

definieren wie die Faltung als

$$f \star g = \int \dots \int_{\mathbb{R}^n} f(x_1 - y_1, \dots, x_n - y_n) g(y_1, \dots, y_n) dy_1 \dots dy_n$$

– sofern dieses Integral existiert.

Auch die anschauliche Interpretation ist dieselbe wie im eindimensionalen Fall: Wenn wir f als eine Gewichtsfunktion auffassen, ist $f \star g$ ein gewichtetes Mittel über Werte von g ; für

$$f(x_1, \dots, x_n) = \frac{1}{\pi^{n/2} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n x_k^2}$$

etwa, die n -dimensionale GAUSS-Funktion, entspricht das im Fall $n = 2$ einem je nach Größe von σ mehr oder weniger defokussierten Bild.

Durch mehrdimensionale Faltungen mit δ -Distributionen lassen sich Verschiebungen realisieren: Beispielsweise wäre, wenn der Satz von FUBINI in einer solchen Situation anwendbar wäre,

$$\iint_{\mathbb{R}^2} \delta(x - a) \delta(y - b) f(x, y) dx dy = f(x - a, y - b),$$

und genau so *definieren* wir die Interpretation der *a priori* sinnlosen linken Seite.

(Man beachte, daß Ausdrücke wie $\delta(x - a) \delta(x - b)$ oder $\delta(x)^2$ weiterhin sinnlos bleiben, egal ob sie unter einem oder mehreren Integralzeichen stehen.)

Ist also $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Funktion, die (z.B. durch Grauwerte) ein Bild definiert und die außerhalb des Bereichs $0 \leq x, y \leq 1$ verschwindet, so ist mit der Distribution

$$\eta(x, y) = \sum_{k=1}^N \sum_{\ell=1}^M \delta(x - k) \delta(y - \ell)$$

die Faltung $\eta \star f$ ein Bilderbogen aus NM Exemplaren dieses Bildes. Abbildung 27 zeigt dies für den Graph einer zweidimensionalen Normalverteilung.

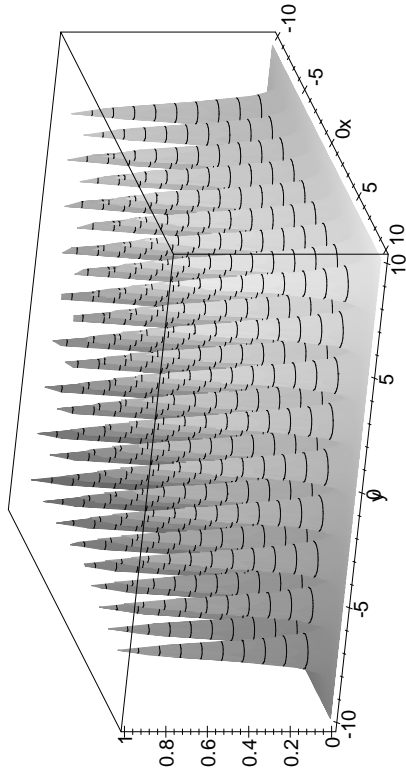


Abb. 27: Eine zweidimensionale Faltung

Auch die FOURIER-Transformation läßt sich in völliger Analogie zum eindimensionalen Fall auf beliebige Dimensionen verallgemeinern: Für $f: \mathbb{R}^n \rightarrow \mathbb{C}$ definieren wir

$$\mathbb{R}^n \rightarrow \mathbb{C} \\ \hat{f}: \left\{ \begin{aligned} (\omega_1, \dots, \omega_n) &\mapsto \int \dots \int_{\mathbb{R}^n} f(x_1, \dots, x_n) e^{-i \sum_{k=1}^n \omega_k x_k} dx_1 \dots dx_n \end{aligned} \right.$$

Da es nur eine Zeit gibt, läßt sich dies nicht als Zerlegung eines zeitlichen Signals in seine Frequenzen interpretieren; die x_i sollte man sich hier als *räumliche* Koordinaten vorstellen. Beispiele dazu folgen im nächsten Abschnitt, wo wir eine Anwendung solcher räumlicher FOURIER-Transformationen betrachten.

Wenigsten kurz sei noch angedeutet, wie man auch die mehrdimensionale FOURIER-Theorie über stark abfallende Funktionen mehrerer Veränderlicher exakt begründen kann:

Eine Funktion $\varphi: \mathbb{R}^n \rightarrow \mathbb{C}$ heißt *stark abfallend*, wenn *alle* Ausdrücke der Form

$$x_1^{e_1} \dots x_n^{e_n} \frac{\partial^{r_1 + \dots + r_n}}{\partial x_1^{r_1} \dots \partial x_n^{r_n}} \varphi(x_1, \dots, x_n)$$

auf ganz \mathbb{R}^n beschränkt sind. Der Vektorraum aller dieser Funktionen ist der SCHWARTZ-Raum $\mathcal{S}(\mathbb{R}^n)$.

Für Funktionen aus diesem Raum ist wieder alles relativ problemlos; zur Verallgemeinerungen auf interessantere Funktionen führt auch hier der Umweg über Distributionen

$$T: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C},$$

die in der naheliegenden Weise als Verallgemeinerungen eindimensionaler Distributionen definiert werden. Beispielsweise kann man dem gerade *ad hoc* betrachteten Produkt $\delta(x-a)\delta(y-b)$ über die Distribution

$$\Delta_{a,b}: \begin{cases} \mathbb{R}^2 \rightarrow \mathbb{C} \\ \varphi \mapsto \varphi(a,b) \end{cases}$$

einen präzisen Sinn geben – solange es in einem sinnvollen Kontext unter zwei Integralzeichen steht.

b) Fraunhofer-Beugung

Wenn Licht auf Strukturen trifft, in Vergleich zu deren Größe seine Wellenlänge nicht mehr vernachlässigbar klein ist, lassen sich die Gesetze der geometrischen Optik bekanntlich nicht mehr anwenden; man beobachtet dann Beugungsphänomene.

Beugung ist ein sehr komplexes Gebiet; für ein Beispiel im Rahmen einer Vorlesung über Höhere Mathematik müssen wir uns auf den allerinfachsten Fall beschränken. Wir gehen daher aus von einem Lichtstrahl, der aus sehr großer Entfernung kommt oder der zumindest (z.B. dank einer Linse, aus deren Brennpunkt er kommt) so aussieht, und beobachten auch die Beugungsfigur in großer Entfernung. Diese Situation bezeichnet man als FRAUNHOFER-Beugung.



JOSEPH VON FRAUNHOFER (1787–1826) wurde in Straubing als elftes und letztes Kind eines Glasermeisters geboren; er machte auch selbst eine Lehre als Glas-schleifer und Spiegelmacher. Daneben besuchte er die Feterabendschule, wo er zumindest primitive Grundkenntnisse im Rechnen erwarb. 1806 kam er an das optische Institut von UTZSCHNEIDER, der ihm Bücher über Optik und Mathematik besorgte. FRAUNHOFER entwickelte Präzisionsmaschinen zur Fertigung optischer Instrumente von bis dahin nicht gekannter Qualität und er fand auch das optische Gitter. Durch seine Versuche zur Lichtbeugung bewies er die Wellennatur des Lichts. 1824 wurde er zum Professor ernannt; er berichtete unter anderem in öffentlichen Sonntagsvorlesungen über seine Arbeit. Im gleichen Jahr wurde er vom bayrischen König LUDWIG I. in den Adelsstand erhoben. Zwei Jahre später starb er an Tuberkulose.

Zur mathematischen Behandlung der optischen Beugung brauchen wir zunächst ein physikalisches Modell für Lichtwellen. Für eine physikalisch korrekte Beschreibung müssen wir Licht als zeitlich veränderliches elektromagnetisches Feld betrachten, d.h. wir brauchen zwei räumlich und zeitlich variable Vektorfelder $\vec{E}(x, y, z; t)$ und $\vec{B}(x, y, z; t)$, die den MAXWELLSchen Gleichungen genügen. Glücklicherweise muß man in der Optik aber nur selten so weit gehen: Zwar hängt die Beugung an einem Spalt theoretisch durchaus von der Leitfähigkeit des verwendeten Materials ab, aber diese Abhängigkeit ist so gering, daß man sie für alle praktischen Zwecke vernachlässigen kann.

Man arbeitet daher in der Wellenoptik gerne mit einer sogenannten *skalaren Welle*, über deren physikalische Bedeutung man sich keine sonderlichen Gedanken macht. Aus rechnerischen Gründen betrachtet man sie als komplexwertige Funktion; falls man sich unbedingt etwas darunter vorstellen will, kann man beispielsweise den Realteil dieser Funktion als die x -Komponente des elektrischen Felds interpretieren, muß dann aber beachten, daß eine skalare Welle im Gegensatz zu einem elektrischen Feld *keine* Wechselwirkung mit Materie egal welcher Leitfähigkeit zeigt – das ist eine der Idealisierungen hinter dem Konzept der skalaren Welle. Wichtig für uns ist nur, daß die Intensität der Welle (also z.B. die Intensität der Beugungslinien, die wir auf einem Schirm beobachten) gleich dem Betragsquadrat der Wellenfunktion sein soll.

Eine Welle hat eine räumliche wie auch zeitliche Periodizität. Zeitlich periodische Vorgänge kennen wir bereits: Das sind Schwingungen, die mathematisch durch Funktionen der Art

$$f(t) = A_0 e^{i\omega t} \quad \text{oder etwas allgemeiner} \quad f(t) = A_0 e^{i(\omega t + \varphi)}$$

beschrieben werden, wobei die Phasenverschiebung φ dafür sorgt, daß wir auch Schwingungen behandeln können, die ihre maximale Auslenkung nicht zur Zeit $t = 0$ erreichen. Wir wollen dies jedoch im folgenden ignorieren und mit der einfacheren ersten Funktion arbeiten.

Für räumlich periodische Vorgänge haben wir entsprechend zur Periode T einer Schwingung eine Wellenlänge λ ; das Analogon zur Kreisfrequenz $\omega = 2\pi/T$ bezeichnen wir als

$$\text{Wellenzahl} \quad k = \frac{2\pi}{\lambda}.$$

Ein eindimensionaler periodischer Vorgang kann somit beschrieben werden durch eine Funktion $g(x) = A_0 e^{ikx}$.

Im Mehrdimensionalen müssen wir die Wellenzahl k ersetzen durch einen Vektor \vec{k} der Länge k , den *Wellenzahlvektor*; und betrachten die Funktion

$$g(\mathbf{x}) = A_0 e^{i\vec{k} \cdot \vec{x}}.$$

(Die Wellenlänge betrachten wir weiterhin nur als Skalar.)

Eine Welle soll zeitlich *und* räumlich periodisch sein; dies leistet die Funktion

$$\psi(\mathbf{x}, t) = A_0 e^{i(\omega t - \vec{k} \cdot \vec{x})}$$

oder natürlich auch die entsprechende Funktion mit einem Pluszeichen im Exponenten. Der Grund, warum wir das Minuszeichen bevorzugen, ist folgender:

Im eindimensionalen Fall ist

$$\psi(x, t) A_0 e^{i(\omega t - kx)} = A_0 e^{ik(\frac{\omega}{k}t - x)},$$

$\psi(x, t)$ hängt also nur ab von $x - \frac{\omega}{k}t$. Dies können wir auch so interpretieren, daß

$$v = \frac{\omega}{k} = \frac{\lambda}{T} = \frac{\lambda\omega}{2\pi}$$

die Ausbreitungsgeschwindigkeit der Welle ist; denn eine Änderung der Zeit um Δt hat denselben Effekt wie eine Änderung des Orts um $v \cdot \Delta t$.

Im Falle mehrerer räumlicher Dimensionen ist alles grundsätzlich genauso, nur die Schreibweise ist etwas komplizierter: Ist \vec{k}_0 ein Einheitsvektor in Richtung von \vec{k} , d.h. $\vec{k} = k \cdot \vec{k}_0$, so ist

$$\psi(\mathbf{x}, t) = A_0 e^{ik(\frac{\omega}{k}t - \vec{k}_0 \cdot \vec{x})};$$

dabei ist $\vec{k}_0 \cdot \vec{x}$ die \vec{x} -Komponente in Richtung von \vec{k} . Damit ist \vec{k}_0 die *Richtung* des Geschwindigkeitsvektors; der Wellenzahlvektor zeigt also in Richtung der Ausbreitungsgeschwindigkeit der Welle, und der Betrag v des Geschwindigkeitsvektors ist durch obige Formel gegeben.

Die Annahme einer konstanten Amplitude A_0 in obigen Formeln ist nur in seltenen Fällen realistisch: Licht kommt meist aus einer (zumindest in erster Näherung) punktförmigen Lichtquelle, und seine Intensität nimmt mit dem Quadrat der Entfernung ab. Da die Intensität das Betragsquadrat der Wellenfunktion sein soll, müssen wir eine solche Kugelwelle also in der Form

$$\psi(\mathbf{x}, t) = \frac{A_0}{|\vec{x}|} e^{i(\omega t - \vec{k} \cdot \vec{x})}$$

ansetzen, sofern die Lichtquelle im Nullpunkt des Koordinatensystems sitzt.

Im Falle einer weit entfernten Lichtquelle, wie wir sie bei der FRAUNHOFER-Beugung annehmen und auch von der Sonne her kennen, spielt allerdings die Ortsabhängigkeit der Amplitude praktisch keine Rolle, so daß wir keinen nennenswerten Fehler machen, wenn wir sie als konstant annehmen. In diesem Fall sprechen wir von einer *ebenen* Welle.

Ausgangspunkt für die Berechnung von Beugungsbildern ist das HUYGENSSCHE Prinzip: Jeder Punkt des Hindernisses ist Quelle einer Kugelwelle, deren Amplitude gleich der Amplitude der einfallenden Welle mal der Durchlässigkeitsfunktion α des Hindernisses im betrachteten Punkt ist. Letztere gibt an, welcher Teil des Lichts durchgelassen wird; sie ist also eins an den Stellen, an denen alles Licht durchkommt, und null dort, wo nichts durchkommt. An Stellen, an denen ein Teil des Lichts durchgelassen wird, kann sie auch Zwischenwerte annehmen.



CHRISTIAAN HUYGENS (1629–1695) kam aus einer niederländischen Diplomatenfamilie. Dadurch und später auch durch seine Arbeit hatte er Kontakte zu führenden europäischen Wissenschaftlern wie DESCARTES und PASCAL. Nach seinem Studium der Mathematik und Juris arbeitete er teilweise auch selbst als Diplomat, interessierte sich aber bald vor allem für Astronomie und den Bau der dazu notwendigen Instrumente. Er entwickelte eine neue Methode zum Schleifen von Linsen und erhielt ein Patent für die erste Pendeluhr. Trotz des französisch-niederländischen Kriegs arbeitete er einen großen Teil seines Lebens an der *Académie Royale des Sciences* in Paris, wo beispielsweise LEIBNIZ viel Mathematik bei ihm lernte. HUYGENS war ein scharfer Kritiker sowohl von NEWTONS Theorie des Lichts als auch seiner Gravitationstheorie, die er für absurd und nutzlos hielt. Gegen Ende seines Lebens beschäftigte er sich mit der Möglichkeit außerirdischen Lebens.

Bei der FRAUNHOFER-Beugung betrachten wir auch die gebeugten Wellen nur aus sehr großer Entfernung und können daher statt von Kugelwellen von ebenen Wellen ausgehen. Außerdem können wir die Zeitabhängigkeit der Welle ignorieren, denn die Frequenzen, mit denen das sichtbare Licht schwingt, liegen um Größenordnungen jenseits sowohl unserer Reaktionszeit als auch der unserer Meßinstrumente, so daß wir nur die Amplituden messen können. Schreiben wir die einfallende Welle als

$$\psi(\mathbf{x}, t) = A_0 e^{-i\mathbf{k}\cdot\mathbf{x}} \cdot e^{i\omega t},$$

ist also der zweite Faktor eine zeitabhängige Phasenvariation, die wir bei der Berechnung des räumlichen Intensitätsverteilung des Beugungsbilds ignorieren können; es reicht also, den Faktor $A_0 e^{-i\mathbf{k}\cdot\mathbf{x}}$ zu betrachten.

Eine weitere Konsequenz des (auch im Vergleich zur Größe des Hindernisses) weit entfernten Betrachtungspunkts ist, daß wir das Hindernis vom Schirm aus praktisch nur als Punkt sehen; was an einer gegebenen Stelle des Schirms ankommt, hängt also im wesentlichen nur ab vom Winkel θ oder (im Zweidimensionalen) den Winkeln θ und φ , unter dem (oder denen) die Strahlen von diesem „Punkt“ ausgehen.

Um die Intensität des Beugungsbilds in einem gegebenen Punkt zu berechnen, müssen wir also alle vom Hindernis in einem festen Winkel

ausgehenden Strahlen aufsummieren, und *hierbei* müssen wir auch die Phasen berücksichtigen, da diese Strahlen miteinander interferieren. Abbildung 28 zeigt, wie sich die Laufwege zweier benachbarter Strahlen unterscheiden, und diese Differenzen können wir nicht vernachlässigen, da sie in der Größenordnung des Hindernisses und damit auch der Wellenlänge des Lichts liegen.

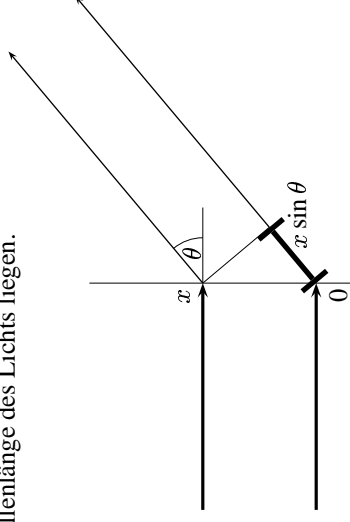


Abb. 28: Laufwegunterschied zweier paralleler Strahlen

Betrachten wir zunächst den (in Abbildung 28 dargestellten) eindimensionalen Fall. Vergleichlich mit dem Strahl, der von einem (irgendwie gewählten) Nullpunkt des Hindernisses ausgeht, hat der Strahl mit Ausgangspunkt in Entfernung x einen Laufwegunterschied von $x \sin \theta$; dies entspricht einem Phasenfaktor von $e^{-ikx \sin \theta}$. Wählen wir also die Phase im Nullpunkt als Referenz (die wir in den zu ignorierenden Phasenfaktor der einfallenden Welle hineinziehen können), ist die Summe aller unter dem Winkel θ abgehenden Strahlen gleich

$$\int_{-\infty}^{\infty} \alpha(x) e^{-ikx \sin \theta} dx;$$

das ist gleich der FOURIER-Transformierten von $\alpha(x)$, ausgewertet im Punkt $u = k \sin \theta$.

Bei einem zweidimensionalen Hindernis müssen entsprechend zwei Winkelvariablen θ und ϕ berücksichtigt werden, und auch die Durchlässigkeitsfunktion α hängt von zwei Variablen x, y ab; außerdem müssen

wir nun vom Wellenzahlvektor sowohl die x - als auch die y -Komponente berücksichtigen. Wir erhalten daher als Summe aller Strahlen unter den beiden gegebenen Winkeln das Integral

$$\iint_{\mathbb{R}^2} \alpha(x, y) e^{-i(k_1 x \sin \theta + k_2 y \sin \phi)} dx dy,$$

d.h. die zweidimensionale FOURIER-Transformierte von α , ausgewertet im Punkt $(u, v) = (k_1 \sin \theta, k_2 \sin \phi)$.

Zur Vereinfachung der Schreibweise drückt man das Beugungsbild meist einfach in der Variablen u bzw. den Variablen u und v aus statt in den Winkelvariablen; dann ist das Beugungsbild eines Hindernisses mit Durchlässigkeitsfunktion α einfach die FOURIER-Transformierte von α .

Die Größen u und v lassen sich zwar als Strecken interpretieren, sind aber *nicht* proportional zu den Strecken, die man auf einem ebenen Schirm messen kann: Deren Längen sind proportional zu $\tan \theta$ und $\tan \varphi$. Für kleine Winkel, auf die man sich bei der FRAUNHOFER-Beugung wegen des großen Abstands zum Schirm notwendigerweise beschränken muß, unterscheiden sich allerdings Sinus, Tangens und Bogenmaß nur wenig, so daß man auch ohne Umrechnung ein gutes Bild des Beugungsmusters erhält.

Als erstes Beispiel wollen wir das Beugungsbild eines eindimensionalen Spalts berechnen. Dieser habe die Breite a ; seine Durchlässigkeitsfunktion kann also beispielsweise geschrieben werden als

$$\alpha: \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} 1 & \text{für } -\frac{a}{2} \leq x \leq \frac{a}{2} \\ 0 & \text{sonst} \end{cases} \end{cases}$$

Damit ist

$$\begin{aligned} \hat{\alpha}(u) &= \int_{-\frac{a}{2}}^{\frac{a}{2}} \alpha(x) e^{-iux} dx = \int_{-\frac{a}{2}}^{\frac{a}{2}} e^{-iux} dx \\ &= \frac{e^{-\frac{ia}{2}u} - e^{\frac{ia}{2}u}}{-iu} = \frac{2 \sin \frac{au}{2}}{u} = a \frac{\sin \frac{au}{2}}{\frac{au}{2}} = a \operatorname{sinc} \frac{au}{2}. \end{aligned}$$

Dies erklärt, warum die Funktion $\operatorname{sinc} x$ auch als *Spaltfunktion* bezeichnet wird.

Die Lichtintensitäten, die man im Beugungsbild beobachtet, sind allerdings *nicht* durch diese Funktion gegeben: $\hat{\alpha}(u)$ ist die Amplitude einer skalaren Welle; die Intensität ist gleich dem Betragsquadrat davon, bei einer reellen Funktion wie hier also einfach das Quadrat

$$\hat{\alpha}(u)^2 = 4 \frac{\sin^2 \frac{au}{2}}{u^2} = a^2 \operatorname{sinc}^2 \frac{au}{2}.$$

Als nächstes Beispiel betrachten wir Beugung an einem regelmäßigem Strichgitter. Der Abstand zweier Striche sei d und es gebe insgesamt $2N + 1$ Striche. Wenn wir in erster Näherung die Breite der Striche vernachlässigen, können wir die Durchlässigkeitsfunktion α als Summe von δ -Distributionen schreiben:

$$\alpha(t) = \sum_{k=-N}^N \delta(x - kd).$$

Das Beugungsbild ist somit gegeben durch

$$\begin{aligned} \hat{\alpha}(u) &= \int_{-\infty}^{\infty} \alpha(x) e^{-iux} dx = \sum_{k=-N}^N \int_{-\infty}^{\infty} \delta(x - kd) e^{-iux} dx \\ &= \sum_{k=-N}^N e^{-iukd} = \sum_{k=-N}^N e^{iukd} = e^{-iuNd} \sum_{k=0}^{2N} e^{iukd} \\ &= e^{-iuNd} \frac{1 - e^{iud(2N+1)}}{1 - e^{iud}} = \frac{e^{-iuNd} - e^{iu(N+1)d}}{1 - e^{iud}} \\ &= \frac{e^{iu(N+\frac{1}{2})d} - e^{-iu(N+\frac{1}{2})d}}{e^{iu\frac{d}{2}} - e^{-iu\frac{d}{2}}} = \frac{\sin u(N + \frac{1}{2})d}{\sin \frac{ud}{2}}. \end{aligned}$$

Abbildung 29 zeigt diese Funktion; man sieht ihr das charakteristische Linienmuster an, das man bei der Beugung am Gitter beobachtet.

Trotzdem wird vielleicht einigen Lesern unwohl sein beim Gedanken an eine Summe von δ -Distributionen als Durchlässigkeitsfunktion. Deshalb wollen wir zur Sicherheit nachrechnen, was sich ändert, wenn wir stattdessen die Striche als Spalte der Breite a annehmen.

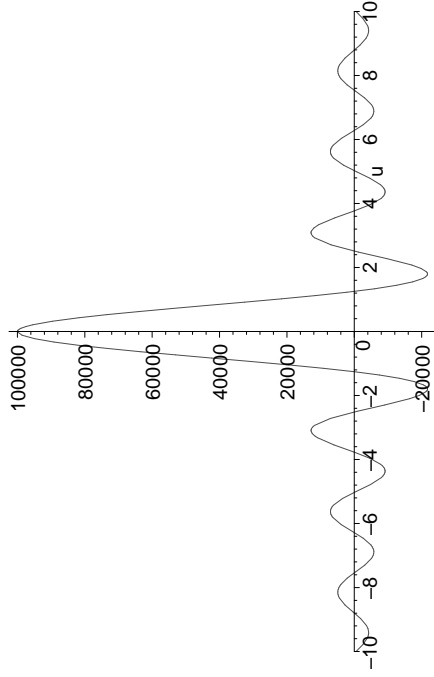


Abb. 29: Beugung am Gitter

Für einen einzelnen solchen Spalt haben wir dann die oben betrachtete Durchlässigkeitsfunktion

$$\alpha_a: \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} 1 & \text{für } -\frac{a}{2} \leq x \leq \frac{a}{2} \\ 0 & \text{sonst} \end{cases} \end{cases}$$

eines Spalts der Breite a , und die Durchlässigkeitsfunktion des gesamten Strichgitters ist die Faltung $\alpha \star \alpha_a$ dieser Funktion mit der oben betrachteten Funktion α . Das Beugungsbild ist somit gegeben durch das Produkt des gerade berechneten Beugungsbilds mit dem Beugungsbild eines Spalts, also durch

$$\frac{\sin u(N + \frac{1}{2})d}{\sin \frac{ud}{2}} \cdot \frac{\sin \frac{au}{2}}{\frac{u}{2}}$$

Da der Abstand zwischen zwei Spaltmitteln gleich d ist, muß die Spaltbreite a echt kleiner als d sein, und die Anzahl N der Striche im Gitter liegt typischerweise bei mindestens einigen Zehntausend. Somit hat der Sinus im zweiten Term eine Kreisfrequenz, die um einen mindestens fünfstelligen Faktor größer ist als die im ersten; der zweite Term zeigt also erst dann eine nennenswerte Variation, wenn wir im ersten Faktor

mehrere Tausend Linien betrachten. In dem Bereich, den wir realistischere beobachten können, ist der zweite Term daher für alle praktischen Zwecke konstant. Der Betrag dieser Konstanten ist irrelevant, denn da wir bei der FRAUNHOFER-Beugung das Beugungsbild in „sehr großer“ Entfernung vom Gitter betrachten, können wir sinnvollerweise ohnehin nur von relativen, nicht aber von absoluten Helligkeiten reden.

Als letztes Beispiel zur eindimensionalen Beugung möchte ich eines betrachten, bei dem das Licht nicht als konstante Wellenfront einfällt: Je nach Wahl der Randbedingungen im optischen Resonator entsteht nicht immer ein Strahl, der näherungsweise als ebene Welle betrachtet werden kann (die sogenannte TEM₀₀-Mode); bringt man Hindernisse in den Strahlengang, können auch höhere TEM-Moden angeregt werden (TEM = *transversal elektromagnetisch*). Bei einem dünnen Hindernis wie etwa einem Haar genau in der Mitte des Strahls beispielsweise entsteht die TEM₀₁-Mode, die aus einem linken und einem rechten Halbstahl besteht, deren Phasen sich um 180° unterscheiden, und die man ansonsten wieder näherungsweise als ebene Wellen betrachten kann. Trifft ein solcher Strahl auf einen Spalt, dessen Mitte mit der Grenze zwischen den beiden Halbstrahlen zusammenfällt, ist also in der linken Hälfte des Spalts die Phase um 180° gegenüber der rechten verschoben; dies können wir formal dadurch beschreiben, daß wir die Durchlässigkeitsfunktion des Spalts multiplizieren mit einer Funktion, die in der linken Hälfte +1 und in der rechten gleich -1 ist. Für einen Spalt der Breite a erhalten wir als Beugungsbild

$$\int_{-\frac{a}{2}}^0 e^{-iux} dx + \int_0^{\frac{a}{2}} -e^{-iux} dx = \frac{1 - e^{iua/2}}{-iu} - \frac{e^{-iua/2} + 1}{-iu} = \frac{2i}{u} \left(1 - \cos \frac{ua}{2}\right)$$

Alternativ läßt sich dies auch über die Beziehung

$$e^{iua/2} + e^{-iua/2} - 2 = (e^{iua/4} - e^{-iua/4})^2 = -4 \sin^2 \frac{ua}{4}$$

als

$$\frac{4i}{u} \sin^2 \frac{ua}{4}$$

schreiben. Daß hier imaginäre Größen auftreten, braucht uns natürlich nicht zu stören: Die beobachteten Intensitäten sind bekanntlich die Betragsquadrate der hier berechneten Funktionen, also reell und positiv.

Zum Abschluß möchte ich noch zumindest ein Beispiel eines zweidimensionalen Beugungsbilds betrachten. Leider sind die zugehörigen FOURIER-Integrale schon in so einfachen Fällen wie dem einer schiefen förmigen Blende nicht mehr elementar auswertbar; wir beschränken uns daher auf den extrem einfachen Fall einer rechteckigen Blende. Deren Durchlässigkeitsfunktion ist

$$\alpha: \begin{cases} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto \begin{cases} 1 & \text{falls } -\frac{a}{2} \leq x \leq \frac{a}{2} \text{ und } -\frac{b}{2} \leq y \leq \frac{b}{2} \\ 0 & \text{sonst} \end{cases} \end{cases},$$

die Beugungsfigur ist also gegeben durch

$$\begin{aligned} \hat{\alpha}(u, v) &= \iint_{\mathbb{R}^2} \alpha(x, y) e^{-i(u x + v y)} dx dy = \iint_{\substack{-\frac{a}{2} \leq x \leq \frac{a}{2} \\ -\frac{b}{2} \leq y \leq \frac{b}{2}}} e^{-i(u x + v y)} dx dy \\ &= \int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} e^{-i u x} e^{-i v y} dx dy = \int_{-\frac{a}{2}}^{\frac{a}{2}} e^{-i u x} \left(\int_{-\frac{b}{2}}^{\frac{b}{2}} e^{-i v y} dy \right) dx \\ &= \int_{-\frac{a}{2}}^{\frac{a}{2}} e^{-i u x} \cdot \frac{\sin \frac{b v}{2}}{v} dx = 4 \cdot \frac{\sin \frac{a u}{2}}{u} \cdot \frac{\sin \frac{b v}{2}}{v}, \end{aligned}$$

da wir das Rechteck als Normalbereich betrachten können und somit das zweidimensionale Integral über zwei eindimensionale Integrationen berechnen können.

Als Beugungsfigur erhalten wir, nicht gerade überraschenderweise, das Produkt einer vertikalen und einer horizontalen Beugungsfigur eines eindimensionalen Spalts.

Kapitel 4 Differentialgleichungen

Differentialgleichungssysteme sind so ziemlich *das* wichtigste mathematische Hilfsmittel der Naturwissenschaften und der Technik. Die dahinterstehende Grundidee ist einfach: Man kann zwar nur selten *a priori* sagen, wie sich ein System über einen längeren Zeitraum hinweg entwickeln wird, aber man hat oft aufgrund von Naturgesetzen eine klare Vorstellung über die Zustandsänderung *im nächsten Augenblick*, d.h. also über den Wert der zeitlichen Ableitung der Zustandsgrößen in Abhängigkeit vom gegenwärtigen Zustand des Systems.

§ 1: Definitionen und erste Beispiele

a) Wurfparabel

Ein einfaches Beispiel hierfür liefert das zweite NEWTONSche Gesetz, wonach die zeitliche Ableitung des Impuls eines Teilchens gleich der auf das Teilchen wirkenden Kraft ist.

Ein in die Luft geworfener Gegenstand bewegt sich unter gewissen Bedingungen näherungsweise auf einer parabelförmigen Bahn. Wir wollen diese etwas vage Aussage präzisieren und mathematisch herleiten.

Es gibt viele Wurftechniken, und nur wenige davon können auf einfache Weise durch ein mathematisches Modell beschrieben werden; wir ignorieren daher den genauen Vorgang des Abwurfs und gehen davon aus, daß der Gegenstand *irgendwie* eine Anfangsgeschwindigkeit \vec{v} erreicht hat im Abwurfpunkt mit Koordinaten (x_0, y_0, z_0) ; den Zeitpunkt des Abwurfs bezeichnen wir mit t_0 .

Als nächstes nehmen wir an, daß wir den Luftwiderstand vernachlässigen können, eine Annahme, die beim Kugelstoßen kaum zu Fehlern führt, die aber beispielsweise für einen Fallschirmspringer (auch mit geschlossenem Fallschirm) oder einen Papierflieger völlig unrealistisch ist. Als nächstes wollen wir auch noch annehmen, daß wir nur relativ geringe Wurfhöhen erreichen, so daß die Erdanziehung als konstant angenommen werden kann.

Die Bewegung des Gegenstandes wird dann durch zwei Naturgesetze bestimmt: Das Gravitationsgesetz beschreibt den Effekt der Erdanziehung, und das zweite NEWTONSche Gesetz sagt uns, wie sich diese Kraft auf die Bewegung des Gegenstands auswirkt. Die Gravitation können wir aufgrund der gemachten Annahmen als konstant annehmen, d.h. auf einen Körper der Masse m wirkt die Kraft gm , wobei $g \approx 9,8 \text{ m/s}^2$ die Gravitationsbeschleunigung an der Erdoberfläche ist; bei „üblicher“ Ausrichtung des Koordinatensystems wirkt sie in Richtung der negativen z -Achse. Diese Gravitationskraft ist nach dem zweiten NEWTONSchen Gesetz gleich der Ableitung des Impulses nach der Zeit; wenn wir die Masse m als konstant voraussetzen, ist das also gleich m mal der Ableitung der Geschwindigkeit oder m mal der zweiten Ableitung des Orts. Wir haben somit das Differentialgleichungssystem

$$\ddot{x}(t) = 0, \quad \dot{y}(t) = 0 \quad \text{und} \quad \ddot{z}(t) = -g.$$

Diese Gleichungen sind erfüllt, wann immer $x(t)$ und $y(t)$ lineare Funktionen von t sind und $z(t)$ eine quadratische Funktion mit führendem Koeffizienten $-g$. Die sechs noch fehlenden Koeffizienten dieser drei Polynomfunktionen geben uns die Anfangsbedingungen: Zum Zeitpunkt $t = t_0$ ist

$$x(t_0) = x_0, \quad y(t_0) = y_0 \quad \text{und} \quad z(t_0) = z_0,$$

und die Geschwindigkeit ist \vec{v} , d.h.

$$\dot{x}(t_0) = v_1, \quad \dot{y}(t_0) = v_2 \quad \text{und} \quad \dot{z}(t_0) = v_3.$$

Also ist

$$x(t) = v_1(t - t_0) + x_0, \quad y(t) = v_2(t - t_0) + y_0$$

und

$$z(t) = -g(t - t_0)^2 + v_3(t - t_0) + z_0.$$

Diese Gleichungen beschreiben in der Tat fast immer eine Parabel: Falls wir die x -Achse des Koordinatensystems so wählen, daß die Anfangsgeschwindigkeit \vec{v} in der (x, z) -Ebene liegt, ist $v_2 = 0$. Falls auch v_1 verschwindet, falls wir den Gegenstand also senkrecht nach oben (oder gar unten) werfen, sind $x(t) = x_0$ und $y(t) = y_0$ konstant und nur

$$z(t) = -g(t - t_0)^2 + v_3(t - t_0) + z_0$$

hängt von der Zeit ab. Andernfalls können wir durch v_1 dividieren; wir erhalten

$$t - t_0 = \frac{x(t) - x_0}{v_1} \quad \text{und} \quad z(t) = \frac{-g}{v_1} (x(t) - x_0)^2 + \frac{v_3}{v_1} (x(t) - x_0) + z_0,$$

die Punkte $(x(t), z(t))$ liegen also in der Tat auf einer Parabel.

b) Radioaktiver Zerfall

Das gerade durchgerechnete Beispiel war insofern untypisch für Differentialgleichungen, als auf den rechten Seite der Gleichungen nur Konstanten standen; üblicherweise wird man dort Funktionen erwarten, die nicht nur von t abhängen (so daß man sie einfach integrieren kann), sondern auch noch von den gesuchten Funktionen. Beim radioaktiven Zerfall etwa ist die pro (kleiner) Zeiteinheit zerfallende Masse proportional zur noch vorhandenen Masse, es gibt also eine Konstante $\lambda > 0$, die sogenannte Zerfallskonstante, so daß die zum Zeitpunkt t vorhandene Masse $m(t)$ der Gleichung

$$\dot{m}(t) = -\lambda m(t)$$

genügt – zumindest, wenn diese Masse hinreichend groß ist. (Im atomaren Bereich muß man auch statistische Effekte berücksichtigen, aber ab etwa 10^{10} Atomen können die für alle praktischen Fälle vernachlässigt werden.)

Wir kennen bereits eine Funktion, die sich so verhält, wie es die obige Differentialgleichung angibt, nämlich die Exponentialfunktion $e^{-\lambda t}$, und natürlich entspricht auch für jedes konstante Vielfache dieser Funktion die Differentiation einfach der Multiplikation mit $-\lambda$. Das sind dann aber bereits alle Funktionen mit dieser Eigenschaft, denn der Quotient

$$q(t) = \frac{m(t)}{e^{-\lambda t}} = m(t) \cdot e^{\lambda t}$$

einer Lösungsfunktion und der Funktion $e^{-\lambda t}$ hat die Ableitung

$$\dot{q}(t) = \dot{m}(t) \cdot e^{-\lambda t} + m(t) \cdot \lambda e^{-\lambda t} = -\lambda m(t) \cdot e^{-\lambda t} + \lambda m(t) \cdot e^{-\lambda t} = 0,$$

ist also gleich einer Konstanten c , so daß

$$m(t) = c \cdot e^{-\lambda t}$$

ist. Indem wir $t = 0$ setzen, sehen wir, daß die Konstante $c = m(0)$ gleich der zum Zeitpunkt 0 vorhandenen Masse ist; falls wir stattdessen die Masse $m_0 = m(t_0)$ zu einem anderen Zeitpunkt t_0 kennen, können wir analog zum obigen Beispiel auch schreiben

$$m(t) = m_0 e^{-\lambda(t-t_0)} = (m_0 e^{\lambda t_0}) \cdot e^{-\lambda t}.$$

c) Differentialgleichungen und Differentialgleichungssysteme

Wir betrachten ein System, das durch n zeitlich veränderliche Größen $y_1(t), \dots, y_n(t)$ beschrieben wird; unter einem System von Differentialgleichungen oder kurz einer Differentialgleichung verstehen wir eine Vorschrift, die die zeitlichen Ableitungen $\dot{y}_1(t), \dots, \dot{y}_n(t)$ aus den Funktionswerten berechnet:

$$\dot{y}_1(t) = f_1(t, y_1(t), \dots, y_n(t))$$

$$\dot{y}_2(t) = f_2(t, y_1(t), \dots, y_n(t))$$

\vdots

$$\dot{y}_n(t) = f_n(t, y_1(t), \dots, y_n(t)).$$

Falls die Funktionen f_i nur von $y_1(t), \dots, y_n(t)$ abhängen und nicht auch noch direkt von der Zeit t spricht man von einem *autonomen* System. Da Naturgesetze nicht von der Zeit abhängen, hat man es in naturwissenschaftlich-technischen Anwendungen meist mit autonomen Systemen zu tun; man kann allerdings auch den Einfluß von Umgebungsgrößen in einem zeitabhängigen Term zusammenfassen und so ein nichtautonomes System erhalten.

Falls wir, wie in den Beispielen aus den vorangegangenen Abschnitten, die Werte der beteiligten Funktionen zu einem festen Zeitpunkt $t = t_0$ kennen, reden wir von einem *Anfangswertproblem*. Solche Probleme treten

typischerweise dann auf, wenn das weitere Verhalten *eines* konkreten Systems vorhergesagt werden soll.

Auch Differentialgleichungen, in denen wie im Beispiel der Wurfparabel höhere Ableitungen vorkommen, lassen sich so interpretieren: Wenn wir dort die drei Komponenten des Geschwindigkeitsvektors als neue Funktionen

$$u(t) = \dot{x}(t), \quad v(t) = \dot{y}(t) \quad \text{und} \quad w(t) = \dot{z}(t)$$

eingeführen, können wir das System schreiben als

$$\begin{aligned} \dot{x}(t) &= u(t), & \dot{y}(t) &= v(t), & \dot{z}(t) &= w(t), \\ \dot{u}(t) &= 0, & \dot{v}(t) &= 0, & \dot{w}(t) &= -g, \end{aligned}$$

und wir kennen für jede der sechs beteiligten Funktionen ihren Wert an der Stelle $t = t_0$.

Ein für die Informationstechnik wichtiger Spezialfall sind Gleichungen der Form

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \dots + a_1\dot{y}(t) + a_0y(t) = b(t),$$

die sogenannten linearen Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten. Auch diese Gleichungen lassen sich leicht auf die obige Form bringen: Wir betrachten n neue Funktionen

$$y_0(t), y_1(t), \dots, y_{n-1}(t)$$

mit der Idee, daß sich $y_i(t)$ so verhalten soll wie die i -te Ableitung von $y(t)$. Dazu bilden wir das Differentialgleichungssystem

$$\dot{y}_0(t) = y_1(t)$$

$$\dot{y}_1(t) = y_2(t)$$

\vdots

$$\dot{y}_{n-2}(t) = y_{n-1}(t)$$

$$\dot{y}_{n-1}(t) = b(t) - a_{n-1}y^{(n-1)}(t) - \dots - a_1\dot{y}(t) - a_0y(t).$$

Für jede Lösung $y(t)$ der obigen Gleichung ist dann das n -tupel

$$(y(t), \dot{y}(t), \ddot{y}(t), \dots, y^{(n-1)}(t))$$

eine Lösung des Systems, und für jede Lösung

$$(y_0(t), y_1(t), y_2(t), \dots, y_{n-1}(t))$$

des Differentialgleichungssystems ist $y_0(t)$ eine Lösung der obigen Gleichung.

Auch wenn das Differentialgleichungssystem als Anfangswertproblem gegeben ist, läßt sich das leicht in Anfangswerte für die Gleichung höherer Ordnung umschreiben: Hier werden die Werte $y(t_0), \dot{y}(t_0)$ usw. bis $y^{(n-1)}(t_0)$ vorgegeben.

Somit beschreiben das System von Differentialgleichungen erster Ordnung und die eine Differentialgleichung höherer Ordnung genau dasselbe Phänomen. Wie wir im vorigen Kapitel gesehen haben, läßt sich die Differentialgleichung höherer Ordnung recht gut mit Hilfe von LAPLACE-Transformationen lösen; in diesem Kapitel werden wir sehen, daß der im letzten Semester entwickelte (und im folgenden noch auszubauende) Apparat der linearen Algebra eine strukturelle Übersicht über die Lösungsmenge des Systems. Erst die Kombination beider Ansätze liefert ein vollständiges Bild.

d) Systeme linearer Differentialgleichungen

Wir betrachten in diesem Abschnitt Systeme von Differentialgleichungen, wie sie zu Beginn dieses Paragraphen definiert wurden, unter der (sehr) einschränkenden Voraussetzung, daß die rechten Seiten linear in den gesuchten Funktionen $y_1(t), \dots, y_n(t)$ sind; wir betrachten also ein System

$$\begin{aligned} \dot{y}_1(t) &= a_{11}(t)y_1(t) + a_{12}(t)y_2(t) + \dots + a_{1n}(t)y_n(t) + b_1(t) \\ \dot{y}_2(t) &= a_{21}(t)y_1(t) + a_{22}(t)y_2(t) + \dots + a_{2n}(t)y_n(t) + b_2(t) \\ &\vdots \\ \dot{y}_n(t) &= a_{n1}(t)y_1(t) + a_{n2}(t)y_2(t) + \dots + a_{nn}(t)y_n(t) + b_n(t) \end{aligned} \quad (*)$$

Eventuell haben wir noch Anfangsbedingungen der Form

$$y_1(t_0) = c_0, \quad y_2(t_0) = c_2, \quad \dots, \quad y_n(t_0) = c_n \quad (**)$$

für ein festes $t_0 \in \mathbb{R}$.

Für das im letzten Kapitel betrachtete Beispiel des elektrischen Schwingkreises mit angelegter Wechselspannung etwa haben wir bei dieser Sicht der Dinge die beiden Funktionen $Q(t)$, die Ladung des Kondensators zum Zeitpunkt t , und $I(t) = \dot{Q}(t)$, die resultierende Stromstärke; das Differentialgleichungssystem (*) ist hier also

$$\begin{aligned} \dot{Q}(t) &= I(t) \\ \dot{I}(t) &= -\frac{R}{L}I(t) - \frac{Q(t)}{LC} + A_0 \cos \omega_0 t \end{aligned}$$

Wir können die Funktionen $y_i(t), b_i(t)$ und die Anfangswerte c_i jeweils zu Vektoren zusammenfassen und die Koeffizientenfunktionen $a_{ij}(t)$ zu einer Matrix: Mit

$$\vec{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \quad \vec{b}(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

und

$$A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \dots & a_{nn}(t) \end{pmatrix},$$

erhalten wir die übersichtlichere Form

$$\dot{\vec{y}}(t) = A(t) \cdot \vec{y}(t) + \vec{b}(t),$$

wobei die Ableitung eines Vektors von Funktionen natürlich der Vektor der abgeleiteten Funktionen sein soll. Falls es Anfangsbedingungen gibt, können sie nun in der kompakte Form $\vec{y}(t_0) = \vec{c}$ geschrieben werden.

In Analogie zu den linearen Gleichungssystemen bezeichnen wir das System (*) als *homogen*, wenn $\vec{b}(t)$ der Nullvektor ist, wenn also alle Funktionen $b_i(t)$ verschwinden; andernfalls bezeichnen wir es als *inhomogen*. Das homogene System zu einem gegebenen inhomogenen System soll einfach dasjenige System sein, in dem alle $b_i(t)$ durch null ersetzt wurden.

Analog zum Fall linearer Gleichungssystemen gilt auch hier

Lemma: a) Die Menge aller Lösungen eines *homogenen* Differentialgleichungssystems der Form (*) ist ein \mathbb{R} -Vektorraum.

b) Ist das System nicht homogen und ist $\vec{y}(t)$ eine feste Lösung, so läßt sich jede andere Lösung $\vec{z}(t)$ schreiben als $\vec{z}(t) = \vec{y}(t) + \vec{x}(t)$ mit einer Lösung $\vec{x}(t)$ des zugehörigen homogenen Systems; die Lösungsmenge ist also ein affiner Raum.

Beweis: a) Wir müssen zeigen, daß für zwei Lösungen $\vec{x}(t)$ und $\vec{y}(t)$ eines homogenen Systems auch jede Linearkombination $\lambda\vec{x}(t) + \mu\vec{y}(t)$ mit $\lambda, \mu \in \mathbb{R}$ wieder eine Lösung ist. Das ist aber klar, denn wenn

$$\vec{x}(t) = A(t) \cdot \vec{x}(t) \quad \text{und} \quad \vec{y}(t) = A(t) \cdot \vec{y}(t)$$

ist, gilt auch

$$\begin{aligned} \frac{d}{dt} (\lambda\vec{x}(t) + \mu\vec{y}(t)) &= \lambda\dot{\vec{x}}(t) + \mu\dot{\vec{y}}(t) = \lambda A(t) \cdot \vec{x}(t) + \mu A(t) \cdot \vec{y}(t) \\ &= A(t) \cdot (\lambda\vec{x}(t) + \mu\vec{y}(t)). \end{aligned}$$

b) Sind $\vec{y}(t)$ und $\vec{z}(t)$ zwei Lösungen von (*), so ist

$$\dot{\vec{y}}(t) = A(t) \cdot \vec{y}(t) + \vec{b}(t) \quad \text{und} \quad \dot{\vec{z}}(t) = A(t) \cdot \vec{z}(t) + \vec{b}(t);$$

die Differenz $x(t) = z(t) - y(t)$ hat somit die Ableitung

$$\begin{aligned} \dot{\vec{x}}(t) &= \dot{\vec{z}}(t) - \dot{\vec{y}}(t) = \left(A(t) \cdot \vec{z}(t) + \vec{b}(t) \right) - \left(A(t) \cdot \vec{y}(t) + \vec{b}(t) \right) \\ &= A(t) \cdot \vec{z}(t) - A(t) \cdot \vec{y}(t) = A(t) \cdot (\vec{z}(t) - \vec{y}(t)) = A(t) \cdot \vec{x}(t) \end{aligned}$$

und $x(t)$ löst also in der Tat das zugehörige homogene System. ■

Um die Lösungsmenge des Differentialgleichungssystems (*) zu verstehen, müssen wir nach diesem Lemma zwei Teilaufgaben lösen:

1.) Wir müssen den Vektorraum der Lösungen des homogenen Systems bestimmen.

2.) Wir müssen uns wenigstens eine Lösung des inhomogenen Systems verschaffen – oder zumindest wissen, daß eine existiert.

Um im einfachsten Fall zu sehen, wie so etwas funktionieren könnte, betrachten wir ein „System“ aus genau einer Gleichung

$$\dot{y}(t) = a(t) \cdot y(t) + b(t);$$

dabei nehmen wir an, daß y eine differenzierbare Funktion von \mathbb{R} nach \mathbb{R} sei und $a, b: \mathbb{R} \rightarrow \mathbb{R}$ stetige Funktionen.

Wir beginnen mit der Lösung des homogenen Systems

$$\dot{y}(t) = a(t) \cdot y(t).$$

Unter der Annahme, daß wir das dürfen, dividieren wir durch $y(t)$ und erhalten

$$\frac{\dot{y}(t)}{y(t)} = a(t).$$

Der Quotient links ist bekanntlich die logarithmische Ableitung von $y(t)$; falls dies nicht mehr bekannt sein sollte, zeigt eine einfache Anwendung der Kettenregel, daß in einem Intervall, in dem $y(t)$ positiv ist,

$$\frac{d}{dt} \ln y(t) = \frac{\dot{y}(t)}{y(t)} = a(t)$$

ist. In einem Intervall, in dem $y(t)$ negativ ist, gilt entsprechend

$$\frac{d}{dt} \ln(-y(t)) = \frac{-\dot{y}(t)}{-y(t)} = \frac{\dot{y}(t)}{y(t)} = a(t),$$

und allgemein haben wir somit

$$\frac{d}{dt} \ln|y(t)| = \frac{\dot{y}(t)}{y(t)} = a(t)$$

in jedem Intervall, in dem $y(t)$ nirgends verschwindet.

Integration beider Seiten führt auf

$$\ln|y(t)| = \int a(t) dt + C \quad \text{oder} \quad y(t) = e^{\int a(t) dt + C} = e^C \cdot e^{\int a(t) dt}$$

oder

$$y(t) = \pm e^C \cdot e^{\int a(t) dt},$$

wobei das Vorzeichen wegen der Stetigkeit von y im gesamten Intervall konstant ist, da die Exponentialfunktion nie null wird.

Damit ist in diesem Fall das erste Problem auf eine einfache Integration zurückgeführt.

Bleibt noch die Frage, was passiert, wenn $y(t)$ an irgendeinem Punkt t_0 eine Nullstelle haben sollte. Wir wollen uns überlegen, daß $y(t)$ dann auch für jedes $t > t_0$ verschwinden muß.

Falls nicht, gibt es einen Punkt $t_1 > t_0$, so daß $y(t_1) \neq 0$ ist. Wegen der Stetigkeit von $y(t)$ ist die Funktion dann auch in einer Umgebung von t_1 von Null verschieden, d.h. dort können wir die obigen Argumente anwenden und sehen, daß $y(t)$ dort die Form $e^{h(t)}$ hat mit irgendeiner Funktion h . Da y als differenzierbare Funktion insbesondere überall stetig sein muß und $e^{h(t)}$ nirgends verschwindet, ist das nicht möglich. Genauso überlegt man sich, daß $y(t)$ für jedes $t < t_0$ verschwinden muß, $y(t)$ ist also gleich der Nullfunktion. Diese ist somit die einzige Lösung, die noch zusätzlich betrachtet werden muß. Insbesondere folgt daraus auch, daß eine Lösungsfunktion, die in irgendeinem Punkt positiv bzw. negativ ist, überall positiv bzw. negativ sein muß, denn eine stetige Funktion kann ihr Vorzeichen nur wechseln, wenn sie in irgendeinem Punkt null wird; wie wir gerade gesehen haben, ist das genau dann der Fall, wenn sie überall verschwindet.

Somit hat jede Lösung die Form

$$y(t) = ae^{\int a(t)dt} \quad \text{mit einem } a \in \mathbb{R}$$

und umgekehrt ist auch jede dieser Funktionen eine Lösung. Insbesondere ist die Lösungsmenge ein eindimensionalen Vektorraum.

Es wäre schön, wenn wir im mehrdimensionalen Fall genauso vorgehen könnten: In Analogie zu

$$\dot{y}(t) = a(t) \cdot y(t) \implies y(t) = ce^{\int a(t)dt} \quad \text{mit } c \in \mathbb{R}$$

könnte vielleicht gelten

$$\dot{\vec{y}}(t) = A(t) \cdot \vec{y}(t) \implies y(t) = e^{\int A(t)dt} \cdot \vec{c} \quad \text{mit } \vec{c} \in \mathbb{R}^n ?$$

Das Problem dabei ist nur, daß wir hier nicht die geringste Ahnung haben, was die rechte Seite bedeuten soll; unser nächstes Ziel wird sein, ihr eine Bedeutung zu geben und uns dann zu überlegen, ob bzw. unter welchen Bedingungen die obige Formel korrekt ist.

e) Die Matrixexponentialfunktion

Wir orientieren uns wieder am Eindimensionalen: Für eine reelle Zahl x ist

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!},$$

also setzen wir analog für eine $n \times n$ -Matrix X

$$e^X \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} \frac{1}{i!} \cdot X^i.$$

Damit ist klar, daß e^X eine $n \times n$ -Matrix sein soll, und das erklärt auch, warum oben der Konstantenvektor \vec{y}_0 auf der rechten Seite steht. Was wir uns noch überlegen müssen, ist die Konvergenz der Reihe.

Dazu müssen wir die Größe der Einträge in den Matrizen X^i abschätzen: Sind allgemein A, B zwei $n \times n$ -Matrizen und sind die Beträge aller Einträge von A kleiner oder gleich a und die von B kleiner oder gleich b , so kann es in AB offensichtlich keinen Eintrag geben, dessen Betrag größer ist als nab : Schließlich ist jeder Eintrag in der Produktmatrix eine Summe von n Summanden, deren jeder Produkt je eines Eintrags von A und von B ist.

Um diese Formel leichter anwenden zu können, machen wir sie mutwillig schlechter und begnügen uns damit, daß jeder Eintrag von AB höchstens den Betrag $(na) \cdot (nb)$ hat.

Ist nun x der Betrag des größten Eintrags in der Matrix X , so folgt induktiv sofort, daß in X^i höchstens Zahlen bis zum Betrag $(nx)^i$ stehen können; in der endlichen Teilsumme

$$\sum_{i=0}^M \frac{1}{i!} X^i$$

hat daher jeder Eintrag einen Betrag kleiner

$$\sum_{i=0}^M \frac{(nx)^i}{i!}.$$

Letztere Summe konvergiert für $M \rightarrow \infty$ gegen e^{nx} , und damit muß auch die Matrixsumme absolut konvergieren, denn die Reihe für e^{nx} ist konvergente Majorante des Betrags eines jeden Eintrags. Insbesondere hat jeder Eintrag von e^x höchstens den Betrag e^{nx} .

Damit wissen wir also, daß die Matrix e^X für jede $n \times n$ -Matrix X existiert; somit ist die Funktion $t \mapsto e^{At}$ wohldefiniert.

f) Eigenschaften der Matrixexponentialfunktion

Wir können natürlich nicht erwarten, daß die Matrixexponentialfunktion alle schönen Eigenschaften der gewöhnlichen Exponentialfunktion erbt. Beispielsweise ist nur schwer vorstellbar, daß für beliebige Matrizen A und B gelten sollte $e^{A+B} = e^A \cdot e^B$. Da für zwei Matrizen A und B stets $A + B = B + A$ ist, müßte dann auch $e^A \cdot e^B = e^B \cdot e^A$ sein, was zumindest unwahrscheinlich aussieht. In der Tat ist etwa für

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

sowohl A^2 als auch B^2 gleich der Nullmatrix, d.h.

$$e^A = E + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad e^B = E + B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

und

$$e^A \cdot e^B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Das Quadrat von $C = A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist aber gleich der Einheitsmatrix und daher ist

$$\begin{aligned} e^{A+B} &= E + C + \frac{1}{2!}C^2 + \frac{1}{3!}C^3 + \frac{1}{4!}C^4 + \frac{1}{5!}C^5 + \dots \\ &= \left(\sum_{i=0}^{\infty} \frac{1}{(2i)!} \right) \cdot E + \left(\sum_{i=0}^{\infty} \frac{1}{(2i+1)!} \right) \cdot C \\ &= \cosh 1 \cdot E + \sinh 1 \cdot C = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e + e^{-1} & e - e^{-1} \\ e - e^{-1} & e + e^{-1} \end{pmatrix}. \end{aligned}$$

Allgemeiner ist

$$e^{At} = E + tA = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad e^{Bt} = E + tB = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$$

und

$$e^{Ct} = \sum_{i=0}^{\infty} \frac{1}{i!} C^i = \left(\sum_{i=0}^{\infty} \frac{1}{(2i)!} \right) \cdot E + \left(\sum_{i=0}^{\infty} \frac{1}{(2i+1)!} \right) \cdot C = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

Zum Glück gilt aber wenigstens

Lemma: Für zwei Matrizen $A, B \in \mathbb{R}^{n \times n}$ mit $AB = BA$ ist

$$e^{A+B} = e^A \cdot e^B = e^B \cdot e^A.$$

Insbesondere ist für $s, t \in \mathbb{R}$

$$e^{A(s+t)} = e^{As} \cdot e^{At}.$$

Beweis: Für zwei reelle Zahlen x, y ist $e^{x+y} = e^x e^y = e^y e^x$; damit gilt dieselbe Formel auch für zwei reellwertige Variablen x und y . Wenn wir in allen Potenzreihen alle x - und y -Potenzen oberhalb der N -ten ignorieren, sagt die Gleichung aus, daß drei Polynome in x und y als *Polynome* identisch sind.

Beim Rechnen mit Polynomen in x und y verwendet man keine speziellen Eigenschaften dieser Variablen *außer*, daß sie *kommutieren*. Damit kann man in so eine Polynomidentität auch kommutierende Matrizen A und B einsetzen: Beispielsweise führt die Polynomidentität

$$(x + y)^2 = x^2 + 2xy + y^2$$

zur Identität

$$(A + B)^2 = A^2 + 2AB + B^2,$$

die wegen der für beliebige Matrizen gültigen Gleichung

$$(A + B)^2 = A^2 + AB + BA + B^2$$

für kommutierende Matrizen in der Tat erfüllt ist.

Damit ist für kommutierende Matrizen A, B speziell stets

$$e^{A+B} = e^A \cdot e^B = e^B \cdot e^A.$$

Da zwei skalare Vielfache derselben Matrix stets miteinander kommutieren, folgt damit auch die letzte Aussage des Lemmas. ■

Die für uns wichtigste Anwendung hiervon ist

Satz: Für jede $n \times n$ -Matrix A ist die Funktion

$$\begin{cases} \mathbb{R} \rightarrow \mathbb{R}^{n \times n} \\ t \mapsto e^{At} \end{cases}$$

stetig differenzierbar mit Ableitung $t \mapsto A \cdot e^{At} = e^{At} \cdot A$.

Beweis: Die Ableitung ist definiert als

$$\lim_{h \rightarrow 0} \frac{e^{A(t+h)} - e^{At}}{h}.$$

Da die Matrizen At und Ah miteinander vertauschbar sind, ist nach dem gerade bewiesenen Lemma $e^{A(t+h)} = e^{At} \cdot e^{Ah}$, also

$$\frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \cdot \frac{e^{Ah} - E}{h}.$$

Dabei ist

$$\frac{e^{Ah} - E}{h} = \frac{1}{h} \cdot \sum_{i=1}^{\infty} \frac{(Ah)^i}{i!} = A + A^2 h \sum_{i=2}^{\infty} \frac{(Ah)^{i-2}}{i!} = A + A^2 h \sum_{i=0}^{\infty} \frac{(Ah)^i}{(i+2)!}.$$

Nach der obigen Diskussion ist jeder Eintrag der Matrix in der rechtenstehenden Summenmatrix höchstens gleich

$$\sum_{i=0}^{\infty} \frac{(ah)^i}{(i+2)!} = \frac{e^{ah} - (1+ah)}{a^2 h^2},$$

bleibt also insbesondere beschränkt. Dies gilt auch für $h \rightarrow 0$, denn wie zweimalige Anwendung der DE L'HOSPITAL'schen Regel oder TAYLOR-Entwicklung zeigen, ist der Grenzwert dann $\frac{1}{2}$. Damit existiert

$$\lim_{h \rightarrow 0} \sum_{i=1}^{\infty} \frac{(Ah)^i}{(i+2)!},$$

und somit ist $\frac{d}{dt} e^{At} = \lim_{h \rightarrow 0} \frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \cdot A$, da der Vorfaktor $A^2 h$ gegen Null geht. Dies ist auch gleich $A \cdot e^{At}$, denn da A mit jeder seiner Potenzen vertauschbar ist, ist es auch mit jeder endlichen Teilsumme der Reihe von e^{At} vertauschbar, also auch mit e^{At} selbst. ■

Am Ende des vorigen Abschnitts hatten wir gehofft, daß vielleicht auch für jede matrixwertige Funktion $A(t)$ gelten könnte, daß

$$\frac{d}{dt} e^{A(t)} = \dot{A}(t) e^{A(t)}$$

ist; dies war offensichtlich zu optimistisch: Da

$$e^{A(t+h)} = e^{A(t)+h\dot{A}(t)+o(h)}$$

ist, bräuchten wir für einen Beweis nach obigem Vorbild, daß $A(t)$ und $\dot{A}(t)$ miteinander kommutieren; dies ist aber im allgemeinen nicht der Fall. Für

$$A(t) = \begin{pmatrix} 1 & t \\ 1 & 1 \end{pmatrix}$$

beispielsweise ist

$$\dot{A}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

also

$$A(t) \cdot \dot{A}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \text{aber} \quad \dot{A}(t) \cdot A(t) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Dies ist natürlich kein Beweis dafür, daß die Ableitung von $e^{A(t)}$ ungleich $\dot{A}(t) \cdot e^{A(t)}$ ist, aber im vorliegenden Fall ist die Ableitung in der Tat verschieden sowohl von $\dot{A}(t)e^{A(t)}$ als auch von $e^{A(t)} \cdot \dot{A}(t)$: Mit den Methoden, die wir im nächsten Abschnitt kennenlernen werden, können wir durch eine (alles andere als angenehme) Rechnung zeigen, daß

$$e^{A(t)} = \begin{pmatrix} \frac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{e^{1+\sqrt{t}} - e^{1-\sqrt{t}}} & \sqrt{t} \frac{e^{1+\sqrt{t}} - e^{1-\sqrt{t}}}{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}} \\ \frac{2}{2\sqrt{t}} & \frac{2}{2} \frac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}} \end{pmatrix},$$

$$\frac{d}{dt} e^{A(t)} = \begin{pmatrix} \frac{(e^{2\sqrt{t}} - 1) e^{1-\sqrt{t}}}{4\sqrt{t}} & \frac{(-1 + \sqrt{t} + \sqrt{t} e^{2\sqrt{t}} + e^{2\sqrt{t}}) e^{1-\sqrt{t}}}{4\sqrt{t}} \\ \frac{(1 - e^{2\sqrt{t}} + \sqrt{t} + \sqrt{t} e^{2\sqrt{t}}) e^{1-\sqrt{t}}}{4t\sqrt{t}} & \frac{(e^{2\sqrt{t}} - 1) e^{1-\sqrt{t}}}{4\sqrt{t}} \end{pmatrix},$$

aber

$$\dot{A}(t) \cdot e^{A(t)} = \begin{pmatrix} e^{1+\sqrt{t}} - e^{1-\sqrt{t}} & e^{1+\sqrt{t}} + e^{1-\sqrt{t}} \\ 2\sqrt{t} & 2 \\ 0 & 0 \end{pmatrix}$$

und

$$e^{A(t)} \cdot \dot{A}(t) = \begin{pmatrix} 0 & \frac{e^{1+\sqrt{t}} + e^{1-\sqrt{t}}}{2} \\ 0 & \frac{e^{1+\sqrt{t}} - e^{1-\sqrt{t}}}{2\sqrt{t}} \end{pmatrix}$$

ist. Wir müssen uns bei diesem Ansatz also begnügen mit linearen homogenen Differentialgleichungen mit *konstanten* Koeffizienten.

Alles was uns zu deren theoretischer Lösung jetzt noch fehlt sind Regeln für den Umgang mit Ableitungen von Matrixfunktionen; für die praktische Lösung fehlen natürlich auch noch Verfahren zur effizienten Berechnung der Matrixexponentialfunktion.

Zumindest für Summen und Produkte gelten, wenn man von der Nichtkommutativität der Multiplikation absieht, für matrixwertige Funktionen die üblichen Regeln:

Lemma: a) $F, G: (a, b) \rightarrow \mathbb{R}^{n \times n}$ seien zwei matrixwertige Funktionen auf dem offenen Intervall (a, b) . Dann ist

$$\frac{d}{dt} (F(t) + G(t)) = \dot{F}(t) + \dot{G}(t).$$

b) Für $F: (a, b) \rightarrow \mathbb{R}^{n \times m}$ und $G: (a, b) \rightarrow \mathbb{R}^{m \times p}$ ist

$$\frac{d}{dt} (F(t)G(t)) = \dot{F}(t) \cdot G(t) + F(t) \cdot \dot{G}(t).$$

c) Für einen konstanten Vektor $\vec{v} \in \mathbb{R}^n$ ist

$$\frac{d}{dt} (F(t) \cdot \vec{v}) = \dot{F}(t) \cdot \vec{v}.$$

Beweis: a) Sind $f_{ij}(t)$ und $g_{ij}(t)$ die Komponenten der Matrizen F und G , so sind die Summen $f_{ij}(t) + g_{ij}(t)$ die Komponenten von $F + G$, und deren Ableitung ist die Summe der Ableitungen.

b) Die (i, j) -Komponente von FG ist die Funktion $\sum_{\nu=1}^m f_{i\nu}(t) \cdot g_{\nu j}(t)$, und deren Ableitung ist

$$\sum_{\nu=1}^m \left(\dot{f}_{i\nu}(t) \cdot g_{\nu j}(t) + f_{i\nu}(t) \cdot \dot{g}_{\nu j}(t) \right),$$

die (i, j) -Komponente von $\dot{F}(t) \cdot G(t) + F(t) \cdot \dot{G}(t)$.

c) Ist der Spezialfall $n + m$ und $p = 1$ von b), wobei zusätzlich noch G eine konstante Funktion ist, so daß alle Terme mit $\dot{G}(t)$ verschwinden. ■

Damit haben wir alles zusammen und können zeigen

Satz: Die sämtlichen Lösungen des homogenen Differentialgleichungssystems $\dot{y} = Ay$ sind genau die Funktionen $t \mapsto e^{At} \vec{y}_0$ mit $\vec{y}_0 \in \mathbb{R}^n$.

Beweis: Da e^{At} die Ableitung Ae^{At} hat, ist die Ableitung der vektorwertigen Funktion $\vec{f}(t) = e^{At} \cdot \vec{y}_0$ nach der zuletzt bewiesenen Formel gleich $A \cdot e^{At} \cdot \vec{y}_0$, also in der Tat gleich $A \cdot \vec{f}(t)$.

Nun sei $\vec{y}: (a, b) \rightarrow \mathbb{R}^n$ irgendeine differenzierbare vektorwertige Funktion mit der Eigenschaft, daß $\vec{y}'(t) = A \cdot \vec{y}(t)$ ist. Wir betrachten die Funktion $\vec{g}(t) = e^{-At} \cdot \vec{y}(t)$. Deren Ableitung ist

$$\vec{g}'(t) = -A \cdot e^{-At} \cdot \vec{y}(t) + e^{-At} \cdot \vec{y}'(t) = -A \cdot e^{-At} \cdot \vec{y}(t) + e^{-At} \cdot A \vec{y}(t) = \vec{0},$$

denn die Matrix A ist mit e^{-At} vertauschbar. Damit haben alle Komponenten von \vec{g} die Ableitung Null, sind also konstant, und somit ist $\vec{g}(t) \stackrel{\text{def}}{=} \vec{y}_0$ ein konstanter Vektor mit der Eigenschaft, daß

$$\vec{y}_0 = e^{-At} \cdot \vec{y}(t), \quad \text{d.h.} \quad \vec{y}(t) = e^{At} \cdot \vec{y}_0,$$

wie behauptet. ■