

Reinhard Weiß

L i c h t

Kohärenz – Polarisation – Wellenoptik

Dieses Skript ist in Bearbeitung. Letzte Aktualisierung am 06.02.2026.

Inhaltsverzeichnis

1	Vorbemerkungen	5
2	Intensität elektromagnetischer Wellen	9
3	Absorption und Emission elektromagnetischer Strahlung	12
3.1	Schalenmodell und Orbitalmodell	12
3.2	Einstein-Koeffizienten und Übergangswahrscheinlichkeiten	16
3.3	Die Idee zum Funktionsprinzip des Lasers	23
4	Brechungsindex und das Verhalten von Licht an Grenzflächen	24
5	Evaneszentes Feld – Goos-Hänchen-Verschiebung	26
6	Die Kohärenz des Lichts	33
6.1	Wellenpaket – Wellenzug – Superpositionsprinzip	33
6.1.1	Teilchenbild	33
6.1.2	Wellenbild	33
6.2	Kohärenz	37
6.3	Zur zeitlichen (longitudinalen) Kohärenz – Kohärenzlänge	42
6.4	Zur räumlichen (lateralen) Kohärenz – räumliche Kohärenzlänge	46
6.5	Zur Kohärenz bei Punktquellen	49
7	Die Polarisation des Lichts	51
7.1	Linear polarisiertes Licht	53
7.2	Zirkular polarisiertes Licht	54
7.3	Elliptisch polarisiertes Licht	56
7.4	Jones-Formalismus	59
7.5	Intensität des polarisierten Lichts im Vakuum	71
8	Polarisation des Lichts in Bra-Ket-Notation	72
8.1	Der Umgang mit komplexen Vektoren und Matrizen in Bra-Ket-Notation . .	72
8.2	Lineare Polarisation	77
8.2.1	VON-Basis, Bra- und Ket-Vektoren	77
8.2.2	Polarisationsfilter mit horizontaler Transmissionsachse	78
8.2.3	Polarisationsfilter mit Transmissionsachse im Winkel ϑ	80
8.3	Zirkulare Polarisation	81
8.4	Photonenspin	81
9	Tabellen: Jones-Vektoren und Jones-Matrizen	83
10	Paradoxien im Teilchenbild	85
11	Grundlegendes zur Wellenoptik	87
11.1	Der Begriff „Wellenoptik“	87
11.2	Fraunhoferbeugung	89
11.2.1	Elementare Herleitung im Reellen	89
11.3	Fresnel'sche Zonen	97
12	Zur (skalaren) Fresnel-Kirchhoff'schen Beugungstheorie	105
12.1	Vorbereitung – Helmholtz-Gleichung	105
12.2	Die Kirchhoff'sche Formel	109

12.3	Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals	112
12.4	Der Neigungsfaktor K	115
12.5	Zur auslaufenden Welle im Unendlichen	116
12.6	Anhang: Fresnel- und Fraunhofer-Beugung	120
13	Optische Abbildung in Fresnel- und Fraunhofer-Näherung	125
13.1	Fresnel-Näherung	126
13.2	Bedingung für die Fresnel-Näherung	128
13.3	Fraunhofer-Näherung	128
13.4	Bedingung für die Fraunhofer-Näherung	129
14	Fraunhofer-Beugung und Auflösungsvermögen	130
14.1	Bestimmung des Auflösungsvermögens bei kreisrunder Öffnung	135
15	Das laterale Auflösungsvermögen des Mikroskops	137
15.1	Laterales Auflösungsvermögen des Mikroskops mit Rayleigh-Kriterium	137
15.2	Laterales Auflösungsvermögen des Mikroskops mit Abbe-Kriterium	139
15.2.1	Senkrecht einfallendes Licht	140
15.2.2	Schräg einfallendes Licht	142
16	Zur Abbe'schen Abbildungstheorie	145
17	Anhang	147
17.1	Hauptachsentransformation von Kegelschnitten zur Bestimmung der Parame- ter von Ellipsen	147
17.1.1	Hauptachsentransformation	147
17.1.2	Spezielle Ellipsen	151
17.1.3	Eigenwerte und Eigenvektoren zur Matrix A	152
17.1.4	Bestimmung der Halbachsenlängen schräger Ursprungsellipsen	154
17.1.5	Bestimmung des Winkels α schräger Ursprungsellipsen	154
17.2	Zu Abschnitt 7.3 – Berechnung von E_{0x} und E_{0y}	156

Quellen

- Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag, Wien, München, 2005.
- F. Pedrotti, L. Pedrotti, W. Bausch, H. Schmidt, *Optik für Ingenieure – Grundlagen*, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, 2005.
- Wolfgang Demtröder, Springer-Lehrbuch, *Experimentalphysik 2 – Elektrizität und Optik*, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2004, Abschnitt 7.4 *Polarisation elektromagnetischer Wellen*, Seite 192 und Seite 193, Kapitel 10. *Interferenz, Beugung und Streuung*, Seite 295 bis Seite 300.
- Rudolf Langkau, Wolfgang Scobel, Gunnar Lindström, Springer-Lehrbuch, *Physik kompakt 2 – Elektrodynamik und Elektromagnetische Wellen*, 2. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2002, Kapitel 3 *Optik*, Seite 401 bis Seite 404.
- Rainer Müller und Hartmut Wiesner, *Das Münchener Unterrichtskonzept zur Quantenmechanik*,
[https://www.didaktik.physik.uni-muenchen.de/archiv/
inhalt_materialien/milq/lehrtext_x_.pdf](https://www.didaktik.physik.uni-muenchen.de/archiv/inhalt_materialien/milq/lehrtext_x_.pdf)
- Josef Oswald, *Grundprinzipien der Quantenphysik*, Vorlesung – WS03/04,
https://oswald.unileoben.ac.at/qpv1_1/qpv1.pdf
- Timo Weigand, Kursvorlesung PTP4 Theoretische Quantenmechanik (digitalisiert von David Breyel), Institut für Theoretische Physik, Universität Heidelberg, 2011,
[https://www.thphys.uni-heidelberg.de/
courses/weigand/Skript-QM2011.pdf](https://www.thphys.uni-heidelberg.de/courses/weigand/Skript-QM2011.pdf)

1 Vorbemerkungen

- Die Maxwell'schen Gleichungen ermöglichen die vollständige Beschreibung elektromagnetischer Wellen und damit von Licht im **Wellenbild** – aber nicht im Teilchenbild.
- Im **Teilchenbild** besteht Licht aus Lichtquanten, den Photonen (Index ph für Photon).
- Im **photoelektrischen Effekt** manifestiert sich der Teilchencharakter des Lichts als **Planck-Einstein-Beziehung**

$$E_{\text{ph}} = h \cdot \nu$$

(E_{ph} Energie eines Photons, h Planck'sches Wirkungsquantum, ν Lichtfrequenz).

- Die Wellenlänge λ bzw. der Wellenvektor \vec{k} oder die Wellenzahl $|\vec{k}| = k = \frac{2\pi}{\lambda}$ sind charakteristische Welleneigenschaften.

Der Impuls $|\vec{p}| = p$ ist eine charakteristische Teilcheneigenschaft.

Die **De Broglie-Beziehung**

$$p = \hbar \cdot k = \frac{h}{2\pi} \cdot \frac{2\pi}{\lambda} = \frac{h}{\lambda}$$

gilt für Teilchen ohne Ruhemasse, also für Photonen, und verknüpft den Wellencharakter (Wellenlänge λ) des Lichts mit dessen Teilchencharakter (Impuls p), wie man am Beispiel des Strahlungsdrucks verifizieren kann:

Dass ein Photon den Impuls p_{ph} besitzt, zeigt sich in der Kraftwirkung beim Auftreffen von Licht auf ein Hindernis, d. h. durch den vom Licht ausgeübten Strahlungsdruck P .

Ausgehend von der Maxwell'schen Gleichung $\nabla \times \vec{E} = -\frac{\partial}{\partial t} \vec{B}$ für das Vakuum und ausgehend von den aus den Maxwell'schen Gleichungen abgeleiteten **Wellengleichungen**

$$\nabla^2 \vec{E} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E} \quad \text{bzw.} \quad \nabla^2 \vec{B} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{B} \quad (1)$$

für das \vec{E} -Feld bzw. für das \vec{B} -Feld im Vakuum erhalten wir die Beziehung

$$|\vec{B}| = \frac{1}{c} |\vec{E}| = \frac{1}{c} E$$

und damit dann den **Strahlungsdruck** P einer elektromagnetischen Welle:¹

$$P = \frac{I}{c}, \quad \text{Intensität } I, \quad \text{Vakuumlichtgeschwindigkeit } c.$$

Für den Strahlungsdruck gilt also sinngemäß

$$\text{Strahlungsdruck} = \frac{\text{Kraft}}{\text{Fläche}} = \frac{\text{Impuls}}{\text{Zeit} \cdot \text{Fläche}} = \frac{\frac{1}{c} \text{Energie}}{\text{Zeit} \cdot \text{Fläche}} = \frac{1}{c} \cdot \text{Intensität}.$$

Ist i die „Teilchenzahlintensität“, also die Anzahl der pro Zeiteinheit und pro Flächeneinheit auf ein Hindernis treffenden Photonen monochromatischen Lichts der Frequenz ν , dann gilt mit dem Impuls p_{ph} und der Energie $h \cdot \nu$ eines Photons

$$P = \frac{I}{c} = \frac{i \cdot h\nu}{c} = i \cdot \frac{h}{\lambda} = i \cdot \hbar k = i \cdot p_{\text{ph}} \Rightarrow$$

$$p_{\text{ph}} = \hbar k. \quad \square$$

¹Herleitungen des Strahlungsdrucks elektromagnetischer Wellen finden sich beispielsweise in: Spektrum-Lehrbuch *Physik* von Paul A. Tipler, 1. Auflage, Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 1994, Seite 1000 bis Seite 1005, Springer-Lehrbuch Gerthsen Physik, H. Vogel, 20. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 1999, Seite 601.

- Für ein Teilchen wie beispielsweise das Elektron mit der Ruhemasse m_0 , der Geschwindigkeit v und dem klassischen Impuls p gilt die *klassische* Beziehung

$$E_{\text{kin}} = \frac{\hbar^2 k^2}{2m_0} = \frac{m_0 \cdot v^2}{2} = \frac{1}{2} p \cdot v, \quad v \ll c. \quad (2)$$

Photonen besitzen keine Ruhemasse, bewegen sich mit Lichtgeschwindigkeit und sind demzufolge relativistische Teilchen mit der Energie

$$E_{\text{ph}} = p_{\text{ph}} \cdot c = h \cdot \nu. \quad (3)$$

(3) resultiert aus

$$E = mc^2 = h\nu \Leftrightarrow m = \frac{h\nu}{c^2} \quad \text{und} \quad p_{\text{ph}} = m \cdot c = \frac{h\nu}{c^2} \cdot c = \frac{E_{\text{ph}}}{c}.$$

- **Photonenspin**

Photonen besitzen also einen **linearen Impuls** $p_{\text{ph}} = \frac{h\nu}{c} = \hbar k$ mit $\vec{p}_{\text{ph}} \uparrow\uparrow \vec{k}$.

„Wenn ein geladenes Teilchen elektromagnetische Strahlung emittiert oder absorbiert, ändert sich neben seiner Energie und seinem linearen Impuls auch sein Drehimpuls um ein **ganzzahliges** Vielfaches von $\pm\hbar$.“² Neben dem linearen Impuls besitzen Photonen demzufolge auch einen Drehimpuls bzw. Eigendrehimpuls, auch intrinsischer Drehimpuls oder kurz **Spin** \vec{s} genannt. Photonen sind folglich Spin-1-Teilchen bzw. Bosonen.

Der Photonenspin ist unabhängig von der Photonenenergie und entspricht einem zusätzlichen Freiheitsgrad des Photons (neben dem linearen Impuls). Die Orientierung des Spins eines Photons ergibt sich aus dessen **Helizität** σ und wird in Bezug auf die „erzwungene“ Rotation *positiver* Ladungen in absorbierenden Medien definiert:

linkszirkulare Polarisation (\mathcal{L} - oder σ^+ -Zustand) : $\vec{s} \uparrow\uparrow \vec{k}$,

rechtszirkulare Polarisation (\mathcal{R} - oder σ^- -Zustand) : $\vec{s} \downarrow\downarrow \vec{k}$.

Bezüglich \mathcal{L} und \mathcal{R} haben wir hier die althergebrachte Konvention aus der Optik verwendet. In der Quantenphysik sind die Eigenschaften links- und rechtszirkular über die Helizität und demzufolge umgekehrt definiert, sodass dort für die **zirkularen Polarisationszustände**

$$\mathcal{L} \longrightarrow |\sigma^+\rangle = |R\rangle \quad \text{und} \quad \mathcal{R} \longrightarrow |\sigma^-\rangle = |L\rangle$$

gilt.

Linear polarisiertes Licht kann man als Überlagerung *gleicher Anteile* von rechts- und linkszirkular polarisiertem Licht auffassen, wobei die beiden Spinzustände $\pm\hbar$ mit der gleichen Wahrscheinlichkeit vorkommen, sodass kein Gesamtspin resultiert.

Dementsprechend handelt es sich bei elliptisch polarisiertem Licht um eine Überlagerung ungleicher Anteile von $|\sigma^+\rangle$ - und $|\sigma^-\rangle$ -Zuständen, sodass die Spins $+\hbar$ und $-\hbar$ mit unterschiedlicher Wahrscheinlichkeit vorkommen und so ein positiver oder negativer Gesamtspin resultiert.

²Zitiert aus: Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag, Wien, München, 2005, Seite 535.

- Analog zu den Wellengleichungen (1) für elektromagnetische Wellen und unter Berücksichtigung von (2) gilt für **nichtrelativistische** (klassische) Teilchen mit einer Ruhemasse m_0 allgemein die **zeitabhängige Schrödinger-Gleichung**

$$\underbrace{-\frac{\hbar^2}{2m_0} \nabla^2 \Psi(\vec{r}, t)}_{\hat{= E_{\text{kin}}}} + \underbrace{V(\vec{r}, t) \Psi(\vec{r}, t)}_{\hat{= E_{\text{pot}}}} = \underbrace{i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t)}_{\hat{= E_{\text{ges}}}} .$$

Für den Fall eines explizit zeitunabhängigen Potentials $V = V(\vec{r})$, d. h. für stationäre Zustände bzw. Gleichgewichtszustände wie beispielsweise beim Atom-Schalenmodell resultiert aus der zeitabhängigen Schrödinger-Gleichung die zeitunabhängige oder **stationäre Schrödinger-Gleichung**

$$-\frac{\hbar^2}{2m_0} \nabla^2 \Psi(\vec{r}) + V(\vec{r}) \cdot \Psi(\vec{r}) = E \cdot \Psi(\vec{r}) .$$

- Die Lösungen Ψ der Schrödinger-Gleichung sind normierte quadratintegrale *komplexe* Wellenfunktionen (Zustände), die auch **Wahrscheinlichkeitsamplituden** genannt werden, nicht beobachtbar bzw. nicht messbar sind und demzufolge keine physikalische Realität besitzen. Die entsprechende beobachtbare Messgröße ist die (*reelle*) **Wahrscheinlichkeitsdichte** $|\Psi|^2$. Die unter den gegebenen physikalischen Bedingungen (insbesondere $V(\vec{r}, t)$) mit Hilfe der Schrödinger-Gleichung ermittelten Wahrscheinlichkeitsamplituden ermöglichen mit ihrem Betragsquadrat die räumliche Darstellung der Wahrscheinlichkeitsverteilung der Energie – ggf. auch in Abhängigkeit von der Zeit. So entsprechen sich im Raum ausbreitende Wellenpakete zeitabhängigen Zuständen, während man sich stationäre Zustände als stehende Wellen veranschaulichen kann. Die Wahrscheinlichkeitsamplitude sagt also etwas aus über die Aufenthalts- oder Antreffwahrscheinlichkeit eines Teilchens im Bereich des ihm entsprechenden Wellenpakets bzw. Wellenzugs.
- **Zusammenhang zwischen Beugung (Wellenbild) und Heisenberg'scher Unschärferelation (Teilchenbild)**

Betrachten wir die **Beugung** von Licht der Wellenlänge λ **am Einfachspalt** der Breite b , so gilt im Wellenbild für den Beugungswinkel α bis zum ersten Minimum:

$$b \cdot \sin \alpha = \lambda \quad \Leftrightarrow \quad \sin \alpha = \frac{\lambda}{b} .$$

Diese Gleichung lässt sich mit Hilfe der De Broglie-Beziehung auch durch den Impuls $\vec{p} = \hbar \vec{k}$, also im Teilchenbild ausdrücken. Dafür nehmen wir an, dass der Spalt längs der y -Achse verläuft, sich die Spaltbreite folglich längs der x -Achse erstreckt, und dass sich das einfallende Licht mit dem Wellenvektors $\vec{k} = (0, 0, k_z)$ längs der z -Achse ausbreitet. Dann gilt bezüglich des ersten Minimums mit dem Wellenvektor $\vec{k} = (k_x, 0, k_z)$ des ausfallenden Lichts in der (z, x) -Ebene hinter dem Spalt:

$$k_x = |\vec{k}| \sin \alpha = k \sin \alpha = \frac{2\pi}{\lambda} \sin \alpha = \frac{2\pi}{\lambda} \frac{\lambda}{b} = \frac{2\pi}{b} .$$

Die daraus resultierende de-Broglie-Beziehung lautet

$$p_x = \hbar k_x = \frac{\hbar}{2\pi} \frac{2\pi}{b} = \frac{h}{b} \quad \Leftrightarrow$$

$$p_x \cdot b = h .$$

„ b kann man als eine Positionsfestlegung der durchgehenden Photonen auf der x -Achse auf den Bereich $\Delta x = b$ interpretieren. Δp_x andererseits kann man als eine Unbestimmtheit des Impulses p_x in Bezug auf die x -Richtung interpretieren, da es ja eine vertikale Ablenkung von der geradlinigen Fortbewegung beschreibt. Der Grund dafür liegt darin, dass jeder Punkt auf dem „Beugungsfleck“ als möglicher Auftreffpunkt eines Photons in Frage kommt. D.h. der Beugungsfleck engt im Wesentlichen den Bereich ein, in dem das Photon statistisch betrachtet auftreten kann. In diesem Sinne kann man obige Gleichung auch so schreiben:“³

$$\Delta p_x \cdot \Delta x \approx h = 2\pi\hbar.$$

Dies entspricht der **Heisenberg’schen Impuls-Orts-Unschärferelation**

$$\Delta p_x \cdot \Delta x \geq \frac{h}{4\pi} = \frac{\hbar}{2}.$$

Die Lichtbeugung im Wellenbild entspricht, wie man sieht, der Heisenberg’schen Unschärferelation im Teilchenbild. Dies zeigt sich in der Tatsache, dass sich das „Beugungsbild“ aufspreizt, wenn der Spalt schmaler gewählt wird. Mit anderen Worten, wenn die Breite $b = \Delta x$ des Spalts und damit auch die Unbestimmtheit der Photonen im Ort kleiner werden, wird im Gegenzug die Unbestimmtheit Δp_x des Impulses der Photonen größer. Die Unbestimmtheit im Ort und im Impuls der Teilchen verhalten sich in bestimmter Weise reziprok zueinander.

Der Vollständigkeit halber geben wir noch die **Heisenberg’sche Energie-Zeit-Unschärferelation** an:

$$\Delta E \cdot \Delta t \geq \frac{\hbar}{2}.$$

³Mit kleinen Änderungen der Notation zitiert aus: Josef Oswald, *Grundprinzipien der Quantenphysik*, Vorlesung – WS03/04, Seite 9, https://oswald.unileoben.ac.at/qpv1_1/qpv1.pdf

2 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 $\omega = c_0 \cdot k \quad \Leftrightarrow \quad \frac{1}{c_0 k} = \frac{1}{\omega}$ 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 die¹

$$\text{Intensität } I = \langle |\vec{S}| \rangle = \varepsilon_0 c_0^2 \langle |\vec{E} \times \vec{B}| \rangle = \varepsilon_0 c_0 \langle E^2 \rangle .$$

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

$$I = \frac{1}{2} \varepsilon_0 c_0 A^2 . \quad (4)$$

Sinngemäß ist die Intensität die pro Zeiteinheit durch eine Flächeneinheit fließende zeitlich gemittelte Energie. Sie besitzt demzufolge die Maßeinheit

$$[I] = 1 \frac{\text{J}}{\text{s} \cdot \text{m}^2} = 1 \text{ W/m}^2 .$$

¹Die Intensität wird in der Optik auch **Bestrahlungsstärke** genannt.

Betrachten wir die Intensität einer elektromagnetischen Welle innerhalb eines homogenen, isotropen Dielektrikums (Mediums) mit dem Brechungsindex n , so wird in (4) 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

- 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 11.2.1) mit den Maxima

$$I_{\max} = \frac{1}{2} \varepsilon_0 c_0 (A_1 + A_2)^2 \text{ für } \delta = m \cdot 2\pi, m = 0, 1, 2, \dots$$

und den Minima

$$I_{\min} = \frac{1}{2} \varepsilon_0 c_0 (A_1 - A_2)^2 \text{ für } \delta = (2m + 1) \cdot \pi, m = 0, 1, 2, \dots$$

- 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 $\vec{k} \uparrow\uparrow \vec{r}$ gilt, können wir statt $\vec{k} \cdot \vec{r}$ vereinfachend kr schreiben. Für die Intensität bzw. Bestrahlungsstärke hatten wir bereits festgestellt:

$$I \propto |\vec{E}|^2 = E^2 . \quad (5)$$

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 Abstandsquadrat-Gesetz gilt, also

$$I \propto \frac{1}{r^2} ,$$

müssen wegen (5) 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.

3 Absorption und Emission elektromagnetischer Strahlung

3.1 Schalenmodell und Orbitalmodell

Literaturtipps:

- Holger Hoffmeister, Freies Lehrbuch „Anorganische Chemie“, Kapitel 19 Das Orbitalmodell als Atommodell, 2020
<https://hoffmeister.it/index.php/chemiebuch-anorganik/227-freies-lehrbuch-anorganische-chemie-19-das-orbitalmodell-als-atommodell>
- 5. Energiezustände von Atomen und Atomorbitale
<https://ethz.ch/content/dam/ethz/special-interest/chab/physical-chemistry/ultrafast-spectroscopy-dam/documents/lectures/allgemeinechemieHS17/script/Kapitel%205.pdf>

Wir betrachten im Folgenden nicht Moleküle sondern vereinfachend und ausgehend vom Periodensystem der Elemente nur Atome. Atome besitzen genauso viele Elektronen wie Protonen, sind also nicht ionisiert und demzufolge nach außen elektrisch neutral. Die **Ordnungszahl** der Elemente repräsentiert die Anzahl der Protonen im Atomkern und damit auch die Anzahl der Elektronen in der zugehörigen Atom-Elektronenhülle. Die (potentielle) **Energie** kernferner Elektronen ist größer als die Energie kernnaher Elektronen. Deshalb lassen sich kernferne Elektronen („Leuchtelektronen“) leichter, d. h. mit weniger Energieaufwand aus der Elektronenhülle herauslösen. Mit anderen Worten, näher am Atomkern befindliche Elektronen sind stärker an den Atomkern gebunden als kernferne. Demzufolge sind für die Absorption und Emission elektromagnetischer Strahlung im sichtbaren Bereich (Licht, Photonen vergleichsweise niedriger Energie) die äußeren Schalen der Elektronenhülle der Atome verantwortlich.

- Das **Periodensystem der Elemente** besteht aus acht Hauptgruppen und sieben Perioden. Im **Schalenmodell** der Atom-Elektronenhülle entspricht jede neue Periode einer zusätzlichen Schale. Elemente der ersten Periode besitzen also nur eine Schale, die K-Schale, Elemente der zweiten Periode besitzen zwei Schalen, die K-Schale und die L-Schale, usw. Die siebente und letzte Periode besitzt demzufolge sieben Schalen: K-, L-, M-, N-, O-, P- und Q-Schale. In die innerste Schale, die K-Schale, passen maximal nur zwei Elektronen. In alle anderen Schalen passen maximal acht Elektronen (Oktettregel).

Betrachtet werden hierbei die Atome der Elemente in ihrem **Grundzustand**, d. h. im Zustand geringstmöglicher Energie, also nicht in irgendeinem angeregten und schon gar nicht im ionisierten Zustand.

- Den Schalen entsprechen bestimmte Energiebereiche, wobei die Schalenenergie prinzipiell von der innersten (K-Schale) zur äußersten (Q-Schale) ansteigt. Allerdings überlappen sich die Energiebereiche der Schalen ab der M-Schale (ab der 3. Schale) aufwärts. So befinden sich die fünf 3d-Orbitale der M-Schale auf einem höheren Energieniveau als das 4s-Orbital der N-Schale.
- Die Besetzung der Schalen mit Elektronen erfolgt auf bestimmte Weise strukturiert in Unterschalen. Diese Unterschalen werden auch Orbitale genannt, woraus das **Orbitalmodell** der Atom-Elektronenhülle resultiert. Orbitale sind die Raumbereiche, in denen sich Elektronen mit 95%iger Wahrscheinlichkeit „aufhalten“. Orbitale repräsentieren Lösungen (Wahrscheinlichkeitswellenfunktionen $\Psi(\vec{r})$) der Schrödinger-Gleichung, denn $|\Psi|^2$ ist die Wahrscheinlichkeitsdichte und somit ein Maß für die Wahrscheinlichkeit, ein Elektron zu einer bestimmten Zeit an einem bestimmten Ort anzutreffen.

Aber Achtung!

Wegen der Heisenberg'schen Unschärferelation besitzen die Aufenthaltsbereiche der Elektronen der jeweiligen Orbitale keine feste bzw. keine scharfe Grenze.

- **Orbitale** entsprechen quantenmechanischen Zuständen der Elektronen und werden durch vier Quantenzahlen charakterisiert: Hauptquantenzahl n , Neben- oder Drehimpulsquantenzahl l , Magnetquantenzahl m und magnetische Spinquantenzahl m_l oder auch s .

1. Den sieben Schalen K bis Q (von innen nach außen) werden die **Hauptquantenzahlen** $n=1$ bis $n=7$ zugeordnet.
2. Die **Nebenquantenzahlen** (Drehimpulsquantenzahlen, azimutale Quantenzahlen) l sind ganze Zahlen von $l = 0$ bis $l = (n - 1)$. Sie bestimmen den Bahndrehimpuls der Elektronen und die (geometrische) Form der Orbitale:

ausgeschrieben	Orbitale	Nebenquantenzahl	Form
sharp	s-Orbital	$l = 0$	Kugel
principal	p-Orbitale	$l = 1$	Hantel
diffuse	d-Orbitale	$l = 2$	gekreuzte Doppelhantel
fundamental	f-Orbitale	$l = 3$	Rosette
	g-Orbitale	$l = 4$	
	h-Orbitale	$l = 5$	
	i-Orbitale	$l = 6$	

3. Die **Magnetquantenzahl** m eines Orbitals kann die ganzen Zahlen von $m = -l$ bis $m = l$ einschließlich $m = 0$ annehmen, also

$$m = -l, -(l-1), \dots, 0, \dots, (l-1), l.$$

Manchmal wird für die Magnetquantenzahl auch das Symbol m_l verwendet. Sie beschreibt die räumliche Ausrichtung, die ein Orbital bezüglich eines äußeren Magnetfeldes einnimmt und tritt demzufolge nur in Erscheinung bzw. ist nur messbar, wenn von außen ein Magnetfeld angelegt wird.

4. Der Spin eines Elektrons entspricht gleichsam einem inneren Drehsinn des Elektrons, wird deshalb auch Eigen- oder intrinsischer Drehimpuls genannt und besitzt die Spinquantenzahl $s = \frac{1}{2}$. Bezüglich eines äußeren Magnetfeldes können Elektronen zwei Drehsinne einnehmen, der eine entspricht dem

Spin-up (Symbol \uparrow) mit der **magnetischen Spinquantenzahl** $m_s = +\frac{1}{2}$, der entgegengesetzte Drehsinn entspricht dem

Spin-down (Symbol \downarrow) mit **magnetischer Spinquantenzahl** $m_s = -\frac{1}{2}$. Wie die Ausrichtung des Bahndrehimpulses (Magnetquantenzahl m) tritt auch die Spinausrichtung (magnetische Spinquantenzahl m_s) nur bei Vorhandensein eines äußeren Magnetfeldes in Erscheinung. Manchmal wird für die magnetische Spinquantenzahl einfach das Symbol s verwendet.

- Leider existiert keine mathematische Formel zur Ermittlung der Elektronenkonfiguration der Atome. Es gibt nur die folgenden Regeln für die sukzessive Besetzung der Schalen und Orbitale mit Elektronen, wobei man vom niedrigsten Energieniveau des einen Elektrons des Wasserstoffatoms ausgeht und dann die Besetzung mit zunehmender Höhe der Energieniveaus der Elektronen der folgenden Elemente fortsetzt. Entscheidend für die Besetzung ist also nicht die ansteigende Ordnungszahl der Elemente sondern die Zunahme der Höhe der Energieniveaus.

Aufbauprinzip:

Jedes neu hinzugefügte Elektron „sucht sich“ den Zustand geringster Energie.

Hund'sche Regel¹ :

Orbitale gleicher Energie werden immer nacheinander zuerst einzeln, d. h. mit nur einem Elektron besetzt. Dabei haben diese ungepaarten Elektronen alle den gleichen Spin. Erst danach erfolgt die Doppelbesetzung.

Pauli'sches Ausschlussprinzip (Pauli-Verbot) :

In einem Atom können sich niemals zwei Elektronen in demselben quantenmechanischen Zustand befinden. Sie müssen sich immer wenigstens in einer Quantenzahl unterscheiden.

Die Anwendung dieser drei Regeln führt zum **Madelung-Schema** der Elektronenkonfigurationen der Elemente

- Gemeinsam haben Aufbauprinzip, Hund'sche Regel und Pauli-Verbot zur Folge, dass Orbitale höchstens mit zwei (gepaarten) Elektronen unterschiedlichen Spins besetzt werden können.

So besitzen die Elemente der zweiten Nebengruppe die K-Schale ($n = 1$) und die L-Schale ($n = 2$). In der K-Schale gibt es nur eine Unterschale, das 1s-Orbital ($n = 1, l = 0$). In der L-Schale gibt es zwei Unterschalen, nämlich das 2s-Orbital ($n = 2, l = 0$) und die 2p-Unterschale ($n = 2, l = 1$).

Die 2p-Unterschale wiederum besteht aus den folgenden drei 2p-Orbitalen :

$$\begin{aligned} 2p_x\text{-Orbital} & \quad (n = 2, l = 1, m = 1), \\ 2p_y\text{-Orbital} & \quad (n = 2, l = 1, m = -1), \\ 2p_z\text{-Orbital} & \quad (n = 2, l = 1, m = 0). \end{aligned}$$

Beispielsweise hat **Stickstoff** mit der Ordnungszahl 7 die **Elektronenkonfiguration**

$$1s^2 2s^2 2p^3 \Rightarrow \begin{cases} 2 \text{ Elektronen im } 1s\text{-Orbital bzw. in der } 1s\text{-Unterschale,} \\ 2 \text{ Elektronen im } 2s\text{-Orbital bzw. in der } 2s\text{-Unterschale,} \\ 3 \text{ Elektronen in der } 2p\text{-Unterschale.} \end{cases}$$

Wie man sieht, sind hierbei die hochgestellten Zahlen keine Exponenten sondern die Anzahl der Elektronen, mit denen die entsprechende Unterschale besetzt ist. 3 Elektronen in der 2p-Unterschale des Stickstoffatoms heißt, dass die drei 2p-Orbitale $2p_x$, $2p_y$ und $2p_z$ jeweils mit nur einem Spin-up-Elektron besetzt sind.

- **Alle Orbitale einer bestimmten Unterschale, d. h. alle Orbitale mit gleicher Hauptquantenzahl n und zusätzlich gleicher Nebenquantenzahl l besitzen das gleiche Energieniveau und können doppelt besetzt sein.** Dieser Sachverhalt führt zu der folgenden Tabelle:

Unterschale	$2l + 1$ Orbitale	maximale Elektronenzahl auf gleichem Energieniveau
s $\Rightarrow l = 0$	1 s-Orbital	2
p $\Rightarrow l = 1$	3 p-Orbitale	6
d $\Rightarrow l = 2$	5 d-Orbitale	10
f $\Rightarrow l = 3$	7 f-Orbitale	14
g $\Rightarrow l = 4$	9 g-Orbitale	18
$\vdots \Rightarrow \vdots$	$\vdots \quad \vdots$	\vdots

¹In der Physik werden vier Hund'sche Regeln verwendet, insbesondere wenn der Gesamtdrehimpuls J eine Rolle spielt.

Die Unterschalen der Elektronenhüllen von Atomen sind bezüglich ihrer **Energieniveaus** entartet. Man spricht von **Entartung** eines quantenmechanischen Systems wie z. B. einer Unterschale, wenn zwei oder mehr

Zustände (linear unabhängige Eigenzustände wie z. B. die Orbitale) zum selben

Messwert (Eigenwert einer Observablen wie z. B. der Energie) existieren.² So existieren beispielsweise in der d-Unterschale fünf d-Orbitale gleicher Energie bzw. mit dem gleichen Energieniveau der Elektronen, wobei jedes dieser fünf Orbitale dem Pauli-Verbot entsprechend mit zwei Elektronen+ unterschiedlichen Spins besetzt werden kann. In der d-Unterschale können sich also maximal zehn Elektronen gleicher Energie aber mit verschiedenem Satz an Quantenzahlen aufhalten. Demzufolge ist die d-Unterschale 10-fach entartet, d. h. sie besitzt den Entartungsgrad bzw. den Entartungsfaktor 10.³

- Die Elektronen in der äußersten Schale (Valenzschale) eines Atoms heißen **Valenzelektronen**. Wie bereits dargestellt besitzt beispielsweise Stickstoff fünf Valenzelektronen: zwei Elektronen im 2s-Orbital und insgesamt drei Elektronen in den 2p-Orbitalen $2p_x$, $2p_y$ und $2p_z$. Weil sich die Elektronen der äußersten Schale leichter durch elektrodynamische Wechselwirkungen beeinflussen lassen, sind die Valenzelektronen wesentlich für die chemischen Eigenschaften der Elemente verantwortlich. Die äußersten Schalen der Edelgase sind mit acht Elektronen voll besetzt: zwei Elektronen im s-Orbital und insgesamt sechs Elektronen in den drei p-Orbitalen p_x , p_y und p_z . Diese Elektronenkonfiguration mit vollständiger Besetzung aller Schalen heißt **Edelgaskonfiguration**. Edelgase sind chemisch besonders stabil. Sie kommen deshalb nicht in Molekülform sondern nur atomar vor.

²Man geht mit dem Entartungsbegriff relativ salopp um. So spricht man auch von entarteten Eigenwerten (Messwerten) und entarteten Zuständen (Eigenzuständen).

³Siehe: Franz Schwabl, Quantenmechanik (QM I), 6. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2002, Tabelle 13.3. Entartungsgrad der ersten Schalen, Seite 252.

3.2 Einstein-Koeffizienten und Übergangswahrscheinlichkeiten

Wir betrachten im Folgenden ein (stationäres) System aus N Teilchen bzw. N Atomen im **thermodynamischen Gleichgewicht** mit einem elektromagnetischen Strahlungsfeld (Photonenfeld, Photonengas oder auch Vakuumfeld) bei konstanter Temperatur T und mit **konstantem Volumen** V . Ein sich daraus ergebendes praktikables Modell zur Herleitung der Einstein-Koeffizienten mit den entsprechenden Schlussfolgerungen sind N Atome im Innern eines **Hohlraumresonators**. Dieses Modell ermöglicht uns die Verwendung der sog. Maxwell-Boltzmann-Verteilung und des Planck'schen Strahlungsgesetzes.⁴

Im Wellenbild wird das Photonenfeld innerhalb des Hohlraumresonators durch optische Eigenschwingungen bestimmter Frequenz ν dargestellt. Diese Eigenschwingungen werden auch **Moden** genannt und lassen sich durch stehende Wellen veranschaulichen. Die Moden können quantenmechanisch als harmonische Oszillatoren mit der Energie

$$E_n = \left(n + \frac{1}{2}\right) h\nu, \quad n = 0, 1, 2, 3, \dots \quad (6)$$

angesehen werden. Dabei ist n die Anzahl der Photonen in der betrachteten Mode. Im thermodynamischen Gleichgewicht ist die mittlere Anzahl \bar{n} der Photonen pro Mode gleich dem Bose-Faktor:

$$\bar{n} = \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}.$$

Im Gegensatz zum klassischen harmonischen Oszillator besitzt der quantenmechanische harmonische Oszillator und somit auch jede Mode für $n = 0$ die Grundzustands- oder **Nullpunktenergie** $E_0 = \frac{1}{2} h\nu$. Und aus (6) folgt, dass diese „Nullpunktsschwingungen“ für alle zulässigen Werte von ν existieren.

- **Maxwell-Boltzmann-Verteilung**

Die klassische Maxwell-Boltzmann-Verteilung besagt, dass in einem System aus N freien Teilchen bei der Temperatur T **im Mittel**

$$N_i = K \cdot e^{-\frac{E_i}{k_B T}}$$

Teilchen jeweils die Energie E_i besitzen. Der Proportionalitätsfaktor $K = K(T)$ ist dabei allgemein eine temperaturabhängige Konstante.

Berücksichtigen wir jetzt bezüglich der Maxwell-Boltzmann-Verteilung das (Elektronen-)Schalenmodell der Atome mit dem sich daraus ergebendem Entartungsgrad g_i , d. h. die (maximal mögliche) Besetzungszahl des i -ten Zustands der Energie E_i mit Elektronen, resultiert das **Boltzmann'sche Energieverteilungsgesetz**

$$N_i \propto g_i \cdot e^{-\frac{E_i}{k_B T}}.$$

Wie man sieht, ist die Anzahl N_i der Atome, die sich im Energiezustand E_i befinden, *proportional* zum Entartungsgrad g_i . Der Entartungsgrad g_i ist also das statistische Gewicht für die Zustände der Energie E_i .

⁴Erläuterungen zur Maxwell-Boltzmann-Verteilung und zum Planck'schen Strahlungsgesetz sowie die zugehörigen Herleitungen finden sich u. a. bei Wikipedia und in meinem Skript

Grundlegendes zur Statistischen Physik – Ensembles und Verteilungen in den Abschnitten 11.1 Maxwell-Boltzmann-Verteilung, 13.2 Was ist der Boltzmann-Faktor $e^{-E_T/(k_B T)}$, 18.2 Hohlraumresonator und Mode – Planck'sches Strahlungsgesetz und 18.6 Strahlungsleistung des schwarzen Körpers.

Mit der kanonischen Zustandssumme $Z_k = \sum_i g_i \cdot e^{-\frac{E_i}{k_B T}}$, erhalten wir schließlich die Wahrscheinlichkeit P_i für das Vorhandensein eines Atoms im Energiezustand E_i wie folgt:

$$P_i = \frac{N_i(E_i)}{N} = \frac{g_i \cdot e^{-\frac{E_i}{k_B T}}}{Z_k}.$$

Für eine bestimmte Temperatur T gilt folglich

$$\frac{N_j}{N_i} = \frac{\tilde{K} \cdot g_j \cdot e^{-\frac{E_j}{k_B T}}}{\tilde{K} \cdot g_i \cdot e^{-\frac{E_i}{k_B T}}} = \frac{g_j}{g_i} \cdot e^{-\frac{\Delta E}{k_B T}}, \quad \Delta E = E_j - E_i \quad (7)$$

mit dem

$$\text{Boltzmann-Faktor} \quad e^{-\frac{\Delta E}{k_B T}}.$$

- Das **Planck'sche Strahlungsgesetz** liefert die **spektrale Energiedichte** u :

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{e^{h\nu/(k_B T)} - 1} \quad \text{bzw.} \quad (8)$$

$$u(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \cdot \frac{1}{e^{\hbar\omega/(k_B T)} - 1} \quad \text{bzw.}$$

$$u(\lambda, T) = \frac{8\pi hc}{\lambda^5} \cdot \frac{1}{e^{hc/(\lambda k_B T)} - 1},$$

$$[u] = \frac{(\text{Energie in J})}{(\text{Volumen in m}^3 \cdot (\text{Frequenz in } \frac{1}{s}))} = \frac{\text{J} \cdot \text{s}}{\text{m}^3} = \frac{\text{kg}}{\text{m} \cdot \text{s}}.$$

Die spektrale Energiedichte *entspricht* also der Photonendichte pro Frequenzeinheit. Die Terme $\frac{1}{e^{h\nu/(k_B T)} - 1}$ und dessen Äquivalente sind der Bose-Faktor.

Wir werden im Folgenden die spektrale Modendichte $G(\nu)$ des Hohlraumresonators verwenden. Diese hängt mit der spektralen Energiedichte auf folgende Weise zusammen:

$$u(\nu) = \underbrace{\frac{8\pi \nu^2}{c^3}}_{G(\nu)} \cdot \underbrace{\frac{h\nu}{e^{h\nu/(k_B T)} - 1}}_{\bar{E}_{\text{mod}}}.$$

Dabei ist \bar{E}_{mod} die mittlere Modenenergie und die **spektrale Modendichte** ist

$$\boxed{G(\nu) = \frac{8\pi \nu^2}{c^3}}. \quad (9)$$

Schließlich benötigen wir auch noch den Zusammenhang zwischen der spektralen Energiedichte u und der **spektralen Strahlungsintensität**

$$\boxed{I'(\nu) = \frac{2\pi h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} = \frac{c}{4} \cdot u(\nu)} \quad (10)$$

des schwarzen Körpers bzw. Hohlraumresonators. Die *spektrale Strahlungsintensität* wird in der Optik auch *spektrale Bestrahlungsstärke* genannt und besitzt die Maßeinheit

$$[I'] = \frac{(\text{Energie in J})}{(\text{Fläche in m}^2 \cdot (\text{Zeit in s}) \cdot (\text{Frequenz in } \frac{1}{s}))} = \frac{\text{J}}{\text{m}^2} = \frac{\text{kg}}{\text{s}^2}.$$

Wenden wir jetzt die Maxwell-Boltzmann-Verteilung auf die N Atome eines Gases in einem Raumbereich an, der die Bedingungen unseres anfangs definierten Modells erfüllt. Dabei betrachten wir die energetischen Übergänge zwischen dem Grundzustand und dem angeregten Zustand eines Atoms. Wir sprechen in diesem Fall von einem **Zwei-Niveau-System**. Den Grundzustand verstehen wir mit dem tiefgestelltem Index 1 und den angeregten Zustand mit dem tiefgestelltem Index 2.

Anmerkung

Man kann statt der Teilchenzahl N in einem festen Volumen auch allgemein die Teilchendichte für die Herleitungen verwenden, was aber letztlich zu den gleichen Ergebnissen führt.

Die Gesamtanzahl der Atome ist also $N = N_1 + N_2$ mit der Anzahl N_1 von Atomen im Grundzustand und der Anzahl N_2 von Atomen im angeregten Zustand.

Im **thermodynamischen Gleichgewicht** ist die Rate der Atome, die durch das Photonenfeld energetisch angeregt werden, genauso groß wie die Rate der Atome, die ihre Anregungsenergie wieder abstrahlen. Dabei spielen sich stets gleichzeitig die folgenden, in der Abbildung 1 veranschaulichten Photonen-Absorptions- und -Emissionsprozesse ab:

- **(Induzierte) Absorption**

Springen Elektronen mit der Energie E_1 auf das höhere Energieniveau E_2 beim Übergang der zugehörigen Atome vom Grundzustand in den angeregten Zustand durch die Zufuhr von jeweils einem Photon der Energie $h\nu_{12}$ pro Atom, so ist mit (7) die Anzahl der angeregten Atome im Mittel

$$N_2 = \tilde{K} \cdot g_2 \cdot e^{-\frac{E_2}{k_B T}}.$$

Die Anzahl der Atome auf dem 1-ten Energieniveau ist dann $N_1 = N - N_2$. Und wegen $E_2 > E_1$ gilt $N_2 < N_1$. Das Verhältnis N_2/N_1 wird auch **relative Besetzungszahl** genannt:

$$\frac{N_1}{N_2} = \frac{\tilde{K} \cdot g_1 \cdot e^{-\frac{E_1}{k_B T}}}{\tilde{K} \cdot g_2 \cdot e^{-\frac{E_2}{k_B T}}} = \frac{g_1}{g_2} \cdot e^{\frac{E_2 - E_1}{k_B T}}, \quad E_2 - E_1 = h\nu_{12} \quad \Rightarrow$$

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \cdot e^{-\frac{h\nu_{12}}{k_B T}} \quad \Leftrightarrow \quad N_2 = \frac{g_2}{g_1} \cdot N_1 \cdot e^{-\frac{h\nu_{12}}{k_B T}}.$$

(11)

Die „Triebkraft“ der induzierten Absorption ist das Photonenfeld. Die Übergangsrate, d. h. die Anzahl der Atome, die dabei pro Zeiteinheit in diesem Fall energetisch angeregt werden, hängt nämlich von der spektralen Energiedichte u des Photonenfeldes im Frequenzbereich des Übergangs ab gemäß

$$\frac{dN_1}{dt} \propto u(\nu).$$

Außerdem gilt: Je größer die Anzahl N_1 der Atome im Grundzustand ist, desto mehr Atome können in den angeregten Zustand (mit der Elektronenenergie E_2) übergehen gemäß

$$\frac{dN_1}{dt} \propto N_1.$$

Mit dem Proportionalitätsfaktor B_{12} erhalten wir daraus schließlich die **Übergangsrate**

$$\left(\frac{dN_1}{dt} \right)_{\text{ind.A.}} = -B_{12} \cdot N_1 u \quad \Rightarrow \quad [B_{12}] = \frac{\text{m}}{\text{kg}}.$$

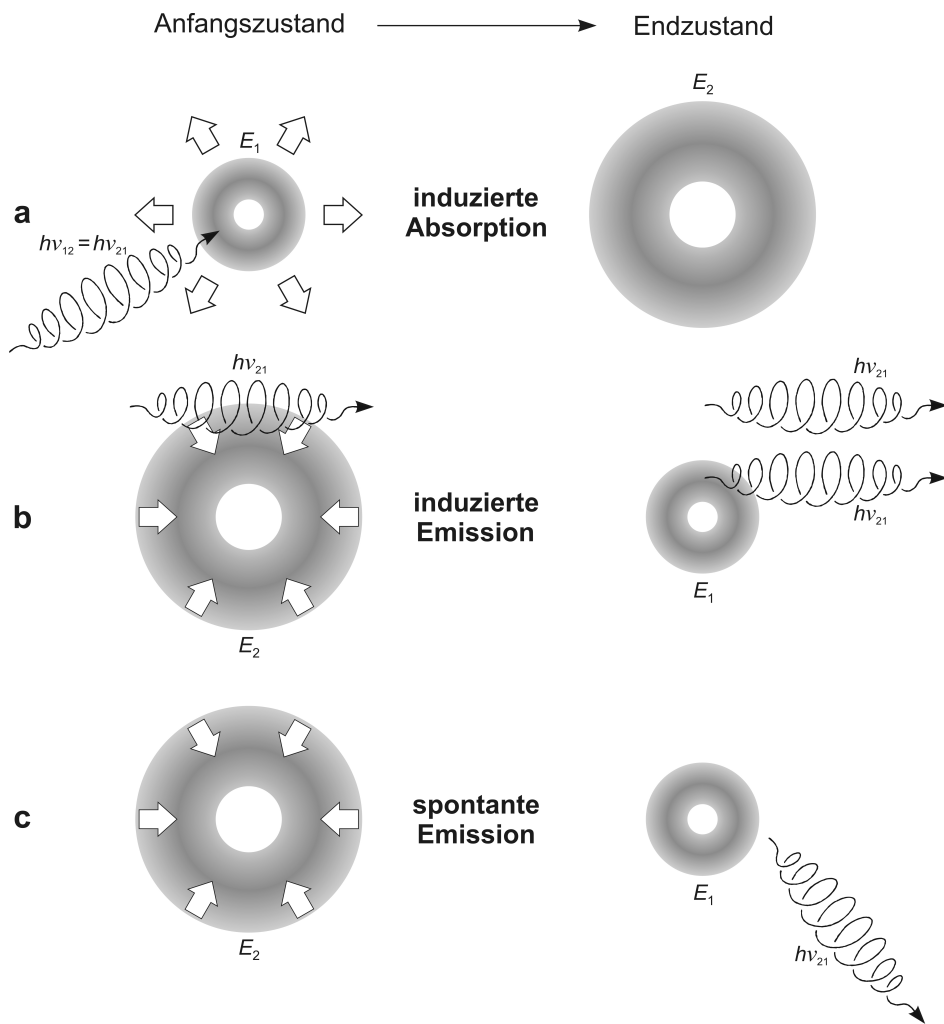


Abb. 1 Veranschaulichung **a)** der induzierten Absorption, **b)** der induzierten Emission und **c)** der spontanen Emission am Beispiel des Übergangs zwischen dem Grundzustand des Wasserstoffatoms mit der Elektronenkonfiguration $1s^1$ (Elektronenenergie E_1) und seinem ersten angeregten Zustand mit der Elektronenkonfiguration $2s^2$ (Elektronenenergie E_2). Es handelt sich dabei um den Übergang zwischen L- und K-Schale aus der Lyman-Serie.

Nach Eugene Hecht, Optik, 4. Auflage, Oldenbourg Verlag, München, Wien, 2005, Seite 943.

B_{12} ist hier der **Einstein-Koeffizient** für die induzierte Absorption und das Minuszeichen steht für die Abnahme von N_1 .

Der Term $B_{12} \cdot u$ besitzt die „Dimension“ $1/s$ und beschreibt die **Wahrscheinlichkeit** für *einen* Absorptionsübergang pro Zeiteinheit (also für *ein* Atom im Grundzustand bei fester Gesamtzahl N der Atome).

- **Induzierte Emission**

Umgekehrt erfolgt der Übergang der angeregten Atome in den nicht angeregten Ausgangszustand bzw. Grundzustand indem die Elektronen des (höheren) 2-ten Energieniveaus jeweils ein Photon der Energie $h\nu_{21} = h\nu_{12}$ abstrahlen und so zurück in das 1-te Energieniveau fallen.

Die induzierte (stimulierte, erzwungene) Emission setzt genau wie die (induzierte) Absorption die Einwirkung eines Photonenfeldes voraus und ist der

Umkehrprozess zur Absorption.

Mit dem entsprechenden **Einstein-Koeffizienten** B_{21} erhalten wir somit für den Übergang vom angeregten Zustand in den Grundzustand die **Übergangsrate**

$$\left(\frac{dN_2}{dt} \right)_{\text{ind.E.}} = -B_{21} \cdot N_2 u \quad \Rightarrow \quad [B_{21}] = \frac{\text{m}}{\text{kg}}.$$

Das aus dem Photonenfeld stammende und die induzierte Emission stimulierende Photon besitzt ebenfalls die Anregungsenergie $h\nu_{12}$ und stellt in der Emissionsbilanz einen „durchlaufenden Posten“ dar. Es nimmt nämlich dabei ein Photon der gleichen Energie $h\nu_{21} = h\nu_{12}$ mit. Außerdem ist es bemerkenswert, dass bei der induzierten Emission das emittierte Photon

- dieselbe Ausbreitungsrichtung (siehe Abbildung 1 b),
- dieselbe Frequenz,
- dieselbe Phase und
- dieselbe Polarisation

besitzt wie das stimulierende Photon. Im Wellenbild sind demzufolge die stimulierende einfallende Welle und die induzierte emittierte Welle **kohärent**. Anders gesagt und beziehend auf den Hohlraumresonator befindet sich das emittierte Photon in *derselben* Strahlungsmoden wie die einfallende Welle.

Der Term $B_{21} \cdot u$ besitzt die „Dimension“ $1/\text{s}$ und beschreibt die **Wahrscheinlichkeit** für *einen* induzierten Emissionsübergang pro Zeiteinheit (also für *ein* angeregtes Atom bei fester Gesamtzahl N der Atome).

• Spontane Emission

Bei der spontanen Emission gibt das angeregte Atom völlig statistisch und **unabhängig vom Photonenfeld** ein Photon der Anregungsenergie $h\nu_{21} = h\nu_{12}$ mit zufälliger Phase, zufälliger Polarisation sowie **beliebiger Ausbreitungsrichtung** ab und fällt dabei in seinen Grundzustand zurück. Die **Übergangsrate** bei der spontanen Emission hängt demzufolge nur von der Anzahl N_2 angeregter Atome ab gemäß

$$\left(\frac{dN_2}{dt} \right)_{\text{sp.E.}} = -A_{21} \cdot N_2 \quad \Rightarrow \quad [A_{21}] = \frac{1}{\text{s}}.$$

Dabei ist A_{21} der **Einstein-Koeffizient** der spontanen Emission.

Der Einstein-Koeffizient A_{21} besitzt die „Dimension“ $1/\text{s}$ und beschreibt die **Wahrscheinlichkeit** für *einen* spontanen Emissionsübergang pro Zeiteinheit (also für *ein* angeregtes Atom).

Das aber bedeutet, dass die mittlere Dauer, bis ein angeregtes Atom spontan (ohne äußere Einwirkung) in den Grundzustand übergeht, durch den Kehrwert von A_{21} bestimmt wird. Diese (mittlere) **Lebensdauer** eines angeregten Zustands ist also

$$\tau = \frac{1}{A_{21}}.$$

Die Lebensdauer τ eines angeregten Zustands unterliegt der Heisenberg’schen Unschärfelation gemäß

$$\Gamma \cdot \tau = \frac{\Gamma}{A_{21}} = \hbar$$

mit der Energieniveaubreite (Spektrallinienebreite) Γ (siehe Abschnitt 6.1.2).

Quantenelektrodynamisch betrachtet ist die Ursache der spontanen Emission die Nullpunktsenergie. Und die Ursache der Nullpunktsenergie wiederum ist die Heisenberg'sche Unschärferelation. Die Energie-Zeit-Unschärfe erlaubt nämlich einer Mode innerhalb eines kurzen Zeitintervalls in einen angeregten Zustand und sofort wieder zurück in den Grundzustand zu springen. Bei diesen sog. Vakuumfluktuationen werden *virtuelle Photonen* emittiert, welche angeregte Atome zu „spontanen“ Übergängen veranlassen können. Insofern kann man die spontane Emission als eine von Vakuumfluktuationen induzierte Emission betrachten. Und weil die Vakuumfluktuationen statistischer Natur sind, ist folglich auch die spontane Emission statistischer Natur.

Während die Boltzmann-Verteilung von der Temperatur abhängt, sind die drei Einstein-Koeffizienten **stoffspezifische Konstanten** und hängen demzufolge *nicht* von äußeren Faktoren wie beispielsweise der Temperatur ab.

Im thermodynamischen Gleichgewicht muss die Gesamtemissionsrate stets gleich der Absorptionsrate sein. Das bedeutet, dass sowohl die Anzahl der angeregten Atome als auch die Anzahl der Atome im Grundzustand zeitlich konstant ist. Wir erhalten demzufolge

$$N_1(t) = N_2(t) = \text{const} \quad \Rightarrow \quad \frac{dN_1}{dt} = -\frac{dN_2}{dt} = 0$$

mit

$$\begin{aligned} \frac{dN_1}{dt} &= \left(\frac{dN_1}{dt} \right)_{\text{ind.A.}} - \left(\frac{dN_2}{dt} \right)_{\text{ind.E.}} - \left(\frac{dN_2}{dt} \right)_{\text{sp.E.}} \\ &= -B_{12} \cdot N_1 u + B_{21} \cdot N_2 u + A_{21} \cdot N_2 = 0 \end{aligned}$$

und

$$\begin{aligned} \frac{dN_2}{dt} &= \left(\frac{dN_2}{dt} \right)_{\text{ind.E.}} + \left(\frac{dN_2}{dt} \right)_{\text{sp.E.}} - \left(\frac{dN_1}{dt} \right)_{\text{ind.A.}} \\ &= -B_{21} \cdot N_2 u - A_{21} \cdot N_2 + B_{12} \cdot N_1 u = 0 . \end{aligned}$$

Durch einfache Äquivalenzumformung resultiert daraus schließlich

$$\boxed{A_{21} \cdot N_2 + B_{21} \cdot N_2 u = B_{12} \cdot N_1 u \quad \Leftrightarrow \quad \frac{N_2}{N_1} = \frac{B_{12} \cdot u}{A_{21} + B_{21} \cdot u}} . \quad (12)$$

Aufgrund der Maxwell-Boltzmann-Verteilung ist im thermodynamischen Gleichgewicht

$$N_2 < N_1 \quad \Rightarrow \quad \frac{N_2}{N_1} < 1 .$$

Wie man an (12) erkennt, ist deshalb die Absorptionswahrscheinlichkeit kleiner als die (Gesamt-)Emissionswahrscheinlichkeit gemäß

$$B_{12} \cdot u < A_{21} + B_{21} \cdot u .$$

Dass trotz kleinerer Absorptionswahrscheinlichkeit die Absorptionsrate gleich der Emissionsrate ist, wird also durch die größere Teilchenzahl N_1 im Grundzustand gewährleistet.

Im thermodynamischen Gleichgewicht verhalten sich Teilchenzahl und zugehörige Übergangswahrscheinlichkeit der verschiedenen Energieniveaus zueinander umgekehrt proportional.

Gleichsetzen von (12) mit (20) und Auflösen nach u liefert

$$\begin{aligned} \frac{N_2}{N_1} &= \frac{B_{12} \cdot u}{A_{21} + B_{21} \cdot u} = \frac{g_2}{g_1} \cdot e^{\frac{-h\nu_{12}}{k_B T}} \quad \Leftrightarrow \\ u = u(\nu) &= \frac{A_{21}}{B_{21}} \cdot \frac{1}{\frac{B_{12}}{B_{21}} \cdot \frac{g_1}{g_2} \cdot e^{\frac{h\nu}{k_B T}} - 1} . \end{aligned} \quad (13)$$

Der Koeffizientenvergleich von (13) mit dem Planck'schen Strahlungsgesetz (8)

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{e^{h\nu/(k_B T)} - 1}$$

zeigt, dass sich (13) durch die gleichzeitige Anwendung der Beziehungen

$$A_{21} = \frac{8\pi h\nu^3}{c^3} B_{21} = G(\nu) \cdot h\nu \cdot B_{21} \quad (14)$$

und

$$B_{12} \cdot g_1 = B_{21} \cdot g_2 \quad \Leftrightarrow \quad B_{21} = \frac{g_1}{g_2} B_{12} \quad (15)$$

in (8) überführen lässt. Diese beiden Beziehungen beschreiben die Zusammenhänge zwischen den Einstein-Koeffizienten und gelten für alle Frequenzen ν und alle Temperaturen T . Bei gleichen statistischen Gewichten (Entartungsgraden) g_1 und g_2 resultiert aus (15)

$$g_1 = g_2 \quad \Rightarrow \quad B_{21} = B_{12} = B . \quad (16)$$

In diesem speziellen Fall sind induzierte Emission und induzierte Absorption gleich wahrscheinlich gemäß $B \cdot u(\nu)$ für *einen* Übergang pro Zeiteinheit.

Betrachten wir jetzt das Verhältnis von induzierter Emission zu spontaner Emission unter Berücksichtigung der spektralen Modendichte $G(\nu) = \frac{8\pi\nu^2}{c^3}$ (9). Mit den Wahrscheinlichkeit $W_{\text{ind.E.}}$ für die induzierte Emission und $W_{\text{spont.E.}}$ für die spontane Emission erhalten wir zunächst

$$\frac{W_{\text{ind.E.}}}{W_{\text{spont.E.}}} = \frac{N_2 \cdot B_{21} \cdot u}{N_2 \cdot A_{21}} = \frac{B_{21} \cdot u}{A_{21}} = \frac{B_{21} \cdot G(\nu) \cdot h\nu \cdot \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}}{A_{21}} . \quad (17)$$

Wenn wir jetzt A_{21} durch (14) ersetzen, so resultiert

$$\frac{W_{\text{ind.E.}}}{W_{\text{spont.E.}}} = \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} . \quad (18)$$

Wie man sieht, verschiebt sich das Verhältnis von induzierter zu spontaner Emission mit zunehmender Temperatur T und bei fester Frequenz ν zugunsten der induzierten Emission, sodass letztere bei hohen Temperaturen klar überwiegt. Zum gleichen Ergebnis kommen wir, wenn wir von der spektralen Strahlungsintensität (10)

$$I'(\nu) = \frac{2\pi h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}$$

des Hohlraumresonators ausgehen und mit der spektralen Energiedichte u vergleichen. Denn wie wir unter Berücksichtigung von (14) und mit (16) feststellen können, sind I' und u proportional gemäß

$$I'(\nu) = \frac{c}{4} \cdot u(\nu) = \frac{c}{4} \cdot \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} = \underbrace{\frac{c}{4} \cdot \frac{A_{21}}{B}}_{=\text{const}} \cdot \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} , \quad g_1 = g_2 .$$

3.3 Die Idee zum Funktionsprinzip des Lasers

Bei der Beschreibung der Idee zum Funktionsprinzip des Lasers folgen wir der Argumentation von M. Alonso und E. J. Finn.⁵

Das Kunstwort „Laser“ bedeutet

„**L**ight **A**mplification by **S**timulated **E**mission of **R**adiation“.

Sinn und Zweck des Lasers ist es, bei kontinuierlicher Zufuhr von Energie in das System „Laser“ eine möglichst kontinuierliche Nettoemission von kohärenter elektromagnetischer Strahlung bzw. von Photonen aus dem System heraus zu erzeugen. Das bedeutet, dass ein Laser **nicht im thermodynamischen Gleichgewicht** „arbeitet“.

Sind Strahlung und Materie im thermodynamischen Gleichgewicht, gilt für unser Zwei-Niveau-System im Hohlraumresonator im Fall $g_1 \stackrel{!}{=} g_2 \Rightarrow B_{21} = B_{12} = B$:

Emissionsrate = Absorptionsrate

$$B_{21} \cdot u(\nu) N_2 + A_{21} N_2 = B_{12} \cdot u(\nu) N_1$$

$$[B \cdot u + A_{21}] N_2 = B \cdot u N_1 \Rightarrow$$

$$\boxed{\frac{\text{Emissionsrate}}{\text{Absorptionsrate}} = \left(1 + \frac{A_{21}}{B \cdot u}\right) \frac{N_2}{N_1} = 1} \quad (19)$$

Ist die Differenz $E_2 - E_1 = h\nu$ zwischen den beiden Niveaus jedoch hinreichend klein, so folgt mit (17) und (18) aus (19)

$$\frac{\text{Emissionsrate}}{\text{Absorptionsrate}} \approx \frac{N_2}{N_1}, \quad (20)$$

denn

$$\frac{W_{\text{spont.E.}}}{W_{\text{ind.E.}}} = \frac{A_{21}}{B \cdot u} = e^{\frac{h\nu}{k_B T}} - 1 \xrightarrow{h\nu \rightarrow 0} \frac{A_{21}}{B \cdot u} \rightarrow 0.$$

Das bedeutet: Wenn im thermodynamischen Gleichgewicht E_2 nur geringfügig größer ist als E_1 , dann ist N_2 nur geringfügig kleiner als N_1 , wobei die Emissionsrate selbstverständlich immer noch gleich der Absorptionsrate ist. Eine Netto-Emission bzw. Lichtverstärkung ist demzufolge im thermodynamischen Gleichgewicht (bis zum Grenzfall $\frac{N_2}{N_1} = 1$) nicht möglich.

Eine Netto-Emission ist also nur im thermodynamischen Ungleichgewicht mit einer **Besetzungsinversion**, d. h. mit einer relativen Besetzungszahl $\frac{N_2}{N_1} > 1$ zu erwarten. Aus (20) folgt dann nämlich

$$N_2 > N_1 \Rightarrow \text{Emissionsrate} > \text{Absorptionsrate}.$$

Diese Besetzungsinversion erreicht man, indem man Atome aus dem Grundzustand (E_1) durch Zufuhr von Energie bzw. Photonen geeigneter Energie in einen angeregten Zustand „pumpt“. Im Fall des Zweiniveaulasers handelt es sich bei diesem angeregten Zustand gleichzeitig um das obere Laserniveau (E_2). Weil Zweiniveaulaser nicht kontinuierlich arbeiten, muss man dafür sorgen, dass durch optisches Pumpen kontinuierlich Atome auf ein Niveau angeregt werden, das oberhalb des oberen Laserniveaus liegt. Im Fall eines Dreiniveaulasers wäre dies der angeregte Zustand (E_3). Der Übergang der angeregten Atome (E_3) in den Grundzustand erfolgt dann zumindest teilweise über die Zwischenstufe, den metastabilen (langsam zerfallenden) Zustand (E_2) bzw. das obere Laserniveau, wodurch sich dann dauerhaft hohe Besetzungszahlen $N_2 > N_1$ mit der daraus resultierenden Netto-Emission realisieren lassen.

⁵Marcelo Alonso und Edward J. Finn, Quantenphysik und Statistische Physik, 4. Auflage, Oldenbourg Verlag, München, Wien, 2005, Seite 577 bis Seite 579.

4 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 $\alpha_1 = \alpha_g$:

$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2 = n_2 \sin 90^\circ = n_2 \quad \Leftrightarrow \quad \sin \alpha_1 = \sin \alpha_g = \frac{n_2}{n_1},$$

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 5).
- 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

$$n_1 \sin \alpha_1 = n_1 \sin \alpha_B = n_2 \sin \alpha_2 = n_2 \sin(90^\circ - \alpha_B) = n_2 \cos \alpha_B \quad \Rightarrow \quad \frac{n_2}{n_1} = \frac{\sin \alpha_B}{\cos \alpha_B} = \tan \alpha_B,$$

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.

5 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**.“¹

Das evaneszente² Feld führt also bei Totalreflexion keine Energie über das optisch dünnere Medium II ab.³

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 innere Totalreflexion genannt.⁴ 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

¹Zitiert aus: ChemgaPedia, Optische Grundlagen der Sensorik – Das evaneszente Feld.

²Evanescere (lat.) bedeutet verschwinden, sich verflüchtigen.

³Dies gilt auch allgemein für die **vollständige Reflexion** an Grenzflächen zwischen zwei Medien.

⁴Im Zusammenhang mit der Transmission bei frustrierter Totalreflexion spricht man auch von einer *Auskopplung* des evaneszenten Feldes.

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:⁵

$$\Psi_I(x) = A \cdot e^{ikx} \quad \Rightarrow \quad \Psi_{II}(x) = C \cdot e^{i(i\kappa)x} = C \cdot e^{-\kappa x}.$$

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 2). 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.

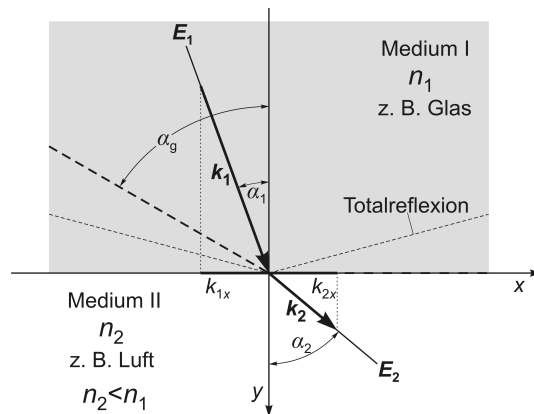


Abb. 2 Die ebene Welle \vec{E}_1 trifft im Medium I mit dem Brechungsindex $n_1 > n_2$ auf die Grenzfläche zum Medium II mit dem Brechungsindex n_2 . Die transmittierte ebene Welle ist dann \vec{E}_2 . Die Grenzfläche ist die Ebene, die von den z, x -Koordinatenachsen aufgespannt wird. Die \vec{E}_1 -Einfallsebene und die \vec{E}_2 -Ausfallsebene liegen in der Ebene der x, y -Koordinatenachsen, sodass die y -Achse das Einfalls- und das Ausfallsot 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 .

⁵Zur Herstellung der Analogie zwischen elektromagnetischen Wellen und den quantenphysikalischen Wahrscheinlichkeitsamplituden setzen wir im Exponenten der elektromagnetischen Wellenfunktion $t = 0$:
 $E(x, t) = A \cdot e^{i(kx - \omega t)} \xrightarrow{t=0} E(x, 0) = A \cdot e^{ikx}.$

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 $E_{1T} + E'_{1T}$ 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):

$$E_{1T} + E'_{1T} = E_{2T} . \quad (21)$$

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

$$\boxed{\omega_1 \equiv \omega'_1 \equiv \omega_2 \equiv \omega} . \quad (22)$$

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 (21) 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:

$$\vec{k}_1 \vec{r} - \omega_1 t = \vec{k}'_1 \vec{r} - \omega'_1 t = \vec{k}_2 \vec{r} - \omega_2 t .$$

Unter Berücksichtigung von (22) folgt daraus die wichtige Beziehung

$$\boxed{\vec{k}_1 \cdot \vec{r} = \vec{k}'_1 \cdot \vec{r} = \vec{k}_2 \cdot \vec{r}} . \quad (23)$$

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

$$\vec{k}_1 = k_{1x} \vec{e}_x + k_{1y} \vec{e}_y , \quad \vec{r} = x \vec{e}_x + z \vec{e}_z . \quad (24)$$

Und weil die Richtungen der Wellenvektoren \vec{k}'_1 und \vec{k}_2 noch unbekannt sind, machen wir den allgemeinen Ansatz

$$\begin{aligned}\vec{k}'_1 &= k'_{1x} \vec{e}_x + k'_{1y} \vec{e}_y + k'_{1z} \vec{e}_z , \\ \vec{k}_2 &= k_{2x} \vec{e}_x + k_{2y} \vec{e}_y + k_{2z} \vec{e}_z .\end{aligned} \quad (25)$$

Einsetzen von (24) und (25) in (23) 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:

$$k'_{1z} = k_{2z} = 0 \quad \Rightarrow$$

$$\boxed{k_{1x} = k'_{1x} = k_{2x}} . \quad (26)$$

Die Wellenvektoren der einfallenden, der reflektierten und der gebrochenen Welle liegen in der Einfallsebene.

Der Abbildung 2 entnehmen wir

$$k_{1x} = |\vec{k}_1| \cdot \sin \alpha_1 = k_1 \cdot \sin \alpha_1 = n_1 \frac{\omega}{c_0} \cdot \sin \alpha_1 , \quad (27)$$

$$k'_{1x} = |\vec{k}'_1| \cdot \sin \alpha'_1 = k'_1 \cdot \sin \alpha'_1 = n'_1 \frac{\omega}{c_0} \cdot \sin \alpha'_1 ,$$

$$k_{2x} = |\vec{k}_2| \cdot \sin \alpha_2 = k_2 \cdot \sin \alpha_2 = n_2 \frac{\omega}{c_0} \cdot \sin \alpha_2 . \quad (28)$$

Mit (26) folgt daraus einerseits

$$k_{1x} = k'_{1x} \quad \wedge \quad n_1 = n'_1 \quad \Rightarrow \quad n_1 \frac{\omega}{c_0} \cdot \sin \alpha_1 = n_1 \frac{\omega}{c_0} \cdot \sin \alpha'_1 \quad \Rightarrow$$

$$\sin \alpha_1 = \sin \alpha'_1 \quad \Rightarrow \quad \alpha_1 = \alpha'_1 , \quad \textbf{Einfallswinkel gleich Reflexionswinkel}$$

und andererseits

$$k_{1x} = k_{2x} \quad \Rightarrow \quad n_1 \frac{\omega}{c_0} \cdot \sin \alpha_1 = n_2 \frac{\omega}{c_0} \cdot \sin \alpha_2 \quad \Rightarrow$$

$$\textbf{Snellius'sches Brechungsgesetz} \quad n_1 \cdot \sin \alpha_1 = n_2 \cdot \sin \alpha_2 .$$

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 (26) und (27) $k_{2x}^2 = k_{1x}^2$ 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 :

$$\begin{aligned} k_{2y}^2 &= \underbrace{\left(n_2 \frac{\omega}{c_0}\right)^2}_{=k_2^2} \cdot (-1) \cdot \underbrace{\left(\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 - 1\right)}_{=\beta^2} \Rightarrow \\ k_{2y} &= \pm k_2 \cdot i \cdot \underbrace{\sqrt{\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 - 1}}_{=\beta}, \\ \boxed{k_{2y} &= \pm i k_2 \beta \quad \text{für} \quad \alpha_1 > \alpha_g} \end{aligned} \quad (29)$$

Jetzt setzen wir die positive Lösung von (29) in die transmittierte ebene Welle \vec{E}_2 ein:

$$\begin{aligned} \vec{E}_2 &= \vec{A}_2 \cdot e^{i(\vec{k}_2 \vec{r} - \omega t)} = \vec{A}_2 \cdot e^{i(k_{2x} \cdot x + k_{2y} \cdot y - \omega t)} = \vec{A}_2 \cdot e^{i(k_{2x} \cdot x + i k_2 \beta \cdot y - \omega t)} \Rightarrow \\ \boxed{\text{evaneszentes Feld} \quad \vec{E}_2 &= \vec{A}_2 \cdot \underbrace{e^{-\beta k_2 \cdot |y|}}_{\text{Dämpfung}} \cdot \underbrace{e^{i(k_{2x} \cdot x - \omega t)}}_{\text{Oszillation}}, \quad \alpha_1 > \alpha_g} \end{aligned}$$

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:⁶

E -Feld:

$$\begin{aligned} e^{-\beta k_2 \cdot |y|} &\stackrel{!}{=} \frac{1}{e} \Rightarrow \ln(e^{-\beta k_2 \cdot |y|}) = -\beta k_2 \cdot |y| = \ln e^{-1} = -1 \Rightarrow |y| \cdot \beta k_2 = 1 \Leftrightarrow \\ |y_{(E/e)}| &= \frac{1}{k_2 \beta}. \end{aligned}$$

Die Intensität I einer ebenen Welle ist proportional zum Betragsquadrat des E -Feldes:

$$\begin{aligned} I &\propto \left(A_2 \cdot e^{-\beta k_2 \cdot |y|}\right)^2 \Rightarrow e^{-2\beta k_2 \cdot |y|} \stackrel{!}{=} \frac{1}{e} \Rightarrow \\ |y_{(I/e)}| &= \frac{1}{2k_2 \beta}. \end{aligned}$$

⁶ $e = 2,718\dots$

Beispiel

Wie groß ist der Abstand $|y_{(I/e)}|$ bei einer Glasplatte in Luft für grünes Licht?

$$\begin{aligned}
 c_0 &= 3 \cdot 10^8 \text{ m s}^{-1}, \\
 f &= 600 \text{ THz} = 6 \cdot 10^{14} \text{ s}^{-1} \text{ (grünes Licht)} \Rightarrow \\
 \omega &= 2\pi \cdot 6 \cdot 10^{14} \text{ s}^{-1}, \\
 n_1 &= 1,5 \text{ (z. B. Glas)}, \\
 n_2 &= 1,0 \text{ (Vakuum oder näherungsweise Luft)}, \\
 \alpha_1 &= 75^\circ \Rightarrow \sin \alpha_1 \approx 0,966 \Rightarrow \sin^2 \alpha_1 \approx 0,93, \\
 \alpha_g &\approx 42^\circ \leftarrow \sin \alpha_g = n_2/n_1 = 1/1,5 = 0,6.
 \end{aligned}$$

$$|y_{(I/e)}| = \frac{1}{2k_2\beta} = \frac{1}{2 \frac{n_2\omega}{c_0} \sqrt{\frac{n_1^2}{n_2^2} \sin^2 \alpha_1 - 1}} \approx \frac{1}{2 \frac{1 \cdot 2\pi \cdot 6 \cdot 10^{14} \text{ s}^{-1}}{3 \cdot 10^8 \text{ m s}^{-1}} \sqrt{\frac{1,5^2}{1^2} 0,93 - 1}},$$

$$|y_{(I/e)}| \approx 3,8 \cdot 10^{-8} \text{ m} = 38 \text{ nm}.$$

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.

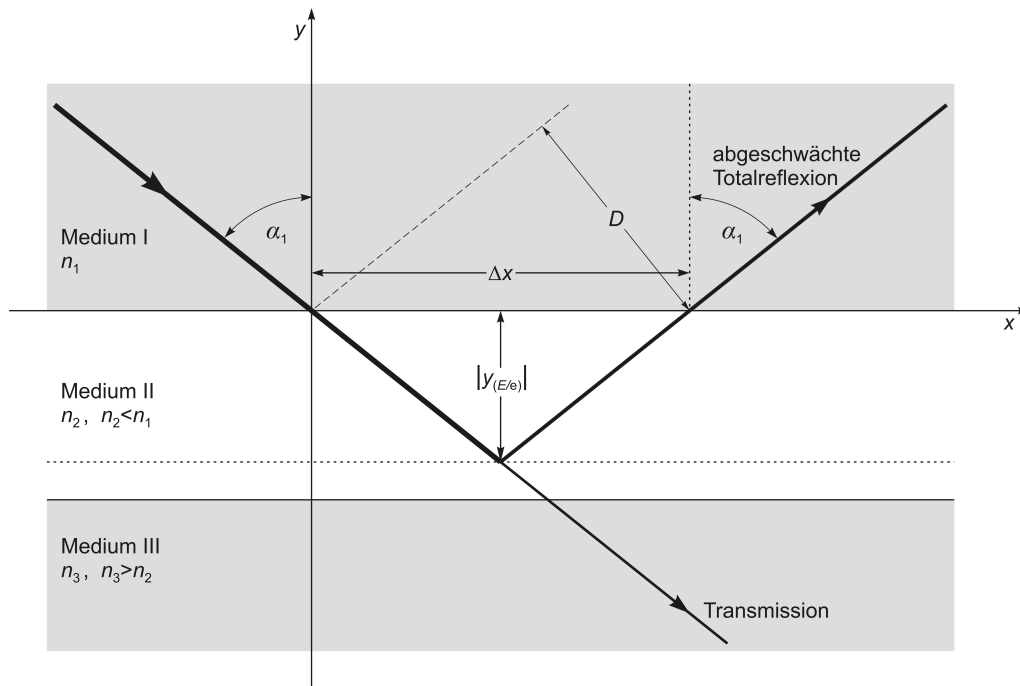


Abb. 3 Goos-Hänchen-Verschiebung Δx bei frustrierter bzw. abgeschwächter Totalreflexion. Gelegentlich wird $D = \Delta x \cos \alpha_1$ als Goos-Hänchen-Effekt bezeichnet.

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 3):

$$\Delta x \approx 2 \cdot |y_{(E/e)}| \cdot \tan \alpha_1 = 2 \frac{1}{k_2 \beta} \tan \alpha_1. \quad (30)$$

Für k_2 schreiben wir

$$k_2 = \frac{\omega}{c_0} \cdot n_2 = \frac{2\pi f}{c_0} \cdot n_2 = \frac{2\pi n_2}{\lambda_0} \quad \Leftrightarrow \quad \frac{1}{k_2} = \frac{\lambda_0}{2\pi n_2} ,$$

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 (30) 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:

$$\Delta x \propto \lambda , \quad \Delta x \propto \alpha_1 .$$

6 Die Kohärenz des Lichts

6.1 Wellenpaket – Wellenzug – Superpositionsprinzip

6.1.1 Teilchenbild

Im **Teilchenbild** bzw. Photonenbild ist natürlich ausgestrahltes Licht, das durch spontane Emission entsteht wie beispielsweise das thermische Licht einer Glühfadenlampe, ein „Gemisch“ aus Photonen verschiedener Energie und verschiedener Ausbreitungsrichtung. Photonen sind Bosonen und können quantenphysikalisch in Raum und Zeit als (räumlich „begrenzte“) Wellenpakete (Wellengruppen) beschrieben werden, die der Heisenberg’schen Energie-Zeit-Unschärfe gehorchen müssen.

Ein Photon mit der Energie $h\nu$ in Gestalt eines **Wellenpakets** beispielsweise $\psi(t)$ oder $\vec{E}(t)$ ist die Superposition von unendlich vielen harmonischen Funktionen verschiedener Periode (deshalb auch die Bezeichnung Wellengruppe) mit der Amplitudenmodulation dieser Wellengruppe durch eine Einhüllende (z. B. eine Gauß-Funktion oder eine Lorentz-Funktion). Die Fourier-Transformierte $F(\omega)$ des Wellenpakets ist dessen **kontinuierliches Frequenzspektrum** mit der Halbwertsbreite $\Delta\omega$ um die mittlere Kreisfrequenz $\omega_0 = 2\pi\nu_0$, die an der Stelle des Maximums der Funktion der Fourier-Transformierten liegt. Jedes Wellenpaket besitzt demzufolge ein kontinuierliches Frequenzspektrum und an einem festen Ort die „Lebensdauer“ Δt . In Übereinstimmung damit und auch in Übereinstimmung mit der Heisenberg’schen Energie-Zeit-Unschärferelation gilt:

Aus der raumzeitlichen Beschränktheit der Wellenpakete folgt eine Frequenzstreuung und damit eine natürliche (unvermeidbare) Linienbreite im Frequenzspektrum. Genauer gesagt: Je länger die Lebensdauer Δt eines Photons ist, desto schmaler ist sein Frequenzspektrum bzw. desto kleiner ist die Halbwertsbreite $\Delta\omega$ und umgekehrt. Je größer Δt ist, desto schärfer ist die zugehörige Spektrallinie. Deshalb besitzt das Frequenzspektrum eines Lasers eine besonders kleine Linienbreite.

6.1.2 Wellenbild

Im **Wellenbild** ist natürlich ausgestrahltes Licht im Allgemeinen eine Überlagerung von statistisch (zufällig) verteilten elementaren sinusförmigen Transversalwellenzügen (kurz: elementaren Wellenzügen) verschiedener mittlerer Kreisfrequenz $\omega_0 = 2\pi\nu_0$, verschiedener Polarisierung, verschiedenem Phasenwinkel ϕ (kurz: verschiedener Phase ϕ) und verschiedener Lebensdauer Δt an einem festen Ort entsprechend einer verschiedenen räumlichen Länge $c \cdot \Delta t$.

Wie Wellenbild und Teilchenbild so sind auch elementare Wellenzüge und Photonen komplementär, das heißt, ein Photon und der ihm entsprechende elementare Wellenzug sind zwei sich gegenseitig ausschließende Erscheinungsformen derselben physikalischen Realität. Anders gesagt, je nach Beobachtungsmethode besitzt Licht entweder Teilchen- bzw. Photonencharakter oder Wellencharakter. Dieser Sachverhalt wird auch als Welle-Teilchen-Dualismus bezeichnet.

Emittiert werden elementare Wellenzüge beispielsweise durch elektronische Übergänge, also wenn Elektronen einer Strahlungsquelle von einem höheren in ein niedrigeres Energieniveau springen bzw. wenn Teilchen wie beispielsweise Atome von einem angeregten in einen weniger angeregten oder in den Grundzustand übergehen. Die Dauer dieser elektronischen Übergänge bzw. Emissionsvorgänge beträgt etwa 10^{-10} s bis 10^{-8} s, liegt also in der Größenordnung um $\Delta t = 10^{-9}$ s = 1 ns entsprechend einer Länge der elementaren Wellenzüge von etwa $c \cdot \Delta t = 30$ cm. In Abhängigkeit von der Emissionsdauer und von der Wellenlänge besitzen dann die elementaren Wellenzüge etwa einhunderttausend bis zu vielen Millionen Perioden.

Die angeregten Zustände besitzen die mittlere Lebensdauer $\overline{\Delta t} = \tau_c$, die unmittelbar

zusammenhängt mit der **mittleren Lebensdauer der elementaren Wellenzüge**.¹ Es gilt nämlich: „Die Breite einer Spektrallinie, die einem gedämpften Wellenzug des einzelnen Emissionsaktes entspricht, ist gleich der reziproken **Lebensdauer des angeregten Zustandes**. Diese Beziehung gilt ganz allgemein, gleichgültig ob der Strahlungsprozeß durch Dämpfung oder anders abgebrochen wird.“² Das heißt aber, dass sich die Halbwertsbreite $\Delta\omega$ einer Spektrallinie um die mittlere Kreisfrequenz ω_0 reziprok verhält zur mittleren Lebensdauer τ des zugehörigen elementaren Wellenzuges gemäß

$$\tau_c = \frac{1}{\Delta\omega} \quad \Leftrightarrow \quad \Delta\omega \cdot \tau_c = 1 \quad . \quad (31)$$

Diese Beziehung heißt **spektrale Kohärenzbedingung** und wird oft auch in der schwächeren Form

$$\Delta\omega \cdot \tau_c \lesssim 2\pi \quad \Leftrightarrow \quad \Delta\nu \cdot \tau_c \lesssim 1$$

verwendet. Sie erfüllt die Heisenberg'sche Energie-Zeit-Unschärferelation, denn

$$\Delta E \cdot \Delta t = \hbar \Delta\omega \cdot \tau_c \geq \frac{\hbar}{2} \quad \Rightarrow \quad \Delta\omega \cdot \tau_c = 1 > \frac{1}{2} .$$

Wir zeigen eine Herleitung der Beziehung (31) nach dem

Springer-Lehrbuch *Physik kompakt 3 – Quantenphysik und Statistische Physik* von Gunnar Lindström, Rudolph Langkau und Wolfgang Scobel, 2. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2002, Abschnitt 5.3 *Beispiel zur Energie-Zeit-Unschärfe* (Resonanz angeregter Atomzustände), Seite 68 bis Seite 77 :

Ein Atom im angeregten Zustand verhält sich wie ein Resonator. Wir erfassen dessen Zustand durch die Wellenfunktion bzw. die Wahrscheinlichkeitsamplitude $\psi(t)$, die charakterisiert wird durch die Eigenfrequenz ω_0 des Zustands und durch die **Dämpfung bei spontaner Emission** von elektromagnetischer Energie $\Delta E = \hbar\Delta\omega$ in „Gestalt“ eines Photons. Erfolgt diese Emission ab dem Zeitpunkt $t = 0$, so beschreibt die Wellenfunktion die

$$\text{freie gedämpfte Schwingung} \quad \psi(t) = A e^{-\frac{1}{2\tau_c} t} e^{i\omega_0 t} . \quad (32)$$

Hierbei ist τ_c die mittlere Lebensdauer des angeregten Zustands. Die Momentanamplitude der emittierten elektromagnetischen Strahlung soll proportional zur Wellenfunktion $\psi(t)$ sein, sodass die Intensität $I(t)$ der Strahlung proportional zur Wahrscheinlichkeitsdichte $\psi^* \psi$ für die Emission eines Photons ist:

$$I(t) \propto \psi^* \psi = \left(A e^{-\frac{1}{2\tau_c} t} e^{-i\omega_0 t} \right) \cdot \left(A e^{-\frac{1}{2\tau_c} t} e^{i\omega_0 t} \right) = A^2 e^{-\frac{1}{\tau_c} t} \quad \Rightarrow$$

$$I(t) = I_0 e^{-\frac{1}{\tau_c} t} . \quad (33)$$

(33) bedeutet aber auch: Befinden sich N_0 Atome zum Zeitpunkt $t = 0$ im angeregten Zustand, so sind es zum Zeitpunkt t nur noch $N(t) = N_0 \cdot e^{-\frac{1}{\tau_c} t}$ und nach Verstreichen der mittleren Lebensdauer des angeregten Zustands schließlich $N(\tau_c) = \frac{1}{e} \cdot N_0$.

Nach Fourier lässt sich die Wellenfunktion (32) durch Überlagerung von unendlich vielen harmonischen Schwingungen F verschiedener Frequenz ω , also als Fourier-Integral darstellen:

$$\psi(t) = \frac{1}{2\pi} \int_0^\infty F(\omega) e^{i\omega t} \cdot d\omega , \quad \omega \geq 0 .$$

$F(\omega)$ ist die Fourier-Transformierte bzw. die Spektralfunktion von $\psi(t)$ und hat die Gestalt

$$F(\omega) = \int_0^\infty \psi(t) e^{-i\omega t} \cdot dt \quad \text{wegen} \quad \psi(t) = 0 \quad \text{für} \quad t < 0 .$$

¹Wir verwenden bei τ_c hier bereits den Index c , weil sich im Abschnitt 6.3 zeigt, dass τ_c die Kohärenzzeit ist.

²Zitiert aus dem Springer-Lehrbuch *Gerthsen Physik*, Helmut Vogel, 20. Auflage, Springer-Verlag, Berlin, Heidelberg, 1999, Abschnitt 12.2.2 *Linienverbreiterung*, Seite 606.

Einsetzen von (32) für $\psi(t)$ liefert

$$\begin{aligned} F(\omega) &= \int_0^{\infty} A e^{-\frac{1}{2\tau_c} t} e^{i\omega_0 t} e^{-i\omega t} \cdot dt \\ &= A \int_0^{\infty} A e^{-at} \cdot dt \quad \text{mit} \quad a = \frac{1}{2\tau_c} + i(\omega_0 - \omega), \\ F(\omega) &= -A \frac{1}{a} e^{-at} \Big|_0^{\infty} = A \frac{1}{a}. \end{aligned}$$

„Entsprechend der statistischen Interpretation von $\psi(t)$ ist $F(\omega)$ die Wahrscheinlichkeitsamplitude für die Emission der Frequenz ω , $F^*(\omega) \cdot F(\omega)$ ist die Wahrscheinlichkeitsdichte, mit der die Frequenz ω emittiert wird.

Wir benutzen jetzt einen hier nicht hergeleiteten Sachverhalt: Die Wahrscheinlichkeit für die Absorption eines Photons $\hbar\omega$ ist gleich der Wahrscheinlichkeit für die Emission eines Photons $\hbar\omega$ derselben Frequenz. Damit wird die Absorptionsintensität für die Anregung des atomaren Niveaus der mittleren Energie $E_0 = \hbar\omega_0$ gegeben durch:

$$I(\omega) \propto F^*(\omega) F(\omega)$$

Und dies muss dann auch gleichzeitig die Streuintensität sein!³

Die Intensität der emittierten Strahlung ist folglich

$$\begin{aligned} I(\omega) &= b \cdot F^*(\omega) F(\omega) = b \cdot A \frac{1}{a^*} \cdot A \frac{1}{a} = \frac{b A^2}{a^* \cdot a} \\ &= \frac{C}{\left[\frac{1}{2\tau_c} - i(\omega_0 - \omega) \right] \left[\frac{1}{2\tau_c} + i(\omega_0 - \omega) \right]} = \frac{C}{\left[\frac{1}{2\tau_c} + i(\omega - \omega_0) \right] \left[\frac{1}{2\tau_c} - i(\omega - \omega_0) \right]}, \\ I(\omega) &= \frac{C}{(\omega - \omega_0)^2 + \frac{1}{4\tau_c^2}}. \end{aligned}$$

Für $\omega = \omega_0$ erhalten wir daraus die Gleichung, in der $I(\omega)$ maximal ist und aus der wir die Konstante C bestimmen können:

$$I(\omega_0) = \frac{C}{\frac{1}{4\tau_c^2}} \quad \Leftrightarrow \quad C = I(\omega_0) \cdot \frac{1}{4\tau_c^2}.$$

Damit haben wir schließlich die Anregungsfunktion bzw.

Resonanzfunktion
$$I(\omega) = I(\omega_0) \frac{\frac{1}{4\tau_c^2}}{(\omega - \omega_0)^2 + \frac{1}{4\tau_c^2}}$$

(34)

gefunden. Den halben Wert gegenüber ihrem Maximum $I(\omega_0)$ besitzt die Resonanzfunktion an den beiden Stellen

$$\begin{aligned} \omega_1 = \omega_0 - \frac{1}{2\tau_c} \quad \text{und} \quad \omega_2 = \omega_0 + \frac{1}{2\tau_c} : \\ I(\omega) = I(\omega_0) \frac{\frac{1}{4\tau_c^2}}{\left[\left(\omega_0 - \frac{1}{2\tau_c} \right) - \omega_0 \right]^2 + \frac{1}{4\tau_c^2}} = I(\omega_0) \frac{\frac{1}{4\tau_c^2}}{\left[\left(\omega_0 + \frac{1}{2\tau_c} \right) - \omega_0 \right]^2 + \frac{1}{4\tau_c^2}} = I(\omega_0) \frac{\frac{1}{4\tau_c^2}}{\frac{1}{4\tau_c^2}} = I(\omega_0) \cdot \frac{1}{2}. \end{aligned}$$

Demzufolge ist die **Halbwertsbreite der Resonanzkurve**

$$\Delta\omega = \omega_2 - \omega_1 = \left(\omega_0 + \frac{1}{2\tau_c} \right) - \left(\omega_0 - \frac{1}{2\tau_c} \right) = \frac{1}{\tau_c}.$$

Die Halbwertsbreite $\Delta\omega$ der Resonanzkurve ist die (spektrale) Linienbreite. Somit lautet die Beziehung zwischen der Linienbreite und der mittleren Lebensdauer eines angeregten Zustands bzw. der mittleren Lebensdauer eines elementaren Wellenzuges

$$\Delta\omega \cdot \tau_c = 2\pi\Delta\nu \cdot \tau_c = 1. \quad \square$$

³Zitiert aus: Gunnar Lindström, Rudolph Langkau, Wolfgang Scobel, Springer-Lehrbuch *Physik kompakt 3 – Quantenphysik und Statistische Physik*, 2. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2002, Seite 72 und Seite 73. Dabei haben wir $F(\omega)$ für $f(\omega)$ geschrieben.

Wir können aber die Resonanzfunktion auch als Funktion der Photonenenergie $E = \hbar\omega$ bzw. $E_0 = \hbar\omega_0$ betrachten, wenn wir in (34) den Bruchterm mit \hbar^2 erweitern und $\frac{1}{\tau_c^2} = (\Delta\omega)^2$ sowie $\hbar \cdot \Delta\omega = \Gamma$ setzen. In Analogie zur Halbwertsbreite $\Delta\omega$ ist dann Γ die Energieniveaubreite:

$$I(E) = I(E_0) \frac{\frac{1}{4\tau_c^2} \cdot \hbar^2}{\left[(\omega - \omega_0)^2 + \frac{1}{4\tau_c^2}\right] \cdot \hbar^2} = I(E_0) \frac{\frac{1}{4} (\Delta\omega)^2 \hbar^2}{(\hbar\omega - \hbar\omega_0)^2 + \frac{1}{4} (\Delta\omega)^2 \hbar^2} ,$$

$$I(E) = I(E_0) \frac{\frac{1}{4} \Gamma^2}{(E - E_0)^2 + \frac{1}{4} \Gamma^2} . \quad (35)$$

(35) entspricht dem **Lorentz-Profil** bzw. einer Breit-Wigner-Kurve.

Zwei weitere Herleitungen der Beziehung (31) finden sich im Vorlesungsskript von Professor Dieter Freude, Universität Leipzig 2006 unter <https://home.uni-leipzig.de/energy/pdf/freusd2.pdf> Abschnitt 2.6 *Lebensdauer und natürliche Linienbreite*, Seite 14.

Die elementaren Wellenzüge sind definiert durch den elektrischen Feldanteil bzw. die elektrische Feldstärke \vec{E}_n , den magnetischen Feldanteil bzw. die magnetische Flussdichte \vec{B}_n , die Vakuumlichtgeschwindigkeit $c_0 = |\vec{c}_0|$, den Wellenvektor oder Propagationsvektor \vec{k}_n in Ausbreitungsrichtung und die Wellenzahl $k_n = |\vec{k}_n| = \frac{2\pi}{\lambda_n}$ gemäß

$$c_0 \frac{\vec{k}_n}{k_n} = \vec{c}_0 .$$

Die Feldvektoren eines elementaren Wellenzuges bilden im Vakuum mit dem zugehörigen Wellenvektor ein orthogonales Rechtssystem in der Reihenfolge

$$\vec{E}_n \xrightarrow{\perp} \vec{B}_n \xrightarrow{\perp} \vec{k}_n \xrightarrow{\perp} \vec{E}_n .$$

Im Folgenden unterdrücken wir vereinfachend den magnetischen Feldanteil und betrachten nur das \vec{E} -Feld, auch optisches Feld genannt. Außerdem legen wir, wenn nicht anders ausgewiesen, die Schwingungsebene des \vec{E} -Feldes in die x, y -Ebene, sodass \vec{k} die Richtung der z -Achse besitzt.

Die elementaren Wellenzüge überlagern sich zum optischen Strahlungs- oder Wellenfeld. Insofern sind die elementaren Wellenzüge Teilwellen des resultierenden optischen Wellenfeldes.

Nach dem **Superpositionsprinzip** erhält man die Gesamtfeldstärke \vec{E} eines optischen Wellenfeldes aus den Feldstärken \vec{E}_n und den Amplituden \vec{A}_n der einzelnen elementaren Wellenzüge mit dem Laufindex n wie folgt:

$$\vec{E}(\vec{r}, t) = \sum_n \vec{E}_n = \sum_n \vec{A}_n(\vec{r}, t) \cdot e^{i\phi_n} , \quad \phi_n(\vec{r}, t) = \vec{k}_n \cdot \vec{r} - \omega_n t .$$

Anders als in der oft verwendeten symbolhaften Darstellung sind elementare Wellenzüge selbstverständlich dreidimensionale physikalische Phänomene, die sich im Ortsraum ausbreiten, sich gegenseitig durchdringen und dabei das Wellenfeld $\vec{E}(\vec{r}, t)$ bilden.

6.2 Kohärenz

Siehe auch: Dr. Wiebke Salzmann, physik.wissenstexte.de, Wissenstexte – Kohärenz.

Kohärent bedeutet zusammenhängend und geht zurück auf die lateinischen Begriffe **cohaerentia** – Zusammenhang bzw. **cohaerere** – zusammenhängen. Wir schreiben im Folgenden statt elementare Wellenzüge kurz Wellenzüge. Die Kohärenz ergibt sich aus der Phasenkorrelation zwischen den Wellenzügen des Lichts. Die **Kohärenzbedingung** ist eine konstante Phasenbeziehung bzw. eine konstante Phasendifferenz $\Delta\phi$ einschließlich $\Delta\phi = 0$ zwischen sich überlagernden Wellenzügen. Man kann jeden Wellenzug für sich genommen als kohärent (interferenzfähig) betrachten, denn prinzipiell wäre mit einem einzelnen Wellenzug beispielsweise Interferenz am Doppelspalt möglich.

Wir beginnen unsere Überlegungen mit zwei Zitaten zur Definition der wichtigsten Begriffe:

„Die **Kohärenz** bestimmt die Interferenzfähigkeit der Teilwellen. Die Kohärenz wird immer in Bezug auf eine gewisse Mess- Beobachtungsdauer T_{mess} definiert. Eine zeitlich stationäre Interferenzstruktur kann nur dann beobachtet werden, wenn sich die Phasenunterschiede $\Delta\phi = \phi_j(\vec{r}_j, t) - \phi_i(\vec{r}_i, t)$ zwischen beliebigen Teilwellen $\vec{E}_i(\vec{r}_i, t)$ und $\vec{E}_j(\vec{r}_j, t)$ während der Beobachtungsdauer oder Messdauer T_{mess} um weniger als 2π ändern. Wenn sich an einem festen Raumpunkt $\vec{r}_i = \vec{r}_j$ die Phasenunterschiede $\Delta\phi$ innerhalb T_{mess} um weniger als 2π ändern, dann nennt man die Teilwellen **zeitlich kohärent**. Andererseits, wenn sich Phasenunterschiede $\Delta\phi$ an zwei beliebigen Orten $\vec{r}_i \neq \vec{r}_j$ innerhalb T_{mess} um weniger als 2π ändern, dann nennt man die Teilwellen **räumlich kohärent**. Zusammenfassend gilt, dass nur mit kohärenten Teilwellen stationäre Interferenzmuster beobachtet werden können.“⁴

„Eine zeitlich stationäre Interferenzstruktur kann nur dann beobachtet werden, wenn sich die Phasendifferenzen $\Delta\varphi = \varphi_j - \varphi_k$ zwischen beliebigen Teilwellen \vec{E}_j, \vec{E}_k im Raumpunkt $P(\vec{r})$ während der Beobachtungsdauer Δt um weniger als 2π ändern. Man nennt die Teilwellen dann **zeitlich kohärent**. ...

Die maximale Zeitspanne Δt_c , während der sich Phasendifferenzen zwischen allen im Punkt P überlagerten Teilwellen um höchstens 2π ändern, heißt **Kohärenzzeit** ...

Ändert sich die räumliche Differenz

$$\Delta_r \varphi_i = \varphi_i(\vec{r}_1) - \varphi_i(\vec{r}_2)$$

der Phase φ_i einer beliebigen Teilwelle \vec{E}_i während der Beobachtungszeit Δt um weniger als 2π , so heißt das Wellenfeld räumlich kohärent ... Die Fläche senkrecht zur Ausbreitungsrichtung, auf der $\Delta_r \varphi_i = 0$ erfüllt ist, heißt **Kohärenzfläche** F_c .

Als Kohärenzlänge $\Delta s_c = c' \cdot \Delta t_c$ wird die Strecke bezeichnet, die das Licht während der Kohärenzzeit zurücklegt. Das Produkt aus Kohärenzfläche und Kohärenzlänge Δs_c heißt **Kohärenzvolumen** ΔV_c ... Nur innerhalb des Kohärenzvolumens können Interferenzstrukturen beobachtet werden.“⁵

Anmerkung zum Sprachgebrauch

Man spricht gelegentlich von der Kohärenz einer Lichtquelle, meint dabei aber selbstverständlich die Kohärenz des von dieser Quelle emittierten Lichts.

⁴Zitiert aus: Prof. Dr. Ursula Keller, ETH Zürich, Quantenelektronik, FS11, Kap 4 Interferenz und Kohärenz, Abschnitt 4.1 Definition von Interferenz und Kohärenz, Seite 1.

⁵Zitiert aus: Wolfgang Demtröder, Springer-Lehrbuch, *Experimentalphysik 2 – Elektrizität und Optik*, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2004, Seite 295 und Seite 296. Dort wird für die Kohärenzzeit Δt_c statt τ_c und für die Kohärenzlänge Δs_c statt l_c geschrieben, was ebenfalls üblich ist.

Man unterscheidet also zwischen **zeitlicher Kohärenz** (auch longitudinale Kohärenz, räumliche Longitudinalkohärenz, Längenkohärenz oder englisch *longitudinal spatial coherence* genannt), die die Spektralverteilung einer Lichtquelle beschreibt und somit ein Maß für die spektrale Reinheit ist, und **räumlicher Kohärenz** (auch laterale Kohärenz, räumliche Transversalkohärenz, Winkelkohärenz oder englisch *lateral spatial coherence* genannt), die von der räumlichen Ausdehnung der Lichtquelle abhängt.

Voraussetzungen für die Interferenz zweier Teilwellen sind

- hinsichtlich der Polarisierung:

Die \vec{E} -Felder der Teilwellen müssen am Ort des Interferenzmusters kollinear sein. Zirkular polarisierte Wellen mit unterschiedlicher Helizität können nicht interferieren.

- hinsichtlich Frequenz und Phase:

Die Teilwellen müssen die gleiche Frequenz besitzen und die Phasendifferenz zwischen den Teilwellen muss zeitlich konstant sein, was allgemein nur erfüllt ist, wenn die Teilwellen aus derselben Quelle stammen.

„Es gilt ganz allgemein: Haben zwei Wellen verschiedene Frequenzen oder ändern sich ihre Phasenunterschiede mit der Zeit beliebig, so beobachtet man kein stationäres Interferenzmuster, und die Wellen werden als inkohärent bezeichnet.“⁶

**Kohärenz bedeutet Interferenzfähigkeit, d. h.
kohärentes Licht ist interferenzfähiges Licht.**

Vollständig kohärente und vollständig inkohärente optische Wellenfelder sind Idealisierungen, weil die dafür erforderlichen Anforderungen an die Quelle physikalisch nicht realisierbar sind:

- Die vollständige Kohärenz erforderte eine punktförmige Quelle (ohne räumliche Ausdehnung), die streng monochromatisches Licht emittiert. Die Wellenzüge streng monochromatischen Lichts wären unendlich lang.
- Die vollständige Inkohärenz erforderte eine Quelle mit räumlich unendlich ausgedehnter Emissionsfläche, von der Licht mit einem unendlich ausgedehnten Frequenzspektrum emittiert wird. Wellenzüge mit einer unendlichen Frequenzbreite wären unendlich kurz.

Damit wären alle von real existierenden (natürlichen und künstlichen) Lichtquellen emittierten Wellenfelder genau genommen immer nur partiell-kohärent bzw. teilweise kohärent oder kurz teilkohärent. Auch kann es kein absolut monochromatisches Licht geben, sondern nur näherungsweise monochromatisches bzw. quasimonochromatisches.

Wir werden uns im Folgenden mit der üblichen *genäherten* Graduierung begnügen. So unterscheiden wir dann nicht-monochromatisches von monochromatischem Licht. Weiterhin betrachten wir **monochromatisches Licht** als ein optisches Wellenfeld, das durch Superposition von Wellenzügen der gleichen Frequenz entsteht, **Laserlicht** als monochromatisch und kohärent und schließlich **natürlich erzeugtes Licht** als nicht-monochromatisch und inkohärent. Damit sind alle übrigen Wellenfelder mehr oder weniger stark bzw. schwach teilkohärent. Im Fall der Teilkohärenz sind die entsprechenden Interferenzmuster nicht so scharf bzw. kontrastreich ausgebildet wie im Fall der Kohärenz. Wie man sieht, ist der Übergang von der Kohärenz zur Inkohärenz in der physikalischen Realität fließend.

Zur Orientierung nehmen wir eine Auflistung einiger Wellenfelder hinsichtlich ihrer Kohärenz vorweg:

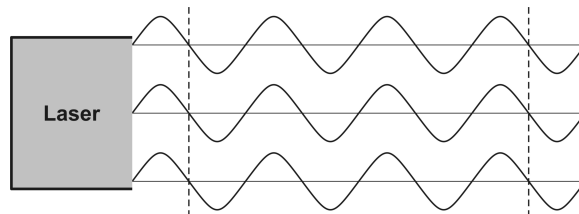
⁶Zitiert aus: Prof. Dr. Ursula Keller, ETH Zürich, Quantenelektronik, FS11, Kap 4 Interferenz und Kohärenz, Abschnitt 4.1 *Definition von Interferenz und Kohärenz*, Seite 9.

- monochromatische ebene Welle **kohärent** –
zeitlich vollständig, räumlich vollständig,
- nicht-monochromatische ebene Welle **partiell-kohärent** –
zeitlich partiell, räumlich vollständig,
- monochromatische Kugelwelle **kohärent** –
zeitlich vollständig, räumlich vollständig,
- nicht-monochromatische Kugelwelle **partiell-kohärent** –
zeitlich partiell, räumlich vollständig,
- natürlich erzeugtes Licht **inkohärent** –
zeitlich partiell, räumlich partiell.

Parallele Wellenzüge bilden durch Überlagerung ein optisches Wellenfeld mit ebenen Wellenfronten (ebene Wellen) und sind **räumlich kohärent**.

Als Standardbeispiel für kohärentes (vollständig zeitlich *und* vollständig räumlich kohärentes) Licht gilt das Bündel paralleler monochromatischer Lichtstrahlen eines Lasers (siehe Abbildung 4). **Laserlicht** entsteht durch **induzierte Emission** und ist vollständig phasenkorreliert.

Abb. 4 (Vollständige) Kohärenz des von einem Laser ausgestrahlten Lichts.

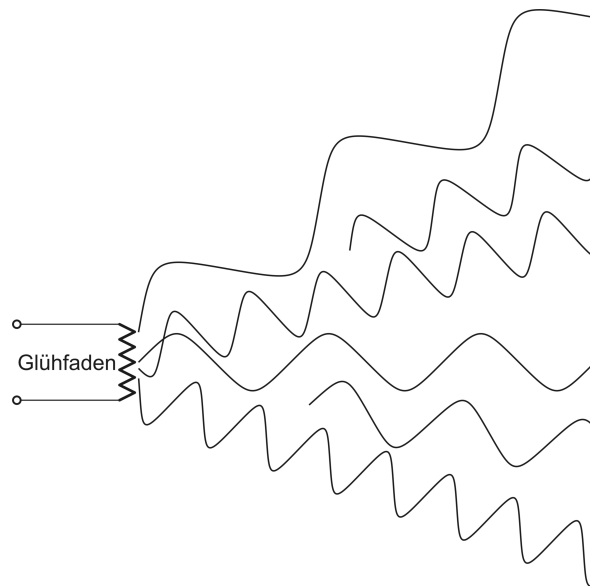


In diesem Fall sind die parallel zueinander verlaufenden Wellenzüge extrem lang und besitzen nahezu die gleiche Frequenz sowie die gleiche Phase (schwingen also im Gleichtakt), sodass die Wellenfronten Ebenen senkrecht zur Ausbreitungsrichtung bilden. Derartiges Laserlicht kommt dem Ideal einer ebenen elektromagnetischen Welle sehr nahe.

Natürlich ausgestrahltes Licht (siehe Abbildung 5) entsteht durch **spontane Emission** und ist **inkohärent**,

- weil die Zeitabstände zwischen den Emissionen der Wellenzüge statistisch (zufällig, willkürlich) verteilt sind,
- weil die Wellenzüge verschieden lang sind,
- weil die Wellenzüge verschiedene Frequenzen besitzen, denn natürlich ausgestrahltes Licht ist nicht monochromatisch,
- weil die Wellenzüge verschiedenen Phasen besitzen und verschieden polarisiert sind bzw. weil Phase und Polarisation schwanken und
- weil die Wellenzüge verschiedene Ausbreitungsrichtungen besitzen. Photonen werden statistisch in alle Richtungen emittiert.

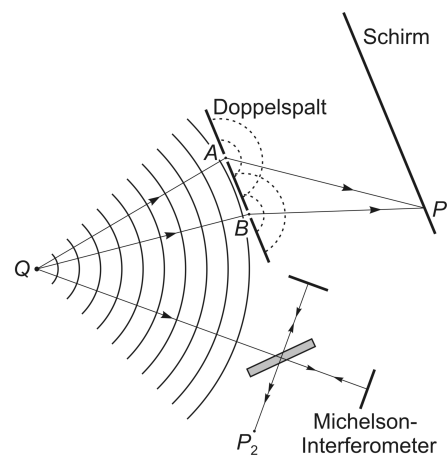
Abb. 5 Von einer Glühfadenlampe natürlich ausgestrahltes Licht. Dass Licht allgemein inkohärent ist, liegt an der diskontinuierlichen Natur der Emissionsprozesse.



Von Punktquellen emittierte Wellenzüge sind Kugelwellen und ihre Wellenfronten bilden Kugelflächen. Das von einer Punktquelle emittierte monochromatische Licht ist kohärent, also vollständig zeitlich und vollständig räumlich kohärent.

Abb. 6 Veranschaulichung der Kohärenz am Beispiel einer Punktquelle Q , die monochromatisches Licht emittiert.

Abbildung nach F. Pedrotti, L. Pedrotti, W. Bausch, H. Schmidt, 2005.



Erläuterung von zeitlicher und räumlicher Kohärenz anhand der Abbildung 6:

- Die **zeitliche (longitudinale) Kohärenz**

resultiert aus der Phasenkorrelation zwischen den Teilwellen bezüglich unterschiedlicher Laufzeiten entsprechend unterschiedlicher (räumlicher) Laufstrecken **in Ausbreitungsrichtung** dieser Teilwellen und zeigt sich beispielsweise beim Nachweis von Interferenz mit dem Michelson-Interferometer:

Beim Michelson-Interferometer fungiert der halbdurchlässige Spiegel als Strahlteiler (für „einen Lichtstrahl“ in Ausbreitungsrichtung). Die Längenänderung eines der beiden Interferometerarme verändert in diesem die Lichtlaufzeit, so dass man die Laufzeitdifferenz zwischen den beiden erzeugten Teilwellen variieren kann. Nach der Wiedervereinigung der beiden Teilwellen stellt man dann, analog zur Intensitätsverteilung des Interferenzmusters beim Doppelspalt, verschiedene Lichtintensitäten in Abhängigkeit von der Laufzeitdifferenz fest.

Monochromatisches Licht ist **zeitlich kohärent**, denn die Summe (Überlagerung) von jeweils zwei (harmonischen) Wellenzügen E_1 und E_2 gleicher Kreisfrequenz ω ist wieder eine harmonische Welle E mit der Kreisfrequenz ω gemäß

$$E_1 + E_2 = A_1 \cdot \sin(\omega t + \phi_1) + A_2 \cdot \sin(\omega t + \phi_2) = A \cdot \sin(\omega t + \phi) = E .$$

Und allgemein gilt: Wellenzüge gleicher und im Beobachtungszeitraum konstanter Frequenz (mit daraus resultierender zeitlich konstanter Phasendifferenz) sind zeitlich kohärent.

- Die **räumliche (transversale) Kohärenz**

resultiert aus der Phasenkorrelation zwischen **unterschiedlichen räumlichen Punkten** im optischen Wellenfeld und zeigt sich beispielsweise bei der Interferenz am Doppelspalt:

Durch den Doppelspalt erzeugt man aus dem primären Wellenfeld, hier der monochromatischen Kugelwelle, zwei sekundäre Teilwellen der gleichen Frequenz. Fällt monochromatisches Licht der Frequenz $\nu = \frac{c}{\lambda}$ auf den Doppelspalt, so interferieren die beiden von den Spalten ausgehenden Sekundärwellen miteinander und bilden in Abhängigkeit vom Gangunterschied $\Delta s = QAP_1 - QBP_1$ der Teilwellen ein für den Doppelspalt charakteristisches Interferenzmuster auf dem Schirm. Der Gangunterschied zwischen den Teilwellen entspricht der Phasendifferenz $\Delta\phi$, denn mit dem Spaltabstand g und dem Richtungswinkel φ zwischen Doppelspalt und Schirm gilt (siehe Abschnitt 19.2 *Fraunhoferbeugung* im Skript *Ausgewählte Themen und Herleitungen aus dem Physik-Grundstudium*, Seite 96)

$$\Delta\phi = \frac{2\pi}{\lambda} \cdot g \cdot \sin\varphi = k \cdot \Delta s .$$

6.3 Zur zeitlichen (longitudinalen) Kohärenz – Kohärenzlänge

Die zeitliche Kohärenz ist charakterisiert durch die Korrelation zweier Wellenfelder an ein und demselben Ort und zu unterschiedlichen Zeiten. Anders gesagt: Wellenzüge sind zeitlich kohärent, wenn ihre Phasen bzw. Wellenfronten bezüglich **eines festen Raumpunktes in Ausbreitungsrichtung** (longitudinal) korreliert sind. Zeitliche Kohärenz von Wellenzügen oder Wellenfeldern liegt also vor, wenn die Phasendifferenz zwischen ihnen *zeitlich* konstant ist.

Dass die vollständige zeitliche Kohärenz genau genommen ein Idealzustand ist, liegt an der endlichen räumlichen Ausdehnung der Wellenzüge in Ausbreitungsrichtung. Denn an einem bezüglich der Quelle festen Raumpunkt erscheinen die vorbeilaufenden Lichtwellen jeweils nur eine begrenzte und allgemein kurze Zeit lang sinusförmig zwischen plötzlich bzw. zufällig (statistisch) auftretenden Phasenänderungen. Diese Tatsache führt uns direkt zum Begriff der Kohärenzzeit:

„Ein einzelner Wellenzug existiert durchschnittlich eine als Kohärenzzeit Δt_c bezeichnete Zeitspanne lang, die gleich dem Kehrwert der Frequenzbreite $\Delta\nu$ ist. ... Die Kohärenzzeit entspricht effektiv derjenigen Zeitspanne, für die man die Phase der Lichtquelle in einem gegebenen Punkt des Raumes noch hinreichend genau voraussagen kann.“⁷

Anders gesagt: Das *durchschnittliche* Zeitintervall, in dem die Wellenzüge einer Strahlung bzw. eines Wellenfeldes ununterbrochen und ohne Phasensprung eine Sinuskurve beschreiben, bezeichnen wir als

$$\text{Kohärenzzeit } \overline{\Delta t} = \tau_c .$$

Im Sinne der Konsistenz insbesondere mit dem Abschnitt 6.1 verwenden wir für die Kohärenzzeit das Symbol τ_c und nicht Δt_c oder t_c .

Je länger die Kohärenzzeit ist, desto ausgeprägter ist die zeitliche Kohärenz des Wellenfeldes. Den Grad der Korrelation der Phasen an zwei in Ausbreitungsrichtung liegenden Punkten bezeichnet man in der englischsprachigen Literatur auch als *longitudinal coherence*.

Die Kohärenzzeit beeinflusst das Auftreten bzw. die Beobachtbarkeit von Interferenzmustern beispielsweise bei der Interferenz am Doppelspalt wie folgt:

„Wenn es sich ... um endliche Wellenzüge handelt, könnte am Beobachtungspunkt ein Wellenzug schon passiert sein, während der zweite noch gar nicht eingetroffen ist. Das kann sich besonders bei großen Ablenkungswinkeln ereignen, wenn der Gangunterschied groß und die Länge der Wellenzüge klein ist. Deshalb hat man den Begriff der Kohärenzlänge eingeführt, der der mittleren Länge der Wellenzüge entspricht.“⁸

Mit der Lichtgeschwindigkeit c ist dann die

$$\text{Kohärenzlänge } l_c = c \cdot \tau_c .$$

Die Kohärenzlänge ist also die *durchschnittliche* (räumliche) Länge der Wellenzüge des Wellenfeldes. Anders gesagt: Die Kohärenzlänge ist die räumliche Länge, über die die Teilwellen eine feste Phasenbeziehung haben.

Veranschaulichung der zeitlichen Kohärenz

Ein Beobachter befindet sich an einem festen Ort im Ruhesystem der Lichtquelle, greift einen der vorbeiziehenden Wellenzüge heraus und betrachtet die übrigen Wellenzüge des Lichts und

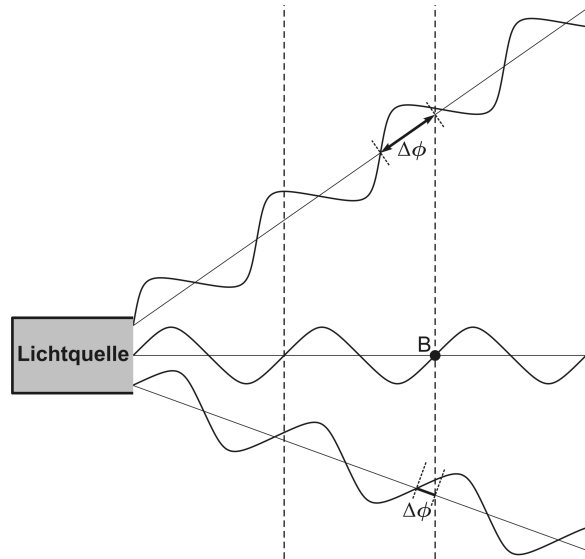
⁷Zitiert aus: Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag, Wien, München, 2005, Seite 901 und Seite 902.

⁸Zitiert aus: Würzburger Quantenphysik-Konzept. G56 Kohärenzlänge, <https://www.forphys.de/Website/qm/gloss/kohlaenge.html>

zwar senkrecht zur Ausbreitungsrichtung des herausgegriffenen Wellenzuges. Findet er dabei feste Phasenbeziehungen (Phasendifferenzen) zwischen dem herausgegriffenen Wellenzug und anderen Wellenzügen, so sind der herausgegriffene und diese anderen Wellenzüge zeitlich kohärent (siehe Abbildung 7).

Abb. 7 Zeitliche Kohärenz.

Der Beobachter B ruht bezüglich der Lichtquelle und „sieht“ die Wellenzüge gleicher Frequenz jeweils mit einer bestimmten zeitlich konstanten Phasendifferenz $\Delta\phi$ an sich vorbeiziehen.



Zusammenfassend und um den Unterschied zur räumlichen Kohärenz zu verdeutlichen, stellen wir fest:

Besteht in einem bezüglich der Quelle festen Raumpunkt über ein beobachtbares Zeitintervall T_{mess} zwischen Wellenzügen eine zeitlich konstante Phasendifferenz, so sind diese Wellenzüge zeitlich kohärent:

$$\tau_c > T_{\text{mess}} \Rightarrow \text{zeitliche Kohärenz, Interferenzmuster beobachtbar ,}$$

$$\tau_c < T_{\text{mess}} \Rightarrow \text{zeitliche Inkohärenz, Interferenzmuster nicht beobachtbar .}$$

Wenn die Wellenzüge zu kurz sind, „überlappen“ sie sich ggf. nicht mehr und es kommt nicht zur Interferenz. Der Idealfall vollständiger zeitlicher Kohärenz würde unendlich lange Wellenzüge gleicher Frequenz voraussetzen.

Im Wellenfrontenkonzept bedeutet zeitliche Kohärenz den konstanten räumlichen Abstand zwischen den Wellenfronten in Ausbreitungsrichtung (longitudinal) beim Vorbeilaufen der Wellenzüge an einem festen Raumpunkt. Alle Wellenzüge mit gleicher Frequenz wie beispielsweise bei monochromatischem Licht sind (zumindest partiell) zeitlich kohärent, aber nicht unbedingt auch räumlich kohärent.

Kurz gesagt:

Wenn Wellenzüge bzw. Teilwellen die gleiche Frequenz und über einen „beobachtbaren“ Zeitraum eine feste Phasenbeziehung (Phasendifferenz) haben, sind sie zeitlich kohärent.

Für das Intensitätsspektrum $I(\omega)$ bzw. die Intensität der harmonischen Anteile eines Wellenzugs der Kreisfrequenz ω_0 gilt:

$$I(\omega) = I_0 \left[\frac{\sin\left(\frac{\omega - \omega_0}{2} \cdot \tau_c\right)}{\frac{\omega - \omega_0}{2} \cdot \tau_c} \right]^2. \quad (36)$$

I_0 ist hierbei die Intensität an der Stelle des zentralen Maximums, d. h. an der Stelle der **Spektrallinie** mit der Kreisfrequenz ω_0 . Schreiben wir jetzt

$$(\omega - \omega_0) \cdot \tau_c = \omega \cdot \tau_c - \omega_0 \cdot \tau_c = \phi - \phi_0,$$

resultiert aus (36)

$$I(\omega) = I_0 \left[\frac{\sin\left(\frac{\phi - \phi_0}{2}\right)}{\frac{\phi - \phi_0}{2}} \right]^2.$$

Die Gleichung (36) stimmt formal mit der Intensitätsformel

$$I_\varphi = I_0 \left[\frac{\sin\left(\frac{\pi}{\lambda} \cdot a \sin \varphi\right)}{\frac{\pi}{\lambda} \cdot a \sin \varphi} \right]^2 \quad (37)$$

für die Interferenz am Einzelspalt überein (siehe Gleichung (81) im Abschnitt 11.2 Fraunhoferbeugung). Hier liegt das zentrale Maximum bzw. das Hauptmaximum bei $\varphi = 0 \Rightarrow \sin \varphi = 0$ und das 1. Minimum zu beiden Seiten des zentralen Maximums liegt bei $\sin \varphi = \frac{\lambda}{a}$, sodass (37) an den Stellen der 1. Minima die Form

$$I_\varphi = I_0 \left[\frac{\sin\left(\frac{\pi}{\lambda} \cdot a \cdot \frac{\lambda}{a}\right)}{\frac{\pi}{\lambda} \cdot a \cdot \frac{\lambda}{a}} \right]^2 = I_0 \left[\frac{\sin \pi}{\pi} \right]^2 = 0 \quad (38)$$

annimmt. Es gilt also für den Einzelspalt:

$$\Delta\phi = \pm\pi \quad \text{Phasendifferenz zwischen 1. Minimum und zentralem Maximum,}$$

$$\Delta\phi = 2\pi = \frac{2\pi}{\lambda} a \sin \varphi \quad \text{Phasendifferenz zwischen den beiden 1. Minima.}$$

In der Herleitung der Interferenz am Strichgitter bzw. am Einzelspalt im Abschnitt 11.2a hatten wir für die Phasendifferenz δ bezüglich des zentralen Maximums bei $\varphi = 0$ mit (73) folgendes gefunden:

$$\frac{\delta}{2\pi} = \frac{\Delta s}{\lambda} = \frac{g \sin \varphi}{\lambda} \Leftrightarrow \delta = \frac{2\pi}{\lambda} \cdot g \sin \varphi.$$

Setzen wir diese Phasendifferenz δ in (74) ein, erhalten wir für die vorläufige Amplitudenformel (74) für die Interferenz am Strichgitter

$$A_\varphi = A' \frac{\sin\left(N \cdot \frac{\pi}{\lambda} \cdot g \sin \varphi\right)}{\sin\left(\frac{\pi}{\lambda} \cdot g \sin \varphi\right)} = A' \frac{\sin\left(N \cdot \frac{\delta}{2}\right)}{\sin \frac{\delta}{2}},$$

sodass aus (77) schließlich

$$A_\varphi = A_0 \frac{\sin\left(\frac{\pi}{\lambda} \cdot a \sin \varphi\right)}{\frac{\pi}{\lambda} \cdot a \sin \varphi} = A_0 \frac{\sin \frac{\delta}{2}}{\frac{\delta}{2}}$$

resultiert. Schreiben wir jetzt ϕ_0 für die Phase ϕ an der Stelle des zentralen Maximums, so gilt

$$\delta = \phi - \phi_0$$

und (77) wird zu

$$A_\varphi = A_0 \frac{\sin \frac{\phi - \phi_0}{2}}{\frac{\phi - \phi_0}{2}} \quad \Rightarrow \quad I_\varphi = I_0 \left[\frac{\sin \frac{\phi - \phi_0}{2}}{\frac{\phi - \phi_0}{2}} \right]^2 . \quad \square$$

Für die 1. Minima im Verlauf von $I(\omega)$ (siehe Abbildung 18) muss gelten

$$\begin{aligned} \frac{(\omega_1 - \omega_0) \cdot \tau_c}{2} &= -\pi \quad \Leftrightarrow \quad \omega_1 = -\frac{2\pi}{\tau_c} + \omega_0 , \\ \frac{(\omega_2 - \omega_0) \cdot \tau_c}{2} &= +\pi \quad \Leftrightarrow \quad \omega_2 = +\frac{2\pi}{\tau_c} + \omega_0 . \end{aligned}$$

Als Breite $\Delta\omega$ des (zentralen) Hauptmaximums bzw. der Spektrallinie wird üblicherweise der *halbe* Abstand zwischen den beiden 1. Minima angenommen, d. h.

$$\frac{1}{2} \cdot |\omega_2 - \omega_1| = \frac{1}{2} \cdot \frac{4\pi}{\tau_c} = \frac{2\pi}{\tau_c} = \Delta\omega .$$

Mit $\omega = 2\pi\nu$ erhalten wir daraus die **spektrale Frequenzbreite** $\Delta\nu$ gemäß

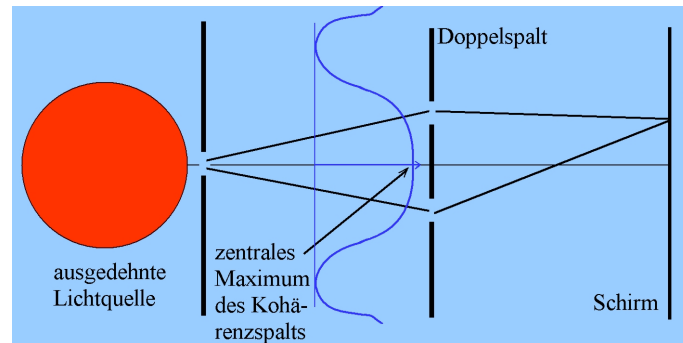
$$\boxed{\Delta\nu = \frac{1}{\tau_c} \quad \Leftrightarrow \quad \Delta\nu \cdot \tau_c = 1} .$$

6.4 Zur räumlichen (lateralen) Kohärenz – räumliche Kohärenzlänge

Die räumliche Kohärenz ist charakterisiert durch die Korrelation zweier Wellenfelder an verschiedenen Orten innerhalb einer zur Ausbreitungsrichtung senkrechten Ebene und zu einer festen Zeit.

Leicht veranschaulichen lässt sich die räumliche Kohärenz beispielsweise mit der Interferenz am Doppelspalt (siehe dazu Abbildung 8):

Abb. 8 Abbildung zur räumlichen Kohärenz aus: Würzburger Quantenphysik-Konzept. G56 Kohärenzlänge.



„Bei einer ausgedehnten Lichtquelle, wie etwa der Sonne, strahlen weit entfernte Atome vollkommen unabhängig voneinander ihr Licht ab. Die Wahrscheinlichkeit, dass sich die ausgesandten Wellenpakete nach dem Doppelspalt auf dem Schirm tatsächlich überlappen, ist also sehr gering. Verkleinert man jedoch die Lichtquelle künstlich durch einen **Kohärenzspalt**, dann entsteht Beugung am Einfachspalt mit einem zentralen Maximum. Wo auch immer die beitragenden Lichtstrahlen herkamen, sie müssen gleichphasig sein, sonst würden sie nicht zum zentralen Maximum beitragen. Sorgt man jetzt dafür, dass das zentrale Maximum vom Kohärenzspalt die beiden Spalte eines Doppelspalts gemeinsam ausleuchtet, so ist Doppelspalt-Interferenz möglich, weil das Licht durch den Kohärenzspalt **räumlich kohärent** gemacht wurde (**„Sichtbarkeitsbedingung“**).“⁹

„Die räumliche Kohärenz oder *räumliche Transversalkohärenz* (engl. *lateral spatial coherence*) ist möglicherweise leichter zu verstehen, weil sie eng mit dem Konzept der Wellenfront zusammenhängt: Befinden sich zu einem bestimmten Zeitpunkt zwei Punkte nebeneinander auf derselben Wellenfront, so bezeichnet man die Felder in diesen Punkten als räumlich kohärent.“¹⁰

Die räumliche Kohärenz hängt von der räumlichen Ausdehnung der Lichtquelle ab. Im Vergleich zur Wellenlänge λ sehr ausgedehnte Lichtquellen sind in der Regel räumlich inkohärent, weil in diesem Fall zwischen zwei transversal zur Ausbreitungsrichtung benachbarten Punkten im weit von der Quelle entfernten Wellenfeld keine Phasenkorrelation besteht. Im Idealfall vollständige räumliche Kohärenz würde eine (unendlich kleine) Punktquelle voraussetzen.

Wir zeigen diesen Sachverhalt anhand der Interferenz am Doppelspalt in der in Abbildung 9 dargestellten Versuchsanordnung und leiten dabei den Begriff der räumlichen Kohärenzlänge l_r her: Eine monochromatische Lichtquelle bestehe aus vielen voneinander unabhängigen Quellenelementen Q_n , sodass die Phasen ϕ_n der von den einzelnen Quellenelementen emittierten Teilwellen statistisch verteilt sind. Die Quelle habe die Querdimension b mit den Randquellen bzw. Randpunkten R_1 und R_2 . Das vom Punkt 0 emittierte Licht besitzt in den Spalten S_1 und S_2 wegen $\overline{OS_1} = \overline{OS_2}$ synchrone Phasenschwankungen und erzeugt deshalb auf dem Beobachtungsschirm ein zeitlich konstantes Interferenzmuster. Wegen der Laufstreckendifferenz bzw. des Gangunterschieds wie beispielsweise

⁹Zitiert aus: Würzburger Quantenphysik-Konzept. G56 Kohärenzlänge,
<https://www.forphys.de/Website/qm/gloss/kohlaenge.html>

¹⁰Zitiert aus: Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag, Wien, München, 2005, Seite 905.

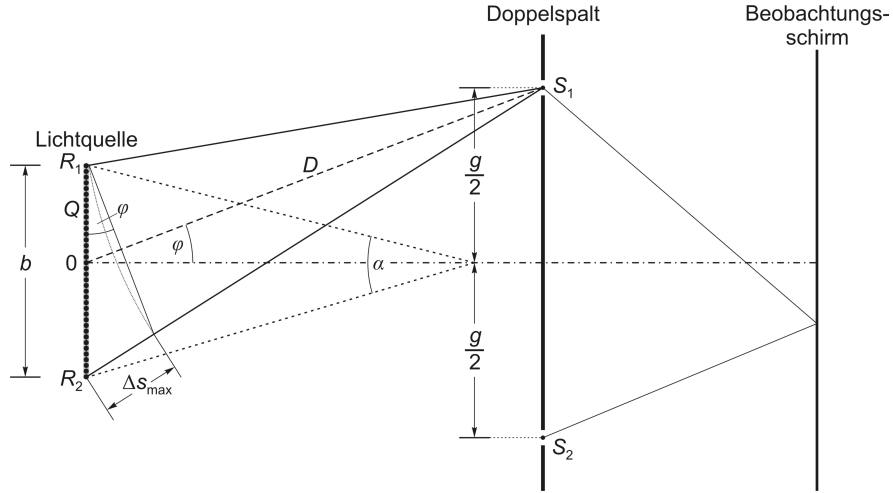


Abb. 9 Beim Young'schen Doppelspaltexperiment zur Herleitung der räumlichen Kohärenzlänge l_r zeigt sich der Einfluss der Quellengröße auf die Kohärenz des Wellenfeldes am Ort der Spalte S_1 und S_2 . Die Versuchsanordnung ist (nahezu) symmetrisch und die Ebenen von Quelle, Doppelspalt und Schirm sind komplanar und erstrecken sich senkrecht zur Mittellinie durch 0, was die Herleitung deutlich vereinfacht. Die Mittellinie kann man hier auch als optische Achse ansehen. Abbildung und Legende nach Wolfgang Demtröder, 2004.

$\Delta s = \overline{QS_1} - \overline{QS_2}$ ist das für alle anderen Quellenpunkte Q_n nicht der Fall, insbesondere auch nicht für die Randpunkte R_1 und R_2 mit dem größten Gangunterschied

$$\Delta s_{\max} = \overline{R_1 S_2} - \overline{R_1 S_1} = \overline{R_2 S_1} - \overline{R_1 S_1} \quad \text{mit} \quad \overline{R_1 S_1} = \overline{R_2 S_2}$$

$$\approx b \cdot \sin \varphi = b \cdot \frac{g}{D} \quad \text{für} \quad D \gg g.$$

Einem Gangunterschied Δs zwischen den Teilwellen entspricht die Phasendifferenz

$$\Delta \phi = k \cdot \Delta s = \frac{2\pi}{\lambda} \cdot \Delta s.$$

Wegen der statistischen Emission der Wellenzüge, ausgehend von den verschiedenen Quellenpunkten Q_n , folgt daraus:

Wenn der Gangunterschied Δs_{\max} größer als $\lambda/2$ wird, kann die Phasendifferenz $\Delta \phi = \phi(S_1) - \phi(S_2)$ um mehr als π schwanken, sodass sich dann das Interferenzmuster in der Beobachtungsebene zeitlich wegmittelt. Damit haben wir die Bedingung für die räumlich kohärente Beleuchtung der beiden Spalte für eine Quelle mit der Querdimension bzw. Breite b gefunden:

$$\Delta s_{\max} \approx \frac{b \cdot g}{2D} < \frac{\lambda}{2} \quad \Leftrightarrow \quad \frac{b \cdot g}{D} < \lambda \quad \Rightarrow \quad (39)$$

räumliche Kohärenzlänge $l_r = g < \frac{D}{b} \cdot \lambda$

(40)

Die Kohärenz und damit der Kontrast des Interferenzmusters werden stärker bei Verkleinerung der Quellengröße bzw. Quellenbreite b oder bei Verkleinerung des Spaltabstands g . Wird der Spaltabstand g größer als die räumliche Kohärenzlänge l_r , verschwindet das Interferenzmuster. Allgemein gilt:

Das monochromatische Licht einer räumlich ausgedehnten Quelle ist zwar zeitlich vollständig kohärent, aber räumlich nur partiell-kohärent, sodass der Kontrast des Interferenzmusters im Vergleich zu einer Punktquelle *gleichmäßig* gedämpft ist.

Wir betrachten jetzt das optische Wellenfeld in extrem großer Entfernung von einer räumlich ausgedehnten Quelle, die kein monochromatisches Licht emittiert. Ein Beispiel dafür ist das auf der Erde ankommende Licht der Fixsterne. Die Wellenzüge verschiedener Frequenz verlaufen dann nahezu parallel und besitzen näherungsweise die Eigenschaften ebener Wellen (siehe Abbildung 10). Man spricht deshalb bei Fixsternen in diesem Zusammenhang auch von Multifrequenzlasern. Das ankommende Wellenfeld ist vollständig räumlich und partiell zeitlich kohärent, also teilkohärent. Wenn wir den Sterndurchmesser bzw. die Quellenbreite b nicht kennen, können wir die räumliche Kohärenzlänge nicht mit (40) bestimmen. Allerdings lässt sich die Winkelausdehnung α der Quelle von der Doppelspaltebene aus leicht messen (siehe Abbildung 9). Ist die Quelle extrem weit entfernt bzw. für $D \gg b$ wird α sehr klein und wir erhalten in guter Näherung

$$\tan \frac{\alpha}{2} \approx \frac{b}{2D} \Rightarrow \frac{\alpha}{2} \approx \frac{b}{2D} \Leftrightarrow \alpha \approx \frac{b}{D}.$$

Dies setzen wir in (39) ein und erhalten so die räumliche Kohärenzbedingung bezüglich der Winkelausdehnung der Quelle:

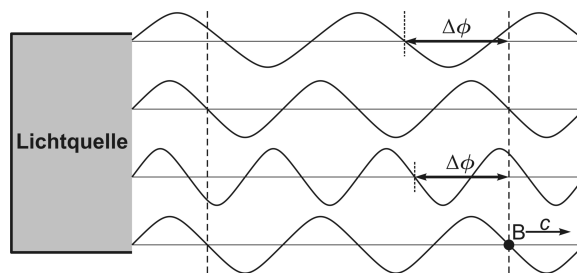
$$\frac{b}{D} \cdot g \approx \alpha \cdot g < \lambda \Leftrightarrow$$

$$\boxed{\alpha < \frac{\lambda}{g}}.$$

Veranschaulichung der räumlichen Kohärenz für den Fall paralleler Wellenzüge

Abb. 10 Räumliche Kohärenz.

Der Beobachter B bewegt sich mit einem der Wellenzüge. Selbstverständlich handelt es sich hierbei um ein Gedankenexperiment, weil sich Körper mit einer Ruhemasse nicht mit Lichtgeschwindigkeit bewegen können.



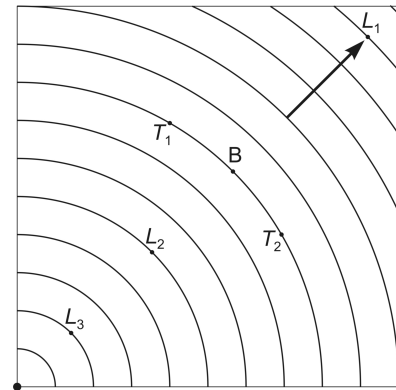
Ein Beobachter greift einen der parallelen Wellenzüge des Wellenfeldes heraus, bewegt sich dann mit ihm und beobachtet dabei die übrigen Wellenzüge des Wellenfeldes und zwar komplanar zur Ebene der Wellenfront, d. h. in diesem Fall senkrecht zur Ausbreitungsrichtung des herausgegriffenen Wellenzuges. Findet er dabei feste Phasenbeziehungen (Phasendifferenzen) zwischen dem herausgegriffenen Wellenzug und anderen Wellenzügen, so sind der herausgegriffene und diese anderen Wellenzüge räumlich kohärent. Wie man an der Abbildung 10 erkennt, besteht bezüglich der Wellenfronten der verschiedenen Wellenzüge eine konstante Phasendifferenz. Das Licht ist in diesem Fall vollständig räumlich kohärent.

Parallele Wellenzüge mit gleicher Ausbreitungsrichtung bzw. Wellenfelder mit ebenen Wellenfronten sind räumlich kohärent, aber nicht unbedingt auch zeitlich kohärent. Unabhängig von der Frequenz besitzen sie im Vergleich zur Phase des Wellenzuges des „mitbewegten“ Beobachters in jedem Raumpunkt senkrecht zur Ausbreitungsrichtung immer die gleiche Phasendifferenz. Verglichen wird also im Fall der räumlichen Kohärenz paralleler Wellenzüge die Phasendifferenz zwischen den Wellenzügen zu Beginn und am Ende von T_{mess} an verschiedenen Orten senkrecht zur Ausbreitungsrichtung.

6.5 Zur Kohärenz bei Punktquellen

Von Punktquellen emittierte Wellenzüge sind Kugelwellen und ihre Wellenfronten bilden Kugelflächen.

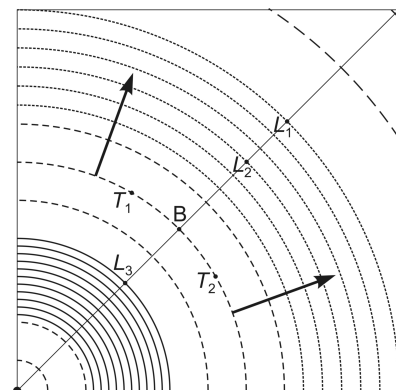
Abb. 11 Vollständige räumliche und vollständige zeitliche Kohärenz des Wellenfeldes einer Punktquelle, die monochromatisches Licht emittiert.
Abbildung nach E. Hecht, 2005.



Betrachten wir zunächst das Wellenfeld einer Punktquelle, die monochromatisches Licht emittiert (siehe Abbildung 11). Ein bezüglich der Quelle ortsfester Beobachter B findet im gesamten Bereich des Wellenfeldes eine konstante Phasendifferenz zu den sich radial ausbreitenden Wellenzügen und es besteht eine starke Phasenkorrelation mit den Punkten L_1 bis L_3 bzw. im gesamten Wellenfeld *längs der Ausbreitungsrichtung*. Weiterhin stellen wir fest, dass nicht nur die Phasendifferenz in den Punkten T_1 und T_2 auf der Wellenfront von B sondern auch die Phasendifferenzen *längs aller Wellenfronten* im gesamten Wellenfeld zeitlich konstant gleich Null sind.

Das von einer Punktquelle emittierte monochromatische optische Wellenfeld ist also vollständig zeitlich und vollständig räumlich kohärent.

Abb. 12 Vollständige räumliche und evl. teilweise zeitliche Kohärenz des Wellenfeldes einer Punktquelle, die Wellenzüge verschiedener Frequenz emittiert.
Abbildung nach E. Hecht, 2005.



Betrachten wir jetzt das optische Wellenfeld einer Punktquelle, die kein monochromatisches Licht emittiert (siehe Abbildung 12). Wie man sieht, sind auch in diesem Fall die Phasendifferenzen *längs aller Wellenfronten* der Wellenzüge verschiedener Frequenz im gesamten

Wellenfeld zeitlich konstant gleich Null, sodass das Wellenfeld vollständig räumlich kohärent ist.

Ein bezüglich der Quelle ortsfester Beobachter B findet keine im gesamten Bereich des Wellenfeldes konstante Phasendifferenzen zwischen den sich radial ausbreitenden Wellenzügen, sodass keine oder nur eine schwache Phasenkorrelation zwischen den Punkten L_1 bis L_3 bzw. im gesamten Wellenfeld *längs der Ausbreitungsrichtung* besteht. Deshalb ist das Wellenfeld in der Abbildung 12 wenn überhaupt nur teilweise zeitlich kohärent. Treffen nämlich die Wellenfronten von Abbildung 12 beispielsweise auf einen Doppelspalt, so kommt es danach in Abhängigkeit von der Länge der nacheinander folgenden Wellenzüge und abhängig von ihrer Frequenz zu mehr oder weniger kurzzeitigen, verschiedenen Interferenzmustern. Wenn bei nicht-monochromatischem Licht überhaupt ein stationäres Interferenzbild entsteht, ist dieses keinesfalls so scharf, wie es bei monochromatischem Licht der Fall wäre. Allgemein gilt: Von einer Punktquelle emittiertes Licht mit einem ausgedehnten Frequenzspektrum ist zwar räumlich vollständig kohärent, aber zeitlich nur partiell-kohärent, sodass der Kontrast des Interferenzmuster mit zunehmendem Abstand von der optischen Achse *zunehmend* gedämpft wird.

7 Die Polarisation des Lichts

Wir beginnen dieses Kapitel mit einem Zitat – in Anführungszeichen gesetzt – aus Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag München Wien, 2005, Seite 533 und Seite 534:

„Natürliches Licht

Eine gewöhnliche Lichtquelle besteht aus einer sehr großen Zahl zufällig ausgerichteter atomarer Strahler. Jedes angeregte Atom sendet etwa 10^{-8} s lang einen polarisierten Wellenzug aus. Alle diese Wellen gleicher Frequenz überlagern sich zu einer einzigen polarisierten Welle, die maximal 10^{-8} s lang stabil bleibt. Ständig werden neue Wellenzüge emittiert, und die Gesamtpolarisation ändert sich in einer vollkommen unvorhersehbaren Art und Weise. Finden diese Veränderungen mit so großer Geschwindigkeit statt, dass man die einzelnen resultierenden Polarisationszustände nicht mehr erkennen kann, so spricht man von **natürlichem Licht** oder *unpolarisiertem Licht*, wobei letztere Bezeichnung etwas irreführend ist, da sich das Licht in Wirklichkeit aus einer schnell veränderlichen Abfolge verschiedener Polarisationszustände zusammensetzt. *Zufällig polarisiertes Licht* ist wahrscheinlich eine treffendere Bezeichnung.

Wir können natürliches Licht mathematisch anhand zweier beliebiger, *inkohärenter*, orthogonaler, linear polarisierter Wellen gleicher Amplitude aufschreiben, also Wellen, deren relativer Phasenunterschied sich schnell und willkürlich verändert.

Dabei wollen wir nicht vergessen, dass eine idealisierte monochromatische Welle als unendlicher Wellenzug dargestellt werden muss. Wird diese Welle in zwei orthogonale Komponenten senkrecht zur Fortpflanzungsrichtung zerlegt, so müssen die Komponenten ihrerseits dieselbe Frequenz besitzen, unendlich ausgedehnt und deshalb wechselseitig kohärent sein ($\varepsilon = \text{konstant}$).¹ Mit anderen Worten: *Eine ideal monochromatische ebene Welle ist immer polarisiert. . . .*

Im Allgemeinen ist Licht, ob „natürlichen“ oder „künstlichen“ Ursprungs, weder vollkommen polarisiert noch unpolarisiert – beides sind Grenzfälle. Der elektrische Feldvektor verändert sich meistens weder völlig unregelmäßig noch völlig regelmäßig, weshalb man solches Licht auch **teilweise polarisiert** nennt. Am einfachsten beschreibt man dieses Verhalten mit einer Überlagerung von natürlichem und polarisiertem Licht.“

Wir werden uns in diesem und in den folgenden Kapiteln mit polarisiertem, insbesondere mit linear polarisiertem und zirkular polarisiertem Licht (in Form *ebener* Wellen) beschäftigen. Dabei ist es uns erlaubt, vereinfachend nur den elektrischen bzw. *optischen Anteil* $\vec{E}(\vec{r}, t)$ der elektromagnetischen Wellen zu verwenden, denn die Polarisation einer elektromagnetischen Welle ist durch die Richtung des elektrischen Feldvektors \vec{E} definiert.

Polarisation des Lichts im Wellenbild

Polarisationszustände

- \mathcal{P} -Zustand: lineare Polarisation,
- \mathcal{L} -Zustand: linkszirkulare Polarisation,
- \mathcal{R} -Zustand: rechtszirkulare Polarisation,
- \mathcal{E} -Zustand: elliptische Polarisation (allgemeiner Fall).

Die lineare und die zirkulare Polarisation kann man als Spezialfälle der elliptischen Polarisation ansehen.

¹Wir werden den relativen Phasenunterschied ε mit $\Delta\phi$ bezeichnen.

Wir gehen aus vom elektrischen bzw. optischen Anteil

$$\vec{E}(\vec{r}, t) = \begin{pmatrix} E_x(\vec{r}, t) \\ E_y(\vec{r}, t) \\ E_z(\vec{r}, t) \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i(\vec{k} \cdot \vec{r} - \omega t + \phi_{0x})} \\ E_{0y} \cdot e^{i(\vec{k} \cdot \vec{r} - \omega t + \phi_{0y})} \\ E_{0z} \cdot e^{i(\vec{k} \cdot \vec{r} - \omega t + \phi_{0z})} \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i\phi_{0x}} \\ E_{0y} \cdot e^{i\phi_{0y}} \\ E_{0z} \cdot e^{i\phi_{0z}} \end{pmatrix} e^{i(\vec{k} \cdot \vec{r} - \omega t)}$$

einer **ebenen elektromagnetischen Welle** in der Darstellung als komplexwertige Wellenfunktion mit **konstanter Amplitude** $\vec{E}_0 = (E_{0x}, E_{0y}, E_{0z})$, mit dem variablen Phasenanteil $(\vec{k} \cdot \vec{r} - \omega t)$ und mit den Phasenkonstanten $\phi_{0x}, \phi_{0y}, \phi_{0z}$. Die Phasenkonstanten können positive und negative Werte annehmen. Das Vorzeichen im variablen Phasenanteil $(\vec{k} \cdot \vec{r} - \omega t)$ bestimmt die Ausbreitungsrichtung der Wellen:

$$\begin{aligned} e^{i(kz - \omega t)} &\Rightarrow \text{Ausbreitung in } (+z)\text{-Richtung,} \\ e^{i(kz + \omega t)} &\Rightarrow \text{Ausbreitung in } (-z)\text{-Richtung.} \end{aligned}$$

Wie allgemein üblich werden auch wir im Folgenden nur Wellen mit der Ausbreitung in die positive Richtung der z -Achse betrachten.

Ohne Beschränkung der Allgemeinheit legt man gern die Ebene der Wellenfront in die (x, y) -Ebene, sodass sich die Welle gemäß $|\vec{k}| = k_z = k$ längs der z -Achse ausbreitet. Die komplexwertige Wellenfunktion erhält dann die einfache Gestalt

$$\vec{E}(z, t) = \begin{pmatrix} E_x(z, t) \\ E_y(z, t) \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i(kz - \omega t + \phi_{0x})} \\ E_{0y} \cdot e^{i(kz - \omega t + \phi_{0y})} \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i\phi_{0x}} \\ E_{0y} \cdot e^{i\phi_{0y}} \\ 0 \end{pmatrix} e^{i(kz - \omega t)}.$$

Weil man den Phasennullpunkt frei wählen kann, sind nicht die Phasenkonstanten selbst im Zusammenhang mit der Polarisation relevant, sondern von Bedeutung ist nur der

$$\textbf{Phasenunterschied } \Delta\phi := \phi_{0y} - \phi_{0x}$$

zwischen den Komponenten E_x und E_y von \vec{E} , welcher dann der y -Komponente zugeordnet wird. Die komplexwertige Wellenfunktion hat damit schließlich die Gestalt

$$\boxed{\vec{E}(z, t) = \begin{pmatrix} E_{0x} \cdot e^{i(kz - \omega t)} \\ E_{0y} \cdot e^{i(kz - \omega t + \Delta\phi)} \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\Delta\phi} \\ 0 \end{pmatrix} e^{i(kz - \omega t)}}. \quad (41)$$

7.1 Linear polarisiertes Licht

Unabhängig vom Amplitudenverhältnis E_{0y}/E_{0x} entsteht linear polarisiertes Licht (\mathcal{P}) unter den Bedingungen, dass entweder die x - und die y -Komponente von \vec{E} die gleiche Phase besitzen oder dass der Phasenunterschied $\Delta\phi$ zwischen den Komponenten ein positives oder negatives ganzzahliges Vielfaches von π beträgt:

$$\Delta\phi = m \cdot \pi, \quad m \in \mathbb{Z}, \quad (E_{0x} = E_{0y}) \wedge (E_{0x} \neq E_{0y})$$

Damit und ausgehend von (41) wird linear polarisiertes Licht (\mathcal{P}) beschrieben durch die komplexwertige Wellenfunktion

$$\vec{E}(z, t) = \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\Delta\phi} \\ 0 \end{pmatrix} e^{i(kz - \omega t)} \quad \text{für} \quad \Delta\phi = m \cdot \pi, \quad m \in \mathbb{Z} \quad (42)$$

Für $m = 0, \pm 2, \pm 4, \dots$ schwingen die x - und die y -Komponente von \vec{E} gleichsinnig beide gleichzeitig in die positive oder negative Richtung. Das bedeutet, dass die Schwingungsebene für gerade m und $m = 0$ graphisch gesehen durch den 1. und 3. Quadranten der (x, y) -Ebene geht.

Für $m = \pm 1, \pm 3, \pm 5, \dots$ schwingt die y -Komponente entgegengesetzt zur x -Komponente von \vec{E} . Das bedeutet, dass die Schwingungsebene für ungerade m graphisch gesehen durch den 2. und 4. Quadranten der (x, y) -Ebene geht.

Das Verhältnis E_{0y}/E_{0x} legt die (konstante) Polarisationsrichtung fest gemäß

$$\tan \alpha = \frac{E_{0y}}{E_{0x}},$$

wobei α der **Neigungswinkel** ist, der zwischen der positiven x -Achse und dem \vec{E} -Vektor und demzufolge im 1. oder 4. Quadranten der (x, y) -Ebene liegt, sodass $-90^\circ \leq \alpha \leq +90^\circ$ gilt.

Mit einem Polarisationsfilter kann man die Komponente \vec{E}_{pol} von \vec{E} herausfiltern. \vec{E}_{pol} ist dabei die Projektion von \vec{E} auf die Polarisationsrichtung des Filters und besitzt demzufolge die Polarisationsrichtung des Filters nach unseren Voraussetzungen in der (x, y) -Ebene. Ist ϑ der Winkel zwischen dem \vec{E} -Vektor und der Polarisationsrichtung, erhalten wir

$$|\vec{E}_{\text{pol}}| = \cos \vartheta \cdot |\vec{E}|.$$

Bei dieser „Filterung“ bzw. Polarisation wird die Phase nicht verändert, sodass \vec{E}_{pol} die gleiche Phase besitzt wie \vec{E} .

7.2 Zirkular polarisiertes Licht

Im speziellen Fall mit

$$\Delta\phi = \phi_{0y} - \phi_{0x} = \pm \frac{\pi}{2} + m \cdot 2\pi, \quad m \in \mathbb{Z} \quad \text{und} \quad E_{0x} = E_{0y} = E_0$$

resultiert die **zirkulare Polarisation**, beschrieben durch die komplexwertige Wellenfunktion

$$\vec{E}(z, t) = \begin{pmatrix} E_x(z, t) \\ E_y(z, t) \\ 0 \end{pmatrix} = \begin{pmatrix} E_0 \\ E_0 \cdot e^{i\Delta\phi} \\ 0 \end{pmatrix} e^{i(kz - \omega t)} \quad \text{für} \quad \Delta\phi = \pm \frac{\pi}{2} + m \cdot 2\pi, \quad m \in \mathbb{Z} \quad . \quad (43)$$

Wir unterscheiden rechtszirkular und linkszirkular polarisiertes Licht.

Rechtszirkular polarisiert (\mathcal{R}) bedeutet, dass der \vec{E} -Feldvektor bei *Blickrichtung zur Lichtquelle* mit konstanter Amplitude E_0 und mit der gleichen Frequenz, mit der er oszilliert, *im Uhrzeigersinn* um die Ausbreitungsachse rotiert. Analog dazu rotiert der \vec{E} -Feldvektor bei **linkszirkularer Polarisation** (\mathcal{L}) entgegengesetzt zum Uhrzeigersinn um die Ausbreitungsachse.²

Eine andere Möglichkeit, die zirkulare Polarisation zu klassifizieren, bietet die Helizität. Die **positive Helizität** σ^+ entspricht einer rechtshändigen Schraubenlinie und die **negative Helizität** σ^- einer linkshändigen Schraubenlinie, wobei der entsprechende Daumen in die Richtung der Vorwärtsbewegung der Schraube zeigt.

Wir führen jetzt einige Beziehungen an, die wir im Folgenden benötigen werden:

- $\cos\left(\frac{\pi}{2} \pm \alpha\right) = \mp \sin \alpha \quad \Rightarrow \quad \cos\left(\alpha \pm \frac{\pi}{2}\right) = \mp \sin \alpha,$
 $\sin\left(\frac{\pi}{2} \pm \alpha\right) = \pm \cos \alpha \quad \Rightarrow \quad \sin\left(\alpha \pm \frac{\pi}{2}\right) = \pm \cos \alpha.$
- $\Delta\phi = -\frac{\pi}{2} + m \cdot 2\pi, \quad m \in \mathbb{Z} \quad \Rightarrow$
 $\cos(kz - \omega t - \frac{\pi}{2}) = \sin(kz - \omega t) \quad \text{eilt} \quad \cos(kz - \omega t) \quad \textbf{voraus}.$
- $\Delta\phi = +\frac{\pi}{2} + m \cdot 2\pi, \quad m \in \mathbb{Z} \quad \Rightarrow$
 $\cos(kz - \omega t + \frac{\pi}{2}) = -\sin(kz - \omega t) \quad \text{läuft} \quad \cos(kz - \omega t) \quad \textbf{nach}.$
- $\Delta\phi = -\frac{\pi}{2}$ entspricht dem Faktor $-i = e^{i(-\frac{\pi}{2})}$:

$$\begin{aligned} e^{i(kz - \omega t - \frac{\pi}{2})} &= e^{i(-\frac{\pi}{2})} \cdot e^{i(kz - \omega t)} = -i \cdot [\cos(kz - \omega t) + i \sin(kz - \omega t)], \\ e^{i(kz - \omega t - \frac{\pi}{2})} &= \sin(kz - \omega t) - i \cos(kz - \omega t). \end{aligned}$$

- $\Delta\phi = +\frac{\pi}{2}$ entspricht dem Faktor $+i = e^{i(+\frac{\pi}{2})}$:

$$\begin{aligned} e^{i(kz - \omega t + \frac{\pi}{2})} &= e^{i(+\frac{\pi}{2})} \cdot e^{i(kz - \omega t)} = +i \cdot [\cos(kz - \omega t) - i \sin(kz - \omega t)], \\ e^{i(kz - \omega t + \frac{\pi}{2})} &= \sin(kz - \omega t) + i \cos(kz - \omega t). \end{aligned}$$

²Diese Konvention hinsichtlich rechts- und linkszirkularer Polarisation ist in der Optik üblich. In Anlehnung an die Helizität der Photonen gilt in der Quantenphysik die umgekehrte Konvention.

Physikalisch relevant ist letztlich nur der **Realteil** der komplexwertigen Wellenfunktion (41), nämlich

$$\operatorname{Re}\left\{\vec{E}(z, t)\right\} = \begin{pmatrix} E_{0x} \cdot \cos(kz - \omega t + \phi_{0x}) \\ E_{0y} \cdot \cos(kz - \omega t + \phi_{0y}) \\ 0 \end{pmatrix}.$$

Im speziellen Fall des **zirkular polarisierten Lichts** ist das in unserer Notation

$$\operatorname{Re}\left\{\vec{E}(z, t)\right\} = E_0 \begin{pmatrix} \cos(kz - \omega t) \\ \cos(kz - \omega t + \Delta\phi) \\ 0 \end{pmatrix} \quad \text{für } \Delta\phi = \pm \frac{\pi}{2} + m \cdot 2\pi, \quad m \in \mathbb{Z}.$$

Weil in der Literatur der Phasenunterschied $\Delta\phi$ manchmal auch der x -Komponente von \vec{E} zugeordnet wird, zeigen wir der Vollständigkeit halber tabellarisch die **Phasen** der x - und y -Komponente von $\vec{E}(z, t)$ für die sich dann ergebenden vier möglichen Fälle

$$\phi_{0y} = \Delta\phi = \pm \frac{\pi}{2} \Rightarrow \phi_{0x} = 0 \quad \text{und} \quad \phi_{0x} = \Delta\phi = \pm \frac{\pi}{2} \Rightarrow \phi_{0y} = 0 :$$

Tabelle: Zirkulare Polarisation

rechtszirkulare Polarisation (\mathcal{R}) $\hat{=}\sigma^-$ negative Helizität	linkszirkulare Polarisation (\mathcal{L}) $\hat{=}\sigma^+$ positive Helizität
E_y eilt E_x um $\frac{\pi}{2}$ voraus.	E_y läuft E_x um $\frac{\pi}{2}$ hinterher.
$\vec{E}(z, t) = E_0 \begin{pmatrix} e^{i(kz - \omega t)} \\ e^{i(kz - \omega t - \frac{\pi}{2})} \\ 0 \end{pmatrix} \xrightarrow{(44)}$ $\begin{pmatrix} \cos(kz - \omega t) \\ \cos(kz - \omega t - \frac{\pi}{2}) \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(kz - \omega t) \\ \sin(kz - \omega t) \\ 0 \end{pmatrix}$ <p>oder</p> $\vec{E}(z, t) = E_0 \begin{pmatrix} e^{i(kz - \omega t + \frac{\pi}{2})} \\ e^{i(kz - \omega t)} \\ 0 \end{pmatrix} \xrightarrow{(45)}$ $\begin{pmatrix} \cos(kz - \omega t + \frac{\pi}{2}) \\ \cos(kz - \omega t) \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin(kz - \omega t) \\ \cos(kz - \omega t) \\ 0 \end{pmatrix}$	$\vec{E}(z, t) = E_0 \begin{pmatrix} e^{i(kz - \omega t)} \\ e^{i(kz - \omega t + \frac{\pi}{2})} \\ 0 \end{pmatrix} \xrightarrow{(45)}$ $\begin{pmatrix} \cos(kz - \omega t) \\ \cos(kz - \omega t + \frac{\pi}{2}) \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(kz - \omega t) \\ -\sin(kz - \omega t) \\ 0 \end{pmatrix}$ <p>oder</p> $\vec{E}(z, t) = E_0 \begin{pmatrix} e^{i(kz - \omega t - \frac{\pi}{2})} \\ e^{i(kz - \omega t)} \\ 0 \end{pmatrix} \xrightarrow{(44)}$ $\begin{pmatrix} \cos(kz - \omega t - \frac{\pi}{2}) \\ \cos(kz - \omega t) \\ 0 \end{pmatrix} = \begin{pmatrix} \sin(kz - \omega t) \\ \cos(kz - \omega t) \\ 0 \end{pmatrix}$

Achtung!

Während der variable Phasenanteil

$$kz - \omega t \quad \text{bzw.} \quad \omega t - kz = -(kz - \omega t)$$

äquivalent verwendet werden kann, beide Varianten beschreiben die Ausbreitung derselben Welle in die positive z -Richtung, ist dies für die Phasen $kz - \omega t + \phi_0$ und $\omega t - kz + \phi_0$ mit dem konstanten Phasenanteil $\phi_0 \neq 0$ nicht der Fall. Wenn wir nämlich statt $kz - \omega t$ als variablen Anteil $\omega t - kz$ benutzten, würden wegen

$$kz - \omega t + \phi_0 \quad \text{und} \quad \omega t - kz + \phi_0 = -(kz - \omega t - \phi_0)$$

die x - und die y -Komponente in den Vektoren der Tabelle „Zirkulare Polarisation“ vertauschen und damit auch die Formeln für rechts- und linkszirkulare Polarisation.

7.3 Elliptisch polarisiertes Licht

Die **elliptische Polarisation** (\mathcal{E}) ist im Grunde genommen der allgemeine Fall der Polarisation. Man kann nämlich die lineare und die zirkulare Polarisation als Spezialfälle der elliptischen Polarisation auffassen, denn mit $m \in \mathbb{Z}$ gilt für die

- lineare Polarisation:

$$\Delta\phi = m \cdot \pi \quad \text{im Fall} \quad (E_{0x} = E_{0y}) \text{ und } (E_{0x} \neq E_{0y}),$$

- zirkulare Polarisation:

$$\Delta\phi = m \cdot \pi \pm \frac{\pi}{2} \quad \text{im Fall} \quad E_{0x} = E_{0y},$$

- elliptische Polarisation, Ellipsen-Hauptachse entlang der x - oder y -Achse orientiert:

$$\Delta\phi = m \cdot \pi \pm \frac{\pi}{2} \quad \text{im Fall} \quad E_{0x} \neq E_{0y}, \quad (46)$$

- elliptische Polarisation, Ellipsen-Hauptachse um den Winkel α gegen die positive x -Achse geneigt:

$$\left. \begin{array}{l} \Delta\phi \neq m \cdot \pi \\ \text{und} \\ \Delta\phi \neq m \cdot \pi \pm \frac{\pi}{2} \end{array} \right\} \quad \text{im Fall} \quad (E_{0x} \neq E_{0y}) \text{ und } (E_{0x} = E_{0y}). \quad (47)$$

(46) und (47) sind also die Bedingungen für das Auftreten der elliptischen Polarisation. In einer ortsfesten (x, y) -Ebene ($z = \text{const}$) rotiert der \vec{E} -Vektor bei elliptischer Polarisation senkrecht zu \vec{k} um die z -Achse und verändert dabei seinen Betrag. Der **Neigungswinkel** α von der positiven x -Achse zur Hauptachse³ der Polarisationsellipse (siehe Abbildung 13) ergibt sich aus

$$\tan 2\alpha = \frac{2 \cdot E_{0x} E_{0y} \cdot \cos \Delta\phi}{E_{0x}^2 - E_{0y}^2} = \frac{2 \frac{E_{0y}}{E_{0x}}}{1 - \left(\frac{E_{0y}}{E_{0x}}\right)^2} \cos \Delta\phi. \quad (48)$$

³Zur Erinnerung: Die **Hauptachse** entspricht dem größten Durchmesser und die **Nebenachse** dem kleinsten Durchmesser einer Ellipse. Die Länge der **großen Halbachse** ist gleich dem größten Radius und die Länge der **kleinen Halbachse** ist gleich dem kleinsten Radius einer Ellipse.

Und den **Elliptizitätswinkel** ε der Polarisationsellipse erhält man aus

$$\sin 2\varepsilon = \frac{2 \cdot E_{0x} E_{0y} \cdot \sin \Delta\phi}{E_{0x}^2 + E_{0y}^2} = \frac{2 \frac{E_{0y}}{E_{0x}}}{1 + \left(\frac{E_{0y}}{E_{0x}}\right)^2} \sin \Delta\phi. \quad (49)$$

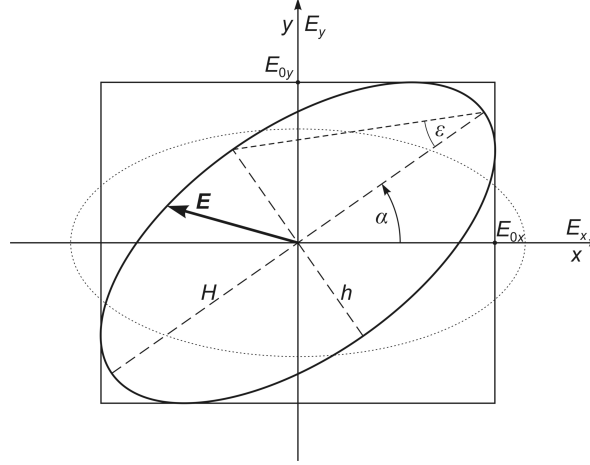


Abb. 13 Polarisationsellipse mit der großen Halbachse H und der kleinen Halbachse h . Mit dieser Notation gilt die Beziehung $H^2 + h^2 = E_{0x}^2 + E_{0y}^2$. Abbildung nach Eugene Hecht, *Optik*, 4. Auflage, Oldenbourg Verlag, Wien, München, 2005, Seite 531.

Ausgehend von

$$(41) \quad \vec{E}(z, t) = \begin{pmatrix} E_x(z, t) \\ E_y(z, t) \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i(kz - \omega t)} \\ E_{0y} \cdot e^{i(kz - \omega t + \Delta\phi)} \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\Delta\phi} \\ 0 \end{pmatrix} e^{i(kz - \omega t)}$$

mit $\phi_{0x} = 0 \Rightarrow \Delta\phi = \phi_{0y} - \phi_{0x} = \phi_{0y}$ werden wir jetzt den **Realteil**

$$\text{Re}\{\vec{E}(z, t)\} = \begin{pmatrix} \text{Re}E_x \\ \text{Re}E_y \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot \cos(kz - \omega t) \\ E_{0y} \cdot \cos(kz - \omega t + \Delta\phi) \\ 0 \end{pmatrix}$$

von (41) unter der Bedingung (47) diskutieren, d. h. für eine Polarisationsellipse mit gegen die x -Achse geneigter Hauptachse. Das Ziel dabei ist, eine Gleichung für die Kurve zu ermitteln, welche die „Spitze“ des \vec{E} -Vektors in einer ortsfesten (x, y) -Ebene (Schwingungsebene mit $z = \text{const}$) beim Durchlauf der Welle beschreibt. Die gesuchte Kurvengleichung sollte am Ende weder von z noch von der Zeit abhängen. Formen wir also

$$\text{Re}E_x = E_{0x} \cdot \cos(kz - \omega t) \quad \text{und} \quad \text{Re}E_y = E_{0y} \cdot \cos(kz - \omega t + \Delta\phi)$$

u. a. mit Hilfe des Additionstheorems $\cos(\alpha \pm \beta) = \cos\alpha \cdot \cos\beta \mp \sin\alpha \cdot \sin\beta$ um:

$$\begin{aligned} \frac{\text{Re}E_x}{E_{0x}} \cos \Delta\phi &= \cos(kz - \omega t) \cdot \cos \Delta\phi, \\ \sin^2(kz - \omega t) &= 1 - \cos^2(kz - \omega t) = 1 - \left(\frac{\text{Re}E_x}{E_{0x}}\right)^2, \end{aligned} \quad (50)$$

$$\frac{\text{Re}E_y}{E_{0y}} = \cos(kz - \omega t) \cdot \cos \Delta\phi - \sin(kz - \omega t) \cdot \sin \Delta\phi.$$

Daraus folgt

$$\begin{aligned}
\frac{\operatorname{Re} E_y}{E_{0y}} - \frac{\operatorname{Re} E_x}{E_{0x}} \cos \Delta\phi &= -\sin(kz - \omega t) \cdot \sin \Delta\phi \quad \xrightarrow{(50)} \\
\left(\frac{\operatorname{Re} E_y}{E_{0y}} - \frac{\operatorname{Re} E_x}{E_{0x}} \cos \Delta\phi \right)^2 &= \left[1 - \left(\frac{\operatorname{Re} E_x}{E_{0x}} \right)^2 \right] \cdot \sin^2 \Delta\phi, \\
\left(\frac{\operatorname{Re} E_y}{E_{0y}} \right)^2 - 2 \left(\frac{\operatorname{Re} E_x}{E_{0x}} \right) \left(\frac{\operatorname{Re} E_y}{E_{0y}} \right) \cos \Delta\phi + \left(\frac{\operatorname{Re} E_x}{E_{0x}} \right)^2 \cos^2 \Delta\phi \\
&= \sin^2 \Delta\phi - \left(\frac{\operatorname{Re} E_x}{E_{0x}} \right)^2 \sin^2 \Delta\phi.
\end{aligned}$$

Umordnen ergibt schließlich die **Ellipsengleichung**

$$\boxed{\left(\frac{\operatorname{Re} E_x}{E_{0x}} \right)^2 + \left(\frac{\operatorname{Re} E_y}{E_{0y}} \right)^2 - 2 \left(\frac{\operatorname{Re} E_x}{E_{0x}} \right) \left(\frac{\operatorname{Re} E_y}{E_{0y}} \right) \cos \Delta\phi = \sin^2 \Delta\phi}$$

für eine **schräge Ursprungsellipse**, d. h. für eine Ellipse, deren Mittelpunkt zwar mit dem Koordinatenursprung zusammenfällt, deren Haupt- und Nebenachse aber nicht parallel zu den Koordinatenachsen verlaufen. Wir sprechen in diesem Fall von einer Ellipse in **nicht achsenparalleler Ursprungslage**.

Achtung!

Weiterführendes zu Ellipsen im Allgemeinen und eine Darstellung der Hauptachsentransformation finden sich im Anhangsabschnitt 17.1 Und die Berechnung von E_{0x} und E_{0y} zeigen wir im Anhangsabschnitt 17.2.

7.4 Jones-Formalismus

Mit Hilfe des Jones-Formalismus⁴, also mit Jones-Vektoren und Jones-Matrizen lässt sich die Gesamtwirkung von nacheinander angeordneten polarisierenden optischen Bauelementen bequemer berechnen.

Zunächst zeigen wir, wie Jones-Vektoren konstruiert werden. Man geht vom elektrischen (optischen) Anteil bzw. dem \vec{E} -Feld einer ebenen elektromagnetischen Welle in der komplexen Darstellung aus und legt den \vec{E} -Vektor in die x, y -Ebene, sodass der Wellenvektor \vec{k} und damit die Ausbreitung der Welle in die positive Richtung der z -Achse zeigt:

$$\begin{aligned}\vec{E}(z, t) &= \begin{pmatrix} E_x(z, t) \\ E_y(z, t) \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i(kz - \omega t + \phi_{0x})} \\ E_{0y} \cdot e^{i(kz - \omega t + \phi_{0y})} \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i\phi_{0x}} \\ E_{0y} \cdot e^{i\phi_{0y}} \\ 0 \end{pmatrix} e^{i(kx - \omega t)} \\ &= \begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \\ 0 \end{pmatrix} e^{i(kx - \omega t)} .\end{aligned}$$

Unterdrückt man jetzt die z -Komponente und den variablen Phasenanteil, so erhält man einen zweidimensionalen Vektor $\tilde{\vec{E}}$, der nur noch die Information über die Amplituden $E_{0x} = \text{const}$ und $E_{0y} = \text{const}$ sowie über die Phasenkonstanten ϕ_{0x} und ϕ_{0y} enthält:

$$\tilde{\vec{E}} = \begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot e^{i\phi_{0x}} \\ E_{0y} \cdot e^{i\phi_{0y}} \end{pmatrix} . \quad (51)$$

„Da der Phasennullpunkt beliebig gewählt werden kann (es kommt nur auf die Differenz $\Delta\phi = \phi_{0y} - \phi_{0x}$ an), können wir $\phi_{0x} = 0$ wählen.“⁵ Demzufolge erhalten wir aus (51)

$$\boxed{\phi_{0x} = 0 \quad \Rightarrow \quad \Delta\phi := \phi_{0y} - \phi_{0x} = \phi_{0y}} \quad \Rightarrow$$

$$\tilde{\vec{E}} = \begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \end{pmatrix} = \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\Delta\phi} \end{pmatrix} .$$

Normiert man schließlich $\tilde{\vec{E}}$ auf (die Länge bzw. das Betragsquadrat) 1, resultiert der

$$\boxed{\text{Jones-Vektor } \vec{J} = \frac{\tilde{\vec{E}}}{|\tilde{\vec{E}}|} = \frac{1}{|\tilde{\vec{E}}|} \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\Delta\phi} \end{pmatrix}}$$

mit

$$\begin{aligned}|\tilde{\vec{E}}| &= \sqrt{E_{0x} e^{i\phi_{0x}} \cdot E_{0x} e^{-i\phi_{0x}} + E_{0y} e^{i\phi_{0y}} \cdot E_{0y} e^{-i\phi_{0y}}} \\ &= \sqrt{E_{0x} \cdot E_{0x} + E_{0y} e^{i\Delta\phi} \cdot E_{0y} e^{-i\Delta\phi}} = \sqrt{E_{0x}^2 + E_{0y}^2} .\end{aligned}$$

Die Jones-Vektoren für die Polarisationszustände \mathcal{E} (allgemeiner Fall), \mathcal{P} , \mathcal{R} und \mathcal{L} sind u. a. unter Berücksichtigung der Tabelle „Zirkulare Polarisation“ :

⁴Siehe auch bei Wikipedia unter dem Suchbegriff *Jones-Formalismus*.

⁵Zitiert aus Wolfgang Demtröder, Springer-Lehrbuch, *Experimentalphysik 2 – Elektrizität und Optik*, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2004, Abschnitt 9.6.7 *Jones-Vektoren*, Seite 286.

- \mathcal{P} -Zustand (lineare Polarisation)

Mit dem Neigungswinkel α zwischen der positiven x -Achse und der nächstgelegenen Schwingungshalbebene gilt für den Amplitudenvektor

$$\vec{E}_0 = \begin{pmatrix} E_{0x} \\ E_{0y} \end{pmatrix} = \begin{pmatrix} E_0 \cos \alpha \\ E_0 \sin \alpha \end{pmatrix} .$$

Und mit $\Delta\phi = 0$ sowie

$$|\tilde{E}| = \sqrt{E_{0x}^2 + E_{0y}^2} = \sqrt{E_0^2 \cos^2 \alpha + E_0^2 \sin^2 \alpha} = E_0$$

ist der zugehörige Jones-Vektor

$$\boxed{\vec{J}_{\mathcal{P}} = \frac{1}{E_0} \begin{pmatrix} E_0 \cdot \cos \alpha \\ E_0 \cdot \sin \alpha \end{pmatrix} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} .}$$

Für

$$\alpha = \pm 45^\circ \quad \Rightarrow \quad \cos(\pm 45^\circ) = 1/\sqrt{2}, \quad \sin(\pm 45^\circ) = \pm 1/\sqrt{2}$$

resultiert dann der Jones-Vektor

$$\boxed{\vec{J}_{\pm 45^\circ} = \frac{1}{E_0} \begin{pmatrix} E_0 \cdot \frac{1}{\sqrt{2}} \\ E_0 \cdot \left(\pm \frac{1}{\sqrt{2}}\right) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} .}$$

Die zwei Sonderfälle der linearen Polarisation sind der horizontale \mathcal{P} -Zustand bzw. die horizontale Polarisation (in x -Richtung) mit dem Index H und der vertikale \mathcal{P} -Zustand bzw. die vertikale Polarisation (in y -Richtung) mit dem Index V. Mit $E_{0x} = E_0 \wedge E_{0y} = 0$ bzw. $E_{0y} = E_0 \wedge E_{0x} = 0$ resultieren also die Jones-Vektoren

$$\boxed{\vec{J}_H = \frac{1}{E_0} \begin{pmatrix} E_0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{J}_V = \frac{1}{E_0} \begin{pmatrix} 0 \\ E_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .}$$

Wir können $\{\vec{J}_H, \vec{J}_V\}$ als **Orthonormalbasis** betrachten, denn

$$\vec{J}_H \cdot \vec{J}_H = \vec{J}_V \cdot \vec{J}_V = 1 \quad \text{und} \quad \vec{J}_H \cdot \vec{J}_V = 0 .$$

Darauf werden wir später noch zurückkommen.

- \mathcal{R} -Zustand (rechtszirkulare Polarisation)

Für den Fall der rechtszirkularen Polarisation erhalten wir unter Berücksichtigung der Tabelle „Zirkulare Polarisation“ und mit (44) bzw. mit $\Delta\phi = -\frac{\pi}{2}$

$$\tilde{\mathbf{E}} = \begin{pmatrix} E_0 \\ E_0 \cdot e^{-i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} E_0 \\ -i E_0 \end{pmatrix}$$

und mit

$$|\tilde{\mathbf{E}}| = \sqrt{E_0^2 + E_0^2} = \sqrt{2} \cdot E_0$$

den Jones-Vektor

$$\boxed{\vec{J}_{\mathcal{R}} = \frac{1}{\sqrt{2} E_0} \begin{pmatrix} E_0 \\ -i E_0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}}.$$

- \mathcal{L} -Zustand (linkszirkulare Polarisation)

Für den Fall der linkszirkularen Polarisation erhalten wir unter Berücksichtigung der Tabelle „Zirkulare Polarisation“ und mit (45) bzw. mit $\Delta\phi = +\frac{\pi}{2}$

$$\tilde{\mathbf{E}} = \begin{pmatrix} E_0 \\ E_0 \cdot e^{+i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} E_0 \\ i E_0 \end{pmatrix}$$

und mit

$$|\tilde{\mathbf{E}}| = \sqrt{E_0^2 + E_0^2} = \sqrt{2} \cdot E_0$$

den Jones-Vektor

$$\boxed{\vec{J}_{\mathcal{L}} = \frac{1}{\sqrt{2} E_0} \begin{pmatrix} E_0 \\ i E_0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}}.$$

- \mathcal{E} -Zustand (elliptische Polarisation, allgemeiner Fall)

Für den Fall der elliptischen Polarisation erhalten wir unter Berücksichtigung von

$$\phi_{0x} = 0 \Rightarrow \Delta\phi = \phi_{0y} - \phi_{0x} = \phi_{0y}, \quad \Delta\phi \neq m \cdot \pi, \quad m \in \mathbb{Z}$$

zunächst den Jones-Vektor in der Form

$$\vec{J}_{\mathcal{E}} = \frac{1}{|\tilde{\mathbf{E}}|} \tilde{\mathbf{E}} = \frac{1}{\tilde{E}} \begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \end{pmatrix} = \frac{1}{\sqrt{E_{0x}^2 + E_{0y}^2}} \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\Delta\phi} \end{pmatrix}.$$

Für die komplexe Zahl $\tilde{E}_y = E_{0y} \cdot e^{i\Delta\phi}$ gilt gemäß der Euler'schen Formel:

$$\begin{aligned} E_{0y} \cdot e^{i\Delta\phi} &= E_{0y} \cdot (\cos \Delta\phi + i \sin \Delta\phi) = \underbrace{E_{0y} \cos \Delta\phi}_{\text{Re}\{\tilde{E}_y\}} + i \cdot \underbrace{E_{0y} \sin \Delta\phi}_{\text{Im}\{\tilde{E}_y\}} \\ &= \text{Re}\{\tilde{E}_y\} + i \cdot \text{Im}\{\tilde{E}_y\} = \text{Re}\tilde{E}_y + i \text{Im}\tilde{E}_y. \end{aligned}$$

Der Jones-Vektor für die elliptische Polarisation ist damit schließlich auch

$$\vec{J}_{\mathcal{E}} = \frac{1}{\sqrt{E_{0x}^2 + E_{0y}^2}} \begin{pmatrix} E_{0x} \\ \operatorname{Re} \tilde{E}_y + i \operatorname{Im} \tilde{E}_y \end{pmatrix}.$$

Für $E_{0x} = E_{0y}$ und $\Delta\phi = m \cdot \pi \pm \frac{\pi}{2}$ resultiert daraus ein Jones-Vektor für die zirkulare Polarisation.

Für $E_{0x} \neq E_{0y}$ und $\Delta\phi = m \cdot \pi \pm \frac{\pi}{2}$ resultiert daraus ein Jones-Vektor für die elliptische Polarisation mit längs der x - oder längs der y -Achse orientierter Hauptachse der Polarisationsellipse.

Bei Jones-Vektoren zur Beschreibung der zirkularen Polarisation sind die Komponenten betragsgleich und ist eine der Komponenten rein imaginär.

Bei Jones-Vektoren zur Beschreibung der elliptischen Polarisation mit längs der x - oder längs der y -Achse orientierter Hauptachse sind die Komponenten nicht betragsgleich und ist eine der Komponenten rein imaginär.

Bei Jones-Vektoren zur Beschreibung der elliptischen Polarisation **mit geneigter Hauptachse** sind die Komponenten allgemein nicht betragsgleich, können aber betragsgleich sein. Immer aber ist in diesem Fall eine der Komponenten komplex mit nicht verschwindendem Realteil und nicht verschwindendem Imaginärteil.

Die Überlagerung von polarisiertem Licht lässt sich durch Addition der entsprechenden Jones-Vektoren beschreiben, wie beispielsweise :

$$\vec{J}_H + \vec{J}_H = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \cdot \vec{J}_H ,$$

$$\vec{J}_H + \vec{J}_V = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \sqrt{2} \cdot \vec{J}_{45^\circ} ,$$

$$\vec{J}_R + \vec{J}_L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \sqrt{2} \cdot \vec{J}_H .$$

Wie man sieht, ist die Summe von Jones-Vektoren nicht normiert. Dass bei der Überlagerung von \vec{J}_R und \vec{J}_L der Jones-Vektor $\sqrt{2} \cdot \vec{J}_H$ (für die horizontale Polarisation) resultiert, liegt an der Art und Weise, wie in (41) die Wellenfunktion $\vec{E}(z, t)$ zur Beschreibung der Polarisation definiert ist. Dort wird nämlich die Phasendifferenz $\Delta\phi$ der y -Komponente zugeordnet gemäß $\phi_{0x} \stackrel{!}{=} 0$.

Man unterscheidet im Wesentlichen drei Arten von polarisierenden optischen Bauelementen:

- **Polarisationsfilter oder Linearpolarisatoren :**

Selektion einer Schwingungsrichtung.

Durch die Unterdrückung aller anderen Schwingungsrichtungen ist die Intensität des ausfallenden Lichts gegenüber dem einfallenden Licht reduziert. Es resultiert also linear polarisiertes Licht, das in der Polarisationssebene bzw. in Richtung der Durchlass- oder Transmissionsachse des Linearpolarisators schwingt. Transmissionsachse und positive x -Achse bilden den Winkel ϑ .

- **Polarisationsdreher :**

Die Schwingungsrichtung bzw. Polarisationssebene linear polarisierten Lichts wird um den Winkel β gedreht bei unveränderter Lichtintensität.

- **Phasenverzögerer :**

Das Licht wird in *einer* Schwingungsrichtung gegenüber der dazu orthogonalen Schwingungsrichtung verzögert. Die Komponente des einfallenden Lichts, die verzögert wird, benötigt mehr Zeit, um das optische Bauelement zu durchlaufen. Phasenverzögerer besitzen also eine durch die Verzögerung langsame Transmissionsachse und folglich orthogonal dazu eine schnelle Achse, wodurch im ausfallenden Licht eine Phasenverschiebung oder **Phasendifferenz** zwischen der Komponente der langsamen Achse und der Komponente der schnellen Achse resultiert.

Polarisierende optische Bauelemente entsprechen **Operatoren** in Gestalt von 2×2 -Matrizen, den **Jones-Matrizen** M . Einfallendes polarisiertes Licht kann durch Jones-Vektoren dargestellt werden, auf die dann die Jones-Matrizen wirken. Jones-Matrizen sind allgemein nicht kommutativ. Werden also mehrere optische Bauelemente nacheinander eingesetzt, müssen die zugehörigen Jones-Matrizen in der gleichen Reihenfolge auf den Jones-Vektor wirken wie die Bauelemente auf das einfallende Licht.

Herleitung einiger Jones-Matrizen

- Ein **Linearpolarisator** habe eine **Transmissionsachse** mit dem Winkel ϑ zur positiven x -Achse. Auf diesen Polarisator mit der Jones-Matrix $M_{\vartheta} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ treffe
 - a) linear polarisiertes Licht mit dem Neigungswinkel $\alpha = \vartheta$, das den Polarisator ungehindert passiert, und
 - b) linear polarisiertes Licht mit dem Neigungswinkel $\alpha + 90^\circ = \vartheta + 90^\circ$, das vollständig blockiert wird.

Mit den entsprechenden Jones-Vektoren

$$\vec{J}_{\alpha} \stackrel{\alpha=\vartheta}{=} \vec{J}_{\vartheta} = \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix} = \vec{J}_{\parallel}$$

und

$$\vec{J}_{(\alpha+90^\circ)} \stackrel{\alpha=\vartheta}{=} \begin{pmatrix} \cos(\vartheta + 90^\circ) \\ \sin(\vartheta + 90^\circ) \end{pmatrix} = \begin{pmatrix} -\sin \vartheta \\ \cos \vartheta \end{pmatrix} = \vec{J}_{\perp}$$

erhalten wir zwei Matrixgleichungen, die uns dann zwei Gleichungssysteme aus jeweils zwei Gleichungen mit zwei Unbekannten zur Bestimmung der Matrixelemente liefern:

$$\begin{aligned} M_{\vartheta} \vec{J}_{\vartheta} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix} = M_{\vartheta} \vec{J}_{\parallel} = \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix}, \\ M_{\vartheta} \vec{J}_{(\vartheta+90^{\circ})} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -\sin \vartheta \\ \cos \vartheta \end{pmatrix} = M_{\vartheta} \vec{J}_{\perp} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \end{aligned}$$

Bestimmung von a und b :

$$\begin{aligned} \text{(I)} : \quad a \cos \vartheta + b \sin \vartheta &= \cos \vartheta, \\ \text{(II)} : \quad a (-\sin \vartheta) + b \cos \vartheta &= 0 \Rightarrow \\ b \cdot \cos \vartheta = a \cdot \sin \vartheta &\Leftrightarrow b = a \frac{\sin \vartheta}{\cos \vartheta}, \quad a \cdot \cos \vartheta + a \frac{\sin \vartheta}{\cos \vartheta} \cdot \sin \vartheta = \cos \vartheta \Rightarrow \\ a &= \cos^2 \vartheta, \\ b &= \cos \vartheta \sin \vartheta. \end{aligned}$$

Bestimmung von c und d :

$$\begin{aligned} \text{(III)} : \quad c \cos \vartheta + d \sin \vartheta &= \sin \vartheta, \\ \text{(IV)} : \quad c (-\sin \vartheta) + d \cos \vartheta &= 0 \Rightarrow \\ d \cdot \cos \vartheta = c \cdot \sin \vartheta &\Leftrightarrow d = c \frac{\sin \vartheta}{\cos \vartheta}, \quad c \cdot \cos \vartheta + c \frac{\sin \vartheta}{\cos \vartheta} \cdot \sin \vartheta = \sin \vartheta \Rightarrow \\ c &= \sin \vartheta \cos \vartheta, \\ d &= \sin^2 \vartheta. \end{aligned}$$

Die Jones-Matrix für den Linearpolarisator mit einer Transmissionsachse im Winkel ϑ zur positiven x -Achse ist also

$$M_{\vartheta} = \begin{pmatrix} \cos^2 \vartheta & \cos \vartheta \sin \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix}.$$

Mit $\sin(\pm 45^{\circ}) = \pm \frac{1}{\sqrt{2}}$ und $\cos(\pm 45^{\circ}) = +\frac{1}{\sqrt{2}}$ erhalten wir damit sofort die entsprechenden Jones-Matrizen für Linearpolarisatoren mit den Transmissionsachsen im Winkel von $\vartheta = +45^{\circ}$ und $\vartheta = -45^{\circ}$:

$$M_{\pm 45^{\circ}} = \begin{pmatrix} \frac{1}{2} & \pm \frac{1}{2} \\ \pm \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}.$$

M_H sei die Jones-Matrix zu einem Linearpolarisator mit horizontaler Transmissionsachse, der also horizontal polarisiertes Licht erzeugt. Wenn die x -Achse horizontal verläuft, gilt $\vartheta = 0^{\circ}$ bzw. $\vartheta = \pm 180^{\circ}$ und damit

$$M_H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

M_V sei die Jones-Matrix zu einem Linearpolarisator mit vertikaler Transmissionsachse, der also vertikal polarisiertes Licht erzeugt. Wenn die y -Achse vertikal verläuft, gilt $\vartheta = \pm 90^\circ$ bzw. $\vartheta = \pm 270^\circ$ und damit

$$M_V = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} .$$

- **Polarisationsdreher** um den Winkel β :

Die Jones-Matrix M_β soll den Jones-Vektor $\vec{J}_\vartheta = \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix}$ durch Drehung der Polarisationssebene um den Winkel β in den Jones-Vektor $\vec{J}_{\vartheta+\beta}$ überführen gemäß $\vec{J}_\vartheta \xrightarrow{M_\beta} \vec{J}_{\vartheta+\beta}$:

$$M_\beta \vec{J}_\vartheta = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix} = \begin{pmatrix} \cos(\vartheta + \beta) \\ \sin(\vartheta + \beta) \end{pmatrix} = \vec{J}_{\vartheta+\beta} \Rightarrow$$

$$\underline{a} \cos \vartheta + \underline{b} \sin \vartheta = \cos(\vartheta + \beta) , \quad \underline{c} \cos \vartheta + \underline{d} \sin \vartheta = \sin(\vartheta + \beta) ,$$

(Anwendung der Additionstheoreme)

$$\underline{\cos \beta} \cos \vartheta - \underline{\sin \beta} \sin \vartheta = \cos(\vartheta + \beta) , \quad \underline{\sin \beta} \cos \vartheta + \underline{\cos \beta} \sin \vartheta = \sin(\vartheta + \beta) .$$

Der Koeffizientenvergleich liefert schließlich

$$a = \cos \beta , \quad b = -\sin \beta , \quad c = \sin \beta , \quad d = \cos \beta ,$$

$$\boxed{M_\beta = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix}} . \quad (52)$$

- **Phasenverzögerer** um den Phasenwinkel bzw. die Phasendifferenz $\Delta\Phi$:

Für das einfallende Licht schreiben wir gemäß (51)

$$\tilde{\mathbf{E}} = \begin{pmatrix} E_{0x} \cdot e^{i\phi_{0x}} \\ E_{0y} \cdot e^{i\phi_{0y}} \end{pmatrix} = \begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \end{pmatrix} .$$

mit den Phasenkonstanten ϕ_{0x} und ϕ_{0y} . Dabei wurde der variable Phasenanteil $e^{i(kz - \omega t)}$ unterdrückt. Ein Phasenverzögerer verändert die Phasen des einfallenden Lichts wie folgt:

$$\begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \end{pmatrix} \longrightarrow \begin{pmatrix} E_{0x} \cdot e^{i(\phi_{0x} + \Phi_x)} \\ E_{0y} \cdot e^{i(\phi_{0y} + \Phi_y)} \end{pmatrix} = \begin{pmatrix} \tilde{E}_x \cdot e^{i\Phi_x} \\ \tilde{E}_y \cdot e^{i\Phi_y} \end{pmatrix} .$$

Es resultiert dabei zwischen den Komponenten \tilde{E}_x und \tilde{E}_y die ggf. zusätzliche

$$\textbf{Phasendifferenz } \Delta\Phi = \Phi_y - \Phi_x ,$$

d. h. eine der beiden Komponenten ist nach der Passage des Phasenverzögerers gegenüber der anderen Komponente verzögert. Die Matrixgleichung dafür lautet

$$M_{\Delta\Phi} \cdot \tilde{\mathbf{E}} = \begin{pmatrix} e^{i\Phi_x} & 0 \\ 0 & e^{i\Phi_y} \end{pmatrix} \begin{pmatrix} \tilde{E}_x \\ \tilde{E}_y \end{pmatrix} = \begin{pmatrix} \tilde{E}_x \cdot e^{i\Phi_x} \\ \tilde{E}_y \cdot e^{i\Phi_y} \end{pmatrix}$$

mit der allgemeinen Jones-Matrix für Phasenverzögerer

$$M_{\Delta\Phi}^{(\dots)} = \begin{pmatrix} e^{i\Phi_x} & 0 \\ 0 & e^{i\Phi_y} \end{pmatrix}. \quad (53)$$

Der phasenverzögerten Komponente entspricht die langsame Achse des Phasenverzögerers. Orthogonal dazu besitzt der Phasenverzögerer demzufolge seine schnelle Achse. Die schnelle Achse der Phasenverzögerer wird in den zugehörigen Jones-Matrizen als hochgestellter Index angegeben. Ausgehend von der Tabelle „Zirkulare Polarisation“ im Abschnitt 7 gilt folglich

$$\Phi_y - \Phi_x = \Delta\Phi > 0 \Rightarrow \Phi_y > \Phi_x \Rightarrow M_{\Delta\Phi}^{(x)}, \text{ } x\text{-Achse ist schnelle Achse,}$$

$$\Phi_y - \Phi_x = \Delta\Phi < 0 \Rightarrow \Phi_y < \Phi_x \Rightarrow M_{\Delta\Phi}^{(y)}, \text{ } y\text{-Achse ist schnelle Achse,}$$

wobei die x -Achse horizontal und die y -Achse vertikal verlaufen soll.

Wichtige Phasenverzögerer sind die **Viertelwellenverzögerungsplatte**, kurz

$$\frac{\lambda}{4}\text{-Platte, entsprechend } \Delta\Phi = \pm\frac{\pi}{2},$$

die **Halbwellenverzögerungsplatte**, kurz

$$\frac{\lambda}{2}\text{-Platte, entsprechend } \Delta\Phi = \pm\pi,$$

und evl. auch die **Achtelwellenverzögerungsplatte**, kurz

$$\frac{\lambda}{8}\text{-Platte, entsprechend } \Delta\Phi = \pm\frac{\pi}{4}.$$

Ausgehend von (53) können wir die zugehörigen Jones-Matrizen beispielsweise wie folgt bestimmen:

$\lambda/4$ -Platte:

Für $\Phi_y = \frac{\pi}{2}, \Phi_x = 0 \Rightarrow \Delta\varphi = \frac{\pi}{2} > 0$ gilt

$$M_{\lambda/4}^{(x)} = \begin{pmatrix} e^0 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}.$$

Für $\Phi_y = 0, \Phi_x = \frac{\pi}{2} \Rightarrow \Delta\varphi = -\frac{\pi}{2} < 0$ gilt

$$M_{\lambda/4}^{(y)} = \begin{pmatrix} e^{i\frac{\pi}{2}} & 0 \\ 0 & e^0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}.$$

Für $\Phi_y = 0, \Phi_x = -\frac{\pi}{2} \Rightarrow \Delta\varphi = \frac{\pi}{2} > 0$ gilt

$$M_{\lambda/4}^{(x)} = \begin{pmatrix} e^{-i\frac{\pi}{2}} & 0 \\ 0 & e^0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & 1 \end{pmatrix}.$$

Für $\Phi_y = -\frac{\pi}{2}, \Phi_x = 0 \Rightarrow \Delta\varphi = -\frac{\pi}{2} < 0$ gilt

$$M_{\lambda/4}^{(y)} = \begin{pmatrix} e^0 & 0 \\ 0 & e^{-i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}.$$

Für $\Phi_y = \frac{\pi}{4}, \Phi_x = -\frac{\pi}{4} \Rightarrow \Delta\varphi = \frac{\pi}{2} > 0$ gilt

$$M_{\lambda/4}^{(x)} = \begin{pmatrix} e^{-i\frac{\pi}{4}} & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} = e^{-i\frac{\pi}{4}} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix}$$

Für $\Phi_y = -\frac{\pi}{4}$, $\Phi_x = \frac{\pi}{4} \Rightarrow \Delta\varphi = -\frac{\pi}{2} < 0$ gilt

$$M_{\lambda/4}^{(y)} = \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & e^{-i\frac{\pi}{4}} \end{pmatrix} = e^{i\frac{\pi}{4}} \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}.$$

Dabei haben wir die folgende Beziehung verwendet:

$$e^{\pm i\frac{\pi}{4}} = \cos(\pm 45^\circ) + i \sin(\pm 45^\circ) = \frac{1}{\sqrt{2}} \pm i \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} (1 \pm i).$$

$\lambda/2$ -Platte:

Für $\Phi_y = \pi$, $\Phi_x = 0 \Rightarrow \Delta\varphi = \pi > 0$ gilt

$$M_{\lambda/2}^{(x)} = \begin{pmatrix} e^0 & 0 \\ 0 & e^{i\pi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Für $\Phi_y = 0$, $\Phi_x = \pi \Rightarrow \Delta\varphi = -\pi < 0$ gilt

$$M_{\lambda/2}^{(y)} = \begin{pmatrix} e^{i\pi} & 0 \\ 0 & e^0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -M_{\lambda/2}^{(x)}.$$

Für $\Phi_y = \frac{\pi}{2}$, $\Phi_x = -\frac{\pi}{2} \Rightarrow \Delta\varphi = \pi > 0$ gilt

$$M_{\lambda/2}^{(x)} = \begin{pmatrix} e^{-i\frac{\pi}{2}} & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

Für $\Phi_y = -\frac{\pi}{2}$, $\Phi_x = \frac{\pi}{2} \Rightarrow \Delta\varphi = -\pi < 0$ gilt

$$M_{\lambda/2}^{(y)} = \begin{pmatrix} e^{i\frac{\pi}{2}} & 0 \\ 0 & e^{-i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = -M_{\lambda/2}^{(x)}.$$

Die Voreilung $\Delta\varphi = +\pi$ und die Nacheilung $\Delta\varphi = -\pi$ sind physikalisch äquivalent.

$\lambda/8$ -Platte:

Für $\Phi_y = \frac{\pi}{4}$, $\Phi_x = 0 \Rightarrow \Delta\varphi = \frac{\pi}{4} > 0$ gilt

$$M_{\lambda/8}^{(x)} = \begin{pmatrix} e^0 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}.$$

Für $\Phi_y = 0$, $\Phi_x = \frac{\pi}{4} \Rightarrow \Delta\varphi = -\frac{\pi}{4} < 0$ gilt

$$M_{\lambda/8}^{(y)} = \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & e^0 \end{pmatrix} = \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & 1 \end{pmatrix}.$$

Anwendungsbeispiele

- Das einfallende Licht sei linear polarisiert mit $\alpha = +45^\circ$ gemäß dem Jones-Vektor \vec{J}_{45° und laufe durch einen Linearpolarisator mit horizontaler Transmissionsachse entsprechend der Jones-Matrix M_H :

$$M_H \vec{J}_{45^\circ} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \vec{J}_H.$$

Das austretende Licht ist die horizontale Komponente des einfallenden Lichts und demzufolge horizontal polarisiert.

- Das einfallende Licht sei horizontal polarisiert gemäß dem Jones-Vektor \vec{J}_H und laufe durch eine $\frac{\lambda}{4}$ -Platte mit schneller y -Achse entsprechend der Jones-Matrix $M_{\lambda/4}^{(y)}$:

$$M_{\lambda/4}^{(y)} \vec{J}_H = \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} i \\ 0 \end{pmatrix} = e^{i\frac{\pi}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e^{i\frac{\pi}{2}} \vec{J}_H .$$

Der Vergleich von einfallendem und austretendem Licht zeigt:

$$\vec{J}_H \hat{=} \vec{E} = \begin{pmatrix} E_0 e^{i(kx-\omega t)} \\ 0 \\ 0 \end{pmatrix} \longrightarrow e^{i\frac{\pi}{2}} \vec{J}_H \hat{=} \vec{E} = \begin{pmatrix} E_0 e^{i(kx-\omega t+\frac{\pi}{2})} \\ 0 \\ 0 \end{pmatrix} .$$

Das austretende Licht ist also ebenfalls horizontal polarisiert und besitzt die gleiche Amplitude wie das einfallende Licht. Allerdings läuft die Phase des austretenden Lichts der Phase des einfallenden Lichts infolge der Verzögerung der x -Komponente durch die $\frac{\lambda}{4}$ -Platte um $\frac{\pi}{2}$ hinterher (siehe Tabelle „Zirkulare Polarisation“).

- Das einfallende Licht sei rechtszirkular polarisiert gemäß dem Jones-Vektor \vec{J}_R und laufe durch einen Linearpolarisator mit vertikaler Transmissionsachse entsprechend der Jones-Matrix M_V :

$$M_V \vec{J}_R = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -i \end{pmatrix} = \frac{e^{-i\frac{\pi}{2}}}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{e^{-i\frac{\pi}{2}}}{\sqrt{2}} \vec{J}_V .$$

Das austretende Licht ist die vertikale Komponente des einfallenden Lichts und demzufolge vertikal polarisiert. In Übereinstimmung mit der Tabelle „Zirkulare Polarisation“ zeigt der Faktor $-i = e^{-i\frac{\pi}{2}}$, dass die herausgefilterte vertikale Komponente der unterdrückten horizontalen Komponente um $\frac{\pi}{2}$ vorausseilt.

- Das einfallende Licht sei linear polarisiert mit dem Neigungswinkel $\alpha = 45^\circ$ und laufe durch eine $\frac{\lambda}{4}$ -Platte mit schneller x -Achse:

$$M_{\lambda/4}^{(x)} \vec{J}_{45^\circ} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \vec{J}_L .$$

Das austretende Licht ist linkszirkular polarisiert.

Wollen wir zirkular polarisiertes Licht erzeugen, brauchen wir also nur mit einem entsprechenden Polarisationsfilter linear polarisiertes Licht mit einem Neigungswinkel von $\pm 45^\circ$ bzw. betragsgleichen Komponenten zu erzeugen und dieses anschließend durch eine $\frac{\lambda}{4}$ -Platte laufen zu lassen.

- Das einfallende Licht sei linkszirkular polarisiert und laufe durch eine $\frac{\lambda}{4}$ -Platte mit schneller x -Achse:

$$M_{\lambda/4}^{(x)} \vec{J}_L = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \vec{J}_{-45^\circ} .$$

Das austretende Licht ist linear polarisiert mit dem Neigungswinkel $\alpha = -45^\circ$

- Das einfallende Licht sei linear polarisiert mit dem Neigungswinkel $\alpha = 45^\circ$ und laufe durch eine $\frac{\lambda}{8}$ -Platte mit schneller x -Achse:

$$M_{\lambda/8}^{(x)} \vec{J}_{45^\circ} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\frac{\pi}{4}} \end{pmatrix} = \vec{J}_E .$$

Wegen $\Delta\phi = \frac{\pi}{4} \neq m\pi + \frac{\pi}{2}$ ist das austretende Licht *nicht* mit längs der x - oder längs der y -Achse ausgerichteter Hauptachse der Polarisationsellipse elliptisch polarisiert, *nicht* zirkular polarisiert und selbstverständlich *nicht* linear polarisiert. Und auch wenn die Komponenten von $\vec{J}_{\mathcal{E}}$ hier gemäß

$$|e^{i\frac{\pi}{4}}| = \sqrt{e^{i\frac{\pi}{4}} \cdot e^{-i\frac{\pi}{4}}} = \sqrt{e^0} = 1$$

betragsgleich sind, so ist das austretende Licht in diesem Fall dennoch elliptisch polarisiert – jedoch mit dem Neigungswinkel α von der positiven x -Achse zur Hauptachse der Polarisationsellipse. Bestimmen wir also α unter Berücksichtigung von

$$\vec{J}_{\mathcal{E}} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \cdot e^{i\frac{\pi}{4}} \end{pmatrix} \Rightarrow \tilde{\vec{E}} = \begin{pmatrix} E_{0x} \\ E_{0y} \cdot e^{i\frac{\pi}{4}} \end{pmatrix} \quad \text{und} \quad E_{0x} = E_{0y} \Rightarrow \frac{E_{0y}}{E_{0x}} = 1$$

mit (48):

$$\tan 2\alpha = \frac{2 \frac{E_{0y}}{E_{0x}} \cdot \cos \Delta\phi}{1 - \left(\frac{E_{0y}}{E_{0x}}\right)^2} = \frac{2 \frac{E_{0y}}{E_{0x}} \cdot \left(+\frac{1}{\sqrt{2}}\right)}{1 - \left(\frac{E_{0y}}{E_{0x}}\right)^2} \Rightarrow$$

$$\lim_{E_{0x} \rightarrow E_{0y}} (\tan 2\alpha) = +\infty \Rightarrow 2\alpha = 90^\circ, \quad \alpha = 45^\circ.$$

Das austretende Licht ist also elliptisch polarisiert mit einem Neigungswinkel der Polarisationsellipse von $\alpha = 45^\circ$.

Es bestätigt sich in diesem Fall, dass bei elliptischer Polarisation mit geneigter Hauptachse der Polarisationsellipse eine der Komponenten von $\vec{J}_{\mathcal{E}}$ komplex ist mit nicht verschwindendem Realteil und nicht verschwindendem Imaginärteil, denn

$$e^{i\frac{\pi}{4}} = \cos 45^\circ + i \sin 45^\circ = \frac{1}{\sqrt{2}} + i \frac{1}{\sqrt{2}}.$$

- Das einfallende Licht sei linkszirkular polarisiert und laufe durch eine $\frac{\lambda}{8}$ -Platte mit schneller x -Achse:

$$M_{\lambda/8}^{(x)} \vec{J}_{\mathcal{L}} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\frac{3\pi}{4}} \end{pmatrix} = \vec{J}_{\mathcal{E}}.$$

Im austretenden Licht ist also $\Delta\phi = \frac{3\pi}{4} = 135^\circ \Rightarrow \cos \Delta\phi = -\frac{1}{\sqrt{2}}$ und die Komponenten des zugehörigen Jones-Vektors sind betragsgleich. Für den Neigungswinkel α der Polarisationsellipse erhalten wir damit:

$$\tan 2\alpha = \frac{2 \frac{E_{0y}}{E_{0x}} \cdot \cos \Delta\phi}{1 - \left(\frac{E_{0y}}{E_{0x}}\right)^2} = \frac{2 \frac{E_{0y}}{E_{0x}} \cdot \left(-\frac{1}{\sqrt{2}}\right)}{1 - \left(\frac{E_{0y}}{E_{0x}}\right)^2} \Rightarrow$$

$$\lim_{E_{0x} \rightarrow E_{0y}} (\tan 2\alpha) = -\infty \Rightarrow 2\alpha = -90^\circ, \quad \alpha = -45^\circ.$$

Das austretende Licht ist elliptisch polarisiert mit einem Neigungswinkel der Polarisationsellipse von $\alpha = -45^\circ$.

- Mit einem Polarisationsfilter erzeugen wir vertikal polarisiertes Licht mit dem Jones-Vektor \vec{J}_V und lassen dieses der Reihe nach zuerst gemäß (52) durch einen Polarisationsdreher mit einem Drehwinkel von $\beta = 45^\circ$, dann durch eine $\frac{\lambda}{4}$ -Platte mit schneller y -Achse und schließlich durch einen Linearpolarisator mit vertikaler Transmissionsachse laufen:

$$\underbrace{M_V \cdot M_{\lambda/4}^{(y)} \cdot M_{(\beta=45^\circ)}}_{\frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}} \cdot \vec{J}_V = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \vec{J}_V.$$

Das austretenden Licht ist vertikal polarisiert, besitzt also nur die y -Komponente des \vec{E} -Feldes. Allerdings ist diese Komponente durch die Polarisationsdrehung mittels $M_{(\beta=45^\circ)}$ um den Faktor $\frac{1}{\sqrt{2}}$ kleiner als im einfallenden Licht. Die vom Polarisationsdreher gebildete x -Komponente wird durch das Polarisationsfilter mit vertikaler Transmissionsachse am Ende unterdrückt, sodass die Intensität des austretenden Lichts gegenüber dem einfallenden Licht um den Faktor $(\frac{1}{\sqrt{2}})^2 = \frac{1}{2}$ kleiner ist.

7.5 Intensität des polarisierten Lichts im Vakuum

Siehe auch Kapitel 2.

Der **Poynting-Vektor**

$$\vec{S} = \vec{E} \times \vec{H} = \varepsilon_0 c_0^2 \vec{E} \times \vec{B}$$

für das Vakuum beschreibt die (momentane) Energiestromdichte eines elektromagnetischen Feldes bzw. des Lichts. Die sich daraus ergebende

$$\text{Intensität } I = \langle |\vec{S}| \rangle = \varepsilon_0 c_0^2 \langle |\vec{E} \times \vec{B}| \rangle = \varepsilon_0 c_0^2 \langle E \cdot \frac{1}{c_0} E \rangle = \varepsilon_0 c_0 \langle E^2 \rangle$$

ist dann das zeitliche Mittel des Betrags des Poynting-Vektors.

Die Intensität einer linear polarisierten ebenen elektromagnetischen Welle

$$\vec{E}(\vec{r}, t) = \vec{E}_0 \begin{pmatrix} \cos(kz - \omega t + \phi_0) \\ \cos(kz - \omega t + \phi_0) \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \cdot \cos(kz - \omega t + \phi_0) \\ E_{0y} \cdot \cos(kz - \omega t + \phi_0) \\ 0 \end{pmatrix},$$

die sich im Vakuum in positiver z -Richtung ausbreitet, ist mit dem Mittelwert $\langle \cos^2 \alpha \rangle = \frac{1}{2}$ und mit

$$E^2(t) = E_{0x}^2 \cos^2(kz - \omega t + \phi_0) + E_{0y}^2 \cos^2(kz - \omega t + \phi_0) = E_0^2 \cos^2(kz - \omega t + \phi_0) :$$

lineare Polarisation : $I = \varepsilon_0 c_0 \langle E^2 \rangle = \frac{\varepsilon_0 c_0}{2} E_0^2$.

Die Intensität der zirkular polarisierten ebenen elektromagnetischen Welle

$$\vec{E}(\vec{r}, t) = \begin{pmatrix} E_0 \cdot \cos(kz - \omega t) \\ E_0 \cdot \sin(kz - \omega t) \\ 0 \end{pmatrix},$$

die sich im Vakuum in positiver z -Richtung ausbreitet, ist mit $\sin^2 \alpha + \cos^2 \alpha = 1$ und mit

$$E^2(t) = E_0^2 \cos^2(kz - \omega t) + E_0^2 \sin^2(kz - \omega t) = E_0^2 = E^2 = \text{const} :$$

zirkulare Polarisation : $I = \varepsilon_0 c_0 \langle E^2 \rangle = \varepsilon_0 c_0 E_0^2 = \varepsilon_0 c_0 E^2$.

8 Polarisation des Lichts in Bra-Ket-Notation

Die Bra-Ket-Notation geht auf Paul Dirac zurück und wird deshalb auch Dirac-Notation genannt. Sie dient insbesondere der bequemen mathematischen Behandlung der Quantenphysik.

8.1 Der Umgang mit komplexen Vektoren und Matrizen in Bra-Ket-Notation

Als Einstieg in dieses Kapitel bzw. zur Erinnerung geben wir kurzgefasst die wichtigsten Regeln zum Umgang mit komplexen Vektoren und Matrizen in der Bra-Ket-Notation an und verweisen auf das diesbezüglich etwas ausführlichere

Kapitel 12 *Rechnen mit komplexen Vektoren und Matrizen* im Skript *Mathematik – Einige ausgewählte Themen für das Physikstudium*.

- **Zustandsvektoren:**

$$\begin{array}{ll} \text{Ket-Vektor (Spaltenvektor)} & |v\rangle := \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \\ \vdots \end{pmatrix}, \\ \text{Bra-Vektor (Zeilenvektor)} & \langle v| := (v_1^* \ v_2^* \ \cdots \ v_i^* \ \cdots). \end{array}$$

- **Adjungiert** (hochgestellter Index \dagger)

heißt komplex konjugiert (hochgestellter Index $*$) und zusätzlich transponiert (hochgestellter Index T), also sinngemäß

$$\dagger = (*)^T = (T)^*.$$

Beispiele:

$$\text{Adjungieren der Matrix } A \quad : \quad A^\dagger = (A^*)^T = (A^T)^*,$$

$$\text{Adjungieren des Spaltenvektors } \vec{v} := \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad : \quad \vec{v}^\dagger = (\vec{v}^*)^T := (v_1^*, v_2^*),$$

$$\text{Adjungieren des Zustands-Ket-Vektors } |v\rangle \quad : \quad |v\rangle^\dagger = (|v\rangle^*)^T = \langle v|.$$

- **Komplexes Standardskalarprodukt** (kurz Skalarprodukt) $\langle \alpha u | \beta v \rangle$ mit $\alpha, \beta \in \mathbb{C}$ ist hermitesch, im ersten Argument antilinear (konjugiert linear) und im zweiten Argument linear:

$$\langle \alpha u | \beta v \rangle = \alpha^* \langle u | \beta v \rangle = \alpha^* \beta \langle u | v \rangle.$$

Ausgeschrieben in Zeilen- und Spaltenvektor erhält dieses Skalarprodukt dann die Gestalt

$$(\alpha^* u_1^* \ \alpha^* u_2^* \ \cdots) \cdot \begin{pmatrix} \beta v_1 \\ \beta v_2 \\ \vdots \end{pmatrix} = \alpha^* \beta (u_1^* v_1 + u_2^* v_2 + \cdots).$$

- **Vollständige Orthonormalbasis** (kurz VON-Basis) oder vollständiges Orthonormalsystem (kurz VONS) wie beispielsweise $\{|a_i\rangle\}$:

Die Basis-Bra-Vektoren $\langle a_i| := (0 \cdots 0 \ a_i^* \ 0 \cdots 0)$ und die

$$\text{Basis-Ket-Vektoren } |a_j\rangle := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_j \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

liefern als Skalarprodukt

$$\langle a_i|a_j\rangle = (\langle a_i|a_j\rangle)^* = \delta_{ij} \Rightarrow \langle a_i|a_i\rangle = 1, \quad \langle a_i|a_j\rangle = 0$$

und als dyadisches Produkt

$$\sum_i |a_i\rangle\langle a_i| = \sum_i \hat{P}_i = \mathbb{1}$$

mit den **Projektionsoperatoren** $\hat{P}_i = |a_i\rangle\langle a_i|$ und der Einheitsmatrix bzw. dem **Identitätsoperator** $\mathbb{1}$.

Beispiel:

$$\sum_{i=1}^2 |a_i\rangle\langle a_i| = \begin{pmatrix} a_1 \\ 0 \end{pmatrix} \cdot (a_1^* \ 0) + \begin{pmatrix} 0 \\ a_2 \end{pmatrix} \cdot (0 \ a_2^*) = \underbrace{\begin{pmatrix} a_1 a_1^* & 0 \\ 0 & 0 \end{pmatrix}}_{P_1} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & a_2 a_2^* \end{pmatrix}}_{P_2},$$

$$\text{mit } a_i a_i^* = a_i^* a_i = \langle a_i|a_i\rangle = 1 \Rightarrow$$

$$\sum_{i=1}^2 |a_i\rangle\langle a_i| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}.$$

- Das Skalarprodukt c_i aus dem Zustandsvektor $|v\rangle$ und dem Basis-Bra-Vektor $|a_i\rangle$, also

$$c_i = \langle a_i|v\rangle,$$

ist die Projektion von $|v\rangle$ auf $|a_i\rangle$, gesprochen: „Skalarprodukt v in a_i “. c_i ist somit die komplexe skalare Vektorkomponente von $|v\rangle$ „in Richtung“ des Basisvektors $|a_i\rangle$.

- **Entwicklungssatz** – Darstellung (Entwicklung) eines (Zustands)vektors $|v\rangle$ in der VON-Basis $\{|a_i\rangle\}$ im \mathbb{C}^n :

$$|v\rangle = \sum_i \underbrace{\langle a_i|v\rangle}_{c_i = v_i} |a_i\rangle = \sum_i v_i |a_i\rangle := \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle a_1|v\rangle \\ \langle a_2|v\rangle \\ \vdots \\ \langle a_i|v\rangle \\ \vdots \end{pmatrix}, \quad (54)$$

$$= \underbrace{\sum_i |a_i\rangle\langle a_i|}_{=\mathbb{1}} |v\rangle = \sum_i \hat{P}_i |v\rangle = \mathbb{1}|v\rangle = |v\rangle. \quad (55)$$

Die Projektionen von $|v\rangle$ auf die Basisvektoren $|a_i\rangle$ bzw. die Skalarprodukte $\langle v|a_i\rangle$ sind die **Entwicklungskoeffizienten** $v_i \in \mathbb{C}$. In (54) steht ganz bewusst nicht das Gleichheitszeichen, sondern das $:=$ -Zeichen („definiert durch“). Die **skalaren** Vektorkomponenten $v_i = \langle a_i|v\rangle$ sind nämlich nicht gleich den zugehörigen (**vektoriellen**) Vektorkomponenten $v_i|a_i\rangle = \langle a_i|v\rangle|a_i\rangle = |a_i\rangle\langle a_i|v\rangle = \hat{P}_i|v\rangle$, wie man bei einem Vergleich von (54) und (55) feststellt und wodurch noch einmal die Bedeutung des Identitätsoperators $\mathbb{1}$ und der Projektionsoperatoren \hat{P}_i deutlich wird.

- Für Operatoren in der **basisfreien** bzw. abstrakten Form schreiben wir beispielsweise \hat{A} .

Im konkreten Fall bezieht man sich jedoch auf eine geschickt gewählte **VON-Basis**. Die dann in dieser Basis dargestellten Operatoren nehmen die Gestalt von Matrizen (A_{ij}) mit den Matrixelementen A_{ij} an. Es gilt deshalb $\hat{A} := (A_{ij})$.

Zur Erläuterung zeigen wir den Übergang von der basisfreien Darstellung zur Darstellung in einer Basis mit einer Analogie am Beispiel des Zusammenhangs zwischen der elektrischen Feldstärke \vec{E} und der elektrischen Flussdichte \vec{D} in anisotropen Materialien, vermittelt durch die Permittivität $\varepsilon = \varepsilon_0 \varepsilon_r$. Dabei ist die Permittivität ein Tensor 2. Stufe bzw. im konkreten Fall eine (3×3) -Matrix mit den Matrixelementen ε_{ij} :

Abstrakte, basisfreie Darstellung: $\vec{D} = \varepsilon \vec{E}$.

Darstellung in der Standardbasis $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$:

$$\vec{D} = \sum_{i=1}^3 D_i \vec{e}_i = \sum_{i=1}^3 \sum_{j=1}^3 \varepsilon_{ij} E_j \vec{e}_j,$$

$$\vec{D} := \begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}.$$

- **Hermiteische Operatoren:**

Observablen werden repräsentiert durch Operatoren und realisieren sich in Gestalt von Messwerten. Messwerte aber können nur reelle Zahlen sein. Deshalb sind alle zu den *üblichen* Observablen in der Quantenmechanik gehörigen Operatoren hermitesch, d. h. selbstadjungiert wie beispielsweise die Matrix

$$A = \begin{pmatrix} 3 & 2+i & i \\ 2-i & 2 & 1-2i \\ -i & 1+2i & 5 \end{pmatrix} = A^\dagger.$$

Hermiteische Operatoren bzw. die zugehörigen hermiteschen Matrizen besitzen nämlich folgende Eigenschaften:

- Ihre Hauptdiagonalelemente sind reell.
- Ihre Eigenwerte sind reell und repräsentieren mögliche Messwerte wie z. B. in den Eigenwertgleichungen $\hat{A}|\psi\rangle = a|\psi\rangle$ bzw. $\hat{A}|\psi_n\rangle = a_n|\psi_n\rangle$.
- Ihre Eigenfunktionen sind orthogonal und bilden ein VONS.

- **Matrixdarstellung von Operatoren** in der VON-Basis $\{|a_i\rangle\}$,
Matrizelement :

Die basisfreie Operatorgleichung

$$|v\rangle = \hat{A}|u\rangle \quad (56)$$

erhält in der VON-Basis $\{|a_i\rangle\}$ die Gestalt

$$\begin{aligned} |v\rangle &= \sum_j \underbrace{\langle a_j|v\rangle}_{v_j} |a_j\rangle = \hat{A}|u\rangle = \hat{A} \sum_j \underbrace{\langle a_j|u\rangle}_{u_j} |a_j\rangle \\ \sum_j v_j |a_j\rangle &= \sum_j u_j \hat{A}|a_j\rangle . \end{aligned} \quad (57)$$

Durch Multiplikation von (57) mit dem Basis-Bra-Vektor $\langle a_i|$ von links erhalten wir

$$\begin{aligned} \sum_j \langle a_i|v_j|a_j\rangle &= \sum_j v_j \underbrace{\langle a_i|a_j\rangle}_{\delta_{ij}} = \sum_j u_j \langle a_i|\hat{A}|a_j\rangle \Rightarrow \\ v_i &= \sum_j \underbrace{\langle a_i|\hat{A}|a_j\rangle}_{A_{ij}} \cdot u_j , \\ v_i &= \sum_j A_{ij} \cdot u_j , \end{aligned} \quad (58)$$

also die skalare Vektorkomponente v_i des Vektors $|v\rangle$ und die

$$\textbf{Matrizelemente} \quad \langle a_i|\hat{A}|a_j\rangle = A_{ij}$$

in der VON-Basis $\{|a_i\rangle\}$. Aus den Matrizelementen A_{ij} ergibt sich die Matrixdarstellung (A_{ij}) des Operators \hat{A} . Mit ihr erhalten wir

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \\ \vdots \end{pmatrix} = (A_{ij}) \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_j \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle a_1|\hat{A}|a_1\rangle & \langle a_1|\hat{A}|a_2\rangle & \cdots & \langle a_1|\hat{A}|a_j\rangle & \cdots \\ \langle a_2|\hat{A}|a_1\rangle & \langle a_2|\hat{A}|a_2\rangle & \cdots & \langle a_2|\hat{A}|a_j\rangle & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ \langle a_i|\hat{A}|a_1\rangle & \langle a_i|\hat{A}|a_2\rangle & \cdots & \langle a_i|\hat{A}|a_j\rangle & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_j \\ \vdots \end{pmatrix} .$$

Die **Matrixdarstellung** eines Operators \hat{A} ist also

$$\boxed{\hat{A} := \sum_{i,j} A_{ij} |a_i\rangle \langle a_j| = (A_{ij})} .$$

In der beispielsweise 2-dimensionalen VON-Basis $\{a_1, a_2\}$ ist dies

$$\begin{aligned} &A_{11}|a_1\rangle\langle a_1| + A_{12}|a_1\rangle\langle a_2| + A_{21}|a_2\rangle\langle a_1| + A_{22}|a_2\rangle\langle a_2| \\ &= A_{11} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \ 0) + A_{12} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (0 \ 1) + A_{21} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (1 \ 0) + A_{22} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \ 1) \\ &= \begin{pmatrix} A_{11} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & A_{12} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ A_{21} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} . \quad \square \end{aligned}$$

Doch wie erhalten wir durch Anwendung des Operators \hat{A} auf den Vektor $|u\rangle$ die Darstellung des Vektors $|v\rangle$ in der Form

$$|v\rangle = v_1|a_1\rangle + v_2|a_2\rangle + \cdots + v_i|a_i\rangle + \cdots = \sum_i v_i|a_i\rangle ?$$

Wir multiplizieren (58) mit dem Basis-Ket-Vektor $|a_i\rangle$ und summieren über i :

$$\begin{aligned} |v\rangle &= \sum_i v_i|a_i\rangle = \sum_i \left(\sum_j \langle a_i|\hat{A}|a_j\rangle \cdot u_j \right) |a_i\rangle \\ &= \sum_i \underbrace{\langle a_i|v\rangle}_{v_i} |a_i\rangle = \sum_{ij} \langle a_i|\hat{A}|a_j\rangle \cdot \underbrace{\langle a_j|u\rangle}_{u_j} |a_i\rangle \\ &= \underbrace{\sum_i |a_i\rangle \langle a_i|}_{\mathbb{1}} v = \underbrace{\sum_{ij} |a_i\rangle \langle a_i|\hat{A}|a_j\rangle \langle a_j|}_{\hat{A}} u, \\ |v\rangle &= \hat{A}|u\rangle. \end{aligned}$$

Wie man sieht, gilt für die Matrixdarstellung (A_{ij}) eines Operators \hat{A} in der VON-Basis $\{|a_i\rangle\}$ auch

$$\boxed{\hat{A} = \mathbb{1}\hat{A}\mathbb{1} := \sum_{ij} |a_i\rangle \langle a_i|\hat{A}|a_j\rangle \langle a_j| = \sum_{ij} A_{ij} |a_i\rangle \langle a_j| = (A_{ij})}.$$

• **Synapse: Adjungierte**

$$[(AB)^T = B^T A^T \Rightarrow (AB)^\dagger = B^\dagger A^\dagger]$$

★ Bra-Vektor $\langle u|$: $\langle u|^\dagger = |u\rangle$.

★ Ket-Vektor $|v\rangle$: $|v\rangle^\dagger = \langle v|$.

★ Skalarprodukt $\langle u|v\rangle$:

$$(\langle u|v\rangle)^\dagger = |v\rangle^\dagger \langle u|^\dagger = \langle v|u\rangle = \sum_i v_i^* u_i = (\langle u|v\rangle)^*.$$

★ dyadisches Produkt $|u\rangle\langle v|$:

$$(|u\rangle\langle v|)^\dagger = \langle v|^\dagger |u\rangle^\dagger = |v\rangle\langle u|.$$

★ Matrix-Vektor-Produkt $A|u\rangle = |v\rangle$:

$$(A|u\rangle)^\dagger = |v\rangle^\dagger = \langle v| = |u\rangle^\dagger A^\dagger = \langle u|A^\dagger,$$

wobei A^\dagger rechts von $\langle u|$ stehen muss, u. a. weil $\langle u|$ ein Zeilenvektor ist.

★ Matrix-Matrix-Vektor-Produkt $A(B|u\rangle) = |v\rangle$:

$$[A(B|u\rangle)]^\dagger = |v\rangle^\dagger = \langle v| = (\langle u|B^\dagger) A^\dagger.$$

Hierbei ist die Reihenfolge von Matrizen und Vektor zu beachten.

8.2 Lineare Polarisation

8.2.1 VON-Basis, Bra- und Ket-Vektoren

Als VON-Basis wählen wir die Jones-Vektoren \vec{J}_H für horizontal bzw. längs der x -Achse polarisiertes und \vec{J}_V für vertikal bzw. längs der y -Achse polarisiertes Licht. In der Bra-Ket-Notation schreiben wir für die Basisvektoren

$$\begin{aligned}\vec{J}_H &\rightarrow |H\rangle \quad \text{und} \quad \vec{J}_H^\dagger \rightarrow \langle H|, \\ \vec{J}_V &\rightarrow |V\rangle \quad \text{und} \quad \vec{J}_V^\dagger \rightarrow \langle V|.\end{aligned}$$

Damit erhält die VON-Basis die Gestalt

$$\begin{aligned}\{|a_i\rangle\} &= \{|a_1\rangle, |a_2\rangle\} \\ \longrightarrow &= \{|H\rangle, |V\rangle\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},\end{aligned}$$

sodass beispielsweise

$$\begin{aligned}\langle a_i | a_j \rangle &= \delta_{ij} \\ \longrightarrow &= \langle H | V \rangle = (1 \ 0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0, \\ \longrightarrow &= \langle V | V \rangle = (0 \ 1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1\end{aligned}$$

und

$$\begin{aligned}\sum_i |a_i\rangle \langle a_i| &= \sum_i P_i = \mathbb{1} \\ \longrightarrow &= |H\rangle \langle H| + |V\rangle \langle V| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \ 0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \ 1) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}\end{aligned}$$

gilt. Wir werden jetzt linear polarisiertes Licht bzw. das linear polarisierte optische Feld

$$\begin{aligned}\tilde{\vec{E}} &:= \begin{pmatrix} E_{0x} \\ E_{0y} \end{pmatrix} = E_0 \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} = E_0 \cdot \vec{J}_P \Rightarrow \\ \tilde{\vec{E}} &= E_{0x} \vec{e}_x + E_{0y} \vec{e}_y = E_0 \cos \alpha \cdot \vec{e}_x + E_0 \sin \alpha \cdot \vec{e}_y = \sum_i \left(\vec{e}_i \cdot \tilde{\vec{E}} \right) \cdot \vec{e}_i, \quad i \in \{x, y\},\end{aligned}$$

mit dem Neigungswinkel α zur positiven x -Achse in der Bra-Ket-Notation und in der oben gewählten VON-Basis darstellen. Damit gilt für den Jones-Vektor \vec{J}_P :

$$\vec{J}_P \rightarrow \cos \alpha |H\rangle + \sin \alpha |V\rangle \quad \text{und} \quad \vec{J}_P^\dagger \rightarrow \cos \alpha \langle H| + \sin \alpha \langle V|.$$

Analog zur Darstellung eines Bra-Vektors

$$|u\rangle = \sum_i \underbrace{\langle a_i | u \rangle}_{= u_i} |a_i\rangle = \sum_i \underbrace{|a_i\rangle \langle a_i | u \rangle}_{= \mathbb{1}}$$

in der VON-Basis $\{|a_i\rangle\}$ wird demzufolge das linear polarisierte optische Feld $\tilde{\mathbf{E}}$ dargestellt durch den Bra-Vektor $|e\rangle$ gemäß

$$E_0 \cdot \vec{\mathbf{J}}_{\mathcal{P}} \longrightarrow |e\rangle = \underbrace{E_0 \cos \alpha}_{\langle H|e\rangle = E_{0x}} |H\rangle + \underbrace{E_0 \sin \alpha}_{\langle V|e\rangle = E_{0y}} |V\rangle ,$$

$$\boxed{\tilde{\mathbf{E}} \longrightarrow |e\rangle = \langle H|e\rangle |H\rangle + \langle V|e\rangle |V\rangle = E_{0x} |H\rangle + E_{0y} |V\rangle} . \quad (59)$$

8.2.2 Polarisationsfilter mit horizontaler Transmissionsachse

Analog zur Wirkung des Operators \hat{A} auf den Bra-Vektor $|u\rangle$ gemäß

$$\text{basisfrei} \quad : \quad |v\rangle = \hat{A} |u\rangle , \quad (60)$$

$$\text{Basis } \{|a_i\rangle\} \quad : \quad |v\rangle = \sum_{ij} \langle a_i | \hat{A} | a_j \rangle \cdot \langle a_j | u \rangle |a_i\rangle , \quad (61)$$

$$v_i = \sum_j \underbrace{\langle a_i | \hat{A} | a_j \rangle}_{A_{ij}} \cdot \underbrace{\langle a_j | u \rangle}_{u_j} = \sum_j A_{ij} \cdot u_j \Rightarrow$$

$$|v\rangle := (A_{ij}) \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_j \\ \vdots \end{pmatrix}$$

lassen wir jetzt die Jones-Matrix M_H für das Polarisationsfilter, das die horizontal polarisierte Komponente des einfallenden Lichts herausfiltert, auf linear polarisiertes Licht (59) entsprechend dem Jones-Vektor $\vec{\mathbf{J}}_{\mathcal{P}}$ mit dem Neigungswinkel α wirken. Das einfallende Licht indizieren wir mit „in“ und das austretende mit „out“:

$$\tilde{\mathbf{E}}_{\text{out}} = M_H \cdot \tilde{\mathbf{E}}_{\text{in}} = M_H \cdot \underbrace{E_0 \vec{\mathbf{J}}_{\mathcal{P}}}_{\tilde{\mathbf{E}}_{\text{in}}} = E_0 \cdot M_H \vec{\mathbf{J}}_{\mathcal{P}} , \quad (62)$$

$$\tilde{\mathbf{E}}_{\text{out}} := E_0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_0 \cdot \cos \alpha \\ E_0 \cdot \sin \alpha \end{pmatrix} = \begin{pmatrix} E_0 \cos \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} E_{0x} \\ 0 \end{pmatrix} . \quad (63)$$

Wie erwartet wird die horizontale x -Komponente von $\tilde{\mathbf{E}}_{\text{in}}$, also die *Projektion* von $\tilde{\mathbf{E}}_{\text{in}}$ auf den Basisvektor $\vec{\mathbf{J}}_H$, herausgefiltert.

Weil wir die Jones-Matrizen im Abschnitt 7.4 bezüglich der VON-Basis $\{\vec{\mathbf{J}}_H, \vec{\mathbf{J}}_V\}$ berechnet hatten,¹ sind die Elemente der Jones-Matrizen gleich den Matrixelementen des Operators \hat{A} bzw. der Matrix (A_{ij}) . Das bedeutet, dass der Operator \hat{A} die Funktion des polarisierenden optischen Elements bzw. hier des Polarisationsfilters repräsentiert. Kurz gesagt:

$$M_H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |H\rangle \langle H|$$

ist der *Projektionsoperator* für die horizontale Projektion und repräsentiert damit das entsprechende Polarisationsfilter.

¹Eine VON-Basis legt man fest. Die Matrixdarstellung eines Operators jedoch muss man „finden“.

Den Übergang von der Gleichung (62) im Jones-Formalismus zur Gleichung

$$|e_{\text{out}}\rangle = M_{\text{H}}|e_{\text{in}}\rangle$$

in Bra-Ket-Notation erläutern wir, indem wir, angepasst an unseren Fall, zunächst die allgemeine basisfreie Gleichung (60) mittels (61) in der speziellen

$$\text{VON-Basis } \{|a_i\rangle\} = \{|a_1\rangle, |a_2\rangle\}$$

darstellen. Dabei verwenden wir für den Operator \hat{A} , ebenfalls angepasst an unseren Fall, die Matrix $|a_1\rangle\langle a_1|$ gemäß

$$\begin{aligned} \hat{A} &:= |a_1\rangle\langle a_1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = M_{\text{H}} = |H\rangle\langle H| : \\ |v\rangle &= \hat{A}|u\rangle \\ &= \sum_{i=1}^2 \sum_{j=1}^2 \langle a_i|\hat{A}|a_j\rangle \cdot \langle a_j|u\rangle |a_i\rangle = \sum_{i=1}^2 \sum_{j=1}^2 \langle a_i|a_1\rangle \langle a_1|a_j\rangle \cdot \langle a_j|u\rangle |a_i\rangle \\ &= \underbrace{\langle a_1|a_1\rangle}_{1} \underbrace{\langle a_1|a_1\rangle}_{1} \cdot \underbrace{\langle a_1|u\rangle}_{u_1} |a_1\rangle + \underbrace{\langle a_1|a_1\rangle}_{1} \underbrace{\langle a_1|a_2\rangle}_{0} \cdot \underbrace{\langle a_2|u\rangle}_{u_2} |a_1\rangle + \\ &\quad \underbrace{\langle a_2|a_1\rangle}_{0} \underbrace{\langle a_1|a_1\rangle}_{1} \cdot \underbrace{\langle a_1|u\rangle}_{u_1} |a_2\rangle + \underbrace{\langle a_2|a_1\rangle}_{0} \underbrace{\langle a_1|a_2\rangle}_{0} \cdot \underbrace{\langle a_2|u\rangle}_{u_2} |a_2\rangle = \underbrace{\langle a_1|u\rangle}_{u_1} |a_1\rangle \quad (64) \\ &= (A_{11} u_1 + A_{12} u_2)|a_1\rangle + (A_{21} u_1 + A_{22} u_2)|a_2\rangle = A_{11} u_1 |a_1\rangle = u_1 |a_1\rangle, \end{aligned}$$

$$|v\rangle := \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} A_{11} u_1 \\ 0 \end{pmatrix} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix}. \quad (65)$$

Ausgehend von (64) und (65) können wir jetzt unter Berücksichtigung von

$$|v\rangle = \hat{A}|u\rangle \Rightarrow \begin{cases} |a_1\rangle = |H\rangle, & |a_2\rangle = |V\rangle, \\ |u\rangle = |e_{\text{in}}\rangle = E_0 (\cos \alpha |H\rangle + \sin \alpha |V\rangle) = E_{0x}|H\rangle + E_{0y}|V\rangle = \tilde{\mathbf{E}}_{\text{in}}, \\ u_1 = E_0 \cos \alpha = E_{0x}, & u_2 = E_0 \sin \alpha = E_{0y}, \\ |v\rangle = |e_{\text{out}}\rangle = \tilde{\mathbf{E}}_{\text{out}} \end{cases} \quad (66)$$

die Wirkung des Polarisationsfilters mit der

$$\text{Jones-Matrix } M_{\text{H}} = (A_{ij}) =: \hat{A}$$

auf linear polarisiertes Licht mit dem Jones-Vektor $\tilde{\mathbf{J}}_{\mathcal{P}}$ in Bra-Ket-Notation zeigen:

$$\begin{aligned} |e_{\text{out}}\rangle &= |H\rangle\langle H|e_{\text{in}}\rangle \\ &= \langle H|e_{\text{in}}\rangle |H\rangle = \langle H|E_0 (\cos \alpha |H\rangle + \sin \alpha |V\rangle) |H\rangle \\ &= \underbrace{E_0 \cos \alpha}_{E_{0x}} \underbrace{\langle H|H\rangle}_1 |H\rangle + \underbrace{E_0 \sin \alpha}_{E_{0y}} \underbrace{\langle H|V\rangle}_0 |H\rangle, \\ |e_{\text{out}}\rangle &= E_{0x}|H\rangle, \\ |e_{\text{out}}\rangle &:= \begin{pmatrix} E_{0x} \\ 0 \end{pmatrix}. \end{aligned}$$

Dieses Ergebnis stimmt mit (63) überein.

8.2.3 Polarisationsfilter mit Transmissionsachse im Winkel ϑ

Die Wirkung des Polarisationsfilters mit einer Transmissionsachse im Winkel ϑ zur positiven x -Achse auf linear polarisiertes Licht entsprechend dem Jones-Vektor $\vec{J}_{\mathcal{P}}$ wird mit Hilfe des Jones-Formalismus beschrieben durch

$$\begin{aligned}\tilde{\vec{E}}_{\text{out}} &= M_{\vartheta} \cdot \tilde{\vec{E}}_{\text{in}} = M_{\vartheta} \cdot \underbrace{E_0 \vec{J}_{\mathcal{P}}}_{\tilde{\vec{E}}_{\text{in}}} = E_0 \cdot M_{\vartheta} \vec{J}_{\mathcal{P}} , \\ \tilde{\vec{E}}_{\text{out}} &:= \begin{pmatrix} \cos^2 \vartheta & \cos \vartheta \sin \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix} \cdot E_0 \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} = \begin{pmatrix} \cos^2 \vartheta & \cos \vartheta \sin \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix} \begin{pmatrix} E_{0x} \\ E_{0y} \end{pmatrix} \\ &:= E_0 \cdot \begin{pmatrix} \cos^2 \vartheta \cos \alpha + \cos \vartheta \sin \vartheta \sin \alpha \\ \sin \vartheta \cos \vartheta \cos \alpha + \sin^2 \vartheta \sin \alpha \end{pmatrix} .\end{aligned}\tag{67}$$

Bei der Darstellung von (67) in der Bra-Ket-Notation unter Berücksichtigung von (66) und mit der

$$\text{Jones-Matrix } M_{\vartheta} = (A_{ij}) =: \hat{A}$$

zeigt sich, wie im Abschnitt 8.2.2 erläutert, dass die Elemente der Jones-Matrix M_{ϑ} gleich den Matrixelementen des entsprechenden Operators \hat{A} sind. Wir wollen dies deshalb lediglich mit der Berechnung *eines* Matrixelements analog zu (64) demonstrieren:

$$\begin{aligned}A_{21} &= \langle a_2 | M_{\vartheta} | a_1 \rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \cos^2 \vartheta & \cos \vartheta \sin \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \cos^2 \vartheta \\ \sin \vartheta \cos \vartheta \end{pmatrix} = \sin \vartheta \cos \vartheta = A_{21} . \quad \square\end{aligned}$$

Damit erhalten wir schließlich

$$\begin{aligned}|e_{\text{out}}\rangle &= \overbrace{\cos^2 \vartheta}^{A_{11}} \overbrace{\langle H|u\rangle}^{u_1} |H\rangle + \overbrace{\cos \vartheta \sin \vartheta}^{A_{12}} \overbrace{\langle V|u\rangle}^{u_2} |H\rangle + \\ &\quad \underbrace{\sin \vartheta \cos \vartheta}_{A_{21}} \underbrace{\langle H|u\rangle}_{u_1} |V\rangle + \underbrace{\sin^2 \vartheta}_{A_{22}} \underbrace{\langle V|u\rangle}_{u_2} |V\rangle \\ &= \cos^2 \vartheta \cdot \overbrace{E_0 \cos \alpha}^{E_{0x}} |H\rangle + \cos \vartheta \sin \vartheta \cdot \overbrace{E_0 \sin \alpha}^{E_{0y}} |H\rangle + \\ &\quad \sin \vartheta \cos \vartheta \cdot \underbrace{E_0 \cos \alpha}_{E_{0x}} |V\rangle + \sin^2 \vartheta \cdot \underbrace{E_0 \sin \alpha}_{E_{0y}} |V\rangle , \\ |e_{\text{out}}\rangle &:= \begin{pmatrix} \cos^2 \vartheta & \cos \vartheta \sin \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix} \begin{pmatrix} E_{0x} \\ E_{0y} \end{pmatrix} . \quad \square\end{aligned}$$

8.3 Zirkulare Polarisation

In Analogie zum Jones-Formalismus (siehe Tabelle 2) schreiben wir unter Verwendung der VON-Basis $\{|H\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |V\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ für linkszirkular(\mathcal{L})-polarisiertes und rechtszirkular(\mathcal{R})-polarisiertes Licht in Bra-Ket-Notation²

$$\begin{aligned}\mathcal{L} : \quad \vec{J}_{\mathcal{L}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} &\longrightarrow |R\rangle = |\sigma^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle) := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \\ \mathcal{R} : \quad \vec{J}_{\mathcal{R}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} &\longrightarrow |L\rangle = |\sigma^-\rangle = \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle) := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.\end{aligned}$$

Dabei sind die Faktoren $\pm i = e^{\pm i\frac{\pi}{2}}$ Phasenfaktoren und stehen für die Unterschiede $\Delta\phi = \pm\frac{\pi}{2}$ der Phase von \tilde{E}_y gegenüber \tilde{E}_x .

Um Verwechslungen zwischen den Bezeichnungen \mathcal{L} , \mathcal{R} und $|R\rangle$, $|L\rangle$ zu vermeiden, schreiben wir in der Bra-Ket-Notation für linkszirkulare Polarisationszustände $|\sigma^+\rangle$ und für rechtszirkulare Polarisationszustände $|\sigma^-\rangle$.

8.4 Photonenspin

Der Begriff Photonenspin wurde im Kapitel 1 eingeführt. Den Spin s_z eines sich in z -Richtung fortbewegenden Photons erhalten wir durch die Anwendung des

$$\textbf{Photonen-Spinoperators} \quad \hat{s}_z := \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

auf die Zustandsvektoren $|\sigma^+\rangle$ oder $|\sigma^-\rangle$ dieses Photons:

$$\hat{s}_z|\sigma^+\rangle = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \hbar \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \hbar|\sigma^+\rangle \quad (68)$$

$$\Rightarrow s_z = \hbar \quad \text{bzw.} \quad \vec{s} = \begin{pmatrix} 0 \\ 0 \\ \hbar \end{pmatrix},$$

$$\hat{s}_z|\sigma^-\rangle = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \hbar \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ i \end{pmatrix} = -\hbar|\sigma^-\rangle \quad (69)$$

$$\Rightarrow s_z = -\hbar \quad \text{bzw.} \quad \vec{s} = \begin{pmatrix} 0 \\ 0 \\ -\hbar \end{pmatrix}.$$

Die Matrix $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ ist wie die Pauli-Matrizen hermitesch, unitär (mit Determinante -1) und spurlos. Deshalb sind ihre Eigenwerte reell mit Betrag 1 und in der Summe gleich Null. Folglich sind ihre Eigenwerte $+1$ und -1 . Und die normierten Eigenvektoren³ sind

$$|\sigma^+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ zum Eigenwert } +1 \quad \text{und} \quad |\sigma^-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ zum Eigenwert } -1.$$

²Siehe dazu: Wolfgang Demtröder, Springer-Lehrbuch, *Experimentalphysik 2 – Elektrizität und Optik*, 3. Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 2004, Seite 193.

³Alle Vielfachen von Eigenvektoren sind auch Eigenvektoren. Deshalb findet man in diesem Fall bei der Berechnung beispielsweise mit einem „Online-Eigenvektoren-Berechner“ auch die Eigenvektoren $\begin{pmatrix} -1 \\ i \end{pmatrix} = -i\sqrt{2} \cdot |\sigma^+\rangle$ zum Eigenwert $+1$ und $\begin{pmatrix} 1 \\ i \end{pmatrix} = i\sqrt{2} \cdot |\sigma^-\rangle$ zum Eigenwert -1 .

Die Anwendung des Photonen-Spinoperators auf die Basis-Zustände $|H\rangle$ und $|V\rangle$ liefert

$$\hat{s}_z \cdot \begin{cases} |H\rangle \\ |V\rangle \end{cases} = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{cases} = \hbar \cdot \begin{cases} \begin{pmatrix} 0 \\ i \end{pmatrix} \\ \begin{pmatrix} -i \\ 0 \end{pmatrix} \end{cases} = \begin{cases} i\hbar |V\rangle \\ -i\hbar |H\rangle \end{cases}$$

und damit keine physikalisch sinnvollen Ergebnisse, weil $|H\rangle$ und $|V\rangle$ keine Eigenzustände zum Operator \hat{s} sind und weil in diesem Fall komplexwertige Messergebnisse für den Spin resultieren würden. Gemäß den Eigenwertgleichungen (68) und (69) gehören nämlich zum Operator \hat{s}_z die Eigenzustände $|\sigma^+\rangle_{(\hbar)}$ und $|\sigma^-\rangle_{(-\hbar)}$ mit den Eigenwerten \hbar und $-\hbar$.

9 Tabellen: Jones-Vektoren und Jones-Matrizen

Tabelle 1 Jones-Vektoren

Polarisation	Jones-Vektor	Bra-Ket-Notation
linear in x -Richtung (horizontal)	$\vec{J}_H = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$ H\rangle$
linear in y -Richtung (vertikal)	$\vec{J}_V = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$ V\rangle$
linear mit Neigungswinkel $\pm 45^\circ$	$\vec{J}_{\pm 45^\circ} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$	$ D\rangle = \frac{1}{\sqrt{2}} (H\rangle + V\rangle)$
linear mit Neigungswinkel α	$\vec{J}_P = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$	$\cos \alpha H\rangle + \sin \alpha V\rangle$
linkszirkular	$\vec{J}_L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$	$ \sigma^+\rangle = \frac{1}{\sqrt{2}} (H\rangle + i V\rangle)$
rechtszirkular	$\vec{J}_R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$	$ \sigma^-\rangle = \frac{1}{\sqrt{2}} (H\rangle - i V\rangle)$
elliptisch allgemein	$\vec{J}_E = \frac{1}{\sqrt{E_{0x}^2 + E_{0y}^2}} \begin{pmatrix} E_{0x} \\ E_{0y} e^{i\Delta\phi} \end{pmatrix}$	

Tabelle 2 Jones-Matrizen, Abkürzung TA für Transmissionsachse.

optisches Element	Jones-Matrix
Polarisationsfilter – TA horizontal	$M_H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
Polarisationsfilter – TA vertikal	$M_V = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
Polarisationsfilter – TA im Winkel $\pm 45^\circ$	$M_{\pm 45^\circ} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}$
Polarisationsfilter – TA im Winkel ϑ	$M_\vartheta = \begin{pmatrix} \cos^2 \vartheta & \cos \vartheta \sin \vartheta \\ \sin \vartheta \cos \vartheta & \sin^2 \vartheta \end{pmatrix}$
Polarisationsdreher um Winkel β	$M_\beta = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix}$
Phasenverzögerer um Phasendifferenz $\Delta\Phi$	$M_{\Delta\Phi} = \begin{pmatrix} e^{i\Phi_x} & 0 \\ 0 & e^{i\Phi_y} \end{pmatrix}$
$\frac{\lambda}{2}$ -Platte mit schneller x -Achse	$M_{\lambda/2}^{(x)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\frac{\lambda}{2}$ -Platte mit schneller y -Achse	$M_{\lambda/2}^{(y)} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
$\frac{\lambda}{4}$ -Platte mit schneller x -Achse	$M_{\lambda/4}^{(x)} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$
$\frac{\lambda}{4}$ -Platte mit schneller y -Achse	$M_{\lambda/4}^{(y)} = \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}$
$\frac{\lambda}{8}$ -Platte mit schneller x -Achse	$M_{\lambda/8}^{(x)} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$
$\frac{\lambda}{8}$ -Platte mit schneller y -Achse	$M_{\lambda/8}^{(y)} = \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & 1 \end{pmatrix}$

10 Paradoxien im Teilchenbild

Man spricht im Zusammenhang mit quantenphysikalischen Phänomenen gelegentlich von sog. Paradoxien, wenn sich diese Phänomene mit den klassischen Vorstellungen von Teilchen bzw. Körpern nicht erklären lassen.

So hat beispielsweise ein quantenphysikalisches Teilchen wie das Elektron (fast) nichts gemein mit einer Billardkugel, außer dass Elektronen ein Bestandteil der Billardkugel sind. Während man (makroskopischen) Körpern auf der makroskopischen Skala ein im Rahmen der Messgenauigkeit „exaktes“ Volumen und einen im Rahmen der Messgenauigkeit „exakten“ Aufenthaltsort zubilligen kann, besitzen quantenphysikalische Teilchen keine fest umrissene Gestalt bzw. keine exakten räumlichen Grenzen und demzufolge auch kein festes Volumen (siehe Heisenberg'sche Impuls-Ortsunschärfe und Energie-Zeitunschärfe). Allgemein kann man deshalb in der Realität keine Aussage über den jeweiligen exakten Aufenthaltsort eines quantenphysikalischen Teilchens treffen.

Man stößt also zwangsläufig auf Widersprüche bei dem Versuch, quantenphysikalische Phänomene im Teilchenbild auf klassische Weise zu erklären. Wir diskutieren diese Problematik am Beispiel zweier Phänomene.

Reflexionsstrahlteiler

Betrachten wir einen Reflexionsstrahlteiler (halbdurchlässigen Spiegel), der die Intensität des einfallenden Lichts bei der Reflexion/Transmission halbiert. Im Wellenbild ist die Beschreibung dieses Phänomens unproblematisch, nicht jedoch im Teilchenbild. Es stellt sich hier die Frage, was mit dem einzelnen Photon (Wellenpaket) der Energie $E = h \cdot \nu$ geschieht, wenn es auf den Strahlteiler trifft. Zu einer Teilung in ein reflektiertes und ein transmittiertes Photon mit jeweils der halben Energie $\frac{1}{2}E = h \cdot \frac{1}{2}\nu$ kommt es nicht, denn die Lichtfrequenz ν bleibt bei der Strahlteilung erhalten. Die Photonen sind unteilbar. Tatsächlich kann man messen, dass jedes einzelne Photon mit einer Wahrscheinlichkeit von in unserem Fall 50 % reflektiert oder transmittiert wird. Das Ergebnis (Reflexion oder Transmission) im Einzelfall ist dabei nicht vorhersehbar.

Das bedeutet folgendes: Im klassischen Sinne sollte es möglich sein, den Vorgang der Strahlteilung so zu beschreiben, dass er bei jeder Messung zumindest im Idealfall in der gleichen Weise abläuft. Das ist hier im Teilchenbild jedoch nicht möglich, was zu dem Schluss führt, dass das Ergebnis nicht unbeeinflusst von der Messung eintritt sondern durch die Messung herbeigeführt bzw. realisiert wird.

Interferenz am Doppelspalt

Die Intensitätsverteilung im Interferenzmuster beim Doppelspalt zeigt deutliche erste Minima neben dem Hauptmaximum, was sich mit dem Wellencharakter des Lichts problemlos erklären lässt. Allerdings gilt dies auch für quantenphysikalische Teilchen mit einer Ruhemasse. Wir betrachten deshalb im Folgenden die Interferenz von Elektronen am Doppelspalt. Es handelt sich hierbei um ein ortsabhängiges quantenphysikalisches Phänomen.

Im Teilchenbild würde man zunächst intuitiv stets die **Überlagerung der Elektrentrefferbilder** der beiden Einzelspalte auf einem Bildschirm erwarten, wobei es jedem einzelnen Elektron, wenn es einen Spalt passiert, „egal“ sein sollte, ob der andere Spalt geöffnet ist oder nicht. Besitzt ein Elektron, das den Spalt 1 passiert hat, die Wellenfunktion $\psi_1(x)$ mit der zugehörigen Wahrscheinlichkeitsdichte $w_1(x) = |\psi_1(x)|^2$ und ein Elektron, das den Spalt 2 passiert hat, die Wellenfunktion $\psi_2(x)$ mit der zugehörigen Wahrscheinlichkeitsdichte $w_2(x) = |\psi_2(x)|^2$, dann entspräche die Überlagerung bzw. die Summe

$$|\psi_1(x)|^2 + |\psi_2(x)|^2 = \tilde{w}(x) \quad (70)$$

der beiden Wahrscheinlichkeitsdichten $w_1(x)$ und $w_2(x)$ dem zu erwartenden Trefferbild. Dieses Bild besitzt allerdings keine ersten Minima und entsteht nur dann, wenn die Elektronen an einem oder an beiden Spalten detektiert werden. (70) beschreibt also nicht die Interferenz der Elektronen am Doppelspalt.

Sind beide Spalte geöffnet und werden die Elektronen beim Passieren der Spalte nicht beobachtet bzw. detektiert, geht jedes Elektron quasi gleichzeitig durch beide Spalte und bildet danach einen

$$\textbf{Überlagerungszustand} \quad \psi(x) = \psi_1(x) + \psi_2(x) \quad (71)$$

aus den beiden Wellenfunktionen $\psi_1(x)$ von Spalt 1 und $\psi_2(x)$ von Spalt 2.

Unbeobachtet lässt sich demzufolge den Elektronen bezüglich der Passage durch die Spalte kein bestimmter Ort zuordnen.

Das Doppelspalt-Interferenzmuster ergibt sich damit schließlich aus der Wahrscheinlichkeitsdichte

$$w(x) = |\psi_1(x) + \psi_2(x)|^2 = |\psi_1(x)|^2 + |\psi_2(x)|^2 + \underbrace{\psi_1(x)\psi_2^*(x) + \psi_1^*(x)\psi_2(x)}_{\text{Interferenzterm}}$$

des Überlagerungszustands (71).

Dem Elektron an dem einen Spalt ist es also nicht egal, ob der andere Spalt geöffnet oder geschlossen ist. Und es ist den Elektronen auch nicht egal, ob sie beobachtet bzw. detektiert werden oder nicht, denn die Detektion (Beobachtung, Messung) **reduziert** die Überlagerungswellenfunktion bzw. den Zustand $\psi(x)$ der Elektronen auf die Wellenfunktionen bzw. Zustände $\psi_1(x)$ und $\psi_2(x)$, was den **Kollaps** des Überlagerungszustands (71) bedeutet und das Verschwinden des Interferenzterms und damit des Interferenzmusters zur Folge hat. Es resultiert nach dem Kollaps des Überlagerungszustands das Überlagerungsbild entsprechend $|\psi_1(x)|^2 + |\psi_2(x)|^2$ aus den Elektronentrefferbildern jedes der beiden Einzelspalte.

Dieser Sachverhalt führt zu einer grundlegenden Annahme in der Quantenphysik bzw. quantenphysikalischen Theorie:

Bereits einzelne quantenphysikalische Teilchen (Elementarteilchen, Photonen, Atome usw.) gehorchen dem Superpositionsprinzip und können sich demzufolge in einem Überlagerungszustand verschiedener quantenphysikalischer Zustände befinden.

11 Grundlegendes zur 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

11.1 Der Begriff „Wellenoptik“

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**¹:

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.

¹Das (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.

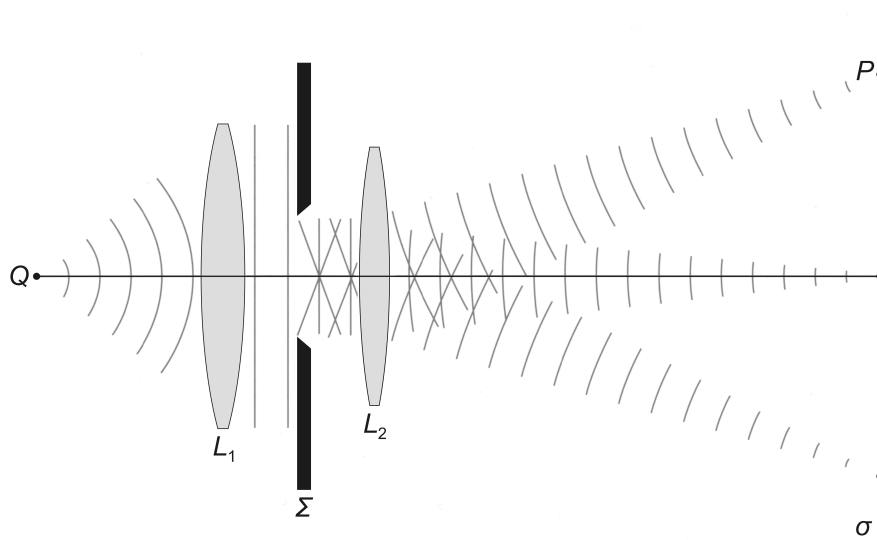


Abb. 14 Fraunhoferbeugung. Die Punktquelle Q liegt im vorderen Brennpunkt der Sammellinse L_1 , die deshalb die von Q ausgehenden Kugelwellen in ebene Wellen transformiert (Kollimator). Die Abbildungsebene σ liegt in der hinteren Brennebene der Sammellinse L_2 , sodass die einfallenden ebenen Wellenzüge 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)

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

$$\text{Fraunhoferbeugung} \quad \Leftrightarrow \quad R > \frac{a^2}{\lambda}.$$

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 14 wird dargestellt, wie sich die Fraunhoferbeugung mit Hilfe von Sammellinsen platzsparend realisieren lässt.

11.2 Fraunhoferbeugung

11.2.1 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 15 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).

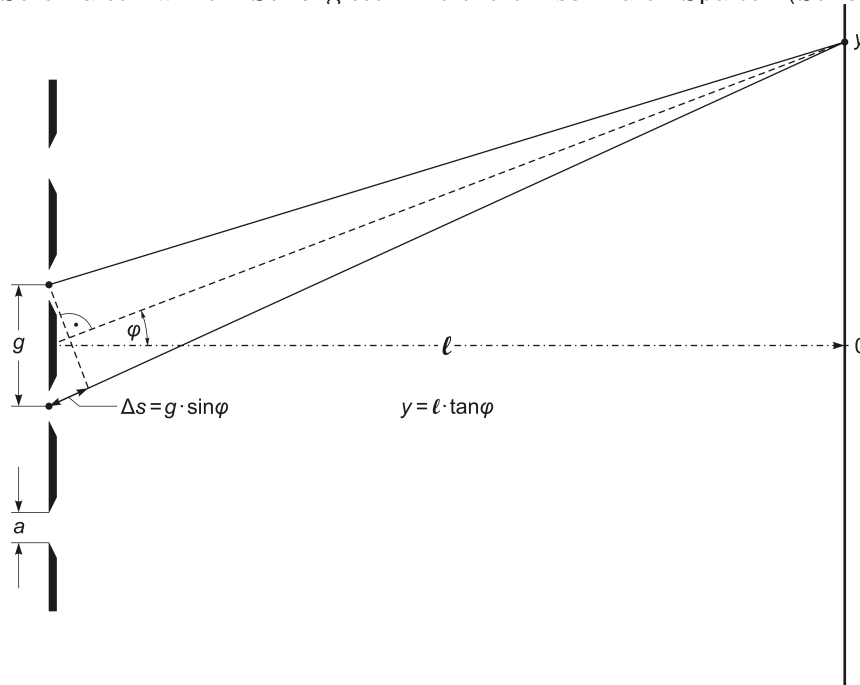


Abb. 15 Strichgitter und Schirm in möglichst großem Abstand ℓ zur Darstellung des Interferenzmusters. Gitterkonstante g , Strichzahl N , Spaltbreite a , Spalt-Schirm-Abstand ℓ mit $\ell \gg g$, Winkel φ zwischen der Wellenfrontnormalen und der Spaltnormalen, Gangunterschied Δs .

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:

$$\frac{\delta}{2\pi} = \frac{\Delta s}{\lambda} = \frac{g \cdot \sin \varphi}{\lambda} \Leftrightarrow \text{Phasendifferenz } \delta = \frac{2\pi}{\lambda} \cdot g \cdot \sin \varphi = k \cdot \Delta s. \quad (72)$$

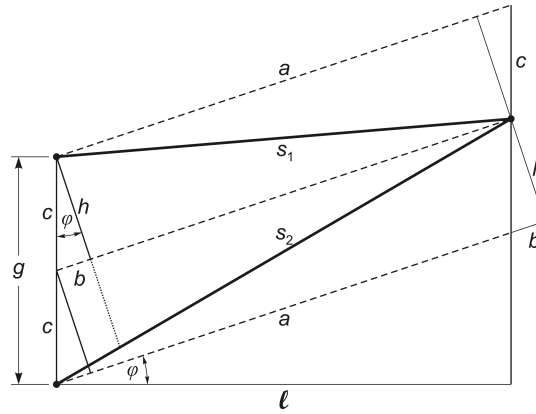


Abb. 16 Der korrekte Gangunterschied $\Delta s = s_2 - s_1$.

Anhand der Abbildung 16 zeigen wir, dass der in Abbildung 15 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:

$$c = \frac{g}{2}, \quad a = \frac{\ell}{\cos \varphi}, \quad b = c \sin \varphi, \quad h = c \cos \varphi,$$

$$\begin{aligned} \Delta s &= s_2 - s_1 = \sqrt{(a+b)^2 + h^2} - \sqrt{(a-b)^2 + h^2} \\ &= \sqrt{\left(\frac{\ell}{\cos \varphi} + c \sin \varphi\right)^2 + c^2 \cos^2 \varphi} - \sqrt{\left(\frac{\ell}{\cos \varphi} - c \sin \varphi\right)^2 + c^2 \cos^2 \varphi} \Rightarrow \\ \lim_{\varphi \rightarrow 0} \Delta s &= \sqrt{\ell^2 + c^2} - \sqrt{\ell^2 + c^2} = 0 = g \sin(0), \quad \square \\ \lim_{\varphi \rightarrow \frac{\pi}{2}} \Delta s &= \lim_{\varphi \rightarrow \frac{\pi}{2}} \left[\sqrt{\left(\frac{\ell}{\cos \varphi} + c\right)^2} - \sqrt{\left(\frac{\ell}{\cos \varphi} - c\right)^2} \right] \\ &= \lim_{\varphi \rightarrow \frac{\pi}{2}} \left[\left(\frac{\ell}{\cos \varphi} + c\right) - \left(\frac{\ell}{\cos \varphi} - c\right) \right] = 2c = g = g \sin\left(\frac{\pi}{2}\right). \quad \square \end{aligned}$$

Hauptmaxima bei konstruktiver Interferenz:

$$\begin{aligned} \Delta s &= g \cdot \sin \varphi \stackrel{!}{=} m \cdot \lambda, \\ \delta &= \frac{2\pi g}{\lambda} \sin \varphi \stackrel{!}{=} m \cdot 2\pi, \quad m = 0, 1, 2, 3, \dots \\ \Rightarrow \sin \varphi &= m \cdot \frac{\lambda}{g}, \quad \sin \varphi \sim \lambda, \quad \sin \varphi \sim \frac{1}{g}. \end{aligned} \tag{73}$$

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:

$$\begin{aligned} \Delta s &= g \cdot \sin \varphi \stackrel{!}{=} \left(m + \frac{1}{2}\right) \lambda, \\ \delta &= \frac{2\pi g}{\lambda} \sin \varphi \stackrel{!}{=} (2m + 1) \pi, \quad m = 0, 1, 2, 3, \dots \end{aligned}$$

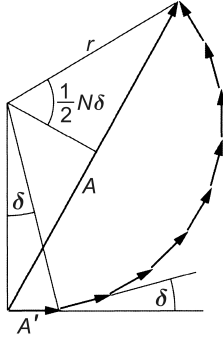


Abb. 17 Überlagerung vieler gleichstarker Wellen mit der Phasenverschiebung δ zwischen je zweien, z. B. der Teilwellen aus den Spalten eines Beugungsgitters. Darstellung dieser Überlagerung durch ein Zeigerdiagramm.

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:

$$\boxed{A_\varphi = A' \frac{\sin \left(\frac{N\pi g}{\lambda} \sin \varphi \right)}{\sin \left(\frac{\pi g}{\lambda} \sin \varphi \right)}}. \quad (74)$$

- Je kleiner g ist, desto breiter wird das Spektrum gespreizt, denn

$$\Delta s = g \cdot \sin \varphi \Rightarrow \sin \varphi = \Delta s \cdot \frac{1}{g} \Rightarrow \sin \varphi \propto \frac{1}{g}. \quad (75)$$

- 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 (74). 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 (72)

$$\sin \frac{\delta}{2} = \sin \left(\frac{\pi g}{\lambda} \sin \varphi \right) = \frac{\pi g}{\lambda} \sin \varphi. \quad (76)$$

Berücksichtigen wir (76) in der vorläufigen Amplitudenformel (74) 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**:

$$\boxed{A_\varphi = A_0 \frac{\sin \left(\frac{\pi a}{\lambda} \sin \varphi \right)}{\frac{\pi a}{\lambda} \sin \varphi}}. \quad (77)$$

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 *del'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 (75) 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

$$\Delta\Phi = \frac{2\pi a}{\lambda} \sin \varphi = \frac{2\pi \sin \varphi}{\lambda} a \quad (78)$$

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 (74) für die Amplitude A' des einzelnen Gitterspalts die Amplitudenformel (77) 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**

$$A_\varphi = A_0 \frac{\sin\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\frac{\pi a}{\lambda} \sin \varphi} \cdot \frac{\sin\left(\frac{N\pi g}{\lambda} \sin \varphi\right)}{\sin\left(\frac{\pi g}{\lambda} \sin \varphi\right)}. \quad (79)$$

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 (79) liefert folglich die Intensitätsformel für die Beugung und Interferenz am **Strichgitter**:

$$I_\varphi = I_0 \frac{\sin^2\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\left(\frac{\pi a}{\lambda} \sin \varphi\right)^2} \cdot \frac{\sin^2\left(\frac{N\pi g}{\lambda} \sin \varphi\right)}{\sin^2\left(\frac{\pi g}{\lambda} \sin \varphi\right)}. \quad (80)$$

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**:

$$I_\varphi = I_0 \frac{\sin^2\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\left(\frac{\pi a}{\lambda} \sin \varphi\right)^2}. \quad (81)$$

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 18).

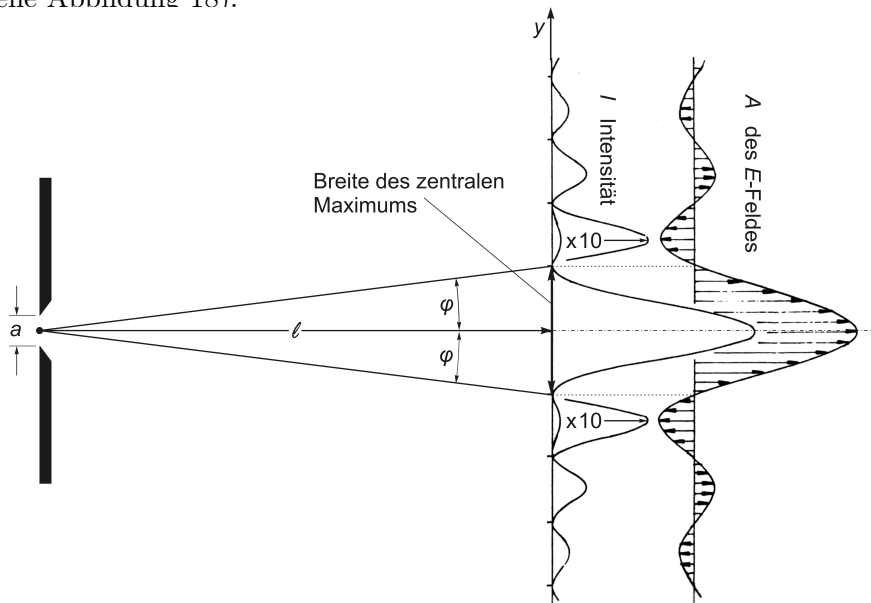


Abb. 18 Die Breite des zentralen Maximums bei der Beugung am Einfachspalt. Weil die Intensität der Nebenmaxima sehr stark abfällt, sind diese zur Verdeutlichung um das 10-fache überhöht. Der Vollständigkeit halber wird der Amplitudenverlauf des E -Feldes längs der y -Achse der Intensität gegenübergestellt. (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)

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 (78) 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).² 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. 18)

$$2 \cdot \ell \cdot \tan \left(\arcsin \frac{\lambda}{a} \right) .$$

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 $\sin \varphi_m = m \frac{\lambda}{a}$, $m = 1, 2, 3, \dots$.

²Der erste Nulldurchgang der Amplitude zu beiden Seiten des zentralen Maximums bildet die ersten Intensitätsminima und bestimmt die Breite des zentralen Maximums.

Minima und Maxima beim Strichgitter:

$$\begin{array}{ll} \text{Maxima} & \text{bei } \sin \varphi_m = m \frac{\lambda}{g}, \quad m = 0, 1, 2, 3, \dots \\ \text{Minima} & \text{bei } \sin \varphi_m = m \frac{\lambda}{N \cdot g}, \quad m = 1, 2, 3, \dots \end{array}$$

f) Spektrales Auflösungsvermögen des Strichgitters

Die Herleitung erfolgt nach dem Lehrbuch von Gerthsen, Seite 522, und auf der Grundlage von Formel (80) sowie speziell auf der Grundlage der vorläufigen Amplitudenformel (74)

$$A_\varphi = A' \cdot \frac{\sin \left(\frac{N\pi g}{\lambda} \sin \varphi \right)}{\sin \left(\frac{\pi g}{\lambda} \sin \varphi \right)},$$

denn insbesondere das breite zentrale Maximum von A' überlagert den Quotienten in (74) nur und ist für das Auflösungsvermögen des Strichgitters nicht relevant.

Der Quotient auf der rechten Seite von (74) 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:

$$\delta = \frac{2\pi g}{\lambda} \cdot \sin \varphi_m \stackrel{!}{=} m \cdot 2\pi \quad \Leftrightarrow \quad \sin \varphi_m = \frac{m\lambda}{g}, \quad m = 0, 1, 2, 3, \dots$$

$$\text{Nenner: } \sin \left(\frac{\pi g}{\lambda} \sin \varphi_m \right) = \sin(m\pi) = 0$$

$$\text{Zähler: } \sin(N \cdot m\pi) = 0$$

$$\text{Maximum m-ter Ordnung: } \varphi_m \approx \frac{m\lambda}{g}. \quad (82)$$

- Die Minima liegen dort, wo nur der Zähler verschwindet.

$$\text{Zähler: } \sin \left(\frac{\pi N g}{\lambda} \sin \varphi \right) \stackrel{!}{=} 0 \quad \text{bei } \sin \varphi \neq 0 \text{ bzw. } \varphi \neq 0, \quad \text{d. h. wenn}$$

$$N \cdot \Delta s = N \cdot g \cdot \sin \varphi_m = \lambda \quad \Rightarrow \quad \sin \varphi_m = m \cdot \frac{\lambda}{N \cdot g}, \quad m = 1, 2, 3, \dots$$

Beim ersten Minimum bilden die Vektoren der N Gitterquellen einen einfachen geschlossenen Kreis und löschen sich somit aus. Es gilt dann

$$N \cdot \delta = N \cdot \frac{2\pi g}{\lambda} \sin \varphi = 2\pi.$$

$$\Delta(\sin \varphi) = (m+1) \frac{\lambda}{N \cdot g} - m \frac{\lambda}{N \cdot g} = \frac{\lambda}{N \cdot g} \text{ liefert somit den } \mathbf{Abstandswinkel}$$

$$\Delta \varphi \approx \frac{\lambda}{N \cdot g}$$

des ersten Minimums von einem Hauptmaximum.

- Die Maxima von Licht mit der Wellenlänge $\lambda + \Delta\lambda$ besitzen gemäß (82) 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.

11.3 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.

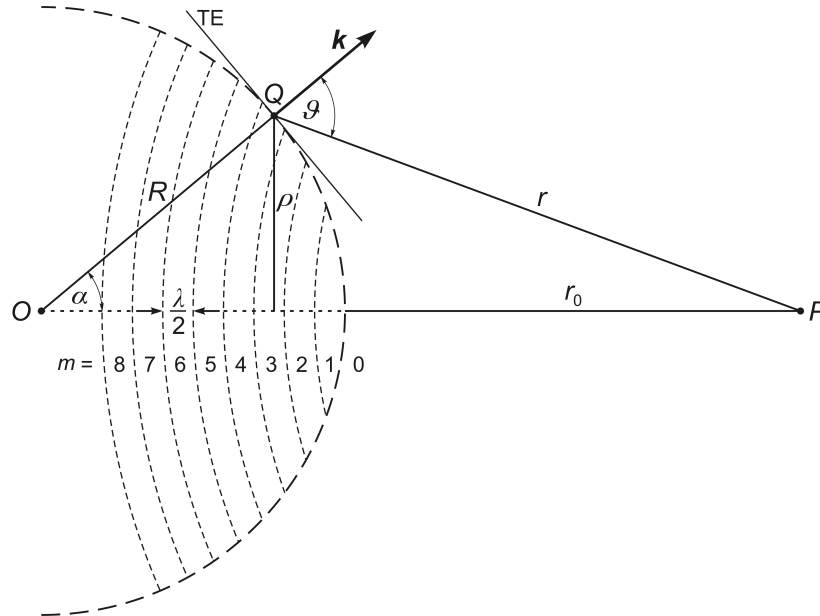


Abb. 19 Fresnel'sche Zonen. Dargestellt sind nur die ersten 8 Zonen. TE steht für Tangentialebene an den Punkt Q der dargestellten Kugelwellenfront.

- Die von der Punktquelle O ausgehende und sich frei im Raum ausbreitende Kugelwelle, die **Primärwelle**, sei

$$E = E_R \cdot e^{i(kR - \omega t)} = \frac{A}{R} \cdot e^{i(kR - \omega t)} . \quad (83)$$

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 19).
- 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

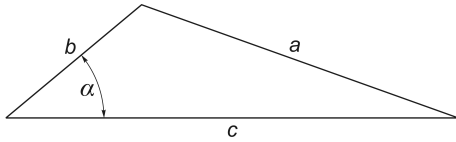


Abb. 20 Kosinussatz: $a^2 = b^2 + c^2 - 2bc \cos \alpha$.

auf das Dreieck OQP liefert

$$\begin{aligned} r^2 &= R^2 + (R + r_0)^2 - 2R(R + r_0) \cos \alpha \\ \Rightarrow \quad \frac{d}{d\alpha} r^2 &= 2r \frac{dr}{d\alpha} = 2R(R + r_0) \sin \alpha \quad \Leftrightarrow \\ R \sin \alpha \, d\alpha &= \frac{1}{R + r_0} r \, dr . \end{aligned} \quad (84)$$

- 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 (83) 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 :

$$E' \propto E_R \quad \Rightarrow \quad E' = q \cdot E_R . \quad (85)$$

- Weiterhin ist der Feldstärkebeitrag durch die Sekundärwellen in P abhängig vom Winkel ϑ (siehe Abbildung 19), d. h. abhängig von der „Gerichtetheit“ der sekundären Quellen auf der primären Kugelwellenfront. Diese „Gerichtetheit“ wird in diesem speziellen Zusammenhang beschrieben durch den³

³Auf 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 12.4 gesondert ein.

$$\text{Neigungsfaktor } K = K(\vartheta) = \frac{1}{2} (1 + \cos \vartheta) .$$

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 19)

$$K(\vartheta) = \frac{1}{2} (1 + \cos \vartheta) \propto E(P) . \quad (86)$$

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 (85) und (86) 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

$$dE(P) = C \cdot K \cdot \frac{E_R}{r} e^{i[k(R+r) - \omega t]} dS . \quad (87)$$

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

$$dS = 2\pi R \sin \alpha \cdot R d\alpha = 2\pi R \cdot R \sin \alpha d\alpha .$$

Aus dem Vergleich mit (84) resultiert

$$dS = \frac{2\pi R}{R + r_0} r dr .$$

Somit ist dS nur noch von r abhängig.

- Durch Integration über (87) 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

$$\begin{aligned} k &= \frac{2\pi}{\lambda} , \\ r_m &= r_0 + m \frac{\lambda}{2} , \\ r_{m-1} &= r_0 + (m-1) \frac{\lambda}{2} = r_0 + m \frac{\lambda}{2} - \frac{\lambda}{2} , \\ e^{i\pi} &= e^{-i\pi} = -1 \end{aligned}$$

ist

$$\begin{aligned} \left[e^{ikr} \right]_{r_{m-1}}^{r_m} &= e^{ikr_m} - e^{ikr_{m-1}} = e^{ik(r_0 + m \frac{\lambda}{2})} - e^{ik(r_0 + m \frac{\lambda}{2} - \frac{\lambda}{2})} \\ &= e^{ikr_0} \cdot e^{ik \frac{\lambda}{2} m} - e^{ikr_0} \cdot e^{ik \frac{\lambda}{2} m} \cdot e^{-ik \frac{\lambda}{2}} = e^{ikr_0} \cdot e^{ik \frac{\lambda}{2} m} \cdot (1 - e^{-ik \frac{\lambda}{2}}) \\ &= e^{ikr_0} \cdot (e^{i\pi})^m \cdot (1 - e^{-i\pi}) , \\ \left[e^{ikr} \right]_{r_{m-1}}^{r_m} &= e^{ikr_0} \cdot (-1)^m \cdot 2 . \end{aligned}$$

$$\begin{aligned} E_m &= C \cdot K_m E_R \frac{2\pi R}{R + r_0} e^{i(kR - \omega t)} \cdot \frac{1}{ik} \cdot e^{ikr_0} \cdot (-1)^m \cdot 2 \\ &= -i(-1)^m \cdot C \cdot K_m E_R \frac{2\pi R}{R + r_0} e^{i[k(R+r_0) - \omega t]} \cdot \frac{\lambda}{\pi} , \\ E_m &= C \cdot (-1)^{m+1} \frac{i \cdot 2\lambda \cdot K_m E_R \cdot R}{R + r_0} e^{i[k(R+r_0) - \omega t]} , \end{aligned}$$

$$\boxed{E_m(P) = C \cdot (-1)^{m+1} \frac{2 \cdot i \lambda \cdot K_m \cdot E_R \cdot R}{R + r_0} e^{i[k(R+r_0) - \omega t]}} . \quad (88)$$

- Die Gesamtfeldstärke im Punkt P ist folglich

$$\begin{aligned} E(P) &= \sum_{m=1}^N E_m = |E_1| - |E_2| + |E_3| - |E_4| + |E_5| - \dots \pm |E_N| \\ &= \left(\frac{1}{2}|E_1| + \frac{1}{2}|E_1| \right) - |E_2| + \left(\frac{1}{2}|E_3| + \frac{1}{2}|E_3| \right) - |E_4| + \left(\frac{1}{2}|E_5| + \frac{1}{2}|E_5| \right) - \dots \pm |E_N| . \quad (89) \end{aligned}$$

Dabei haben wir den Vorzeichenwechsel von E_m beim Zonenwechsel berücksichtigt.

- Mit K_m verändern sich auch die $|E_m|$ nur sehr langsam, sodass

$$|E_m| \approx \frac{1}{2}(|E_{m-1}| + |E_{m+1}|) \quad \Leftrightarrow \quad \frac{1}{2}|E_{m-1}| - |E_m| + \frac{1}{2}|E_{m+1}| \approx 0 .$$

Dem entsprechend können wir (89) umordnen und erhalten

$$E(P) = \frac{1}{2}|E_1| + \underbrace{\left(\frac{1}{2}|E_1| - |E_2| + \frac{1}{2}|E_3|\right)}_{\approx 0} + \underbrace{\left(\frac{1}{2}|E_3| - |E_4| + \frac{1}{2}|E_5|\right)}_{\approx 0} + \dots + \frac{1}{2}|E_N|. \quad (90)$$

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 (88) für E_1 und mit $E_R = A/R$

$$E(P) \approx \frac{1}{2} E_1 = \frac{1}{2} \cdot \frac{C \cdot 2 \cdot i \lambda \cdot K_1 \cdot E_R \cdot R}{R + r_0} e^{i[k(R+r_0) - \omega t]},$$

$$\boxed{E(P) \approx C \cdot K_1 \cdot i \lambda \cdot \frac{A}{R + r_0} e^{i[k(R+r_0) - \omega t]}}. \quad (91)$$

- 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 (91) 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 Primärwelle⁴ (83) erzeugt wird:

$$C \cdot K_1 \cdot i \lambda \cdot \frac{A}{R + r_0} = C \cdot i \lambda \cdot \frac{A}{R + r_0} = \frac{A}{R + r_0} \Rightarrow$$

$$\boxed{C = \frac{1}{i \lambda}}.$$

- 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

$$\begin{aligned} r^2 &= \varrho^2 + z^2 = \varrho^2 + c^2 + R^2 - \varrho^2 - 2c\sqrt{R^2 - \varrho^2}, \\ r^2 &= c^2 + R^2 - 2c\sqrt{R^2 - \varrho^2} \Leftrightarrow \\ \sqrt{R^2 - \varrho^2} &= \frac{c^2 + R^2 - r^2}{2c}. \end{aligned}$$

⁴Der in (91) 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.

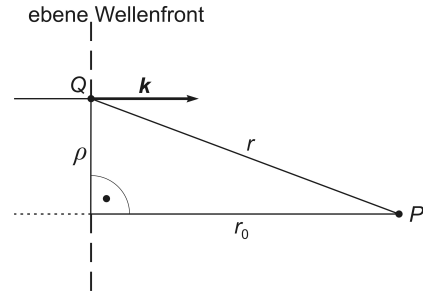
Quadrieren und Auflösen nach ϱ liefert schließlich

$$\varrho = \frac{\sqrt{4c^2R^2 - (c^2 + R^2 - r^2)^2}}{2c}.$$

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 19 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 21):

Abb. 21 Zur Berechnung von ϱ bzw. ϱ_m für $R \rightarrow \infty$.



bzw.

$$\varrho^2 = r^2 - r_0^2 \Rightarrow \varrho = \sqrt{r^2 - r_0^2}$$

$$\varrho_m^2 = \left(r_0 + m \frac{\lambda}{2}\right)^2 - r_0^2 = \frac{m^2}{4} \lambda^2 + m r_0 \lambda \Rightarrow$$

$$\boxed{\varrho_m = \sqrt{\frac{m^2}{4} \lambda^2 + m r_0 \lambda} \quad \text{für } R \rightarrow \infty}.$$

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 19):

$$\varrho_m \approx \sqrt{r_m^2 - r_0^2}.$$

Wir nähern weiter mittels Taylor-Entwicklung des Radikanden bis zur 1. Ordnung an der Stelle $\lambda = 0$:

$$\begin{aligned} r_m^2 - r_0^2 &= \left(r_0 + m \frac{\lambda}{2}\right)^2 - r_0^2 = r_0^2 + \left(\frac{m}{2} \lambda\right)^2 + r_0 m \lambda - r_0^2 \\ &= \frac{m^2}{4} \lambda^2 + r_0 m \lambda \Rightarrow \end{aligned}$$

$$r_m^2 - r_0^2 \approx \left[\frac{m^2}{2} \lambda + r_0 m \right]_{\lambda=0} \cdot \lambda = m r_0 \lambda \Rightarrow$$

$$\boxed{\varrho_m \approx \sqrt{m \cdot r_0 \cdot \lambda} \quad \text{für } r_0 \gg \lambda} \quad (92)$$

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, nämlich⁵

$$D = 2 \cdot \varrho_1 \approx 2 \cdot \sqrt{r_1^2 - r_0^2} = 2 \cdot \sqrt{\left(r_0 + 1 \cdot \frac{\lambda}{2}\right)^2 - r_0^2} .$$

Somit erreicht nur das Licht aus der 1. Fresnel'schen Zone den Punkt P . Gemäß (88) 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 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 (90) 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 (88) 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 (Kreisringe) 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).

⁵Beispielsweise 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

$$\varrho \approx \sqrt{r^2 - r_0^2} \Rightarrow \varrho_m \approx \sqrt{\left(r_0 + m \cdot \frac{\lambda}{2}\right)^2 - r_0^2} \Rightarrow \varrho_1 \approx \sqrt{\left(r_0 + \frac{\lambda}{2}\right)^2 - r_0^2} \approx 0,32 \text{ mm} .$$

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.

12 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 11.3 *Fresnel'sche Zonen*.

12.1 Vorbereitung – Helmholtz-Gleichung

Mit dem

$$\text{Potentialansatz} \quad \vec{E} = -\text{grad } \Phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \text{rot } \vec{A} \quad (93)$$

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

$$\Delta \Phi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \Phi = 0, \quad (94)$$

$$\Delta \vec{A} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} = \vec{0}. \quad (95)$$

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. Jellito¹.

¹Rainer 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 (94) 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 (94) für Φ die Kugelwellenfunktion

$$E(r, t) = \frac{1}{r} \cdot \sin(kr - \omega t) \quad (96)$$

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 (95) 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 (94) und (95) zurück zum \vec{E} -Feld kommen wir, indem wir zunächst den Potentialansatz (93) in die homogene (quellenfreie) Maxwell'sche Gleichung

$$\text{rot } \vec{E} + \frac{\partial}{\partial t} \vec{B} = \vec{0}$$

einsetzen:

$$\text{rot } \vec{E} + \frac{\partial}{\partial t} \text{rot } \vec{A} = \text{rot } \vec{E} + \text{rot } \frac{\partial}{\partial t} \vec{A} = \vec{0} \quad \Leftrightarrow \quad \vec{E} = -\frac{\partial \vec{A}}{\partial t} . \quad (97)$$

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 (97)

$$\vec{E}(\vec{r}, t) = \underbrace{-i\omega \vec{A}_0}_{=\vec{E}_0} e^{i(\vec{k}\vec{r} - \omega t)} = \vec{E}_0 e^{i(\vec{k}\vec{r} - \omega t)} ,$$

ebenfalls eine Wellenfunktion, wie wir oben bereits gezeigt hatten.

In der Optik sind vor allem das \vec{E} -Feld, d. h. das **optische Feld** und folglich die skalare homogene Wellengleichung (94) von Bedeutung, weil bei Abwesenheit zeitlich-veränderlicher Magnetfelder gemäß des Potentialansatzes (93) allein

$$\vec{E} = -\text{grad } \Phi \quad \longleftrightarrow \quad \Phi = - \int \vec{E} \cdot d\vec{r}$$

gilt. Setzen wir dies in (94) ein, resultiert

$$\underbrace{\Delta \left(- \int \vec{E} d\vec{r} \right)}_{=\Phi} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \underbrace{\left(- \int \vec{E} d\vec{r} \right)}_{=\Phi} \Rightarrow - \int \Delta \vec{E} d\vec{r} = - \int \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E} d\vec{r} \Rightarrow$$

$$\Delta \vec{E}(\vec{r}, t) = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E}(\vec{r}, t) \quad \text{bzw.} \quad \begin{pmatrix} \Delta E_x \\ \Delta E_y \\ \Delta E_z \end{pmatrix} = \begin{pmatrix} \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_x \\ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_y \\ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_z \end{pmatrix}.$$

Wie das skalare Potential Φ so erfüllt also auch das aus Φ hervorgehende elektrische Feld \vec{E} die Wellengleichung (94).

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 (94). 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 Maxwellschen 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 Maxwellschen 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.“²

Verwenden wir also an Stelle des Potentials Φ in der skalaren Wellengleichung (94) als Lösungsfunktion die Funktion des \vec{E} -Feldes einer monochromatischen elektromagnetischen Welle, z. B. die Kugelwellenfunktion (96)

$$E(r, t) = \frac{1}{r} e^{i(kr - \omega t)} = \frac{1}{r} e^{ikr} \cdot e^{-i\omega t}.$$

Dann können wir (94) mit dem Produktansatz nach den Koordinaten r und t in zwei gewöhnliche Differentialgleichungen separieren, die voneinander unabhängig und jeweils nur

²Zitiert aus: Rainer J. Jellito, Studententext – Elektrodynamik, Theoretische Physik, 3. Auflage, Aula-Verlag, Wiesbaden, 1994, Seite 131.

von einer Variablen abhängig sind:

$$\textbf{Produktansatz : } E(r, t) = \mathcal{E}(r) \cdot T(t) \Rightarrow \begin{cases} \text{Ortsanteil } \mathcal{E}(r) = \frac{1}{r} e^{ikr} \\ \text{Zeitanteil } T(t) = e^{-i\omega t} \end{cases} .$$

Setzen wir also den Produktansatz in die Wellengleichung (94) 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

$$\Delta_r \mathcal{E}(r) - \kappa \cdot \mathcal{E}(r) = 0 . \quad (98)$$

Mit der Lösungsfunktion $\mathcal{E}(r) = \frac{1}{r} e^{ikr}$ in (98) bestimmen wir die Separationskonstante κ bezüglich unserer Kugelwellenfunktion. Wir erhalten als Ergebnis

$$-\frac{1}{r} \cdot k^2 \cdot e^{ikr} - \kappa \cdot \frac{1}{r} \cdot e^{ikr} = 0 \Leftrightarrow \kappa = -k^2$$

und damit schließlich die

Helmholtz-Gleichung $\Delta \mathcal{E} + k^2 \mathcal{E} = 0$

 .

12.2 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

$$E(\vec{r}, t) = E_0(\vec{r}) \cdot e^{i(kr - \omega t)} \quad \text{bzw. im Fall einer Kugelwelle} \quad E(r, t) = E_0(r) \cdot e^{i(kr - \omega t)} .$$

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

$$\begin{aligned} \oint_S (\mathcal{E} \text{grad } \Psi - \Psi \text{grad } \mathcal{E}) \cdot d\vec{S} &= \int_V (\mathcal{E} \Delta \Psi - \Psi \Delta \mathcal{E}) dV \\ &= \int_V [\mathcal{E}(-k^2 \Psi) - \Psi(-k^2 \mathcal{E})] dV \\ &= \int_V k^2 \underbrace{(\Psi \mathcal{E} - \mathcal{E} \Psi)}_{=0} dV = 0 . \end{aligned} \tag{99}$$

Als Funktion Ψ wählen wir den Ortsanteil der vom Punkt P des Integrationsgebiets ausgehenden Kugelwelle (siehe Abbildung 22)

$$\Psi = \frac{e^{ikr}}{r} .$$

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 (99) nur dann mit Sicherheit

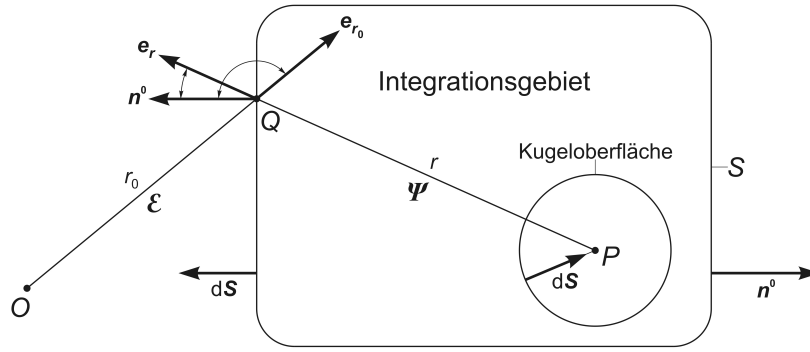


Abb. 22 Zur Herleitung der Kirchhoff'schen Formel. Das Integrationsgebiet, das von der geschlossenen Fläche S umhüllt wird, ist von beliebiger Gestalt! Weil wir hier die **freie Wellenausbreitung** ohne irgendwelche Hindernisse betrachten, ist die Hülle S allein mathematischer Natur, also nur gedacht 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

$$\vec{PQ} = \vec{r} \text{ und } \vec{PO} = \vec{r}' \Rightarrow r_0 = |\vec{r} - \vec{r}'|.$$

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 (99) 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 (99) 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:

$$\begin{aligned} \oint_S \dots dS + \oint_{\text{Kugel}} \dots dS &= 0 \Leftrightarrow \oint_S \dots dS = - \oint_{\text{Kugel}} \dots dS \Rightarrow \\ \lim_{\text{Kugelradius} \rightarrow 0} \underbrace{\oint_S \dots dS}_{= \oint_S \dots dS \text{ einschließlich } P} &= \lim_{\text{Kugelradius} \rightarrow 0} - \oint_{\text{Kugel}} \dots dS. \end{aligned} \quad (100)$$

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:³

$$\oint_S \left(\mathcal{E} \text{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \text{grad} \mathcal{E} \right) \cdot d\vec{S} + \oint_{\text{Kugel}} \left(\mathcal{E} \text{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \text{grad} \mathcal{E} \right) \cdot d\vec{S} = 0. \quad (101)$$

³Um 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.

Berechnung des Grenzwertes des Integrals über die kleine Kugel für $r \rightarrow 0$:

$$\oint_{\text{Kugel}} \left(\mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} \right) \cdot d\vec{S} \quad (102)$$

- 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 Kugelflächelement

$$d\vec{S} = -\vec{e}_r \cdot dS = -\vec{e}_r \cdot r^2 \sin \vartheta \, d\vartheta \, d\varphi .$$

- 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 (102) für $r \rightarrow 0$ gemäß

$$\oint_{\text{Kugel}} - \underbrace{(\vec{e}_r \cdot \operatorname{grad} \mathcal{E}) \cdot e^{ikr} \cdot r \cdot \sin \vartheta \, d\vartheta \, d\varphi}_{=0 \text{ für } \lim_{r \rightarrow 0}}$$

- Mit

$$\operatorname{grad} \Psi = \operatorname{grad} \frac{e^{ikr}}{r} = \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r$$

verbleibt von (102) nur das Integral

$$\begin{aligned} \oint_{\text{Kugel}} \mathcal{E} \cdot \operatorname{grad} \frac{e^{ikr}}{r} \cdot d\vec{S} &= \oint_{\text{Kugel}} \mathcal{E} \cdot \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r \cdot (-\vec{e}_r) \cdot r^2 \sin \vartheta \, d\vartheta \, d\varphi \\ &= \oint_{\text{Kugel}} \mathcal{E} \cdot (1 - ikr) e^{ikr} \cdot \sin \vartheta \, d\vartheta \, d\varphi \end{aligned}$$

mit dem Grenzwert

$$\lim_{\substack{r \rightarrow 0 \\ \mathcal{E} \rightarrow \mathcal{E}_p}} \oint_{\text{Kugel}} \mathcal{E} \cdot (1 - ikr) e^{ikr} \cdot \sin \vartheta \, d\vartheta \, d\varphi = \mathcal{E}_p \cdot \oint_{\text{Kugel}} \sin \vartheta \, d\vartheta \, d\varphi = \mathcal{E}_p \cdot 4\pi .$$

Das Hüllenintegral (101) über die geschlossene Oberfläche S des gesamten Integrationsgebiets *einschließlich der Singularität im Punkt P* ist mit (100)

$$\oint_S \left(\mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} \right) \cdot d\vec{S} = -\mathcal{E}_p \cdot 4\pi . \quad (103)$$

Durch Äquivalenzumformung (Achtung Vorzeichen!) folgt daraus schließlich die

Kirchhoff'sche Formel $\mathcal{E}_p = \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} \right) \cdot d\vec{S}$

(104)

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.

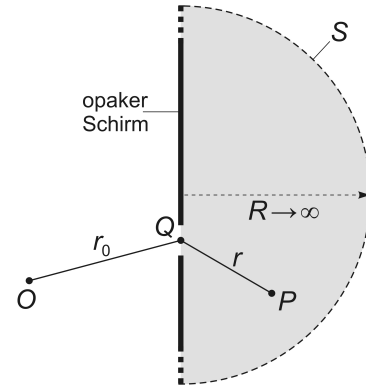
12.3 Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals

Zur Herleitung des Fresnel-Kirchhoff'schen Beugungsintegrals wenden wir die im Abschnitt 12.2 hergeleitete Kirchhoff'scher Formel (104) 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 23).

Abb. 23 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 22 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 12.5 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 12.5 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 22 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 (104) 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 22 zeigt, dass $\angle(\vec{n}^0, \vec{e}_{r_0})$ ein stumpfer Winkel ist und $\angle(\vec{n}^0, \vec{e}_r)$ 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:

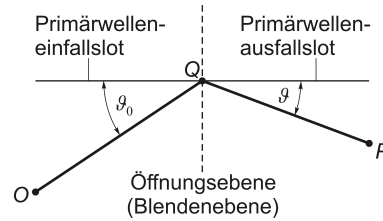
$$\mathcal{E}_p = \frac{-i\epsilon_0}{\lambda} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\frac{\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0})}{2} \right] dS. \quad (105)$$

Gemäß Abbildung 22 und Abbildung 24 ist der Term in eckigen Klammern der im folgenden Abschnitt 12.4 diskutierte **Neigungsfaktor** K :

$$\cos(\vec{n}^0, \vec{e}_r) = \cos \vartheta \quad \text{und} \quad -\cos(\vec{n}^0, \vec{e}_{r_0}) = \cos \vartheta_0 \quad \Rightarrow$$

$$K = \frac{1}{2} \left[\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0}) \right] = \frac{1}{2} (\cos \vartheta + \cos \vartheta_0).$$

Abb. 24 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 (105) sowohl r und r_0 im Nenner als auch K als nahezu konstant über die Öffnung betrachten und vor das Integral ziehen:

$$\mathcal{E}_p(r) = \frac{-i\epsilon_0}{\lambda} \frac{K}{r \cdot r_0} \int_{\text{Öffnung}} e^{ik(r+r_0)} \cdot dS. \quad (106)$$

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 (105) 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.

12.4 Der Neigungsfaktor K

Allgemein ist der **Neigungsfaktor** $K = \frac{1}{2}(\cos \vartheta_0 + \cos \vartheta)$ (siehe Abbildung 24).

Wir betrachten zunächst den speziellen Fall der freien Ausbreitung einer Kugelwelle:

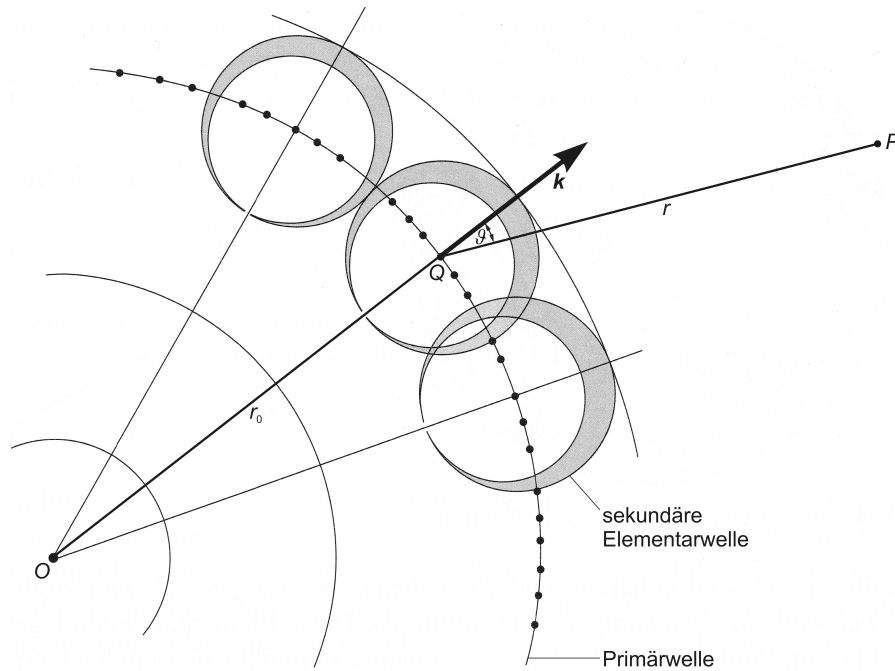


Abb. 25 Die **freie Ausbreitung einer Kugelwelle** und das Abstrahlungsverhalten der sekundären Elementarwellen. \vec{k} ist der Wellenvektor der betrachteten primären Kugelwellenfront.

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.

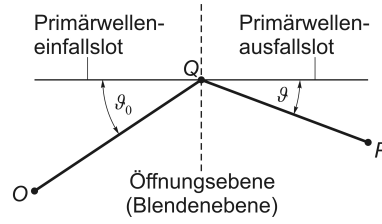
(Abbildung nach: Eugene Hecht, Optik, 4. Auflage, Oldenbourg Verlag München Wien, 2005, Seite 781)

Bei **freier Wellenausbreitung** wie in Abbildung 25 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 23 und 26.

Abb. 26 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 26 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.

12.5 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 25 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 25 gilt auch für die im Unendlichen liegende Halbsphäre in Abbildung 23

$$\boxed{\lim_{R \rightarrow \infty} K = 0},$$

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 23 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 (104)

$$\mathcal{E}_p = \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \text{grad } \mathcal{E} - \mathcal{E} \text{grad } \frac{e^{ikr}}{r} \right) \cdot d\vec{S},$$

d. h. von der freien Wellenausbreitung gemäß der Abbildung 22 bzw. von der freien Kugelwelle gemäß der Abbildung 25 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 Kugelflächelement 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 Raumwinkelement, 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 Normalkomponenten der Gradienten von Ψ und \mathcal{E} sind

$$\begin{aligned} \text{grad } \Psi \cdot \vec{n}^0 &= \frac{\partial \Psi}{\partial n} = \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \cdot \vec{e}_r \cdot \vec{n}^0, \\ \text{grad } \mathcal{E} \cdot \vec{n}^0 &= \frac{\partial \mathcal{E}}{\partial n} = \left(\frac{ik}{r_0} - \frac{1}{r_0^2} \right) e^{ikr_0} \cdot \vec{e}_{r_0} \cdot \vec{n}^0. \end{aligned} \quad (107)$$

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 (107) in die Kirchhoff'sche

Formel (104) ein:

$$\begin{aligned}
\mathcal{E}_p &= \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} \operatorname{grad} \frac{e^{ikr}}{r} \right) \cdot d\vec{S} \\
&\approx \frac{1}{4\pi} \oint_S \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} ik \frac{e^{ikr}}{r} \vec{e}_r \right) \cdot d\vec{S} \\
&\approx \frac{1}{4\pi} \int_{\mathbb{R}^3} \left(\frac{e^{ikr}}{r} \operatorname{grad} \mathcal{E} - \mathcal{E} ik \frac{e^{ikr}}{r} \vec{e}_r \right) \cdot \vec{n}^0 \cdot r^2 d\Omega \\
&\approx \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{e^{ikr}}{r} \left(\frac{\partial \mathcal{E}}{\partial n} - ik \mathcal{E} \right) r^2 d\Omega, \\
\boxed{\mathcal{E}_p \approx \frac{1}{4\pi} \int_{\mathbb{R}^3} e^{ikr} \left(\frac{\partial \mathcal{E}}{\partial n} - ik \mathcal{E} \right) r d\Omega} &. \tag{108}
\end{aligned}$$

Das Feld auf der Oberfläche einer unendlich großen Kugel leistet im Beobachtungspunkt P keinen Beitrag, wenn das Integral (108) verschwindet bzw. wenn für dessen Integranden in isotroper Weise die

Sommerfeld'sche Strahlungsbedingung $\lim_{r \rightarrow \infty} \left(\frac{\partial \mathcal{E}}{\partial n} - ik \mathcal{E} \right) r = 0$

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 12.3 für die im Unendlichen liegende Halbsphäre (siehe Abbildung 23).

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

$$\Rightarrow \quad \vec{e}_r \uparrow \vec{n}^0 \wedge \vec{e}_{r_0} \uparrow \vec{n}^0 \quad \Rightarrow \quad \vec{e}_r \cdot \vec{n}^0 = \vec{e}_{r_0} \cdot \vec{n}^0 = 1.$$

Für den Grenzfall $r \rightarrow \infty$ dürfen wir also in der Kirchhoff'schen Formel (104)

$$r \rightarrow \infty \quad \Rightarrow \quad r_0 = r \quad \text{und} \quad \vec{e}_r \cdot \vec{n}^0 = \vec{e}_{r_0} \cdot \vec{n}^0 = 1 \tag{109}$$

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 (109) \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 .

12.6 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 13 noch einmal abhandeln.

Um es im Umgang mit dem Fresnel-Kirchhoff'schen Beugungsintegral in der Form (106) 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 27 dargestellt.

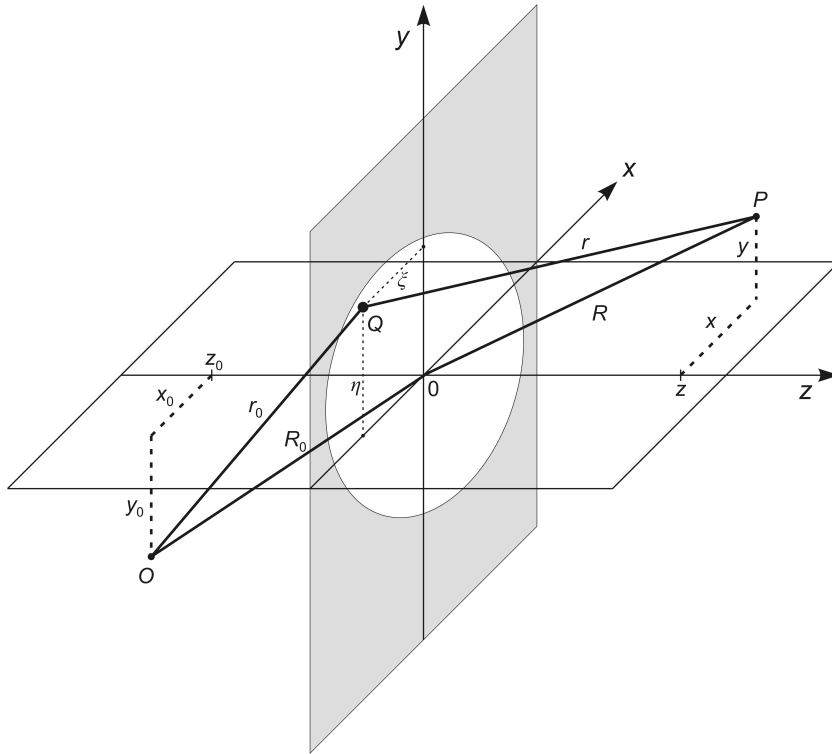


Abb. 27 Zur Beugung in einer ebenen Schirmöffnung beliebiger Gestalt. Nach Joos, Lehrbuch der Theoretischen Physik, Aula-Verlag, Wiesbaden, 1989, 15. Auflage, Seite 367.

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

$$r^2 = (x - \xi)^2 + (y - \eta)^2 + z^2, \tag{110}$$

$$r_0^2 = (x_0 - \xi)^2 + (y_0 - \eta)^2 + z_0^2.$$

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 (106) zunächst

$$\mathcal{E}_p = \frac{-i\epsilon_0}{\lambda} \frac{K}{R \cdot R_0} \int_{\text{Öffnung}} e^{ik(r+r_0)} \cdot dS \quad (111)$$

wird. In der Exponentialfunktion unter dem Integral hängen r und r_0 gemäß (110) 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

$$r(0, 0) = \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{\frac{1}{2}} \Big|_{(0,0)} = (x^2 + y^2 + z^2)^{\frac{1}{2}} = R.$$

1. Ordnung

$$\begin{aligned} & \frac{\partial}{\partial \xi} r(\xi, \eta) \Big|_{(0,0)} \cdot \xi + \frac{\partial}{\partial \eta} r(\xi, \eta) \Big|_{(0,0)} \cdot \eta \\ &= \left\{ \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{1}{2}} \cdot 2(x - \xi) \cdot (-1) \right\} \Big|_{(0,0)} \cdot \xi \\ &+ \left\{ \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{1}{2}} \cdot 2(y - \eta) \cdot (-1) \right\} \Big|_{(0,0)} \cdot \eta \\ &= \frac{-x\xi}{R} + \frac{-y\eta}{R} = -\frac{x\xi + y\eta}{R}. \end{aligned}$$

2. Ordnung

$$\begin{aligned} & \frac{1}{2} \left(\frac{\partial^2 r}{\partial \xi^2} \Big|_{(0,0)} \cdot \xi^2 + 2 \frac{\partial^2 r}{\partial \xi \partial \eta} \Big|_{(0,0)} \cdot \xi \eta + \frac{\partial^2 r}{\partial \eta^2} \Big|_{(0,0)} \cdot \eta^2 \right) \\ &= \frac{1}{2} \frac{\partial^2 r}{\partial \xi^2} \Big|_{(0,0)} \cdot \xi^2 + \frac{\partial^2 r}{\partial \xi \partial \eta} \Big|_{(0,0)} \cdot \xi \eta + \frac{1}{2} \frac{\partial^2 r}{\partial \eta^2} \Big|_{(0,0)} \cdot \eta^2 \\ &= \frac{1}{2} \left\{ \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{1}{2}} + (x - \xi) \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{3}{2}} \cdot 2(x - \xi) \cdot (-1) \right\} \Big|_{(0,0)} \cdot \xi^2 \\ &\quad + (x - \xi) \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{3}{2}} \cdot 2(y - \eta) \cdot (-1) \Big|_{(0,0)} \cdot \xi \eta \\ &\quad + \frac{1}{2} \left\{ \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{1}{2}} + (y - \eta) \frac{1}{2} \left[(x - \xi)^2 + (y - \eta)^2 + z^2 \right]^{-\frac{3}{2}} \cdot 2(y - \eta) \cdot (-1) \right\} \Big|_{(0,0)} \cdot \eta^2 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \left\{ \frac{1}{R} - \frac{x^2}{R^3} \right\} \cdot \xi^2 - \frac{xy}{R^3} \cdot \xi \eta + \frac{1}{2} \left\{ \frac{1}{R} - \frac{y^2}{R^3} \right\} \cdot \eta^2 \\
&= \frac{1}{2} \frac{\xi^2 + \eta^2}{R} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} .
\end{aligned}$$

Damit haben wir $r + r_0$ entwickelt:

$$\begin{aligned}
r + r_0 &= R - \frac{x\xi + y\eta}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} \\
&+ R_0 - \frac{x_0\xi + y_0\eta}{R_0} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3} + \dots .
\end{aligned}$$

Abschließend reihen wir die Glieder entsprechend ihrer Ordnung aneinander:

$$\begin{aligned}
r + r_0 &= R + R_0 - \frac{x\xi + y\eta}{R} - \frac{x_0\xi + y_0\eta}{R_0} \\
&+ \frac{1}{2} \frac{\xi^2 + \eta^2}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3} + \dots ,
\end{aligned}$$

$$\begin{aligned}
r + r_0 &= R + R_0 - \left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi - \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \\
&+ \frac{1}{2} \frac{\xi^2 + \eta^2}{R} + \frac{1}{2} \frac{\xi^2 + \eta^2}{R_0} - \frac{1}{2} \frac{(x\xi + y\eta)^2}{R^3} - \frac{1}{2} \frac{(x_0\xi + y_0\eta)^2}{R_0^3} + \dots .
\end{aligned}$$

(112)

Im Folgenden schreiben wir (112) in einer kompakteren Form und benutzen das Gleichheitszeichen, obwohl wir auf die Fortsetzungspunkte am Ende der Entwicklung bis zur 2. Ordnung verzichten:

$$r + r_0 = R + R_0 + \phi(\xi, \eta) \quad (113)$$

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 (113) einschließlich der zweiten Ordnung in das bereits teilweise genäherte Fresnel-Kirchhoffsche Beugungsintegral (111) ein, so erhalten wir das Integral für die **Fresnel-Beugung**

$$\mathcal{E}_p = \frac{-i\epsilon_0}{\lambda} \frac{K}{R \cdot R_0} e^{ik(R+R_0)} \int_{\text{Öffnung}} e^{ik\phi(\xi, \eta)} \cdot dS$$

(114)

und sprechen von der Nahfeldnäherung, mit der auch schon das Nahfeld bzw. das Fresnel-Gebiet hinter der Öffnung beschrieben werden kann.⁴

⁴Um das Gebiet *unmittelbar* hinter der Öffnung erfassen zu können, sind auch Entwicklungsglieder der Ordnung > 2 erforderlich. Die Beschreibung dieses Gebiets kann recht mühsam sein.

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

$$e^{ik \cdot \frac{1}{2} \frac{\xi^2 + \eta^2}{R}}, \quad (115)$$

$$e^{-ik \cdot \frac{1}{2} \frac{(x \xi + y \eta)^2}{R^3}} \quad (116)$$

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 (115), also

$$e^{ik \cdot \frac{1}{2} \frac{\xi^2 + \eta^2}{R}} \stackrel{!}{=} e^{i\pi} = -1.$$

Mit dem Maximum⁵

$$\hat{\varrho} = \max \left(\sqrt{\xi^2 + \eta^2} \right) \Rightarrow \max (\xi^2 + \eta^2) = \hat{\varrho}^2$$

der Öffnung bedeutet das

$$k \cdot \frac{1}{2} \frac{\hat{\varrho}^2}{R} = \frac{2\pi}{\lambda} \cdot \frac{1}{2} \frac{\hat{\varrho}^2}{R} = \frac{\pi}{\lambda} \frac{\hat{\varrho}^2}{R} \stackrel{!}{=} \pi \Leftrightarrow \frac{\hat{\varrho}^2}{\lambda \cdot R} \stackrel{!}{=} 1 \Leftrightarrow R_{\max} = \frac{\hat{\varrho}^2}{\lambda}. \quad (117)$$

Es gilt also Faustregel⁶

Fresnel-Gebiet $R \leq \frac{\hat{\varrho}^2}{\lambda}$

Schätzen wir abschließend noch den Exponenten von (116) für R_{\max} ab. Dazu vereinfachen wir $\hat{\varrho}$, indem wir den Fall $\eta \equiv 0$ und $y \equiv 0$ betrachten, sodass $\hat{\varrho}^2 = \xi^2$. Damit erhalten wir

$$k \cdot \frac{1}{2} \frac{(x \xi + y \eta)^2}{R^3} = \frac{\pi}{\lambda} \frac{(x \xi + y \eta)^2}{R^3} \Rightarrow \frac{\pi}{\lambda} \frac{x^2 \hat{\varrho}^2}{R^3} = \pi x^2 \frac{\lambda^2}{\hat{\varrho}^4}.$$

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 (116) 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**

$$\mathcal{E}_p = \frac{-i \epsilon_0}{\lambda} \frac{K}{R \cdot R_0} e^{ik(R+R_0)} \int_{\text{Öffnung}} e^{-ik \left[\left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi + \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \right]} \cdot dS, \quad (118)$$

⁵ $\hat{\varrho}$ ist der maximale „Radius“ der Öffnung, also eine feste Größe, und darf nicht verwechselt werden mit der Variablen ϱ im Kapitel 14.

⁶ Mit der Apertur D findet man in der Literatur als Faustregel für das Nahfeld auch $L < \frac{D^2}{\lambda}$, wobei $L = z$ und $z \leq R$:

$$\hat{\varrho} \stackrel{!}{=} \frac{1}{2} D \Rightarrow R \leq \frac{\hat{\varrho}^2}{\lambda} \Rightarrow L \leq R \leq \frac{1}{4} \cdot \frac{D^2}{\lambda} \Rightarrow L < \frac{D^2}{\lambda}. \quad \square$$

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 $\hat{\varrho}^2/\lambda$:

$$\textbf{Fraunhofer-Gebiet} \quad \min\left(\{R_0, R\}\right) \gg \frac{\hat{\varrho}^2}{\lambda}, \quad \hat{\varrho} = \max\left(\sqrt{\xi^2 + \eta^2}\right).$$

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.

13 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 12.6 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 28).

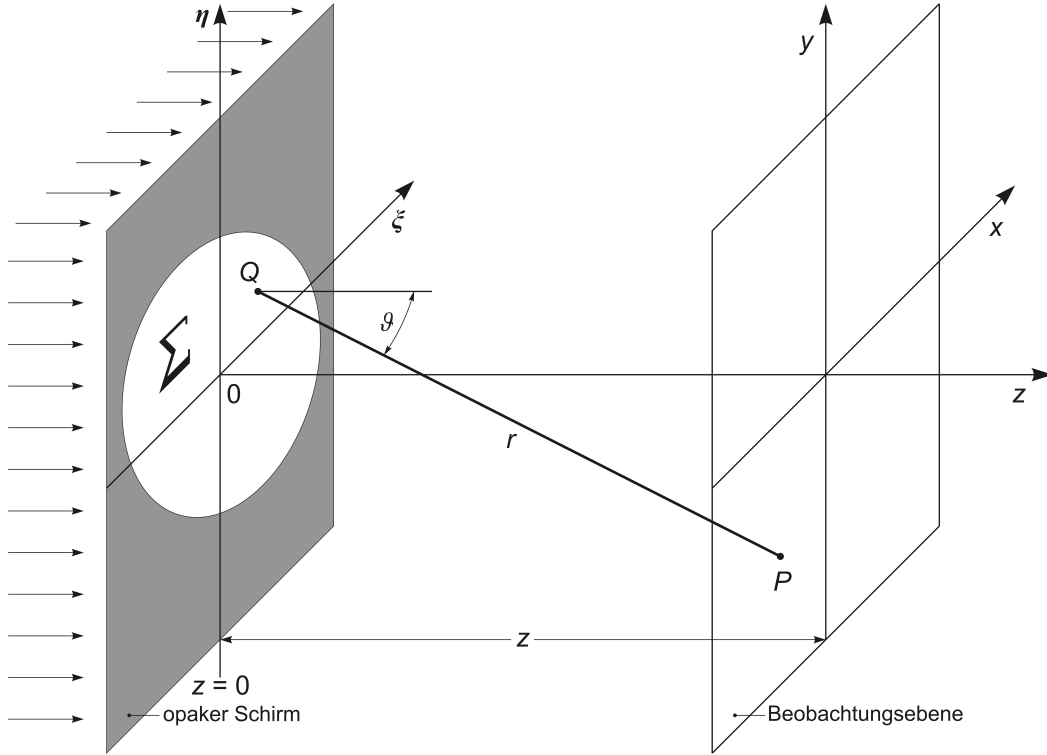


Abb. 28 Zur optischen Abbildung. Die Beobachtungsebene liegt im Abstand z planparallel zur Apertur Σ im lichtundurchlässigen Schirm. Σ ist also u. a. definiert durch $z = 0$.

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 (105)

$$\mathcal{E}_p = \frac{-i\epsilon_0}{\lambda} \int_{\text{Öffnung}} \frac{e^{ik(r+r_0)}}{r \cdot r_0} \left[\frac{\cos(\vec{n}^0, \vec{e}_r) - \cos(\vec{n}^0, \vec{e}_{r_0})}{2} \right] dS.$$

Allerdings werden wir diese Formel im Folgenden deutlich vereinfachen. Und auch die Fresnel- und die Fraunhofer-Näherung werden weitreichender ausfallen als im Abschnitt 12.6. 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“

$$E(x, y, z) = \frac{1}{i\lambda} \iint_{\Sigma} u(\xi, \eta) \frac{z}{r} \frac{e^{ikr}}{r} d\xi d\eta .$$

13.1 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

$$E(x, y, z) = \frac{1}{i\lambda \cdot z} \iint_{\Sigma} u(\xi, \eta) e^{ikr(\xi, \eta)} d\xi d\eta . \quad (119)$$

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

$$q = \left(\frac{\xi - x}{z}\right)^2 + \left(\frac{\eta - y}{z}\right)^2 = \frac{(\xi - x)^2 + (\eta - y)^2}{z^2} . \quad (120)$$

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 $q = 0$.

$$\begin{aligned} f(q) &= f(q) \Big|_{q=0} + f'(q) \Big|_{q=0} \cdot q + \frac{1}{2} \cdot f''(q) \Big|_{q=0} \cdot q^2 + \dots \\ &= (1 + q)^{\frac{1}{2}} \Big|_{q=0} + \frac{1}{2} (1 + q)^{-\frac{1}{2}} \Big|_{q=0} \cdot q + \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \frac{1}{2} (1 + q)^{-\frac{3}{2}} \Big|_{q=0} \cdot q^2 \pm \dots , \end{aligned}$$

$$f(q) = 1 + \frac{1}{2}q - \frac{1}{8}q^2 \pm \dots$$

Wir vernachlässigen im Folgenden die Entwicklungsglieder ab der 2. Ordnung und verwenden folglich im Exponenten von e^{ikr} die Näherung

$$r \approx z \cdot \left[1 + \frac{1}{2}q \right] \Rightarrow \quad (121)$$

$$r \approx z \cdot \left[1 + \frac{1}{2} \left(\frac{\xi - x}{z} \right)^2 + \frac{1}{2} \left(\frac{\eta - y}{z} \right)^2 \right] = z + \frac{(\xi - x)^2}{2z} + \frac{(\eta - y)^2}{2z}.$$

Das Fresnel-Kirchhoff'sche Beugungsintegral (119) ist damit

$$\begin{aligned} E(x, y, z) &= \frac{1}{i\lambda \cdot z} \iint_{\Sigma} u(\xi, \eta) e^{ik \left[z + (\xi - x)^2/2z + (\eta - y)^2/2z \right]} d\xi d\eta \\ &= \frac{e^{ikz}}{i\lambda \cdot z} \iint_{\Sigma} u(\xi, \eta) e^{ik \left[(\xi - x)^2/2z + (\eta - y)^2/2z \right]} d\xi d\eta \end{aligned}$$

und mit $ik \cdot \frac{1}{2z} = i \frac{2\pi}{\lambda} \cdot \frac{1}{2z} = \frac{i\pi}{\lambda z}$

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \iint_{\Sigma} u(\xi, \eta) e^{\frac{i\pi}{\lambda z} \left[(\xi - x)^2 + (\eta - y)^2 \right]} d\xi d\eta.$$

Quadrieren der Binome im Exponenten, also

$$\begin{aligned} \frac{i\pi}{\lambda z} \left[(\xi - x)^2 + (\eta - y)^2 \right] &= \frac{i\pi}{\lambda z} \left[\xi^2 - 2\xi x + x^2 + \eta^2 - 2\eta y + y^2 \right] \\ &= \frac{i\pi}{\lambda z} (\xi^2 + \eta^2) + \frac{i\pi}{\lambda z} (x^2 + y^2) - \frac{i \cdot 2\pi}{z \cdot \lambda} (x\xi + y\eta), \end{aligned}$$

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**

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \cdot e^{\frac{i\pi}{\lambda z} (x^2 + y^2)} \iint_{\Sigma} u(\xi, \eta) \cdot e^{-\frac{ik}{z} (x\xi + y\eta)} \cdot e^{\frac{i\pi}{\lambda z} (\xi^2 + \eta^2)} d\xi d\eta \quad (122)$$

Zur Bedeutung der Terme¹ in (122):

$$\frac{e^{ikz}}{i\lambda z} \equiv \text{const} \quad \textbf{Intensitätsfaktor} \text{ mit Phasenkonstante } \frac{1}{i} = -i = e^{-i\frac{\pi}{2}},$$

$$e^{\frac{i\pi}{\lambda z} (x^2 + y^2)} = g(x, y, z) \quad \textbf{Phasenfaktor},$$

$$e^{\frac{i\pi}{\lambda z} (\xi^2 + \eta^2)} = f(\xi, \eta, z) \quad \textbf{Chirp-Funktion}, \text{ engl. } to \textit{ chirp} - \text{ zwitschern, zirpen, wird in der Fraunhofer-Näherung vernachlässigt }.$$

¹Verallgemeinert 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 $|f(x)| \equiv 1$.

Fourier-Darstellung von (122) mit

$$\frac{x}{\lambda z} = \mu, \quad \frac{y}{\lambda z} = \nu, \quad \frac{e^{ikz}}{\lambda z} \cdot g(x, y, z) = \frac{e^{ikz}}{\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2+y^2)} = \frac{e^{ikz}}{\lambda z} \cdot e^{i\pi\lambda z(\mu^2+\nu^2)} = \psi(\mu, \nu, z) :$$

$$\begin{aligned} E(\mu, \nu, z) &= \frac{1}{i} \psi(\mu, \nu, z) \iint u(\xi, \eta) \cdot f(\xi, \eta, z) \cdot e^{i2\pi(\mu\xi + \nu\eta)} d\xi d\eta \\ &= \frac{1}{i} \psi(\mu, \nu, z) \cdot \mathcal{F}\{u(\xi, \eta) \cdot \text{chirp}(\xi, \eta)\} \\ E(\mu, \nu, z) &= \frac{1}{i} \psi(\mu, \nu, z) \cdot \mathcal{F}\{u(\xi, \eta)\} \otimes \mathcal{F}\{\text{chirp}(\xi, \eta)\} . \end{aligned}$$

13.2 Bedingung für die Fresnel-Näherung

Mit (120) und (121) verwenden wir in der Fresnel-Näherung

$$kr \approx kz \cdot \left[1 + \frac{1}{2}q\right] = kz \cdot \left[1 + \frac{1}{2} \frac{(\xi - x)^2 + (\eta - y)^2}{z^2}\right] \quad (123)$$

unter Vernachlässigung der Entwicklungsglieder ab der 2. Ordnung, d. h. unter der Bedingung

$$\boxed{kz \cdot \frac{1}{8}q^2 = kz \cdot \frac{1}{8} \left[\frac{(\xi - x)^2 + (\eta - y)^2}{z^2} \right]^2 \stackrel{!}{\ll} 1} .$$

Ermitteln wir also die Werte von z , welche diese Bedingung erfüllen:

$$1 \gg \frac{2\pi}{\lambda} z \cdot \frac{1}{8} \frac{[(\xi - x)^2 + (\eta - y)^2]^2}{z^4} = \frac{\pi}{4\lambda} \frac{[(\xi - x)^2 + (\eta - y)^2]^2}{z^3} \Leftrightarrow$$

$$\boxed{\text{Fresnel-Bedingung} \quad z \gg \sqrt[3]{\frac{\pi}{4\lambda} [(\xi - x)^2 + (\eta - y)^2]^2}} .$$

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: $(\xi - x) = (\eta - y) = 1 \text{ cm} = 10^{-2} \text{ m}$, $\lambda = 500 \text{ nm} = 5 \cdot 10^{-7} \text{ m}$,

$$\begin{aligned} z &\gg \sqrt[3]{\frac{\pi \cdot [(10^{-2})^2 + (10^{-2})^2]^2 \text{ m}^4}{4 \cdot 5 \cdot 10^{-7} \text{ m}}} = \sqrt[3]{\frac{\pi \cdot 4 \cdot 10^{-8} \text{ m}^4}{20 \cdot 10^{-7} \text{ m}}} = \sqrt[3]{\frac{\pi}{50} \text{ m}^3} , \\ z &\gg 0,40 \text{ m} = 40 \text{ cm} . \end{aligned}$$

13.3 Fraunhofer-Näherung

Ausgehend von der Fresnel-Näherung (122) 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

$$\boxed{E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2+y^2)} \iint_{\Sigma} u(\xi, \eta) \cdot e^{-\frac{ik}{z}(x\xi + y\eta)} d\xi d\eta} . \quad (124)$$

Fourier-Darstellung von (124) mit

$$\frac{x}{\lambda z} = \mu, \quad \frac{y}{\lambda z} = \nu, \quad \frac{e^{ikz}}{\lambda z} \cdot g(x, y, z) = \frac{e^{ikz}}{\lambda z} \cdot e^{\frac{i\pi}{\lambda z}(x^2+y^2)} = \frac{e^{ikz}}{\lambda z} \cdot e^{i\pi\lambda z(\mu^2+\nu^2)} = \psi(\mu, \nu, z) :$$

$$E(\mu, \nu, z) = \frac{1}{i} \psi(\mu, \nu, z) \iint u(\xi, \eta) \cdot e^{i2\pi(\mu\xi + \nu\eta)} d\xi d\eta$$

$$E(\mu, \nu, z) = \frac{1}{i} \psi(\mu, \nu, z) \cdot \mathcal{F}\{u(\xi, \eta)\} .$$

13.4 Bedingung für die Fraunhofer-Näherung

Ausgehend von (123), 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.

14 Fraunhofer-Beugung und 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 14, bei der man sich die Linse L_2 in der Öffnungsebene Σ positioniert vorstellen sollte, insbesondere die folgende Abbildung 29:

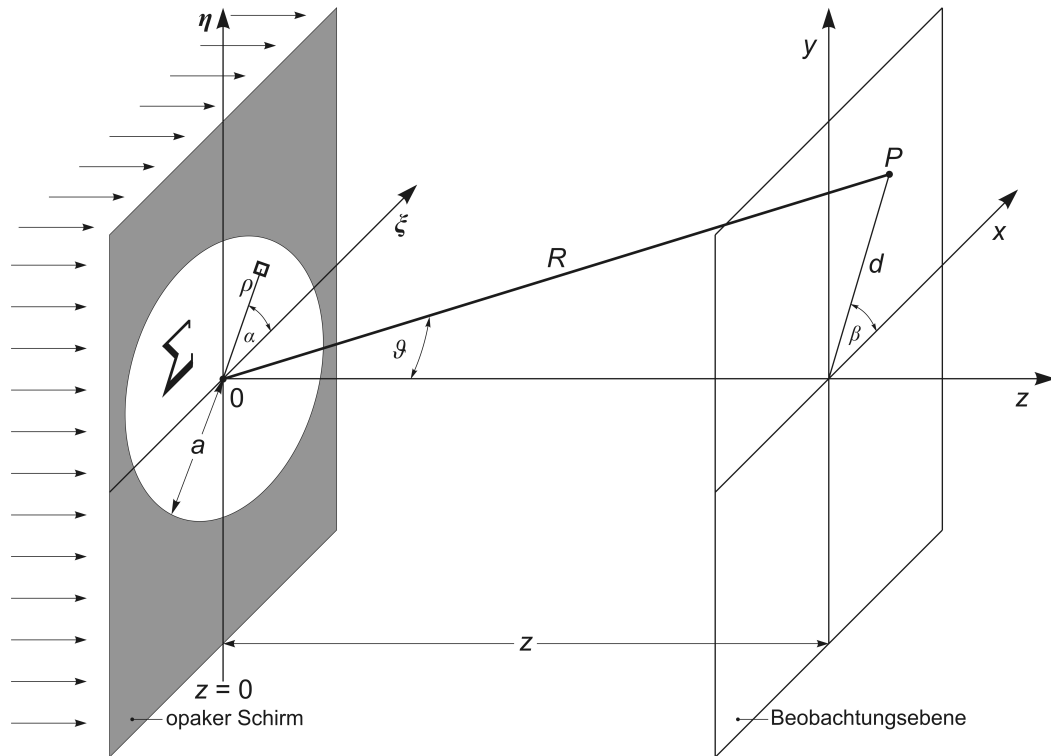


Abb. 29 Zur Fraunhofer-Beugung an einer kreisrunden Öffnung bzw. Apertur Σ mit dem Durchmesser $D = 2 \cdot a$.

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.“¹

Die Öffnung sei gleichmäßig ausgeleuchtet, d. h., für die optische Feldstärke in der Ebene Σ gelte

$$u(\xi, \eta) \equiv \text{const.}$$

¹Zitiert aus Eugene Hecht, Optik, Oldenbourg-Verlag, München, Wien, 4. Auflage, 2002, Seite 752.

Wir gehen aus von der im Abschnitt 12.6 hergeleiteten Formel (118)

$$\begin{aligned}
\mathcal{E}_p &= \frac{-i\epsilon_0}{\lambda} \frac{K}{R \cdot R_0} e^{ik(R+R_0)} \int_{\text{Öffnung}} e^{-ik \left[\left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi + \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \right]} \cdot dS \\
&= \frac{1}{i\lambda} \cdot \frac{\epsilon_0 e^{ikR_0}}{R_0} \cdot K \cdot \frac{e^{ikR}}{R} \int_{\text{Öffnung}} e^{-ik \left[\left(\frac{x}{R} + \frac{x_0}{R_0} \right) \xi + \left(\frac{y}{R} + \frac{y_0}{R_0} \right) \eta \right]} \cdot dS
\end{aligned} \tag{125}$$

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 (125) 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:

$$\begin{aligned}
\Sigma : \quad \xi &= \varrho \cdot \cos \alpha, & \eta &= \varrho \cdot \sin \alpha, \\
\sigma : \quad x &= d \cdot \cos \beta, & y &= d \cdot \sin \beta, \\
dS &= d\xi d\eta = \varrho d\varrho d\alpha.
\end{aligned}$$

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:

$$E(d, \beta = 0, z) = \frac{\epsilon_A e^{ikR}}{R} \int_{\varrho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k \varrho d}{R} \cdot \cos \alpha} \cdot \varrho d\varrho d\alpha . \quad (126)$$

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

$$J_n(w) = \frac{1}{2\pi \cdot i^n} \int_0^{2\pi} e^{i(n \cdot \alpha + w \cdot \cos \alpha)} d\alpha$$

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 (126)

$$E(d, \beta = 0, z) = \frac{\epsilon_A e^{ikR}}{R} \int_{\varrho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k \varrho d}{R} \cdot \cos \alpha} \cdot \varrho d\varrho d\alpha .$$

Wir verwenden die Substitution

$$-\frac{k\rho d}{R} = w \quad \Leftrightarrow \quad \rho = -\frac{R}{kd} w \quad \Rightarrow \quad d\rho = -\frac{R}{kd} dw$$

mit den daraus folgenden Integralgrenzen

$$\rho = 0 \quad \Rightarrow \quad w = 0 \quad \text{und} \quad \rho = a \quad \Rightarrow \quad w = -\frac{kad}{R}$$

und lassen den Vorfaktor des Flächenintegrals zunächst unberücksichtigt. Der Winkelanteil des Integrals ist damit

$$\begin{aligned} \int_0^{2\pi} e^{i(-k\rho d/R) \cdot \cos \alpha} d\alpha &= \int_0^{2\pi} e^{i w \cdot \cos \alpha} d\alpha = 2\pi \left(\frac{1}{2\pi} \int_0^{2\pi} e^{i w \cdot \cos \alpha} d\alpha \right) \\ &= 2\pi \cdot J_0(w) . \end{aligned}$$

Vom Flächenintegral verbleibt folglich das Integral längs des Radius ρ :

$$\begin{aligned} \int_{\rho=0}^a 2\pi \cdot J_0(w) \cdot \rho \cdot d\rho &= 2\pi \int_{w=0}^{-kad/R} J_0(w) \cdot \left(-\frac{R}{kd} w\right) \cdot \left(-\frac{R}{kd} dw\right) \\ &= 2\pi \left(\frac{R}{kd}\right)^2 \int_{w=0}^{-kad/R} J_0(w) \cdot w \cdot dw \\ &= 2\pi \left(\frac{R}{kd}\right)^2 \left[w \cdot J_1(w) \right]_{w=0}^{-kad/R} . \end{aligned}$$

Da $J_1(w)$ eine ungerade Funktion ist, gilt

$$w = -\frac{kad}{R} \quad \Rightarrow \quad J_1(-kad/R) = -J_1(kad/R) \quad \Rightarrow \quad \left[w \cdot J_1(w) \right] \Big|_{w=-kad/R} = \frac{kad}{R} J_1(kad/R) ,$$

sodass

$$\begin{aligned} \int_{\rho=0}^a 2\pi \cdot J_0(w) \cdot \rho \cdot d\rho &= 2\pi \left(\frac{R}{kd}\right)^2 \cdot \frac{kad}{R} J_1(kad/R) , \\ &= 2\pi a^2 \cdot \frac{R}{kad} \cdot J_1(kad/R) . \end{aligned}$$

Mit dem Flächeninhalt

$$\pi a^2 = A$$

der Öffnung Σ ist dann das Flächenintegral insgesamt, also auch mit dem Vorfaktor,

$$\frac{\epsilon_A e^{ikR}}{R} \int_{\rho=0}^a \int_{\alpha=0}^{2\pi} e^{-i \frac{k\rho d}{R} \cdot \cos \alpha} \cdot \rho d\rho d\alpha = \frac{\epsilon_A e^{ikR}}{R} \cdot 2 \cdot A \cdot \frac{R}{kad} \cdot J_1(kad/R) . \quad (127)$$

Zur besseren Übersicht in den folgenden Rechnungen ordnen wir die Terme in (127) um und schreiben für die Lösung des Flächenintegrals (126) schließlich

$$\boxed{E(d, z) = \frac{2\epsilon_A}{R} e^{ikR} \cdot A \cdot \frac{J_1(kad/R)}{kad/R}} .$$

Wegen der Achsensymmetrie des Beugungsmusters, das im Fall der kreisrunden Öffnung auch als **Airy-Muster**² bezeichnet wird, ist die Bestrahlungsstärke bzw. der zeitliche Mittelwert der Intensität *in der Beobachtungsebene* auf folgende Weise abhängig vom Radius d :

$$\bar{I}(d) = \left\langle (\text{Re}\{E\})^2 \right\rangle_t = \frac{1}{2} \cdot E \cdot E^* = \frac{1}{2} \cdot \frac{4\epsilon_A^2 A^2}{R^2} \left[\frac{J_1(kad/R)}{kad/R} \right]^2 ,$$

²Sir George Bidell Airy (1801–1892): königlich-englischer Astronom.

$$\boxed{\bar{I}(d) = \frac{2 \epsilon_A^2 A^2}{R^2} \left[\frac{J_1(kad/R)}{kad/R} \right]^2} . \quad (128)$$

Daraus und mit dem Ergebnis (130) aus der folgenden Nebenrechnung erhalten wir die Bestrahlungsstärke im Zentrum des Beugungsmusters, also für $d = 0 \Rightarrow w = 0$, wie folgt:

Nebenrechnung

$$\begin{aligned} \frac{J_1(0)}{0} &= \frac{0}{0} \text{ nicht definiert} \Rightarrow \text{mit L'Hospital :} \\ \lim_{w \rightarrow 0} \frac{J_1(w)}{w} &= \lim_{w \rightarrow 0} \frac{\frac{d}{dw} J_1(w)}{\frac{d}{dw} w} = \lim_{w \rightarrow 0} \frac{\frac{d}{dw} J_1(w)}{1} , \\ \lim_{w \rightarrow 0} \frac{J_1(w)}{w} &= \lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) . \end{aligned} \quad (129)$$

$$\begin{aligned} \frac{d}{dw} [w \cdot J_1(w)] &= w \cdot J_0(w) = w \cdot \frac{d}{dw} J_1(w) + J_1(w) \\ \Leftrightarrow J_0(w) &= \frac{d}{dw} J_1(w) + \frac{J_1(w)}{w} \\ \Rightarrow J_0(0) = 1 &= \lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) + \lim_{w \rightarrow 0} \frac{J_1(w)}{w} \\ &\stackrel{(129)}{\Rightarrow} 1 = 2 \cdot \lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) \Leftrightarrow \\ &\boxed{\lim_{w \rightarrow 0} \frac{d}{dw} J_1(w) = \frac{1}{2}} . \end{aligned} \quad (130)$$

$$\bar{I}(d=0) = \bar{I}_0 = \frac{2 \epsilon_A^2 A^2}{R^2} \cdot \lim_{d \rightarrow 0} \left[\frac{J_1(kad/R)}{kad/R} \right]^2 \stackrel{(130)}{=} \frac{2 \epsilon_A^2 A^2}{R^2} \cdot \left(\frac{1}{2} \right)^2 ,$$

$$\boxed{\bar{I}_0 = \frac{\epsilon_A^2 A^2}{2 R^2}} .$$

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 29) angegeben wird. Dazu klammern wir \bar{I}_0 aus (128) aus und führen die Substitution

$$\frac{d}{R} = \sin \vartheta$$

durch:

$$\begin{aligned} \bar{I}(d) &= \frac{\epsilon_A^2 A^2}{2 R^2} \cdot 4 \left[\frac{J_1(kad/R)}{kad/R} \right]^2 = \bar{I}_0 \cdot \left[2 \cdot \frac{J_1(ka \frac{d}{R})}{ka \frac{d}{R}} \right]^2 , \\ &\boxed{\bar{I}(\vartheta) = \bar{I}_0 \cdot \left[\frac{2 \cdot J_1(ka \sin \vartheta)}{ka \sin \vartheta} \right]^2} . \end{aligned} \quad (131)$$

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.

14.1 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**³ (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 **auf lösbare 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 (131) 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 Zentralstrahlen⁴ 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 29) bekannt, kann man den (gerade noch auflösbaren) Mindestabstand d_{\min} **in der Beobachtungsebene** berechnen mit Hilfe von

$$\sin \theta_{\min} = \frac{d_{\min}}{R} = 1,22 \cdot \frac{\lambda}{D} \quad \Leftrightarrow$$

$$d_{\min} = 1,22 \cdot \frac{R \lambda}{D}. \quad (132)$$

- $\sin \theta_{\min}$ ist die **Winkelauflösungsgrenze**.
- θ_{\min} ist der **minimale Winkelabstand**.
- d_{\min} ist der **bildseitige** (auflösbare) **Mindestabstand**.

³ John William Strutt „Lord Rayleigh“ (1842–1919): englischer Physiker und Nobelpreisträger.

⁴ Als Zentralstrahl bezeichnen wir hier den Strahl von der Punktquelle zum (zentralen) Hauptmaximum des Airy-Musters.

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 (132) 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 „Bildpunkt Abstand“

$$\Delta y'_{\min} = \gamma \cdot \frac{\lambda_0}{n' \cdot \sin \alpha'} , \quad 0,47 \leq \gamma \leq 1,22 .$$

- Das Auflösungsvermögen des Mikroskops *entspricht* dem objektseitigen Mindestabstand, d. h. dem kleinsten erfassbaren Objektpunkt Abstand

$$\Delta y_{\min} = \gamma \cdot \frac{\lambda_0}{n \cdot \sin \alpha} , \quad 0,47 \leq \gamma \leq 1,22 .$$

λ_0 ist die Vakuumwellenlänge, $n \cdot \sin \alpha$ ist die objektseitige und $n' \cdot \sin \alpha'$ die bildseitige numerische Apertur NA . Der Faktor $\gamma = 0,47$ resultiert aus dem Sparrow-Kriterium, auf das wir nicht eingehen werden.

15 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 34). 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} .

15.1 Laterales Auflösungsvermögen des Mikroskops mit 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 .

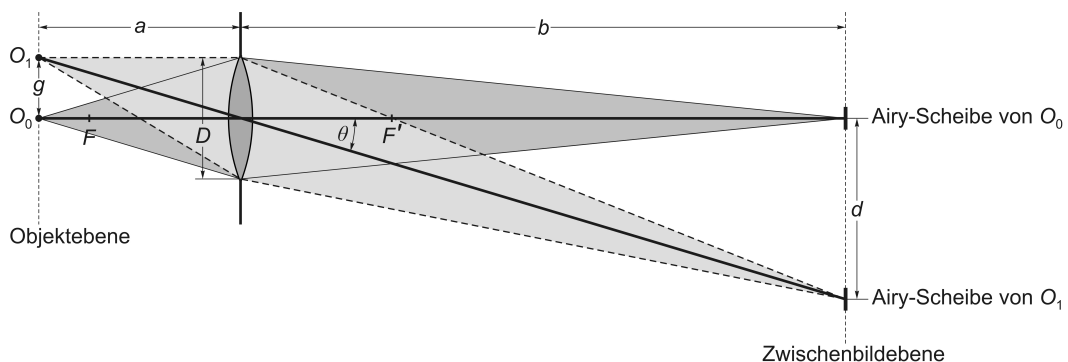


Abb. 30 Zum Auflösungsvermögen des Mikroskops mit dem Rayleigh-Kriterium.

Von zwei beleuchteten Objektpunkten O_0 und O_1 (siehe Abbildung 30) 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 (132) zur Berechnung des bildseitigen (auflösbaren) Mindestabstands d_{\min} heranziehen können:

$$d_{\min} = 1,22 \cdot \frac{b \lambda}{D} . \quad (133)$$

Diesem bildseitigen Mindestabstand d_{\min} entspricht ein objektseitiger (gerade noch auflösbarer) Mindestabstand g_{\min} nach dem Strahlensatz wie folgt:

$$\frac{g_{\min}}{a} = \frac{d_{\min}}{b} = 1,22 \cdot \frac{\lambda}{D} \quad \Leftrightarrow \quad g_{\min} = 1,22 \cdot a \cdot \frac{\lambda}{D} . \quad (134)$$

Weil allgemein in der Durchlichtmikroskopie die Objektebene praktisch in der vorderen Brennebene des Objektivs liegt, setzen wir $a = f$ und erhalten aus (134) 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 (15.1) ein, so erhalten wir

$$g_{\min} \approx 1,22 \cdot \frac{\lambda}{2 \cdot \sin \alpha} = 1,22 \cdot \frac{\lambda_0}{2 \cdot n \cdot \sin \alpha} ,$$

$$\boxed{g_{\min} \approx 0,61 \cdot \frac{\lambda_0}{n \cdot \sin \alpha} = 0,61 \cdot \frac{\lambda_0}{NA} .}$$

15.2 Laterales Auflösungsvermögen des Mikroskops mit 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:

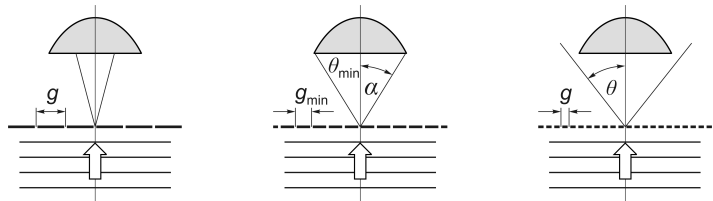
$$I(\theta) = \left[A_1(\theta) + A_2(\theta) \right]^2.$$

1873 beantwortete Ernst Abbe¹ 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 31 und Abbildung 34). Man spricht in diesem Zusammenhang folglich vom

Gitterauflösungsvermögen.

Abb. 31 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.

¹Ernst Carl Abbe (1840–1905): deutscher Physiker, Optiker und Industrieller.

15.2.1 Senkrecht einfallendes Licht

Wir betrachten zunächst den einfachen Fall, dass das Objekt (Strichgitter) mit senkrecht einfallendem kohärentem Licht beleuchtet (durchleuchtet) wird (siehe Abbildung 32).

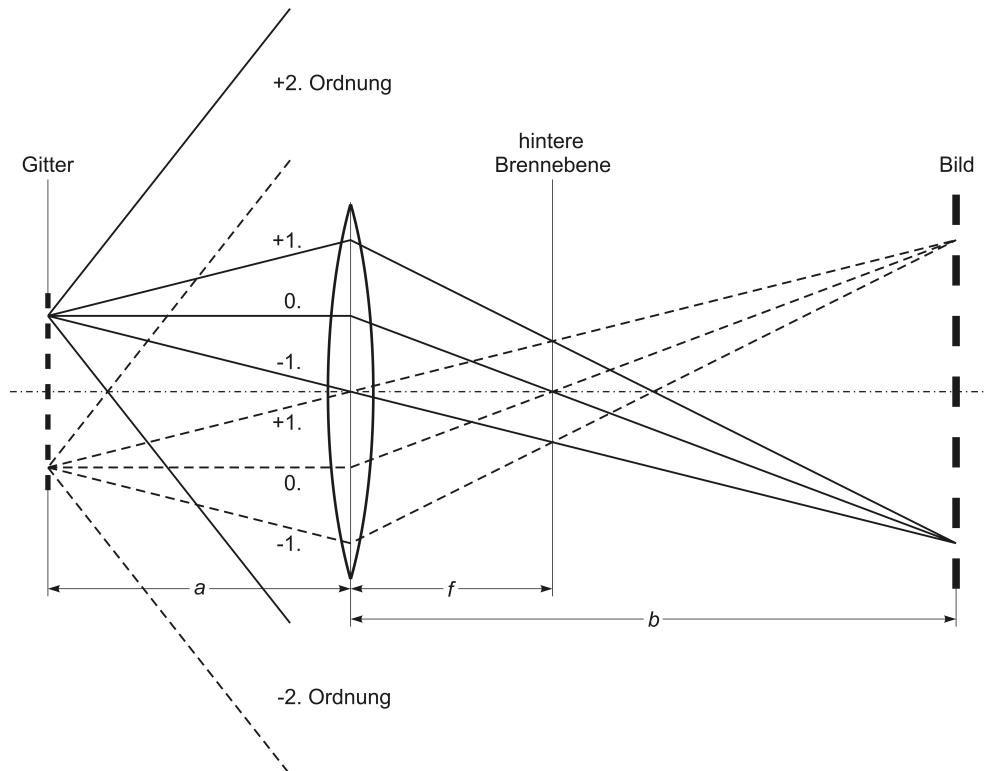


Abb. 32 Zum Auflösungsvermögen des Mikroskops nach dem Abbe-Kriterium. Eine ebene kohärente Welle falle von links senkrecht auf das Strichgitter. Die Sammellinse generiert in ihrer hinteren Brennebene aus dem am Gitter gebeugten Licht das Fraunhofer-Beugungsmuster und erzeugt schließlich in der Bildebene das umgekehrte Bild des Strichgitters in einem bestimmten Abbildungsmaßstab.

Mit (72) 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 (73) 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 (140), also mit dem Zusammenhang zwischen der Vakuumwellenlänge λ_0 und der Wellenlänge im Medium mit dem Brechungsindex n , ergibt dies

$$g_{\min} = \frac{\lambda}{\sin \theta_{\min}} \xrightarrow{(140)} g_{\min} = \frac{\lambda_0}{n \cdot \sin \theta_{\min}}. \quad (135)$$

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:

$$\delta_1 = 2\pi = \frac{2\pi}{\lambda} \cdot g \sin \theta_1 \quad \Leftrightarrow \quad \sin \theta_1 = \frac{1}{g} \cdot \lambda$$

$$\Rightarrow \sin \theta_1 \propto \lambda ,$$

$$\xrightarrow{\lambda = \frac{\lambda_0}{n}} \sin \theta_1 \propto \frac{1}{n} .$$

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.

15.2.2 Schräg einfallendes Licht

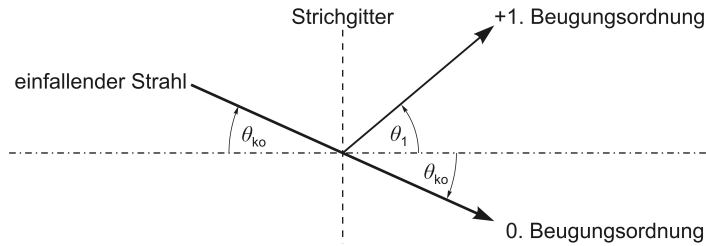


Abb. 33 Zur „schiefen Beleuchtung“ des Strichgitters. Der Index ko steht für kondensorseitig und der Index 1 für objektivseitig und/oder 1. Beugungsordnung. Ein vom Kondensor ausgehendes kohärentes Lichtbündel (einfallender Strahl) fällt unter dem Winkel θ_{ko} auf die Beugungsebene (Strichgitter). Der 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

$$\Delta s = \Delta s_1 - \Delta s_{ko} = g \sin \theta_1 - g \sin \theta_{ko} \quad \text{für} \quad \theta_{ko} < 0, \quad \theta_1 > 0 \quad (136)$$

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 33 mit $\sin(-\theta_{ko}) = -\sin \theta_{ko}$ und im Hinblick auf die folgende Herleitung des Auflösungsvermögens des Mikroskops schreiben wir für (136)

$$\Delta s = g \cdot |\sin \theta_{ko}| + g \cdot |\sin \theta_1| .$$

Daraus resultiert dann die Phasendifferenz bei konstruktiver Interferenz in der 1. Beugungsordnung entsprechend $m = 1$ und mit

$$\lambda_{ko} = \frac{\lambda_0}{n_{ko}}, \quad \lambda_1 = \lambda = \frac{\lambda_0}{n} :$$

$$\begin{aligned} \delta &= \frac{2\pi}{\lambda_{ko}} g |\sin \theta_{ko}| + \frac{2\pi}{\lambda_1} g |\sin \theta_1| \\ &= \frac{2\pi}{\lambda_0} g n_{ko} |\sin \theta_{ko}| + \frac{2\pi}{\lambda_0} g n |\sin \theta_1| , \\ \delta &= \frac{2\pi}{\lambda_0} g \left(n_{ko} |\sin \theta_{ko}| + n |\sin \theta_1| \right) = m \cdot 2\pi , \\ \delta_{(m=1)} &= \frac{2\pi}{\lambda_0} g \left(n_{ko} |\sin \theta_{ko}| + n |\sin \theta_1| \right) = 2\pi . \end{aligned} \quad (137)$$

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 (137) ermitteln:

$$\boxed{g_1 = \frac{\lambda_0}{n_{ko} |\sin \theta_{ko}| + n |\sin \theta_1|}} . \quad (138)$$

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 34).

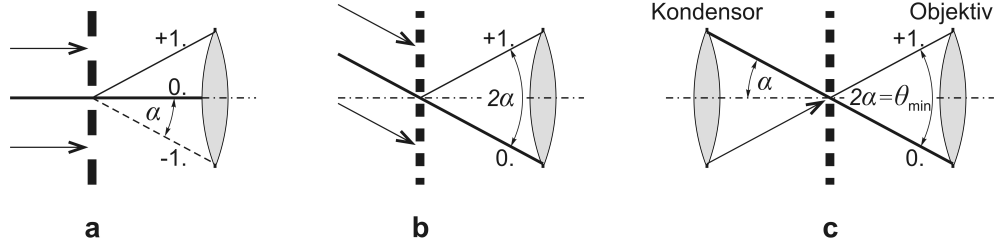


Abb. 34 Zum Auflösungsvermögen des Mikroskops nach dem Abbe-Kriterium. Als Objekt dient hier ein Strichgitter.

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 Kondensor, 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 Kondensors $2\alpha_{ko}$ und gilt

$$|\sin \theta_{ko}| = \sin \alpha_{ko} \quad \text{und} \quad |\sin \theta_1| = \sin \alpha ,$$

so definiert (138) den gerade noch auflösbaren Abstand bzw. den Mindestabstand zwischen benachbarten Gitterstrichen für den allgemeinen Fall, dass die Brechungsindizes und die Öffnungswinkel kondensor- und objektivseitig verschieden sind:

$$g_{\min} = \frac{\lambda_0}{n_{ko} |\sin \theta_{ko}| + n |\sin \theta_1|} = \frac{\lambda_0}{n_{ko} \sin \alpha_{ko} + n \sin \alpha} ,$$

$$g_{\min} = \frac{\lambda_0}{NA_{Kondensor} + NA_{Objektiv}} .$$

(139)

Das Produkt aus halbem Öffnungswinkel von Kondensor bzw. Objektiv mit dem Brechungsindex des jeweils umgebenden Mediums bezeichnet man als numerische Apertur NA des Kondensors bzw. Objektivs.

Praktisch realisiert wird die „schiefe Beleuchtung“ im einfachsten Fall durch einen Kondensor, der den gleichen Öffnungswinkel 2α wie das Objektiv besitzt (siehe Abbildung 34 c). Die allgemeine Formel (139) vereinfacht sich dann wegen $\sin \alpha_{ko} = \sin \alpha$ zu

$$g_{\min} = \frac{\lambda_0}{(n_{ko} + n) \cdot \sin \alpha} .$$

Wenn das Medium kondensorseitig Luft gemäß $n_{ko} = n = 1$ ist, was fast immer der Fall ist, erhalten wir

$$g_{\min} = \frac{\lambda_0}{(1 + n) \cdot \sin \alpha} .$$

Und wenn das Medium sowohl kondensor- 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

$$\textbf{Grenzauflösung } g_{\min} = 0,51 \frac{\lambda_0}{n \cdot \sin \alpha} ,$$

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: $g_{\min} = 0,51 \cdot \frac{\lambda_0}{n \cdot \sin \alpha} .$

16 Zur Abbe'schen Abbildungstheorie

• Abbe'sche Sinusbedingung¹

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:

$$c = \frac{c_0}{n} = \lambda \cdot f \quad \Leftrightarrow \quad c_0 = n \lambda \cdot f = \lambda_0 \cdot f \quad \Rightarrow$$

$$\text{Vakuumwellenlänge } \lambda_0 = n \lambda \quad \Rightarrow \quad \lambda = \frac{\lambda_0}{n} \propto \frac{1}{n} . \quad (140)$$

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.

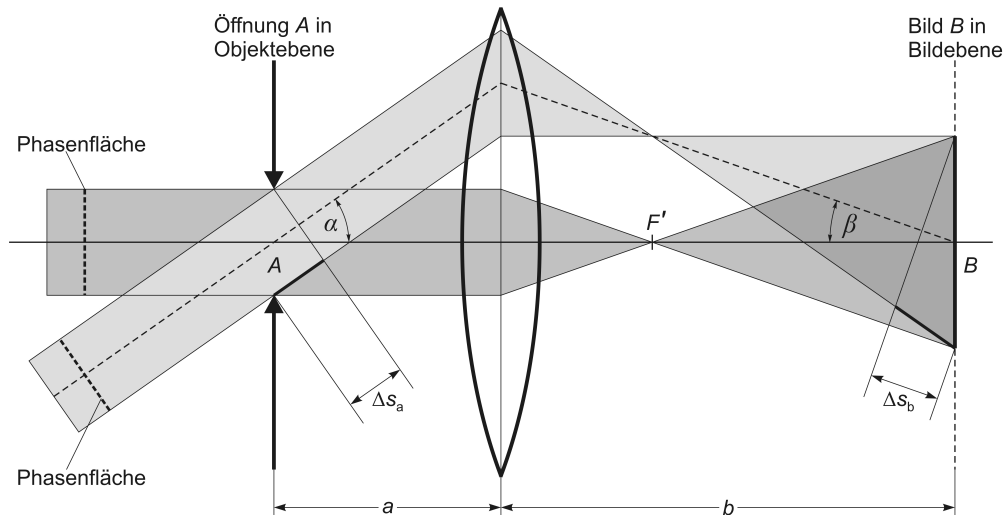


Abb. 35 Zur Abbe'schen Sinusbedingung. (Vergleiche mit den Abbildungen 15 und 16 im Abschnitt 11.2.1.)

Wie in der Abbildung 35 zu sehen ist, beleuchten wir eine Kreisblende, das Objekt A in der Objektelebene, 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 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

¹Siehe 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.

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**

$$\delta = k \cdot \Delta s = \frac{2\pi}{\lambda} \cdot \Delta s \stackrel{(140)}{=} \frac{2\pi}{\frac{\lambda_0}{n}} \cdot \Delta s = \frac{2\pi}{\lambda_0} \cdot n \Delta s$$

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 Objektpfeil-Spitze zu Bildpfeil-Spitze gleich der Änderung der Phase im Strahlenverlauf Objektpfeil-Basis zu Bildpfeil-Basis 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 $\frac{n_a \sin \alpha}{n_b \sin \beta} = \frac{ B }{ A } = M = \text{const}$.

17 Anhang

17.1 Hauptachsentransformation von Kegelschnitten zur Bestimmung der Parameter von Ellipsen

Siehe dazu:

- Wikipedia, Suchbegriff *Hauptachsentransformation*.
- Friedrich W. Buckel, Internetbibliothek für Schulmathematik und Studium, 2022, Text Nr. 54303, Suchbegriff *mathe-cd Schräge Ellipsen Hauptachsentransformation*.
- Suchbegriff *46 Eigenwerte und Eigenvektoren symmetrischer Matrizen*, <https://resources.mpi-inf.mpg.de/departments/d1/teaching/ss10/MFI2/kap46.pdf>
- mein Skript *Mathematik – Einige ausgewählte Themen für das Physikstudium*, Kapitel 13 *Eigenwertgleichung einer 2-reihigen reellen Matrix – Verallgemeinerung für n-reihige Matrizen* und Abschnitt 15.3 *Diagonalisierung von Matrizen (Operatoren)*.

Kegelschnitte sind Kreise, Ellipsen, Parabeln und Hyperbeln. Unser Ziel ist es, im Allgemeinen zu zeigen, wie man mit dem Verfahren der Hauptachsentransformation die Parameter von Kegelschnitten bestimmen kann, die nicht in ihrer Normalform gegeben sind.

Die Lösungsmenge einer quadratischen Gleichung mit mehreren Unbekannten (Variablen) ist eine Punktmenge, die Quadrik (Plural: Quadriken). Im \mathbb{R}^2 bzw. in der Ebene, beispielsweise dargestellt durch ein kartesisches (x, y) -Koordinatensystem, sind dies 2-dimensionale Quadriken, welche in der allgemeinen Form durch die quadratische Gleichung

$$\begin{aligned} a x^2 + b y^2 + c x y + d x + e y + f &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + d x + e y + f \\ &= \vec{r}^T A \vec{r} + d x + e y + f = 0 \end{aligned} \quad (141)$$

mit den Koeffizienten a, b, c, d, e, f definiert sind. Es handelt sich bei dieser Gleichung um eine implizite Funktion der Form $F(x, y) = 0$ und im Prinzip um ein unterbestimmtes nichtlineares Gleichungssystem, das nur aus einer Gleichung aber mit zwei Unbekannten besteht.

Die symmetrische 2-reihige (quadratische) Matrix A in (141) besitzt mit den Koeffizienten a, b, c die Form

$$A = \begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix}. \quad (142)$$

Die linke Seite von (141) wird quadratische Form genannt. Typische 2-dimensionale Quadriken sind beispielsweise die Kegelschnitte. Aber nicht alle Lösungen von (141) sind Kegelschnitte, jedoch sind alle Kegelschnitte Lösungen von (141). Achsenparallele Kegelschnitte befinden sich in **Normalform**. Ist ein achsenparalleler Kegelschnitt auch noch in Nullpunkts- oder Ursprungslage, so befindet er sich in seinem **Hauptachsensystem**, d. h. die Hauptachsen des Kegelschnitts und die Achsen des Koordinatensystems fallen zusammen.

17.1.1 Hauptachsentransformation

Wir können einen durch (141) beschriebenen Kegelschnitt in die Normalform überführen, indem wir den Kegelschnitt im (x, y) -Koordinatensystem so drehen, dass er im (x, y) -Koordinatensystem achsenparallel liegt, oder indem wir das ursprüngliche (x, y) -Koordinatensystem (dazu entgegengesetzt) drehen, sodass das neue (ξ, η) -Koordinatensystem achsenparallel zum Kegelschnitt liegt. Dabei verschwindet der Koeffizient c bzw. der für die Verdrehung des Kegelschnitts aus der Normalform zuständige gemischte Term cxy und aus der symmetrischen Matrix A wird mittels der Transformations- bzw. **Drehmatrix** S die

$$\text{Diagonalmatrix } D_A = S^T A S. \quad (143)$$

Hilfreich ist hier, dass Drehmatrizen (reelle) orthogonale Matrizen sind und folglich die Eigenschaft

$$S = S^T = S^{-1} \quad \text{und im } \mathbb{R}^2 \quad \det S = +1 \quad \text{oder} \quad \det S = -1$$

besitzen. S^T und S sind also die Matrizen für die Diagonalisierung $A \rightarrow D_A$ und gleichzeitig ist S aber auch die Transformationsmatrix für den Übergang der Koordinaten $(\xi, \eta) \rightarrow (x, y)$ gemäß

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = S \begin{pmatrix} \xi \\ \eta \end{pmatrix} = S \vec{q}.$$

Erläuterungen zur Verwendung der Drehmatrix S :

Das kartesische (ξ, η) -Koordinatensystem sei im mathematisch positiven Drehsinn gegenüber dem kartesischen (x, y) -Koordinatensystem um den gemeinsamen Koordinatenursprung um den Winkel α gedreht. Dann wird die zugehörige Drehmatrix M üblicherweise wie folgt notiert:

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow M =: \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = S^{-1} = S^T.$$

In Anlehnung an den Wikipedia-Bertrag zur Hauptachsentransformation verwenden wir für unsere Drehmatrix S :

$$\begin{pmatrix} x \\ y \end{pmatrix} = M^{-1} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \Rightarrow M^{-1} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = S \Rightarrow S \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Bei der Diagonalisierung von A mittels S verändern sich die Koeffizienten a, b, c bzw. sie werden so transformiert, dass der Kegelschnitt im (x, y) -Koordinatensystem in seine achsenparallele Form gedreht wird. Die anschließende Umbenennung der Koordinaten $x \rightarrow \xi, y \rightarrow \eta$ liefert dann die gewünschte achsenparallele Form des Kegelschnitts im (ξ, η) -Koordinatensystem.

Da die Gleichung (141) keine Vektorgleichung ist, sind die Koeffizienten d und e der linearen Terme (Glieder) dx und ey keine skalaren Vektorkomponenten und werden deshalb beim Übergang zum gedrehten (ξ, η) -Koordinatensystem nicht „direkt“ transformiert. In diesen Gliedern betrifft die Transformation im Grunde nur die Koordinaten x, y , die durch die Koordinaten ξ, η ersetzt werden müssen, indem man $x = x(\xi, \eta)$ und $y = y(\xi, \eta)$ verwendet. Die resultierenden Terme werden dann geordnet und schließlich zu den linearen Termen im (ξ, η) -Koordinatensystem zusammengefasst. Die Transformation der linearen Terme geschieht also mittels S wie folgt:

$$\begin{pmatrix} x \\ y \end{pmatrix} = S \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi \cos \alpha - \eta \sin \alpha \\ \xi \sin \alpha + \eta \cos \alpha \end{pmatrix} = \begin{pmatrix} x(\xi, \eta) \\ y(\xi, \eta) \end{pmatrix},$$

$$\begin{aligned} d \cdot x + e \cdot y &= d \cdot (\xi \cos \alpha - \eta \sin \alpha) + e \cdot (\xi \sin \alpha + \eta \cos \alpha) \\ &= \underbrace{(d \cos \alpha + e \sin \alpha)}_{\tilde{d}} \xi + \underbrace{(-d \sin \alpha + e \cos \alpha)}_{\tilde{e}} \eta, \end{aligned}$$

$$\boxed{d \cdot x + e \cdot y = \tilde{d} \cdot \xi + \tilde{e} \cdot \eta}$$

mit

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix} = S^{-1} \begin{pmatrix} d \\ e \end{pmatrix} = \begin{pmatrix} \tilde{d} \\ \tilde{e} \end{pmatrix}.$$

Probe mit $\begin{pmatrix} \xi \\ \eta \end{pmatrix} = S^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$, $(S^{-1})^T = (S^T)^T = S$ und $(AB)^T = B^T A^T$:

$$\tilde{d} \xi + \tilde{e} \eta = \begin{pmatrix} \tilde{d} \\ \tilde{e} \end{pmatrix}^T \cdot \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \left[S^{-1} \begin{pmatrix} d \\ e \end{pmatrix} \right]^T \cdot S^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} d \\ e \end{pmatrix}^T \underbrace{S \cdot S^{-1}}_{=1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} d \\ e \end{pmatrix}^T \cdot \begin{pmatrix} x \\ y \end{pmatrix} = dx + ey. \quad \square$$

Symmetrische $(n \times n)$ -Matrizen wie beispielsweise die Matrix A in (142) lassen sich immer mit Hilfe einer orthogonalen Matrix S diagonalisieren. So entspricht, wie wir später noch sehen werden, beispielsweise die Drehung einer Ellipse von (147) nach (146) der Diagonalisierung von A (nach D_A).

Wir erhalten die Drehmatrix S durch Bestimmung der Eigenwerte λ_i von A sowie durch Bestimmung der zugehörigen normierten Eigenvektoren \vec{v}_i aus der quadratischen charakteristischen Gleichung wie folgt:

- Aufstellen der **charakteristischen Gleichung**¹ von A mit der Einheitsmatrix $\mathbb{1}$:

$$\begin{aligned} \det(A - \lambda \cdot \mathbb{1}) &= \det \left[\begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \begin{vmatrix} a - \lambda & c/2 \\ c/2 & b - \lambda \end{vmatrix} \\ &= (a - \lambda)(b - \lambda) - (c/2)^2 \\ &= \lambda^2 - \lambda(a + b) + ab - (c/2)^2 \stackrel{!}{=} 0. \end{aligned} \quad (144)$$

- Lösen der (quadratischen) charakteristischen Gleichung (144):
Die beiden Lösungen sind die zur Matrix A gehörenden Eigenwerte λ_1 und λ_2 .
- Bestimmen der normierten Eigenvektoren \vec{v}_1 und \vec{v}_2 aus den zwei Gleichungssystemen

$$(A - \lambda_1 \cdot \mathbb{1}) \cdot \vec{r} = \begin{pmatrix} a - \lambda_1 & c/2 \\ c/2 & b - \lambda_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \vec{0} \Rightarrow \tilde{\vec{v}}_1 \Rightarrow \text{Normierung} \Rightarrow \vec{v}_1,$$

$$(A - \lambda_2 \cdot \mathbb{1}) \cdot \vec{r} = \begin{pmatrix} a - \lambda_2 & c/2 \\ c/2 & b - \lambda_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \vec{0} \Rightarrow \tilde{\vec{v}}_2 \Rightarrow \text{Normierung} \Rightarrow \vec{v}_2.$$

- Zusammensetzen der normierten Eigenvektoren zur Diagonalisierungs- bzw. Transformationsmatrix

$$S = (\vec{v}_1 \mid \vec{v}_2),$$

wobei die Orientierung der Eigenvektoren so gewählt werden muss, dass $\det S = +1$ ist.

Die zu symmetrischen (reellen) Matrizen A gehörenden Eigenvektoren \vec{v}_i stehen orthogonal aufeinander und sind normierte Basisvektoren. Sie bilden also in unserem zweidimensionalen Fall die **Orthonormalbasis** $\{\vec{v}_1, \vec{v}_2\}$.

- Mit der (orthogonalen) Transformationsmatrix $S \Rightarrow S^{-1} = S^T$ können wir jetzt die Diagonalmatrix $D_A = S^T A S$ berechnen, die den Koeffizienten des Kegelschnitts in Normalform entspricht und in der, wie bereits diskutiert wurde, der Koeffizient c verschwunden ist. Weiterhin stellen wir fest, dass die Hauptdiagonalelemente von D_A die Eigenwerte von A sind gemäß

$$D_A = \begin{pmatrix} \tilde{a} & 0 \\ 0 & \tilde{b} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Durch das Verschwinden des gemischten Terms cxy legen allein die neuen Koeffizienten $\tilde{a} = \lambda_1$ und $\tilde{b} = \lambda_2$ den quadratischen Anteil der ursprünglichen Gleichung (141) für das (ξ, η) -Koordinatensystem fest, denn dort erscheint der ursprüngliche Kegelschnitt infolge der Drehung mittels S jetzt in Normalform, dessen quadratische Terme jetzt nicht mehr durch die symmetrische Matrix A sondern durch die Diagonalmatrix D_A dargestellt werden.

- Wegen $\tilde{f} = f = \text{const}$ brauchen wir nur noch mit Hilfe der Transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = S \begin{pmatrix} \xi \\ \eta \end{pmatrix} = (\vec{v}_1 \mid \vec{v}_2) \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} x(\xi, \eta) \\ y(\xi, \eta) \end{pmatrix}$$

die Koordinaten x, y durch die Koordinaten ξ, η darzustellen und erhalten aus (141)

$$\lambda_1 \xi^2 + \lambda_2 \eta^2 + \underbrace{d x(\xi, \eta) + e y(\xi, \eta)}_{= \tilde{d} \xi + \tilde{e} \eta} + f = 0 \Rightarrow$$

$$\boxed{\lambda_1 \xi^2 + \tilde{d} \xi + \lambda_2 \eta^2 + \tilde{e} \eta = -f}. \quad (145)$$

Kegelschnitte, die sich im (ξ, η) -Koordinatensystem in achsenparalleler Lage bzw. in Normalform befinden, erfüllen die Gleichung (145).

¹Die Matrix $(A - \lambda \mathbb{1})$ ist die **charakteristische Matrix** von A und die Gleichung $\det(A - \lambda \cdot \mathbb{1}) \stackrel{!}{=} 0$ ist die **charakteristische Gleichung** mit dem charakteristischen Polynom $\det(A - \lambda \cdot \mathbb{1})$.

- Nach Ausklammern von λ_1 aus den ξ -abhängigen Termen und von λ_2 aus den η -abhängigen Termen lässt sich die Gleichung (145) durch

$$\text{quadratische Ergänzung von } \xi^2 + \frac{\tilde{d}}{\lambda_1}\xi \quad \text{und} \quad \eta^2 + \frac{\tilde{e}}{\lambda_2}\eta$$

sowie anschließender entsprechender Umformung auf die Form

$$\frac{(\xi - \xi_0)^2}{h_\xi^2} + \frac{(\eta - \eta_0)^2}{h_\eta^2} = 1 \quad \text{oder} \quad \pm \frac{(\xi - \xi_0)^2}{h_\xi^2} \mp \frac{(\eta - \eta_0)^2}{h_\eta^2} = 1$$

bringen, aus der man dann die Mittelpunktslage (ξ_0, η_0) sowie die Halbachsenlängen h_ξ und h_η ablesen kann.

- Durch die Koordinatentransformation $\vec{\varrho} \xrightarrow{S} \vec{r}$, d. h. durch

$$\begin{pmatrix} x \\ y \end{pmatrix} = S \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

lassen sich schließlich auch noch die Mittelpunkts- und Scheitelkoordinaten der Kegelschnitte im (x, y) -Koordinatensystem berechnen

17.1.2 Spezielle Ellipsen

Ellipsen in achsenparalleler Ursprungslage:

Diese befinden sich in der einfachsten Form und erfüllen die Gleichung

$$a x^2 + b y^2 + f = (x \ y) \underbrace{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}}_{A \text{ diagonal}} \begin{pmatrix} x \\ y \end{pmatrix} + f = \vec{r}^T D_A \vec{r} + f = 0 \quad (146)$$

mit der Diagonalmatrix $A = D_A$ und mit $\frac{a}{-f} = \frac{1}{h_x^2}$ sowie $\frac{b}{-f} = \frac{1}{h_y^2}$ schließlich die übliche Gleichung

$$\boxed{\frac{x^2}{h_x^2} + \frac{y^2}{h_y^2} = 1}$$

für eine Ellipse in ihrem **Hauptachsensystem**. Dabei ist h_x die Länge der Halbachse in x -Richtung und h_y ist die Länge der Halbachse in y -Richtung. Entsprechend der Länge von h_x und h_y spricht man von kleiner und großer Halbachse. Die Ellipsenachsen, manchmal auch Hauptachsen genannt, mit den Längen $2h_x$ und $2h_y$ stehen senkrecht aufeinander. Die kleine Ellipsenachse bezeichnet man als **Nebenachse** und die große Ellipsenachse als **Hauptachse**.

Ellipsen in achsenparalleler Lage und mit dem Mittelpunkt (x_0, y_0) :

Diese erfüllen die Gleichung

$$\frac{(x - x_0)^2}{h_x^2} + \frac{(y - y_0)^2}{h_y^2} = 1.$$

Ausmultiplizieren der Binome liefert

$$\begin{aligned} & \frac{1}{h_x^2} (x^2 - 2xx_0 + x_0^2) + \frac{1}{h_y^2} (y^2 - 2yy_0 + y_0^2) = 1 \\ \Rightarrow & \frac{1}{h_x^2} x^2 + \frac{1}{h_y^2} y^2 - \frac{2x_0}{h_x^2} x - \frac{2y_0}{h_y^2} y + \underbrace{\frac{x_0^2}{h_x^2} + \frac{y_0^2}{h_y^2} - 1}_{=0} = 0 \quad \Rightarrow \\ & a x^2 + b y^2 + d x + e y + f = 0. \end{aligned}$$

Wie wir sehen, stehen die linearen Terme dx und ey für die Verschiebung des Mittelpunkts der Ellipse aus dem Koordinatenursprung.

Ellipsen in nicht achsenparalleler Ursprungslage (schräge Ursprungsellipsen):

Diese erfüllen die Gleichung

$$a x^2 + b y^2 + c xy + f = (x \ y) \underbrace{\begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix}}_{A \text{ symmetrisch}} \begin{pmatrix} x \\ y \end{pmatrix} + f = \vec{r}^T A \vec{r} + f = 0. \quad (147)$$

Der gemischte Term cxy steht für die Verdrehung der Ellipse im (x, y) -Koordinatensystem. Die Ellipse ist folglich nicht achsenparallel. Das Fehlen der linearen Terme dx und ey zeigt, dass sich die Ellipse in Ursprungslage befindet.

17.1.3 Eigenwerte und Eigenvektoren zur Matrix A

Wir berechnen jetzt die Eigenwerte λ_i und die Eigenvektoren $\tilde{\vec{v}}_i$ sowie die normierten Eigenvektoren \vec{v}_i der reellen **symmetrischen Matrix**

$$A \Rightarrow \left(\begin{array}{cc} a & c/2 \\ c/2 & b \end{array} \right)$$

über die **charakteristische Gleichung** (144)

$$\lambda^2 - \lambda(a+b) + ab - (c/2)^2 \stackrel{!}{=} 0$$

und das **homogene lineare Gleichungssystem**

$$\begin{pmatrix} a-\lambda & c/2 \\ c/2 & b-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \vec{0} \quad \Rightarrow \quad \begin{array}{l} \text{I.} \quad (a-\lambda)x + \frac{c}{2}y = 0, \\ \text{II.} \quad \frac{c}{2}x + (b-\lambda)y = 0. \end{array} \quad (148)$$

Für lineares Gleichungssystem schreiben wir im Folgenden kurz LGS. Mit Hilfe der sog. pq -Formel

$$\lambda_{1,2} = -\frac{p}{2} \mp \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

zur Lösung quadratischer Gleichungen und mit den Ersetzungen

$$-(a+b) = p \quad \text{und} \quad ab - \left(\frac{c}{2}\right)^2 = q$$

erhalten wir die **Eigenwerte**

$$\begin{aligned} \lambda_1 &= \frac{a+b}{2} - \sqrt{\left(\frac{a+b}{2}\right)^2 - ab + \left(\frac{c}{2}\right)^2} = \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right], \\ \lambda_2 &= \frac{a+b}{2} + \sqrt{\left(\frac{a+b}{2}\right)^2 - ab + \left(\frac{c}{2}\right)^2} = \frac{1}{2} \left[(a+b) + \sqrt{(a-b)^2 + c^2} \right]. \end{aligned}$$

Mit diesen Eigenwerten gehen wir in das homogene LGS (148), dessen Lösungen dann die Eigenvektoren sind.

Hinsichtlich **homogener LGS** müssen wir dabei folgendes berücksichtigen:

- Homogene LGS besitzen (immer) die triviale bzw. Nullvektor-Lösung.
- Ist die Anzahl der Variablen in der zu einem homogenen LGS gehörenden Koeffizientenmatrix gleich ihrem Rang, so ist der Nullvektor die einzige Lösung.
- Triviale Lösungen bzw. Nullvektoren kommen als Eigenvektoren nicht in Betracht: $\tilde{\vec{v}}_i \stackrel{!}{\neq} \vec{0}$.
- Ist der Rang der zu einem homogenen LGS gehörenden Koeffizientenmatrix kleiner als die Anzahl der Variablen, so existieren **unendlich viele Lösungen**, die als Eigenvektoren in Betracht kommen. Eine derartige Matrix ist **singulär** (nicht invertierbar) und ihre Zeilen sind linear abhängig. Ein zu einer singulären Koeffizientenmatrix gehörendes *homogenes* LGS besitzt also außer der „obligatorischen“ trivialen Lösung immer noch mehr als eine Lösung.
- Das bedeutet:
Das zur Koeffizientenmatrix gehörende homogene LGS darf nicht eindeutig lösbar sein, um Eigenvektoren zu liefern. Alle Vielfachen von Eigenvektoren sind ebenfalls Eigenvektoren.
Die Ursache dafür ist, dass das charakteristische Polynom, also die Determinante der charakteristischen Matrix zur Bestimmung der zu den Eigenvektoren gehörenden Eigenwerte, gleich Null gesetzt wurde. Ist aber die Determinante einer Matrix gleich Null, so ist diese Matrix singulär.
- Reelle symmetrische $(n \times n)$ -Matrizen und allgemein auch hermitesche $(n \times n)$ -Matrizen besitzen nur reelle Eigenwerte und genau n **orthogonale Eigenvektoren** zu verschiedene Eigenwerten.

Bestimmung der Eigenvektoren zur reellen symmetrischen Matrix A

$$A = \begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix}, \quad \lambda_1 = \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right], \quad \lambda_2 = \frac{1}{2} \left[(a+b) + \sqrt{(a-b)^2 + c^2} \right].$$

Wir verwenden jetzt den Gauß-Algorithmus und schreiben dabei (148) wegen der besseren Übersicht als (zweizeilige) erweiterte Koeffizientenmatrix

$$\begin{pmatrix} \text{I} \\ \text{II} \end{pmatrix} := \left(\begin{array}{cc|c} a - \lambda & c/2 & 0 \\ c/2 & b - \lambda & 0 \end{array} \right).$$

Bestimmung von $\tilde{v}_1(\lambda_1)$ aus

$$\left(\begin{array}{cc|c} a - \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] & c/2 & 0 \\ c/2 & b - \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] & 0 \end{array} \right) :$$

Wir dividieren die Zeile I durch $a - \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right]$:

$$\left(\begin{array}{cc|c} 1 & c / \left\{ 2a - \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] \right\} & 0 \\ c/2 & b - \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] & 0 \end{array} \right). \quad (149)$$

Von Zeile II subtrahieren wir jetzt

$$\frac{c}{2} \cdot \text{Zeile I von (149)} = \left(\begin{array}{cc|c} \frac{c}{2} & \frac{c^2}{4a - 2 \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right]} & 0 \end{array} \right)$$

und erhalten schließlich die mit Hilfe des Gauß-Algorithmus vereinfachte erweiterte singuläre Koeffizientenmatrix

$$\left(\begin{array}{cc|c} 1 & c / \left\{ 2a - \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] \right\} & 0 \\ 0 & 0 & 0 \end{array} \right).$$

Wegen (144) muss im letzten Schritt tatsächlich

$$\left\{ b - \frac{1}{2} \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] \right\} - \frac{c^2}{4a - 2 \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right]} = 0$$

resultieren, was wir durch akribisches Umformen und Ausmultiplizieren mit dem Zwischenergebnis

$$\begin{aligned} c^2 &= 4ab - 2a \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] \\ &\quad - 2b \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right] + \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right]^2 \end{aligned}$$

zeigen können:

$$c^2 = \underbrace{4ab - 2(a+b)^2 + (a+b)^2 + (a-b)^2}_{=0} + c^2 = c^2. \quad \square$$

Wir dürfen jetzt eine der beiden Variablen x_1 und y_1 frei wählen und erhalten mit $y_1 \stackrel{!}{=} 1$ aus der Zeile I

$$1 \cdot x_1 + \frac{c}{2a - \left[(a+b) - \sqrt{(a-b)^2 + c^2} \right]} \cdot y_1 = 0$$

$$\xrightarrow{y_1 \stackrel{!}{=} 1} x_1 = \frac{c}{b - a - \sqrt{(a-b)^2 + c^2}}.$$

Damit ist der nicht normierte Eigenvektor zum Eigenwert λ_1 :

$$\boxed{\tilde{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \frac{c}{(b-a) - \sqrt{(a-b)^2 + c^2}} \\ 1 \end{pmatrix}}.$$

Völlig analog ist die **Bestimmung von $\tilde{\mathbf{v}}_2(\lambda_2)$** aus

$$\begin{pmatrix} a - \frac{1}{2} \left[(a+b) + \sqrt{(a-b)^2 + c^2} \right] & c/2 \\ c/2 & b - \frac{1}{2} \left[(a+b) + \sqrt{(a-b)^2 + c^2} \right] \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} :$$

$$\boxed{\tilde{\mathbf{v}}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{c}{(b-a) + \sqrt{(a-b)^2 + c^2}} \\ 1 \end{pmatrix}} .$$

Durch Normierung der Eigenvektoren $\tilde{\mathbf{v}}_i$ gemäß

$$\mathbf{v}_i = \frac{\tilde{\mathbf{v}}_i}{|\tilde{\mathbf{v}}_i|} = \frac{\tilde{\mathbf{v}}_i}{\sqrt{x_i^2 + y_i^2}}$$

erhalten wir die **normierten Eigenvektoren**

$$\mathbf{v}_1 = \frac{1}{\sqrt{1 + \left[\frac{c}{(b-a) - \sqrt{(a-b)^2 + c^2}} \right]^2}} \begin{pmatrix} \frac{c}{(b-a) - \sqrt{(a-b)^2 + c^2}} \\ 1 \end{pmatrix}, \quad (150)$$

$$\mathbf{v}_2 = \frac{1}{\sqrt{1 + \left[\frac{c}{(b-a) + \sqrt{(a-b)^2 + c^2}} \right]^2}} \begin{pmatrix} \frac{c}{(b-a) + \sqrt{(a-b)^2 + c^2}} \\ 1 \end{pmatrix}. \quad (151)$$

17.1.4 Bestimmung der Halbachsenlängen schräger Ursprungsellipsen

Zunächst überführen wir die Gleichung $ax^2 + by^2 + cxy + f = \tilde{\mathbf{r}}^T A \tilde{\mathbf{r}} + f = 0$ für eine schräge Ursprungsellipse durch Diagonalisierung von A , also durch

$$A = \begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix} \xrightarrow{\text{Diagonalisierung}} D_A = \begin{pmatrix} \tilde{a} & 0 \\ 0 & \tilde{b} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

in die Gleichung

$$\tilde{a}x^2 + \tilde{b}y^2 + f = \tilde{\mathbf{r}}^T D_A \tilde{\mathbf{r}} + f = \lambda_1 x^2 + \lambda_2 y^2 + f = 0$$

für die entsprechende achsenparallele Ursprungsellipse. Aus dieser Gleichung erhalten wir durch Äquivalenzumformung die übliche Form der Ellipsengleichung

$$\frac{x^2}{\frac{-f}{\lambda_1}} + \frac{y^2}{\frac{-f}{\lambda_2}} = \frac{x^2}{h_x^2} + \frac{y^2}{h_y^2} = 1, \quad f \neq 0$$

mit den Halbachsenlängen

$$h_x = \sqrt{\frac{-f}{\lambda_1}} \quad \text{und} \quad h_y = \sqrt{\frac{-f}{\lambda_2}} .$$

17.1.5 Bestimmung des Winkels α schräger Ursprungsellipsen

Die Ellipsengleichung (147)

$$ax^2 + by^2 + cxy + f = (x, y) \underbrace{\begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix}}_A \begin{pmatrix} x \\ y \end{pmatrix} + f = 0$$

mit der **symmetrischen** Matrix A bildet eine schräge (nicht achsenparallele) Ellipse in Ursprungslage auf das (x, y) -Koordinatensystem bzw. deren Standardbasis $\{\tilde{\mathbf{e}}_x, \tilde{\mathbf{e}}_y\}$ ab. Zwischen der (großen) Hauptachse der schrägen (ursprünglichen) Ellipse und der x -Achse des (x, y) -Koordinatensystem liegt dann der Winkel α ,

den wir aus den Koeffizienten bzw. Parametern der schrägen Ellipse berechnen wollen. Dafür benötigen wir die Drehmatrix S .

Die (passive) Drehung des kartesischen (x, y) -Koordinatensystems K um den Winkel α im mathematisch **positiven** Drehsinn gemäß

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = S^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

liefert das gedrehte (x', y') -Koordinatensystem K' und ist gleichbedeutend mit der (aktiven) Drehung des Ortsvektors $\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ innerhalb des (x, y) -Koordinatensystems K um den Winkel α im mathematisch **negativen** Drehsinn. Die Umkehroperation dazu bzw. die Drehung in umgekehrter Richtung wird beschrieben durch die Drehmatrix

$$S = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \Rightarrow S^T = S^{-1}.$$

Bei der Diagonalisierung von A gemäß (143)

$$S^T A S = D_A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

mit der (orthogonalen) Drehmatrix

$$S = (\vec{v}_1 \mid \vec{v}_2),$$

deren Spalten die **orthonormalen** Eigenvektoren \vec{v}_1 und \vec{v}_2 zu A sind, resultiert die Diagonalmatrix D_A , die der Abbildung der ursprünglichen Ellipse in ihrem (x, y) -**Hauptachsensystem** entspricht. Anders gesagt, mit Hilfe der Drehmatrix S wird die ursprünglich schräge Ellipse um den Winkel α so gedreht, dass sie schließlich im (x, y) -Koordinatensystem achsenparallel liegt. Der Übergang von der schrägen zur achsenparallelen Ellipse durch Drehung um den Winkel α stellt sich ausgehend von (147) wie folgt dar:

$$(x, y) A \begin{pmatrix} x \\ y \end{pmatrix} \longrightarrow (x, y) S^T A S \begin{pmatrix} x \\ y \end{pmatrix} = (x, y) D_A \begin{pmatrix} x \\ y \end{pmatrix}.$$

Das aber bedeutet, dass wir den Drehwinkel α aus der Drehmatrix

$$S = (\vec{v}_1 \mid \vec{v}_2) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \Rightarrow \vec{v}_1 = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$$

ermitteln können, denn die normierten Eigenvektoren \vec{v}_1 und \vec{v}_2 zur Matrix A hatten wir bereits mit (150) und (151) berechnet.

Anmerkungen

Die Vertauschung der normierten Eigenvektoren bzw. der Spalten in der Drehmatrix S bewirkt eine Vertauschung der Eigenwerte in der Diagonalmatrix D_A und damit eine Änderung der Schrägstellung der Ellipse um 90° , also eine Vertauschung der Richtung von Haupt- und Nebenachse. Weil von vornherein aber nicht klar ist, welcher der beiden Eigenvektoren die Richtung der Hauptachse und welcher die Richtung der Nebenachse besitzt und da man der Ellipsengleichung (147) die Richtung der Ellipsenachsen nicht unbedingt ansieht, kann es ratsam sein, sich diesbezüglich vorab Klarheit zu verschaffen.

Dafür bestimmen wir in (147) die beiden y -Werte für einen möglichst kleinen festen Wert $x_0 > 0$, indem wir $ax_0^2 + by^2 + cx_0y + f = 0$ in die quadratische Gleichung $y^2 + py + q = 0$ umformen und diese mittels der sog. pq -Formel lösen. Wir erhalten die beiden Lösungen $y_+(x_0) > 0$ und $y_-(x_0) < 0$, die uns zeigen, welche der Ellipsenachsen durch den ersten Quadranten verläuft:

$$|y_+| > |y_-| \Rightarrow \text{(große) Hauptachse liegt im 1. Quadranten,}$$

$$|y_+| < |y_-| \Rightarrow \text{Nebenachse liegt im 1. Quadranten.}$$

Dabei stellen wir fest, dass der kleinere der beiden Eigenwerte λ den Eigenvektor liefert, welcher die Hauptachse beschreibt. Wählen wir beispielsweise $\lambda_1 < \lambda_2$, so beschreibt der zu λ_1 gehörende Eigenvektor \vec{v}_1 die Richtung der Hauptachse und der zu λ_2 gehörende Eigenvektor \vec{v}_2 die Richtung der Nebenachse.

17.2 Zu Abschnitt 7.3 – Berechnung von E_{0x} und E_{0y}

Achtung!

Die Länge der Halbachsen bezeichnen wir in diesem Abschnitt wie allgemein üblich mit a und b . Abweichend davon hatten wir in der Abbildung 13 (wegen der Notation bei der Hauptachsentransformation im Abschnitt 17.1) die Länge der großen Halbachse mit H und die Länge der kleinen Halbachse mit h bezeichnet.

Das (x', y') -Koordinatensystem K' sei das Hauptachsensystem der Ellipse

$$\mathcal{E}' := \frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1 \quad \Leftrightarrow \quad \begin{cases} x' = x'(y') = \frac{a}{b} \sqrt{b^2 - y'^2}, \\ y' = y'(x') = \frac{b}{a} \sqrt{a^2 - x'^2}. \end{cases} \quad (152)$$

K' sei gegenüber dem (x, y) -Koordinatensystem K um den gemeinsamen Koordinatenursprung im mathematisch positiven Sinn um den Winkel α gedreht, sodass \mathcal{E}' in K als die schräge Ursprungsellipse \mathcal{E} erscheint. Die Substitution der K' -Koordinaten durch die zugehörigen K -Koordinaten in (152) liefert dann die Gleichung von \mathcal{E} in K . Die dafür erforderlichen in (x, y) -Koordinaten ausgedrückten (x', y') -Koordinaten erhalten wir mit der Drehmatrix S^{-1} gemäß

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = S^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \alpha + y \sin \alpha \\ -x \sin \alpha + y \cos \alpha \end{pmatrix} = \begin{pmatrix} x'(x, y) \\ y'(x, y) \end{pmatrix}.$$

Damit resultiert aus (152)

$$\mathcal{E} := \frac{1}{a^2} (x \cos \alpha + y \sin \alpha)^2 + \frac{1}{b^2} (-x \sin \alpha + y \cos \alpha)^2 = 1 = F(x, y). \quad (153)$$

Die Extrema (Index E) der Ellipse \mathcal{E} in K sind

$$y_E = y(x_0) = \pm |y_E| \quad \text{an den Stellen} \quad x_0 = \pm |x_0| \quad \text{oder} \quad x_0 = \mp |x_0|,$$

$$x_E = x(y_0) = \pm |x_E| \quad \text{an den Stellen} \quad y_0 = \pm |y_0| \quad \text{oder} \quad y_0 = \mp |y_0|.$$

Diese bestimmen wir, indem wir zunächst (153) implizit nach x bzw. nach y differenzieren und anschließend $\frac{dy}{dx}$ bzw. $\frac{dx}{dy}$ gleich Null setzen. Dabei verwenden wir

$$\frac{d}{dx} F(x, y) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{dy}{dx} = 0 \quad \Leftrightarrow \quad \frac{dy}{dx} = -\frac{F_x}{F_y} \quad \text{und analog} \quad \frac{dx}{dy} = -\frac{F_y}{F_x} :$$

$$\begin{aligned} \frac{dy}{dx} &= -\frac{\frac{1}{a^2} 2(x \cos \alpha + y \sin \alpha) \cos \alpha + \frac{1}{b^2} (-x \sin \alpha + y \cos \alpha)(-\sin \alpha)}{\frac{1}{a^2} 2(x \cos \alpha + y \sin \alpha) \sin \alpha + \frac{1}{b^2} (-x \sin \alpha + y \cos \alpha) \cos \alpha} \stackrel{!}{=} 0 \\ &\Rightarrow \quad x = x_0 \quad \Rightarrow \quad y(x_0) = y_E \quad \Rightarrow \end{aligned}$$

$$\frac{1}{a^2} 2(x_0 \cos \alpha + y_E \sin \alpha) \cos \alpha + \frac{1}{b^2} 2(-x_0 \sin \alpha + y_E \cos \alpha)(-\sin \alpha) = 0 \quad \left| \cdot \frac{a^2 b^2}{2} \right.$$

$$b^2 (x_0 \cos \alpha + y_E \sin \alpha) \cos \alpha - a^2 (-x_0 \sin \alpha + y_E \cos \alpha) \sin \alpha = 0 \quad \Leftrightarrow$$

$$x_0 (a^2 \sin^2 \alpha + b^2 \cos^2 \alpha) = y_E \sin \alpha \cos \alpha (a^2 - b^2),$$

$$\boxed{x_0 = y_E \underbrace{\frac{(a^2 - b^2) \sin \alpha \cos \alpha}{a^2 \sin^2 \alpha + b^2 \cos^2 \alpha}}_{k_x} = k_x \cdot y_E}.$$

Dies setzen wir in die Ellipsengleichung (153) für $F(x, y) \longrightarrow F(x_0, y_E)$ ein und erhalten

$$\begin{aligned} \frac{1}{a^2} \left(k_x y_E \cos \alpha + y_E \sin \alpha \right)^2 + \frac{1}{b^2} \left(-k_x y_E \sin \alpha + y_E \cos \alpha \right)^2 &= 1, \\ y_E^2 \frac{(k_x \cos \alpha + \sin \alpha)^2}{a^2} + y_E^2 \frac{(\cos \alpha - k_x \sin \alpha)^2}{b^2} &= 1 \quad \Leftrightarrow \\ y_E^2 &= \frac{a^2 b^2}{b^2 (k_x \cdot \cos \alpha + \sin \alpha)^2 + a^2 (\cos \alpha - \sin \alpha \cdot k_x)^2}, \\ y_E^2 &= \frac{1}{\frac{1}{a^2} \left[\frac{(a^2 - b^2) \sin \alpha \cos \alpha}{a^2 \sin^2 \alpha + b^2 \cos^2 \alpha} \cdot \cos \alpha + \sin \alpha \right]^2 + \frac{1}{b^2} \left[\cos \alpha - \sin \alpha \cdot \frac{(a^2 - b^2) \sin \alpha \cos \alpha}{a^2 \sin^2 \alpha + b^2 \cos^2 \alpha} \right]^2}, \end{aligned} \quad (154)$$

$$y_E = \frac{1}{\sqrt{\frac{1}{a^2} \left[\frac{(a^2 - b^2) \sin \alpha \cos^2 \alpha}{a^2 \sin^2 \alpha + b^2 \cos^2 \alpha} + \sin \alpha \right]^2 + \frac{1}{b^2} \left[\cos \alpha - \frac{(a^2 - b^2) \sin^2 \alpha \cos \alpha}{a^2 \sin^2 \alpha + b^2 \cos^2 \alpha} \right]^2}}.$$

y_E entspricht dabei E_{0y} . Wie man sofort sieht, gilt wie erwartet

$$\alpha = 0^\circ \Rightarrow y_E = \pm b, \quad \alpha = 90^\circ \Rightarrow y_E = \pm a.$$

Völlig analog erhalten wir aus dem Ansatz $\frac{dx}{dy} \stackrel{!}{=} 0 \Rightarrow F(x, y) \longrightarrow F(x_E, y_0)$ und $x(y_0) = x_E$ mit

$$y_0 = x_E \underbrace{\frac{(a^2 - b^2) \sin \alpha \cos \alpha}{a^2 \cos^2 \alpha + b^2 \sin^2 \alpha}}_{k_y} = k_y \cdot x_E$$

schließlich

$$x_E^2 = \frac{1}{\frac{1}{a^2} \left[\cos \alpha + \sin \alpha \cdot \frac{(a^2 - b^2) \sin \alpha \cos \alpha}{a^2 \cos^2 \alpha + b^2 \sin^2 \alpha} \right]^2 + \frac{1}{b^2} \left[\frac{(a^2 - b^2) \sin \alpha \cos \alpha}{a^2 \cos^2 \alpha + b^2 \sin^2 \alpha} \cdot \cos \alpha - \sin \alpha \right]^2}, \quad (155)$$

$$x_E = \frac{1}{\sqrt{\frac{1}{a^2} \left[\cos \alpha + \frac{(a^2 - b^2) \sin^2 \alpha \cos \alpha}{a^2 \cos^2 \alpha + b^2 \sin^2 \alpha} \right]^2 + \frac{1}{b^2} \left[\frac{(a^2 - b^2) \sin \alpha \cos^2 \alpha}{a^2 \cos^2 \alpha + b^2 \sin^2 \alpha} - \sin \alpha \right]^2}}.$$

x_E entspricht dabei E_{0x} . Wie man sofort sieht, gilt wie erwartet

$$\alpha = 0^\circ \Rightarrow x_E = \pm a, \quad \alpha = 90^\circ \Rightarrow x_E = \pm b.$$

Weiterhin gilt

$$a^2 + b^2 = x_E^2 + y_E^2,$$

was man mit (154) und (155) beispielsweise für $\alpha = 45^\circ \Rightarrow \sin 45^\circ = \cos 45^\circ = \frac{1}{\sqrt{2}}$ leicht zeigen kann, denn in diesem Fall gilt

$$x_E^2 + y_E^2 = \frac{1}{2}(a^2 + b^2) + \frac{1}{2}(a^2 + b^2) = a^2 + b^2.$$

Und für $a = 4$, $b = 3$, $\alpha = 30^\circ \Rightarrow \sin 30^\circ = \frac{1}{2}$, $\cos 30^\circ = \frac{\sqrt{3}}{2}$ erhalten wir

$$x_E^2 = 14, 25, \quad y_E^2 = 10, 75 \quad \Rightarrow \quad x_E^2 + y_E^2 = 25 = a^2 + b^2.$$