

Reinhard Weiß

Ausgewählte Themen und Herleitungen
aus dem Grundstudium

04.05.2024

Inhaltsverzeichnis

1	Spickzettel Fehlerrechnung	5
2	Differentialbegriff und Infinitesimalrechnung	10
	2.1 Differentialrechnung	10
	2.2 Integralrechnung	11
	2.3 Beispiele	13
3	Totales Differential und thermodynamisches Potential, totale Ableitung	15
4	Lösung der Integrale $\int_0^{\infty} x^q \cdot e^{-ax^2} dx$, $q \in \mathbb{N}$, $a \in \mathbb{R}^{>0}$	19
	4.1 Für ungerade q	19
	4.2 Für $q = 0$	21
	4.3 Für gerade q	21
5	Das Wichtigste zu reellen Matrizen	24
	5.1 Assoziativgesetz für die Matrizenmultiplikation	24
	5.2 Die Transponierte einer Matrix	25
	5.3 Multiplikation transponierter Matrizen	25
	5.4 Quadratische Matrizen	26
	5.4.1 Invertieren von Matrizen	28
	5.4.2 Orthogonale Matrizen – Drehmatrizen	30
6	Polar-, Zylinder- und Kugelkoordinaten auf einen Blick	32
	6.1 Polarkoordinaten (r, φ)	32
	6.2 Zylinderkoordinaten (r, φ, z)	34
	6.3 Kugelkoordinaten (r, ϑ, φ)	36
7	Linien-, Flächen-, Volumenelement	38
	7.1 Linienelement	38
	7.1.1 Linienelement in kartesischen Koordinaten	38
	7.1.2 Linienelement in Polarkoordinaten	39
	7.2 Flächenelement	43
	7.3 Volumenelement	44
8	Green'sche Identitäten	48
9	Differentialoperatoren in krummlinig-orthogonalen Koordinatensystemen	49
	9.1 Gradient	49
	9.2 Divergenz	50
	9.3 Laplace-Operator	52
	9.4 Rotation	53
10	Konservatives Vektorfeld, Potentialfunktion, Gravitationsfeld	55
11	Der Gay-Lussac'sche Versuch	59
	11.1 Einleitung	59
	11.2 Versuchsbeschreibung	62
	11.3 Der Gay-Lussac'sche Versuch irreversibel in zwei reversibel ge- führten Schritten	64
	11.4 Der Gay-Lussac'sche Versuch reversibel gestaltet in einem Schritt	67

11.5	Schlussfolgerungen	68
12	Kirchhoff'sches Strahlungsgesetz	69
13	Green'sche Funktion	71
13.1	Lösung der Poisson-Gleichung als Standardbeispiel aus der Elek- trostatik	73
14	Zur Ausbreitungsrichtung ebener Wellen	75
15	Superposition ebener Wellen	77
15.1	Überlagerung zweier harmonischer ebener Wellen gleicher Frequenz	77
15.2	Überlagerung zweier Wellenzüge verschiedener Amplitude und verschiedener Frequenz	80
16	Brechungsindex und das Verhalten von Licht an Grenzflächen	83
17	Intensität elektromagnetischer Wellen	84
18	Evaneszentes Feld – Goos-Hänchen-Verschiebung	87
19	Wellenoptik	94
19.1	Einleitung	94
19.2	Fraunhoferbeugung	96
19.2.1	Elementare Herleitung im Reellen	96
19.3	Fresnel'sche Zonen	105
20	Zur (skalaren) Fresnel-Kirchhoff'schen Beugungstheorie	113
20.1	Vorbereitung – Helmholtz-Gleichung	113
20.2	Die Kirchhoff'sche Formel	116
20.3	Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals	120
20.4	Der Neigungsfaktor K	123
20.5	Zur auslaufenden Welle im Unendlichen	124
20.6	Anhang: Fresnel- und Fraunhofer-Beugung	128
21	Optische Abbildung in Fresnel- und Fraunhofer-Näherung	133
21.1	Fresnel-Näherung	134
21.2	Bedingung für die Fresnel-Näherung	136
21.3	Fraunhofer-Näherung	137
21.4	Bedingung für die Fraunhofer-Näherung	138
22	Fraunhofer-Beugung an kreisrunder Öffnung – Auflösungsvermögen	139
22.1	Bestimmung des Auflösungsvermögens bei kreisrunder Öffnung	145
23	Das laterale Auflösungsvermögen des Mikroskops	147
23.1	Laterales Auflösungsvermögen des Mikroskops mit dem Rayleigh- Kriterium	147
23.2	Laterales Auflösungsvermögen des Mikroskops mit dem Abbe- Kriterium	149
23.2.1	Senkrecht einfallendes Licht	150
23.2.2	Schräg einfallendes Licht	152
24	Zur Abbe'schen Abbildungstheorie	155

Literaturtip!

Zumindest für das Grundstudium lohnt es sich im Internet zu recherchieren unter:
WMI – Lecture Notes wmi.badw.de

1 Spickzettel Fehlerrechnung

Quellen:

- Horst Kuchling, Taschenbuch der Physik, 16. Auflage, Fachbuchverlag Leipzig im Carl Hanser Verlag, München, Wien, 1999, Teil F, Fehlerrechnung, Seite 594 bis Seite 603.
- I. Rückmann, S. Glüge, C. Windzio, Hinweise zum Praktikum und zur Auswertung von Messergebnissen, Fachbereich Physik Universität Bremen, 2011, https://www.uni-bremen.de/fileadmin/user_upload/fachbereiche/fb1/fb1/Physika/Regeln-und-Infos/hyper_skript_26_10_11.pdf

Dieser Spickzettel enthält nur das Wichtigste aus der für das physikalische Grundpraktikum erforderlichen Fehlerrechnung. Einige im Folgenden verwendete Größen sind:

Messgrößen x, y, \dots ; Anzahl n der Messwerte x_i, y_i, \dots ; Funktion $F(x, y, \dots)$.

Messfehler sind Abweichungen der Messwerte einer Messgröße von ihrem wahren Wert. Statt des wahren Wertes verwendet man in der Praxis den Mittelwert oder den bezüglich systematischer Fehler korrigierten Mittelwert, weil der wahre Wert der Mittelwert aus unendlich vielen Messungen wäre und somit ein ideeller Wert ist. Die Messunsicherheit u hängt ab einerseits von statistischen (zufälligen) Messfehlern und andererseits von den leider nie ganz vermeidbaren systematischen Fehlern. Messfehler bzw. Messunsicherheiten, d. h. im Folgenden die Abweichungen vom Mittelwert, können als Absolut- und als Relativfehler angegeben werden.

(1) Darstellung des Messergebnisses einer Messgröße:

Messergebnis = korrigierter Mittelwert \pm Messunsicherheit ,

d. h. Messergebnis für eine Messgröße $x = \bar{x}_c \pm u_x$.

Beispiel: Messgröße $x =$ Volumen V :

Messergebnis $V =$ korrigierter Mittelwert $\bar{V}_c \pm$ Absolutfehler

z. B. $V = 1,762 \text{ m}^3 \pm 0,036 \text{ m}^3$ bzw. $(1,762 \pm 0,036) \text{ m}^3$.

Messergebnis $V =$ korrigierter Mittelwert $\bar{V}_c \pm$ Relativfehler

z. B. $V = 1,762 \text{ m}^3 (1 \pm 2,1\%)$ bzw. leider üblich aber nicht korrekt $1,762 \text{ m}^3 \pm 2,1\%$.

(2) (arithmetischer) Mittelwert, wahrscheinlichster Wert:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i .$$

(3) Standardabweichung, mittlerer Fehler der Einzelmessung:

$$s = \pm \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \Delta x .$$

68,3% der Messwerte liegen im Intervall $[\bar{x} - s, \bar{x} + s]$.

(4) Vertrauensbereich, mittlerer Fehler des Mittelwertes:

$$\bar{s} = \pm \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} = \Delta \bar{x} , \quad n \stackrel{!}{\geq} 6 .$$

(5) Varianz (Streuung):

$$s^2 = \pm \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 .$$

(6) Grobe Fehler entstehen z. B. durch Irrtümer oder gröbere, störende äußere Einflüsse auf die Messung. Sie sind nicht reproduzierbar und werden nicht berücksichtigt.

(7) **Systematische Fehler** können z. B. Eichfehler oder Beobachtungsfehler sein. Sie sind reproduzierbar und allgemein auch quantifizierbar und sollten möglichst behoben werden. Andernfalls müssen sie im Messergebnis als Korrektur berücksichtigt und ggf. auch diskutiert werden. Aus der bezüglich einer Messgröße x immer existierenden Messunsicherheit durch technisch nicht vermeidbare Ungenauigkeiten der **Messinstrumente** resultiert der **systematische Restfehler** Δx_s .

(8) **Zufällige Fehler** sind statistische Fehler und werden gemäß der Fehlerrechnung angegeben.

(9) Der **absolute Fehler** eines Messwertes x_i ist in der Praxis der Unterschied Δx_i zwischen dem Messwert x_i und dem Mittelwert \bar{x} , also

$$\Delta x_i = x_i - \bar{x}.$$

Er hat die Dimension der zugehörigen Messgröße x .

(10) Für den **relativen Fehler** eines Messwertes x_i gilt in der Praxis

$$\text{relativer Fehler} = \frac{\text{absoluter Fehler}}{\text{Mittelwert}} = \frac{\Delta x_i}{\bar{x}}.$$

Er wird meistens als prozentualer Fehler, also in % angegeben.

(11) Der **Maximalfehler** (Größtfehler) Δx_{\max} einer Messreihe mit den Messwerte x_i ist

$$\Delta x_{\max} = \pm \max(|x_i - \bar{x}|).$$

(12) **Prozedere für die Ermittlung eines Messergebnisses:**

1. Messung von n einzelnen Messwerten x_i .
2. Berechnung des Mittelwertes \bar{x} auf eine Stelle mehr als die Messwerte aufweisen.
3. Korrektur des Mittelwertes \bar{x} durch Berücksichtigung berechenbarer systematischer Fehler bzw. Abweichungen (errors) $e_c = \text{const}$ gemäß

$$\bar{x} + e_c = \bar{x}_c.$$

Ob man den korrigierten Mittelwert \bar{x}_c oder den „ursprünglichen“ Mittelwert \bar{x} verwendet, ändert prinzipiell nichts am Formalismus. Deshalb wird oft in den Formeln vereinfachend für \bar{x}_c nur \bar{x} geschrieben. Ob \bar{x} für die Fehlerrechnung ausreicht oder ob \bar{x} zu \bar{x}_c korrigiert werden muss, ist dann jeweils zu entscheiden.

Weil die letzte Stelle des Messergebnisses unter dem Einfluss der Messunsicherheit steht, wird der Mittelwert \bar{x} (bzw. der korrigierte Mittelwert \bar{x}_c) nicht genauer angegeben, als es die Messunsicherheit zulässt.

4. Bestimmung des durch das Messgerät bedingten systematischen Restfehlers Δx_s .

5. Berechnung des zufälligen Fehlers

$$\begin{aligned} \text{für } n \geq 6 & \Rightarrow \text{Vertrauensbereich } \bar{s} = \Delta \bar{x}, \\ \text{für } n < 6 & \Rightarrow \text{Maximalfehler } \Delta x_{\max}. \end{aligned}$$

6. Berechnung der Messunsicherheit u_x :

$$\text{für } n \geq 6 \quad \text{mit Vertrauensbereich } \bar{s} = \Delta \bar{x} \text{ und } \bar{x}_c \Rightarrow u_x = \sqrt{(\Delta \bar{x})^2 + (\Delta x_s)^2},$$

$$\text{für } 1 < n < 6 \quad \text{mit Maximalfehler } \Delta x_{\max} \text{ und } \bar{x}_c \Rightarrow u_x = |\Delta x_{\max}| + |\Delta x_s|,$$

$$\text{für } n = 1 \quad \text{mit } x_c \text{ statt } \bar{x}_c \Rightarrow u_x = |\Delta x_s|.$$

Dabei wurde berücksichtigt, dass $\sqrt{a^2 + b^2} < |a| + |b|$ und dass man im Fall $n = 1$ keinen Mittelwerte \bar{x} bilden und deshalb, wenn möglich, nur den *einen* vorliegenden Messwert x zu x_c korrigieren kann.

7. Darstellung des Messergebnisses der Messgröße x wie unter (1):

$$x = \bar{x}_c \pm u_x.$$

- (13) **Mittlerer Fehler** bzw. Messunsicherheit u_F **des Funktionswertes** $F(\bar{x}, \bar{y}, \dots) = \bar{F}$, wobei die Messgrößen x, y, \dots voneinander unabhängig seien und in Messreihen mit $n \geq 6$ gemessen wurden:

\bar{F} ist nicht der Mittelwert sondern im Allgemeinen der wahrscheinlichste Wert der Funktion $F(x, y, \dots)$ an der Stelle $(x = \bar{x}, y = \bar{y}, \dots)$ mit den dazugehörigen Messunsicherheiten u_x, u_y, \dots . Folglich ist u_F die Messunsicherheit bzw. der mittlere Fehler von \bar{F} infolge der Messunsicherheiten u_x, u_y, \dots der Einzelmessgrößen x, y, \dots an der Stelle $(x = \bar{x}, y = \bar{y}, \dots)$.

Gauß'sches Fehlerfortpflanzungsgesetz :

$$\text{Absolutfehler } u_F = \pm \sqrt{\left(\left.\frac{\partial F}{\partial x}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_x\right)^2 + \left(\left.\frac{\partial F}{\partial y}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_y\right)^2 + \dots},$$

$$\text{Relativfehler } \frac{u_F}{\bar{F}} = \pm \sqrt{\left(\left.\frac{\partial F/\partial x}{F}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_x\right)^2 + \left(\left.\frac{\partial F/\partial y}{F}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_y\right)^2 + \dots}.$$

- (14) **Größtfehlerabschätzung**, d. h. Bestimmung des Maximalfehlers oder der maximalen Messunsicherheit u_F **von Funktionswerten** $F(\bar{x}, \bar{y}, \dots) = \bar{F}$, wobei die Messgrößen x, y, \dots voneinander unabhängig seien:

$$\text{Absolutfehler } u_F \approx \pm \left\{ \left| \left.\frac{\partial F}{\partial x}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_x \right| + \left| \left.\frac{\partial F}{\partial y}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_y \right| + \dots \right\}, \quad (1.1)$$

$$\text{Relativfehler } \frac{u_F}{\bar{F}} \approx \pm \left\{ \left| \left.\frac{\partial F/\partial x}{F}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_x \right| + \left| \left.\frac{\partial F/\partial y}{F}\right|_{\bar{x}, \bar{y}, \dots} \cdot u_y \right| + \dots \right\}. \quad (1.2)$$

Die Größtfehlerabschätzung erfolgt bei Messreihen mit $n \geq 6$ aus Gründen der Bequemlichkeit, wenn eine strenge Fehlerbetrachtung nach dem Gauß'schen Fehlerfortpflanzungsgesetz nicht notwendig ist.

Falls wegen $n < 6$ der Messreihe einer Messgröße x kein Vertrauensbereich $\Delta\bar{x}$ berechnet werden kann, wird bei der Größtfehlerabschätzung statt $\Delta\bar{x}$ der geschätzte Maximalfehler Δx_{\max} verwendet:

$$1 < n < 6 \quad \Rightarrow \quad u_x = |\Delta x_{\max}| + |\Delta x_s|.$$

Und liegt sogar nur ein einziger Messwert von x vor, braucht man bei u_x nur den systematischen Restfehler Δx_s zu verwenden: $n = 1 \quad \Rightarrow \quad u_x = |\Delta x_s|$.

Bei der Angabe des Maximalfehlers einer Funktion F ist die **größte Fehlerbreite** bzw. Messunsicherheit $\pm u_F$ aus dem untersuchten Wertebereich der Funktion im Ergebnis anzugeben.

- (15) **Abschätzung der durch die Messgeräte bedingten Größtfehler:**

Bei Messgeräten wird die Genauigkeit in % vom Messbereich angegeben.

Beispiel: Genauigkeitsklasse 0,5 bei einem Messbereich der Stromstärke I von 250 mA heißt

$$\Delta I = \pm 1,25 \text{ mA für den gesamten Mess- und Skalenbereich.}$$

ΔI ist hier ein absoluter systematischer Größtfehler, den wir deshalb auch mit ΔI_s bezeichnen können.

Wenn keine Genauigkeitsklasse angegeben ist, gilt für Messgeräte mit Analoganzeige und nicht zu großem Skalenstrichabstand z. B. für die Messgröße x

$$\Delta x_s \hat{=} \pm 1 \text{ Skalenstrichabstand.}$$

Achtung! Bei **digitalen Messgeräten** entspricht die **Mindestgröße** des systematischen Restfehlers ± 1 der letzten angezeigten Ziffer. Für den maximalen systematischen Restfehler einer Messgröße x gilt in diesem Fall beispielsweise

$$\Delta x_s \hat{=} \pm 2 \text{ der letzten angezeigten Ziffer.}$$

(16) **Genauigkeit der Angabe von Mess- bzw. Funktionswerten und ihren Messunsicherheiten:**

Die Ergebnisse einer Rechnung dürfen nicht mit mehr Stellen (unabhängig vom Komma) angegeben werden als die Messwerte.

Messunsicherheiten (Fehler) werden 1-stellig bis maximal 2-stellig (unabhängig vom Komma) berechnet und **immer aufgerundet**.

Messwerte bzw. Funktionswerte werden so gerundet, dass sie genauso viele Stellen nach dem Komma besitzen wie ihre Messunsicherheiten (Fehler).

(17) **Mittelwert und Runden:** Beispiel: $x_1 = 20,0$, $x_2 = 20,5$, $\bar{x} = 20,25$.

Bei Mittelwerten sind die Angabe einer zusätzlichen Dezimalstelle und auch das Weiterrechnen mit dieser zusätzlichen Dezimalstelle notwendig.

(18) **Quadrate und Dezimalstellen:** Beispiel: $x = 0,02$, $x^2 = 0,0004$.

Die Angabe der beim Quadrieren zusätzlich entstandenen Dezimalstellen und das Weiterrechnen mit diesen zusätzlichen Dezimalstellen ist notwendig.

(19) Um die Messunsicherheit (den Fehler) z. B. bei der Zeitmessung von periodischen Vorgängen zu minimieren, lässt sich die zeitliche Periode T aus der **Messung ihrer Vielfachen** nT ermitteln.

Es handelt sich bei der Ermittlung von T aus $n \cdot T$ mit $T = \frac{n \cdot T}{n}$ um ein anderes Versuchsmodell als bei der direkten Ermittlung von $1 \cdot T$. Deshalb bleibt der Umstand der Messung von Vielfachen bei der statistischen Fehlerermittlung unberücksichtigt.

(20) Ausgleichsgeraden (**Regressionsgeraden**) werden geschätzt. Dabei werden grobe Fehler vernachlässigt.

(21) **Bestimmung des Steigungsfehlers einer Geraden:**

Durch den Schwerpunkt der Regressionsgeraden werden zwei Geraden gelegt, die eine so durch die Enden der Fehlerbalken, dass sich eine minimale Steigung m_{\min} ergibt, die andere derart durch die Enden der Fehlerbalken, dass sich eine maximale Steigung m_{\max} ergibt. Die sich daraus ergebende Steigungsdifferenz liefert die **Steigungsfehlerbreite**

$$\pm \Delta m = \pm \frac{m_{\max} - m_{\min}}{2} .$$

Aus (1.1) bzw. (1.2) hergeleitete einfache Formeln zur Größtfehlerabschätzung von Funktionswerten

• **Additive Funktionen:**

$$F(x, y) = ax \pm by, \quad a, b = \text{const}$$

$$\Rightarrow u_F = \pm \left\{ |au_x| + |bu_y| \right\}$$

Die mit den Konstanten multiplizierten absoluten Messunsicherheiten der Variablen werden addiert.

• **Multiplikative Funktionen:**

$$F(x, y, z) = ax \cdot y \cdot \frac{b}{z}, \quad a, b = \text{const}$$

$$\Rightarrow \frac{u_F}{F} = \pm \left\{ \left| \frac{u_x}{x} \right| + \left| \frac{u_y}{y} \right| + \left| \frac{u_z}{z} \right| \right\}$$

Die relativen Messunsicherheiten werden addiert.

• **Multiplikative Funktionen mit Potenzen:** $F(x, y, z) = ax^l \cdot y^m \cdot \frac{b}{z^n}, \quad a, b = \text{const}$

$$\Rightarrow \frac{u_F}{F} = \pm \left\{ \left| l \frac{u_x}{x} \right| + \left| m \frac{u_y}{y} \right| + \left| n \frac{u_z}{z} \right| \right\}$$

Die nach den Potenzen gewichteten relativen Messunsicherheiten werden addiert.

Beispiele:

$$\bullet F(x, y) = ax + by \Rightarrow u_F = \pm \left\{ |au_x| + |bu_y| \right\} \Rightarrow \frac{u_F}{F} = \pm \left\{ \left| \frac{a}{ax + by} u_x \right| + \left| \frac{b}{ax + by} u_y \right| \right\},$$

$$\bullet F(x, y, z) = ax^2 \cdot y^3 \cdot \frac{b}{z^4}$$

$$\Rightarrow u_F = \pm \left\{ \left| 2ax \cdot y^3 \cdot \frac{b}{z^4} \cdot u_x \right| + \left| ax^2 \cdot 3y^2 \cdot \frac{b}{z^4} \cdot u_y \right| + \left| ax^2 \cdot y^3 \cdot (-4) \frac{b}{z^5} \cdot u_z \right| \right\}$$

$$\Rightarrow \frac{u_F}{F} = \pm \left\{ \left| 2 \frac{u_x}{x} \right| + \left| 3 \frac{u_y}{y} \right| + \left| 4 \frac{u_z}{z} \right| \right\}.$$

2 Differentialbegriff und Infinitesimalrechnung

Siehe auch:

Springer-Lehrbuch von Klaus Weltner, Mathematik für Physiker 1, 12. Aufl., Springer-Verlag, Berlin, Heidelberg, New York, 2001, Kapitel 5 Differentialrechnung und Kapitel 6 Integralrechnung.

Ohne Anspruch auf mathematische Vollständigkeit wollen wir in diesem Kapitel die Begriffe Differential, Differenzenquotient und Differentialquotient anhand der Abbildung 1 und anschließend die Begriffe bestimmtes und unbestimmtes Integral skizzieren. Dabei wird deutlich werden, dass Differentiale nicht unbedingt infinitesimal kleine Größen sein müssen, auch wenn diese Betrachtungsweise sehr praktisch sein kann und deshalb in der Physik üblich ist gemäß $dx = \lim_{x \rightarrow x_0} (x - x_0) \hat{=} x - x_0 = \Delta x \rightarrow 0$.

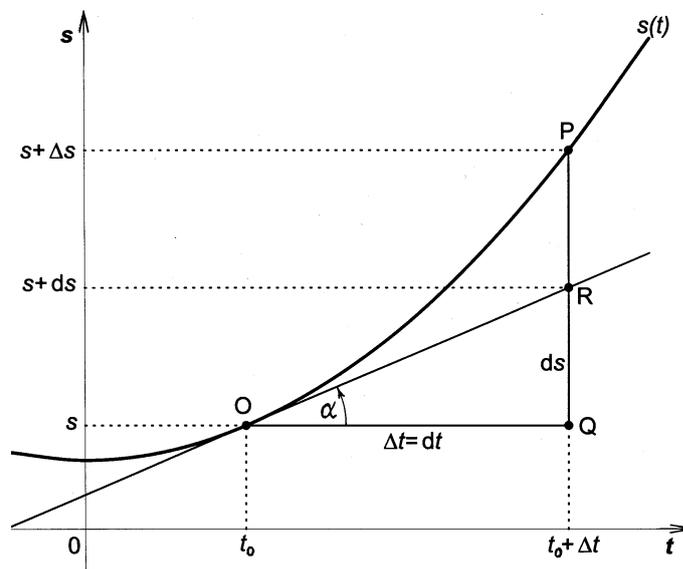


Abb. 1 Geometrische Veranschaulichung der Differentiale ds und dt sowie des Differentialquotienten $ds/dt = \tan \alpha$ im Punkt $s(t_0)$ der Funktion $s(t)$.

2.1 Differentialrechnung

Gegeben sei eine glatte Funktion $s(t)$, also eine stetige, von t abhängige Funktion ohne „Knick“. Z. B. kann die unabhängige Variable t die Zeit und die abhängige Variable s die in der Zeit zurückgelegte Weglänge sein. Während die Zeit t ausgehend von t_0 um Δt (Strecke \overline{OQ}) fortschreitet, ändert sich die Funktion $s(t)$ um $\Delta s = s(t_0 + \Delta t) - s(t_0)$ (Strecke \overline{QP}). Jetzt legen wir im Punkt $s(t_0)$ (Punkt O) eine Tangente an die Kurve $s(t)$. Die Tangente schneidet die Strecke \overline{OP} im Punkt R. Die Strecke \overline{QR} ist die geometrische Darstellung des abhängigen Differentials ds .

Merke:

Das **abhängige Differential** ds (das Differential der Funktion) ist die lineare Näherung für die Änderung Δs . Wir haben dabei das **unabhängige Differential** $dt = \Delta t$ (das Differential des Arguments, Strecke \overline{OQ}) zugrunde gelegt.

Die Näherung ist um so besser, je mehr sich der Punkt Q dem Punkt O nähert bzw. je kleiner wir das unabhängige Differential $dt = \Delta t$ wählen. Bei linearen Funktionen $s(t)$ ist die Tangente an jeden Punkt der Kurve identisch mit der Kurve selbst, so dass in diesem Fall für das abhängige Differential $ds = \Delta s$ gilt.

Der Quotient aus den Differentialen ds und dt bezüglich $s(t_0)$, also der *Differentialquotient* ds/dt an der Stelle t_0 , ist geometrisch die Steigung $\tan \alpha$ der Tangente an den Kurvenpunkt $s(t_0)$ (s. Abbildung 1). Allgemein gilt, je kleiner wir im *Differenzenquotienten* $\Delta s/\Delta t$ das Intervall Δt wählen, desto mehr nähert sich dieser dem Differentialquotienten an, so dass wir für den Differentialquotienten von $s(t)$ an der Stelle $t = t_0$ folgendes schreiben können:

$$\left. \frac{ds}{dt} \right|_{t_0} = \lim_{\Delta t \rightarrow 0} \left. \frac{\Delta s}{\Delta t} \right|_{t_0} = \lim_{\Delta t \rightarrow 0} \frac{s(t_0 + \Delta t) - s(t_0)}{\Delta t} = s'(t_0). \quad (2.1)$$

Diesen Grenzwert können wir für beliebige t bilden, so dass wir eine Funktion der Differentialquotienten von t erhalten. Diese Funktion heißt Ableitung

$$s'(t) = \frac{ds(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{s(t + \Delta t) - s(t)}{\Delta t} = v(t) \quad (2.2)$$

der Funktion $s(t)$. Die Ableitung $s'(t)$ der Funktion $s(t)$ liefert uns für jeden Zeitpunkt t die momentane Änderung des Weges s pro zugehöriger Änderung der Zeit t , also die Geschwindigkeit $s'(t) = v(t) = ds/dt$ und mit (2.2) schreiben wir für das abhängige Differential

$$ds = s'(t)dt = v(t)dt. \quad (2.3)$$

2.2 Integralrechnung

Wir wollen jetzt aus der Funktion $s'(t)$ die Änderung Δs in den Grenzen von t_0 und t berechnen. Dazu teilen wir Δt in N Abschnitte Δt_i , so dass $\sum_{i=1}^N \Delta t_i = \Delta t$. Wenn die Teilbereiche Δt_i klein im Vergleich zu Δt gewählt werden, erhält man mit je einem Funktionswert $s'(t_i)$ aus jedem Teilbereich Δt_i die Näherung

$$\sum_{i=1}^N s'(t_i) \cdot \Delta t_i = \sum_{i=1}^N \left. \frac{ds}{dt} \right|_{t_i} \cdot \Delta t_i = \sum_{i=1}^N v(t_i) \cdot \Delta t_i \approx \Delta s. \quad (2.4)$$

Je kleiner wir die Intervalle Δt_i wählen, desto größer wird ihre Anzahl N in den Grenzen von t_0 bis $t_0 + \Delta t$ und desto besser wird die Näherung. Im Grenzfalle für $\Delta t_i \rightarrow 0$ bzw. $N \rightarrow \infty$ erhalten wir aus der Näherungsgleichung (2.4) die exakte Änderung Δs :

$$\lim_{\Delta t_i \rightarrow 0} \sum_{i=1}^N s'(t_i) \cdot \Delta t_i = \lim_{\Delta t_i \rightarrow 0} \sum_{i=1}^N v(t_i) \cdot \Delta t_i = \quad (2.5)$$

$$= \int_{t_0}^{t_0 + \Delta t} s'(t) dt = \int_{t_0}^{t_0 + \Delta t} v(t) dt = \int_{s(t_0)}^{s(t_0 + \Delta t)} ds = s(t_0 + \Delta t) - s(t_0) = \Delta s. \quad (2.6)$$

Der Ausdruck

$$\int_{t_0}^{t_0 + \Delta t} s'(t) dt = \Delta s \quad (2.7)$$

heißt **bestimmtes Integral**¹. Es hat nämlich wegen seiner Integrationsgrenzen einen eindeutig bestimmbar Wert.

Wie wir sehen, ist das Integralzeichen der Befehl zur Summation über eine unendliche Anzahl von *infinitesimal kleinen Differentialen* $ds = s'(t)dt = v(t)dt$. Geometrisch ist das bestimmte Integral (2.6) der Flächeninhalt A unter dem Graphen der Funktion $s'(t)$ in den Grenzen von t_0 und $t_0 + \Delta t$ (s. Abb. 2 und Abb. 3).

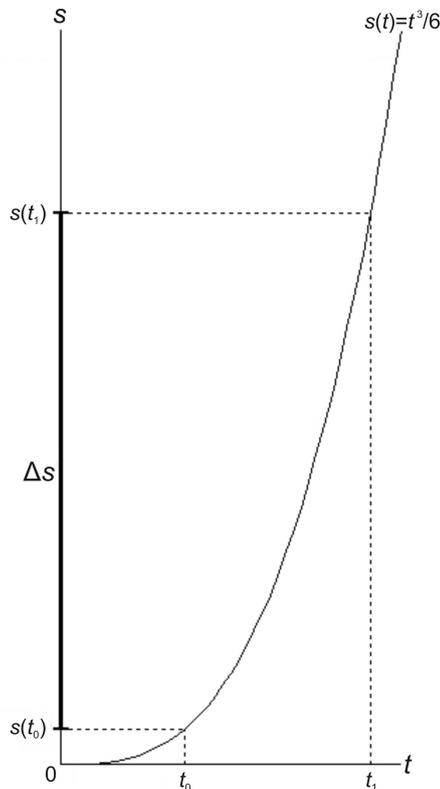


Abb. 2 Beispiel Weg-Zeit-Funktion

$$s(t) = \frac{1}{6} t^3.$$

Gemäß der Weg-Zeit-Funktion $s(t)$ ist $\Delta s = s(t_1) - s(t_0)$ der im Zeitintervall $t_1 - t_0$ zurückgelegte Weg. Der Anstieg des Graphen der Funktion $s(t)$ in Abhängigkeit von t wird beschrieben durch die Ableitung

$$s'(t) = \frac{ds(t)}{dt} = v(t).$$

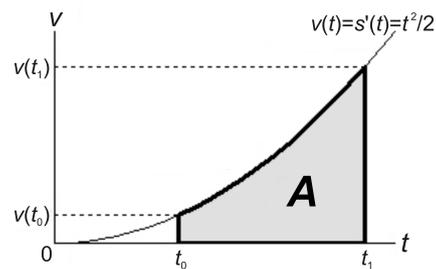


Abb. 3 Ableitung von $\frac{1}{6} t^3$ nach t :

$$s'(t) = \frac{1}{2} t^2.$$

Für den Flächeninhalt A unter dem Graphen der Ableitungsfunktion $s'(t)$ von t_0 bis t_1 gilt

$$A = \Delta s = \int_{t_0}^{t_1} \frac{ds(t)}{dt} dt = \int_{t_0}^{t_1} s'(t) dt.$$

Das **unbestimmte Integral** über $s'(t)$ hat keine Integrationsgrenzen und liefert uns die Stammfunktion zu $s'(t)$:

$$\int s'(t) dt = s(t) + C. \quad (2.8)$$

Die Integrationskonstante C hängt von den Anfangsbedingungen ab. Bilden wir die Ableitung der Stammfunktion, so verschwindet die Integrationskonstante, weil die

¹Dieses bestimmte Integral als Beispiel wird gelesen: „Integral über $s'(t) dt$ in den Grenzen von t_0 bis $t_0 + \Delta t$ gleich Δs “ oder kurz: „Integral $s'(t) dt$ von t_0 bis $t_0 + \Delta t$ gleich Δs “.

Änderung oder Ableitung einer Konstanten gleich Null ist. Wir erhalten deshalb aus (2.8) wieder

$$\frac{d}{dt}(s(t) + C) = \frac{ds(t)}{dt} + \frac{dC}{dt} = s'(t) + 0 = s'(t) . \quad (2.9)$$

Differentiation und Integration sind zueinander Umkehroperationen. Sinngemäß liefern uns die Differentiation einen „Quotienten aus Differenzen“ und die Integration eine „Summe von Produkten“.

2.3 Beispiele

Im Gegensatz zur Tatsache, dass die Integration über unendlich kleine Differentiale erfolgt, sind Differentiale allgemein keine unendlich kleinen Größen. Dessen ungeachtet ist es in der Physik üblich, Differentiale als unendlich kleine Änderungen von Größen anzusehen und mit ihnen zu rechnen, d. h., auf sie die Grundrechenarten anzuwenden. Das führt oft auf bequeme Weise zu exakten Ergebnissen, wenn man die Differenzierbarkeit bzw. die Integrierbarkeit der betrachteten Funktionen voraussetzt. Dies werde für die folgenden veranschaulichenden Beispiele angenommen.

- Gegeben sei die Weg-Zeit-Funktion $x(t) = t^2$. Die Differentiationsregel für Potenzfunktionen ist

$$\frac{d}{dx}(a \cdot x^n + C) = a \cdot nx^{n-1} . \quad (2.10)$$

Dann ist die Geschwindigkeit-Zeit-Funktion

$$v(t) = \frac{dx}{dt} = \frac{d}{dt}t^2 = 2t \quad (2.11)$$

und das abhängige Differential $dx = v(t)dt = 2t dt$.

- Gegeben sei die verkettete Funktion $v \circ x = v(x(t))$ mit $v(x) = x/2$ und $x(t) = t^2$. Die Ableitung von $v(x(t)) = t^2/2$ nach t erhalten wir durch Anwendung der Kettenregel:

$$\frac{dv(x)}{dx} \cdot \frac{dx}{dt} = \frac{1}{2} \cdot 2t = \frac{dv(x(t))}{dt} = t . \quad (2.12)$$

Man kommt direkt zum Ergebnis, wenn man dx „herauskürzt“ und für v die Funktion $v(x(t))$ einsetzt.

- Für das Geschwindigkeitsquadrat im kartesischen Koordinatensystem gilt mit (2.2)

$$v^2 = v_x^2 + v_y^2 + v_z^2 \quad (2.13)$$

$$= \left(\lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} \right)^2 + \left(\lim_{\Delta t \rightarrow 0} \frac{\Delta y}{\Delta t} \right)^2 + \left(\lim_{\Delta t \rightarrow 0} \frac{\Delta z}{\Delta t} \right)^2 \quad (2.14)$$

$$= \lim_{\Delta t \rightarrow 0} \frac{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}{(\Delta t)^2} \quad (2.15)$$

$$v^2 = \frac{(dx)^2 + (dy)^2 + (dz)^2}{(dt)^2} . \quad (2.16)$$

Es ist nicht ganz korrekt aber bequem und deshalb üblich, bei Potenzen der Koordinatendifferentiale die Klammern zu vernachlässigen. So erhalten wir aus (2.16) für das Geschwindigkeitsquadrat

$$v^2 = \frac{dx^2 + dy^2 + dz^2}{dt^2} \quad (2.17)$$

und für das Differential des räumlichen Abstandsquadrats

$$dx^2 + dy^2 + dz^2 = v^2 dt^2 . \quad (2.18)$$

- Ein Körper werde ab dem Zeitpunkt $t = 0$ mit $a(t) = dv/dt = t$ beschleunigt. Bei $t = 0$ habe der Körper die Anfangsgeschwindigkeit $C = v_0 = 4$. Welche Geschwindigkeit hat er dann bei $t_1 = 2$ und bei $t_2 = 8$?

Die Integrationsregel für Potenzfunktionen ist

$$\int a \cdot x^n dx = a \cdot \frac{1}{n+1} x^{n+1} + C . \quad (2.19)$$

Aus dem Geschwindigkeitsdifferential $dv(t) = a(t)dt$ erhalten wir

$$v(t) = \int dv(t) = \int a(t)dt = \int t dt = \frac{1}{2}t^2 + v_0 = \frac{1}{2}t^2 + 4 . \quad (2.20)$$

Die gesuchten Geschwindigkeiten sind also $v(t_1) = v_1 = 6$ und $v(t_2) = v_2 = 36$. Der Körper hat also seine Geschwindigkeit in der Zeit von t_1 bis t_2 von 6 auf 36 erhöht. Diese Geschwindigkeitsänderung um 30 liefert uns das bestimmte Integral:

$$v_2 - v_1 = \int_{v_1}^{v_2} dv = \int_{t_1=2}^{t_2=8} t dt = \left. \frac{1}{2}t^2 \right|_{t_1=2}^{t_2=8} = 32 - 2 = 30 . \quad (2.21)$$

Die Integrationskonstante $v_0 = 4$ brauchen wir beim bestimmten Integral nicht zu berücksichtigen, denn sie fällt bei der Subtraktion $v_2 - v_1 = (32 + 4) - (2 + 4)$ heraus.

3 Totales Differential und thermodynamisches Potential, totale Ableitung

Mit dem Gradienten (siehe Abschnitt 10) eines skalaren Feldes $f(\vec{r}) = f(x, y, z)$ in unmittelbarem Zusammenhang steht das vollständige oder totale Differential auf folgende Weise (totale Differenzierbarkeit von f vorausgesetzt):

Gradient $\frac{f(\vec{r} + d\vec{r}) - f(\vec{r})}{d\vec{r}} = \frac{df(\vec{r})}{d\vec{r}} = \text{grad } f(\vec{r}) := \vec{e}_x \frac{\partial f}{\partial x} + \vec{e}_y \frac{\partial f}{\partial y} + \vec{e}_z \frac{\partial f}{\partial z} \Rightarrow$

totales Differential $df(\vec{r}) = f(\vec{r} + d\vec{r}) - f(\vec{r}) = \frac{df(\vec{r})}{d\vec{r}} \cdot d\vec{r} = \text{grad } f(\vec{r}) \cdot d\vec{r}$
 $= \left(\vec{e}_x \frac{\partial f(x, y, z)}{\partial x} + \vec{e}_y \frac{\partial f(x, y, z)}{\partial y} + \vec{e}_z \frac{\partial f(x, y, z)}{\partial z} \right) \cdot (dx \vec{e}_x + dy \vec{e}_y + dz \vec{e}_z),$

$$\boxed{df(\vec{r}) = \text{grad } f(\vec{r}) \cdot d\vec{r} = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz} \quad (3.1)$$

Die Funktion, von der das totale Differential gebildet wird, muss nicht von den drei Raumkoordinaten abhängen, sondern kann von beliebig vielen Koordinaten (Größen) x_i ($i = 1, 2, 3, \dots$) abhängen. Deshalb lautet die allgemeine Definition des totalen Differentials

$$df := \sum_i \frac{\partial f}{\partial x_i} dx_i.$$

Wenn in einem Punkt \vec{r}_0 alle partiellen Ableitungen einer Funktion $f(\vec{r})$ existieren und wenn dann dort auch noch alle partiellen Ableitungen stetig sind, dann ist f in \vec{r}_0 total differenzierbar. Geometrisch bedeutet das, dass sich die total differenzierbare Funktion lokal durch eine lineare Abbildung (Tangentialebene) approximieren lässt.²

Die geometrische Bedeutung des totalen Differentials lässt sich für eine Funktion von zwei unabhängigen Variablen wie beispielsweise $f(\vec{r}) = f(x, y) = z$ graphisch veranschaulichen (siehe Abbildung 4).

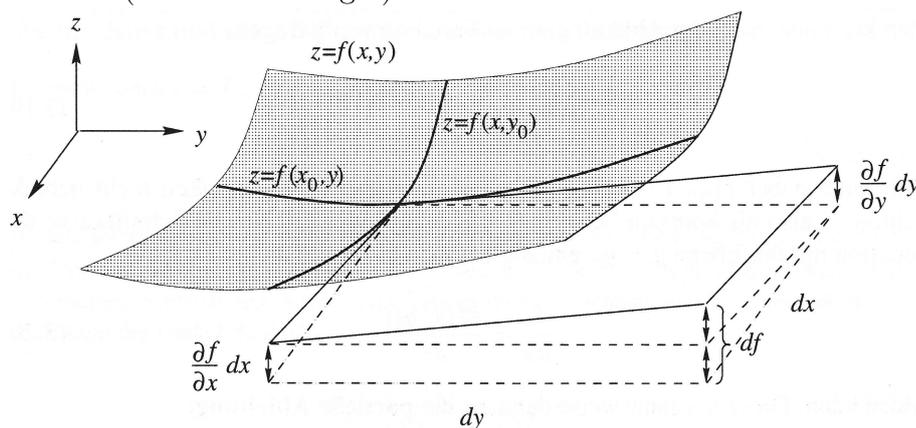


Abb. 4 Veranschaulichung der geometrischen Bedeutung des totalen Differentials der Funktion $f(x, y) = z$. Die Funktionswerte z bilden die punktierte gekrümmte Fläche über der (x, y) -Koordinatenebene. Abbildung aus dem Hochschultaschenbuch von Christian B. Lang und Norbert Pucker, Spektrum Akademischer Verlag, Heidelberg, Berlin, 1998, Seite 84.

²Genauereres dazu findet man unter dem Suchbegriff *Totale Differenzierbarkeit* – Wikipedia.

Mit $dx = x - x_0$, $dy = y - y_0$, $dz = z - z_0 = df(x_0, y_0)$ und $z_0 = f(x_0, y_0)$ liefert das totale Differential die Ebenengleichung für die Tangentialebene an den Graphen der Funktion $f(x, y)$ im Punkt (x_0, y_0) :

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \Rightarrow$$

$$df(x_0, y_0) = \left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)} dx + \left. \frac{\partial f}{\partial y} \right|_{(x_0, y_0)} dy \Rightarrow$$

$$z - z_0 = \frac{\partial f(x_0, y_0)}{\partial x} (x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y} (y - y_0) \Leftrightarrow$$

$$\text{Tangentialebene} := z = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x} (x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y} (y - y_0) .$$

Die Differentiale dx und dy bzw. die Differenzen $x - x_0 = \Delta x$ und $y - y_0 = \Delta y$ stehen hier in der ersten Potenz, beschreiben also einen linearen Zusammenhang, eine Ebene. Vergleichen wir die Tangentialebenengleichung mit der Taylor-Entwicklung von Funktionen mehrerer Variablen, so stellen wir fest, dass das totale Differential (die Ebenengleichung) eine lineare Näherung von f ist, d. h. eine Taylor-Entwicklung bis zur ersten Ordnung.

Handelt es sich bei f um eine Funktion von nur einer Variablen wie beispielsweise $f(x) = y$, so entspricht das zugehörige totale Differential einer Geraden, nämlich der Tangente an die Funktion f im Punkt $f(x_0)$ mit der Tangentengleichung

$$y = f(x) = f(x_0) + \left. \frac{df(x)}{dx} \right|_{x_0} \cdot (x - x_0) .$$

Diese Tangentengleichung ist aber gerade die Taylor-Entwicklung der Funktion $f(x)$ an der Stelle x_0 bis zum linearen Glied, also bis zur ersten Ordnung. Und wie man hier sieht, stimmen im 1-dimensionalen Fall und nur im 1-dimensionalen Fall die klassische reelle, die partielle und die totale Differenzierbarkeit überein.

Ist f jedoch von mehr als zwei Variablen abhängig, so spricht man im Zusammenhang mit der geometrischen Bedeutung des zugehörigen totalen Differential trotzdem von einer Tangentialebene. Ein **linearer** Zusammenhang entspricht immer einer Ebene bzw. **ebenen** Verhältnissen.

Eine wichtige Rolle spielen totale Differentiale in der **klassischen Thermodynamik**. Im Gegensatz zu den **Prozessgrößen** Wärmeenergie Q und Arbeit W mit ihren „infinitesimalen“ Änderungen δQ und δW besitzen die **Zustandsgrößen**³ totale Differentiale. So besitzt beispielsweise die Zustandsgröße (die Zustandsfunktion, das thermodynamische Potential) innere Energie U das totale Differential

$$dU = \left(\frac{\partial U(V, T)}{\partial V} \right)_{(T)} dV + \left(\frac{\partial U(V, T)}{\partial T} \right)_{(V)} dT .$$

³**Intensive Zustandsgrößen:** z. B. Druck p und Temperatur T .

Extensive Zustandsgrößen: z. B. Volumen V , Teilchenzahl N , Stoffmenge, Entropie und thermodynamische Potentiale.

Thermodynamische Potentiale: innere Energie U , freie Energie, Gibbs-Energie und großkanonisches Potential.

Die tiefgestellten Indizes in Klammern, also (T) und (V) , bedeuten hier, dass T bzw. V jeweils festgehalten werden. In Analogie zu (3.1) und zu den Ausführungen im Abschnitt 10 können wir $U(V, T)$ als thermodynamisches **Potential** bzw. als skalares Feld über den unabhängigen Variablen V und T ansehen. Der Gradient von U liefert demzufolge das Gradienten- bzw. Vektorfeld

$$\vec{V}(V, T) = \text{grad } U(V, T) = \begin{pmatrix} \frac{\partial U(V, T)}{\partial V} \\ \frac{\partial U(V, T)}{\partial T} \end{pmatrix} .$$

Damit erhält das totale Differential der inneren Energie U die Form

$$dU = \text{grad } U(V, T) \cdot \begin{pmatrix} dV \\ dT \end{pmatrix} = \vec{V} \cdot \begin{pmatrix} dV \\ dT \end{pmatrix} \quad (3.2)$$

Schließlich integrieren wir (3.2) in den Grenzen von V_1, T_1 bis V_2, T_2 :

$$\begin{aligned} \int_{V_1, T_1}^{V_2, T_2} dU &= \int_{V_1, T_1}^{V_2, T_2} \vec{V} \cdot \begin{pmatrix} dV \\ dT \end{pmatrix} \\ &= \int_{V_1, T_1}^{V_2, T_2} \left[\left(\frac{\partial U(V, T)}{\partial V} \right)_{(T)} dV + \left(\frac{\partial U(V, T)}{\partial T} \right)_{(V)} dT \right] \\ &= U(V_2, T_2) - U(V_1, T_1) = \Delta U . \end{aligned}$$

Wir können feststellen, dass das zur inneren Energie gehörende Gradientenfeld $\vec{V}(V, T)$ dem konservativen Vektorfeld $\vec{V}(\vec{r})$ im Abschnitt 10 entspricht. Deshalb ist das Integral über $\vec{V}(V, T)$ wegunabhängig und somit das totale Differential dU bzw. die Änderung ΔU der inneren Energie nur vom Anfangs- und Endpunkt der Änderung abhängt. Oft vereinfachen sich dadurch Problemlösungen in der klassischen Thermodynamik oder werden dadurch erst möglich.

Explizite und implizite Zeitabhängigkeit, totale Zeitableitung

Insbesondere in der Mechanik spielen die verschiedenen Zeitabhängigkeiten physikalischer Größen und die sich daraus ergebenden verschiedenen **Ableitungen nach der Zeit** eine große Rolle. Betrachten wir also eine physikalische Größe f als Funktion von der Zeit und von den kartesischen Koordinaten x, y, z im Ortsraum:

- $f = f(\vec{r}, t) = f(x, y, z, t)$

f hängt von den Ortskoordinaten und **explizit** von der Zeit ab, wobei die Ortskoordinaten nicht von der Zeit abhängen. Dabei werden die Orts- und die Zeitabhängigkeit von f an einem **festen** Ort betrachtet. Weil die Ortskoordinaten zeitunabhängig sind, ist die Zeitableitung von f in diesem Fall ausschließlich partiell:

$$\frac{\partial f(\vec{r}, t)}{\partial t} = \frac{\partial f(x, y, z, t)}{\partial t} . \quad (3.3)$$

- $f = f(\vec{r}(t)) = f(x(t), y(t), z(t))$

f hängt über die zeitabhängigen Ortskoordinaten ausschließlich **implizit** von der Zeit ab. Wir betrachten f also an einem mit der Geschwindigkeit \vec{v} **bewegten** Raumpunkt. Die Zeitableitung erfolgt hier nach der Kettenregel:

$$\begin{aligned} \frac{df(\vec{r}(t))}{dt} &= \frac{\partial f(x, y, z)}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f(x, y, z)}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f(x, y, z)}{\partial z} \cdot \frac{dz}{dt} \quad (3.4) \\ &= \text{grad } f \cdot \frac{d\vec{r}}{dt} = \vec{v} \cdot \nabla f = (\vec{v} \cdot \vec{\nabla}) f \end{aligned}$$

$$\text{mit } \vec{v} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \quad \text{und dem Operator } (\vec{v} \cdot \vec{\nabla}).$$

Formal liefert die Multiplikation von df/dt mit dt das totale Differential df .

- $f = f(\vec{r}(t), t) = f(x(t), y(t), z(t), t)$

f ist sowohl implizit über die Zeitabhängigkeit der Ortskoordinaten als auch explizit zeitabhängig. Man spricht dann auch von einer zusammengesetzten Funktion. Die Ableitung von f nach der Zeit setzt sich folglich zusammen aus (3.3) und (3.4) und heißt **totale Ableitung**:

$$\begin{aligned} \frac{df(\vec{r}(t), t)}{dt} &= \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dt} + \frac{\partial f}{\partial t} \\ &= \text{grad } f \cdot \frac{d\vec{r}}{dt} + \frac{\partial f}{\partial t} = (\vec{v} \cdot \nabla) f + \frac{\partial f}{\partial t}. \end{aligned}$$

4 Lösung der Integrale $\int_0^{\infty} x^q \cdot e^{-ax^2} dx$, $q \in \mathbb{N}$, $a \in \mathbb{R}^{>0}$

Siehe auch Integrationsmethode nach Feynman oder Feynman-Trick.

\mathbb{N} sei die Menge der natürlichen Zahlen **einschließlich der Null**. Wir werden im Folgenden eine Fallunterscheidung vornehmen und betrachten zuerst den Fall mit

$$\text{ungeraden } q \in \mathbb{N} \Rightarrow m = 2k + 1 \text{ mit } k = 0, 1, 2, 3, \dots$$

und anschließend den Fall mit

$$\text{geraden } q \in \mathbb{N} \Rightarrow n = 2k \text{ mit } k = 0, 1, 2, 3, \dots$$

Außerdem ist zu berücksichtigen, dass die Integranden $x^q \cdot e^{-ax^2}$ mit ungeradem q bzw. m ungerade Funktionen und mit geradem q bzw. n gerade Funktionen von x sind.

4.1 Für ungerade q

Die Substitution

$$\begin{aligned} x^2 = u &\Rightarrow x^m = u^{\frac{m}{2}}, \\ &\Rightarrow x^{2k+1} = u^{k+\frac{1}{2}} = u^k \cdot u^{\frac{1}{2}}, \\ &\Rightarrow \frac{du}{dx} = 2x, \quad dx = \frac{1}{2x} du \end{aligned}$$

liefert

$$\begin{aligned} \int_0^{\infty} x^m e^{-ax^2} dx &= \int_0^{\infty} x^{2k+1} \cdot \frac{1}{2x} \cdot e^{-au} du = \frac{1}{2} \int_0^{\infty} x^{2k} \cdot e^{-au} du = \frac{1}{2} \int_0^{\infty} u^k \cdot e^{-au} du \Rightarrow \\ \int_0^{\infty} x^m e^{-ax^2} dx &= \frac{1}{2} \int_0^{\infty} \left(-\frac{\partial}{\partial a}\right)^k e^{-au} du, \quad k = \frac{m-1}{2} \text{ für ungerade } m \in \mathbb{N}, \quad u = x^2, \end{aligned} \tag{4.1}$$

denn

$$\begin{aligned} m = 1 &\Rightarrow k = 0 \Rightarrow \left(-\frac{\partial}{\partial a}\right)^0 e^{-au} = e^{-au}, \\ m = 3 &\Rightarrow k = 1 \Rightarrow \left(-\frac{\partial}{\partial a}\right)^1 e^{-au} = u \cdot e^{-au}, \\ m = 5 &\Rightarrow k = 2 \Rightarrow \left(-\frac{\partial}{\partial a}\right)^2 e^{-au} = u^2 \cdot e^{-au}, \\ m = 7 &\Rightarrow k = 3 \Rightarrow \left(-\frac{\partial}{\partial a}\right)^3 e^{-au} = u^3 \cdot e^{-au}, \\ &\vdots \end{aligned}$$

$$\left(-\frac{\partial}{\partial a}\right)^k e^{-au} = (-1)^k \cdot \left(\frac{\partial}{\partial a}\right)^k e^{-au} = u^k \cdot e^{-au}. \tag{4.2}$$

Damit kann man für (4.1)

$$\frac{1}{2} \int_0^{\infty} \left(-\frac{\partial}{\partial a}\right)^k e^{-au} du = \frac{1}{2} \cdot (-1)^k \cdot \frac{\partial^k}{\partial a^k} \int_0^{\infty} e^{-au} du$$

schreiben und mit

$$\int_0^{\infty} e^{-au} du = -\frac{1}{a} \cdot e^{-au} \Big|_{u=0}^{\infty} = \frac{1}{a} = a^{-1}$$

schließlich

$$\int_0^{\infty} x^m e^{-ax^2} dx = \frac{1}{2} (-1)^k \frac{\partial^k}{\partial a^k} a^{-1} = \frac{k!}{2 \cdot a^{k+1}} \text{ für ungerade } m \in \mathbb{N} \text{ und } k = \frac{m-1}{2} .$$

(4.3)

Beispiel $m = 1 \Rightarrow k = 0 :$

$$\int_0^{\infty} x \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^0 \cdot a^{-1} = \frac{1}{2} \cdot 0! \cdot \frac{1}{a^1} ,$$

$$\int_0^{\infty} x \cdot e^{-ax^2} dx = \frac{1}{2a} .$$

(4.4)

Beispiel $m = 3 \Rightarrow k = 1 :$

$$\int_0^{\infty} x^3 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^1 \cdot (-1) \cdot a^{-2} = \frac{1}{2} \cdot 1! \cdot \frac{1}{a^2} ,$$

$$\int_0^{\infty} x^3 \cdot e^{-ax^2} dx = \frac{1}{2a^2} .$$

Beispiel $m = 5 \Rightarrow k = 2 :$

$$\int_0^{\infty} x^5 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^2 \cdot (-1) \cdot (-2) \cdot a^{-3} = \frac{1}{2} \cdot 2! \cdot \frac{1}{a^3} ,$$

$$\int_0^{\infty} x^5 \cdot e^{-ax^2} dx = \frac{1}{a^3} .$$

Beispiel $m = 7 \Rightarrow k = 3 :$

$$\int_0^{\infty} x^7 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^3 \cdot (-1) \cdot (-2) \cdot (-3) \cdot a^{-4} = \frac{1}{2} \cdot 3! \cdot \frac{1}{a^4} ,$$

$$\int_0^{\infty} x^7 \cdot e^{-ax^2} dx = \frac{3}{a^4} .$$

4.2 Für $q = 0$

Mit $\int_{-\infty}^{\infty} x^0 e^{-ax^2} dx = \int_{-\infty}^{\infty} y^0 e^{-ay^2} dy = \int_{-\infty}^{\infty} e^{-ax^2} dx$ erhalten wir

$$\int_{-\infty}^{\infty} x^0 e^{-ax^2} dx = \sqrt{\int_{-\infty}^{\infty} e^{-ax^2} dx \cdot \int_{-\infty}^{\infty} e^{-ay^2} dy} = \sqrt{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dx dy}.$$

Die Verwendung von ebenen Polarkoordinaten ermöglicht die Substitutionen

$$x^2 + y^2 = r^2, \quad \text{Flächenelement } dx dy = r dr d\varphi,$$

sodass

$$\int_{-\infty}^{\infty} x^0 e^{-ax^2} dx = \sqrt{\int_0^{2\pi} d\varphi \cdot \int_0^{\infty} r e^{-ar^2} dr}.$$

Mit $\int_0^{\infty} r e^{-ar^2} dr = \frac{1}{2a}$ gemäß (4.4) und mit $\int_0^{2\pi} d\varphi = 2\pi$ resultiert daraus schließlich

$$\int_{-\infty}^{\infty} x^0 e^{-ax^2} dx = \sqrt{2\pi \cdot \frac{1}{2a}},$$

$$\boxed{\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \Rightarrow \int_0^{\infty} e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}}, \quad \text{für } n = 0} \quad (4.5)$$

Hierbei haben wir berücksichtigt, dass e^{-ax^2} eine gerade Funktion ist.

4.3 Für gerade q

Mit der Substitution $x^2 = u \Rightarrow x^n = u^{\frac{n}{2}} = u^k$ gilt zunächst

$$\int_0^{\infty} x^n e^{-ax^2} dx = \int_0^{\infty} u^k e^{-au} dx.$$

Nach Äquivalenzumformung von (4.2) gemäß

$$u^k \cdot e^{-au} = (-1)^k \cdot \left(\frac{\partial}{\partial a}\right)^k e^{-au} \Leftrightarrow u^k = e^{au} \cdot (-1)^k \cdot \left(\frac{\partial}{\partial a}\right)^k e^{-au}$$

können wir u^k im Integranden ersetzen und erhalten

$$\begin{aligned} \int_0^{\infty} x^n e^{-ax^2} dx &= \int_0^{\infty} e^{au} \cdot (-1)^k \cdot \left[\left(\frac{\partial}{\partial a}\right)^k e^{-au}\right] \cdot e^{-au} dx \\ &= (-1)^k \cdot \int_0^{\infty} \underbrace{e^{au} \cdot e^{-au}}_{=1} \cdot \left(\frac{\partial}{\partial a}\right)^k e^{-au} dx. \end{aligned}$$

Wir ziehen wieder den Differentiationsoperator vor das Integral und verwenden für das verbleibende Integral die Lösung (4.5):

$$\begin{aligned} \int_0^{\infty} x^n e^{-ax^2} dx &= (-1)^k \cdot \left(\frac{\partial}{\partial a}\right)^k \int_0^{\infty} e^{-ax} dx = (-1)^k \cdot \left(\frac{\partial}{\partial a}\right)^k \int_0^{\infty} e^{-ax^2} dx \\ &= (-1)^k \cdot \left(\frac{\partial}{\partial a}\right)^k \frac{1}{2} \sqrt{\frac{\pi}{a}}, \end{aligned}$$

$$\int_0^{\infty} x^n e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^k \cdot \sqrt{\pi} \cdot \left(\frac{\partial}{\partial a}\right)^k a^{-\frac{1}{2}} \text{ für } \mathbf{\text{gerade}} \ n \in \mathbb{N} \text{ und } k = \frac{n}{2} .$$

Beispiel $n = 2 \Rightarrow k = 1$:

$$\int_0^{\infty} x^2 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^1 \cdot \sqrt{\pi} \cdot \left(\frac{\partial}{\partial a}\right)^1 a^{-\frac{1}{2}} = \frac{1}{2} \cdot (-1) \cdot \sqrt{\pi} \cdot \left(-\frac{1}{2}\right) \cdot a^{-\frac{3}{2}},$$

$$\int_0^{\infty} x^2 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot \sqrt{\pi} \cdot \frac{1}{2} a^{-\frac{3}{2}} \Rightarrow \int_{-\infty}^{\infty} x^2 \cdot e^{-ax^2} dx = \sqrt{\pi} \cdot \frac{1}{2} a^{-\frac{3}{2}} .$$

Beispiel $n = 4 \Rightarrow k = 2$:

$$\int_0^{\infty} x^4 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^2 \cdot \sqrt{\pi} \cdot \left(\frac{\partial}{\partial a}\right)^2 a^{-\frac{1}{2}} = \frac{1}{2} \cdot \sqrt{\pi} \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot a^{-\frac{5}{2}},$$

$$\int_0^{\infty} x^4 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot \sqrt{\pi} \cdot \frac{3}{4} a^{-\frac{5}{2}} \Rightarrow \int_{-\infty}^{\infty} x^4 \cdot e^{-ax^2} dx = \sqrt{\pi} \cdot \frac{3}{4} a^{-\frac{5}{2}} .$$

Beispiel $n = 6 \Rightarrow k = 3$:

$$\int_0^{\infty} x^6 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot (-1)^3 \cdot \sqrt{\pi} \cdot \left(\frac{\partial}{\partial a}\right)^3 a^{-\frac{1}{2}} = \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot \sqrt{\pi} \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot \left(-\frac{5}{2}\right) \cdot a^{-\frac{7}{2}},$$

$$\int_0^{\infty} x^6 \cdot e^{-ax^2} dx = \frac{1}{2} \cdot \sqrt{\pi} \cdot \frac{15}{8} a^{-\frac{7}{2}} \Rightarrow \int_{-\infty}^{\infty} x^6 \cdot e^{-ax^2} dx = \sqrt{\pi} \cdot \frac{15}{8} a^{-\frac{7}{2}} .$$

Die letzten drei Beispiele lassen sich verallgemeinern zu

$$\int_0^{\infty} x^n e^{-ax^2} dx = \frac{1}{2} \sqrt{\pi} \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (n-1)}{2^{\frac{n}{2}}} a^{-\frac{n+1}{2}} \text{ für } \mathbf{\text{gerade}} \ n \in \mathbb{N} \setminus 0 .$$

Der Vollständigkeit halber zeigen wir noch die Lösungen von häufig vorkommenden Integralen, die eine gewisse Ähnlichkeit zu den bisher dargestellten Integralen der Form $\int_0^\infty x^q e^{-ax^2} dx$ besitzen.

- e^{-ax} statt e^{-ax^2} und $q = N \in \mathbb{N}$:

$$\int_0^\infty x^N e^{-ax} dx = \frac{N!}{a^{N+1}}, \quad N \in \mathbb{N}, 0! = 1, a \in \mathbb{R}^{>0}$$

lässt sich zeigen mit der Substitution

$$-ax = u \quad \Rightarrow \quad dx = -\frac{1}{a} du,$$

mit den Grenzen

$$\begin{aligned} x = 0 &\quad \Rightarrow \quad u = 0, \\ x \rightarrow \infty &\quad \Rightarrow \quad u \rightarrow -\infty \end{aligned}$$

und durch ggf. wiederholte Anwendung der partiellen Integration.

- e^{-ax} statt e^{-ax^2} und $q = \frac{m}{2}$ mit $m = 1, 3, 5, \dots$:

$$\int_0^\infty x^{\frac{1}{2}} e^{-ax} dx = \frac{1}{2} \sqrt{\frac{\pi}{a^3}} \quad \text{und} \quad \int_0^\infty x^{\frac{3}{2}} e^{-ax} dx = \frac{3}{4} \sqrt{\frac{\pi}{a^5}}, \quad a \in \mathbb{R}^{>0}$$

kann man zeigen mit der Substitution

$$x = u^2 \quad \text{bzw.} \quad x^{\frac{1}{2}} = u \quad \text{und} \quad x^{\frac{3}{2}} = u^3 \quad \Rightarrow \quad dx = 2u du$$

und durch anschließende Anwendung von (4.3).

5 Das Wichtigste zu reellen Matrizen

Notation:

Wenn wir im Folgenden von Matrizen sprechen, sind nur reelle Matrizen gemeint. Wir verwenden für die Notation von Matrizen einen Mix aus verschiedenen gebräuchlichen Notationen und schreiben für eine gesprochen „**m-Kreuz-n-Matrix A**“ bzw. für eine Matrix **A** vom Typ (m, n)

$$(m \times n)\text{-Matrix } \mathbf{A} = \mathbf{A}^{(m,n)} = (a_{ij})^{(m,n)}$$

mit

$i \in \mathbb{N}$	Zeilenindex $i = \{1, \dots, m\}$,
$j \in \mathbb{N}$	Spaltenindex $j = \{1, \dots, n\}$,
a_{ij}	Elemente (Matrizelemente) oder Koeffizienten von A ,
$m \times n$	Dimension der Matrix A = Anzahl ihrer Matrizelemente.

Für die Determinante einer Matrix **A** schreiben wir

$$\det \mathbf{A} = |\mathbf{A}| = \det (a_{ij}) .$$

5.1 Assoziativgesetz für die Matrizenmultiplikation

Wenn die Formate (Anzahl von Zeilen und Spalten) der Matrizen es ermöglichen, gilt für die Matrizenmultiplikation

$$\boxed{\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}} .$$

Beweis mit $\mathbf{A} = (a_{ij})^{(m,n)}$, $\mathbf{B} = (b_{jk})^{(n,r)}$, $\mathbf{C} = (c_{kl})^{(r,s)}$:

$$\begin{aligned} \mathbf{A}(\mathbf{BC}) &= \left(\sum_{j=1}^n a_{ij} \left(\sum_{k=1}^r b_{jk} c_{kl} \right) \right) \\ &= \left(\sum_{k=1}^r \left(\sum_{j=1}^n a_{ij} b_{jk} \right) c_{kl} \right) = (\mathbf{AB})\mathbf{C} \\ &= (d_{il})^{(m,s)} = \mathbf{D} \end{aligned}$$

oder bei Summation über doppelt auftretende Indizes

$$(a_{ij}) \cdot \underbrace{[(b_{jk}) \cdot (c_{kl})]}_{(g_{jl})} = (d_{il}) = \underbrace{[(a_{ij}) \cdot (b_{jk})]}_{(h_{ik})} \cdot (c_{kl}) = (d_{il}) . \quad \square$$

Die Faktoren, d. h. die Matrizen kommutieren dabei im Allgemeinen nicht. Sie dürfen deshalb in ihrer Reihenfolge allgemein nicht vertauscht werden.

5.2 Die Transponierte einer Matrix

Beim Transponieren einer Matrix werden ihre Zeilen und Spalten miteinander vertauscht, d. h. die erste Zeile wird zur ersten Spalte, die zweite Zeile wird zur zweiten Spalte usw. Aus einer

$$(m \times n)\text{-Matrix } \mathbf{A} = (a_{ij})^{(m,n)}$$

wird so deren Transponierte, die

$$(n \times m)\text{-Matrix } \mathbf{B} = (b_{ij})^{(n,m)} = (a_{ji})^{(n,m)} = (a_{ij})^T = \mathbf{A}^T .$$

Für das einzelne Matrixelement bedeutet die Transposition T

$$a_{ij} \xrightarrow{T} a_{ji} = a_{ij}^T = b_{ij} .$$

Obwohl sich ein Matrixelement allein nicht transponieren lässt, schreiben wir hier *symbolisch* a^T für b .

Beispiel:

$$\begin{aligned} \mathbf{A} &= (a_{ij})^{(2,3)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \xrightarrow{T} \\ \mathbf{A}^T &= (a_{ji})^{(3,2)} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = (b_{ij})^{(3,2)} = \mathbf{B} . \end{aligned}$$

5.3 Multiplikation transponierter Matrizen

Wenn die Formate der Matrizen es ermöglichen, gilt

$$\boxed{(\mathbf{A} \cdot \mathbf{B})^T = \mathbf{B}^T \cdot \mathbf{A}^T} .$$

Beweis mit

$$\begin{aligned} \mathbf{A} &= (a_{ij})^{(m,n)} \Rightarrow \mathbf{A}^T = (a_{ji})^{(n,m)} , \\ \mathbf{B} &= (b_{jk})^{(n,r)} \Rightarrow \mathbf{B}^T = (b_{kj})^{(r,n)} , \\ \mathbf{A} \cdot \mathbf{B} &= \left(\sum_{j=1}^n a_{ij} \cdot b_{jk} \right) = (c_{ik})^{(m,r)} \Rightarrow (\mathbf{A} \cdot \mathbf{B})^T = (c_{ki})^{(r,m)} : \\ \mathbf{B}^T \cdot \mathbf{A}^T &= \left(\sum_{j=1}^n b_{kj} \cdot a_{ji} \right) = (c_{ki})^{(r,m)} = (\mathbf{A} \cdot \mathbf{B})^T . \quad \square \end{aligned}$$

5.4 Quadratische Matrizen

Als Beispiel verwenden wir o.B.d.A. eine (4×4) -Matrix $\mathbf{A} = (a_{ij})^{(4,4)}$ mit der zugehörigen 4-reihigen Determinante bzw. Determinante 4. Ordnung $|\mathbf{A}| = \det(a_{ij})$.

- Eine quadratische Matrix ist beispielsweise die $n \times n$ -Matrix $\mathbf{A} = (a_{ij})^{(n,n)}$.
- Eine quadratische Matrix \mathbf{A} ist **singulär**, wenn $\det \mathbf{A} = 0$.
- Eine quadratische Matrix ist **regulär**, wenn ihre Determinante von Null verschieden ist:

$$\det \mathbf{A} \neq 0 \quad \Rightarrow \quad \mathbf{A} = (a_{ij})^{(n,n)} \text{ regulär .}$$

Die Begriffe „singuläre Matrix“ und „reguläre Matrix“ sind nur für quadratische Matrizen definiert.

- **Minor** (synonym: **Unterdeterminante**) D_{ij} :

Durch Streichen der i -ten Zeile und der j -ten Spalte einer Determinante $|\mathbf{A}|$ erhalten wir den Minor bzw. die Unterdeterminante D_{ij} zum Element a_{ij} :
beispielsweise

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} \quad \Rightarrow \quad D_{ij} = D_{32} = \begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} .$$

- **Adjunkte** (synonym: **algebraisches Komplement**) A_{ij} :

Die Adjunkte bzw. das algebraische Komplement A_{ij} erhalten wir durch Multiplikation der Unterdeterminante D_{ij} mit dem Faktor $(-1)^{i+j}$ entsprechend dem sogenannten

$$\text{Vorzeichen-Schachbrettmuster} \quad \begin{array}{cccc} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \vdots & \vdots & \vdots & \ddots \end{array} \quad \text{für Determinanten :}$$

$$\boxed{A_{ij} = (-1)^{i+j} \cdot D_{ij}} .$$

- Mit den Adjunkten A_{ij} einer Matrix \mathbf{A} lassen sich n -reihige Determinanten bzw. Determinanten n -ter Ordnung berechnen nach dem **Laplace'schen Entwicklungssatz**:

$$\text{Entwicklung nach der } i\text{-ten Zeile:} \quad \det \mathbf{A} = \sum_{j=1}^n a_{ij} A_{ij} \quad (i = 1, 2, \dots, n) ,$$

$$\text{Entwicklung nach der } j\text{-ten Spalte:} \quad \det \mathbf{A} = \sum_{i=1}^n a_{ij} A_{ij} \quad (j = 1, 2, \dots, n) .$$

Der Wert der 1-reihigen Determinante ist gleich dem Wert des einzigen Matrixelements.

2-reihige Determinanten werden wie folgt berechnet:

$$\begin{aligned}\det \mathbf{A} &= \det (a_{ij})^{(2,2)} = a_{11} a_{22} - a_{12} a_{21} \\ &= a_{11} a_{22} - a_{21} a_{12} .\end{aligned}$$

Nicht nur für 2-reihige Determinanten, wie man hier sehen kann, sondern auch allgemein gilt folglich

$$\det \mathbf{A} = \det \mathbf{A}^T \quad (\text{Stürzen der Determinante } |\mathbf{A}|) .$$

3-reihige Determinanten werden beispielsweise durch Entwicklung nach der 1. Zeile wie folgt berechnet:

$$\begin{aligned}\det \mathbf{A} &= \det (a_{ij})^{(3,3)} \\ &= a_{11} D_{11} - a_{12} D_{12} + a_{13} D_{13} \quad (D_{1j} \text{ Unterdeterminanten zur 1. Zeile}) \\ &= a_{11} A_{11} + a_{12} A_{12} + a_{13} A_{13} \quad (A_{1j} \text{ Adjunkten zur 1. Zeile}) .\end{aligned}$$

4-reihige Determinanten werden beispielsweise durch Entwicklung nach der 1. Zeile wie folgt berechnet:

$$\det \mathbf{A} = \det (a_{ij})^{(4,4)} = \sum_{j=1}^4 a_{1j} A_{1j} = a_{11} A_{11} + a_{12} A_{12} + a_{13} A_{13} + a_{14} A_{14} ,$$

wobei sich die Adjunkten zur 1. Zeile wie folgt aus den 3-reihigen Unterdeterminanten zur 1. Zeile ergeben:

$$A_{11} = D_{11}, \quad A_{12} = -D_{12}, \quad A_{13} = D_{13}, \quad A_{14} = -D_{14} .$$

- **adjungierte Matrix \mathbf{A}_{adj} :**

Die zu einer Matrix \mathbf{A} adjungierte Matrix \mathbf{A}_{adj} ist die Transponierte $(A_{ij})^T$ der Matrix (A_{ij}) , gebildet aus den Adjunkten A_{ij} von \mathbf{A} , also beispielsweise

$$\mathbf{A}_{\text{adj}} = (A_{ij})^T = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & A_{31} & A_{41} \\ A_{12} & A_{22} & A_{32} & A_{42} \\ A_{13} & A_{23} & A_{33} & A_{43} \\ A_{14} & A_{24} & A_{34} & A_{44} \end{pmatrix} .$$

5.4.1 Invertieren von Matrizen

Invertieren kann man nur quadratische Matrizen. Eine (quadratische) Matrix \mathbf{A} besitzt genau dann eine Inverse, wenn sie regulär ist, d. h. wenn $|\mathbf{A}| \neq 0$ ist. Daraus folgt mit der **Einheitsmatrix** $\mathbb{1}$ und mit dem **Multiplikationstheorem** für Determinanten $|\mathbf{C}| = |\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|$:

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbb{1} \quad \Rightarrow$$

$$|\mathbf{A}^{-1} \cdot \mathbf{A}| = |\mathbb{1}| = |\mathbf{A}^{-1}| \cdot |\mathbf{A}| = 1 \quad \Leftrightarrow \quad |\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|} \neq 0.$$

Zunächst zeigen wir am Beispiel der 4×4 -Matrix $\mathbf{A} = (a_{ij})^{(4,4)}$, dass

$$\boxed{\mathbf{A} \cdot \mathbf{A}_{\text{adj}} = \mathbf{A} \cdot (A_{ij})^T = |\mathbf{A}| \cdot \mathbb{1}} : \quad (5.1)$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} A_{11} & A_{21} & A_{31} & A_{41} \\ A_{12} & A_{22} & A_{32} & A_{42} \\ A_{13} & A_{23} & A_{33} & A_{43} \\ A_{14} & A_{24} & A_{34} & A_{44} \end{pmatrix} = \begin{pmatrix} |\mathbf{A}| & 0 & 0 & 0 \\ 0 & |\mathbf{A}| & 0 & 0 \\ 0 & 0 & |\mathbf{A}| & 0 \\ 0 & 0 & 0 & |\mathbf{A}| \end{pmatrix} = |\mathbf{A}| \cdot \mathbb{1} . \quad (5.2)$$

Die Matrixelemente $|\mathbf{A}|$ der Diagonalen gehen gemäß des Laplace'schen Entwicklungssatzes beispielsweise zurück auf die Determinante, gebildet aus dem Produkt der 2. Zeile von \mathbf{A} mit der 2. Spalte von \mathbf{A}_{adj} :

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix} \cdot \begin{pmatrix} A_{21} \\ A_{22} \\ A_{23} \\ A_{24} \end{pmatrix} = a_{21} A_{21} + a_{22} A_{22} + a_{23} A_{23} + a_{24} A_{24} = |\mathbf{A}|$$

ist die Entwicklung nach der 2. Zeile: $|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} .$

Dass die Matrixelemente außerhalb der Diagonalen alle gleich Null sind, ergibt sich aus der Rechenregel für Determinanten, nach der eine Determinante dann den Wert Null hat, wenn zwei Parallelreihen miteinander übereinstimmen. So geht beispielsweise das Produkt der 1. Zeile von \mathbf{A} mit der 3. Spalte von \mathbf{A}_{adj} , also

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \end{pmatrix} \cdot \begin{pmatrix} A_{31} \\ A_{32} \\ A_{33} \\ A_{34} \end{pmatrix} = a_{11} A_{31} + a_{12} A_{32} + a_{13} A_{33} + a_{14} A_{34} = 0 , \quad (5.3)$$

zurück auf die Determinante

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = 0 ,$$

weil bei der Bildung der Adjunkten nicht die 1. Zeile der Matrix \mathbf{A} „gestrichen“ wird, wie es für die Entwicklung nach der 1. Zeile erforderlich wäre, sondern die 3. Zeile, sodass im Produkt (5.3) statt der 3. Zeile die 1. Zeile noch einmal und damit doppelt auftritt.

Wer dies nicht sofort überschaut, sollte sich die Mühe machen, eine 3-reihige Determinante $\det(a_{ij})^{(3,3)}$ mit dem Laplace'schen Entwicklungssatz beispielsweise nach der ersten Zeile zu entwickeln, aber nicht mit den Adjunkten zur 1. Zeile sondern mit den Adjunkten zur 2. Zeile. Man stellt dann fest, dass die auf diese eigentlich falsche Weise entwickelte Determinante den Wert Null hat.

Anschließend ist in der gleichen Determinante $\det(a_{ij})^{(3,3)}$ die 2. Zeile durch die 1. Zeile zu ersetzen, sodass die 1. Zeile doppelt auftritt, und dann die Determinante mit der Regel von Sarrus nach der 1. Zeile zu berechnen. Wieder hat die Determinante den Wert Null.

Aber **Achtung!** Dem Vorzeichen-Schachbrettmuster entsprechend ist die Reihenfolge der Vorzeichen der Adjunkten zur 2. Zeile gegenüber der Vorzeichenreihenfolge der Adjunkten zur 1. Zeile umgekehrt. Diese Vorzeichenumkehr tritt immer dann auf, wenn sich die Entwicklungsreihe für die Adjunkten um eine *ungerade* Zahl von der „richtigen“ Entwicklungsreihe für die Determinante unterscheidet, was aber keinen Einfluss auf das Ergebnis in (5.2) hat.

- Um jetzt die Formel zur Berechnung der Inversen von \mathbf{A} mit Hilfe der adjungierten Matrix \mathbf{A}_{adj} zu erhalten, multiplizieren wir (5.1) von links mit \mathbf{A}^{-1} :

$$\underbrace{\mathbf{A}^{-1} \cdot \mathbf{A}}_{= \mathbb{1}} \cdot \mathbf{A}_{\text{adj}} = \mathbf{A}^{-1} \cdot |\mathbf{A}| \cdot \mathbb{1} = |\mathbf{A}| \cdot \mathbf{A}^{-1} \Leftrightarrow$$

$$\boxed{\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \mathbf{A}_{\text{adj}}}$$

- Invertieren von transponierten Matrizen:

$$\boxed{(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T},$$

denn

$$\underbrace{(\mathbf{A}^T)^{-1} \cdot \mathbf{A}^T}_{= \mathbb{1}} = (\mathbf{A}^{-1})^T \cdot \mathbf{A}^T = (\mathbf{A} \cdot \mathbf{A}^{-1})^T = \mathbb{1}^T = \mathbb{1} . \quad \square$$

- Invertieren von Matrizenprodukten:

$$\boxed{(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}},$$

denn

$$(\mathbf{A} \cdot \mathbf{B})(\mathbf{B}^{-1} \cdot \mathbf{A}^{-1}) = \mathbf{A} \cdot \underbrace{(\mathbf{B}\mathbf{B}^{-1})}_{= \mathbb{1}} \cdot \mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbb{1} . \quad \square$$

5.4.2 Orthogonale Matrizen – Drehmatrizen

Orthogonale Matrizen \mathbf{O} sind reguläre quadratische Matrizen mit der Eigenschaft

$$\det \mathbf{O} = \begin{cases} -1 & \text{für Spiegelungsmatrizen ,} \\ +1 & \text{für Drehungs-oder Drehmatrizen } \mathbf{D} \text{ im Rechtssystem .} \end{cases}$$

Uns interessieren hier nur die **Drehmatrizen** \mathbf{D} . Sie beschreiben Drehungen von einem Rechtssystem in ein Rechtssystem und haben die folgenden Eigenschaften:

- **Orthogonalität:** $\mathbf{D}^T = \mathbf{D}^{-1} \Leftrightarrow \mathbf{D}\mathbf{D}^T = \mathbf{D}^T\mathbf{D} = \mathbb{1}$.
Die Orthogonalität der Drehmatrizen impliziert folglich ihre Reihen-Orthonormalität, d. h. ihre Zeilen- und Spalten-Orthonormalität, was man am einfachsten durch Drehung einer Standardbasis $\{\vec{e}_i\}$ in die Standardbasis $\{\vec{e}_i'\}$ mittels \mathbf{D}^{-1} und die anschließende Rückdrehung mittels \mathbf{D} zeigen kann.
- Produkte orthogonaler Matrizen sind wieder orthogonal.

Nacheinander ausgeführte Drehungen mit den Drehmatrizen \mathbf{A} , \mathbf{B} , \mathbf{C} :

$$\text{erste Drehung : } \mathbf{A} \cdot \vec{x} = \vec{x}' ,$$

$$\text{zweite Drehung : } \mathbf{B} \cdot \vec{x}' = \vec{x}'' ,$$

$$\text{Gesamtdrehung : } \mathbf{C} \cdot \vec{x} = \mathbf{B} \cdot (\mathbf{A} \cdot \vec{x}) = (\mathbf{B} \cdot \mathbf{A}) \cdot \vec{x} .$$

6 Polar-, Zylinder- und Kugelkoordinaten auf einen Blick

(Siehe Lehrbuch aus der Teubner Studienbücherei Physik: Siegfried Großmann, Mathematischer Einführungskurs für die Physik, Teubner-Verlag, Stuttgart, Leipzig, 8. Auflage, 2000.)

Es ist üblich, weil es am einfachsten und am anschaulichsten ist, die verschiedenen Koordinatensysteme in Beziehung zu kartesischen Koordinaten darzustellen. Das kartesische (x, y, z) -Koordinatensystem ist nämlich ausgezeichnet durch seine globale Orthonormalbasis $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$. Die kartesische Orthonormalbasis wird Standardbasis, natürliche Basis oder kanonische Basis genannt. Die Darstellung von Ortsvektoren \vec{r} in nichtkartesischen Koordinaten u, v, w erfolgt meistens in der Form

$$\underbrace{\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\text{kartesisch}} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z \quad \longrightarrow \quad \vec{r} = \begin{pmatrix} x(u, v, w) \\ y(u, v, w) \\ z(u, v, w) \end{pmatrix} = x(u, v, w)\vec{e}_x + y(u, v, w)\vec{e}_y + z(u, v, w)\vec{e}_z.$$

Besitzt das (u, v, w) -Koordinatensystem die normierte Basis $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$, so kann die Darstellung eines Ortsvektors \vec{r} in der Form

$$\vec{r} = u\vec{e}_u + v\vec{e}_v + w\vec{e}_w \quad \longrightarrow \quad \vec{r} = u(x, y, z)\vec{e}_u(x, y, z) + v(x, y, z)\vec{e}_v(x, y, z) + w(x, y, z)\vec{e}_w(x, y, z)$$

erfolgen. Dabei haben wir berücksichtigt, dass die Standardbasis unabhängig von \vec{r} , also ortsunabhängig bzw. **global** ist, während die Basisvektoren $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$ allgemein abhängig von den Ortskoordinaten und somit **lokal** sind (wie z. B. bei krummlinigen Koordinatensystemen wie den Polarkoordinatensystemen). Wenn man von Polarkoordinaten im Allgemeinen spricht, sind sphärische Polarkoordinaten (Kugelkoordinaten), Zylinderkoordinaten und ebene Polarkoordinaten (kurz Polarkoordinaten) gemeint. **Diese drei Polarkoordinatensysteme sind lokal orthogonal und teilweise krummlinig.**

6.1 Polarkoordinaten (r, φ)

- Polarkoordinaten (für $0 \leq \varphi < 2\pi$), dargestellt in kartesischen Koordinaten:

$$r = \sqrt{x^2 + y^2},$$

$$\varphi = \begin{cases} \arctan\left(\frac{y}{x}\right) & , \text{ wenn } x > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & , \text{ wenn } x < 0 \wedge y \geq 0, \\ \arctan\left(\frac{y}{x}\right) - \pi & , \text{ wenn } x < 0 \wedge y < 0, \\ \arccos\frac{x}{r} & , \text{ wenn } y \geq 0, \\ 2\pi - \arccos\frac{x}{r} & , \text{ wenn } y < 0. \end{cases}$$

- Ortsvektor \vec{r} :

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cdot \cos \varphi \\ r \cdot \sin \varphi \end{pmatrix} = r \cos \varphi \cdot \vec{e}_x + r \sin \varphi \cdot \vec{e}_y = r \cdot \vec{e}_r \quad \Rightarrow$$

$$|\vec{g}_r| = \left| \frac{\partial \vec{r}}{\partial r} \right| = g_r = 1, \quad |\vec{g}_\varphi| = \left| \frac{\partial \vec{r}}{\partial \varphi} \right| = g_\varphi = r.$$

- Basiseinheitsvektoren:

$$\vec{e}_r = \frac{\frac{\partial \vec{r}}{\partial r}}{\left| \frac{\partial \vec{r}}{\partial r} \right|} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad \vec{e}_\varphi = \frac{\frac{\partial \vec{r}}{\partial \varphi}}{\left| \frac{\partial \vec{r}}{\partial \varphi} \right|} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}.$$

- Jacobi-Matrix J :

$$J = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = (\vec{g}_r, \vec{g}_\varphi) \Leftrightarrow J^{-1} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{1}{r} \sin \varphi & \frac{1}{r} \cos \varphi \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}.$$

- Drehmatrix (Rotationsmatrix) D :

$$J = D \cdot h = \underbrace{\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}}_D \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix}}_h = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

D ist orthogonal $\Rightarrow D^{-1} = D^T$.

- Funktionaldeterminante $\det J = r$.
- Metriktensor: $(g_{kl}) = J^T \cdot J = h^2 = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$.

- Transformation der Basiseinheitsvektoren:

$$(\vec{e}_x, \vec{e}_y) \cdot D = (\vec{e}_r, \vec{e}_\varphi) \Leftrightarrow (\vec{e}_r, \vec{e}_\varphi) \cdot D^T = (\vec{e}_x, \vec{e}_y) \quad \text{mit } D^T = D^{-1} \text{ (} D \text{ ist orthogonal)}.$$

- Transformation von Differentialen:

$$J^{-1} \cdot \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} dr \\ d\varphi \end{pmatrix} \Leftrightarrow J \cdot \begin{pmatrix} dr \\ d\varphi \end{pmatrix} = \begin{pmatrix} dx \\ dy \end{pmatrix}.$$

- Transformation partieller Ableitungen:

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \cdot J = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \varphi} \right) \Leftrightarrow \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \varphi} \right) \cdot J^{-1} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right).$$

Mit $\cos \varphi = \frac{x}{r} = \frac{x}{\sqrt{x^2+y^2}}$ und $\sin \varphi = \frac{y}{r} = \frac{y}{\sqrt{x^2+y^2}}$ ergibt das

$$\begin{aligned} \frac{\partial}{\partial r} &= \frac{x}{\sqrt{x^2+y^2}} \frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2+y^2}} \frac{\partial}{\partial y}, & \frac{\partial}{\partial \varphi} &= x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}, \\ \frac{\partial}{\partial x} &= \cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi}, & \frac{\partial}{\partial y} &= \sin \varphi \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi}. \end{aligned}$$

- Transformation eines Vektorfeldes $\vec{F}(\vec{r})$:

$$D^T \cdot \begin{pmatrix} F_x \\ F_y \end{pmatrix} = \begin{pmatrix} F_r \\ F_\varphi \end{pmatrix} \Leftrightarrow D \cdot \begin{pmatrix} F_r \\ F_\varphi \end{pmatrix} = \begin{pmatrix} F_x \\ F_y \end{pmatrix},$$

$$\vec{F} = F_x \vec{e}_x + F_y \vec{e}_y = F_r \vec{e}_r + F_\varphi \vec{e}_\varphi.$$

- Flächenelement: $dS(r, \varphi) = |\det J| \cdot dr d\varphi = r dr d\varphi$.

- Linienelement ds :

$$(d\vec{r})^2 = (ds)^2 = g_r^2 (dr)^2 + g_\varphi^2 (d\varphi)^2 = (dr)^2 + r^2 (d\varphi)^2.$$

- Gradient: $\text{grad } f(r, \varphi) = \frac{\partial f}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} \vec{e}_\varphi$.

- Divergenz: $\text{div } \vec{F}(r, \varphi) = \frac{1}{r} \frac{\partial}{\partial r} (r \cdot F_r) + \frac{1}{r} \frac{\partial F_\varphi}{\partial \varphi}$.

- Rotation: $[\text{rot } \vec{F}(r, \varphi)]_z = \frac{1}{r} \frac{\partial}{\partial r} (r \cdot F_\varphi) - \frac{1}{r} \frac{\partial F_r}{\partial \varphi}$.

- Laplacian: $\Delta f(r, \varphi) = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \varphi^2}$,

wobei der Laplace-Operator oder Laplacian $\Delta := \nabla \cdot \nabla = \nabla^2 = \text{div grad}$.

6.2 Zylinderkoordinaten (r, φ, z)

- Zylinderkoordinaten (für $0 \leq \varphi < 2\pi$), dargestellt in kartesischen Koordinaten:

$$r = \sqrt{x^2 + y^2},$$

$$\varphi = \begin{cases} \arctan\left(\frac{y}{x}\right) & , \text{ wenn } x > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & , \text{ wenn } x < 0 \wedge y \geq 0, \\ \arctan\left(\frac{y}{x}\right) - \pi & , \text{ wenn } x < 0 \wedge y < 0, \\ \arccos\frac{x}{r} & , \text{ wenn } y \geq 0, \\ 2\pi - \arccos\frac{x}{r} & , \text{ wenn } y < 0, \end{cases}$$

$$z = z.$$

- Ortsvektor \vec{r} :

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cdot \cos \varphi \\ r \cdot \sin \varphi \\ z \end{pmatrix} = r \cos \varphi \cdot \vec{e}_x + r \sin \varphi \cdot \vec{e}_y + z \cdot \vec{e}_z = r \cdot \vec{e}_r + z \cdot \vec{e}_z \Rightarrow$$

$$\vec{g}_r = \frac{\partial \vec{r}}{\partial r} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \quad \vec{g}_\varphi = \frac{\partial \vec{r}}{\partial \varphi} = \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \\ 0 \end{pmatrix}, \quad \vec{g}_z = \frac{\partial \vec{r}}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$|\vec{g}_r| = \left| \frac{\partial \vec{r}}{\partial r} \right| = g_r = 1, \quad |\vec{g}_\varphi| = \left| \frac{\partial \vec{r}}{\partial \varphi} \right| = g_\varphi = r, \quad |\vec{g}_z| = \left| \frac{\partial \vec{r}}{\partial z} \right| = g_z = 1.$$

- Basiseinheitsvektoren:

$$\vec{e}_r = \frac{\frac{\partial \vec{r}}{\partial r}}{\left| \frac{\partial \vec{r}}{\partial r} \right|} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \quad \vec{e}_\varphi = \frac{\frac{\partial \vec{r}}{\partial \varphi}}{\left| \frac{\partial \vec{r}}{\partial \varphi} \right|} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \quad \vec{e}_z = \frac{\frac{\partial \vec{r}}{\partial z}}{\left| \frac{\partial \vec{r}}{\partial z} \right|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

- Jacobi-Matrix J :

$$J = \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = (\vec{g}_r, \vec{g}_\varphi, \vec{g}_z) \Leftrightarrow$$

$$J^{-1} = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\frac{1}{r} \sin \varphi & \frac{1}{r} \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} & 0 \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

- Drehmatrix (Rotationsmatrix) D :

$$J = D \cdot h = \underbrace{\begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}}_D \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{pmatrix}}_h.$$

D ist orthogonal $\Rightarrow D^{-1} = D^T$.

- Funktionaldeterminante $\det J = r$.

- Metriktensor: $(g_{kl}) = J^T \cdot J = h^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

- Transformation der Basiseinheitsvektoren:

$$(\vec{e}_x, \vec{e}_y, \vec{e}_z) \cdot D = (\vec{e}_r, \vec{e}_\varphi, \vec{e}_z) \Leftrightarrow (\vec{e}_r, \vec{e}_\varphi, \vec{e}_z) \cdot D^T = (\vec{e}_x, \vec{e}_y, \vec{e}_z).$$

- Transformation von Differentialen:

$$J^{-1} \cdot \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} dr \\ d\varphi \\ dz \end{pmatrix} \Leftrightarrow J \cdot \begin{pmatrix} dr \\ d\varphi \\ dz \end{pmatrix} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}.$$

- Transformation partieller Ableitungen:

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \cdot J = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \varphi}, \frac{\partial}{\partial z} \right) \Leftrightarrow \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \varphi}, \frac{\partial}{\partial z} \right) \cdot J^{-1} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

- Transformation eines Vektorfeldes $\vec{F}(\vec{r})$:

$$D^T \cdot \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} F_r \\ F_\varphi \\ F_z \end{pmatrix} \Leftrightarrow D \cdot \begin{pmatrix} F_r \\ F_\varphi \\ F_z \end{pmatrix} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix},$$

$$\vec{F} = F_x \vec{e}_x + F_y \vec{e}_y + F_z \vec{e}_z = F_r \vec{e}_r + F_\varphi \vec{e}_\varphi + F_z \vec{e}_z.$$

- Volumenelement: $dV = |\det J| \cdot dr d\varphi dz = r \cdot dr d\varphi dz$.
- Flächenelement auf dem Zylindermantel mit $r = \text{const}$:

$$dS = \frac{dV}{dr} = |\det J| \cdot d\varphi dz = r \cdot d\varphi dz.$$

Das Flächenelement in den Ebenen $z = \text{const}$ ist gleich dem Flächenelement in ebenen Polarkoordinaten.

- Linienelement ds :

$$(d\vec{r})^2 = (ds)^2 = g_r^2 (dr)^2 + g_\varphi^2 (d\varphi)^2 + g_z^2 (dz)^2 = (dr)^2 + r^2 (d\varphi)^2 + (dz)^2.$$

- Gradient: $\text{grad } f(r, \varphi, z) = \frac{\partial f}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} \vec{e}_\varphi + \frac{\partial f}{\partial z} \vec{e}_z$.

- Divergenz: $\text{div } \vec{F}(r, \varphi, z) = \frac{1}{r} \frac{\partial}{\partial r} (r \cdot F_r) + \frac{1}{r} \frac{\partial F_\varphi}{\partial \varphi} + \frac{\partial F_z}{\partial z}$.

- Rotation: $\text{rot } \vec{F}(r, \varphi, z) = \left(\frac{1}{r} \frac{\partial F_z}{\partial \varphi} - \frac{\partial F_\varphi}{\partial z} \right) \vec{e}_r + \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r} \right) \vec{e}_\varphi + \frac{1}{r} \left(\frac{\partial}{\partial r} (r \cdot F_\varphi) - \frac{\partial F_r}{\partial \varphi} \right) \vec{e}_z$.

- Laplacian: $\Delta f(r, \varphi, z) = \frac{1}{r} \frac{\partial}{\partial r} \left(r \cdot \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2}$.

6.3 Kugelkoordinaten (r, ϑ, φ)

- Kugelkoordinaten (für $0 \leq \vartheta < \pi$ und $0 \leq \varphi < 2\pi$), dargestellt in kartesischen Koordinaten:

$$r = \sqrt{x^2 + y^2 + z^2},$$

$$\vartheta = \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} = \operatorname{arccot} \frac{z}{\sqrt{x^2 + y^2}},$$

$$\varphi = \begin{cases} \arctan\left(\frac{y}{x}\right) & , \text{ wenn } x > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & , \text{ wenn } x < 0 \wedge y \geq 0, \\ \arctan\left(\frac{y}{x}\right) - \pi & , \text{ wenn } x < 0 \wedge y < 0, \\ \arccos \frac{x}{r} & , \text{ wenn } y \geq 0, \\ 2\pi - \arccos \frac{x}{r} & , \text{ wenn } y < 0. \end{cases}$$

- Ortsvektor \vec{r} :

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cdot \sin \vartheta \cos \varphi \\ r \cdot \sin \vartheta \sin \varphi \\ r \cdot \cos \vartheta \end{pmatrix} = r \sin \vartheta \cos \varphi \cdot \vec{e}_x + r \sin \vartheta \sin \varphi \cdot \vec{e}_y + r \cos \vartheta \cdot \vec{e}_z$$

$$= r \cdot \vec{e}_r \Rightarrow$$

$$\vec{g}_r = \frac{\partial \vec{r}}{\partial r} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}, \quad \vec{g}_\vartheta = \frac{\partial \vec{r}}{\partial \vartheta} = \begin{pmatrix} r \cos \vartheta \cos \varphi \\ r \cos \vartheta \sin \varphi \\ -r \sin \vartheta \end{pmatrix}, \quad \vec{g}_\varphi = \frac{\partial \vec{r}}{\partial \varphi} = \begin{pmatrix} -r \sin \vartheta \sin \varphi \\ r \sin \vartheta \cos \varphi \\ 0 \end{pmatrix},$$

$$|\vec{g}_r| = \left| \frac{\partial \vec{r}}{\partial r} \right| = g_r = 1, \quad |\vec{g}_\vartheta| = \left| \frac{\partial \vec{r}}{\partial \vartheta} \right| = g_\vartheta = r, \quad |\vec{g}_\varphi| = \left| \frac{\partial \vec{r}}{\partial \varphi} \right| = g_\varphi = r \sin \vartheta.$$

- Basiseinheitsvektoren:

$$\vec{e}_r = \frac{\frac{\partial \vec{r}}{\partial r}}{\left| \frac{\partial \vec{r}}{\partial r} \right|} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}, \quad \vec{e}_\vartheta = \frac{\frac{\partial \vec{r}}{\partial \vartheta}}{\left| \frac{\partial \vec{r}}{\partial \vartheta} \right|} = \begin{pmatrix} \cos \vartheta \cos \varphi \\ \cos \vartheta \sin \varphi \\ -\sin \vartheta \end{pmatrix}, \quad \vec{e}_\varphi = \frac{\frac{\partial \vec{r}}{\partial \varphi}}{\left| \frac{\partial \vec{r}}{\partial \varphi} \right|} = \begin{pmatrix} -\sin \vartheta \sin \varphi \\ \sin \vartheta \cos \varphi \\ 0 \end{pmatrix}. \quad (6.1)$$

- Jacobi-Matrix $J = (\vec{g}_r, \vec{g}_\vartheta, \vec{g}_\varphi)$:

$$J = \begin{pmatrix} \frac{\partial(x, y, z)}{\partial(r, \vartheta, \varphi)} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \vartheta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \vartheta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \vartheta} & \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \sin \vartheta \cos \varphi & r \cos \vartheta \cos \varphi & -r \sin \vartheta \sin \varphi \\ \sin \vartheta \sin \varphi & r \cos \vartheta \sin \varphi & r \sin \vartheta \cos \varphi \\ \cos \vartheta & -r \sin \vartheta & 0 \end{pmatrix},$$

$$J^{-1} = \begin{pmatrix} \frac{\partial(r, \vartheta, \varphi)}{\partial(x, y, z)} \end{pmatrix} = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} & \frac{\partial r}{\partial z} \\ \frac{\partial \vartheta}{\partial x} & \frac{\partial \vartheta}{\partial y} & \frac{\partial \vartheta}{\partial z} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} & \frac{\partial \varphi}{\partial z} \end{pmatrix} = \begin{pmatrix} \sin \vartheta \cos \varphi & \sin \vartheta \sin \varphi & \cos \vartheta \\ \frac{1}{r} \cos \vartheta \cos \varphi & \frac{1}{r} \cos \vartheta \sin \varphi & -\frac{1}{r} \sin \vartheta \\ -\frac{1}{r} \frac{\sin \varphi}{\sin \vartheta} & \frac{1}{r} \frac{\cos \varphi}{\sin \vartheta} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{x}{r} & \frac{y}{r} & \frac{z}{r} \\ \frac{xz}{r^2 \sqrt{x^2 + y^2}} & \frac{yz}{r^2 \sqrt{x^2 + y^2}} & \frac{-(x^2 + y^2)}{r^2 \sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} & 0 \end{pmatrix}.$$

- Drehmatrix (Rotationsmatrix) D :

$$J = D \cdot h = \underbrace{\begin{pmatrix} \sin \vartheta \cos \varphi & \cos \vartheta \cos \varphi & -\sin \varphi \\ \sin \vartheta \sin \varphi & \cos \vartheta \sin \varphi & \cos \varphi \\ \cos \vartheta & -\sin \vartheta & 0 \end{pmatrix}}_D \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & r \sin \vartheta \end{pmatrix}}_h.$$

- Funktionaldeterminante $\det J = r^2 \sin \vartheta$.

- Metriktensor: $(g_{kl}) = J^T \cdot J = h^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \vartheta \end{pmatrix}$.

- Transformation der Basiseinheitsvektoren:

$$(\vec{e}_x, \vec{e}_y, \vec{e}_z) \cdot D = (\vec{e}_r, \vec{e}_\vartheta, \vec{e}_\varphi) \Leftrightarrow (\vec{e}_r, \vec{e}_\vartheta, \vec{e}_\varphi) \cdot D^T = (\vec{e}_x, \vec{e}_y, \vec{e}_z).$$

- Transformation von Differentialen:

$$J^{-1} \cdot \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} dr \\ d\vartheta \\ d\varphi \end{pmatrix} \Leftrightarrow J \cdot \begin{pmatrix} dr \\ d\vartheta \\ d\varphi \end{pmatrix} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}.$$

- Transformation partieller Ableitungen:

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \cdot J = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \vartheta}, \frac{\partial}{\partial \varphi} \right) \Leftrightarrow \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \vartheta}, \frac{\partial}{\partial \varphi} \right) \cdot J^{-1} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

- Transformation eines Vektorfeldes $\vec{F}(\vec{r})$:

$$D^T \cdot \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} F_r \\ F_\vartheta \\ F_\varphi \end{pmatrix} \Leftrightarrow D \cdot \begin{pmatrix} F_r \\ F_\vartheta \\ F_\varphi \end{pmatrix} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix},$$

$$\vec{F} = F_x \vec{e}_x + F_y \vec{e}_y + F_z \vec{e}_z = F_r \vec{e}_r + F_\vartheta \vec{e}_\vartheta + F_\varphi \vec{e}_\varphi.$$

- Volumenelement: $dV = dS \cdot dr = |\det J| \cdot dr d\vartheta d\varphi = r^2 \sin \vartheta \cdot dr d\vartheta d\varphi$.

- Flächenelement auf der Kugeloberfläche mit Radius $r = \text{const}$:

$$dS = \frac{dV}{dr} = |\det J| \cdot d\vartheta d\varphi = r^2 \sin \vartheta \cdot d\vartheta d\varphi.$$

- Linienelement ds :

$$(d\vec{r})^2 = (ds)^2 = g_r^2 (dr)^2 + g_\vartheta^2 (d\vartheta)^2 + g_\varphi^2 (d\varphi)^2 = (dr)^2 + r^2 (d\vartheta)^2 + r^2 \sin^2 \vartheta (d\varphi)^2.$$

- Gradient: $\text{grad } f(r, \vartheta, \varphi) = \frac{\partial f}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \vartheta} \vec{e}_\vartheta + \frac{1}{r \sin \vartheta} \frac{\partial f}{\partial \varphi} \vec{e}_\varphi$.

- Divergenz: $\text{div } \vec{F}(r, \vartheta, \varphi) = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \cdot F_r) + \frac{1}{r \sin \vartheta} \frac{\partial}{\partial \vartheta} (\sin \vartheta \cdot F_\vartheta) + \frac{1}{r \sin \vartheta} \frac{\partial F_\varphi}{\partial \varphi}$.

- Rotation: $\text{rot } \vec{F}(r, \vartheta, \varphi) = \left[\frac{1}{r \sin \vartheta} \left(\frac{\partial}{\partial \vartheta} (\sin \vartheta \cdot F_\varphi) - \frac{\partial F_\vartheta}{\partial \varphi} \right) \right] \vec{e}_r$
 $+ \left[\frac{1}{r \sin \vartheta} \frac{\partial F_r}{\partial \varphi} - \frac{1}{r} \frac{\partial}{\partial r} (r \cdot F_\varphi) \right] \vec{e}_\vartheta$
 $+ \left[\frac{1}{r} \frac{\partial}{\partial r} (r \cdot F_\vartheta) - \frac{1}{r} \frac{\partial F_r}{\partial \vartheta} \right] \vec{e}_\varphi$.

- Laplacian: $\Delta f(r, \vartheta, \varphi) = \underbrace{\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \cdot \frac{\partial f}{\partial r} \right)}_{\text{Radialableitung}} + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \cdot \frac{\partial f}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2 f}{\partial \varphi^2}$,

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \cdot \frac{\partial f}{\partial r} \right) = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r} \frac{\partial f}{\partial r} = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r \cdot f).$$

7 Linien-, Flächen-, Volumenelement

7.1 Linienelement

Weil Linien oder Kurven indimensionale geometrische Objekte sind, benötigt man für ihre Beschreibung nur *einen* freien Parameter, den Kurvenparameter, auch wenn es sich beispielsweise um eine Raumkurve (im 3-dimensionalen Raum) handelt.

7.1.1 Linienelement in kartesischen Koordinaten

Im kartesischen Koordinatensystem kann folglich jeder Punkt einer Linie durch den zugehörigen parametrisierten Ortsvektor dargestellt werden, z. B.

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \text{ mit Parameter } t, \quad \vec{r}(x) = \begin{pmatrix} x \\ y(x) \\ z(x) \end{pmatrix} \text{ mit Parameter } x.$$

Der Parameter t kann z. B. die Zeit sein und der Parameter x ist die Ortskoordinate x selbst. Aus dem Differentialquotienten

$$\frac{d\vec{r}(t)}{dt} = \begin{pmatrix} dx(t)/dt \\ dy(t)/dt \\ dz(t)/dt \end{pmatrix}$$

nach dem Parameter t resultiert das

$$\text{vektorielle Linienelement } d\vec{r}(t) = \frac{d\vec{r}(t)}{dt} \cdot dt = \begin{pmatrix} dx(t)/dt \\ dy(t)/dt \\ dz(t)/dt \end{pmatrix} \cdot dt.$$

Die Arbeit W , die eine Kraft $\vec{F}(\vec{r})$ längs eines Weges $\vec{r}(t)$ verrichtet ist damit das **Arbeitsintegral**

$$\begin{aligned} \int \vec{F} \cdot d\vec{r} &= \int \begin{pmatrix} F_x(x(t), y(t), z(t)) \\ F_y(x(t), y(t), z(t)) \\ F_z(x(t), y(t), z(t)) \end{pmatrix} \cdot \begin{pmatrix} dx(t)/dt \\ dy(t)/dt \\ dz(t)/dt \end{pmatrix} \cdot dt \\ &= \int F_x(\vec{r}(t)) \frac{dx(t)}{dt} \cdot dt + \int F_y(\vec{r}(t)) \frac{dy(t)}{dt} \cdot dt + \int F_z(\vec{r}(t)) \frac{dz(t)}{dt} \cdot dt. \end{aligned}$$

Interessant und wichtig ist, dass $d\vec{r}(t)/dt$, d. h. die Ableitung der Kurvenfunktion nach dem Kurvenparameter, der Tangentenvektor an die Kurve im Punkt $\vec{r}(t)$ ist.⁴

Das (**skalare**) **Linienelement** $ds = ds(t)$ resultiert aus der Ableitung der Kurvenlängenfunktion $s(t)$ nach dem Parameter t wie folgt:

$$\begin{aligned} \text{Kurvenstücklänge} &= s(t) \Big|_{t_1}^{t_2} \approx \sum_i \sqrt{(\Delta x_i(t))^2 + (\Delta y_i(t))^2 + (\Delta z_i(t))^2} \Big|_{t_1}^{t_2} \\ &= \lim_{\Delta t \rightarrow 0} \sum_i \sqrt{\left(\frac{\Delta x_i(t)}{\Delta t}\right)^2 + \left(\frac{\Delta y_i(t)}{\Delta t}\right)^2 + \left(\frac{\Delta z_i(t)}{\Delta t}\right)^2} \cdot \Delta t \Big|_{t_1}^{t_2} \\ &= \int_{t_1}^{t_2} \sqrt{\left(\frac{dx(t)}{dt}\right)^2 + \left(\frac{dy(t)}{dt}\right)^2 + \left(\frac{dz(t)}{dt}\right)^2} \cdot dt = \int_{t_1}^{t_2} ds(t). \end{aligned}$$

⁴Siehe dazu: Lothar Papula, Mathematik für Ingenieure und Naturwissenschaftler Band 3, Vieweg-Verlag, Braunschweig, Wiesbaden, 4. Auflage, 2001, Abschnitt 1.2 *Differentiation eines Vektors nach einem Parameter*, Seite 4 bis Seite 20.

Mit $\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} = \sqrt{\frac{d\vec{r}}{dt} \cdot \frac{d\vec{r}}{dt}} = \left|\frac{d\vec{r}}{dt}\right|$ folgt daraus

$$\boxed{ds = \left|\frac{d\vec{r}}{dt}\right| \cdot dt = |d\vec{r}| \Leftrightarrow \frac{ds(t)}{dt} = \left|\frac{d\vec{r}(t)}{dt}\right| .}$$

Das Linien- oder Kurvenintegral längs $\vec{r}(t)$ in einem skalaren Feld $f(x, y, z)$ ist damit

$$\int f \cdot ds = \int f(x(t), y(t), z(t)) \cdot \left|\frac{d\vec{r}(t)}{dt}\right| \cdot dt .$$

Allgemein, also beispielsweise für krummlinig-schiefwinklige Koordinatensysteme, wird das Linienelement ds wie folgt im Tensoralkül dargestellt:⁵

$$(ds)^2 = g_{kl} du^k du^l .$$

Dabei ist g_{kl} der metrische Tensor oder Metriktensor, der sich aus dem verwendeten Koordinatensystem mit der Jakobi-Matrix J durch

$$J^T \cdot J = h^2 = (g_{kl}) \longleftrightarrow g_{kl}$$

ergibt. Wenn man die Matrix (g_{kl}) in einen Tensor umschreibt, erhält man den Metriktensor g_{kl} , der die Metrik des Raums beschreibt.

7.1.2 Linienelement in Polarkoordinaten

Im Folgenden verwenden wir gelegentlich die Laufindizes $i \in \{1, 2, 3\}$ und $j \in \{1, 2, 3\}$ am Koordinatensymbol u , sodass $u \equiv u_1, v \equiv u_2, w \equiv u_3$.

Jetzt leiten wir das Linienelement für krummlinig-orthogonale Koordinatensysteme her. Dies betrifft insbesondere die in der Praxis wichtigen **Polarkoordinaten**. Dazu gehören insbesondere

- ebene Polarkoordinaten oder kurz Polarkoordinaten,
- Zylinderkoordinaten,
- sphärische Polarkoordinaten oder Kugelkoordinaten.

Hierbei handelt es sich um spezielle (u, v, w) - bzw. (u_1, u_2, u_3) -Koordinatensysteme, deren Koordinatenlinien $(u, v, w) \equiv (u_1, u_2, u_3)$ und folglich auch ihre Basiseinheitsvektoren

$$\vec{e}_{u_1} \equiv \vec{e}_u = \frac{\frac{\partial \vec{r}}{\partial u}}{\left|\frac{\partial \vec{r}}{\partial u}\right|}, \quad \vec{e}_{u_2} \equiv \vec{e}_v = \frac{\frac{\partial \vec{r}}{\partial v}}{\left|\frac{\partial \vec{r}}{\partial v}\right|}, \quad \vec{e}_{u_3} \equiv \vec{e}_w = \frac{\frac{\partial \vec{r}}{\partial w}}{\left|\frac{\partial \vec{r}}{\partial w}\right|}$$

stets senkrecht aufeinander stehen, sodass gilt:

$$\vec{e}_{u_i} \cdot \vec{e}_{u_j} = \delta_{ij} = \begin{cases} 0 & \text{für } i \neq j, \\ 1 & \text{für } i = j. \end{cases}$$

Außerdem bilden die Basiseinheitsvektoren in der Reihenfolge u, v, w bzw. u_1, u_2, u_3 ein Rechtssystem, also insgesamt ein orthogonales Rechtssystem bzw. ein rechtwinkliges rechtshändiges Dreibein gemäß

$$\vec{e}_i \cdot (\vec{e}_{u_j} \times \vec{e}_{u_k}) = \varepsilon_{ijk} .$$

⁵Im Abschnitt 6.8 *Der metrische Tensor (Metriktensor)* meines Skripts *Grundlegendes zur Elektrodynamik und Quantenmechanik* ist das Linienelementquadrat mit dem Metriktensor vollständig ausgeschrieben. Um derartige „Monstren“ zu handhaben, ist der Tensoralkül entwickelt worden. Es ist gut zu wissen, was sich hinter derartigen Tensoren verbirgt.

Aus

$$d\vec{r} = \frac{\partial \vec{r}}{\partial u} du + \frac{\partial \vec{r}}{\partial v} dv + \frac{\partial \vec{r}}{\partial w} dw = \vec{e}_u \left| \frac{\partial \vec{r}}{\partial u} \right| du + \vec{e}_v \left| \frac{\partial \vec{r}}{\partial v} \right| dv + \vec{e}_w \left| \frac{\partial \vec{r}}{\partial w} \right| dw$$

und mit den metrischen Koeffizienten

$$g_u = \left| \frac{\partial \vec{r}}{\partial u} \right|^2, \quad g_v = \left| \frac{\partial \vec{r}}{\partial v} \right|^2, \quad g_w = \left| \frac{\partial \vec{r}}{\partial w} \right|^2$$

folgt das **vektorielle Linienelement** (7.1)

$$d\vec{r} = g_u du \cdot \vec{e}_u + g_v dv \cdot \vec{e}_v + g_w dw \cdot \vec{e}_w \quad \Rightarrow \quad (7.1)$$

$$(d\vec{r})^2 = (ds)^2 = g_u^2 du^2 + g_v^2 dv^2 + g_w^2 dw^2. \quad (7.2)$$

Durch Radizieren des Linienelementquadrates (7.2) resultiert das (**skalare**) **Linienelement**

$$ds = \sqrt{g_u^2 du^2 + g_v^2 dv^2 + g_w^2 dw^2}. \quad (7.3)$$

Die Darstellung einer Kurve erfolgt beispielsweise durch die Parametrisierung – hier mit dem Parameter t – des Ortsvektors gemäß

$$\vec{r}(u, v, w) \quad \longrightarrow \quad \vec{r}(u(t), v(t), w(t)).$$

Bei der Bildung von Kurvenintegralen längs der Kurve $\vec{r}(t)$ benötigen wir das parametrisierte Linienelement. Dabei ist zu beachten, dass hier die metrischen Koeffizienten $g_{u_i}(t)$ längs der zugehörigen Koordinaten $u_i(t)$ konstant sind und folglich hinsichtlich der Ableitung der Differentiale $du_i(t)$ nach dem Parameter t als konstante Faktoren betrachtet werden können. Zwar sind die Basiseinheitsvektoren bei krummlinig-orthogonalen Koordinatensystemen ortsabhängig und werden deshalb parametrisiert, doch ergeben sie im Skalarprodukt unter dem Integral $\vec{e}_{u_i}(t) \cdot \vec{e}_{u_j}(t) = 0$ für $i \neq j$ und $\vec{e}_{u_i}(t) \cdot \vec{e}_{u_i}(t) = 1$. Im Übrigen erfolgt die Parametrisierung völlig analog zu der in kartesischen Koordinaten:

$$d\vec{r}(t) = \left(g_u(t) \frac{du(t)}{dt} \cdot \vec{e}_u(t) + g_v(t) \frac{dv(t)}{dt} \cdot \vec{e}_v(t) + g_w(t) \frac{dw(t)}{dt} \cdot \vec{e}_w(t) \right) \cdot dt,$$

$$ds(t) = \sqrt{g_u^2(t) \left(\frac{du(t)}{dt} \right)^2 + g_v^2(t) \left(\frac{dv(t)}{dt} \right)^2 + g_w^2(t) \left(\frac{dw(t)}{dt} \right)^2} \cdot dt.$$

Beispiel:

Betrachten wir wieder das **Arbeitsintegral**, aber diesmal in Kugelkoordinaten. Ist das Kraftfeld $\vec{F}(\vec{r})$ in kartesischen Koordinaten gegeben, müssen wir es zunächst in Kugelkoordinaten transformieren. Dies zeigen wir am Kraftfeld

$$\vec{F}(x, y, z) = \begin{pmatrix} x - yz \\ y + xz \\ z \end{pmatrix}.$$

Mit

$$\begin{aligned} x &= r \cdot \sin \vartheta \cdot \cos \varphi, \\ y &= r \cdot \sin \vartheta \cdot \sin \varphi, \\ z &= r \cdot \cos \vartheta \end{aligned}$$

ersetzen wir die kartesischen Koordinaten in $\vec{F}(x, y, z)$ und erhalten die folgende Darstellung in

Kugelkoordinaten:

$$\begin{aligned}\vec{F}(r, \vartheta, \varphi) &= \begin{pmatrix} r \sin \vartheta \cos \varphi - r \sin \vartheta \sin \phi \cdot r \cos \vartheta \\ r \sin \vartheta \sin \varphi + r \sin \vartheta \cos \phi \cdot r \cos \vartheta \\ r \cos \vartheta \end{pmatrix} \\ &= r \cdot \underbrace{\begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}}_{=\vec{e}_r} + r^2 \sin \vartheta \cdot \underbrace{\begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}}_{=\vec{e}_\varphi} \implies\end{aligned}\quad (7.4)$$

$$\begin{aligned}\vec{F}(r, \vartheta, \varphi) &= (r \sin \vartheta \cos \varphi - r^2 \sin \vartheta \sin \phi \cos \vartheta) \vec{e}_x + (r \sin \vartheta \sin \varphi + r^2 \sin \vartheta \cos \phi \cos \vartheta) \vec{e}_y + r \cos \vartheta \vec{e}_z \\ &= r \cdot \vec{e}_r + r^2 \sin \vartheta \cdot \vec{e}_\varphi \\ &= F_r \cdot \vec{e}_r + F_\varphi \cdot \vec{e}_\varphi.\end{aligned}\quad (7.5)$$

An diesem Beispiel lässt sich leicht überprüfen, dass man die skalaren Vektorkomponenten F_r , F_ϑ , F_φ von \vec{F} (in Kugelkoordinaten) entweder durch

$$D^T \cdot \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} \sin \vartheta \cos \vartheta & \sin \vartheta \sin \varphi & \cos \vartheta \\ \cos \vartheta \cos \vartheta & \cos \vartheta \sin \varphi & -\sin \vartheta \\ -\sin \varphi & \cos \varphi & 0 \end{pmatrix} \cdot \begin{pmatrix} r \sin \vartheta \cos \varphi - r \sin \vartheta \sin \phi \cdot r \cos \vartheta \\ r \sin \vartheta \sin \varphi + r \sin \vartheta \cos \phi \cdot r \cos \vartheta \\ r \cos \vartheta \end{pmatrix} = \begin{pmatrix} F_r \\ F_\vartheta \\ F_\varphi \end{pmatrix}$$

oder durch die Projektion des Vektors $\vec{F}(r, \vartheta, \varphi)$ in Darstellung (7.4) auf die Basiseinheitsvektoren \vec{e}_r , \vec{e}_ϑ , \vec{e}_φ in Darstellung (6.1) berechnen kann:

$$\vec{F}(r, \vartheta, \varphi) \cdot \vec{e}_r = F_r, \quad \vec{F}(r, \vartheta, \varphi) \cdot \vec{e}_\vartheta = F_\vartheta, \quad \vec{F}(r, \vartheta, \varphi) \cdot \vec{e}_\varphi = F_\varphi.$$

Mit dem (7.1) entsprechenden vektoriellen Linienelement für Kugelkoordinaten

$$\begin{aligned}d\vec{r}(u, v, w) &= g_u \cdot du \cdot \vec{e}_u + g_v \cdot dv \cdot \vec{e}_v + g_w \cdot dw \cdot \vec{e}_w \implies \\ d\vec{r}(r, \vartheta, \varphi) &= 1 \cdot dr \cdot \vec{e}_r + r \cdot d\vartheta \cdot \vec{e}_\vartheta + r \sin \vartheta \cdot d\varphi \cdot \vec{e}_\varphi\end{aligned}$$

können wir jetzt das Arbeitsintegral in Kugelkoordinaten bilden:

$$\begin{aligned}\int \vec{F} \cdot d\vec{r} &= \int (F_r \cdot \vec{e}_r + F_\vartheta \cdot \vec{e}_\vartheta + F_\varphi \cdot \vec{e}_\varphi) \cdot (1 \cdot dr \cdot \vec{e}_r + r \cdot d\vartheta \cdot \vec{e}_\vartheta + r \sin \vartheta \cdot d\varphi \cdot \vec{e}_\varphi) \\ &= \int (r \cdot \vec{e}_r + 0 \cdot \vec{e}_\vartheta + r^2 \sin \vartheta \cdot \vec{e}_\varphi) \cdot (1 \cdot dr \cdot \vec{e}_r + r \cdot d\vartheta \cdot \vec{e}_\vartheta + r \sin \vartheta \cdot d\varphi \cdot \vec{e}_\varphi).\end{aligned}$$

Die Parametrisierung mit dem Kurvenparameter t liefert daraus

$$\begin{aligned}\int \vec{F}(\vec{r}(t)) \cdot d\vec{r}(t) &= \int (r(t) \cdot \vec{e}_r(t) + 0 \cdot \vec{e}_\vartheta(t) + (r(t))^2 \sin \vartheta(t) \cdot \vec{e}_\varphi(t)) \\ &\quad \cdot \left(1 \cdot \frac{dr(t)}{dt} \cdot \vec{e}_r(t) + r(t) \cdot \frac{d\vartheta(t)}{dt} \cdot \vec{e}_\vartheta(t) + r(t) \sin \vartheta(t) \cdot \frac{d\varphi(t)}{dt} \cdot \vec{e}_\varphi(t)\right) \cdot dt.\end{aligned}$$

Weil im Skalarprodukt (unter dem Integral) stets $\vec{e}_{u_i}(t) \cdot \vec{e}_{u_i}(t) = 1$ gilt, ist das Arbeitsintegral in unserem Beispielfeld für ein Kurvenstück von $\vec{r}(t_1)$ bis $\vec{r}(t_2)$

$$\int_{\vec{r}(t_1)}^{\vec{r}(t_2)} \vec{F} \cdot d\vec{r} = \int_{t_1}^{t_2} \left(r(t) \frac{dr(t)}{dt} + (r(t))^3 \sin^2 \vartheta(t) \frac{d\varphi(t)}{dt} \right) \cdot dt.$$

In Polarkoordinaten (u, v, w) ist das mit t parametrisierte Kurvenintegral in einem Vektorfeld $\vec{F}(u, v, w)$

$$\begin{aligned}\int \vec{F} \cdot d\vec{r} &= \int \vec{F}(u(t), v(t), w(t)) \cdot d\vec{r}(u(t), v(t), w(t)) \\ &= \left(F_u(t) \cdot g_u \frac{du(t)}{dt} + F_v(t) \cdot g_v \frac{dv(t)}{dt} + F_w(t) \cdot g_w \frac{dw(t)}{dt} \right) \cdot dt.\end{aligned}$$

In Polarkoordinaten ist das Linien- oder Kurvenintegral längs $\vec{r}(t)$ in einem skalaren Feld $f(u, v, w)$ mit dem skalaren Linienelement (7.3) nach Parametrisierung mit t

$$\begin{aligned}\int f \cdot ds &= \int f(u(t), v(t), w(t)) \cdot ds(u(t), v(t), w(t)) \\ &= \int f(u(t), v(t), w(t)) \cdot \sqrt{g_u^2 \left(\frac{du(t)}{dt} \right)^2 + g_v^2 \left(\frac{dv(t)}{dt} \right)^2 + g_w^2 \left(\frac{dw(t)}{dt} \right)^2} \cdot dt.\end{aligned}$$

7.2 Flächenelement

Flächen S werden durch 2 freie Parameter beschrieben, z. B. durch die Parameter bzw. Koordinaten u und v :

$$S := \vec{r}(u, v).$$

Allgemein ist damit das **vektorielle Flächenelement**

$$d\vec{S}(u, v) = \underbrace{\left(\frac{\partial \vec{r}(u, v)}{\partial u} \times \frac{\partial \vec{r}(u, v)}{\partial v} \right)}_{\vec{n}} \cdot \underbrace{du dv}_{dA} = \vec{n} \cdot dA.$$

\vec{n} ist der Normalenvektor auf S und es gilt $\vec{S} \uparrow \vec{n}$. Das (**skalare**) **Flächenelement** ist folglich

$$dS(u, v) = \left| \frac{\partial \vec{r}(u, v)}{\partial u} \times \frac{\partial \vec{r}(u, v)}{\partial v} \right| \cdot du dv = |\vec{n}| \cdot dA.$$

Weiterhin gilt mit dem Einheitsnormalenvektor \vec{n}^0

$$d\vec{S} = \frac{\vec{n}}{|\vec{n}|} \cdot |\vec{n}| dA = \vec{n}^0 \cdot dS$$

Flächenelement auf der Kugeloberfläche mit Radius R in kartesischen Koordinaten:

Als freie Parameter wählen wir die beiden kartesischen Koordinaten x und y :

$$\begin{aligned} \text{Kugelfläche } S := \vec{r}(x, y) &= \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix} \Rightarrow \\ \vec{n} = \frac{\partial \vec{r}(x, y)}{\partial x} \times \frac{\partial \vec{r}(x, y)}{\partial y} &= \begin{pmatrix} x \cdot (R^2 - x^2 - y^2)^{-\frac{1}{2}} \\ y \cdot (R^2 - x^2 - y^2)^{-\frac{1}{2}} \\ 1 \end{pmatrix}, \quad |\vec{n}| = \frac{R}{\sqrt{R^2 - x^2 - y^2}}, \\ \vec{n}^0 = \frac{\vec{n}}{|\vec{n}|} &= \frac{1}{R} \cdot \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix}. \\ dS = |\vec{n}| \cdot dA &= \frac{R}{\sqrt{R^2 - x^2 - y^2}} \cdot dx dy, \\ d\vec{S} = \vec{n}^0 \cdot dS = \vec{n} \cdot dA &= \begin{pmatrix} x \cdot (R^2 - x^2 - y^2)^{-\frac{1}{2}} \\ y \cdot (R^2 - x^2 - y^2)^{-\frac{1}{2}} \\ 1 \end{pmatrix} \cdot dx dy. \end{aligned}$$

Flächenelement auf der Kugeloberfläche mit Radius R in Kugelkoordinaten:

Die freien Parameter der Kugeloberfläche sind die Kugelkoordinaten ϑ und φ :

$$\begin{aligned} \text{Kugelfläche } S := \vec{r}(\vartheta, \varphi) &= \begin{pmatrix} R \cdot \sin \vartheta \cos \varphi \\ R \cdot \sin \vartheta \sin \varphi \\ R \cdot \cos \vartheta \end{pmatrix} \Rightarrow \\ \vec{n} = \frac{\partial \vec{r}(\vartheta, \varphi)}{\partial \vartheta} \times \frac{\partial \vec{r}(\vartheta, \varphi)}{\partial \varphi} &= R^2 \begin{pmatrix} \sin^2 \vartheta \cos \varphi \\ \sin^2 \vartheta \sin \varphi \\ \sin \vartheta \cos \vartheta \end{pmatrix}, \quad |\vec{n}| = R^2 \sin \vartheta, \\ \vec{n}^0 = \frac{\vec{n}}{|\vec{n}|} &= \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}. \\ dS = |\vec{n}| \cdot dA &= R^2 \sin \vartheta \cdot d\vartheta d\varphi, \\ d\vec{S} = \vec{n}^0 \cdot dS = \vec{n} \cdot dA &= \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} R^2 \sin \vartheta \cdot d\vartheta d\varphi. \end{aligned}$$

7.3 Volumenelement

Siehe auch:

Siegfried Großmann, Mathematischer Einführungskurs für die Physik, Teubner-Verlag, Stuttgart, Leipzig, 8. Auflage, 2000, Abschnitt 5.3.4. Wechsel der Variablen, Abschnitt 5.3.4.2. Die Funktionaldeterminante und Abschnitt 5.3.4.3. Die Transformation von Flächenelementen, Seite 203 bis Seite 207.

https://vhm.mathematik.uni-stuttgart.de/Vorlesungen/Mehrdimensionale_Integration/Folien_Mehrdimensionale_Integration.pdf

Volumina besitzen 3 freie Parameter. Es ist sofort klar, dass das Volumenelement

$$\text{in kartesischen Koordinaten } (x, y, z) : dV = dx dy dz$$

ist. Und schaut man sich die Abbildungen zum Volumenelement im „Großmann“ oder im „Weltner“⁶ an, so ist ebenfalls intuitiv klar, dass gilt

$$\text{in Kugelkoordinaten } (r, \vartheta, \varphi) : dV = dS \cdot dr = r^2 \sin \vartheta \cdot dr d\vartheta d\varphi .$$

Allgemein kann man sich das Volumenelement als infinitesimales Parallelepiped, auch Spat genannt, vorstellen, dessen Kanten gebildet werden von den infinitesimalen Tangentenvektoren an die Koordinatenlinien des betrachteten Koordinatensystems. Im (geradlinigen, global orthogonalen) kartesischen (x, y, z) -Koordinatensystem fallen die infinitesimalen Tangentenvektoren mit den Koordinatenlinien zusammen, sodass das Volumenelement dV ein Kubus mit dem Volumen $dV = dx \cdot dy \cdot dz$ ist.

Die Tangentenvektoren an die Koordinatenlinien nichtkartesischer (u, v, w) -Koordinatensysteme erhalten wir, falls die Koordinatentransformation $(x, y, z) = g(u, v, w)$ bzw.

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(u, v, w) \\ y(u, v, w) \\ z(u, v, w) \end{pmatrix}$$

bekannt ist, mit der Jacobi-Matrix J wie folgt:

$$J = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix} = \left(\frac{\partial \vec{r}(u, v, w)}{\partial u}, \frac{\partial \vec{r}(u, v, w)}{\partial v}, \frac{\partial \vec{r}(u, v, w)}{\partial w} \right),$$

$$J = (\vec{g}_u, \vec{g}_v, \vec{g}_w) .$$

Die Vektoren $\vec{g}_u, \vec{g}_v, \vec{g}_w$ sind die Tangentenvektoren an die Koordinatenlinien u, v, w und bilden die Spalten in der Jacobi-Matrix. Der Betrag des Spatprodukts, also

$$\left| (\vec{g}_u \times \vec{g}_v) \cdot \vec{g}_w \right| = |\det J| ,$$

aus diesen Tangentenvektoren ergibt das Volumen des Parallelepipeds im Verhältnis zum Volumen des kartesischen Kubus' $(\partial \vec{r} / \partial x \times \partial \vec{r} / \partial y) \cdot \partial \vec{r} / \partial z$ mit $\vec{r}(x, y, z) = (x, y, z)$:

$$\left| \det \left(\frac{\partial \vec{r}(x, y, z)}{\partial x}, \frac{\partial \vec{r}(x, y, z)}{\partial y}, \frac{\partial \vec{r}(x, y, z)}{\partial z} \right) \right| = \left| \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| = 1 .$$

Der Betrag der Determinante der Jacobi-Matrix, d. h. **der Betrag der Funktionaldeterminante ist der Skalierungsfaktor für das Volumenelement** gemäß

$$dV = 1 \cdot dx dy dz = dx dy dz = |\det J| \cdot du dv dw ,$$

⁶Klaus Weltner, Springerlehrbuch, Mathematik für Physiker 2, Springer-Verlag, Berlin, Heidelberg, New York, 12. Auflage, 2001, Abschnitt 15.4.3 Kugelkoordinaten, Seite 51.

oder ausführlich geschrieben:

$$\begin{aligned} dV &= \left| \left(\frac{\partial \vec{r}}{\partial x} dx \times \frac{\partial \vec{r}}{\partial y} dy \right) \cdot \frac{\partial \vec{r}}{\partial z} dz \right| = 1 \cdot dx dy dz \\ &= \left| \left(\frac{\partial \vec{r}}{\partial u} du \times \frac{\partial \vec{r}}{\partial v} dv \right) \cdot \frac{\partial \vec{r}}{\partial w} dw \right| = |\det J| \cdot du dv dw . \end{aligned}$$

Zu beachten ist, dass in den Volumenelementen (und auch in den Flächenelementen) der **Betrag der Funktionaldeterminante** verwendet werden muss, um ein orientierungsabhängiges Vorzeichen zu vermeiden.

Eine vereinfachende Besonderheit gibt es bei lokal orthogonalen Koordinatensystemen. In diesem Fall sind die Spalten der Jacobi-Matrix orthogonal zueinander, sodass für den Betrag der Funktionaldeterminante

$$|\det J| = |\vec{g}_u| \cdot |\vec{g}_v| \cdot |\vec{g}_w| = g_u \cdot g_v \cdot g_w$$

gilt. Angewandt auf Kugelkoordinaten ergibt dies mit

$$g_r = \left| \frac{\partial}{\partial r} \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix} \right| = 1, \quad g_\vartheta = \left| \frac{\partial}{\partial \vartheta} \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix} \right| = r, \quad g_\varphi = \left| \frac{\partial}{\partial \varphi} \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix} \right| = r \sin \vartheta$$

das Volumenelement

$$dV = |\vec{g}_r| |\vec{g}_\vartheta| |\vec{g}_\varphi| \cdot dr d\vartheta d\varphi = 1 \cdot r \cdot r \sin \vartheta \cdot dr d\vartheta d\varphi = r^2 \sin \vartheta \cdot dr d\vartheta d\varphi .$$

Gelegentlich hilfreiche Ergänzungen

- **Zur Funktionaldeterminante bei Flächenelementen im \mathbb{R}^2 :**

Flächen sind 2-parametrische geometrische Objekte, besitzen also zwei freie Parameter. Die Analogie zum Volumenelement im 3-dimensionalen Raum ist folglich das Flächenelement im 2-dimensionalen Raum bzw. in der Ebene mit der quadratischen **Jacobi-Matrix**

$$J = \begin{pmatrix} \frac{\partial x(u,v)}{\partial u} & \frac{\partial x(u,v)}{\partial v} \\ \frac{\partial y(u,v)}{\partial u} & \frac{\partial y(u,v)}{\partial v} \end{pmatrix} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} = \begin{pmatrix} \frac{\partial(x,y)}{\partial(u,v)} \end{pmatrix}$$

bzw. der zugehörigen **Inversen**

$$J^{-1} = \begin{pmatrix} \frac{\partial u(x,y)}{\partial x} & \frac{\partial u(x,y)}{\partial y} \\ \frac{\partial v(x,y)}{\partial x} & \frac{\partial v(x,y)}{\partial y} \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} \frac{\partial(u,v)}{\partial(x,y)} \end{pmatrix} .$$

Das Flächenelement im \mathbb{R}^2 ist dann

$$dS(u,v) = \left| \frac{\partial \vec{r}(u,v)}{\partial u} \times \frac{\partial \vec{r}(u,v)}{\partial v} \right| \cdot du dv = |\det J| \cdot du dv = \left| \det \begin{pmatrix} \frac{\partial(x,y)}{\partial(u,v)} \end{pmatrix} \right| \cdot du dv ,$$

was man sofort für den Fall ebener Polarkoordinaten verifizieren kann (siehe Abschnitt 6.1).

- **Beziehung zwischen Funktionaldeterminante und ihrer Inversen:**

Es gilt

$$dV(x,y,z) = dx dy dz = \left| \det \begin{pmatrix} \frac{\partial(x,y,z)}{\partial(u,v,w)} \end{pmatrix} \right| \cdot du dv dw = dV(u,v,w) . \quad (7.6)$$

Und mit

$$d\tilde{V}(u,v,w) = du dv dw = \left| \det \begin{pmatrix} \frac{\partial(u,v,w)}{\partial(x,y,z)} \end{pmatrix} \right| \cdot dx dy dz$$

resultiert aus (7.6)

$$\begin{aligned} dx \, dy \, dz &= \underbrace{\left| \det \left(\frac{\partial(x,y,z)}{\partial(u,v,w)} \right) \right| \cdot \left| \det \left(\frac{\partial(u,v,w)}{\partial(x,y,z)} \right) \right|}_{=1} \cdot dx \, dy \, dz = dx \, dy \, dz \\ \Rightarrow \left| \det \left(\frac{\partial(u,v,w)}{\partial(x,y,z)} \right) \right| &= \frac{1}{\left| \det \left(\frac{\partial(x,y,z)}{\partial(u,v,w)} \right) \right|} \end{aligned}$$

in Übereinstimmung mit

$$A \text{ invertierbar} \Rightarrow \det A^{-1} = (\det A)^{-1} = \frac{1}{\det A}.$$

Angewandt auf die Funktionaldeterminante $\det J$ hat dann diese Beziehung in vereinfachter Notation die folgende Gestalt:

$$\det J^{-1} = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = \frac{1}{\begin{vmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{vmatrix}} = \frac{1}{\det J}.$$

Außerdem sieht man, dass

$$d\tilde{V}(u, v, w) = du \, dv \, dw \neq dx \, dy \, dz = |\det J| \cdot du \, dv \, dw = dV(u, v, w).$$

• **Zum Skalierungsfaktor bei Flächenelementen im \mathbb{R}^3 :**

Weil Flächen nur zwei freie Parameter besitzen, gibt es für sie im \mathbb{R}^3 keine Funktionaldeterminante in der bereits diskutierten Form.⁷ Man kann aber auch für Flächen im \mathbb{R}^3 eine Determinante konstruieren, die den richtigen Skalierungsfaktor liefert, wenn wir davon ausgehen, dass der Skalierungsfaktor der Betrag eines Spatprodukts

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = (\vec{b} \times \vec{c}) \cdot \vec{a} = (\vec{c} \times \vec{a}) \cdot \vec{b}$$

sein soll, und wenn wir berücksichtigen, dass man ein Spatprodukt als Determinante 3. Ordnung schreiben kann:

$$\det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} = \det(\vec{a}, \vec{b}, \vec{c}) = (\vec{a} \times \vec{b}) \cdot \vec{c}.$$

Die Elemente der Determinante sind hier skalare Vektorkomponenten und dürfen nicht mit partiellen Ableitungen nach den Koordinaten x, y, z verwechselt werden. Wenn jetzt $\frac{\partial \vec{r}}{\partial u} = \vec{a}$ und $\frac{\partial \vec{r}}{\partial v} = \vec{b}$ die Tangentenvektoren an die Koordinatenlinien der Fläche sind und wenn \vec{n}^0 der nach außen gerichtete Einheitsnormalenvektor auf dieser Fläche ist gemäß

$$\vec{a} \times \vec{b} = \vec{n} \Rightarrow \vec{n}^0 = \frac{\vec{n}}{|\vec{n}|} = \frac{\vec{a} \times \vec{b}}{|\vec{a} \times \vec{b}|}, \quad |\vec{n}^0| = 1,$$

dann können wir damit ein Spatprodukt konstruieren, das uns den Skalierungsfaktor für das Flächenelement liefert:

$$\left| (\vec{a} \times \vec{b}) \cdot \vec{n}^0 \right| = |\vec{n} \cdot \vec{n}^0| = |\vec{n}| \cdot |\vec{n}^0| = |\vec{n}| = |\vec{a} \times \vec{b}| \Rightarrow$$

$$\boxed{\left| \frac{\partial \vec{r}(u, v)}{\partial u} \times \frac{\partial \vec{r}(u, v)}{\partial v} \right| = \left| \left(\frac{\partial \vec{r}(u, v)}{\partial u} \times \frac{\partial \vec{r}(u, v)}{\partial v} \right) \cdot \vec{n}^0 \right| = \left| \det \left(\frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v}, \vec{n}^0 \right) \right|}.$$

⁷Siehe Siegfried Großmann, Mathematischer Einführungskurs für die Physik, Teubner-Verlag, Stuttgart, Leipzig, 8. Auflage, 2000, Abschnitt 5.3.4.2. Die Funktionaldeterminante, Seite 205.

Sehr bequem lässt sich diese Determinante für die Kugeloberfläche ermitteln, denn Kugelkoordinaten sind lokal orthogonal. Deshalb steht der Vektor \vec{g}_r stets senkrecht auf den Tangentenvektoren \vec{g}_ϑ und \vec{g}_φ an die Koordinatenlinien ϑ und φ der Kugeloberfläche:

$$\vec{g}_r = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} = \vec{n}^0, \quad |\vec{g}_r| = g_r = 1.$$

Der Skalierungsfaktor für das Flächenelement der Kugeloberfläche mit $r = \text{const} = R$ ist damit

$$\begin{aligned} \left| \frac{\partial \vec{r}(\vartheta, \varphi)}{\partial \vartheta} \times \frac{\partial \vec{r}(\vartheta, \varphi)}{\partial \varphi} \right| &= \left| \det \left(\vec{g}_\vartheta, \vec{g}_\varphi, \vec{g}_r \right) \right| \\ &= \left| \det \begin{pmatrix} R \cos \vartheta \cos \varphi & -R \sin \vartheta \sin \varphi & \sin \vartheta \cos \varphi \\ R \cos \vartheta \sin \varphi & R \sin \vartheta \cos \varphi & \sin \vartheta \sin \varphi \\ -R \sin \vartheta & 0 & \cos \vartheta \end{pmatrix} \right| = R^2 \sin \vartheta. \end{aligned}$$

- Das **Raumwinkelement** $d\Omega$ ist

$$d\Omega := \sin \vartheta \, d\vartheta \, d\varphi,$$

sodass das Integral über den gesamten Raum den **vollen Raumwinkel**

$$\int_{\vartheta=0}^{\pi} \int_{\varphi=0}^{2\pi} d\Omega = \int_{\vartheta=0}^{\pi} \int_{\varphi=0}^{2\pi} \sin \vartheta \, d\vartheta \, d\varphi = 4\pi$$

ergibt. Der Raumwinkel Ω ist nämlich über die Gesamtoberfläche A_{OK} einer Kugel mit dem Radius R und über deren Teiloberflächenstücke A wie folgt definiert:

$$\Omega(\text{gesamter Raum}) = \frac{A_{\text{OK}}}{R^2} = \frac{4\pi R^2}{R^2} = 4\pi \quad \Rightarrow \quad \Omega := \frac{A}{R^2}.$$

8 Green'sche Identitäten

Die Green'schen Identitäten werden auch Green'sche Theoreme genannt. Der Vollständig halber geben wir zu Beginn den

$$\text{Satz von Stokes} \quad \int_S \text{rot} \vec{F} \cdot d\vec{S} = \oint_{\partial S} \vec{F} \cdot d\vec{r}$$

und den

$$\text{Gauß'schen Integralsatz} \quad \int_V \text{div} \vec{F} \, dV = \oint_{\partial V} \vec{F} \cdot d\vec{S}$$

an.

Die Funktionen Φ und Ψ beschreiben Skalarfelder. Φ sei mindestens ein mal und Ψ mindestens zwei mal stetig differenzierbar. Dann ergibt sich die 1. Green'sche Identität aus dem Gauß'schen Integralsatz wie folgt:

$$\oint_{\partial V} (\Phi \nabla \Psi) \cdot d\vec{S} = \int_V \nabla \cdot (\Phi \nabla \Psi) \, dV = \int_V (\Phi \nabla^2 \Psi + \nabla \Phi \cdot \nabla \Psi) \, dV \Rightarrow$$

$$\text{1. Green'sche Identität} \quad \oint_{\partial V} (\Phi \nabla \Psi) \cdot d\vec{S} = \int_V (\Phi \Delta \Psi + \nabla \Phi \cdot \nabla \Psi) \, dV .$$

Jetzt seien sowohl Φ als auch Ψ mindestens zwei mal stetig differenzierbar. Dann ergibt sich die 2. Green'sche Identität aus der 1. Green'schen Identität wie folgt:

$$\oint_{\partial V} (\Phi \nabla \Psi) \cdot d\vec{S} = \int_V (\Phi \Delta \Psi + \nabla \Phi \cdot \nabla \Psi) \, dV , \quad (8.1)$$

$$\oint_{\partial V} (\Psi \nabla \Phi) \cdot d\vec{S} = \int_V (\Psi \Delta \Phi + \nabla \Psi \cdot \nabla \Phi) \, dV \Rightarrow \quad (8.2)$$

Wir subtrahieren (8.2) von (8.1) und erhalten mit $\nabla \Phi \cdot \nabla \Psi = \nabla \Psi \cdot \nabla \Phi$ die

$$\text{2. Green'sche Identität} \quad \oint_{\partial V} (\Phi \nabla \Psi - \Psi \nabla \Phi) \cdot d\vec{S} = \int_V (\Phi \Delta \Psi - \Psi \Delta \Phi) \, dV .$$

9 Differentialoperatoren in krummlinig-orthogonalen Koordinatensystemen

(Nach dem Lehrbuch aus der Teubner Studienbücherei Physik: Siegfried Großmann, Mathematischer Einführungskurs für die Physik, Teubner-Verlag, Stuttgart, Leipzig, 8. Auflage, 2000, Abschnitt 7.2. *Differentialoperatoren in krummlinig-orthogonalen Koordinaten*, Seite 256 bis Seite 258.)

In diesem Kapitel verallgemeinern wir die Differentialoperatoren des geradlinig-orthogonalen (kartesischen) (x, y, z) -Koordinatensystems auf die Differentialoperatoren in krummlinig-orthogonalen (u, v, w) -Koordinatensystemen.

9.1 Gradient

In kartesischen Koordinaten ist der Gradient einer skalaren Funktion bzw. eines skalaren Feldes $f = f(\vec{r}) = f(x, y, z)$ definiert durch

$$\text{grad } f = \frac{\partial f}{\partial x} \cdot \vec{e}_x + \frac{\partial f}{\partial y} \cdot \vec{e}_y + \frac{\partial f}{\partial z} \cdot \vec{e}_z =: \frac{df}{d\vec{r}}. \quad (9.1)$$

Bei $\frac{df}{d\vec{r}}$ handelt es sich um eine symbolische Schreibweise, denn sie ist inkonsistent mit dem Vektorkalkül. Das Skalarprodukt

$$\text{grad } f \cdot d\vec{r} = df$$

beschreibt die Feldänderung bei Verschiebung $d\vec{r}$.

Der Gradient ist also ein Vektor, der die Änderung der Feldgröße f pro Verschiebung (Länge) beschreibt und in die Richtung der maximalen Änderung von f zeigt. Das aber bedeutet, dass die u -Komponente grad_u die Projektion des Gradienten(vektors) auf den Basiseinheitsvektor \vec{e}_u ist:

$$\text{grad}_u = \text{grad } f \cdot \vec{e}_u. \quad (9.2)$$

Mit

$$\vec{e}_u = \frac{\frac{\partial \vec{r}}{\partial u}}{\left| \frac{\partial \vec{r}}{\partial u} \right|} = \frac{1}{g_u} \frac{\partial \vec{r}}{\partial u} = \frac{1}{g_u} \left(\frac{\partial x}{\partial u} \vec{e}_x + \frac{\partial y}{\partial u} \vec{e}_y + \frac{\partial z}{\partial u} \vec{e}_z \right)$$

resultiert aus (9.2) das Skalarprodukt bzw. die skalare Vektorkomponente des Gradienten von f in Richtung der u -Koordinate:

$$\begin{aligned} \text{grad}_u &= \left(\frac{\partial f}{\partial x} \cdot \vec{e}_x + \frac{\partial f}{\partial y} \cdot \vec{e}_y + \frac{\partial f}{\partial z} \cdot \vec{e}_z \right) \cdot \frac{1}{g_u} \left(\frac{\partial x}{\partial u} \vec{e}_x + \frac{\partial y}{\partial u} \vec{e}_y + \frac{\partial z}{\partial u} \vec{e}_z \right) \\ &= \frac{1}{g_u} \cdot \underbrace{\left(\frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial u} \right)}_{= \frac{\partial f}{\partial u}}. \end{aligned}$$

Der unterklammerte Term ergibt nicht $3 \frac{\partial f}{\partial u}$ sondern $\frac{\partial f}{\partial u}$ wie beim totalen Differential gemäß der **Kettenregel**, sodass wir für die skalare Vektorkomponente in u -Richtung

$$\text{grad}_u = \frac{1}{g_u} \cdot \frac{\partial f}{\partial u}$$

erhalten. Die skalaren Gradientenkomponenten in v - und in w -Richtung ergeben sich völlig analog. Mit diesen drei Komponenten ist dann der Gradient in den orthogonalen Koordinaten u, v, w

$$\text{grad } f(\vec{r}(u, v, w)) = \frac{1}{g_u} \frac{\partial f}{\partial u} \vec{e}_u + \frac{1}{g_v} \frac{\partial f}{\partial v} \vec{e}_v + \frac{1}{g_w} \frac{\partial f}{\partial w} \vec{e}_w.$$

Wählen wir für die allgemeinen orthogonalen Koordinaten u, v, w speziell die kartesischen Koordinaten x, y, z mit $g_x = g_y = g_z = 1$, so resultiert erwartungsgemäß (9.1).

9.2 Divergenz

Die folgende Herleitung ist analog zu der Herleitung der Divergenz für kartesische Koordinaten im Kapitel 15 meines *Basisskripts zur Infinitesimalrechnung*. Im Folgenden verwenden wir gelegentlich den Laufindex $i \in \{1, 2, 3\}$ am Koordinatensymbol u , sodass $u \equiv u_1, v \equiv u_2, w \equiv u_3$.

Wenn der infinitesimale Abstandsvektor in orthogonalen Koordinaten

$$d\vec{r} = \frac{\partial \vec{r}}{\partial u} du + \frac{\partial \vec{r}}{\partial v} dv + \frac{\partial \vec{r}}{\partial w} dw$$

ist und für die entsprechenden orthogonalen Basiseinheitsvektoren

$$\vec{e}_{u_i} = \frac{1}{g_{u_i}} \frac{d\vec{r}}{du_i} \Leftrightarrow \frac{d\vec{r}}{du_i} = \vec{e}_{u_i} g_{u_i}$$

gilt, dann können wir für den infinitesimalen Abstandsvektor unter Verwendung der Basiseinheitsvektoren

$$d\vec{r} = \vec{e}_u g_u du + \vec{e}_v g_v dv + \vec{e}_w g_w dw$$

schreiben und das Volumenelement ist

$$dV = g_u du \cdot g_v dv \cdot g_w dw = g_u g_v g_w \cdot du dv dw .$$

Und für das Volumen bei kleinen Abständen gilt dann

$$\Delta V \approx g_u g_v g_w \cdot \Delta u \Delta v \Delta w .$$

Für den Fluss des Vektors $\vec{F}(u, v, w) = (F_u(u, v, w), F_v(u, v, w), F_w(u, v, w))$ durch die einhüllende Fläche $S = \partial V$ eines quaderanalogen Bereichs V , dessen Kanten gebildet werden von den krummlinig-orthogonalen (u, v, w) -Koordinatenlinien, gilt

$$\begin{aligned} \oint_S \vec{F} \cdot \vec{n}^0 dS &= \int_{v_0}^{v_0+\Delta v} \int_{w_0}^{w_0+\Delta w} [F_u(u_0 + \Delta u, v, w) - F_u(u_0, v, w)] g_v dv g_w dw \\ &+ \int_{w_0}^{w_0+\Delta w} \int_{u_0}^{u_0+\Delta u} [F_v(u, v_0 + \Delta v, w) - F_v(u, v_0, w)] g_w dw g_u du \\ &+ \int_{u_0}^{u_0+\Delta u} \int_{v_0}^{v_0+\Delta v} [F_w(u, v, w_0 + \Delta w) - F_w(u, v, w_0)] g_u du g_v dv \\ &= \int_{v_0}^{v_0+\Delta v} \int_{w_0}^{w_0+\Delta w} [g_v g_w F_u(u_0 + \Delta u, v, w) - g_v g_w F_u(u_0, v, w)] dv dw \quad (9.3) \\ &+ \int_{w_0}^{w_0+\Delta w} \int_{u_0}^{u_0+\Delta u} [g_w g_u F_v(u, v_0 + \Delta v, w) - g_w g_u F_v(u, v_0, w)] dw du \\ &+ \int_{u_0}^{u_0+\Delta u} \int_{v_0}^{v_0+\Delta v} [g_u g_v F_w(u, v, w_0 + \Delta w) - g_u g_v F_w(u, v, w_0)] du dv . \end{aligned}$$

Die Taylor-Entwicklung beispielsweise des Integranden in (9.3) an der Stelle u_0 liefert

$$\begin{aligned} &[g_v g_w F_u(u_0 + \Delta u, v, w) - g_v g_w F_u(u_0, v, w)] \\ &= g_v g_w F_u(u_0, v, w) + \frac{\partial}{\partial u} [g_v g_w F_u(u_0, v, w)] \Delta u + \frac{1}{2} \frac{\partial^2}{\partial u^2} [g_v g_w F_u(u_0, v, w)] \Delta u^2 + \dots \\ &\quad - g_v g_w F_u(u_0, v, w) , \end{aligned}$$

$$\left[g_v g_w F_u(u_0 + \Delta u, v, w) - g_v g_w F_u(u_0, v, w) \right] = \frac{\partial}{\partial u} \left[g_v g_w F_u(u_0, v, w) \right] \Delta u \dots$$

Die Glieder höherer Ordnung haben wir deshalb nicht mehr aufgeführt, weil alle Glieder der Ordnung > 1 bei der später erfolgenden Grenzwertbildung mit $\Delta u_i \rightarrow 0$ verschwinden. Diese Taylor-Entwicklung, die analog auch für die beiden übrigen Teilintegranden gilt, setzen wir in das Flussintegral ein:

$$\begin{aligned} \oint_S \vec{F} \cdot \vec{n}^0 dS &= \int_{v_0}^{v_0 + \Delta v} \int_{w_0}^{w_0 + \Delta w} \left\{ \frac{\partial}{\partial u} \left[g_v g_w F_u(u_0, v, w) \right] \Delta u \dots \right\} dv dw \\ &+ \int_{w_0}^{w_0 + \Delta w} \int_{u_0}^{u_0 + \Delta u} \left\{ \frac{\partial}{\partial v} \left[g_w g_u F_v(u, v_0, w) \right] \Delta v \dots \right\} dw du \\ &+ \int_{u_0}^{u_0 + \Delta u} \int_{v_0}^{v_0 + \Delta v} \left\{ \frac{\partial}{\partial w} \left[g_u g_v F_w(u, v, w_0) \right] \Delta w \dots \right\} du dv . \end{aligned}$$

Darauf wenden wir den **Mittelwertsatz der Integralrechnung**

$$\int_{x_0}^{x_1} f(x) dx = (x_1 - x_0) f(\tilde{x}) = f(\tilde{x}) \Delta x$$

an und erhalten

$$\begin{aligned} \oint_S \vec{F} \cdot \vec{n}^0 dS &= \frac{\partial}{\partial u} \left[g_v g_w F_u(u_0, \tilde{v}, \tilde{w}) \right] \Delta u \Delta v \Delta w \dots \\ &+ \frac{\partial}{\partial v} \left[g_w g_u F_v(\tilde{u}, v_0, \tilde{w}) \right] \Delta u \Delta v \Delta w \dots \\ &+ \frac{\partial}{\partial w} \left[g_u g_v F_w(\tilde{u}, \tilde{v}, w_0) \right] \Delta u \Delta v \Delta w \dots . \end{aligned}$$

Die mittlere Quelledichte von \vec{F} für den quaderanalogen Bereich V ist folglich

$$\begin{aligned} \frac{1}{\Delta V} \oint_S \vec{F} \cdot \vec{n}^0 dS &\approx \frac{1}{g_u g_v g_w \cdot \Delta u \Delta v \Delta w} \oint_S \vec{F} \cdot \vec{n}^0 dS \approx \frac{1}{g_u g_v g_w} \frac{\partial}{\partial u} \left[g_v g_w F_u(u_0, \tilde{v}, \tilde{w}) \right] \dots \\ &+ \frac{1}{g_u g_v g_w} \frac{\partial}{\partial v} \left[g_w g_u F_v(\tilde{u}, v_0, \tilde{w}) \right] \dots \\ &+ \frac{1}{g_u g_v g_w} \frac{\partial}{\partial w} \left[g_u g_v F_w(\tilde{u}, \tilde{v}, w_0) \right] \dots . \end{aligned}$$

Wenn wir jetzt das Volumen ΔV gegen Null gehen lassen gemäß

$$\Delta V \rightarrow 0 \Rightarrow \begin{cases} \Delta u \rightarrow 0 & \tilde{u} \rightarrow u_0 \\ \Delta v \rightarrow 0 & \tilde{v} \rightarrow v_0 \\ \Delta w \rightarrow 0 & \tilde{w} \rightarrow w_0 \end{cases} ,$$

erhalten wir **exakt** die (lokale) Quelledichte bzw. die Divergenz von \vec{F} für den Raumpunkt $\vec{r}_0 = (u_0, v_0, w_0)$:

$$\begin{aligned} \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \oint_S \vec{F} \cdot \vec{n}^0 dS &= \operatorname{div} \vec{F}(\vec{r}_0) \\ &= \frac{1}{g_u g_v g_w} \left\{ \frac{\partial}{\partial u} \left[g_v g_w F_u(u_0, v_0, w_0) \right] + \frac{\partial}{\partial v} \left[g_w g_u F_v(u_0, v_0, w_0) \right] + \frac{\partial}{\partial w} \left[g_u g_v F_w(u_0, v_0, w_0) \right] \right\} . \end{aligned}$$

Weil aber diese Divergenzformel für jeden Raumpunkt \vec{r} im Vektorfeld $\vec{F}(\vec{r})$ gelten soll, können wir \vec{r}_0 durch \vec{r} ersetzen und erhalten schließlich

$$\operatorname{div} \vec{F}(\vec{r}(u, v, w)) = \frac{1}{g_u g_v g_w} \left[\frac{\partial}{\partial u} (g_v g_w F_u) + \frac{\partial}{\partial v} (g_w g_u F_v) + \frac{\partial}{\partial w} (g_u g_v F_w) \right] .$$

9.3 Laplace-Operator

In kartesischen Koordinaten x, y, z ist der Laplace-Operator die Anwendung des Divergenz-Operators auf den Nabla-Operator („Gradient-Operator“) gemäß⁸

$$\Delta := \operatorname{div} \operatorname{grad} = \vec{\nabla}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Dabei darf man das Laplace-Operator-Symbol Δ nicht mit dem Differenz-Delta verwechseln. Auf den Vektorpfeil über dem Nabla-Symbol ∇ für den Gradienten wird meistens verzichtet, weil von vornherein klar ist, dass es sich beim Gradienten um einen Vektor handelt. Der Laplace-Operator beinhaltet also zuerst die Bildung des Gradienten $\vec{F}(\vec{r})$ einer Funktion bzw. eines skalaren Feldes⁹ $f(\vec{r})$ und daran anschließend die Bildung der Divergenz des Gradienten bzw. Vektorfeldes $\vec{F}(\vec{r})$. Das gilt in dieser Weise natürlich auch für krummlinig-orthogonale Koordinaten:

$$\operatorname{grad} f(u, v, w) = (F_u, F_v, F_w) = \vec{F}(u, v, w) \quad \longrightarrow \quad \operatorname{div} \vec{F}(u, v, w) = \Delta f(u, v, w).$$

Um den Laplace-Operator in krummlinig-orthogonalen Koordinaten zu erhalten, brauchen wir also nur die Ergebnisse der Abschnitte (9.1) und (9.2) zu kombinieren:

$$\begin{aligned} \operatorname{grad} f &= \frac{1}{g_u} \frac{\partial f}{\partial u} \vec{e}_u + \frac{1}{g_v} \frac{\partial f}{\partial v} \vec{e}_v + \frac{1}{g_w} \frac{\partial f}{\partial w} \vec{e}_w \\ &= F_u \vec{e}_u + F_v \vec{e}_v + F_w \vec{e}_w, \end{aligned}$$

$$\operatorname{div} \vec{F} = \frac{1}{g_u g_v g_w} \left[\frac{\partial}{\partial u} (g_v g_w F_u) + \frac{\partial}{\partial v} (g_w g_u F_v) + \frac{\partial}{\partial w} (g_u g_v F_w) \right] \Rightarrow$$

$$\Delta f = \operatorname{div} \operatorname{grad} f(\vec{r}(u, v, w)) = \frac{1}{g_u g_v g_w} \left[\frac{\partial}{\partial u} \left(\frac{g_v g_w}{g_u} \frac{\partial f}{\partial u} \right) + \frac{\partial}{\partial v} \left(\frac{g_w g_u}{g_v} \frac{\partial f}{\partial v} \right) + \frac{\partial}{\partial w} \left(\frac{g_u g_v}{g_w} \frac{\partial f}{\partial w} \right) \right].$$

⁸ $\vec{\nabla}^2 = \vec{\nabla} \cdot \vec{\nabla} = \left(\frac{\partial}{\partial x} \vec{e}_x + \frac{\partial}{\partial y} \vec{e}_y + \frac{\partial}{\partial z} \vec{e}_z \right) \cdot \left(\frac{\partial}{\partial x} \vec{e}_x + \frac{\partial}{\partial y} \vec{e}_y + \frac{\partial}{\partial z} \vec{e}_z \right) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$ \square

⁹Es gibt auch den Vektorgradienten, d. h. den Gradienten eines Vektorfeldes. Das schon an dieser Stelle zu diskutieren, würde zu weit führen.

9.4 Rotation

Die folgende Herleitung ist analog zu der Herleitung der Rotation für kartesische Koordinaten im Kapitel 16 meines *Basisskripts zur Infinitesimalrechnung*. Im Folgenden verwenden wir gelegentlich den Laufindex $i \in \{1, 2, 3\}$ am Koordinatensymbol u , sodass $u \equiv u_1, v \equiv u_2, w \equiv u_3$.

Die Rotation ist ein Vektor bzw. ein Vektorfeld. In krummlinig-orthogonalen Koordinaten u, v, w besitzt die Rotation $\text{rot } \vec{F}$ des Vektorfeldes $\vec{F}(\vec{r}(u, v, w))$ die

$$\text{Vektorkomponenten } \vec{e}_{u_i} \cdot \text{rot}_{u_i} \vec{F}.$$

Das bedeutet, dass die

$$\text{rot}_{u_i} \vec{F} = (\text{rot } \vec{F}) \cdot \vec{e}_{u_i}$$

die skalaren Vektorkomponenten sind. Weil die Herleitung der einzelnen skalaren Vektorkomponenten der Rotation völlig analog zueinander ist, werden wir nur die w -Komponente der Rotation herleiten und die beiden übrigen Komponenten schlussfolgern.

Mit dem vektoriellen Linienelement

$$d\vec{r} = \vec{e}_u \cdot g_u du + \vec{e}_v \cdot g_v dv + \vec{e}_w \cdot g_w dw$$

gilt in krummlinig-orthogonalen Koordinaten für das vektorielle Linienintegral in einem Vektorfeld $\vec{F}(\vec{r})$ (siehe Arbeitsintegral im Abschnitt 7.1.2)

$$\int \vec{F} \cdot d\vec{r} = \int (F_u \cdot g_u du + F_v \cdot g_v dv + F_w \cdot g_w dw).$$

Betrachten wir in einem krummlinig-orthogonalen (u, v, w) -Koordinatensystem ein rechteckanalages Flächenstück S_w . Dieses werde bei $w_0 = \text{const}$ im mathematisch positiven Umlaufsinn umrandet von der geschlossenen Kurve $\partial S_w = C_{uv}$, die wiederum aus den folgenden vier Teilkurven besteht:

von (u_0, v_0, w_0)	längs u bis	$(u_0 + \Delta u, v_0, w_0)$,
von $(u_0 + \Delta u, v_0, w_0)$	längs v bis	$(u_0 + \Delta u, v_0 + \Delta v, w_0)$,
von $(u_0 + \Delta u, v_0 + \Delta v, w_0)$	längs u bis	$(u_0, v_0 + \Delta v, w_0)$,
von $(u_0, v_0 + \Delta v, w_0)$	längs v bis	(u_0, v_0, w_0) .

Damit bilden wir das Linienintegral oder genauer gesagt das Umlaufintegral bzw. die Zirkulation von \vec{F} längs C_{uv} :

$$\begin{aligned} \oint_{C_{uv}} \vec{F} \cdot d\vec{r} &= \int_{u_0}^{u_0 + \Delta u} F_u(u, v_0, w_0) \cdot g_u du + \int_{v_0}^{v_0 + \Delta v} F_v(u_0 + \Delta u, v, w_0) \cdot g_v dv \\ &+ \int_{u_0 + \Delta u}^{u_0} F_u(u, v_0 + \Delta v, w_0) \cdot g_u du + \int_{v_0 + \Delta v}^{v_0} F_v(u_0, v, w_0) \cdot g_v dv, \end{aligned}$$

$$\begin{aligned} \oint_{C_{uv}} \vec{F} \cdot d\vec{r} &= \int_{v_0}^{v_0 + \Delta v} \left[g_v F_v(u_0 + \Delta u, v, w_0) - g_v F_v(u_0, v, w_0) \right] dv \\ &- \int_{u_0}^{u_0 + \Delta u} \left[g_u F_u(u, v_0 + \Delta v, w_0) - g_u F_u(u, v_0, w_0) \right] du. \end{aligned}$$

Analog zur Herleitung der Divergenz liefert die Taylor-Entwicklung der Integranden

$$\oint_{C_{uv}} \vec{F} \cdot d\vec{r} = \int_{v_0}^{v_0 + \Delta v} \left\{ \frac{\partial}{\partial u} [g_v F_v(u_0, v, w_0)] \Delta u + \dots \right\} dv - \int_{u_0}^{u_0 + \Delta u} \left\{ \frac{\partial}{\partial v} [g_u F_u(u, v_0, w_0)] \Delta v + \dots \right\} du. \quad (9.4)$$

Auch bei dieser Taylor-Entwicklung brauchen wir die Glieder der Ordnung > 1 nicht zu berücksichtigen, weil diese am Ende bei der Grenzwertbildung $\Delta S \rightarrow 0$ verschwinden. Mit dem Mittelwertsatz der Integralrechnung schreiben wir für (9.4)

$$\oint_{C_{uv}} \vec{F} \cdot d\vec{r} = \left\{ \frac{\partial}{\partial u} [g_v F_v(u_0, \tilde{v}, w_0)] \Delta u + \dots \right\} \Delta v - \left\{ \frac{\partial}{\partial v} [g_u F_u(\tilde{u}, v_0, w_0)] \Delta v + \dots \right\} \Delta u.$$

Die Division der Zirkulation durch den Flächeninhalt¹⁰

$$\Delta S_w \approx g_u g_v \cdot \Delta u \Delta v$$

des rechteckanalogen Flächenstücks S_w ergibt die mittlere Zirkulationsflächendichte von \vec{F} in S_w :

$$\frac{1}{\Delta S_w} \oint_{C_{uv}} \vec{F} \cdot d\vec{r} \approx \frac{1}{g_u g_v} \left\{ \frac{\partial}{\partial u} [g_v F_v(u_0, \tilde{v}, w_0)] + \dots - \frac{\partial}{\partial v} [g_u F_u(\tilde{u}, v_0, w_0)] + \dots \right\}.$$

Wenn wir schließlich den Flächeninhalt ΔS_w gegen Null gehen lassen bzw. die geschlossene Kurve C_{uv} auf den Punkt $\vec{r}_0 = (u_0, v_0, w_0)$ zusammenziehen gemäß

$$\left. \begin{array}{l} \Delta u \\ \Delta v \end{array} \right\} \rightarrow 0, \quad \begin{array}{l} \tilde{u} \rightarrow u_0 \\ \tilde{v} \rightarrow v_0 \end{array},$$

erhalten wir **exakt** die (lokale) Zirkulationsflächendichte in S_w bzw. die skalare w -Komponente der Rotation von \vec{F} im Raumpunkt $(u_0, v_0, w_0) = \vec{r}_0$:

$$\lim_{\Delta S_w \rightarrow 0} \frac{1}{\Delta S_w} \oint_{C_{uv}} \vec{F} \cdot d\vec{r} = \frac{1}{g_u g_v} \left\{ \frac{\partial}{\partial u} [g_v F_v(u_0, v_0, w_0)] - \frac{\partial}{\partial v} [g_u F_u(u_0, v_0, w_0)] \right\} = \text{rot}_w \vec{F}(\vec{r}_0).$$

Weil die Rotation für das Vektorfeld \vec{F} in der gesamten von C_{uv} umrandeten Fläche gelten soll, ersetzen wir wieder \vec{r}_0 durch \vec{r} . Durch zyklische Vertauschung von u, v, w erhalten wir schließlich auch die u - und v -Komponente der Rotation und somit das Rotationsvektorfeld

$$\begin{aligned} \text{rot } \vec{F}(\vec{r}(u, v, w)) &= \frac{1}{g_v g_w} \left[\frac{\partial}{\partial v} (g_w \cdot F_w) - \frac{\partial}{\partial w} (g_v \cdot F_v) \right] \vec{e}_u \\ &+ \frac{1}{g_w g_u} \left[\frac{\partial}{\partial w} (g_u \cdot F_u) - \frac{\partial}{\partial u} (g_w \cdot F_w) \right] \vec{e}_v \\ &+ \frac{1}{g_u g_v} \left[\frac{\partial}{\partial u} (g_v \cdot F_v) - \frac{\partial}{\partial v} (g_u \cdot F_u) \right] \vec{e}_w. \end{aligned}$$

Dies lässt sich auch kompakt als Determinante schreiben:

$$\text{rot } \vec{F} = \begin{vmatrix} \vec{e}_u & \vec{e}_v & \vec{e}_w \\ g_v g_w & g_w g_u & g_u g_v \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ g_u \cdot F_u & g_v \cdot F_v & g_w \cdot F_w \end{vmatrix}.$$

¹⁰Weil wir den Mittelwertsatz wie üblich mit dem „Differenz- Δ “ in ΔS_w anwenden, haben wir von der Benutzung des Flächenelements $dS_w = g_u g_v \cdot du dv$ Abstand genommen.

10 Konservatives Vektorfeld, Potentialfunktion, Gravitationsfeld

Siehe auch *Ableitungen von skalaren Feldern – Der Gradient* unter https://hp.physnet.uni-hamburg.de/pfannkuche/Physik_1_WS0607/vorlesung/vorlesung10.pdf

In diesem Abschnitt beschränken wir uns in der allgemeinen Darstellung der Problematik auf kartesische Koordinaten. Weiterhin steht dieser Abschnitt inhaltlich in engem Zusammenhang mit dem Abschnitt 3 zum totalen Differential.

Die Bildung des Gradienten (grad) eines skalaren Feldes, beispielsweise $\Phi(\vec{r})$, erfolgt mit dem Nabla-Operator $\vec{\nabla}$, vereinfacht ∇ . Dabei resultiert das Gradientenfeld, ein spezielles Vektorfeld wie beispielsweise $\vec{V}(\vec{r})$, auf folgende Weise:

$$\text{grad } \Phi(\vec{r}) = \nabla \Phi = \left(\vec{e}_x \cdot \frac{\partial}{\partial x} + \vec{e}_y \cdot \frac{\partial}{\partial y} + \vec{e}_z \cdot \frac{\partial}{\partial z} \right) \Phi(\vec{r}) = \begin{pmatrix} \frac{\partial \Phi}{\partial x} \\ \frac{\partial \Phi}{\partial y} \\ \frac{\partial \Phi}{\partial z} \end{pmatrix} = \vec{V}(\vec{r}), \quad (10.1)$$

$$\vec{V}(\vec{r}) = \text{grad } \Phi(\vec{r}) = \nabla \Phi. \quad (10.2)$$

Der Gradient ist ein Vektor, der in die Richtung des stärksten Wachstums im zugehörigen skalaren Feld zeigt. **Alle Gradientenfelder sind konservativ**, denn ausgehend von (10.1) gilt mit¹¹

$$\vec{V} = \nabla \Phi = \begin{pmatrix} \frac{\partial \Phi}{\partial x} \\ \frac{\partial \Phi}{\partial y} \\ \frac{\partial \Phi}{\partial z} \end{pmatrix} = \begin{pmatrix} \Phi_x \\ \Phi_y \\ \Phi_z \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix}$$

und mit den gemischten zweiten partiellen Ableitungen von Φ wie beispielsweise $\frac{\partial^2 \Phi}{\partial z \partial y} = \Phi_{zy} = \frac{\partial w}{\partial y}$:

$$\left. \begin{array}{l} \Phi_{zy} = \Phi_{yz} \\ \Phi_{xz} = \Phi_{zx} \\ \Phi_{yx} = \Phi_{xy} \end{array} \right\} \Rightarrow \left. \begin{array}{l} \frac{\partial w}{\partial y} = \frac{\partial v}{\partial z} \\ \frac{\partial u}{\partial z} = \frac{\partial w}{\partial x} \\ \frac{\partial v}{\partial x} = \frac{\partial u}{\partial y} \end{array} \right\} \Leftrightarrow \left. \begin{array}{l} \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} = 0 \\ \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} = 0 \\ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0 \end{array} \right\} \Rightarrow$$

$$\begin{pmatrix} \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \\ \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \\ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \vec{V} = \nabla \times \vec{V} = \text{rot } \vec{V} = \vec{0}.$$

Zusammenfassend können wir feststellen:

$$\boxed{\vec{V}(\vec{r}) \text{ konservativ} \Leftrightarrow \text{rot } \vec{V} = \vec{0}}.$$

„Wenn Φ mindestens zweimal stetig differenzierbar ist, dann sind die“ gemischten zweiten „Ableitungen vertauschbar; ... und ... $\nabla \times \vec{V}$ verschwindet. Konservative Vektorfelder sind dann also wirbelfrei.“¹² Umgekehrt sind **Wirbelfelder**, d. h. Vektorfelder $\vec{V}(\vec{r})$ mit $\text{rot } \vec{V} \neq \vec{0}$, **nicht konservativ**. So existiert beispielsweise zum Wirbelfeld

$$\vec{V}(\vec{r}) = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} \Rightarrow \text{rot } \vec{V} = 2$$

keine skalare Funktion $\Phi(\vec{r})$, für die $\nabla \Phi = \vec{V}$ gilt.

¹¹**Achtung!** Beispielsweise für die skalare Vektorkomponente $\Phi_y = V_y$ kennzeichnet der Index y bei Φ_y die partielle Ableitung von Φ nach y , bei V_y jedoch die y -Komponente von \vec{V} .

¹²Zitiert aus dem Spektrum-Hochschultaschenbuch von Christian B. Lang, Norbert Pucker, Mathematische Methoden in der Physik, Spektrum Akademischer Verlag, Heidelberg, Berlin, 1998, Seite 293.

Aus der Konservativität der Gradientenfelder folgt, dass Linienintegrale längs eines Weges von einem Anfangspunkt \vec{r}_1 zu einem Endpunkt \vec{r}_2 **wegunabhängig** sind und dass Linienintegrale längs eines geschlossenen Weges, also für $\vec{r}_1 \circlearrowright \vec{r}_2 = \vec{r}_1$, verschwinden:

$$\oint \nabla\Phi(\vec{r}) \cdot d\vec{r} = \oint \vec{V}(\vec{r}) \cdot d\vec{r} = 0 .$$

Das erkennt man sofort wie folgt: Das Linienintegral längs eines beliebigen Weges von \vec{r}_1 nach \vec{r}_2 liefert

$$\int_{\vec{r}_1}^{\vec{r}_2} \vec{V} \cdot d\vec{r} = \Phi(\vec{r}_2) - \Phi(\vec{r}_1)$$

und das Linienintegral längs eines beliebigen Rückwegs von \vec{r}_2 nach \vec{r}_1 liefert dann

$$\int_{\vec{r}_2}^{\vec{r}_1} \vec{V} \cdot d\vec{r} = \Phi(\vec{r}_1) - \Phi(\vec{r}_2) .$$

In der Summe, also für den geschlossenen Weg, verschwindet folglich das Integral.

Weil das Linienintegral nur von den Integrationsgrenzen und nicht vom Integrationsweg abhängt, vereinfacht sich die Umkehroperation zur Gradientenbildung. Wegen der Wegunabhängigkeit können wir dabei nämlich einen beliebigen Integrationsweg längs der Koordinatenlinien x, y, z wählen und kommen so ohne Parametrisierung des Integrationsweges aus:

$$\int \vec{V}(\vec{r}) \cdot d\vec{r} = \int \left(V_x(\vec{r})\vec{e}_x + V_y(\vec{r})\vec{e}_y + V_z(\vec{r})\vec{e}_z \right) \cdot \left(dx\vec{e}_x + dy\vec{e}_y + dz\vec{e}_z \right) ,$$

$$\int \vec{V} \cdot d\vec{r} = \int \left(V_x dx + V_y dy + V_z dz \right) = \Phi .$$

Achtung Vorzeichen!

$\Phi(\vec{r})$ ist gemäß (10.2) das skalare Feld zum Gradientenfeld $\vec{V}(\vec{r})$.

$\Phi(\vec{r})$ ist aber das **Potentialfeld** bzw. die **Potentialfunktion** wenn im Vergleich zu (10.2) gilt:

$$\vec{V}(\vec{r}) = -\text{grad } \Phi(\vec{r}) = -\nabla\Phi$$

\Rightarrow Der Gradient zeigt demzufolge in Richtung der stärksten Potential**abnahme**.

Wie wir gleich im *Beispiel Gravitationsfeld* sehen werden, ergibt sich hier das negative Vorzeichen aus der Definition der potentiellen Energie E_{pot} .

Standardbeispiele für Gradientenfelder mit den zugehörigen Potentialfeldern sind das (**repulsive**) elektrostatische Feld¹³

$$\vec{E}(\vec{r}) = -\text{grad } \varphi(\vec{r})$$

mit der elektrischen Feldstärke \vec{E} und dem elektrostatischen Potential φ und das (**attraktive**) Gravitationsfeld

$$\vec{g}(\vec{r}) = -\text{grad } \phi(\vec{r})$$

mit der Gravitationsfeldstärke \vec{g} und dem Gravitationspotential ϕ . Die Gravitationsfeldstärke ist äquivalent zur Beschleunigung \vec{g} von Massen im Schwerfeld. Folgende Konventionen sind üblich:

$$\begin{aligned} \lim_{|\vec{r}'| \rightarrow \infty} \varphi(\vec{r}') \stackrel{!}{=} 0 &\Rightarrow \int_{\vec{r}}^{\infty} \vec{E}(\vec{r}') \cdot d\vec{r}' \stackrel{!}{=} -\Delta\varphi = -\left[\underbrace{\varphi(\infty)}_{=0} - \varphi(\vec{r}) \right] = \varphi(\vec{r}), \\ \lim_{|\vec{r}'| \rightarrow \infty} \phi(\vec{r}') \stackrel{!}{=} 0 &\Rightarrow \int_{\vec{r}}^{\infty} \vec{g}(\vec{r}') \cdot d\vec{r}' \stackrel{!}{=} -\Delta\phi = -\left[\underbrace{\phi(\infty)}_{=0} - \phi(\vec{r}) \right] = \phi(\vec{r}). \end{aligned} \quad (10.3)$$

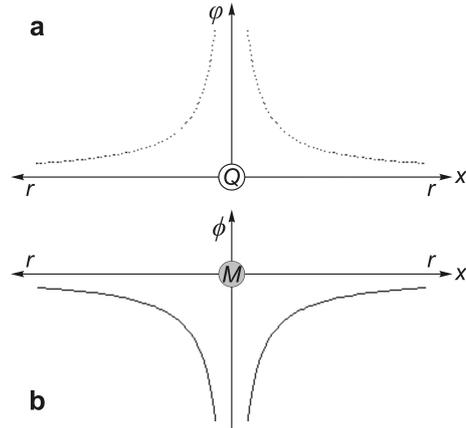
¹³Bei den Definitionen im Zusammenhang mit dem elektrostatischen Feld wird konventionsgemäß von positiven Feldladungen und positiven Probeladungen ausgegangen. Deshalb nennt man das elektrostatische Feld repulsiv. Durch Berücksichtigung der Ladungsvorzeichen resultieren daraus die entsprechenden Zusammenhänge anderer Ladungskonstellationen. So wirkt eine positive Feldladung auf eine negative Probeladung attraktiv – wie die Massen im Gravitationsfeld.

Abb. 5

a Potential φ einer im Koordinatenursprung liegenden positiven elektrischen Punktladung Q .

b Potential ϕ einer im Koordinatenursprung liegenden Punktmasse M .

Im Gegensatz zum stets positiven Potential im elektrostatischen Feld einer positiven Ladung ist das Potential im Gravitationsfeld stets negativ, weil das elektrostatische Feld im Fall gleichnamiger Ladungen *repulsiv* wirkt während das Gravitationsfeld immer *attraktiv* ist. Wegen der Radialsymmetrie der Probleme konnten die y -Achse und die z -Achse unterdrückt werden.



Beispiel Gravitationsfeld einer im Koordinatenursprung liegenden Punktmasse M mit einer Probepunktmasse m im Raumpunkt $\vec{r} = |\vec{r}|\vec{e}_r = r\vec{e}_r$:

Es handelt sich hier um ein radialsymmetrisches Problem. Mit der Gravitationskonstante G lautet das dem Beispiel entsprechende Gravitations(kraft)gesetz

$$\vec{F}_G(\vec{r}) = -G \frac{m \cdot M}{|\vec{r}|^3} \vec{r} \quad \text{bzw. in Kugelkoordinaten} \quad \vec{F}_G(r) = -G \frac{m \cdot M}{r^2} \vec{e}_r .$$

Das Minuszeichen vor dem G bringt zum Ausdruck, dass die Massen einander anziehen bzw. dass die Gravitationskraft \vec{F}_G entgegengesetzt gerichtet ist zu \vec{e}_r . Wir rechnen weiter in Kugelkoordinaten und so resultiert die Gravitationsfeldstärke

$$\vec{g}(r) = \frac{\vec{F}_G(r)}{m} = -G \frac{M}{r^2} \vec{e}_r .$$

Mit zunehmendem Abstand r vom Gravitationszentrum, d. h. von der Masse M im Koordinatenursprung, nimmt der Betrag $|\vec{g}(r)| = g(r) = GM/r^2$ der Gravitationsfeldstärke entsprechend $1/r^2$ ab. **Die Änderung der potentiellen Energie, also ΔE_{pot} , ist gleich der zu verrichtenden Verschiebungsarbeit W an der Probemasse m auf dem Weg von \vec{r}_1 nach \vec{r}_2 (von r_1 nach r_2).** Daraus folgt mit

$$\Delta\phi = \frac{\Delta E_{\text{pot}}}{m} ,$$

dass mit zunehmender Entfernung vom Gravitationszentrum gemäß $r_2 > r_1$ auch das Gravitationspotential wächst, weil dabei Verschiebungsarbeit $W > 0$ gegen die Gravitationskraft verrichtet wird:

$$r_2 > r_1 \Rightarrow W > 0 \Rightarrow \Delta E_{\text{pot}} > 0 \Rightarrow \Delta\phi > 0 .$$

Damit erhalten wir für das Gravitationspotential unter Berücksichtigung der Konvention (10.3) und mit $d\vec{r} = dr \vec{e}_r$

$$\int_r^\infty \vec{g}(r') \cdot dr' \vec{e}_r = \int_r^\infty -G \frac{M}{r'^2} \vec{e}_r \cdot dr' \vec{e}_r = -GM \int_r^\infty \frac{1}{r'^2} dr' = -GM \left[-\frac{1}{r} \right]_r^\infty = -G \frac{M}{r} = \phi(r) ,$$

$$\boxed{\lim_{r \rightarrow \infty} \phi \stackrel{!}{=} 0 \Rightarrow \phi(r) = -G \frac{M}{r} , \quad -\infty < \phi < 0 .}$$

Wählen wir als Nullpotential des Gravitationsfeldes das Potential im Unendlichen, dann ist das Gravitationspotential **stets negativ** (siehe Abbildung 5b). Außerdem besitzt das Gravitationspotential im Gravitationszentrum, d. h. in der felderzeugenden Punktmasse M , eine Singularität. Dann gilt für das Potential eines radialsymmetrischen Gravitationsfeldes: $\sup \phi = 0$, (uneigentliches) $\inf \phi = -\infty$.

Probe mit der Umkehroperation:

$$\begin{aligned} \vec{g}(r) &= -\nabla\phi(r) \\ &= -\frac{d}{dr} \left(-G \frac{M}{r} \right) \vec{e}_r = GM \vec{e}_r \frac{d}{dr} r^{-1} = -G \frac{M}{r^2} \vec{e}_r = \vec{g}(r) . \quad \square \end{aligned}$$

11 Der Gay-Lussac'sche Versuch

11.1 Einleitung

Wir beginnen mit einem Zitat aus dem Taschentextbuch von Hans-Werner Kammer und Kurt Schwabe, Thermodynamik irreversibler Prozesse – Eine Einführung, VCH Verlagsgesellschaft – Physik-Verlag, Weinheim, 1985, Seite 9 und Seite 10:

„Die klassische Thermodynamik basiert auf experimentell gesicherten Gesetzen, den Hauptsätzen. Sie beschäftigt sich mit der Untersuchung von Gleichgewichtszuständen der Materie. Aus diesem Grunde muß im Rahmen der klassischen Thermodynamik jeder Übergang eines Systems von einem Zustand in den anderen als Folge von Gleichgewichtszuständen angesehen werden. Die Zeit spielt dabei keine Rolle. Eine solche Folge von Gleichgewichtszuständen kann von dem System auch in umgekehrter Richtung durchlaufen werden, ohne daß Änderungen in der Umgebung zurückbleiben; man sagt, daß sich das System auf einem reversiblen Weg bewegt.

In der klassischen Thermodynamik spielt die Reversibilität eine fundamentale Rolle. Das ist eine Idealisierung, denn jeder reale Übergang eines Systems von einem Anfangs- in einen Endzustand kann sich nur vollziehen, wenn sich das System gerade nicht im Gleichgewicht befindet. Die phänomenologische Thermodynamik vermag daher nur aus den Veränderungen von Größen beim Übergang vom Anfangs- in den Endzustand die Schlußfolgerung abzuleiten, ob der Prozeß entsprechend den Hauptsätzen möglich ist oder nicht. Das wird erreicht, da der **2. Hauptsatz** eine Aussage über die Richtung des Ablaufs von makroskopischen Prozessen macht:

In einem abgeschlossenen System verlaufen alle Prozesse so, daß die Entropie zunimmt.

Es erhebt sich die Frage, ob man die Thermodynamik auf Nichtgleichgewichtssituationen ausdehnen kann, ob man mit thermodynamischen Methoden Aussagen über den Ablauf von realen Prozessen gewinnen kann. ... Man kann nun den Weg skizzieren: Ausgangspunkt muß die GIBBSSche Fundamentalgleichung sein, die den 1. und 2. Hauptsatz in sich vereint:

$$TdS = dU + PdV - \sum_i \mu_i dX_i \quad (11.1)$$

... Die in (11.1) auftretenden Größen sind thermodynamische Variable, definiert für das Gleichgewicht. Sie sind zeitunabhängig. Um den Zeitablauf von Vorgängen erfassen zu können, muß man annehmen, daß (11.1) in jedem Zeitpunkt erfüllt ist, also

$$T \frac{dS}{dt} = \frac{dU}{dt} + P \frac{dV}{dt} - \sum_i \mu_i \frac{dX_i}{dt} \quad (11.2)$$

gilt. Damit wird den thermodynamischen Variablen ein dynamischer Charakter verliehen. Das ist für die thermodynamische Theorie irreversibler Prozesse fundamental. Der Übergang von (11.1) nach (11.2) ist unumgänglich für eine thermodynamische Theorie irreversibler Prozesse. Es ist jedoch folgendes dabei zu beachten: Mit (11.2) wird postuliert, daß die thermodynamischen Variablen und die Beziehungen zwischen ihnen auch außerhalb des Gleichgewichtes gelten. Dem liegt das Prinzip des lokalen Gleichgewichtes zugrunde. Man setzt voraus, daß jedes Volumenelement im Ablauf des Prozesses nur Zustände durchläuft, die das isolierte Volumenelement im Gleichgewicht annehmen würde. Die Ausdehnung der Thermodynamik auf Nichtgleichgewichtssituationen erfordert somit, daß man die globale Reversibilität durch eine lokale ersetzt. Ein Beweis dieses Prinzips kann nur durch die kinetische Theorie erbracht werden. Der Vergleich der auf diesem Prinzip aufgebauten Theorie mit dem Experiment zeigt jedoch seine Richtigkeit.“

- Die Formeln der klassischen Thermodynamik besitzen keine Zeitabhängigkeit und beschreiben nur Systeme, die sich im thermodynamischen Gleichgewicht befinden.
- Bei einem **quasistatischen Prozess** bzw. **Gleichgewichtsprozess** werden die Parameter so langsam verändert, dass sich das System stets im thermodynamischen Gleichgewicht befindet. Quasistatische Prozesse sind idealisierte Grenzfälle, denn sie bestehen aus einer unendlichen Folge infinitesimal kleiner Zustandsänderungen und dauern unendlich lange. Wenn sich ein System stets im Gleichgewicht befindet, dann ist es reversibel, d. h. zeitlich umkehrbar. Deshalb spricht man bei quasistatischen Prozessen auch von reversiblen Prozessen oder **reversiblen Zustandsänderungen**. Streng genommen laufen in der Realität Prozesse niemals reversibel ab. Schließlich zeigt der (thermodynamische) „Zeitpfeil“ immer in dieselbe Richtung.
- Für die extensiven, unabhängigen Größen bzw. Zustandsfunktionen Entropie S und innere Energie U gilt:

1. Bei jeder reversiblen Zustandsänderung vom Zustand 1 in den Zustand 2 sind

$$\Delta S = \int_1^2 \frac{dQ}{T} \quad \text{und} \quad \Delta U = U_2 - U_1$$

wegunabhängig, sodass für einen ganzen Zyklus eines reversiblen Kreisprozesses gilt:

$$\Delta S = \oint \frac{dQ_{\text{reversibel}}}{T} = 0 \quad \text{und} \quad \Delta U = 0 .$$

2. dS und dU sind daher vollständige Differentiale.

- Die innere Energie U ist eine Zustandsfunktion und nur abhängig von den Zustandsgrößen P, V, T, n .
- Die Wärmeenergie Q und die Arbeit W sind keine Zustandsgrößen und demzufolge sind dQ und dW keine vollständigen Differentiale, d. h. die zugehörigen Integrale sind allgemein **wegabhängig**. Allgemein gilt also

$$\oint dQ \neq 0 \quad \text{und} \quad \oint dW \neq 0 .$$

- In der klassischen Thermodynamik sind nur der Anfangs- und der Endzustand von Prozessen von Interesse, weil die Änderungen von Zustandsfunktionen bzw. Zustandsgrößen wie z. B. der inneren Energie U und der Entropie S wegunabhängig sind.
- Um Zustandsänderungen bequemer berechnen zu können, ist man deshalb stets bemüht, sie gemäß der klassischen Thermodynamik quasistatisch bzw. reversibel zu „gestalten“.

Regeln zur reversiblen Gestaltung thermodynamischer Prozesse

- Fixiere den Zustand des Systems. Dies geschieht durch Einteilung in Untersysteme.
- Die Entropie des Gesamtsystems ist gleich der Summe der Entropien der Untersysteme.
- Ändere sämtliche Untersysteme reversibel so, dass insgesamt der interessierende Endzustand erreicht wird.
- Der Wärmeaustausch mit Hilfsreservoirien muss immer so erfolgen, dass die Temperaturen von System oder Untersystem und Hilfsreservoir gleich sind. Dieser Wärmeaustausch muss also reversibel erfolgen.

Thermodynamik reversibler Prozesse mit idealen Gasen

Vorzeichenkonvention: Die dem System zugeführte Wärmeenergie ΔQ ist positiv.
Die vom System verrichtete Arbeit ΔW ist negativ.

	isochor: $V = \text{const}$	isotherm: $T = T_k = \text{const}$	isobar: $P = \text{const}$	isentrop: $\Delta Q = 0$
$\frac{P_1}{T_1} = \frac{P}{T} = \text{const}$	$P_1 \cdot V_1 = P \cdot V = \text{const}$	$P_1 \cdot V_1 = P \cdot V = \text{const}$	$\frac{V_1}{T_1} = \frac{V}{T} = \text{const}$	$\frac{P_1 \cdot V_1}{T_1} = \frac{P \cdot V}{T} = \text{const}$
2 variable Größen: P, T $dU = dQ$ $dQ = dU = c_v \cdot n \cdot dT$	2 variable Größen: P, V $dU = 0$ $dQ = -dW = P \cdot dV$	2 variable Größen: V, T $dU = dQ + dW = dQ - P \cdot dV$ $dQ = dU - dW$ $dQ = c_v \cdot n \cdot dT + P \cdot dV = c_p \cdot n \cdot dT$	2 variable Größen: P, V, T $dU = dQ + dW = dQ - P \cdot dV$ $dQ = dU - dW$ $dQ = c_v \cdot n \cdot dT + P \cdot dV = 0$	3 variable Größen: P, V, T $dU = dQ = -P \cdot dV$ $dQ = dU - dW = 0$ $dQ = c_v \cdot n \cdot dT + P \cdot dV = 0$
$\Delta W = 0$	Wenn $P(V)$ als Isotherme gegeben ist: $P(V) \cdot V = n \cdot R \cdot T_k$ $P(V) = n \cdot R \cdot T_k \cdot \frac{1}{V}$ $dQ = P(V) \cdot dV = n \cdot R \cdot T_k \cdot \frac{dV}{V}$	$P \cdot dV = c_p \cdot n \cdot dT - c_v \cdot n \cdot dT = (c_p - c_v) \cdot n \cdot dT$ $= R \cdot n \cdot dT$ $\Delta Q = c_v \cdot n \cdot \Delta T + n \cdot R \cdot \Delta T$ $\Delta W = -\int P(V) \cdot dV = -P \cdot \Delta V = -n \cdot R \cdot \Delta T$	$c_p \cdot n \cdot dT = -P \cdot dV = -\frac{n \cdot R \cdot T}{V} \cdot dV$ $= -n \cdot (c_p - c_v) \cdot T \cdot \frac{dV}{V}$ $\frac{dT}{T} = -\frac{c_p - c_v}{c_p} \cdot \frac{dV}{V}$ $dW = -P(V) \cdot dV = c_v \cdot n \cdot dT$ $\Delta W = c_v \cdot n \cdot \Delta T$	$\Delta U = \Delta W = c_v \cdot n \cdot \Delta T$ $PV^\kappa = \text{const}, TV^{\kappa-1} = \text{const}, PT^{\frac{\kappa}{\kappa-1}} = \text{const}$ $\kappa = \frac{c_p}{c_v} = 1 + \frac{R}{c_v}$
$\Delta Q = \Delta U = c_v \cdot n \cdot \Delta T$	$\Delta Q = -\Delta W = n \cdot R \cdot T_k \cdot \ln \frac{V_2}{V_1}$	$\Delta Q = c_v \cdot n \cdot \Delta T + n \cdot R \cdot \Delta T$	$\Delta Q = c_v \cdot n \cdot \Delta T + n \cdot R \cdot \Delta T$	$\Delta Q = c_v \cdot n \cdot \Delta T - c_v \cdot n \cdot \Delta T = 0$

1. Hauptsatz:

11.2 Versuchsbeschreibung

Wir diskutieren den **Gay-Lussac'schen Versuch** bzw. Gay-Lussac'schen Überströmversuch als Beispiel für die Anwendung der phänomenologischen Thermodynamik. Er wurde 1807 von dem französischen Chemiker/Physiker Joseph Louis Gay-Lussac (1778 – 1850) durchgeführt:

Ein luftgefüllter Behälter ist durch ein zunächst geschlossenes Ventil von einem evakuierten Behälter getrennt. Dann wird das Ventil geöffnet, sodass die Luft in den evakuierten Behälter überströmt, bis sie sich gleichmäßig auf beide Behälter verteilt hat. Die Versuchsanordnung sorgt dafür, dass kein Wärmeenergieaustausch mit der Umgebung erfolgen kann und somit die innere Energie der Luft in den Behältern konstant bleibt. Die Lufttemperatur in den Behältern wird vor und nach dem Druckausgleich gemessen. Man stellt keinen Temperaturunterschied fest.¹⁴

Das bedeutet, dass die innere Energie eines idealen Gases bei gegebener Stoffmenge nur von seiner Temperatur abhängt und nicht von seinem Volumen und auch nicht vom Druck.



Der Gay-Lussac'sche Versuch mit einem idealen Gas ist ein **irreversibler** thermodynamischer Prozess mit charakteristischen Merkmalen der isothermen und der adiabatischen Zustandsänderung:

1. Durch thermische Isolierung findet kein Wärmeenergieaustausch mit der Umgebung statt:

$$\Delta Q = 0 \quad (\text{charakteristisch für } \mathbf{adiabatische} \text{ Zustandsänderung}) .$$

2. Bei der Expansion in das Vakuum hinein ändert sich die Temperatur eines idealen Gases nicht:

$$\Delta U = 0 \quad \Rightarrow \quad T = \text{const} \quad (\text{charakteristisch für } \mathbf{isotherme} \text{ Zustandsänderung}) .$$

3. Ein Gas verrichtet nur bei *extrem* schneller Expansion in ein Vakuum hinein keine Arbeit:

$$\Delta W = 0 \quad \text{und} \quad \Delta Q = 0 \quad \Rightarrow \quad \Delta U = 0 \quad \text{und} \quad T = \text{const} .$$

Dieser Prozessverlauf ist irreversibel (nicht quasistatisch) und wird im Abschnitt 11.3 diskutiert und berechnet.

Beim Gay-Lussac'schen Versuch befindet sich das System nur im Anfangszustand und im Endzustand im thermodynamischen Gleichgewicht.

¹⁴Da Luft nur näherungsweise ein ideales Gas ist, stellt man bei genauer Messung fest, dass die Luft sich durch die Expansion beim Druckausgleich ein wenig abkühlt.

4. Die Entropie innerhalb eines abgeschlossenen Systems kann niemals abnehmen, als Grenzfall bleibt sie konstant, allgemein nimmt sie zu.
Es ist intuitiv klar, dass beim Gay-Lussac'schen Versuch die **Entropie zunimmt**, weil sich das Gas irreversibel ausdehnt, die Moleküle sich auf ein größeres Volumen verteilen und somit in einen Zustand größerer thermodynamischer Wahrscheinlichkeit übergehen.
5. Wir werden den Gay-Lussac'schen Versuch im Abschnitt 11.3 insgesamt irreversibel gestalten, d. h. ohne zusätzliches Wärmereservoir und ohne die Möglichkeit, nach außen Volumenarbeit zu übertragen. Dabei werden wir den Prozess zweiteilen, damit wir die quasistatischen (reversiblen) Zustandsfunktionen der klassischen Thermodynamik für die Berechnungen verwenden können. Im Abschnitt 11.4 werden wir den Prozess jedoch in einem Schritt reversibel (als Gleichgewichtsprozess) gestalten und berechnen.
6. Wenn wir im Zusammenhang mit dem Gay-Lussac'schen Versuch vom **System** sprechen, ist nicht die Luft sondern stattdessen eine **bestimmte, in den Versuchsbehältern eingeschlossene Stoffmenge eines idealen Gases** gemeint.

11.3 Der Gay-Lussac'sche Versuch irreversibel in zwei reversibel geführten Schritten

$$\left\{ \begin{array}{l} \text{Zustand 1} \\ P_1 \quad V_1 \quad T_1 \end{array} \right\} \xrightarrow[\text{Expansion}]{\text{isentrop}} \left\{ \begin{array}{l} \text{Zustand 2'} \\ P_{2'} \quad V_2 \quad T_{2'} \end{array} \right\} \xrightarrow[\text{Wärmeenergiezufuhr}]{\text{isochore}} \left\{ \begin{array}{l} \text{Zustand 2} \\ P_2 \quad V_2 \quad T_1 \end{array} \right\}$$

Für adiabatische Zustandsänderungen, die quasistatisch bzw. reversibel geführt werden, gilt mit dem 1. Hauptsatz

$$\Delta U = \Delta Q + \Delta W \quad \text{und} \quad \Delta Q = 0 \quad \Rightarrow$$

$$dU = dW = -P dV .$$

Bei adiabatischen Prozessen ändert sich die innere Energie also nur dadurch, dass das System Expansionsarbeit verrichtet oder dass Kompressionsarbeit in das System hineingesteckt wird. Bei der Expansion kommt es gleichzeitig zur Abkühlung und bei der Kompression gleichzeitig zur Erwärmung des idealen Gases gemäß

$$dW = -P dV = c_v n dT .$$

Weil bei adiabatischen Zustandsänderungen(, die quasistatisch bzw. reversibel geführt werden,) kein Wärmeenergieaustausch mit der Umgebung erfolgt, bleibt die Entropie konstant:

$$\Delta Q = \Delta Q_{\text{rev}} = 0 \quad \Rightarrow \quad \Delta S = \frac{1}{T} \Delta Q_{\text{rev}} = 0 .$$

Wir werden deshalb im Folgenden nicht von adiabatischen sondern von isentropen Zustandsänderungen sprechen. Bei **isentropen Zustandsänderungen**¹⁵ variieren alle drei Zustandsgrößen P , T , V und erfüllen die Zustandsgleichungen

$$P \cdot V^\kappa = \text{const} \quad \text{und} \quad T \cdot V^{\kappa-1} = \text{const} .$$

Die Einteilung des Prozesses in zwei quasistatische Schritte ist selbstverständlich willkürlich. Denn in der Realität laufen die isentrope Expansion und die isochore Wärmeenergiezufuhr gleichzeitig ab. Für den Gay-Lussac'schen Versuch resultiert aus dem **1. Hauptsatz**

$$dU = dQ + dW = dQ - P dV$$

wegen $U = \text{const}$:

$$dU = 0 \quad \Rightarrow \quad dQ = -dW = P dV .$$

Weil kein Wärmeenergieaustausch mit der Umgebung erfolgt, kann man folglich annehmen, dass der Wärmeenergievorrat für die isochore Wärmeenergiezufuhr im 2. Schritt durch die theoretisch am System selbst verrichtete „Volumenarbeit“ des 1. Schrittes geliefert wird.

In der kinetischen Gastheorie wäre die Argumentation anders. Dort würden die sich statistisch bewegenden Gasteilchen das evakuierte Volumen „erobern“, ohne in ihrer Gesamtheit an thermischer (kinetischer) Energie zu verlieren. Die Summe der

¹⁵Bei allen anderen Zustandsänderungen wird jeweils eine der drei Zustandsgrößen festgehalten, während die anderen beiden variieren.

thermischen Energie aller Teilchen im System ist aber die innere Energie U des Systems und die wäre somit konstant. Weil dann aber auch die mittlere kinetische Energie der Teilchen konstant ist, muss auch die Temperatur des idealen Gases konstant sein. Wie man sieht, kommt diese Betrachtungsweise der Realität recht nahe.

Wir werden im Folgenden die spezifischen Wärmekapazitäten c_v für konstantes Volumen und c_p für konstanten Druck sowie den Adiabatenexponenten $\kappa = c_p/c_v > 1$ verwenden. Weiterhin werden wir auf die

ideale Gasgleichung bzw. die **thermische Zustandsgleichung** idealer Gase

$$PV = Nk_B T = nRT, \quad n = \frac{N}{N_A}, \quad R = k_B N_A$$

zurückgreifen. Die dabei verwendeten Größen sind:

Teilchenzahl N , Avogadro-Zahl N_A (Teilchenzahl pro Mol), Molzahl n , Boltzmann-Konstante $k_B = 1,380\,649 \cdot 10^{-23}$ J/K, allgemeine Gaskonstante $R = c_p - c_v$.

1. Schritt: Zustand 1 \rightarrow Zustand 2',
(reversible) **isentrop Expansion** eines idealen Gases:

Das Volumen nimmt zu, die Temperatur nimmt ab und es gilt

$$\Delta Q = 0, \quad T_1 V_1^{\kappa-1} = T_{2'} V_2^{\kappa-1}.$$

Gesucht wird $T_{2'}$ in Abhängigkeit von V :

$$T_{2'} = T_1 \left(\frac{V_1}{V_2} \right)^{\kappa-1}. \quad (11.3)$$

2. Schritt: Zustand 2' \rightarrow Zustand 2,
(reversible) **isochore Wärmeenergiezufuhr**:

Es ist $V_2 = \text{const}$, die Temperatur $T_{2'}$ steigt durch Wärmeenergiezufuhr ΔQ wieder auf den Anfangswert $T_2 = T_1$ und wegen $V_2 = \text{const}$ wird im 2. Schritt keine Volumenarbeit verrichtet. Mit dem 1. Hauptsatz der Thermodynamik erhalten wir ΔQ wie folgt:

$$dU = dQ = n \cdot c_v \cdot dT \quad \Rightarrow \quad (11.4)$$

$$\Delta U = \Delta Q = n \cdot c_v \cdot T \Big|_{T_{2'}}^{T_1}. \quad (11.5)$$

(11.3) in (11.5) eingesetzt liefert

$$\Delta Q = n \cdot c_v \cdot T_1 \left(1 - \frac{V_1^{\kappa-1}}{V_2^{\kappa-1}} \right).$$

Die Expansionsarbeit, die das System beim Gay-Lussac'schen Versuch an sich selbst verrichtet, steckt es *gleichzeitig* in Form von Wärmeenergie wieder in sich hinein, wodurch innere Energie und Temperatur konstant bleiben. Gegenüber dem Anfangszustand hat sich im Endzustand neben einer Volumenzunahme und einer Druckabnahme die Entropie des Systems erhöht. Diese Entropie wurde im und vom Gas selbst erzeugt (nicht von außen zugeführt) und entspricht der soeben erwähnten (in der Berechnung *reversibel*) hineingesteckten Wärmeenergie.

Die Entropieänderung beim Gay-Lussac'schen Versuch berechnen wir mit $dQ_{\text{rev}} = TdS$ und mit (11.4):

$$dS = \frac{1}{T} dQ_{\text{rev}} = \frac{1}{T} \cdot n \cdot c_v \cdot dT \quad \Rightarrow$$

$$\Delta S = n \cdot c_v \cdot \int_{T_{2'}}^{T_1} \frac{1}{T} dT = n \cdot c_v \cdot \ln \frac{T_1}{T_{2'}} .$$

Durch Einsetzen von (11.3) erhalten wir daraus

$$\Delta S = n \cdot c_v \cdot \ln \frac{T_1}{T_1 \left(\frac{V_1}{V_2}\right)^{\kappa-1}} = n \cdot c_v \cdot \ln \left(\frac{V_2}{V_1}\right)^{\kappa-1} = n \cdot c_v \cdot (\kappa - 1) \cdot \ln \left(\frac{V_2}{V_1}\right)$$

und mit $\kappa = 1 + \frac{R}{c_v}$ resultiert schließlich

$$\Delta S = n \cdot R \cdot \ln \frac{V_2}{V_1}, \quad V_2 > V_1 \Rightarrow \Delta S > 0 .$$

Betrachtungen zu den „Nichtzustandsgrößen“ Q und W

Bei isentroper Zustandsänderung im 1. Schritt gilt

$$dU = dW = -PdV \quad \Rightarrow \quad \Delta W = \Delta U = n \cdot c_v \cdot \Delta T \quad \Rightarrow$$

$$\Delta W = - \int_{V_1}^{V_2} P dV = n \cdot c_v \cdot T \Big|_{T_1}^{T_{2'}} = n \cdot c_v \cdot T_1 \left(\frac{V_1^{\kappa-1}}{V_2^{\kappa-1}} - 1 \right) . \quad (11.6)$$

Die vom System verrichtete isentrope Expansionsarbeit von V_1 bis V_2 ist negativ, weil $V_2 > V_1$.

Die im 2. Schritt zugeführte Wärmeenergie, die erforderlich ist, um das System auf isochorem Wege von der Temperatur $T_{2'}$ zur Anfangstemperatur T_1 zurückzuführen, hatten wir bereits berechnet:

$$\Delta Q = n \cdot c_v \cdot T \Big|_{T_{2'}}^{T_1} = n \cdot c_v \cdot T_1 \left(1 - \frac{V_1^{\kappa-1}}{V_2^{\kappa-1}} \right) . \quad (11.7)$$

Die dem System zugeführte Wärmeenergie ist positiv, weil $T_1 > T_{2'}$. Wie wir sehen ist $\Delta Q = -\Delta W$, sodass sich (11.6) und (11.7) widerspruchsfrei zur Beschreibung des Gay-Lussac'schen Versuchs zusammenfassen lassen:

$$W \Big|_{V_1}^{V_2} + Q \Big|_{T_{2'}}^{T_1} = 0 .$$

11.4 Der Gay-Lussac'sche Versuch reversibel gestaltet in einem Schritt

$$\left\{ \begin{array}{l} \text{Zustand 1} \\ P_1 \quad V_1 \quad T_1 \end{array} \right\} \xrightarrow[\text{Expansion}]{\text{isotherme}} \left\{ \begin{array}{l} \text{Zustand 2} \\ P_2 \quad V_2 \quad T_1 \end{array} \right\}$$

Wir werden jetzt den thermodynamischen Prozess des Gay-Lussac'schen Versuchs als quasistatisch-isotherme Expansion eines idealen Gases von V_1 nach V_2 führen. Weil bei dieser Zustandsänderung dem System Energie von außen zugeführt wird, müssen wir bei der Berechnung der Zustandsgrößen für den Anfangs- und den Endzustand die Forderung nach thermischer Isolierung des Systems aufgeben. Das ist möglich, weil die Zustandsfunktionen wegunabhängig sind, sodass wir unter Einhaltung der Plausibilität Gestaltungsfreiheit besitzen. Also bringen wir das System in Kontakt mit einem Wärmeenergieservoir der konstanten Temperatur T_1 , das dem System die für die Temperaturerhaltung nötige Wärmeenergie ΔQ_{rev} liefert. Weiterhin nehmen wir an, dass die Volumenzunahme nicht über ein Ventil sondern über einen reibungsfrei beweglichen Kolben erfolgt, der Expansionsarbeit dW nach außen abgeben kann. Wie kommt das System unter diesen, für die Berechnung neu angesetzten Bedingungen vom Anfangs- zum Endzustand? Mit der idealen Gasgleichung für $T_1 = \text{const}$

$$PV = nRT_1 = \text{const} \quad \Leftrightarrow \quad P(V) = \frac{n \cdot R \cdot T_1}{V}$$

ist dann die vom System verrichtete isotherme Expansionsarbeit

$$\Delta W = \int_{V_1}^{V_2} -P(V) dV = -nRT_1 \int_{V_1}^{V_2} \frac{1}{V} dV = -nRT_1 \cdot \ln \frac{V_2}{V_1}.$$

Wegen $T_1 = \text{const}$ ist hier $\Delta U = \Delta Q_{\text{rev}} + \Delta W = 0$ und

$$\Delta Q_{\text{rev}} = -\Delta W = nRT_1 \cdot \ln \frac{V_2}{V_1}.$$

Damit lässt sich jetzt die Entropieänderung für den Gay-Lussac'schen Versuch angeben:

$$\Delta S = \frac{1}{T_1} \Delta Q_{\text{rev}} = \frac{1}{T_1} \cdot nRT_1 \cdot \ln \frac{V_2}{V_1},$$

$$\Delta S = n \cdot R \cdot \ln \frac{V_2}{V_1}, \quad V_2 > V_1 \Rightarrow \Delta S > 0.$$

Aus dem Wärmeenergieservoir strömen während des Expansion ständig *reversibel* Wärmeenergie in das System und mit ihr vergesellschaftet auch die Entropie. Während die von außen aus dem Wärmeenergieservoir zugeführte Entropie im System verbleibt, wird die Wärmeenergie sofort wieder als Expansionsarbeit $-P dV$ über den Kolben nach außen abgegeben. Die quasistatisch geführte *isotherme* Expansion des Systems von V_1 nach V_2 verläuft also in der Bilanz *isoenergetisch* – die Entropie des Systems jedoch nimmt dabei zu.

11.5 Schlussfolgerungen

1. Beim Gay-Lussac'schen Versuch wird die Entropie bei irreversibler Expansion im Gas selbst erzeugt und nicht von außen zugeführt.
2. Bei reversibler Gestaltung mittels Wärmeenergieservoir und Kolben wird die Entropie jedoch von außen zugeführt.
3. Für das Endergebnis ist es aber egal, ob die Entropie aus dem Wärmeenergieservoir stammt oder ob sie vor Ort erzeugt wurde. Entscheidend ist nur, dass die innere Energie und damit die Temperatur im System am Ende konstant geblieben sind.

12 Kirchhoff'sches Strahlungsgesetz

Siehe auch: Springer-Lehrbuch Gerthsen Physik, 20. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 1999, Abschnitt 11.2.1 *Wärmestrahlung und thermisches Gleichgewicht*, Seite 569 bis Seite 571.

Wir bringen einen **Probekörper** K mit der Anfangstemperatur T in einen Hohlraumstrahler S, auch **schwarzer Strahler** genannt.

Im Folgenden verwenden wir den **Absorptionsgrad** bzw. das Absorptionsvermögen $0 < \alpha \leq 1$ und den **Emissionsgrad** bzw. das Emissionsvermögen $0 < \varepsilon \leq 1$. Physikalische Größen bezüglich S indizieren wir mit S, Größen bezüglich K sind indexfrei.

S habe die Anfangstemperatur $T_S \neq T$ und strahle die Leistung $\tilde{P}_S(T_S)$ auf K, während K gleichzeitig von \tilde{P}_S den Leistungsanteil $\tilde{P}(T) = \alpha \cdot \tilde{P}_S$ absorbiert. Umgekehrt strahlt K die Leistung $\tilde{P}(T)$ auf S, während S gleichzeitig von \tilde{P} den Leistungsanteil $\tilde{P}_S(T) = \alpha_S \cdot \tilde{P}$ absorbiert. Dabei kommt es auf der Grundlage des 2. Hauptsatzes der Thermodynamik zu einem Angleichen der Temperaturen von K und S, bis sich schließlich das thermische Gleichgewicht mit $T = T_S$ eingestellt hat.

Im thermischen Gleichgewicht gilt:

1. Wenn mit der Umgebung kein Energieaustausch erfolgt, ist gemäß dem 1. und dem 2. Hauptsatz der Thermodynamik $T = T_S = \text{const}$.
2. Sowohl S als auch K müssen jeweils stets genauso viel Leistung emittieren wie sie gleichzeitig absorbieren:

$$\underbrace{\alpha \cdot P_S}_{\text{Absorption bei K}} = \underbrace{\varepsilon \cdot P_S}_{\text{Emission bei K}} \quad \text{und} \quad \underbrace{\alpha_S \cdot P}_{\text{Absorption bei S}} = \underbrace{\varepsilon_S \cdot P}_{\text{Emission bei S}} .$$

Daraus folgt

$$\alpha = \varepsilon \quad \text{und} \quad \alpha_S = \varepsilon_S .$$

3. S strahlt die konstante Leistung P_S auf K und K strahlt die konstante Leistung P auf S, sodass wegen $T = T_S = \text{const}$ gelten muss

$$\boxed{\underbrace{\alpha \cdot P_S}_{\text{Absorption bei K}} = \underbrace{\alpha_S \cdot P}_{\text{Absorption bei S}} \quad \text{und} \quad \underbrace{\varepsilon \cdot P_S}_{\text{Emission bei K}} = \underbrace{\varepsilon_S \cdot P}_{\text{Emission bei S}}} . \quad (12.1)$$

4. Ein schwarzer Strahler absorbiert alles und emittiert alles gemäß

$$\alpha_S = \varepsilon_S = 1 .$$

Damit resultiert aus (12.1) das **Kirchhoff'sche Strahlungsgesetz**

$$\boxed{P = \alpha \cdot P_S \quad \text{bzw.} \quad P = \varepsilon \cdot P_S} .$$

Im thermischen Gleichgewicht zwischen einem Probekörper K und einem schwarzen Strahler S ist die von K abgestrahlte Leistung P proportional zu seinem Absorptionsgrad α bzw. zu seinem Emissionsgrad ε . Je besser ein Körper absorbiert, desto stärker strahlt er.

Für den **Reflektionsgrad** ϱ gilt dann

$$1 - \varrho = \varepsilon = \alpha \quad \Leftrightarrow \quad \varrho = 1 - \varepsilon = 1 - \alpha .$$

Ein idealer Spiegel (Metallspiegel) reflektiert alles, d. h. er absorbiert und emittiert nichts, also

$$\varrho = 1 \quad \Rightarrow \quad \alpha = \varepsilon = 0 .$$

13 Green'sche Funktion

Quellen:

- YouTube unter dem Suchbegriff TheNilsor-SPECIAL-Green'sche Funktion,
- Greensche Funktion – Wikipedia,
- Rainer J. Jelitto, Studententext Physik, Elektrodynamik (Theoretische Physik 3), 3. Auflage, Aula-Verlag, Wiesbaden, 1994, Seite 57 und folgende.

Green'sche Funktionen G dienen vorwiegend der Lösung inhomogener linearer gewöhnlicher und partieller Differentialgleichungen (Dgln) der Form

$$D\psi = f$$

mit D ein- oder mehrdimensionaler linearer Differentialoperator

z. B. $D_t = \frac{d}{dt}$, $D_x = \frac{d^2}{dx^2}$ oder $D_{\vec{r}} = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$,

ψ Lösungsfunktion (kurz Lösung) der Dgl

z. B. $\psi(t)$, $\psi(x)$ oder $\psi(\vec{r})$,

f Inhomogenität in Gestalt einer Funktion f

z. B. $f(t)$, $f(x)$ oder $f(\vec{r})$.

Ist $f \equiv 0$, so handelt es sich nicht um eine inhomogene sondern um eine homogene Dgl.

Abweichend von der in der Mathematik für Dgln gebräuchlichen Notation $Dy(x) = f(x)$ schreiben wir ψ statt y , weil hier x , y und z bereits zur Bezeichnung der kartesischen Koordinaten vergeben sind. Die allgemeine Lösung ψ der inhomogenen Dgl setzt sich zusammen aus der allgemeinen homogenen Lösung ψ_0 für den homogenen Fall $D\psi_0 = 0$, also für $f \equiv 0$, und aus der partikulären Lösung ψ_p :

$$\psi = \psi_0 + \psi_p.$$

Die partikuläre („einzelne“) Lösung ist in diesem Zusammenhang die der Inhomogenität f zugeordnete Lösung. Bei der Erläuterung der Idee, die zur Green'schen Funktion führt, wählen wir o.B.d.A. den dreidimensionalen Fall mit der unabhängigen Variablen $\vec{r} = (x, y, z)$, also

$$D_{\vec{r}}\psi(\vec{r}) = f(\vec{r}). \quad (13.1)$$

Wir suchen also die partikuläre Lösung ψ_p der Dgl (13.1). Die Anwendung eines zu D inversen Operators D^{-1} gemäß

$$D^{-1}D\psi_p = D^{-1}f \Rightarrow \psi_p = D^{-1}f \quad (13.2)$$

führt nicht zum Ziel, weil (13.2) für nichttriviale Lösungen ψ_p im Fall $f = 0$ nicht erfüllt ist.

Der folgende Ansatz mit der Dirac'schen δ -Funktion, die „definiert“ ist durch die Faltung¹⁶

$$\delta * f = \begin{cases} \int_{-\infty}^{\infty} \delta(\vec{r} - \vec{r}') d\vec{r}' & = 1 & \text{für Ausgangsfunktion } f \equiv 1, \\ \int_{-\infty}^{\infty} \delta(\vec{r} - \vec{r}') f(\vec{r}') d\vec{r}' & = f(\vec{r}) & \text{für Ausgangsfunktion } f = f(\vec{r}), \end{cases}$$

mit der Green'schen Funktion¹⁷ $G = G(x - x')$ und mit

$$\boxed{DG \stackrel{!}{=} \delta \quad \text{bzw.} \quad DG(\vec{r} - \vec{r}') \stackrel{!}{=} \delta(\vec{r} - \vec{r}')} \quad (13.3)$$

führt jedoch zum Ziel:

$$\text{a) } D\psi_p = f = \delta * f = (DG) * f = D \underbrace{(G * f)}_{=\psi_p} \quad \text{bzw.}$$

$$\text{b) } D\psi_p = f \quad \Rightarrow \quad (DG) * \psi_p = G * f \quad \Rightarrow$$

$$\boxed{\psi_p = G * f := \int G(\vec{r} - \vec{r}') f(\vec{r}') d\vec{r}'} \quad (13.4)$$

In Faltungsintegrale ausformuliert ergibt das:

$$\text{a) } \int \delta(\vec{r} - \vec{r}') f(\vec{r}') d\vec{r}' = \int D_{\vec{r}} G(\vec{r} - \vec{r}') f(\vec{r}') d\vec{r}' = D_{\vec{r}} \underbrace{\int G(\vec{r} - \vec{r}') f(\vec{r}') d\vec{r}'}_{=\psi_p} = f(\vec{r})$$

bzw.

$$\text{b) } \int D_{\vec{r}} G(\vec{r} - \vec{r}') \psi_p(\vec{r}') d\vec{r}' = \int \delta(\vec{r} - \vec{r}') \psi_p(\vec{r}') d\vec{r}' = \psi_p(\vec{r}) = \int G(\vec{r} - \vec{r}') f(\vec{r}') d\vec{r}'.$$

Wie man sieht, ist die Green'sche Funktion mit dem Differentialoperator verknüpft. Das heißt, dass zu jedem Differentialoperator eine bestimmte Green'sche Funktion gehört. Hat man also die Green'sche Funktion G zu einem Differentialoperator D gefunden, kann man alle inhomogenen Dgln der Form $D\psi_p = f$ mit dem Faltungsintegral (13.4) lösen. Beispiele für Green'sche Funktionen finden sich im Internet u. a. unter dem Suchbegriff *Greensche Funktion* – *Wikipedia*

¹⁶Die **Faltung** $f * g$ zweier Funktionen $f(x)$ und $g(x)$ ist **kommutativ** und über das **Faltungsintegral** wie folgt definiert: $f(x) * g(x) := \int f(x')g(x - x') dx' = \int g(x - x')f(x') dx'$
 $= g(x) * f(x) := \int g(x')f(x - x') dx' = \int f(x - x')g(x') dx'$.

Für die Ableitung einer Faltung $f * g$ mit dem Differentialoperator D gilt: $D(f * g) = (Df) * g = f * Dg$. Das **neutrale Element** der Faltung ist die δ -Funktion, denn $\delta * f = f$ und $\int \delta(x - x') dx' = 1$.

¹⁷Für die Green'sche Funktion G gilt $G(\vec{r}, \vec{r}') = G(\vec{r} - \vec{r}')$, weil G in den beiden Variablen \vec{r} und \vec{r}' symmetrisch ist und weil die δ -Funktion nur von $\vec{r} - \vec{r}'$ abhängt.

13.1 Lösung der Poisson-Gleichung als Standardbeispiel aus der Elektrostatik

	Für elektrische Punktladung q im Raumpunkt \vec{r}	Für elektrische Ladungsverteilung mit Ladungsdichte $\varrho(\vec{r}')$
elektrische Ladung	$q(\vec{r}) = \int q(\vec{r}') \delta(\vec{r} - \vec{r}') dV'$	$q = \int \varrho(\vec{r}') dV'$
elektrische Feldstärke	$\vec{E}(\vec{r}) = \frac{q}{4\pi\epsilon_0} \frac{\vec{r} - \vec{r}'}{ \vec{r} - \vec{r}' ^3}$	$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\vec{r} - \vec{r}'}{ \vec{r} - \vec{r}' ^3} \varrho(\vec{r}') dV'$
elektrostatisches Potential	$\Phi(\vec{r}) = \frac{q}{4\pi\epsilon_0} \frac{1}{ \vec{r} - \vec{r}' }$	$\Phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{ \vec{r} - \vec{r}' } \varrho(\vec{r}') dV'$

Tabelle 1 Elektrostatik

Die Quelle eines elektrostatischen Feldes der Feldstärke $\vec{E}(\vec{r})$ ist eine endlich ausgedehnte Ladungsverteilung, die beschrieben wird durch die elektrische Ladungsdichte $\varrho(\vec{r})$. Kann sich dieses Feld ungehindert im gesamten Raum (bis ins Unendliche) ausbreiten, sprechen wir von einem freien Feld. Mit dem elektrostatischen Potential Φ sind dann die zugehörigen Feldgleichungen

$$\vec{E} = -\nabla\Phi, \quad \operatorname{div} \vec{E} = \frac{1}{\epsilon_0} \varrho \quad \Rightarrow$$

$$\operatorname{div} \vec{E} = \operatorname{div}(-\nabla\Phi) = -\operatorname{div} \operatorname{grad} \Phi = -\Delta\Phi = \frac{1}{\epsilon_0} \varrho \quad \Rightarrow$$

Poisson-Gleichung: $\Delta\Phi = -\frac{1}{\epsilon_0} \varrho$

(13.5)

Die Poisson-Gleichung¹⁸ (Poisson'sche Potentialgleichung) ist eine lineare inhomogene Dgl 2. Ordnung, für die folglich das Superpositionsprinzip gilt. Die Potentialgleichung für einen Raumbereich außerhalb einer Ladungsverteilung, also für den quellenfreien Raum, ist die

$$\text{Laplace-Gleichung:} \quad \Delta\Phi = 0.$$

Die Lösung der Poisson-Gleichung (13.5) gemäß (13.4), also mit $\psi_p = \Phi$ und $f = -\frac{1}{\epsilon_0} \varrho$, ist hier sehr einfach, denn ihre Lösung steht bereits in der Tabelle 1, sodass sich die Green'sche Funktion G zum Laplace-Operator Δ aus Tabelle 1 wie folgt ergibt:

$$\begin{aligned} \Phi(\vec{r}) &= \frac{1}{4\pi\epsilon_0} \int \frac{1}{|\vec{r} - \vec{r}'|} \varrho(\vec{r}') dV' = \int \underbrace{\frac{1}{4\pi} \frac{1}{|\vec{r} - \vec{r}'|}}_{=G} \left[-\frac{1}{\epsilon_0} \varrho(\vec{r}') \right] dV' \\ &= \int G(\vec{r} - \vec{r}') \left[-\frac{1}{\epsilon_0} \varrho(\vec{r}') \right] dV'. \end{aligned}$$

¹⁸Die Poisson-Gleichung für den unendlich ausgedehnten Raum, also für „Randbedingungen im Unendlichen“, ist die freie Poisson-Gleichung.

Die nicht eindeutig bestimmte bzw. freie Green'sche Funktion zum Laplace-Operator (mit „Randbedingungen im Unendlichen“) ist also

$$G(\vec{r} - \vec{r}') = -\frac{1}{4\pi} \frac{1}{|\vec{r} - \vec{r}'|}.$$

Die **vollständig bestimmte Green'sche Funktion zum Laplace-Operator** ist dann

$$G_{\Delta}(\vec{r}, \vec{r}') = -\frac{1}{4\pi} \frac{1}{|\vec{r} - \vec{r}'|} + g(\vec{r}, \vec{r}') \quad \text{mit} \quad \Delta g(\vec{r}, \vec{r}') = 0.$$

Die Funktion g muss im Integrationsvolumen V die Bedingung $\Delta g(\vec{r}, \vec{r}') = 0$ erfüllen, ist aber unter dieser Bedingung frei wählbar, sodass sie zur Realisierung von evtl. existierenden Randbedingungen genutzt werden kann. Weiterhin kann sie in \vec{r} und \vec{r}' symmetrisch sein. Wichtig ist in diesem Zusammenhang auch die Beziehung¹⁹

$$\Delta_{\vec{r}} \frac{1}{|\vec{r} - \vec{r}'|} = \Delta_{\vec{r}'} \frac{1}{|\vec{r} - \vec{r}'|} = \Delta \frac{1}{|\vec{r} - \vec{r}'|} = -4\pi \delta(\vec{r} - \vec{r}'). \quad (13.6)$$

Wie hier dargestellt, ist der Laplace-Operator symmetrisch bezüglich \vec{r} und \vec{r}' . Die Beziehung (13.6) lässt sich jetzt leicht verifizieren, denn nach Voraussetzung (13.3) ist

$$DG = \Delta \left(-\frac{1}{4\pi} \frac{1}{|\vec{r} - \vec{r}'|} \right) = -\frac{1}{4\pi} \Delta \frac{1}{|\vec{r} - \vec{r}'|} = -\frac{1}{4\pi} \left[-4\pi \delta(\vec{r} - \vec{r}') \right] = \delta(\vec{r} - \vec{r}'). \quad \square$$

Ausgehend vom elektrostatischen Potential einer Punktladung q im Raumpunkt \vec{r}' (siehe Tabelle 1) und mit (13.6) ist die Poisson-Gleichung für diese Punktladung

$$\Delta \Phi(\vec{r}) = \Delta \left(\frac{q}{4\pi\epsilon_0} \frac{1}{|\vec{r} - \vec{r}'|} \right) = \frac{q}{4\pi\epsilon_0} \Delta \frac{1}{|\vec{r} - \vec{r}'|} = \frac{q}{4\pi\epsilon_0} \left[-4\pi \delta(\vec{r} - \vec{r}') \right] \Rightarrow$$

$$\text{Poisson-Gleichung für eine Punktladung } q: \quad \Delta \Phi(\vec{r}) = -\frac{q}{\epsilon_0} \delta(\vec{r} - \vec{r}').$$

¹⁹Siehe Herleitung im Abschnitt 5.7 meines Skripts *Grundlegendes zur Elektrodynamik und Quantenmechanik*.

14 Zur Ausbreitungsrichtung ebener Wellen

Siehe auch Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 1 – Mechanik und Wärme, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2003, Abschnitt 11.9.3 *Allgemeine Darstellung beliebiger Wellen. Wellengleichung*, Seite 364 bis Seite 365.

Aus Gründen der Bequemlichkeit betrachten wir die sich längs der z -Achse ausbreitende **harmonische ebene Welle**

$$E(z, t) = A \cdot \cos \phi(z, t)$$

mit der **Phase** (dem Phasenwinkel, dem Argument der Winkelfunktion) $\phi = \phi(z, t)$. Mit

Wellenzahl	$k = \frac{2\pi}{\lambda},$
Phasengeschwindigkeit	$v = \frac{\Delta z}{\Delta t} = \lambda \cdot \frac{1}{T} = \lambda \cdot f \Leftrightarrow \frac{v}{\lambda} = f,$
Kreisfrequenz	$\omega = 2\pi f$

erhalten wir ϕ wie folgt:

$$\phi = 2\pi \left(\frac{z}{\lambda} - \frac{t}{T} \right) = 2\pi \left(\frac{z}{\lambda} - ft \right) = 2\pi f \left(\frac{z}{v} - t \right) = kz - \omega t$$

bzw.

$$z - vt \xrightarrow{\cdot k} kz - kvt = kz - 2\pi \frac{v}{\lambda} t = kz - 2\pi f t = kz - \omega t.$$

Ausbreitung in positiver Richtung der z -Achse:

$$z = z_0 + vt \Rightarrow z_0 = z - vt \Rightarrow \phi(z, t) = kz - \omega t,$$

Ausbreitung in negativer Richtung der z -Achse:

$$z = z_0 - vt \Rightarrow z_0 = z + vt \Rightarrow \phi(z, t) = kz + \omega t.$$

Diese beiden Feststellungen hinsichtlich der Wellenausbreitungsrichtung lassen sich formal am einfachsten verifizieren, wenn wir die Bewegung einer **Wellenfront** verfolgen, also einer zur Ausbreitungsrichtung senkrechten Wellenebene **gleicher Phase**. Für die Phase einer Wellenfront gilt nämlich

$$\phi(z, t) = \text{const} \quad \text{unter den Bedingungen} \quad k = \text{const}, \quad \omega = \text{const}, \quad \Delta t \geq 0.$$

Die Zunahme von z bei Ausbreitung in positiver Richtung muss folglich kompensiert werden durch die „Abnahme“ von t mittels des negativen Vorzeichens von ωt . Umgekehrt erfordert die Ausbreitung der Welle in negativer Richtung verbunden mit der Abnahme von z das positive Vorzeichen von ωt .

Weil es sich im Komplexen oft leichter rechnet, verwendet man allgemein für harmonische ebene Wellen im Euklidischen Raum die komplexe Darstellung

$$\vec{E}(\vec{r}, t) = \vec{A} \cdot e^{i\phi} = \vec{A} \cdot e^{i(\vec{k} \cdot \vec{r} - \omega t + \phi_0)} = \vec{A} \cdot \left[\cos(\vec{k} \cdot \vec{r} - \omega t + \phi_0) + i \sin(\vec{k} \cdot \vec{r} - \omega t + \phi_0) \right]$$

mit dem **Wellenvektor** \vec{k} , der in die **Ausbreitungsrichtung** der Welle bzw. der Wellenfronten zeigt, und mit der **Phasenkonstante** ϕ_0 , d. h. dem Nullphasenwinkel bei $\vec{r} = \vec{0}$ und $t = 0$. Die Phasenkonstante darf nicht verwechselt werden mit der Phasendifferenz $\Delta\phi$ und dem dazu entsprechenden Gangunterschied (der Wegdifferenz) Δs zwischen den benachbarten Wellenfronten zweier Wellenzüge gleicher Kreisfrequenz ω . Physikalisch relevant sind am Ende aber nur der Realteil, ggf. der Imaginärteil oder das Betragsquadrat der Darstellung. Man zieht den reellen Anteil aus dem Ergebnis heraus und der Rest ist gleichsam der mathematischen Methode geschuldetes Beiwerk. Beobachtbar bzw. messbar sind nämlich nur reelle Größen.²⁰ So gilt bei Verwendung der komplexen Darstellung

$$\operatorname{Re}\{e^{i\phi}\} = \operatorname{Re}\{\cos\phi + i\sin\phi\} = \cos\phi,$$

$$\operatorname{Re}\{e^{-i\phi}\} = \operatorname{Re}\{\cos(-\phi) + i\sin(-\phi)\} = \operatorname{Re}\{\cos\phi - i\sin\phi\} = \cos\phi.$$

Wie man sieht, betrifft die Vorzeichenänderung im Exponenten nur den Imaginärteil, die Sinusfunktion. Und Bezüglich der Phasenkonstante ist beispielsweise folgendes zu berücksichtigen:

$$\phi_0 = \pm \frac{\pi}{2} \Rightarrow \begin{cases} \cos(\vec{k}\vec{r} - \omega t \pm \frac{\pi}{2}) = \mp \sin(\vec{k}\vec{r} - \omega t) \\ \sin(\vec{k}\vec{r} - \omega t \pm \frac{\pi}{2}) = \pm \cos(\vec{k}\vec{r} - \omega t) \end{cases} \Rightarrow$$

$$\begin{aligned} i = e^{i\frac{\pi}{2}} : \quad e^{i(\vec{k}\vec{r} - \omega t + \frac{\pi}{2})} &= e^{i\frac{\pi}{2}} \cdot e^{i(\vec{k}\vec{r} - \omega t)} = i \cdot e^{i(\vec{k}\vec{r} - \omega t)} \\ &= \cos(\vec{k}\vec{r} - \omega t + \frac{\pi}{2}) + i \sin(\vec{k}\vec{r} - \omega t + \frac{\pi}{2}) \\ &= -\sin(\vec{k}\vec{r} - \omega t) + i \cos(\vec{k}\vec{r} - \omega t), \end{aligned}$$

$$\begin{aligned} -i = e^{-i\frac{\pi}{2}} : \quad e^{i(\vec{k}\vec{r} - \omega t - \frac{\pi}{2})} &= e^{-i\frac{\pi}{2}} \cdot e^{i(\vec{k}\vec{r} - \omega t)} = -i \cdot e^{i(\vec{k}\vec{r} - \omega t)} \\ &= \cos(\vec{k}\vec{r} - \omega t - \frac{\pi}{2}) + i \sin(\vec{k}\vec{r} - \omega t - \frac{\pi}{2}) \\ &= \sin(\vec{k}\vec{r} - \omega t) - i \cos(\vec{k}\vec{r} - \omega t). \end{aligned}$$

Oft findet man auch andere, äquivalente komplexe Darstellungen wie beispielsweise

$$\begin{aligned} \vec{E}(\vec{r}, t) &= \vec{A} \cdot e^{-i(\omega t - \vec{k}\vec{r} + \phi_0)} = \vec{A} \cdot \left[\cos(\omega t - \vec{k}\vec{r} + \phi_0) - i \sin(\omega t - \vec{k}\vec{r} + \phi_0) \right] = \\ &= \vec{A} \cdot e^{+i(\vec{k}\vec{r} - \omega t - \phi_0)} = \vec{A} \cdot \left[\cos(\vec{k}\vec{r} - \omega t - \phi_0) + i \sin(\vec{k}\vec{r} - \omega t - \phi_0) \right] \end{aligned}$$

oder

$$\begin{aligned} \vec{E}(\vec{r}, t) &= \vec{A} \cdot e^{+i(\omega t - \vec{k}\vec{r} + \phi_0)} = \vec{A} \cdot \left[\cos(\omega t - \vec{k}\vec{r} + \phi_0) + i \sin(\omega t - \vec{k}\vec{r} + \phi_0) \right] = \\ &= \vec{A} \cdot e^{-i(\vec{k}\vec{r} - \omega t - \phi_0)} = \vec{A} \cdot \left[\cos(\vec{k}\vec{r} - \omega t - \phi_0) - i \sin(\vec{k}\vec{r} - \omega t - \phi_0) \right]. \end{aligned}$$

„Da die Maxwell'schen Gleichungen linear sind, werden der Real- und der Imaginärteil komplexer Lösungen in ihnen nicht gemischt. Finden wir also, daß bereits die *komplexen Wellen* Lösungen dieser Gleichungen sind, so können wir – ebenso wie früher bei der Behandlung des harmonischen Oszillators – durchgängig mit ihnen rechnen und den Übergang zum Realteil erst am Endergebnis vollziehen.“²¹

²⁰Das ist sogar in der Quantenmechanik so, obwohl diese a priori komplexer Natur ist. So sind dort die Wahrscheinlichkeitsamplituden (Wahrscheinlichkeitsfunktionen) Ψ nicht nur um einer besseren mathematischen Handhabung willen komplex, sondern weil die Physik dies erfordert. Messbar ist aber nicht das komplexe Ψ sondern nur dessen (reelles) Betragsquadrat, die Wahrscheinlichkeitsdichte $|\Psi|^2 = \Psi \cdot \Psi^*$.

²¹Zitat: Rainer J. Jellito, Studententext – Elektrodynamik, Theoretische Physik, 3. Auflage, Aula-Verlag, Wiesbaden, 1994, Seite 131.

15 Superposition ebener Wellen

Ebene Wellen können beschrieben werden z. B. im Reellen durch

$$E(\vec{r}, t) = A \cdot \cos(\omega t - \vec{k} \cdot \vec{r}), \quad (E, A) \in \mathbb{R}$$

oder allgemein im Komplexen durch

$$\psi(\vec{r}, t) = \alpha \cdot e^{i(\omega t - \vec{k} \cdot \vec{r})}, \quad (\psi, \alpha) \in \mathbb{C},$$

wobei es sich in diesem Fall außerdem auch noch um eine harmonische Welle handelt. Vereinfachend betrachten wir nur Wellen mit der gleichen linearen Polarisierung, sodass wir auf die vektorielle Darstellung der Wellenfelder verzichten können.

Wir werden im Folgenden die Herleitungen nur im Reellen zeigen. Das ist zwar meistens etwas „umständlich“, aber dafür etwas näher an der physikalischen Realität.

Mit E bezeichnen wir die **momentane Auslenkung** der (linear polarisierten) Welle vom Nulldurchgang und mit A bezeichnen wir die **Amplitude**, d. h. die Schwingungsbreite bzw. den Maximalwert der Auslenkung der Welle von ihrem Nulldurchgang.²²

Wir behandeln im Folgenden nur zwei einfache (spezielle) aber für das Verständnis wichtige Fälle der Überlagerung zweier Wellenzüge:

15.1 Überlagerung zweier harmonischer ebener Wellen gleicher Frequenz

(nach: Wolfgang Demtröder, Experimentalphysik 1, Springer-Lehrbuch, 2003, 3. Auflage, Seite 376 bis Seite 379)

Stationäre Interferenzstrukturen am gleichen Ort und zur gleichen Zeit entstehen nur bei der Überlagerung **kohärenter Wellen**. Kohärente Wellen sind räumlich kohärent (eben) und zeitlich kohärent (besitzen dieselbe *konstante* Kreisfrequenz ω).

Es gibt zwei Möglichkeiten zwei kohärente Teilwellen zu erzeugen:

- die phasenstarre Kopplung zweier Schwingungserreger und
 - die Erzeugung einer zweiten kohärenten Teilwelle durch Reflektion oder Brechung.
- So erhält man an jedem Ort eine zeitlich konstante, d. h. stationäre Phasendifferenz. Die Überlagerung zweier harmonischer Wellen $E_1(\vec{r}, t)$ und $E_2(\vec{r}, t)$ aus zwei phasenstarr gekoppelten Quellen liefert an einem festen Ort \vec{r}_0 die Gesamtwelle

$$\begin{aligned} E &= E_1 + E_2 \\ &= A_1 \cdot \cos(\omega t - \vec{k}_1 \cdot \vec{r}_0 + \varphi_{01}) + A_2 \cdot \cos(\omega t - \vec{k}_2 \cdot \vec{r}_0 + \varphi_{02}). \end{aligned}$$

φ_{01} und φ_{02} sind die Phasen der Wellen am Ort der Quellen zur Zeit $t = t_0$. Mit

$$\begin{aligned} \varphi_1 &= \vec{k}_1 \cdot \vec{r}_1 - \varphi_{01}, \\ \varphi_2 &= \vec{k}_2 \cdot \vec{r}_2 - \varphi_{02}, \\ \Delta\varphi &= \varphi_1 - \varphi_2 \end{aligned}$$

²²Wegen der gleichen linearen Polarisierung der in den Herleitungen verwendeten Wellenzüge, können wir die Wellenfeldgrößen wie Skalare behandeln. Andernfalls oder allgemein müssten wir statt E und A die Vektoren $\vec{E} = (E_x(\vec{r}, t), E_y(\vec{r}, t), E_z(\vec{r}, t))$ und $\vec{A} = (A_x(\vec{r}, t), A_y(\vec{r}, t), A_z(\vec{r}, t))$ verwenden.

ergibt dies

$$E = E_1 + E_2 = A_1 \cdot \cos(\omega t - \varphi_1) + A_2 \cdot \cos(\omega t - \varphi_2) . \quad (15.1)$$

Die Summe (Überlagerung) der harmonischen Wellen E_1 und E_2 mit der Kreisfrequenz ω ist ebenfalls eine harmonische Welle mit der Kreisfrequenz ω , nämlich

$$\begin{aligned} E &= E_1 + E_2 \\ &= A_1 \cdot \cos(\omega t - \varphi_1) + A_2 \cdot \cos(\omega t - \varphi_2) = A \cdot (\omega t - \varphi) . \end{aligned}$$

Daraus folgt mit dem Additionstheorem $\cos(\alpha \pm \beta) = \cos \alpha \cdot \cos \beta \mp \sin \alpha \cdot \sin \beta$

$$\left. \begin{aligned} &A_1 \cos(\omega t) \cos \varphi_1 + A_1 \sin(\omega t) \sin \varphi_1 \\ &+ A_2 \cos(\omega t) \cos \varphi_2 + A_2 \sin(\omega t) \sin \varphi_2 \end{aligned} \right\} = A \cos(\omega t) \cos \varphi + A \sin(\omega t) \sin \varphi . \quad (15.2)$$

Die Subtraktion der rechten Seite von der linken Seite der Gleichung (15.2) und geeignetes Ausklammern führt zu

$$\underbrace{(A_1 \cos \varphi_1 + A_2 \cos \varphi_2 - A \cos \varphi)}_{=0} \cdot \cos(\omega t) - \underbrace{(A_1 \sin \varphi_1 + A_2 \sin \varphi_2 - A \sin \varphi)}_{=0} \cdot \sin(\omega t) = 0 .$$

Weil dies für alle Zeiten t gelten soll, müssen die Koeffizienten bei $\cos(\omega t)$ und $\sin(\omega t)$ (beide) gleich Null sein. Weiterhin können A und φ stets so gewählt werden, dass

$$A_1 \cos \varphi_1 + A_2 \cos \varphi_2 - A \cos \varphi = 0 \quad \Leftrightarrow \quad A \cos \varphi = A_1 \cos \varphi_1 + A_2 \cos \varphi_2 , \quad (15.3)$$

$$A_1 \sin \varphi_1 + A_2 \sin \varphi_2 - A \sin \varphi = 0 \quad \Leftrightarrow \quad A \sin \varphi = A_1 \sin \varphi_1 + A_2 \sin \varphi_2 . \quad (15.4)$$

Die Division von (15.4) durch (15.3) ergibt

$$\boxed{\frac{A \cdot \sin \varphi}{A \cdot \cos \varphi} = \tan \varphi = \frac{A_1 \sin \varphi_1 + A_2 \sin \varphi_2}{A_1 \cos \varphi_1 + A_2 \cos \varphi_2} .}$$

Durch Quadrieren von (15.3) und (15.4) und anschließendes Addieren erhält man A^2 in Abhängigkeit von den Amplituden A_1 und A_2 sowie von den Phasen φ_1 und φ_2 der beiden ursprünglichen Wellen:

$$\begin{aligned} A^2 \cos^2 \varphi + A^2 \sin^2 \varphi &= A_1^2 \cos^2 \varphi_1 + A_2^2 \cos^2 \varphi_2 + 2A_1 A_2 \cos \varphi_1 \cos \varphi_2 \\ &+ A_1^2 \sin^2 \varphi_1 + A_2^2 \sin^2 \varphi_2 + 2A_1 A_2 \sin \varphi_1 \sin \varphi_2 . \end{aligned}$$

Mit $\sin^2 \varphi + \cos^2 \varphi = 1$ und mit dem bereits verwendeten Additionstheorem $\cos \varphi_1 \cdot \cos \varphi_2 + \sin \varphi_1 \cdot \sin \varphi_2 = \cos(\varphi_1 - \varphi_2) = \cos(\varphi_2 - \varphi_1) = \cos(\Delta\varphi)$ resultiert daraus schließlich die **Grundgleichung der Interferenz**

$$\boxed{A^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\Delta\varphi)} .$$

Die momentane Intensität $I(t)$ einer Welle ist proportional zum Quadrat ihrer momentanen Auslenkung $E(t)$. Für die Überlagerung zweier ebener Wellen mit der gleichen Frequenz gilt folglich mit (15.1)

$$I(t) \propto E^2 = (E_1 + E_2)^2 = [A_1 \cos(\omega t - \varphi_1) + A_2 \cos(\omega t - \varphi_2)]^2 \quad \Rightarrow$$

$$I(t) \propto A_1^2 \cos^2(\omega t - \varphi_1) + A_2^2 \cos^2(\omega t - \varphi_2) + 2A_1A_2 \cos(\omega t - \varphi_1) \cos(\omega t - \varphi_2). \quad (15.5)$$

Den zeitlichen Mittelwert \bar{I} der Intensität erhält man aus (15.5) mit

$$\begin{aligned} \overline{\cos^2(\omega t - \alpha)} &= \frac{1}{2}, \\ \overline{\cos(\omega t - \alpha)} &= 0, \\ \cos \alpha \cdot \cos \beta &= \frac{1}{2} \left[\cos(\alpha + \beta) + \cos(\alpha - \beta) \right], \\ \cos(-\alpha) &= \cos \alpha \end{aligned}$$

wie folgt:

$$\begin{aligned} \cos(\omega t - \varphi_1) \cdot \cos(\omega t - \varphi_2) &= \frac{1}{2} \left[\cos(\omega t - \varphi_1 + \omega t - \varphi_2) + \cos(\omega t - \varphi_1 - \omega t + \varphi_2) \right] \\ &= \frac{1}{2} \left[\cos(2\omega t - \varphi_1 - \varphi_2) + \underbrace{\cos(-\varphi_1 + \varphi_2)}_{=\cos(\Delta\varphi)} \right] \Rightarrow \\ \bar{I} &\propto A_1^2 \underbrace{\overline{\cos^2(\omega t - \varphi_1)}}_{=\frac{1}{2}} + A_2^2 \underbrace{\overline{\cos^2(\omega t - \varphi_2)}}_{=\frac{1}{2}} + 2A_1A_2 \cdot \frac{1}{2} \left[\underbrace{\overline{\cos(2\omega t - \varphi_1 - \varphi_2)}}_{=0} + \cos(\Delta\varphi) \right], \end{aligned}$$

$$\boxed{\bar{I} \propto \frac{1}{2} (A_1^2 + A_2^2) + A_1A_2 \cos(\Delta\varphi) = \frac{1}{2} A^2}.$$

Im Komplexen wird die Überlagerung der beiden Wellenfunktionen gleicher Frequenz beschrieben durch

$$\psi = \psi_1 + \psi_2 = A_1 \cdot e^{i(\omega t - \varphi_1)} + A_2 \cdot e^{i(\omega t - \varphi_2)}, \quad A \in \mathbb{R}.$$

Der zeitliche Mittelwert der Intensität ist dann die Hälfte des Produktes aus ψ und der komplex Konjugierten ψ^* , d. h. die **Hälfte des Betragsquadrats** von ψ :

$$\begin{aligned} \bar{I} &\propto \frac{1}{2} \cdot \psi \cdot \psi^* = \frac{1}{2} \cdot [\psi_1 + \psi_2] \cdot [\psi_1^* + \psi_2^*] \\ &\propto \frac{1}{2} \cdot [A_1 \cdot e^{i(\omega t - \varphi_1)} + A_2 \cdot e^{i(\omega t - \varphi_2)}] \cdot [A_1 \cdot e^{-i(\omega t - \varphi_1)} + A_2 \cdot e^{-i(\omega t - \varphi_2)}] \\ &\propto \frac{1}{2} \cdot e^{i\omega t} [A_1 e^{-i\varphi_1} + A_2 e^{-i\varphi_2}] \cdot e^{-i\omega t} [A_1 e^{i\varphi_1} + A_2 e^{i\varphi_2}] \\ &\propto \frac{1}{2} \cdot e^0 [A_1^2 + A_1A_2 e^{-i(\varphi_1 - \varphi_2)} + A_2A_1 e^{i(\varphi_1 - \varphi_2)} + A_2^2] \\ &\propto \frac{1}{2} \cdot \left[A_1^2 + A_2^2 + A_1A_2 \underbrace{\left(e^{i(\varphi_1 - \varphi_2)} + e^{-i(\varphi_1 - \varphi_2)} \right)}_{=2 \cos(\varphi_1 - \varphi_2)} \right] \\ \bar{I} &\propto \frac{1}{2} (A_1^2 + A_2^2) + A_1A_2 \cos(\Delta\varphi) = \frac{1}{2} A^2. \quad \square \end{aligned}$$

15.2 Überlagerung zweier Wellenzüge verschiedener Amplitude und verschiedener Frequenz

(nach: Friedrich Herrmann, Physik IV Optik, Universität Karlsruhe, 1997, Seite 14)

Wir betrachten zwei ebene und in der gleichen Richtung linear polarisierte ebene Wellenzüge, die sich mit der Geschwindigkeit

$$c = \lambda \cdot f = \frac{2\pi}{k} \cdot \frac{\omega}{2\pi} = \frac{\omega}{k} = \text{const}$$

in die gleiche Richtung, nämlich in z -Richtung ausbreiten:

$$E_1 = A_1 \cos(\omega_1 t - k_1 z),$$

$$E_2 = A_2 \cos(\omega_2 t - k_2 z).$$

Wegen $c = \text{const}$ ist $f \propto \frac{1}{\lambda} \Rightarrow \omega \propto k$, sodass keine Dispersion²³ vorliegt.

Für die Kreisfrequenzen der beiden Wellenzüge gelte $\omega_2 > \omega_1$, sodass sich mit der **Mittelfrequenz** ω_m die folgenden Zusammenhänge ergeben:

$$\Delta\omega = \omega_2 - \omega_1, \quad \omega_m = \frac{\omega_1 + \omega_2}{2}, \quad \Rightarrow \begin{cases} \omega_1 = \omega_m - \frac{\Delta\omega}{2}, \\ \omega_2 = \omega_m + \frac{\Delta\omega}{2}. \end{cases}$$

Mit der Schwingungsdauer (Periode) T gilt für die Wellenzahl k

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{T} \cdot \frac{T}{\lambda} = \omega \cdot \frac{1}{c} = \frac{\omega}{c} \quad \Rightarrow \quad k \cdot z = \frac{\omega}{c} \cdot z,$$

$$\text{mittlere Wellenzahl } k_m = \frac{\omega_m}{c} \quad \Rightarrow \quad k_m \cdot z = \frac{\omega_m}{c} \cdot z.$$

Angewandt auf unseren Beispielfall erhalten wir damit

$$\frac{\omega_1}{c} \cdot z = \frac{\omega_m - \frac{\Delta\omega}{2}}{c} \cdot z = \left(\omega_m - \frac{\Delta\omega}{2} \right) \cdot \frac{z}{c} = k_1 \cdot z,$$

$$\frac{\omega_2}{c} \cdot z = \frac{\omega_m + \frac{\Delta\omega}{2}}{c} \cdot z = \left(\omega_m + \frac{\Delta\omega}{2} \right) \cdot \frac{z}{c} = k_2 \cdot z,$$

²³Beispielsweise ist die Vakuumlichtgeschwindigkeit $v_{\text{Ph}} = c = \lambda f = \frac{\omega}{k} = \text{const}$, d. h. bei Licht im Vakuum liegt wegen $\omega(k) = c \cdot k$ keine Dispersion vor.

Ist $v_{\text{Ph}} = v_{\text{Ph}}(\omega) \neq \text{const}$ bzw. $\frac{d\omega(k)}{dk} \neq \frac{\omega(k)}{k} = v_{\text{Ph}}$, so liegt Dispersion vor, d. h. Wellen verschiedener Wellenlänge breiten sich unterschiedlich schnell aus. Die Funktion $\omega = \omega(k)$ wird deshalb auch **Dispersionsrelation** genannt. Bezüglich des Zusammenhangs zwischen Wellenlänge und Frequenz bei elektromagnetischen Wellen siehe Kapitel 18 und insbesondere Gleichung (18.2).

$$\begin{aligned}
E_1 &= A_1 \cos(\omega_1 t - k_1 z) = A_1 \cos \left[\left(\omega_m - \frac{\Delta\omega}{2} \right) t - \left(\omega_m - \frac{\Delta\omega}{2} \right) \frac{z}{c} \right] \\
&= A_1 \cos \left[\left(\omega_m - \frac{\Delta\omega}{2} \right) \left(t - \frac{z}{c} \right) \right] \\
E_1 &= A_1 \cos \left[\omega_m \left(t - \frac{z}{c} \right) - \frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right], \quad (15.6)
\end{aligned}$$

$$\begin{aligned}
E_2 &= A_2 \cos(\omega_2 t - k_2 z) = A_2 \cos \left[\left(\omega_m + \frac{\Delta\omega}{2} \right) t - \left(\omega_m + \frac{\Delta\omega}{2} \right) \frac{z}{c} \right] \\
&= A_2 \cos \left[\left(\omega_m + \frac{\Delta\omega}{2} \right) \left(t - \frac{z}{c} \right) \right] \\
E_2 &= A_2 \cos \left[\omega_m \left(t - \frac{z}{c} \right) + \frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right]. \quad (15.7)
\end{aligned}$$

Die Überlagerung der beiden Wellenzüge (15.6) und (15.7) liefert

$$E = E_1 + E_2 = A_1 \cos \left[\omega_m \left(t - \frac{z}{c} \right) - \frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] + A_2 \cos \left[\omega_m \left(t - \frac{z}{c} \right) + \frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right].$$

Dies können wir durch Anwendung des Additionstheorems

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta$$

umformen: $E =$

$$\begin{aligned}
&A_1 \left\{ \cos \left[\omega_m \left(t - \frac{z}{c} \right) \right] \cdot \cos \left[\frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] + \sin \left[\omega_m \left(t - \frac{z}{c} \right) \right] \cdot \sin \left[\frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] \right\} \\
&+ A_2 \left\{ \cos \left[\omega_m \left(t - \frac{z}{c} \right) \right] \cdot \cos \left[\frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] - \sin \left[\omega_m \left(t - \frac{z}{c} \right) \right] \cdot \sin \left[\frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] \right\},
\end{aligned}$$

$$\begin{aligned}
E &= (A_1 + A_2) \left\{ \cos \left[\omega_m \left(t - \frac{z}{c} \right) \right] \cdot \cos \left[\frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] \right\} \\
&+ (A_1 - A_2) \left\{ \sin \left[\omega_m \left(t - \frac{z}{c} \right) \right] \cdot \sin \left[\frac{\Delta\omega}{2} \left(t - \frac{z}{c} \right) \right] \right\}.
\end{aligned}$$

Mit $\omega_m \cdot \frac{z}{c} = k_m \cdot z$ und $\frac{\Delta\omega}{c} = \Delta k \Rightarrow \frac{\Delta\omega}{2} \cdot \frac{z}{c} = \frac{\Delta k}{2} \cdot z$ erhält man schließlich

$$\boxed{
\begin{aligned}
E = E_1 + E_2 &= (A_1 + A_2) \cdot \cos(\omega_m t - k_m z) \cdot \cos \left(\frac{\Delta\omega}{2} t - \frac{\Delta k}{2} z \right) \\
&+ (A_1 - A_2) \cdot \sin(\omega_m t - k_m z) \cdot \sin \left(\frac{\Delta\omega}{2} t - \frac{\Delta k}{2} z \right).
\end{aligned}
} \quad (15.8)$$

Im Fall $A_1 = A_2$ verschwindet der zweite Summand auf der rechten Seite von (15.8).

Die Überlagerung zweier Wellenzüge verschiedener Frequenz ergibt eine **Schwebungswelle**²⁴ mit der Kreisfrequenz ω_m und der Wellenzahl k_m , die moduliert wird durch eine Welle mit der Kreisfrequenz $\Delta\omega/2$ und der Wellenzahl $\Delta k/2$. Graphisch liefert die modulierende Welle die **Einhüllende** der Schwebungswelle.

Die **Phasengeschwindigkeit** der Schwebungswelle ist

$$v_{\text{Ph}} = \frac{\omega_m}{k_m} = c = \text{const.}$$

Die Ausbreitungsgeschwindigkeit der Einhüllenden bzw. der modulierenden Welle (die Geschwindigkeit der Wellengruppen) ist die **Gruppengeschwindigkeit** v_G , die Kreisfrequenzänderung pro Wellenzahländerung:

$$v_G = \frac{d\omega(k)}{dk} = \frac{d}{dk}(c \cdot k) = c = v_{\text{Ph}}.$$

Die Gleichung (15.8) können wir aber auch so interpretieren, dass die Modulationswelle die Kreisfrequenz $\frac{\Delta\omega}{2}$ und die Wellenzahl $\frac{\Delta k}{2}$ besitzt. Die Gruppengeschwindigkeit ist damit

$$v_G = \frac{\frac{\Delta\omega}{2}}{\frac{\Delta k}{2}} = \frac{\omega_2 - \omega_1}{k_2 - k_1} = \frac{c \cdot k_2 - c \cdot k_1}{k_2 - k_1} = c = v_{\text{Ph}}.$$

Wie man sieht, ist in unserem einfachen Beispiel die Phasengeschwindigkeit gleich der Gruppengeschwindigkeit, sodass keine Dispersion vorliegt. Je dichter ω_1 und ω_2 beieinander liegen, desto kleiner ist die **Frequenz der Modulationswelle**:

$$\omega = 2\pi f \quad \Rightarrow \quad \frac{\Delta\omega}{2} \hat{=} \frac{\Delta f}{2} = \frac{f_2 - f_1}{2}.$$

Die **Schwebungsfrequenz** f_S ist jedoch

$$f_S = 2 \cdot \frac{\Delta f}{2} = \Delta f = f_2 - f_1.$$

Dieser Sachverhalt lässt sich leicht demonstrieren:

Wir lassen zwei im Handel erhältliche Stimmgabeln erklingen, die eine mit $f_1 = 442$ Hz, die andere mit $f_2 = 443$ Hz, und halten sie dicht nebeneinander. Mit etwas Geschick hören wir eine Schwebung, d. h. ein periodisches An- und Abschwelen der Lautstärke, und zwar mit einer Frequenz von *einer* Schwebung pro Sekunde. So haben wir die Schwebungsfrequenz $f_S = f_2 - f_1 = 1$ Hz ermittelt. Die Modulationswelle hingegen hat die Frequenz $\frac{\Delta f}{2} = 0,5$ Hz, also die Hälfte der Schwebungsfrequenz. Das liegt daran, dass die Schwebungswelle sowohl bei den Maxima als auch bei den Minima der Modulationswelle ein Maximum besitzt, sodass auf eine Periode der Modulationswelle zwei Maxima der Schwebungswelle (Schwebungen) kommen.

²⁴Im Internet findet man Videos von Versuchen zur Schwebung und schöne graphische Darstellungen und Animationen der Überlagerung von zwei harmonischen Schwingungen verschiedener Frequenz z. B. unter dem Suchbegriff *Schwebung* - *Wikipedia*.

16 Brechungsindex und das Verhalten von Licht an Grenzflächen

- Brechungsindex-Definition:

Für die Ausbreitungsgeschwindigkeit oder exakt die **Phasengeschwindigkeit** von elektromagnetischen Wellen wie z. B. Licht gilt:

$c_0 = \text{const}$ Vakuumlichtgeschwindigkeit, **physikalische Fundamentalkonstante**,
 c Phasengeschwindigkeit in einem Medium mit dem **Brechungsindex** n .

$$\text{Brechungsindex } n = \frac{c_0}{c} \Leftrightarrow c = c(n) = c_0 \frac{1}{n} \Rightarrow c \propto \frac{1}{n}, \quad c_0 > c \Rightarrow n > 1.$$

Um die Phasengeschwindigkeit im Vakuum von der in Medien zu unterscheiden, sind verschiedene Indizierungen üblich. Oft wird die Geschwindigkeit im Vakuum nicht indiziert und die Geschwindigkeit in Medien gestrichen.

- Reflektion und Transmission ändern nicht die Frequenz f bzw. die Kreisfrequenz $\omega = 2\pi f$ der einfallenden elektromagnetischen Wellen, auch wenn sich bei der Transmission der Brechungsindex ändert. Die Änderung des Brechungsindex hat nur eine Änderung der Wellenlänge λ bzw. eine Änderung der Wellenzahl $k = \frac{2\pi}{\lambda}$ zur Folge und damit auch eine Änderung der Lichtgeschwindigkeit (im Medium):

$$c = \frac{c_0}{n} = \lambda \cdot f = \frac{\omega}{k}, \quad c_0 = \text{const}, \quad \omega = \text{const} \Rightarrow$$

$$k = k(n) = \frac{\omega}{c_0} \cdot n \Rightarrow k \propto n.$$

- Zusammenhang zwischen Permittivität, magnetischer Permeabilität und Brechungsindex:

$\varepsilon = \varepsilon_r \cdot \varepsilon_0$ mit

- ε Permittivität des Mediums,
- ε_r relative Permittivität (Permittivitätszahl) des Mediums, dimensionslos,
- ε_0 elektrische Feldkonstante (Dielektrizitätskonstante).

$\mu = \mu_r \cdot \mu_0$ mit

- μ magnetische Permeabilität des Mediums,
- μ_r relative magnetische Permeabilität (Permeabilitätszahl) des Mediums, dimensionslos,
- μ_0 magnetische Feldkonstante (magnetische Permeabilität des Vakuums).

Im Vakuum gilt $\varepsilon_r = \mu_r = 1$. Der Brechungsindex eines Mediums ergibt sich aus seiner Permittivität und seiner magnetischen Permeabilität wie folgt:

$$\left. \begin{aligned} c_0 &= \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \\ c &= \frac{1}{\sqrt{\varepsilon \mu}} \end{aligned} \right\} \Rightarrow n = \sqrt{\varepsilon_r \mu_r}.$$

- Reflektion:

Licht falle in das Medium I mit dem Brechungsindex n_1 ein, treffe dann auf die Grenzfläche zum Medium II mit dem Brechungsindex n_2 und werde dann in das Medium I reflektiert. Allgemein wird das einfallende Licht nicht vollständig reflektiert, sondern es wird bei bestimmten Gegebenheiten auch durch die Grenzfläche in das Medium II hinein transmittiert, z. B. bei der Lichtbrechung an Grenzflächen. Betrachten wir den *innerhalb des Mediums I* auf die Grenzfläche einfallenden Lichtanteil und den von der Grenzfläche reflektierten Lichtanteil, so ergeben sich für die Reflektion zwei Fälle:

$n_1 < n_2 \Rightarrow$ **äußere Reflektion**,

$n_1 > n_2 \Rightarrow$ **innere Reflektion**.

Die vollständige innere Reflektion nennt man auch kurz Totalreflektion.

- Grenzwinkel α_g der Totalreflektion:

Beim Lichtübergang an einer Grenzfläche von einem Medium mit größerem Brechungsindex n_1 in ein Medium mit kleinerem Brechungsindex n_2 ,

$$n_1 > n_2,$$

wird das Licht mit dem Ausfallswinkel α_2 vom Ausfallslot weg gebrochen, wenn der Einfallswinkel α_1 den Grenzwinkel der Totalreflektion α_g nicht überschreitet. Der **Grenzwinkel der Totalreflektion** ist der zum Ausfallswinkel $\alpha_2 = 90^\circ$ korrespondierende Einfallswinkel $\alpha_1 = \alpha_g$:

$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2 = n_2 \sin 90^\circ = n_2 \quad \Leftrightarrow \quad \sin \alpha_1 = \sin \alpha_g = \frac{n_2}{n_1},$$

$$\text{Grenzwinkel der Totalreflektion } \alpha_g = \arcsin \frac{n_2}{n_1}, \quad n_1 > n_2,$$

d. h., die Totalreflektion an der Grenzfläche tritt ein bei einem Einfallswinkel α_1 entsprechend $\sin \alpha_1 > \frac{n_2}{n_1}$.

- Evaneszentes Wellenfeld – optischer Tunneleffekt (siehe Abschnitt 18).
- Brewsterwinkel:

Der Grenzwinkel der Totalreflektion darf nicht mit dem **Brewsterwinkel** oder Polarisationswinkel α_B verwechselt werden. Dieser spielt bei der Reflektion eine Rolle, die beim Lichtübergang von einem optisch dünneren Medium mit n_1 in ein optisch dichteres Medium mit n_2 auftritt, und muss die Bedingung

$$\alpha_1 + \alpha_2 \stackrel{!}{=} 90^\circ \quad \Leftrightarrow \quad \alpha_1 = \alpha_B \quad \Rightarrow \quad \alpha_2 = 90^\circ - \alpha_B$$

erfüllen. Dabei ist α_1 sowohl der Einfallswinkel als auch der Reflektionswinkel und α_2 ist der Brechungswinkel. Das Snellius'sche Brechungsgesetz liefert schließlich

$$n_1 \sin \alpha_1 = n_1 \sin \alpha_B = n_2 \sin \alpha_2 = n_2 \sin(90^\circ - \alpha_B) = n_2 \cos \alpha_B \quad \Rightarrow \quad \frac{n_2}{n_1} = \frac{\sin \alpha_B}{\cos \alpha_B} = \tan \alpha_B,$$

$$\text{Brewsterwinkel } \alpha_B = \arctan \frac{n_2}{n_1}, \quad n_1 < n_2.$$

Trifft unpolarisiertes oder zirkular polarisiertes Licht mit dem Einfallswinkel α_B auf eine Glasoberfläche, so wird nur der senkrecht zur Einfallsebene (Ein-Ausfallsebene) polarisierte Anteil oder die senkrecht zur Einfallsebene polarisierte Komponente des Lichts reflektiert, weil die Abstrahlung elektromagnetischer Wellen nur senkrecht zum Dipol und nicht in Dipolrichtung erfolgt.

- Auch bei der Beschreibung der Reflektion von elektromagnetischen Wellen an Metalloberflächen spielt der Brechungsindex eine Rolle. Allerdings ist er dann komplex, also $n(\lambda) \in \mathbb{C}$, wobei sein Realteil durchaus kleiner als 1 werden kann.

17 Intensität elektromagnetischer Wellen

Siehe auch: Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 2, Elektrizität und Optik, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2004, Abschnitt 7.5 *Das Magnetfeld elektromagnetischer Wellen* und Abschnitt 7.6 *Energie- und Impulstransport durch elektromagnetische Wellen*, Seite 193 bis Seite 195.

Die (momentane) Energieflussdichte, auch Energiestromdichte genannt, ist definiert durch den

$$\text{Poynting-Vektor } \vec{S} = \vec{E} \times \vec{H} = \frac{1}{\mu_0} (\vec{E} \times \vec{B}) = \varepsilon_0 c_0^2 (\vec{E} \times \vec{B}).$$

\vec{H} ist die magnetische Feldstärke und \vec{B} die magnetische Flussdichte. Die Vektoren \vec{S} , \vec{E} , \vec{H} bilden ein rechtshändiges, rechtwinkliges Dreiein. **Im Vakuum** propagiert die damit

beschriebene elektromagnetische Welle in Richtung des Wellenvektors \vec{k} , wobei dann $\vec{k} \uparrow\uparrow \vec{S}$ ist und folglich auch \vec{k} , \vec{E} , \vec{B} ein rechtshändiges Dreiein bilden gemäß

$$\vec{B} = \frac{1}{\omega} (\vec{k} \times \vec{E}) \quad \Rightarrow \quad B = \frac{1}{\omega} (k \cdot E).$$

Dabei haben wir $\omega = c_0 \cdot k \Leftrightarrow \frac{1}{c_0 k} = \frac{1}{\omega}$ verwendet.

Für eine ebene elektromagnetische Welle **im Vakuum** gilt folglich:

Die Feldvektoren \vec{E} und \vec{B} stehen senkrecht aufeinander, schwingen in Phase und stehen senkrecht zur Ausbreitungsrichtung, definiert durch den Wellenvektor \vec{k} , sodass für den Betrag von \vec{B} gilt

$$|\vec{B}| = \frac{1}{c_0} |\vec{E}| \quad \Rightarrow \quad B = \frac{1}{c_0} E.$$

Achtung! Bei Anwesenheit von elektrischen Strömen oder/und Ladungen stehen \vec{E} und \vec{B} nicht unbedingt senkrecht aufeinander und in anisotropen Medien können \vec{S} und \vec{k} in verschiedene Richtungen zeigen.

Der zeitliche Mittelwert des Betrags des Poynting-Vektors (im Vakuum) über ein Zeitintervall $\Delta t \gg T = \frac{1}{\nu} = \frac{\lambda}{c_0}$ ist die²⁵

$$\text{Intensität } I = \langle |\vec{S}| \rangle = \varepsilon_0 c_0^2 \langle |\vec{E} \times \vec{B}| \rangle = \varepsilon_0 c_0 \langle E^2 \rangle.$$

Mit dem Betrag A der Amplitude \vec{A} des elektrischen Feldes $\vec{E} = \vec{A} \cdot \cos(\vec{k}\vec{r} - \omega t)$ ist der zeitliche Mittelwert von E^2 :

$$\langle E^2 \rangle = A^2 \cdot \underbrace{\langle \cos^2(\vec{k}\vec{r} - \omega t) \rangle}_{= \frac{1}{2} \text{ für } \Delta t \gg T} = \frac{1}{2} A^2.$$

Damit ist die Intensität einer ebenen elektromagnetischen Welle im Vakuum

$$I = \frac{1}{2} \varepsilon_0 c_0 A^2. \quad (17.1)$$

Sinngemäß ist die Intensität die pro Zeiteinheit durch eine Flächeneinheit fließende zeitlich gemittelte Energie. Sie besitzt demzufolge die Maßeinheit

$$[I] = 1 \frac{\text{J}}{\text{s} \cdot \text{m}^2} = 1 \text{ W/m}^2.$$

Betrachten wir die Intensität einer elektromagnetischen Welle innerhalb eines homogenen, isotropen Dielektrikums (Mediums) mit dem Brechungsindex n , so wird in (17.1) aus der elektrischen Feldkonstante bzw. der Permittivität des Vakuums ε_0 die Permittivität des Mediums

$$\varepsilon = \varepsilon_r \cdot \varepsilon_0$$

und aus der Vakuumlichtgeschwindigkeit c_0 die Lichtgeschwindigkeit im Medium

$$c = \frac{c_0}{n}.$$

Die Intensität einer ebenen elektromagnetischen Welle in einem homogenen, isotropen Medium ist damit schließlich

$$I = \frac{1}{2} \varepsilon_r \varepsilon_0 \frac{c_0}{n} A^2 = \frac{1}{2} \varepsilon c A^2.$$

Achtung! Überlagerte kohärente Wellenzüge unterscheiden sich in ihrer Intensität von überlagerten inkohärenten Wellenzügen:

Siehe dazu die Abschnitte *Kohärente Streuung und Interferenz* sowie *Inkohärente Streuung* unter www.ieap.uni-kiel.de/et/people/wimmer/teaching/Phys_II/P2_V8.pdf

²⁵Die Intensität wird in der Optik auch **Bestrahlungsstärke** genannt.

- Überlagerung kohärenter Wellenzüge (kohärente Streuung):
Die Intensität entspricht dem Quadrat der Summe aller Amplituden der Wellenzüge (aller Streuamplituden).

In Abhängigkeit von der Phasendifferenz δ bzw. dem Gangunterschied Δs zwischen den Wellenzügen liefert die Intensität ein Interferenzmuster (siehe auch Abschnitt 19.2.1) mit den Maxima

$$I_{\max} = \frac{1}{2} \varepsilon_0 c_0 (A_1 + A_2)^2 \quad \text{für } \delta = m \cdot 2\pi, \quad m = 0, 1, 2, \dots$$

und den Minima

$$I_{\min} = \frac{1}{2} \varepsilon_0 c_0 (A_1 - A_2)^2 \quad \text{für } \delta = (2m + 1) \cdot \pi, \quad m = 0, 1, 2, \dots$$

- Überlagerung inkohärenter Wellenzüge (inkohärente Streuung):
Die Intensität ist gleich der Summe der Intensitäten der einzelnen Wellenzüge (der gestreuten Wellen):

$$I = I_1 + I_2 = \frac{1}{2} \varepsilon_0 c_0 (A_1^2 + A_2^2).$$

Wie man sieht, entsteht bei Inkohärenz der Wellenzüge kein Interferenzmuster.

Kugelwellen

Beim Umgang mit (zentralsymmetrischen) Kugelwellen ist es bequemer, Polarkoordinaten zu verwenden. So bezeichnet die Koordinate $|\vec{r}| = r \in \mathbb{R}$ den räumlichen Abstand eines Punktes im Kugelwellenfeld von der Quelle. Weil bei Kugelwellen $\vec{k} \uparrow \vec{r}$ gilt, können wir statt $\vec{k} \cdot \vec{r}$ vereinfachend kr schreiben. Für die Intensität bzw. Bestrahlungsstärke hatten wir bereits festgestellt:

$$I \propto |\vec{E}|^2 = E^2. \quad (17.2)$$

Im Gegensatz zur ebenen Welle $\vec{E}(\vec{r}, t) = A e^{i\vec{k}\vec{r} - \omega t}$ ist bei der Kugelwelle die Phase neben t insbesondere von r abhängig und die Intensität und folglich auch die Amplitude A hängen nur vom Abstand r zur Punktquelle ab. Aus der vektoriellen Darstellung $\vec{E}(\vec{r}, t)$ in kartesischen Koordinaten wird somit $E(r, t)$ in Polarkoordinaten. Weil für die Intensität der Kugelwelle das Abstandsgesetz gilt, also

$$I \propto \frac{1}{r^2},$$

müssen wegen (17.2) für die **Kugelwelle** die folgenden Beziehungen gelten:

$$\boxed{E \propto \frac{1}{r} \quad \Rightarrow \quad E(r, t) = A(r) e^{i(kr - \omega t)} = \frac{A_0}{r} e^{i(kr - \omega t)}}.$$

Dabei ist die Amplitude $A(r) = \frac{A_0}{r}$ umgekehrt proportional zum Abstand r von der Quelle gemäß

$$A(r) \propto \frac{1}{r}.$$

$A_0 = \text{const}$ ist hier also nicht die Amplitude, was man auch an der Dimension von A_0 sieht.

18 Evaneszentes Feld – Goos-Hänchen-Verschiebung

Quellen:

Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 2, Elektrizität und Optik, Springer-Verlag, Berlin, Heidelberg, New York, 3. Auflage, 2004, Abschnitt 8.4 *Wellen an Grenzflächen zwischen zwei Medien*, Seite 228 bis Seite 231.

ChemgaPedia, Optische Grundlagen der Sensorik – Das evaneszente Feld,

http://www.chemgapedia.de/vsengine/vlu/vsc/de/ch/13/vlu/sensorik/opt_grundlagen.vlu/Page/vsc/de/ch/13/pc/sensorik/optik/evaneszent.vscml.html

Wird beim Übergang des Lichts an der Grenzfläche von einem optisch dichteren Medium I (mit dem Brechungsindex n_1) zu einem optisch dünneren Medium II (mit dem Brechungsindex $n_2 < n_1$) der Grenzwinkel α_g der Totalreflexion erreicht bzw. überschritten, so tritt innere Totalreflexion ein, die kurz **Totalreflexion** genannt wird.

„Goos und Hänchen stellten in Experimenten zur Totalreflexion fest, dass der reflektierte Strahl nicht am Ort des Auftreffens, sondern erst etwas versetzt zurückläuft.

Mit der Annahme, dass das Licht etwas in das optisch dünnere Medium eindringt, lässt sich dies mit Hilfe der geometrischen Optik verdeutlichen.

Abb. 1 Goos-Hänchen-Verschiebung

Offenbar fällt das elektromagnetische Feld der totalreflektierten Welle an der Grenzfläche nicht abrupt auf Null ab. Das folgt aus den Stetigkeitsbedingungen für das elektromagnetische Feld bei Totalreflexion vom optisch dichteren ins optisch dünnere Medium. Für $\alpha_1 > \alpha_g$ entsteht eine entlang der Grenzfläche propagierende Welle, die exponentiell mit der Eindringtiefe gedämpft wird. Diese Welle im optisch dünneren Medium wird als **evaneszentes Feld** bezeichnet. Der Abklingfaktor beschreibt, wie weit das evaneszente Feld in das optisch dünnere Medium eindringt: ... nur strömt keine Energie über – durch die Grenzfläche hindurchtretende Energie strömt direkt wieder zurück. Gibt man der Strahlungsenergie allerdings innerhalb der Eindringtiefe die Möglichkeit zur Wechselwirkung, kann sie zurückgehalten werden, z. B. durch Absorption, Anregung von Fluoreszenz oder durch ein weiteres Prisma oder einen Wellenleiter, in dem das Licht propagieren kann. Man bezeichnet dies als **abgeschwächte Totalreflexion** bzw. **frustrierte Totalreflexion**.²⁶

Das evaneszente²⁷ Feld führt also bei Totalreflexion keine Energie über das optisch dünnere Medium II ab.²⁸

Führen wir jedoch ein drittes und wieder optisch dichteres Medium III (mit dem Brechungsindex $n_3 > n_2$) mit seiner Grenzfläche in das evaneszente Feld hinein, dann dringt das evaneszente Feld in dieses Medium ein und transportiert fortwährend Energie in Form von elektromagnetischen Wellen (Licht) in das Medium III. Um diesen transmittierten Lichtanteil reduziert sich der innerhalb des Mediums I reflektierte Lichtanteil, sodass wir dann nicht mehr von Totalreflexion sprechen, sondern nur noch von **frustrierter Totalreflexion** – auch abgeschwächte, verhinderte oder gestörte

²⁶Zitiert aus: ChemgaPedia, Optische Grundlagen der Sensorik – Das evaneszente Feld.

²⁷Evanescere (lat.) bedeutet verschwinden, sich verflüchtigen.

²⁸Dies gilt auch allgemein für die **vollständige Reflexion** an Grenzflächen zwischen zwei Medien.

innere Totalreflexion genannt.²⁹ Folglich lässt sich das ursprünglich einfallende Licht durch entsprechende optische Vorrichtungen aufteilen in einen reflektierten und einen transmittierten Anteil (Strahlenteiler).

Die bei frustrierter Totalreflexion auftretende Transmission elektromagnetischer Wellen vom Medium I – durch die Barriere des optisch dünneren Mediums II – hinein in das Medium III lässt sich quantenphysikalisch und völlig analog zum Tunneleffekt erklären (siehe Abschnitt 6.1 in meinem Skript: Quantenmechanik – Schrödinger-Gleichung – Tunneleffekt). Man bezeichnet dieses Phänomen deshalb auch als den **optischen Tunneleffekt**.

Bei quantenphysikalischer Interpretation entspricht dort das Medium I dem Gebiet I, das Medium II dem Gebiet II, d. h. der Potentialbarriere, und das Medium III dem Gebiet III. Im klassisch nicht erlaubten Gebiet II (Medium II) ist demzufolge der Wellenvektor bzw. dort die Wellenzahl k komplexwertig, nämlich $k \equiv i\kappa$, $\kappa \in \mathbb{R}$. Dadurch wird der komplexwertige Exponent der einfallenden Welle (Wahrscheinlichkeitsamplitude) beim Übergang in das Gebiet II reell und es resultiert eine abklingende Exponentialfunktion, völlig analog zum evaneszenten Feld.³⁰

$$\Psi_I(x) = A \cdot e^{ikx} \quad \Rightarrow \quad \Psi_{II}(x) = C \cdot e^{i(i\kappa)x} = C \cdot e^{-\kappa x} .$$

Die folgende Plausibilisierung des evaneszenten Feldes erfolgt im klassischen Wellenbild. Für den Betrag von Vektorkomponenten beispielsweise $|\vec{E}_{1T}|$ schreiben wir kurz E_{1T} . Zur Vereinfachung und ohne Einschränkung der Allgemeinheit drehen wir das kartesische Koordinatensystem so, dass die Einfallsebene des Lichts mit der Ebene $:= (x, y, z = 0)$ zusammenfällt und gleichzeitig die Grenzfläche zwischen den Medien I und II in der Ebene $:= (x, y = 0, z)$ liegt (siehe Abbildung 6). Folglich und wie wir gleich sehen werden, liegt dann sowohl der Wellenvektor \vec{k}_i als auch dessen Tangentialkomponente k_{iT} von einfallendem, von reflektiertem und von transmittiertem Licht in der Ebene $:= (x, y, 0)$. Alle Größen im Medium I indizieren wir mit 1 und alle Größen im Medium II mit 2. Größen, die sich auf das reflektierte Licht im Medium I beziehen, versehen wir zusätzlich mit einem Strichindex. Die Tangentialkomponenten der elektrischen Felder und der Wellenvektoren bzw. Wellenzahlen erhalten den Index T, weil sie tangential zur Grenzfläche verlaufen.

Wir betrachten die drei ebenen Wellen

$$\begin{aligned} \vec{E}_1 &= \vec{A}_1 \cdot e^{i(\vec{k}_1 \vec{r} - \omega_1 t)} && \text{einfallendes Licht ,} \\ \vec{E}'_1 &= \vec{A}'_1 \cdot e^{i(\vec{k}'_1 \vec{r} - \omega'_1 t)} && \text{reflektiertes Licht ,} \\ \vec{E}_2 &= \vec{A}_2 \cdot e^{i(\vec{k}_2 \vec{r} - \omega_2 t)} && \text{transmittiertes bzw. gebrochenes Licht .} \end{aligned}$$

Aus der Stetigkeit des E -Feldes an Grenzflächen folgt, dass die Tangentialebene des E -Feldes auf der Einfallseite der Grenzfläche (Medium I), also die Summe $E_{1T} + E'_{1T}$

²⁹Im Zusammenhang mit der Transmission bei frustrierter Totalreflexion spricht man auch von einer *Auskopplung* des evaneszenten Feldes.

³⁰Zur Herstellung der Analogie zwischen elektromagnetischen Wellen und den quantenphysikalischen Wahrscheinlichkeitsamplituden setzen wir im Exponenten der elektromagnetischen Wellenfunktion $t = 0$:

$$E(x, t) = A \cdot e^{i(kx - \omega t)} \xrightarrow{t=0} E(x, 0) = A \cdot e^{ikx} .$$

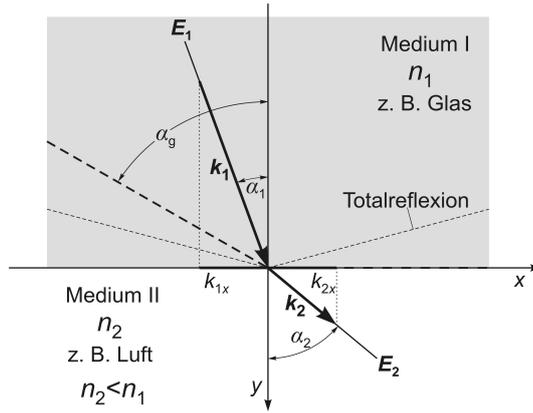


Abb. 6 Die ebene Welle \vec{E}_1 trifft im Medium I mit dem Brechungsindex $n_1 > n_2$ auf die Grenzfläche zum Medium II mit dem Brechungsindex n_2 . Die transmittierte ebene Welle ist dann \vec{E}_2 . Die Grenzfläche ist die Ebene, die von den z, x -Koordinatenachsen aufgespannt wird. Die \vec{E}_1 -Einfallsebene und die \vec{E}_2 -Ausfallsebene liegen in der Ebene der x, y -Koordinatenachsen, sodass die y -Achse das Einfallslot bildet. Wie man sieht, sind die Komponenten k_{1x} von \vec{k}_1 und k_{2x} von \vec{k}_2 gleich lang. Hingegen ist der Wellenvektor \vec{k}_2 kürzer als der Wellenvektor \vec{k}_1 , weil die Wellenlänge im Medium mit dem kleineren Brechungsindex n_2 größer ist als die Wellenlänge im Medium mit dem größeren Brechungsindex n_1 .

der Tangentialkomponenten des E -Feldes von einfallendem und reflektiertem Licht, gleich sein muss der Tangentialkomponente E_{2T} des E -Feldes auf der Ausfallsseite der Grenzfläche (Medium II):

$$E_{1T} + E'_{1T} = E_{2T} . \quad (18.1)$$

Betrachten wir jetzt die drei Tangentialkomponenten in dem festen Raumpunkt $\vec{r} = 0$, so erhalten wir

$$\begin{aligned} E_{iT} &= A_{iT} \cdot e^{i(k_{iT} \cdot 0 - \omega_i t)} = A_{iT} \cdot e^{-i\omega_i t} \Rightarrow \\ A_{1T} \cdot e^{-i\omega_1 t} + A'_{1T} \cdot e^{-i\omega'_1 t} &= A_{2T} \cdot e^{-i\omega_2 t} . \end{aligned}$$

Diese Gleichung muss für beliebige Zeiten t gelten und hat deshalb nur dann nichttriviale Lösungen, wenn

$$\boxed{\omega_1 \equiv \omega'_1 \equiv \omega_2 \equiv \omega} . \quad (18.2)$$

Beim Übergang einer elektromagnetischen Welle in ein Medium mit anderem Brechungsindex kann sich nur die Wellenlänge ändern. Die Frequenz bzw. Kreisfrequenz ändert sich dabei nicht.

Die Bedingung (18.1) gilt für beliebige Punkte \vec{r} der Grenzfläche. Deshalb müssen die Phasen der drei Wellen E_{1T} , E'_{1T} und E_{2T} in jedem Punkt \vec{r} der Grenzfläche gleich sein:

$$\vec{k}_1 \cdot \vec{r} - \omega_1 t = \vec{k}'_1 \cdot \vec{r} - \omega'_1 t = \vec{k}_2 \cdot \vec{r} - \omega_2 t .$$

Unter Berücksichtigung von (18.2) folgt daraus die wichtige Beziehung

$$\boxed{\vec{k}_1 \cdot \vec{r} = \vec{k}'_1 \cdot \vec{r} = \vec{k}_2 \cdot \vec{r}} . \quad (18.3)$$

Da der Wellenvektor \vec{k}_1 in der Einfallsebene, d. h. in der $(x, y, 0)$ -Ebene liegt und \vec{r} in der Grenzfläche, d. h. in der $(x, 0, z)$ -Ebene, erhalten wir in Komponentendarstellung

$$\vec{k}_1 = k_{1x}\vec{e}_x + k_{1y}\vec{e}_y, \quad \vec{r} = x\vec{e}_x + z\vec{e}_z. \quad (18.4)$$

Und weil die Richtungen der Wellenvektoren \vec{k}'_1 und \vec{k}_2 noch unbekannt sind, machen wir den allgemeinen Ansatz

$$\begin{aligned} \vec{k}'_1 &= k'_{1x}\vec{e}_x + k'_{1y}\vec{e}_y + k'_{1z}\vec{e}_z, \\ \vec{k}_2 &= k_{2x}\vec{e}_x + k_{2y}\vec{e}_y + k_{2z}\vec{e}_z. \end{aligned} \quad (18.5)$$

Einsetzen von (18.4) und (18.5) in (18.3) liefert die Gleichung

$$k_{1x}x = k'_{1x}x + k'_{1z}z = k_{2x}x + k_{2z}z.$$

Diese Gleichung muss für beliebige Werte x und z der Grenzfläche gelten, woraus folgt:

$$k'_{1z} = k_{2z} = 0 \quad \Rightarrow$$

$$\boxed{k_{1x} = k'_{1x} = k_{2x}}. \quad (18.6)$$

Die Wellenvektoren der einfallenden, der reflektierten und der gebrochenen Welle liegen in der Einfallsebene.

Der Abbildung 6 entnehmen wir

$$k_{1x} = |\vec{k}_1| \cdot \sin \alpha_1 = k_1 \cdot \sin \alpha_1 = n_1 \frac{\omega}{c_0} \cdot \sin \alpha_1, \quad (18.7)$$

$$k'_{1x} = |\vec{k}'_1| \cdot \sin \alpha'_1 = k'_1 \cdot \sin \alpha'_1 = n'_1 \frac{\omega}{c_0} \cdot \sin \alpha'_1,$$

$$k_{2x} = |\vec{k}_2| \cdot \sin \alpha_2 = k_2 \cdot \sin \alpha_2 = n_2 \frac{\omega}{c_0} \cdot \sin \alpha_2. \quad (18.8)$$

Mit (18.6) folgt daraus einerseits

$$k_{1x} = k'_{1x} \quad \wedge \quad n_1 = n'_1 \quad \Rightarrow \quad n_1 \frac{\omega}{c_0} \cdot \sin \alpha_1 = n_1 \frac{\omega}{c_0} \cdot \sin \alpha'_1 \quad \Rightarrow$$

$$\sin \alpha_1 = \sin \alpha'_1 \quad \Rightarrow \quad \alpha_1 = \alpha'_1, \quad \text{Einfallswinkel gleich Reflexionswinkel}$$

und andererseits

$$k_{1x} = k_{2x} \quad \Rightarrow \quad n_1 \frac{\omega}{c_0} \cdot \sin \alpha_1 = n_2 \frac{\omega}{c_0} \cdot \sin \alpha_2 \quad \Rightarrow$$

$$\text{Snellius'sches Brechungsgesetz} \quad n_1 \cdot \sin \alpha_1 = n_2 \cdot \sin \alpha_2.$$

Der Betrag des Wellenvektors \vec{k}_2 der transmittierten Welle \vec{E}_2 ist

$$|\vec{k}_2| = k_2 = n_2 \frac{\omega}{c_0} = \sqrt{k_{2x}^2 + k_{2y}^2}.$$

Quadrieren dieser Gleichung und Umstellung nach k_{2y}^2 ergibt

$$k_{2y}^2 = \left(n_2 \frac{\omega}{c_0} \right)^2 - k_{2x}^2 .$$

Setzen wir jetzt gemäß (18.6) und (18.7) $k_{2x}^2 = k_{1x}^2$ ein, so erhalten wir die Funktion $k_{2y}(\alpha_1)$, also die Abhängigkeit der y -Komponente des Wellenvektors der transmittierten Welle vom Einfallswinkel:

$$k_{2y}^2 = \left(n_2 \frac{\omega}{c_0} \right)^2 - \left(n_1 \frac{\omega}{c_0} \right)^2 \cdot \sin^2 \alpha_1 = \underbrace{\left(n_2 \frac{\omega}{c_0} \right)^2}_{=k_2^2} \cdot \left(1 - \frac{n_1^2}{n_2^2} \sin^2 \alpha_1 \right) .$$

Der erste Faktor k_2^2 auf der rechten Seite dieser Gleichung ist positiv. Wie aber verhält sich der zweite Faktor in Abhängigkeit vom Einfallswinkel α_1 ?

Wenn der Einfallswinkel α_1 gleich dem Grenzwinkel der Totalreflexion α_g ist, also

$$\sin \alpha_1 = \sin \alpha_g = \frac{n_2}{n_1} \quad \Rightarrow \quad \sin^2 \alpha_1 = \frac{n_2^2}{n_1^2} ,$$

verschwindet der zweite Faktor, sodass, wie für α_g zu erwarten war, $k_{2y} = 0$ resultiert. Wenn aber der Einfallswinkel größer ist als der Grenzwinkel der Totalreflexion, also

$$\sin \alpha_1 > \sin \alpha_g \quad \Rightarrow \quad \sin^2 \alpha_1 > \frac{n_2^2}{n_1^2} ,$$

so wird

$$\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 > 1$$

und der zweite Faktor wird negativ:

$$\left(1 - \frac{n_1^2}{n_2^2} \sin^2 \alpha_1 \right) := -\beta^2 < 0 .$$

Daraus folgt für die y -Komponente von \vec{k}_2 :

$$k_{2y}^2 = \underbrace{\left(n_2 \frac{\omega}{c_0} \right)^2}_{=k_2^2} \cdot (-1) \cdot \underbrace{\left(\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 - 1 \right)}_{=\beta^2} \quad \Rightarrow$$

$$k_{2y} = \pm k_2 \cdot i \cdot \underbrace{\sqrt{\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 - 1}}_{=\beta} ,$$

$$\boxed{k_{2y} = \pm i k_2 \beta \quad \text{für} \quad \alpha_1 > \alpha_g} . \quad (18.9)$$

Jetzt setzen wir die positive Lösung von (18.9) in die transmittierte ebene Welle \vec{E}_2 ein:

$$\vec{E}_2 = \vec{A}_2 \cdot e^{i(\vec{k}_2 \vec{r} - \omega t)} = \vec{A}_2 \cdot e^{i(k_{2x} \cdot x + k_{2y} \cdot y - \omega t)} = \vec{A}_2 \cdot e^{i(k_{2x} \cdot x + i k_2 \beta \cdot y - \omega t)} \quad \Rightarrow$$

evaneszentes Feld $\vec{E}_2 = \vec{A}_2 \cdot \underbrace{e^{-\beta k_2 \cdot |y|}}_{\text{Dämpfung}} \cdot \underbrace{e^{i(k_2 x \cdot x - \omega t)}}_{\text{Oszillation}}, \quad \alpha_1 > \alpha_g$.

Wie man sieht, oszilliert das evaneszente Feld in x -Richtung und mit der Zeit. Es ist also eine sich längs der Grenzfläche ausbreitende **Longitudinalwelle** und wird deshalb auch evaneszente Welle genannt. Gleichzeitig wird das evaneszente Feld in y -Richtung exponentiell gedämpft, d. h., das transmittierte Feld „verflüchtigt“ sich exponentiell mit zunehmendem Abstand $|y|$ von der Grenzfläche.

Abschließend berechnen wir aus dem Dämpfungs- oder Abklingfaktor den Abstand $|y|$ für die Abschwächung von E -Feld und Intensität auf $1/e \approx 1/3$ der ursprünglichen Größe.³¹

E -Feld:

$$e^{-\beta k_2 \cdot |y|} \stackrel{!}{=} \frac{1}{e} \Rightarrow \ln(e^{-\beta k_2 \cdot |y|}) = -\beta k_2 \cdot |y| = \ln e^{-1} = -1 \Rightarrow |y| \cdot \beta k_2 = 1 \Leftrightarrow$$

$$|y_{(E/e)}| = \frac{1}{k_2 \beta} .$$

Die Intensität I einer ebenen Welle ist proportional zum Betragsquadrat des E -Feldes:

$$I \propto (A_2 \cdot e^{-\beta k_2 \cdot |y|})^2 \Rightarrow e^{-2\beta k_2 \cdot |y|} \stackrel{!}{=} \frac{1}{e} \Rightarrow$$

$$|y_{(I/e)}| = \frac{1}{2k_2 \beta} .$$

Beispiel

Wie groß ist der Abstand $|y_{(I/e)}|$ bei einer Glasplatte in Luft für grünes Licht?

$$c_0 = 3 \cdot 10^8 \text{ m s}^{-1} ,$$

$$f = 600 \text{ THz} = 6 \cdot 10^{14} \text{ s}^{-1} \text{ (grünes Licht)} \Rightarrow$$

$$\omega = 2\pi \cdot 6 \cdot 10^{14} \text{ s}^{-1} ,$$

$$n_1 = 1,5 \text{ (z. B. Glas)} ,$$

$$n_2 = 1,0 \text{ (Vakuum oder näherungsweise Luft)} ,$$

$$\alpha_1 = 75^\circ \Rightarrow \sin \alpha_1 \approx 0,966 \Rightarrow \sin^2 \alpha_1 \approx 0,93 ,$$

$$\alpha_g \approx 42^\circ \leftarrow \sin \alpha_g = n_2/n_1 = 1/1,5 = 0,6 \bar{6} .$$

$$|y_{(I/e)}| = \frac{1}{2k_2 \beta} = \frac{1}{2 \frac{n_2 \omega}{c_0} \sqrt{\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 - 1}} \approx \frac{1}{2 \frac{1 \cdot 2\pi \cdot 6 \cdot 10^{14} \text{ s}^{-1}}{3 \cdot 10^8 \text{ m s}^{-1}} \sqrt{\frac{1,5^2}{1^2} 0,93 - 1}} ,$$

$$|y_{(I/e)}| \approx 3,8 \cdot 10^{-8} \text{ m} = 38 \text{ nm} .$$

Der Abstand von der Glasplatte, bei dem die Lichtintensität bzw. die Intensität des evaneszenten Feldes näherungsweise auf 1/3 der ursprünglichen Intensität abgefallen ist, beträgt in diesem Fall ca. 38 nm. Dieser Abstand ist ungefähr eine Größenordnung kleiner als die Wellenlänge des verwendeten Lichts.

³¹ $e = 2,718 \dots$

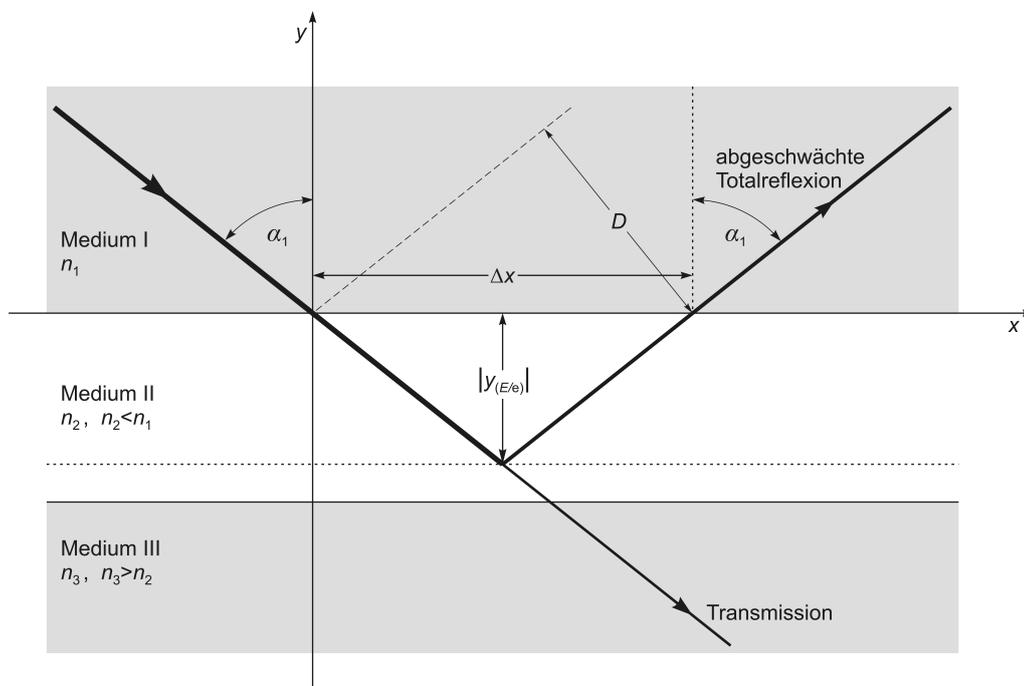


Abb. 7 Goos-Hänchen-Verschiebung Δx bei frustrierter bzw. abgeschwächter Totalreflexion. Gelegentlich wird $D = \Delta x \cos \alpha_1$ als Goos-Hänchen-Effekt bezeichnet.

Weil das Feld des einfallenden Strahls trotz Totalreflexion ein wenig in das Medium II eindringt, verlagert sich der Reflexionspunkt ebenfalls von der Grenzfläche ein wenig in das Medium II hinein. Er liegt dann konventionsgemäß etwa in der Ebene, in der die Feldstärke auf $1/e$ abgefallen ist. Der reflektierte Strahl durchdringt dann aber die Grenzfläche näherungsweise um eine Strecke Δx , die Goos-Hänchen-Verschiebung, gegenüber dem einfallenden Strahl versetzt (siehe Abbildung 7):

$$\Delta x \approx 2 \cdot |y_{(E/e)}| \cdot \tan \alpha_1 = 2 \frac{1}{k_2 \beta} \tan \alpha_1 . \quad (18.10)$$

Für k_2 schreiben wir

$$k_2 = \frac{\omega}{c_0} \cdot n_2 = \frac{2\pi f}{c_0} \cdot n_2 = \frac{2\pi n_2}{\lambda_0} \Leftrightarrow \frac{1}{k_2} = \frac{\lambda_0}{2\pi n_2} ,$$

wobei λ_0 die zur Frequenz f gehörende Vakuumwellenlänge ist. Außerdem formen wir den Term β zur Vereinfachung noch etwas um:

$$\beta = \sqrt{\left(\frac{n_1}{n_2}\right)^2 \sin^2 \alpha_1 - 1} = \sqrt{\left(\frac{n_1}{n_2}\right)^2 \left[\sin^2 \alpha_1 - \left(\frac{n_2}{n_1}\right)^2\right]} = \frac{n_1}{n_2} \sqrt{\sin^2 \alpha_1 - \left(\frac{n_2}{n_1}\right)^2} .$$

Damit ist

$$\frac{1}{k_2 \beta} = \frac{\lambda_0}{2\pi n_2} \cdot \frac{1}{\frac{n_1}{n_2} \sqrt{\sin^2 \alpha_1 - \left(\frac{n_2}{n_1}\right)^2}} = \frac{\lambda_0}{2\pi n_1 \sqrt{\sin^2 \alpha_1 - \left(\frac{n_2}{n_1}\right)^2}} .$$

Dies setzen wir in (18.10) ein und erhalten so die

Goos-Hänchen-Verschiebung $\Delta x \approx \frac{\lambda_0 \cdot \tan \alpha_1}{\pi n_1 \sqrt{\sin^2 \alpha_1 - \left(\frac{n_2}{n_1}\right)^2}} .$

Wie man sieht, ist die Goos-Hänchen-Verschiebung proportional sowohl zur Wellenlänge als auch zum Einfallswinkel bzw. Reflektionswinkel:

$$\Delta x \propto \lambda , \quad \Delta x \propto \alpha_1 .$$

19 Wellenoptik

Quellen und weiterführende Literatur:

- Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 2, Elektrizität und Optik, Springer-Verlag, Berlin, Heidelberg, New York, 3. Auflage, 2004.
- Eugene Hecht, Optik, Oldenbourg-Verlag, München, 4. Auflage, 2005.
- Rudolf Gross, WMI – Lecture Notes – Kapitel 5, Beugung und Interferenz,
https://www.wmi.badw.de/teaching/LectureNotes/Physik3/Gross_Physik_III_Kap_5.pdf
- Falko Jahn, Masterarbeit,
https://monami.hs-mittweida.de/frontdoor/deliver/index/docId/2219/file/Masterarbeit_Falko_Jahn_Bibliotheksexemplar.pdf

19.1 Einleitung

Die **geometrische Optik** (Strahlenoptik) beschäftigt sich im Wesentlichen mit der Refraktion (Brechung) und der Reflektion von Licht auf der Grundlage eines Minimalprinzips, dem **Fermat'schen Prinzip** :

Bei Reflektion bzw. Brechung laufen die Lichtstrahlen von der Quelle zum Empfänger so, dass die Lichtlaufzeit minimal ist.

In der **Wellenoptik** ist die Wellennatur des Lichtes dominierend. Dabei wird Licht als ein Feld elektromagnetischer Wellen (elektromagnetisches Wellenfeld) betrachtet. Licht setzt sich also zusammen aus einem Magnetfeld $\vec{B}(\vec{r}, t)$ und einem elektrischen Feld $\vec{E}(\vec{r}, t)$. Weil das elektrische Feld effektiver auf elektrische Ladungen wirkt als das Magnetfeld, bezeichnet man $\vec{E}(\vec{r}, t)$ auch als **optisches Feld**. Die Wellenoptik beschäftigt sich im Wesentlichen mit Diffraktion (Beugung) und Interferenz (Überlagerung, Superposition) von Licht auf der Grundlage des **Fresnel-Huygens-Prinzips**³² :

Jeder Punkt einer Wellenfront (der Primärwelle) ist Quelle sekundärer Elementarwellen mit der gleichen Frequenz wie die der Primärwelle. In jedem nachfolgenden Punkt ergibt sich die Amplitude des optischen Feldes durch die Interferenz *aller* dieser sekundären Elementarwellen.

Weil Beugung und Interferenz allgemein vergesellschaftet auftreten, ist es um so wichtiger, sie begrifflich voneinander klar zu trennen:

Interferenz tritt ein bei der Überlagerung von mindestens zwei notwendigerweise kohärenten Teilwellen (d. h. ursprünglich voneinander getrennten kohärenten Wellen).

Beugung tritt ein bei der räumlichen Begrenzung von Wellenfronten (z. B. an einer Kante oder durch eine Blende). Von den Rändern der Begrenzungen bzw. Hindernissen werden zur einlaufenden Primärwelle phasenkohärente Teilwellen (Sekundärwellen) in alle möglichen Richtungen ausgesandt. Diese Teilwellen interferieren dann miteinander.

³²Das (ursprüngliche) **Huygens'sche Prinzip** besagt, dass jeder Punkt einer primären Wellenfront Ausgangspunkt kugelförmiger sekundärer Elementarwellen ist, die in der Folge eine Wellenfront (Einhüllende) bilden. Damit allein lässt sich das frequenzabhängig-unterschiedliche Verhalten bzw. die Interferenz der Wellenfronten leider nicht ausreichend erklären.

Von einer Punktquelle Q gehe eine *Kugelwelle* aus. Diese treffe auf einen undurchsichtigen, ebenen Schirm Σ mit einer Öffnung (Aperturblende) im Abstand $\overline{Q\Sigma}$. Der durch die Öffnung hindurchtretende (auslaufende) Anteil der Kugelwelle breitet sich dann hinter der Öffnung aus und treffe dann schließlich auf die Beobachtungsebene σ . Die Ebenen Σ und σ sollen parallel zueinander verlaufen. Der Abstand zwischen der Mitte der Öffnung auf Σ und einem Beobachtungspunkt P auf σ sei $\overline{\Sigma P}$. Das Licht wird beim Passieren der Öffnung gebeugt. Abhängig von den Abständen $\overline{Q\Sigma}$ und $\overline{\Sigma P}$ können wir jetzt zwei Arten der Beugung unterscheiden. Wenn Q und P sehr weit von Σ entfernt sind, dann sind die auf die Öffnung einlaufende Welle und die von der Öffnung auslaufende Welle als nahezu eben anzusehen, sodass die Weglängenunterschiede in Richtung P lineare Funktionen von den Öffnungsvariablen sind. Wir sprechen dann von **Fraunhoferbeugung** oder Fernfeldbeugung (Linearität). Sind Q oder/und P jedoch so nahe bei Σ , dass ein- oder/und auslaufende Welle nicht mehr als eben betrachtet werden können und die Weglängenunterschiede nicht mehr linear von den Öffnungsvariablen abhängen, so spricht man von **Fresnelbeugung** oder Nahfeldbeugung (Nichtlinearität). Als Faustregel für den Fraunhofer-Grenzfall beim Übergang von der Fresnel- zur Fraunhoferbeugung gilt

$$\text{Fraunhoferbeugung} \Leftrightarrow R > \frac{a^2}{\lambda}.$$

Dabei ist R der kleinere der beiden Abstände $\overline{Q\Sigma}$ und $\overline{\Sigma P}$, a ist die größte Ausdehnung der Öffnung und λ die Wellenlänge des Lichts. Die Fraunhoferbeugung kann man als den „einfach“ herzuleitenden Spezialfall der Fresnelbeugung betrachten. In der Abbildung 8 wird dargestellt, wie sich die Fraunhoferbeugung mit Hilfe von Sammellinsen platzsparend realisieren lässt.

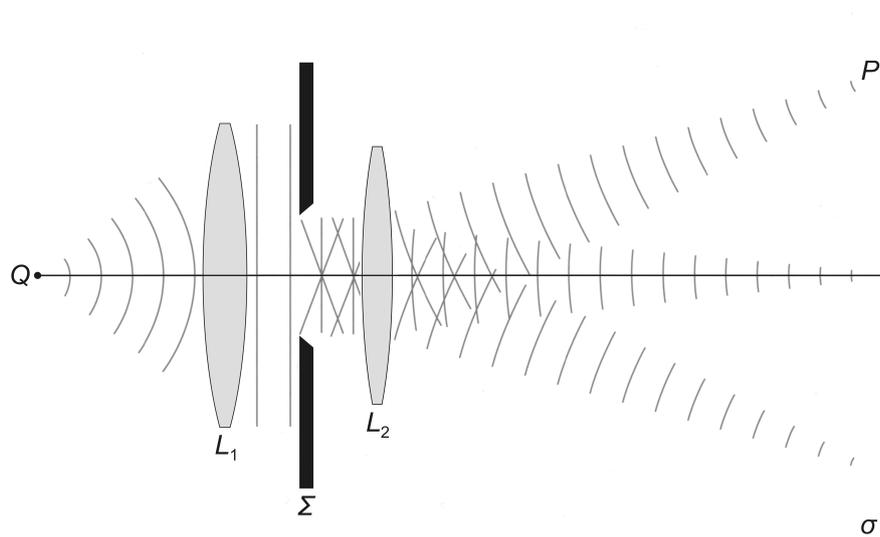


Abb. 8 Fraunhoferbeugung. Die Punktquelle Q liegt im vorderen Brennpunkt der Sammellinse L_1 , die deshalb die von Q ausgehenden Kugelwellen in ebene Wellen transformiert (Kollimator). Die Abbildungsebene σ liegt in der hinteren Brennebene der Sammellinse L_2 , sodass die einfallenden ebenen Wellenzüge auf σ konvergieren, in der Abbildung beispielsweise in P für einen bestimmten Wellenzug. Deshalb entsteht das Interferenzmuster, beispielsweise ein Streifenmuster bei der Beugung am Einfachspalt, in praktikabler Entfernung auf σ und nicht erst nahezu im Unendlichen. (Abbildung nach E. Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag, München, Wien, 2005, Seite 723)

19.2 Fraunhoferbeugung

19.2.1 Elementare Herleitung im Reellen

Kurzdarstellung in Anlehnung an

Gerthsen, Physik, Springer-Verlag, 20. Auflage, 1999, Seite 517 bis Seite 522 und

Paul A. Tipler, Physik, Spektrum-Verlag, 1. deutsche Auflage, 2000, Seite 1120 bis Seite 1131.

Wir betrachten in diesem Abschnitt der Einfachheit halber nur den Fall, dass die Fronten einer ebenen Lichtwelle komplanar auf einen ebenen Schirm treffen, der lichtdurchlässige, schmale, gerade, parallel verlaufende Spalten besitzt (Einfachspalt, Doppelspalt, Strichgitter). Somit haben alle Elementarwellen in der Ebene der Spalten die gleiche Phase Φ .

a) Vorläufige Amplitudenformel

Auf der Grundlage von Abbildung 9 nehmen wir an, dass die Breite a der Spalten so klein ist, dass diese Spalten Reihen von kohärenten **Punktquellen** mit dem Reihenabstand g darstellen. So erhalten wir ein Strichgitter mit extrem schmalen Spalten (Strichen).

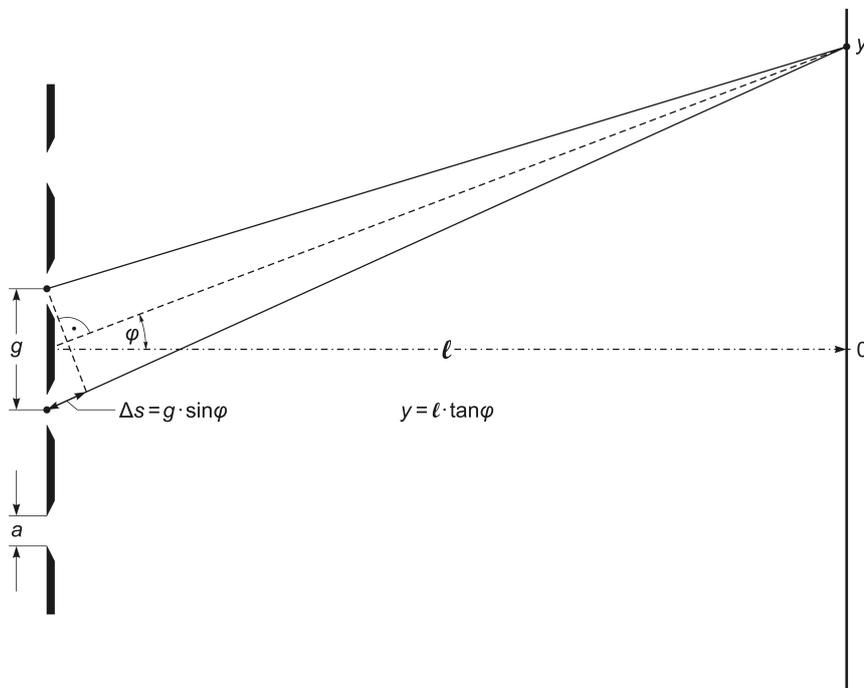


Abb. 9 Strichgitter und Schirm in möglichst großem Abstand ℓ zur Darstellung des Interferenzmusters. Gitterkonstante g , Strichzahl N , Spaltbreite a , Spalt-Schirm-Abstand ℓ mit $\ell \gg g$, Winkel φ zwischen der Wellenfrontnormalen und der Spaltnormalen, Gangunterschied Δs .

Im Fall $\ell \rightarrow \infty$ verlaufen die von den beiden dargestellten Punktquellen ausgehenden ebenen Wellenzüge parallel zueinander und in Richtung zum Punkt y des Interferenzmusters auf dem Schirm.

Die Phasendifferenz δ zwischen den Wellenzügen benachbarter Spalten ergibt sich im Bogenmaß [rad] aus dem Laufwegunterschied oder kurz **Gangunterschied** Δs wie folgt:

$$\frac{\delta}{2\pi} = \frac{\Delta s}{\lambda} = \frac{g \cdot \sin \varphi}{\lambda} \Leftrightarrow \text{Phasendifferenz } \delta = \frac{2\pi}{\lambda} \cdot g \cdot \sin \varphi = k \cdot \Delta s. \quad (19.1)$$

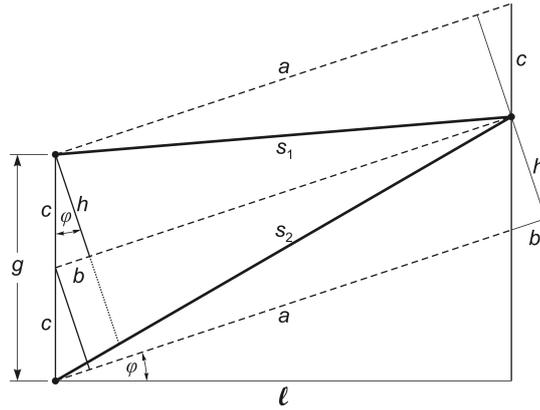


Abb. 10 Der korrekte Gangunterschied $\Delta s = s_2 - s_1$.

Anhand der Abbildung 10 zeigen wir, dass der in Abbildung 9 dargestellte Gangunterschied

$$\Delta s = g \cdot \sin \varphi \quad \text{für } \ell \gg g$$

eine gute Näherung ist und an den Grenzen $\varphi = 0$ und $\varphi = \frac{\pi}{2}$ mit $g \sin \varphi$ übereinstimmt:

$$c = \frac{g}{2}, \quad a = \frac{\ell}{\cos \varphi}, \quad b = c \sin \varphi, \quad h = c \cos \varphi,$$

$$\begin{aligned} \Delta s &= s_2 - s_1 = \sqrt{(a+b)^2 + h^2} - \sqrt{(a-b)^2 + h^2} \\ &= \sqrt{\left(\frac{\ell}{\cos \varphi} + c \sin \varphi\right)^2 + c^2 \cos^2 \varphi} - \sqrt{\left(\frac{\ell}{\cos \varphi} - c \sin \varphi\right)^2 + c^2 \cos^2 \varphi} \Rightarrow \end{aligned}$$

$$\lim_{\varphi \rightarrow 0} \Delta s = \sqrt{\ell^2 + c^2} - \sqrt{\ell^2 + c^2} = 0 = g \sin(0), \quad \square$$

$$\begin{aligned} \lim_{\varphi \rightarrow \frac{\pi}{2}} \Delta s &= \lim_{\varphi \rightarrow \frac{\pi}{2}} \left[\sqrt{\left(\frac{\ell}{\cos \varphi} + c\right)^2} - \sqrt{\left(\frac{\ell}{\cos \varphi} - c\right)^2} \right] \\ &= \lim_{\varphi \rightarrow \frac{\pi}{2}} \left[\left(\frac{\ell}{\cos \varphi} + c\right) - \left(\frac{\ell}{\cos \varphi} - c\right) \right] = 2c = g = g \sin\left(\frac{\pi}{2}\right). \quad \square \end{aligned}$$

Hauptmaxima bei konstruktiver Interferenz:

$$\Delta s = g \cdot \sin \varphi \stackrel{!}{=} m \cdot \lambda,$$

$$\delta = \frac{2\pi g}{\lambda} \sin \varphi \stackrel{!}{=} m \cdot 2\pi, \quad m = 0, 1, 2, 3, \dots \quad (19.2)$$

$$\Rightarrow \sin \varphi = m \cdot \frac{\lambda}{g}, \quad \sin \varphi \sim \lambda, \quad \sin \varphi \sim \frac{1}{g}.$$

Für konstruktive Interferenz gilt: $\left. \begin{array}{l} \text{Je größer } \lambda, \\ \text{Je kleiner } g, \end{array} \right\} \text{desto größer ist } \varphi.$

Minima bei destruktiver Interferenz:

$$\Delta s = g \cdot \sin \varphi \stackrel{!}{=} \left(m + \frac{1}{2}\right) \lambda,$$

$$\delta = \frac{2\pi g}{\lambda} \sin \varphi \stackrel{!}{=} (2m + 1) \pi, \quad m = 0, 1, 2, 3, \dots$$

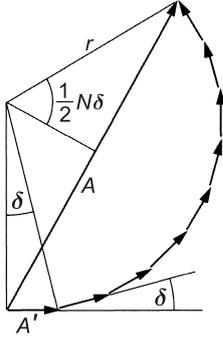


Abb. 11 Überlagerung vieler gleichstarker Wellen mit der Phasenverschiebung δ zwischen je zweien, z. B. der Teilwellen aus den Spalten eines Beugungsgitters. Darstellung dieser Überlagerung durch ein Zeigerdiagramm.

Vervollständigt man das Diagramm durch Aneinanderfügen weiterer Zeiger (Vektoren), so erhält man einen kreisförmigen Polygonzug mit dem Radius r .

(Abbildung und Legende nach Gerthsen, Physik, Springer-Verlag, 20. Auflage, 1999, Seite 518)

Von jedem der N Gitterspalten (Striche) geht eine Reihe von Kugelwellen aus, d. h., jeder Spalt liefert eine Welle mit der Amplitude A' . Bei $\varphi = 0 \Rightarrow \delta = 0$ gilt infolge konstruktiver Interferenz aller A'

$$A_{(\varphi=0)} = A_0 = N \cdot A' = A_{max} .$$

Bei $\varphi \neq 0$ ist wegen der Phasendifferenz $\delta = \frac{2\pi g}{\lambda} \sin \varphi$ die Phasendifferenz $\Delta\Phi$ zwischen dem ersten und dem letzten Spalt des Gitters $\Delta\Phi = (N - 1) \cdot \delta$ und für eine große Gitterstrichzahl N näherungsweise

$$\Delta\Phi = N \cdot \delta = N \cdot \frac{2\pi g}{\lambda} \sin \varphi .$$

Entsprechend dem Zeigerdiagramm erhält man daraus die von φ abhängige Amplitude A_φ (Betrag der Vektorsumme bzw. Länge der Sehne des Polygonzuges für einen bestimmten Winkel φ) wie folgt:

$$\sin\left(\frac{1}{2} \Delta\Phi\right) = \sin\left(\frac{1}{2} N\delta\right) = \frac{\frac{A_\varphi}{2}}{r} = \frac{A_\varphi}{2r} \Rightarrow A_\varphi = 2r \cdot \sin \frac{N\delta}{2} \Leftrightarrow r = \frac{A_\varphi}{2 \sin \frac{N\delta}{2}} .$$

Für $N = 1$ erhält man $r = \frac{A'}{2 \sin \frac{\delta}{2}}$. Dieses in $A_\varphi = 2r \cdot \sin \frac{N\delta}{2}$ eingesetzt ergibt

$$A_\varphi = 2 \left(\frac{A'}{2 \sin \frac{\delta}{2}} \right) \sin \frac{N\delta}{2} = A' \frac{\sin \frac{N\delta}{2}}{\sin \frac{\delta}{2}} .$$

Substitution von δ durch $\frac{2\pi g}{\lambda} \sin \varphi$ liefert schließlich die **vorläufige Amplitudenformel** für die Interferenz am Strichgitter:

$$\boxed{A_\varphi = A' \frac{\sin\left(\frac{N\pi g}{\lambda} \sin \varphi\right)}{\sin\left(\frac{\pi g}{\lambda} \sin \varphi\right)}} . \quad (19.3)$$

- Je kleiner g ist, desto breiter wird das Spektrum gespreizt, denn

$$\Delta s = g \cdot \sin \varphi \Rightarrow \sin \varphi = \Delta s \cdot \frac{1}{g} \Rightarrow \sin \varphi \propto \frac{1}{g} . \quad (19.4)$$

- Aus $\sin \varphi \propto \lambda$ folgt, dass Licht mit großer Wellenlänge (z. B. rotes Licht) stärker gebeugt wird als Licht mit kleiner Wellenlänge (z. B. blaues Licht).
- Ein Gitter macht um so schärfere Spektrallinien, je mehr Striche es hat.
- Zwischen 2 Hauptmaxima liegen stets $N - 1$ Minima und $N - 2$ Nebenmaxima.

b) Amplitudenformel für die Beugung und Interferenz am Einfachspalt

Die Herleitung der Amplitudenformel für die Beugung und Interferenz am Einfachspalt erfolgt auf der Basis der vorläufigen Amplitudenformel (19.3). Diese wurde mit Hilfe des Konzepts der Vektoraddition bzw. des Zeigerdiagramms entwickelt, wobei die Spaltbreite des Strichgitters als extrem klein und somit der einzelne Gitterspalt als eine Reihe von Punktquellen angenommen wurde. A' war dabei die aus nur *einer* dieser Reihen von Punktquellen (aus nur *einem* Strich) resultierende Amplitude.

Den Einfachspalt erhält man jetzt, indem beim Strichgitter solange die Gitterkonstante g verkleinert und gleichzeitig die Strichzahl N erhöht wird, bis schließlich ein Einfachspalt mit der endlichen Spaltbreite a resultiert. Wir können also sinngemäß oder besser gesagt symbolisch schreiben:

$$\lim_{\substack{g \rightarrow 0 \\ N \rightarrow \infty}} \sum \text{Gitterspalten} \approx N \cdot g = a \quad \Leftrightarrow \quad g = \frac{a}{N}.$$

Aus dem Strichgitter ist jetzt ein Einfachspalt geworden, der gleichsam aus nahezu unendlich vielen, nahezu unendlich schmalen Gitterstrichen besteht, die insgesamt die Einfachspaltbreite a ergeben. Damit wird aber auch die Phasendifferenz $\delta = \frac{2\pi g}{\lambda} \sin \varphi$ zwischen den Wellenzügen aus benachbarten Gitterstrichen verschwindend klein. Für kleine δ können wir aber

$$\sin \frac{\delta}{2} = \frac{\delta}{2}$$

schreiben und erhalten so mit (19.1)

$$\sin \frac{\delta}{2} = \sin \left(\frac{\pi g}{\lambda} \sin \varphi \right) = \frac{\pi g}{\lambda} \sin \varphi. \quad (19.5)$$

Berücksichtigen wir (19.5) in der vorläufigen Amplitudenformel (19.3) und setzen wir für g dort außerdem $\frac{a}{N}$ ein, so resultiert

$$A_\varphi = A' \frac{\sin \left(\frac{N\pi \cdot \frac{a}{N}}{\lambda} \sin \varphi \right)}{\frac{\pi \cdot \frac{a}{N}}{\lambda} \sin \varphi} = A' \frac{\sin \left(\frac{\pi a}{\lambda} \sin \varphi \right)}{\frac{\pi a}{N \cdot \lambda} \sin \varphi} = N A' \frac{\sin \left(\frac{\pi a}{\lambda} \sin \varphi \right)}{\frac{\pi a}{\lambda} \sin \varphi}$$

und daraus mit $N A' = A_{max} = A_0$ schließlich die **Amplitudenformel** für die Beugung und Interferenz am **Einfachspalt**:

$$\boxed{A_\varphi = A_0 \frac{\sin \left(\frac{\pi a}{\lambda} \sin \varphi \right)}{\frac{\pi a}{\lambda} \sin \varphi}}. \quad (19.6)$$

A_φ ist der Betrag der Vektorsumme, also die Länge der Sehne des Polygonzuges. A_0 ist die resultierende Amplitude für die Ausbreitungsrichtung mit $\varphi = 0$, also das Maximum von A_φ , denn nach der Regel von *de l'Hospital* ist $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$.

- Durch das Anwachsen des Sinus im Nenner nehmen die Maxima bei der Beugung am Einfachspalt zur Seite hin sehr schnell an Höhe ab.
- Mit (19.4) und wegen $g \rightarrow 0$ sowie $N \rightarrow \infty$ existiert nur *ein* zentrales (Haupt-) Maximum. Alle anderen Hauptmaxima wandern unendlich weit seitwärts, sodass nur unendlich viele Nebenmaxima übrigbleiben.

- Die Phasendifferenz (zwischen der Phase an dem einen Spaltenrand und der Phase am anderen Spaltenrand) für einen bestimmten Winkel φ beim Einfachspalt ist

$$\Delta\Phi = \frac{2\pi a}{\lambda} \sin \varphi = \frac{2\pi \sin \varphi}{\lambda} a \quad (19.7)$$

Es gilt also $\Delta\Phi \sim a$, d. h., dass eine bestimmte Phasendifferenz $\Delta\Phi$ mit abnehmender Spaltbreite a erst bei einem größeren Winkel φ erreicht wird. Das Beugungsspektrum wird also bei abnehmender Spaltbreite gespreizt und das zentrale Maximum wird dabei gleichzeitig breiter und niedriger.

c) Amplitudenformel für die Beugung und Interferenz am Strichgitter

Die endgültige Formel für die Beugung und Interferenz am Strichgitter erhält man, wenn in (19.3) für die Amplitude A' des einzelnen Gitterspalts die Amplitudenformel (19.6) des Einfachspalts einsetzt, denn

$$A' = A_0 \frac{\sin\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\frac{\pi a}{\lambda} \sin \varphi}.$$

Dadurch wird in der Amplitude A_φ jeder einzelne Gitterspalt als Einzelspalt mit der endlichen Spaltbreite a berücksichtigt. Insgesamt ist dann die endgültige **Amplitudenformel** für das **Strichgitter**

$$A_\varphi = A_0 \frac{\sin\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\frac{\pi a}{\lambda} \sin \varphi} \cdot \frac{\sin\left(\frac{N\pi g}{\lambda} \sin \varphi\right)}{\sin\left(\frac{\pi g}{\lambda} \sin \varphi\right)}. \quad (19.8)$$

Wie man sieht, resultiert für $N = 1$ die Amplitudenformel für den Einfachspalt, für $N = 2$ die Amplitudenformel für den Doppelspalt usw.

d) Intensitätsformel für die Beugung und Interferenz am Strichgitter

Die Intensität einer ebenen Lichtwelle ist proportional zum Quadrat ihrer (Feldstärken-) Amplitude:

$$A_\varphi^2 \propto I_\varphi, \quad A_0^2 = N^2 \cdot A'^2 \propto I_0 \quad \Rightarrow \quad I_0 \propto N^2.$$

Quadrieren von (19.8) liefert folglich die Intensitätsformel für die Beugung und Interferenz am **Strichgitter**:

$$I_\varphi = I_0 \frac{\sin^2\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\left(\frac{\pi a}{\lambda} \sin \varphi\right)^2} \cdot \frac{\sin^2\left(\frac{N\pi g}{\lambda} \sin \varphi\right)}{\sin^2\left(\frac{\pi g}{\lambda} \sin \varphi\right)}. \quad (19.9)$$

Bei der Anwendung dieser Intensitätsformel kann man den Doppelspalt als ein Strichgitter mit $N = 2$ und den Einfachspalt als ein Strichgitter mit $N = 1$ betrachten.

e) Intensitätsformel für die Beugung und Interferenz am Einfachspalt

Somit ist die Intensitätsformel für die Beugung und Interferenz am **Einfachspalt**:

$$I_{\varphi} = I_0 \frac{\sin^2 \left(\frac{\pi a}{\lambda} \sin \varphi \right)}{\left(\frac{\pi a}{\lambda} \sin \varphi \right)^2} .$$

Das zentrale Maximum wird beim Einfachspalt auch Hauptmaximum genannt. Daneben existieren nur noch Nebenmaxima der Ordnung $m = 1, 2, 3, \dots$, deren Intensität mit ansteigender Ordnung stark abnimmt.

Der Abstand zwischen dem ersten Intensitätsminimum links und dem ersten Intensitätsminimum rechts vom zentralen Maximum kann als Breite des zentralen Maximums angesehen werden (siehe Abbildung 12).

Dem Zeigerdiagramm entsprechend liegen der erste Nulldurchgang der Amplitude links und der erste Nulldurchgang der Amplitude rechts vom zentralen Maximum bei demjenigen Winkel φ , der im Zeigerdiagramm einen einfachen geschlossenen Vektorkreis liefert, d. h. gemäß (19.7) bei

$$\Delta\Phi = 2\pi \cdot \frac{a \cdot \sin \varphi}{\lambda} \stackrel{!}{=} 2\pi .$$

Bildet das Zeigerdiagramm einen Vollkreis, so ist die Vektorsumme nämlich gleich null, d. h. die Teilamplituden löschen sich in ihrer Gesamtheit aus (destruktive Interferenz).³³ Der zum ersten Nulldurchgang der Amplitude gehörende Winkel φ ist somit

$$\Leftrightarrow \sin \varphi = \frac{\lambda}{a} \quad \Rightarrow \quad \varphi = \arcsin \frac{\lambda}{a} .$$

Damit beträgt die Breite des zentralen Maximums auf dem Schirm (s. Abb. 12)

$$2 \cdot \ell \cdot \tan \left(\arcsin \frac{\lambda}{a} \right) .$$

³³Der erste Nulldurchgang der Amplitude zu beiden Seiten des zentralen Maximums bildet die ersten Intensitätsminima und bestimmt die Breite des zentralen Maximums.

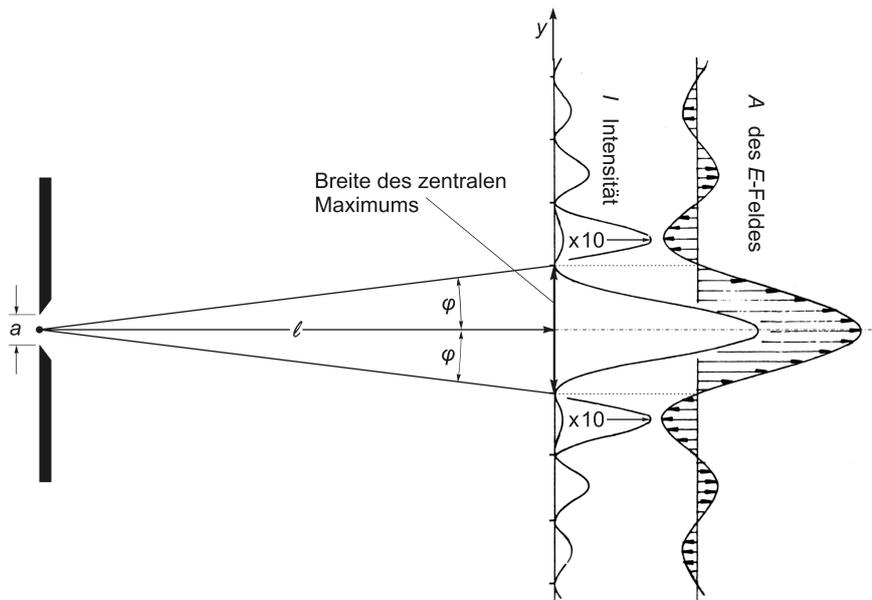


Abb. 12 Die Breite des zentralen Maximums bei der Beugung am Einfachspalt. Weil die Intensität der Nebenmaxima sehr stark abfällt, sind diese zur Verdeutlichung um das 10-fache überhöht. Der Vollständigkeit halber wird der Amplitudenverlauf des E -Feldes längs der y -Achse der Intensität gegenübergestellt. (Abbildung nach dem Springer-Lehrbuch Langkau, Scobel, Lindström, Physik kompakt 2, Elektrodynamik und Elektromagnetische Wellen, Springer-Verlag, Berlin, Heidelberg, New York, 2. Auflage, 2002, Seite 250)

Minima und Maxima beim Einfachspalt:

Hauptmaximum bei $\varphi = 0$.

Nebenmaxima bei $\sin \varphi_m = \left(m + \frac{1}{2}\right) \frac{\lambda}{a}$, $m = 1, 2, 3, \dots$

Minima bei $\sin \varphi_m = m \frac{\lambda}{a}$, $m = 1, 2, 3, \dots$

Minima und Maxima beim Strichgitter:

Maxima bei $\sin \varphi_m = m \frac{\lambda}{g}$, $m = 0, 1, 2, 3, \dots$

Minima bei $\sin \varphi_m = m \frac{\lambda}{N \cdot g}$, $m = 1, 2, 3, \dots$

f) Spektrales Auflösungsvermögen des Strichgitters

Die Herleitung erfolgt nach dem Lehrbuch von Gerthsen, Seite 522, und auf der Grundlage von Formel (19.9) sowie speziell auf der Grundlage der vorläufigen Amplitudenformel (19.3)

$$A_\varphi = A' \cdot \frac{\sin\left(\frac{N\pi g}{\lambda} \sin \varphi\right)}{\sin\left(\frac{\pi g}{\lambda} \sin \varphi\right)},$$

denn insbesondere das breite zentrale Maximum von A' überlagert den Quotienten in (19.3) nur und ist für das Auflösungsvermögen des Strichgitters nicht relevant.

Der Quotient auf der rechten Seite von (19.3) ist verantwortlich für die Anordnung der Hauptmaxima und Minima im Gitterspektrum:

- Die Maxima liegen dort, wo sowohl Zähler also auch Nenner verschwinden bzw. gegen Null gehen:

$$\delta = \frac{2\pi g}{\lambda} \cdot \sin \varphi_m \stackrel{!}{=} m \cdot 2\pi \quad \Leftrightarrow \quad \sin \varphi_m = \frac{m\lambda}{g}, \quad m = 0, 1, 2, 3, \dots$$

$$\text{Nenner: } \sin\left(\frac{\pi g}{\lambda} \sin \varphi_m\right) = \sin(m\pi) = 0$$

$$\text{Zähler: } \sin(N \cdot m\pi) = 0$$

$$\text{Maximum m-ter Ordnung: } \varphi_m \approx \frac{m\lambda}{g}. \quad (19.10)$$

- Die Minima liegen dort, wo nur der Zähler verschwindet.

$$\text{Zähler: } \sin\left(\frac{\pi N g}{\lambda} \sin \varphi\right) \stackrel{!}{=} 0 \quad \text{bei} \quad \sin \varphi \neq 0 \quad \text{bzw.} \quad \varphi \neq 0, \quad \text{d. h. wenn}$$

$$N \cdot \Delta s = N \cdot g \cdot \sin \varphi_m = \lambda \quad \Rightarrow \quad \sin \varphi_m = m \cdot \frac{\lambda}{N \cdot g}, \quad m = 1, 2, 3, \dots$$

Beim ersten Minimum bilden die Vektoren der N Gitterquellen einen einfachen geschlossenen Kreis und löschen sich somit aus. Es gilt dann

$$N \cdot \delta = N \cdot \frac{2\pi g}{\lambda} \sin \varphi = 2\pi.$$

$$\Delta(\sin \varphi) = (m+1) \frac{\lambda}{N \cdot g} - m \frac{\lambda}{N \cdot g} = \frac{\lambda}{N \cdot g} \quad \text{liefert somit den **Abstandswinkel**}$$

$$\Delta \varphi \approx \frac{\lambda}{N \cdot g}$$

des ersten Minimums von einem Hauptmaximum.

- Die Maxima von Licht mit der Wellenlänge $\lambda + \Delta\lambda$ besitzen gemäß (19.10) die Winkel

$$\varphi_m \approx \frac{m(\lambda + \Delta\lambda)}{g} = \frac{m\lambda}{g} + \frac{m \cdot \Delta\lambda}{g}$$

und liegen damit gegenüber Licht mit der Wellenlänge λ um den Winkel $\frac{m \cdot \Delta\lambda}{g}$ versetzt.

- Die Maxima (Spektrallinien) des Lichts mit λ und des Lichtes mit $(\lambda + \Delta\lambda)$ lassen sich trennen, wenn das Maximum (der Ordnung m) des Lichts mit $(\lambda + \Delta\lambda)$ in das erste Minimum des Lichts mit λ fällt oder sogar vom Maximum (der Ordnung m) des Lichts mit λ noch weiter entfernt ist, also wenn gilt

$$\frac{m \cdot \Delta\lambda}{g} \geq \Delta\varphi \approx \frac{\lambda}{N \cdot g} \quad \Leftrightarrow$$

spektrales Auflösungsvermögen $m \cdot N \geq \frac{\lambda}{\Delta\lambda}$.

- Das Auflösungsvermögen $\frac{\lambda}{\Delta\lambda}$ eines Gitters wächst proportional mit der Anzahl der Striche (Strichzahl N) des Gitters und der Ordnung m des verwendeten Maximums. Das Auflösungsvermögen eines Gitters hängt nicht von der Gitterkonstante g ab, auch wenn gilt:
Je kleiner g , desto weiter wird das Spektrum gespreizt.
- $\frac{\lambda}{\Delta\lambda}$ darf nicht größer als das Gitterauflösungsvermögen $m \cdot N$ werden, damit das Spektrum des Lichtes mit der Wellenlänge $\lambda + \Delta\lambda$ noch aufgelöst werden kann. Je größer also das Auflösungsvermögen $m \cdot N$ ist, desto kleiner darf $\Delta\lambda$ sein, d. h., desto besser lassen sich kleine Frequenzunterschiede des Lichtes spektral auflösen.

19.3 Fresnel'sche Zonen

Quellen:

Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 2, Elektrizität und Optik, Springer-Verlag, Berlin, Heidelberg, New York, 3. Auflage, 2004, Abschnitt 10.6.1 *Fresnelsche Zonen*, Seite 322 bis Seite 325,

Eugene Hecht, Optik, Oldenbourg-Verlag, München, 4. Auflage, 2005, Abschnitt 10.3.1 *Die freie Ausbreitung einer Kugelwelle*, Seite 779 bis Seite 785.

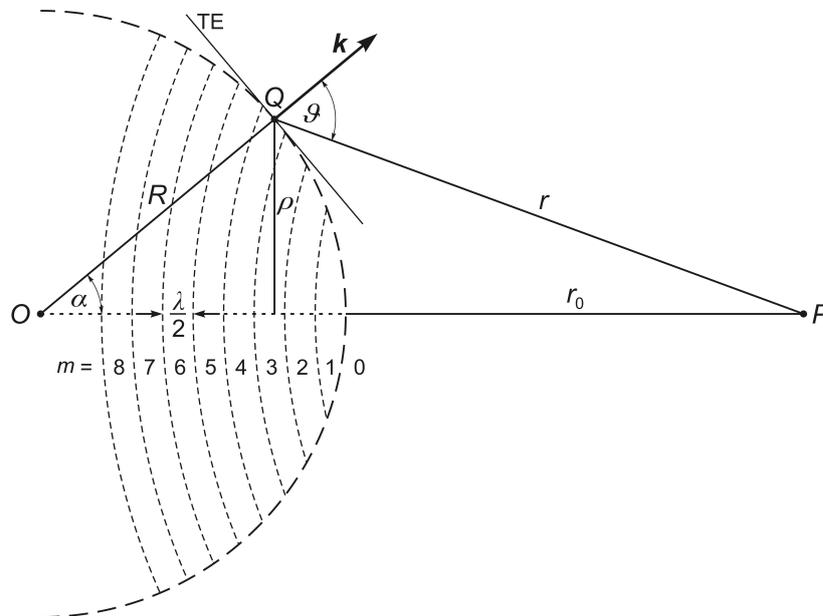


Abb. 13 Fresnel'sche Zonen. Dargestellt sind nur die ersten 8 Zonen. TE steht für Tangentialebene an den Punkt Q der dargestellten Kugelfront.

- Die von der Punktquelle O ausgehende und sich frei im Raum ausbreitende Kugelwelle, die **Primärwelle**, sei

$$E = E_R \cdot e^{i(kR - \omega t)} = \frac{A}{R} \cdot e^{i(kR - \omega t)}. \quad (19.11)$$

Die Amplitude $E_R = \frac{A}{R}$ ist also umgekehrt proportional zum Radius R der betrachteten Kugelwelle. Wie man sieht, kann A nicht die Feldstärkenamplitude sein, weil A nicht die Maßeinheit (Dimension) von E besitzt. Weiterhin besitzt die Funktion $E_R(R)$ bei $R = 0$ eine Singularität.

- Nach dem **Huygens'schen Prinzip** sind alle Punkte Q der Kugelfläche mit Radius R um O Ausgangspunkte neuer Kugelwellen, den **Sekundärwellen**.
- Mit diesem Ansatz berechnen wir das E -Feld in einem Punkt P , der längs \overline{OP} den Abstand r_0 von der betrachteten Kugelfläche besitzt (siehe Abbildung 13).
- Die Abstände zwischen den Punkten Q auf der Kugelfläche und dem Punkt P bezeichnen wir mit r .
- Die Strecke \overline{OP} besitzt die Länge

$$R + r_0 \Rightarrow r(\alpha = 0) = r_0.$$

- Der Kreis, der von einem bestimmten $r = r(\alpha)$ auf der Kugelfläche gebildet wird, besitzt den Radius

$$\varrho = R \cdot \sin \alpha .$$

- Die **Fresnel'sche Zonen** bzw. Halbperiodenzonen auf der Kugeloberfläche mit dem Radius R erhalten wir – abhängig von r – wie folgt:

$$\left. \begin{array}{l} r_0 = r_0 + 0 \cdot \frac{\lambda}{2} = r_0 \\ r_1 = r_0 + 1 \cdot \frac{\lambda}{2} \\ r_2 = r_0 + 2 \cdot \frac{\lambda}{2} \\ r_3 = r_0 + 3 \cdot \frac{\lambda}{2} \\ \vdots \end{array} \right\} \Rightarrow r_m = r_0 + m \cdot \frac{\lambda}{2}, \quad m = 1, 2, 3, \dots .$$

Die 1. Zone erstreckt sich demnach von r_0 bis $r_1 = r_0 + \frac{\lambda}{2}$ und die

$$m\text{-te Fresnel'sche Zone} \left\{ \begin{array}{l} \text{von } r_{m-1} = r_0 + (m-1) \cdot \frac{\lambda}{2} \\ \text{bis } r_m = r_0 + m \cdot \frac{\lambda}{2} . \end{array} \right.$$

- Die Anwendung des Kosinussatzes

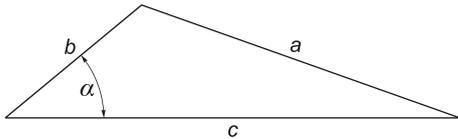


Abb. 14 Kosinussatz: $a^2 = b^2 + c^2 - 2bc \cos \alpha$.

auf das Dreieck OQP liefert

$$\begin{aligned} r^2 &= R^2 + (R + r_0)^2 - 2R(R + r_0) \cos \alpha \\ \Rightarrow \frac{d}{d\alpha} r^2 &= 2r \frac{dr}{d\alpha} = 2R(R + r_0) \sin \alpha \quad \Leftrightarrow \\ R \sin \alpha d\alpha &= \frac{1}{R + r_0} r dr . \end{aligned} \quad (19.12)$$

- Wir betrachten ein ring- bzw. kreisförmiges und konzentrisch um \overline{OP} gelegenes Flächenelement dS innerhalb einer Fresnel'schen Zone. Alle Punktquellen in dS emittieren kohärent und phasengleich mit der zugehörigen Primärwelle. Die dabei entstehenden sekundären Elementarwellen (Sekundärwellen) kommen demzufolge nach Zurücklegen des Weges r zum Zeitpunkt t mit der Phase $[k \cdot (R + r) - \omega t]$ phasengleich im Punkt P an. Gemäß (19.11) ist $E_R = A/R$ die Amplitude der Primärwelle auf dS .

E' sei der *Feldstärkebeitrag* von E_R pro Flächeneinheit auf dS , also sinngemäß dE_R/dS , für die *Sekundärwellen*. Offensichtlich ist der Beitrag von E_R zur Feldstärke der Sekundärwellen in P proportional zu E' mit der Proportionalitätskonstante q :

$$E' \propto E_R \quad \Rightarrow \quad E' = q \cdot E_R . \quad (19.13)$$

- Weiterhin ist der Feldstärkebeitrag durch die Sekundärwellen in P abhängig vom Winkel ϑ (siehe Abbildung 13), d. h. abhängig von der „Gerichtetheit“ der sekundären Quellen auf der primären Kugelwellenfront. Diese „Gerichtetheit“ wird in diesem speziellen Zusammenhang beschrieben durch den³⁴

$$\text{Neigungsfaktor } K = K(\vartheta) = \frac{1}{2}(1 + \cos \vartheta) .$$

Für den Feldstärkebeitrag der Sekundärwellen im Punkt P gilt

$$E(P) \propto K .$$

Für unsere Kugelwellenfront mit $\vartheta_0 = 0 \Rightarrow \cos \vartheta_0 = 1$ bedeutet das (siehe Abbildung 13)

$$K(\vartheta) = \frac{1}{2}(1 + \cos \vartheta) \propto E(P) . \quad (19.14)$$

Und bei $\vartheta_0 = 0$ und $\vartheta = 0$ ist dann $K = 1$ und Q liegt innerhalb der Strecke \overline{OP} . Weiterhin bedeutet das, dass antiparallel zu \vec{k} laufende Sekundärwellen (bei entsprechend rückwärtiger Lage von Q) wegen

$$\vartheta = \pi \Rightarrow \cos \vartheta = -1 \Rightarrow K(\vartheta) = \frac{1}{2}(1 - 1) = 0$$

keinen Feldstärkebeitrag in P leisten. O liegt dann innerhalb der Strecke \overline{QP} .

- Mit (19.13) und (19.14) machen wir jetzt den folgenden Ansatz:

Das um \overline{OP} ring- bzw. kreisförmige Flächenelement dS mit dem zugehörigen Abstand r zum Punkt P liefert zur Feldstärke in P den Beitrag

$$dE(P) = \tilde{C} \cdot K \cdot \frac{E'}{r} e^{i[k(R+r) - \omega t]} dS = \tilde{C} \cdot K \cdot \frac{q \cdot E_R}{r} e^{i[k(R+r) - \omega t]} dS$$

und mit der Zusammenfassung der Konstanten \tilde{C} und q zu $\tilde{C} \cdot q = C$ schließlich

$$dE(P) = C \cdot K \cdot \frac{E_R}{r} e^{i[k(R+r) - \omega t]} dS . \quad (19.15)$$

Der Neigungsfaktor $K = \frac{1}{2}(1 + \cos \vartheta)$ bezüglich der primären Kugelwellenfront beschreibt die Abhängigkeit der von dS in Richtung P abgestrahlten Amplitude vom Winkel ϑ . Da $K(\vartheta)$ eine nur langsam veränderliche Funktion ist, nehmen wir K über jede einzelne Fresnel'sche Zone als konstant an und schreiben K_m für den Neigungsfaktor der m -ten Fresnel'schen Zone.

- Das Flächenelement dS besitzt den Kreisumfang $2\pi \cdot \varrho = 2\pi \cdot R \sin \alpha$, die Breite $R \cdot d\alpha$ und folglich die Fläche

$$dS = 2\pi R \sin \alpha \cdot R d\alpha = 2\pi R \cdot R \sin \alpha d\alpha .$$

³⁴Auf den ersten Blick würde man meinen, dass der Feldstärkebeitrag in P durch die Sekundärwellen proportional zu einer Richtungsfunktion $f(\vartheta) = \cos \vartheta$ sei, wovon auch Fresnel zunächst in seiner Theorie ausging. Warum das nicht so ist, wird in der skalaren Beugungstheorie von Kirchhoff deutlich. Auf den Neigungsfaktor in allgemeiner Form gehen wir im Abschnitt ?? gesondert ein.

Aus dem Vergleich mit (19.12) resultiert

$$dS = \frac{2\pi R}{R + r_0} r dr .$$

Somit ist dS nur noch von r abhängig.

- Durch Integration über (19.15) können wir jetzt den Beitrag E_m der m -ten Fresnel'schen Zone zum E -Feld im Punkt P berechnen:

$$E_m = \int_{r_{m-1}}^{r_m} dE_m = \int_{r_{m-1}}^{r_m} C \cdot K_m \cdot \frac{E_R}{r} e^{i[k(R+r) - \omega t]} \cdot \frac{2\pi R}{R + r_0} r dr ,$$

$$E_m = C \cdot K_m E_R \frac{2\pi R}{R + r_0} e^{i(kR - \omega t)} \cdot \int_{r_{m-1}}^{r_m} e^{ikr} dr ,$$

$$E_m = C \cdot K_m E_R \frac{2\pi R}{R + r_0} e^{i(kR - \omega t)} \cdot \frac{1}{ik} \cdot \left[e^{ikr} \right]_{r_{m-1}}^{r_m} ,$$

Zwischenrechnung

Mit

$$k = \frac{2\pi}{\lambda} ,$$

$$r_m = r_0 + m \frac{\lambda}{2} ,$$

$$r_{m-1} = r_0 + (m-1) \frac{\lambda}{2} = r_0 + m \frac{\lambda}{2} - \frac{\lambda}{2} ,$$

$$e^{i\pi} = e^{-i\pi} = -1$$

ist

$$\begin{aligned} \left[e^{ikr} \right]_{r_{m-1}}^{r_m} &= e^{ikr_m} - e^{ikr_{m-1}} = e^{ik(r_0 + m \frac{\lambda}{2})} - e^{ik(r_0 + m \frac{\lambda}{2} - \frac{\lambda}{2})} \\ &= e^{ikr_0} \cdot e^{ik \frac{\lambda}{2} m} - e^{ikr_0} \cdot e^{ik \frac{\lambda}{2} m} \cdot e^{-ik \frac{\lambda}{2}} = e^{ikr_0} \cdot e^{ik \frac{\lambda}{2} m} \cdot (1 - e^{-ik \frac{\lambda}{2}}) \\ &= e^{ikr_0} \cdot (e^{i\pi})^m \cdot (1 - e^{-i\pi}) , \end{aligned}$$

$$\left[e^{ikr} \right]_{r_{m-1}}^{r_m} = e^{ikr_0} \cdot (-1)^m \cdot 2 .$$

$$E_m = C \cdot K_m E_R \frac{2\pi R}{R + r_0} e^{i(kR - \omega t)} \cdot \frac{1}{ik} \cdot e^{ikr_0} \cdot (-1)^m \cdot 2$$

$$= -i(-1)^m \cdot C \cdot K_m E_R \frac{2\pi R}{R + r_0} e^{i[k(R+r_0) - \omega t]} \cdot \frac{\lambda}{\pi} ,$$

$$E_m = C \cdot (-1)^{m+1} \frac{i \cdot 2\lambda \cdot K_m E_R \cdot R}{R + r_0} e^{i[k(R+r_0) - \omega t]} ,$$

$$\boxed{E_m(P) = C \cdot (-1)^{m+1} \frac{2 \cdot i \lambda \cdot K_m \cdot E_R \cdot R}{R + r_0} e^{i[k(R+r_0) - \omega t]} } . \quad (19.16)$$

- Die Gesamtfeldstärke im Punkt P ist folglich

$$\begin{aligned}
 E(P) &= \sum_{m=1}^N E_m = |E_1| - |E_2| + |E_3| - |E_4| + |E_5| - \dots \pm |E_N| \\
 &= \left(\frac{1}{2}|E_1| + \frac{1}{2}|E_1|\right) - |E_2| + \left(\frac{1}{2}|E_3| + \frac{1}{2}|E_3|\right) - |E_4| + \left(\frac{1}{2}|E_5| + \frac{1}{2}|E_5|\right) - \dots \pm |E_N|.
 \end{aligned} \tag{19.17}$$

Dabei haben wir den Vorzeichenwechsel von E_m beim Zonenwechsel berücksichtigt.

- Mit K_m verändern sich auch die $|E_m|$ nur sehr langsam, sodass

$$|E_m| \approx \frac{1}{2}(|E_{m-1}| + |E_{m+1}|) \Leftrightarrow \frac{1}{2}|E_{m-1}| - |E_m| + \frac{1}{2}|E_{m+1}| \approx 0.$$

Dem entsprechend können wir (19.17) umordnen und erhalten

$$E(P) = \frac{1}{2}|E_1| + \underbrace{\left(\frac{1}{2}|E_1| - |E_2| + \frac{1}{2}|E_3|\right)}_{\approx 0} + \underbrace{\left(\frac{1}{2}|E_3| - |E_4| + \frac{1}{2}|E_5|\right)}_{\approx 0} + \dots + \frac{1}{2}|E_N|. \tag{19.18}$$

Der Beitrag E_N der N -ten Zone, d. h. der letzten Zone um $\vartheta = \pi$, kann wegen $K_N = 0$ vernachlässigt werden. Es verbleibt als Gesamtfeldstärke im Punkt P mit (19.16) für E_1 und mit $E_R = A/R$

$$E(P) \approx \frac{1}{2} E_1 = \frac{1}{2} \cdot \frac{C \cdot 2 \cdot i \lambda \cdot K_1 \cdot E_R \cdot R}{R + r_0} e^{i[k(R+r_0) - \omega t]},$$

$$\boxed{E(P) \approx C \cdot K_1 \cdot i \lambda \cdot \frac{A}{R + r_0} e^{i[k(R+r_0) - \omega t]}}. \tag{19.19}$$

- Für die Bestimmung von C benutzen wir einerseits, dass für E_1 bzw. in der ersten Zone $\cos(\vartheta \approx 0) \approx 1 \Rightarrow K_1 = 1$ gilt, und andererseits den Vergleich von (19.19) mit der Feldstärke

$$\boxed{E(P) = \frac{A}{R + r_0} \cdot e^{i[k(R+r_0) - \omega t]}}.$$

die im Punkt P , d. h. im Abstand $R + r_0$ von O , durch die von O ausgehende Primärwelle³⁵ (19.11) erzeugt wird:

$$C \cdot K_1 \cdot i \lambda \cdot \frac{A}{R + r_0} = C \cdot i \lambda \cdot \frac{A}{R + r_0} = \frac{A}{R + r_0} \Rightarrow$$

$$\boxed{C = \frac{1}{i \lambda}}.$$

³⁵Der in (19.19) auftretende Faktor $i = e^{i\frac{\pi}{2}}$ kann in der Exponentialfunktion berücksichtigt werden und verändert diese dann zu $e^{i[k(R+r_0) - \omega t + \frac{\pi}{2}]}$. Das bedeutete eine Phasenverschiebung der Sekundärwellen gegenüber der Primärwelle um $\frac{\pi}{2}$. Allerdings ist eine Diskussion dieses Problems in diesem Zusammenhang nicht notwendig.

- Berechnung von ϱ :

Mit

$$\begin{aligned} c &= R + r_0, \\ \sin \alpha &= \frac{\varrho}{R}, \\ R \cdot \cos \alpha &= R \cdot \sqrt{1 - \sin^2 \alpha} = \sqrt{R^2 - \varrho^2}, \\ z &= c - R \cos \alpha = c - \sqrt{R^2 - \varrho^2} \end{aligned}$$

erhält man

$$\begin{aligned} r^2 &= \varrho^2 + z^2 = \varrho^2 + c^2 + R^2 - \varrho^2 - 2c\sqrt{R^2 - \varrho^2}, \\ r^2 &= c^2 + R^2 - 2c\sqrt{R^2 - \varrho^2} \Leftrightarrow \\ \sqrt{R^2 - \varrho^2} &= \frac{c^2 + R^2 - r^2}{2c}. \end{aligned}$$

Quadrieren und Auflösen nach ϱ liefert schließlich

$$\varrho = \frac{\sqrt{4c^2R^2 - (c^2 + R^2 - r^2)^2}}{2c}.$$

Dies lässt sich leicht überprüfen. $R = 3$, $r = 4$ und $c = 5$ beispielsweise ergeben $\varrho = 2,4$ und $\sin \alpha = 0,8$ bzw. $\cos \alpha = 0,6$ sowie $R \cos \alpha = 1,8$.

Für sehr große R bzw. für $R \rightarrow \infty$ geht die Kugelfläche in Abbildung 13 über in eine Ebene, d. h. die Kugelwellenfront der Primärwelle geht über in eine ebene Wellenfront. Das führt zu einer wesentlichen Vereinfachung der Berechnung von ϱ (siehe Abbildung 15):

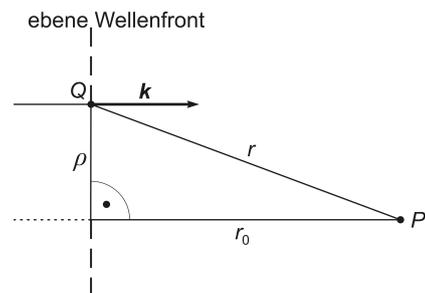


Abb. 15 Zur Berechnung von ϱ bzw. ϱ_m für $R \rightarrow \infty$.

$$\varrho^2 = r^2 - r_0^2 \Rightarrow \varrho = \sqrt{r^2 - r_0^2}$$

bzw.

$$\varrho_m^2 = \left(r_0 + m \frac{\lambda}{2}\right)^2 - r_0^2 = \frac{m^2}{4} \lambda^2 + m r_0 \lambda \Rightarrow$$

$$\boxed{\varrho_m = \sqrt{\frac{m^2}{4} \lambda^2 + m r_0 \lambda} \quad \text{für } R \rightarrow \infty}.$$

Wie man sieht, hängen Breite und Lage der Fresnel'schen Zonen neben λ in jedem Fall insbesondere von r_0 ab.

- Näherung von ϱ_m für $\lambda \ll r_0$ und für kleine m (siehe Abbildung 13):

$$\varrho_m \approx \sqrt{r_m^2 - r_0^2}.$$

Wir nähern weiter mittels Taylor-Entwicklung des Radikanden bis zur 1. Ordnung an der Stelle $\lambda = 0$:

$$\begin{aligned} r_m^2 - r_0^2 &= \left(r_0 + m \frac{\lambda}{2}\right)^2 - r_0^2 = r_0^2 + \left(\frac{m}{2} \lambda\right)^2 + r_0 m \lambda - r_0^2 \\ &= \frac{m^2}{4} \lambda^2 + r_0 m \lambda \quad \Rightarrow \end{aligned}$$

$$r_m^2 - r_0^2 \approx \left[\frac{m^2}{2} \lambda + r_0 m \right]_{\lambda=0} \cdot \lambda = m r_0 \lambda \quad \Rightarrow$$

$$\boxed{\varrho_m \approx \sqrt{m \cdot r_0 \cdot \lambda} \quad \text{für } r_0 \gg \lambda}. \quad (19.20)$$

Mit dem Konzept der Fresnel'schen Zonen lassen sich einige Phänomene erklären, die auf den ersten Blick merkwürdig erscheinen, weil sie sich durch die (anschauliche) geometrische Optik nicht beschreiben lassen. Wir zeigen dies im Folgenden an drei Beispielen:

1. kreisförmige Blende

Zwischen O und P positionieren wir einen lichtundurchlässigen Schirm mit einem zentralen kreisförmigen Loch, d. h. einer Blende, im Abstand r_0 vor dem Punkt P . Der Durchmesser D der Blende sei gerade so groß wie die 1. Fresnel'sche Zone, nämlich³⁶

$$D = 2 \cdot \varrho_1 \approx 2 \cdot \sqrt{r_1^2 - r_0^2} = 2 \cdot \sqrt{\left(r_0 + 1 \cdot \frac{\lambda}{2}\right)^2 - r_0^2}.$$

Somit erreicht nur das Licht aus der 1. Fresnel'schen Zone den Punkt P . Gemäß (19.16) und mit $K_1 = \frac{1}{i\lambda}$ sowie $E_R = \frac{A}{R}$ ist die Feldstärke in P folglich

$$E(P) = E_1 = 2 \cdot \frac{A}{R + r_0} e^{i[k(R+r_0) - \omega t]},$$

also doppelt so groß wie ohne Schirm. Und die daraus resultierende Intensität in P ist viermal so groß wie ohne Schirm.

Ursache dafür ist die fehlende (insbesondere destruktive) Interferenz mit dem Licht aus den übrigen Fresnel'schen Zonen.

2. kreisförmige Scheibe

Zwischen O und P positionieren wir eine lichtundurchlässige kreisförmige Scheibe im Abstand r_0 vor dem Punkt P . Der Durchmesser D der Scheibe sei gerade so

³⁶Beispielsweise beträgt der Radius ϱ_1 der 1. Fresnel'schen Zone für $r_0 = 20$ cm und grünes Licht mit der Wellenlänge $\lambda = 500$ nm = $5,0 \cdot 10^{-7}$ m

$$\varrho \approx \sqrt{r^2 - r_0^2} \quad \Rightarrow \quad \varrho_m \approx \sqrt{\left(r_0 + m \cdot \frac{\lambda}{2}\right)^2 - r_0^2} \quad \Rightarrow \quad \varrho_1 \approx \sqrt{\left(r_0 + \frac{\lambda}{2}\right)^2 - r_0^2} \approx 0,32 \text{ mm}.$$

groß wie die 1. Fresnel'sche Zone, sodass das Licht von der 1. Fresnel'schen Zone ausgeblendet wird. Wenn wir also $E_1 = 0$ in (19.18) berücksichtigen, erhalten wir

$$E(P) = 0 + \underbrace{\left(0 - |E_2| + \frac{1}{2}|E_3|\right)}_{\approx 0} + \underbrace{\left(\frac{1}{2}|E_3| - |E_4| + \frac{1}{2}|E_5|\right)}_{\approx 0} + \cdots + \frac{1}{2}|E_N|.$$

Mit $|E_1| \approx |E_2| \approx |E_3|$, unter Vernachlässigung von $|E_N|$ und unter Beachtung, dass gemäß (19.16) E_2 negativ und E_3 positiv ist, resultiert daraus schließlich

$$E(P) \approx -|E_2| + \frac{1}{2}|E_3| + \underbrace{\left(\frac{1}{2}|E_3| - |E_4| + \frac{1}{2}|E_5|\right)}_{\approx 0} + \cdots + 0 = -E_1 + \frac{1}{2}E_1,$$

$$E(P) \approx -\frac{1}{2}E_1.$$

Obwohl das Licht der 1. Fresnel'schen Zone durch eine Scheibe ausgeblendet wird, ist die Lichtintensität in P genauso groß wie ohne Scheibe.

3. Fresnel'sche Zonenplatte

Werden die zu einer bestimmten „Bildweite“ r_0 und für eine bestimmte Wellenlänge λ ermittelten geraden *oder* ungeraden Fresnel'schen Zonen (Kreisinge) auf einer Glasplatte vollständig oder teilweise lichtundurchlässig gemacht und treffen dann näherungsweise ebene Wellen (z. B. Sonnenlicht) auf diese Platte, so kommt es dahinter im Punkt P zu einer Intensitätsverstärkung des Lichts. Diese Fresnel'sche Zonenplatte wirkt also ähnlich wie eine Sammellinse. Ursache für die Verstärkung ist die konstruktive Interferenz der hindurchgelassenen und in P gleichphasig ankommenden Wellenanteile bei gleichzeitiger Verhinderung der destruktiven Interferenz durch die gegenphasigen Anteile (mit entgegengesetztem Vorzeichen von E).

Wenn der Durchmesser D der kreisförmigen Blende, die wie oben beschrieben im Abstand r_0 von P positioniert ist, wesentlich größer ist als der Durchmesser der 1. Fresnel'sche Zone, also

$$D \gg 2\varrho_1 \approx 2\sqrt{r_0\lambda},$$

tragen viele Fresnel'sche Zonen zum Feld im Punkt P bei und wir sprechen von Fresnel-Beugung.

Wenn aber r_0 so groß wird, dass der Durchmesser D der Blende gleich groß oder kleiner ist als der Durchmesser der 1. Fresnel'schen Zone, also

$$D \leq 2\varrho_1 \approx 2\sqrt{r_0\lambda},$$

trägt nur die 1. Fresnel'sche Zone zum Feld im Punkt P bei und wir sprechen von Fraunhofer-Beugung.

20 Zur (skalaren) Fresnel-Kirchhoff'schen Beugungstheorie

Wir stützen uns in diesem Kapitel auf die leicht „lesbaren“ Abhandlungen in:

Joos, Lehrbuch der Theoretischen Physik, Aula-Verlag Wiesbaden, 15. Auflage, 1989, Drittes Buch, Elftes Kapitel *Elektromagnetische Wellen. IV. Teil. Der Einfluß der Begrenzung (Theorie der Beugung)*, Seite 363 bis Seite 368 und

Eugene Hecht, Optik, Oldenbourg-Verlag, München, Wien, 4. Auflage, 2002, Abschnitt 10.4 *Die skalare Beugungstheorie von Kirchhoff*, Seite 819 bis Seite 824.

Ziel dieses Kapitels ist die Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals über das 2. Green'sche Theorem aus den Maxwell'schen Gleichungen, den grundlegenden Gleichungen der Elektrodynamik. **Achtung!** Aus praktischen Erwägungen verwenden wir in diesem Kapitel eine andere Notation als im Abschnitt 19.3 *Fresnel'sche Zonen*.

20.1 Vorbereitung – Helmholtz-Gleichung

Mit dem

$$\text{Potentialansatz} \quad \vec{E} = -\text{grad } \Phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \text{rot } \vec{A} \quad (20.1)$$

und der

$$\text{Lorenz-Eichung} \quad \text{div } \vec{A} + \frac{1}{c_m^2} \frac{\partial \Phi}{\partial t} \stackrel{!}{=} 0$$

liefern die Maxwell'schen Gleichungen die skalare inhomogene Wellengleichung

$$\frac{1}{c_m^2} \frac{\partial^2}{\partial t^2} \Phi - \Delta \Phi = \frac{1}{\varepsilon} \varrho$$

(Quellenterm $\frac{1}{\varepsilon} \varrho$, Dichte freier Ladungen bzw. elektrische Ladungsdichte ϱ) und die vektorielle inhomogene Wellengleichung

$$\frac{1}{c_m^2} \frac{\partial^2}{\partial t^2} \vec{A} - \Delta \vec{A} = \mu \vec{j}$$

(Quellenterm $\mu \vec{j}$, elektrische Stromdichte \vec{j}).

Die Lichtgeschwindigkeit in einem Medium ist c_m . Die Quellenterme verschwinden in einem Bereich ohne freie elektrische Ladungen und ohne elektrische Ströme – z. B. in einem Dielektrikum. Betrachten wir das elektromagnetische Wechselfeld außerdem in einem Bereich ohne Materie, d. h. im **Vakuum**, so müssen wir statt c_m die **Vakuumlichtgeschwindigkeit** c verwenden und erhalten die (**quellenfreien**) homogenen Wellengleichungen

$$\Delta \Phi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \Phi = 0, \quad (20.2)$$

$$\Delta \vec{A} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} = \vec{0}. \quad (20.3)$$

Dabei handelt es sich um lineare partielle Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten. Warum und in welcher Weise es sich zudem um Wellengleichungen handelt, wird sehr ausführlich beschrieben von Rainer J. Jellito³⁷.

³⁷Rainer J. Jellito, Studententext – Elektrodynamik, Theoretische Physik, 3. Auflage, Aula-Verlag, Wiesbaden, 1994, Abschnitt 5.2 *Die Wellengleichung*, Seite 119 bis Seite 130.

Wir zeigen dies an drei Beispielen. Zunächst setzen wir in (20.2) für Φ die „verallgemeinerte“ Wellenfunktion

$$F(x, t) = a \cdot f(x - vt)$$

mit der Ausbreitungsgeschwindigkeit v ein:

$$\begin{aligned} \frac{\partial^2}{\partial x^2} [a \cdot f(x - vt)] &= \frac{1}{v^2} \frac{\partial^2}{\partial t^2} [a \cdot f(x - vt)] \\ \Rightarrow a \cdot f''(x - vt) &= \frac{1}{v^2} a \cdot f''(x - vt) \cdot (-v)^2 = a \cdot f''(x - vt) . \quad \square \end{aligned}$$

Jetzt setzen wir in (20.2) für Φ die Kugelwellenfunktion

$$E(r, t) = \frac{1}{r} \cdot \sin(kr - \omega t) \quad (20.4)$$

mit $\omega = ck$ und der Ausbreitungsgeschwindigkeit $c = \frac{\omega}{k}$ ein. Allerdings müssen wir dabei den Laplace-Operator in Kugelkoordinaten

$$\Delta_r = \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial}{\partial r} \right]$$

verwenden:

$$\begin{aligned} \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial}{\partial r} \left(\frac{1}{r} \cdot \sin(kr - \omega t) \right) \right] &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \left(\frac{1}{r} \cdot \sin(kr - \omega t) \right) \\ \Rightarrow -\frac{k^2}{r} \sin(kr - \omega t) &= \frac{1}{c^2} \left(-\frac{\omega^2}{r} \sin(kr - \omega t) \right) \\ &= -\frac{k^2}{r} \sin(kr - \omega t) . \quad \square \end{aligned}$$

Schließlich setzen wir die Funktion

$$\vec{E}(\vec{r}, t) = -i\omega \vec{A}_0 e^{i(\vec{k}\vec{r} - \omega t)} ,$$

die wir weiter unten herleiten und benötigen werden, in die vektorielle Wellengleichung (20.3) ein und finden mit

$$\omega = c \cdot |\vec{k}| \quad \Rightarrow \quad \omega^2 = c^2 k^2$$

dass auch diese Funktion die Wellengleichung erfüllt und somit eine Wellenfunktion ist:

$$ik^2 \omega \vec{A}_0 e^{i(\vec{k}\vec{r} - \omega t)} = \frac{1}{c^2} \omega^2 i\omega \vec{A}_0 e^{i(\vec{k}\vec{r} - \omega t)} = ik^2 \omega \vec{A}_0 e^{i(\vec{k}\vec{r} - \omega t)} . \quad \square$$

Von den Potentialgleichungen (20.2) und (20.3) zurück zum \vec{E} -Feld kommen wir, indem wir zunächst den Potentialansatz (20.1) in die homogene (quellenfreie) Maxwell'sche Gleichung

$$\text{rot } \vec{E} + \frac{\partial}{\partial t} \vec{B} = \vec{0}$$

einsetzen:

$$\text{rot } \vec{E} + \frac{\partial}{\partial t} \text{rot } \vec{A} = \text{rot } \vec{E} + \text{rot } \frac{\partial}{\partial t} \vec{A} = \vec{0} \quad \Leftrightarrow \quad \vec{E} = -\frac{\partial \vec{A}}{\partial t} . \quad (20.5)$$

Setzen wir jetzt für \vec{A} die Wellenfunktion $\vec{A}(\vec{r}, t) = \vec{A}_0 e^{i(\vec{k}\vec{r} - \omega t)}$ an, wobei die Amplitude \vec{A}_0 nicht von t abhängt, resultiert aus (20.5)

$$\vec{E}(\vec{r}, t) = \underbrace{-i\omega \vec{A}_0}_{= \vec{E}_0} e^{i(\vec{k}\vec{r} - \omega t)} = \vec{E}_0 e^{i(\vec{k}\vec{r} - \omega t)} ,$$

ebenfalls eine Wellenfunktion, wie wir oben bereits gezeigt hatten.

In der Optik sind vor allem das \vec{E} -Feld oder auch **optische Feld** und folglich die skalare homogene Wellengleichung (20.2) von Bedeutung, weil bei Abwesenheit zeitlich-veränderlicher Magnetfelder gemäß des Potentialansatzes (20.1) allein

$$\vec{E} = -\text{grad } \Phi \quad \longleftrightarrow \quad \Phi = - \int \vec{E} \cdot d\vec{r}$$

gilt. Setzen wir dies in (20.2) ein, resultiert

$$\Delta \underbrace{\left(- \int \vec{E} d\vec{r} \right)}_{=\Phi} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \underbrace{\left(- \int \vec{E} d\vec{r} \right)}_{=\Phi} \Rightarrow - \int \Delta \vec{E} d\vec{r} = - \int \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E} d\vec{r} \Rightarrow$$

$$\Delta \vec{E}(\vec{r}, t) = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E}(\vec{r}, t) \quad \text{bzw.} \quad \begin{pmatrix} \Delta E_x \\ \Delta E_y \\ \Delta E_z \end{pmatrix} = \begin{pmatrix} \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_x \\ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_y \\ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_z \end{pmatrix}.$$

Wie das skalare Potential Φ so erfüllt also auch das aus Φ hervorgehende elektrische Feld \vec{E} die Wellengleichung (20.2).

In der skalaren Beugungstheorie von Kirchhoff sind vereinfachend nur monochromatische Wellen mit $\omega = c|\vec{k}| = \text{const}$ von Interesse. Außerdem beschränkt sich diese Theorie auf ebene Wellen oder Kugelwellen. Ebene Wellen $\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\vec{k}\vec{r} - \omega t)}$ hängen nur in einer Richtung vom Ort ab, nämlich in ihrer Ausbreitungsrichtung, der Richtung des Wellenvektors \vec{k} . Kugelwellen $\vec{E}(r, t) = \frac{\vec{E}_0}{r} e^{i(kr - \omega t)}$, hier praktischerweise dargestellt in Kugelkoordinaten, lösen für $r \neq 0$ ebenfalls die Wellengleichung (20.2). Im Punkt $r = 0$ liefern sie jedoch mit der Wellengleichung die Beziehung

$$(\Delta_r + k^2) \frac{e^{\pm i(kr - \omega t)}}{r} = \left(\Delta_r + \frac{\omega^2}{c^2} \right) \frac{e^{\pm i(kr - \omega t)}}{r} = -4\pi \delta\vec{r},$$

„Wichtig ist nun für uns, zu sehen, daß die von uns untersuchten ebenen Wellen auch die Maxwell'schen Gleichungen selbst lösen, sofern die Quellterme ρ und \vec{j} identisch verschwinden. Das zeigt uns nämlich die *Existenz elektromagnetischer Wellen*, welche sich, einmal angeregt, im Vakuum ungedämpft fortpflanzen. . . .

Da die Maxwell'schen Gleichungen linear sind, werden der Real- und der Imaginärteil komplexer Lösungen in ihnen nicht gemischt. Finden wir also, daß bereits die *komplexen Wellen* Lösungen dieser Gleichungen sind, so können wir . . . durchgängig mit ihnen rechnen und den Übergang zum Realteil erst am Endergebnis vollziehen. Dieses Verfahren erleichtert unsere Untersuchungen beträchtlich.“³⁸

Verwenden wir also an Stelle des Potentials Φ in der skalaren Wellengleichung (20.2) als Lösungsfunktion die Funktion des \vec{E} -Feldes einer monochromatischen elektromagnetischen Welle, z. B. die Kugelwellenfunktion (20.4)

$$E(r, t) = \frac{1}{r} e^{i(kr - \omega t)} = \frac{1}{r} e^{ikr} \cdot e^{-i\omega t}.$$

³⁸Zitiert aus: Rainer J. Jellito, Studententext – Elektrodynamik, Theoretische Physik, 3. Auflage, Aula-Verlag, Wiesbaden, 1994, Seite 131.

Dann können wir (20.2) mit dem Produktansatz nach den Koordinaten r und t in zwei gewöhnliche Differentialgleichungen separieren, die voneinander unabhängig und jeweils nur von einer Variablen abhängig sind:

$$\text{Produktansatz : } E(r, t) = \mathcal{E}(r) \cdot T(t) \quad \Rightarrow \quad \begin{cases} \text{Ortsanteil } \mathcal{E}(r) = \frac{1}{r} e^{ikr} \\ \text{Zeitanteil } T(t) = e^{-i\omega t} \end{cases} .$$

Setzen wir also den Produktansatz in die Wellengleichung (20.2) ein:

$$\Delta_r \mathcal{E}(r) \cdot T(t) - \frac{1}{c^2} \frac{d^2 T(t)}{dt^2} \cdot \mathcal{E}(r) = 0 .$$

Mittels Division dieser Gleichung durch $\mathcal{E}(r) \cdot T(t)$ erfolgt dann die Separation:

$$\frac{\Delta_r \mathcal{E}(r)}{\mathcal{E}(r)} = \frac{1}{c^2} \frac{\ddot{T}(t)}{T(t)} = \text{const} = \kappa .$$

κ ist die Separationskonstante. Uns interessiert im Folgenden nur die daraus resultierende ortsabhängige gewöhnliche lineare homogene Differentialgleichung 2. Ordnung

$$\Delta_r \mathcal{E}(r) - \kappa \cdot \mathcal{E}(r) = 0 . \quad (20.6)$$

Mit der Lösungsfunktion $\mathcal{E}(r) = \frac{1}{r} e^{ikr}$ in (20.6) bestimmen wir die Separationskonstante κ bezüglich unserer Kugelwellenfunktion. Wir erhalten als Ergebnis

$$-\frac{1}{r} \cdot k^2 \cdot e^{ikr} - \kappa \cdot \frac{1}{r} \cdot e^{ikr} = 0 \quad \Leftrightarrow \quad \kappa = -k^2$$

und damit schließlich die

Helmholtz-Gleichung $\Delta \mathcal{E} + k^2 \mathcal{E} = 0$

20.2 Die Kirchhoff'sche Formel

Die Fresnel-Kirchhoffsche Beugungstheorie ist eine skalare Theorie, weil nur eine der beiden zur Wellennormalen \vec{k}/k senkrechten Komponenten von

$$\vec{E}(\vec{r}, t) = \vec{E}_0(\vec{r}) \cdot e^{i(\vec{k} \cdot \vec{r} - \omega t)} = \vec{E}_0(\vec{r}) \cdot e^{i(kr - \omega t)} = \vec{\mathcal{E}}(\vec{r}) \cdot e^{-i\omega t}$$

betrachtet wird. Wenn also das \vec{E} -Feld beispielsweise genau in z -Richtung propagiert, betrachten wir die E_x - oder die E_y -Komponente. Für die im Folgenden betrachtete Komponente schreiben wir mit der Amplitude E_0 und dem Phasenwinkel $\phi = kr - \omega t$ kurz

$$E(\vec{r}, t) = E_0(\vec{r}) \cdot e^{i(kr - \omega t)} \quad \text{bzw. im Fall einer Kugelwelle} \quad E(r, t) = E_0(r) \cdot e^{i(kr - \omega t)} .$$

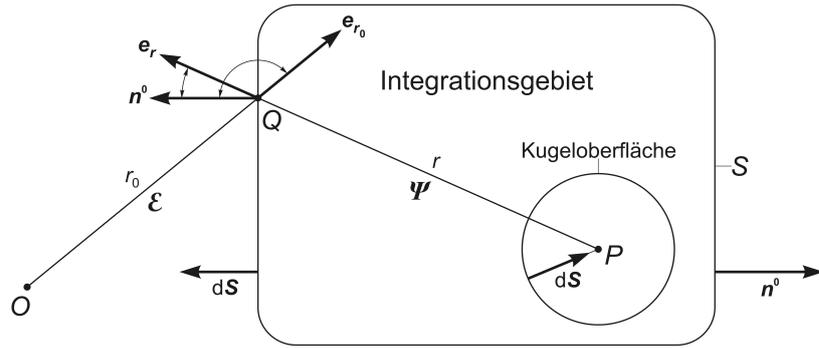


Abb. 16 Zur Herleitung der Kirchhoff'schen Formel. Das Integrationsgebiet, das von der geschlossenen Fläche S umhüllt wird, ist von beliebiger Gestalt! Weil wir hier die **freie Wellenausbreitung** ohne irgendwelche Hindernisse betrachten, ist die Hülle S allein mathematischer Natur, also nur gedacht und nicht „materiell“. Der Abstand r_0 wird von O aus und der Abstand r wird von P aus gemessen. \vec{n}^0 ist der Einheitsnormalenvektor auf S . Weiterhin gilt $\Psi = \Psi(r)$ und $\mathcal{E} = \mathcal{E}(r_0)$ mit $\vec{PQ} = \vec{r}$ und $\vec{PO} = \vec{r}' \Rightarrow r_0 = |\vec{r} - \vec{r}'|$.

Und weil uns im Folgenden nur der ortsabhängige Anteil \mathcal{E} der monochromatischen Kugelwellenfunktion $E(r_0, t)$ interessiert, machen wir hinsichtlich der Zeitabhängigkeit den Ansatz

$$E(r_0, t) = \mathcal{E}(r_0) \cdot e^{-i\omega t} \quad \text{mit} \quad \omega = |\vec{k}| \cdot c = k \cdot c = \text{const} \quad \text{und} \quad k = \frac{2\pi}{\lambda}.$$

Die Funktion $\mathcal{E}(r_0)$ löst die bereits gezeigte Helmholtz-Gleichung

$$\Delta \mathcal{E} + k^2 \mathcal{E} = 0 \quad \Leftrightarrow \quad \Delta \mathcal{E} = -k^2 \mathcal{E}.$$

Angenommen wir haben außer der Funktion \mathcal{E} eine weitere Funktion Ψ , die ebenfalls die Wellengleichung und damit die Helmholtz-Gleichung erfüllt. Unter Berücksichtigung der Helmholtz-Gleichung ergibt dann das 2. Green'sche Theorem mit diesen beiden Funktionen bei der Integration über ein beliebiges Raumgebiet und dortiger Stetigkeit der Integranden

$$\begin{aligned} \oint_S (\mathcal{E} \text{grad} \Psi - \Psi \text{grad} \mathcal{E}) \cdot d\vec{S} &= \int_V (\mathcal{E} \Delta \Psi - \Psi \Delta \mathcal{E}) dV \\ &= \int_V [\mathcal{E}(-k^2 \Psi) - \Psi(-k^2 \mathcal{E})] dV \quad (20.7) \\ &= \int_V k^2 \underbrace{(\Psi \mathcal{E} - \mathcal{E} \Psi)}_{=0} dV = 0. \end{aligned}$$

Als Funktion Ψ wählen wir den Ortsanteil der vom Punkt P des Integrationsgebiets ausgehenden Kugelwelle (siehe Abbildung 16)

$$\Psi = \frac{e^{ikr}}{r}.$$

Weil die Kugelwellenfunktion und auch ihr Ortsanteil in P eine Unendlichkeitsstelle bzw. eine Singularität besitzen, können wir nicht ohne weiteres über das gesamte Integrationsgebiet integrieren. Zumindest ist die Beziehung (20.7) nur dann mit Sicherheit

gleich Null, wenn die Singularität nicht zum Integrationsgebiet gehört. Die Idee zur Lösung dieses Problems ist folgende Vorgehensweise:

1. Wir schließen die Singularität aus dem Integrationsgebiet aus, indem wir P als Mittelpunkt für eine kleine Kugel innerhalb des Integrationsgebiets wählen.
2. Das Integral über die gesamte Oberfläche des Integrationsgebiets setzt sich dann zusammen aus dem Integral über die Außenfläche S und dem Integral über die Innenfläche bzw. die Kugeloberfläche und ist gemäß (20.7) aber immer noch gleich Null.
3. Schließlich lassen wir den Radius r der kleinen Kugel gegen Null gehen, d. h. wir lassen die Kugel auf den Punkt P schrumpfen und schauen, welchen Beitrag P zum Integral leisten kann. Wenn überhaupt, kann ein Beitrag zum Integral nur aus dem Punkt P stammen, weil das Integral (20.7) unter Ausschluss von P stets verschwindet. Anders gesagt, der Grenzwert des Oberflächenintegrals der kleinen Kugel für $r \rightarrow 0$ liefert uns den Wert des Volumenintegrals über das gesamte Integrationsgebiet einschließlich des Punktes P in symbolischer Schreibweise wie folgt:

$$\oint_S \cdots dS + \oint_{\text{Kugel}} \cdots dS = 0 \Leftrightarrow \oint_S \cdots dS = - \oint_{\text{Kugel}} \cdots dS \implies$$

$$\underbrace{\lim_{\text{Kugelradius} \rightarrow 0} \oint_S \cdots dS}_{= \oint_S \cdots dS \text{ einschließlich } P} = \lim_{\text{Kugelradius} \rightarrow 0} - \oint_{\text{Kugel}} \cdots dS. \quad (20.8)$$

Wenn das vektorielle Flächenelement $d\vec{S}$ der äußeren Oberfläche S nach außen zeigt, muss das der Kugeloberfläche im Innern radial nach innen zum Punkt P zeigen, damit wir die beiden Flächenteilintegrale wie folgt addieren können:³⁹

$$\oint_S \left(\mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} \right) \cdot d\vec{S} + \oint_{\text{Kugel}} \left(\mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} \right) \cdot d\vec{S} = 0. \quad (20.9)$$

Berechnung des Grenzwertes des Integrals über die kleine Kugel für $r \rightarrow 0$:

$$\oint_{\text{Kugel}} \left(\mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} \right) \cdot d\vec{S} \quad (20.10)$$

- Wir rechnen in Kugelkoordinaten, d. h. der Einheitsnormalenvektor \vec{e}_r zeigt ausgehend vom Kugelmittelpunkt nach außen. Folglich gilt für das nach innen zeigende vektorielle Kugelflächenelement

$$d\vec{S} = -\vec{e}_r \cdot dS = -\vec{e}_r \cdot r^2 \sin \vartheta \, d\vartheta \, d\varphi.$$

³⁹Um das einzusehen, kann man einen Schnitt derart durch das gesamte Integrationsgebiet (es handelt sich hier um einen doppelt zusammenhängenden Bereich) legen, dass der Schnitt auch die Kugel zweiteilt. Setzt man dann die Hüllenintegrale beider Teile zusammen, heben sich die Schnittflächenintegrale gegenseitig auf.

- Die Funktion \mathcal{E} und der Gradient von \mathcal{E} seien im gesamten Integrationsgebiet stetig, weshalb \mathcal{E} im Punkt P den endlichen Wert \mathcal{E}_p annimmt.
- Wegen der Stetigkeit von \mathcal{E} und weil dS für $r \rightarrow 0$ mit r^2 abnimmt, im Nenner von $\Psi = e^{ikr}/r$ das r aber nur in der ersten Potenz steht, verschwindet der Subtrahend im Integranden von (20.10) für $r \rightarrow 0$ gemäß

$$\oint_{\text{Kugel}} - \underbrace{(\vec{e}_r \cdot \text{grad } \mathcal{E}) \cdot e^{ikr} \cdot r \cdot \sin \vartheta \, d\vartheta \, d\varphi}_{=0 \text{ für } \lim_{r \rightarrow 0}}$$

- Mit

$$\text{grad } \Psi = \text{grad } \frac{e^{ikr}}{r} = \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r$$

verbleibt von (20.10) nur das Integral

$$\begin{aligned} \oint_{\text{Kugel}} \mathcal{E} \cdot \text{grad } \frac{e^{ikr}}{r} \cdot d\vec{S} &= \oint_{\text{Kugel}} \mathcal{E} \cdot \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r \cdot (-\vec{e}_r) \cdot r^2 \sin \vartheta \, d\vartheta \, d\varphi \\ &= \oint_{\text{Kugel}} \mathcal{E} \cdot (1 - ikr) e^{ikr} \cdot \sin \vartheta \, d\vartheta \, d\varphi \end{aligned}$$

mit dem Grenzwert

$$\lim_{\substack{r \rightarrow 0 \Rightarrow \\ \mathcal{E} \rightarrow \mathcal{E}_p}} \oint_{\text{Kugel}} \mathcal{E} \cdot (1 - ikr) e^{ikr} \cdot \sin \vartheta \, d\vartheta \, d\varphi = \mathcal{E}_p \cdot \oint_{\text{Kugel}} \sin \vartheta \, d\vartheta \, d\varphi = \mathcal{E}_p \cdot 4\pi .$$

Das Hüllenintegral (20.9) über die geschlossene Oberfläche S des gesamten Integrationsgebiets *einschließlich der Singularität im Punkt P* ist mit (20.8)

$$\oint_S \left(\mathcal{E} \text{ grad } \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \text{ grad } \mathcal{E} \right) \cdot d\vec{S} = -\mathcal{E}_p \cdot 4\pi . \quad (20.11)$$

Durch Äquivalenzumformung (Achtung Vorzeichen!) folgt daraus schließlich die

Kirchhoff'sche Formel $\mathcal{E}_p = \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \text{ grad } \mathcal{E} - \mathcal{E} \text{ grad } \frac{e^{ikr}}{r} \right) \cdot d\vec{S}$

(20.12)

zur Berechnung des Ortsanteils \mathcal{E} des E -Feldes im Punkt P , wenn die Funktionen $\mathcal{E}(r_0)$ und $\Psi(r)$ auf der geschlossenen Oberfläche des Integrationsgebiets bekannt sind. Die Kirchhoff'sche Formel wird auch Kirchhoff'sche allgemeine Beugungsformel genannt.

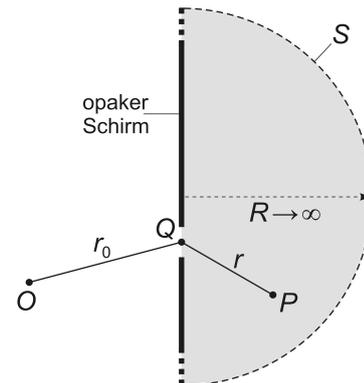
20.3 Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals

Zur Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals wenden wir die im Abschnitt 20.2 hergeleitete Kirchhoff'scher Formel (20.12) auf den Fall an, dass sich im Punkt O die primäre Lichtquelle befindet, welche die Kugelwelle

$$\mathcal{E} = \frac{e^{ikr_0}}{r_0}$$

mit der gleichen Wellenzahl k wie in Ψ emittiert (siehe Abbildung 17).

Abb. 17 Zur Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals. Vereinfachend ist der lichtundurchlässige Schirm hier eben dargestellt. Schirm und Öffnung im Schirm dürfen aber beliebig geformt sein. Während mit Abbildung 16 die freie Wellenausbreitung betrachtet wird, dort die Hülle S des Integrationsgebiets also nicht „materiell“ sondern nur gedacht ist, soll hier der Schirm bis auf die Öffnung „materiell“ sein. Die Halbsphäre spielt bei diesen Überlegungen keine Rolle, weil sie im Unendlichen liegt und keinen Beitrag zum Flächenintegral leistet (siehe Abschnitt 20.5 Zur auslaufenden Welle im Unendlichen).



Weiterhin sind für die Herleitung noch einige Annahmen und Näherungen erforderlich:

- Die geschlossene Außenfläche (Integrationsfläche) S , die das Integrationsgebiet mit dem Punkt P umhüllt, setze sich zusammen beispielsweise aus einer Halbsphäre mit unendlichem Radius und ihrer Basisfläche, die nicht eben sein muss. Die folglich ebenfalls bis ins Unendliche ausgedehnte Basisfläche sei ein opaker Schirm mit einer **lichtdurchlässigen Öffnung** beliebiger Gestalt.
- O befinde sich im Abstand r_0 vor und P in endlichem Abstand r hinter dieser Öffnung.
- Der lichtundurchlässige Schirm liefert keinen Beitrag zum Integral, was intuitiv klar ist.
- Und auch die im Unendlichen liegende Halbsphäre von S soll keinen Beitrag zum Integral leisten (siehe Abschnitt 20.5 Zur auslaufenden Welle im Unendlichen).
- Weiterhin nehmen wir an, dass innerhalb der Öffnung die Feldausbreitung ungestört sei, d. h., die Störung infolge der Feldbegrenzung durch den opaken Schirm soll nicht berücksichtigt werden.
- Innerhalb der Öffnung gelte also für die einfallende Welle

$$\mathcal{E} = \frac{e^{ikr_0}}{r_0} \quad \text{mit} \quad \text{grad } \mathcal{E} ,$$

wobei

$$r_0 \vec{e}_{r_0} \quad \text{der von } O \text{ zu einem Punkt } Q$$

in der Öffnung gezogene Radiusvektor und

$$r \vec{e}_r \quad \text{der von } P \text{ zu einem Punkt } Q$$

in der Öffnung gezogene Radiusvektor ist.

Achtung! Während bei der Herleitung der Kirchhoff'schen Formel die Integrationsfläche die (geschlossene) Hülle S war, beschränkt sich jetzt die Integrationsfläche nur noch auf die Fläche der Öffnung im Schirm.

Der Einheitsnormalenvektor auf der Öffnungsebene sei \vec{n}^0 . Er zeige wie in Abbildung 16 vom Integrationsgebiet ausgehend nach außen. Die folgenden Terme

$$\mathcal{E} = \frac{e^{ikr_0}}{r_0} \quad \Rightarrow \quad \text{grad } \mathcal{E} = \left(\frac{ik}{r_0} e^{ikr_0} - \frac{1}{r_0^2} e^{ikr_0} \right) \cdot \vec{e}_{r_0} = \left(ik - \frac{1}{r_0} \right) \frac{e^{ikr_0}}{r_0} \cdot \vec{e}_{r_0},$$

$$\Psi = \frac{e^{ikr}}{r} \quad \Rightarrow \quad \text{grad } \Psi = \left(\frac{ik}{r} e^{ikr} - \frac{1}{r^2} e^{ikr} \right) \cdot \vec{e}_r = \left(ik - \frac{1}{r} \right) \frac{e^{ikr}}{r} \cdot \vec{e}_r,$$

$$\vec{e}_{r_0} \cdot d\vec{S} = \vec{e}_{r_0} \cdot \vec{n}^0 dS = \cos(\vec{n}^0, \vec{e}_{r_0}) dS,$$

$$\vec{e}_r \cdot d\vec{S} = \vec{e}_r \cdot \vec{n}^0 dS = \cos(\vec{n}^0, \vec{e}_r) dS.$$

sind also in (20.12) einzusetzen. Dabei müssen wir beachten, dass das Flächenintegral nur über die Öffnung läuft:

$$\begin{aligned} \mathcal{E}_p &= \frac{1}{4\pi} \int_{\text{Öffnung}} \left[\frac{e^{ikr}}{r} \left(ik - \frac{1}{r_0} \right) \frac{e^{ikr_0}}{r_0} \cdot \vec{e}_{r_0} - \frac{e^{ikr_0}}{r_0} \left(ik - \frac{1}{r} \right) \frac{e^{ikr}}{r} \cdot \vec{e}_r \right] \cdot \vec{n}^0 dS \\ &= \frac{1}{4\pi} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\left(ik - \frac{1}{r_0} \right) \cos(\vec{n}^0, \vec{e}_{r_0}) - \left(ik - \frac{1}{r} \right) \cos(\vec{n}^0, \vec{e}_r) \right] dS. \end{aligned}$$

Als weitere Näherung vernachlässigen wir in den runden Klammern die Quotienten $1/r_0$ und $1/r$ gegenüber $ik = i2\pi/\lambda$, weil in der Praxis die Abstände zwischen Lichtquelle O bzw. Beobachtungspunkt P und der Öffnung sehr groß sind im Vergleich zur Wellenlänge λ :

$$r_0 \gg \lambda \quad \text{und} \quad r \gg \lambda \quad \Rightarrow \quad \frac{1}{r_0} \ll \frac{2\pi}{\lambda} = k \quad \text{und} \quad \frac{1}{r} \ll \frac{2\pi}{\lambda} = k.$$

Außerdem können wir dann $ik = i2\pi/\lambda$ vor das Integral ziehen:

$$\mathcal{E}_p = \frac{i}{2\lambda} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\cos(\vec{n}^0, \vec{e}_{r_0}) - \cos(\vec{n}^0, \vec{e}_r) \right] dS.$$

Ein Vergleich mit Abbildung 16 zeigt, dass $\sphericalangle(\vec{n}^0, \vec{e}_{r_0})$ ein stumpfer Winkel ist und $\sphericalangle(\vec{n}^0, \vec{e}_r)$ ein spitzer Winkel. Somit gilt:

$$\cos(\vec{n}^0, \vec{e}_{r_0}) < 0, \quad \cos(\vec{n}^0, \vec{e}_r) > 0.$$

Wir vertauschen deshalb die Kosinusse in der eckigen Klammer und ziehen das dabei resultierende Minuszeichen vor das Integral. Dann ziehen wir den Faktor $1/2$ in die eckige Klammer und erhalten

$$\mathcal{E}_p = \frac{-i}{\lambda} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\frac{\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0})}{2} \right] dS.$$

Wie wir sehen, besitzt die rechte Seite dieser Gleichung nicht die Dimension (Maßeinheit) der elektrischen Feldstärke \mathcal{E} . Wir führen deshalb den Faktor ϵ_0 ein, welcher der Feldstärkeamplitude der Punktquelle in O entspricht und nicht mit der elektrischen Feldkonstante ϵ_0 verwechselt werden darf. Damit haben wir das

Fresnel-Kirchhoff'sche Beugungsintegral hergeleitet:

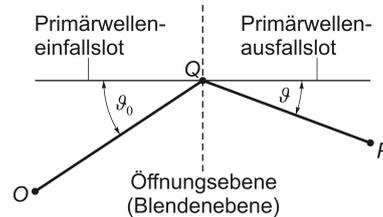
$$\boxed{\mathcal{E}_p = \frac{-i \epsilon_0}{\lambda} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\frac{\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0})}{2} \right] dS}. \quad (20.13)$$

Gemäß Abbildung 16 und Abbildung 18 ist der Term in eckigen Klammern der im folgenden Abschnitt 20.4 diskutierte **Neigungsfaktor** K :

$$\cos(\vec{n}^0, \vec{e}_r) = \cos \vartheta \quad \text{und} \quad -\cos(\vec{n}^0, \vec{e}_{r_0}) = \cos \vartheta_0 \quad \Rightarrow$$

$$\boxed{K = \frac{1}{2} \left[\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0}) \right] = \frac{1}{2} (\cos \vartheta + \cos \vartheta_0)}.$$

Abb. 18 Neigungsfaktor $K = \frac{1}{2}(\cos \vartheta_0 + \cos \vartheta)$ im Fall einer Aperturblende bzw. einer Öffnung.



Wenn r und r_0 groß sind gegen die Ausdehnung der Öffnung, so können wir in (20.13) sowohl r und r_0 im Nenner als auch K als nahezu konstant über die Öffnung betrachten und vor das Integral ziehen:

$$\mathcal{E}_p(r) = \frac{-i \epsilon_0}{\lambda} \frac{K}{r \cdot r_0} \int_{\text{Öffnung}} e^{ik(r+r_0)} \cdot dS. \quad (20.14)$$

Weil der Faktor $k = 2\pi/\lambda$ im Exponenten groß ist, können wir dort r und r_0 nicht als konstant annehmen und dürfen folglich die Exponentialfunktion $e^{ik(r+r_0)}$ nicht vor das Integral ziehen.

Wenn wir dann noch das Fresnel-Kirchhoff'sche Beugungsintegral (20.13) mit dem Zeitanteil $T(t) = e^{-i\omega t}$ kombinieren und zudem den konstanten „Phasenfaktor“ $-i = e^{-i\frac{\pi}{2}}$ berücksichtigen, resultiert das E -Feld im Beobachtungspunkt P :

$$E_p(r, t) = \mathcal{E}_p(r) \cdot T(t) = \frac{\epsilon_0}{\lambda} \frac{K}{r \cdot r_0} \int_{\text{Öffnung}} e^{i[k(r+r_0) - \omega t - \frac{\pi}{2}]} \cdot dS.$$

Die aus dem konstanten „Phasenfaktor“ resultierende **Phasenverschiebung** des Feldes im Beobachtungspunkt P um $-\frac{\pi}{2}$ gegenüber der einfallenden Primärwelle ergibt sich aus der Störung der Wellenausbreitung durch die Apertur. Die Wellenausbreitung von den Punkten Q auf der „Aperturfläche“ zum Punkt P wird nämlich bestimmt durch die zeitliche Änderung bzw. Ableitung der Feldstärke innerhalb der Apertur. Und da der Zeitanteil der primären monochromatischen Welle beschrieben wird durch $e^{-i\omega t}$, ist dessen zeitliche Ableitung proportional zur Kreisfrequenz $\omega = 2\pi\nu$ und insbesondere proportional zu $-i = \frac{1}{i}$. Eine zum Phänomen der Phasenverschiebung analoge Diskussion findet sich in E. Hecht, *Optik*, 4. Auflage, Oldenbourg-Verlag, München, Wien, 2005, Abschnitt 4.2.3 *Transmission und Brechungsindex*, Seite 160 bis Seite 164.

20.4 Der Neigungsfaktor K

Allgemein ist der **Neigungsfaktor** $K = \frac{1}{2}(\cos \vartheta_0 + \cos \vartheta)$ (siehe Abbildung 18).

Wir betrachten zunächst den speziellen Fall der freien Ausbreitung einer Kugelwelle:

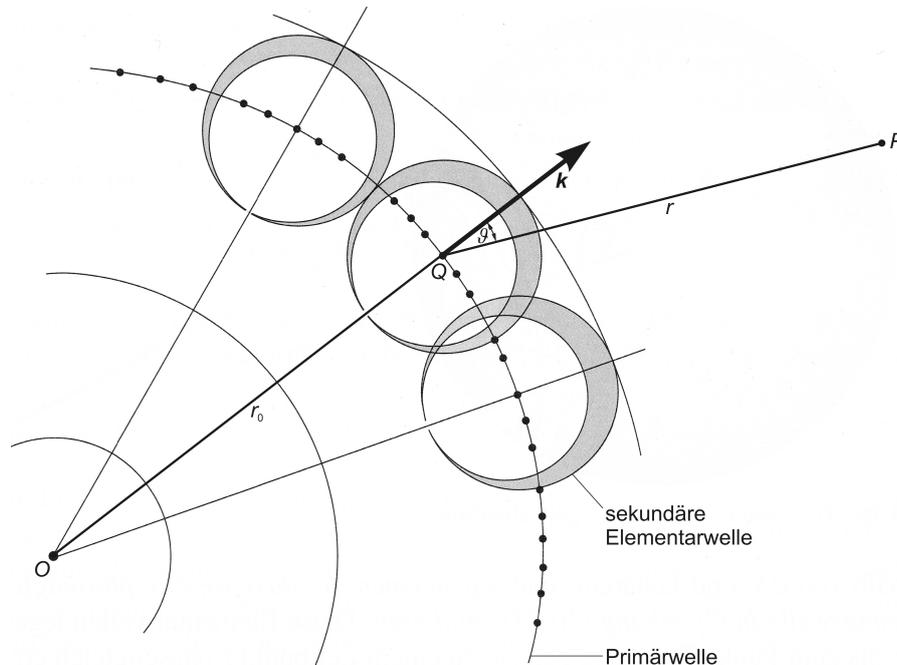


Abb. 19 Die freie Ausbreitung einer Kugelwelle und das Abstrahlungsverhalten der sekundären Elementarwellen. \vec{k} ist der Wellenvektor der betrachteten primären Kugelwellenfront. Das Abstrahlungsverhalten der Sekundärwellen bei freier Ausbreitung der Primärwelle wird beschrieben durch den sog. Neigungsfaktor

$$K(\vartheta) = \frac{1}{2}(1 + \cos \vartheta) \quad \text{mit } 0 \leq K \leq 1.$$

Der Feldstärkebeitrag der Sekundärwellen ist in Ausbreitungsrichtung der Primärwelle maximal entsprechend $K = 1$. Der Feldstärkebeitrag der Sekundärwellen verschwindet entgegengesetzt zur Ausbreitungsrichtung der Primärwelle entsprechend $K = 0$. Diese Theorie stimmt mit dem experimentellen Ergebnis überein, wonach die sekundären Elementarwellen einer primären Kugelwelle insgesamt keine zentripetalen sondern nur zentrifugale Wellenfronten bilden.

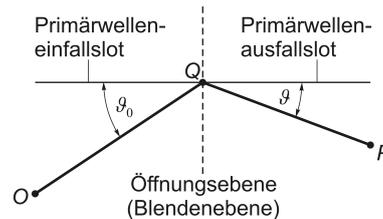
(Abbildung nach: Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag München Wien, 2005, Seite 781)

Bei **freier Wellenausbreitung** wie in Abbildung 19 gilt:

- ϑ_0 ist der Winkel $\vartheta_0 = 0$ zwischen der Ausbreitungsrichtung der Primärwelle und dem Primärwelleneinfallslot auf die Tangentialebene an den Punkt Q der Kugelwellenfront.
- ϑ ist der Winkel zwischen der Ausbreitungsrichtung einer Sekundärwelle und dem Primärwellenausfallslot auf die Tangentialebene an den Punkt Q der Kugelwellenfront.

In der Realität bzw. allgemein erfolgt die Wellenausbreitung jedoch gestört, z. B. durch die Begrenzung mittels einer Aperturblende wie in den Abbildungen 17 und 20.

Abb. 20 Neigungsfaktor $K = \frac{1}{2}(\cos \vartheta_0 + \cos \vartheta)$ im Fall einer Aperturblende bzw. einer Öffnung.



Im Fall einer **Aperturblende** in einem opaken Schirm gemäß Abbildung 20 gilt:

- ϑ_0 ist der Winkel zwischen der Ausbreitungsrichtung der Primärwelle und dem Primärwelleneinfallslot auf die Öffnungsebene.
- ϑ der Winkel zwischen der Ausbreitungsrichtung der Sekundärwelle und dem Primärwellenausfallslot auf die Öffnungsebene.

20.5 Zur auslaufenden Welle im Unendlichen

Wir diskutieren die Aussage:

Eine auslaufende Welle im Unendlichen leistet keinen Betrag zum Feldstärkeintegral im Beobachtungspunkt P .

- Argumentation mit dem Neigungsfaktor K :

Betrachten wir bezüglich Abbildung 19 eine Kugelwellenfront im Unendlichen, also bei $r_0 = R \rightarrow \infty$. Liegt dann (bei festen Punkten O und P) der Punkt Q auf dieser Wellenfront im Unendlichen, so gilt

$$R \rightarrow \infty \Rightarrow \vartheta \rightarrow \pi \Rightarrow K \rightarrow 0.$$

In diesem Fall würden also alle Sekundärwellen genau entgegengesetzt zu den zugehörigen Primärwellen bei P einlaufen. In Analogie zu Abbildung 19 gilt auch für die im Unendlichen liegende Halbsphäre in Abbildung 17

$$\boxed{\lim_{R \rightarrow \infty} K = 0},$$

sodass auch dort die auslaufenden Wellen von der im Unendlichen liegenden Halbsphäre her keinen Beitrag zum Feldstärke-Flächenintegral für den Punkt P leisten. Und weil außerdem der Schirm opak ist, braucht in dem von der Abbildung 17 beschriebenen Modell das Flächenintegral schließlich nur über die Öffnung im Schirm zu laufen.

- Argumentation mit der Sommerfeld'schen Strahlungsbedingung :
(Quelle: Joseph W. Goodman, Introduction to Fourier Optics, 2. Auflage, McGraw-Hill, USA, 1996, Abschnitt 3.4 THE KIRCHHOFF FORMULATION OF DIFFRACTION BY A PLANAR SCREEN, Seite 42 bis Seite 46.)

Wir gehen aus von der Kirchhoff'schen Formel (20.12)

$$\mathcal{E}_p = \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \text{grad } \mathcal{E} - \mathcal{E} \text{grad } \frac{e^{ikr}}{r} \right) \cdot d\vec{S},$$

d. h. von der freien Wellenausbreitung gemäß der Abbildung 16 bzw. von der freien Kugelwelle gemäß der Abbildung 19 mit den Kugelwellenfunktionen

$$\begin{aligned} \Psi = \frac{e^{ikr}}{r} &\Rightarrow \text{grad } \Psi = \text{grad } \frac{e^{ikr}}{r} = \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r, \\ \mathcal{E} = \frac{e^{ikr_0}}{r_0} &\Rightarrow \text{grad } \mathcal{E} = \text{grad } \frac{e^{ikr_0}}{r_0} = \left(\frac{ik}{r_0} - \frac{1}{r_0^2} \right) e^{ikr_0} \cdot \vec{e}_{r_0}. \end{aligned}$$

Zur Vereinfachung nehmen wir jetzt an, dass die geschlossene Integrationsfläche S eine Kugeloberfläche mit dem Radius r und dem nach außen zeigenden Einheitsnormalenvektor \vec{n}^0 ist. Das zugehörige vektorielle Kugelflächenelement ist dann

$$d\vec{S} = \vec{n}^0 \cdot dS = \vec{n}^0 \cdot r^2 \sin \theta \, d\theta \, d\varphi = \vec{n}^0 \cdot r^2 \, d\Omega.$$

Hierbei ist $d\Omega = \sin \theta \, d\theta \, d\varphi$ das Raumwinkelelement, welches bei Integration über den gesamten Raum den

$$\text{vollen Raumwinkel } \Omega_{\mathbb{R}^3} = \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \sin \theta \, d\theta \, d\varphi = \int_{\mathbb{R}^3} d\Omega = 4\pi$$

liefert. \mathbb{R}^3 ist hier der Index für den gesamten (dreidimensionalen) Raum. Die Normalenkomponenten der Gradienten von Ψ und \mathcal{E} sind

$$\begin{aligned} \text{grad } \Psi \cdot \vec{n}^0 &= \frac{\partial \Psi}{\partial n} = \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r \cdot \vec{n}^0, \\ \text{grad } \mathcal{E} \cdot \vec{n}^0 &= \frac{\partial \mathcal{E}}{\partial n} = \left(\frac{ik}{r_0} - \frac{1}{r_0^2} \right) e^{ikr_0} \cdot \vec{e}_{r_0} \cdot \vec{n}^0. \end{aligned} \quad (20.15)$$

Für die Argumentation mit der Sommerfeld'schen Strahlungsbedingung machen wir jetzt die Näherungen

$$\text{für große } r \Rightarrow \begin{cases} \text{grad } \Psi = \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r \approx ik \frac{e^{ikr}}{r} \cdot \vec{e}_r, \\ \vec{e}_r \cdot \vec{n}^0 \approx 1. \end{cases}$$

Diese Näherungen setzen wir unter Berücksichtigung von (20.15) in die Kirchhoff'sche Formel (20.12) ein:

$$\begin{aligned}
\mathcal{E}_p &= \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} \right) \cdot d\vec{S} \\
&\approx \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} ik \frac{e^{ikr}}{r} \vec{e}_r \right) \cdot d\vec{S} \\
&\approx \frac{1}{4\pi} \int_{\mathbb{R}^3} \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} ik \frac{e^{ikr}}{r} \vec{e}_r \right) \cdot \vec{n}^0 \cdot r^2 d\Omega \\
&\approx \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{e^{ikr}}{r} \left(\frac{\partial \mathcal{E}}{\partial n} - ik \mathcal{E} \right) r^2 d\Omega,
\end{aligned}$$

$$\boxed{\mathcal{E}_p \approx \frac{1}{4\pi} \int_{\mathbb{R}^3} e^{ikr} \left(\frac{\partial \mathcal{E}}{\partial n} - ik \mathcal{E} \right) r d\Omega} \quad (20.16)$$

Das Feld auf der Oberfläche einer unendlich großen Kugel leistet im Beobachtungspunkt P keinen Beitrag, wenn das Integral (20.16) verschwindet bzw. wenn für dessen Integranden in isotroper Weise die

$$\boxed{\text{Sommerfeld'sche Strahlungsbedingung} \quad \lim_{r \rightarrow \infty} \left(\frac{\partial \mathcal{E}}{\partial n} - ik \mathcal{E} \right) r = 0}$$

gilt. Diese Bedingung ist erfüllt, wenn \mathcal{E} mindestens so schnell verschwindet wie eine Kugelwelle, was hier wegen

$$\mathcal{E}(r_0) = \frac{e^{ikr_0}}{r_0} = \frac{e^{ikr_0}}{|\vec{r} - \vec{r}'|} \quad \text{mit} \quad \vec{r}' = \vec{PO} = \text{const}, \quad |\vec{r}'| = r' \ll r$$

der Fall ist. Diese Argumentation gilt analog auch im Abschnitt 20.3 für die im Unendlichen liegende Halbsphäre (siehe Abbildung 17).

Wir hätten aber auch ganz einfach wie folgt argumentieren können:

Wenn der Radius r der Kugelwellenfront gegen Unendlich geht, dann geht auch r_0 gegen Unendlich. Daraus folgt

$$\Rightarrow \vec{e}_r \uparrow \vec{n}^0 \wedge \vec{e}_{r_0} \uparrow \vec{n}^0 \quad \Rightarrow \quad \vec{e}_r \cdot \vec{n}^0 = \vec{e}_{r_0} \cdot \vec{n}^0 = 1.$$

Für den Grenzfall $r \rightarrow \infty$ dürfen wir also in der Kirchhoff'schen Formel (20.12)

$$r \rightarrow \infty \quad \Rightarrow \quad r_0 = r \quad \text{und} \quad \vec{e}_r \cdot \vec{n}^0 = \vec{e}_{r_0} \cdot \vec{n}^0 = 1 \quad (20.17)$$

setzen:

$$\begin{aligned}\mathcal{E}_p &= \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} \right) \cdot d\vec{S} \\ &= \frac{1}{4\pi} \oint_S \left[\frac{e^{ikr}}{r} \left(\frac{ik}{r_0} - \frac{1}{r_0^2} \right) e^{ikr_0} \cdot \vec{e}_{r_0} - \frac{e^{ikr_0}}{r_0} \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r \right] \cdot \vec{n}^0 dS\end{aligned}$$

\Rightarrow mit (20.17) \Rightarrow

$$\lim_{r \rightarrow \infty} \mathcal{E}_p = \frac{1}{4\pi} \oint_S \underbrace{\left[\frac{e^{ikr}}{r} \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} - \frac{e^{ikr}}{r} \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \right]}_{=0} dS = 0. \quad \square$$

Im Grenzfall $r \rightarrow \infty$ verschwinden Integrand und Integral, d. h., die im Unendlichen liegende Wellenfront einer Kugelwelle liefert keinen Feldstärkebeitrag im Beobachtungspunkt P .

20.6 Anhang: Fresnel- und Fraunhofer-Beugung

Hinweis!

Dieser Abschnitt, ebenfalls auf der Grundlage des Lehrbuchs der Theoretischen Physik von Joos, ist die direkte Fortsetzung der vorausgegangenen Abschnitte und eher von theoretischer als von praktischer Bedeutung. Er kann übersprungen werden, weil wir den gleichen Inhalt praxisnah aufbereitet im Kapitel 21 nocheinmal abhandeln.

Um es im Umgang mit dem Fresnel-Kirchhoff'schen Beugungsintegral in der Form (20.14) bequemer zu haben, nehmen wir im Folgenden an, dass der opake Schirm mit seiner lichtdurchlässigen Öffnung eben sei und in der (x, y) -Ebene eines kartesischen Koordinatensystems liege. Außerdem soll der Koordinatenursprung innerhalb der Öffnungsebene liegen, sodass die z -Achse so verläuft, wie in Abbildung 21 dargestellt.

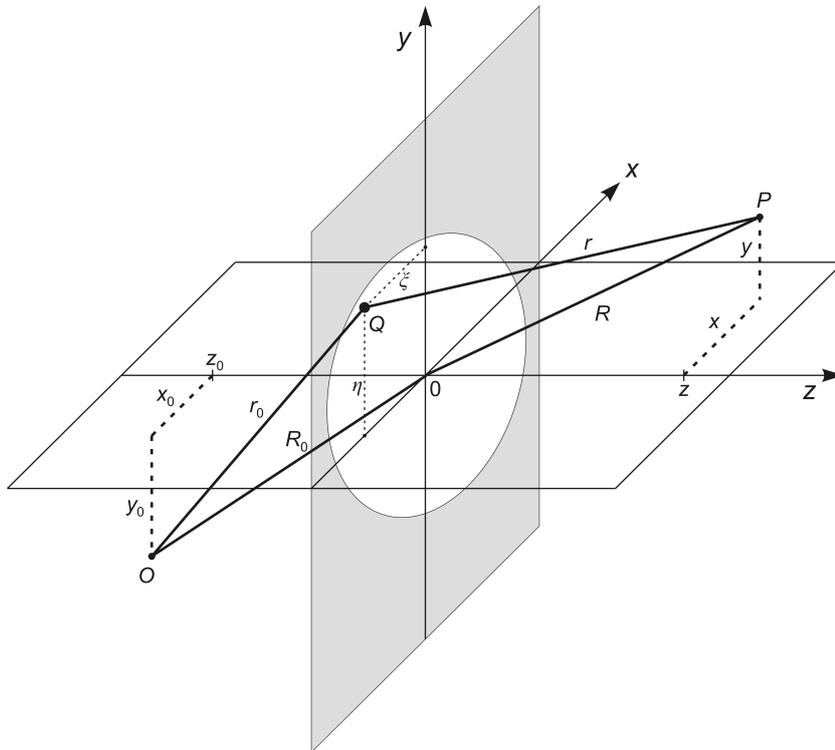


Abb. 21 Zur Beugung in einer ebenen Schirmöffnung beliebiger Gestalt. Nach Joos, Lehrbuch der Theoretischen Physik, Aula-Verlag, Wiesbaden, 1989, 15. Auflage, Seite 367.

Dadurch besitzen die Punkte O , P und Q die Koordinaten

$$O := (x_0, y_0, z_0), \quad P := (x, y, z), \quad Q := (\xi, \eta, 0),$$

sodass für die zugehörigen Abstände gilt

$$\begin{aligned} r^2 &= (x - \xi)^2 + (y - \eta)^2 + z^2, \\ r_0^2 &= (x_0 - \xi)^2 + (y_0 - \eta)^2 + z_0^2. \end{aligned} \tag{20.18}$$

Und die Abstände zum Koordinatenursprung, den wir hier kurz mit 0 bezeichnen, sind

$$R = \overline{P0} = \sqrt{x^2 + y^2 + z^2},$$

$$R_0 = \overline{O0} = \sqrt{x_0^2 + y_0^2 + z_0^2}.$$

Damit ist dann

$$K = \frac{1}{2} \left[\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0}) \right] = \frac{1}{2} \left(\frac{\sqrt{x^2 + y^2}}{R} + \frac{\sqrt{x_0^2 + y_0^2}}{R_0} \right)$$

und r und r_0 im Nenner können wir ebenfalls durch R und R_0 ersetzen, sodass aus (20.14) zunächst

$$\mathcal{E}_p = \frac{-i \epsilon_0}{\lambda} \frac{K}{R \cdot R_0} \int_{\text{Öffnung}} e^{ik(r+r_0)} \cdot dS \quad (20.19)$$

wird. In der Exponentialfunktion unter dem Integral hängen r und r_0 gemäß (20.18) von ξ und η ab. Wir werden deshalb in der folgenden Nebenrechnung r und r_0 bis zur 2. Ordnung nach ξ und η an der Stelle $\xi = \eta = 0$ entwickeln und anschließend diskutieren, was die sich daraus ergebenden Näherungen bedeuten.

Nebenrechnung: (Binomiale) Taylor-Entwicklung bzw. Binomialentwicklung von $r + r_0 = r(\xi, \eta) + r_0(\xi, \eta)$ an der Stelle $\xi = \eta = 0$ bis zur 2. Ordnung.

Es genügt, nur r zu entwickeln, weil die Entwicklung von r_0 völlig analog ist. Für $\xi = \eta = 0$ schreiben wir symbolisch $(0, 0)$.

0. Ordnung

$$r(0, 0) = \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]_{(0,0)}^{\frac{1}{2}} = (x^2 + y^2 + z^2)^{\frac{1}{2}} = R.$$

1. Ordnung

$$\begin{aligned} & \left. \frac{\partial}{\partial \xi} r(\xi, \eta) \right|_{(0,0)} \cdot \xi + \left. \frac{\partial}{\partial \eta} r(\xi, \eta) \right|_{(0,0)} \cdot \eta \\ &= \left\{ \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{1}{2}} \cdot 2(x - \xi) \cdot (-1) \right\}_{(0,0)} \cdot \xi \\ &+ \left\{ \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{1}{2}} \cdot 2(y - \eta) \cdot (-1) \right\}_{(0,0)} \cdot \eta \\ &= \frac{-x \xi}{R} + \frac{-y \eta}{R} = -\frac{x \xi + y \eta}{R}. \end{aligned}$$

2. Ordnung

$$\begin{aligned} & \frac{1}{2} \left(\left. \frac{\partial^2 r}{\partial \xi^2} \right|_{(0,0)} \cdot \xi^2 + 2 \left. \frac{\partial^2 r}{\partial \xi \partial \eta} \right|_{(0,0)} \cdot \xi \eta + \left. \frac{\partial^2 r}{\partial \eta^2} \right|_{(0,0)} \cdot \eta^2 \right) \\ &= \frac{1}{2} \left. \frac{\partial^2 r}{\partial \xi^2} \right|_{(0,0)} \cdot \xi^2 + \left. \frac{\partial^2 r}{\partial \xi \partial \eta} \right|_{(0,0)} \cdot \xi \eta + \frac{1}{2} \left. \frac{\partial^2 r}{\partial \eta^2} \right|_{(0,0)} \cdot \eta^2 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \left\{ \left[(x-\xi)^2 + (y-\eta)^2 + z^2 \right]^{-\frac{1}{2}} + (x-\xi) \frac{1}{2} \left[(x-\xi)^2 + (y-\eta)^2 + z^2 \right]^{-\frac{3}{2}} \cdot 2(x-\xi) \cdot (-1) \right\} \Big|_{(0,0)} \cdot \xi^2 \\
&\quad + (x-\xi) \frac{1}{2} \left[(x-\xi)^2 + (y-\eta)^2 + z^2 \right]^{-\frac{3}{2}} \cdot 2(y-\eta) \cdot (-1) \Big|_{(0,0)} \cdot \xi \eta \\
&+ \frac{1}{2} \left\{ \left[(x-\xi)^2 + (y-\eta)^2 + z^2 \right]^{-\frac{1}{2}} + (y-\eta) \frac{1}{2} \left[(x-\xi)^2 + (y-\eta)^2 + z^2 \right]^{-\frac{3}{2}} \cdot 2(y-\eta) \cdot (-1) \right\} \Big|_{(0,0)} \cdot \eta^2 \\
&= \frac{1}{2} \left\{ \frac{1}{R} - \frac{x^2}{R^3} \right\} \cdot \xi^2 - \frac{xy}{R^3} \cdot \xi \eta + \frac{1}{2} \left\{ \frac{1}{R} - \frac{y^2}{R^3} \right\} \cdot \eta^2 \\
&= \frac{1}{2} \frac{\xi^2 + \eta^2}{R} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3}.
\end{aligned}$$

Damit haben wir $r + r_0$ entwickelt:

$$\begin{aligned}
r + r_0 &= R - \frac{x\xi + y\eta}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} \\
&\quad + R_0 - \frac{x_0\xi + y_0\eta}{R_0} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3} + \dots
\end{aligned}$$

Abschließend reihen wir die Glieder entsprechend ihrer Ordnung aneinander:

$$\begin{aligned}
r + r_0 &= R + R_0 - \frac{x\xi + y\eta}{R} - \frac{x_0\xi + y_0\eta}{R_0} \\
&\quad + \frac{1}{2} \frac{\xi^2 + \eta^2}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3} + \dots,
\end{aligned}$$

$$\boxed{
\begin{aligned}
r + r_0 &= R + R_0 - \left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi - \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \\
&\quad + \frac{1}{2} \frac{\xi^2 + \eta^2}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3} + \dots
\end{aligned}
} \tag{20.20}$$

Im Folgenden schreiben wir (20.20) in einer kompakteren Form und benutzen das Gleichheitszeichen, obwohl wir auf die Fortsetzungspunkte am Ende der Entwicklung bis zur 2. Ordnung verzichten:

$$r + r_0 = R + R_0 + \phi(\xi, \eta) \tag{20.21}$$

mit

$$\begin{aligned}
\phi(\xi, \eta) &= \\
&\underbrace{- \left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi - \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta}_{1. \text{ Ordnung}} + \underbrace{\frac{1}{2} \frac{\xi^2 + \eta^2}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3}}_{2. \text{ Ordnung}}.
\end{aligned}$$

Setzen wir die Entwicklung (20.21) einschließlich der zweiten Ordnung in das bereits teilweise genäherte Fresnel-Kirchhoffsche Beugungsintegral (20.19) ein, so erhalten wir

das Integral für die **Fresnel-Beugung**

$$\mathcal{E}_p = \frac{-i\epsilon_0}{\lambda} \frac{K}{R \cdot R_0} e^{ik(R+R_0)} \int_{\text{Öffnung}} e^{ik\phi(\xi,\eta)} \cdot dS \quad (20.22)$$

und sprechen von der Nahfeldnäherung, mit der auch schon das Nahfeld bzw. das Fresnel-Gebiet hinter der Öffnung beschrieben werden kann.⁴⁰

Jetzt stellt sich die Frage, bis zu welchem Abstand R hinter der Öffnung das Nahfeld reicht? Anders gefragt, bis zu welchem *maximalen endlichen* Abstand R_{\max} hinter der Öffnung sind die Entwicklungsglieder 2. Ordnung bei gegebener Öffnung relevant?

Wir erreichen das gesuchte R_{\max} , wenn die Exponentialterme 2. Ordnung ausgehend von R im Unendlichen erstmalig gleich ± 1 werden, also bei

$$e^{\pm i n \cdot \pi} = \begin{cases} +1 & \text{für } n = 0, 2, 4, \dots, \\ -1 & \text{für } n = 1, 3, 5, \dots. \end{cases}$$

Schauen wir uns den ersten der beiden in Frage kommenden Exponentialterme

$$e^{ik \cdot \frac{1}{2} \frac{\xi^2 + \eta^2}{R}}, \quad (20.23)$$

$$e^{-ik \cdot \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3}} \quad (20.24)$$

an, so stellen wir fest, dass $n = 0 \Rightarrow n \cdot \pi = 0$ nicht in Frage kommen kann, weil dann R gegen Unendlich gehen müsste. Die (danach) nächstliegende Möglichkeit für ein maximales endliches R ist $n = 1 \Rightarrow n \cdot \pi = \pi$ im Exponentialterm (20.23), also

$$e^{ik \cdot \frac{1}{2} \frac{\xi^2 + \eta^2}{R}} \stackrel{!}{=} e^{i\pi} = -1.$$

Mit dem Maximum⁴¹

$$\hat{\rho} = \max\left(\sqrt{\xi^2 + \eta^2}\right) \Rightarrow \max(\xi^2 + \eta^2) = \hat{\rho}^2$$

der Öffnung bedeutet das

$$k \cdot \frac{1}{2} \frac{\hat{\rho}^2}{R} = \frac{2\pi}{\lambda} \cdot \frac{1}{2} \frac{\hat{\rho}^2}{R} = \frac{\pi}{\lambda} \frac{\hat{\rho}^2}{R} \stackrel{!}{=} \pi \Leftrightarrow \frac{\hat{\rho}^2}{\lambda \cdot R} \stackrel{!}{=} 1 \Leftrightarrow R_{\max} = \frac{\hat{\rho}^2}{\lambda}. \quad (20.25)$$

Es gilt also Faustregel⁴²

$$\boxed{\text{Fresnel-Gebiet } R \leq \frac{\hat{\rho}^2}{\lambda}}.$$

Schätzen wir abschließend noch den Exponenten von (20.24) für R_{\max} ab. Dazu vereinfachen wir $\hat{\rho}$, indem wir den Fall $\eta \equiv 0$ und $y \equiv 0$ betrachten, sodass $\hat{\rho}^2 = \xi^2$. Damit erhalten wir

$$k \cdot \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} = \frac{\pi}{\lambda} \frac{(x\xi + y\eta)^2}{R^3} \Rightarrow \frac{\pi}{\lambda} \frac{x^2 \hat{\rho}^2}{R^3} = \pi x^2 \frac{\lambda^2}{\hat{\rho}^4}.$$

⁴⁰Um das Gebiet *unmittelbar* hinter der Öffnung erfassen zu können, sind auch Entwicklungsglieder der Ordnung > 2 erforderlich. Die Beschreibung dieses Gebiets kann recht mühsam sein.

⁴¹ $\hat{\rho}$ ist der maximale „Radius“ der Öffnung, also eine feste Größe, und darf nicht verwechselt werden mit der Variablen ρ im Kapitel 22.

⁴²Mit der Apertur D findet man in der Literatur als Faustregel für das Nahfeld auch $L < \frac{D^2}{\lambda}$, wobei $L = z$ und $z \leq R$:

$$\hat{\rho} \stackrel{!}{=} \frac{1}{2} D \Rightarrow R \leq \frac{\hat{\rho}^2}{\lambda} \Rightarrow L \leq R \leq \frac{1}{4} \cdot \frac{D^2}{\lambda} \Rightarrow L < \frac{D^2}{\lambda}. \quad \square$$

Berücksichtigen wir jetzt, dass unter den o. g. Voraussetzungen

$$x = \sin \vartheta \cdot R_{\max} = \sin \vartheta \cdot \frac{\hat{\varrho}^2}{\lambda} \Rightarrow x^2 = \sin^2 \vartheta \cdot \frac{\hat{\varrho}^4}{\lambda^2}$$

gilt, so ist der Exponent bei R_{\max} schließlich

$$\pi \cdot \sin^2 \vartheta = \begin{cases} 0 & \text{für } \vartheta = 0, \\ \pi & \text{für } \vartheta = \frac{\pi}{2} \end{cases} \text{ entsprechend } x = R \text{ und } z = 0.$$

Zumindest bei kleinen Winkeln ϑ geht der Exponentialterm (20.24) gegen 1.

Die weitergehende Näherung im Exponenten, d. h. wenn wir auf die quadratischen Glieder in ξ und η verzichten und nur noch die konstanten und die linearen Glieder verwenden, ist die Fernfeldnäherung und ergibt die **Fraunhofer-Beugung**

$$\mathcal{E}_p = \frac{-i \epsilon_0}{\lambda} \frac{K}{R \cdot R_0} e^{ik(R+R_0)} \int_{\text{Öffnung}} e^{-ik \left[\left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi + \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \right]} \cdot dS, \quad (20.26)$$

mit der sich nur noch das Fernfeld oder Fraunhofer-Gebiet beschreiben lässt. Das Fernfeld beginnt nach einem Übergangsbereich etwa dort, wo der kleinere der beiden Abstände R und R_0 um etwa eine Größenordnung größer wird als $\hat{\varrho}^2/\lambda$:

$$\text{Fraunhofer-Gebiet} \quad \min \left(\{R_0, R\} \right) \gg \frac{\hat{\varrho}^2}{\lambda}, \quad \hat{\varrho} = \max \left(\sqrt{\xi^2 + \eta^2} \right).$$

Das Symbol \gg bedeutet „mindestens eine Größenordnung größer als“. Und salopp ausgedrückt gilt für die Fraunhofer-Beugung:

Je kleiner die Öffnung bei gegebener Wellenlänge bzw.
je größer die Wellenlänge bei gegebener Öffnung,
desto eher beginnt hinter der Öffnung das Fraunhofer-Gebiet.

21 Optische Abbildung in Fresnel- und Fraunhofer-Näherung

Im Folgenden verwenden wir die Konvention

» heißt: „mindestens um 1 Größenordnung bzw. um den Faktor 10 größer als“.

Auch in diesem Kapitel werden wir die optische Feldstärke in einem Beobachtungspunkt P in Fresnel- und in Fraunhofer-Näherung diskutieren. Während P im Abschnitt 20.6 aber als ein Punkt im radialen Abstand R vom Koordinatenursprung innerhalb der Schirmöffnung betrachtet wurde, definieren wir jetzt die Punkte P als Punkte in der (x, y) -**Beobachtungsebene**. Diese liegt im **Abstand** z planparallel zur Schirmebene mit der (ebenen) Öffnung bzw. Apertur Σ (siehe Abbildung 22).

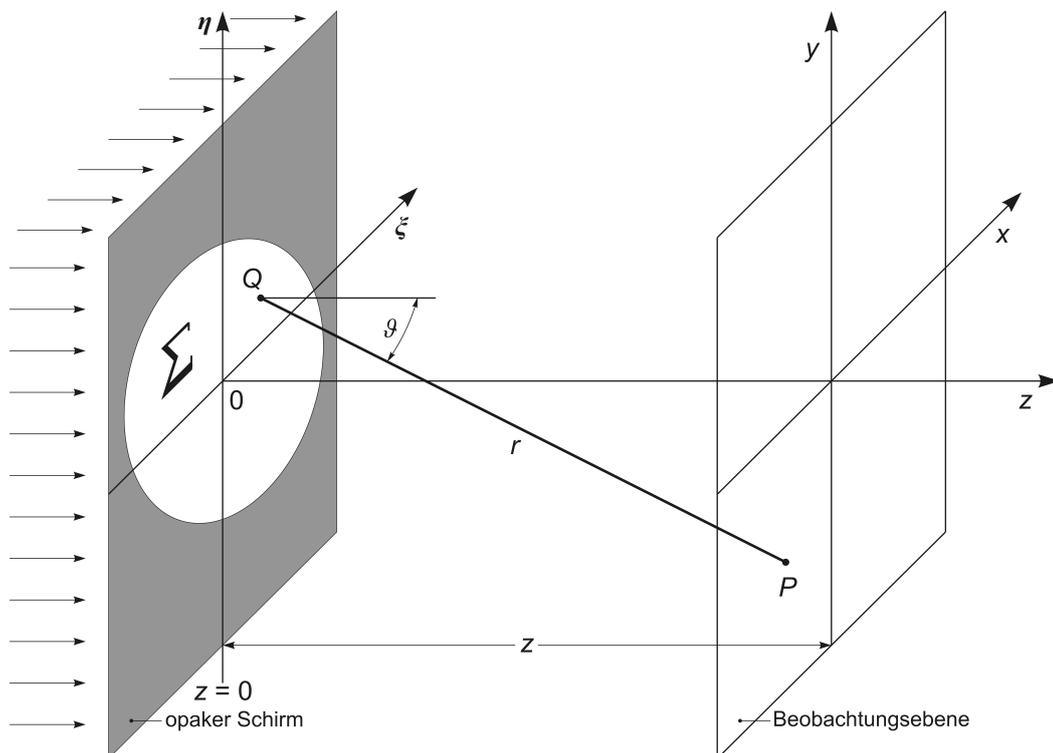


Abb. 22 Zur optischen Abbildung. Die Beobachtungsebene liegt im Abstand z planparallel zur Apertur Σ im lichtundurchlässigen Schirm. Σ ist also u. a. definiert durch $z = 0$.

Auch wenn der Ansatz in diesem Kapitel zugunsten von mehr Praxisnähe etwas abgewandelt wurde, so gehen wir trotzdem wieder aus vom Fresnel-Kirchhoff'schen Beugungsintegral (20.13)

$$\mathcal{E}_p = \frac{-i \epsilon_0}{\lambda} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\frac{\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0})}{2} \right] dS .$$

Allerdings werden wir diese Formel im Folgenden deutlich vereinfachen. Und auch die Fresnel- und die Fraunhofer-Näherung werden weitreichender ausfallen als im Abschnitt 20.6.

Für die optische Feldstärke \mathcal{E}_p schreiben wir $E(x, y, z)$ und wir gehen davon aus, dass die Feldstärkefunktion $u = u(\xi, \eta)$ in der Apertur bekannt ist. Das heißt, wir

„verstecken“ alle Größen, die sich auf die Quellenseite links vom Schirm beziehen, in der Funktion $u(\xi, \eta)$:

$$\epsilon_0 \cdot \frac{e^{ikr_0}}{r_0} \cdot \frac{-\cos(\vec{n}^0, \vec{e}_{r_0})}{2} \longrightarrow u(\xi, \eta).$$

Für den verbleibenden abbildungsseitigen Anteil des Neigungsfaktors K schreiben wir genähert

$$\frac{\cos(\vec{n}^0, \vec{e}_r)}{2} \longrightarrow \cos \vartheta = \frac{z}{r}.$$

Und weil das Flächenintegral über die Apertur Σ und somit über ξ und η läuft, ist das Fresnel-Kirchhoff'sche Beugungsintegral mit diesen „Anpassungen“

$$E(x, y, z) = \frac{1}{i\lambda} \iint_{\Sigma} u(\xi, \eta) \frac{z}{r} \frac{e^{ikr}}{r} d\xi d\eta.$$

21.1 Fresnel-Näherung

Jetzt erfolgt eine wichtige Näherung, die Näherung von r *im Nenner*:

$$z \gg \text{laterale Abmessungen } \xi, \eta, x, y \Rightarrow \Sigma \text{ und } \vartheta \text{ vergleichsweise klein} \Rightarrow \begin{cases} r \approx z, \\ \frac{z}{r} \approx 1. \end{cases}$$

Damit erhält das Fresnel-Kirchhoff'sche Beugungsintegral für unsere Anwendung die einfache Form

$$E(x, y, z) = \frac{1}{i\lambda \cdot z} \iint_{\Sigma} u(\xi, \eta) e^{ikr(\xi, \eta)} d\xi d\eta. \quad (21.1)$$

Im Exponenten von $e^{ikr(\xi, \eta)}$ durften wir r nicht einfach durch z ersetzen, weil die Wellenzahl k allgemein groß ist im Vergleich zu r und deshalb kleine Änderungen von r große Veränderungen des Exponenten bewirken:

$$k = \frac{2\pi}{\lambda} \text{ mit } \lambda \approx \frac{1}{2} \cdot 10^{-6} \text{ m} \Rightarrow k \approx 10^5 \text{ m}^{-1}.$$

Wenn wir r *im Exponenten* durch z nähern wollen, dann mit Hilfe der Taylorentwicklung. Wir drücken also zunächst r in den Koordinaten ξ, η, x, y, z aus:

$$r = \sqrt{z^2 + (\xi - x)^2 + (\eta - y)^2} = z \sqrt{1 + \left(\frac{\xi - x}{z}\right)^2 + \left(\frac{\eta - y}{z}\right)^2} = z \cdot \sqrt{1 + q}$$

mit

$$q = \left(\frac{\xi - x}{z}\right)^2 + \left(\frac{\eta - y}{z}\right)^2 = \frac{(\xi - x)^2 + (\eta - y)^2}{z^2}. \quad (21.2)$$

Wegen

$$z \gg \text{laterale Abmessungen von } \xi, \eta, x, y \Rightarrow 0 < \left(\frac{\xi - x}{z}\right)^2 + \left(\frac{\eta - y}{z}\right)^2 = q \ll 1$$

entwickeln wir

$$f(q) = \sqrt{1+q} = (1+q)^{\frac{1}{2}}$$

für kleine q bzw. an der Stelle $q = 0$.

Nebenrechnung: Taylorentwicklung bis zur 2. Ordnung von $f(q) = (1+q)^{\frac{1}{2}}$ an der Stelle $q = 0$.

$$\begin{aligned} f(q) &= f(q)\Big|_{q=0} + f'(q)\Big|_{q=0} \cdot q + \frac{1}{2} \cdot f''(q)\Big|_{q=0} \cdot q^2 + \dots \\ &= (1+q)^{\frac{1}{2}}\Big|_{q=0} + \frac{1}{2} (1+q)^{-\frac{1}{2}}\Big|_{q=0} \cdot q + \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \frac{1}{2} (1+q)^{-\frac{3}{2}}\Big|_{q=0} \cdot q^2 \pm \dots, \end{aligned}$$

$$\boxed{f(q) = 1 + \frac{1}{2}q - \frac{1}{8}q^2 \pm \dots}$$

Wir vernachlässigen im Folgenden die Entwicklungsglieder ab der 2. Ordnung und verwenden folglich im Exponenten von e^{ikr} die Näherung

$$r \approx z \cdot \left[1 + \frac{1}{2}q\right] \Rightarrow \quad (21.3)$$

$$r \approx z \cdot \left[1 + \frac{1}{2} \left(\frac{\xi-x}{z}\right)^2 + \frac{1}{2} \left(\frac{\eta-y}{z}\right)^2\right] = z + \frac{(\xi-x)^2}{2z} + \frac{(\eta-y)^2}{2z}.$$

Das Fresnel-Kirchhoff'sche Beugungsintegral (21.1) ist damit

$$\begin{aligned} E(x, y, z) &= \frac{1}{i\lambda \cdot z} \iint_{\Sigma} u(\xi, \eta) e^{ik[z + (\xi-x)^2/2z + (\eta-y)^2/2z]} d\xi d\eta \\ &= \frac{e^{ikz}}{i\lambda \cdot z} \iint_{\Sigma} u(\xi, \eta) e^{ik[(\xi-x)^2/2z + (\eta-y)^2/2z]} d\xi d\eta \end{aligned}$$

und mit $ik \cdot \frac{1}{2z} = i \frac{2\pi}{\lambda} \cdot \frac{1}{2z} = \frac{i\pi}{\lambda z}$

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \iint_{\Sigma} u(\xi, \eta) e^{\frac{i\pi}{\lambda z} [(\xi-x)^2 + (\eta-y)^2]} d\xi d\eta.$$

Quadrieren der Binome im Exponenten, also

$$\begin{aligned} \frac{i\pi}{\lambda z} [(\xi-x)^2 + (\eta-y)^2] &= \frac{i\pi}{\lambda z} [\xi^2 - 2\xi x + x^2 + \eta^2 - 2\eta y + y^2] \\ &= \frac{i\pi}{\lambda z} (\xi^2 + \eta^2) + \frac{i\pi}{\lambda z} (x^2 + y^2) - \frac{i \cdot 2\pi}{z \cdot \lambda} (x\xi + y\eta), \end{aligned}$$

liefert schließlich mit $2\pi/\lambda = k$

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \iint_{\Sigma} u(\xi, \eta) e^{\frac{i\pi}{\lambda z} (\xi^2 + \eta^2) + \frac{i\pi}{\lambda z} (x^2 + y^2) - \frac{ik}{z} (x\xi + y\eta)} d\xi d\eta$$

und damit die **Fresnel-Näherung**

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2+y^2)} \iint_{\Sigma} u(\xi, \eta) \cdot e^{-\frac{ik}{z}(x\xi+y\eta)} \cdot e^{\frac{i\pi}{\lambda z}(\xi^2+\eta^2)} d\xi d\eta \quad (21.4)$$

Zur Bedeutung der Terme⁴³ in (21.4):

$$\frac{e^{ikz}}{i\lambda z} \equiv \text{const} \quad \text{Intensitätsfaktor mit Phasenkonstante } \frac{1}{i} = -i = e^{-i\frac{\pi}{2}},$$

$$e^{\frac{i\pi}{\lambda z}(x^2+y^2)} = g(x, y, z) \quad \text{Phasenfaktor},$$

$$e^{\frac{i\pi}{\lambda z}(\xi^2+\eta^2)} = f(\xi, \eta, z) \quad \text{Chirp-Funktion, engl. to chirp - zwitschern, zirpen, wird in der Fraunhofer-Näherung vernachlässigt}.$$

Fourier-Darstellung von (21.4) mit

$$\frac{x}{\lambda z} = \mu, \quad \frac{y}{\lambda z} = \nu, \quad \frac{e^{ikz}}{\lambda z} \cdot g(x, y, z) = \frac{e^{ikz}}{\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2+y^2)} = \frac{e^{ikz}}{\lambda z} \cdot e^{i\pi\lambda z(\mu^2+\nu^2)} = \psi(\mu, \nu, z) :$$

$$\begin{aligned} E(\mu, \nu, z) &= \frac{1}{i} \psi(\mu, \nu, z) \iint u(\xi, \eta) \cdot f(\xi, \eta, z) \cdot e^{i2\pi(\mu\xi+\nu\eta)} d\xi d\eta \\ &= \frac{1}{i} \psi(\mu, \nu, z) \cdot \mathcal{F}\{u(\xi, \eta) \cdot \text{chirp}(\xi, \eta)\} \end{aligned}$$

$$E(\mu, \nu, z) = \frac{1}{i} \psi(\mu, \nu, z) \cdot \mathcal{F}\{u(\xi, \eta)\} \otimes \mathcal{F}\{\text{chirp}(\xi, \eta)\}.$$

21.2 Bedingung für die Fresnel-Näherung

Mit (21.2) und (21.3) verwenden wir in der Fresnel-Näherung

$$kr \approx kz \cdot \left[1 + \frac{1}{2}q\right] = kz \cdot \left[1 + \frac{1}{2} \frac{(\xi-x)^2 + (\eta-y)^2}{z^2}\right] \quad (21.5)$$

unter Vernachlässigung der Entwicklungsglieder ab der 2. Ordnung, d. h. unter der Bedingung

$$kz \cdot \frac{1}{8} q^2 = kz \cdot \frac{1}{8} \left[\frac{(\xi-x)^2 + (\eta-y)^2}{z^2} \right]^2 \ll 1.$$

Ermitteln wir also die Werte von z , welche diese Bedingung erfüllen:

$$1 \gg \frac{2\pi}{\lambda} z \cdot \frac{1}{8} \frac{[(\xi-x)^2 + (\eta-y)^2]^2}{z^4} = \frac{\pi}{4\lambda} \frac{[(\xi-x)^2 + (\eta-y)^2]^2}{z^3} \Leftrightarrow$$

⁴³Verallgemeinert dargestellt ist die Chirp-Funktion $f(x) = e^{i\pi\kappa x^2}$.

Damit sind hier $x = \xi$ bzw. $x = \eta$ und $\kappa = \frac{1}{\lambda z}$. Die Chirp-Funktion ist eine reine Phasenfunktion mit konstantem Betrag $|f(x)| \equiv 1$.

$$\text{Fresnel-Bedingung} \quad z \gg \sqrt[3]{\frac{\pi}{4\lambda} [(\xi - x)^2 + (\eta - y)^2]^2}.$$

Das Fresnel-Gebiet oder Nahfeld erstreckt sich folglich über den z -Bereich, der die Fresnel-Bedingung erfüllt, und geht dann bei sehr großen z über in das Fernfeld.

Beispiel: $(\xi - x) = (\eta - y) = 1 \text{ cm} = 10^{-2} \text{ m}$, $\lambda = 500 \text{ nm} = 5 \cdot 10^{-7} \text{ m}$,

$$z \gg \sqrt[3]{\frac{\pi \cdot [(10^{-2})^2 + (10^{-2})^2]^2 \text{ m}^4}{4 \cdot 5 \cdot 10^{-7} \text{ m}}} = \sqrt[3]{\frac{\pi \cdot 4 \cdot 10^{-8} \text{ m}^4}{20 \cdot 10^{-7} \text{ m}}} = \sqrt[3]{\frac{\pi}{50} \text{ m}^3},$$

$$z \gg 0,40 \text{ m} = 40 \text{ cm}.$$

21.3 Fraunhofer-Näherung

Ausgehend von der Fresnel-Näherung (21.4) vernachlässigen wir bei der Fraunhofer-Näherung auch noch den bezüglich ξ und η quadratischen Term im Exponenten, d. h. die Chirp-Funktion $e^{\frac{i\pi}{\lambda z}(\xi^2 + \eta^2)}$ wird gleich 1 gesetzt. So erhalten wir die

Fraunhofer-Näherung

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2 + y^2)} \iint_{\Sigma} u(\xi, \eta) \cdot e^{-\frac{ik}{z}(x\xi + y\eta)} d\xi d\eta. \quad (21.6)$$

Fourier-Darstellung von (21.6) mit

$$\frac{x}{\lambda z} = \mu, \quad \frac{y}{\lambda z} = \nu, \quad \frac{e^{ikz}}{\lambda z} \cdot g(x, y, z) = \frac{e^{ikz}}{\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2 + y^2)} = \frac{e^{ikz}}{\lambda z} \cdot e^{i\pi\lambda z(\mu^2 + \nu^2)} = \psi(\mu, \nu, z) :$$

$$E(\mu, \nu, z) = \frac{1}{i} \psi(\mu, \nu, z) \iint u(\xi, \eta) \cdot e^{i2\pi(\mu\xi + \nu\eta)} d\xi d\eta$$

$$E(\mu, \nu, z) = \frac{1}{i} \psi(\mu, \nu, z) \cdot \mathcal{F}\{u(\xi, \eta)\}.$$

21.4 Bedingung für die Fraunhofer-Näherung

Ausgehend von (21.5), d. h. von kr in Fresnel-Näherung gemäß

$$\begin{aligned} kr &\approx kz \cdot \left[1 + \frac{1}{2} q \right] = kz \cdot \left[1 + \frac{1}{2} \frac{(\xi - x)^2 + (\eta - y)^2}{z^2} \right], \\ &\approx kz \cdot \left[1 + \frac{1}{2} \frac{\xi^2 - 2\xi x + x^2 + \eta^2 - 2\eta y + y^2}{z^2} \right], \\ kr &\approx kz + k \frac{x^2 + y^2}{2z} - k \frac{\xi x + \eta y}{z} + \underbrace{k \frac{\xi^2 + \eta^2}{2z}}, \end{aligned}$$

wird in Fraunhofer-Näherung der bezüglich ξ und η quadratische (unterklammerte) Term vernachlässigt, d. h.

$$k \frac{\xi^2 + \eta^2}{2z} \rightarrow 0.$$

Ermitteln wir also die Werte von z , welche diese Bedingung erfüllen:

$$1 \gg k \frac{\xi^2 + \eta^2}{2z} = \frac{\pi}{\lambda} \frac{\xi^2 + \eta^2}{z} \Leftrightarrow$$

Fraunhofer-Bedingung $z \gg \frac{\pi}{\lambda} (\xi^2 + \eta^2)$

Beispiel: $(\xi - x) = (\eta - y) = 1 \text{ cm} = 10^{-2} \text{ m}$, $\lambda = 500 \text{ nm} = 5 \cdot 10^{-7} \text{ m}$,

$$z \gg \frac{\pi \cdot [(10^{-2})^2 + (10^{-2})^2] \text{ m}^2}{5 \cdot 10^{-7} \text{ m}} = \frac{\pi \cdot 2 \cdot 10^{-4} \text{ m}^2}{5 \cdot 10^{-7} \text{ m}},$$

$$z \gg 1,257 \cdot 10^3 \text{ m} = 1257 \text{ m}.$$

Wie man sieht, beginnt das Fraunhofer-Gebiet oder Fernfeld erst bei viel größeren Abständen z von der Apertur als das Nahfeld.

22 Fraunhofer-Beugung an kreisrunder Öffnung – Auflösungsvermögen

Nach Eugene Hecht, *Optik*, Oldenbourg-Verlag, München, Wien, 4. Auflage, 2002, Abschnitt 10.2.5 *Beugung an einer kreisrunden Öffnung* und Abschnitt 10.2.6 *Das Auflösungsvermögen abbildender Systeme*, Seite 752 bis Seite 762.

Weil die optischen Linsen und folglich die Öffnungen der meisten optischen Instrumente kreisrund sind, werden wir in diesem Abschnitt das Auflösungsvermögen optischer Instrumente mit kreisrunder Öffnung berechnen. Grundlage unserer Überlegungen ist neben der Abbildung 8, bei der man sich die Linse L_2 in der Öffnungsebene Σ positioniert vorstellen sollte, insbesondere die folgende Abbildung 23:

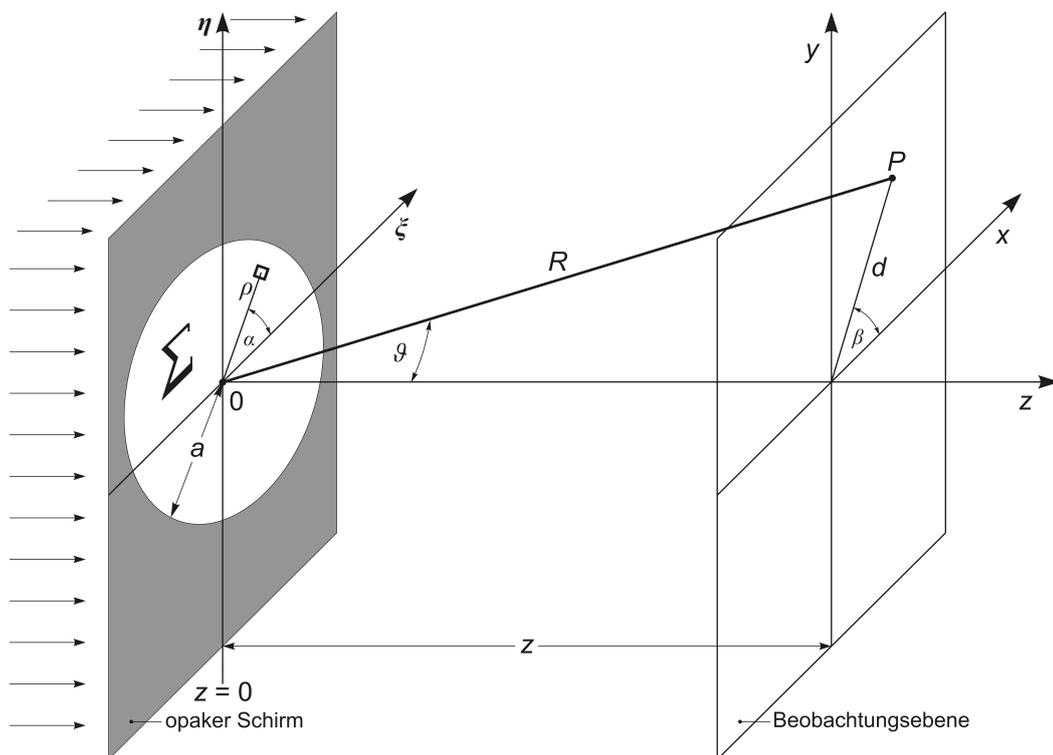


Abb. 23 Zur Fraunhofer-Beugung an einer kreisrunden Öffnung bzw. Apertur Σ mit dem Durchmesser $D = 2 \cdot a$.

Eine **ebene, monochromatische Welle** treffe von links so auf die Öffnung Σ , dass ihr Wellenvektor \vec{k} in Richtung der z -Achse verläuft.

„ Von der Lichtwelle, die auf Σ trifft, wird auf jeden Fall nur ein kreisförmiger Ausschnitt verwendet – jener, der sich durch die Linse L_2 ausbreitet –, um in der Brennebene das Bild zu erzeugen. Nichts anderes passiert offensichtlich im Auge, einem Teleskop, Mikroskop oder einer Kameraoptik. Das Bild einer fernen Punktquelle, erzeugt mit einer ideal aberrationsfreien Sammellinse, ist daher nie ein Punkt, sondern stets eine Art Beugungsbild. Da wir immer nur einen Teil der auftreffenden Wellenfront verwenden, können wir nicht erwarten, ein ideales Bild zu erhalten.“⁴⁴

⁴⁴Zitiert aus Eugene Hecht, *Optik*, Oldenbourg-Verlag, München, Wien, 4. Auflage, 2002, Seite 752.

Die Öffnung sei gleichmäßig ausgeleuchtet, d. h., für die optische Feldstärke in der Ebene Σ gelte

$$u(\xi, \eta) \equiv \text{const} .$$

Wir gehen aus von der im Abschnitt 20.6 hergeleiteten Formel (20.26)

$$\begin{aligned} \mathcal{E}_p &= \frac{-i\epsilon_0}{\lambda} \frac{K}{R \cdot R_0} e^{ik(R+R_0)} \int_{\text{Öffnung}} e^{-ik \left[\left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi + \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \right]} \cdot dS \\ &= \frac{1}{i\lambda} \cdot \frac{\epsilon_0 e^{ikR_0}}{R_0} \cdot K \cdot \frac{e^{ikR}}{R} \int_{\text{Öffnung}} e^{-ik \left[\left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi + \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \right]} \cdot dS \end{aligned} \quad (22.1)$$

zur Beschreibung der Fraunhofer-Beugung, müssen diese aber an die hier vorliegenden speziellen Verhältnisse anpassen. Weil wir nur jene gebeugten Strahlen berücksichtigen wollen, die nur gering von der Richtung parallel zur z -Achse abweichen, können wir $K = 1$ setzen. Bei einer ebenen einfallenden Welle gilt $R_0 \rightarrow \infty$, sodass die Quotienten $\frac{x_0}{R_0}$ und $\frac{y_0}{R_0}$ im Exponenten unter dem Integral verschwinden. Der Term

$$\frac{\epsilon_0 e^{ikR_0}}{R_0}$$

im Vorfaktor beschreibt die Feldstärke in der Öffnung für den Fall einer Kugelwelle, die ihre Punktquelle im endlichen Abstand R_0 vor der Öffnung besitzt. Da wir jetzt aber von einer einfallenden ebenen Welle ausgehen, können wir diesen Term durch $\text{const} \equiv u$ ersetzen. Und weil der Zeitanteil $e^{-i\omega t}$ für unsere Betrachtungen nicht relevant ist, vernachlässigen wir ihn im Folgenden wieder. Damit erhält (22.1) die Form

$$E(x, y, z) = \underbrace{\frac{1}{i\lambda} \cdot u}_{\epsilon_A \equiv \text{const}} \cdot \frac{e^{ikR}}{R} \int_{\text{Öffnung}} e^{-ik \left[\frac{x}{R} \xi + \frac{y}{R} \eta \right]} \cdot dS .$$

Aus Bequemlichkeit fassen wir die beiden konstanten Vorfaktoren $\frac{1}{i\lambda}$ und u zu ϵ_A zusammen, sodass kurz

$$E(x, y, z) = \frac{\epsilon_A e^{ikR}}{R} \int_{\text{Öffnung}} e^{-i \frac{k}{R} (\xi x + \eta y)} \cdot dS$$

resultiert. Weil die einfallende ebene Welle senkrecht auf die kreisrunde Öffnung Σ trifft, handelt es sich hier um ein axialsymmetrisches Problem (bezüglich der z -Achse). Es ist deshalb sinnvoll, auf Σ und der Beobachtungsebene σ (ebene) Polarkoordinaten einzuführen:

$$\begin{aligned} \Sigma : \quad \xi &= \varrho \cdot \cos \alpha , \quad \eta = \varrho \cdot \sin \alpha , \\ \sigma : \quad x &= d \cdot \cos \beta , \quad y = d \cdot \sin \beta , \\ dS &= d\xi d\eta = \varrho d\varrho d\alpha . \end{aligned}$$

Mit dem Additionstheorem

$$\cos \alpha \cdot \cos \beta \pm \sin \alpha \cdot \sin \beta = \cos(\alpha \mp \beta)$$

ergibt das

$$\xi x + \eta y = \varrho \cos \alpha \cdot d \cos \beta + \varrho \sin \alpha \cdot d \sin \beta = \varrho d \cdot \cos(\alpha - \beta)$$

und damit schließlich das Flächenintegral über die Öffnung Σ in (ebenen) Polarkoordinaten:

$$E(d, \beta, z) = \frac{\epsilon_A e^{ikR}}{R} \int_{\varrho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k}{R} \varrho d \cdot \cos(\alpha - \beta)} \cdot \varrho d\varrho d\alpha .$$

E ist aber wegen der Axialsymmetrie unabhängig von β , d. h., für jeden festen Radius d besitzt E unabhängig von β stets den gleichen Wert und erzeugt somit in der Beobachtungsebene ein um die z -Achse kreisförmig angeordnetes Beugungsmuster. Wir können deshalb vereinfachend und o.B.d.A. $\beta = 0$ setzen:

$$E(d, \beta = 0, z) = \frac{\epsilon_A e^{ikR}}{R} \int_{\varrho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k\varrho d}{R} \cdot \cos \alpha} \cdot \varrho d\varrho d\alpha . \quad (22.2)$$

Die Berechnung dieses Flächenintegrals erfordert den Umgang mit der Bessel-Funktion. Wir lösen es deshalb in der folgenden Nebenrechnung.

Bessel-Funktion

Die **Integraldarstellung** der Bessel-Funktion J erster Gattung und der Ordnung n (n ganzzahlig, d. h. $n \in \mathbb{Z}$) ist

$$J_n(w) = \frac{1}{2\pi \cdot i^n} \int_0^{2\pi} e^{i(n\alpha + w \cdot \cos \alpha)} d\alpha$$

mit dem Argument w und der **Rekursionsformel**

$$\begin{aligned} \frac{d}{dw} [w \cdot J_1(w)] &= w \cdot J_0(w) , \\ \frac{d}{dw} [w^n \cdot J_n(w)] &= w^n \cdot J_{n-1}(w) \\ \Rightarrow w^n \cdot J_n(w) &= \int_0^w w^n \cdot J_{n-1}(w) dw \Rightarrow \end{aligned}$$

$$J_n(w) = \frac{1}{w^n} \int_0^w w^n \cdot J_{n-1}(w) dw .$$

Das heißt

$$J_1(w) = \frac{1}{w} \int_0^w w \cdot J_0(w) dw .$$

Weiterhin gilt

$$\begin{aligned}
 J_0(w) &= \frac{1}{2\pi} \int_0^{2\pi} e^{i w \cdot \cos \alpha} d\alpha, & J_0(0) &= 1, \\
 J_1(w) &= \frac{1}{2\pi i} \int_0^{2\pi} e^{i(\alpha + w \cdot \cos \alpha)} d\alpha, & J_1(0) &= 0.
 \end{aligned}$$

Nullstellen von $J_1(w)$, also

$$J_1(w) = 0 \quad \text{bei} \quad \begin{cases} w_0 = 0 \\ w_1 = 3,832 \\ w_2 = 7,016 \\ \vdots \end{cases}$$

Nebenrechnung: Lösung des Flächenintegrals (22.2)

$$E(d, \beta = 0, z) = \frac{\epsilon_A e^{ikR}}{R} \int_{\varrho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k \varrho d}{R} \cdot \cos \alpha} \cdot \varrho d\varrho d\alpha.$$

Wir verwenden die Substitution

$$-\frac{k \varrho d}{R} = w \quad \Leftrightarrow \quad \varrho = -\frac{R}{kd} w \quad \Rightarrow \quad d\varrho = -\frac{R}{kd} dw$$

mit den daraus folgenden Integralgrenzen

$$\varrho = 0 \quad \Rightarrow \quad w = 0 \quad \text{und} \quad \varrho = a \quad \Rightarrow \quad w = -\frac{kad}{R}$$

und lassen den Vorfaktor des Flächenintegrals zunächst unberücksichtigt. Der Winkelanteil des Integrals ist damit

$$\begin{aligned}
 \int_0^{2\pi} e^{i(-k \varrho d/R) \cdot \cos \alpha} d\alpha &= \int_0^{2\pi} e^{i w \cdot \cos \alpha} d\alpha = 2\pi \left(\frac{1}{2\pi} \int_0^{2\pi} e^{i w \cdot \cos \alpha} d\alpha \right) \\
 &= 2\pi \cdot J_0(w).
 \end{aligned}$$

Vom Flächenintegral verbleibt folglich das Integral längs des Radius ϱ :

$$\begin{aligned}
 \int_{\varrho=0}^a 2\pi \cdot J_0(w) \cdot \varrho \cdot d\varrho &= 2\pi \int_{w=0}^{-kad/R} J_0(w) \cdot \left(-\frac{R}{kd} w\right) \cdot \left(-\frac{R}{kd} dw\right) \\
 &= 2\pi \left(\frac{R}{kd}\right)^2 \int_{w=0}^{-kad/R} J_0(w) \cdot w \cdot dw \\
 &= 2\pi \left(\frac{R}{kd}\right)^2 \left[w \cdot J_1(w) \right]_{w=0}^{-kad/R}.
 \end{aligned}$$

Da $J_1(w)$ eine ungerade Funktion ist, gilt

$$w = -\frac{kad}{R} \quad \Rightarrow \quad J_1(-kad/R) = -J_1(kad/R) \quad \Rightarrow \quad \left[w \cdot J_1(w) \right]_{w=-kad/R} = \frac{kad}{R} J_1(kad/R),$$

sodass

$$\begin{aligned} \int_{\varrho=0}^a 2\pi \cdot J_0(w) \cdot \varrho \cdot d\varrho &= 2\pi \left(\frac{R}{kd}\right)^2 \cdot \frac{kad}{R} J_1(kad/R), \\ &= 2\pi a^2 \cdot \frac{R}{kad} \cdot J_1(kad/R). \end{aligned}$$

Mit dem Flächeninhalt

$$\pi a^2 = A$$

der Öffnung Σ ist dann das Flächenintegral insgesamt, also auch mit dem Vorfaktor,

$$\frac{\epsilon_A e^{ikR}}{R} \int_{\varrho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k\varrho d}{R} \cdot \cos \alpha} \cdot \varrho d\varrho d\alpha = \frac{\epsilon_A e^{ikR}}{R} \cdot 2 \cdot A \cdot \frac{R}{kad} \cdot J_1(kad/R). \quad (22.3)$$

Zur besseren Übersicht in den folgenden Rechnungen ordnen wir die Terme in (22.3) um und schreiben für die Lösung des Flächenintegrals (22.2) schließlich

$$E(d, z) = \frac{2\epsilon_A}{R} e^{ikR} \cdot A \cdot \frac{J_1(kad/R)}{kad/R}.$$

Wegen der Achsensymmetrie des Beugungsmusters, das im Fall der kreisrunden Öffnung auch als **Airy-Muster**⁴⁵ bezeichnet wird, ist die Bestrahlungsstärke bzw. der zeitliche Mittelwert der Intensität *in der Beobachtungsebene* auf folgende Weise abhängig vom Radius d :

$$\bar{I}(d) = \left\langle (\operatorname{Re}\{E\})^2 \right\rangle_t = \frac{1}{2} \cdot E \cdot E^* = \frac{1}{2} \cdot \frac{4\epsilon_A^2 A^2}{R^2} \left[\frac{J_1(kad/R)}{kad/R} \right]^2,$$

$$\bar{I}(d) = \frac{2\epsilon_A^2 A^2}{R^2} \left[\frac{J_1(kad/R)}{kad/R} \right]^2. \quad (22.4)$$

Daraus und mit dem Ergebnis (22.6) aus der folgenden Nebenrechnung erhalten wir die Bestrahlungsstärke im Zentrum des Beugungsmusters, also für $d = 0 \Rightarrow w = 0$, wie folgt:

Nebenrechnung

$$\frac{J_1(0)}{0} = \frac{0}{0} \text{ nicht definiert} \Rightarrow \text{mit L'Hospital :}$$

$$\lim_{w \rightarrow 0} \frac{J_1(w)}{w} = \lim_{w \rightarrow 0} \frac{\frac{d}{dw} J_1(w)}{\frac{d}{dw} w} = \lim_{w \rightarrow 0} \frac{\frac{d}{dw} J_1(w)}{1},$$

$$\lim_{w \rightarrow 0} \frac{J_1(w)}{w} = \lim_{w \rightarrow 0} \frac{d}{dw} J_1(w). \quad (22.5)$$

⁴⁵Sir George Bidell Airy (1801–1892): königlich-englischer Astronom.

$$\begin{aligned}
\frac{d}{dw} [w \cdot J_1(w)] &= w \cdot J_0(w) = w \cdot \frac{d}{dw} J_1(w) + J_1(w) \\
\Leftrightarrow J_0(w) &= \frac{d}{dw} J_1(w) + \frac{J_1(w)}{w} \\
\Rightarrow J_0(0) = 1 &= \lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) + \lim_{w \rightarrow 0} \frac{J_1(w)}{w} \\
&\stackrel{(22.5)}{\Rightarrow} 1 = 2 \cdot \lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) \Leftrightarrow
\end{aligned}$$

$$\boxed{\lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) = \frac{1}{2}}. \quad (22.6)$$

$$\bar{I}(d=0) = \bar{I}_0 = \frac{2 \epsilon_A^2 A^2}{R^2} \cdot \lim_{d \rightarrow 0} \left[\frac{J_1(kad/R)}{kad/R} \right]^2 \stackrel{(22.6)}{=} \frac{2 \epsilon_A^2 A^2}{R^2} \cdot \left(\frac{1}{2} \right)^2,$$

$$\boxed{\bar{I}_0 = \frac{\epsilon_A^2 A^2}{2 R^2}}.$$

Für einige praktische Anwendungen ist es nützlich, wenn die Bestrahlungsstärke des Beugungsmusters als „Vielfaches“ von \bar{I}_0 und als Funktion des Winkels ϑ (siehe Abbildung 23) angegeben wird. Dazu klammern wir \bar{I}_0 aus (22.4) aus und führen die Substitution

$$\frac{d}{R} = \sin \vartheta$$

durch:

$$\bar{I}(d) = \frac{\epsilon_A^2 A^2}{2 R^2} \cdot 4 \left[\frac{J_1(kad/R)}{kad/R} \right]^2 = \bar{I}_0 \cdot \left[2 \cdot \frac{J_1(ka \frac{d}{R})}{ka \frac{d}{R}} \right]^2,$$

$$\boxed{\bar{I}(\vartheta) = \bar{I}_0 \cdot \left[\frac{2 \cdot J_1(ka \sin \vartheta)}{ka \sin \vartheta} \right]^2}. \quad (22.7)$$

Das Airy-Muster zeichnet sich durch einen kreisrunden Bereich hoher Intensität um das (zentrale) Hauptmaximum aus. Dieser Bereich ist die **Airy-Scheibe** und wird begrenzt durch das erste Minimum im Airy-Muster.

22.1 Bestimmung des Auflösungsvermögens bei kreisrunder Öffnung

Eine allgemeine Übersicht zur Thematik „Auflösungsvermögen“ findet man im Internet unter: Spektrum.de – Auflösungsvermögen – Lexikon der Optik

Die Begriffe Auflösungsvermögen und Auflösung werden synonym verwendet. Bei der Bestimmung des Auflösungsvermögens optischer Instrumente stützt man sich meistens auf das **Rayleigh-Kriterium**⁴⁶ (1874):

Zwei Punktquellen werden dann gerade noch getrennt wahrgenommen („aufgelöst“), wenn das Hauptmaximum des Fraunhofer-Beugungsmusters der einen Quelle und das erste Minimum des Fraunhofer-Beugungsmusters der anderen Quelle zusammenfallen. Man spricht in diesem Zusammenhang folglich vom **Doppelpunktauflösungsvermögen**.

Der sich daraus ergebende **auflösbare Minimal- oder Mindestabstand** der beiden Quellen *entspricht* bei kreisrunder Apertur dem Abstand des ersten Minimums vom Zentrum des Airy-Musters, also dem **Radius der Airy-Scheibe**. Es ist meistens sinnvoll, nicht mit dem Abstand bzw. der Strecke zwischen den Quellen zu rechnen, sondern mit dem Winkel θ , unter dem die beiden Punktquellen beobachtet werden.

Wir brauchen in unserem Fall folglich nur den zum ersten Minimum von $\bar{I}(\vartheta)$ gehörenden Winkel ϑ_1 zu berechnen. Dieses Minimum befindet sich dort, wo der Zähler im Klammerterm von (22.7) verschwindet bzw. wo $J_1(ka \sin \vartheta) = J_1(w)$ seine erste Nullstelle hat. Und das ist bei $w_1 = 3,832$ der Fall:

$$J_1(ka \sin \vartheta_1) = J_1(w_1) = 0 \quad \Leftrightarrow$$

$$w_1 = 3,832 = ka \sin \vartheta_1 \quad \Leftrightarrow \quad \sin \vartheta_1 = 3,832 \cdot \frac{1}{ka} = 3,832 \cdot \frac{\lambda}{2\pi a} = \frac{3,832}{\pi} \cdot \frac{\lambda}{2a}.$$

Mit dem Durchmesser $D = 2a$ der kreisrunden Apertur und unter Berücksichtigung der Tatsache, dass der Winkel ϑ_1 zwischen den Zentralstrahlen⁴⁷ zweier verschiedener Punktquellen liegt und wir ihn deshalb mit θ_{\min} bezeichnen, erhalten wir daraus schließlich

$$\sin \theta_{\min} = 1,22 \cdot \frac{\lambda}{D}.$$

Ist der Abstand R (siehe Abbildung 23) bekannt, kann man den (gerade noch auflösbaren) Mindestabstand d_{\min} **in der Beobachtungsebene** berechnen mit Hilfe von

$$\sin \theta_{\min} = \frac{d_{\min}}{R} = 1,22 \cdot \frac{\lambda}{D} \quad \Leftrightarrow$$

$$d_{\min} = 1,22 \cdot \frac{R \lambda}{D}. \quad (22.8)$$

⁴⁶John William Strutt „Lord Rayleigh“ (1842–1919): englischer Physiker und Nobelpreisträger.

⁴⁷Als Zentralstrahl bezeichnen wir hier den Strahl von der Punktquelle zum (zentralen) Hauptmaximum des Airy-Musters.

- $\sin \theta_{\min}$ ist die **Winkelauflösungsgrenze**.
- θ_{\min} ist der **minimale Winkelabstand**.
- d_{\min} ist der **bildseitige** (auflösbare) **Mindestabstand**.

Beispiel aus der Fotooptik

Das Objektiv sei auf die Beobachtungsebene bzw. Bildebene fokussiert. Bei einer **Land-schaftsaufnahme** ist die Gegenstandsweite a sehr groß bzw. geht gegen Unendlich, sodass dann die Bildweite b nur vernachlässigbar größer ist als die Objektivbrennweite f . Unter Berücksichtigung von (22.8) können wir also von $R = b = f$ ausgehen und erhalten für den bildseitigen (auflösbaren) Mindestabstand

$$d_{\min} = 1,22 \cdot \frac{f \lambda}{D} .$$

Im Fall von Nahaufnahmen ist die Bildweite deutlich größer als die Objektivbrennweite, sodass wir dann b statt f verwenden müssten, um ein Ergebnis mit hinreichender Genauigkeit zu erhalten.

Bei der Durchlichtmikroskopie ist die Gegenstandsweite sogar nur geringfügig größer als die Objektivbrennweite und die Bildweite ist sehr viel größer als die Objektivbrennweite. Außerdem interessiert uns in der Mikroskopie nicht der bildseitige sondern der objektseitige Mindestabstand, d. h. der gerade noch auflösbare Abstand zwischen zwei Objektpunkten. Weil also bei der Mikroskopie die Verhältnisse teilweise umgekehrt sind, werden wir das Auflösungsvermögen des Mikroskops gesondert betrachten. Allgemein und vereinfacht gilt die folgende Übersichtsdarstellung, in der die bildseitigen Größen mit einem Strich indiziert sind:

- Das Doppelpunktauflösungsvermögen *entspricht* dem bildseitigen Mindestabstand, d. h. dem kleinsten erfassbaren „Bildpunktabstand“

$$\Delta y'_{\min} = \gamma \cdot \frac{\lambda_0}{n' \cdot \sin \alpha'} , \quad 0,47 \leq \gamma \leq 1,22 .$$

- Das Auflösungsvermögen des Mikroskops *entspricht* dem objektseitigen Mindestabstand, d. h. dem kleinsten erfassbaren Objektpunktabstand

$$\Delta y_{\min} = \gamma \cdot \frac{\lambda_0}{n \cdot \sin \alpha} , \quad 0,47 \leq \gamma \leq 1,22 .$$

λ_0 ist die Vakuumwellenlänge, $n \cdot \sin \alpha$ ist die objektseitige und $n' \cdot \sin \alpha'$ die bildseitige numerische Apertur \mathcal{NA} . Der Faktor $\gamma = 0,47$ resultiert aus dem Sparrow-Kriterium, auf das wir nicht eingehen werden.

23 Das laterale Auflösungsvermögen des Mikroskops

Quellen und Literaturhinweise:

- Rudolf Gross, WMI – Lecture Notes – Kapitel 7, Abbildungstheorie, https://www.wmi.badw.de/teaching/Lecturenotes/Physik3/Gross_Physik_III_Kap_7.pdf
- Friedrich-Schiller-Universität Jena, Physikalisches Grundpraktikum, 404 – Mikroskop, www.uni-jena.de/pafmedia/Studium/Grundpraktikum/V_404.pdf

In der Mikroskopie ist der objektseitige Öffnungswinkel ω_{Ob} des Objektivs von großer Bedeutung (siehe Abbildung 28). Er ist wie folgt definiert:

$$\omega_{\text{Ob}} := 2 \cdot \arctan\left(\frac{D/2}{a}\right)$$

mit der Gegenstandsweite a und dem Objektivdurchmesser D . Und es gilt folglich:

$$\text{halber Öffnungswinkel } \alpha = \frac{\omega_{\text{Ob}}}{2} = \arctan\left(\frac{D/2}{a}\right).$$

Für den (gerade noch auflösbaren) **Mindestabstand** zwischen zwei Objektpunkten oder zwei benachbarten Gitterstrichen schreiben wir im Rahmen der Durchlichtmikroskopie einheitlich g_{min} .

23.1 Laterales Auflösungsvermögen des Mikroskops mit dem Rayleigh-Kriterium

Nach Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 2, Elektrizität und Optik, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2004, Abschnitt 11.3.3 *Auflösungsvermögen des Mikroskops*, Seite 353 und Seite 354.

Das Auflösungsvermögen mit dem Rayleigh-Kriterium ist ein sog. Doppelpunktauflösungsvermögen und bezieht sich auf selbstleuchtende Objekte wie z. B. bei der Fluoreszenzmikroskopie und auf **inkohärent** durchleuchtete Objekte. Bei einzelnen inkohärenten Punktquellen sind die Intensitäten der jeweiligen Beugungsmuster dieser Punktquellen zu addieren, um das Bild zu erhalten:

$$I(\theta) = I_1(\theta) + I_2(\theta) = A_1^2(\theta) + A_2^2(\theta).$$

Inkohärente Beleuchtung führt *allgemein* zu einer besseren Auflösung als kohärente Beleuchtung.

Der Abstand Objektebene–Objektiv ist die Gegenstandsweite a , der Abstand Objektivöffnung–Zwischenbildebene die Bildweite b .

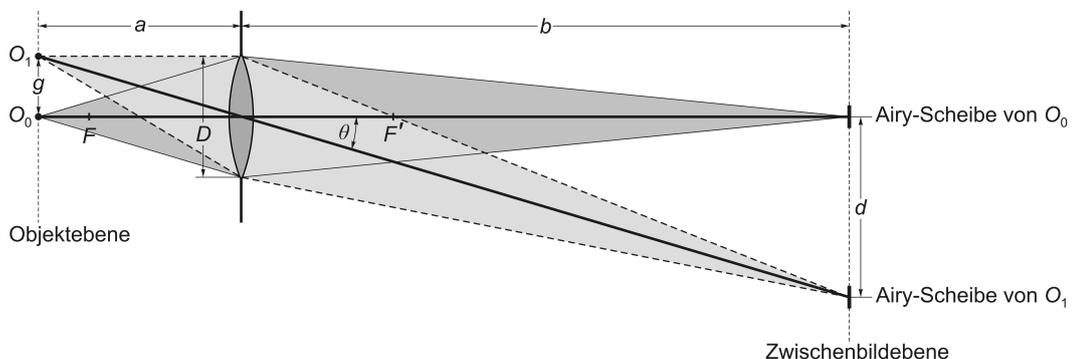


Abb. 24 Zum Auflösungsvermögen des Mikroskops mit dem Rayleigh-Kriterium.

Von zwei beleuchteten Objektpunkten O_0 und O_1 (siehe Abbildung 24) verlaufen die „Zentralstrahlen“ mit dem Winkel θ zueinander durch das Öffnungszentrum des Objektivs zur Beobachtungs- bzw. Zwischenbildebene. Dort erzeugen sie nicht ihren (geometrischen) Bildpunkt sondern jeweils ihr Airy-Scheibchen, sodass wir mit $R = b$ die Formel (22.8) zur Berechnung des bildseitigen (auflösbaren) Mindestabstands d_{\min} heranziehen können:

$$d_{\min} = 1,22 \cdot \frac{b \lambda}{D}. \quad (23.1)$$

Diesem bildseitigen Mindestabstand d_{\min} entspricht ein objektseitiger (gerade noch auflösbarer) Mindestabstand g_{\min} nach dem Strahlensatz wie folgt:

$$\frac{g_{\min}}{a} = \frac{d_{\min}}{b} = 1,22 \cdot \frac{\lambda}{D} \Leftrightarrow g_{\min} = 1,22 \cdot a \cdot \frac{\lambda}{D}. \quad (23.2)$$

Weil allgemein in der Durchlichtmikroskopie die Objektebene praktisch in der vorderen Brennebene des Objektivs liegt, setzen wir $a = f$ und erhalten aus (23.2) schließlich

$$\boxed{g_{\min} = 1,22 \cdot \frac{f \lambda}{D}}.$$

Definitionsgemäß ist damit das Auflösungsvermögen A des Mikroskops

$$A = \frac{1}{g_{\min}} = 0,82 \cdot \frac{D}{f \lambda}.$$

Der Öffnungswinkel 2α eines Objektivs wird manchmal angegeben mittels der Näherung

$$2 \cdot \sin \alpha \approx \frac{D}{f}.$$

Allerdings ist diese Näherung nur für Winkel α bis ca. 30° eine brauchbare Näherung. Für weiter anwachsende Winkel α wird diese Näherung sehr schnell unbrauchbar. Dessen ungeachtet wird sie verwendet, um die Abhängigkeit des Auflösungsvermögens von der numerischen Apertur $NA = n \cdot \sin \alpha$ des Objektivs darzustellen. Setzen wir diese Näherung nämlich in (23.1) ein, so erhalten wir

$$g_{\min} \approx 1,22 \cdot \frac{\lambda}{2 \cdot \sin \alpha} = 1,22 \cdot \frac{\lambda_0}{2 \cdot n \cdot \sin \alpha},$$

$$\boxed{g_{\min} \approx 0,61 \cdot \frac{\lambda_0}{n \cdot \sin \alpha} = 0,61 \cdot \frac{\lambda_0}{NA}}.$$

23.2 Laterales Auflösungsvermögen des Mikroskops mit dem Abbe-Kriterium

Das Abbe-Kriterium bezieht sich auf die Durchlichtmikroskopie mit **kohärent** beleuchteten Objekten.

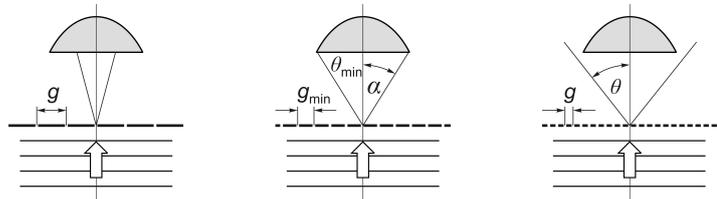
Bei kohärenten Quellen sind die Einzelamplituden zu addieren und die Amplitudensumme anschließend zu quadrieren. Es resultiert das Intensitätsmuster als Bild:

$$I(\theta) = \left[A_1(\theta) + A_2(\theta) \right]^2.$$

1873 beantwortete Ernst Abbe⁴⁸ die Frage, bis zu welcher kleinsten Gitterkonstante g_{\min} mit einem Mikroskop die Abbildung eines **Strichgitters** möglich ist. Er fand dabei das nach ihm benannte Kriterium für das Auflösungsvermögen des Mikroskops, das **Abbe-Kriterium**:

Um eine mikroskopische Abbildung eines Objekts zu erhalten bzw. ein Objekt aufzulösen, muss neben dem Beugungsmaximum 0. Ordnung mindestens ein Fraunhofer-Beugungsmaximum 1. Ordnung (+1. Ordnung *oder* -1. Ordnung) in das Objektiv fallen (siehe Abbildung 25 und Abbildung 28). Man spricht in diesem Zusammenhang folglich vom **Gitterauflösungsvermögen**.

Abb. 25 Zum Abbe-Kriterium. Mit kleiner werdender Gitterkonstante g nimmt die Spreizung der Beugungsordnungen zu, sodass dann die 1. Beugungsordnung nicht mehr in das Objektiv fällt.



Beim Fraunhofer-Beugungsspektrum des Strichgitters bezeichnet man die Hauptmaxima als Beugungsmaxima m -ter Ordnung. Das Fraunhofer-Beugungsmaximum 0. Ordnung *allein* liefert keine Bildinformation und somit auch keine Abbildung des Objekts. Durch jede weitere höhere Beugungsordnung kommen Bildinformationen hinzu, sodass die Auflösung bzw. die Bildschärfe zunimmt. Je größer der Öffnungswinkel 2α des Objektivs ist, desto mehr Beugungsordnungen können vom Objektiv zum Aufbau des Bildes „herausgefiltert“ werden. Wird jedoch nur die 0. Ordnung herausgefiltert, so entsteht in der hinteren Brennebene (Fourier-Ebene) nur ein Punkt bzw. eine Linie des Fraunhofer-Beugungsspektrums, das Beugungsmaximum 0. Ordnung, und in der Bildebene nur ein Lichtfleck gleichmäßiger Helligkeit ohne weitere Bildinformationen.

Das Auflösungsvermögen des Mikroskops *entspricht* also der Gitterkonstante g_{\min} , d. h. dem Abstand zweier benachbarter Striche, bei dem gerade noch die 0. und die 1. Beugungsordnung vom Objektiv durchgelassen werden. Bei weiterer Verringerung des Strichabstands (der Gitterkonstante) g , wird das Spektrum der Beugungsordnungen räumlich weiter gespreizt, sodass nur noch die 0. Ordnung in das Objektiv gelangt und somit ein Bildaufbau nicht mehr möglich ist.

⁴⁸Ernst Carl Abbe (1840–1905): deutscher Physiker, Optiker und Industrieller.

23.2.1 Senkrecht einfallendes Licht

Wir betrachten zunächst den einfachen Fall, dass das Objekt (Strichgitter) mit senkrecht einfallendem kohärentem Licht beleuchtet (durchleuchtet) wird (siehe Abbildung 26).

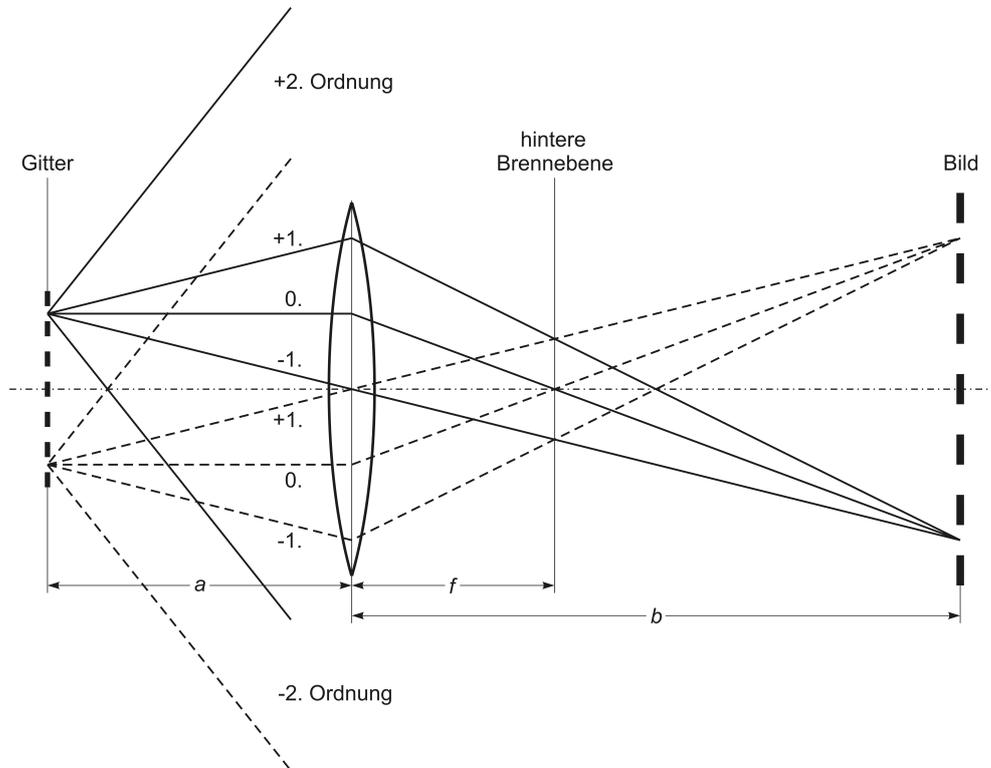


Abb. 26 Zum Auflösungsvermögen des Mikroskops nach dem Abbe-Kriterium. Eine ebene kohärente Welle falle von links senkrecht auf das Strichgitter. Die Sammellinse generiert in ihrer hinteren Brennebene aus dem am Gitter gebeugten Licht das Fraunhofer-Beugungsmuster und erzeugt schließlich in der Bildebene das umgekehrte Bild des Strichgitters in einem bestimmten Abbildungsmaßstab.

Mit (19.1) und dem Gangunterschied Δs zwischen den benachbarten Strahlenbündeln beim Strichgitter hatten wir für die zugehörige **Phasendifferenz**

$$\delta = k \cdot \Delta s = \frac{2\pi}{\lambda} \cdot \underbrace{g \sin \theta}_{=\Delta s}$$

gefunden. Gemäß (19.2) gilt dann für die Hauptmaxima der Ordnung m beim Strichgitter

$$\delta_m = \frac{2\pi}{\lambda} \cdot g \sin \theta_m = m \cdot 2\pi \quad \Leftrightarrow \quad g = m \frac{\lambda}{\sin \theta_m} \quad \Rightarrow \quad g \propto \frac{1}{\sin \theta_m} .$$

Nach dem Abbe-Kriterium, d. h. mit $m = 1$, und mit (24.1), also mit dem Zusammenhang zwischen der Vakuumwellenlänge λ_0 und der Wellenlänge im Medium mit dem Brechungsindex n , ergibt dies

$$g_{\min} = \frac{\lambda}{\sin \theta_{\min}} \xrightarrow{(24.1)} g_{\min} = \frac{\lambda_0}{n \cdot \sin \theta_{\min}} . \quad (23.3)$$

Durch den Objektivdurchmesser ist dessen Öffnungswinkel 2α vorgegeben. Und weil für $\theta_1 \leq \alpha$ die 1. Beugungsordnung in das Objektiv fallen kann, erhalten wir schließlich

$$g_{\min} = \frac{\lambda_0}{n \cdot \sin \alpha} .$$

Dass sich mit dem Größerwerden des Brechungsindex n das Auflösungsvermögen vergrößert (Verkleinerung von g_{\min}), sehen wir auch bei Betrachtung der Phasendifferenz

$$\delta_m = m \cdot 2\pi \xrightarrow{m=1} \delta_1 = 2\pi$$

für die 1. Beugungsordnung:

$$\delta_1 = 2\pi = \frac{2\pi}{\lambda} \cdot g \sin \theta_1 \Leftrightarrow \sin \theta_1 = \frac{1}{g} \cdot \lambda$$

$$\Rightarrow \sin \theta_1 \propto \lambda ,$$

$$\xrightarrow{\lambda = \frac{\lambda_0}{n}} \sin \theta_1 \propto \frac{1}{n} .$$

Damit wird deutlich, warum man bei Verwendung von Immersionsöl als Medium zwischen Objekt und Objektiv das Auflösungsvermögen und damit die **förderliche Vergrößerung** eines Mikroskops erhöhen kann. Vergrößerungen durch ein Mikroskop mittels stark vergrößernder Okulare über die förderliche Vergrößerung hinaus liefern keine zusätzlichen Bildinformationen bzw. keine neuen Objektdetails, weshalb man dann von **leerer Vergrößerung** spricht. Während Luft nur einen Brechungsindex $n \approx n_0 = 1$ besitzt, hat Immersionsöl einen Brechungsindex von $n = 1,5$. Der gerade noch auflösbare Abstand zwischen zwei Gitterstrichen bzw. zwei Punktquellen wird bei Verwendung von Immersionsöl mit dem Faktor $\frac{1}{n} = \frac{1}{1,5} = 2/3$ verringert, d. h. *um* 1/3 verkleinert.

23.2.2 Schräg einfallendes Licht

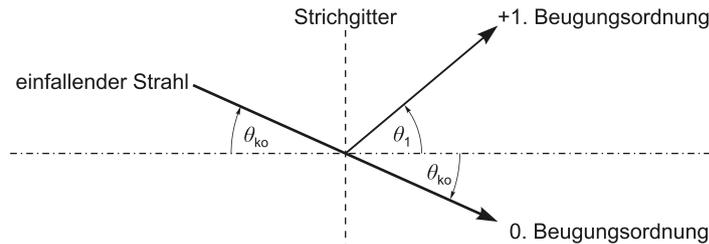


Abb. 27 Zur „schiefen Beleuchtung“ des Strichgitters. Der Index ko steht für kondensorseitig und der Index 1 für objektivseitig und/oder 1. Beugungsordnung. Ein vom Kondensator ausgehendes kohärentes Lichtbündel (einfallender Strahl) fällt unter dem Winkel θ_{ko} auf die Beugungsebene (Strichgitter). Der einfallende Strahl und die 0. Beugungsordnung verlaufen in der gleichen Richtung. Im Folgenden ist zu berücksichtigen, dass entsprechend dieser Abbildung $\sin \theta_{ko} < 0$ und $\sin \theta_1 > 0$ gilt.

Man kann sich überlegen, dass für den Gangunterschied zwischen der 0. und der +1. Beugungsordnung

$$\Delta s = \Delta s_1 - \Delta s_{ko} = g \sin \theta_1 - g \sin \theta_{ko} \quad \text{für } \theta_{ko} < 0, \theta_1 > 0 \quad (23.4)$$

gilt, und nicht $\Delta s = g \cdot \sin (|\theta_{ko}| + |\theta_1|)$, weil wegen des schrägen Einfalls eines kohärenten Lichtbündels am Gitter bereits der Gangunterschied Δs_{ko} (innerhalb dieses Bündels) besteht. Durch die Beugung dieses einfallenden Lichtbündels am Strichgitter entsteht dann der zusätzliche Gangunterschied Δs_1 . Unter Berücksichtigung von Abbildung 27 mit $\sin(-\theta_{ko}) = -\sin \theta_{ko}$ und im Hinblick auf die folgende Herleitung des Auflösungsvermögens des Mikroskops schreiben wir für (23.4)

$$\Delta s = g \cdot |\sin \theta_{ko}| + g \cdot |\sin \theta_1| .$$

Daraus resultiert dann die Phasendifferenz bei konstruktiver Interferenz in der 1. Beugungsordnung entsprechend $m = 1$ und mit

$$\lambda_{ko} = \frac{\lambda_0}{n_{ko}}, \quad \lambda_1 = \lambda = \frac{\lambda_0}{n} :$$

$$\begin{aligned} \delta &= \frac{2\pi}{\lambda_{ko}} g |\sin \theta_{ko}| + \frac{2\pi}{\lambda_1} g |\sin \theta_1| \\ &= \frac{2\pi}{\lambda_0} g n_{ko} |\sin \theta_{ko}| + \frac{2\pi}{\lambda_0} g n |\sin \theta_1| , \\ \delta &= \frac{2\pi}{\lambda_0} g \left(n_{ko} |\sin \theta_{ko}| + n |\sin \theta_1| \right) = m \cdot 2\pi , \\ \delta_{(m=1)} &= \frac{2\pi}{\lambda_0} g \left(n_{ko} |\sin \theta_{ko}| + n |\sin \theta_1| \right) = 2\pi . \end{aligned} \quad (23.5)$$

Hierbei ist die Gitterkonstante g vorgegeben. Wenn wir aber g nicht kennen, dafür aber die Winkel θ_{ko} und θ_1 , welche die 1. Beugungsordnung liefern, so können wir die

zugehörige Gitterkonstante, also den Strichabstand g_1 durch Äquivalenzumformung aus (23.5) ermitteln:

$$g_1 = \frac{\lambda_0}{n_{\text{ko}} |\sin \theta_{\text{ko}}| + n |\sin \theta_1|} . \quad (23.6)$$

Bei vorgegebenem Öffnungswinkel 2α des Objektivs lässt sich das Auflösungsvermögen des Mikroskops durch „schiefe Beleuchtung“ oder genauer gesagt durch schiefe oder schräge Durchleuchtung des Objekts vergrößern (siehe Abbildung 28).

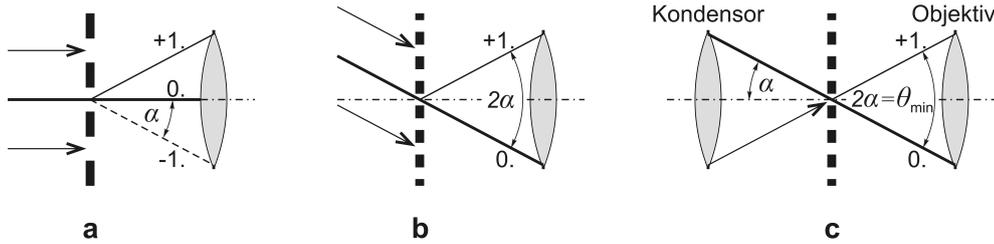


Abb. 28 Zum Auflösungsvermögen des Mikroskops nach dem Abbe-Kriterium. Als Objekt dient hier ein Strichgitter.

- a** Bei senkrechtem Lichteinfall bewirkt hier die vorgegebene Gitterkonstante g , dass sowohl die 0. und +1. als auch die 0. und -1. Beugungsordnung gerade noch in das Objektiv fallen.
- b** „Schiefe Beleuchtung“ des Strichgitters führt hier dazu, dass die 0. und +1. Beugungsordnung auch bei kleinerer Gitterkonstante g als in Abbildung a gerade noch in das Objektiv fallen.
- c** „Schiefe Beleuchtung“ durch einen Kondensator, der den gleichen Öffnungswinkel 2α wie das Objektiv besitzt. Der Öffnungswinkel 2α ist hier gleich dem kleinsten Winkel θ_{\min} entsprechend der minimalen Gitterkonstante $g_1 = g_{\min}$, die es ermöglicht, dass sowohl die 0. und +1. als auch die 0. und -1. Beugungsordnung gerade noch in das Objektiv fallen. Infolge des schrägen Lichteinfalls ist g_{\min} hier kleiner als g in Abbildung a.

Das ist möglich, weil zur Bildentstehung neben der 0. Beugungsordnung nur mindestens eine der 1. Beugungsordnungen (+1. oder -1. Ordnung) erforderlich ist. Ist der Öffnungswinkel des Kondensators $2\alpha_{\text{ko}}$ und gilt

$$|\sin \theta_{\text{ko}}| = \sin \alpha_{\text{ko}} \quad \text{und} \quad |\sin \theta_1| = \sin \alpha ,$$

so definiert (23.6) den gerade noch auflösbaren Abstand bzw. den Mindestabstand zwischen benachbarten Gitterstrichen für den allgemeinen Fall, dass die Brechungsindizes und die Öffnungswinkel kondensator- und objektivseitig verschieden sind:

$$g_{\min} = \frac{\lambda_0}{n_{\text{ko}} |\sin \theta_{\text{ko}}| + n |\sin \theta_1|} = \frac{\lambda_0}{n_{\text{ko}} \sin \alpha_{\text{ko}} + n \sin \alpha} ,$$

$$g_{\min} = \frac{\lambda_0}{NA_{\text{Kondensator}} + NA_{\text{Objektiv}}} . \quad (23.7)$$

Das Produkt aus halbem Öffnungswinkel von Kondensator bzw. Objektiv mit dem Brechungsindex des jeweils umgebenden Mediums bezeichnet man als numerische Apertur NA des Kondensators bzw. Objektivs.

Praktisch realisiert wird die „schiefe Beleuchtung“ im einfachsten Fall durch einen Kondensator, der den gleichen Öffnungswinkel 2α wie das Objektiv besitzt (siehe Abbildung 28 c). Die allgemeine Formel (23.7) vereinfacht sich dann wegen $\sin \alpha_{\text{ko}} = \sin \alpha$ zu

$$g_{\min} = \frac{\lambda_0}{(n_{\text{ko}} + n) \cdot \sin \alpha} .$$

Wenn das Medium kondensatorseitig Luft gemäß $n_{\text{ko}} = n = 1$ ist, was fast immer der Fall ist, erhalten wir

$$g_{\min} = \frac{\lambda_0}{(1 + n) \cdot \sin \alpha} .$$

Und wenn das Medium sowohl kondensator- als auch objektivseitig Luft ist, so resultiert schließlich die oft für das maximal mögliche „Auflösungsvermögen“ des Mikroskops angegebene einfache Formel

$$g_{\min} = \frac{\lambda_0}{2 \cdot \sin \alpha} = 0,5 \cdot \frac{\lambda_0}{\sin \alpha} .$$

Das laterale Auflösungsvermögen des Mikroskops mit dem Abbe-Kriterium bei „schiefer Beleuchtung“ ist folglich fast identisch mit der

$$\text{Grenzauflösung } g_{\min} = 0,51 \frac{\lambda_0}{n \cdot \sin \alpha} ,$$

bei der sich die benachbarten „Bildmittelpunkte“ bzw. Mittelpunkte der Airy-Scheiben im Abstand der Halbwertsbreite der zugehörigen Fraunhofer-Hauptmaxima befinden.

Auflösbare Mindestabstände g_{\min} bei verschiedenen Kriterien:

- Rayleigh-Kriterium: $g_{\min} = 0,61 \cdot \frac{\lambda_0}{n \cdot \sin \alpha} ,$
- Abbe-Kriterium bei senkrechter Beleuchtung: $g_{\min} = \frac{\lambda_0}{n \cdot \sin \alpha} ,$
- Abbe-Kriterium bei „schiefer Beleuchtung“: $g_{\min} = \frac{\lambda_0}{n_{\text{ko}} \cdot \sin \alpha_{\text{ko}} + n \cdot \sin \alpha} ,$
 für $\alpha_{\text{ko}} = \alpha$ und $n_{\text{ko}} = 1$: $g_{\min} = \frac{\lambda_0}{(1 + n) \cdot \sin \alpha} ,$
 für $\alpha_{\text{ko}} = \alpha$ und $n_{\text{ko}} = n = 1$: $g_{\min} = 0,50 \cdot \frac{\lambda_0}{\sin \alpha} ,$
- Grenzauflösung: $g_{\min} = 0,51 \cdot \frac{\lambda_0}{n \cdot \sin \alpha} .$

24 Zur Abbe'schen Abbildungstheorie

- Abbe'sche Sinusbedingung⁴⁹

Tatsache ist, dass die Lichtgeschwindigkeit c vom Brechungsindex n des Ausbreitungsmediums abhängt gemäß

$$c = \frac{c_0}{n} \quad \text{mit der Vakuumlichtgeschwindigkeit } c_0 .$$

Außerdem ist die Frequenz f des Lichts unabhängig vom Medium, sodass für den Zusammenhang zwischen der Wellenlänge λ und dem Brechungsindex n des zugehörigen Mediums folgendes gilt:

$$c = \frac{c_0}{n} = \lambda \cdot f \quad \Leftrightarrow \quad c_0 = n \lambda \cdot f = \lambda_0 \cdot f \quad \Rightarrow$$

$$\text{Vakuummwellenlänge } \lambda_0 = n \lambda \quad \Rightarrow \quad \lambda = \frac{\lambda_0}{n} \propto \frac{1}{n} . \quad (24.1)$$

Mit der Herleitung der Abbe'schen Sinusbedingung wird die Frage beantwortet, unter welcher Bedingung eine weitestgehend aplanatische Abbildung, d. h. eine Abbildung ohne nennenswerte sphärische Aberration, mit einer Sammellinse möglich ist.

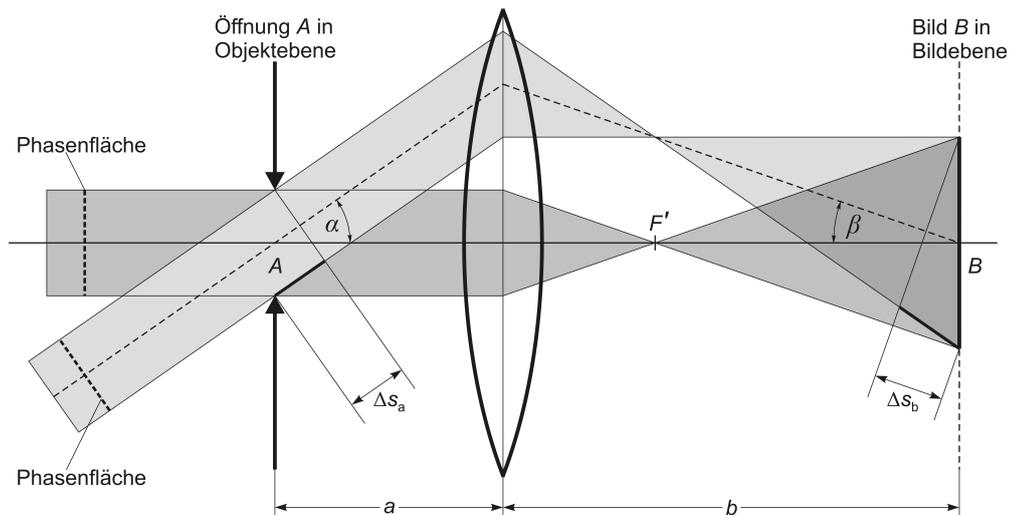


Abb. 29 Zur Abbe'schen Sinusbedingung. (Vergleiche mit den Abbildungen 9 und 10 im Abschnitt 19.2.1.)

Wie in der Abbildung 29 zu sehen ist, beleuchten wir eine Kreisblende, das Objekt A in der Objektebene, mit einer sehr weit entfernten und sehr ausgedehnten Lichtquelle und greifen zwei hinter der Kreisblende resultierende Parallelbündel mit ihren ebenen Phasenflächen heraus. Das eine Bündel verlaufe senkrecht zur Kreisblende und somit entlang der optischen Achse, das andere verlaufe schräg zur Kreisblende bzw. im Winkel α zur optischen Achse. Mittels einer Sammellinse

⁴⁹Siehe Wolfgang Demtröder, Springer-Lehrbuch Experimentalphysik 2 – Elektrizität und Optik, Springer-Verlag, Berlin, Heidelberg, New York, 2004, 3. Auflage, Abschnitt 9.5.7 Die aplanatische Abbildung, Seite 281 bis Seite 282.

entsteht aus den beiden Lichtbündeln in der Bildebene das umgekehrte Abbild B der Kreisblende im Abbildungsmaßstab

$$M = \frac{|B|}{|A|}.$$

Das schräge Lichtbündel verläuft auf der Bildseite mit dem Winkel β zur optischen Achse. Wenn wir jetzt die Strahlen vom Objekt zum Bild verfolgen, können wir unter Berücksichtigung der Tatsachen, dass die hintere Brennebene der Sammellinse durch F' gleichzeitig die Fraunhofer-Beugungsebene ist, folgendes feststellen:

Von Pfeilbasis zu Pfeilbasis entsteht (objektseitig) vor der Fraunhofer-Beugungsebene durch Wegverlängerung der Gangunterschied

$$\Delta s_a = |A| \cdot \sin \alpha.$$

Von Pfeilspitze zu Pfeilspitze entsteht (bildseitig) hinter der Fraunhofer-Beugungsebene durch Wegverlängerung der Gangunterschied

$$\Delta s_b = |B| \cdot \sin \beta.$$

Die aus diesen Gangunterschieden resultierenden **Phasendifferenzen**

$$\delta = k \cdot \Delta s = \frac{2\pi}{\lambda} \cdot \Delta s \stackrel{(24.1)}{=} \frac{2\pi}{\frac{\lambda_0}{n}} \cdot \Delta s = \frac{2\pi}{\lambda_0} \cdot n \Delta s$$

sind

$$\delta_a = \frac{2\pi}{\lambda_0} \cdot n_a \Delta s_a \quad \text{und} \quad \delta_b = \frac{2\pi}{\lambda_0} \cdot n_b \Delta s_b.$$

Wenn wir bei *vorgegebener Gegenstandsweite* a dafür sorgen, dass diese beiden Phasendifferenzen gleich sind, erhalten wir eine aplanatische Abbildung, weil dann die Änderung der Phase im Strahlenverlauf Objektspitze zu Bildspitze gleich der Änderung der Phase im Strahlenverlauf Objektbasis zu Bildbasis ist:

$$\delta_a = \delta_b \quad \Rightarrow \quad n_a \Delta s_a = n_b \Delta s_b$$

$$n_a \cdot |A| \cdot \sin \alpha = n_b \cdot |B| \cdot \sin \beta \quad \Leftrightarrow$$

Abbe'sche Sinusbedingung $\frac{n_a \sin \alpha}{n_b \sin \beta} = \frac{ B }{ A } = M = \text{const}$.
