1280	99.9	IRLS, MUSE, BUUV
NOAA-1	11 Dec 70	114.8	1422	1472	102.0	2 APT, 2 AVCS, 2 SR, FPR
NOAA-2	15 Oct 72	114.9	1451	1458	98.6	2 VHRR, 2 VTPR, 2 SR, SPM
Nimbus-5	11 Dec 72	107.1	1093	1105	99.9	SCMR, ITPR, NEMS, ESMR, THIR
NOAA-3	06 Nov 73	116.1	1502	1512	101.9	2 VHRR, 2 VTPR, 2 SR, SPM
SMS-1	17 May 74	1436.4	35,605	35,975	0.6	VISSR, DCS, WEFAX, SEM (EPS,MAG,XRS)
NOAA-4	15 Nov 74	101.6	1447	1461	114.9	2 VHRR, 2 VTPR, 2 SR, SPM,
SMS-2	06 Feb 75	1436.5	35,482	36,103	0.4	VISSR, DCS, WEFAX, SEM
Nimbus-6	12 Jun 75	107.4	1101	1115	99.9	ERB, ESMR, HIRS, LRIR, T+DRE, SCAMS, THIR, TWERLE, PMR
NOES-1	16 Oct 75	1436.2	35,728	35,847	0.8	VISSR, DCS, WEFAX, SEM (EPS,MAG,XRS)
NOAA-5	29 Jul 76	116.2	1504	1518	102.1	2 VHRR, 2 VTPR, 2 SR, SPM
NOES-2	16 Jun 77	1436.1	35,600	36,200	0.5	VISSR, DCS, WEFAX, SEM (EPS,MAG,XRS)
NOES-3	15 Jun 78	1436.1	35,600	36,200	0.5	VISSR, DCS, WEFAX, SEM (EPS,MAG,XRS)
TIROS-N	13 Oct 78	98.9	849	864	102.3	AVHRR, HIRS/2, SSU, MSU, HEPAD, MEPED, TED
Nimbus-7	24 Oct 78	99.3	943	955	104.1	LIMS, SAMS, SAM-II, SBUV/TOMS, , ERB, SMMR, THIR, CZCS, SMMR
NOAA-6	27 Jun 79	101.3	807.5	823	98.7	AVHRR, HIRS/2, SSU, MSU, HEPAD, MEPED, TED
NOES-4	09 Sep 80	1436.1	35,600	35,600	0.5	VAS, DCS, SEM (EPS,MAG,XRS,HEPAD), WEFAX
NOES-5	22 May 81	1436.1	35,600	35,600	0.5	VAS, DCS, SEM (EPS,MAG,XRS,HEPAD), WEFAX
NOAA-7	23 Jun 81	101.9	852	869	98.9	AVHRR, HIRS/2, SSU, MSU, HEPAD, MEPED, TED
NOAA-8	28 Mar 83	101.2	801	826	98.2	AVHRR, HIRS/2, SSU, MSU, MEPED, TED, S&R
NOES-6	28 Apr 83	1436.1	35,803	35,771	0.1	VAS, DCS, WEFAX, SEM (EPS,MAG,XRS)
ERBS	05 Oct 84	96.8	393	608	57	ERBE, SAGE-II
NOAA-9	12 Dec 84	102.0	842	862	98.9	AVHRR, (HIRS/2, SSU, MSU)=TOVS, S&R, SBUV/2, ERBE
NOAA-10	17 Sep 86	101.3	803	824	98.7	AVHRR, (HIRS/2, SSU, MSU)=TOVS, MEPED, TED, S&R, SBUV, ERBE
NOES-7	26 Feb 87	1436.1	35,759	35,826	0.1	VAS, DCS, WEFAX, SEM (EPS,MAG,XRS), S&R
NOAA-11	24 Sep 88	102.1	845	863	98.9	AVHRR, (HIRS/2, SSU, MSU)=TOVS, S&R, SBUV/2
NOAA-12	14 May 91	101.3	806	825	98.7	AVHRR, HIRS/2, MSU, SEM (MEPED, TED)
NOAA-13	09 Aug 93	a S/C failure (power loss) occurred on Aug. 21, 1993				AVHRR, TOVS, SBUV, SEM, ARGOS, S&R, EHC, MAXIE
NOES-8	13 Apr 94	Geosynchronous orbit				Imager, Sounder, SEM, DCS, Wefax, S&R
NOAA-14	30 Dec 94	102.1	847.7	860.8	98.9	AVHRR, SBUV/2, HIRS/2I, MSU, SSU, SEM, ARGOS (DCS), S&R
NOES-9	23 May 95	Geosynchronous orbit				Imager, Sounder, SEM, DCS, Wefax, S&R
NOAA-15	13 May, 98	101.3	803.7	817.8	98.70	AVHRR/3, HIRS/3, AMSU-A (A1+A2), AMSU-B, SEM-2, ARGOS (DCS-2), S&R
NOAA-16	21 Sept, 00	102.1	845.4	860.4	98.74	AVHRR/3, HIRS/3, AMSU-A (A1+A2), AMSU-B, SBUV/2, SEM-2, ARGOS (DCS-2), S&R

Tabelle 4.3: Übersicht amerikanischer Umwelt- und Wettersatelliten von TIROS bis NOAA [62]

einem Blickfeld von $1200 \times 1200 \text{ km}^2$ und einem Auflösungsvermögen von 3 km. Die Verbesserungen betrafen die technischen Details wie Datenspeicherung und Übertragung sowie das radiometrische Auflösungsvermögen. Ab TOS-3 wird zur Untersuchung der Strahlungsbilanz (reflektierte solare Einstrahlung) ein oder mehrere FPRs (Flat Plate Radiometer, 3 kg, 4 W) geflogen (da wird 1966 das begonnen, was die ESA erst mit der zweiten MeteoSat-Generation angehen wird). Die Serie endet mit TOS-9.

Ab 1970 wird von NOAA die Serie als ITOS (Improved TIROS Operational System) betrieben, der erste Satellit NOAA-1 (auch ITOS-A, ITOS-1 oder TIROS-M genannt) startet Ende 1970. Neben dem Videosystem AVCS und dem Radiometer FPRs wird auf diesen Satelliten eine Vielzahl anderer Instrumente geflogen, dazu gehören ein Solar Proton Monitor (SPM), ein einfaches Scanning Radiometer (SR) mit einem Kanal im Sichtbaren und einem im thermischen Infrarot und ab NOAA-2 auch VHRR (Very High Resolution Radiometer, 16 kg, 25 W), mit zwei Spektralbändern und einem Auflösungsvermögen von 0.87 km bei

einer Streifenbreite von 2580 km.¹

Die nächste Generation ab TIROS-N ist die vierte Generation, sie lebt auch heute noch mit NOAA-16 als ihrem aktuellsten Vertreter. Hier hat sich der Wettersatellit zum Umweltmonitor mutiert. Das bereits aus der ITOS-Serie bekannte VHRR wurde zum AVHRR (Advanced Very High Resolution Radiometer) weiterentwickelt, ein System TVOS (TIROS Operational Vertical Sounder) von Soundern misst vertikale Profile von Druck, Temperatur und Wasserdampf von der Erdoberfläche bis zur Oberkante der Atmosphäre, und der Space Environment Monitor (SEM) misst energiereiche Teilchen über einen weiten Energiebereich. Außerdem befasst sich ERBE (Earth Radiation Budget Experiment) auf NOAA-9 und NOAA-10 mit der Strahlungsbilanz der Erde.

AVHRR ist in der Abbildungsgeometrie VHRR vergleichbar und misst in 5 Kanälen; ein panchromatischer Kanal im sichtbaren für die Wolkenbedeckung (tagsüber) und Oberflächenkartierung, ein Kanal im nahen IR zur Identifikation von Oberflächengewässern und Vegetation, ein Kanal im atmosphärischen Fenster des IR bei 3.8 μm für die Temperatur der Meeresoberfläche und zum Nachweis von Feuern, und ein Kanal bei 11 μm für Meeresoberflächentemperaturen und nächtliche Wolkenbedeckung. Bei den ersten Versionen des AVHRR ist dieser Kanal doppelt vorhanden (Kalibrierung), bei den späteren Versionen ist er durch einen Kanal bei 12 μm ersetzt, mit dessen Hilfe Oberflächentemperaturen und Tag/Nacht-Unterschiede in der Wolkenbedeckung nachgewiesen werden können. Über das AVHRR gibt es verschiedene gute Dokumentationen, teilweise auch mit Datenbeispielen, z.B. <http://edc.usgs.gov/guides/avhrr.html>, <http://edc.usgs.gov/products/satellite/avhrr.html> und <http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html>.

Da das AVHRR eine relativ grobe Bodenauflösung hat, wird es nicht für die Kartographie verwendet sondern eher zur Untersuchung großskaligerer Phänomene. Viele Anwendungsbeispiele (auch aktuelle Daten) stellt die Uni Bern als JPEG-Dateien unter <http://saturn.unibe.ch/rsbern/noaa/dw/realtime/> (aktuelle Daten) und <http://www.geo.mtu.edu/rs/avhrr/> (Beispiele für Anwendungen zu Fragen des global change) zur Verfügung.

4.4 Der Generalist für Umweltfragen: EnviSat

EnviSat (Environmental Satellite) ist ESA's Schwergewicht (Gesamtmasse 8140 kg, Nutzlastmasse 2150 kg, 6.5 kW Spacecraft Power, 1.9 kW für die Nutzlast) in Sachen Umwelt. Die Nutzlast besteht aus 9 Instrumenten, die sich gegenseitig ergänzen und teilweise überlappen (notwendig zur Kalibrierung). Jedes Instrument ist spezifischen Fragestellungen angepasst. Die Homepage von EnviSat ist zu finden unter <http://envisat.esa.int/>, die Dokumentationen zu den einzelnen Instrumenten sowie Broschüren zu EnviSat und seinen Instrumenten können über die Seite http://envisat.esa.int/services/esa_doc/doc_env.html erreicht werden.

MERIS (Medium Resolution Imaging Spectrometer) beobachtet zwischen 0.29 μm und 1.04 μm mit 15 Sampling-Intervallen, die ganz gezielt auf bestimmte Stoffe bzw. Fragestellungen angepasst sind, vgl. Tab. 4.4. Das Bodenaufklärungsvermögen beträgt $1040 \times 1200 \text{ m}^2$ über den Ozeanen bzw. $260 \times 300 \text{ m}^2$ über Land, die Aufnahme des Bildes erfolgt auf einem CCD-Array, d.h. das Instrument enthält keine beweglich Komponente zum Scannen sondern bildet wieder wie beim Film auf eine Fläche ab. Die Kombination von mittlerem Bodenaufklärungsvermögen mit sehr gutem spektralen Auflösungsvermögen soll das Monitoring von Ozeanen, Atmosphäre und Land ermöglichen, insbesondere die Messung biophysikalischer und biochemischer Parameter der Ozeanoberfläche (Chlorophyll, Gelbstoff, gelöste

¹Das Grundkonzept von VHRR ist natürlich gar nicht so verschieden von dem von TM oder MSS auf Landsat, der wesentliche Unterschied liegt im Bodenaufklärungsvermögen und in der Streifenbreite: Landsat ist für Kartographie und Detailklassifizierung gebaut, d.h. das Bodenaufklärungsvermögen muss sehr gut sein, aber die Aufnahmezeit kann recht lang sein, da sich Berge und Kontinente eher stationär sind. VHRR dagegen betrachtet globale und relativ schnell veränderliche Prozesse wie Wolken, Wettersystem oder im thermischen IR Meeresströmungen. Deren Zeitskalen erfordern jedoch eine häufige Beobachtung während ihre räumlichen Skalen so groß sind, dass ein Auflösungsvermögen im Bereich von einigen zehn Metern übertrieben ist.

Band No.	Band center (nm)	Bandwidth (nm)	Application
1	412.5	10	Yellow substance and turbidity
2	442.5	10	Chlorophyll absorption maximum
3	490	10	Chlorophyll and other pigments
4	510	10	Turbidity, suspended sediment and red tides
5	560	10	Chlorophyll, suspended sediment
6	620	10	Suspended sediment
7	665	10	Chlorophyll absorption
8	681.25	7.5	Chlorophyll fluorescence, red edge
9	705	10	Aerosol, red edge transition
10	753.75	7.5	Oxygen absorption reference band, vegetation
11	760	2.5	O ₂ absorption R-branch
12	775	15	Aerosol, vegetation
13	865	20	Aerosols correction over ocean
14	890	10	Water vapor absorption reference
15	900	10	Water vapor absorption, vegetation

Tabelle 4.4: Spektralbänder von MERIS

Teilchen), die Beobachtung von Meeresverschmutzung und Erosion, die Vermessung von atmosphärischem Wasserdampf und Aerosol, die Beobachtung der Wolkenbedeckung, die Kartierung von Schnee- und Eisflächen an Land, die Beobachtung der Bioproduktivität sowie Untersuchungen zur Pflanzengesundheit. Eine genauere Instrumentbeschreibung ist unter <http://esapub.esrin.esa.it/bulletin/bullet103/besy103.pdf> zu finden, die Beschreibung der Datenprodukte und Dokumentation unter <http://envisat.esa.int/dataproducts/meris/CNTR.htm> (gibt es, wie bei den anderen Instrumenten auch, auch als pdf zum Herunterladen).

MIPAS (Michelson Interferometer for Passive Atmospheric Sounding) ist ein Limb-Emission-Sounder im Spektralbereich zwischen 4.15 μm und 14.6 μm zur Untersuchung der Chemie der Stratosphäre (insbesondere die Messung von O₃, NO, NO₂, HNO₃, HNO₄, N₂O₅, ClONO₂, COF₂, HOCl), zur Klimaforschung (insbesondere Messung der globalen Verteilung relevanter Parameter wie Wolken, Aerosole, O₃, CH₄, H₂O, N₂O, CFCs (F11, F12, F22, CCl₄, CF₄), CO, OCs), zur Untersuchung von Transportprozessen über die Tropopause (insbesondere Austausch von O₃, CH₄, N₂O, C₂H₂, C₂H₆ oder SF₆ zwischen Troposphäre und Stratosphäre) sowie die Chemie der oberen Troposphäre (insbesondere Messung von NO, CO, CH₄, O₃ und HNO₃). MIPAS misst in einem Höhenbereich von 8 km bis 150 km mit einer Höhenauflösung von 3 km. MIPAS alleine hat eine Masse von 327 kg und eine Leistungsaufnahme von 210 W. Die Instrumentbeschreibung findet sich unter <http://esapub.esrin.esa.it/bulletin/bullet101/endemann.pdf>, das Handbuch zu den Datenprodukten unter <http://envisat.esa.int/dataproducts/mipas/CNTR.htm>.

RA-2 (Radar-Altimeter-2) mit einem zusätzlichen Laser Retro-Reflector (LRR) ist ein typisches Altimeter mit der Fragestellung nach der genauen Form des Geoids und der bereits im Zusammenhang mit dem SeaSat Altimeter diskutierten Messung von Wellenhöhen. Der LRR dient der zusätzlichen Überprüfung und Kalibrierung des RA-2, so dass gegenüber früheren Instrumenten eine größere Genauigkeit erwartet wird. Für die Instrumentbeschreibung siehe <http://esapub.esrin.esa.it/bulletin/bullet98/RESTI.PDF>, das Datenhandbuch findet sich unter <http://envisat.esa.int/dataproducts/ra2-mwr/CNTR.htm> RA-2. L

MWR (Microwave Radiometer) ist eher ein Service-Instrument: seine wichtigste Aufgabe ist die Messung des Wasserdampfgehalts der Atmosphäre. Damit lässt sich das Signal von RA-2 korrigieren.

ASAR (Advanced SAR) ist ein fortgeschrittenes bildgebendes Radarsystem, das das von der Erdoberfläche zurück gestreute Radarsignal in fünf verschiedenen Polarisations-Moden misst. Das Bodenaufklärungsvermögen hängt von der gewählten Breite des Abtaststreifens ab; es liegt bei 28 \times 28 m² bei einer Streifenbreite von 100 km bis zu 950 \times 980 m² bei einer Streifenbreite von mehr als 400 km. Diese Flexibilität erlaubt die Anpassung an die räumlichen Skalen einer gerade aktuellen Fragestellung oder Beobachtungskampagne. Die wichtigsten mit ASAR zu untersuchenden Fragestellungen auf globalen Skalen betreffen den Zustand des Meereises auf verschiedenen Skalen, die Kartographie der Eisschichten und ihrer Dynamic, die Kartographie des Meereises und seiner Dynamik, die Beobachtung

natürlicher und anthropogener Verschmutzungen über den Ozeanen sowie der Nachweis großskaliger Veränderungen in der Vegetation (z.B. Desertifikation). Die Modi mit hoher Bodenauflösung erlauben auch die Untersuchung regionaler Fragestellungen, z.B. Küstenschutz, Monitoring des Schiffsverkehrs sowie alle klassischen Fragen wie land- und forstwirtschaftliches Monitoring, Monitoring der Bodenfeuchte und geologische Exploration. Eine Instrumentbeschreibung findet sich unter http://earth.esa.int/pub/ESA_DOC/SP_1225.pdf und <http://esapub.esrin.esa.it/bulletin/bullet102/Desnos102.pdf>, das Datenhandbuch unter <http://envisat.esa.int/dataproducts/asar/CNTR.htm>.

GOMOS (Global Ozone Monitoring by Occultation of Stars) ist ein Limb-Sounding Spektrometer, das das Spektrum vom UV bis zum nahen Infrarot abdeckt. Das Instrument arbeitet im Absorptionsmodus, als Lichtquellen werden nicht wie sonst üblich Sonne oder Mond sondern Sterne verwendet. Dadurch wird eine permanente Datenüberdeckung erreicht – insbesondere sind Messungen in der Polarnacht möglich. Die wichtigsten zu messenden Parameter sind die Höhenprofile von Aerosolen, Temperatur, O₃, H₂O, NO₂ und NO₃ in der Stratosphäre mit einer vertikalen Auflösung von 1.7 km. Die wichtigste Aufgabe von GOMOS ist die langfristige Beobachtung von stratosphärischem und mesosphärischem Ozon und Wasserdampf sowie der Änderungen dieser Parameter. Eine Instrumentbeschreibung steht unter <http://esapub.esrin.esa.it/bulletin/bullet97/ratier.pdf>, das Datenhandbuch unter <http://envisat.esa.int/dataproducts/gomos/CNTR.htm>.

AATSR (Advanced Along Track Scanning Radiometer) ist ein bildgebendes Radiometer, das in 4 Kanälen im IR (1.6 μm für Wolkenbedeckung und -schichtung, sowie 3.7, 10.8 und 12.0 μm jeweils für die Temperatur der Meeresoberfläche) sowie drei Kanälen im sichtbaren bzw. nahen IR (0.55 μm für den Nachweis von Chlorophyll, 0.66 und 0.87 μm für den Vegetationsindex) arbeitet. Das räumliche Auflösungsvermögen beträgt 1 km² bei einer Streifenbreite von 500 km, das Temperaturauflösungsvermögen für die Temperatur der Meeresoberfläche soll besser als 0.3 K sein. Eine Instrumentbeschreibung findet sich unter <http://esapub.esrin.esa.it/bulletin/bullet103/besy103.pdf>, das Datenhandbuch unter <http://envisat.esa.int/dataproducts/aatsr/CNTR.htm>.

SCIAMACHY (Scanning Imaging Absorption Spectrometer for atmospheric Cartography) ist ein passives, auf der Basis von GOME entwickeltes Spektrometer zur Messung von Spurengasen in der Troposphäre und Stratosphäre. Es überdeckt den Spektralbereich von 240 bis 2400 nm und verwendet parallel zwei Sounding Techniken: Limb-Sounding unter Verwendung von Sonne bzw. Mond als Lichtquellen und Messung des zurück gestreuten Lichts. Der Spektralbereich ist in acht, den verschiedenen zu untersuchenden Spurengasen angepasste Bänder unterteilt. Im Absorptionsmodus wird eine differentielle Technik verwendet wie sie auch für die Ozonmessung vom Erdboden aus benutzt wird: man vergleicht die Intensitäten in zwei Spektralbändern, von denen das eine durch Absorption an dem zu untersuchenden Spurengas nicht beeinflusst ist, das andere dagegen wohl. Die wissenschaftlichen Ziele von SCIAMACHY umfassen die globale Konzentration und Verteilung von Spurengasen in der Tropo- und Stratosphäre, die Messung von Aerosolen, Wolkenhöhen und des spektralen Reflektionsvermögens von Oberflächen, die Messung der troposphärischen Spurengase O₃, O₂, O₄, NO₂, N₂O, CO, CO₂, CH₄ und H₂O sowie unter bestimmten Bedingungen HCHO, SO₂ und NO₃, sowie die Messung der stratosphärischen Spurengase O₃, O₂, NO, NO₂, NO₃, N₂O, CO, CO₂, CH₄ und H₂O sowie im Ozonloch OCIO und ClO. Die Datenprodukte werden beschrieben unter <http://envisat.esa.int/instruments/sciamachy/data-app/dataprod.html>, die SCIAMACHY-Homepage der Uni Bremen ist unter <http://www.iup.physik.uni-bremen.de/sciamachy/index.html> zu finden.

4.5 Small is beautiful: CHAMP und Diamant

CHAMP (Challenging Minisatellite Payload, http://op.gfz-potsdam.de/champ/index_CHAMP.html) mit dessen Produkten wir, im Gegensatz zu denen der bisher beschriebenen Missionen, als Normalverbraucher wahrscheinlich kaum in Berührung kommen werden. CHAMP

Band No	Band center (nm)	Commercial applications	Science applications
1	415	Water quality mapping road detection, land cover, urban areas	Desert soils, biogenic crusts, dissolved organic material
2	445	Real natural blue color for mass media, water quality, urban areas, landcover,	Low chlorophyll concentration, maximum chlorophyll-A absorption at 450 nm
3	490	Water monitoring, road detection, lithology and soils, crops and farming	Non-linear chlorophyll, medium chlorophyll concentration
4	510	Soils, landcover, surveying, water quality, coastal zones	Low suspended sediment, high pigment concentration
5	555	Real natural green color for mass media, landcover, crops and farming	Green peak of chlorophyll, null band of chlorophyll variation
6	645	Real natural red color for mass media, landcover, lithology and soils, forestry	Chlorophyll absorption, vegetation indices, high suspended sediment
7	670	Crops and farming, pest control, security	Max chlorophyll absorption, vegetation indices, 2nd peak of chlorophyll-A absorption
8	700	Vegetation general, nature conservation	Red edge of vegetation
9	740	Vegetation general, nature conservation	Red edge of vegetation
10	790	Vegetation general, nature conservation	Red edge of vegetation
11	865	Lithology and soils, crops and farming, forestry, landcover, nature conservation	Max. chlorophyll reflectance, vegetation indices, water vapor, atmospheric correction
12	905	Lithology and soils, nature conservation, wetlands, coastal waters	Max. chlorophyll reflectance, atmospheric correction

Tabelle 4.5: Spektralkanäle MSRS und Anwendungsbeispiele

ist ein Satellit des Geoforschungszentrums Potsdam, der sich mit der genauen Vermessung des Schwerefeldes der Erde und des geomagnetischen Feldes befasst. Mit Hilfe seines GPS Empfängers führt er ferner Okkultationsexperimente durch, d.h. er untersucht auch die Atmosphäre, insbesondere die Elektronendichte der Thermosphäre. Das Teilcheninstrument auf CHAMP ist leider während des Starts beschädigt worden und hat keine sinnvollen Daten geliefert.

CHAMP ist interessant insofern, als dass er am anderen Ende des Spektrums von Satellitengrößen liegt: während alleine der Scanner des ETM+ auf LandSat 7 eine Masse von 318 kg und eine Leistungsaufnahme von 590 W hat, hat der gesamte Satellit CHAMP lediglich eine Masse von 522 kg und eine Leistungsaufnahme der Nutzlast von nur 50 W.

Auch ein Spektrometer lässt sich mit einem kleinen und billigen Satelliten realisieren, z.B. das MSRS (Multi-Spectral High Resolution System, http://www.fuchs-gruppe.com/eo/msrs/0_MSRS.html) auf Diamant. Das Instrument soll Anwender-orientierte Daten liefern. Dazu werden 12 Spektralkanäle im sichtbaren und nahen IR ausgewertet mit einer Bodenauflösung von 5 m, einem Blickfeld von $26 \times 40 \text{ km}^2$ und hoher radiometrischer Auflösung. Die einzelnen Spektralkanäle und Anwendungsbeispiele sind in Tabelle 4.5 gegeben. Statt eines Scanners wird ein Array aus CCDs verwendet, das Scannen erfolgt dann wie bei der Kehrbesentechnik.

Das ganze Raumfahrzeug ist mit einer Masse von weniger als 250 kg deutlich leichter als nur der Scanner des ETM+ und hat auch nur eine mittlere Leistungsaufnahme von 55 W. Die großen Unterschiede zwischen diesen beiden Instrumenten haben zwei Ursachen: das größere Blickfeld von LandSat 7 auf Grund des Scanners statt der CCD ist ein wesentlicher instrumenteller Unterschied. Damit zusammen hängt der zweite Grund: HSRS basiert auf neuer Technologie (vulgo: die fliegen einfach kleine Digitalkameras) während LandSat eine Weiterentwicklung einer eher fossilen Technik trägt. Die unterschiedlichen Ansätze haben ihre Gründe: Diamant ist ein kommerzielles, anwendungsorientiertes System für ganz spezielle aktuelle Fragen eines Anwenders (z.B. Ernteüberwachung), während LandSat als Monitoring-System die Kontinuität erhalten muss: heutige Datensätze können nur Veränderungen gegenüber älteren Datensätzen zeigen, wenn sie mit vergleichbaren Instrumenten aufgenommen wurden oder wenn zumindest die Instrumente zur wechselseitigen Kalibrierung eine gewisse Zeit gemeinsam im Orbit waren.

4.6 Lost and found: GPS-NAVSTAR

Das Global Positioning System (GPS) ist nicht nur für die Positionsbestimmung auf der Erde wichtig sondern auch für die Positionsbestimmung von Satelliten: was nützt die Angabe

der Höhe von SeaSat über dem Ozean zur Bestimmung des Geoids mit einer Genauigkeit von 1 cm wenn die Position von SeaSat nur mit einer Genauigkeit von 150 m bekannt ist. Außerdem ist GPS vielleicht auch das am häufigsten zweckentfremdete System: da Radiosignale durch den Wasserdampfgehalt der Atmosphäre beeinflusst werden und gleichzeitig die GPS-Signale genau bekannt sind (insbesondere ihr Timing) lässt sich ein GPS-Empfänger auch zur Messung atmosphärischer Bedingungen missbrauchen.

GPS wird durch ein Netz von 21 NAVSTAR (NAVigations System with Time And Ranging) Satelliten realisiert mit Flughöhen von 20 200 km und einer Inklination von 55° : die ursprünglich geplante Inklination von 63° hätte eine Überdeckung auch der hohen Breiten ermöglicht und das System damit wirklich global gemacht, die 55° wurden gewählt, damit die NAVSTARs auch vom Shuttle gestartet werden können.

Das NAVSTAR-Programm wurde 1973 vom DoD initiiert, der erste Satellit startete 1978, erst 1992 war das System vollständig. Seitdem werden immer wieder neue NAVSTAR-Satelliten gestartet (Anfang 2001 der 54.), da alte Satelliten ausfallen bzw. durch neuere Modelle ersetzt werden.

Den ursprünglichen GPS-Signale für den zivilen Gebrauch war eine Störung überlagert, die die Genauigkeit in der Ortsbestimmung auf ca. 100 m reduzierte. Dieses Störsignal wurde Mitte der 90er Jahre entfernt, so dass heutzutage auch für zivile Anwender eine Genauigkeit im Bereich von einigen Metern (optimale Empfangsbedingungen) bis zu wenigen 10 m (ungünstiger Empfang) erreicht werden kann.² Das gilt allerdings nur für die horizontale Position, bei der vertikalen Position haben normale Empfänger Ungenauigkeiten im Bereich von etlichen 10 bis ca. 100 m – die Koordinate wird mit der geringsten Genauigkeit bestimmt, da man davon ausgeht, dass sich der Anwender letztendlich nicht im 3D bewegt sondern auf einer gekrümmten Fläche, d.h. sein Aufenthaltsort durch zwei Koordinaten bestimmt ist.

Die Genauigkeit eines GPS-Empfängers lässt sich auf Ortsgenauigkeiten im cm-Bereich verbessern, wenn man zusätzlich zu den Satelliten eine ortsfeste Referenzstation verwendet. Dieses Verfahren wird als differentielles GPS (DGPS) bezeichnet, es wird u.a. bei den Vermessungsämtern verwendet sowie bei GPS-basierten Navigationssystemen von Rettungsdiensten.

4.7 Mission to Planet Earth

Bei dem Projekt ‘Mission to Planet Earth’ (http://www.earth.nasa.gov/flash_top.html und <http://spacelink.nasa.gov/Instructional.Materials/NASA.Educational.Products/Our.Mission.to.Planet.Earth/>) handelt es sich um einen Versuch der NASA (und ursprünglich auch der ESA) möglichst viele Satelliten-Beobachtungen der Erde zusammenzufassen. Dabei sollen die Instrumente aufeinander abgestimmt sein, so dass sie einander möglichst perfekt ergänzen und auch in ihren zeitlichen und räumlichen Datenrahmen aufeinander abgestimmt sind, so dass eine Synchronisation verschiedener Instrumente überhaupt erst möglich wird. Ziel dieser Mission ist die Entwicklung eines möglichst vollständigen Bildes der dynamischen Prozesse auf der Erde. Dabei hat auch die Frage nach dem Einfluss des Menschen auf seine Umwelt eine zentrale Bedeutung. Die Instrumente, die bei diesem Projekt berücksichtigt werden, umfassen das komplette Spektrum von Instrumenten, die für ‘remote sensing’ entwickelt wurden.

In einem gewissen Sinne ist Mission to Planet Earth auch Konkurrenzprojekt zu dem Einzelsatelliten EnviSat der ESA. Während EnviSat erfolgreich gestartet wurde und Daten liefert, leiden viele Satelliten des Mission to Planet Earth Programms unter Unterfinanzierung bzw. es wurde ein Satellit gestartet aber der Satellit, der die ergänzenden Instrumente trägt, nicht oder nicht rechtzeitig: teilweise auf Grund politischer Probleme (wie viel Umwelt braucht man eigentlich), teilweise auf Grund der fehlenden Fokussierung der einzelnen Projekte und Satelliten, da viele Instrumente nur deshalb vorgeschlagen wurden, weil auch

²Wofern die USA nicht gerade damit beschäftigt sind, einen Schurkenstaat in seinem Benehmen einzunorden und allen anderen Schurken die Möglichkeit nehmen wollen, sich mit Hilfe von GPS genau zu orientieren oder gar einzumischen.

diese bestimmte Gruppe ein Instrument bauen wollte. Die folgende Übersicht hat daher eher historische Bedeutung, die Grundidee des Projekts ist populär in Baker [3] zusammen gefasst, die Grundidee und der aktuelle Stand findet sich unter <http://spacelink.nasa.gov/Instructional.Materials/NASA.Educational.Products/Our.Mission.to.Planet.Earth/>

4.7.1 Zielsetzung

Grob gesagt soll ‘Mission to Planet Earth’ alle Fragestellungen bezüglich der Erde, ihrer Untersysteme wie Biosphäre, Hydrosphäre oder Geosphäre, und die Wechselwirkungen zwischen diesen Teilsystemen umfassen. Dazu sind die Zielsetzungen in zwei Zeitskalen unterteilt: geologische Zeitskalen im Bereich von Tausenden und Millionen von Jahren. Das sind im wesentlichen die Prozesse, die die Erde geformt haben bzw. externe Einflüsse, die sich auf derartig langen Zeitskalen abspielen, z.B. klimatische Veränderungen durch Schwankungen der Erdbahnparameter (Milankovich-Theorie) oder Veränderungen des solaren Energieoutputs. Die andere Zeitskala umfasst Perioden von Jahrzehnten oder Jahrhunderten. Dazu gehören auch die anthropogenen Einflüsse auf das Klima und die Biosphäre (Spurengase, CO₂, Ozonloch, Abholzung des tropischen Regenwaldes, Flächenversiegelung).

Die wesentliche Zielsetzung ist also, die Erde als etwas veränderliches zu erkennen, die Zeitskalen dieser Veränderungen zu beschreiben und zu den durch den Menschen bewirkten Veränderungen in Beziehung zu setzen.

Die wesentlichen Anstöße zu den langfristigen Forschungen kommen aus der Geophysik und Geologie, als wesentliche Stichworte seien hier erwähnt Prozesse im Kern und Mantel der Erde, die ihrerseits wieder Auswirkungen auf das Erdmagnetfeld (dessen Veränderungen wir an den Ablagerungen in der Nähe der Tiefssegräben verfolgen können) und durch Ausgasungsprozesse auch auf die chemische Zusammensetzung der Atmosphäre haben (heutige Vorstellungen über die Entwicklung des Planetensystems gehen davon aus, dass Atmosphäre und Hydrosphäre durch Vulkanismus aus der noch heißen, jungen Erdkugel entstanden sind; die sich dabei ausbildende ‘Uratmosphäre’ war ursprünglich völlig lebensfeindlich, zumindest für heutiges Leben, die ersten Lebensformen begannen dann jedoch, die Atmosphäre zu verändern, so dass sich in einem ca. 3.8 Milliarden Jahre dauernden Wechselspiel zwischen Leben³ und Atmosphäre die heutige Atmosphäre entwickelt hat). Auch die Ausbildung der Topographie ist über die Kontinentaldrift und Vulkanismus mit den Vorgängen im Erdinnern verbunden. Die Bedeutung der Kontinentaldrift für das Klima auf der Erde und auch die Entwicklung von Leben ist nicht zu vernachlässigen, so können z.B. Eiszeiten nur dann entstehen, wenn mindestens eine Polregion landbedeckt ist oder, wie im Falle des Norpols, so von Land umschlossen ist, dass die warmen Meeresströmungen nicht bis in diese Bereiche gelangen können [80, 107, 113]. Veränderungen der Lage der Kontinente bewirken natürlich auch Veränderungen in den globalen Zirkulationsmustern (sowohl Meeres- als auch atmosphärische Zirkulation). Zusätzlich sind auf diesen langen Zeitskalen auch Veränderungen externer Parameter wie z.B. der Einstrahlung der Sonne zu berücksichtigen

Die meisten anderen Wissenschaftsbereiche befassen sich eher mit den kurzzeitigen Veränderungen, über die Abb. 4.1 einen Überblick gibt. Dabei sind die Größen $n(x)$ und $\phi(x)$ jeweils Konzentrationen bzw. Flüsse einer Größe x . Beispiele sind die Glazialogie, die das Schnee- und Eisvorkommen auf der Erde kartographiert, dazu gehören auch Veränderungen in der Eisbedeckung der Antarktis sowie Gletscherschwankungen. Diese Untersuchungen geben Aufschlüsse über lokale und globale Veränderungen von Lufttemperatur und Zirkulationssystemen. Ein weiterer Forschungspunkt ist in diesem Zusammenhang die Dynamik schneller Gletschervorstöße. Alle diese Fragen sind auch mit der Frage nach klimatischen Änderungen und den möglichen anthropogenen Einflüssen verbunden, so können Gletscher z.B. als eine

³Die größte Umweltverschmutzung war der Übergang des Lebens vom Meer auf das Land: die Uratmosphäre enthielt keinen Sauerstoff und damit auch kein Ozon, so dass biologische Materie sofort durch die solare UV-Strahlung zerstört wurde. Erst als die Pflanzen im Wasser soviel Sauerstoff ausgeatmet hatten, dass sich eine dünne Ozonschicht bilden konnte, konnten sie auch die Landflächen besiedeln und durch erhöhte Bioproduktivität den Sauerstoff in die Atmosphäre eintragen und damit auch die Ozonschicht ermöglichen, die zur Entwicklung höherer Lebewesen erforderlich war.

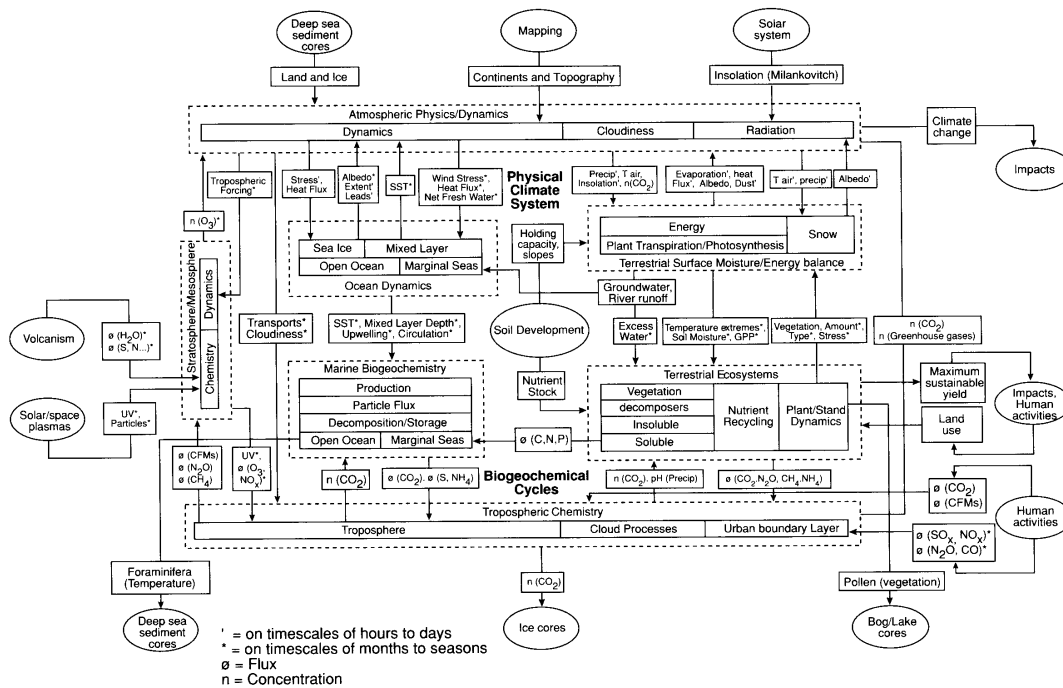


Abbildung 4.1: Modell des Systems Erde und der auf Zeitskalen von Jahrzehnten und Jahrhunderten wichtigen Prozesse [96]

Art globales Thermometer betrachtet werden, haben jedoch ihrerseits durch Veränderungen der lokalen Strahlungsbilanz auch wieder Rückwirkungen auf die Prozesse, die sie zum Wachsen oder Schrumpfen brachten (positive und negative Feedback-Mechanismen [55]).

Nicht nur mit den Gletschern sondern ebenso mit der Biosphäre verbunden ist der Wasserkreislauf als Forschungsgebiet der Hydrologie. Hierzu gehören die regionale und globale Bestimmung von Niederschlag und Verdunstung (Feuchtebilanz), insbesondere auch die Frage, welche Merkmale diese Niederschlagsmuster bestimmen. Interessant ist hier der tropische Regenwald, da er über dem Festland die größten Niederschlagsmengen erhält. Der Wasserkreislauf beeinflusst jedoch über die regionale Niederschlagsverteilung nicht nur die Biosphäre (Versteppung bei Wassermangel, erhöhte Produktion an Biomasse bei ausreichend Wasser, aber auch Schäden bei Wasserüberschuss) sondern auch über die Erosion den geochemischen Kreislauf und über die Bildung von Schnee-, Eis-, und Wasserflächen die Strahlungsbilanz und damit auch die Temperatur der Erde, letzteres gilt ebenso für die Bewölkung und die Bedeutung der unterschiedlichen Wolken für die Strahlungsbilanz der Erde [42, 90]. Verdunstung und Niederschlag sind ebenfalls für einen Teil des Wärmetransports verantwortlich.

In direktem Zusammenhang damit steht die Atmosphärenforschung. Wesentliche Ziele sind das Verständnis der großräumigen Zirkulationsmuster, sowie der Bewölkung und der Luftfeuchtigkeit. Zusätzliche Bedeutung gewinnen in der Atmosphärenforschung immer mehr die Abschätzungen der Auswirkungen anthropogener Einflüsse. Dazu zählen insbesondere Messungen der Konzentration und Höhenverteilung von Ozon, Aerosolen und CO_2 sowie anderer klimarelevanter Spurengase. Gerade die Atmosphäre kann nicht als isoliertes System betrachtet werden, sondern steht in vielfältiger Wechselwirkung mit allen anderen Systemen, z.B. durch Eintrag von Aerosolen bei Vulkanausbrüchen (Wechselwirkung mit der Lithosphäre), klimatische Auswirkungen des Eintrages von Wasser in Form von Wolken oder Wasserdampf (WeWi mit der Hydrosphäre), Veränderung von Verdunstung und Zirkulation in Gegenwart großer Eisflächen (Kryosphäre), und nicht zuletzt die Wechselwirkung mit der Biosphäre (Senken für CO_2 , Quellen für O_2 , Veränderung der bodennahen Winde durch Veränderung des Bodenbewuchses).

Damit kommen wir zum letzten entscheidenden Bestandteil des Systems Erde, der Biosphäre. Die direktesten Anwendungen von Satellitenbeobachtungen auf die Biosphäre liegen in der Forst- und Landwirtschaft, hier insbesondere in Überblicken über Flächennutzung und den Zustand der Vegetation (Reifezustand ebenso wie Schädigungen, s.o.). Die wichtigste Anwendung von Satellitenbeobachtungen der Biosphäre dürfte in den globalen Fragen liegen, z.B. nach der Biomasse insgesamt, insbesondere auch dem Plankton im Meer, das für die Speicherung von CO₂ eine nicht unerhebliche Rolle spielt. Aber auch einzelne Organismenfamilien können vom Satelliten aus kartographiert werden, so dass diese eventuell als Bioindikatoren für Auswirkungen von Klimaänderungen dienen können oder durch auffälliges Verhalten auf bisher nicht beachtete Zusammenhänge im System Erde hinweisen können. Als Beispiel für diese Indikatorfunktion sei der Rückgang des Phytoplanktons in der Nähe der Antarktis aufgrund des Ozonloches erwähnt (vgl. Geoskop, April 1992). Dies ist einerseits der erste direkte Hinweis auf die Folgen des Ozonlochs für die Biosphäre, andererseits wird hieran auch eine Kopplung der vielen Systeme und ihrer Einflüsse auf das Klima deutlich: Das Ozonloch scheint das Phytoplankton zu reduzieren, damit wird aber eine CO₂-Senke geschwächt, d.h. der Kohlendioxidgehalt in der Atmosphäre steigt relativ stark → das Ozon trägt demnach (abgesehen von möglichen direkten Eingriffen in die Strahlungsbilanz der Erde) auch sehr indirekt zum Treibhauseffekt bei.

Die Kenntnisse derartiger Zusammenhänge und die Messungen von Konzentrationen und Flüßen sind für die Güte unserer Klimamodelle von entscheidender Bedeutung, vgl. dazu den Artikel von Bolle [11], der ausführlich auf die Probleme und Defizite von Klimamodellen eingeht, ebenso wie die Artikel von Schneider [105] und White [127]. Diese Klimasimulationen rücken zunehmend in das Blickfeld, da zwar allmählich allgemein akzeptiert wird, dass ein CO₂-Anstieg in der Atmosphäre zu globaler Erwärmung führen wird, jedoch können die Folgen einer derartigen Erwärmung möglicherweise völlig unerwartet ausfallen. So sind in letzter Zeit einige neuere Modellrechnungen aufgetaucht, die z.B. für Europa eine neue Eiszeit infolge des Treibhauseffektes prophezeien, da die anfangs geringe Erhöhung der Temperatur die globalen Zirkulationssysteme derart verändern würde, dass der Golfstrom nach Süden abgelenkt wird. Damit wird Europa aber eine entscheidende Wärmequelle entzogen, der Schnee der dann kälteren Winter würde länger liegen bleiben, damit aber auch die Albedo erhöhen und so zu einer weiteren Abkühlung führen (es wird mehr reflektiert, also weniger absorbiert). Der Effekt wäre, dass Europa innerhalb größenordnungsmäßig eines Jahrhunderts vergletschern könnte (z.B. [50, 57] und Zitate darin).

Earth Observing System EOS

Das hierfür erforderliche System von Beobachtungssatelliten und Instrumenten gliedert sich in zwei Teile, die geplanten EOS-Satelliten (Earth Observing System [3]), die alleine jeweils 28 Instrumente tragen sollen, sowie viele Einzelmisionen unter verschiedenen Gesichtspunkten, dazu gehören auch Beobachtungen vom Shuttle aus.

Das amerikanische Kernstück der Mission to Planet Earth ist das Earth Observing System EOS. EOS soll aus zwei Satelliten EOS-A und EOS-B bestehen, beide sollen in polare Umlaufbahnen gestartet werden. Die Plattformen werden eine Länge von 16 m und eine Breite von 4 m haben, das Gewicht wird bei 14 t liegen. Die Instrumentation für EOS-A (<http://terra.nasa.gov/>) ist festgelegt, die Instrumentation von EOS-B ist noch nicht festgelegt, je nach den Erfahrungen mit EOS-A kann der zweite Satellit entweder eine identische Instrumentierung bekommen oder aber nach völlig anderen Fragestellung ausgerüstet werden. Die Instrumente auf EOS-A sind ausgelegt auf die Untersuchung der folgende Fragestellungen:

- hydrologischer Zyklus (Verdunstung, Niederschlag, Transport von Wasserdampf, Wolken)
- Strahlungsbilanz und Oberflächentemperaturen,
- Atmosphärische Zirkulation,
- biologische Aktivität und Ökosystem,
- Oberflächenmineralogie.

Für den Fall, dass EOS-B nicht eine zu A identische Instrumentierung erhält, sind für ihn Instrumente mit folgenden Messaufgaben geplant:

- Stratosphäre: Ozon, Chemie, und Wind
- globale Beobachtung der troposphärischen Chemie,
- ozeanische und atmosphärische Zirkulation, insbesondere Oberflächenwinde,
- Glaziologie,
- Erdbeben.

Das Gesamtsystem zur Erdbeobachtung wird noch um europäische und japanische Plattformen erweitert, die Informationen bereitstellen sollen über

- atmosphärische Zusammensetzung, Wasserdampf, Wolken und Aerosole
- Vorkommen von Mineralien und Pflanzenarten für Geologie, Land- und Forstwirtschaft,
- Temperatur der Meeresoberfläche und Windsysteme.

Die auf diesen Satelliten zu verwendenden Instrumente gehören einer ganz anderen Generation an als die bisher diskutierten Instrumente (z.B. LandSat). Für einige der für EOS-A geplanten Instrumente wurde in Abschn. 3.2.9 eine kurze Übersicht über die Aufgaben und Fähigkeiten gegeben.

Literatur

Alles was kreucht und fleucht ist in Kramer [62] beschrieben, für die einzelnen Missionen finden sich auf den entsprechenden Homepages weitere Literaturhinweise.

Kapitel 5

Kommunikation

Die Kommunikation zwischen einem Satelliten und seiner Bodenstation zerfällt in zwei Bereiche: (a) operationale Daten und (b) Nutzdaten. Die operationalen Daten umfassen die Vitalparameter des Satelliten, wie Temperaturen, Lage, Stromversorgung und Bahnparameter, sowie von Seite der Bodenstation die Befehle zur Lagekontrolle, zum An- und Ausschalten von Instrumenten, zum Senden der Daten usw. Operationale Daten müssen in Echtzeit und möglichst ohne Fehler transmittiert werden, die zu sendenden Datenmengen sind gering. Der größte Teil der Kommunikation ist einseitig, das Senden der Nutzdaten vom Satelliten zum Boden. Hierbei fallen sehr große Datenmengen an, die möglichst fehlerfrei versendet werden sollen, aber in der Regel nicht notwendigerweise in Echtzeit – was dann jedoch Datenspeicher erfordert.

Der Schwerpunkt dieses Kapitel liegt auf der Übermittlung von Daten über unsichere Kommunikationskanäle mit geringer Datenübertragungskapazität, die wichtigsten Themenbereiche sind Datenreduktion bzw. Kompression und Platz sparende Kodierung der Daten, bei der gleichzeitig eine Fehlererkennung und möglichst auch -korrektur möglich ist.

5.1 Historisches

Bei interplanetaren Missionen und Missionen zu anderen Planeten wurden die Messverfahren so gewählt, dass die Daten in digitalisierter Form vorlagen und damit direkt an die Bodenstation übermittelt werden konnten. Dadurch ist die Qualität des Datenendprodukts häufig begrenzt, zum einen durch die geringe Datenübertragungsrate, zum anderen durch die Notwendigkeit, Instrumente zu verwenden, die elektronische Daten produzieren obwohl diese nicht zwingend die besten sein müssen (vgl. Videokamera und konventionelle Kamera wie in Abschn. 3.2.7 diskutiert).

In der Erdfernerkundung wurden zwei Sonderwege beschritten, um möglichst optimale Datensätze zu erhalten. Dabei sollte nicht vergessen werden, dass die militärische Aufklärung eine der wesentlichen Motivationen der Erdfernerkundung war. Das bedeutet einerseits eine hohe Anforderung an das Datenprodukt, andererseits aber auch große finanzielle, materielle und personelle Ressourcen.

5.1.1 Rückführung des Filmmaterials

In der Frühzeit der Erdfernerkundung von Raumfahrzeugen aus war das dominierende Instrument die konventionelle Photokamera mit einem Schwarz-Weiß-Film, als Erweiterung die Kombination von SW-Kameras zur Multispektralaufnahme oder die Verwendung von Farbfilmen. In jedem Fall war das Endprodukt auf dem Raumfahrzeug ein belichteter Film. Bei einer bemannten Raummission (z.B. Apollo) war dies kein Problem, da der Film mit den Astronauten zur Erde zurück gebracht wurde. Bei unbemannten Missionen adoptierte man

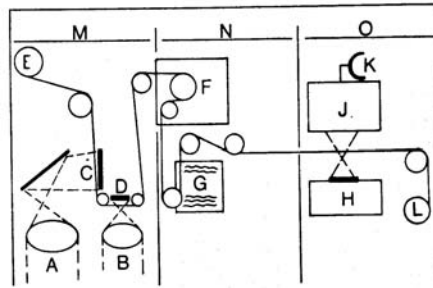


Abbildung 5.3: Die Daten werden weiterhin mit Film und Kamera gewonnen, der Film wird an Bord des Satelliten entwickelt und gescannt und die so gewonnenen digitalen Daten werden transmittiert [19]

5.1.2 Onboard-Scans

Beide Verfahren, Abwurf der Filmkapsel und kontrollierter Absturz des Satelliten, sind nicht nur kostenintensiv sondern haben einen weiteren Nachteil, der für die Militärs noch wesentlich gravierender ist als für zivile Nutzer: Echtzeit-Aufklärung ist mit dieser Methode der Datenrückführung nicht möglich.

Auf Grund der hohen Qualität der photographischen Aufnahmen hat man daher relativ früh ein hybrides Verfahren entwickelt: die eigentliche Aufnahme (der Sensor) erfolgt weiterhin analog (Kamera und Film), die Datenübertragung dagegen digital. Voraussetzung dafür ist die Umwandlung des analogen Signals in ein elektronisches Signal an Bord des Satelliten.

Dazu wurde das in Abb. 5.3 dargestellte Verfahren entwickelt: der Film wird an Bord des Satelliten entwickelt und anschließend gescannt. Damit liegt die Information in digitalisierter Form vor und kann zur Bodenstation transmittiert werden. Damit ist nahezu Echtzeit-Aufklärung möglich, allerdings ist die Datenqualität schlechter als die des Originalfilms, so dass dieser zumindest bei Bedarf dennoch mit der Kapsel zurück geführt wurde.

5.1.3 Technische Anmerkungen

Heutige Erdfernerkundungssatelliten erzeugen nur noch digitale Daten, die transmittiert werden.³ Die technischen Probleme bei der Kommunikation mit Satelliten lassen sich daher reduzieren auf die Punkte:

- die anfallenden Datenmengen sind groß, die Übertragungsraten dagegen klein. Damit ergibt sich die Notwendigkeit der Datenkompression (Quellencodierung).
- die Kommunikationskanäle sind gestört, z.B. Störungen der atmosphärischen Ionisation durch solare Eruptionen, Dämpfung in der Troposphäre durch starken Regen oder hohen Wasserdampfgehalt). Damit ergibt sich die Notwendigkeit, die Daten so zu codieren, dass eine Fehlererkennung und gegebenenfalls -korrektur möglich wird (Kanalcodierung).
- Verfügbarkeit einer Boden- oder Relais-Station oder gegebenenfalls Datenspeicherung.

Der Schwerpunkt dieses Kapitels wird in den Bereichen Datenkompression und Codierung mit Fehlererkennung und -beseitigung liegen.

5.2 Information

Kommunikation lässt sich definieren als der Austausch von Information zwischen Systemem, Teilsystemen und Systemelementen, die in der Lage sind, Informationen zu erzeugen, aufzunehmen, zu verarbeiten und zu speichern. Ein wesentlicher Bestandteil der Kommunikation

³Es geht zwar das Gerücht, dass zumindest die amerikanischen Späher wie Big Bird und die KH (Keyhole)-Serie immer noch mit der Möglichkeit ausgerüstet sind, Filme zurück zu senden – selbst wenn dem so ist, ist es für unsere Betrachtungen unerheblich, da es eine Sonderlösung ist.

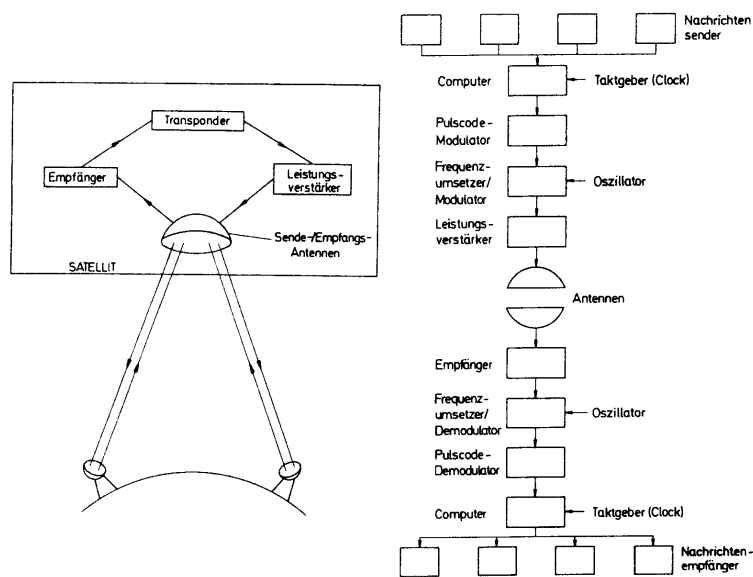


Abbildung 5.4: Nachrichtenstrecke bei Nachrichtensatelliten (links) und Blockdiagramm einer typischen Telemetrie-/Telekommando-Strecke (rechts, [95])

ist daher der Umgang mit Information. Daher ist es sinnvoll, sich erst einmal mit den Grundlagen der Informationstheorie zu beschäftigen, bevor wir zur Kommunikation übergehen.

5.2.1 Kommunikation

In diesem Übersichtsabschnitt möchte ich einige Begriffe einführen und einige generelle Anmerkungen zur Kommunikation machen. Dies ist wichtig, da einige umgangssprachliche Bedeutungen nicht immer vollständig mit der informationstheoretischen Definition der Begriffe in Einklang stehen müssen. Beginnen wir dazu mit dem Begriff der Kommunikation. Bei Kommunikation handelt es sich, wie bereits eingangs festgestellt, um die Übertragung von Information. Dieser Informationsaustausch kann auf zwei prinzipiell unterschiedliche Weisen erfolgen:

1. Passive Übertragung von Information (z.B. Kommunikationssatelliten, Telefon, Fernsehen, TDRS: Tracking and Data Relay System): Hierbei wirkt der übertragende Satellit als Relaisstation, interpretiert die Signale aber nicht, sondern empfängt sie und sendet sie nach einer Frequenzumsetzung an eine andere Station zurück (vgl. Abb. 5.4, links). Diese Frequenzumsetzung mittels eines Transponders ist notwendig, damit die Empfangsdaten trotz des großen Pegelunterschiedes zu den Sendedaten vom Satelliten einwandfrei empfangen werden können. Ein wesentlich einfacheres Beispiel wären die Echo-Ballon-Satelliten, bei denen die Signale lediglich passiv vom Satelliten reflektiert wurden (also ohne Frequenzumsetzung und Verstärkung).
2. Datenübertragung (z.B. Messdaten von Instrumenten ebenso wie Telekommandos an Satelliten, vgl. Abb 5.4, rechts), beides setzt Ver- und Entschlüsselung von Informationen voraus.

Die passive Übertragung ist eine rein technische Frage. Die die Information übertragenden Satelliten werden zwar als Kommunikationssatelliten bezeichnet, sie nehmen jedoch im eigentlichen Sinne gar nicht an der Kommunikation teil, da sie den Inhalt der zu übertragenden Information nicht erkennen, sondern nur kopieren. Im informationstheoretischen Kontext sind Kommunikationssatelliten lediglich ein Bestandteil des Übertragungskanal (s.u.).

Zum besseren Verständnis dieser Aussage ist in Abb. 5.5 ein Kommunikationssystem symbolisch dargestellt. Die Grundbestandteile eines Kommunikationssystems sind eine Nachricht-

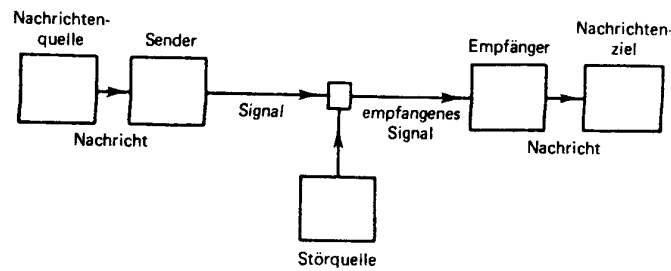


Abbildung 5.5: Kommunikationssystem [112]

tenquelle, ein Übertragungskanal und eine Nachrichtensenke. Nachrichtenquelle und -senke werden mit den Mitteln der Informationstheorie beschrieben, wobei die ursprüngliche Nachricht vor der Abgabe an den Sender bzw. nach dem Auffangen durch den Empfänger noch codiert bzw. decodiert werden kann (Quellencodierung). Der Übertragungskanal selbst sollte die Nachricht unmodifiziert weitergeben, hier führen jedoch Störquellen zu einer Modifikation. Um die durch diese Störquellen bewirkten unsystematischen Änderungen der Nachricht zu identifizieren und die Fehler teilweise sogar wieder zu kompensieren, verwendet man spezielle Codierungsverfahren (Kanalcodierung, vgl. Abschn. 5.3.2).

Betrachten wir das Kommunikationssystem in Abb. 5.5, so können wir daran sofort die Fragen verdeutlichen, die wir in den folgenden beiden Kapiteln behandeln werden:

- Was ist Information? Wie misst man den Betrag der Information?
- Wie misst man die Kapazität eines Übertragungskanals?
- Die Übersetzung der Nachricht in das Signal durch den Sender beinhaltet einen Codiervorgang. Was sind die charakteristischen Merkmale eines effizienten Codiervorganges? Und wenn die Codierung so effizient wie möglich ist, mit welcher Übertragungsrate kann der Kanal Informationen weiterleiten?
- Was sind die allgemeinen Merkmale der Störungen? Wie beeinflussen diese Störungen die Genauigkeit der Nachricht, die schließlich das Ziel erreicht? Wie kann man die unerwünschten Effekte der Störungen auf ein Minimum beschränken und bis zu welchem Grad können sie ausgeschaltet werden?
- Wenn das zu übertragene Signal kontinuierlich ist (z.B. Musik) und nicht aus diskreten Zeichen besteht, wie beeinflusst diese Tatsache das Problem?

Bisher haben wir Kommunikation eher unter dem technischen Aspekt betrachtet. Eine vollständige Beschreibung der Kommunikation muss jedoch alle drei Ebenen der Kommunikation berücksichtigen [112]:

- (A) Wie genau können die Zeichen der Kommunikation übertragen werden (technisches Problem)?
- (B) Wie genau entsprechen die übertragenen Zeichen der gewünschten Bedeutung (semantisches Problem)?
- (C) Wie effektiv beeinflusst die empfangene Nachricht das Verhalten des Empfängers in der gewünschten Weise (Effektivitätsproblem)?

Die technische Ebene (A) wird in den folgenden Abschnitten erläutert werden. Das semantische Problem betrifft die Übereinstimmung der Interpretation der Nachricht beim Empfänger verglichen mit der Bedeutung, die der Sender der Nachricht geben wollte. Selbst in dem alltäglichen Problem der Kommunikation durch Sprache kann dies zu verwickelten Situationen und Missverständnissen führen. Shannon und Weaver [112] illustrieren diesen Punkt mit Hilfe der Personen X und Y: Wenn X nicht zu verstehen scheint, was Y sagt, so ist es theoretisch nicht möglich, dass diese Situation, solange Y weiterhin nur mit X redet, geklärt werden kann. Wenn Y fragt, ‘Verstehen Sie mich jetzt?’ und X erwidert ‘Ja, natürlich’, so ist das nicht unbedingt ein Beweis dafür, dass die Verständigung erreicht wurde. Es kann

ja sein, dass X die Frage nicht verstanden hat. Falls das so lächerlich klingt, stellen Sie sich folgenden Dialog vor: ‘Czy pañ mnie rozumie?’ mit der Antwort ‘Hai wakkate imasu.’.

Das Effektivitätsproblem ist in einem gewissen Sinne mit dem semantischen Problem verbunden. Grundsätzlich handelt es sich dabei um die Frage nach dem Erfolg, mit dem eine vom Empfänger erhaltene Nachricht sein Verhalten in der vom Sender gewünschten Weise beeinflusst. Auf den ersten Blick mag die Annahme, der einzige Zweck der Kommunikation sei es, das Verhalten des Empfängers zu beeinflussen, als eine unerwünschte Beschränkung erscheinen. Mit einer einigermaßen weiten Auslegung des Begriffes ‘Verhalten’ ist es jedoch klar, dass Kommunikation entweder das Verhalten beeinflusst oder aber ohne irgendeine ersichtliche Wirkung bleibt.

Beginnen wir jetzt aber mit der Diskussion unseres in Abb. 5.5 dargestellten Kommunikationssystems und wenden uns als erstes der Frage zu, was Information ist und wie sie sich darstellen lässt.

5.2.2 Was ist Information?

Die Organismen und Systeme, die unsere Umwelt bilden, stehen in einem dauernden Austausch miteinander. Offensichtlich und schon lange bekannt und beschrieben ist der Austausch von Materie und Energie. Eine dritte Art des Austausches besteht im Austausch von *Nachrichten* oder *Information*. Solche Nachrichten können Reize, Befehle, Fragen, Antworten, Beobachtungen, Behauptungen usw. sein [35].

Der Begriff der Information ist auch in der normalen Umgangssprache vorhanden. Eine mit seiner gewohnten Interpretation in Einklang stehende Definition ist im folgenden in drei verschiedenen Formulierungen wiedergegeben [64]:

- Information ist das, was wir nicht wissen.
- Information ist die Eigenschaft einer Nachricht, beim Empfänger Kenntniszuwachs zu erzeugen.
- Information wird als Wissenszuwachs oder Abbau von Unsicherheiten über einen interessierenden Sachverhalt erklärt.

Information hat also den Charakter von etwas Unbekanntem oder Unerwartetem. Alles, was wir bereits wissen, enthält keine Information mehr. Hierbei ist nur der Aspekt des Unbekannten wichtig, nicht die ‘Inhaltsschwere’ oder Bedeutsamkeit der Nachricht. Damit enthält die Nachricht ‘Die Kuh XYZ auf der Alm QRS hat heute gekalbt’ den gleichen Informationsgehalt wie ‘Sie haben den Jackpott mit 42 Mio. DM geknackt’. Die Nachricht ‘Irgendwo auf der Welt hat heute irgendeine beliebige Kuh gekalbt’ hat dagegen einen Informationsgehalt von Null: zwar irritiert mich die Meldung vielleicht, da ich nicht weiß, was ich mit ihr anfangen soll. Die Information in dieser Nachricht ist jedoch trivial.

Zur Darstellung von Information dienen Signale und Zeichen. Ein Signal ist die physikalische Repräsentation einer Nachricht, das können z.B. die Farbe einer Verkehrsampel, Intensität oder Frequenz einer Licht- oder Schallquelle, die Spannung auf einer elektrischen Leitung oder der zeitliche Verlauf dieser Spannung sein. Kleine Variationen dieser Signale ändern nicht notwendigerweise auch deren Bedeutung: die akustischen Signale der Sprache z.B. werden gleichermaßen verstanden, wenn ein Kind in heller Stimmlage oder ein Erwachsener in einer ganz anderen Stimmlage spricht. Auch Handschriften ergeben völlig unterschiedliche Signale, selbst wenn sie denselben Text darstellen. Bei Spannungssignalen wird in der TTL-Logik jedes Signal unterhalb 0.8 V in eine Signalklasse 0 eingeordnet, jedes Signal oberhalb 2 V in eine andere Signalklasse 1. Eine derartige Signalklasse bezeichnet man auch als Zeichen. Weitere Beispiele für Zeichen sind die Buchstaben des Alphabets, die Ziffern und die Satzzeichen. Diese Zeichen lassen sich ihrerseits wieder zu Zeichenketten (Hyperzeichen) zusammenfassen, es entstehen damit Zeichenhierarchien: in einem Text wären diese Hierarchien gegeben durch Buchstaben – Wörter – Sätze – Abschnitte usw.

Physikalisch gesehen sind Nachrichten also dadurch gegeben, dass Zeichen in einem bestimmten räumlichen und zeitlichen Zusammenhang auftreten. Wir werden uns in den folgenden Betrachtungen stets auf diskrete Signale oder Zeichen beschränken, kontinuierliche

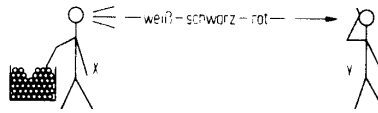


Abbildung 5.6: Ungestörter Informationskanal [44]

Signale (also z.B. Zeitverläufe von Spannungen) werden, außer im Zusammenhang mit dem Abtasttheorem, nicht betrachtet.

5.2.3 Informationsgehalt

In der von Shannon [110, 111] begründeten Informationstheorie werden die Formen, die Nachrichten annehmen können, und die Beziehungen zwischen ihnen quantitativ untersucht. Die Beschreibungsmittel der Informationstheorie sind die Elemente der Wahrscheinlichkeitsrechnung.

Informationsquelle

Weiter oben hatten wir bereits Information als das Neue, Unbekannte bezeichnet, hier soll an einem Beispiel erläutert werden, dass diese Definition sich durchaus innerhalb des umgangssprachlichen Bedeutungsbereiches von Information befindet. Auch wird die statistische Beschreibung der Information deutlich. Wir wollen hier einen ungestörten Informationskanal betrachten, um das Wesen der Information deutlich zu machen. Störungen und den Umgang mit ihnen wollen wir in Abschn. 5.3.2 betrachten. Unser Informationskanal ist in Abb. 5.6 dargestellt: In einer Urne befindet sich eine große Menge Kugeln, davon seien 70% weiß, 20% schwarz und 10% rot. X zieht jetzt jeweils eine Kugel und ruft Y das Ergebnis zu. Dann ist X zusammen mit der Urne eine Nachrichtenquelle, die die Nachrichten 'weiß', 'rot' und 'schwarz' erzeugt. Y ist ein Nachrichtenempfänger, der sich genau für diese Informationen interessiert. Welche der drei Nachrichten enthält nun für Y die größte Information? Nach der Ziehung und Übermittlung von vielen Kugeln kennt der Empfänger Y die Wahrscheinlichkeit p_x für das Auftreten der einzelnen Farben: $p_w=0.7$, $p_s=0.2$ und $p_r=0.1$. Damit kennt er aber auch den Informationsgehalt dieser Nachrichten: Die Nachricht 'weiß' enthält die geringste Information (sie ist aufgrund ihrer großen Wahrscheinlichkeit das Standardsignal, interessant sind erst Abweichungen von 'weiß'), die Nachricht 'rot' enthält am meisten Information. Das ist anschaulich, wenn man sich vorstellt, dass die Urne nur Kugeln einer Farbe enthält, z.B. weiß. Dann ist $p_w=1$ und der Empfänger Y weiß bereits vorher, welche Nachricht ihm X übermitteln wird. Da es sich dann um eine sichere Nachricht handelt, ist ihr Informationsgehalt gleich Null. Aufgrund dieser statistischen Beschreibung wird im folgenden auch manchmal der Begriff Versuchsausgang anstelle von Signal, Zeichen oder Information verwendet.

Aus diesem Experiment ergeben sich zwei Kriterien zur Definition des Informationsgehaltes einer Nachricht:

- K1 Der *Informationsgehalt* I_x einer Nachricht ist umso größer, je kleiner die Wahrscheinlichkeit p_x ihres Auftretens ist, d.h. je größer ihr 'Überraschungswert' ist. Damit wird, wie bereits oben vorgeschlagen, Information als der Neuigkeitsgehalt aber nicht die Inhaltsschwere einer Nachricht definiert.
- K2 Eine Nachricht mit der Wahrscheinlichkeit $p_x=1$, d.h. das sichere Ereignis, muss den Informationsgehalt $I_x=0$ haben.

Ein drittes Kriterium zur Definition des Informationsgehalts erhält man bei der Verwendung mehrerer Nachrichten:

- K3 Der Informationsgehalt verschiedener voneinander unabhängiger Nachrichten soll sich addieren.

Diese Definition ist unmittelbar einsichtig aus der Interpretation der Information mittels wahrscheinlichkeitstheoretischer Begriffe. Auch für die Wahrscheinlichkeiten gilt dort, dass sich bei voneinander unabhängigen Ereignissen die Gesamtwahrscheinlichkeit eines Ereignisses durch die Multiplikation der Wahrscheinlichkeiten der Einzelereignisse ergibt. Der Unterschied in der Rechenoperation, Multiplikation statt Addition, wird deutlich werden, wenn wir zur formalen Definition der Information übergehen.

Diskrete Quelle, gleichwahrscheinliche unabhängige Symbole

Nach dieser anschaulichen Darstellung wollen wir uns nun dem einfachsten Fall zuwenden, einer Signalquelle, die n gleichwahrscheinliche, unabhängige Symbole erzeugt. Ein Beispiel für eine derartige Quelle wäre ein Würfel ($n = 6$) oder die Münze beim Münzwurf ($n = 2$). Die Wahrscheinlichkeit p_x für das Auftreten eines Symbols ist dann gegeben durch

$$p_x = \frac{1}{n}. \quad (5.1)$$

Da alle Symbole mit der gleichen Wahrscheinlichkeit p_x auftreten, hat jedes Symbol den gleichen Informationsgehalt I_x . In Übereinstimmung mit den oben gegebenen drei Kriterien lässt sich dann die folgende Definition des *Informationsgehaltes einer Nachricht* angeben:

$$I_x = \text{ld} \frac{1}{p_x} = -\text{ld}(p_x), \quad (5.2)$$

wobei ld (log dualis) den Logarithmus zur Basis 2 bezeichnet. Die Einheit ist bit (binary digit) – nicht zu verwechseln mit Bit für Binärzeichen. Diese Definition ist einsichtig insofern, als dass das erste Kriterium K1 die Verwendung der Kehrwerte der Wahrscheinlichkeiten nahe legt, das zweite Kriterium K2 die Verwendung einer logarithmischen Definition nahe legt und das dritte Kriterium K3 diese sogar erzwingt (Addition der Informationen durch Umwandlung des Produkts der reziproken Wahrscheinlichkeiten in eine Summe der Informationsgehalte der einzelnen Zeichen). Man beachte, dass der so definierte Informationsgehalt, z.B. einer Auswahl eines Zeichens aus einem Zeichensatz, einfach durch die Statistik, also durch die Wahrscheinlichkeit des Auftretens eines Zeichens, bestimmt wird, nicht aber durch die Semantik, d.h. die Bedeutung des Zeichens.

Die Verwendung des Logarithmus zur Basis 2 ist eine willkürliche Festlegung. Der logarithmus dualis wurde gewählt, damit für den einfachsten Fall einer symmetrischen, binären Nachricht mit $p_x=0.5$ (d.h. gleichviele rote und weiße Kugeln in der Urne, keine schwarzen; oder alternatives Beispiel Münzwurf) der Informationsgehalt eines Zeichens 1 bit wird, d.h.

Definition 2 Der Informationsgehalt eines Versuches mit zwei gleichwahrscheinlichen Ausgängen beträgt 1 bit.

Diese Definition lässt sich auch auf Versuche mit einer größeren Zahl gleichwahrscheinlicher Ausgänge erweitern [20, 21]. Betrachten wir dazu einen Versuch mit $N = 2^n$ gleichwahrscheinlichen Ausgängen. Diese Ausgänge können wir zuerst in zwei gleichgroße Mengen aufteilen und dann untersuchen, in welcher der beiden Hälften der Versuchsausgang liegt. Da beide Hälften gleichwahrscheinlich sind, erfragen wir damit ein erstes bit an Information. Die Hälfte, in der sich der Versuchsausgang befindet, halbieren wir wieder, und wiederholen dieses Verfahren solange, bis die letzte Halbierung dann genau einen Ausgang ergibt. Damit haben wir

$$I = n = \text{ld}(N) \quad (5.3)$$

bit an Information erfragt. Diese Methode lässt sich auch dann anwenden, wenn N keine Zweierpotenz ist.

Die Größe dieser Einheit bit lässt sich am Beispiel des Zahlenlotto ‘6 aus 49’ veranschaulichen: Dabei gibt es

$$N = \binom{49}{6} = 13\,983\,816 \quad (5.4)$$

Möglichkeiten, 6 Zahlen aus 49 auszuwählen. Der Informationsgehalt einer solchen Auswahl beträgt also

$$I = \text{ld} \binom{49}{6} = 23.73 \text{ bit} . \quad (5.5)$$

Zum Vergleich: Ein Würfel stellt eine Quelle dar, bei der jedes Signal mit einer Wahrscheinlichkeit von $1/6$ erzeugt wird. Damit ergibt sich ein Informationsgehalt von $\text{ld} 6 = 2.6$ bit.

Diskrete Quelle, unabhängige Symbole, unterschiedliche Wahrscheinlichkeiten

Eine Quelle mit unabhängigen aber nicht gleichwahrscheinlichen Symbolen lässt sich leicht erzeugen durch Würfeln mit zwei Würfeln. Addieren wir dann die Augenzahl, so ergeben sich Versuchsausgänge mit Werten zwischen 2 und 12, wobei aber die Wahrscheinlichkeiten für diese Versuchsausgänge unterschiedlich sind und damit nach (5.2) auch ihr Informationsgehalt. Für dieses Beispiel würden wir für die Wahrscheinlichkeiten erhalten $p(2) = p(12) = 1/36$; $p(3) = p(11) = 2/36$; ...; $p(6) = p(8) = 5/36$ und $p(7) = 1/6$. Damit ergeben sich für die einzelnen Versuchsausgänge Informationsgehalte von $I(2) = I(12) = 5.2$ bit, $I(3) = I(11) = 4.2$ bit, $I(6) = I(8) = 2.9$ bit und $I(7) = 2.8$ bit.

Die Sprache ist kein Beispiel für eine Quelle unabhängiger nicht gleichwahrscheinlicher Symbole. Zwar sind auch hier die Wahrscheinlichkeiten für jeden einzelnen Buchstaben unterschiedlich, so kommen z.B. die Vokale häufiger vor als Konsonanten und auch die einzelnen Vokale bzw. Konsonanten treten mit unterschiedlicher Wahrscheinlichkeit auf, aber die Zeichen sind nicht unabhängig: so gibt es bestimmte Buchstabenkombinationen, die immer wieder auftreten und es gibt sogar im Deutschen eine sichere Folge: auf ein 'q' folgt stets ein 'u'.

Diskrete Quelle, kontinuierliche Symbole

Wenn man eine Größe messen will, deren Wert in einem kontinuierlichen Bereich liegen kann, z.B. bei einer Längenmessung, so scheint es zunächst, als ob dabei zwischen unendlich vielen Messwerten zu unterscheiden wäre und in einer solchen Messung daher unendlich viel Information steckt. Tatsächlich ist aber bei jeder Messung die Genauigkeit wie auch der erfassbare Bereich begrenzt, und es ist nicht sinnvoll, diesen Bereich in feinere Zellen zu unterteilen als im Rahmen der Messgenauigkeit noch aufgelöst werden können. Erstreckt sich der Messbereich von y_1 bis y_2 und ist die absolute Auflösung im ganzen Bereich gleich Δy , so können

$$N = \frac{y_2 - y_1}{\Delta y} \quad (5.6)$$

unterschiedliche Signale erzeugt werden. Eine Messung würde dann $\text{ld}(N)$ bit Information liefern.

Alternativ kann es auch vorkommen, dass nicht die absolute, sondern die relative Auflösung $a = \Delta y/y$ konstant ist. Dann würde man den Bereich logarithmisch abbilden mit $z = \log y$. Jeder y -Wert ist dann um einen Faktor $1 - p$ größer als der vorangehende, und für die Beziehung zwischen den Grenzen des Messbereichs gilt $y_2 = y_1 (1 + p)^n$. Daher ergibt sich für die Zahl der Intervalle

$$n = \frac{\log \left(\frac{y_2}{y_1} \right)}{\log(1 + p)} \quad (5.7)$$

und die $\log y$ -Skala wäre in Zellen der einheitlichen Breite $\Delta z = \log(1 + p)$ eingeteilt. Der Informationsgehalt ergibt sich dann selbstverständlich wieder als $\text{ld}(n)$.

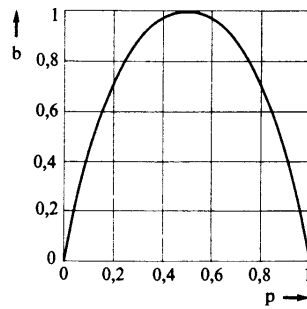


Abbildung 5.7: Shannon-Funktion

5.2.4 Shannon-Funktion und Entropie

Wir haben bisher den Informationsgehalt I_x eines einzelnen Zeichens betrachtet. Wenden wir uns jetzt jedoch auch der Frage nach dem mittleren Informationsgehalt eines Zeichens aus einem Zeichenvorrat zu. Letztendlich erlaubt uns diese Betrachtung ja, die Wichtigkeit bzw. den Informationsgehalt der einzelnen Zeichen besser zu beurteilen. Der mittlere Informationsgehalt eines Zeichensatzes ergibt sich durch Mittelwertbildung über alle Zeichen, wobei der Informationsgehalt jedes einzelnen Zeichens mit der Wahrscheinlichkeit des Auftretens dieses Zeichens gewichtet werden muss:

$$H = \bar{I}_x = - \sum_{x=1}^n p_x \cdot \text{ld}(p_x) . \quad (5.8)$$

Dieser mittlere Informationsgehalt H wird als die Entropie bezeichnet:

Definition 3 *Der mittlere Informationsgehalt je Zeichen, die Entropie, gibt an, wie viele Binärzeichen man bei der binären Codierung von Nachrichten für die einzelnen Zeichen im Mittel wenigstens aufwenden muss. Andererseits sind bei geschickter Codierung auch kaum mehr Zeichen erforderlich [111].*

Der Begriff der Entropie wurde aus der Thermodynamik entlehnt, wo er ein Maß für die statistische Unordnung oder Unsicherheit in einem System ist.

Die Entropie eines Binärzeichens lässt sich mit Hilfe der *Shannon-Funktion* (vgl. Abb. 5.7) angeben. Betrachte dazu wieder einen Versuch mit zwei Ausgängen, diese haben die Wahrscheinlichkeiten p und $q = (1 - p)$. Der mittlere Informationsgehalt eines dieser Zeichen ist dann gegeben zu

$$H = b(p) = -p \cdot \text{ld}(p) - (1 - p) \cdot \text{ld}(1 - p) . \quad (5.9)$$

Diese Funktion $b(p)$ wird als die Shannon-Funktion bezeichnet, sie ist in Abb. 5.7 wiedergegeben. Für $p = 0$, d.h. es tritt nur eines der beiden Zeichen auf, ist der Informationsgehalt gleich Null (sicheres Ereignis), mit wachsendem p steigt der mittlere Informationsgehalt je Zeichen an bis zu einem Maximum bei $p = 0.5$ (beide Ausgänge sind dann gleich wahrscheinlich, damit ist das Ergebnis am schlechtesten vorherzusagen). Für weiter wachsendes p nimmt die Entropie wieder ab, bis sie für $p = 1$ Null wird (sicheres Ereignis des anderen Signals). Dieses Ergebnis lässt sich verallgemeinern zu:

Definition 4 *Die Entropien eines Versuches mit den möglichen Ausgängen x_i ist am größten, wenn alle Ausgänge gleich wahrscheinlich sind [21].*

5.2.5 Redundanz

Anhand der Shannon-Funktion bzw. ihrer Verallgemeinerung auf Zeichensätze mit einem größeren Zeichenvorrat können wir erkennen, dass die maximale Entropie und damit der

#	Buchstabe	$f = p_i$	$F = \sum p_i$	$S_i = p_i \lg \frac{1}{p_i}$
1	-	0.151490	0.1514	0.41251
2	E	0.147004	0.2985	0.40661
3	N	0.088351	0.3869	0.30927
4	R	0.068577	0.4554	0.26512
5	I	0.063770	0.5192	0.25232
6	S	0.053881	0.5731	0.22795
7	T	0.047310	0.6204	0.20824
8	D	0.043854	0.6642	0.19783
9	H	0.053554	0.7078	0.19691
10	A	0.043309	0.7511	0.19616
11	U	0.031877	0.7830	0.15847
12	L	0.029312	0.8123	0.14927
13	C	0.026733	0.8390	0.13968
14	G	0.026672	0.8657	0.13945
15	M	0.021336	0.8870	0.11842
16	O	0.017717	0.9047	0.10389
17	B	0.015972	0.9207	0.09585
18	Z	0.014225	0.9349	0.08727
19	W	0.014201	0.9491	0.08716
20	F	0.013598	0.9637	0.08431
21	K	0.009558	0.9723	0.06412
22	V	0.007350	0.9796	0.05209
23	Ü	0.005799	0.9854	0.04309
24	P	0.004992	0.9904	0.03817
25	Ä	0.004907	0.9953	0.03764
26	Ö	0.002547	0.9979	0.02194
27	J	0.001645	0.9995	0.01521
28	Y	0.000173	0.9997	0.00217
29	Q	0.000142	0.9999	0.00181
30	X	0.000129	1.0000	0.00167
Σ		1.000000	4.11461	

Tabelle 5.1: Wahrscheinlichkeiten und mittlerer Informationsgehalt von Buchstaben der deutschen Sprache

maximale Informationsgehalt einer Nachricht nur dann erreicht werden kann, wenn die einzelnen Nachrichtenelemente gleich verteilt und statistisch voneinander unabhängig sind. Ist dies nicht der Fall, so benötige ich zur Darstellung einer Information eine wesentlich größere Menge an Zeichen.

Als Beispiel für eine ungleich verteilte und nicht völlig unabhängige Signalquelle sei hierzu die Sprache betrachtet. Würden in der deutschen Sprache alle 26 Buchstaben mit der gleichen Wahrscheinlichkeit $p_x=1/26$ vorkommen und unabhängig voneinander sein, so ergäbe sich für den Informationsgehalt eines Buchstabens (d.h. den mittleren Informationsgehalt eines Zeichens meines Zeichenvorrats)

$$H_0 = I = \lg(26) = 4.7 \text{ bit} . \quad (5.10)$$

Betrachtet man aber einen Text, so erkennt man, dass die Buchstaben mit unterschiedlicher Häufigkeit auftreten. In einem deutschen Text z.B. tritt der Buchstabe 'e' mit einer Häufigkeit von 14.4% auf (nur das Leerzeichen ist mit 14.42% geringfügig häufiger), die Buchstaben 'n' ('s', 'i', 'm') haben Häufigkeiten von 8.65% (6.46%, 6.28%, 1.72%), die seltensten Buchstaben sind 'x' und 'q' mit Häufigkeiten von 0.8% bzw. 0.5% (vgl. Tabelle 5.1). Berücksichtigt man

diese Verteilungen der Häufigkeiten, so ergibt sich ein mittlerer Informationsgehalt H von nur 4.1 bit/Zeichen alleine aus der Tatsache, dass die Zeichen nicht gleich wahrscheinlich sind.

Zusätzlich müssen wir aber auch noch berücksichtigen, dass die Zeichen in der Sprache nicht unabhängig voneinander sind. Betrachtet man Folgen von Zeichen in einem Text, so zeigt sich, dass z.B. auf ein 'a' wesentlich häufiger ein 'n' folgt als ein 'o', dass ein 'c' meist in Kombination mit einem 'h' auftritt, häufig sogar in der Form 'sch', und dass auf ein 'q' stets ein 'u' folgt. Nach Küpfmüller [65] ergibt sich dann unter Berücksichtigung dieser Kombinationen und der unterschiedlichen Häufigkeiten der einzelnen Zeichen ein mittlerer Informationsgehalt H von 1.3 bit/Zeichen (verglichen mit den 4.7 bit/Zeichen bei gleich wahrscheinlichen und unabhängigen Zeichen).

Jede Abweichung von der Gleichverteilung bzw. von der statistischen Unabhängigkeit verringert also den mittleren Informationsgehalt einer Nachricht. Diese Verringerung der Entropie wird als Redundanz bezeichnet, manchmal auch als Weitschweifigkeit übersetzt. Formal definiert sich die Redundanz als die Differenz zwischen der maximal möglichen Entropie H_0 (5.10) und der in einer realen Zeichenkette steckenden Entropie H

$$R = H_0 - H . \quad (5.11)$$

Die relative Redundanz ist dann gegeben durch

$$r = \frac{H_0 - H}{H_0} . \quad (5.12)$$

Kehren wir noch einmal zu dem obigen Beispiel der deutschen Sprache zurück, so erhalten wir eine Redundanz R von 3.4 bit/Zeichen bzw. eine relative Redundanz r von 0.73. Bezüglich des Informationsgehaltes der deutschen Sprache bedeutet das, dass 73% der Sprache redundant oder überflüssig sind und nur 27% Information tragen. Wir könnten also mit einer anderen Sprachstruktur aber dem gleichen Alphabet die zur Übermittlung einer Nachricht notwendigen Materialien auf etwas über 1/4 reduzieren (Bücher würden dünner, Morsebotschaften schneller etc.).

Ökonomisch würde die Verminderung der Redundanz daher attraktiv erscheinen. Aber wohl auch nur erscheinen. Redundanz in der Sprache ist wichtig: Zwar wird die Nachricht und damit auch die Übertragungszeit verlängert, andererseits ist jedoch das Erkennen von Fehlern und deren Korrektur möglich (so kann man in der Regel auch bei stark gestörter Telefonübertragung den Sender noch verstehen, zumindest wenn er in der Muttersprache spricht, da man aus den verstandenen Teilen die unverstandenen ergänzen kann, es aufgrund der Redundanz eben nicht auf jeden einzelnen Buchstaben ankommt). Steinbuch [114] gibt dazu verschiedene Beispiele, von denen einige im folgenden wiedergegeben werden. Ganz deutlich wird dabei, dass die gegenseitige Beziehung zwischen Zeichen (Wörtern, Satzfragmenten, Sätzen) für unser Verständnis von entscheidender Bedeutung ist. Liest man z.B. in diesem Text die sechs Buchstaben 'Inf.....ion', so besteht kein Zweifel, dass die fehlenden fünf Buchstaben 'ormat' sind und damit das ganze Wort 'Information' ist. Die Abhängigkeit der Zeichenergänzung vom Satzzusammenhang wird in den folgenden Beispielen deutlich. Beispielsweise werden wir im Satz 'Es gibt konkave und kon.... Linsen' das verstümmelte Signal 'kon....' als 'konvexe' deuten. Dagegen werden wir dasselbe verstümmelte Signal 'kon.....' im Satz 'Es gibt abstrakte und kon..... Malerei' als 'konkrete' deuten.

Die Fähigkeit des Empfängers, Ergänzungen an unvollständigen Signalen vorzunehmen, hängt wesentlich von zwei Komponenten ab: (1) der Kenntnis des Empfängers über den Sender und sein Repertoire an Nachrichten und (2) der Fähigkeit des Empfängers, die verschiedenen empfangenen Nachrichten logisch zu verknüpfen. Haben wir beispielsweise die Liste

- (1) kon....
- (2) konkordant
- (3) kon....
- (4) konsekutiv
- (5) konsequent
- (6) konservativ
- (7) konsiderabel
- (8) konsistent
- (9) konstant
- (10) konträr
- (11) kon....

so können wir bei oberflächlicher Betrachtung nicht wissen, aus welchen vollständigen Worten die verstümmelten Buchstabengruppen (1), (3) und (11) entstanden sind. Bei genauerer Betrachtung stellen wir jedoch fest, dass es sich um eine alphabetisch geordnete Liste von Adjektiven handelt. Wollen wir mit dieser Kenntnis die verstümmelten Buchstabengruppen ergänzen, so greifen wir z.B. zum Duden und ergänzen ohne großes Risiko: (1) konkav, (3) konkret und (11) konvex. Da diese Fähigkeit des Empfängers, die Signale zu ergänzen, vom Empfänger und seinen Kenntnissen über den Sender abhängt, können unterschiedliche Empfänger aus exakt denselben Signalen ein unterschiedliches Maß an Informationen entnehmen.

Diese Bedeutung von Redundanz und Zusammenhängen in der Sprache für die Informationsaufnahme des Menschen macht Wibberenz (1990) an einem Beispiel deutlich. Dazu sind vier Zeichenketten aus je acht Buchstaben vorgegeben, der Betrachter hat jeweils eine Sekunde Zeit, sich diese Buchstabenfolgen einzuprägen:

- (1) AAAABBBB
- (2) ABCBEEPS
- (3) FGKWOSLM
- (4) TOMORROW

Dabei hat (1) die größte Redundanz, (2) und (4) sind teilweise redundant und haben, zumindest was die Häufigkeitsverteilungen der einzelnen Buchstaben angeht, gleiche Redundanz, (3) ist nicht redundant und ist gleichzeitig auch am schwersten aufzunehmen (man muss sich wirklich jeden Buchstaben einzeln merken). (4) bildet in diesen Zeichenketten eine Sonderstellung, da es ein Hyperzeichen ist und insofern nur ein einzelnes Zeichen (aber aus einem anderen Zeichenvorrat, in diesem Falle den Wörtern der englischen Sprache stammt) aufgenommen wird und daher auch die einzelnen Buchstaben nicht wie bei Beispiel (3) 'gespeichert' werden müssen. (Anmerkung: bei (2) 'beträgt' man beim Speichern natürlich auch: kaum jemand wird sich die Buchstabenfolge so merken, eher wird man sie sich einprägen in der Form 'die ersten drei Buchstaben des Alphabets + beeps', da letzters schon wieder ein Hyperzeichen der englischen Sprache wäre).

Die Kenntnis der Redundanz ist für die platzsparende Codierung von Information wichtig. Als einfachstes Beispiel seien hier Versuche erwähnt, das Alphabet in einem Binärcode darzustellen (vgl. Abb. 5.8). Der älteste 'Binärcode' für Buchstaben geht auf Bacon (1623) zurück und basiert auf der Verwendung von 5 Zeichen pro Buchstaben. Das Morse-Alphabet (1844) dagegen bedient sich bereits einer zumindest teilweise Redundanz verminderten Codierung: je häufiger ein Zeichen auftritt, umso kürzer ist es (das stimmt bei Morse nur näherungsweise, war aber in seinem Ansatz enthalten). In Abschn. 5.3.1 werden wir weitere Codierungsverfahren diskutieren, die die Redundanz einer Signalquelle verringern sollen (Quellencodierung).

Einschub: Künstliche Sprachen

Am Beispiel der Sprache kann man sich sehr viele Eigenschaften von Information und diskreten Quellen klar machen, wie wir es ja auch schon im Zusammenhang mit der Redundanz gesehen haben. Insbesondere lässt sich der stochastische Aspekt verdeutlichen. Shannon

Buchstabe	Häufigkeit p_i / %	Bacon 1623 Baudot 1874	Morse 1844	Huffman 1952
┌	14,42	00100	00	000
E	14,40	10000	100	001
N	8,65	00110	01100	010
S	6,46	10100	11100	0110
I	6,28	01100	1100	0111
R	6,22	01010	101100	1000
⋮				
M	1,72	00111	010100	111010
⋮				
X	0,08	10111	01110100	1111111110
Q	0,05	11101	010110100	11111111110
H_c / bit/Buchstabe		5	4,79	4,13

Abbildung 5.8: Binärcodes für alphabetischen Text [74]

und Weaver [112] folgend sei hier als ein Beispiel eine künstliche Sprache eingeführt. Diese Annäherung an eine Sprache umfasst drei Stufen:

- (A) Angenommen, wir haben 5 Buchstaben A, B, C, D und E, die jeweils mit einer Wahrscheinlichkeit von 0.2 gewählt sind, wobei aufeinander folgende Auswahlen unabhängig voneinander sind (diskrete Quelle mit gleich wahrscheinlichen, unabhängigen Symbolen). Das würde z.B. zu einer Folge

BDCBCECCADCBDAAECEEAAABBDAEECACEEBAEECBCEAD

führen.

- (B) Nun werden dieselben 5 Buchstaben benutzt mit den Wahrscheinlichkeiten 0.4, 0.1, 0.2, 0.2, und 0.1, wobei aufeinander folgende Auswahlen ebenfalls wieder unabhängig sind (diskrete Quelle mit unabhängigen, nicht gleich wahrscheinlichen Symbolen). Eine typische Folge wäre dann:

AAACDCBDCEAADADACEDAEADCABEDADDCECAAAAAD.

- (C) Eine kompliziertere Struktur ergibt sich, wenn, wie in der gewöhnlichen Sprache, die Zeichen nicht mehr voneinander unabhängig sind, sondern ihre Wahrscheinlichkeiten von dem vorangegangenen Zeichen abhängen. Im einfachsten Fall hängt eine Auswahl dann nur vom direkt vorangegangenen Zeichen ab, nicht jedoch auch noch von früheren Zeichen. Der statistische Aufbau kann dann durch einen Satz von Übergangswahrscheinlichkeiten $p_i(j)$ beschrieben werden, die Wahrscheinlichkeit, dass dem Buchstaben i der Buchstabe j folgt. Ein alternativer Weg wäre die Angabe der Digramm-Wahrscheinlichkeiten (zweidimensionale Verbundwahrscheinlichkeit) $p(i, j)$, d.h. die relative Häufigkeit des Digramms ij . Die Buchstabenhäufigkeiten $p(i)$ (die Wahrscheinlichkeit des Buchstabens i), die Übergangswahrscheinlichkeiten $p_i(j)$, und die Verbundwahrscheinlichkeiten $p(i, j)$ hängen zusammen über

$$p(i) = \sum_j p(i, j) = \sum_i p(j, i) = \sum_i p(j)p_j(i), \quad (5.13)$$

$$p(i, j) = p(i)p_j(j) \quad (5.14)$$

und

$$\sum_i p_i(j) = \sum_i p(i) = \sum_{i,j} p(i, j) = 1. \quad (5.15)$$

Als ein Beispiel nehmen wir an, es gibt drei Buchstaben A, B, und C mit den Wahrscheinlichkeitstabellen

$p_i(j)$		j	
		A	B
		C	
		A	0
		B	$\frac{4}{5}$
		C	$\frac{1}{5}$
i		A	$\frac{1}{2}$
		B	$\frac{1}{2}$
		C	0

i	$p(i)$
	A
	B
	C
	$\frac{2}{7}$
	$\frac{16}{7}$
	$\frac{2}{7}$

$p(i, j)$		j	
		A	B
		C	
		A	0
		B	$\frac{4}{15}$
		C	$\frac{1}{15}$
i		A	$\frac{2}{7}$
		B	$\frac{2}{7}$
		C	0
		A	$\frac{1}{27}$
		B	$\frac{4}{135}$
		C	$\frac{1}{135}$

Ein Beispiel für eine Nachricht dieser Quelle wäre

ABBABABABABABABBBBABBBBABBABABABABBBACACABBABBBBABBAB.

Dieses Verfahren lässt sich erweitern auf Trigramm-Häufigkeiten, bei denen die Auswahl eines Buchstabens von den zwei vorhergehenden abhängt. Entsprechend müssten Trigramm-Häufigkeiten $p(i, j, k)$ angegeben werden. Erweiterungen auf den allgemeinen Fall der n -gramm-Häufigkeiten sind ebenfalls möglich. Hierbei ergeben sich zwar kompliziertere Zusammenhänge aber keine neuen Strukturen.

- (D) Eine Erweiterung ergibt sich, wenn wir stochastische Prozesse definieren, die als Text eine Folge von Wörtern hervorbringen. In dieser Sprache gibt es wieder die 5 Buchstaben A, B, C, D und E und 16 Wörter mit den zugehörigen Wahrscheinlichkeiten:

0,10 A	0,16 BEBE	0,11 CABED	0,04 DEB
0,04 ADEB	0,04 BED	0,05 CEED	0,15 DEED
0,05 ADEE	0,02 BEED	0,08 DAB	0,01 EAB
0,01 BADD	0,05 CA	0,04 DAD	0,05 EE

Werden aufeinander folgende Wörter unabhängig voneinander ausgewählt und durch einen Zwischenraum getrennt, so ergibt sich eine typische Nachricht in der Form

DAB EE A BEBE DEED DEB ADEE ADEE EE DEB BEBE BEBE ADEE BED DEED
DEED CEED ADEE A DEED DEED BEBE CABED BEBE BED DAB DEED ADEB

Wenn alle Wörter die gleiche Länge haben, ist dieser Prozess äquivalent zu einem der vorangegangenen, allerdings kann die Beschreibung mit Hilfe der Wortstruktur und der Wortwahrscheinlichkeiten einfacher sein. Natürlich kann man jetzt auch hier vereinfachen und Übergangswahrscheinlichkeiten zwischen Wörtern einführen usw.

Shannon [110] hat diese Abhängigkeiten in der Sprache und ihren Zusammenhang mit stochastischen Prozessen folgendermaßen illustriert: Je weiter man die Einwirkung vorhergehender Buchstaben einer Zufallsfolge von Zeichen statistisch an diejenige der wirklichen Sprache angleicht, um so mehr nähert sich ihre Struktur der der Sprache an, wenn auch der Text sinnlos bleibt. Unter Verwendung der 26 Buchstaben des englischen Alphabets und des Zwischenraums hat Shannon die folgenden Beispiele konstruiert:

- Näherung nullter Ordnung (die Zeichen sind unabhängig voneinander und gleich wahrscheinlich):

XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAKBZAACIBZLHJQD

Damit ergibt sich ein 'Text', den man nur buchstabieren, nicht aber lesen kann. Sie können sich dieses veranschaulichen, wenn Sie die Zeit bestimmen, die Sie benötigen, um diese Zeichenkette laut zu lesen und sie mit der Zeit vergleichen, die Sie benötigen, um eine gleichlange Zeichenkette mit normalem Text zu lesen.

- Näherung erster Ordnung (die Zeichen sind weiterhin unabhängig, aber die Häufigkeit der einzelnen Zeichen entspricht der in der englischen Sprache):

OCRO HLI RGWR NMIELSWIS EU LL NBNESBEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL

Ein gälischer Text sieht für unsere Augen auch nicht sehr viel anders aus: Thig sealbh a'dichioll gabh beachd fair an t-sreath dhiamhair Cheilteach seo. Cum oirre dho A gu B.

- Näherung zweiter Ordnung (Digramm-Struktur wie im Englischen, Shannon hat hierzu ein Buch zufällig aufgeschlagen, auf der Seite einen Buchstaben aufgepickt, das Buch geschlossen, wieder zufällig geöffnet und dann auf der Seite bis zum ersten Auftreten dieses Buchstabens gelesen und den Buchstaben zusammen mit dem darauf folgenden rausgeschrieben, Wiederholung des Verfahrens mit dem zweiten Buchstaben usw.):

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE
TOCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE

Im Gegensatz zu den beiden vorangegangenen Versuchen kann man diesen Text schon fast 'lesen': man muss nicht mehr alle Zeichen einzeln buchstabieren und kann damit eine längere Zeichenkette als im ersten Beispiel in einer kürzeren Zeit lesen.

- Näherung dritter Ordnung (Trigramm-Struktur wie im Englischen, wie oben, aber mit drei Buchstaben):

IN NO IST LAT WHEY CRACTICT FROURE BIRS GROCID PONDENOME OF DE-
MONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE

- Näherung erster Ordnung, auf Wörter bezogen (Wörter treten hierbei nur mit ihren natürlichen Häufigkeiten auf):

REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT
NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO
FURNISHES THE LINE MESSAGE HAD BE THESE

- Näherung zweiter Ordnung, auf Wörter bezogen (wie die Digramm-Struktur bei den Buchstaben erzeugt, einzelne Folgen von Worten lassen sich hierbei ohne Problem in sinnvolle Sätze einbauen, die längste Folge umfasst die elf unterstrichenen Wörter, die sich mit etwas Mühe in dieser Reihenfolge noch in einen sinnvollen Satz einbauen lassen):

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE
CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE
LETTERS THAT THE TIME WHO EVER TOLD THE PROBLEM FOR AN UNEX-
PECTED

Markoff- und ergodische Prozesse

Stochastische Prozesse der beschriebenen Art werden als Markoff-Prozesse bezeichnet. Der allgemeine Fall lässt sich folgendermaßen beschreiben: Es existiert eine bestimmte Anzahl S_1, S_2, \dots, S_n von möglichen Zuständen des Systems. Ferner gibt es einen Satz von Übergangswahrscheinlichkeiten $p_i(j)$, die die Wahrscheinlichkeit angeben, mit der das System von einem Zustand S_i in einen Zustand S_j übergeht. Um aus einem solchen Markoff-Prozess eine Informationsquelle zu machen, müssen wir annehmen, dass bei jedem Übergang von einem Zustand in einen anderen ein Zeichen erzeugt wird. Diese Zustände entsprechen dann dem noch verbleibenden Einfluss der vorhergehenden Buchstaben.

Abbildung 5.9 zeigt dazu grafische Darstellungen der Quellen (B), (C) und (D) aus dem obigen Beispiel. Die Zustände sind die Knotenpunkte, die Wahrscheinlichkeiten und Buchstaben, die bei einem Übergang entstehen, sind an den Linien angegeben. Im linken oberen Teil (Quelle (B)) gibt es nur einen Zustand, da die aufeinander folgenden Buchstaben unabhängig sind. Im linken unteren Teil (Quelle (C)) gibt es so viele Zustände wie Buchstaben. Wenn ein Trigramm-Beispiel konstruiert würde, gäbe es höchstens n^2 Zustände, entsprechend den möglichen Buchstabenpaaren, die einem einzelnen gewählten Buchstaben vorangehen. Das rechte Diagramm entspricht dem Beispiel der Wortstruktur in Quelle (D), hier entspricht 's' dem Leerzeichen oder Wortzwischenraum.

Anhand der Diagramm-Strukturen in Abb 5.9 lässt sich das Konzept der ergodischen oder statistisch homogenen Quellen einführen. Alle in einem ergodischen Prozess erzeugten Folgen haben dieselben statistischen Eigenschaften. Daher werden sich die von bestimmten Folgen erhaltenen Buchstabenhäufigkeiten, Digrammhäufigkeiten usw. bei einer Vergrößerung der Länge der Folgen festgesetzten Grenzen nähern, und zwar unabhängig von der Folge. Dadurch

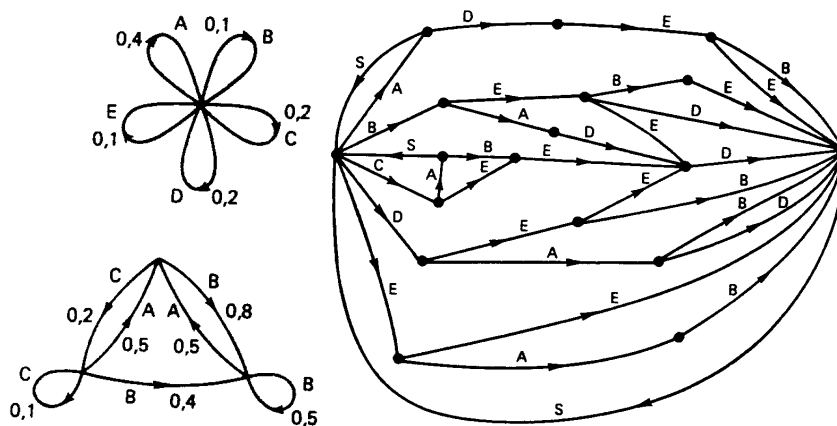


Abbildung 5.9: Diagramme der Quellen B, C und D aus dem Beispiel der künstlichen Sprachen [112]

können wir beim ‘Buchstabenzählen’ auch wirklich auf die Eigenschaften der Quelle zurück schließen und nicht nur auf die Eigenschaften der Folge.

Alle hier betrachteten Beispiele künstlicher Sprache sind ergodisch. Diese Besonderheit hängt mit der Struktur des dazu gehörigen Diagramms zusammen. Ein Prozess ist ergodisch, wenn das Diagramm die beiden folgenden Eigenschaften hat:

- (E1) Das Diagramm zerfällt nicht in zwei getrennte Teile A und B, so dass es unmöglich wäre, von Knotenpunkten im Teil A über Linien des Diagramms in Pfeilrichtung zu Knotenpunkten im Teil B zu gelangen und umgekehrt.
- (E2) Eine geschlossene Linie im Diagramm mit Pfeilen, die in die gleiche Umlaufrichtung zeigen, wird ein ‘Umlauf’ genannt. Die Länge eines Umlaufs ist durch die Anzahl der in ihm vorkommenden Linien gegeben. Die Folge BEBES im unteren Teil von Abb. 5.9 ist ein Umlauf mit der Länge 5. Eine zweite notwendige Eigenschaft ist, dass der größte gemeinsame Teiler der Länge aller Umläufe im Diagramm 1 ist.

Falls die letzte Bedingung verletzt ist, d.h. es gibt einen größten gemeinsamen Teiler größer als 1, besitzen die Folgen eine gewisse periodische Struktur.

5.2.6 Abtasttheorem (Signal als kontinuierliche Funktion der Zeit)

Haben wir bisher nur diskrete Zeichen betrachtet, so soll hier an einem Beispiel der Informationsgehalt eines kontinuierlichen Signals betrachtet werden. Bei einem kontinuierlichen Signal scheint es, ebenso wie bei einem Signal aus einem kontinuierlichen Zeichenvorrat, als hätten wir zwischen unendlich vielen möglichen Werten zu unterscheiden und damit auch unendlich viel Information in diesem Signal. Tatsächlich sind aber bei einer Messung wieder sowohl die Genauigkeit Δx als auch der erfassbare Bereich (x_1, x_2) begrenzt.

Betrachten wir dazu ein Signal als eine kontinuierliche Funktion der Zeit. Dann müssen wir im Prinzip zu jedem Zeitpunkt die Information über die Amplitude übertragen. Das ist jedoch aufgrund der Fehler nicht möglich, wir können uns aber die Frage stellen, wie oft wir einen Messwert übertragen müssen, um die Funktion zuverlässig rekonstruieren zu können. Hat die Zeitfunktion $s(t)$ ein Spektrum, das ausschließlich im Frequenzbereich f von 0 bis B liegt, lässt sie sich durch Abtastproben vollständig beschreiben, wenn für das Abtastintervall gilt

$$T \leq \frac{1}{2B} \quad (5.16)$$

(Shannon’sches Sampling Theorem, Abtasttheorem). Bei äquidistanten Proben muss dann die

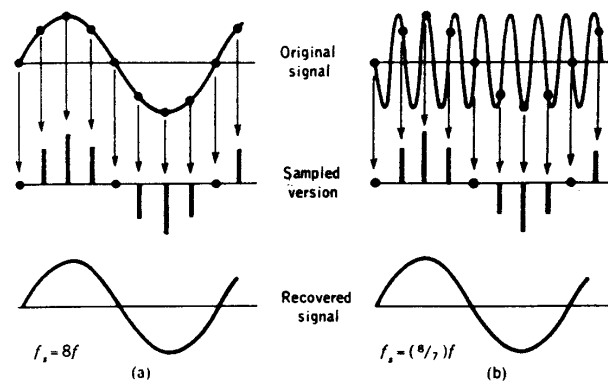


Abbildung 5.10: Aliasing: Falsche Rekonstruktion eines Signals durch zu geringe Abtastfrequenz [116]

Abtastfrequenz zu

$$f \geq 2B \quad (5.17)$$

gewählt werden. Beweise dieses Theorems finden sich z.B. [111] oder [22].

Das Gleichheitszeichen in (5.17) erscheint auf den ersten Blick uneinsichtig, da wir bei einer Abtastfrequenz, die genau das Doppelte der Signalfrequenz ist, im ungünstigsten Falle immer die Nulldurchgänge des Signals erfassen und daher kein Signal rekonstruieren können. Dieses Problem wird auf technische Weise dadurch umgangen, dass das um 90° phasenverschobene Signal ebenfalls vorliegt, von dem in diesem Falle dann die Maxima abgetastet würden, so dass das Originalsignal wieder rekonstruiert werden kann [44].

Ist diese Voraussetzung nicht erfüllt, kann es zu Fehlinformationen (Aliasing) kommen. Abbildung 5.10 zeigt dazu das Beispiel zweier sinusförmigen Signale. Die Abtastfrequenz sei gegeben durch f_s , im linken Beispiel ist $f=8f_s$ und das Signal kann vollständig wieder hergestellt werden, im rechten Beispiel ist $f_s=(8/7)f$, die Abtastung erfolgt zu selten und das rekonstruierte Signal ist falsch.

Eine Vorstellung über den Informationsgehalt in einem kontinuierlichen Signal liefert die digitale Speicherung von Musik. Die Bandbreite erstreckt sich dabei bis zu 16 kHz, d.h. nach dem Abtasttheorem sind 32 000 Stützstellen zum Abtasten dieses Signals erforderlich. Geht man davon aus, dass an jeder Stützstelle die Amplitude in 256 Werte eingeteilt wird (das entspricht 8 Bit), so erhält man eine Gesamtinformationsmenge von 256 000 bit/sec. Der Mensch dagegen kann nur ca. 20 bit/sec aufnehmen (vgl. Tabelle 5.6, [125]). Die Informationsaufnahme des Menschen lässt sich im Prinzip mit Tests wie den oben gegebenen Achtzeichenketten bestimmen [114, 125] oder [124]. Da wir aber andererseits Musik wahrnehmen und auch relativ feine Unterschiede zwischen verschiedenen Aufnahmen erkennen können, muss der Mensch bei der Informationsaufnahme und -auswertung relativ stark mit Mustern und Hyperzeichen arbeiten.

5.2.7 Wahrnehmungspsychologie: Verwendung von Hyperzeichen im Erkennen

Ein sehr interessanter Seitenaspekt im Zusammenhang mit der Informationstheorie betrifft die Informationsverarbeitung durch den Menschen. Dazu gehören sowohl die Aufnahme von Information durch sehr beschränkte Kanäle als auch die Verarbeitung der Information, insbesondere im Bereich des Lernens. Vor ca. 30 Jahren war das Gebiet der Informationspsychologie sehr populär, heute wird Lernen eher vom Aspekt der Motivation als dem der Information her untersucht. Gute, allerdings nicht mehr in allen Punkten aktuelle Bücher sind Steinbuch [114], von Cube [125], Hassenstein [41] oder Vester [124].



Abbildung 5.11: Verschiedenen Schallquellen: Unterscheiden Sie zwischen dem Geräusch, das das Publikum beim Betreten des Hörsaales macht, einem Violinkonzert und dem Stimmen der Instrumente in einem Sinfonieorchester [120]

Ich möchte hier nur einige sehr wenige (und leider unzusammenhängende) Beispiele vorstellen, um einen ‘menschlichen’ Aspekt in die Informationstheorie mit einzubringen. Dazu gehört insbesondere die Mustererkennung zur Reduktion des Informationsflusses. Wir hatten im Zusammenhang mit dem Abtasttheorem erfahren, dass wir beim Musikhören einen Informationsfluss von ca. 250 000 bps verarbeiten müssten, verglichen mit den maximal 20 bps, die der Mensch aufzunehmen vermag. Betrachten wir dazu Abb. 5.11. Dort sind Oszillogramme von drei verschiedenen Geräuschen gezeigt. Eines gibt das Geräusch wieder, das entsteht, wenn das Publikum den Hörsaal betritt. Ein anderes Oszillogramm stammt von einem Sinfonieorchester am Ende des ersten Satzes von Mendelssohns Violinkonzert und das dritte von einem Sinfonieorchester beim Stimmen vor dem Konzert. Die Reihenfolge der Kurven ist zufällig. Welche Kurve gehört zu welchem Geräusch? (Lösung weiter unten im Text)

Ganz offensichtlich arbeitet das Ohr in einer Weise, die es ihm erlaubt, aus diesen für das Auge ununterscheidbaren Signalen genau die wichtigen Merkmale herauszuarbeiten. Andererseits kann es aber auch nicht wie ein AD-Wandler arbeiten, da dann ein vom Menschen nicht mehr beherrschbarer Informationsfluss entstehen würde, d.h. das Ohr muss ein anderes Verfahren haben, als es bei der digitalen Schallaufzeichnung angewendet wird.

Betrachten wir dazu den Verlauf eines Signals im Ohr (nach [131], zur Struktur des Ohres vgl. Abb. 5.12). Von der Ohrmuschel gelangt der Schall in den Gehörgang, wo er durch Resonanz verstärkt wird. Am Ende des Gehörganges trifft er auf das Trommelfell und löst damit eine Bewegung der drei Gehörknöchelchen im Mittelohr (Hammer, Amboss und Steigbügel) aus. Diese leiten das Schallsignal an das mit Flüssigkeit gefüllte Innenohr weiter, wo sich Hör- und Gleichgewichtsorgan befinden. Das Hörorgan wird aufgrund seiner Form als Schnecke (Cochlea) bezeichnet (vgl. Abb. 5.12). Stellt man sich die Cochlea entrollt vor, so erhält man einen Schlauch, der durch die Basilarmembran longitudinal in

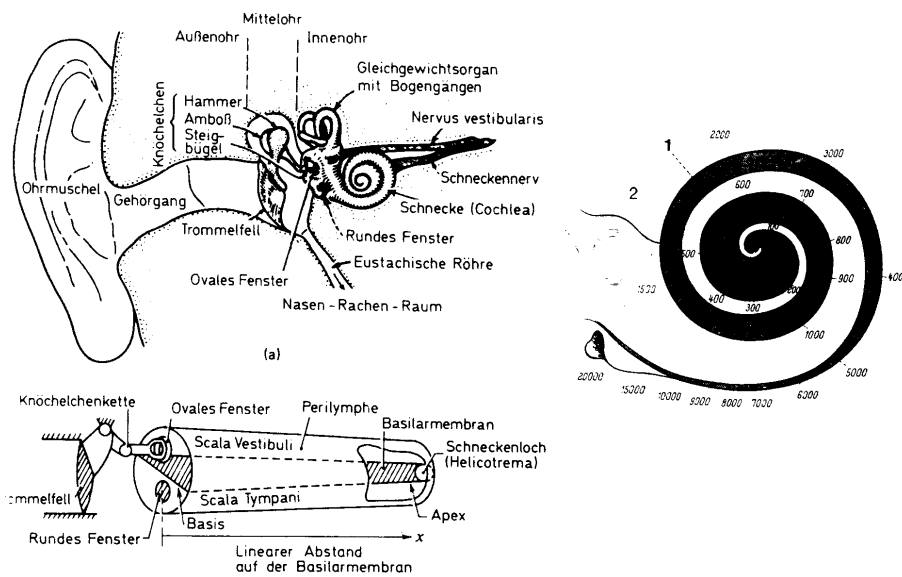


Abbildung 5.12: Aufbau des menschlichen Ohres (links, Teil a) mit schematischer Darstellung der Schnecke (links, Teil b; [97]). Im rechten Teil Schema der Basilar membran in der Schnecke mit Angabe der dazugehörigen Töne in Hz [70]

ein oberes und ein unteres Kompartiment unterteilt ist. Beide Räume werden von der Perilymphe, einer extrazellulären Flüssigkeit durchspült. Die cochleäre Trennwand zwischen den beiden Flüssigkeitskompartimenten ist kompliziert aufgebaut: In der Mitte befindet sich ein flüssigkeitsgefüllter Hohlraum (Scala media), der von oben von einer dünnen Membran (Reissnersche Membran) und nach unten vom Cortischen Organ begrenzt wird.

Ein Schallsignal wird vom Steigbügel zuerst an die Perilymphe des oberen und dann an die des unteren Cholea-Schlauches weitergegeben. Dadurch ergibt sich eine Wellenbewegung der cochleären Trennwand von der Basis zur Spitze (Wanderwelle). Dies ist so ähnlich, wie wenn jemand mit der Handkante auf ein gespanntes Seil schlägt: man sieht eine Welle wandern, die früher oder später abebbt. Im Gegensatz zur Seilwelle nimmt die Wanderwelle im Innenohr jedoch verblüffenderweise zuerst dramatisch zu und verschwindet dann plötzlich, sobald sie an einem charakteristischen Ort entlang der cochleären Trennwand ihr Maximum erreicht hat. Dieser Punkt hängt von der Frequenz des Tones ab: bei hohen Tönen liegt er in der Nähe der Cochleaspitze, bei tiefen Tönen befindet er sich bereits kurz hinter der Steigbügelußplatte. Mittlere Töne bilden sich in der Mitte der cochleären Trennwand ab, d.h. es ergibt sich eine Frequenzdispersion. Damit ist, wie auf der Tastatur eines Klaviers, jeder Ort entlang der Trennwand in der Hörschnecke einem bestimmten Ton zugeordnet. Allerdings ist der Tonumfang nicht so begrenzt wie auf der Klaviertastatur: die Cochlea kann auf begrenztem Raum rund 7 000 Tonhöhen auseinanderhalten. Diese Frequenzselektivität geht ganz wesentlich darauf zurück, dass das Amplitudenmaximum der Wanderwelle nicht rund, sondern sehr spitz ist. Sowohl das Amplitudenmaximum als auch seine präzise Spitze werden von hoch spezialisierten Zellen, den äußeren Haarzellen, im Cortischen Organ erzeugt.

Damit hat das Ohr aber eine ganz andere Funktionsweise als ein AD-Wandler. Während dieser eine Zeitfunktion $f(t)$, also z.B. eines der in Abb. 5.11 gezeigten Oszillogramme, analysieren müsste, unterscheidet das Ohr 'nur' das Vorkommen bestimmter Frequenzen. Die einkommende Information wird also von der Zeitserie $f(t)$ auf das Auftreten bestimmter Töne verdichtet (Bestimmung eines Frequenzspektrums eines Tones). Eine derartige *Informationsverdichtung* ist ein natürlicher Vorgang, der es einem Nervensystem mit begrenzter Kapazität gestattet, mit einer sehr komplexen Umwelt fertig zu werden. Ohne die Verdichtung von Informationseingaben würde das Zentralnervensystem überwältigt werden [125]. Diese Informationsverdichtung beschränkt sich aber nicht nur auf die Reduktion des Zeichenvorrates

*WWWWW WWWWW WWHWH	Sullivan	The lost chord (Lied)
*WWWWW WWWWW WWHHH	Sibelius	Valse triste, 2. Takt, orch. op. 44
*WWWWW WWWWW WHTTT	Strawinski	Petruschka: Tableau, 2. Takt
*WWWWW WWWWW WHTTT	Beethoven	Sinfonie Nr. 5 c-moll, 3. Satz 2. Takt
*WWWWW WWWWW WHTTT	Hummel	Trompetenkonzert E-Dur, 3. Takt
*WWWWW WWWWW WHTWT	Verdi	Requiem aeternam
*WWWWW WWWWW WHHTW	Berlioz	Requiem Nr. 8, Hostias
*WWWWW WWWWW WHHWT	J. Strauss jr.	Geschichten a. d. Wienerwald Nr. 5, 1. Takt
*WWWWW WWWWW WHHWW	Beethoven	Sinfonie Nr. 7 A-Dur, 2. Satz 1. Takt
*WWWWW WWWWW HTTWT	Mozart	Serenade E-Dur, 1. Satz 1. Takt, K 375

*THTHH TTHHT THTHH	Strawinski	Petruschka: Tanz der Zigeunerinnen
*THTHH TTHHT THHTH	Strawinski	Sinfonie in 3 Sätzen, 2. Satz 3. Takt
*THTHH TTHHT HTHHT	Händel	Cembalo-Suite Nr. 5 E-Dur, 3. Satz
*THTHH TTHHT HTHHT	D'Indy	Sinfonie ü. Hirtengesang, 3. Satz 1. Takt
*THTHH TTHHT HHHHT	Schubert	Sinfonie Nr. 8 h-moll, „Unvollendete“, 1. Satz 2. Takt, D 759
*THTHH TTHHW TWTWT	Schubert	Klaviersonate A-Dur, 2. Satz 1. Takt, D 959
*THTHH TWHTT HHTHT	Mahler	Lieder & Gesänge aus der Jugendzeit, Nr. 2, Erinnerung
*THTHH TWTHH TWHTH	Mozart	Flötenkonzert Nr. 2 D-Dur (Oboe C-Dur), 2. Satz, K 314
*THTHH TWTHH WTTHT	Sibelius	Rückkehr v. Limminkainen, 3. Takt, op. 27/4
*THTHH TWWTH THHTW	Beethoven	Bläser-Oktet E-Dur, allegro, op. 103

Abbildung 5.13: Auszug aus Parsons Dictionary of Tunes [120]

(eben 7 000 Töne statt der Zeitserie), sondern umfasst auch eine Vielzahl anderer Vorgänge im Zentralnervensystem, dazu gehören die Bildung von Autokorrelationen zum Erkennen von Intervallen, die Bildung von Kreuzkorrelationen zum räumlichen Hören und weiteres [97].

Die Bedeutung der Musik als Informationsträger lässt sich auch in anderem Zusammenhang darstellen. Die volle Zeitserie $f(t)$ enthält die größte Informationsmenge. Dennoch lassen sich Musikstücke (unter Verwendung von Tönen bzw. Tonsprüngen, also nicht durch einfache Angabe des Titels) durch sehr geringe Zeichenfolgen eindeutig identifizieren. Ein Beweis dafür, dass mit wenigen Tonfolgen große Informationsmengen überliefert werden können, ist das Dictionary of Tunes von Parson [84]. Stellen Sie sich die Situation vor: Sie hören eine Melodie oder summen sie vor sich hin, können aber beim besten Willen weder auf den Titel noch den Komponisten des Stückes kommen. Um sie in Parsons Melodienverzeichnis zu finden, müssen sie nur die Abfolge der ersten 15 Noten kennen. Dabei kommt es aber nicht auf die musikalische Notierung an, sondern Sie müssen nur wissen, ob die folgende Note tiefer, höher oder gleich hoch ist. Diese Codierung ist Kontext abhängig. Für die Melodie von ‘He-jo, spannen den Wagen an’ würde sich dann die Folge *THWWWTHWHWHWWWWT ergeben, wobei der * den ersten Ton angibt und H, T, W, ob der folgende Ton tiefer, höher oder eine Wiederholung des vorhergehenden ist. Diese kurze Kette ist ausreichend, um ein bestimmtes Stück aus einem Repertoire von mehr als 10 000 klassischen und 6 000 populären Themen zu identifizieren (wenn alle Tonfolgen auftreten könnten, könnte man mit 15 Zeichen aus einem drei-elementigen Zeichenvorrat $3^{15} = 14.35$ Mio verschiedene Kombinationen darstellen). Abbildung 5.13 zeigt zwei Ausschnitte aus diesem Verzeichnis. Auch wenn die Mächtigkeit dieses einfachen Ordnungs- und Identifikationssystems sicherlich beeindruckend ist, ein Blick auf Abb. 5.13 ist ähnlich unbefriedigend wie einer auf Abb. 5.11 (hier waren übrigens oben Mendelssohns Violinkonzert, in der Mitte das Stimmen im Sinfonieorchester und im unteren Teilbild das Geräusch des Publikums beim Betreten des Saales gezeigt): zwar sind die Informationen vollständig (im Falle von Abb. 5.11) bzw. sehr geschickt verpackt (im Falle von Abb. 5.13), aber in beiden Fällen ist die bewusste Wahrnehmung im Gehirn des Betrachters eine völlig andere als im Falle des Hörens des Geräusches bzw. Musikstückes. Insbesondere ist ein ‘Wiedererkennen’ entweder praktisch unmöglich (Abb. 5.11) oder sehr erschwert (Abb. 5.13). Allerdings sind in letzterem Code auch praktisch alle Redundanzen beseitigt, die ja auch beim normalen Lesen die Aufnahme des Textes erleichtern (vgl. die

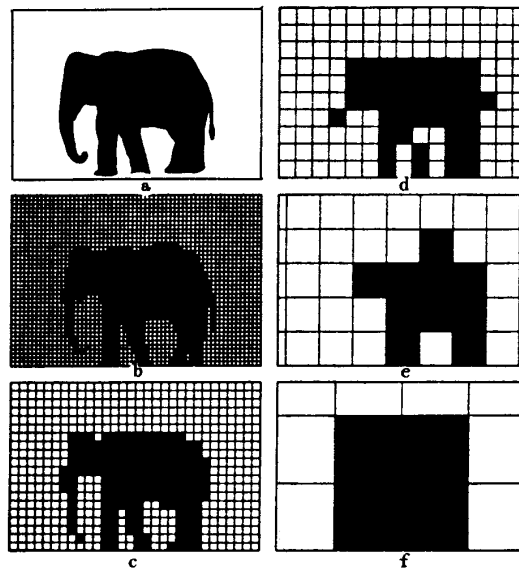


Abbildung 5.14: Bildvorlage mit unterschiedlicher Rasterung: (a) durch das Druckverfahren bestimmte, sehr feine Rasterung mit einer großen Zahl von Bildpunkten, (b) 2400 Pixel, (c) 600 Pixel, (d) 150 Pixel, (e) 37.5 Pixel und (f) 9.4 Pixel [114]

Beispiele zur künstlichen Sprache).

Aber nicht nur das Ohr komprimiert die einfallende Information. Bevor diese Information bewusst aufgenommen werden kann, wird sie teilweise analysiert, insbesondere unter der Fragestellung, ob sich eine Entschlüsselung der Botschaft überhaupt lohnt. Besonders deutlich wird dies z.B. auf einer Party oder in anderen Situationen mit starkem akustischem Hintergrund. Dabei kann es geschehen, dass beide Ohren unterschiedliche Informationen empfangen. Die mit dem einem Ohr empfangene Information kann dann zurückgewiesen werden, obwohl sie technisch ‘gehört’ wurde. Der Unterschied zwischen den beiden unterschiedlichen Nachrichten besteht also nicht darin, dass sie von den Ohren unterschiedlich empfangen wurden, sondern dass die eine ins Gedächtnis ‘übertragen’ wurde, die andere aber nicht. Das bedeutet: wir können Nachrichten empfangen, ohne dass sie in das Gedächtnis übergehen. Man bezeichnet diesen Vorgang auch als selektive Wahrnehmung. Das Ohr, das nicht ‘erhört’ wird, scheint aber von einem Teil des Gehirn relativ genau ausgewertet zu werden. Nur so lässt sich z.B. verstehen, warum der Filter der selektiven Wahrnehmung in der Muttersprache wesentlich besser funktioniert als in einer Fremdsprache (und mit wachsendem Gewohnheitsgrad der Fremdsprache besser wird).

Eine noch wesentlich stärkere Selektion der einströmenden Information muss vom Sehsystem vorgenommen werden. Der auf die Netzhaut einströmende Informationsfluss beträgt ungefähr 10 Millionen bps, von denen das Gehirn aber nur ca. 40 bps verarbeiten und nur wenige bit in das Langzeitgedächtnis überführen kann. Ähnlich dem Hörsinn erfolgt auch beim Gesichtssinn eine erhebliche Nachrichtenkompression und -selektion bereits in der ersten Station des Sehsystems, in der Netzhaut.

Die Informationsreduktion des Auges lässt sich leicht anhand von Abb. 5.14 illustrieren, die auch gleichzeitig als Illustration zum Thema Mustererkennung verwendet werden kann. Im linken oberen Teilbild ist ein Elefant in der durch die Druckqualität bedingten Auflösung gezeigt. Ein Film z.B. würde für die dieser Auflösung entsprechende Pixelzahl jeweils die Information schwarz oder weiß zu übertragen haben, d.h. die Informationsmenge entspräche dem Logarithmus dualis der Pixelzahl. Geht man zu Abbildung (b), so ist die Pixelzahl auf 2 400 ($\lg 2\,400 = 11.1$) verringert, der Elefant aber immer noch eindeutig zu erkennen. Verringern wir die Auflösung weiter auf 600 Pixel (Abbildung (c), $\lg 600 = 9.1$), so ist der Elefant ebenfalls noch zu erkennen. Verringern wir die Pixelzahl weiter auf 150 Pixel (Ab-

bildung (d), ld 150 = 7.1), so ist es bereits schwieriger, den Elefanten zu erkennen. Haben wir aber zusätzlich Informationen, z.B. dass das Signal ein Tier darstellen soll, so kann eine Identifikation weiterhin möglich sein. Erst beim Übergang zu noch größeren Gittern ist der Informationsgehalt zu gering, um die Nachricht 'Elefant' noch zu decodieren, in (e) könnte man auch versucht sein, die Form durch einen Turnierritter mit angelegter Lanze zu beschreiben. Allerdings sind die Abbildungen (a), (b), (c) und (d) nur im Hinblick auf die Nachricht 'Elefant' identisch. Eine Unterscheidung zwischen indischem und afrikanischem Elefanten oder genauere Informationen über den speziellen Elefanten in der Abbildung (Größe, Geschlecht, auffällige Merkmale) setzen eine hohe Auflösung (und damit einen größeren Informationsgehalt) voraus.

Allerdings ist die Informationstheorie nicht nur wichtig, wenn es um den Nachrichtenempfänger 'Mensch' und die Aufnahme und Verarbeitung der Zeichen durch die menschlichen Sinne und das Gehirn geht. Die eben angesprochenen Bereiche des Erkennens, sowohl im akustischen als auch im optischen Kanal, werden auch in der Mustererkennung behandelt. Hierbei geht es darum, einem Computer durch optische oder akustische Eingabe Anweisungen zu erteilen. Während die optische Mustererkennung bereits recht gut entwickelte Verfahren hat, die auch teilweise in die Bearbeitung der von einem Satelliten gemachten Bilder Eingang findet, ist die Spracherkennung noch sehr rudimentär. Das Problem liegt darin, dass das menschliche Ohr zwar einen geringeren Informationsfluss empfängt als das Auge, diese Informationen aber weniger einfach zu verarbeiten sind. Insbesondere sind in der Sprache die gleichen Buchstaben von Sprecher zu Sprecher unterschiedlich (mit starken regionalen Unterschieden), wobei unterschiedliche Buchstaben in ihren Signaturen soweit voneinander abweichen, dass die individuellen Unterschiede zwischen den verschiedenen Darstellungen des Buchstaben meistens kleiner sind als die Unterschiede zwischen zwei verschiedenen Buchstaben. Daher ist es kein Problem, für einen PC oder anderen Rechner einen halbwegs anständig arbeitenden Scanner mit Texterkennung zu finden. Spracheingabe in den Computer dagegen hat sich bisher, trotz intensiver Bemühungen, noch nicht als hinreichend zuverlässiges System entwickeln können.

5.2.8 Zusammenfassung

In diesem Abschnitt haben wir Information als das Unbekannte bzw. Unerwartete kennengelernt. Die 'Bedeutungsschwere' einer Nachricht ist dabei unerheblich, es zählt nur, was sie beim Empfänger an Unsicherheit oder Unwissen beseitigt. Formal wird Information mit Hilfe der Wahrscheinlichkeitsrechnung oder Statistik beschrieben, eine mathematische Formulierung, die den Aspekt des Unerwarteten in der Information noch unterstreicht. Allerdings zeigt diese Definition auch, dass der Informationsgehalt einer Nachricht eine subjektive Größe ist, die einerseits vom Kenntnisstand des Empfängers abhängt, andererseits aber auch von den physikalischen Eigenschaften des Empfängers. So wird zur digitalen Speicherung eines Musikstückes eine wesentlich größere Zahl von bits benötigt, d.h. ein größerer Informationsfluss erzeugt, als das Ohr beim Hören der gleichen Zeitreihe, allerdings unter anderen Verarbeitungsprinzipien, erkennen könnte. Nicht alle Zeichen, die wahrgenommen werden, tragen auch Information bzw. den vollen Informationsgehalt. Dies wird durch das Konzept der Redundanz oder Weitschweifigkeit beschrieben. Ist die Codierung der Information optimal, so wird die Redundanz Null. Sprache dagegen ist hochgradig redundant, nur 27% der Sprache tragen wirklich Information. Was einerseits als Verschwendung erscheinen mag, hat andererseits aber auch den Vorteil, dass eine Verständigung auch bei gestörten Kommunikationskanälen noch möglich ist, da die fehlenden Teile aus dem Zusammenhang ergänzt werden können.

5.3 Codierung

Codierung hat zwei Aufgaben: bei der Quellencodierung geht es darum, die Signale einer Quelle in Abhängigkeit von den Eigenschaften dieser Quelle in einen möglichst sparsamen (Binär-)Code umzusetzen. 'Sparsam' bedeutet dabei, mit einer geringst möglichen Zahl an

000	NUL	033	!	066	B	099	e	132	ä	165	ñ	198	ä	231	þ
001	Start Of Header	034	"	067	C	100	d	133	å	166	*	199	Å	232	þ
002	Start Of Text	035	#	068	D	101	e	134	ä	167	°	200	Æ	233	Û
003	End Of Text	036	\$	069	E	102	f	135	ç	168	¿	201	ƒ	234	Ü
004	End Of Transmission	037	%	070	F	103	g	136	è	169	®	202	‰	235	Ù
005	Enquiry	038	&	071	G	104	h	137	e	170	~	203	™	236	Ý
006	Acknowledge	039	*	072	H	105	i	138	è	171	¼	204	€	237	Ÿ
007	Bell	040	(073	I	106	j	139	í	172	½	205	=	238	~
008	Backspace	041)	074	J	107	k	140	î	173	¾	206	≠	239	·
009	Horizontal Tab	042	*	075	K	108	l	141	ï	174	◀	207	≠	240	-
010	Line Feed	043	+	076	L	109	m	142	Ä	175	▶	208	δ	241	±
011	Vertical Tab	044	,	077	M	110	n	143	Å	176	∴	209	∅	242	—
012	Form Feed	045	-	078	N	111	o	144	É	177	∴	210	É	243	¼
013	Carriage Return	046	.	079	O	112	p	145	œ	178	■	211	Ê	244	¶
014	Shift Out	047	/	080	P	113	q	146	Æ	179		212	Ë	245	§
015	Shift In	048	0	081	Q	114	r	147	ö	180	↓	213	ì	246	•
016	Delete	049	1	082	R	115	s	148	ó	181	À	214	í	247	*
017	-- frei --	050	2	083	S	116	t	149	ô	182	Á	215	î	248	*
018	-- frei --	051	3	084	T	117	u	150	ú	183	Â	216	ï	249	--
019	-- frei --	052	4	085	U	118	v	151	û	184	Û	217	ª	250	.
020	-- frei --	053	5	086	V	119	w	152	ÿ	185	Ü	218	«	251	*
021	Negative Acknowledge	054	6	087	W	120	x	153	Ó	186	Ý	219	■	252	*
022	Synchronous Idle	055	7	088	X	121	y	154	Ü	187	Þ	220	■	253	*
023	End Of Transmission Block	056	8	089	Y	122	z	155	ø	188	ß	221	¿	254	■
024	Cancel	057	9	090	Z	123	{	156	é	189	¸	222	À	255	
025	End Of Medium	058	:	091	[124		157	ø	190	¥	223	■		
026	Substitute	059	;	092	\	125	}	158	×	191	Œ	224	Ó		
027	Escape	060	<	093]	126	~	159	f	192	Ł	225	ß		
028	File Separator	061	=	094	^	127	o	160	á	193	Ł	226	Ó		
029	Group Separator	062	>	095	_	128	ç	161	í	194	Œ	227	Ó		
030	Record Separator	063	?	096	`	129	u	162	ó	195	Œ	228	ó		
031	Unit Separator	064	@	097	a	130	é	163	ú	196	—	229	Ó		
032		065	A	098	b	131	ä	164	ñ	197	†	230	µ		

Tabelle 5.2: ASCII-Code

Bits auszukommen, d.h. die Zahl der benötigten Bits soll den Informationsgehalt der Quelle (in bits) nicht überschreiten. Diese komprimierende Codierung erfolgt durch Verringerung bzw., wenn möglich, vollständige Entfernung der Redundanzen.

Eine andere Form der Codierung, die Kanalcodierung, expandiert diesen komprimierten Code wieder: im optimal komprimierten Code können Fehler weder erkannt noch ausgebessert werden. Hier führt die Kanalcodierung, in Abhängigkeit von den Eigenschaften des Übertragungskanal, gezielt Redundanzen ein, die bei minimaler Expansion des Codes eine maximale Fehlererkennung und -korrektur ermöglichen sollen.

5.3.1 Quellencodierung

Definition 5 Bei der Quellencodierung geht es darum, die Signale einer Quelle in Abhängigkeit von den Eigenschaften dieser Quelle so in einen Code umzusetzen, dass mit möglichst wenigen Bit möglichst viele bit ausgedrückt werden.

Präfix-Code und Decodierbarkeit

Ein Codierungsverfahren, das allgemein bekannt ist, ist der ASCII-Code (vgl. Tab. 5.2), der mit 7 (bzw. in der neueren Form 8) Bits (Bit mit großem B als Binärzeichen, nicht zu verwechseln mit bit mit kleinem b als Maß für den Informationsgehalt) die Zeichen einer Schreibmaschinen- bzw. Computertastatur darstellt. Hierbei handelt es sich um ein typisches Beispiel für einen Blockcode, d.h. die einzelnen Zeichen werden unabhängig von ihrer Häufigkeit stets mit der gleichen Anzahl von Bits codiert. Blockcodierung ist für die Datenverarbeitung aufgrund der stets gleichen Länge von Zeichen besonders gut geeignet. Denkt man jedoch an die Speicherung und Übertragung von Information, so zeigt die Blockcodierung eine deutliche Schwäche: im Mittel sind wesentlich mehr Bits zur Informationsdarstellung aufzuwenden, als vom Informationsgehalt her notwendig wären.

Daher versucht man, Codes zu entwickeln, in denen die einzelnen Zeichen unterschiedliche Längen haben, wobei die häufigsten Zeichen die geringste Länge haben. Zur Illustration betrachten wir dazu den vierelementigen Zeichenvorrat A, wie er in Tabelle 5.3 dargestellt ist. Bei den beiden dort vorgeschlagenen Codierungsmöglichkeiten handelt es sich um einen Blockcode der Länge zwei (Code 1) und einen Code mit unterschiedlichen Längen der Co-

Zeichen	Wahrscheinlichkeit	Code 1	Code 2
a	1/2	00	0
b	1/4	01	10
c	1/8	10	110
d	1/8	11	111

Tabelle 5.3: Zwei Codes für einen vierelementigen Zeichenvorrat

dewörter (Code 2). Der mittlere Informationsgehalt der Zeichen ergibt sich nach (5.10) zu

$$H = \frac{1}{2}\text{ld}(2) + \frac{1}{4}\text{ld}(4) + \frac{1}{8}\text{ld}(8) + \frac{1}{8}\text{ld}(8) = 1.75 \text{ bit} . \quad (5.18)$$

Das ist nach Shannon und Weaver [111] auch die minimale Anzahl von Bits, die bei einer Codierung pro Zeichen benötigt werden. Im Code 1 (Blockcode) ist die Menge der Bits, die für die Darstellung eines Zeichens benötigt werden (Elementarzeichenbedarf) für jedes Zeichen gleich 2, im Code 2 liegt der Elementarzeichenbedarf je nach Zeichen zwischen 1 und 3. So wie sich der mittlere Informationsgehalt durch Mittelwertbildung über alle Zeichen ergibt, so errechnet sich der mittlere Elementarzeichenbedarf als der Mittelwert des Elementarzeichenbedarfs über alle Zeichen. Dieser mittlere Elementarzeichenbedarf m ist bei Code 1 gleich 2, bei Code 2 dagegen ist der Elementarzeichenbedarf eines Zeichens gleich seinem Informationsgehalt. Der mittlere Elementarzeichenbedarf ist daher auch gleich dem mittleren Informationsgehalt, d.h. $m=H=1.75$ bit. Diese Form der Codierung wird als Quellencodierung bezeichnet, da die Zeichen so behandelt werden, wie sie von der Datenquelle geliefert werden, d.h. die Datenquelle und deren Symbolstatistik bestimmen den Code.

Die Güte eines Codierungsverfahrens lässt sich durch die Redundanz bestimmen. Nach der Definition der Redundanz in (5.11) und Einsetzen des mittleren Elementarzeichenbedarfs m lässt sich der *mittlere Mehraufwand je Zeichen* beschreiben als die Redundanz des Codes, d.h.

$$R = m - H . \quad (5.19)$$

Die relative Redundanz ist dann der Mehraufwand des Codes bezogen auf den tatsächlichen Gesamtaufwand:

$$r = \frac{R}{m} = \frac{m - H}{m} = 1 - \frac{H}{m} = 1 - e_f , \quad (5.20)$$

worin der Quotient H/m auch als die Effizienz e_f eines Codes bezeichnet wird.

Definition 6 *Verlust freie Datenkompression bedeutet also Redundanzreduktion.*

Zwar haben wir mit Code 2 die Redundanz reduziert, jedoch sind die einzelnen Zeichen jetzt unterschiedlich lang. Was passiert aber in einem solchen Code, wenn wir verschiedene Zeichen miteinander verknüpfen, d.h. hintereinander schreiben? Bei Blockcodes sind die Zeichen aufgrund der konstanten Zeichenlänge bei der Decodierung einfach voneinander zu trennen, bei unterschiedlichen Wortlängen dagegen ist dieses Decodierungsverfahren nicht sofort eindeutig. Betrachten wir dazu nochmals den Code 2 und geben dem Zeichen b statt des Codes 10 den Code 01. Dann wären zwar weiterhin alle Codezeichen unterschiedlich, allerdings wäre die Decodierbarkeit nicht mehr gegeben. Betrachte dazu die Zeichenfolge 011110: diese könnte mit dem veränderten Code sowohl als 'ada' decodiert werden als auch als 'bc'. Diese Codierung wäre also nicht eindeutig, da der Empfänger die gesendete Information nicht vollständig zurückgewinnen könnte (vgl. Aufgabe 3).

Code 2 in der in Tabelle 5.4 gegebenen Form ist daher so konstruiert, dass derartige Mehrdeutigkeiten nicht auftreten können. Der Code ist ein Präfix-Code, d.h. kein Codewort ist der Anfang (Präfix) eines anderen. Diese Eigenschaft zieht die eindeutige Decodierbarkeit unmittelbar nach sich. Die Umkehrung gilt nicht, d.h. die Präfix-Eigenschaft ist hinreichend aber nicht notwendig für die eindeutige Decodierbarkeit. Am Codierbaum (vgl. Abb. 5.15) zeigt

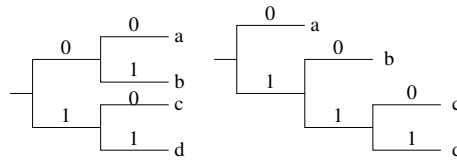


Abbildung 5.15: Codierbäume für die Codes aus Tabelle 5.3

Zeichen	p	P	Block	Shannon	Huffman	Fano
a	0.4	0	000	00	0	00
b	0.2	0.4	001	011	100	01
c	0.15	0.6	010	100	101	10
d	0.15	0.75	011	110	110	110
e	0.05	0.9	100	11100	1110	1110
f	0.05	0.95	101	11110	1111	1111
$H = 2.246$			$m = 3$	$m = 2.8$	$m = 2.3$	$m = 2.35$

Tabelle 5.4: Beispiele für Redundanzreduktion durch verschiedene Codes

sich die Eigenschaft der Decodierbarkeit dadurch, dass die Zeichen nur an den Zweigenden des Baumes sitzen, nicht jedoch in den Verzweigungen.

Präfix-Codes unterschiedlicher Länge werden in der Praxis häufig zur Redundanzreduktion angewendet. So wird z.B. bei der Festlegung von Befehlssätzen (z.B. Maschinensteuerung, Betriebssystem eines Computers) darauf geachtet, dass die häufig vorkommenden Befehle möglichst kurz sind, während seltener auftretende Befehle länger sein dürfen. Die Festlegung von Steuerzeichen in der Datenübertragung folgt den gleichen Regeln.

Redundanzsparende Codes: Shannon, Huffman und Fano

Für die drei folgenden Codierungsverfahren verwenden wir als Beispiel die in Tabelle 5.4 angegebenen sechs Zeichen a, b, c, d, e und f. In der allgemeinen Form besteht der zugrunde gelegte Zeichenvorrat jedoch aus n Zeichen, die jeweils mit den Wahrscheinlichkeiten p_1, p_2, \dots, p_n auftreten. Allen in diesem Abschnitt zu besprechenden Codierungsverfahren ist gemeinsam, dass sie Präfix-Codes sind, bei denen die Wortlänge mit abnehmender Wahrscheinlichkeit des Auftretens des Zeichens zunimmt.

Für die *Codierung nach Shannon* werden diese Zeichen nach abfallenden Wahrscheinlichkeiten ihres Auftretens geordnet:

$$p_1 \geq p_2 \geq p_3 \geq \dots \geq p_n . \tag{5.21}$$

Mit P_k sei nun die kumulierte Wahrscheinlichkeit aller Zeichen, die vor dem k-ten Zeichen liegen, bezeichnet, d.h.

$$P_k = p_1 + p_2 + \dots + p_{k-1} . \tag{5.22}$$

Die Dualzahldarstellung von P_k liefert das Code-Wort für das k-te Zeichen. Die genaue Festlegung des Code-Wortes geschieht dabei nach den folgenden Regeln:

1. Die Codewortlänge m_k des k-ten Zeichens genügt den Ungleichungen

$$- \text{ld}(p_k) \leq m_k < 1 - \text{ld}(p_k) \tag{5.23}$$

oder anders ausgedrückt: m_k ist die kleinste ganze Zahl, für die gilt

$$\frac{1}{2^{m_k}} \leq p_k . \tag{5.24}$$

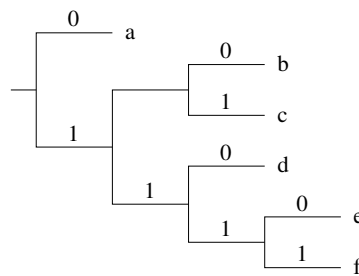


Abbildung 5.16: Codierbaum zum Huffman-Code der in Tabelle 5.3 gegebenen Zeichen

2. Die Entwicklung der Zahl P_k in eine Dualzahl wird nach der m_k -ten Stelle abgebrochen, der Koeffizient der Entwicklung stellt das gesuchte Codewort dar. Das gesuchte Codewort für das k -te Zeichen ist dann durch die Binärzeichen $b_1, b_2, b_3 \dots b_m$ in der Form $b_1b_2b_3\dots b_m$ gegeben, so dass die Zahl

$$Z_k = b_1 \cdot 2^{-1} + b_2 \cdot 2^{-2} + \dots + b_m \cdot 2^{-m} \quad (5.25)$$

möglichst gross wird, aber immer noch kleiner als P_k ist, d.h. es gilt

$$P_k - 2^{-m} < Z_k \leq P_k. \quad (5.26)$$

Der Shannon-Code ist ein Präfix-Code. Damit ist er auch eindeutig decodierbar. In Tabelle 5.3 ist die Codierung nach Shannon auf das Beispiel unseres eingangs erwähnten Zeichenvorrats mit sechs Zeichen angewandt. Die Redundanzreduktion ist erkennbar, die beiden anderen in der Tabelle angegebenen Codierungsverfahren weisen jedoch eine größere Redundanzreduktion auf. Dennoch erlaubt der Shannon-Code eine beliebig gute Approximation an die ideale und damit redundanzfreie Codierung (s.u.).

Für die *Codierung nach Huffman* bedient man sich ebenfalls der Ordnung der Zeichen nach fallenden Wahrscheinlichkeiten. Die Codewörter werden dann mit Hilfe eines Codierbaumes bestimmt. Da es sich um einen Präfix-Code handelt entspricht jedes Zeichen einem der n Endzweige des Codierbaumes.

Die Codierung beginnt mit den Zeichen geringster Häufigkeit (vgl. den Codierungsbaum in Abb. 5.16, der sich auf das Beispiel in Tabelle 5.4 bezieht). Wähle zunächst die beiden seltensten Zeichen aus (im Beispiel e und f). Die zugehörigen Codewörter werden dann in der letzten Stelle durch die Bits 0 und 1 unterschieden. Die beiden seltensten Zeichen werden damit zu den Endpunkten einer Verzweigung im Codierbaum. Diese Verzweigung entspringt in einem Zweig, der von nun an als Endzweig gilt. Diesem Zweig ordnen wir die Wahrscheinlichkeit zu, die sich aus der Summe der Wahrscheinlichkeiten der beiden abgehenden Zweige ergibt. Die zwei seltensten Buchstaben (e und f) werden dabei also zu einem neuen Zeichen (ef) zusammengefasst, das eine Wahrscheinlichkeit von $0.05 + 0.05 = 0.1$ hat. Die obige Betrachtung wird jetzt für einen Codierbaum wiederholt, der nur noch $n - 1$ Endzweige hat. Dabei fasst man nun wieder die beiden seltensten Endzweige zusammen (das sind in dem Beispiel d und der bereits zusammengefasste ef-Zweig). Auf diese Weise arbeitet man sich von den Endzweigen bis zur Wurzel des Codierbaumes vor.

Die *Codierung nach Fano* ist von der Idee her eine Umkehrung der Codierung nach Huffman: hier geht man von der Baumwurzel aus und hangelt sich zu den Zweigen vor. Die erste Verzweigung entspricht dabei einer Aufteilung des Zeichenvorrates in zwei Mengen mit möglichst gleich großen Wahrscheinlichkeiten. Die eine Menge enthält dabei die wahrscheinlicheren Zeichen, die andere die unwahrscheinlicheren. Auf diese Weise ist das am weitesten links stehende Bit der Codierung festgelegt. Die beiden Äste der Verzweigung werden nun als Wurzeln für Codierbäume der beiden Untermengen aufgefasst.

Betrachtet man für die Codierung nach Fano noch einmal das Beispiel in Tabelle 5.4, so ergeben sich zwei Möglichkeiten der Codierung: Die erste Aufspaltung in je eine Unterliste mit wahrscheinlicheren und unwahrscheinlicheren Zeichen kann wahlweise zwischen dem b und dem c erfolgen oder zwischen dem a und dem b (in beiden Fällen ist die Wahrscheinlichkeit in

der einen Untermenge 0.6 und in der anderen 0.4). Im ersten Falle ergibt sich der in Tabelle 5.4 gegebene Fano-Code, im zweiten Falle hätte sich der ebenfalls in Tabelle 5.4 gegebene Huffman-Code ergeben.

Quellencodierung im allgemeineren Fall: Fundamentalsatz der Quellencodierung

Die Fano-Codierung beseitigt die Redundanz einer Zeichenquelle nur dann völlig, wenn sich die Zeichen der Quelle zur eindeutigen Identifizierung immer in zwei gleichwahrscheinliche Hälften teilen lassen (das war im Beispiel in Tabelle 5.3 der Fall, nicht jedoch in dem Beispiel in Tabelle 5.4). Lässt sich die Redundanz auch in den Fällen beseitigen, in denen diese Voraussetzung nicht gegeben ist?

Auch in diesem Falle lässt sich die Redundanz durch geeignete Codierung beliebig reduzieren [30, 110], die entscheidende Idee dabei ist, nicht die einzelnen Zeichen zu codieren, sondern Zeichen x_i zu Zeichenketten X der Länge N zusammenzufassen. Diese Zeichenketten werden dann codiert. Die Quelle habe dazu n voneinander unabhängige Zeichen, die mit den Wahrscheinlichkeiten p_i erzeugt werden. Die Wahrscheinlichkeit p einer aus dieser Quelle erzeugten zusammengesetzten Zeichenkette ist dann näherungsweise gegeben durch

$$p = \prod_{i=1}^n p_i^{N p_i} . \quad (5.27)$$

Für den in der Zeichenkette enthaltenen Informationsgehalt gilt damit

$$-\text{ld}(p) = -\sum_{i=1}^n N p_i \text{ld}(p_i) = -N \sum_{i=1}^n p_i \text{ld}(p_i) = N H(X) . \quad (5.28)$$

Entsprechend ist die Wahrscheinlichkeit jeder dieser Zeichenketten ungefähr gegeben durch

$$p = 2^{-N H(X)} , \quad (5.29)$$

wobei die Wahrscheinlichkeiten für jede Zeichenkette ungefähr gleich groß sind.

Die Gesamtwahrscheinlichkeit dieser Zeichenketten kann durch eine genügend große Länge N beliebig nahe an 1 gebracht werden, daher muss die Anzahl dieser annähernd gleichwahrscheinlichen Zeichenketten ungefähr gleich $2^{N H(X)}$ sein, es werden also $N H(X)$ bit benötigt, um anzugeben, welche von diesen Zeichenketten vorliegt. Auf jedes der N Zeichen einer derart codierten Zeichenkette entfallen dann nur $H(X)$ bit anstelle der $H_0(X) = \text{ld}N$ bit.

Andererseits sind diese $2^{N H(X)}$ annähernd gleich wahrscheinlichen Zeichenketten nur ein kleiner Bruchteil aller n^N Zeichenketten der Länge N . Ihr relativer Anteil ist gegeben durch

$$\frac{n^N}{2^{N H(X)}} = 2^{N(\text{ld}(n) - H(X))} . \quad (5.30)$$

Diese zahlreicheren fast n^N Zeichenketten lassen sich mit jeweils $N \text{ld}(n) \geq N H(X)$ bit codieren. Da aber die Gesamtwahrscheinlichkeit dieser vielen Zeichenketten durch eine ausreichend lange Wortlänge N beliebig klein gemacht werden kann, tragen sie nur vernachlässigbar wenig zum mittleren Zeichenbedarf bei. Daraus ergibt sich der *Fundamentalsatz der Quellencodierung*:

Definition 7 Fasst man eine zu codierende Nachricht in genügend langen Zeichenfolgen zusammen, dann lassen sich diese so codieren, dass die Redundanz beliebig klein wird.

Die hierbei eingeführten Zeichenketten müssen ziemlich lang sein, damit ihre Gesamtwahrscheinlichkeit ausreichend gewährleistet ist. Damit müssen viele Quellzeichen zu einer Zeichenkette zusammengefasst werden, was auf lange Codewörter führt. Damit zeigt sich auch eine Voraussetzung des Verfahrens: Die Originalnachrichten müssen eine ausreichende Länge haben.

Lempel-Ziv-Quellencodierung

Huffman-Codes Huffman-Codierung sind optimale Codes – allerdings nur unter der Voraussetzung, dass die Wahrscheinlichkeiten der Signale der Quelle (Quellenstatistik) bekannt sind. Will man einen üblichen Text aus einer bekannten Sprache codieren, so ist dessen Quellenstatistik bekannt.

Anders sieht die Situation jedoch bei unbekanntem Daten aus. Eine mit einem Imager aufgenommene Szene ordnet jedem Pixel einen von 255 möglichen Helligkeitswerten zu – aber welche Verteilung haben die Helligkeitswerte? Meeresoberfläche bei Nacht, Prärie im Mittelwesten oder die Sahara werden sicherlich jeweils unterschiedliche Histogramme der Helligkeitswerte liefern, also auch unterschiedliche Quellenstatistik haben. Dann benötigt jede dieser Szenen einen eigenen Huffman-Code – oder man muss sehr verschwenderisch über alle möglichen Szenen mitteln und aus dieser Quellenstatistik einen zwar für den Gesamtdatensatz optimierten, für jede einzelne Szene jedoch suboptimalen Code entwickeln.

In solchen Fällen ist ein von der Quellenstatistik unabhängiger Codierungsalgorithmus hilfreich. Der Lempel-Ziv-Algorithmus ist ein Beispiel für einen derartigen universellen Quellencodierungsalgorithmus. Lempel-Ziv ist ein *Variabel-zu-Festlänge-Codierungssystem*: eine Sequenz der Quellen wird in Phrasen variierender Länge unterteilt, die ihrerseits jedoch in Codewörter gleicher Länge codiert werden.

Was bestimmt aber die Länge einer Phrase? Die Gliederung erfolgt durch die Identifikation von Phrasen der kleinsten Länge, die bisher noch nie aufgetreten sind. Nach Beginn einer Phrase wird erst dann wieder eine neue Phrase begonnen, wenn das gerade ausgegebene Zeichen dazu führt, dass die so entstandene Phrase nicht mehr mit einer bereits bekannten Phrase übereinstimmt. Führt das letzte Zeichen dazu, dass sich die so entstandene Phrase von den vorher bekannten unterscheidet, so wird sie als neue Phrase erkannt: eine neue Phrase ist also die Verkettung aus einer bekannten Phrase und einem neuen Quellenausgang. Um diese zu codieren werden binäre Expansionen der lexikographischen Ordnung der bisherigen Phrase mit dem neuen Bit verkettet. Auf diese Weise wachsen die Phrasen.

Betrachten wir das Ausgangssignal

```
0100001100001010000010100000110000010100001001001
```

Dieses Signal wird in Phrasen gegliedert:

```
0||1||00||001||10||000||101||0000||01||010||00001||100||0001||0100||0010||01001
```

Zur Codierung dieser 16 Phrasen benötigen wir jeweils 5 Bit (4 Bit für die Phrase sowie ein zusätzliches Bit für den neuen Quellenausgang). Die Codierung ist in Tabelle 5.5 gegeben.

Aus diesem Beispiel wird der Vorteil des Algorithmus nicht deutlich, da die ursprüngliche Sequenz von 49 Zeichen (entsprechend 49 Bit) in 16 Phrasen zerlegt und daher mit 80 Bit codiert wurde. Dieses Phänomen erklärt sich aus der Kürze der Testsequenz, bei einer längeren Testsequenz komprimiert der Algorithmus wesentlich besser.

Lempel-Ziv ist die Grundlage vieler Datenpacker, z.B. ZIP, ZOO, ARJ usw. Modifikationen von Lempel-Ziv werden auch bei der Kompression von Bildern verwendet, vgl. Abschn. 6.4.1.

Beispiel: Redundanzreduktion beim Fernkopieren

Fernkopieren (Faxen) ist eine Form der Datenübertragung, die sich sehr leicht und sehr platzsparend codieren lässt. Die vollständige Information des Bildes ergibt sich aus der Zahl der zu seiner Darstellung benötigten Pixel (Punktauflösung, vgl. Abb. 5.14) und der Angabe, ob der jeweilige Bildpunkt schwarz oder weiß ist. Die Informationsmenge wäre also insgesamt gleich der Zahl der Bildpunkte (die Information schwarz oder weiß entspricht ja genau der Informationsmenge von einem bit pro Pixel). Im Gegensatz zu einer farbigen Darstellung hat man beim Schwarz-Weiß-Bild aber den Vorteil, dass die Farbe von Bildpunkt zu Bildpunkt

Wörterbuchposition	Phrase	Codewort
1	0001	0 0000 0
2	0010	1 0000 1
3	0011	00 0001 0
4	0100	001 0011 1
5	0101	10 0010 0
6	0110	000 0011 0
7	0111	101 0101 1
8	1000	0000 0110 0
9	1001	01 0001 1
10	1010	010 1001 0
11	1011	00001 1000 1
12	1100	100 0101 0
13	1101	0001 0110 1
14	1110	0100 1010 0
15	1111	0010 0100 0
16		1110 1

Tabelle 5.5: Beispiel zum Lempel-Ziv-Algorithmus [88]

nur relativ selten wechselt. Das gilt für sehr feine Auflösung (Raster) natürlich noch stärker als für grobe Raster (vgl. Abb. 5.14).

Bei der Lauflängencodierung wird jede Punktfolge bestimmter Farbe und Länge (run length, Lauflänge) als Wort eines Zeichenvorrates aufgefasst. Jede Bildzeile ist dann eine Folge solcher Lauflängen. Die Lauflängen werden binäre codiert, wobei jede schwarze und jede weiße Lauflänge ein eigenes Codewort erhält. Die Wahl der Codewörter geschieht so, dass der mittlere Binärzeichenbedarf je Lauflänge möglichst gering wird. Dafür gibt es redundanzsparende Codes, z.B. die Codierung nach Huffman oder eine Variante davon, die sogenannte modifizierte Huffman-Codierung. Je nach Bildvorlage sind dann nur noch 5 bis 20% des Aufwandes gegenüber der Punkt-für-Punkt-Übertragung nötig. Einige Beispiele für den modifizierten Huffman-Code zur Bildübertragung sind [103]:

```
s1 → 010
s2 → 11
s3 → 10
s4 → 011
... → ...
w6 → 1110
w7 → 1111
w8 → 10011
w9 → 10100
... → ...
```

Stellen wir uns jetzt eine aus 33 Zeichen bestehende Zeile vor. Die Schwarz-Weiß-Folge dieser Zeile ist gegeben durch

```
000000001100000001111000000000111
```

wobei die 0 einen weißen, die 1 einen schwarzen Bildpunkt markiert. In Lauflängen lässt sich die Folge dann beschreiben als w8, s2, w7, s4, w9 und s3. Mit dem oben gegebenen modifizierten Huffman-Code lässt sich dies verschlüsseln zu der Bitfolge

```
100111111110111010010.
```

Im Verhältnis zur ursprünglichen Folge fällt die Verkürzung hier natürlich noch nicht besonders beeindruckend aus (21 Bit statt der ursprünglich 33 Bit), das liegt aber im wesentlichen

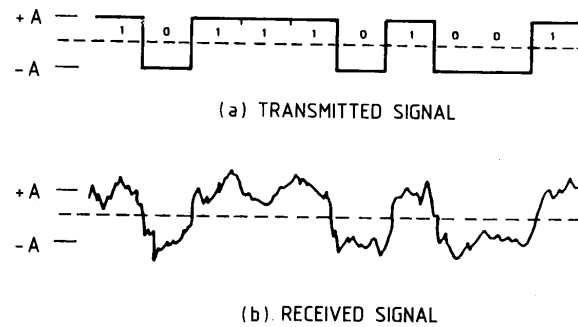


Abbildung 5.17: Binärsignal 1011101001 am Sender (oben) und beim Empfänger (unten) für einen gestörten Übertragungskanal [8]

daran, dass im Verhältnis zu einer ‘realen’ Fax-Vorlage die Lauflängen im obigen Beispiel recht kurz waren.

5.3.2 Kanalcodierung

Bisher haben wir uns damit befasst, wie man Information möglichst ökonomisch verpacken kann, um bei einer Speicherung oder Übertragung dieser Informationen Platz bzw. Zeit zu sparen. Im Interesse der Datenkompression wurde also Redundanz so weit wie möglich ausgeschaltet. Im Folgenden werden wir lernen, dass man andererseits dem so codierten Signal wieder Redundanzen zufügt, um eine Fehlererkennung und möglichst auch -korrektur zu ermöglichen.

Informationsübertragung

Dabei stellt sich die Frage: Lohnt sich der Aufwand, Redundanzen zur Fehlererkennung zuzufügen (um den Preis einer wesentlich aufwendigeren = teureren Datenübertragung), oder ist die Fehlerwahrscheinlichkeit bzw. -häufigkeit so gering, dass wir die vereinzelt auftretenden Fehler in Kauf nehmen können? Diese Frage wird durch die Eigenschaften des Übertragungskanals zwischen einem Raumfahrzeug, insbesondere einer planetaren Mission (z.B. Voyager), unmittelbar deutlich: große Abstände zwischen Sender und Empfänger bei gleichzeitiger geringer Sendeleistung führen zu einem geringen Signal-Rausch-Verhältnis, d.h. die Wahrscheinlichkeit eines Übertragungsfehlers wird entsprechend groß. Zusätzlich ist die Atmosphäre mit ihren hochgradig variablen Eigenschaften ein Bestandteil des Übertragungskanals. Dabei kann die Störung der Ionosphäre durch die zusätzliche Ionisation in der Folge eines Flares die Datenübertragung ausgerechnet zu den Zeiten stark beeinträchtigen, in denen die für die Sonnenforscher interessantesten Daten von den Satelliten gesendet werden.

Ein Beispiel für die Relation zwischen gesendeten und empfangenen Daten ist in Abb. 5.17 gezeigt. Im oberen Teil ist ein Binärsignal in ungestörter Form aufgetragen, im unteren Teil das sich bei einem gestörten Kommunikationskanal am Empfänger ergebende Signal. Man erkennt dabei das Risiko, dass bei starken Störungen einzelne Bits oder gar Folgen von Bits falsch identifiziert werden. Insbesondere starke Störungen durch erhöhte Ionisation oder starken Regen sind im Verhältnis zu den Übertragungszeiten sehr lange andauernd, so dass es hierbei häufig zu einer Verfälschung langer Folgen von Bits kommen kann [89].

Wie lässt sich Informationsübertragung beschreiben? Im einfachsten Falle haben der Sender und der Empfänger den gleichen Zeichensatz und für jedes vom Sender ausgewählte, codierte und gesendete Zeichen kann der Empfänger dieses Zeichen unverfälscht erhalten (ungestörte Übertragung), decodieren und identifizieren. Im realen Fall dagegen wird diese Übertragung nicht ungestört erfolgen, so dass die Decodierung und Identifizierung schwierig oder gar unmöglich werden kann.

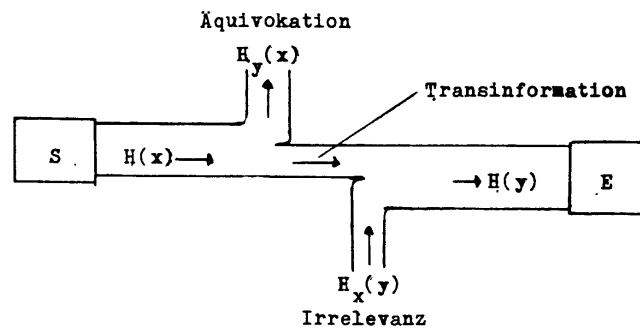


Abbildung 5.18: Entropiefluss im gestörten Übertragungskanal [114]

Abbildung 5.18 stellt dazu den Entropiefluss in einem gestörten Nachrichtenkanal dar. Der Sender gibt aus seinem Zeichenvorrat X Signale mit einer mittleren Entropie $H(X)$ ab, auch als Sender-Entropie $H(X)$ bezeichnet. Der Empfänger wählt daraufhin aus seinem Zeichenvorrat Y ein Zeichen mit der Empfänger-Entropie $H(Y)$. Wegen möglicher Störungen verbleibt ihm dabei aber eine mittlere Unkenntnis $H_Y(X)$ über das Gesendete. Diese Unkenntnis wird als Äquivokation bezeichnet. Andererseits lassen Störungen aber auch eine mittlere Unsicherheit bei der Auswahl der Empfängerzeichen für ein beliebiges gesendetes Zeichen zurück. Diese Unvorhersagbarkeit wird als Irrelevanz $H_X(Y)$ bezeichnet. Die im Mittel pro Zeichen übertragene Information wird dann als mittlere Transinformation bezeichnet.

Kanalkapazität

Bisher haben wir die Signale als diskrete Zeichen betrachtet und uns nur mit ihrer Codierbarkeit und Decodierbarkeit beschäftigt. Bei realen Nachrichten handelt es sich jedoch um Signalfolgen, d.h. Signale, die in einer bestimmten Zeit gesendet werden. Bei der Nachrichtenübertragung ist daher gerade die Geschwindigkeit, mit der Information übertragen werden kann, eine wichtige Kenngröße. Diese bildet z.B. die Motivation zur Quellencodierung, d.h. zur Verringerung der Redundanzen in einer Nachricht, bildet aber gleichzeitig auch eine Grenze für die Kanalcodierung, d.h. das Einfügen von Redundanzen zur Fehlererkennung und -korrektur.

Die Geschwindigkeit der Nachrichtenübertragung lässt sich definieren als der Informationsfluss (Einheit bps, bit pro Sekunde)

$$F = \frac{H}{T_m} . \quad (5.31)$$

Darin ist

$$T_m = \sum p_x T_x \quad (5.32)$$

die mittlere Zeit, die für die Übertragung eines Nachrichtensymbols benötigt wird (Einheit s/Symbol).

Unter der Kanalkapazität verstehen wir nun den maximalen Informationsfluss, der über einen gegebenen Nachrichtenkanal fehlerfrei übertragen werden kann:

$$C = F_{\max} = \left(\frac{H}{T_m} \right)_{\max} . \quad (5.33)$$

Die maximale Kanalkapazität ist alleine durch die Eigenschaften des Übertragungskanals gekennzeichnet aber nicht durch die Eigenschaften der Signalquelle, da wir deren redundanzfreie Codierung bereits vorausgesetzt haben (diese können wir redundanzfrei codieren, wie im vorangegangenen Abschnitt gezeigt; wir müssen sogar redundanzfrei codieren, da wir sonst nur Kanalkapazität verschwenden würden).

Für einen realen, d.h. gestörten Kanal, gilt der *Fundamentalsatz der Kanalcodierung* (Shannon):

Kanal	B [kHz]	ϱ [dB]	C [bit/s]
Telefon	3.1	40	$4.1 \cdot 10^4$
Rundfunk UKW	15	60	$3 \cdot 10^5$
Fernsehen	5000	45	$7.5 \cdot 10^7$
InfoAufnahme Mensch			20

Tabelle 5.6: Übertragungskapazitäten einiger Kanäle

Definition 8 *Es gibt eine Kanalkapazität C , bei der man bei geeigneter Codierung Informationen mit beliebig kleiner Fehlerrate übertragen kann, solange der Informationsfluss kleiner als C ist. Wenn der Informationsfluss der Quelle die Kanalkapazität übersteigt, ist Informationsverlust unvermeidlich.*

Dieser Satz ist nicht offensichtlich, da man erwarten würde, dass mit zunehmenden Störungen eines Kanals auch die Übertragungsfehler größer werden. Das ist auch solange der Fall, wie man keine besonderen Codierungsmaßnahmen ergreift. Shannons Satz besagt jedoch, dass es stets eine Codierung gibt, mit der man die Fehlerrate beliebig verringern kann, unabhängig davon, wie die Störungen geartet sind. Die Suche nach einem geeigneten Code kann dann allerdings schwierig werden. Auch geringe Kanalkapazitäten schränken die Möglichkeit, einen Code für verschwindende Fehlerraten zu finden, ein.

Betrachten wir zum Verständnis dieser letzteren Einschränkung einen Übertragungskanal der Bandbreite B , der durch ein Rauschen mit Gaußscher Amplitudenverteilung gestört ist. P_S sei die Leistung des Signals, P_N die des Rauschens am Empfängereingang. Dann ergibt sich für die Kanalkapazität [40]

$$C = B \operatorname{ld} \left(1 + \frac{P_S}{P_N} \right). \quad (5.34)$$

Der Informationsfluss in einem normalen Übertragungskanal (z.B. Telefon, Rundfunk) bleibt in der Praxis meistens weit hinter der Kanalkapazität zurück. Geht man dann in (5.34) auf den Zehnerlogarithmus über, so ergibt sich

$$C \approx \frac{B}{3} 10 \lg \left(1 + \frac{P_S}{P_N} \right) = \frac{B}{3} \varrho. \quad (5.35)$$

Für $P_S \gg P_N$ ist dann

$$\varrho \approx 10 \lg \left(\frac{P_S}{P_N} \right) \quad (5.36)$$

das Signal-Rausch-Verhältnis. Gerade bei den planetaren Missionen, wie z.B. Pioneer und Voyager, ist dieses Verhältnis aufgrund der großen Abstände gering, so dass die Kanalkapazitäten gering werden und sich damit auch die Möglichkeiten, einen fehlerunempfindlichen Code zu finden, verringern.

Tabelle 5.6 gibt die Kanalkapazitäten für einige Übertragungskanäle. Dabei fällt auf, dass der Mensch das langsamste Glied in der Kette ist, d.h. Information kann wesentlich schneller übertragen werden, als der Mensch sie verarbeiten kann.

Doch wenden wir uns wieder der Übertragungskapazität eines Nachrichtenkanals zu. Ist bei binärer Übertragung die Wahrscheinlichkeit, dass ein Zeichen verfälscht wird (also 0 statt 1 oder umgekehrt) durch p' gegeben, so ist nach Shannon die Kapazität eines derartigen Kanals gegeben durch

$$C = 1 - b(p') \text{ bit/s} \quad (5.37)$$

mit b als der Shannon-Funktion wie in (5.9) gegeben.

Wir können in (5.34) die Leistung P_S des Signals und die Leistung P_N des Rauschens umschreiben, indem wir die Bitzeit T_b , die Bitrate $R_b = 1/T_b$, die spektrale Rauschleistungsdichte N_0 und die Bitenergie E_b einführen [40]. Dann ist

$$\frac{P_S}{P_N} = \frac{P_S T_B R_b}{N_0 B}. \quad (5.38)$$

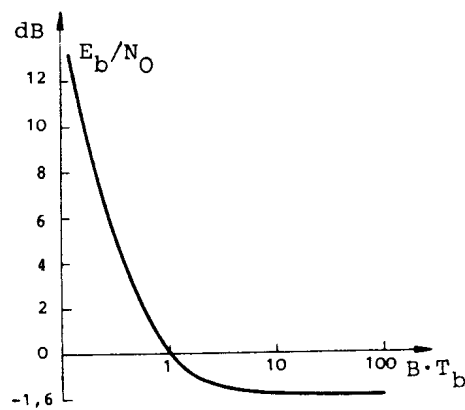


Abbildung 5.19: Shannon-Grenze für fehlerfreien Empfang [40]

Für $C = R_b$, d.h. den optimalen Kanal, wird dann

$$\frac{R_b}{B} = \text{ld} \left(1 + \frac{E_b}{N_0} \frac{R_b}{B} \right). \quad (5.39)$$

Diese Beziehung zwischen der Energie/Bit und der Bandbreite/Bit ist in Abb. 5.19 dargestellt. Mit zunehmender Bandbreite B lässt sich dann bei vorgegebenem Verhältnis E_b/N_0 von Bitenergie zu Rauschleistungsdichte eine größere Bitrate R_b erreichen. Im Extremfall geht dabei $B/R_b \rightarrow \infty$ und es wird

$$\frac{E_b}{N_0} = -1.6 \text{ db} \cong 0.69. \quad (5.40)$$

Das ist die Mindestenergie, die pro Bit und Rauschleistungsdichte aufgebracht werden muss. Für geringere Energien ist eine fehlerfreie Übertragung ausgeschlossen. Dieser Grenzwert gilt übrigens nur für den idealen Codierungsfall (den es zwar gibt, aber der nicht immer bekannt sein muss); für die meisten realen Codierungsverfahren sind diese Anforderungen größer, d.h. es muss mehr Energie zur Übertragung aufgebracht werden.

Aus der Bandbreite und dieser Mindestenergie lässt sich ein Kommunikationswirkungsgrad η_k definieren als

$$\eta_k = \frac{E_b}{N_0 B}. \quad (5.41)$$

Dieser Wirkungsgrad erlaubt den Vergleich verschiedener Codierungs- und Modulationsverfahren und die Auswahl eines optimalen Verfahrens. Marko [76] gibt einen guten und anschaulichen Überblick über den hier dargestellten Zusammenhang von Information und Energie.

Kanalcodierung

Bei der Behandlung von Fehlern bei der Informationsübertragung gibt es zwei prinzipiell unterschiedliche Ansätze:

- eine Fehlererkennung ist ausreichend, bei einem erkannten Fehler wird der Sender zur Wiederholung aufgefordert. Dieses Verfahren ist bei vielen Satelliten nicht praktikabel, da die Daten häufig nicht zwischengespeichert sondern direkt ausgelesen werden.
- Eine Fehlererkennung alleine ist nicht ausreichend, es wird eine zusätzliche Fehlerkorrektur gefordert. Dieses Verfahren ist zwar aufwendiger, wenn aber keine Wiederholung der als fehlerhaft erkannten Nachricht möglich ist, ist es die einzige Möglichkeit, den Informationsverlust gering zu halten.

Für beide Ansätze gibt es Codierungsverfahren, wobei die Codierung zur Fehlererkennung ohne Korrektur selbstverständlich viel einfacher ist (z.B. durch Verwendung eines Paritätsbits

oder einer Prüfzahl, z.B. der Quersumme des Signals) als die Codierung für Fehlererkennung und -korrektur.

Haben wir bisher die Behauptung aufgestellt, dass es stets eine Codierung gibt, mit der die Übertragungsfehler möglichst gering werden können, wollen wir in diesem Abschnitt einige Beispiele für derartige Codierungen betrachten. Diese Kanalcodierung führt zusätzliche Redundanzen ein, wobei gilt:

Definition 9 Ziel der redundanten Codierung ist es, die zuverlässige und sichere Kommunikation über unzuverlässige Kanäle zu ermöglichen.

Bei den folgenden Betrachtungen werden wir uns auf Blockcodes beschränken. Das bedeutet, dass die Quellencodierung nicht wie in den obigen Beispielen mit Code-Wörtern variabler Länge erfolgt, sondern mit festen Blocklängen. Für viele Anwendungen bei Weltraumexperimenten kann die Codierung der Quelle durch einen Blockcode durchaus ausreichend sein. Zusätzlich lassen sich aber auch Codes mit Codewörtern variabler Länge in einem Blockcode unterbringen: Man schreibt die Codewörter einfach hintereinander und trennt nach einer gewissen Länge einen Block ab, unabhängig davon, ob dabei ein Codewort zertrennt wird oder nicht. Zwischen diesen Blöcken befinden sich dann die im Rahmen der Kanalcodierung eingeführten Redundanzen. Da der Empfänger aber sowohl die Länge der Blöcke als auch die der eingefügten Prüfbits kennt, kann er aus dem empfangenen Signal erst das quellencodierte Signal zurückgewinnen und daraus die Originalnachricht decodieren. Daher stellt die Betrachtung von Blockcodes in diesem Abschnitt keine allgemeine Einschränkung oder gar einen Rückschritt dar.

Fehlererkennung

Beginnen wir mit einfachen Möglichkeiten zur Fehlererkennung. Dazu betrachten wir wieder unseren vierelementigen Zeichenvorrat aus Tabelle 5.3, und zwar in der Blockcodierung wie durch Code 1 gegeben. Die einfachste Methode zur Erzeugung eines Codes, in dem Fehler erkannt werden können, besteht darin, jedes Zeichen einfach noch einmal zu senden (Signalverdoppelung). Durch die Verdoppelung unserer Signale benötigen wir dann pro Zeichen statt 2 Bit 4 Bit. Gehen wir davon aus, dass bei der Übertragung nur jeweils 1 Bit verfälscht wird, so können wir jeden Fehler erkennen. Fehlermöglichkeiten sind dann jeweils gegeben durch

Zeichen	Code	Codierung	F_1	F_2	F_3	F_4
a	00	0000	1000	0100	0010	0001
b	01	0101	1101	0001	0110	0100
c	10	1010	0010	1110	1000	1011
d	11	1111	0111	1011	1101	1110

wobei die Spalten F_i jeweils den Fehler im i -ten Bit symbolisieren sollen. Jeder einfache Fehler (d.h. es wird nur ein Bit falsch übertragen) wird erkannt, da das fehlerhafte Signal sich nicht mehr aus zwei gleichen Teilen zusammensetzt. Jedoch kann der Fehler in diesem Code noch nicht korrigiert werden: Erhält der Empfänger das Signal 0010, so kann es sich dabei um ein a handeln (drittes Bit umgeklappt) oder ein c (zweites Bit umgeklappt). Werden mehr als ein Bit fehlerhaft übertragen, so ist mit diesem Verfahren auch die Fehlererkennung nicht mehr gewährleistet, da dann statt a (0000) z.B. b (0101) empfangen werden könnte. Fehlererkennung ist hierbei nur dann möglich, wenn ein Zeichen wie z.B. 1100 empfangen wurde, das weder zum Zeichenvorrat der Codewörter noch zu dem der um 1 Bit verfälschten Codewörter gehört.

Das betrachtete Codierungsbeispiel ermöglicht also bei Verdopplung der Bits nur eine einfache Fehlererkennung, aber keine Korrektur. Fehlererkennung ließe sich für dieses Beispiel (also mit nur einem umklappenden Bit) auch einfacher einführen. Dazu wird an den ursprünglichen Code ein Paritätsbit angehängt. Dieses Paritätsbit ist so gewählt, dass jedes Zeichen eine gerade Anzahl von Einsen enthält. Unsere Codierungstabelle wird dann zu

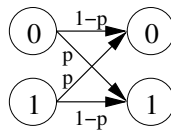


Abbildung 5.20: Symmetrischer Binärkanal

Zeichen	Code	Codierung
a	00	000
b	01	011
c	10	101
d	11	110

Um einzusehen, dass bei dieser Codierung Fehler erkannt werden können, ist es nicht einmal notwendig, sich eine Tabelle aller möglicher Fehler anzulegen. Betrachtet man die Codierung, so zeigt sich, dass sich Zeichen stets durch zwei Bit voneinander unterscheiden (sonst könnte ja die Zahl der Einsen nicht immer gerade sein), d.h. wenn nur ein Bit falsch übertragen wird, ist das Zeichen immer ‘illegal’ insofern, als dass dann plötzlich die Zahl der Bits ungerade wird. Damit gehört das Zeichen aber nicht mehr zum Zeichenvorrat der Quelle und ist damit auch als falsches Zeichen erkannt. Dieser Abstand zwischen zwei Codewörtern (im obigen Beispiel 2 Bit) wird als Hamming-Distanz d bezeichnet. Die Hamming-Distanz ist ein zur Bestimmung der Korrekturmöglichkeiten eines Codes wichtiger Begriff [37].

Um diese Begriffe für Hamming-Distanz und Korrekturfähigkeit quantitativ etwas besser erfassen zu können, betrachten wir den in Abb. 5.20 dargestellten symmetrischen Binärkanal: Der symmetrische Binärkanal verfälscht jeweils mit der gleichen Wahrscheinlichkeit p eine 0 in eine 1 oder eine 1 in eine 0, die Verfälschung wirkt so, dass aus einem Binärzeichen stets wieder ein Binärzeichen wird und kein Zeichen verloren geht oder dazu kommt (die Länge des Codewortes bleibt also erhalten). Die Einzelfehlerwahrscheinlichkeit p ist damit ein Maß für die Störungsintensität des Kanals (beim Telefon ist p ungefähr 0.2 Promille, bei höherwertigen Verbindungen wie Datex-P oder Datex-L liegt die Fehlerwahrscheinlichkeit zwischen 1 und 10 pro Million übertragener Bits).

Über diesen binärsymmetrischen Kanal aus Abb. 5.20 werden nun Nachrichten als Binärzeichen der Länge n übertragen. Nur eine Auswahl aus allen möglichen Wörtern der Länge n wird vom Codierer erzeugt und dient somit der Informationsübertragung. Im obigen Beispiel der Codierung mit Paritätsbit werden die ersten 2 von 3 Bits zur Informationsübertragung benötigt, das dritte Bit wird in Abhängigkeit von den ersten beiden Bits gewählt. Es gibt noch jeweils vier weitere Worte, die aber nicht verwendet werden, da dann die gerade Anzahl der Einsen nicht gegeben ist, d.h. nur die Hälfte der mit drei Bit erzeugten Zeichen wird auch als Code verwendet). Von der geschickten Auswahl der Codewörter hängt die Qualität eines solchen Codes ab:

Definition 10 Die durch die Kanalcodierung eingebrachte Redundanz ist dann gut genutzt, wenn kein Paar von Codewörtern sich zu ähnlich ist, d.h. wenn die Hamming-Distanz möglichst groß wird.

Wenn wir die minimale Hamming-Distanz d in einem Code gefunden haben, so wissen wir auch, wie viele der Bits noch als falsch erkannt werden können. Mit f_e als der Zahl der erkennbaren Binärzeichenfehler gilt dann

$$f_e = d - 1 \quad (5.42)$$

oder in Worten

Definition 11 Ein Blockcode mit der minimalen Hamming-Distanz d gestattet die Erkennung aller Fehler, solange weniger als d Stellen je Wort verfälscht werden.

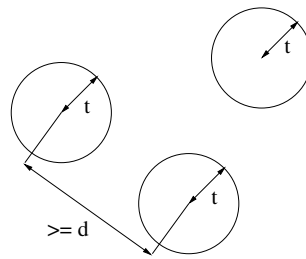


Abbildung 5.21: Ähnlichkeitsdecodierung

Fehlererkennung und -korrektur

Bisher haben wir uns auf die Fehlererkennung beschränkt, wir wollen jetzt versuchen die Möglichkeiten zur Fehlerkorrektur in Abhängigkeit von der Hamming-Distanz zu beschreiben. Betrachten wir dazu als Beispiel einen vierstelligen Code, aus dem aber nur zwei Wörter ausgewählt werden, nämlich 0011 und 1100. Die Codierung erfolgt dabei über die Vorschrift

$$\begin{aligned} 0 &\rightarrow 0011 \\ 1 &\rightarrow 1100 \end{aligned}$$

Alle Worte, die sich gegenüber dem ersten Codewort um höchstens ein Bit verändert haben, fassen wir jetzt in einer Menge A zusammen. Die Menge B enthält alle Wörter, die sich vom zweiten Codewort nur um höchstens ein Bit unterscheiden. Die verbliebenen Wörter fassen wir in der Menge C zusammen. Diese drei Mengen sind dann gegeben durch

$$\begin{aligned} A &= \{0011, 0001, 0010, 0111, 1011\} \\ B &= \{1100, 0100, 1000, 1110, 1101\} \\ C &= \{0000, 0101, 0110, 1001, 1010, 1111\} \end{aligned}$$

Für die Decodierung eines Zeichens betrachtet man jetzt die Menge, in die das Zeichen gefallen ist. Alle Zeichen, die in A fallen, werden so interpretiert, dass das ursprünglich gesendete Zeichen 0011 war, alle Zeichen in B werden dem Zeichen 1100 zugeordnet. Zeichen, die in C einfallen, werden als falsch erkannt, können aber nicht korrigiert werden, sondern werden verworfen. Die Decodierungsvorschrift ist dann

$$\begin{aligned} A &\rightarrow 0 \\ B &\rightarrow 1 \\ C &\rightarrow \text{Fehler} \end{aligned}$$

Dieses Codierungsverfahren erlaubt es also, alle einfachen Fehler zu erkennen und zu korrigieren. Zusätzlich kann ein Teil der Mehrfachfehler erkannt werden. Dieses Verfahren wird als Ähnlichkeitsdecodierung bezeichnet, da das empfangene Zeichen dem Codewort zugeordnet wird, dem es am ähnlichsten ist, von dem es sich also in der geringsten Zahl von Bits unterscheidet. Die Zahl t der korrigierbaren Stellen in einem solchen Verfahren ist gegeben durch

$$2t + 1 \leq d. \tag{5.43}$$

Der größte Wert von t , der mit dieser Bedingung verträglich ist, wird als die *maximale Korrekturfähigkeit eines Codes* bezeichnet.

Den Beweis von (5.43) kann man sich mit Hilfe von Abb. 5.21 veranschaulichen (indirekter Beweis, [30]): Dazu nehmen wir an, dass bis zu t Stellen korrigierbar sein. Wenn die Ungleichung (5.43) verletzt ist, bedeutet das aber, dass es zwei Codewörter gibt, die sich in maximal $2t$ Stellen unterscheiden. Dann gibt es aber bei der Veränderung von t Stellen des einen Codewortes ein Wort, das sich auch von dem anderen Codewort nur um maximal t Stellen unterscheidet, d.h. es ist keine eindeutige Zuordnung des falschen Wortes mehr möglich, im Widerspruch zur vorausgesetzten Korrigierbarkeit des Codes.

	Wort				Par.Bit
	0	0	0	0	0
	0	1	1	1	1
	0	0	1	0	1
	1	1	0	1	1
Par.Bit	1	0	0	0	

Abbildung 5.22: Fehlererkennung und -korrektur bei der Blockcodierung in Matrixform

In dem obigen Beispiel haben wir, um ein Zeichen sicher übertragen zu können, drei zusätzliche Zeichen verwendet, d.h. eine hohe Redundanz erzeugt. Generell besteht ein Wort der Länge n also aus einem Information tragenden Teil der Länge k und einer Zahl m von Prüfzeichen. Die relative Redundanz eines solchen Codes ist dann gleich

$$r = \frac{m}{n}, \quad (5.44)$$

die Redundanz ist

$$R = m. \quad (5.45)$$

Für den oben betrachteten Code erhalten wir also eine relative Redundanz von 75%, wobei bei dieser großen Redundanz aber nur ein Bitfehler zuverlässig erkannt und korrigiert werden kann.

Für die Redundanz $R = m$ eines Korrektur fähigen Codes gilt in Abhängigkeit von der Korrekturfähigkeit t die Abschätzung

$$2^m \geq \sum_{j=0}^t \binom{n}{j}. \quad (5.46)$$

Anschaulich lässt sich die Gültigkeit dieser Hamming-Schranke so zeigen: Es gibt $\binom{n}{j}$ Wörter, die von einem bestimmten Codewort in genau j Stellen abweichen. Der Summenausdruck ist dann genau die Zahl der Wörter, die in höchstens t Stellen von einem bestimmten Codewort abweichen. Dann ist der Code aber nach Voraussetzung (5.43) decodierbar.

Für große Blocklängen n lässt sich die Redundanz [86] dann näherungsweise in der Form

$$r = \frac{m}{n} \geq b \left(\frac{t}{n} \right). \quad (5.47)$$

angeben, mit b als der Shannon-Funktion (5.9).

Wir hatten weiter oben die Einführung eines Paritätsbits zur Fehlererkennung kennen gelernt. Mit Hilfe von geeignet eingeführten Paritätsbits lässt sich in der Blockcodierung in Matrixform (Produktcode: Längs- und Querprüfung) außer der Fehlererkennung auch, zumindest in begrenztem Rahmen, eine Fehlerkorrektur vornehmen. Abbildung 5.22 zeigt dazu ein Beispiel: Der Nachrichtenstrom wird dabei zu Wörtern gleicher Länge (in diesem Falle 4 Bits) zusammengefasst, jedem dieser Wörter wird ein Paritätsbit angefügt. Ordnet man mehrere Wörter zu untereinanderstehenden Blöcken an, so lässt sich zu jeder der so entstehenden Spalten ebenfalls ein Paritätsbit erzeugen und angeben. Im Gegensatz zum einfachen Paritätsbit ist hierbei dann sogar eine Fehlerkorrektur möglich: Würde das eingekreiste Bit falsch übertragen, so würden sowohl in der zweiten Zeile als auch in der dritten Spalte die Paritätsbits nicht stimmen. Damit lässt sich nicht nur bestimmen, dass ein Fehler aufgetreten ist, sondern auch welches Bit das fehlerhafte ist. Damit ist dann aber auch Korrektur möglich. Voraussetzung für die Korrektur ist hierbei aber, dass im ganzen Block nur 1 Bit falsch übertragen ist, bei mehreren falschen Bits kann zwar (muss aber nicht!) die Fehlererkennung noch funktionieren, die Korrektur ist jedoch keinesfalls mehr möglich.

5.3.3 Zusammenfassung

Die Codierung von Nachrichten lässt sich in zwei Bereiche unterteilen: Das Ziel der Quellencodierung ist eine möglichst vollständige Beseitigung der Redundanzen, d.h. die Nachricht soll in einen (Binär-Code) mit maximaler Effizienz umgewandelt werden. Der Quellencodierung schließt sich die Kanalcodierung an, die dem Code gezielt Redundanzen zugefügt, um eine Fehlererkennung und -korrektur zu ermöglichen. Dadurch wird aus jeder Nachricht ein Code erzeugt, der aus einem Nachrichten- und einem Prüfteil besteht. Die häufigste Form der Kanalcodierung ist die Verwendung von Paritätsbits.

5.4 Technische Anmerkungen

Dieser Abschnitt enthält einen kurzen Überblick über einige technische Probleme, die bei der Kommunikation mit bzw. über Satelliten auftreten können. Es muss deutlich darauf hingewiesen werden, dass die Auswahl keinesfalls repräsentativ ist, sondern nur das Ziel hat, einen kleinen (subjektiven) Einblick zu geben.

5.4.1 Überblick

Die technischen Probleme bei der Kommunikation mit Raumfahrzeugen sind wesentlich größer als die Probleme bei der passiven Datenübertragung (Telefon, Fernsehen). Die wichtigsten Probleme sind:

- (A) sehr große Abstände des Senders von der Erde führen zu einem geringen Signal-Rausch-Verhältnis. Zur Illustration: interplanetare Raumsonden wie Voyager und Pioneer befinden sich in Abständen von einigen zehn AU (entsprechend etlichen Milliarden Kilometern) von der Erde. Dadurch ist die Signalleistung extrem gering, d.h. bordseitig muss viel Leistung pro Bit investiert werden. Kann die Leistung bei zunehmendem Abstand nicht erhöht werden, muss entsprechend die Bitrate verkleinert werden, d.h. es kann (bei vorhandenem Zwischenspeicher) u.U. sehr lange dauern, bis die volle Information eines Bildes zur Erde übertragen ist.
- (B) Das schlechte Signal-Rausch-Verhältnis ist aus Gewichts- und Leistungsgründen auch bestimmt durch die Begrenzung der bordseitigen Sendeleistung sowie der bordseitigen Antennengröße. Die sehr kleinen Antennen sind auf Satellitenseite auch für den Empfang von Steuerbefehlen ein Problem, allerdings kann hier durch eine Erhöhung der Sendeleistung auf der Erde gegen gesteuert werden.
- (C) Die Leistungsbegrenzung führt sowohl bord- als auch bodenseitig auf die Notwendigkeit von Richtantennen.
- (D) Die Begrenzung der Sendeleistung erzwingt eine optimale Nutzung der zur Verfügung stehenden begrenzten Übertragungsrate (Bitrate). Daher besteht die Notwendigkeit zur Datenkompression, d.h. Entfernung von Redundanz aus den Daten.
- (E) Das schlechte Signal-Rausch-Verhältnis ebenso wie die Störung des Übertragungskanals durch meteorologische oder ionosphärische Ereignisse bergen das Risiko von Übertragungsfehlern. Daher ist die Möglichkeit der Fehlererkennung und -beseitigung bei der Kommunikation mit Satelliten von Bedeutung.
- (F) Die Satelliten sind gegenüber der Erde als Ganzem ebenso wie gegenüber den Empfangsstationen auf der Erde in Bewegung. Dadurch kann es sich ergeben, dass sich nicht zu allen Zeiten eine terrestrische Empfangsstation im 'Blickfeld' des Satelliten befindet. Dadurch ergibt sich die Notwendigkeit der Speicherung von Daten, in der Regel durch ein Magnetband. Hierbei handelt es sich übrigens um ein interessantes technisches Problem, da das Einschalten eines Magnetbandes aufgrund der Drehimpulserhaltung auf eine Rotation des Satelliten in entgegengesetzter Richtung führt. Dieser Effekt ist aber aufgrund der genau ausgerichteten Instrumente und Antennen unerwünscht und daher zu vermeiden.

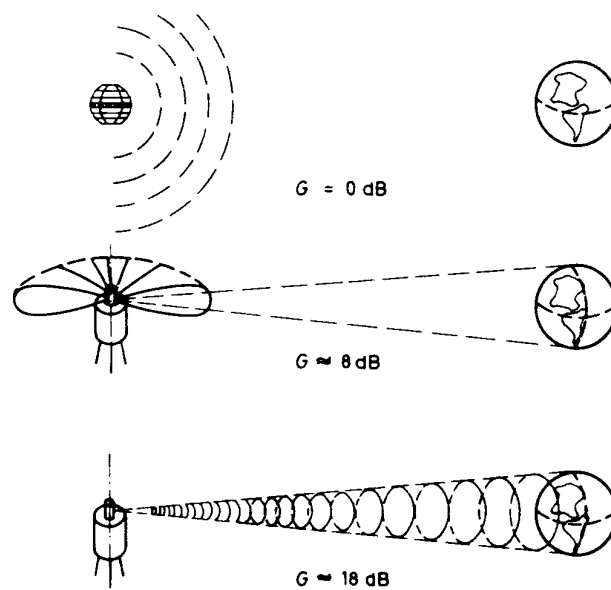


Abbildung 5.23: Antennengewinn bei Bündelung [43]

(G) Beim Versenden von Telekommandos (d.h. Anweisungen an den Satelliten) ist insbesondere bei den planetaren Missionen die lange Laufzeit des Signals zu berücksichtigen. So erreicht der Befehl eine Raumsonde am Jupiter erst nach ca. 45 min, eine am Neptun dagegen erst nach 4 Stunden. Die Rückmeldung der Raumsonde ist entsprechend genauso lange unterwegs, d.h. alle Manöver müssen sehr sorgfältig im voraus geplant werden. Kommandos sind gerade in der Nähe der Planeten wichtig, z.B. zur Kurskorrektur für ein geplantes Swingby-Manöver oder zum genaueren Ausrichten der Kameras bei Planeten- oder Mondaufnahmen, aber auch zum Abschalten einzelner Instrumente, da sich in den Strahlungsgürteln der Planeten so große Teilchenflüsse ergeben können, dass es bei eingeschalteten Instrumenten zu einer Beschädigung von Sensoren oder anderen Instrumentbestandteilen kommen kann. Ist also ein Manöverbedarf erkannt, so dauert es erst einmal die einfache Signallaufzeit, bis der Satellit diesen Befehl erhält. Die Rückmeldung, ob dieser Befehl ausgeführt werden konnte oder nicht, kommt dann erst Stunden, nachdem der Befehl gegeben wurde. Andererseits haben z.B. die Voyager-Sonden Geschwindigkeiten, die im Bereich von Hunderttausend km/h liegen, d.h. falls der Befehl zu einer Kurskorrektur gegeben wurde, muss er schon im voraus sehr sorgfältig geplant werden, spontane Reaktionen auf kleine Fehler sind aufgrund der langen Signallaufzeiten nachher nicht mehr möglich (das ist ungefähr so, als würde man sich hier in ein Auto setzen und schon beim Starten wissen müssen, wie man in München das Einparkmanöver genau durchzuführen hat). Klingholz [60] beschreibt diese Probleme recht anschaulich.

Diese Auflistung enthält die wichtigsten Probleme, wie sie bei der Kommunikation mit Satelliten bzw. insbesondere mit Raumsonden im interplanetaren Medium auftreten. Für einige der hier angesprochenen Probleme soll ein Lösungsansatz skizziert werden.

5.4.2 Lösungsansätze

Richtantennen

Wir hatten das geringe Signal-Rausch-Verhältnis ebenso wie die geringe Sendeleistung als Motivation für die Verwendung von Richtantennen angesprochen. Während die Verwendung von Richtantennen bei interplanetaren Missionen sofort einleuchtet, ist sie aber bereits bei den vergleichsweise geringen Abständen der geostationären Kommunikationssatelliten

unerlässlich. Würden wir in diesem geringen Abstand einen Kugelstrahler als Sendeantenne verwenden, so würde bereits der größte Teil der Leistung verlorengehen (oberer Teil in Abb. 5.23). Daher versucht man, die Antennenleistung in einen schmalen Kegel zu bündeln, der gerade den gesamten vom Satelliten sichtbaren Bereich der Erdoberfläche ausleuchtet (Mitte und Unten). Ist der Satellit jedoch spinstabilisiert, so wird aufgrund der Rotation des Satelliten immer noch ein großer Teil der Leistung sinnlos in den Raum abgestrahlt. Erst wenn man die Antenne entsprechend entgegengesetzt zum Satelliten rotieren lässt (entdrallte Antenne), wird immer nur ein schmaler Kegelausschnitt auf die Erdoberfläche gerichtet und somit bei fester Senderleistung die maximale Signalleistung beim Empfänger erreicht.

Leistungsbilanz

Der wichtigste Parameter zur Beschreibung des Verhältnisses zwischen eingesetzter und empfangener Leistung ist die Raumdämpfung (space loss). Die Energiebilanz bei der Kommunikation mit Raumsonden hängt dabei stark von der Bündelung der Antennen sowohl auf dem Raumfahrzeug als auch bei der Bodenstation ab. Die Bündelung kann durch den Antennengewinn G beschrieben werden. Dieser wird für eine bestimmte Richtung definiert als das Verhältnis aus dem tatsächlich in diese Richtung emittierten Energiefluss zu dem sich bei der isotropen Emission einer ungebündelten Antenne ergebenden Energiefluss. Der maximale Antennengewinn ergibt sich normalerweise entlang der Symmetrieachse der Antenne, so dass der Antennengewinn in der Regel entlang dieser Richtung bestimmt wird. In einem Abstand d von der Antenne erzeugt eine emittierte Leistung P dann einen Energiefluss

$$\Phi = \frac{P \cdot G}{4\pi d^2} . \quad (5.48)$$

Ingenieure messen den Gewinn (ebenso wie Verluste) in einer logarithmischen Skala

$$g = 10 \cdot \log_{10} G \quad (5.49)$$

mit g in Dezibel (dB). Dabei entspricht 1dB einem Antennengewinn G von 1.25.

Um einen hohen Antennengewinn zu erreichen, werden Parabolspiegel benutzt mit einem Oszillator im Brennpunkt. Könnte man diesen Oszillator als eine Punktquelle mit verschwindend kleiner Wellenlänge bauen, so würde sich ein exakt paralleler Strahl ergeben mit einem unendlichen Antennengewinn. Unter realen Verhältnissen ergibt sich jedoch eine Aufweitung des Antennenstrahls, die einen Winkel einnimmt, der ungefähr durch das Verhältnis von Wellenlänge λ und Durchmesser der Antenne gegeben ist. Der Raumwinkel, in den die Strahlung emittiert wird, ist dann von der Größenordnung des Verhältnisses λ^2 zu Fläche A der Antenne. Der Antennengewinn (4π multipliziert mit dem Kehrwert des Raumwinkels) kann dann geschrieben werden als

$$G = \frac{4\pi A_{\text{eff}}}{\lambda^2} . \quad (5.50)$$

Für einen perfekten Parabolspiegel wäre die effektive Fläche A_{eff} gleich der Fläche A der Antenne. In der Praxis kann die effektive Fläche bis zu 0.8 oder 0.9 A erreichen. Mit sehr großen Antennen (die Teleskope des Deep Space Network DSN haben Durchmesser bis zu 70 m) können im Mikrowellenbereich dann Antennengewinne von 60 bis 80 dB erreicht werden, entsprechend Winkeln von 0.001 bis 0.0001 rads. Diese hohen Antennengewinne erfordern umgekehrt aber eine extrem genaue Ausrichtung der Antennen!

Die effektive Fläche A_{eff} der Antenne ist gleichzeitig auch deren Querschnitt im Empfangsmodus. An Bord des Satelliten würde dann bei einem Telekommando eine Leistung

$$P' = \frac{PGA'_{\text{eff}}}{4\pi d^2} \quad (5.51)$$

empfangen werden, wobei das Hochkomma Größen bezeichnet, die mit dem Satelliten zu tun haben. Wenn wir mit G' den Gewinn des Empfängers beschreiben, so lässt sich (5.51) umschreiben zu

$$P' = \frac{PGG'\lambda^2}{(4\pi d)^2} = PGG'L . \quad (5.52)$$

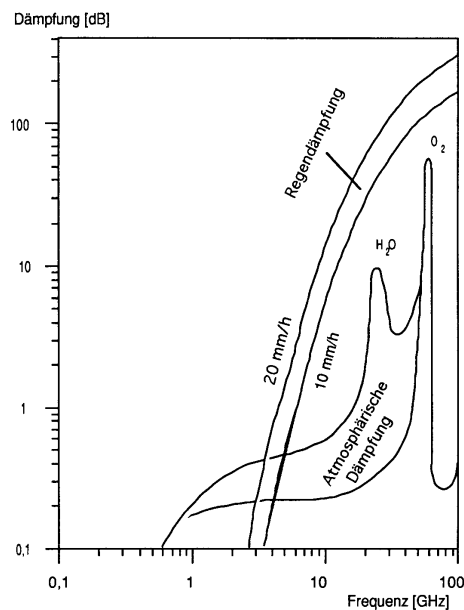


Abbildung 5.24: Atmosphärische Dämpfung und Regendämpfung in Abhängigkeit von der Frequenz (Dodel und Baumgart, 1986, in [89])

Darin ist die Raumdämpfung $L = (\lambda/4\pi d)^2$ ein Maß für den Verlust, der sich alleine durch die geometrische Aufweitung des Antennenstrahls während seiner Ausbreitung ergibt. Üblicherweise stellt man diese Energiebilanz logarithmisch dar:

$$p = 10 \cdot \log_{10}(P/1mW) . \quad (5.53)$$

Dann lässt sich die Leistungsverstärkung in dBm (dB bezogen auf mW) im Bezug auf eine nominelle Leistung von 1 mW angeben als

$$p' = p + g + g' + \ell . \quad (5.54)$$

Darin ist die logarithmische Raumdämpfung ℓ negativ. Zusätzliche Verluste durch atmosphärische Streuung und ohmsche Verluste in Sender und Empfänger sind in dieser einfachen Betrachtung nicht berücksichtigt.

Seems it never rains in southern california

Auf seinem letzten Stückchen zwischen Raumsonde und Satellit hat das Funksignal noch die Atmosphäre zu durchqueren. Die Atmosphäre verändert das Signal durch

- atmosphärische Dämpfung: Diese wird durch verschiedene Moleküle bewirkt, im wesentlichen jedoch Wasserdampf und Sauerstoff. Die atmosphärische Dämpfung hängt vom Wasserdampfgehalt der Atmosphäre ab und nimmt für unterschiedliche Frequenzen unterschiedliche Werte an (vgl. Abb. 5.24).
- Regendämpfung: Der Einfluss der Regendämpfung steigt mit zunehmender Frequenz an und steigt ebenfalls mit der Regenrate bzw. Tropfendichte. Die Regendämpfung steigt nahezu linear mit der Frequenz an (vgl. Abb. 5.24). Es ist dabei zwischen konvektivem Regen, der kräftig und lokal begrenzt ist, und Restregen mit niedriger Regenrate aber großer räumlicher Ausdehnung zu unterscheiden.
- Zusätzliches Rauschen durch Drehung der Polarisationssebene in Folge von Niederschlägen.

Extrem starke Regenfälle und insbesondere Gewitter können, allerdings in der Regel nur kurzzeitig und lokal begrenzt, zu einem vollständigen Ausfall von Satellitenverbindungen führen. Selbst wenn es nicht zum vollständigen Zusammenbruch der Verbindung kommt,

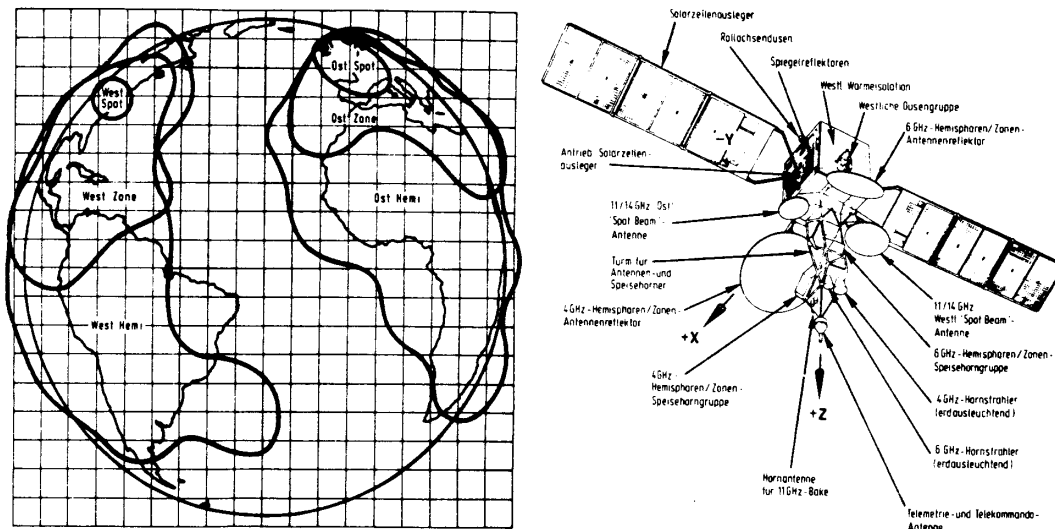


Abbildung 5.25: Ausleuchtungszonen von Intelsat V, über dem Atlantik stationiert (links) sowie Aufbau des Satelliten (rechts; [43])

können die Störungen aber so groß werden, dass die Fehlerrate Werte erreicht, bei denen eine Korrektur nicht mehr möglich ist.

5.4.3 Kommunikation über Satelliten

Alle bisher angesprochenen Problemem gelten natürlich auch für die Kommunikationssatelliten in erdnahen Bahnen. Um eine ungefähre Vorstellung über Kommunikationssatelliten zu vermitteln, zeigt Abb. 5.25 den Aufbau und die Ausleuchtungszonen von Intelsat V. Aus dem Aufbau des Satelliten kann man erkennen, dass die Kommunikation auf Wellenlängen im Bereich einiger GHz über verschiedene Antennen erfolgt. Eine dieser Antennen leuchtet praktisch die gesamte sichtbare Fläche der Erdkugel aus, die anderen Antennen dienen jeweils der Ausleuchtung der den Atlantik im Osten bzw. Westen begrenzenden Kontinente insgesamt, der kontinentalen Bereiche auf der Nordhalbkugel, sowie einer eher lokalen Verbindung zwischen Mitteleuropa und der mittleren Ostküste der USA.

Literatur

Eine ausgezeichnete und aktuelle Einführung in die Grundlagen der Kommunikationstechnik geben Proakis und Salehi [88]. Das Buch betrachtet zwar nicht speziell die Kommunikation zwischen Satellit und Bodenstation, da jedoch Kommunikation heute zu einem großen Teil über Satelliten erfolgt, sind die dort diskutierten Probleme und Verfahren für die Kommunikation mit Satelliten ebenso anwendbar.

Sehr gut ist die Darstellung der informationstheoretischen Grundbegriffe in Shannon und Weaver [112] oder Fano [21]. Die Darstellungen in v. Cube [125] und insbesondere Steinbuch [114, 115] sind zwar kürzer, enthalten aber neben den Grundlagen der Informationstheorie sehr gute und informative Betrachtungen zur Informationsverarbeitung durch den Menschen.

Als gute (und preisgünstige) Literatur zur Codierung ist zu empfehlen Grams [30]. In dem Buch werden auch weitere Codierungsverfahren und kompliziertere Annahmen über die Eigenschaften des Übertragungskanal diskutiert. Grundlagen der Datenkompression durch Redundanzreduktion werden auch in Wayner [126] und Witten et al. [129] diskutiert.

Technische Aspekte zur Kommunikation mit Satelliten und über Kommunikationssatelliten werden behandelt in Bertotti und Farinella [8], Hartl [40], Herter und Rupp [43], Quern-

heim [89] und Renner et al. [95]. Technische Aspekte der Signalübertragung allgemein werden diskutiert in Herter und Lörcher [44] und Lüke [74].

Fragen

1. Wie kann man einen Informationskanal formal beschreiben?
2. Geben Sie ein einfaches Modell für einen Informationskanal.
3. Was ist Information? Geben Sie Definitionen und Beispiele.
4. Wie misst man Information?
5. Wie bestimmt man den Informationsgehalt eines einzelnen Zeichens, wie den mittleren Informationsgehalt eines Zeichensatzes?
6. Erläutern Sie die Information, die in einem Versuch mit zwei möglichen Ausgängen stecken kann?
7. Was ist die Shannon'sche Funktion?
8. Was versteht man unter Redundanz? Wie kann man sie formal definieren.
9. Welche Bedeutung hat Redundanz in der normalen Kommunikation?
10. Welche Aufgabe hat die Quellenkodierung?
11. Nennen Sie Beispiele für Quellenkodierung.
12. Vergleichen Sie unter informationstheoretischen Gesichtspunkten die Codierung des Alphabets im ASCII- und im Morse-Code.
13. Was versteht man unter Kanalcodierung? Welche Ziele werden mit der Kanalcodierung verfolgt?
14. Erläutern Sie einfache Verfahren zur Fehlererkennung.
15. Welche Voraussetzungen muss ein Codierungsverfahren erfüllen, um nicht nur die Fehlererkennung sondern auch die Korrektur zu ermöglichen?
16. Was versteht man unter einem Präfix-freien Code?
17. Erläutern Sie den Lempel–Ziv-Algorithmus.

Aufgaben

1. Zum Beispiel in Abb. 5.4. Bestimmen Sie den Informationsgehalt der einzelnen Farben sowie den mittleren Informationsgehalt (Entropie).
2. Die beiden Binärzeichen mögen mit den Wahrscheinlichkeiten $1/6$ bzw. $5/6$ auftreten. Wie groß ist die Entropie H ? Beachten Sie, dass 1 Bit jetzt nicht mehr den Informationsgehalt 1 bit hat! Berechnen Sie die Anzahl N_{12} aller möglichen Folgen aus 12 Binärzeichen mit genau 2 Einsen. Die relativen Häufigkeiten der Zeichen entsprechen in dieser Folge genau den Wahrscheinlichkeiten der Zeichen. Berechnen Sie analog dazu N_{30} für Folgen aus 30 Zeichen mit genau 5 Einsen und N_{60} für Folgen der Länge 60 mit genau 10 Einsen. Wie verhalten sich die Werte $\text{ld}(N_k)/k$ zur Entropie H ?
3. Zeichnen Sie den Codier-Baum für die nicht decodierbare Variante von Code 2 in Tabelle 5.3.
4. Geben Sie einen entzifferbaren Code ohne Präfix-Eigenschaften an. Zeichnen Sie den Codier-Baum dazu.
5. Die Zeichen a,b,c,d,e,f und g treten mit den Wahrscheinlichkeiten 0.3, 0.2, 0.2, 0.1, 0.1, 0.05 und 0.05 auf. Bilden Sie die Codes analog zu denen in Tabelle 8.2 und berechnen Sie die relative Redundanz.
6. Codieren Sie den Satz 'Quellenkodierung ist Redundanzreduktion' mit (a) einer Huffman-Codierung, die das Auftreten der einzelnen Buchstaben in der deutschen Sprache berücksichtigt (vgl. Tabelle 5.1), (b) einem Lempel-Ziv-Algorithmus angewandt auf den in ASCII (8-bit) codierten Text.

Kapitel 6

Datenaufbereitung

Datenaufbereitung von Satellitendaten ist lange Zeit im wesentlichen digitale, teilweise auch analoge Bildbearbeitung gewesen. Heutzutage sind viele Satelliteninstrumente zwar keine bildgebenden Instrumente mehr, die Bildbearbeitung nimmt jedoch auch hier einen großen Teil der Datenaufbereitung ein, da (a) die auch die Daten primär nicht bildgebender Verfahren bildhaft dargestellt werden, (b) die Darstellung in Bildern häufig Fehlererkennung und -korrektur erlauben, und (c) bei der automatischen Klassifikation nur der Übergang vom Pixel zum Voxel (Volumenelement) gemacht werden muss, ansonsten aber die gleichen Regeln gelten. . Daher werde ich mich in diesem Abschnitt auf die Bildbearbeitung beschränken.

Die Bearbeitung von Satellitendaten am Boden erfolgt unter vier wesentlichen Kriterien:

1. Korrektur fehlender oder verrauschter Bildpunkte oder -zeilen, d.h. die Identifikation dieser Pixel und ihre anschließende Ersetzung durch einen sinnvollen Wert;
2. Verfahren zur Bildverbesserung wie Kantenanreicherung oder Kontrasterhöhung, d.h. Verfahren zur verbesserung der Erkennung von Strukturen;
3. Verfahren zur automatischen Klassifikation von Bildmerkmalen, d.h. die automatische Erzeugung von Nutzer interessierenden Daten;
4. Verfahren zur effizienten und möglichst verlustfreien Speicherung der Daten, d.h. Datenkompression nach Regeln wie bereits bei der Quellencodierung in Abschn. 5.3.1 diskutiert.

Einen weiteren wichtigen Aspekt der Bildverarbeitung werden wir nicht betrachten: die geometrische Korrektur der Abbildungsfehler. So sind die vom Scanner gelieferten Streifen nicht exakt senkrecht zur Flugrichtung sondern auf Grund der Eigenbewegung des Satelliten leicht gegen die Senkrechte geneigt. Auch sind die Pixel zu den Ende der Scanbahn durch die schräge Blickrichtung gegenüber denen direkt auf der Bodenspur verzerrt.¹ Beide Effekte sind rein geometrisch, treten in jeder Szene auf und können durch Projektionen korrigiert werden, vgl. z.B. [16, 17, 52, 101].

6.1 Korrekturverfahren

Korrekturverfahren zerfallen, wie bereits in Abschn. 5.3.2 diskutiert, in zwei Teile: die Fehlererkennung und die eigentliche Fehlerkorrektur. Einige Fehler sind einfach zu erkennen,

¹Systematische Verzerrungen entstehen durch das Aufnahmeverfahren, als Beispiel sei hier der LandSat TM betrachtet. Der Zeitunterschied zwischen der Aufnahme des ersten und letzten Bildelements einer Zeilenreihe beträgt 60 ms, der Unterschied zwischen der ersten und letzten Zeile einer Szene 25 s. Da sich der Satellit in diesen Zeiten bewegt (mit einer Geschwindigkeit von ungefähr 7 bis 8 km/s bzw. 480 m während einer Zeilenabtastung), wird kein exakt rechtwinkliges Raster der Erdoberfläche abgetastet, wie es uns in der Abbildung vorgespiegelt wird. Durch die Erdkrümmung und die panoramische Abtastung quer zur Flugrichtung erhält das Bild außerdem einen von der Bildmitte zum rechten und linken Rand hin kontinuierlich abnehmenden Maßstab. Diese Fehler sind, da sie systematisch sind, rechnerisch relativ einfach zu erfassen und zu korrigieren. Unsystematische Fehler können nur mit Hilfe von vorher festgelegten Referenzpunkten aus der Aufnahme herauskorrigiert werden.

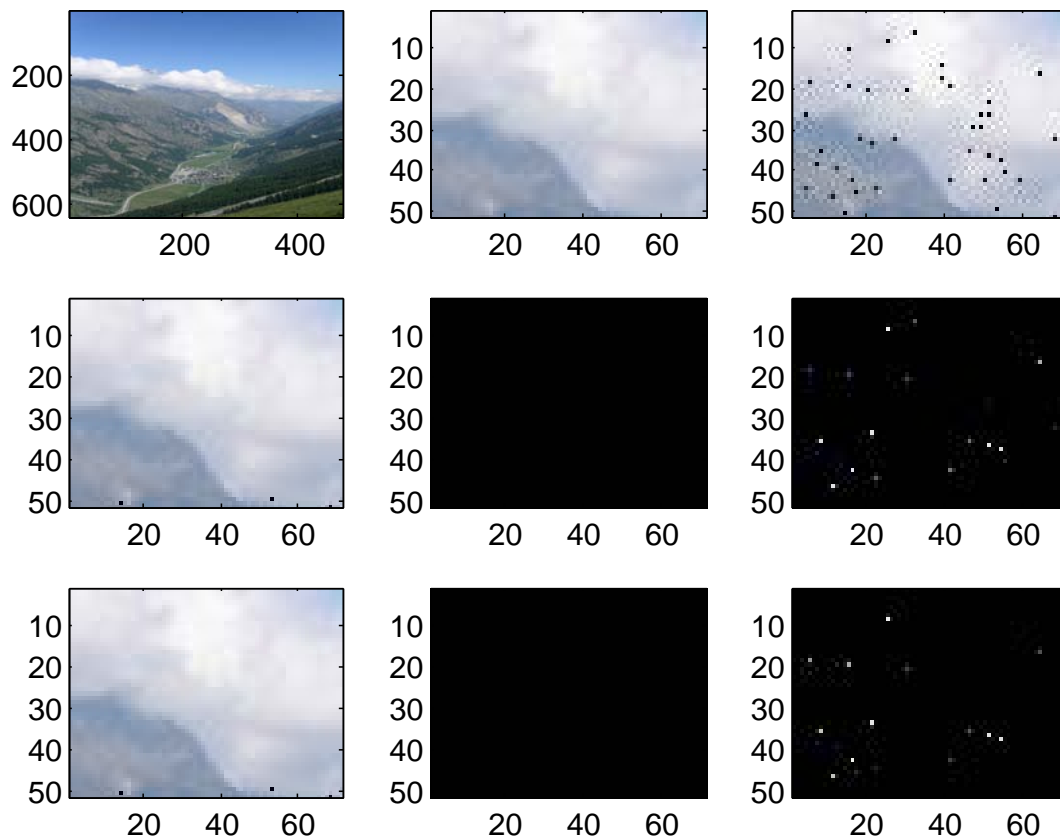


Abbildung 6.1: Korrektur fehlender Pixel: in der oberen Reihe links das vollständige Bild, in der Mitte ein Ausschnitt und rechts ein Bild, in dem einige Pixel nach Zufallsprinzip entfernt wurden. Die Zeilen darunter geben jeweils links ein korrigiertes Bild, in der Mitte die Abweichung zwischen korrigiertem Bild und Original und rechts die um einen Faktor 100 überhöht dargestellte Abweichung für einen 3×3 -Kernel in der mittleren Zeile und für einen 5×5 -Kernel in der unteren

insbesondere Pixel mit dem Wert 0 oder gar fehlende oder durch einen Füllwert ersetzte Zeilen. Andere Fehler, insbesondere verrauschte Pixel, dagegen sind nicht unbedingt so einfach zu erkennen.

Eine echte Korrektur der so identifizierten fehlerhaften Pixel ist allerdings in der Regel nicht möglich: im Zusammenhang mit der Kanalcodierung haben wir bereits festgestellt, dass eine Codierung, die eine Fehlerkorrektur erlaubt, das zu übertragende Datenvolumen deutlich erhöht. In der Fernerkundung geht man daher einen anderen Weg: fehlerhafte Pixel werden durch den Mittelwert der sie umgebenden Pixel ersetzt, da man davon ausgeht, dass zwischen benachbarten Bildpunkten ein gewisser Zusammenhang besteht (Kontext bezogene Korrektur).

6.1.1 Fehlende Pixel

Als erstes wollen wir den Fall betrachten, dass für einzelne Bildpunkte der Intensitätswert nicht oder als Null übertragen wurde. Befinden sich diese Pixel nicht gerade in einer sehr dunklen Szene, so sind sie leicht zu erkennen.

Abbildung 6.1 gibt ein Beispiel für in einem Bild fehlende Pixel (rechts oben). Die Fehlstellen wurden aus dem Original (Mitte oben) durch Nullsetzen zufällig ausgewählter Pixel erzeugt.

Für die Korrektur derartiger einzelner Fehlstellen ersetzt man den fehlenden Wert durch

34	39	49	36	31	59	34	42
54	0	32	43	34	39	49	38
39	45	47	38	54	0	32	52
			46	39	45	47	45
			34	54	43	31	68

Abbildung 6.2: Kernel verschiedener Größe zur Korrektur fehlerhafter Pixel

den Mittelwert der umgebenden Pixel. Für die Bestimmung dieses Mittelwerts kann man einen 3×3 -Kernel verwenden oder einen 5×5 -Kernel, vgl. Abb. 6.2.

Bei einem 3×3 -Kernel werden alle direkt benachbarten Bildpunkte berücksichtigt, vgl. linken Teil von Abb. 6.2. Ist $p(i, j)$ der fehlerhafte Bildpunkt, so wird er ersetzt durch den Mittelwert

$$\bar{p}(i, j) = \frac{1}{8} \left(\sum_{j'=j-1}^{j+1} \sum_{i'=i-1}^{i+1} p(i', j') - p(i, j) \right). \quad (6.1)$$

Im Beispiel in Abb. 6.2 erhalten wir damit $\bar{p} = 42.4$.

Mit diesem Verfahren können keine Pixel korrigiert werden, die sich am Rand einer Szene befinden, da diese nicht von der notwendigen Zahl anderer Pixel umgeben sind.

In der mittleren Zeile von Abb. 6.1 ist ein 3×3 -Kernel zur Korrektur der fehlerhaften Pixel verwendet. Das Bild links ist das korrigierte Bild, in der Mitte ist die Differenz zwischen korrigiertem Bild und Original gezeigt, rechts die um einen Faktor 100 überhöhte Differenz. Die Korrektur liefert bei rein visueller Inspektion bereits recht gute Ergebnisse (insbesondere lenken die fehlenden Pixel nicht mehr die Aufmerksamkeit auf sich), das Differenzbild zeigt nur marginale Abweichungen zum Original.

Eine Erweiterung des Verfahrens ist die Verwendung eines 5×5 -Kernels, vgl. rechtes Teilbild in Abb. 6.2. Auch hier wird der fehlende Pixel durch den Mittelwert der umgebenden Pixel ersetzt:

$$\bar{p}(i, j) = \frac{1}{24} \left(\sum_{j'=j-2}^{j+2} \sum_{i'=i-2}^{i+2} p(i', j') - p(i, j) \right). \quad (6.2)$$

In diesem Beispiel beträgt das Ergebnis $\bar{p} = 43.0$.

Die Anwendung auf die Daten ist in der unteren Zeile von Abb. 6.1 gezeigt. Rein optisch bringt das Verfahren gegenüber dem kleineren Kernel keine Verbesserung, lediglich im hundertfach überhöhten Differenzbild erkennt man, dass einige weitere Pixel keine Abweichung vom Original mehr zeigen.

Beide Kernel haben ihre Vor- und Nachteile: bei stark verrauschten Daten aus einer relativ homogenen Szene ist der größere Kernel geeigneter, da er auf Grund der Mittelung über die große Pixelzahl das Rauschen besser unterdrückt. Allerdings hat der große Kernel den Nachteil einer höheren Rechenzeit und den Verlust einer weiteren Pixelreihe am Rand, die nicht mehr korrigierbar ist. Außerdem ist ein großer Kernel in einer stark variablen Szene und insbesondere in der Nähe einer Grenze ungünstig, da Pixel von verschiedenen Objekten gemittelt werden.

Fehlerhafte Pixel durch einen CCD-Fehler kann man auch durch den Vergleich verschiedener Spektral-Kanäle erkennen: diese treten in allen Kanälen an gleicher Stelle auf. Meistens entstehen fehlende Pixel jedoch durch Übertragungsfehler.

6.1.2 Periodisch fehlende Streifen

Während fehlende Pixel meistens durch stochastische Einflüsse entstehen, sind periodisch fehlende Streifen ein instrumentelles Problem. Für ihre Korrektur werden ebenfalls die Kernel

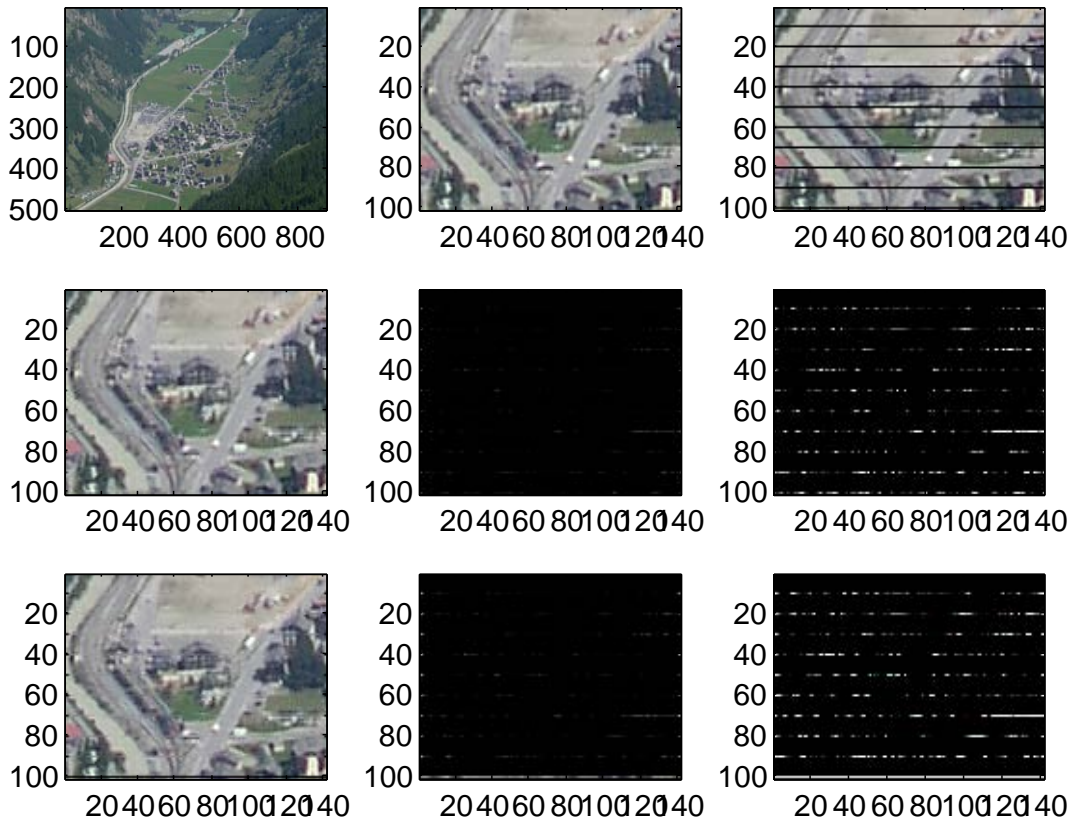


Abbildung 6.3: Fehlende Zeilen und deren Korrektur. Aufbau wie Abb. 6.1, das Differenzbild jedoch nur um einen Faktor 10 erhöht

aus Abb. 6.2 verwendet, allerdings werden jetzt natürlich die Bildpunkte in der fehlenden Zeile nicht mehr berücksichtigt. Für den 3×3 -Kernel ergibt sich damit

$$\bar{p}(i, j) = \frac{1}{6} \left(\sum_{j'=j-1}^{j+1} p(i-1, j') + \sum_{j'=j-1}^{j+1} p(i+1, j') \right), \quad (6.3)$$

für den 5×5 -Kernel

$$\bar{p}(i, j) = \frac{1}{20} \left(\sum_{j'=j-2}^{j+2} \sum_{i'=i-2}^{i+2} p(i', j') - \sum_{j'=j-2}^{j+2} p(i, j') \right). \quad (6.4)$$

Abbildung 6.3 zeigt, in einer Struktur ähnlich Abb. 6.1, ein Beispiel für eine Szene mit periodisch fehlenden Streifen. Die Korrektur erfolgt mit dem 3×3 -Kernel (mittlere Zeile) und dem 5×5 -Kernel (untere Zeile). Die Differenz in der rechten Spalte ist jeweils um einen Faktor 10 überhöht.

Bei der visuellen Inspektion sind keine Relikte der Streifen zu erkennen (bei einer genauen Inspektion fällt natürlich auf, dass die Randpunkte einer Zeile nicht korrigierbar sind). Auch die Differenzbilder zeigen selbst in der zehnfachen Überhöhung (rechts) eine recht gut gelungene Korrektur – da die Linien nicht durchgängig sind, sind etliche Pixel vollständig korrigiert.

Die Verwendung des 5×5 -Kernels (untere Zeile von Abb. 6.3) liefert im überhöhten Differenzbild sogar noch etwas größere Abweichungen – da die Szene sehr stark wechselnde Pixelwerte hat, erweist sich die Mittelung über den großen Bereich als Nachteil. Außerdem ist der untere Streifen fehlender Pixel nicht korrigierbar, da er nur einen Streifen Abstand

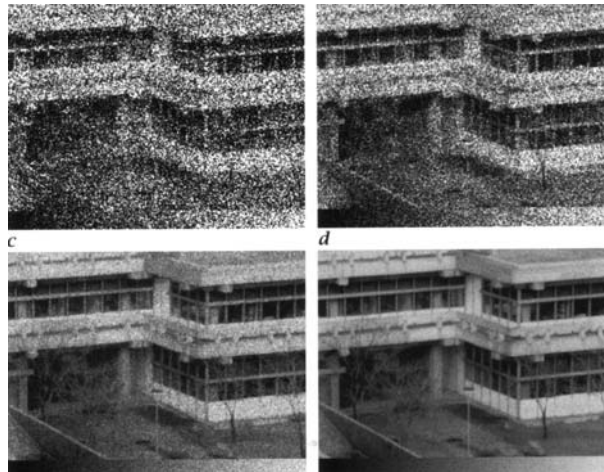


Abbildung 6.4: Simulation des Poisson-Rauschens [52]

vom Rand hat und damit für einen 5×5 -Kernel nicht genug benachbarte Pixel vorhanden sind.

Das Verfahren wurde hier für periodische Streifen demonstriert. Diese entstehen häufig durch nicht korrekt synchronisierte Abtastung. Einzelne fehlende Streifen können ebenfalls auftreten: erkennt man z.B. innerhalb einer Zeile eine mangelhafte Synchronisation der Spalten, so wird die Zeile durch Füllwerte ersetzt. Die Korrektur einzelner fehlender Zeilen folgt dem gleichen Verfahren wie die periodisch fehlender Zeilen.

6.1.3 Periodische Streifen

Periodische Streifen sind den periodisch fehlenden Zeilen äquivalent – der einzige Unterschied besteht darin, dass der Wert des Pixels nicht Null ist sondern in der ganzen Zeile eine Konstante ist. Dies macht einzelne fehlerhafte Zeilen vielleicht nicht unbedingt visuell aber zumindest für den Rechner leicht erkennbar – schon auf Grund des statistischen Rauschens kann selbst eine vollständig homogene Szene nie einen Streifen mit konstanten Pixelwerten liefern. Das Korrekturverfahren entspricht dem für fehlende Streifen.

6.1.4 Random Noise

Während fehlerhafte Streifen aus ihrer Anordnung einfach zu erkennen sind, sind fehlerhafte Pixel wesentlich schwieriger zu identifizieren. Ihre Behandlung ist eng verbunden mit der Behandlung des Rauschens der Bildpunkte.

Die Aufnahme eines Bildes, sei es mit einer konventionellen photographischen Emulsion oder mit einer CCD, basiert auf der Wechselwirkung der von der Szene emittierten oder reflektierten Photonen mit dem Detektor: je größer die Zahl der auftreffenden Photonen, um so größer die Intensität und damit der Helligkeitswert. Vom Messprinzip her basieren bildgebende Verfahren also auf der Zählung von Photonen in diskreten räumlichen Intervallen (Pixel) und diskreten Energiebändern (Spektralbereich, Kanal).

Daten, die zählend erhoben werden, gehorchen der Poisson-Statistik:

$$f(x) = \frac{\mu^x}{x!} e^{-\mu} \quad (6.5)$$

mit μ als dem Erwartungs- oder Mittelwert. Die Standardabweichung σ ist gleich der Wurzel aus dem Mittelwert

$$\sigma = \sqrt{\mu} . \quad (6.6)$$

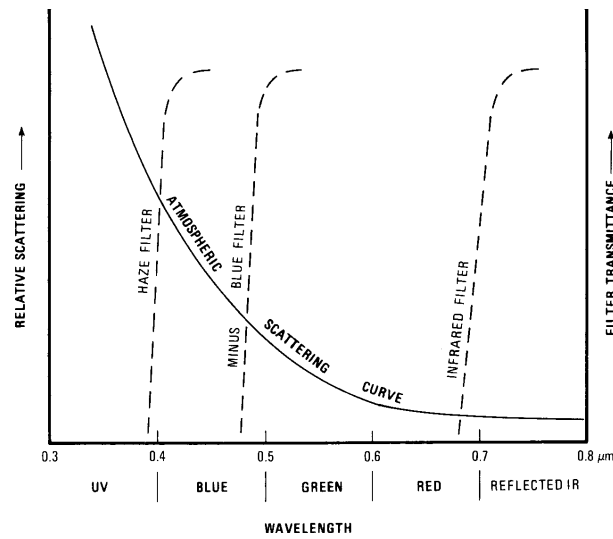


Abbildung 6.5: Atmosphärische Streuung und Bereiche, in denen Filter arbeiten [101]

Der relative Fehler ist damit $\sigma/\mu = 1/\sqrt{\mu}$, d.h. bei kleinen Zählraten ist der relative Fehler sehr groß, mit zunehmender Zählrate nimmt er ab.

Für große Zählraten geht die Poisson-Verteilung über in die Gauß-Verteilung

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}. \quad (6.7)$$

Abbildung 6.4 zeigt ein Beispiel für ein simuliertes Poisson-Rauschen. Die Einzelbilder entsprechen einer Maximalzahl von Photonen, die während der Belichtung auf ein Pixel fallen konnten von 3, 10, 100 bzw. 1000. Am unteren Bildrand ist jeweils ein linearer Graukeil angegeben. Im oberen linken Teilbild sind nur die Bereiche mit großem Kontrast unterscheidbar, Bereiche mit geringem Kontrast (linker unterer Bildteil) dagegen erscheinen homogen. Der schlechte radiometrische Kontrast reduziert damit auch das räumliche Auflösungsvermögen. Da benachbarte Pixel selbst bei homogener Szene auf Grund des Rauschens sehr unterschiedliche Werte annehmen können ($\sqrt{3}$ ist nahezu 2, d.h. der relative Fehler ist ca. 60%), erscheint das Bild sehr körnig. Im rechten unteren Teilbild dagegen ist der relative Fehler wesentlich geringer ($\sqrt{1000} \approx 32$), so dass bei homogener Szene benachbarte Pixel auf der Grauskala auch ähnliche Werte annehmen und damit bei einfacher visueller Inspektion nicht unterschieden werden.

Rauschen sorgt also für Variationen der Helligkeitswerte benachbarter Pixel selbst wenn diese eine homogene Szene betrachten. Umgekehrt können fehlerhafte Pixel leicht im Rauschen untergehen. Zu ihrer Identifikation kann man einen Filter auf der Basis der in Abb. 6.2 gezeigten Kernel verwenden: aus allen Punkten des Kernels wird der Mittelwert gebildet und mit dem Wert des zentralen Punktes verglichen. Liegt dieser mehr als ein vorgegebener Wert, z.B. das Dreifache der Standardabweichung, vom Mittelwert entfernt, so wird der Pixel als fehlerhaft identifiziert. Die Korrektur dieses Pixels erfolgt wie oben für fehlerhafte Pixel beschrieben unter Verwendung des gleichen Kernels.

Ein Glättungsfilter zur Reduktion des Rauschens funktioniert ähnlich. Bei einem linearen Glättungsfilter wird der 3×3 -Kernel (bzw. bei Bedarf der 5×5 -Kernel) über die gesamte Szene geschoben und jeweils das zentrale Pixel durch den Mittelwert des Kernels ersetzt. Statt eines linearen Filters kann man auch einen Gauß-Filter verwenden, bei dem entsprechend der zwei-dimensionalen Normalverteilung

$$f(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (6.8)$$

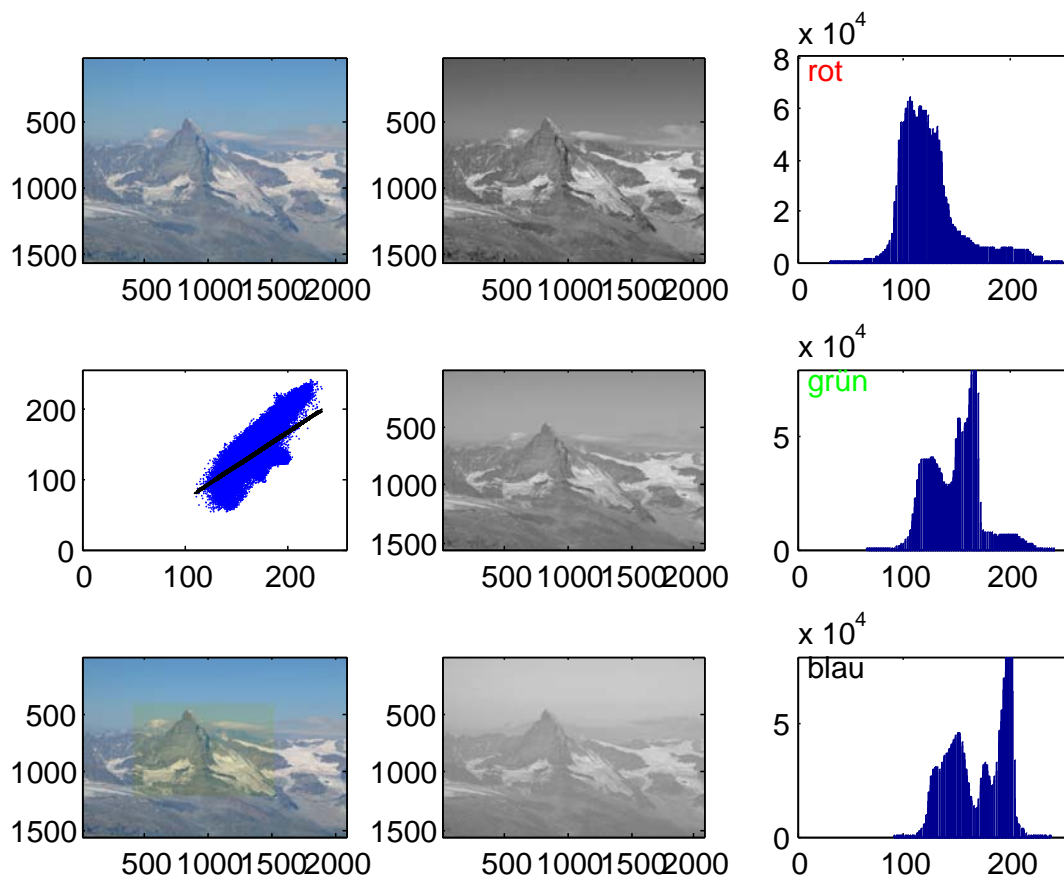


Abbildung 6.6: Dunststreuung

gefiltert wird. Als Filtermatrizen ergeben sich für den 3×3 bzw. 5×5 -Kernel

$$G_3 = \frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix} \quad \text{bzw.} \quad G_5 = \frac{1}{256} \begin{vmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{vmatrix}. \quad (6.9)$$

Im Gegensatz zum Mittelwertfilter wichtet der Gauß'sche Filter die einzelnen Pixel in Abhängigkeit von ihrem Abstand zum Mittelwert. Wir können an Stelle eines linearen Filters auch einen Gauß-Filter bei der Korrektur fehlender oder verrauschter Bildpunkte verwenden, der Kernel ist entsprechend dem in (6.9).

6.1.5 Dunstkorrektur

Bei Multispektralaufnahmen bieten sich noch weitere Verfahren zur Verbesserung der Bildqualität an, da jeder der betrachteten Frequenzbereiche seine eigenen Vorzüge und Schwächen hat (in Abhängigkeit vom verwendeten Filmmaterial, aber auch bedingt durch die atmosphärische Streuung). Durch geschickte Kombination der einzelnen Frequenzbereiche lassen sich ebenfalls Bildverbesserungen erzielen [101].

Abbildung 6.5 zeigt die Frequenzabhängigkeit der atmosphärischen Streuung und die Transmissionskurven verschiedener Filter. Die Streuung ist am stärksten im Ultraviolett bzw. blauen Bereich. Sie nimmt mit zunehmender Wellenlänge ab. Bei Luftbildaufnahmen wird die atmosphärische Streuung durch Verwendung eines Dunstfilters verringert, bei Falsch-

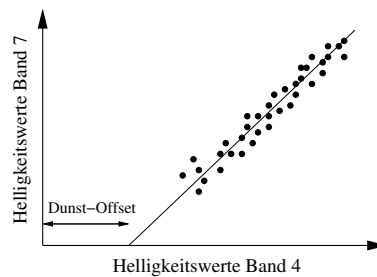


Abbildung 6.7: Dunstkorrektur durch Vergleich der Helligkeitswerte verschiedener Kanäle, hier am Beispiel des MSS auf LandSat

farbenfilmen wird die blaue Schicht auf Grund der starken Beeinflussung durch die Dunststreuung zu Gunsten der IR-empfindlichen Schicht weggelassen.

Einen Eindruck von der Dunststreuung in verschiedenen Kanälen des LandSat TM liefert Abb. 3.24. Dunststreuung stört jedoch nicht nur in Satellitenaufnahmen, wir können den Effekt bereits bei einer normalen terrestrischen Aufnahme über eine größere Distanz sehen. Dazu zeigt Abb. 6.6 im oberen linken Teil eine Aufnahme des Matterhorn vom Strahlhorn aus.² In der mittleren Spalte sind die einzelnen Schichten dieser Aufnahme gezeigt, von oben nach unten Rot, Grün und Blau. Die Dunststreuung ist im blauen Kanal gegenüber dem roten deutlich zu erkennen, auch im grünen Kanal zeigt sich noch Dunststreuung. In der rechten Spalte sind die Histogramme der Helligkeitswerte in den einzelnen Kanälen aufgetragen: der Peak bei großen Helligkeitswerten im blauen (und auch, wenn auch in etwas abgeschwächter Form) grünen Kanal entsteht durch die Dunststreuung. Die beiden unteren Aufnahmen in der linken Spalte geben eine Dunstkorrektur für einen Teil der Abbildung, das Verfahren wird im folgenden beschrieben.

Bei einem Multispektralscanner dagegen wird häufig trotz der starken Beeinträchtigung durch Dunst auch ein blauer Kanal verwendet (z.B. LandSat MSS, vgl. Abb 3.22). Bei LandSat wird der Kanal 7 (nahes IR, 0.8–1.1 μm) zur Korrektur von Kanal 4 (blau, 0.5–0.6 μm) verwendet. Atmosphärische Streuung reduziert den Kontrast, d.h. das Histogramm des entsprechenden Kanals wird gestaucht, insbesondere fehlen die Pixel mit sehr geringem Helligkeitswert: selbst bei der Aufnahme einer schwarzen Fläche werden durch die Atmosphäre immer Photonen in den Kanal gestreut. Dieses Fehlen der niedrigen Helligkeitswerte macht man sich bei der Korrektur zu Nutze.

Trägt man die Helligkeitswerte von Band 7 (IR) gegen die von Band 4 (Blau) auf, so findet sich grob ein linearer Zusammenhang, allerdings um einen Offset gegen den Ursprung verschoben. Dieser Offset entspricht den auf Grund der Dunststreuung in der Verteilung der Helligkeitswerte fehlenden Werten in Band 4. Dieser Offset wird anschließend von den Helligkeitswerten von Band 4 abgezogen, so dass ein korrigierter Kanal entsteht. Zur Kontrolle wird das Verfahren auch in Kombination von Kanal 4 mit Kanal 5 bzw. Kanal 6 vorgenommen, da die Fitgerade in Abb. 6.7 nicht nur durch Dunst sondern auch durch das unterschiedliche spektrale Reflektionsvermögen der verschiedenen Bildelemente entsteht. Das Verfahren funktioniert nur dann zuverlässig, wenn die Szene hinreichend dunkle Bereiche enthält.

Ein etwas mickriger Versuch einer Dunstkorrektur ist in Abb. 6.8 für einen Ausschnitt aus Abb. 6.1 gegeben. Links oben die vollständige Szene. Im rechten oberen Teilbild ist die Intensität des roten gegen die des blauen Kanals aufgetragen. Der Dunst bedingte Offset ist deutlich zu erkennen, er beträgt für diese Szene 20 Helligkeitswerte. Der korrigierte blaue Kanal (links unten) und das korrigierte Bild (rechts unten) zeigen natürlich nicht mehr Kontrast, da die Verteilung der Helligkeitswerte nicht gedehnt sondern nur verschoben wird. Allerdings

²Die horizontale Entfernung beträgt ca. 18 km bei einer Höhe von 4 000 m, entsprechend einer mittleren Dichte der Atmosphäre von 55% der Dichte am Boden, d.h. 0.71 kg/m^3 . damit ergibt sich eine Luftsäule zwischen Beobachter und Matterhorn-Ostwand von 12.7 t/m^2 . Der Druck der Atmosphäre am Boden entspricht dem einer 10 m Wassersäule, d.h. wir erhalten 10 t/m^2 , d.h. die beiden Luftsäulen und damit die mit ihnen verbundene Dunststreuung sind vergleichbar.

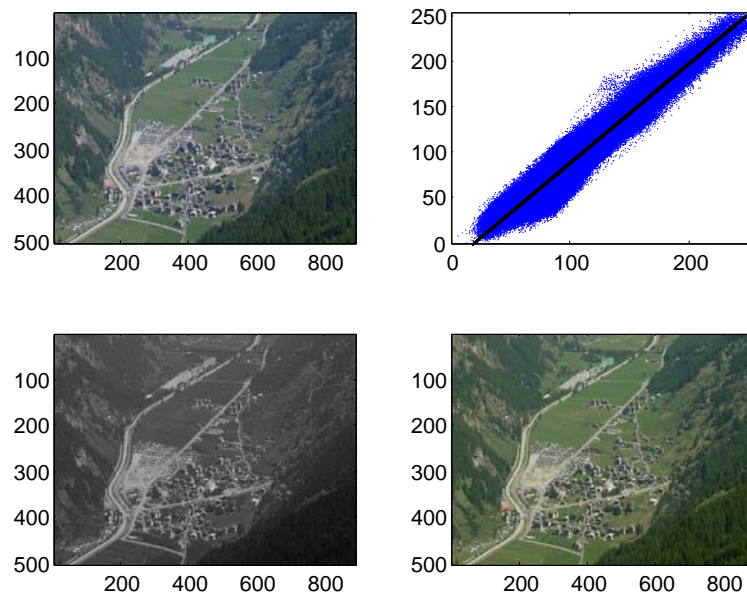


Abbildung 6.8: Dunstkorrektur an einem Ausschnitt aus Abb. 6.1

sind die absoluten Helligkeitswerte im Gesamtbild besser getroffen, so erscheint die Vegetation grün und nicht mehr bläulich. Damit ist sie für ein automatisches Klassifikationsverfahren auch besser zu erkennen als im nicht korrigierten Bild.

Die Korrektur in diesem Beispiel ist sicherlich schlechter als in vielen Satellitendatensätzen, da (a) die betrachteten Kanäle relativ dicht beieinander liegen (es wird der rote Kanal zur Korrektur verwendet, nicht ein Kanal im nahen IR) und (b) im roten Kanal nur relativ wenige wirklich dunklen Pixel, d.h. Helligkeitswerte von Null, auftreten. Damit ist die Bestimmung des Offsets relativ ungenau.

6.2 Bildverbesserung

Verfahren zur Bildverbesserung lassen sich grob einteilen in

1. Kontrastanreicherung: bei eng benachbarten Bildwerten (z.B. Gletscher) lassen sich Unterschiede optisch kaum erkennen. Kontrastanreicherung dient dazu, diese Unterschiede hervor zu heben und damit Strukturen erkennbar zu machen.
2. Kantenanreicherung: der Kontrast an einer Grenze zwischen zwei unterschiedlichen Bildbereichen wird hervor gehoben um diese Grenze deutlicher heraus treten zu lassen. Das Verfahren ist insbesondere zur Identifikation von durch den Menschen errichteten Strukturen hilfreich.
3. Erkennung linearer Strukturen.

6.2.1 Kontrastanreicherung

Die Originalaufnahmen der Satelliten sind häufig schwer zu interpretieren, das Hauptproblem ist der geringe Kontrast. Ein wichtiges Verfahren zur Verbesserung der Bildqualität ist die Kontrastanreicherung. Kontrastanreicherung kann z.B. dadurch erfolgen, dass man den Teil der Grauskala, in dem der größte oder interessanteste Teil des Bildes liegt, heraus nimmt und auf die gesamte Grauskala abbildet, d.h. unmittelbar benachbarte Grauwerte werden etwas auseinander gezogen, so dass sie für das Auge leichter zu unterscheiden sind. Für diese Abbildung gibt es verschiedene Verfahren, die in Abb. 6.9 zusammengefasst sind. Im oberen Teil der Abbildung sind jeweils Histogramme angegeben, die das Auftreten der verschiedenen

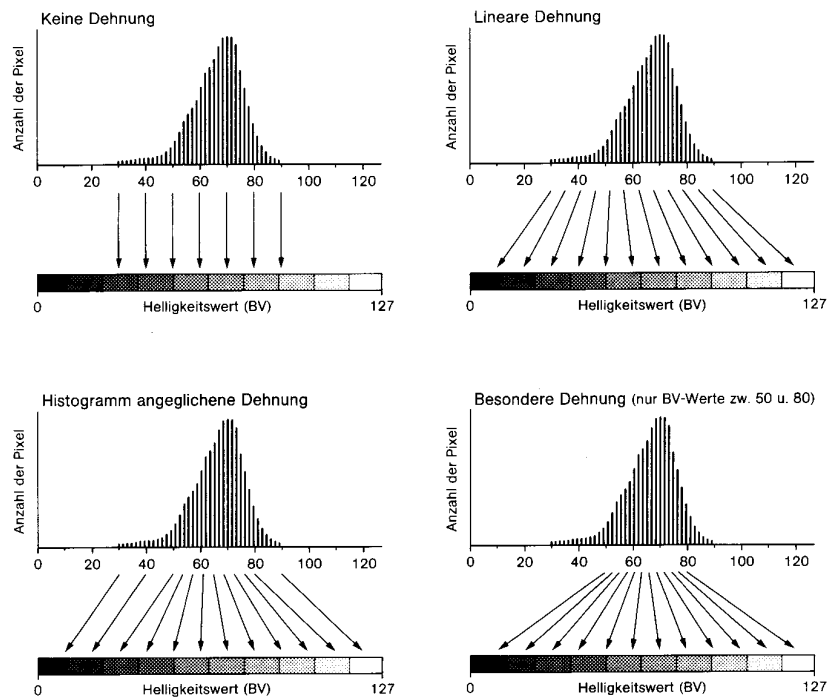


Abbildung 6.9: Dehnungsverfahren zur Kontrastanreicherung [71]

Grauwerte in einem Bild beschreiben. Von der gesamten Skala treten nur Grauwerte im Bereich von 30 bis 90 auf, d.h. wir können diesen Bereich der Grauskala auf die komplette Skala abbilden. Die lineare Abbildung ist dabei der einfachste Fall. Sie berücksichtigt jedoch nicht, dass die Grauwerte nicht gleich verteilt sind, sondern dass der größte Teil der Grauwerte in diesem Beispiel im Bereich zwischen 54 und 76 liegt. Daher kann man auch versuchen, diesen Bereich besonders stark zu dehnen (Histogramm angegliche Dehnung) oder nur diesen Bereich zu dehnen und die anderen (wenigen) Werte zu vernachlässigen (besondere Dehnung).

Der Algorithmus für die lineare Dehnung ist

$$p' = (p - lb) \cdot \frac{ub - lb}{w} \quad (6.10)$$

mit lb und ub als der unteren bzw. oberen Grenze des ausgewählten Bereichs und w als der Bandbreite des ursprünglichen Bereichs. Negative Werte von p' werden auf Null gesetzt, Werte größer W auf w .

Um ein Gefühl für die Histogramme der Helligkeitswerte zu erhalten, betrachten wir nochmals die Szene aus Abb. 6.3. Abbildung 6.10 zeigt links oben die Szene, in der Mitte die drei Spektralbänder, aus denen die Szene zusammengesetzt ist, und rechts die Histogramme für die einzelnen Kanäle. Helligkeitswerte liegen im Bereich zwischen 0 und 255, in allen drei Kanälen wird praktisch nur die untere Hälfte der Helligkeitswerte ausgenutzt, im oberen Bereich liegen nur noch wenige Werte. Während die Verteilungen im roten und grünen Bereich eine ähnliche Struktur mit einem Doppelpeak haben, ist die Verteilung im blauen Bereich um einen dominanten Einzelpeak konzentriert.

Da die Verarbeitung der Bilder kanalweise erfolgt, kann für jeden Kanal eine eigene Kontrastdehnung vorgenommen werden. Die Kanäle können dann anschließend zur Szene kombiniert werden.

In Abb. 6.11 wird nur die Dehnung des blauen Kanals aus Abb. 6.10 betrachtet. In der oberen Zeile sind die Originaldaten gegeben, in der zweiten Zeile wurden die Helligkeitswerte im blauen Kanal linear auf den Bereich zwischen 20 und 210 abgebildet, d.h. es wurde nur

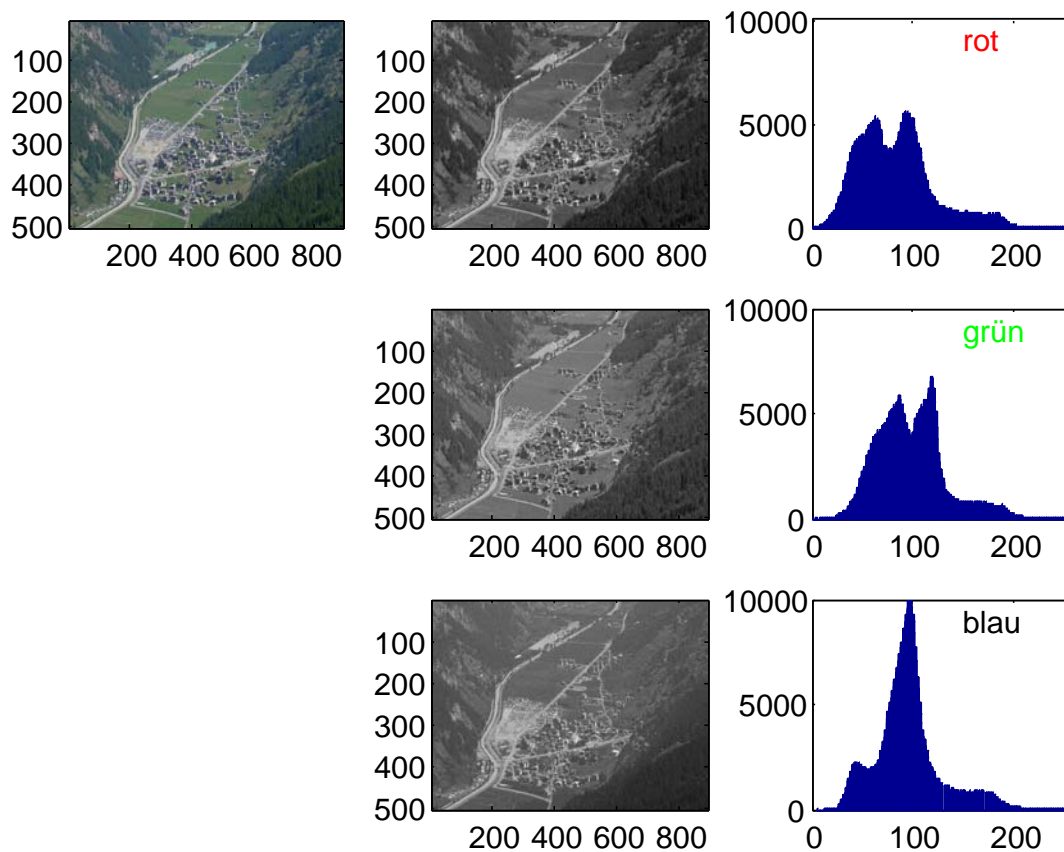


Abbildung 6.10: Die Szene oben links ist aus drei Teilbildern (rot, grün, blau, Mitte) zusammengesetzt, die Histogramme für die Pixelwerte sind jeweils in der rechten Zeile gegeben

der Bereich weg geschnitten, in dem sich ohnehin kaum Helligkeitswerte befinden. Das sich ergebende Histogramm ist im rechten Teil gezeigt, das sich ergebende Bild des blauen Kanals in der Mitte. Die Erhöhung des Kontrastes und damit die Erkennbarkeit von Strukturen wird deutlich. Bei Kombination mit den beiden (unbehandelten) Kanälen ergibt sich die links gezeigt Szene. Eine weitere lineare Streckung des Histogramms auf den Bereich zwischen 50 und 180 (der Originaldaten) ist in der unteren Zeile gezeigt.

Für eine allgemeine Bildverbesserung ist bereits die relativ moderate Kontrastanreicherung in der mittleren Zeile ausreichend, bei der nur die extremen Pixel unterdrückt wurden. Für spezielle Fragestellung (z.B. Identifikation bestimmter Objekte, Rauschunterdrückung) kann eine weitere Kontrastanreicherung auch auf relativ enge Bereiche der Verteilung sinnvoll sein.

Eine andere Situation entsteht in einer Szene mit starkem Kontrast, z.B. bei der Aufnahme eines Gletschers und des umgebenden Felsens. Abbildung 6.12 zeigt ein Beispiel, wieder die Gesamtszene links oben, die einzelnen Teilbilder in der mittleren Spalte und die dazu gehörigen Histogramme rechts. In allen drei Teilbildern sind die Histogramme ähnlich: eine bimodale Verteilung mit einem Maximum bei niedrigen Intensitätswerten (Fels) und einem zweiten Maximum bei hohen Werten (Eis).

Eine Kontrastanreicherung ähnlich der in Abb. 6.11, bei der nur die nicht auftretenden Helligkeitswerte abgeschnitten werden, verändert zwar das Bild, liefert aber keine zusätzliche Information: die Unterscheidung zwischen Eis und Gletscher ist auf Grund der bimodalen Verteilungen auch im unbearbeiteten Bild bereits eindeutig. Die Erkennung von Details im Fels und/oder auf dem Gletscher dagegen verbessert sich kaum, da die beiden Peaks auf Grund ihrer geringen Breite kaum gestreckt werden.

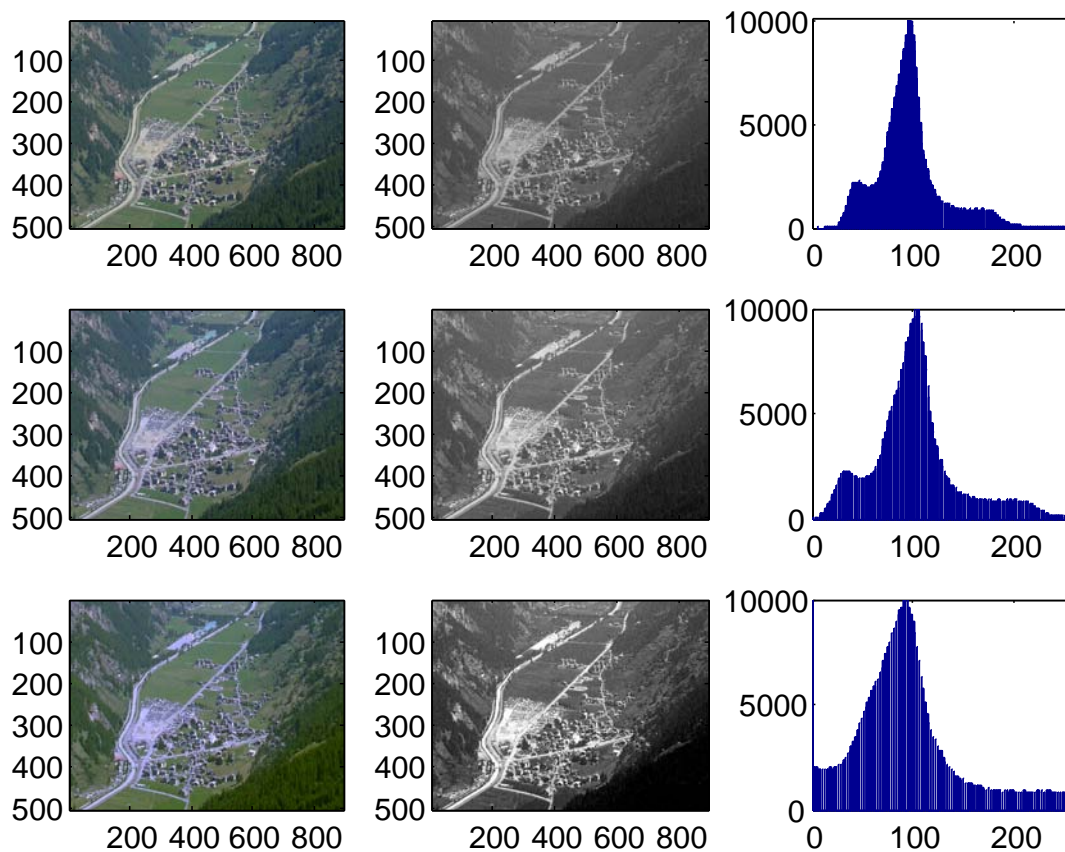


Abbildung 6.11: Für den blauen Kanal der Szene aus Abb. 6.10 wurde eine lineare Kontrastanreicherung vorgenommen: obere Zeile Originaldaten, mittlere Zeile Dehnung auf den Bereich von 20 bis 210, untere von 50 bis 180. In der linken Spalte wurde das gedehnte Bild im Blauen mit den ungedehnten Bildern in Rot und Grün überlagert

Abbildung 6.13 gibt in der oberen Zeile noch einmal die Originalszene, den blauen Kanal sowie das Histogramm des blauen Kanals. Die beiden unteren Zeilen enthalten eine lineare Dehnung: in der Gesamtszene (links) wurden alle Kanäle identisch gedehnt, in der Mitte ist das Ergebnis dieser Dehnung für den blauen Kanal gezeigt, rechts das sich für diesen Kanal ergebende Histogramm.

In der mittleren Zeile wurde der Bereich von Helligkeitswerten zwischen 140 und 215 linear auf den Gesamtbereich gedehnt. Dadurch verringern sich die Kontraste in den dunklen Bildpartien während die Kontraste im hellen Bereich erhöht werden. Insbesondere tritt das Höhenprofil des Gletschers deutlicher hervor: die unterschiedlichen Neigungen der Bildelemente führen dazu, dass das reflektierte Licht mit unterschiedlicher Intensität auf den Detektor trifft – allerdings immer mit großer Intensität. Durch die Dehnung werden diese Kontraste hervorgehoben und damit auch die Struktur.

In der unteren Zeile wurde der Bereich zwischen 10 und 100 auf den gesamten Helligkeitsbereich gedehnt. Daher verschwinden die Kontraste im Bereich des Gletschers während sie im Bereich der Felsen deutlicher hervortreten und die Unterkante des Bildes als Grasbewuchs statt Fels kenntlich wird.

Je nach Fragestellung an das Bild kann sich eine andere Form der Kontrastdehnung als sinnvoll erweisen: ein Glaziologe wird mit dem unteren Verfahren vielleicht nicht so zufrieden sein wie mit dem mittleren, ein Geoökologe dagegen bevorzugt vielleicht eher das untere Verfahren.

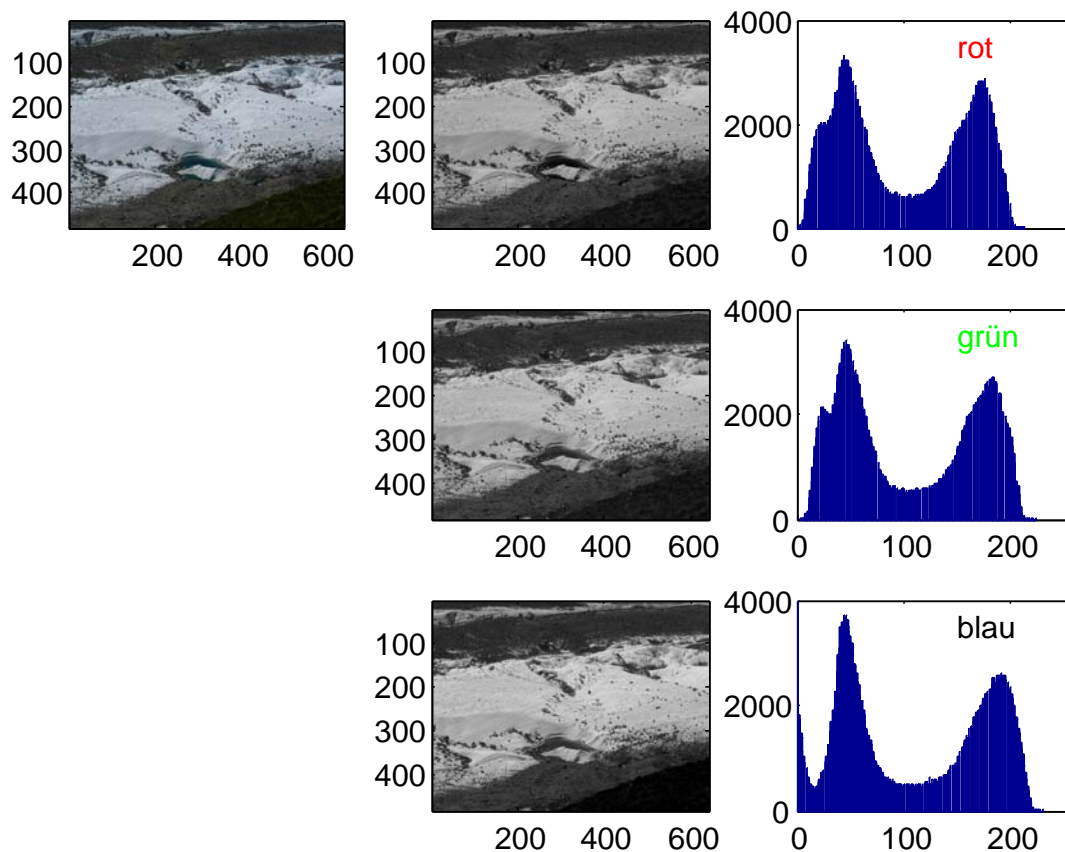


Abbildung 6.12: Teilbilder und Histogramme für eine Szene mit Gletscher und Fels, d.h. starkem Kontrast

6.2.2 Kantenanreicherung

Ein weiteres Verfahren ist die Kantenanreicherung ('edge enhancement'). Hat man einen weißen und einen schwarzen Streifen direkt nebeneinander (wie in dem Teststreifensystem zur Bestimmung des Bodenaufschließungsvermögens), so kann man die Lichtintensität als eine Stufenfunktion darstellen. Auf dem Film wird diese Stufe jedoch in einen kontinuierlichen Übergang transformiert, d.h. die ursprünglich scharfe Begrenzung zwischen den beiden Streifen verwischt etwas. Da aber für jeden Film diese Übertragungsfunktion bekannt ist, kann man eine Rücktransformation vornehmen, die die verwischte Kante wieder in eine scharfe Begrenzung überführt. Ein gutes Beispiel ist in Abbildung 9-13 in Cracknell und Hayes [16] gegeben.

Für digitalisierte Aufnahmen gibt es einige einfache Verfahren zur Kantenanreicherung, die ebenfalls mit Hilfe der in Abb. 6.2 durchgeführt werden kann. Die formale Grundlage zur Kantenanreicherung ist eine Laplace-Transformation [16, 52].

Bei der technischen Realisierung der Kantenanreicherung sind zwei Verfahren zu unterscheiden: gerichtete und nicht-gerichtete. Gerichtete Kantenanreicherung kann verwendet werden, um z.B. in Nord-Süd-Richtung verlaufende Strukturen zu erkennen. Ungerichtete Kantenanreicherung dagegen ist in allen Bereichen sinnvoll, in denen keine Vorinformationen über die Richtung einer möglichen Kante vorhanden sind oder diese sehr variable Richtung haben (z.B. Verkehrsstrukturen).

Ein ungerichteter Filter ist im linken oberen Teil von Abb. 6.14 gegeben: der Helligkeitswert des zentralen Pixels wird vierfach gewichtet, die Helligkeitswerte der direkt benachbarten Pixel jeweils einfach gewichtet abgezogen. Der so gewonnene Filterwert wird zum zentralen

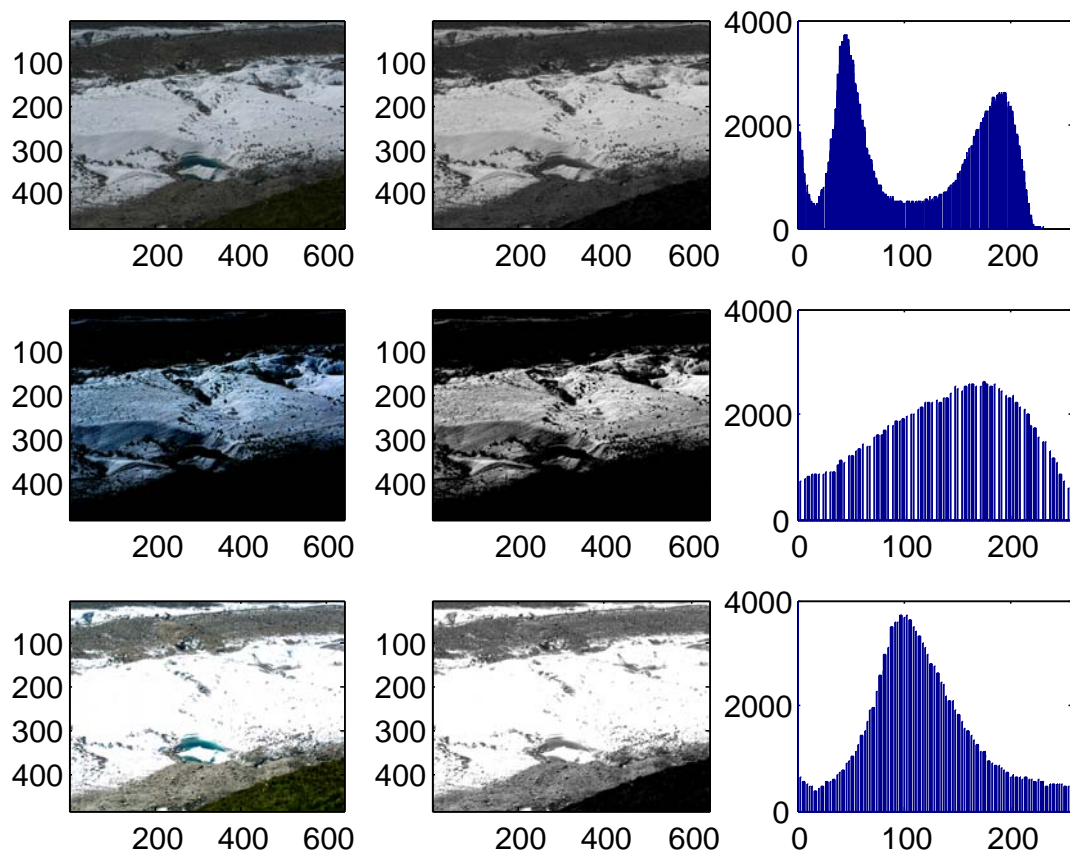


Abbildung 6.13: Lineare Kontrastanreicherung für Abb. 6.12; in der mittleren Zeile wurde der rechte Peak gedehnt, in der unteren der linke

Pixel addiert. Die obere Zeile von Abb. 6.14 zeigt einen Testdatensatz, die Zeile darunter gibt das Profil der Helligkeitswerte des originaldatensatzes entlang einer horizontalen Linie. Die folgende Zeile gibt das Profil nach der Kantenanreicherung, darunter der Datensatz nach Kantenanreicherung. In den Originaldaten beträgt der Kontrastunterschied an der linken Kante $80/60 = 1.33$; nach der Kantenanreicherung ist dieses Verhältnis auf $100/40 = 2.5$ erhöht, d.h. nahezu verdoppelt.

Abbildung 6.15 gibt ein Beispiel für eine Kantenanreicherung mit dem Kernel aus Abb. 6.14. Die Kontrasterhöhung an Kanten ist insbesondere im Bereich des Bahnhofs (links unten) und am Flusslauf zu erkennen. An letzterem wird auch deutlich, dass der Kernel eine ungerichtete Kantenanreicherung bewirkt.

Die Kantenanreicherung für Kanten unter einem beliebigen Winkel α beruht ebenfalls auf der Verwendung von zwei 3×3 -Kernen. Der erste Kernel

$$P_y = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{vmatrix} \quad (6.11)$$

hilft bei der Identifikation horizontaler Strukturen, Kernel

$$P_x = \begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix} \quad (6.12)$$

bei der Identifikation vertikaler Strukturen. Strukturen unter beliebigen Winkeln α werden

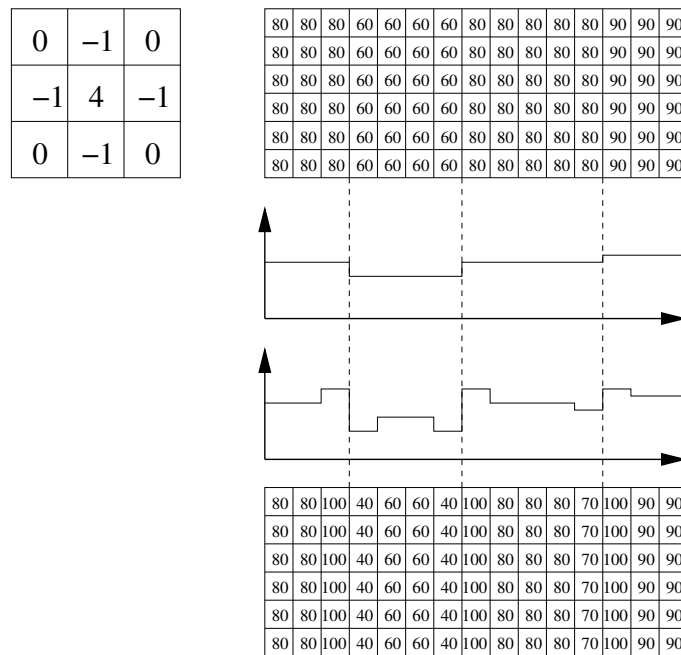


Abbildung 6.14: Ungerichtete Kantenanreicherung: Kernel; Originaldaten, Profil der Originaldaten und der gefilterten Daten und gefilterte Daten

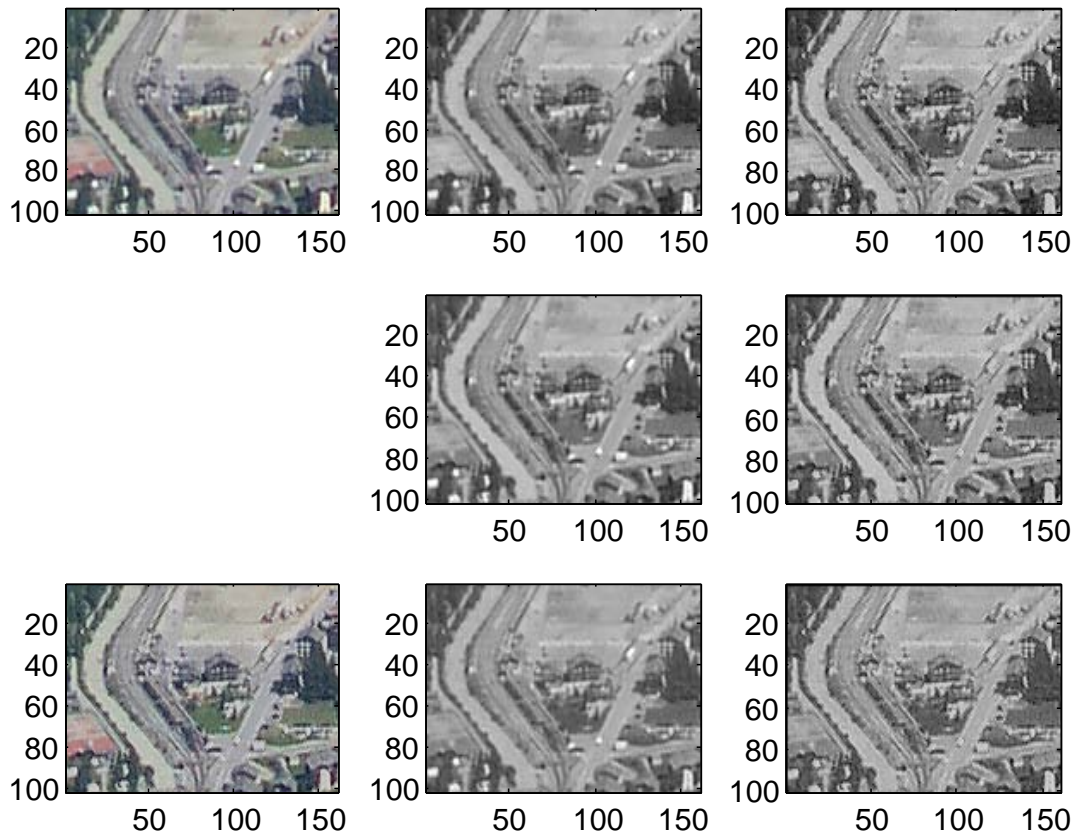


Abbildung 6.15: Szene aus Abb. 6.3 ohne (oben links) und mit (unten links) Kantenanreicherung mit dem Kernel aus Abb. 6.14. Die mittlere Spalte gibt die einzelnen Kanäle im Original, die rechte nach Kantenanreicherung, jeweils oben Rot, Mitte Grün und unten Blau

angereichert mit einem Filter

$$P = \cos \alpha P1 + \sin \alpha P2 = \cos \alpha \begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix} + \sin \alpha \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{vmatrix}. \quad (6.13)$$

Eine diagonal verlaufende Struktur wird entsprechend durch die Kernel

$$P_{45} = \begin{vmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{vmatrix} \quad \text{bzw.} \quad P_{-45} = \begin{vmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} \quad (6.14)$$

Auch hier wird der Filterwert zum Originalpixel addiert.

Die Operatoren P werden als Prewitt-Operatoren bezeichnet. Da sie nicht nur die beiden direkt der Kante benachbarten Pixel in einer Linie senkrecht zur Kante berücksichtigen sondern auch die jeweils parallelen Linien, bewirken sie nicht nur eine Kantenanreicherung sondern gleichzeitig auch eine Glättung. Beim Prewitt-Operator ist diese Glättung linear, beim Sobel-Operator wird eine Gauß-Wichtung vorgenommen:

$$S_x = \begin{vmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{vmatrix}, S_y = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{vmatrix}, \quad (6.15)$$

$$S_{45} = \begin{vmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{vmatrix} \quad \text{bzw.} \quad S_{-45} = \begin{vmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{vmatrix}. \quad (6.16)$$

Derartige Filter können z.B. zur Identifikation linearer Strukturen verwendet werden, wie auch im Zusammenhang mit Abb. 3.26 diskutiert.

6.3 Automatische Klassifikation

Falschfarbenaufnahmen mit IR-Filmen und später Multispektralaufnahmen und -scanner sollen dazu dienen, Objekte besser klassifizieren zu können als es bei einer Aufnahme im sichtbaren Bereich der Fall wäre. Die einzelnen Spektralkanäle der Instrumente und ihr Bezug zu den Eigenschaften der betrachteten Szene sind bereits im Zusammenhang mit den Instrumenten diskutiert.

Da die Identifikation die Kombination verschiedener Spektralkanäle voraussetzt, ist diese nicht mehr auf visuellem Wege vornehmbar sondern muss auf der Basis der digitalisierten Daten erfolgen – möglichst automatisch.

6.3.1 Distanzverfahren

Die einfachsten automatischen Klassifikationsverfahren sind 2D-Verfahren: hier werden zwei verschiedene Spektralkanäle gegen einander aufgetragen. In Abb. 6.16 wurden dazu die Kanäle 3 (rot) und 4 (nahes IR) des LandSat TM verwendet. In dieser Datenmatrix werden Bereiche markiert, die für bestimmte Aufnahmeobjekte typisch sind: Wasser (W), Stadtgebiet (U), Sandfläche (S), Getreidefelder (C), Wiesen (H) und Wald (F). Da einerseits Beleuchtungsverhältnisse und atmosphärische Transmission von Aufnahme zu Aufnahme wechseln und andererseits die betrachteten Objekte nicht homogen sind, bilden sich in der Datenmatrix relativ große Bereiche aus, in denen die Grauwerte für ein bestimmtes Objekt liegen können.

Mit Hilfe einer derartigen Matrix lässt sich eine automatische Klassifikation unbekannter Szenen vornehmen. Dazu gibt es verschiedene Verfahren:

- bei der *Methode der geringsten Distanz* (Teilbild a) wird der Schwerpunkt Verteilung der Zustandsvektoren eines jeden Objekts gebildet. Für ein zu klassifizierendes Pixel wird der

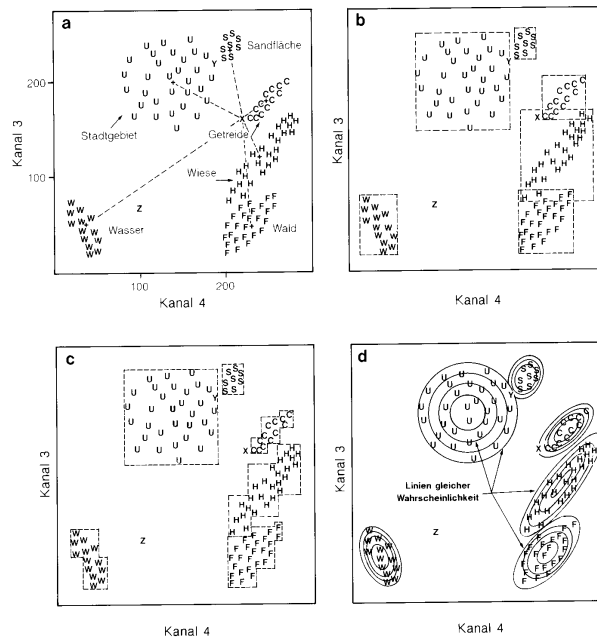


Abbildung 6.16: Zweidimensionale Klassifikationsverfahren [71]

Abstand zu jedem der Schwerpunkte gebildet, das Pixel wird dann dem Schwerpunkt zugeordnet, zu dem es den geringsten Abstand hat. Der maximale Abstand eines zu klassifizierenden Pixels von den Schwerpunkten kann angegeben werden, um bei großen Abständen eine Fehlklassifikation zu verhindern, so bei Pixel Z. Das Verfahren ist rechentechnisch sehr einfach.

- das *Quader-Verfahren* (Teilbild b) ist rechentechnisch ebenfalls einfach: hier wird jeder klassifizierte Bereich der Datenmatrix mit einem Rechteck umschrieben. Zu klassifizierende Pixel, die in ein derartiges Rechteck fallen, werden dieser Objektklasse zugeordnet; Bildpunkte, die in keines der Rechtecke fallen, werden nicht klassifiziert. Das Verfahren hat den Nachteil, dass Merkmalsgruppen sich nicht zwingend geschickt in ein Rechteck einordnen lassen (z.B. die Wiesen (H)) und sich Rechtecke überlappen können obwohl sich die Merkmalsklassen nicht überlappen (Wiese H und Getreidefeld C).
- das *detaillierte Quaderverfahren* (Teilbild c) ist eine Verfeinerung des Quaderverfahrens bei dem die Merkmalsklasse nicht in ein Rechteck sondern in eine aus mehreren kleinen Rechtecken gebildete Fläche eingeschlossen wird.
- das *Verfahren der größten Wahrscheinlichkeit* (Maximum Likelihood Method) unterscheidet sich von den Quaderverfahren dadurch, dass die Punkte einer Merkmalsklasse nicht gleichberechtigt behandelt werden sondern gewichtet werden. Dies basiert auf der Annahme, dass es einen Mittelwert für die Merkmalsklasse gibt, um den die Punkte in Form einer Normalverteilung schwanken. Dann lassen sich neben dem Mittelwert Linien gleicher Wahrscheinlichkeit festlegen. Dies ermöglicht nicht nur eine Klassifikation wie bei der Methode der geringsten Distanz sondern insbesondere bei überlappenden Merkmalsbereichen auch eine Aussage über die Wahrscheinlichkeit, mit der ein zu klassifizierender Bildpunkt einer oder mehreren Merkmalsklassen angehört. Das Verfahren der größten Wahrscheinlichkeit wird daher, obwohl rechentechnisch aufwendiger, häufig angewendet.

Insbesondere bei Untersuchungen der Vegetation kann eine multitemporale Klassifikation bessere Ergebnisse liefern: eine Birke im Frühjahr mit frischen Blättern erzeugt ein anderes Signal als die gleiche Birke im Sommer mit ausgewachsene Blättern oder gar später im Jahr in Herbstfärbung. Eine Verbesserung der automatischen Klassifikation ergibt sich, wenn die Trainingsmatrix nicht für einen festen Zeitpunkt aufgenommen wird sondern zu verschiede-

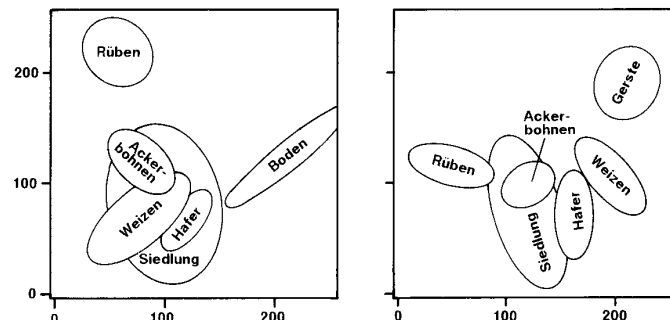


Abbildung 6.17: Unitemporale (links) und multitemporale (rechts) 2D-Klassifikation [71]

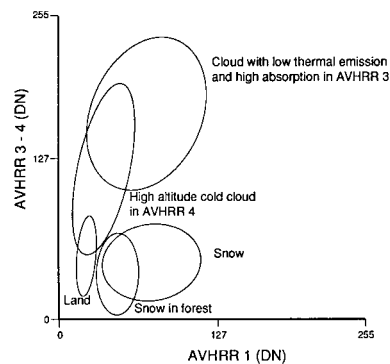


Abbildung 6.18: Klassifikation 'weißer' Pixel in den AVHRR-Daten zur Unterscheidung zwischen Schnee und Bewölkung [72]

nen Zeiten über die Vegetationsperiode verteilt. Als Beispiel zeigt Abb. 6.17 die Streuellipsen (Kurven fester Wahrscheinlichkeit beim Verfahren größter Wahrscheinlichkeit) im Merkmalsraum für verschiedene Landnutzungsklassen in unitemporaler (links) und multitemporaler (rechts) Auswertung. Generell ist bei der multitemporalen Auswertung eine bessere Trennung der verschiedenen Nutzungsklassen möglich – offensichtlich in der geringeren Überlappung der Streuellipsen.

Beispiel: Klassifikation von Schnee

Auf den ersten Blick sollte die Identifikation der Schneebedeckung in Satellitenbildern ein triviales Problem sein: Schnee ist weiß und reflektiert den größten Teil der auf ihn auftreffenden solaren Strahlung (Albedo zwischen 75% und 90%). Damit sollte er sich im sichtbaren Spektrum hell abzeichnen, d.h. durch einfache Betrachtung einer Schwarzweiß- oder Farbaufnahme erkennbar sein. Das ist korrekt, solange man nur die Bodenbedeckung betrachtet. Dass die Identifikation von Schneeflächen in der Realität jedoch schwieriger ist, kann man bei einem Blick auf eine normale MeteoSat-Aufnahme sofort nachvollziehen: außer Schnee haben auch Wolken ein sehr hohes Reflektionsvermögen, d.h. auch diese zeichnen sich auf den Aufnahmen als sehr helle Bereiche ab. Handelt es sich bei den Wolken um stark strukturierte Wolkengebilde (weit in die Troposphäre aufragende Gewitterwolken oder Wolkenwirbel im Zusammenhang mit einem Tiefdruckgebiet) so sind sie als solche zu identifizieren. Handelt es sich jedoch um eine gleichmäßige Bedeckung mit niedrigen Wolken, so ist eine Unterscheidung zwischen Wolken und Schnee kaum möglich. Hier kann man sich wieder, wie im Falle der Bodennutzung, mit Klassifikationsbereichen (in Abb. 6.19 dargestellt als Cluster, in Abb. 6.18 als Ellipsen) weiterbehelfen. Verwendet wird dabei ein im sichtbaren Bereich arbeitender Kanal (Kanal 1) des AVHRR sowie die Differenz aus zwei im thermischen Infrarot arbeitenden Kanälen. Der Kanal 3 ($3.55 \mu\text{m} - 3.93 \mu\text{m}$) ist empfindlich sowohl für die reflektierte solare

Strahlung als auch für die emittierte terrestrische Wärmestrahlung. Hier erscheint Schnee extrem dunkel. Zwar erscheinen auch Wolken dunkel, allerdings mit einer leichten Abweichung. Auch im Kanal 4 bestehen geringe Unterschiede zwischen Wolken und Schnee. Durch Differenzbildung zwischen den beiden Kanälen werden diese Unterschiede deutlicher hervorgehoben. Daher wird auch die Differenz für die Klassifikationsellipsen verwendet. Betrachtet man diese Klassifikationsellipsen in Abb. 6.18, so wird die Ähnlichkeit von Wolken und Schnee im sichtbaren Bereich (Kanal 1) sofort deutlich, während die Unterschiede zwischen Wolken und Schnee durch den Differenzkanal im Infraroten definiert werden.

6.3.2 Mehrdimensionale Klassifikationsverfahren

Mehrdimensionale Klassifikationsverfahren basieren auf den bereits von den 2D-Verfahren bekannten Ideen. Die Identifikation von Schnee in den AVHRR-Daten ist ein Hybrid-Verfahren: die Klassifikationsmatrix ist zweidimensional, die Informationen stammen jedoch aus drei Kanälen, da auf einer Achse nicht die Helligkeitswerte eines Kanals sondern die Differenz der Helligkeitswerte zweier Kanäle betrachtet wird.

Ein Beispiel für die Bestimmung der Bodennutzung in einem relativ eng begrenzten Gebiet zeigt Abb. 6.19. Die Zielsetzung ist dabei eine einfachere gewesen als bei der Bestimmung der globalen Vegetationsverteilung: Untersucht werden sollte ein Teil des Imperial Valley und Salton Sea in Kalifornien im Hinblick auf die Bodennutzung. Im oberen linken Teil von Abb. 6.19 ist zur Orientierung eine Aufnahme des LandSat MSS im Band 5 (grün) gezeigt, im unteren Teil die Klassifizierung entsprechend der Bodennutzung als (A) landwirtschaftlich genutzte Fläche, (D) Wüste, (M) Berge und (W) Wasser. Nichtmarkierte Pixel ließen sich nicht identifizieren. Durch Vergleich mit der Aufnahme darüber erkennt man, dass es sich bei diesen unidentifizierten Pixeln praktisch ausschließlich um Pixel in der Nähe von Begrenzungen handelt, d.h. sie werden wahrscheinlich Mischpixel sein.

Wie nimmt man aber eine solche Klassifikation vor? Wir hatten im Zusammenhang mit dem Vegetationsindex gesehen, dass dieser aus den Intensitäten im Roten und im Infraroten Kanal definiert wird. Zur Bestimmung der Flächennutzung beim LandSat MSS bedient man sich eines etwas genaueren Verfahrens: Es werden die Intensitäten der drei Bänder 4, 5 und 6 (entsprechend grün, rot und einem Kanal im nahen Infrarot) gegeneinander aufgetragen (Abb. 6.19 rechts unten). Entsprechend der unterschiedlichen spektralen Reflektionskurven sammeln sich die typischen Punkte von Vegetation, Wasser und anderen Geländetypen jeweils in Clustern. Zu jedem einzelnen Pixel der Scanner-Aufnahme wird nun in diesem Diagramm das Cluster gesucht, zu dem der Punkt mit der größten Wahrscheinlichkeit gehört bzw. dem er am nächsten liegt. Verschiedene mögliche Verfahren dieser Zuordnung sind für den vereinfachten zweidimensionalen Fall oben beschrieben.

Der in Abb. 6.19 gezeigte Bereich ist einfach zu klassifizieren, da er aus relativ großen Flächen mit einheitlichen Merkmalen besteht. Insbesondere die kleinräumigen rechteckigen Strukturen im Tal lassen sich unschwer als vom Menschen manipulierte Bereiche erkennen. Damit kann man die Klassifizierung in landwirtschaftlich genutzte Fläche aber auch ohne genaue Untersuchung 'von Hand' bzw. mit dem Auge vornehmen. Die Wasserläufe innerhalb der landwirtschaftlich genutzten Flächen lassen sich auf der Aufnahme im sichtbaren Bereich allerdings nur schwer erkennen, sie zeigen sich erst im Vergleich der verschiedenen Kanäle eindeutig. Die sehr unterschiedlichen Helligkeitswerte der verschiedenen landwirtschaftlich genutzten Flächen in der MSS Aufnahme zeigen aber auch, dass die Cluster im Klassifikationsdiagramm nicht gerade klein werden und man wirklich mehrere Kanäle verwenden muss, um eine einigermaßen zuverlässige Klassifikation zu erreichen.

Vegetationsuntersuchungen müssen nicht notwendigerweise so kompliziert durchgeführt werden. Um sich z.B. einen Eindruck über das Ausmaß des Abholzens tropischer Regenwälder zu verschaffen, ist eine einfache Betrachtung der Aufnahmen z.B. in einem Kanal im nahen Infrarot ausreichend, ein Beispiel hierfür gibt Abb. 3.11.

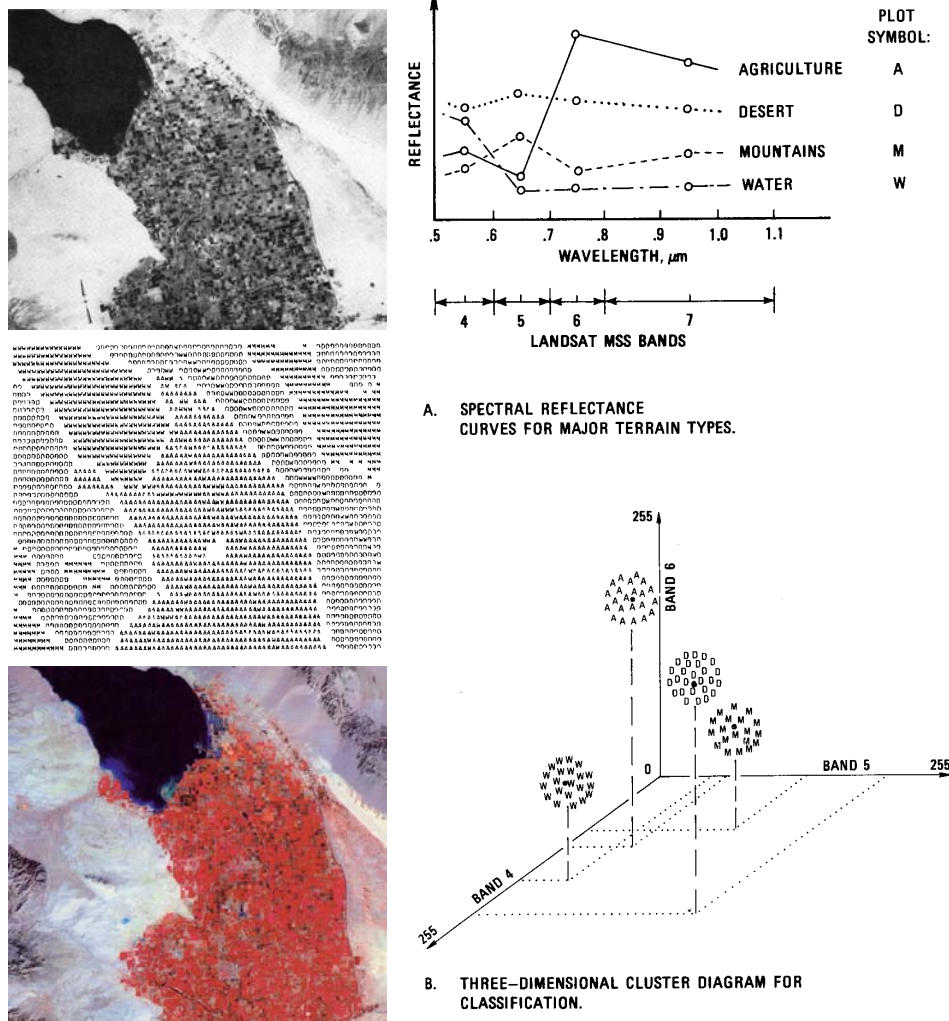


Abbildung 6.19: Klassifikation der Bodennutzung mit LandSat MSS Daten für das Imperial Valley und Salton Sea in Kalifornien. Links oben MSS Band 5 (grün), Mitte und unten die sich durch Vergleich verschiedener Spektralkanäle ergebende Geländeklassifikation. Rechts: Spektrales Reflektionsvermögen für verschieden genutzte Flächen (oben) sowie ein Clusterdiagramm zur Klassifizierung der einzelnen Bildpunkte [101]

6.4 Managing Gigabytes

Bisher haben wir Satellitendaten als Images betrachtet derart, dass jedem Pixel pro Kanal ein Helligkeitswert zugeordnet ist. Mit einer bescheidenen Bildgröße von 2560 mal 1920 Pixel mit einer Farbtiefe von 24 Bit bzw. 8 Bit für jeden der drei Kanäle ergibt sich ein Bitmap von 117 MBit oder knapp 15 MB. Bei einer Szene pro Minute entspricht das unkomprimiert einer Datenmenge von knapp 1 GB, d.h. eine Speicherung als Bitmap ist nicht sinnvoll.

Die Kompression von Bilddaten kann unter verschiedenen Gesichtspunkten erfolgen, eine Übersicht gibt [129]. Kompressionsverfahren können unterteilt werden in verlustfreie und verlustbehaftete Verfahren: in ersteren wird die vollständige, im Bild enthaltene Information codiert und entsprechend auch decodiert, in letzterem wird ein Teil der enthaltenen Information selektiv nicht codiert – entsprechend ist das Bild nach der Decodierung mit Verlusten behaftet.

6.4.1 Verlustfreie Bildformate

Standardformate zur Verlust freien Speicherung von Bilddaten sind GIF (Graphics Interchange Format) und PNG (Portable Network Graphics). Beide Verfahren basieren auf einem modifizierten Lempel–Ziv-Algorithmus, d.h. die Kompression ist ähnlich, als würde man ZIP auf ein Bitmap-Bild anwenden.

Im Gegensatz zu GIF erlaubt PNG optional ein weiteres Element der Kompression, die Kontext-basierte Kodierung. Dazu werden vor der Komprimierung einige Filter verwendet, z.B. ein horizontaler Differenzfilter, bei dem statt des Pixelwerts nur die Differenz zum Wert des Nachbarpixels bestimmt und kodiert wird; das Verfahren verfolgt also ähnliche Ansätze wie die Lauflängencodierung beim Fax. Da normalerweise in einer Szene benachbarte Pixel auch ähnliche Werte haben, treten viele kleine und wenige große Differenzen auf. Bei entsprechende Codierung, z.B. Huffman, ist das über die Differenzen gebildete Bild in codierter Form kleiner als das Original. Statt einer horizontalen Filters lässt sich auch ein vertikaler Filter verwenden oder ein mittlerer Filter, der die Differenz zum darüber und zum links davon liegenden Pixel bestimmt.

Für Grauskalenbilder ist FELICS ein weiteres Verlust freies Komprimierungsverfahren. Auch FELICS ist Kontext-basiert, berücksichtigt aber im Gegensatz zu PNG nicht nur den direkt benachbarten Pixel sondern 2 benachbarte Pixel. Da FELICS auf einem festen Codierungsschema basiert, kann es niemals besser komprimieren als 1 Bit pro Pixel

6.4.2 Und JPEG?

JPEG (Joint Photographic Experts Group) ist ein weit verbreiteter Standard der Bildkomprimierung – allerdings ist JPEG für unsere Fragestellungen nicht geeignet, da JPEG ein Verlust behaftetes Verfahren zur Datenkompression ist. Eine Verlust freie Variante, JPEG-LS, wird in Witten et al. [129] diskutiert.

JPEG ist ein Transformations-Codierungsverfahren, d.h. zuerst wird das Bild transformiert, anschließend wird die Transformation codiert. Die Decodierung erfolgt entsprechend: die eigentliche Decodierung liefert die Transformation des Bildes, dieses wird anschließend rücktransformiert.

Als Transformation wird bei JPEG eine diskrete Kosinustransformation auf ein $N \times N$ -Pixel großes Bild mit den Helligkeitswerten x :

$$\begin{aligned} X(0,0) &= \frac{1}{N} \sum_{i=0}^N \sum_{j=0}^N x(i,j) \\ X(u,v) &= \frac{2}{N} \sum_{i=0}^N \sum_{j=0}^N x(i,j) \cos\left(\frac{2(i+1)u\pi}{2N}\right) \cos\left(\frac{2(j+1)v\pi}{2N}\right) \end{aligned} \quad (6.17)$$

mit $X(0,0)$ als dem Gleichstromanteil und $X(u,v)$ mit $u, v \neq 0$ als dem Wechselstromanteil.

Ein reales Bild wird zuerst in 8×8 Pixel große Untermatrizen zerlegt.³ Auf diese Untermatrix wird die Kosinustransformation angewandt, so dass sich für jede Untermatrix 64 Koeffizienten $X(0,0)$ bis $X(7,7)$ ergeben. Da JPEG komprimieren soll, wird nur den niederfrequenten Anteilen der Transformation eine Bedeutung beigemessen. Bis zu welcher Ordnung diese Anteile gerücksichtigt werden sollen, bestimmt einerseits die Bildqualität und andererseits den Grad der Komprimierung. Diese unteren Koeffizienten werden mit einem Huffman-Verfahren codiert.

Abbildung 6.20 zeigt die Abhängigkeit des reproduzierten JPEG-Bildes von der Zahl der berücksichtigten Koeffizienten. Dazu wurde ein 512×512 -Pixel großes Bild in ein 8×8 -Raster von 64×64 -Pixel großen Untermatrizen zerlegt. Diese relativ großen Untermatrizen wurden gewählt, da sich mit ihnen die Abweichungen vom Original besser verdeutlichen lassen.

³Ist die Kantenlänge kein Vielfaches von 8, so wird zum Auffüllen der fehlenden Zeilen bzw. Spalten die jeweils letzte Zeile oder Spalte wiederholt.

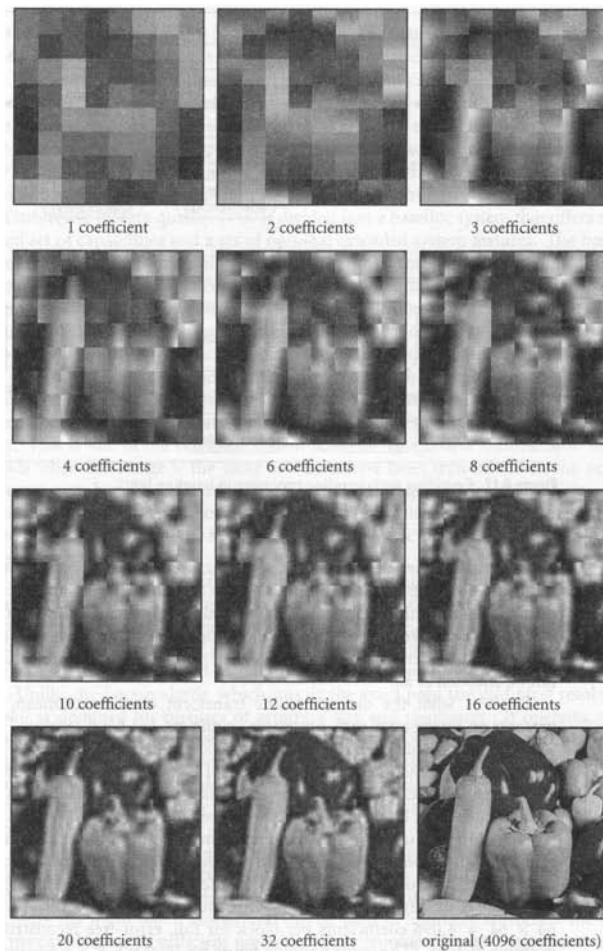


Abbildung 6.20: JPEG-Codierung und Decodierung für unterschiedliche Zahl der berücksichtigten Koeffizienten [129]

Aus dem Vergleich der Teilbilder wird deutlich, dass der Großteil der Information in den untersten Koeffizienten steckt – bereits mit 20 der 4096-Koeffizienten lässt sich ein Bild erzeugen (links unten), das die relevante Struktur des Originalbildes wieder gibt. Die höheren Koeffizienten verbessern nur noch den Detailreichtum.

Während die Helligkeitswerte im Originalbild mit mindestens 8 bit/Pixel und Kanal codiert sind, lassen sich bei einer JPEG-Codierung mit Raten von 0.2–0.5 Bit/Pixel bereits mäßige bis gute Bildqualitäten erreichen. Bei einer Codierung mit 0.5–0.75 Bit/Pixel ist die Bildqualität bereits gut bis sehr gut, bei 1.5–2 Bit/Pixel unterscheidet sich das Bild nicht mehr vom Original.

Die Wirksamkeit aller Komprimierungs- aber auch Korrekturverfahren beruht darauf, dass die Helligkeitswerte einzelner Pixel nicht völlig unabhängig sind und damit die Redundanz im Datensatz relativ groß ist. Ein Bild, in dem die Pixel statistisch erzeugt wurden, lässt sich weder sinnvoll korrigieren noch komprimieren.

JPEG wird genauer beschrieben in [88], Witten et al. [129] und [126]; letztere beschreiben auch andere gebräuchliche (und ebenfalls verlustbehaftete) Kompressionsalgorithmen wie MPEG und MP3.

JPEG-Codierung ist für Fernerkundungsinstrumente unter den Umständen geeignet, in denen das räumliche Auflösungsvermögen nicht voll ausgenutzt werden muss – allerdings kann dann eine initiale Reduktion der Daten durch Mittelung sinnvoller sein.

Auch wenn verlustbehaftete Verfahren wie JPEG bei der Speicherung von Bitmaps zwar

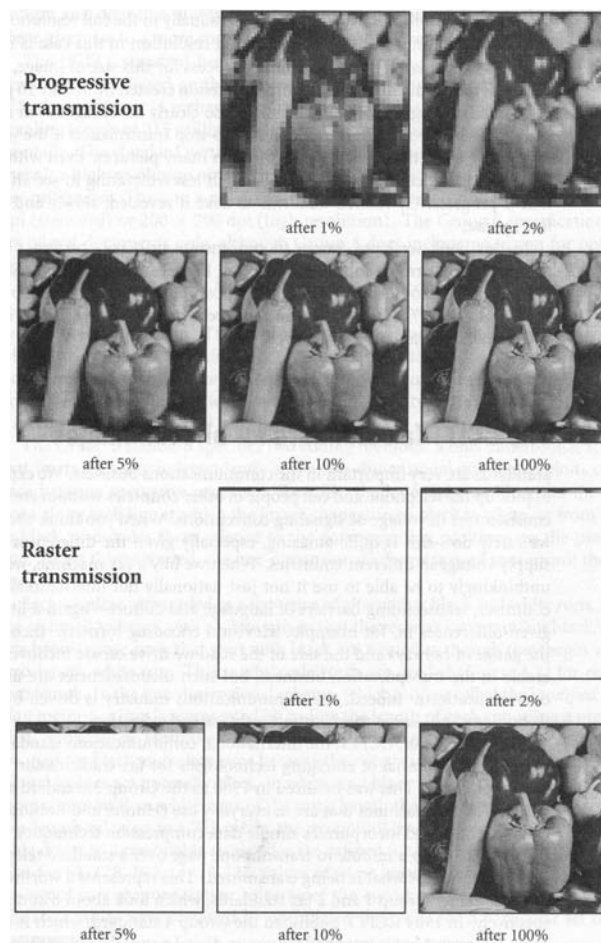


Abbildung 6.21: Progressive und Rasterdatenübertragung

eine hohe Komprimierung aber dafür einen Verlust an Information bedeuten, können derartige Komprimierungsverfahren für die Übertragung von Bildern auf Kanälen mit geringer Übertragungsrate sinnvoll sein. Ein als Bitmap gespeichertes Bild wird als Raster Punkt für Punkt übertragen, es baut sich also langsam auf, in dem jeweils ein Bildpunkt neben den anderen gesetzt wird, Zeile für Zeile (vgl. Abb. 6.21 unten). Diese Übertragungstechnik wird als raster transmission oder Rasterübertragung bezeichnet; bei geringer Übertragungsrate sieht man dem Bild beim Wachsen zu. Die Daten eines JPEG-Bildes dagegen erzeugen einen ganz anderen Bildaufbau. Da bei der progressive transmission bzw. progressiven Datenübertragung nicht Pixel sondern Koeffizienten einer Transformation übertragen werden, entsteht bereits nach der Übertragung des ersten Koeffizienten ein vollständiges, wenn auch extrem grobes Bild (Abb. 6.21 oben). Mit der Übertragung der weiteren Koeffizienten verändert sich nicht mehr die Geometrie des Bildes sondern seine Detailtiefe. Das hat z.B. für die schnelle Darstellung von Bildern im Internet Vorteile.

Bei der Übertragung über unsichere Kanäle, insbesondere die Unterbrechung einer Übertragung, liefern beide Verfahren völlig unterschiedliche Resultate: Bei der Rasterübertragung sind die bisher übertragenen Pixel alle exakt wieder gegeben – möglicherweise wissen wir dann über einen völlig uninteressanten Teil der Szene sehr genau Bescheid. Bei der progressiven Übertragung dagegen wird die Szene insgesamt erst einmal grob übertragen – für den Überblick sicherlich gut, auch wenn die Details noch fehlen.

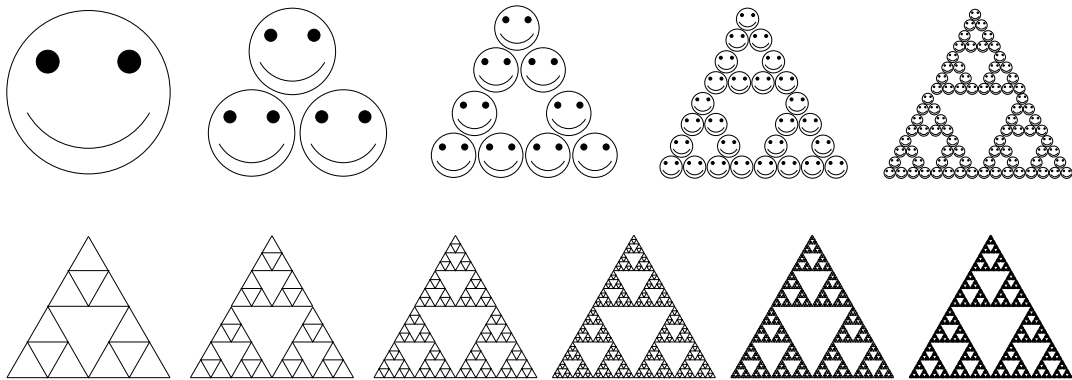


Abbildung 6.22: Fraktale Bilderzeugung: das Ausgangsbild (links) wird beim Kopieren verkleinert und es werden drei Kopien wie ein Dreieck zusammen gesetzt (2tes Bild von links). Der Vorgang wird mit der so entstandenen Kopie wiederholt. Die untere Zeile das Sierpinsky-Dreieck

6.4.3 Fraktale Bildkompression

Die bisher betrachteten Verlust freien Codierungsverfahren basieren auf dem Kontext zwischen benachbarten Pixeln sowie einer geschickten Huffman- oder Lempel-Ziv-Codierung dieser Ergebnisse. Fraktale Bildkompression wählt einen radikaleren Ansatz bei der Erzeugung der zu codierenden Sequenz: hier wird der Bezug zwischen größeren Teilen des Bildes hergestellt, in dem selbst-ähnliche Abschnitte identifiziert werden. Die Grundzüge dieser Verfahren sind in Fisher [23] dargestellt.

Fraktale Bildkompression setzt die Existenz selbstähnlicher Bereiche voraus. Selbstähnliche Bereiche zeigen auf allen räumlichen Skalen die gleiche Struktur: vergrößern wir eine Küstenlinie immer weiter, so zeigt jeder neue Ausschnitt eine ähnliche Struktur. Selbstähnliche Abbildungen haben keine natürliche Größe und lassen sich daher auf beliebige Formate vergrößern ohne dass, wie bei GIF und den anderen Formaten, die Pixel als solche sichtbar werden.

Die Grundidee der fraktalen Bildkompression lässt sich am leichtesten verstehen, wenn wir von der Bildrekonstruktion starten. Abbildung 6.22 illustriert das Abbildungsprinzip: das Ausgangsbild wird beim Kopieren skaliert und nach einem festen Schema wieder zusammengesetzt. In diesem Beispiel wird jeweils drei um einen Faktor 2 verkleinerte Kopien erstellt und in Form eines Dreiecks zusammen gesetzt. Mit jeder weiteren Kopie bleibt die Grundform, das Dreieck, unverändert während sich innerhalb der Struktur immer feinere Details ausbilden. Mit zunehmender Zahl der Kopien sind diese Details jedoch kaum noch wahrnehmbar: jedes der Sierpinsky-Dreiecke im rechten Teil der unteren Zeile zieht gleich aus – und vergrößern wir eines der Dreiecke, so finden wir jeweils seinen Vorgänger darin wieder.

Fraktale Bildkonstruktion basiert daher auf einem einfachen Satz von Regeln:

1. von einem Bild wird eine vorgegebene Zahl von Kopien erstellt,
2. diese Kopien werden nach einem vorgegebenen Regelwerk zusammengesetzt, wobei für jede Kopie Position, Skalierungsfaktoren (Verkleinerung, Verzerrung) und Rotationswinkel vorgegeben sind.⁴

Aus dem Prinzip der Bildkonstruktion können wir die Anforderung für die Bildkompression ableiten: es müssen Abschnitte des Bildes identifiziert werden, die selbstähnlich sind. Für diese Abschnitte muss die Grundfigur bestimmt werden und die Konstruktionsregeln. Idealerweise finden sich in einem Bild mehrere selbstähnliche Abschnitte, die aus der gleichen Grundfunktion konstruiert werden können. Daher muss zusätzlich zu den oben gegebenen

⁴Die Kopien müssen nicht so einfach zusammengesetzt werden wie in Abb. 6.22 sondern können bei geeigneter Skalierung auch Gebilde erzeugen wie z.B. den Barnsley Farn, hier in einer Abbildung von <http://obelix.polito.it/Tesi/Lug99/Mereu/default.htm>

Kopierregeln eine Anpassung von Kontrast und Helligkeit gegeben werden sowie eine Maske, die die Bereiche des Originals markiert, die durch diese Kopien ersetzt werden können.

Fraktale Bildkompression unterscheidet sich von den anderen Kompressionsverfahren durch eine Eigenschaft der Fraktale: da Fraktale auf jeder Längenskala Details zeigen, hat das decodierte Bild keine eigene Größe sondern kann in beliebiger Größe erzeugt werden. Mit zunehmender Größe des decodierten Bildes werden mehr Abbildungsschritte durchlaufen und damit mehr Details hinzu gefügt. Allerdings sind diese Details nicht real, sie erzeugen jedoch einen besseren optischen Eindruck als die bei der Vergrößerung eines Bitmaps entstehende körnige Struktur.

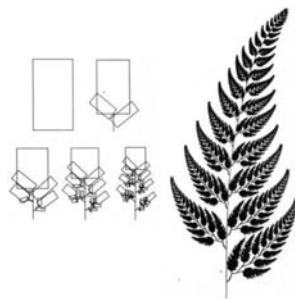
Fraktale Bildkompression ist ein verlustbehaftetes Verfahren, allerdings sind die Kompressionsraten wesentlich größer als z.B. bei JPEG und auf Grund der nicht vorhandenen Körnigkeit der optische Eindruck besser. Bisher hat sich fraktale Bildkompression nicht durchgesetzt, die hohen Kompressionsraten bei im Vergleich zu JPEG geringem Informationsverlust könnten das Verfahren jedoch auch für die Erdfernerkundung attraktiv machen: zum einen haben z.B. Küstenlinien, Flussläufe und Gebirge selbstähnliche Strukturen, zum anderen ist eines der Hauptprobleme der fraktalen Bildkompression, die Identifikation der selbstähnlichen Bereiche in einem Bild nicht ganz so gravierend, da in der Erdfernerkundung die gleichen Szenen immer wieder aufgenommen werden und damit Masken vorgegeben werden können bzw. Netze gut auf eine automatische Identifikation trainiert werden können.

Literatur

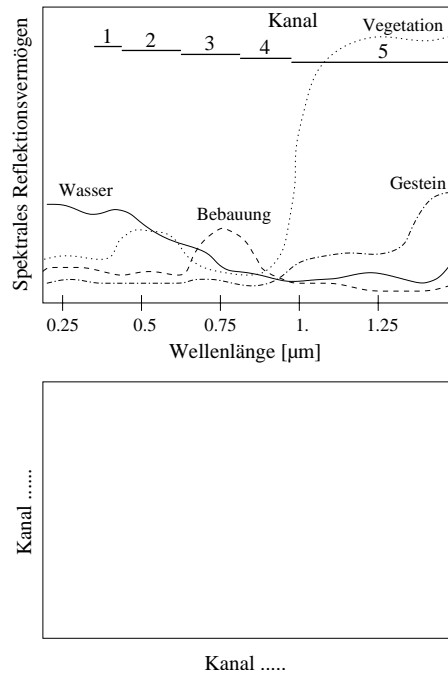
Eine gute Darstellung der theoretischen Grundlagen der Verfahren zur digitalen Bildverarbeitung gibt Jähne [52], Bildbearbeitung für Satelliteninstrumente wird diskutiert in Sabins [101], Cracknell und Hayes [16] und teilweise auch Löffler [71]. Eine gute Einführung, wenn auch mit Anwendungsbeispielen aus der Medizin, gibt auch Handels [38].

Fragen

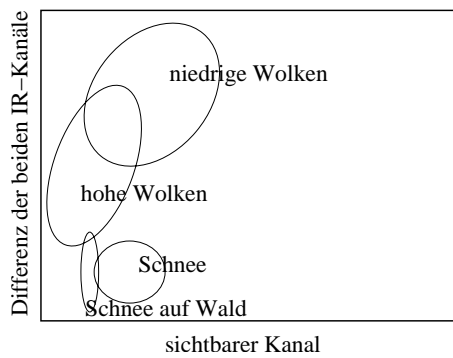
- Was versteht man unter Kontrastanreicherung? Beschreiben Sie Möglichkeiten zur Kontrastanreicherung.
- Satelliten liefern sehr große Datenmengen. Daher ist eine automatische Dateninterpretation erforderlich. Erläutern Sie Methoden, wie aus den Daten eines Multispektralscanners (z.B. Thematic Mapper oder MultiSpectralScanner auf LandSat) zumindest eine Grobunterteilung (z.B. in Wasser, landwirtschaftlich genutzte Fläche usw.) erfolgen kann.
- In der folgenden Abbildung ist das spektrale Reflektionsvermögen für verschiedene Objekte auf dem Boden gezeigt. Konstruieren Sie daraus ein Cluster-Diagramm, das eine Identifikation von Wasser, Gestein, Bebauung und Grünfläche auf der Basis der Kombination von drei Kanälen erlaubt (drei-dimensionales Zeichnen ist ein Problem, tragen Sie zwei Kanäle in das untere 2D-Diagramm ein und erläutern Sie den dritten verbal – z.B. die beiden Cluster liegen in der Projektion auf 2D zwar aufeinander, sind aber entlang der dritten Dimension eindeutig getrennt). Welche Kanäle wählen Sie? Ist eine Unterscheidung zwischen



Wasser und Bebauung unter Verwendung von nur zwei Kanälen möglich?



- Die folgende Abbildung zeigt ein Clusterdiagramm zur Klassifikation weißer Punkte in den AVHRR-Daten zur Unterscheidung von Schnee und Bewölkung. Verwendet wurden Informationen aus einem sichtbaren Kanal sowie die Differenz der beiden Kanäle im thermischen Infrarot. Welche weiteren Informationen (auch von anderen Instrumenttypen) könnte man verwenden, um eine genauere Unterscheidung zu erhalten?



Aufgaben

- Bei der Nachbearbeitung der Bilder am Boden müssen fehlerhafte Pixel erkannt und entfernt werden. Beschreiben Sie das Verfahren und führen Sie es an folgendem Beispiel durch (Hinweis: die Fehler sind ‘offensichtlich’):

34	24	36	30	28	27	30	25	33	32	26	33	32	23	29	32	51	25	30	34
26	33	25	33	35	24	33	23	26	29	30	31	33	27	29	28	25	27	25	36
31	23	27	26	23	35	35	28	33	24	36	32	30	26	34	36	23	24	28	35
33	31	30	28	26	26	30	35	27	23	55	27	28	31	26	32	34	36	33	25
29	27	34	24	35	32	29	25	28	24	26	34	29	35	35	26	32	31	23	36
31	24	29	32	25	34	32	35	23	25	34	24	24	29	27	29	25	28	24	29
36	35	32	33	36	24	33	36	25	32	24	36	32	28	31	36	23	33	31	30
25	26	44	29	32	25	31	29	27	28	32	34	33	28	32	30	76	24	29	34
31	28	24	36	28	35	36	31	26	31	36	24	28	30	29	28	34	34	36	33
33	23	34	28	33	28	24	33	32	28	26	34	27	33	27	32	30	36	24	26

- Bei der Fehlerkorrektur in folgendem Beispiel (99 bezeichnet die fehlerhaften Werte) können Sie wahlweise einen 3x3 oder einen 5x5 Kernel verwenden. Welcher produziert die besseren Ergebnisse und warum?

34	24	36	30	28	27	30	25	33	32	26	33	32	23	29	32	51	25	30	33
26	33	25	33	35	24	33	23	26	29	30	31	33	27	29	28	25	27	25	49
31	23	27	26	23	35	35	28	33	24	36	32	30	26	34	36	23	62	55	56
33	31	30	99	26	26	30	26	27	23	31	27	28	31	26	49	48	50	62	52
29	27	34	24	35	32	29	25	28	32	26	28	29	58	56	53	66	63	54	53
31	24	29	32	25	34	32	35	23	25	34	99	51	60	55	62	50	58	48	63
36	35	32	33	36	24	33	36	25	66	48	54	48	55	54	52	54	49	53	49
25	26	35	29	32	25	31	57	54	55	63	65	49	52	49	57	52	99	51	52
31	28	24	36	28	52	63	49	65	51	55	61	50	63	65	63	48	55	60	58
33	23	34	63	51	58	57	64	49	54	49	60	54	54	57	60	56	65	56	54

- Bei der Fehlerkorrektur in folgendem Beispiel können Sie wahlweise einen 3x3 oder einen 5x5 Kernel verwenden (fehlerhafte Pixel sind bereits markiert). Welcher produziert die besseren Ergebnisse und warum?

34	24	36	30	28	27	30	25	33	32	26	33	32	23	29	32	51	25	30	34
26	33	25	33	35	24	33	23	26	29	30	31	33	27	29	28	25	27	25	36
31	23	27	26	23	35	35	28	33	24	36	32	30	26	34	36	23	24	28	35
33	31	30	28	26	26	30	69	27	23	55	27	28	31	26	32	34	36	33	25
29	27	34	24	35	32	29	25	28	78	26	76	29	35	35	26	32	31	23	36
31	24	29	32	25	34	32	44	23	25	34	24	24	29	27	29	25	28	24	29
36	35	32	33	36	24	33	36	25	32	24	36	32	28	31	36	23	33	31	30
25	26	35	29	32	25	31	29	27	28	32	34	33	28	32	30	34	24	29	34
31	28	24	36	28	35	36	31	26	31	36	24	28	30	29	28	34	34	36	33
33	23	34	28	33	28	24	33	32	28	26	34	27	33	27	32	30	36	24	26

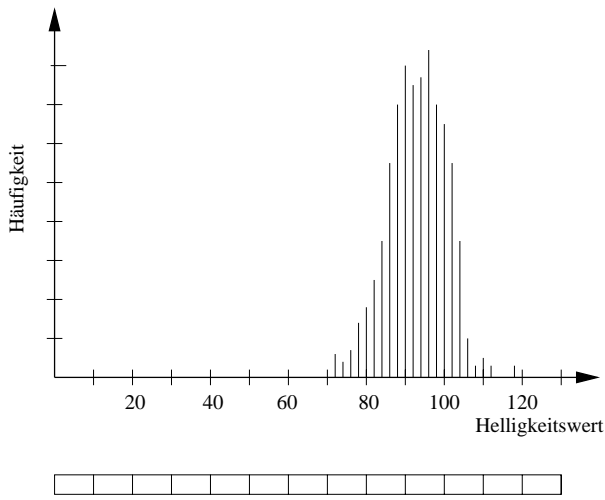
- Korrigieren Sie die folgenden, mit einem Spaltenfehler behafteten ‘Rohdaten’ (Schema erläutern und für die ersten vier Zeilen durchführen).

34	24	36	30	28	27	30	25	33	32	26	33	32	23	29	32	99	25	30	34
26	33	25	33	35	24	33	23	26	29	30	31	33	27	29	28	99	27	25	36
31	23	27	26	23	35	35	28	33	24	36	32	30	26	34	36	99	24	28	35
33	31	30	28	26	26	30	35	27	23	33	27	28	31	26	32	99	36	33	25
29	27	34	24	35	32	29	25	28	24	26	34	29	35	35	26	99	31	23	36
31	24	29	32	25	34	32	35	23	25	34	24	24	29	27	29	99	28	24	29
36	35	32	33	36	24	33	36	25	32	24	36	32	28	31	36	99	33	31	30
25	26	28	29	32	25	31	29	27	28	32	34	33	28	32	30	99	24	29	34
31	28	24	36	28	35	36	31	26	31	36	24	28	30	29	28	99	34	36	33
33	23	34	28	33	28	24	33	32	28	26	34	27	33	27	32	99	36	24	26

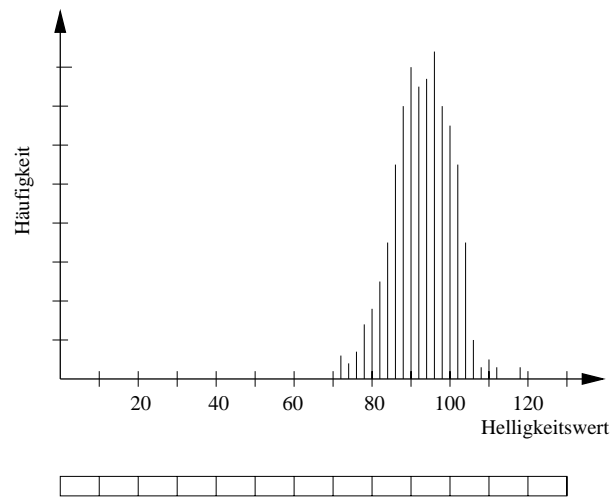
- In der folgenden Abbildung ist eine relativ scharfe Begrenzung (zwischen 5. und 6. Zeile) zwischen zwei Bereichen zu erkennen. Arbeiten Sie diese Grenze mit Hilfe der Kantenanreicherung stärker heraus und tragen Sie ihre Ergebnisse im fett umrandeten Bereich des rechten Teilbild ein.

34	24	36	30	28	27	30	25	33	32	34	24	36	30	28	27	30	25	33	32
26	33	25	33	35	24	33	23	26	29	26	33	25	33	35	24	33	23	26	29
31	23	27	26	23	35	35	28	33	24	31	23	27	26	23	35	35	28	33	24
33	31	30	28	26	26	30	35	27	23	33									23
29	27	34	24	35	32	29	25	28	24	29									24
52	61	54	57	54	59	55	50	48	51	52									51
52	60	49	50	53	49	50	59	57	61	52									61
56	48	60	55	62	57	49	61	55	52	56	48	60	55	62	57	49	61	55	52
54	50	57	52	60	59	62	49	57	53	54	50	57	52	60	59	62	49	57	53
49	54	52	61	49	55	51	57	62	54	49	54	52	61	49	55	51	57	62	54

- Die Auswertung einer Szene in einem verschneiten Gebirge gibt die folgende Verteilung von Helligkeitswerten. Wenden Sie eine lineare Kontrastdehnung (a) auf den vollständigen Datensatz an und (b) auf den Bereich der Helligkeitswerte, der mindestens 95% der Pixel umfasst.



- Konstruieren Sie für die folgende Verteilung von Helligkeitswerten eine Histogramm angeglichene Kontrastanreicherung.



Anhang A

Anhang

A.1 Liste der verwendeten Symbole

β	Ausströmrates
γ	universelle Gravitationskonstante
δ	Vorwärtskompensation
ε	Exzentrizität
ε_λ	spektrales Emissionsvermögen
$\bar{\varepsilon}$	mittleres Emissionsvermögen
η	wahre Anomalie
η_k	Kommunikationswirkungsgrad
$\vec{\lambda}_c$	Laplace-Vektor
λ	Wellenlänge
μ	Gesamtmasse im Gravitationsgesetz
μ	Absorptionskoeffizient
Ω	Rektaszension
ω	Epoche
ϱ	Dichte
ϱ	Signal-Rausch-Verhältnis
σ	Stefan-Boltzmann Konstante
σ_c	Bahnkonstante
θ	Winkelaufklärungsvermögen
a	große Halbachse
a_λ	spektrales Absorptionsvermögen
\vec{a}_s	Störbeschleunigung
A	Fläche
A_{eff}	effektive Fläche
b	Bildweite
$b(p)$	Shannon-Funktion
B	Bandbreite
B	Bildgröße
$B_\lambda(T)$	Planck-Funktion
c_D	Widerstandsbeiwert
c_k	Konstante im 3. Kepler'schen Gesetz
C	Kanalkapazität
D	Öffnungsdurchmesser Objektiv
d	Bildverschiebung
d	Hamming-Distanz
e_f	Effizienz eines Codes

E_b	Bitenergie
E_{kin}	kinetische Energie
f	Brennweite
f_e	Zahl erkennbarer Binärzeichenfehler
F	Informationsfluss
\vec{F}	Kraft
\vec{F}_{Reib}	Reibungskraft
F_B	Blickfeld
g	Gravitationsbeschleunigung
g	Gegenstandweite
G	Gegenstandsgröße
G	Antennengewinn
h	(Flug)Höhe
h	Planck'sches Wirkungsquantum
h_c	Energiekonstante
H	Entropie
H_0	maximale Entropie
i	Inklination
I	Intensität
I	Informationsgehalt
k	Kantenlänge Film
L	Raumdämpfung
m	Masse
m	mittlerer Mehraufwand pro Zeichen
m_0	Startmasse
m_B	Masse bei Brennschluss
m_N	Masse der Nutzlast
m_T	Masse des Treibstoffs
M_E	Masse der Erde
M_M	Masse des Mondes
M_P	Masse des Planeten
M_{\odot}	Masse der Sonne
M_Z	Masse des Zentralkörpers
p	Bahnparameter
p	Wahrscheinlichkeit
\vec{p}	Impuls
P	Umlaufperiode
P_N	Leistung des Rauschens
P_S	Leistung des Signals
q	Strahlungsstrom
\vec{r}	Ortsvektor
\vec{r}_{MM}	Ortsvektor zum Massenmittelpunkt
r	relative Redundanz
R	Redundanz
R_b	Bitrate
R_b	Bodenauflösungsvermögen
R_E	Erdradius
R_M	Radius des Mondes
R_{\odot}	Radius der Sonne
$r_{\odot,E}$	Abstand Sonne–Erde (astronomische Einheit AE, astronomical unit AU, 149 Mio km)
r_{EM}	Abstand Erde–Mond
S_{λ}	Ergiebigkeit
T	Umlaufzeit
T	Temperatur

T_b	Bitzeit
T_{eff}	Effektivtemperatur
t	Zeit
t	maximale Korrekturfähigkeit eines Codes
v, \vec{v}	Geschwindigkeit
v_A	Geschwindigkeit im Apozenrum
v_B	Geschwindigkeit bei Brennschluss
v_P	Geschwindigkeit im Perizentrum
v_k	Kreisbahngeschwindigkeit

A.2 Nützliche Konstanten

γ	$6.67259 \cdot 10^{-11} \text{ N m}^2/\text{kg}^2$	universelle Gravitationskonstante
σ	$5.67051 \cdot 10^{-8} \text{ W}/(\text{m}^2 \text{ K}^4)$	Stefan-Boltzmann-Konstante
c	$2.99792458 \cdot 10^8 \text{ m/s}$	Vakuum-Lichtgeschwindigkeit
e	$1.60217733 \cdot 10^{-19} \text{ C}$	Elementarladung
h	$6.6260755 \cdot 10^{-34} \text{ Js}$	Planck'sches Wirkungsquantum
M_E	$5.9742 \cdot 10^{24} \text{ kg}$	Masse der Erde
M_M	$7.3483 \cdot 10^{22} \text{ kg}$	Masse des Mondes
M_\odot	$1.9891 \cdot 10^{30} \text{ kg}$	Masse der Sonne
R_E	6370 km	Erdradius
R_M	1740 km	Radius des Mondes
R_\odot	697 000 km	Radius der Sonne
$r_{\odot,E}$	$1.49 \cdot 10^6 \text{ km}$	Abstand Sonne-Erde (astronomische Einheit AU)
r_{EM}	384 400 km	Abstand Erde-Mond

A.3 Abkürzungen

AATSR	Advanced Along Track Scanning Radiometer
ARGOS	Advanced Research and Global Observation Satellite
AIRS	Atmospheric Infrared Sounder
ALT	Altimeter
AMSR	Advanced Microwave Scanning Radiometer
AMSU	Advanced Microwave Sounding Unit
ASAR	Advanced SAR
ATLID	Atmospheric Lidar
AVCS	Advanced Vidicon Camera System
AVHRR	Advanced Very High Resolution Radiometer
BUV	Backscattered UV
CCD	Charge Coupled device
CGRO	Compton Gamma Ray Observatory
CHAMP	Challenging Minisatellite Ppayload
CZCS	Coastal Zone Color Scanner
DMSP	Defense Meteorological Satellite Program
DoD	Department of Defense
EnviSat	Environmental Satellite
EOS	Earth Observing System
ERBE	Earth Radiation Budget Experiment
EROS	Earth Resources Observation Systems
ERS	Earth Resource Satellite
ERTS	Earth Resources Technology Satellite
ESA	European Space Agency
ESSA	Environmental Science and Services Administration

ETM+	Enhanced Thematic Mapper+
FPR	Flat Plate Radiometer
GEOS	Geostationary Satellite
GERB	Geostationary Earth Radiation Budget
GIF	Graphics Interchange Format
GLONASS	Global Orbiting and Navigation Satellite System
GLRS	Geoscience Laser Ranging System
GOES	Geostationary Operational Environmental Satellite
GOME	Global Ozone Monitoring Experiment
GOMOS	Global Ozone Monitoring by Occultation of Stars
GPS	Global Positioning System
GRO	Gamma Ray Observatory
HCCM	Heat Capacity Mapping Mission
HIRIS	High-Resolution Imaging Spectrometer
HRIS	High-Resolution Imaging Spectrometer
HRV	High Resolution Visible
HST	Hubble Space Telescope
IFOV	Instantaneous Field of View
IR	Infrarot
ISEE	International Sun Earth Explorer
ITIR	Intermediate and Thermal Infrared Radiometer
ITOS	Improved TIROS Operational System
JPEG	Joint Photographic Experts Group
LAGEOS	Laser Geodynamics Satellite
LandSat	Land (Remote Sensing) Satellite
LAWS	Laser Atmospheric Wind Sounder
LRR	Laser Retro-Reflector
MAPS	Measurement of Air Pollution from Satellites
MERIS	Medium Resolution Imaging Spectrometer
MeteoSat	Meteorological Satellite
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MODIS	Moderate-Resolution Imaging Spectrometer
MOMS	Modularer Optoelektronischer MultispektralScanner
MSG	MeteoSat Second Generation
MSR	Multispectral Radiometer
MSRS	Multi-Spectral High Resolution System
MSS	Multi-Spectral Scanner
MWR	Microwave Radiometer
MVIR	Meteosat Visible and Infrared Radiometer
NAVSTAR	NAVigations System with Time And Ranging
NASA	National Aeronautics and Space Administration
NDVI	normalized difference vegetation index
NOAA	National Oceanographic and Atmospheric Administration
PNG	Portable Network Graphics
POES	Polar-orbiting Operational Environmental Satellite
RA-2	Radar-Altimeter-2
RBV	Return Beam Vidicon
SAR	Synthetic Aperture Radar
SAM	Stratospheric Aerosol Measurement
SCIAMACHY	Scanning Imaging Absorption Spectrometer for atmospheric Cartography
SeaSat	Sea (Remote Sensing) Satellite
SEM	Space Environment Monitor
SEVIRIO	Spinning Enhanced Visible and Infrared Imager
SMM	Solar Maximum Mission

SMMR	Scanning Multichannel Microwave Radiometer
SOHO	Solar and Heliospheric Observatory
SPM	Solar Proton Monitor
SR	Scanning Radiometer
SWH	significant wave height
TDRS	Tracking and Data Relay System
THIR	Temperature Humidity Infrared Radiometer
TIR	thermal infrared, thermisches Infrarot
TIROS	Television Infrared Observation Satellite
TM	Thematic Mapper
TOMS	Total Ozon Mapping Spectrometer
TOS	TIROS Operational Satellite
TRACER	Tropospheric Radiometer for Atmospheric Chemistry and Environmental Research
TVOS	TIROS Operational Vertical Sounder
VHRR	Very High Resolution Radiometer
VIS	visible
VISSR	Visible and Infrared Spin Scan Radiometer

A.4 Spickzettel

$$\text{Ellipse: } r = \frac{p}{1 + \varepsilon \cos \tau}, \quad r_{\text{Peri}} = \frac{p}{1 + \varepsilon}, \quad r_{\text{Apo}} = \frac{p}{1 - \varepsilon},$$

$$\varepsilon = \frac{\sqrt{a^2 - b^2}}{a}, \quad p = \frac{b^2}{a}$$

$$\text{Binet'sche Gleichung: } v = \sqrt{\mu \left(\frac{2}{r} - \frac{1}{a} \right)} \quad \text{mit } \mu = \gamma(m_1 + m_2)$$

$$\text{Winkelaufloesungsvermoegen: } \vartheta = \frac{a}{f} = \frac{\lambda}{D}$$

$$\text{Wien'sches Verschiebungsgesetz: } \lambda_{\text{max}} \cdot T = \text{const} = 2884 \mu\text{m/K}$$

$$M_{\text{Erde}} = 6 \cdot 10^{24} \text{ kg}$$

$$M_{\text{Sonne}} = 1.9 \cdot 10^{30} \text{ kg}$$

$$r_{\text{Erde}} = 6378 \text{ km}$$

$$r_{\text{Sonne}} = 695\,990 \text{ km}$$

$$\gamma = 6.67 \cdot 10^{-11} \text{ Nm}^2/\text{kg}^2$$

$$\sigma = 5.67 \cdot 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$$

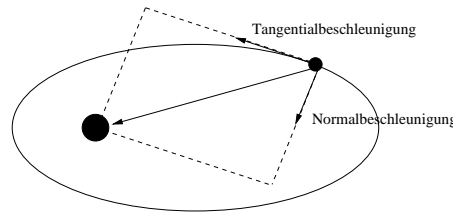
Anhang B

Hinweise zur Lösung einiger Aufgaben

B.1 Kapitel 2

B.1.1 Fragen

- (1) Die Planeten bewegen sich auf Ellipsen, in deren gemeinsamen Brennpunkt die Sonne steht. (2) Der von der Sonne zum Planeten gezogene Fahrstrahl (Radiusvektor \vec{r}) überstricht in gleichen Zeiten Δt gleiche Flächen ΔA : $\Delta A/\Delta t = \text{const.}$ (3) Die Quadrate der Umlaufzeiten T_1, T_2 zweier Planeten verhalten sich wie die Kuben der großen Halbachsen a_1 und a_2 : $T_1^2/T_2^2 = a_1^3/a_2^3$.
- Das Newton'sche Gravitationsgesetz $F_g = \gamma m_1 m_2 / r^2$ liefert eine anziehende Kraft zwischen den beiden Körpern. Die Kreisbahn (bzw. Ellipse) kann nur dann entstehen, wenn einer der Körper eine Geschwindigkeitskomponente senkrecht zur Verbindungsachse der beiden besitzt: sonst ziehen sich die Körper entlang dieser Achse an und bewegen sich aufeinander zu bis sie kollidieren. Die Geschwindigkeitskomponente senkrecht zur Verbindungsachse bedeutet nach Newtons erstem Axiom (ein Körper verharrt im Zustand der gleichförmigen gradlinigen Bewegung bzw. in Ruhe), dass der Körper sich in dieser Richtung weiter bewegen will. Dies wird von einem Beobachter auf dem zweiten Körper als eine nach außen gerichtete Scheinkraft, die Zentrifugalkraft $F_Z = m\omega^2 r$ interpretiert. Das Gleichgewicht aus Anziehungskraft und Zentrifugalkraft liefert eine Kreisbahn; die Geschwindigkeit kann durch Gleichsetzen $F_g = F_Z$ bestimmt werden zu $v_K = \sqrt{\gamma m_2 / r}$. Zur Herleitung der Ellipsenbahn muss die Bewegungsgleichung vollständig gelöst werden.
- An der Gravitationskraft ändert sich nichts. Der Unterschied ergibt sich aus der Scheinkraft, die sich aus dem 1. Newton'sche Axiom ergibt und als Zentrifugalkraft bezeichnet wird: aufgrund der hohen Bahngeschwindigkeit ist diese Scheinkraft im Perigäum größer als es die Zentrifugalkraft wäre, es bleibt eine scheinbare Kraftkomponente in Bewegungsrichtung. Daher entfernt sich der Planet bzw. Satellit von einer Kreisbahn auf eine Bahn mit größerem Radius. Dabei wird kinetische Energie in potentielle Energie umgewandelt und die Geschwindigkeit nimmt ab. Die Abnahme der Geschwindigkeit erfolgt so lange, wie durch die Trägheit eine Bewegungskomponente vom Zentralkörper weg erhalten bleibt. Ist diese aufgebraucht, so befindet sich der Satellit im Apogäum mit einer Geschwindigkeit kleiner der dort erforderlichen Kreisbahngeschwindigkeit – die Zentrifugalkraft kann die Anziehungskraft nicht kompensieren und der Körper fällt wieder auf den Planeten zu. Eine mögliche Skizze, allerdings zu einer verbal etwas anderen Argumentation, wäre:



4. Das eingeschränkte Zweikörperproblem entspricht unserer umgangssprachlichen Beschreibung von Planeten- bzw. Satellitenbahnen: die Planeten bewegen sich auf Ellipsenbahnen um die Sonne bzw. die Satelliten kreisen um die Erde, d.h. wir haben die Vorstellung einer feststehenden Zentralmasse und eines darum kreisenden Probekörpers. In der allgemeinsten Situation (beide Massen sind vergleichbar) kreist jedoch nicht der eine Körper um den anderen sondern beide um ihren gemeinsamen Schwerpunkt. Das ist auch beim Planetensystem bzw. bei einem Satelliten in einer Erdumlaufbahn der Fall. Allerdings fällt bei sehr ungleichen Massen der Schwerpunkt $m_1(\vec{r}_1 - \vec{r}_S) + m_2(\vec{r}_2 - \vec{r}_S) = 0$ praktisch in die schwerere Masse, so dass diese als der ruhende Zentralkörper betrachtet werden kann. Die Bewegungsgleichung wird dann zu

$$\frac{d^2\vec{r}}{dt^2} = -\frac{\gamma(m_1 + m_2)}{r^3} \vec{r} = -\frac{\mu}{r^3} \vec{r}.$$

5. Die Erhaltungssätze sind Energieerhaltung und Drehimpulserhaltung. Die Energieerhaltung ergibt sich aus der Bewegungsgleichung

$$\frac{d^2\vec{r}}{dt^2} = -\frac{\mu}{r^2} \vec{e}_r$$

durch skalare Multiplikation mit $d\vec{r}/dt$. Dann ergibt sich auf der linken Seite ein Ausdruck der Form

$$\frac{d^2\vec{r}}{dt^2} \cdot \frac{d\vec{r}}{dt} = \frac{1}{2} \frac{d}{dt} \left(\frac{d\vec{r}}{dt} \right)^2 = \frac{1}{2} \frac{dv^2}{dt},$$

ein Ausdruck also, der ein Maß für die kinetische Energie ist. Auf der rechten Seite haben wir ein Skalarprodukt aus $\vec{r} \cdot \vec{r}$, das wir entsprechend zusammenfassen können, d.h. wir erhalten

$$\frac{dv^2}{dt} = -2\mu \frac{dr^{-1}}{dt}.$$

Integration der Gleichung liefert das Energieintegral

$$v^2 - \frac{2\mu}{r} = h_c,$$

d.h. die Summe aus kinetischer und potentieller Energie ist in jedem Punkt entlang der Bahn konstant.

Die Drehimpulserhaltung ergibt sich aus der Bewegungsgleichung durch vektorielle Multiplikation mit \vec{r} :

$$\vec{r} \times \frac{d^2\vec{r}}{dt^2} = -\frac{\mu}{r^3} \vec{r} \times \vec{r} = 0.$$

Multiplikationsregeln für Vektorprodukte erlauben das Umschreiben der linken Seite in die Form

$$\frac{d(\vec{r} \times \vec{v})}{dt} = 0,$$

Integration liefert das Momentintegral

$$\vec{r} \times \vec{v} = \vec{\sigma}.$$

Zusatzpunkt: durch skalare Multiplikation mit \vec{r} ergibt sich daraus (wieder unter Berücksichtigung der Regeln für Mehrfachprodukte) die Gleichung der Bahnebene $\vec{\sigma} \cdot \vec{r} = 0$.

6. wahr, falsch, falsch, falsch, falsch, wahr
 7. Die erste kosmische Geschwindigkeit ist die Kreisbahngeschwindigkeit. Sie ergibt sich für eine große Halbachse a gleich dem Radius r der Kreisbahn zu

$$v_{\text{kosmisch1}} = v_{\text{kreis}} = \sqrt{\frac{\mu}{r}}.$$

Einsetzen der Zahlenwerte (Spickzettel) liefert $v_{\text{Kreis}} = 7.92 \text{ km/s}$. Die zweite kosmische Geschwindigkeit beschreibt das Entweichen aus dem Schwerefeld des Körpers, d.h. es entsteht keine geschlossene Bahn und die große Halbachse a wird ∞ . Damit ergibt sich

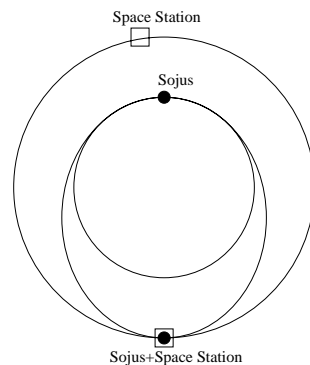
$$v_{\text{kosmisch2}} = \sqrt{\frac{2\mu}{r}} = \sqrt{2} v_{\text{Kreis}} = \sqrt{2} v_{\text{kosmisch1}}.$$

8. Beschreibung der Bahn als Ellipse durch die beiden Parameter *große Halbachse* und *Exzentrizität*. Beschreibung der Lage der Bahn im Raum durch die *Inklination*, d.h. die Neigung der Bahnebene gegenüber der Äquatorebene. Die Lage der Bahnebene im Raum wird durch die *Rektaszension*, d.h. den Winkel zwischen der Knotenlinie aus Bahnebene und Äquatorebene und einer festen Bezugsrichtung, z.B. Richtung des Frühlingspunktes, beschrieben. Damit ist die Bahnebene im Raum festgelegt, nicht jedoch die Bahn der Ellipse in dieser Ebene. Für letztere wird zusätzlich benötigt der Winkel ω zwischen der Knotenlinie und der Richtung zum Perigäum (*Apsidenlinie*). Damit ist die Lage der Bahnebene und der Bahn in dieser Ebene durch die genannten fünf Parameter bestimmt. Zur Festlegung des Satellitenortes wird zusätzlich die Angabe einer *Epoche*, d.h. eines Punktes zu einer bestimmten Zeit (z.B. Perigäumsdurchgang, Kreuzen der Knotenlinie) benötigt.
9. Die Inklination gibt die Neigung der Bahnebene eines Satelliten gegenüber der Äquatorebene. Für einen antriebslosen Satelliten ist die niedrigste Inklination seines Orbits durch die geographische Breite des Startorts bestimmt, da der Massenmittelpunkt des Planeten in der Bahnebene liegen muss. Daher ist ein geostationärer Satellit über Moskau nicht möglich.
10. Anwendung der Binet'schen Gleichung (Spickzettel) und Verwendung der dort gegebene Ausdrücke für die Orte von Apogäum und Perigäum liefert

$$v_{\text{Apo}} = \sqrt{\frac{\mu}{a} \frac{1-\varepsilon}{1+\varepsilon}} \quad \text{und} \quad v_{\text{Peri}} = \sqrt{\frac{\mu}{a} \frac{1+\varepsilon}{1-\varepsilon}}.$$

11. Die Erde zwingt den Mond auf eine Bahn mit dem Radius R und der Periode T . Die Sonne zwingt die Erde auf eine Bahn mit dem Radius $400R$ und der (ungefähren) Periode $13T$. Die Kreisbahnbedingung fordert $\omega^2 R \sim R/T^2 \sim M/R^2$ oder $M \sim R^3/T^2$ mit M als der Masse des jeweiligen Zentralkörpers (damit haben Sie Kepler's 3. Gesetz wiederentdeckt). Die Masse des anderen Körpers fällt jeweils heraus, Division der beiden Gleichungen liefert $M_{\odot}/M_{\text{Erde}} = 400^2/13^3 = 3.8 \cdot 10^5$. Auf diese Weise lässt sich aus bekannter (oder zumindest gut abgeschätzter) Erdmasse die Sonnenmasse bestimmen.
12. Auf Grund der relativ hohen geographischen Breite hat die Bahn des Kleinsatelliten eine Inklination von mindestens 50° . Diese ist für Untersuchungen in niedrigen Breiten nicht geeignet, da die Bahn für zu lange Zeiträume außerhalb des interessierenden Breitenbereiches verläuft. Daher lieber in ein polares Orbit einschließen und mit dem Flüssigwasser-Instrument das Abschmelzen der Polkappen studieren.
13. wahr (zumindest, wenn die Bahn geschlossen bleibt), falsch, wahr, wahr, falsch
 14. Die Idee lässt sich aus den Transferbahnen ableiten: der Sojus-Kommander muss einmal beschleunigen, um auf eine Übergangselipse zu kommen, die den Orbit der Space-Station berührt und dort ein zweites Mal um mit der entsprechenden Bahngeschwindigkeit in diesem Orbit zu bleiben. Damit wird die Bedingung, dass die Relativgeschwindigkeit verschwinden soll, automatisch erfüllt. Im Gegensatz zur Transferbahn ist das Problem hier, dass nicht nur der Orbit der Space Station sondern diese selber erreicht werden soll. Generell gilt: mit zunehmender Höhe nimmt zwar die Bahngeschwindigkeit eines Satelliten

in einem kreisförmigen Orbit zu, seine Winkelgeschwindigkeit jedoch ab (kann man auch über Keplers drittes Gesetz erhalten). Der Sojus-Kommander muss sein Bahnmanöver daher so einleiten, dass er nach dem erfolgten zweiten Schub seinerseits (Übergang von der Transfer- auf die Kreisbahn der Raumstation) am gleichen Ort ist wie die Raumstation und deren Geschwindigkeit hat. Dazu muss er die Bahn der Raumstation so treffen, dass sich die Raumstation hinter ihm befindet und während seiner Beschleunigungsphase zu ihm aufschließt. Nach dem dritten Kepler'schen Gesetz gilt $T_{\text{Sojus}} < T_{\text{Übergang}} < T_{\text{ISS}}$, da die große Halbachse der Übergangselipse zwischen den beiden Kreisbahnen liegt. Erfolgt die Beschleunigung in unendlich kurzer Zeit, muss die Sojus-Kapsel hinter der Space Station in einen Transfer-Orbit gehen (vergl. Abbildung), mit zunehmender Beschleunigungszeit wird der 'Vorlauf' der Space Station geringer, eventuell auch negativ.



15. Aufschlag wird zum Problem, da der Ball nicht wie gewohnt hoch geworfen und dann geschlagen werden kann: beim simplen Hochwerfen bleibt die lineare Geschwindigkeit des Balls erhalten, d.h. seine Winkelgeschwindigkeit nimmt gegenüber des Shuttles ab und er fällt zurück. Ein modifizierter Aufschlag könnte das Schlagen aus der Hand sein (Aua, das kann weh tun) oder ein konventioneller Aufschlag, wobei aber nicht nach oben sondern schräg nach vorne geworfen und ungefähr ein halbes Shuttle-Orbit bis zum Aufschlag gewartet werden müsste. Geht ein Ball ins Aus, so hängt die weitere Bewegung davon ab, wo der Ball ins aus geht. Verfehlt der Spieler am Bug den Ball, so handelt es sich um einen Ball, der eine größere Geschwindigkeit hat als das Shuttle und damit in ein elliptisches Orbit geht, dessen Perigäum in der Shuttle-Bahn liegt. Da verloren gegangener Ball und Shuttle unterschiedliche Umlaufgeschwindigkeiten haben (die großen Halbachsen der Orbits sind verschieden) treffen sich beide erst nach dem kleinsten gemeinsamen Vielfachen der beiden Umlaufzeiten wieder – was insbesondere bei kleinen Differenzen der Umlaufzeiten verdammt lang sein kann. Verfehlt der Spieler am Heck den Ball, so geht dieser ebenfalls in ein elliptisches Orbit, allerdings ist die Shuttlebahn in diesem Falle das Apogäum. Liegt das Perigäum recht niedrig, so wird der Ball durch Reibung in der Atmosphäre abgebremst und stürzt ab. In diesem Fall kann der Ball nicht wieder eingefangen werden, da er auch dann, wenn er nicht sofort abstürzt, nicht mehr sein altes Apogäum erreicht. Wird der übers Heck entweichende Ball nicht in der Atmosphäre abgebremst, so gelten die gleichen Überlegungen wie für den über den Bug entweichenden Ball. Die Fälle senkrecht nach oben und unten sind, wie für oben bereits diskutiert, der nach- oder vorlaufende Ball. Bälle, die senkrecht zur Flugbahn entweichen, haben eine Bahnebene, die gegenüber der Ausgangsbahnebene gedreht ist (Winkel ist durch das Verhältnis der Geschwindigkeiten bestimmt) und gelangen in eine Übergangselipse, deren Perigäum mit der Bahnhöhe des Shuttle übereinstimmt
16. Bei den verschiedenen Ellipsen mit gleicher großer Halbachse steht das Zentralgestirn mit zunehmender Exzentrizität immer weiter vom Schnittpunkt der beiden Halbachsen entfernt.
17. bei einer sonnensynchronen Bahn steht die Bahnebene stets in einem festen Winkel zur Achse Sonne–Erde, d.h. der Satellit überquert den Äquator stets zur gleichen lokalen

Zeit. Dadurch verändern sich die Lichtverhältnisse nicht, was einen Vergleich der Beobachtungen von Orbits verschiedener Daten erleichtert.

18. Die *Bahnhöhe* wird so gewählt, dass eine lange Lebensdauer gewährt wird (keine Reibung), gleichzeitig aber ein gutes Bodenaufklärungsvermögen erreicht werden kann. Typische Werte liegen im Bereich von 700 km bis 1000 km. Ausnahme: geostationäre Wettersatelliten. Aus der Flughöhe ergibt sich automatisch die Umlaufzeit, die Wiederholfrequenz wird durch die Breite des Streifens entlang der Bodenspur bestimmt, in dem beobachtet wird. Die Bahnen sind *kreisförmig*, da alle Punkte aus gleichem Abstand zu beobachten sind (Optimierung von oben als Randbedingung). Die *Inklination* wird so gewählt, dass der interessierende Breitenbereich abgedeckt wird: Erdfernerkundungssatelliten haben daher meistens Inklinationen in der Nähe von 90° . Ausnahme: Satelliten in geostationären Orbits oder Satelliten, die nur tropische Bereiche beobachten sollen. Die Bahn wird *sonnensynchron* gewählt, um gleichbleibende Lichtverhältnisse zu erhalten.
19. Bei der Spinstabilisierung ist der Satellit selbst ein großer Kreisel; das Verfahren ist jedoch für die Erdfernerkundung nicht geeignet, da die Ausrichtung der Instrumente auf einem Spinner nicht trivial ist. Bei der Dreiaxsenstabilisierung verfügt das Raumfahrzeug über drei Kreisel mit senkrecht aufeinander stehenden Achsen, die eine genaue Ausrichtung erlauben. Dies Verfahren ist für die Erdfernerkundung geeignet. Für Forschungssatelliten gibt es ferner eine Magnetfeldstabilisierung, die einem Erdfernerkundungssatelliten jedoch wieder nicht die richtige Richtung eröffnen würde.
20. Die beiden gezielt eingesetzten Störungen ergeben sich aus der Abplattung der Erde. Die Bahndrehung kann zur Erzeugung einer sonnensynchronen Bahn verwendet werden, die entscheidenden Parameter sind die Inklination i und der Bahnparameter p des Satellitenorbits. Bei geeigneter Wahl lässt sich eine Drehung der Bahnebene um 360° innerhalb eines Jahres erreichen, d.h. eine sonnensynchrone Bahn. Die zweite Störung ist die Drehung der Bahnellipse (Apsidendrehung). Diese hängt ebenfalls von der Inklination ab; die Apsidendrehung verschwindet für eine Inklination von 63.4° . Diese Inklination ist für alle Satelliten zu wählen, die sich in einem stark elliptischen Orbit befinden und deren Bahn sich in der Bahnebene nicht drehen soll (z.B. Molniya-Satelliten).
21. Die wichtigsten Störprozesse sind: (1) die Reibung in der Erdatmosphäre, die zu einer Abbremsung von Satelliten führt und damit letztendlich zu deren Absturz. Die Reibung in der Hochatmosphäre hängt von deren Dichte ab und nimmt daher mit zunehmender Bahnhöhe und abnehmender solarer Aktivität ab. (2) Bahnstörungen durch die Abplattung der Erde, dazu gehören die Drehung der Apsidenlinie (d.h. die Drehung der Bahnellipse in der Bahnebene) und die Drehung der Bahnebene selbst.. Letzterer Effekt wird zur Erzeugung einer sonnensynchronen Bahn ausgenutzt. Die Bahnstörungen werden mit zunehmender Bahnhöhe kleiner, da dann die Erde aus der Sicht des Satelliten immer mehr zu einer Punktmasse zusammenschrumpft. (3) Hier lassen sich verschiedene, alle eher mindere Einflüsse auflisten wie (i) gravitative Störungen durch Sonne oder Mond (insbesondere in hohen Orbits), (ii) Strahlungsdruck (für Ballonsatelliten), (iii) Sonnenwind, (iv) thermische Emissionen, (v) atmosphärische Gezeiten und (vi) zufällige Einflüsse.
22. Formal muss in der Bewegungsgleichung zusätzlich zur Gewichtskraft die Reibungskraft als Störbeschleunigung berücksichtigt werden:

$$\frac{d^2\vec{r}}{dt^2} = -\mu \frac{\vec{r}^3}{r} + \vec{a}_s .$$

Die Reibungskeft wird proportional zur Geschwindigkeit angenommen,

$$F = -Dv = \frac{c_D \rho v^2 A}{2} ,$$

wobei in der Proportionalitätskonstante $D = c_D \rho v A / 2$ die Querschnittsfläche A , der aerodynamische Beiwerte c_D , die Dichte ρ und die Geschwindigkeit v enthalten sind.

Das formale Verfahren zur Herleitung kann entweder durch Lösen der Bewegungsgleichung erfolgen (eher aufwendig) oder aus einer einfacheren Energiebetrachtung. Aus dem

Gleichgewicht von Zentripetal- und Gravitationskraft lässt sich für den Zusammenhang zwischen kinetischer und potentieller Energie die Beziehung $E_p = 2 E_k$ finden. Aus der Gleichung für die Kraft lässt sich durch Multiplikation mit der Geschwindigkeit eine Beziehung für die zeitliche Änderung der kinetischen Energie herleiten:

$$\frac{d}{dt} \frac{mv^2}{2} = - \frac{\rho v^3 A c_D}{2}.$$

Die zeitliche Änderung der kinetischen Energie lässt sich auch schreiben als

$$\frac{d}{dt} E_k = \frac{d}{dr} E_k \frac{dr}{dt}.$$

Durch Kombination der beiden Gleichungen ergibt sich die Änderung der Flughöhe zu

$$-\frac{dr}{dt} = \frac{\rho v A c_D r}{m}.$$

Die Beobachtungsgröße ist die (Veränderung der) Umlaufzeit, da sich diese genauer feststellen lässt als die Veränderung der Flughöhe.

23. die Abbremsung erfolgt hauptsächlich im Perigäum, da hier die Dichte der Atmosphäre am größten ist. Solange die Geschwindigkeit noch deutlich über der Kreisbahngeschwindigkeit liegt (was sie zumindest anfangs sein sollte, da sonst das Orbit nicht elliptisch wäre), ändert sich die Perigäumshöhe kaum. Allerdings erreicht der Satellit nicht mehr die Apogäumshöhe, d.h. die Exzentrizität nimmt ab, die Bahn wird immer kreisförmiger. Die Umlaufzeit entlang der Bahn nimmt ebenfalls ab. Erst wenn die Geschwindigkeit im Perigäum in die Nähe der Kreisbahngeschwindigkeit abgesunken ist, verringert sich auch die Perigäumshöhe und der Satellit stürzt ab.
24. Die untere Grenze beträgt ca. 87 min (Faustregel knapp 90 min), entsprechend einer Höhe von 150 km (Faustregel knapp 200 km). Eigentlich ist es nicht sinnvoll, zu genaue Werte anzugeben, da die Dichte der Atmosphäre mit der solaren Aktivität schwankt und ein Orbit im solaren Minimum noch stabil sein kann, im solaren Maximum dagegen zum sofortigen Absturz führt.
25. Ja, in dem Moment, in dem die Gravitationskraft auf den Satelliten so schwach ist, dass dieser durch andere Himmelskörper in seiner Bahn gestört wird bzw. aus seiner Bahn herausgeschleudert wird. Die Abschätzung ist am einfachsten für ein Kräftegleichgewicht zwischen Sonne und Erde (Lagrange-Punkt) oder Erde und Mond.
26. Die Störungen sind die Drehung der Bahnebene, die zur Erzeugung sonnensynchroner Bahnen verwendet wird, und die Drehung der Apsidenlinie, d.h. der Lage der Ellipse in der Bahn, die bei den Molniya-Satelliten zur Inklination von 63° führt.
27. Die Übergangsellipse und damit die Geschwindigkeitsänderungen sind bei beiden Übergängen betragsmäßig gleich – aber eben nur betragsmäßig. Beim Übergang in ein höheres Orbit erfolgt die Beschleunigung in Bewegungsrichtung, beim umgekehrten Übergang in entgegengesetzter Richtung. Die Gesamtenergie der Bahnen ist jeweils die Summe aus potentieller und kinetischer Energie, d.h. $E_{\text{ges}} = v^2 - 2\mu/r$. Aus der Binet'schen Gleichung gilt für die Kreisbahngeschwindigkeit $v = \sqrt{\mu/r}$ und damit für die Gesamtenergie $E_{\text{ges}} = -\mu/r$, d.h. die Gesamtenergie wird mit zunehmendem Abstand immer größer – beim Übergang auf eine niedrigere Bahn muss also Energie abgeführt werden, eben durch Abbremsen.
28. Die Reibung zwischen Flugkörper und Atmosphäre ist durch die Geschwindigkeit und die Dichte bestimmt. Beim Start wird zwar viel Energie zugeführt, jedoch primär zur Erhöhung der potentiellen Energie während die kinetische Energie und damit auch die Geschwindigkeit eher niedrig sind. Damit sind die Reibungskräfte und die Erwärmung gering. Beim Absturz (bzw. beim Eintreten in die Atmosphäre) dagegen hat der Flugkörper eine hohe Geschwindigkeit (ungefähr die Kreisbahngeschwindigkeit seines letzten Orbits). Diese steigt durch Umwandlung von potentieller in kinetische Energie während des Absinkens des Satelliten noch weiter an, d.h. mit zunehmender Annäherung an die dichtere Atmosphäre steigt die Geschwindigkeit und damit auch Reibung und Erwärmung.

B.1.2 Aufgaben

1. hier Perigäum und Apogäum direkt gegeben. Daher muss entweder in

$$v_{\text{Apo}} = \sqrt{\frac{\mu}{a} \frac{1-\varepsilon}{1+\varepsilon}} \quad \text{und} \quad v_{\text{Peri}} = \sqrt{\frac{\mu}{a} \frac{1+\varepsilon}{1-\varepsilon}}$$

ein Ausdruck für die Exzentrizität und die große Halbachse gefunden werden, oder zur Verwendung der Binet'schen Gleichung die große Halbachse bestimmt werden. Letzterer Ansatz ist einfacher, da

$$a = \frac{r_{\text{Apo}} + r_{\text{Peri}}}{2} = \frac{40\,400 \text{ km}}{2} = 20\,200 \text{ km} .$$

Einsetzen in die Binet'sche Gleichung liefert

$$v_{\text{Apo}} = \sqrt{\mu \left(\frac{2}{r_{\text{Apo}}} - \frac{2}{r_{\text{Apo}} + r_{\text{Peri}}} \right)} = \sqrt{\mu \frac{2r_{\text{Peri}}}{r_{\text{Apo}}(r_{\text{Apo}} + r_{\text{Peri}})}} = 14.08 \text{ km/s}$$

sowie

$$v_{\text{Peri}} = \sqrt{\mu \frac{2r_{\text{Apo}}}{r_{\text{Peri}}(r_{\text{Apo}} + r_{\text{Peri}})}} = 1407 \text{ km/s} .$$

2. Der Geschwindigkeitsvektor muss dazu um 90° gedreht werden. Als einfache Abschätzung erhalten wir

$$\Delta \vec{p} = m \Delta \vec{v} = m \begin{pmatrix} -v \\ v \end{pmatrix}$$

und damit für die Kraft

$$\vec{F} = \frac{\Delta p}{\Delta t} = \frac{m}{\Delta t} \begin{pmatrix} -v \\ v \end{pmatrix} .$$

Für die Kreisbahngeschwindigkeit gilt nach Binet'scher Gleichung $v_{\text{Kreis}} = \sqrt{\mu/r} = 7684 \text{ m/s}$ und damit für die Kraft $\vec{F} = (-768, 768) \text{ kN}$, bzw. $|\vec{F}| = 1086 \text{ kN}$.

3. Nein, da in die Konstante die Masse des Zentralgestirns eingeht. Die Konstante ist allerdings für alle Jupitermonde die gleiche. Zentralmasse aus Kräftegleichgewicht

$$\gamma \frac{M_{\text{Jupiter}} m_{\text{Io}}}{r^2} = m_{\text{Io}} \omega^2 r \quad \Rightarrow \quad M_{\text{Jupiter}} = \frac{\omega^2 r^3}{\gamma} = \frac{4\pi^2 r^3}{\gamma T^2} = 2.31 \cdot 10^{27} \text{ kg} .$$

4. Kräftebilanz $mv^2/r = \gamma m M/r^2$ liefert $v_{\text{Mond}} = \sqrt{\gamma M/r} = 5.15 \text{ km/s}$ und für die Erde $v_{\text{Erde}} = 7.86 \text{ km/s}$. Alternativ direkt unter Verwendung der Binet'schen Gleichung.
5. da Wiederkehr des Kometen beobachtet, ist die Bahn geschlossen, also fallen Parabel und Hyperbel flach, d.h. wir haben eine elliptische Bahn, von der die Umlaufzeit und das Perihel bekannt sind. Die Umlaufzeit hängt nach Keplers 3. Gesetz nur von der großen Halbachse ab, d.h. es ist egal, ob der Körper auf einer Kreisbahn oder einer Ellipse umläuft. Gleichsetzen von Gravitations- und Zentrifugalkraft wie in ?? liefert

$$\gamma \frac{M_{\odot} m_{\text{Hal}}}{r^2} = m_{\text{Hal}} \omega^2 r \quad \Rightarrow \quad r = \sqrt[3]{\frac{\gamma M_{\odot} T^2}{4\pi^2}} = 2.64 \cdot 10^{12} \text{ m} .$$

Dies entspricht auch der Länge der großen Halbachse. Da das Perihel gegeben ist, ist das Aphel wegen $2a = r_{\text{Peri}} + r_{\text{Aph}}$ direkt zu bestimmen: $a = 2.67 \cdot 10^{12} \text{ m}$ und $r_{\text{Aph}} = 5.25 \cdot 10^{12} \text{ m}$. Die Exzentrizität ergibt sich aus den großen und kleinen Halbachse gemäß Spickzettel zu $\varepsilon = 0.967$. Die Bahngeschwindigkeiten lassen sich mit Hilfe der Binet'schen Gleichung gemäß Spickzettel bestimmen zu

$$v_{\text{Apo}} = \sqrt{\frac{\mu}{a} \frac{1-\varepsilon}{1+\varepsilon}} \quad \text{und} \quad v_{\text{Peri}} = \sqrt{\frac{\mu}{a} \frac{1+\varepsilon}{1-\varepsilon}}$$

und damit

$$\frac{v_{\text{Peri}}}{v_{\text{Apo}}} = \frac{\sqrt{\frac{\mu}{a} \frac{1-\varepsilon}{1+\varepsilon}}}{\sqrt{\frac{\mu}{a} \frac{1+\varepsilon}{1-\varepsilon}}} = 60.16$$

und für die Winkelgeschwindigkeiten

$$\frac{\omega_{\text{Peri}}}{\omega_{\text{Apo}}} = \frac{\frac{v_{\text{Peri}}}{r_{\text{Peri}}}}{\frac{v_{\text{Apo}}}{r_{\text{Apo}}}} = 3619.$$

[Die Rechnung ist irgendwie nicht ganz konsistent!!!!]

6. Gleichgewicht der Gravitationskräfte fordert

$$\begin{aligned} \frac{\gamma M_{\odot} m_{\text{Sat}}}{r_{\text{Sat}}^2} &= \frac{\gamma M_{\text{Erde}} m_{\text{Sat}}}{(1 \text{ AU} - r_{\text{Sat}})^2} \\ \Rightarrow r_{\text{Sat}} &= -\frac{1 \text{ AU } M_{\odot}}{M_{\odot} - M_{\text{Erde}}} \pm \sqrt{\left(\frac{1 \text{ AU } M_{\odot}}{M_{\odot} - M_{\text{Erde}}}\right)^2 - M_{\odot} (1 \text{ AU})^2} \\ &= 1 \text{ AU} \left(-\frac{M_{\odot}}{M_{\odot} - M_{\text{Erde}}} \pm \sqrt{\left(\frac{M_{\odot}}{M_{\odot} - M_{\text{Erde}}}\right)^2 - M_{\odot}} \right) = -0.998 \text{ AU} \end{aligned}$$

oder auf den Abstand von der Erde bezogen $r_{\text{SE}} = 264 \text{ 310 km}$. Die andere Wurzel ergibt einen Punkt hinter der Erde, an dem beide Kräfte gleich groß sind. Die Rechnung führt auf ein Orbit, in dem gemäß Kepler 3 die Umlaufzeiten kleiner sind als im Erdorbit, d.h. die so berechnete Konfiguration kann nicht stabil sein, da der Satellit aus der Erde-Sonne Linie herauslaufen würde und damit das geforderte Kräftegleichgewicht nicht mehr erfüllt ist. Der Zusatzpunkt: Fehler liegt im Ansatz: in einem raumfesten System aus zwei Massen ist diese Herleitung ok und der Satellit würde auf der Verbindungslinie bleibe. Da wir jedoch ein rotierendes System betrachten, müsste die Zentrifugalkraft berücksichtigt werden, wodurch sich eine andere Herleitungsform ergibt.

7. Der Treibstoffvorrat kann in eine Impulsänderung umgerechnet werden (wobei die Abnahme der Shuttle-Masse bei Verbrennen des Treibstoffs nicht berücksichtigt wird). Für die Geschwindigkeiten in den drei Kreisbahnen gilt $v_{\text{K278}} = \sqrt{\mu/r_1} = 7.75 \text{ km/s}$, $v_{\text{K120}} = \sqrt{\mu/r_2} = 7.85 \text{ km/s}$ und $v_{\text{K385}} = 7.69 \text{ km/s}$. Für die Geschwindigkeitsänderungen ergibt sich beim Abstieg

$$\Delta v_1 = v_{\text{K278}} \sqrt{\frac{2}{r_1/r_2 + 1}} = -97 \text{ m/s} \quad \text{und} \quad \Delta v_2 = v_{\text{K120}} \sqrt{\frac{2}{1 + r_2/r_1}} = -98 \text{ m/s},$$

die Gesamtänderung im Impuls/Masse ist -196 m/s . Zuzüglich der Sicherheitsreserve gibt das 235 m/s . Das Problem des Aufstiegs zerlegt sich in zwei Teile: zum einen muss die höhere Flugbahn erreicht werden, zum anderen die Inklination verändert werden. Für das reine Aufstiegsmanöver ergibt sich $\Delta v_1 = 31 \text{ m/s}$ und $\Delta v_2 = 31 \text{ m/s}$, also insgesamt ein Impuls pro Masse von 62 m/s . Das Erreichen der Höhe (allerdings mit dem Verlust der Möglichkeit, dass Shuttle wieder in ein niedrigeres Orbit zu bringen) wäre kein Problem. Der Aufwand zur Veränderung der Inklination der Bahn, d.h. zur Drehung des Geschwindigkeitsvektors um 20.6° dagegen wäre mit den begrenzten Ressourcen nicht zu leisten gewesen:

$$|\Delta \vec{v}| = \left| \begin{pmatrix} 6.0 \\ 4.9 \end{pmatrix} - \begin{pmatrix} 3.9 \\ 6.7 \end{pmatrix} \right| = \left| \begin{pmatrix} 2.1 \\ -1.8 \end{pmatrix} \right| = 2.8 \text{ km/s}.$$

Zur Drehung der Bahnebene hätte ungefähr die zehnfache Menge an Treibstoff zur Verfügung stehen müssen.

8. Binet'sche Gleichung liefert $v = \sqrt{\gamma m/r} = 7.45$ km/s für die Geschwindigkeit entlang der Bahn und $v_{\text{Boden}} = 6.59$ km/s für die Geschwindigkeit entlang der Bodenspur, da die Winkelgeschwindigkeit konstant sein muss.
9. Die Grundidee einer Transferbahn ist eine elliptische Übergangsbahn. Dazu wird an einem Punkt der Ausgangsbahn ein Schub gegeben, der den Satelliten in die Transferbahn injiziert. Tangiert die Transferbahn (in ihrem Apogäum) die Zielbahn, so ist die Geschwindigkeit des Satelliten für eine Kreisbahn zu gering und es muss erneut ein Schub gegeben werden, damit der Satellit nicht wieder auf der Transferbahn in Richtung auf seine Ausgangsbahn fliegt. Die relevanten Geschwindigkeiten sind jeweils die Kreisbahngeschwindigkeiten auf der Start- und Zielbahn, bestimmt aus der Binet'schen Gleichung, und die große Halbachse a der Übergangselipse:

$$v_{K1} = \sqrt{\frac{\mu}{r_1}} = 7.68 \text{ km/s}, \quad v_{K2} = \sqrt{\frac{\mu}{r_2}} = 97.12 \text{ km/s}$$

und

$$a = \frac{r_1 + r_2}{2} = 24578 \text{ km}.$$

Unter Verwendung des letzteren Wertes lassen sich die Geschwindigkeiten im Perigäum und Apogäum des Transferorbits bestimmen:

$$v_P = v_{K1} \sqrt{\frac{2}{r_1/r_2 + 1}} \quad \text{und} \quad v_A = v_{K2} \sqrt{\frac{2}{1 + r_2/r_1}}.$$

Die Geschwindigkeitsänderungen ergeben sich zu

$$\Delta v_1 = v_{K1} \left(\sqrt{\frac{2}{r_1/r_2 + 1}} - 1 \right) = 2.40 \text{ km/s}$$

und

$$\Delta v_2 = v_{K2} \left(1 - \sqrt{\frac{2}{1 + r_2/r_1}} \right) = 46.2 \text{ km/s}.$$

Für die aufzuwendenden Kräfte gilt daher

$$F_1 = m\Delta v_1/\Delta t = 24 \text{ kN} \quad \text{und} \quad F_2 = m\Delta v_2/\Delta t = 462 \text{ kN}.$$

10. Bahn als Übergangselipse, es muss zweimal Schub gegeben werden: auf der inneren Kreisbahn zum Einleiten des Übergangs auf die Ellipse, auf der äußeren zum Übergang von der Ellipse zur Kreisbahn. Die Ausgangsgeschwindigkeiten sind jeweils die Kreisbahngeschwindigkeiten (aus der Binet'schen Gleichung) auf den beiden Orbits:

$$v_{K1} = \sqrt{\frac{\mu}{r_1}} = 7.80 \text{ km/s}, \quad v_{K2} = \sqrt{\frac{\mu}{r_2}} = 7.44 \text{ km/s}.$$

Die Übergangselipse hat eine große Halbachse $a = (r_1 + r_2)/2 = 6903$ km, die dafür erforderlichen Geschwindigkeiten in Perigäum und Apogäum ergeben sich zu

$$v_P = v_{K1} \sqrt{\frac{2}{r_1/r_2 + 1}} \quad \text{und} \quad v_A = v_{K2} \sqrt{\frac{2}{1 + r_2/r_1}}.$$

Die Geschwindigkeitsänderungen ergeben sich zu

$$\Delta v_1 = v_{K1} \left(\sqrt{\frac{2}{r_1/r_2 + 1}} - 1 \right) = 0.18 \text{ km/s}$$

und

$$\Delta v_2 = v_{K2} \left(1 - \sqrt{\frac{2}{1 + r_2/r_1}} \right) = 0.18 \text{ km/s} .$$

Die aufzuwendende Energie ergibt sich damit zu

$$E_{\text{add}} = E_{\text{add1}} + E_{\text{add2}} = \frac{m}{2} (v_{\text{P}}^2 - v_{\text{K1}}^2 + v_{\text{K2}}^2 - v_{\text{A}}^2) = 1.1 \cdot 10^{10} \text{ J} .$$

11. Bahn als Übergangsellipse, es muss zweimal Schub gegeben werden: auf der inneren Kreisbahn zum Einleiten des Übergangs auf die Ellipse, auf der äußeren zum Übergang von der Ellipse zur Kreisbahn. Die Ausgangsgeschwindigkeiten sind jeweils die Kreisbahngeschwindigkeiten (aus der Binet'schen Gleichung) auf den beiden Orbits:

$$v_{K1} = \sqrt{\frac{\mu}{r_1}} = 29.16 \text{ km/s} , \quad v_{K2} = \sqrt{\frac{\mu}{r_2}} = 46.9[12.8] \text{ km/s} .$$

Die Übergangsellipse hat eine große Halbachse $a = (r_1 + r_2)/2 = 0.96[3.1]$ AU, die dafür erforderlichen Geschwindigkeiten in Perigäum und Apogäum ergeben sich zu

$$v_{\text{P}} = v_{K1} \sqrt{\frac{2}{r_1/r_2 + 1}} = 59.56[37.8] \text{ km/s} \quad \text{und} \quad v_{\text{A}} = v_{K2} \sqrt{\frac{2}{1 + r_2/r_1}} = 21.8[7.3] \text{ km/s} .$$

Die Geschwindigkeitsänderungen ergeben sich zu

$$\Delta v_1 = v_{K1} \left(\sqrt{\frac{2}{r_1/r_2 + 1}} - 1 \right) = 12.7[8.6] \text{ km/s}$$

und

$$\Delta v_2 = v_{K2} \left(1 - \sqrt{\frac{2}{1 + r_2/r_1}} \right) = 7.4[5.5] \text{ km/s} .$$

Die aufzuwendende Energie ergibt sich damit zu

$$E_{\text{add}} = E_{\text{add1}} + E_{\text{add2}} = \frac{m}{2} (v_{\text{P}}^2 - v_{\text{K1}}^2 + v_{\text{K2}}^2 - v_{\text{A}}^2) = 2.2 \cdot 10^{12}[3.45 \cdot 10^{11}] \text{ J} .$$

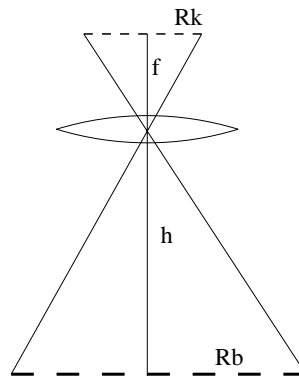
B.2 Kapitel 3

B.2.1 Fragen

1. nein, nein, nein, ja, nein, nein, ja, ja, nein
2. Farbaufnahmen enthalten mehr Information als Schwarz-Weiß- Aufnahmen, allerdings ist das Auflösungsvermögen eines Farbfilms aufgrund der drei übereinander liegenden Filmschichten schlechter als das eines Schwarz-Weiß-Films. Bei der Multispektralkamera verwendet man daher ein System aus mehreren synchronisierten Kameras, die jede mit einem Schwarz-Weiß-Film bestückt sind und durch ein Filter unterschiedliche Bereiche des sichtbaren (und nahes IR) Spektrum betrachten. Der Vorteil ist die Kombination aus hoher Bodenauflösung und Informationen aus verschiedenen Spektralbereichen, der Nachteil sind die technischen Probleme der Synchronisation und Gleichheit der Kameras und der höhere Filmbedarf, daher eher für bemannte Missionen als für Satelliten geeignet.
3. Beides sind Filme aus drei lichtempfindlichen Schichten. Beim konventionellen Farbfilm sind diese empfindlich für Blau, Grün und Rot (zwischen der Blau- und der Grünschicht befindet sich ein Gelbfilter), beim Farb-IR-Film für IR, Grün und Rot, d.h. die blauempfindliche Schicht entfällt. Oberhalb der IR-Schicht sitzt ein Gelbfilter, so dass keine Strahlung aus dem blauen Bereich des Spektrums auf eine der empfindlichen Schichten

fällt. Der blaue Bereich wird geopfert, da er am stärksten durch atmosphärische Streuung (Dunst) beeinträchtigt ist und damit die Informationen am wenigsten zuverlässig sind; der Gelbfilter wird zum Ausblocken benötigt, da Blau gleichzeitig der energiereichste Teil der Strahlung ist. Die Darstellung beim konventionellen Farbfilm erfolgt derart, dass die einzelnen Schichten wieder die ihrer Farbe entsprechenden Informationen liefern. Beim IR-Farbfilm macht diese Zuordnung keinen Sinn, da das IR nicht sichtbar ist, d.h. der Film würde zwar die Information enthalten, der Betrachter sie aber nicht erkennen. Stattdessen werden die drei Schichten wie bei einem konventionellen Farbfilm ausgewertet: die grüne Schicht wird als Blau abgebildet, die rote als Grün und die IR-Schicht als Rot. Daher erscheint Vegetation, die besonders stark im nahen IR reflektiert, auf diesen Falschfarbenaufnahmen rot.

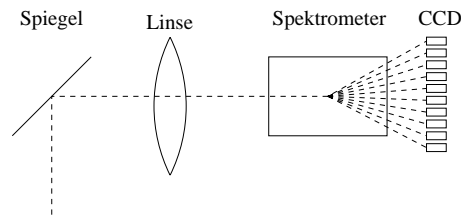
4. Die Identifikation verschiedener beobachteter Objekte basiert auf deren unterschiedlichen spektralen Reflektionsvermögen. Dazu gehört insbesondere das hohe Reflektionsvermögen von Vegetation im nahen IR. Daher wurde anfangs die Verwendung von 3 verschiedenen Spektralbereichen (Farbfilm bzw. IR-Farbfilm, Multispektralkamera) angestrebt, später das Spektrum sogar entsprechend der wissenschaftlichen Fragestellung in deutlich feinere Bereiche zerlegt.
5. Das Bodenaufhebungsvermögen bezeichnet die Fähigkeit eines photographischen Systems, die Streifen eines auf dem Boden liegenden Musters noch getrennt darzustellen. Das Bodenaufhebungsvermögen gibt man in Linien/Länge an. Es bestimmt sich aus dem Abbildungsmaßstab $B/b = G/g$ wobei man für die Gegenstandsweite g die Flughöhe h verwendet kann und für die Bildweite B aufgrund des großen Abstandes zwischen Gegenstand und Objektiv die Brennweite f . Dann gilt für das Bodenaufhebungsvermögen $R_B = h/(fR_K)$ mit R_K als dem Auflösungsvermögen des Systems Kamera-Film.



6. auf die blaue Schicht wird verzichtet, da im blauen Bereich des Spektrums die Informationen durch Streuung am stärksten beeinträchtigt sind.
7. ja, nein, ja, ja, ja, ja, ja, nein, ja
8. Die Höhe geht (Abbildungsgesetz) linear in beide Größen ein: das Auflösungsvermögen verbessert sich um einen Faktor 5, das Blickfeld verkleinert sich um diesen Faktor.
9. Zur Identifikation der Vegetation ist das nahe IR entscheidend, da dort das spektrale Reflektionsvermögen von Pflanzen sehr groß ist. Dann muss auch mindestens ein Kanal im sichtbaren Bereich liegen, da sonst das hohe IR-Signal nicht eindeutig identifiziert werden kann. Hier ist die Wahl Rot sinnvoller als Blau, da der blaue Bereich zu stark durch Streuung beeinträchtigt wird. Ein weiterer Kanal sollte im grünen Bereich liegen, da dort das spektrale Reflektionsvermögen der Vegetation ebenfalls sehr groß ist. Das Verhältnis aus Grün und IR erlaubt eine genauere Klassifikation der verschiedenen Pflanzen. Steht der vierte Kanal zur Verfügung, so wäre eine Unterteilung des nahen IR in zwei Kanäle sinnvoll.
10. Die Vorwärtsbewegung (insbesondere bei hochauflösendem optischen System und niedrig fliegendem Satellit), Streuung in der Atmosphäre (insbesondere im blauen Bereich des Spektrums, daher diesen bei Multispektralaufnahmen zur Dunstkorrektur verwenden

oder beim IR-Farbfilm weglassen), zufällige Dichteschwankungen der Atmosphäre (Szintillation oder Pulsation) lässt sich bei astrophysikalischen Aufnahmen leichter korrigieren.

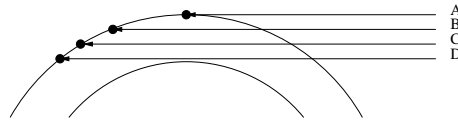
11. (1) Bodenaufklärungsvermögen, d.h. die Fähigkeit, ein am Boden befindliches Streifenmuster als getrennte Linien aufzulösen. Wird bestimmt durch die Flughöhe, die Brennweite, das Auflösungsvermögen des aufzeichnenden Mediums und die Blendenöffnung des Objektivs (Beugungsbildchen)! Das Bodenaufklärungsvermögen ist zur Identifikation kleiner Bildelemente und zur Verringerung der Zahl der Mixed Pixels zu optimieren. (2) Spektrales Auflösungsvermögen, d.h. die Fähigkeit eines optischen Systems, verschiedene Spektralbereiche zu unterscheiden bzw. einen Verlauf des spektralen Reflektanzvermögens der beobachteten Objekte zu erzeugen. Optimierung zur Identifikation erforderlich, allerdings nur dann sinnvoll, wenn der relative Anteil des Mixed Pixels nicht zu groß wird. (c) Radiometrisches Auflösungsvermögen, d.h. die Fähigkeit, Helligkeitswerte zu unterscheiden. Zwar können dicht beieinander liegende Helligkeitswerte durch Dehnungsverfahren getrennt werden (Kontrastanreicherung), das Instrument muss aber die unterschiedlichen Helligkeitswerte als solche erkannt haben. Ein schlechtes radiometrisches Auflösungsvermögen verschlechtert automatisch das Bodenaufklärungsvermögen. (4) Das zeitliche Auflösungsvermögen (Wiederholfrequenz) gibt die Zeit, die zwischen zwei aufeinander folgenden Überflügen eines Zielgebiets vergeht. Die Wahl des zeitlichen Auflösungsvermögens ist durch die Fragestellung mitbestimmt: Kartographie oder Kontinentalverschiebung haben es nicht so eilig, ein militärischer Späher wartet ungerne einige Jahre, bis er ein Krisengebiet nochmals überfliegen kann. (5) Bei thermischen Infrarotsensoren interessiert das thermale Auflösungsvermögen, d.h. die Temperaturunterschiede, die das Instrument noch wahrnehmen kann.
12. ein hohes Bodenaufklärungsvermögen erlaubt im Prinzip die Trennung eng benachbarter Punkte oder Objekte – aber nur, wenn diese auf dem Film auch ein unterschiedliches Signal in Form verschiedener Helligkeitswerte und/oder verschiedener Farben (d.h. spektraler Signaturen) erzeugen. Das gute spektrale Auflösungsvermögen erlaubt die Identifikation oder zumindest Klassifikation verschiedener Objekte (Beispiel Pflanzen im nahen IR), ist jedoch erst dann wirklich sinnvoll, wenn außerdem ein gutes Bodenaufklärungsvermögen zur Verfügung steht, da sonst zu viele Mixed Pixels entstehen.
13. ja, nein, ja, ja, ja, ja, ja
14. Die Kehrbesentechnik ist eine Entwicklung zwischen der Abbildung auf einem konventionellen 2D-Film und der auf einen Punkt gerichteten Abbildung mit Hilfe eines Sammers. Das Aufnahmemedium besteht aus einer Zeile von CCDs (eindimensionaler Film), die durch die Bewegung des Satelliten über die aufzunehmende Szene geschoben wird. Die Technik wird außer auf der MOMS-Doppeloptik-Kamera für das Shuttle kaum verwendet: die Zahl der Bildpunkte der CCD ist mit ca. 2000 für ein gutes Bodenaufklärungsvermögen deutlich zu gering, daher wird in der MOMS eine Gruppe von vier Zeilen kombiniert. Die dazu erforderliche Doppeloptik macht das Instrument jedoch unhandlich. Historisch hat die Kehrbesentechnik jedoch den Vorteil, dass sie die notorisch unzuverlässige Videokamera als erstes nicht film-basiertes System ablösen konnte und gegenüber der Videokamera ein besseres photometrisches Auflösungsvermögen liefert.
15. Das abbildende System besteht im wesentlichen aus dem senkrecht zur Flugrichtung scannenden Spiegel. Dessen Signal wird durch ein Linsensystem auf ein Spektrometer abgebildet, durch das einzelne Spektralbereiche auf verschiedene CCDs abgebildet werden. Auf diese Weise entsteht das spektrale Auflösungsvermögen. Die technischen Realisierungen variieren im Hinblick auf die Blickrichtung des Spiegels (SPOT kann z.B. zur Seite gucken), mehr als einen Streifen auf ein kurzes CCD-Array abbilden (MSS auf LandSat), statt des Spektrometers verschiedene Filter verwenden bzw. kein Spektrometer verwenden, da im panchromatischen Modus (SPOT).



16. Bei den erwähnten Strukturen handelt es sich um Gegenstände, die in einer Dimension zwar deutlich geringere Abmessungen haben als dem Bodenaufklärungsvermögen entspricht, in der anderen Dimension jedoch weit größer und relativ linear sind. Dadurch wird nicht ein einzelner Pixel in seinem Helligkeitswert und spektralen Reflektionsvermögen beeinflusst, sondern es wird eine Serie von aneinander grenzenden Pixeln modifiziert – und diese Veränderung ist in den Aufnahmen zu erkennen.

34	24	36	30	28	27	30	25	33	51	26	33	32	23	29	57	51	25	30	34
26	33	25	33	35	24	33	23	26	29	54	31	33	49	52	28	25	27	25	36
31	23	27	26	23	35	35	28	33	24	36	53	57	26	34	36	23	24	28	35
33	31	30	28	26	26	30	35	27	54	51	27	47	31	26	32	34	36	33	25
29	27	34	24	35	32	29	51	47	24	26	34	29	57	35	26	32	31	23	36
31	24	29	32	25	52	56	35	23	25	34	24	24	29	55	29	25	28	24	29
36	35	32	53	47	24	33	36	25	32	24	36	32	28	31	49	23	33	31	30
25	55	49	29	32	25	31	29	27	28	32	34	33	28	32	30	52	24	29	34
52	28	24	36	28	35	36	31	26	31	36	24	28	30	29	28	34	47	36	33
33	23	34	28	33	28	24	33	32	28	26	34	27	33	27	32	30	36	53	26

17. In einem Pixel sind Informationen von verschiedenen Gegenständen und damit auch unterschiedlichen spektralen Reflektionsvermögen überlagert. Damit ist eine eindeutige Identifikation der beobachteten Objekte nicht mehr möglich bzw. drastisch erschwert. Das Auftreten von Mixed Pixels wird um so geringer, je homogener das Beobachtungsgebiet im Verhältnis zum Bodenaufklärungsvermögen ist, d.h. bei geringem Bodenaufklärungsvermögen in weiträumig homogenen Bereichen oder bei sehr gutem Bodenaufklärungsvermögen. Eine anschließende Korrektur wäre nur für den Fall möglich, dass sehr viele Spektralkanäle vorliegen, in denen zumindest für typische Mixed Pixels Clusterdiagramme erzeugt werden können.
18. Es stehen im wesentlichen drei Messverfahren zur Verfügung: Emission, Streuung und Absorption. Bei letzterem Verfahren wird eine definierte Quelle benötigt, es wird in der Regel als Limb Sounding angewendet. Die anderen Verfahren können auch unter anderen Blickwinkeln betrieben werden (insbesondere Aufsicht). Bei Absorption und Emission handelt es sich um spektroskopische Verfahren, d.h. ein entsprechend gutes spektrales Auflösungsvermögen ist gefordert.
19. Limb-Sounding ist ein Verfahren, bei dem die Veränderung der Absorption von Sonnenlicht während des (scheinbaren) Sonnenauf- bzw. -untergangs aus der Sicht des Satelliten betrachtet wird. Das Verfahren hat den Vorteil, dass das Instrument gegen die Lichtquelle kalibriert wurde, da der Satellit die Sonne am höchsten Punkt seines Orbits (A in der folgenden Abbildung) praktisch ohne störende Atmosphäre sieht. Der Satellit ‘sinkt’ dann unter die bisherige Sichtlinie und es werden Absorptionskoeffizienten in immer dichteren Schichten der Atmosphäre bestimmt (Linien B, C, ...). Die erste Linie liefert den Absorptionskoeffizienten für die oberste Atmosphärenschicht. Mit ihrer Hilfe kann der Absorptionskoeffizient der nächsten Schicht korrigiert werden usw. Als Resultat ergibt sich ein Höhenprofil der Absorptionskoeffizienten. Diese werden in verschiedenen Wellenlängenbereichen bestimmt: insbesondere in denen, in denen die gesuchte Substanz absorbiert, aber auch bei Kontrollwellenlängen, mit deren Hilfe sich die allgemeine Abschwächung durch andere Effekte (Dunst, Vulkanausbruch usw.) abschätzen lässt. Für die einzelnen Frequenzen gilt das Bouguer-Lambert-Beer’sche Gesetz: $I_\nu(x) = I_\nu(0) e^{-\mu x}$, die Spezies werden durch geeignete Kombination verschiedener Frequenzen identifiziert.



20. Das Messprinzip beruht auf der Absorption, beschrieben durch das Bouguer-Lambert-Beer'sche Gesetz: $I_\nu(x) = I_\nu(0) e^{-\mu x}$. Das Absorptionsgesetz wird für jedes Frequenzintervall getrennt betrachtet. Für den Ozonnachweis sind zwei Wellenlängenbereiche im UV (Absorption durch Ozon) sowie mindestens ein weiterer Wellenlängenbereich (keine Absorption durch Ozon) zur Korrektur atmosphärischer Einflüsse erforderlich.
21. nein, nein, nein, ja, ja, nein
22. Die für die Erdfernerkundung interessanten Aerosole sind relativ große Moleküle, die bevorzugt Licht im Bereich um $1 \mu\text{m}$ streuen. Gleichzeitig ist in diesem Bereich die Absorption durch atmosphärische Bestandteile praktisch zu vernachlässigen, so dass keine Störgrößen gemessen werden.
23. Sensoren im thermischen Infrarot haben ein Bodenaufklärungsvermögen von einigen hundert Metern (z.B. TIR auf HCMM hat 600 m). Eine Annäherung an das Bodenaufklärungsvermögen im sichtbaren ist nicht weiter möglich, da das prinzipielle Instrumentendesign sehr ähnlich ist, die Wellenlängen im thermischen Infrarot jedoch gut eine Größenordnung über denen im sichtbaren Bereich liegt. Damit begrenzt die Beugung das Auflösungsvermögen auf entsprechend schlechte Werte.
24. Aufbau ähnlich dem eines optischen Scanners, d.h. Spiegel zum Abscannen senkrecht zur Flugbahn, fokussierendes System zur Abbildung auf den Detektor und Detektor, letzterer gekühlt (das ist auch schon der entscheidende Unterschied zum optischen Scanner). Spektrale Zerlegung erfolgt bei Instrumenten für originär thermische Anwendungen nicht bzw. nur in die den beiden Fenstern der Atmosphäre entsprechenden Anteile (muss bei Instrumenten, die Infrarot-Linien zur Bestimmung von Spurengasen in der Atmosphäre vermessen natürlich erfolgen). Messziel ist der integrierte Strahlungsstrom im IR und damit ein Maß für die Temperatur.
25. Gemessen wird der Gesamtstrahlungsstrom q innerhalb der entsprechenden beiden Fenster der Atmosphäre im IR. Der Gesamtstrahlungsstrom wird nach Stefan-Boltzmann als ein Maß für die effektive Temperatur interpretiert mit $q = \sigma T_{\text{eff}}^4$. Die Effektivtemperatur gibt die Temperatur, die ein schwarzer Körper haben müsste, um den gleichen Gesamtstrahlungsstrom zu emittieren wie beobachtet.
26. Die Effektivtemperatur ist die Temperatur, die einem Körper aufgrund des von ihm emittierten Gesamtstrahlungsstroms zugeschrieben wird. Es ist die Temperatur, die ein schwarzer Strahler haben müsste, um den gleichen Gesamtstrahlungsstrom zu emittieren. Letzterer wird beschrieben nach dem Stefan-Boltzmann-Gesetz $q = \sigma T_{\text{eff}}^4$. Ein grauer Strahler mit einem mittleren Emissionsvermögen ε emittiert einen Gesamtstrahlungsstrom $q = \sigma \varepsilon T^4$, d.h. das Verfahren, die Temperatur aus dem Gesamtstrahlungsstrom zu bestimmen unterschätzt die reale Temperatur des Körpers, kann sie aber nie überschätzen.
27. Passive Mikrowellendetektoren nutzen das langwellige Ende der terrestrischen Ausstrahlung, d.h. sie können, wie Infrarotsensoren, auch während der Nacht eingesetzt werden. Allerdings ist die Wellenlänge noch größer als im IR, so dass sich ein Bodenaufklärungsvermögen im Bereich von Kilometern ergibt. Der Vorteil gegenüber allen anderen Instrumenten: Mikrowellen durchdringen auch Wolken, d.h. es ist das einzige Instrument, das auch bei Bewölkung arbeiten kann. Regen schränkt den Anwendungsbereich eines Mikrowellendetektors ein, da die Wassertröpfchen zu einer Streuung des Signals führen. Beispiele sind insbesondere die Beobachtungen von Meer- und Festlandeis. Da die Polarisation des Signals Informationen über den Flüssigwassergehalt am Sender liefert, können passive Mikrowelleninstrumente zur Bestimmung der Bodenfeuchte (auch wichtig für Waldbrandgefährdung) verwendet werden ebenso wie zur Bestimmung des Abtauens von Eisflächen.
28. Kurzbezeichnung für das aktive Mikrowelleninstrument ist Radar, das Verfahren beruht

auf der Laufzeitmessung eines an der Meeresoberfläche reflektierten Signals. Aus der Laufzeit bekommt man nur die Höhe des Satelliten über der Meeresoberfläche, aus der Verformung des Signals (insbesondere seiner ansteigenden Flanke) ein Maß für die Rauigkeit des Reflektors, also für die mittlere Wellenhöhe.

29. Die Effektivtemperatur unterschätzt die reale Temperatur um einen Faktor $\sqrt[4]{\varepsilon}$ da die Emissivität des Körpers überschätzt wurde (er wird ja als schwarzer Körper betrachtet).

B.2.2 Aufgaben

- zwei Abbildungsgleichungen, $\frac{1}{f} = \frac{1}{b} + \frac{1}{g}$ liefert $f \approx b$. Einsetzen in Abbildungsmaßstab liefert $\frac{B}{G} = \frac{b}{g} = \frac{f}{h}$ und damit für das Blickfeld $F_B = \frac{hk}{f} = 53.3$ km mit k als der Kantenlänge des Films. Das Bodenaufklärungsvermögen ergibt sich mit gleicher Überlegung zu $R_B = \frac{h}{fR_K} = 3$ Linien/m mit R_K als dem Auflösungsvermögen des Systems Kamera-Film.
- Das Winkelaufklärungsvermögen (vergl. Spickzettel) liefert $a = \lambda f/D = 10^{-4}$ cm oder 1/1000 mm. Damit sind die Beugungsscheibchen halb so groß wie das Auflösungsvermögen des Films, die Konfiguration ist gerade noch sinnvoll – allerdings muss man dabei bedenken, dass die etwas langwelligere Strahlung aus dem roten oder nahen IR bereits Beugungsbildchen in der Größe des Filmaufklärungsvermögens erzeugt. Eine etwas größere Blendenöffnung (oder eine kleinere Brennweite) würden den Fehler aufgrund der Beugung auf eine dem Film angemessenere Größe reduzieren.
- Winkelaufklärungsvermögen laut Spickzettel, aufgelöst nach Blendenöffnung ergibt sich mit $a = 0.2 \mu\text{m}$ (Filmaufklärungsvermögen) $D = \lambda f/a = 100$ mm. Da ein Beugungsscheibchen allerdings vom nächsten durch einen Streifenabstand getrennt sein sollte, muss D größer als der so bestimmte Wert sein – mit 200 mm läge man auf der sicheren Seite.
- (a) Bodenaufklärungsvermögen $R_B = h/(fR_K)$ mit R_K als dem Auflösungsvermögen des Systems Kamera-Film. In diesem Fall ist R_K für die CCD 2500 Pixel pro 36 mm oder 69.44 Pixel/mm und damit $R_B = 2.16$ m aus dem LandSat-Orbit bzw. 0.43 m (also ungefähr 10 cm) für ein militärisches Orbit. Für den konventionellen Film ergibt sich aufgrund der höheren Streifenzahl/mm ein deutlich besseres Bodenaufklärungsvermögen von 0.3 m bereits aus dem LandSat-Orbit und 6 cm aus dem militärischen Orbit. (b) Winkelaufklärungsvermögen ergibt sich nach Spickzettel zu $a = \lambda f/D$ bzw. für die Blendenöffnung D muss gelten $D = \lambda f/a$. Für den Film ergibt sich mit $a = 2 \mu\text{m}$ eine Blendenöffnung $D = 1.5$ m, mit der CCD eine von ca. 10 cm. (c) Bodenbereich ist unabhängig vom Film bzw. von der CCD, d.h. nur ein Wert $F_N = hk/f$ mit $k = 36$ mm als der Kantenlänge des Films liefert $F_N = 5.4$ km bzw. für die kurze Seite 3.6 km aus dem LandSat-Orbit und nur 1/5 davon (0.72 km bzw. 1.08 km) aus dem militärischen Orbit. Damit sieht man also alles sehr genau, aber das Instrument guckt nie dahin, wo es vielleicht interessant ist. (d) Aufgrund des nur geringen Blickfeldes von 5.4 km werden 7421 Überflüge benötigt. Bei sonnensynchroner Bahn (keine Datenverluste durch Überflüge auf der Nachtseite) ließe sich die Zahl halbieren (3711 Überflüge), dafür werden bei 90 min pro Orbit allerdings 231 Tage benötigt – nicht wirklich prickelnd. (e) Der Satellit bewegt sich mit nahezu 8 km/s entlang seiner Bahn, d.h. während der 1/500 s Aufnahmezeit legt er eine Strecke von 16 m zurück (entlang der Bodenspur etwas weniger). Damit ist das instrumentelle Bodenaufklärungsvermögen wesentlich besser als die Verzerrung aufgrund der Vorwärtsbewegung. (f) Kompensationsmethoden wären (deutlich) kürzere Belichtungszeiten oder ein Nachführen der Kamera (was aber aufgrund der großen Maße technisch nicht einfach ist – gibt dem Satelliten entweder eine Nickbewegung (bei Gegensteuern) oder bringt ihn ins Rotieren).
- Bewegung entlang der Bahn bzw. der Bodenspur knapp 8 km/s (als Rechenwert verwenden), d.h. für einen 30 m breiten Streifen steht eine Zeit von $t = s/v = 0.00375$ s zur Verfügung. In dieser Zeit muss der Spiegel auf dem anderen Ende des Streifens eine neue reflektierende Fläche zur Verfügung gestellt haben. Da dies in diesem Fall die Spiegelrückseite sein soll, hat der Spiegel insgesamt 0.0075 s Zeit zur Rotation bzw. es wird

eine Rotationsfrequenz von 133.3 Hz benötigt. Während der 0.00375 s Scanzeit müssen 185 km/30 m = 6166 Pixel belichtet werden, d.h. pro Pixel steht eine Zeit von $6 \cdot 10^{-7}$ s zur Verfügung. Die Gesamtzahl der Pixel ergibt sich zu $3.49 \cdot 10^7$.

Die zweite Dimension der Aufnahme ergibt sich aus der Vorwärtsbewegung, ist also nicht von der Rotation des Spiegels abhängig und damit vom ersten Teil der Aufgabe. Hier gelten wieder die 8 km/s, d.h. der Satellit benötigt ca. 21.25 s für das Abscannen des Blickfeldes in Vorwärtsrichtung.

6. Bahngeschwindigkeit aus Binet'scher Gleichung (Spickzettel) $v = \sqrt{\mu/r} = 7.45$ km/s, herunterskaliert auf Bodenspur (Winkelgeschwindigkeit muss konstant sein) $v_{\text{Boden}} = 6.59$ km/s. Problem: Verzerrung der Aufnahmen durch die Bewegung des Satelliten insbesondere bei sehr hohem Bodenauflösungsvermögen des optischen Systems; Kompensationsmöglichkeiten durch Schwenk der Kamera (setzt aber kleines und leichtes System voraus, da sich der Satellit sonst ebenfalls dreht). Sehr kurze Belichtungszeiten, wie sie sich beim Scanner eher realisieren lassen als bei einer Filmkamera sind ebenfalls von Vorteil.
7. Inklination in der Nähe von 90° , da sonst die polare Bereich nicht abgescannt werden. Die abgedeckte Bildbreite ergibt sich zu $b = 2h \arctan(\alpha/2) = 8.8 \cdot 10^{-3} \cdot h$, d.h. für eine Flughöhe von 900 km ergibt sich $b_{900} = 7.9$ km, für ein geostationäres Orbit $b_{36000} = 314$ km. Die Zahl der benötigten Umläufe ist $Z = 40\,000/b = 4.6 \cdot 10^6/h$. Nur bei hohen Orbits ergibt sich überhaupt eine sinnvolle Wiederholfrequenz, allerdings um den Preis eines relativ schlechten Auflösungsvermögens. Bei niedrigen Orbits verliert sich der Satellit in Details.
8. Die für die Aufnahme zur Verfügung stehende Zeit ergibt sich aus der Geschwindigkeit des Satelliten entlang der Bodenspur. Aus der Binet'schen Gleichung erhalten wir für die Bahngeschwindigkeit $v = \sqrt{\mu/r} = 7.41$ km und die Geschwindigkeit entlang der Bodenspur $v_{\text{Boden}} = 6.5$ km/s. Die für das Zurücklegen von 476 m benötigte Zeit beträgt $t = s/v = 7.3 \cdot 10^{-2}$ s, die Rotationsfrequenz eines doppelten Spiegels muss entsprechend 27 Hz betragen. Die Bildzeile zerfällt in 2 342 Elemente, d.h. pro Element stehen 32 μs zur Verfügung. Würde die Aufteilung nicht erfolgen, so würde sich das Bodenauflösungsvermögen um den Faktor 6 verschlechtern, zur Aufnahme stünde die sechsfache Zeit zur Verfügung und der Spiegel müsste genauso schnell bzw. langsam rotieren wie vorher.
9. Der Gesamtstrahlungsstrom eines grauen Strahlers ergibt sich nach Stefan-Boltzmann zu $q = \sigma \varepsilon T^4 = 484$ W/m². Verwendet man die Idee der Effektivtemperatur, so ist $q = \sigma T_{\text{eff}}^4$ und damit $T_{\text{eff}} = 304$ K = 31°C .
10. Das Maximum ergibt sich aus dem Wien'schen Verschiebungsgesetz (Spickzettel) zu $\lambda_{\text{max}} = 11.3$ μm . Der eingeschränkte Messbereich ergibt sich durch die atmosphärischen Fenster.
11. Ölfilm und Wasser haben zwar die gleiche Temperatur, emittieren aufgrund des unterschiedlichen spektralen Reflektionsvermögens ε jedoch einen unterschiedlichen Gesamtstrahlungsstrom $q = \sigma \varepsilon T^4$. Letzterer wird gemäss $q = \sigma T_{\text{eff}}^4$ in eine Effektivtemperatur umgewandelt. Gleichsetzen der beiden Gleichungen liefert $T_{\text{eff}} = \sqrt[4]{\varepsilon} T$ und damit $T_{\text{eff,Wasser}} = 19.5^\circ\text{C}$ und $T_{\text{eff,Oel}} = 17.9^\circ\text{C}$. Der Temperaturunterschied beträgt also 1.6 K, d.h. der IR-Scanner sollte mindestens eine Temperaturauflösung von 0.8 K haben (das ist kein Problem, selbst einfache Modelle haben ein Temperaturauflösungsvermögen von 0.4 K oder besser

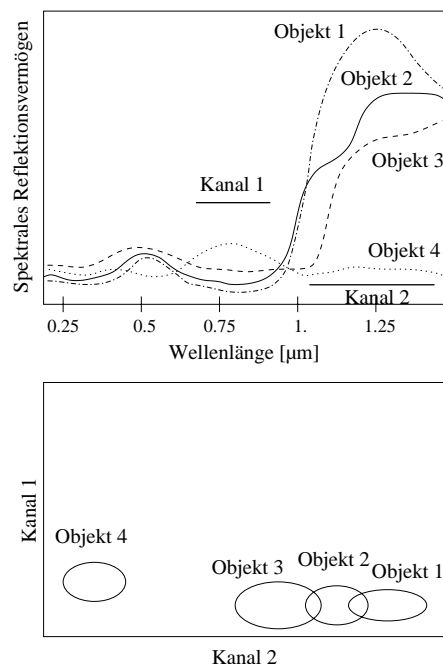
B.3 Kapitel 6

B.3.1 Fragen

1. Das Bodenauflösungsvermögen eines optischen Systems lässt sich nur dann ausnutzen, wenn man die einzelnen Bildpunkte auch gut unterscheiden kann. Bei homogenen Flächen (Polkappen, Schnee im Gebirge, ausgedehnte Vegetationsflächen) dagegen liegen die Hel-

ligkeitswerte der Szene oft in einem relativ engen Bereich. Um dennoch verschiedene Helligkeitswerte auch optisch unterscheiden zu können, verwendet man Verfahren der Kontrastareicherung. Dazu geht man von einer Häufigkeitsverteilung der Helligkeitswerte aus, identifiziert den Bereich, in dem ein Großteil der Helligkeitswerte liegt, und bildet diesen Bereich auf die vollständige Skala ab. Dazu gibt es verschiedene Abbildungsverfahren, die beiden wichtigsten sind: der Teilbereich kann einfach linear auf die gesamte Skala transformiert werden (lineare Dehnung) oder die Transformation erfolgt derart, dass in gleich großen Bereichen der Helligkeitsskala gleiche Mengen an Pixeln liegen (Histogramm-angeglichene Dehnung). Das Verfahren hängt von der Fragestellung ab, es gibt auch die Möglichkeit, unabhängig von der Verteilung der Helligkeitswerte eine lineare Dehnung eines bestimmten Bereiches auf die gesamte Skala vorzunehmen

2. die Grundidee für automatische Klassifizierungsverfahren beruht auf dem unterschiedlichen spektralen Reflektionsvermögen verschiedener Objekte:



Die Spektralkanäle des Scanners müssen dazu so gewählt sein, dass charakteristische Unterschiede im spektralen Reflektionsvermögen erkannt werden können (z.B. der berüchtigte Kanal im nahen IR zur Identifikation von Vegetation). Die Klassifizierung erfolgt mit Hilfe von Cluster-Diagrammen, bei denen die Helligkeitswerte von zwei (oder mehreren) Kanälen gegeneinander aufgetragen werden und Bereiche markiert sind, in denen sich bestimmte Objekte sammeln.

3. am geschicktesten ist es, den IR-Kanal in die dritte Dimension zu legen. Damit ist die Vegetation bereits eindeutig identifiziert (nichts anderes erzeugt ein so hohes IR-Signal) und das Gestein ist bereits ein wenig herausgehoben.
4. Die Differenz der thermischen IR-Kanäle hilft, über den Gesamtstrahlungsstrom die Temperatur zu bestimmen. Dadurch unterscheiden sich die Wolken vom Schnee. Als zusätzliche Informationen ließen sich als 'Fremddaten' passive Mikrowellendaten verwenden, da diese die Wolken durchdringen können, d.h. deren Signal ist in jedem Fall ein Boden-Signal. Außerdem würden alle möglichen Draufsicht-Instrumente zur Bestimmung von Spurengasen zusätzliche Informationen liefern können (Druckverbreiterung der Linien liefert eine Höheninformation).

B.3.2 Aufgaben

- aus der Häufigkeitsverteilung der Pixel bzw. aus der Standard-Abweichung der Pixel in einem 3 mal 3 (oder 5 mal 5)-Kernel lassen sich die Fehler identifizieren. In diesem Fall die 44 in C7, die 55 in K4, und die 76 in Q8; außerdem die 51 in Q1, die sich aber nicht korrigieren lässt, da sich kein vollständiger Kernel um sie bilden lässt. Alle drei sind auf Basis der Häufigkeitsverteilung eindeutige Ausreißer, die genaue Überprüfung zeigt, dass selbst im 3 mal 3-Kernel auch die 44 ein Ausreißer ist (die Umgebung liefert 30 ± 4 , selbst bei Berücksichtigung der 44 ergibt sich 32 ± 6). Die Korrektur erfolgt durch Auswahl eines 3 mal 3 oder 5 mal 5 Gitters, das um diesen Wert zentriert ist. Der Mittelwert der Gitterpunkte der Umgebung liefert den Ersatzwert für den fehlerhaften Wert. C7 muss dabei ersetzt werden durch 30 in beiden Kernen; für K4 ergibt sich ebenfalls 30 im 5 mal 5 Kernel und 29 im kleineren Kernel; für Q8 ergibt sich 30 in beiden Kernen.
- In diesem Fall läuft diagonal durch das Bild eine Trennkante, die beiden Bereiche haben homogene Helligkeitswerte, die drei Ausreißer 99 sind in beiden Bereichen ebenso wie auf der Kante gut zu identifizieren: D4, L6 und R8. Für D4 ergibt sich 28 im 3 mal 3 Kernel und 29 im 5 mal 5 Kernel; für R8 ergibt sich 53 in beiden Kernen und für L6 ergibt sich 40 im kleinen und 43 im großen Kernel. Hier ist die Korrektur schwieriger, da der Wert auf der Grenze zwischen den beiden Gebieten liegt – in allen drei Fällen sind jedoch die Korrekturen sowohl mit dem großen als auch mit dem kleinen Kernel sinnvoll, der Wert in L8 gibt eben den Mittelwert, was er auch in beiden Verfahren tut. Insofern ist die Nachfrage eine ‘Fangfrage’ gewesen, da nicht mein vermitteltes Vorurteil abgefragt wurde sondern wirklich anhand der Daten geprüft werden musste.
- Der Datensatz unterscheidet sich von den bisherigen in relativ gemeiner Weise, da die Fehler zwar geclustert aber recht subtil zu finden sind (69 H4, 55 in K4, 78 in J5, 76 in L5, 44 in H6). Für die Korrektur ergeben sich in H4 29 (bzw. 31 im 5 mal 5, da hier die beiden Fehler in J5 und H6 mit in die Korrektur eingehen); in K4 40 bzw. 32 im 5 mal 5 Kernel (hier ist der letztere besser, da die beiden Fehler in J5 und L5 in beide Kernel eingehen aber im größeren Kernel eine geringere Verfälschung der Ergebnisse bewirken); in J5 30 im 3 mal 3 Kernel (hier macht sich der Fehler aus K4 nicht bemerkbar, da die anderen Werte ungewöhnlich niedrig liegen) bzw. 34 im großen Kernel (obwohl hier alle anderen Fehler mit in die Korrektur eingehen! nachrechnen!); in L5 31 (der Fehler durch den einen fehlerhaften Wert ist nur klein, da der fehlerhafte Wert selbst recht klein ist und die anderen Werte eher niedrig liegen) bzw. 32 (hier kommt zwar ein weiterer fehlerhafter Wert hinzu, aber die anderen Werte sind recht klein); in H6 29 (hier wird in den kleinen Kernel auch kein fehlerhafter Wert einbezogen) bzw. 32 im großen Kernel, der zwei fehlerhafte Werte enthält. Fazit: Mehrfachfehler machen jede Korrektur problematisch. Zwar ist beim 5 mal 5 Kernel das Risiko, weitere Fehler in die Korrektur einzubeziehen größer, allerdings ist der relative Anteil jedes einzelnen Fehlers am korrigierten Wert in einem größeren Kernel geringer. Daher bei Mehrfachfehlern erst Fehlermarkierung und diese dann gegebenenfalls mit einem größeren Kernel gar nicht zur Korrektur verwenden.
- Mit einem 3 mal 3-Kernel (5 mal 5-Kernel) erhält man für die Werte in der Spalte 99(99), 29(99), 30.5(30), 32.5(30), 30(31), 30.5(30.5), 30(30), 31(31), 31(99), 99(99), wobei die 99 jeweils ein Füllwert ist, der nicht korrigiert werden kann, da sich um ihn herum kein Kernel bilden lässt.
- Die Kantenanreicherung ist ein Verfahren, das sich ebenfalls auf einen, in der Regel 3 mal 3-Kernel stützt. Allerdings wird der zentrale Wert dort nicht durch den Mittelwert der umgebenden Kernelemente ersetzt sondern die vorhandenen Werte werden mit dem Kernel multipliziert: die Eckwerte mit einer Null, der zentrale Wert mit einer 4, die mitten der Kanten mit -1:

$$K = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix}.$$

Diese Werte werden aufaddiert und mit dem Zentralwert des Kernels kombiniert. Der so

erhaltene Wert ersetzt den ursprünglichen Zentralwert des Kernels. Das sich ergebende Lösungsmuster ist

$$R = \begin{pmatrix} 34 & 24 & 36 & 30 & 28 & 27 & 30 & 25 & 33 & 32 \\ 26 & 33 & 25 & 33 & 35 & 24 & 33 & 23 & 26 & 29 \\ 31 & 23 & 27 & 26 & 23 & 53 & 35 & 28 & 33 & 24 \\ 33 & 44 & 30 & 34 & 18 & 7 & 25 & 65 & 16 & 23 \\ 29 & -20 & 35 & -24 & 39 & 11 & 3 & -17 & 16 & 24 \\ 52 & 112 & 69 & 103 & 66 & 105 & 87 & 63 & 54 & 51 \\ 52 & 80 & 21 & 36 & 50 & 26 & 38 & 77 & 62 & 61 \\ 56 & 48 & 60 & 55 & 62 & 57 & 49 & 61 & 55 & 52 \\ 54 & 50 & 57 & 52 & 60 & 59 & 62 & 49 & 57 & 53 \\ 49 & 54 & 52 & 61 & 49 & 55 & 51 & 57 & 62 & 54 \end{pmatrix}.$$

Die Kontrastanreicherung funktioniert an den Stellen relativ gut, in denen der Pixel in der Nähe des Mittelwertes für den entsprechenden Bereich liegt – bei starken Abweichungen ist das Resultat nicht so gut.

6. einfach in die Abbildung reinmalen, gesamte Skala linear ausnutzen (im zweiten Fall unter Vernachlässigung der kleinen Werte am Rand).
7. einfach in die Abbildung einzeichnen.

Abbildungsverzeichnis

1.1	Wüstenlandschaft des Fessan (Lybien)	2
1.2	Osnabrück vom Ikonos	2
1.3	Ozonloch in Gesamtozonsäule (TOMS und GOME)	3
1.4	Erdfernerkundung ohne Aufzeichnung	5
1.5	Charakteristische Skalen	7
2.1	Kepler'sche Gesetze	13
2.2	Bewegung um den gemeinsamen Schwerpunkt	14
2.3	Kenngößen einer elliptischen Bahn	18
2.4	Festlegung der Bahnparameter einer Satellitenbahn	21
2.5	Beispiele für Satellitenbahnen	21
2.6	Bodenspur LandSat	23
2.7	Bahnen von Kommunikationssatelliten	24
2.8	Konfiguration der GPS-Satelliten	25
2.9	Höhen und Umlaufzeiten verschiedener Satelliten	27
2.10	Stabilisierungssysteme	28
2.11	Polare und sonnensynchrone Bahn	30
2.12	... und wirft seinen Schatten voraus	30
2.13	Tägliche Dichtevariation in der Hochatmosphäre	33
2.14	Luftwiderstand und Cosmos 1097	35
2.15	Änderung der Bahnparameter durch Luftwiderstand	35
2.16	SMM-Absturz	36
2.17	Bahnstörung in Folge der Abplattung der Erde	37
2.18	Präzession der Bahneben	38
2.19	Apsidendrehung	39
2.20	Hohmann-Bahn, Übergangsellipse	46
2.21	Übergangsbahn aus Ellipsenbögen	48
2.22	Lagrange-Punkte im rotierenden Bezugssystem	51
2.23	Bahnen der Voyager- und Pioneer-Sonden	54
2.24	Swingby an einem Planeten	55
2.25	VEEGA	57
3.1	Aufbau der Atmosphäre	62
3.2	Atmosphärische Transmission	63
3.3	Sensorklassen und ihre Beziehung	66
3.4	Farbmischprojektor	68
3.5	Panchromatisch vs. Infrarotfilm	70
3.6	Vegetation panchromatisch und im Infrarotfilm	70
3.7	Frabfilm	71
3.8	IR-Farbfilm	71
3.9	Farb- und Falschfarbenaufnahme	72
3.10	Spektrale Empfindlichkeiten	73
3.11	Amazonas-Abholzung	75

3.12	Bodenauflösungsvermögen	76
3.13	Bodenauflösungsvermögen	77
3.14	Winkelauflösungsvermögen	77
3.15	Einfluss des Winkelauflösungsvermögens	78
3.16	Charakteristische Kurve eines Films	79
3.17	Erreichtes Bodenauflösungsvermögen	80
3.18	LandSat Videokameras	83
3.19	CCD	84
3.20	Kehrbesentechnik und MOMS-Doppeloptik	84
3.21	HRV auf SPOT	85
3.22	MSS auf LandSat	87
3.23	LandSat MSS Wellenlängenbereiche	87
3.24	TM-Einzelbilder	89
3.25	Mixed Pixels	89
3.26	Linieneffekt	90
3.27	Prinzip eines IR-Scanners	94
3.28	Spektrale Signaturen von Gestein	95
3.29	Tschernobyl im IR, 1985 und 1986	96
3.30	Globale Infrarotaufnahme	97
3.31	Unterirdisches Wasser	97
3.32	Golfstrom im thermischen Infrarot	98
3.33	Mikrowellenradiometer	99
3.34	Desertifikation: Polarisationsunterschiede bei 37 GHz	100
3.35	Abschmelzendes Eis: Polarisationsunterschiede bei 18 GHz	100
3.36	SeaSat Altimeter	102
3.37	Geoid	102
3.38	Sea Floor Mapping	103
3.39	Altimeter-Signal	103
3.40	Scatterometer: Windgeschwindigkeiten	104
3.41	Prinzip des SAR	105
3.42	SAR	106
3.43	Limb-Sounder	107
3.44	Vulkan-Aerosol	110
3.45	Struktur eines Pflanzenblatts	115
3.46	Spektrales Reflektionsvermögen Pflanzenblatt	116
3.47	Globale Landbedeckung	117
3.48	Untergrund und Pflanzenwuchs	119
3.49	Seewasser und Öl: spektrales Reflektionsvermögen	120
3.50	Oberflächentemperatur Meere	122
3.51	El Niño	123
3.52	Die Farbe des Meeres	125
4.1	System Erde	140
5.1	Rückführung des Filmmaterials	144
5.2	Lebensdauer Photosatellit	144
5.3	Rückführung analoger Daten auf digitalem Weg	145
5.4	Nachrichtenstrecke	146
5.5	Kommunikationssystem	147
5.6	Ungestörter Informationskanal	149
5.7	Shannon-Funktion	152
5.8	Binärcodes Alphabet	156
5.9	Diagramme von Quellen	159
5.10	Aliasing	160
5.11	Schallquellen	161

5.12	Menschliches Ohr	162
5.13	Dictionary of Tunes	163
5.14	Bild und Rasterung	164
5.15	Codierbaum	168
5.16	Codierbaum zum Huffman-Code	169
5.17	Ideales und reales Binärsignal	173
5.18	Entropiefluss ungestörter Übertragungskanal	174
5.19	Shannon-Grenze für fehlerfreien Empfang	176
5.20	Symmetrischer Binärkanal	178
5.21	Ähnlichkeitsdecodierung	179
5.22	Blockcodierung in Matrixform	180
5.23	Antennengewinn Richtantenne	182
5.24	Regendämpfung	184
5.25	IntelSat V	185
6.1	Fehlende Pixel	188
6.2	Kernel zur Korrektur fehlender Pixel	189
6.3	Fehlende Zeilen	190
6.4	Poisson-Rauschen	191
6.5	Atmosphärische Streuung	192
6.6	Dunststreuung (terrestrisch)	193
6.7	Dunstkorrektur	194
6.8	Dunstkorrektur Bsp.	195
6.9	Kontrastanreicherung	196
6.10	Helligkeitswerte	197
6.11	Kontrastanreicherung in einem Kanal	198
6.12	Histogramm Gletscher und Fels	199
6.13	Kontrastanreicherung Fels-Gletscher	200
6.14	Ungerichtete Kantenanreicherung	201
6.15	Kantenanreicherung, Beispiel	201
6.16	Zweidimensionale Klassifikationsverfahren	203
6.17	Multitemporale 2D-Klassifikation	204
6.18	Klassifikation weißer Pixel	204
6.19	Regionale Landnutzung	206
6.20	JPEG-Kodierung und Zahl der Koeffizienten	208
6.21	Progressive und Raster-Datenübertragung	209
6.22	Fraktale Bilderzeugung	210

Tabellenverzeichnis

1.1	Wichtige Satelliten(instrumente)	4
1.2	Remote Sensing Instrumente	6
2.1	Kegelschnitte	17
2.2	Flugbahnen bei verschiedenen Abschussgeschwindigkeiten	19
2.3	Bahnparameter als Funktion der großen Halbachse	20
2.4	Übersicht über Bahnstörungen	31
2.5	Äußerer Wirkungsgrad einer Rakete	41
2.6	Ein- und Dreistufenrakete	43
2.7	Hohmann-Übergänge zu anderen Planeten	47
3.1	Übersicht Fernerkundungssensoren	65
3.2	Spektrale Empfindlichkeit MSS	74
3.3	Charakteristika SPOT HRV	86
3.4	Spektralbänder TM (LandSat)	88
3.5	Charakteristika von SARs	107
3.6	Treibhauswirksame Spurengase	108
4.1	Übersicht LandSat	130
4.2	LandSat MSS, TM und ETM-Merkmale	131
4.3	Übersicht TIROS bis NOAA	133
4.4	Spektralbänder auf MERIS	135
4.5	Spektralkanäle MSRS	137
5.1	Informationsgehalt Sprache	153
5.2	ASCII-Code	166
5.3	Beispielcodes	167
5.4	Redundanzreduktion verschiedener Codes	168
5.5	Beispiel zum Lempel–Ziv-Algorithmus	172
5.6	Übertragungskapazitäten	175

Literaturverzeichnis

- [1] R.E.W. Adams, W.E. Brown, and J. Culbert: Radar mapping, archaeology and ancient Maya land use, *Science* 213, 1457, 1981 [113](#)
- [2] G. Asras, M. Fuchs, E.T. Kanemasu, and J.L. Hartfield: Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat, *Agron. J.* 76, 300, 1984 [115](#)
- [3] J.D. Baker: *Planet Earth - the view from space*, Havard University Press, Cambridge, 1990 [80](#), [90](#), [126](#), [139](#), [141](#)
- [4] E.C. Barrett und L.F. Curtis: *Introduction to environmental remote sensing*, Chapman and Hall, London, 1982 [63](#), [74](#), [95](#), [113](#), [115](#), [116](#)
- [5] S.J. Bauer: *Physics of planetary ionospheres*, Springer, Berlin, 1973 [34](#)
- [6] J.K. Beatty, B. O'Leary, and A. Chaikin: *The New Solar System*, Cambridge University Press, Cambridge, 1990 [54](#), [57](#)
- [7] Bergmann-Schaefer: *Lehrbuch der Experimentalphysik III: Optik*, de Gruyter, 1978 [77](#)
- [8] B. Bertotti und P. Farinella: *Physics of the Earth and the solar system*, Kluwer, 1990 [15](#), [31](#), [38](#), [53](#), [57](#), [173](#), [185](#)
- [9] J. Bodechtel und H.-G. Gierloff-Emden: *Weltraumbilder der Erde*, List, 1969 [1](#), [2](#)
- [10] A. Bohrmann: *Bahnen künstlicher Satelliten*, BI, 1966 [15](#), [21](#), [34](#), [35](#), [37](#), [38](#), [39](#), [57](#)
- [11] H.-J. Bolle: Führt der Anstieg atmosphärischer Spurengaskonzentration zum "Klimakollaps"?, *Phys. Blätter*, **43/7**, 240, 1987 [141](#)
- [12] Bormann: *Erdfernerkundung*, Veröffentlichungen des Zentralinstituts für Physik der Erde # 65, Potsdam, 1980 [79](#), [81](#), [105](#)
- [13] O. Bschorr und A. Leibold: Das Einfang- und Katapultvermögen von Planet/Mond-Systemen, *Raumfahrtforschung* **13**, 132, 1968 [54](#)
- [14] N. Calder: *Raumschiff Erde*, vgs, Köln, 1992 [72](#), [75](#), [96](#), [103](#)
- [15] B.J. Choudhuri: *Desertification*, in [\[33\]](#), 1993 [100](#)
- [16] A.P. Cracknell und L.W.B. Hayes: *Introduction to remote sensing*, Taylor und Francis, 1991 [4](#), [6](#), [11](#), [63](#), [87](#), [90](#), [94](#), [96](#), [97](#), [99](#), [102](#), [103](#), [104](#), [110](#), [126](#), [187](#), [199](#), [211](#)
- [17] P.J. Curran: *Principles of remote sensing*, Longmann, New York, 1985 [82](#), [83](#), [113](#), [126](#), [187](#)
- [18] R.S. DeFries, and J.R.G. Townshend: Global land cover: comparison of ground-based datat sets to classifications with AVHRR data, in [\[24\]](#), 1994 [116](#), [117](#)
- [19] C. Drewniok: *Der Einsatz von Satelliten zur Erdbeobachtung*, CENSIS – Center for Science and International Security, Univ. Hamburg, FB Informatik, 1991 [23](#), [30](#), [145](#)
- [20] R.M. Fano: *Transmission of information*, Wiley, New York, 1962 [150](#)
- [21] R.M. Fano: *Informationsübertragung - eine statistische Theorie der Nachrichtenübertragung*, Oldenbourg, München, 1966 [150](#), [152](#), [185](#)
- [22] F.A. Fischer: *Einführung in die statistische Übertragungstheorie*, Bibliographisches Institut, Mannheim, 1969 [160](#)
- [23] Y. Fisher (Hrsg.): *Fractal image compression*, Springer, 1996 [210](#)
- [24] G. Foody und P. Curran: *Environmental remote sensing from regional to global scales*, Wiley, 1994 [7](#), [80](#), [90](#), [244](#), [246](#)
- [25] R.H. Giese: *Weltraumforschung*, BI, 1966 [28](#), [41](#), [43](#), [48](#)

- [26] R.H. Giese: *Einführung in die Astrophysik*, BI-Wissenschaftsverlag, Mannheim, 1981 [54](#), [55](#)
- [27] Goldstein: *Klassische Mechanik* [15](#)
- [28] T.T. Graedel and P.J. Crutzen: *Chemie der Atmosphäre*, Spektrum - Akademischer Verlag, Weinheim, 1994 [108](#), [110](#)
- [29] T.T. Graedel and P.J. Crutzen: The changing atmosphere, in *Managing planet Earth*, W.H. Freeman and Co., New York, 13 - 24, 1989 [110](#)
- [30] T. Grams: *Codierungsverfahren*, BI Hochschultaschenbücher 625, Mannheim, 1986 [170](#), [179](#), [185](#)
- [31] W. Greiner: *Theoretische Physik: Mechanik II*, 2. Aufl. Harri Deutsch, Thun, 1977 [15](#)
- [32] H. Großmann: *Satelliten-Thermalbilder - ein neues Hilfsmittel für die Umweltforschung?*, Bundesforschungsanstalt für Landeskunde und Raumordnung, Bonn, ISBN 3-87994-016-9, 1984 [95](#)
- [33] R.J. Gurney, J.L. Foster, und C.L. Parkinson (Hrsg.): *Atlas of satellite observations related to global change*, Cambridge, 1993 [244](#), [245](#), [246](#), [247](#)
- [34] M. Gurtner: *Karten lesen - Handbuch zu den Landeskarten*, Gemeinschaftsausgabe Bundesamt für Landestopographie und SAC, 1998 [102](#)
- [35] H. Häberle: *Informationstheoretische Grundlagen der Digitaltechnik*, VDI-Bildungswerk, 1966 [148](#)
- [36] W. Hallmann und W. Ley: *Handbuch der Raumfahrttechnik*, Hanser, München, 1988 [28](#)
- [37] R.W. Hamming: *Coding and information theory*, Prentice-Hall, New York, 1980 [178](#)
- [38] H. Handels: *Medizinische Bildverarbeitung*, Teubner, Leipzig, 2000 [211](#)
- [39] R.C. Harriss: Tropospheric chemistry, in [33], 1993 [109](#)
- [40] Ph. Hartl: *Fernwirktechnik der Raumfahrt*, Springer, Berlin, 1977 [175](#), [176](#), [185](#)
- [41] B. Hassenstein: *Biologische Kybernetik*, Quelle und Meyer, Heidelberg, 1965 [160](#)
- [42] T. Hauf, Wendling, P., and K.-T. Kriebel: Wolken, Klima, und Umwelt, *Physik in unserer Zeit*, **22**, 59, 1991 [140](#)
- [43] E. Herter und H. Rupp: *Nachrichtenübertragung über Satelliten*, Springer, Berlin, 1979 [182](#), [185](#)
- [44] E. Herter und W. Lörcher: *Nachrichtentechnik: Übertragung - Vermittlung - Verarbeitung*, Hanser, München, 1992 [149](#), [160](#), [186](#)
- [45] H. Heuseler (Hrsg.): *Europa aus dem All*, dva/Westermann, 1974 [1](#)
- [46] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins: *GPS - theory and practice*, Springer, Wien, 1992 [26](#)
- [47] J.T. Houghton, F.W. Taylor, and C.D. Rodgers: *Remote sounding of atmospheres*, Cambridge University Press, Cambridge, 1984 [106](#), [107](#)
- [48] J.T. Houghton: *The physics of atmospheres*, Cambridge University Press, Cambridge, 1986 [108](#)
- [49] J.T. Houghton: *Global warming*, Cambridge University Press, Cambridge, 1997 [108](#)
- [50] P. Hupfer (Hrsg.): *Das Klimasystem der Erde*, Akademie-Verlag, Leipzig, 1991 [141](#)
- [51] K. Jacob: Das Meer unter der Wüste, *Bild der Wissenschaft* 1/1995. 18, 1995 [97](#), [118](#)
- [52] B. Jähne: *Digitale Bildverarbeitung*, Springer, 2002 [187](#), [191](#), [199](#), [211](#)
- [53] P. Janle: Das Bild der Venus nach der Magellan-Mission, Teil 1 -3, *Sterne und Weltraum* 33, No. 10 - 12, S. 686, 1994 [106](#)
- [54] J. Jansen: *Kriegsschauplatz Kiel*, Wachholtz, Neumünster, 1989 [5](#)
- [55] R.S. Kandel: *Earth and Cosmos*, Pergamon Press, 1980 [140](#)
- [56] W.W. Kelley: *Der Heimatplanet*, Zweitausendeins, Frankfurt, 1988 [1](#)
- [57] D. Kerner und I. Kerner: *Der Klima-Report*, Kiepenheuer und Witsch, Köln, 1990 [141](#)
- [58] W. Kertz: *Einführung in die Geophysik, I und II*, BI Hochschultaschenbücher, Bibliographisches Institut, Mannheim, 1971, Neuauflage 1989 [33](#), [38](#)
- [59] C.R. Kitchin: *Astrophysical Techniques*, Adam Hilger, Bristol, 1984 [79](#), [82](#), [83](#), [84](#), [126](#)
- [60] R. Klingholz: *Marathon im All*, Ullstein, Berlin, 1992 [182](#)

- [61] A. Kowal und L. Dessinow: *In den Weltraum zum Nutzen der Menschheit*, Verlag Progress Moskau und Staatsverlag der Deutschen Demokratischen Republik, 1987 [68](#), [74](#), [81](#)
- [62] H.J. Kramer: *Observations of the Earth and its Environment – Survey of missions and sensors*, Springer, 2002 [61](#), [65](#), [78](#), [80](#), [126](#), [130](#), [131](#), [133](#), [142](#)
- [63] A. Krass: *Verification - How much is enough?*, Lexington Books, D.C. Heath and Co., Lexington MA, 1985 [25](#), [35](#)
- [64] D. Kreß: *Theoretische Grundlagen der Signal- und Informationsübertragung*, Vieweg, Braunschweig, 1977 [148](#)
- [65] K. Küpfmüller: *Einführung in die theoretische Elektrotechnik*, Springer, Berlin, 1990 [154](#)
- [66] T.G. Kyle: *Atmospheric transmission*, Pergamon, Oxford, 1991 [81](#)
- [67] J. Lean: Variations in the Sun's radiative output, *Rev. of Geophys.* **29**, 4, 505, 1991 [36](#)
- [68] J.S. Levine (ed.): *Space Opportunities for tropospheric chemistry*, NASA Conf. Publ. 2450, Washington, DC, 1985 [109](#)
- [69] G.H. Liljequist und K. Cihak: *Allgemeine Meteorologie*, Vieweg, Braunschweig, 1977 [62](#)
- [70] H. Lippert: *Lehrbuch Anatomie*, Urban und Schwarzenberg, München, 1990 [162](#)
- [71] E. Löffler: *Geographie und Fernerkundung*, Teubner, 1994 [4](#), [30](#), [68](#), [69](#), [70](#), [71](#), [82](#), [83](#), [84](#), [85](#), [87](#), [89](#), [90](#), [126](#), [196](#), [203](#), [204](#), [211](#)
- [72] R.M. Lucas and A.R. Harrison: Snow monitoring in the United Kingdom using NOAA AVHRR imagery, in [\[24\]](#), 1994 [119](#), [204](#)
- [73] K. Luchner: *Physik ist überall*, Ehrenwirth, München, 1994 [54](#)
- [74] H.D. Lüke: *Signalübertragung*, Springer, 1992 [156](#), [186](#)
- [75] A.F. Marfeld: *Astronautik*, Safari-Verlag, Berlin (Sonderausgabe zur Apollo-Mondlandung), 1969 [48](#), [49](#)
- [76] H. Marko: Grundlagen der Information und Kommunikation, *Verhandlungen der Ges. dt. Naturforscher und Ärzte*, **113**, 1984 [176](#)
- [77] W. Markwitz, und R. Winter (Hrsg.): *Fernerkundung - Daten und Anwendungen*, Wichmann, Karlsruhe, 1989 [89](#)
- [78] C.R. McClain, G. Feldman und W. Esaias: *Oceanic biological productivity*, in [\[33\]](#) [125](#)
- [79] A.C. Millington, J. Wellens, J.J. Settle, and R.J. Saull: Explaining and monitoring land cover dynamics in drylands using multi-temporal analysis of NOAA AVHRR imagery, in [\[24\]](#), 1994 [116](#)
- [80] B.J. Murphy and R.D. Nance: Mountain belts and the supercontinent cycle, *Scientific American* **266/4**, 34, 1992 [139](#)
- [81] G. Neukum: *High Resolution Stereo Camera (HRSC) - Mars 94 mission*, Deutsche Forschungsanstalt für Luft- und Raumfahrt, Oberpfaffenhofen, DLR-IB 90/3, 1990 [90](#)
- [82] R.E. Newell, H.G. Reichle, Jr., and W. Seller: Carbon monoxide from the burning earth, *Scient. Amer.* **260**, 82, 1989 [109](#)
- [83] E.G. Njoku und O.B. Brown: *Sea surface temperature*, in [\[33\]](#) [122](#), [123](#)
- [84] D. Parson: *Dictionary of tunes and musical themes*, Spencer and Brown, Cambridge, 1975 [163](#)
- [85] Pelkham: DuMont, Köln [5](#)
- [86] W.W. Peterson and E.J. Weldon: *Error correcting codes*, MIT Press, Cambridge, 1972 [180](#)
- [87] J.R. Primack: Gamma-ray observations of orbiting nuclear reactors, *Science* **244**, 397, 1989 [101](#)
- [88] J.G. Proakis und M. Salehi: *Grundlagen der Kommunikationstechnik*, Pearson, 2004 [172](#), [185](#), [208](#)
- [89] U. Quernheim: *Satellitenkommunikation – Kanalmodellierung und Protokollbewertung*, ABI Bd. 1, Verlag der Augustinus Buchhandlung, 1993 [173](#), [184](#), [186](#)
- [90] V. Ramanathan, B.R. Barkstrom, and E.F. Harrison: Climate and the Earth's radiation budget, *Physics Today* /**5**, 22, 1989 [140](#)

- [91] W.R. Rossow: *Clouds*, in [33], 1993 97
- [92] A.E. Roy: *Orbital motion*, Adam Hilger, Bristol, 1982 38
- [93] W.G. Rees: *Physical principles of remote sensing*, Cambridge University Press, Cambridge, 1990 79, 99, 126
- [94] H.G. Reichle, Jr., V.S. Conners, J.A. Holland, W.D. Hypes, H.A. Wallio, J.C. Casas, B.B. Gormsen, M.S. Saylor, and W.D. Heske: Middle and upper atmospheric carbon monoxide mixing ratios as measured by a satellite-born remote sensing sensor during November 1981, *J. Geophys. Res.* 91, 10865, 1986 109
- [95] U. Renner, J. Nauck und N. Balteas: *Satellitentechnik*, Springer, 1988 24, 146, 186
- [96] N. Roberts: The global environmental future, in *The changing global environment* (ed. N. Roberts), Blackwell, Cambridge, 1994 140
- [97] J.G. Roederer: *Introduction to the physics and psychophysics of music*, Springer, Berlin, 1975; neue deutschsprachige Auflage 1993, ebenfalls bei Springer, unter dem Titel *Physikalische und psychoakustische Grundlagen der Musik* 162, 163
- [98] Roedel: *Physik der Atmosphäre*, Springer 108
- [99] F.S. Rowland and I.S.A. Isaksen (eds.): *The changing atmosphere*, Wiley and Sons, New York, 1988 110
- [100] M. Rycroft (Hrsg.): *The Cambridge encyclopedia of space*, Cambridge University Press, Cambridge, 1990 48, 86, 126
- [101] F.F. Sabins, Jr.: *Remote sensing – Principles and interpretation*, Freeman, 1996 11, 68, 69, 70, 72, 73, 74, 76, 79, 81, 86, 87, 88, 89, 97, 105, 106, 107, 114, 118, 119, 120, 126, 129, 187, 192, 193, 206, 211
- [102] P. Sagirov: *Satellitendynamik*, BI-Hochschultaschenbücher, 1970 15, 48, 57
- [103] H. Schönfelder: *Bildkommunikation*, Springer, Berlin, 1983 172
- [104] M. Schneider: *Satellitengeodäsie*, BI, 1988 15, 40
- [105] S.H. Schneider: Klimamodelle, *Spektrum der Wissenschaft* 7, 52, 1987 141
- [106] M.R. Schoeberl: Stratospheric ozone depletion, in [33], 1993 110
- [107] C:D. Schönwiese: *Klima im Wandel*, DVA Stuttgart, 1992 139
- [108] A. Schubert: Water takes a peculiar turn in the gulf stream, *EOS* 75, 473, 1994 122
- [109] P.J. Sellers: Canopy reflectance, photosynthesis and transpiration, *Int. J. remote sensing*, 12, 1113, 1985 115
- [110] C.E. Shannon: *A mathematical theory of communication*, Bell Syst. Tech. J. 27, 379-423 und 623-656, 1948 149, 157, 170
- [111] C.E. Shannon and W. Weaver: *The mathematical theory of information*, Univers. Ill. Press, Urbana, 1949 149, 152, 160, 167
- [112] C.E. Shannon and W. Weaver: *Mathematische Grundlagen der Informationstheorie*, Oldenbourg, München, 1976 147, 156, 159, 185
- [113] S.M. Stanley: *Earth and Life through Time*, Freeman, New York, 1989 139
- [114] K. Steinbuch: *Automat und Mensch*, Springer, Berlin, 1962 154, 160, 164, 174, 185
- [115] K. Steinbuch: *Taschenbuch der Nachrichtenverarbeitung*, Springer, Berlin, 1967 185
- [116] H.L. Stiltz (ed.): *Aerospace telemetry*, Prentice-Hall, New York, 1961 160
- [117] P. Strain and F. Engle: *Portät unseres Planeten*, Westermann, Braunschweig, 1993 4, 5, 72, 114
- [118] C.H. Strandberg: Photoarchaeology, *Photogrammetric Engeneering* 33, 1152, 1967 113
- [119] C.P. Summerhayes und S.A. Thorpe: *Oceanography*, Manson, 1998 98
- [120] C. Taylor: *Der Ton macht die Musik*, Vieweg, Braunschweig, 1994 161, 163
- [121] R.H. Thomas: Ice sheets, in [33], 1993 100
- [122] J.R.G. Townshend, C.J. Tucker, and S.N. Goward: Global vegetation mapping, in [33], 1993 116
- [123] C.J. Tucker, B.N. Holben, and T.E. Goff: Intensive forrest clearing in Rondonia, Brazil, as detected by remote sensing, *Remote sensing of environment* 15, 255, 1984
- [124] F. Vester: *Denken, lernen und vergessen*, dtv 160
- [125] I. von Cube: *Kybernetische Grundlagen des Lernens*, Klett, Stuttgart, 1968 160, 162, 185

- [126] P. Wayner: *Compression algorithms for real programmers*, Morgan Kaufmann, 2000 [185](#), [208](#)
- [127] R.M. White: The Great Climate Debate, *Scientific American*, **263(1)** (July), 18, 1990 [141](#)
- [128] H. Wittbrodt, H. Mielke, G. Narimanow, und J. Saizew: *Raumfahrt für die Erde I*, transpress VEB Verlag für Verkehrswesen, Berlin, 1975 [68](#), [69](#), [78](#), [81](#)
- [129] I.H. Witten, A. Moffat, und T.C. Bell: *Managing gigabytes*, Morgan Kaufmann, 1999 [185](#), [206](#), [207](#), [208](#)
- [130] D.O.A. Wolf, H.M. Hoose, und M.A. Dausen: *Die Militarisierung der Weltraums*, Bernard und Graefe, Karlsruhe, 1983 [144](#)
- [131] H.P. Zenner and E. Zrenner: Einführung in *Physiologie der Sinne*, Spektrum - Akademischer Verlag, Heidelberg, 1994 [161](#)
- [132] R. Zirkle: *Returning to space: priorities and choices in US launch policy*, Union of Concerned Scientists, Cambridge, MA, 1987 [21](#)

Index

- Ähnlichkeitsdecodierung, 179
- Äquivokation, 174
- Übergangsellipse, 46
- Übertragungsfehler, 173
- Übertragungskanal, 146, 147, 173
 - binärsymmetrischer, 178
 - Entropiefluss, 174
 - Entropifluss, 174
 - gestört, 173
- Übertragungskapazität, 175
- Übertragungsrate, 181

- AATSR, 136
- Abbildungsmaßstab, 69
- Abplattung, 37
- Absorptionskoeffizient, 64
- Absorptionsvermögen
 - spektrales, 93
- Abtastfrequenz, 160
- Abtasttheorem, 159
- Advanced Along Track Scanning Radiometer, 136
- Advanced Microwave Scanning Radiometer, 92
- Advanced Microwave Sounding Unit, 92
- Advanced SAR, 135
- Advanced Very High Resolution Radiometer, 22, 90, 116, 132, 134
- Advanced Vidicon Camera System, 132
- AIRS, 91
- Aliasing, 160
- ALT, 92
- Altimeter, 92, 101, 102
- AMSR, 92
- AMSU, 92
- Anfangsgeschwindigkeit
 - elliptische, 17
 - hyperbolische, 17
 - parabolische, 17
- Anomalie
 - wahre, 18
- Antenne
 - entdrallt, 183
- Antennengewinn, 183
- Aphel, 18
- Apogäum, 18

- Apollo, 69
- Apozentrum, 18
- Apsidendrehung, 39
- Apsidenlinie, 18
- ARGOS, 6
- ARJ, 171
- ASAR, 135
- ASCII-Code, 166
- ASTER, 1
- ATLID, 92
- Atmosphäre
 - Stockwerkstruktur, 61, 62
- Atmospheric Infrared Sounder, 91
- Atmospheric Lidar, 92
- Aufklärungssatellit, 22
- Auflösungsvermögen, 66, 88
 - Boden, 22
 - Boden-, 66
 - Film, 79
 - Kamera-Film, 76
 - Mikrowellen, 99
 - Objektiv, 76, 77
 - photographische Emulsion, 76, 78
 - räumliches, 7, 75, 76
 - IR, 95
 - MSS, 87
 - SAR, 105
 - TM, 87
 - radiometrisches, 66, 76, 80
 - IR, 95
 - spektrales, 76, 80
 - TM, 87
 - thermales, 76
 - thermisches, 66, 80
 - Winkel, 77
 - zeitliches, 66, 76
- AVCS, 132
- AVHRR, 1, 3, 22, 90, 116, 117, 132, 134, 204
- Azur, 29

- Bahn
 - äquatoriale, 21
 - geostationäre, 21
 - polare, 21
 - sonnensynchron, 23

- sonnensynchron, 21, 29, 39
- stark exzentrische, 21
- Bahndrehung, 38
- Bahnebene
 - Drehung, 38
 - Gleichung, 16
- Bahnellipse
 - Drehung, 39
- Bahngleichung, 17
- Bahnparameter, 17, 18
- Bahnschrumpfung, 35
- Bandbreite, 175
- Barnsley Farn, 210
- Big Bird, 22
- Bildelement, 76
- Bilderzeugung
 - fraktale, 210
- Bildkomprimierung
 - FELICS, 207
 - fraktale, 210
 - GIF, 207
 - JPEG, 207
 - PNG, 207
- Bildpunkt, 171
- Bildverschiebung, 81
- Binärzeichen, 150, 166
- binary digit, 150
- Binet'sche Gleichung, 19, 46, 56, 58
- Bit, 150, 166
- bit, 150, 166
- Bitenergie, 175
- Bitrate, 175, 181
- Bitzeit, 175
- Blendenöffnung, 77
- Blickfeld, 69
- Blickwinkel, 69
- Blockcode, 166, 177, 178
- Blockcodierung
 - Matrixform, 180
- Bodenaufklärungsvermögen, 7, 22, 66, 69, 75–77
 - IR, 95
 - MSS, 87
 - TM, 87
- Bodenfeuchte, 100
- Bodennutzung, 206
- Bodenspur, 22, 23
- Bouguer–Lambert–Beer'sches Gesetz, 64
- BUV, 108, 109

- Cartwheel, 29, 132
- CCD, 83, 84
- CGRO, 27
- Challenging Minisatellite Payload, 136
- CHAMP, 130, 136

- Charged Coupled Device, 83
- climate change, 1, 3, 67, 132
- Cluster, 205
- Cluster-II, 26
- Coastal Zone Color Scanner, 124
- Codierbaum, 167, 168
- Codierung
 - Fano, 169
 - Huffman, 169
 - ideale, 169
 - redundanzfreie, 169
 - Shannon, 168, 169
- Coloumbia, 60
- Compton Gamma Ray Observatory CGRO, 27
- CORONA, 67
- Cosmos 1097, 36
- CZCS, 124

- Dämpfung
 - Atmosphäre, 184
- Datenübertragung
 - progressive, 209
 - Raster, 209
- dBm, 184
- Decodierbarkeit, 167, 180
- Dehnung, 196
 - besondere, 196
 - Histogramm angeglichen, 196
 - lineare, 196
- Deklassifizierung, 67, 79
- DGPS, 26, 138
- Diamant, 84
- Dichte
 - Hochatmosphäre, 33
- differentielles GPS, 138
- Digramm-Struktur, 158
- Digramm-Wahrscheinlichkeit, 156
- Discoverer-14, 67
- Distanzverfahren, 202
- diurnal bulge, 34
- DMSP, 67
- Doppelsternsystem, 53
- Downlink, 65
- Drehung der Bahnebene, 38
- Drehung der Bahnellipse, 39
- Drei-Achsen-Stabilisierung, 28, 29
- Dreikörperproblem, 49, 50
 - eingeschränktes, 50
- Dunstfilter, 81
- Dunstkorrektur, 81, 193–195
- Dunststreuung, 73, 192–194

- Earth Observing System, 141
- Earth Radiation Budget Experiment, 134

- Earth Resource Satellite, 101
- Earth Resources Technology Satellite, 74, 130
- Earth Terrain Camera, 69, 79, 128
- Echo-Ballon-Satellit, 32, 146
- edge enhancement, 199
- Effektivtemperatur, 94, 121
- Effizienz, 167
- El Niño, 122, 123
- Elementarzeichenbedarf, 167
- Ellipsenbahn
 - Gesamtenergie, 19
- Emission, 106
- Emissionsvermögen
 - mittleres, 93
 - spektrales, 93
- Empfänger-Entropie, 174
- Energieintegral, 15
- Enhanced Thematic Mapper+, 86, 131
- Entropie, 152
 - Übertragungskanal, 174
 - Empfänger, 174
 - Sender, 174
- Environmental Satellite, 134
- Environmental Science and Services Administration, 132
- EnviSat, 64, 65, 109, 130, 134
- EOS, 141
- Epoche, 21
- ERBE, 134
- Erderkundungssatellit, 22
- Ergiebigkeit, 93
- ergodisch, 158, 159
- EROS, 1
- ERS-1, 101, 102
- erste kosmische Geschwindigkeit, 13
- ERTS-1, 74, 130
- ESSA, 67, 132
- ETM+, 86, 131
- Exosphäre, 63
- Explorer 7, 38
- Explorer-6, 67
- Exzentrizität, 17, 18, 38
- Falschfarbenbild, 68
 - Beispiel, 72
 - Farbzuordnung, 72
 - Vegetation, 72
- Falschfarbenfilm, 194
- Fano-Codierung, 169, 170
- Farbe des Meeres, 124, 125
- Farbfilm, 71
- Farbmischprojektor, 68
- Fax, 171
- Fehlererkennung, 176–178, 180
- Fehlerkorrektur, 176, 179
- FELICS, 207
- Film
 - charakteristische Kurve, 78, 79
 - Kontrast, 78
 - Korngröße, 79
 - räuml. Auflösung, 79
 - Rückführung, 143
- Flächensatz, 12, 14, 16
- Flat Plate Radiometer, 133
- Fluchtgeschwindigkeit, 19
- Forschungssatellit, 22, 26
- FPR, 133
- fraktale Bilderzeugung, 210
- Fundamentalsatz
 - Kanalcodierung, 174
 - Quellencodierung, 170
- Galileo, 56, 57
- Gamma Ray Observatory GRO, 27
- Gauß-Filter, 192
- Gauß-Verteilung, 192
- Gauß-Wichtung, 202
- Geländeklassifikation, 206
- Gemini, 68
- Gemini XI, 2
- Geoid, 101, 102
- Geokorona, 63
- GEOS-3, 101
- Geoscience Laser Ranging System, 92
- geostationärer Satellit, 24
- Geostationary Earth Radiation Budget, 132
- GERB, 132
- Gesamtenergie
 - Ellipsenbahn, 19
- Geschwindigkeitsgewinn, 44
- GIF, 207
- Giotto, 86
- Glättung, 202
 - Gauß, 192
 - linear, 192
- Glättungsfilter, 192
- Gleichstromanteil, 207
- global change, 1, 10, 67, 134
- Global Ozone Monitoring by Occultation of Stars, 136
- Global Positioning System, 25, 137
- GLRS, 92
- Golfstrom, 98
- GOME, 3, 136
- GOMOS, 136
- GPS, 25, 27, 131, 137
 - differentielles, 26
- GPS-Satelliten, 25
- Graphics Interchange Format, 207

- Gravitationsgesetz, 13
 Gravitationsumlenkung, 54
 gravity assisted navigation, 54
 GRO, 27
 Großkreis, 18

 Halbachse, 18
 Hamming-Distanz, 178
 Hamming-Schranke, 180
 HCMM, 91, 95
 Heat Capacity Mapping Mission, 95
 Helios 1, 29
 Heterosphäre, 63
 High Resolution Visible, 85, 86
 High-Resolution Imaging Spectrometer, 80, 91
 Hill's Zero-Velocity Curves, 51
 HIRIS, 80, 91
 Hochatmosphäre
 Dichte, 33
 Dichtevariation, 33
 Reibung, 33
 Temperaturmessung, 33
 Hohmann-Übergänge, 47
 Hohmann-Bahn, 40, 45
 Hohmann-Bahnen, 45
 Hohmann-Ellipse, 46, 47
 Homosphäre, 63
 HRIS, 91
 HRV, 85, 86
 HST, 27
<http://nsidc.org/daac/amr/>, 92
 Hubble Space Telescope HST, 26, 27
 Huffman, 172
 Huffman-Codierung, 169, 171, 207
 Hydrometeor, 111
 Hyperzeichen, 148, 160

 IFOV, 76
 Ignosphäre, 62
 Ikonos, 1, 2, 79, 80, 86
 Imperial Valley, 205, 206
 Improved TIROS Operational System, 133
 Information, 145, 148
 Definition, 148
 Informationsaufnahme
 Mensch, 160
 Informationsfluss, 174
 Informationsgehalt, 149, 150, 166
 mittlerer, 152
 Informationskanal, 149
 Informationsquelle, 149
 Informationsverdichtung, 162
 Infrarotfarbfilm, 71
 Schichtstruktur, 71
 spektrale Empfindlichkeit, 73
 Infrarotfilm, 70
 Infrarotscanner, 94
 Inklination, 20, 22, 38, 39
 Instantaneous Field of View, 76
 Intelsat V, 185
 Intermediate and Thermal Infrared Radiometer, 91
 Ionosphäre, 63
 Irrelevanz, 174
 IRS, 1
 ISEE-1,2, 26
 ISEE-3, 26, 49, 53
 ISS, 58, 60
 ITOS, 133

 Jacobi-Konstante, 51, 53
 Joint Photographic Experts Group, 207
 JPEG, 207, 208

 Kamera
 mehrinsige, 68
 Kanalcodierung, 145, 147, 166, 173, 177
 Fundamentalsatz, 174
 Kanalkapazität, 174, 175
 maximale, 174
 Kantenanreicherung, 199, 201
 gerichtete, 199
 ungerichtete, 199
 Kegelschnitte, 17
 Kehrbesentechnik, 84, 85
 Kepler'sche Gesetze, 12, 13
 Kernel, 189
 Kantenanreicherung, 199
 Keyhole, 67
 KH, 67
 KH-11, 22, 27
 KH-12, 27
 Kirchoff-Planck'sches Gesetz, 93
 Klassifikation
 automatische, 202
 multitemporal, 204
 unitemporal, 204
 Klassifikationsdiagramm, 205
 Kommunikation, 145, 147
 Kommunikationsproblem
 Effektivitätsproblem, 147
 semantisch, 147
 technisch, 147
 Kommunikationssatellit, 22, 24, 146
 Kommunikationssystem, 147
 Kommunikationswirkungsgrad, 176
 Kontrast, 78
 Kontrastanreicherung, 81, 195–198, 200
 Korrekturfähigkeit, 178, 180

- maximale, 179
- Kosinustransformation, 207
- kosmische Geschwindigkeit
 - erste, 13, 18, 19, 42, 58
 - zweite, 19, 58
- Kosmos 1042, 66
- Kosmos 1402, 101
- Kosmos 954, 66, 101
- Kreisbahngeschwindigkeit, 13, 18, 58

- LAGEOS, 27
- Lagrange-Punkt, 32, 50, 51
- Lander, 49
- LandSat, 22, 23, 27, 60, 67, 76, 79, 81, 83, 85–87, 95, 96, 119, 130, 131, 194, 205, 206
- LandSat 1, 74, 82, 130
- LandSat 1-3, 82
- LandSat 2, 130
- LandSat 4, 130
- LandSat 6, 131
- LandSat 7, 1, 131
- LandSat 3, 82
- Laplace-Integral, 16
- Laplace-Transformation, 199
- Laplace-Vektor, 16
- Laser Atmospheric Wind Sounder, 92
- Laser Retro-Reflector, 135
- Laufhängencodierung, 172, 207
- LAWS, 92
- Lebensdauer Photosatellit, 144
- Leistungsverstärkung, 184
- Lempel–Ziv-Algorithmus, 171, 172, 207
- Librationspunkt, 49
- Limb-Viewing, 106
- logarithmus dualis, 150
- LRR, 135
- Luftschauer, 64

- Magellan, 49, 106
- Manlifter, 6
- MAPS, 109
- Mariner, 54, 83
- Mariner 10, 56, 83
- Markoff-Prozess, 158
- Massenverhältnis
 - totales, 45
- Maximum Likelihood Method, 203
- Medium Resolution Imaging Spectrometer, 91, 134
- Mehraufwand je Zeichen, 167
- Mehrstufenrakete, 43
- MERIS, 91, 134, 135
- Mesosphäre, 62
- Meteor, 67
- MeteoSat, 6, 7, 110, 131, 204
- MeteoSat 8, 131
- MeteoSat Second Generation, 132
- Meteosat Visible and Infrared Radiometer, 132
- Michelson Interferometer for Passive Atmospheric Sounding, 135
- Microwave Radiometer, 135
- Mikrowellenradiometer, 99
- MIPAS, 135
- MIR, 27
- Mischpixel, 89, 116, 189, 205
- Mission to Planet Earth, 90, 109, 138, 141
- Mixed Pixels, 89, 116, 189, 205
- MKF-6, 67
- Moderate-Resolution Imaging Spectrometer, 91
- MODIS, 91
- MODIS-N, 91
- MODIS-T, 91
- Molniya, 19, 21, 24, 25, 28, 39
- Molniya-1, 67
- Molnyia, 59
- Momentintegral, 16
- MOMS, 84, 85
- MSG, 132
- MSR, 132
- MSRS, 137
- MSS, 67, 74, 81, 86, 87, 119, 130, 131, 194, 205, 206
- Multi-Kanal-Fotografie-6, 67
- Multi-Spectral High Resolution System, 137
- Multispectral Radiometer, 132
- Multispektralkamera, 68, 69
- Multispektralscanner, 86, 87
- Mustererkennung, 161, 164
- MVIR, 132
- MWR, 135

- Nachricht, 148
- Nachrichtenempfänger, 149
- Nachrichtenquelle, 147, 149
- Nachrichtensenke, 147
- Nachrichtenstrecke, 146
- nachtleuchtende Wolken, 62
- NASA-Hasselblad, 69
- NAVIGATIONS System with Time And Ranging, 138
- Navigationssatellit, 22, 25
- NAVSTAR, 25, 27, 138
- NDVI, 115, 116
- Newton, 13
- Nimbus, 22, 27
- Nimbus-7, 95, 99, 100, 107, 124
- NOAA, 3, 6, 8, 90, 116, 130, 133

- NOAA-16, 134
 NOAA-2, 133
 normalized difference vegetation index, 115
- odenauflösung
 SAR, 105
- Okkultation, 106
 Okkultationsinstrument, 108
 optische Tiefe, 110
 Ozonschicht, 62
- panchromatisch, 70
 Panoramakamera, 68
 Paritätsbit, 176–178, 180
 passive Stabilisierung, 28, 29
 Perigäum, 18
 Perihel, 18
 Perizentrum, 18
 Pflanzengesundheit, 74
 Pioneer, 54
 Pixel, 76, 171, 188
 Planck'sches Strahlungsgesetz, 93
 PNG, 207
 POES, 3
 Poisson-Rauschen, 191, 192
 Poisson-Statistik, 191
 Polarisationsunterschiede, 99, 100
 Polarlicht, 62
 Potentialfunktion, 50
 Präfix-Code, 167–169
 Präzession, 38
 Prüfwahl, 177
 Prüfzeichen, 180
 Prewitt-Operatoren, 202
 Produktcode, 180
 progressive Datenübertragung, 209
 progressive transmission, 209
 Pulsation, 81
- Quader-Verfahren, 203
 Quaderverfahren, 203
 detailliertes, 203
 Quellencodierung, 145, 147, 165, 167
 Definition, 166
 Fundamentalsatz, 170
 Quellenstatistik, 171
 Quellfunktion, 93
 QuickBird, 1
- RA-2, 135
 Radar, 104
 Radar-Altimeter-2, 135
 radiometrisches Auflösungsvermögen, 66
 Rakete
 Dreistufenr., 43
 Einstufenr., 42, 43
 Mehrstufenr., 43
 mehrstufige, 42
 Wirkungsgrad, 41
 Zweistufenr., 43
- Raketenformel, 41, 42, 45
 Raketengleichung, 40, 42
 raster transmission, 209
 Rasterübertragung, 209
 Raumdämpfung, 183, 184
 Raumsonde, 26
 Rauschen, 175, 191
 Rauschleistungsdichte, 175
 RBV, 82, 83, 86, 87, 130
 Redundanz, 152, 154, 166, 180
 Code, 167
 relative, 154
 Redundanzreduktion, 167, 168
 Regendämpfung, 184
 Reibung
 Hochatmosphäre, 33
 Rektaszension, 20
 relative Redundanz, 154
 remote sensing, 4
 retrograd, 38
 Return Beam Vidicon, 82
 Richtantenne, 181, 182
 RoSat, 27
 run length, 172
- Salton Sea, 205, 206
 SAM, 107, 108, 110
 SAR, 104, 106
 Charakteristika, 107
 reflektiertes Signal, 106
- Satellit
 Erderkundungss., 22
 Forschungs, 22
 Kommunikationss., 22
 Navigationss., 22
 Wetters., 22
- Satellitenbahn
 Aufklärungss., 22
 Erderkundungss., 22
 Forschungss., 26
 Kommunikations., 24
 Kommunikationss., 24
 Navigationss., 25
 Wetters., 24
- Scanning Multi-Channel Radiometer, 99, 100
 Scanning Multichannel Microwave Radiometer, 99
 Scanning Radiometer, 132, 133
 Scatterometer, 104

- Schubkraft, 41
schwarzer Körper, 93
Schwarzweißfilm, 69
 panchromatisch, 70
SCIAMACHY, 109, 136
Sea Floor Mapping, 103
SeaSat, 22, 27, 101–103
SEC-Vidicon, 82
SEM, 3, 22, 134
Sender-Entropie, 174
Sensorklassen, 66
SEVIRIO, 132
Shannon Codierung, 168, 169
Shannon'sches Sampling Theorem, 159
Shannon-Funktion, 152, 175
Shannon-Grenze, 176
Sierpinsky-Dreieck, 210
Signal, 148
Signal-Rausch-Verhältnis, 173, 175, 181, 182
Signalverdoppelung, 177
SkyLab, 36, 68, 69, 79, 101
SMM, 26, 36, 37
SMMR, 99, 100
Sobel-Operator, 202
SOHO, 26
Sojus, 58
Solar A, 27
Solar Maximum Mission, 26, 36
Solar Proton Monitor, 133
SolarMax, 36
Solwind P78-1, 26
sonnensynchron, 21
sonnensynchrone Bahn, 29, 39
Sonder, 106
Space Environment Monitor, 134
space loss, 183
Space-Shuttle, 68
Spacelab, 68, 69
spektrale Empfindlichkeit
 Farbfilm, 73
 Infrarotarbfilm, 73
 MSS, 87
 TM, 87, 88
Spiegelobjektiv, 69, 78
Spinning Enhanced Visible and Infrared
 Imager, 132
Spinstabilisierung, 28, 29
SPM, 133
SPOT, 22, 27, 60, 79, 81, 85, 86
SPOT 5, 1, 79, 80
Spurengase
 treibhauswirksame, 108
Sputnik 1, 6
Sputnik 2, 35, 38
SR, 132, 133
Störbeschleunigung, 33
Störungsintensität, 178
Stabilisierung
 Drei-Achsen-S., 28, 29
 passive, 28, 29
 Spin, 28, 29
 Spins., 29
Startfenster, 49
Stefan-Boltzmann Gesetz, 93
Strahlungsgesetze, 93
Strahlungstemperatur, 121
Stratopause, 62
Stratospheric Aerosol Measurement, 107,
 108
Streifenkamera, 68, 85
Streuung, 73, 106
 Atmosphäre, 192, 193
SWH, 104
Swingby, 54, 56
Symbolstatistik, 167
Syncom, 29
Synthetic Aperture Radar, 104
Szintillation, 81

TDRS, 146
Telekonnektion, 123
Telemetrie-/Telekommando-Strecke, 146
Television Infrared Observation Satellite,
 132
Temperatur
 Hochatmosphäre, 33
Temperature Humidity Infrared Radiometer,
 95
Thematic Mapper, 85–87, 95
thermisches Auflösungsvermögen, 66
Thermosphäre, 62
THIR, 95
TIR, 91, 95
TIROS, 6, 67, 132, 133
TIROS I, 29
TIROS Operational Satellite, 132
TIROS Operational Vertical Sounder, 134
TIROS-N, 90, 134
TM, 81, 86–88, 131, 194
 Einzelbilder, 89
TOMS, 3, 109
TOS, 132
Total Ozon Mapping Spectrometer, 109
TRACER, 109
Tracking and Data Relay System, 146
Transferbahn, 40, 45
Transformations-Codierungsverfahren, 207
Transinformation, 174
treibhauswirksame Spurengase, 64

- Trigramm-Häufigkeiten, [157](#)
- Trigramm-Struktur, [158](#)
- Trojaner, [53](#)
- Tropopause, [62](#)
- Troposphäre, [61](#)
- Tropospheric Radiometer for Atmospheric Chemistry and Environmental Research, [109](#)
- Tschernobyl, [96](#)
- TV-NA, [132](#)
- TV-WA, [132](#)
- TVOS, [134](#)

- Ulysses, [56](#)
- Umlaufzeit
 - untere Grenze, [36](#)
- Uplink, [65](#)
- UV-Photographie, [70](#)

- Vanguard 2, [33](#), [34](#)
- VEEGA, [57](#)
- Vegetation
 - Flaschfarben, [72](#)
- Vegetationsindex, [115](#), [116](#)
- Venus-Earth-Earth gravity assist, [57](#)
- Verfahren der größten Wahrscheinlichkeit, [203](#)
- Very High Resolution Radiometer, [133](#)
- VHRR, [133](#)
- Videokamera, [82](#)
 - LandSat, [82](#), [83](#)
- Vidicon, [82](#)
- Viking, [83](#)
- Visible and Infrared Spin Scan Radiometer, [132](#)
- VISSR, [132](#)
- Volkszählung, [72](#)
- Vorwärtskompensation, [81](#)
- Voyager, [54](#), [83](#)

- Wüstenbildung, [100](#)
- Wasserdampf, [61](#)
- Wechselstromanteil, [207](#)
- Weitschweifigkeit, [154](#)
- Wellenhöhe
 - signifikante, [104](#)
- Wettersatellit, [22](#), [24](#)
- Wetterschicht, [61](#)
- Widerstandsbeiwert, [34](#)
- Wien'sches Verschiebungsgesetz, [94](#)
- WIND, [26](#)
- Winkelauflösungsvermögen, [77](#), [78](#)
- Wirkungsgrad
 - äußerer, [41](#), [42](#)
 - innerer, [41](#)
- Yokoh, [27](#)
- Zeichen, [148](#)
- Zeichenhierarchie, [148](#)
- Zeichenkette, [170](#)
- Zeiss-Kamera, [69](#)
- zeitliches Auflösungsvermögen, [66](#)
- Zenit, [67](#)
- ZIP, [171](#), [207](#)
- ZOO, [171](#)