

Material zur Vorlesung
Höhere Mathematik IV
für Elektro-Techniker

Technische Universität Berlin
Fachbereich Mathematik
Betreuungsgruppe Elektrotechnik

KAPITEL 14

INTEGRALTRANSFORMATIONEN

- 14.1 LAPLACETRANSFORMATION
- 14.2 FOURIERTRANSFORMATION
- 14.3 Z - TRANSFORMATION

14.1 Die LAPLACETRANSFORMATION

Motivation. In diesem und den folgenden Abschnitten wollen wir gewisse Integraltransformationen betrachten. Diese transformierten Funktionen $f(t)$ in neue Funktionen $F(s)$ und zwar auf folgende Weise:

$$F(s) = \int_a^{\beta} f(t) k(s, t) dt,$$

wobei $k(s, t)$ eine feste, von $f(t)$ unabhängige Funktion ist, die sogenannte Kernfunktion.

Das Interesse an solchen Transformationen rührt vor allem aus den beiden nachstehend beschriebenen Sachverhalten: Sei $f(t)$ die gesuchte Lösung einer Differentialgleichung. Kann man diese Differentialgleichung durch eine äquivalente Gleichung für die transformierte Funktion $F(s)$ ersetzen, die einfacher zu lösen ist, und beherrscht man die Transformation $f(t) \leftrightarrow F(s)$ in beiden Richtungen, so hat man ein Hilfsmittel zur Lösung von Differentialgleichungen. Wir werden sehen, daß man dadurch insbesondere bei inhomogenen Differentialgleichungen mit konstanten Koeffizienten das mühsame Verfahren der Variation der Konstanten ersetzen kann.

Zum anderen: Die FOURIERREIHEN, zum Beispiel

$$f(t) = \sum_n b_n \sin nt,$$

kann man interpretieren als Überlagerung von Eigenschwingungen $\sin nt$ eines Systems mit gewichtenden Intensitäten b_n , die sich wie folgt berechnen:

$$b_n = \int_0^{2\pi} f(\tau) \frac{1}{\pi} \sin n\tau d\tau.$$

Besitzt das System aber nicht nur abzählbar viele Eigenschwingun-

gen, sondern ein kontinuierliches Spektrum, ist also für jedes s - entsprechend dem obigen n - $k(s, t) = \frac{1}{\pi} \sin st$ eine Eigenschwingung, so erhält man deren Intensitätsanteil als

$$b(s) = \int f(t) \frac{1}{\pi} \sin st dt.$$

Zur Darstellung von $f(t)$ muß man dann - statt der Summation über n - über s integrieren. Vgl. auch p. 14.2.7/8.

Definition der LAPLACETRANSFORMATION. Bei der Lösungstheorie linearer Gleichungen haben wir gesehen, daß man der Einfachheit halber auch komplexwertige Lösungen (einer reellen Variablen) zulassen sollte. Wir definieren die LAPLACETRANSFORMATION deshalb für komplexwertige Funktionen $f(t)$ einer reellen Variablen $t \in [0, \infty)$ durch

$$\mathcal{L}\{f\}(z) = \int_0^{\infty} f(t) e^{-zt} dt,$$

falls das Integral existiert, d.h. falls

$$\lim_{\beta \rightarrow \infty} \int_0^{\beta} f(t) e^{-zt} dt$$

existiert. Die neue Funktion $\mathcal{L}\{f\}(z)$ einer komplexen Variablen heißt die LAPLACETRANSFORMIERTE von $f(t)$.

Bemerkung. Bevor wir die LAPLACETRANSFORMIERTE spezieller Funktionen berechnen, wollen wir an die Motivation anknüpfend folgende "naive" Rechnung durchführen:

$$\begin{aligned} \mathcal{L}\{f'\}(z) &= \int_0^{\infty} f'(t) e^{-zt} dt \\ &= f(t) e^{-zt} \Big|_0^{\infty} - \int_0^{\infty} f(t) (-z) e^{-zt} dt \end{aligned}$$

(partielle Integration)

$$= 0 - f(0) + z \int_0^{\infty} f(t) e^{-zt} dt$$

(falls $\text{Re } z > 0$ und z.B. f beschränkt)

$$= z \mathcal{L}\{f\}(z) - f(0)$$

Sie sehen, daß der *Differentiation* von $f(t)$ im wesentlichen die *Multiplikation* von $\mathcal{L}\{f\}(z) = F(z)$ mit z entspricht: Die LAPLACE-Transformation verwandelt *Differentialgleichungen* in *algebraische Gleichungen*! Das ist das Geheimnis ihres Erfolges.

Erste Beispiele. Sei a eine komplexe Zahl und $f(t) = e^{at}$. Dann hat man

$$\begin{aligned} \int_0^\beta f(t) e^{-zt} dt &= \int_0^\beta e^{(a-z)t} dt \\ &= \frac{1}{a-z} e^{(a-z)t} \Big|_0^\beta \\ &= \frac{1}{a-z} e^{(a-z)\beta} - \frac{1}{a-z} \\ &= \frac{1}{a-z} e^{x\beta} e^{iy\beta} + \frac{1}{z-a} \end{aligned}$$

wobei $a-z = x+iy$ mit x, y reell. Der Grenzwert für $\beta \rightarrow +\infty$ existiert also genau dann, wenn $x = \text{Re}(a-z) < 0$. Wir erhalten

$$\mathcal{L}\{e^{at}\}(z) = \frac{1}{z-a} \quad \text{für} \quad \text{Re } z > \text{Re } a$$

Die LAPLACE-transformierte von e^{at} ist also auf einer Halbebene definiert.

Aus der Definition von \mathcal{L} ergibt sich die ebenso triviale wie wichtige Linearität: $\mathcal{L}\{f+g\} = \mathcal{L}\{f\} + \mathcal{L}\{g\}$ und $\mathcal{L}\{af\} = a\mathcal{L}\{f\}$ für komplexe Skalare a . Daher ergeben sich aus unserem obigen Resultat sofort noch die Transformierten

$$\begin{aligned} \mathcal{L}\{\sin at\}(z) &= \mathcal{L}\left\{\frac{1}{2i} \left(e^{iat} - e^{-iat} \right)\right\}(z) \\ &= \frac{1}{2i} \cdot \left\{ \frac{1}{z-ia} - \frac{1}{z+ia} \right\} \end{aligned}$$

$$\mathcal{L}\{\sin at\}(z) = \frac{a}{z^2 + a^2}, \quad \text{Re } z > | \text{Im } a |,$$

und analog

$$\mathcal{L}\{\cos at\}(z) = \frac{z}{z^2 + a^2}, \quad \text{Re } z > | \text{Im } a |,$$

Für $a = 0$ ergibt sich

$$\mathcal{L}\{1\}(z) = \frac{1}{z}, \quad \text{Re } z > 0,$$

und wir beweisen durch Induktion über n noch

$$\mathcal{L}\{t^n\}(z) = \frac{n!}{z^{n+1}}, \quad \text{Re } z > 0.$$

Für $n=0$ ist das gerade die vorangegangene Formel. Für $f(t) = t^n$ läßt sich aber die "naive" Rechnung in der Bemerkung exakt durchführen und liefert

$$\mathcal{L}\{nt^{n-1}\}(z) = z \mathcal{L}\{t^{n-1}\} - 0,$$

also

$$\mathcal{L}\{t^n\} = \frac{n}{z} \mathcal{L}\{t^{n-1}\}.$$

Daraus folgt die Induktionsbehauptung.

Eine Klasse von LAPLACE-transformierbaren Funktionen.

Wir untersuchen nun die Frage der Existenz der LAPLACE-transformierten genauer. Damit

$$\lim_{\beta \rightarrow \infty} \int_0^\beta f(t) e^{-zt} dt$$

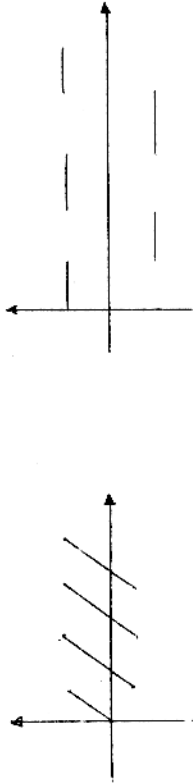
existiert, müssen die Integrale $\int_0^\beta f(t)e^{-zt} dt$ alle existieren, d.h. $f(t)e^{-zt}$ muß integrierbar über jedes Intervall $[0, \beta]$ sein; und der Grenzwert muß endlich sein, d.h. $f(t)$ "darf nicht zu schnell wachsen". Wir formulieren hinreichende Kriterien dafür.

Eine Funktion $f(t)$ heißt *stückweise stetig* auf $[0, \infty)$, wenn für jedes $t \geq 0$ die Grenzwerte

$$f(t+) = \lim_{\tau \rightarrow t} f(\tau) \quad \tau > t$$

$$f(t-) = \lim_{\tau \rightarrow t} f(\tau) \quad \tau < t \quad (\text{nur bei } t \neq 0)$$

existieren und endlich sind, und wenn $f(t)$ in jedem beschränkten Intervall $[0, \beta]$ bis auf endlich viele Punkte definiert und stetig ist. (Also braucht $f(t)$ nicht für *alle* Punkte $t > 0$ definiert zu sein, in jedem beschränkten Intervall dürfen endlich viele Punkte fehlen.) Alle stetigen Funktionen sind natürlich stückweise stetig, aber z.B. auch die bekannten Sprungfunktionen:



Weiter sagen wir $f(t)$ ist von *exponentieller Ordnung* auf $[0, \infty)$, wenn es positive Konstanten C und γ gibt, so daß für alle $t \geq 0$, in denen $f(t)$ definiert ist

$$|f(t)| \leq Ce^{\gamma t}.$$

Beispiele: t^n, e^{at} für komplexes a , und $t^n e^{at}$. Damit haben insbesondere alle Lösungen homogener linearer Differentialgleichungen mit konstanten Koeffizienten exponentielle Ordnung auf $[0, \infty)$.

Satz: Ist $f(t)$ auf $[0, \infty)$ stückweise stetig und von exponentieller Ordnung, so gibt es ein $\gamma \in [-\infty, \infty)$, so daß $\mathcal{L}[f](z)$ für alle z mit $\text{Re}(z) > \gamma$ existiert.
Weiter gilt für reelles x

$$\lim_{x \rightarrow +\infty} \mathcal{L}[f](x) = 0.$$

Beweis: Nach Voraussetzung existieren die Integrale

$$\int_0^\beta f(t)e^{-zt} dt$$

für alle $\beta > 0$. Weiter gibt es positive Konstanten C und γ mit

$$|f(t)| \leq Ce^{\gamma t} \quad \text{für alle } t \geq 0.$$

Dann ist aber mit $z = x + iy, x, y$ reell,

$$\begin{aligned} \int_0^\beta |f(t)e^{-zt}| dt &\leq C \int_0^\beta e^{(\gamma-x)t} |e^{-iyt}| dt \\ &= \frac{C}{\gamma-x} e^{(\gamma-x)t} \Big|_0^\beta \\ &= \frac{C}{x-\gamma} (1 - e^{-(x-\gamma)\beta}). \end{aligned}$$

Für $\beta \rightarrow +\infty$ ist die rechte Seite beschränkt, falls

$$x > \gamma.$$

Also ist $f(t)e^{-zt}$ über $[0, \infty)$ integrierbar, ja sogar absolut integrierbar. Weiter hat man nach obiger Rechnung

$$|\mathcal{L}[f](x)| \leq \frac{C}{x-\gamma}$$

und das geht für $x \rightarrow +\infty$ gegen Null.

Wir präzisieren nun die naive Rechnung betreffend die LAPLACE-transformierte einer Ableitung

Satz: Sei $f(t)$ auf $[0, \infty)$ stetig und $f'(t)$ stückweise stetig von exponentieller Ordnung. Dann ist auch $f(t)$ von exponentieller Ordnung und, falls $\text{Re}(z)$ hinreichend groß ist, gilt

$$\mathcal{L}[f'](z) = z\mathcal{L}[f](z) - f(0).$$

Sind allgemeiner $f^{(1)}(t), \dots, f^{(n-1)}(t)$ stetig und ist $f^{(n)}(t)$ stückweise stetig und von exponentieller Ordnung, so gilt, falls $\text{Re} z$ hinreichend groß ist,

$$\begin{aligned} \mathcal{L}[f^{(n)}](z) &= z^n \mathcal{L}[f](z) - z^{n-1} f(0) - z^{n-2} f'(0) - \dots - f^{(n-1)}(0). \end{aligned}$$

Beweis: Es ist $|f'(t)| \leq C e^{\gamma t}$ mit positiven Konstanten C und γ . Dann ist aber

$$\begin{aligned} |f(t)| &= |f(0) + \int_0^t f'(s) ds| \\ &\leq |f(0)| + \int_0^t C e^{\gamma s} ds \\ &\leq |f(0)| + \frac{C}{\gamma} e^{\gamma t} - \frac{C}{\gamma} \\ &\leq (|f(0)| + \frac{C}{\gamma}) e^{\gamma t} = \tilde{C} e^{\gamma t}. \end{aligned}$$

Daher ist auch $f(t)$ von exponentieller Ordnung. Weiter ist

$$\begin{aligned} \mathcal{L}[f'](z) &= \int_0^\infty f'(t) e^{-zt} dt \\ &= \lim_{\beta \rightarrow \infty} \int_0^\beta f'(t) e^{-zt} dt \\ &= \lim_{\beta \rightarrow \infty} f(t) e^{-zt} \Big|_0^\beta + \lim_{\beta \rightarrow \infty} z \int_0^\beta f(t) e^{-zt} dt \\ &= z \int_0^\infty f(t) e^{-zt} dt + \lim_{\beta \rightarrow \infty} f(t) e^{-z\beta} - f(0). \end{aligned}$$

Aber $|f(t) e^{-zt}| \leq \tilde{C} e^{(\gamma - \text{Re} z)t} \rightarrow 0$ für $\text{Re} z > \gamma$.

Bevor wir ein ausführliches Beispiel zur Anwendung der LAPLACE-Transformation geben, brauchen wir noch zwei weitere Formeln:

Dämpfungssatz. Sei $f(t)$ stückweise stetig und von exponentieller Ordnung auf $[0, \infty)$. Sei a eine komplexe Zahl. Dann gilt

$$\mathcal{L}[e^{at} f(t)](z) = \mathcal{L}[f](z-a)$$

Beweis:

$$\begin{aligned} \mathcal{L}[e^{at} f(t)](z) &= \int_0^\infty e^{at} f(t) e^{-zt} dt \\ &= \int_0^\infty f(t) e^{-(z-a)t} dt \\ &= \mathcal{L}[f](z-a). \end{aligned}$$

Satz: Unter denselben Voraussetzungen ist

$$\mathcal{L}[t f(t)](z) = -\frac{d}{dz} \mathcal{L}[f(t)](z).$$

Die folgende Lösungsmethode mittels der LAPLACETRANSFORMATION ist - so aufwendig sie erscheinen mag - viel einfacher. Wenn Sie's nicht glauben, probieren Sie den alten Weg!

Zunächst ist wegen der Anfangsbedingungen

$$\begin{aligned} \mathcal{L}[y'] &= z\mathcal{L}[y] - y(0) = z\mathcal{L}[y] \\ \mathcal{L}[y''] &= z^2\mathcal{L}[y] - zy(0) - y'(0) = z^2\mathcal{L}[y] + 1 \\ \mathcal{L}[2t] &= \frac{2}{z^2} \end{aligned}$$

$$\begin{aligned} \mathcal{L}[e^{-2t} \cos 3t](z) &= \mathcal{L}[\cos 3t](z+2) \\ &= \frac{z+2}{(z+2)^2+9} \end{aligned}$$

Die Differentialgleichung liefert damit

$$\begin{aligned} z^2\mathcal{L}[y] + 1 + 4z\mathcal{L}[y] + 13\mathcal{L}[y] &= \frac{2}{z^2} + 3\frac{z+2}{(z+2)^2+9} \\ \mathcal{L}[y] &= -\frac{1}{z^2+4z+13} + \frac{2}{z^2(z^2+4z+13)} + \frac{3(z+2)}{(z^2+4z+13)^2} \end{aligned}$$

und die transformierte Gleichung für y ist bereits gelöst: Aber wir brauchen natürlich y(t) und nicht L[y](z).

Unglücklicherweise ist bei unseren bisherigen Formeln aber kein y(t), dessen LAPLACETRANSFORMIERTE die rechte Seite ist: die wird man überhaupt in keiner Tafel finden, aber das ist auch nicht nötig: wir werden mit etwas Arbeit ein solches y(t) konstruieren. (Dies ist die einzige Stelle, wo wir etwas arbeiten müssen!)

Zunächst ist

$$\begin{aligned} -\frac{1}{z^2+4z+13} &= -\frac{1}{3} \frac{3}{(z+2)^2+9} \\ &= \mathcal{L}\left[-\frac{1}{3} e^{-2t} \sin 3t\right] \end{aligned}$$

Beweis:

$$\begin{aligned} \mathcal{L}\{t f(t)\}(z) &= \int_0^\infty f(t) t e^{-zt} dt \\ &= - \int_0^\infty f(t) \frac{\partial}{\partial z} e^{-zt} dt \\ &= - \frac{d}{dz} \int_0^\infty f(t) e^{-zt} dt \\ &= - \frac{d}{dz} \mathcal{L}\{f\}(z). \end{aligned}$$

Anwendung der LAPLACETRANSFORMATION: Ein Beispiel.

Wir betrachten die Differentialgleichung für eine erzwungene Schwingung:

$$L_i'' + R i' + \frac{1}{C} i = U'(t), \quad i(0) = 0, \quad i'(0) = -1$$

mit

$$L = 1, \quad R = 4, \quad \frac{1}{C} = 13, \quad u'(t) = 2t + 3e^{-2t} \cos 3t.$$

Wir schreiben y statt i:

$$\begin{aligned} y'' + 4y' + 13y &= 2t + 3e^{-2t} \cos 3t \\ Y(0) = 0, \quad Y'(0) &= -1. \end{aligned}$$

Beachten Sie: Wir können das lösen! Wir müssen dazu zunächst die charakteristische Gleichung lösen, um ein Fundamentalsystem der homogenen Gleichung zu erhalten, dann Variation der Konstanten durchführen, um eine spezielle Lösung der inhomogenen Gleichung zu berechnen, und schließlich mit einem linearen Gleichungssystem die richtige Kombination zur Erfüllung der Anfangswerte ermitteln.

und mit einem kleinen Trick

$$\begin{aligned} \frac{3(z+2)}{(z^2+4z+13)^2} &= \frac{3(z+2)}{((z+2)^2+9)^2} \\ &= -\frac{3}{2} \frac{d}{dz} \frac{1}{(z+2)^2+9} \\ &= -\frac{3}{2} \frac{d}{dz} \mathcal{L}\left[\frac{1}{3} e^{-2t} \sin 3t\right] \\ &= \mathcal{L}\left[\frac{1}{2} t e^{-2t} \sin 3t\right]. \end{aligned}$$

Mühsamer ist der mittlere Term. Wir verwenden Partialbruchzerlegung:

$$\begin{aligned} \frac{2}{z^2(z^2+4z+13)} &= \frac{A}{z} + \frac{B}{z^2} + \frac{Cz+D}{z^2+4z+13} \\ \rightarrow \quad Az(z^2+4z+13) + B(z^2+4z+13) + (Cz+D)z^2 &= 2 \end{aligned}$$

$$A+C=0$$

$$4A+B+D=0$$

$$13A+4B=0$$

$$13B=2$$

$$A = -\frac{8}{169}, \quad B = \frac{2}{13}, \quad C = \frac{8}{169}, \quad D = \frac{6}{169}$$

Damit wird

$$\begin{aligned} \frac{2}{z^2(z^2+4z+13)} &= -\frac{8}{169} \cdot \frac{1}{z} + \frac{2}{13} \cdot \frac{1}{z^2} + \frac{1}{169} \cdot \frac{8z+6}{z^2+4z+13} \\ &= -\frac{8}{169} \cdot \frac{1}{z} + \frac{2}{13} \cdot \frac{1}{z^2} + \frac{8}{169} \cdot \frac{z+2}{(z+2)^2+9} \\ &\quad - \frac{10}{3 \cdot 169} \cdot \frac{3}{(z+2)^2+9} \\ &= -\frac{8}{169} \mathcal{L}[1] + \frac{2}{13} \mathcal{L}[t] + \frac{8}{169} \mathcal{L}[e^{-2t} \cos 3t] \\ &\quad - \frac{10}{507} \mathcal{L}[e^{-2t} \sin 3t] \end{aligned}$$

Setzen wir also

$$\begin{aligned} Y(t) &= -\frac{1}{3} e^{-2t} \sin 3t + \frac{1}{2} t e^{-2t} \sin 3t \\ &\quad - \frac{8}{169} + \frac{2}{13} t + \frac{8}{169} e^{-2t} \cos 3t - \frac{10}{507} e^{-2t} \sin 3t \\ &= -\frac{179}{507} e^{-2t} \sin 3t + \frac{8}{169} e^{-2t} \cos 3t + \frac{1}{2} t e^{-2t} \sin 3t \\ &\quad + \frac{2}{13} t - \frac{8}{169}. \end{aligned}$$

so ist $\mathcal{L}[y](z)$ gerade die richtige Funktion, und wir haben die gesuchte Lösung. *Haben wir sie wirklich? Kann nicht $\mathcal{L}[y](z) = \mathcal{L}[\tilde{Y}](z)$ für verschiedene Funktionen $y(t)$ und $\tilde{Y}(t)$ sein? Wenn das der Fall ist - wie können wir sicher sein, mit dem oben mühsam konstruierten $y(t)$ das richtige, nämlich die Lösung unseres Anfangswertproblems gefunden zu haben?*

Wir werden im folgenden Abschnitt sehen, daß die LAPLACETRANSFORMATION im wesentlichen eineindeutig ist, so daß wir diese Sorge nicht haben müssen. Die Transformation und Rücktransformation von Funktionen ist das Hauptproblem bei der obigen wie bei den meisten Anwendungen der LAPLACETRANSFORMATION. Einige wichtige weitere Sätze (=Formeln) dazu geben wir weiter unten. Zur effektiven praktischen Arbeit braucht man darüber hinaus aber eine gute Tafel und einige Übung.

Eindeutigkeit. Wir zitieren ohne (den ziemlich aufwendigen) Beweis den folgenden

Satz von LERCH. Seien $f(t)$ und $g(t)$ stückweise stetig und von exponentieller Ordnung auf $[0, \infty)$. Gibt es ein reelles α mit

$$\mathcal{L}[f](z) = \mathcal{L}[g](z)$$

für alle z mit $\operatorname{Re} z > \alpha$, dann gilt

$$f(t) = g(t)$$

in allen $t \geq 0$, in denen beide Funktionen definiert und stetig sind.

Die Bedeutung dieses Satzes sollte oben deutlich geworden sein.

Die Rücktransformation rationaler Funktionen. Bei dem vorstehenden expliziten Beispiel trat das Problem der Rücktransformation einer Funktion der Form $F(z) = \frac{P(z)}{Q(z)}$ mit Polynom $p(z)$ und $q(z)$ auf. Wir wollen annehmen, daß diese Darstellung "ausgekürzt" ist, d.h. daß $p(z)$ und $q(z)$ keine gemeinsamen Nullstellen haben. Ist $F(z)$ die LAPLACETRANSFORMIERTE einer Funktion $f(t)$ von exponentieller Ordnung, so muß für reelles x $\lim_{x \rightarrow +\infty} F(x) = 0$ gelten, d.h. der Grad von $q(z)$ muß höher als der Grad von $p(z)$ sein. Dann ist $F(z)$ nach Partialbruchzerlegung aber eine Summe von Termen der Form

$$\frac{a_i}{(z-z_i)^k}$$

mit komplexen Zahlen a_i, z_i und mit $1 \leq k \leq m_i$. Dabei sind die z_i die Nullstellen von $q(z)$, und m_i ist die Vielfachheit der Nullstellen z_i :

$$q(z) = A(z-z_1)^{m_1} \dots (z-z_n)^{m_n}$$

Man sagt auch: z_i ist ein Pol der Ordnung m_i von $F(z)$. Wegen

$$\mathcal{L} \left[\frac{a_i}{(k-1)!} z_i^t t^{k-1} \right] = \frac{a_i}{(z-z_i)^k}$$

ist das Problem der Rücktransformation von $F(z)$ bis auf die Durchführung der Partialbruchzerlegung erledigt.

Das Wachstumsverhalten von

$$\frac{a_i}{(k-1)!} t^{k-1} z_i^t = \frac{a_i}{(k-1)!} x_i^t (\cos y_i t + i \sin y_i t)$$

für $t \rightarrow +\infty$ wird wesentlich von $x_i = \operatorname{Re}(z_i)$ bestimmt: Ist der Realteil > 0 , so "explodiert" die Funktion, ist er < 0 , so klingt sie ab.

Die Imaginärteile y_i hingegen bestimmen das Frequenzverhalten. Die Kenntnis der Pole von $F(z)$ erlaubt also bereits qualitative Rückschlüsse auf das Verhalten der Zeitfunktion $f(t)$.

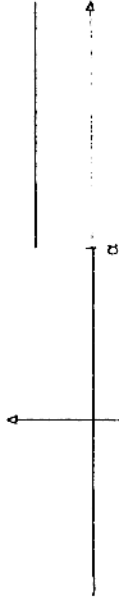
Es ist übrigens nicht immer nötig, ja nicht einmal wünschenswert, die Partialbruchzerlegung im Komplexen bis zu Nennerpolynom $(z-z_i)^k$ durchzuführen, da man z.B. für reelles a, b mit $a^2 < b$ hat

$$\frac{1}{z^2 + 2az + b} = \frac{1}{(z+a)^2 + b-a^2} = \mathcal{L} \left[e^{-at} \frac{\sin \sqrt{b-a^2} t}{\sqrt{b-a^2}} \right]$$

Die Transformierte der Sprungfunktion. Zur Beschreibung von un stetigen Vorgängen, zum Beispiel der Einschaltung eines Stromkreises zur Zeit $t = \alpha$ ist die Funktion

$$u_\alpha(t) := \begin{cases} 0 & t \leq \alpha \\ 1 & t > \alpha \end{cases}$$

sehr nützlich.



Beginnt zu diesem Zeitpunkt α etwa eine Sinus-Wechselstromschwingung, so ist sie gegeben durch

$$u_\alpha(t) \sin(t - \alpha).$$

Wir wollen stets $\alpha > 0$ voraussetzen. Dann gilt allgemein der

Verschiebungssatz. Ist $f(t) = u_\alpha(t)g(t - \alpha)$, $\alpha > 0$, stückweise stetig und von exponentieller Ordnung, so gilt

$$\mathcal{L}\{f\}(z) = e^{-\alpha z} \mathcal{L}\{g\}(z)$$

Beweis:

$$\begin{aligned} \mathcal{L}\{f\}(z) &= \int_0^\infty u_\alpha(t)g(t - \alpha)e^{-zt} dt \\ &= \int_\alpha^\infty g(t - \alpha)e^{-zt} dt \end{aligned}$$

und mit $t - \alpha = s$

$$\begin{aligned} &= \int_0^\infty g(s)e^{-z(s + \alpha)} ds \\ &= \int_0^\infty g(s)e^{-zs}e^{-\alpha z} ds \\ &= e^{-\alpha z} \mathcal{L}\{g\}(z). \end{aligned}$$

Beispiel:

$$\begin{aligned} \mathcal{L}\{u_\alpha(t) \cos t\} &= \mathcal{L}\{u_\alpha(t) \cos(t + \alpha - \alpha)\} \\ &= e^{-\alpha z} \mathcal{L}\{\cos(t + \alpha)\} \\ &= e^{-\alpha z} \mathcal{L}\{\cos \alpha \cos t - \sin \alpha \sin t\} \\ &= e^{-\alpha z} \left\{ \cos \alpha \frac{z}{z^2 + 1} - \sin \alpha \frac{1}{z^2 + 1} \right\} \\ &= \frac{e^{-\alpha z}}{z^2 + 1} (z \cos \alpha - \sin \alpha) \end{aligned}$$

Der Faltungssatz. Der folgende Satz ist aus theoretischen Gründen sehr wichtig. Er hilft auch bei der Rücktransformation von Produktfunktionen, wird dafür aber wegen der erforderlichen Integration wenig verwendet.

Faltungssatz. Seien $f(t)$, $g(t)$ stückweise stetig und von exponentieller Ordnung auf $[0, \infty)$. Dann gilt

$$\mathcal{L}\left\{\int_0^t f(t-s)g(s)ds\right\} = \mathcal{L}\{f\}\mathcal{L}\{g\}.$$

Das Integral auf der linken Seite heißt auch die *Faltung* und wird mit $f * g$ bezeichnet:

$$f * g(t) = \int_0^t f(t-s)g(s)ds.$$

Mit dieser Notation schreibt sich die obige Formel als

$$\mathcal{L}\{f * g\} = \mathcal{L}\{f\}\mathcal{L}\{g\}.$$

Beweis des Faltungssatzes:

$$\begin{aligned} \mathcal{L}\left[\int_0^t f(t-s)g(s)ds\right] &= \int_0^\infty \left(\int_0^t f(t-s)g(s)ds\right) e^{-zt} dt \\ &= \int_0^\infty \int_0^\infty e^{-zt} f(t-s)g(s)ds dt. \end{aligned}$$

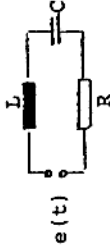
Nach der folgenden Skizze

so folgt aus dem Faltungssatz

$$y(t) = \int_0^t g(t-s)h(s)ds.$$

Ist $n=2$, so ist die Bestimmung von $g(t)$ sehr leicht und führt auf dieselbe Formel für $y(t)$, wie man sie auch nach der Methode der Variation der Konstanten findet. Die Funktion $g(t-s)=:K(t,s)$ heißt auch eine GREENsche Funktion für das Anfangswertproblem.

Beispiel: Netzwerktheorie. Nach der KIRCHHOFFSchen Regel gilt für einen einfachen Stromkreis mit dem Strom $i(t)$ die Gleichung



$$L i'(t) + R i(t) + \frac{1}{C} \int i(\tau) d\tau = e(t)$$

Nimmt man an, daß

$$i(t) = 0 \quad \text{für} \quad t \leq 0,$$

so folgt durch Anwendung der LAPLACETRANSFORMATION

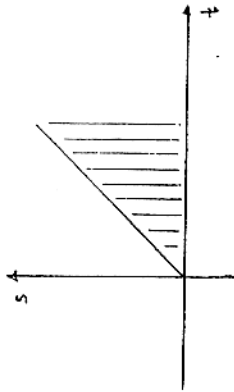
$$Lz \mathcal{L}\{i\} + R \mathcal{L}\{i\} + \frac{1}{Cz} \mathcal{L}\{i\} = \mathcal{L}\{e\},$$

wobei wir für das Integral den Faltungssatz mit $f(t) = 1$ und $g(t) = i(t)$ bemüht haben. Es ist üblich, die neue Variable s statt z zu nennen und $\mathcal{L}\{i\} =: I, \mathcal{L}\{e\} =: E$ zu setzen. Dann folgt

$$E(s) = (Ls + R + \frac{1}{sC}) I(s)$$

$$= : Z(s) I(s)$$

$Z(s)$ heißt die Impedanz des Stromkreises. Bei komplizierten Netzwerken erhält man ein sehr einfaches System Kirchhoffscher Gleichungen



ist das aber nach Vertauschung der Integrationsreihenfolge

$$\begin{aligned} &= \int_0^\infty \int_0^\infty e^{-zt} f(t-s) g(s) dt ds \\ &= \int_0^\infty g(s) \int_0^\infty e^{-z(s+u)} f(u) du ds \quad \text{mit} \quad t-s =: u \\ &= \int_0^\infty g(s) e^{-zs} ds \int_0^\infty f(u) e^{-zu} du \\ &= \mathcal{L}\{f\} \mathcal{L}\{g\}. \end{aligned}$$

Bemerkung: Wir haben nicht gezeigt, daß die Vertauschung der Integrationsreihenfolge zulässig ist. Sie ist's!

Beispiel. Lineare Differentialgleichungen: Wir betrachten die inhomogene lineare Differentialgleichung

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = h(t)$$

mit konstanten Koeffizienten und einer stückweise stetigen Funktion $h(t)$ von exponentieller Ordnung auf $[0, \infty)$. Dazu betrachten wir die Anfangsbedingungen

$$y(0) = y'(0) = \dots = y^{(n-1)}(0) = 0.$$

Anwendung der LAPLACETRANSFORMATION ergibt dann sofort

$$\mathcal{L}\{y\} = (z^n + a_1 z^{n-1} + \dots + a_n)^{-1} \mathcal{L}\{h\}.$$

Findet man also ein $g(t)$ mit

$$\mathcal{L}\{g\} = (z^n + a_1 z^{n-1} + \dots + a_n)^{-1},$$

Die LAPLACETRANSFORMATION VON BESSEL-FUNKTIONEN.

Das vorstehend geübte Verfahren, die LAPLACETRANSFORMATION auf Differentialgleichungen anzuwenden, kann man bei schon bekannten Lösungen auch zur Berechnung von transformierten verwenden. Zum Beispiel erhält man aus

$$y'' + y = 0$$

$$z^2 \mathcal{L}[y] + \mathcal{L}[y] = zy(0) + y'(0)$$

$$\mathcal{L}[y] = \frac{zy(0) + y'(0)}{z^2 + 1}$$

oder

Wählt man die Anfangsbedingungen $y(0) = 0, y'(0) = 1$, so folgt

$$\mathcal{L}[\sin t](z) = \frac{1}{1+z^2}$$

Dieses Verfahren läßt sich insbesondere auf die BESSEL-Gleichung der Ordnung 0 anwenden:

Aus

$$t^2 y'' + ty' + (t^2 - 0)y = 0$$

oder

$$t(y'' + y) + y' = 0$$

erhält man

$$0 = -\frac{d}{dz} \mathcal{L}[y'' + y] + z \mathcal{L}[y] - y(0)$$

$$= -\frac{d}{dz} (z^2 \mathcal{L}[y] - zy(0) - y'(0) + \mathcal{L}[y])$$

$$+ z \mathcal{L}[y] - y(0)$$

$$= - (1 + z^2) \frac{d}{dz} \mathcal{L}[y] - z \mathcal{L}[y]$$

$$E_1 = Z_{11} I_1 + \dots + Z_{1n} I_n$$

$$E_n = Z_{n1} I_1 + \dots + Z_{nn} I_n$$

wobei $E_k(s)$ die (LAPLACETRANSFORMIERTE DER) SPANNUNGSQUELLE in der k-ten Masche ist, $I_k(s)$ der - willkürlich gerichtete - Strom in dieser Masche und $Z_{kl}(s)$ die Impedanz am gemeinsamen Zweig der k-ten und l-ten Masche (mit positivem Vorzeichen, falls I_k und I_l dort gleichgerichtet sind, andernfalls mit negativem Vorzeichen). $Z_{kk}(s)$ ist die gesamte Impedanz der k-ten Masche.

Aus obigem Gleichungssystem kann man die Ströme $I_k(s)$ auf algebraischem Wege ausrechnen*). (Man hat kein Differentialgleichungssystem wie für die $i_k(t)$). Sie ergeben sich in der Form

$$I_k(s) = G_{1k}(s) E_1(s) + \dots + G_{nk}(s) E_n(s)$$

mit sogenannten Übertragungsfunktionen $G_{ik}(s)$ **).

Die Rücktransformation ist für viele Untersuchungen der Netzwerktheorie gar nicht erforderlich. Wird sie doch benötigt, so ist zu berücksichtigen, daß der Gleichung

$$I(s) = G(s) E(s)$$

nach dem Faltungssatz die Gleichung

$$i(t) = g(t) * e(t) = \int_0^t g(t-\tau) e(\tau) d\tau$$

entspricht, also nicht einfach

$$i(t) = g(t) e(t).$$

*) Dabei setzen wir voraus, daß das lineare Gleichungssystem eine eindeutige Lösung besitzt, d.h. daß die Matrix (Z_{kl}) invertierbar ist. Für den anderen, sog. anomalen Fall vgl. DOETSCH, Seite 81 folgende

**) Dabei ist (G_{ik}) die zu (Z_{ik}) inverse Matrix.

Also hat man

$$\frac{d}{dz} \mathcal{L}[Y] = -\frac{z}{1+z^2} \mathcal{L}[Y]$$

und

$$\mathcal{L}[Y] = \frac{c}{\sqrt{1+z^2}}, \quad \text{Re}(z) > \alpha$$

für gewisse Konstanten c und α .

Wir betrachten nun $y(t) = J_0(t)$, vgl. den Abschnitt über die BESSELfunktionen.

Aus dem Satz p. 4.1.6 folgt durch Anwendung auf $y'(t) = J_0'(t)$:

$$\begin{aligned} 0 &= \lim_{x \rightarrow +\infty} \mathcal{L}[y'](x) = \lim_{x \rightarrow +\infty} (x \mathcal{L}[y] - y(0)) \\ &= \lim_{x \rightarrow +\infty} \left(-\frac{xc}{\sqrt{1+x^2}} - 1 \right) = c - 1 \end{aligned}$$

Es folgt $c = 1$ und

$$\boxed{\mathcal{L}[J_0(t)] = \frac{1}{\sqrt{1+z^2}}}$$

Mit Hilfe von Rekursionsformeln erhält man daraus für natürliches k

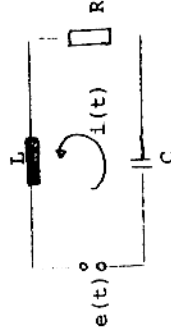
$$\boxed{\mathcal{L}[J_k(t)] = \frac{(\sqrt{1+z^2} - z)^k}{\sqrt{1+z^2}}}$$

vgl. Übungen.

Lineare Systeme, δ -Funktion. Wir wollen nun noch den Anwendungsbereich der LAPLACETRansformation erweitern: auf eine größere Klasse von Gleichungen und, verbunden damit, auf eine allgemeinere Klasse von Funktionen.

Mathematisch betrachtet ist diese Erweiterung ein ziemlich komplizierter Abstraktionsprozeß, auf dessen Einzelheiten wir nicht eingehen können. Die Resultate sind dagegen für die Praxis der Ingenieurmathematik von großem Nutzen gerade weil sie so "abstrakt" sind; sie erlauben es, Systemkomponenten (z.B. Filter, Verstärker etc.) ganzheitlich zu erfassen, ohne ihr Innenleben physikalisch (als Schaltungen) oder mathematisch (als Differential- oder Integralgleichungen) genauer zu analysieren.

Betrachten Sie den einfachen Schaltkreis:



Das mathematische Modell ist eine Integrodifferentialgleichung

$$e(t) = Li'(t) + Ri(t) + \frac{1}{C} \int i(\tau) d\tau.$$

Das Problem sei die Bestimmung von $i(t)$ bei Vorgabe von $e(t)$. Wir haben dann die Situation, daß jedem Eingang (input, Erregung) $e(t)$ auf physikalischem bzw. mathematischem Wege ein Ausgang (output, Antwort) $i(t)$ zugeordnet ist:

$$S : e(t) \rightarrow i(t)$$

Dafür verwendet man in der anwendungsorientierten Literatur auch die Notation:



Die Abbildung (= Zuordnung) S nennen wir auch System oder Übertragungsglied. Wie solche Systeme auch im Detail aussehen mögen, wir verlangen von ihnen zwei Eigenschaften, die z.B. für obiges konkrete Beispiel erfüllt sind: Sie sollen linear und zeitunabhängig sein. Linear bedeutet

$$S[f + g] = S[f] + S[g]$$

und

$$S[\alpha f] = \alpha S[f],$$

ja darüber hinaus wollen wir verlangen, daß unter gewissen, hier nicht näher spezifizierten Konvergenz Voraussetzungen nicht nur

$$S \left[\sum_{k=0}^{\infty} \alpha_k f_k \right] = \sum_{k=0}^{\infty} \alpha_k S[f_k],$$

sondern sogar

$$S \left[\int_a^b f(\tau) g(\tau, t) d\tau \right] = \int_a^b f(\tau) S[g(\tau, t)] d\tau$$

gilt, wobei S stets auf die Variable t wirkt. Genauer ist S also "stetig" und linear.

Zeitunabhängig bedeutet

$$S[f(t-t_0)](\tilde{t}) = S[f(t)](\tilde{t}-t_0)$$

für alle t_0 . Die physikalische Interpretation besagt, daß das System bis auf eine Zeitverschiebung auf den Eingang $f(t)$ in gleicher Weise reagiert, wie auf "denselben" Eingang zu einer späteren Zeit.

Der Mathematiker muß im konkreten Fall natürlich den Definitionsbereich des Systems S angeben und die verlangten Eigenschaften nachweisen. Der Ingenieur muß aufgrund von Einsicht

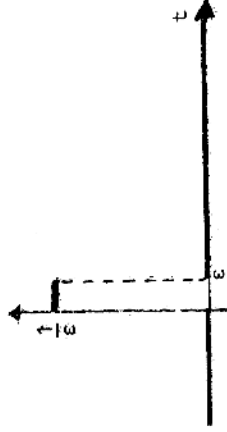
in die physikalisch-mathematische Wirkungsweise oder aufgrund experimentell gestützter Erfahrung wissen, ob ein vorgelegtes System linear und zeitunabhängig arbeitet.

Wir wollen nun die zentrale Eigenschaft solcher Systeme beschreiben und zeigen, daß sie sich für die Behandlung mit der LAPLACE-Transformation anbieten.

Wir betrachten dazu ein lineares, zeitunabhängiges System S, das für "hinreichend anständige" Funktionen $f(t)$ auf $[0, \infty)$ definiert ist.

Für positives ϵ sei

$$v_\epsilon(t) = \begin{cases} \frac{1}{\epsilon} & \text{für } 0 \leq t \leq \epsilon \\ 0 & \text{für } t < 0 \text{ oder } t > \epsilon \end{cases}$$



Diese Funktion beschreibt physikalisch einen (bei kleinem ϵ) zeitlich konzentrierten Impuls.

Sei für beliebiges stückweise stetiges $f(t)$

$$\begin{aligned} f_\epsilon(t) &:= f * v_\epsilon(t) \\ &= \int_0^t f(\tau) v_\epsilon(t-\tau) d\tau \\ &= \int_0^\epsilon f(\tau) v_\epsilon(t-\tau) d\tau, \end{aligned}$$

wobei die letzte Identität aus $v_\epsilon(t) = 0$ für $t < 0$ folgt.

Dann gilt

$$\begin{aligned} S[f_\epsilon] &= \int_0^\infty f(\tau) S[v_\epsilon(t-\tau)] d\tau \\ &= \int_0^\infty f(\tau) S[v_\epsilon](t-\tau) d\tau. \end{aligned}$$

Daher kennt man die Antwort von S auf alle Funktionen $f_\epsilon(t)$, wenn man nur die Antwort $S[v_\epsilon]$ auf $v_\epsilon(t)$ kennt. Natürlich interessiert uns nicht $S[f_\epsilon]$ sondern $S[f]$. Beachten Sie aber, daß

$$\begin{aligned} f_\epsilon(t) &= \int_0^t f(\tau) v_\epsilon(t-\tau) d\tau \\ &= \int_{t-\epsilon}^t f(\tau) \frac{1}{\epsilon} d\tau \\ &= \frac{1}{\epsilon} \int_{t-\epsilon}^t f(\tau) d\tau \\ &= \text{Mittelwert von } f \text{ auf } [t-\epsilon, t]. \end{aligned}$$

Ist $f(t)$ in t_0 stetig, so hat man deshalb

$$\lim_{\epsilon \rightarrow 0} f_\epsilon(t_0) = f(t_0).$$

Wegen der stetigen Linearität von S gilt dann für "hinreichend anständiges" f

$$\lim_{\epsilon \rightarrow 0} S[f_\epsilon] = S[f].$$

Nehmen wir an, daß

$$\lim_{\epsilon \rightarrow 0} S[v_\epsilon](t) =: h(t)$$

existiert, so folgt

$$S[f] = \int_0^\infty f(\tau) h(t-\tau) d\tau.$$

Man nennt $h(t)$ aus ersichtlichen Gründen die *Impulsantwort* des Systems. Sie allein bestimmt also die Reaktion auf (arbiträr) beliebige Eingangsfunktionen $f(t)$.

Um die LAPLACETRANSFORMATION anwenden zu können, verlangen wir noch

$$(*) \quad h(t) = 0 \quad \text{für } t < 0.$$

Dann erhält man

$$S[f](t) = \int_0^t f(\tau) h(t-\tau) d\tau,$$

wobei die Integration nur noch bis t geht, d.h.

$$S[f] = f * h.$$

Ist $f(t) = 0$ für $t \leq 0$, so folgt auch $S[f](t) = 0$ für $t \leq 0$, d.h. das System rührt sich nicht, solange kein Eingang vorhanden ist. Deshalb nennt man Systeme mit (*) auch *kausal*. Die Wirkung eines solchen Systems wird durch die Faltung mit der Impulsantwort gegeben.

Die LAPLACETRANSFORMATION gestaltet diese Situation nach dem Faltungssatz noch einfacher:

$$\mathcal{L}[S[f]] = \mathcal{L}[h] \mathcal{L}[f].$$

Im Bereich der LAPLACETRANSFORMIERTEN FUNKTIONEN WIRKT DAS SYSTEM EINFACH DURCH MULTIPLIKATION MIT DER (LAPLACETRANSFORMIERTE) IMPULSWERT, DER ERGEBTEN ÜBERGANGSFUNKTION.

Wir kommen noch einmal zurück auf den oben vorgenommenen Grenzübergang für $\epsilon \rightarrow 0$:

Für alle $t \neq 0$ ist

$$\lim_{\epsilon \rightarrow 0} v_\epsilon(t) = 0.$$

Da aber der Funktionswert an einer einzelnen Stelle bei der Integration keine Rolle spielt, folgt

$$\int_0^{\infty} f(\tau) \lim_{\epsilon \rightarrow 0} v_{\epsilon}(t-\tau) d\tau = 0,$$

und das ist im allgemeinen verschieden von

$$\lim_{\epsilon \rightarrow 0} \int_0^{\infty} f(\tau) v_{\epsilon}(t-\tau) d\tau = f(t).$$

Prozdem hat man sich angewöhnt, von der "Grenzfunktion"

$$\delta(t) = \lim_{\epsilon \rightarrow 0} v_{\epsilon}(t) \text{ zu sprechen, und ihr die Eigenschaft}$$

$$\int_0^{\infty} f(\tau) \delta(t-\tau) d\tau = f(t)$$

zuzuschreiben, obwohl $\delta(t) = 0$ für alle $t \neq 0$ ist: so groß ist $\delta(0)$! Wählt man insbesondere $f(t) = 1$, so folgt

$$\int_0^{\infty} \delta(\tau) d\tau = 1$$

oder wegen $\delta(-\tau) = \delta(\tau)$ auch

$$\int_{-\infty}^{\infty} \delta(\tau) d\tau = 1.$$

Natürlich gibt es keine solche Funktion, ebenso wie es keine Zahl x mit $x^2 + 1 = 0$ gibt, solange man unter Zahl eben "reelle Zahl" versteht. Wie man aber die reellen Zahlen erweitern kann zu den komplexen, so kann man auch die Funktionen zu den sogenannten verallgemeinerten Funktionen oder Distributionen erweitern. Der Grenzwert $\delta(t)$ der Impulse $v_{\epsilon}(t)$ ist eine solche verallgemeinerte Funktion, DIRAC'sche Deltafunktion oder Einheitsimpuls genannt.

Für eine knappe Darstellung der Distributionen und ihrer fundamentalen Eigenschaften im Zusammenhang mit der LAPLACEtransformation vgl. den Anhang des unten angegebenen Buches von DOETSCH.

Man schreibt

$$S[\delta] = h$$

und erhält für kausale Systeme

$$S[f] = S[f * \delta] = f * S[\delta].$$

Das zeigt noch einmal eindrücklich, wie die Impulsantwort die Reaktion des Systems komplett beschreibt.

Beispiel. Der Frequenzgang eines Systems. Sei $h(t)$ die Impulsantwort des kausalen Systems S . Dann hat man für reelles ω

$$\begin{aligned} S[e^{i\omega t}] &= \int_0^t e^{i\omega \tau} h(t-\tau) d\tau \\ &= \int_0^t e^{i\omega(t-\tau)} h(\tau) d\tau \\ &= e^{i\omega t} \int_0^t e^{-i\omega \tau} h(\tau) d\tau. \end{aligned}$$

Ist das Integral rechts für $t \rightarrow \infty$ konvergent, so folgt weiter

$$\begin{aligned} &= e^{i\omega t} \int_0^{\infty} e^{-i\omega \tau} h(\tau) d\tau - e^{i\omega t} \int_t^{\infty} e^{-i\omega \tau} h(\tau) d\tau \\ &= e^{i\omega t} \mathcal{L}[h](i\omega) - e^{i\omega t} \int_t^{\infty} e^{-i\omega \tau} h(\tau) d\tau. \end{aligned}$$

Das Integral rechts geht für $t \rightarrow \infty$ gegen Null, beschreibt also den Einschwingvorgang des Systems bei aufgeschalteter harmonischer Erregung $e^{i\omega t}$. Der erste Term

$$\mathcal{L}[h](i\omega) e^{i\omega t}$$

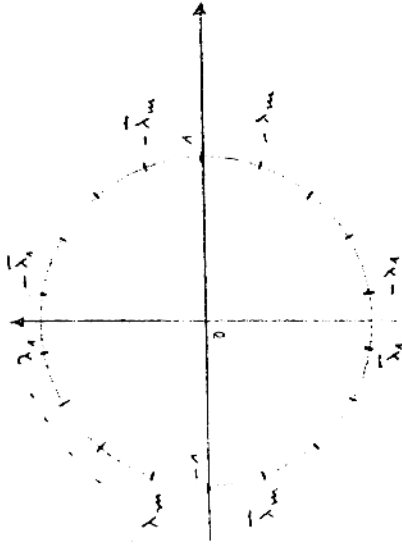
beschreibt den stationären Zustand (große t -Werte) des Systems. Die Antwort ist wieder eine harmonische Schwingung der Frequenz ω . Die Funktion

$$\mathcal{L}[h](i\omega)$$

heißt der Frequenzgang des Systems. Er ist also gerade die LAPLACEtransformierte Impulsantwort (also die Übertragungsfunktion) eingeschränkt auf die imaginäre Achse. Damit hat man auch eine physikalische Interpretation der LAPLACEtransformation!

Beispiel. (BUTTERWORTHfilter)

Sei $n = 2m+1$ eine ungerade natürliche Zahl. Die Nullstellen des Polynoms $1 - z^{2n}$ liegen dann auf dem Einheitskreis im Abstand $\frac{\pi}{n}$:



Weil n ungerade ist, liegt keine Nullstelle auf der imaginären Achse. Sind $\lambda_1, \dots, \lambda_m$ die Nullstellen im zweiten Quadranten, so sind $-\lambda_1, \lambda_1, \dots, \lambda_m, \lambda_m$ und deren Negative gerade die sämtlichen Nullstellen. Setzt man deshalb

$$p(z) := (z+1)(z-\lambda_1)(z-\lambda_1\bar{...})(z-\lambda_m)(z-\lambda_m\bar{...}),$$

so folgt

$$p(z)p(-z) = 1 - z^{2n}.$$

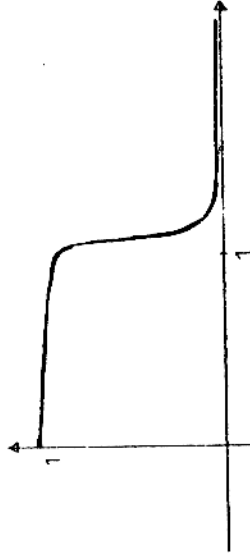
Diese Überlegung benutzen wir, um den Funktionsverlauf von $1/|p(i\omega)|$ zu ermitteln. Es ist

$$\begin{aligned} |p(i\omega)|^2 &= p(i\omega)\overline{p(i\omega)} \\ &= p(i\omega)p(-i\omega) \quad (\text{warum?}) \\ &= p(i\omega)p(-i\omega) \\ &= 1 - (i\omega)^{2n} \\ &= 1 - (-1)^n \omega^{2n} \\ &= 1 + \omega^{2n}, \end{aligned}$$

da n ungerade. Der Graph von

$$\frac{1}{|p(i\omega)|} = \frac{1}{\sqrt{1+\omega^{2n}}}$$

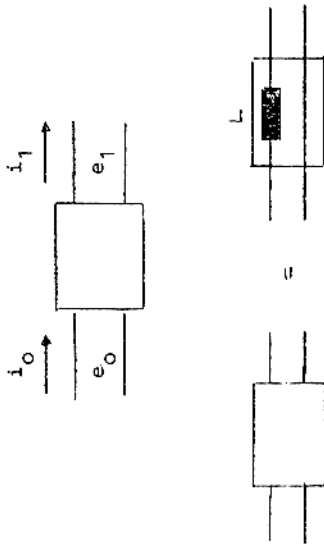
sieht damit so aus



wobei die Stufe um so ausgeprägter ist, je größer n ist. Ein System mit dem Frequenzgang $1/p(i\omega)$ ist deshalb ein ziemlich ideales Tießpaßfilter, ein sogenanntes BUTTERWORTHfilter n -ter Ordnung.

Wir gehen zum Schluß noch auf die Frage ein, wie man ein solches Filter realisieren kann, also auf eine Frage der Netzwerk-synthese. Allgemeiner wollen wir - mit Methoden der Vierpol-theorie - versuchen, eine beliebig vorgegebene Übertragungsfunktion $1/p(s)$ mit $p(s) = a_0 s^n + a_1 s^{n-1} + \dots + a_n$ durch eine (verlustfreie) LC-Kettenschaltung mit einer abschließenden Last zu realisieren.

Wir betrachten zunächst Spannung und Strom an einem einfachen Vierpol:



Für

erhalten wir

$$\begin{aligned} e_0 &= e_1 + Li_1' \\ i_0 &= i_1 \end{aligned}$$

und durch LAPLACEtransformation

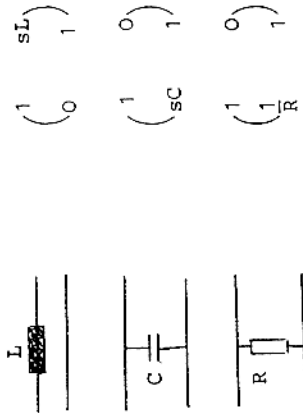
$$\begin{aligned} E_0 &= E_1 + sLI_1 \\ I_0 &= I_1 \end{aligned}$$

Schreiben wir das in Matrixschreibweise, so erhalten wir

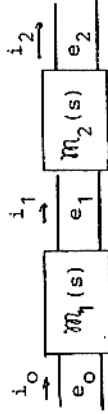
$$\begin{pmatrix} E_0 \\ I_0 \end{pmatrix} = \begin{pmatrix} 1 & sL \\ 0 & 1 \end{pmatrix} \begin{pmatrix} E_1 \\ I_1 \end{pmatrix}$$

Der L-Vierpol ist also durch die Schaltmatrix $M(s) = \begin{pmatrix} 1 & sL \\ 0 & 1 \end{pmatrix}$ gekennzeichnet. (Beachten Sie: Wir haben hier eine Verallgemeinerung linearer Übertragungsglieder. Es werden vektorwertige Funktionen transformiert, und an die Stelle der Übertragungsfunktion tritt jetzt eine Matrix.)

Ähnlich lassen sich die Schaltmatrizen anderer einfacher Vierpole angeben:



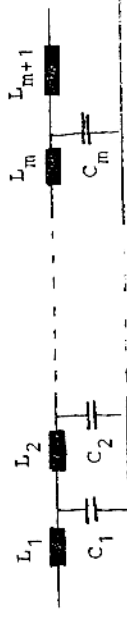
Bei der Hintereinanderschaltung von Vierpolen



multiplizieren sich die Schaltmatrizen:

$$\begin{pmatrix} E_0 \\ I_0 \end{pmatrix} = M_1(s) \begin{pmatrix} E_1 \\ I_1 \end{pmatrix} = M_1(s) M_2(s) \begin{pmatrix} E_2 \\ I_2 \end{pmatrix}$$

Deshalb hat die LC-Kettenschaltung



die Schaltmatrix

$$M(s) = \begin{pmatrix} A(s) & B(s) \\ C(s) & D(s) \end{pmatrix}$$

wird. $A(s)$ und $B(s)$ ermitteln sich durch Matrizenmultiplikation der Schaltmatrizen für die einzelnen L- und C-Glieder, deren man insgesamt so viele braucht, wie der Grad von $p(s)$ angibt.

Ist $\begin{pmatrix} \tilde{A}(s) & \tilde{B}(s) \\ \tilde{C}(s) & \tilde{D}(s) \end{pmatrix}$ die Schaltmatrix der Kette bis einschließlich L_{m+1} , so erhält man für die Kette einschließlich L_{m+1} :

$$\begin{pmatrix} A(s) & B(s) \\ C(s) & D(s) \end{pmatrix} = \begin{pmatrix} \tilde{A}(s) & \tilde{B}(s) \\ \tilde{C}(s) & \tilde{D}(s) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ sC_m & 1 \end{pmatrix} \begin{pmatrix} 1 & sL_{m+1} \\ 0 & 1 \end{pmatrix}$$

Durch Ausmultiplizieren findet man

$$A(s) = \tilde{A}(s) + sC_m \tilde{B}(s)$$

$$B(s) = (\tilde{A}(s) + sC_m \tilde{B}(s)) sL_{m+1} + \tilde{B}(s)$$

und daraus

$$\frac{B(s)}{A(s)} = sL_{m+1} + \frac{\tilde{B}(s)}{\tilde{A}(s) + sC_m \tilde{B}(s)}$$

$$= sL_{m+1} + \frac{1}{sC_m + \frac{1}{\frac{\tilde{B}(s)}{\tilde{A}(s)}}}}$$

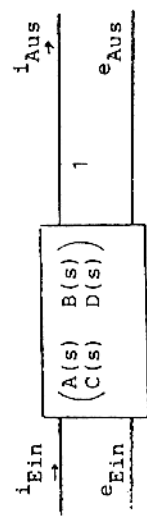
$$= sL_{m+1} + \frac{1}{sC_m + \frac{1}{\dots + \frac{1}{sC_1 + \frac{1}{sL_1}}}}$$

Aus dieser Kettenbauentwicklung kann man also die beteiligten Induktivitäten und Kapazitäten direkt ablesen. Die Herstellung der Kettenbruchentwicklung bei gegebenem $A(s)$ und $B(s)$ ist mittels sukzessiver Polynomdivision sehr einfach (vgl. Beispiele unten). Wir haben also folgendes Verfahren: Man gibt $A(s)$ (y gerade Potenzen $A(s)$) und ungerade Potenzen $B(s)$, entwickelt $A(s)$ in einen Kettenbruch und liest daraus die LC-Kette ab.

mit gewissen Polynomen $A(s)$, $B(s)$, $C(s)$, $D(s)$. Überlegen Sie sich durch Induktion, daß folgendes gilt:

- (i) $A(s)$ und $D(s)$ enthalten nur gerade, $B(s)$ und $C(s)$ nur ungerade Potenzen von s .
- (ii) Der konstante Term von $A(s)$ ist 1.
- (iii) Die höchste in $A(s)$ oder $B(s)$ auftretende Potenz von s ist gleich der Anzahl der von Null verschiedenen Terme $L_1, \dots, L_{m+1}, C_1, \dots, C_m$.

Belasten wir die obige LC-Kette rechts mit einem 1Ω -Widerstand



so wird die Schaltmatrix

$$\begin{pmatrix} A(s) & B(s) \\ C(s) & D(s) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ C(s) + D(s) & 1 \end{pmatrix} = \begin{pmatrix} A(s) + B(s) & B(s) \\ C(s) + D(s) & D(s) \end{pmatrix}$$

Daraus erhält man bei unendlich großen Ausgangswiderstand, also mit $I_{Aus} = 0$:

$$E_{Ein} = (A(s) + B(s)) E_{Aus}$$

Die Übertragungsfunktion (für die Spannung) ist also

$$\frac{1}{p(s)} = \frac{E_{Aus}}{E_{Ein}} = \frac{1}{A(s) + B(s)}$$

Zurück zu unserem Problem: Gegeben sei $p(s)$. Wir wollen einfache L- oder C-Glieder so wählen, daß für die obige Kettenentwicklung mit 1 Ω -Abschluß

$$A(s) + B(s) = p(s)$$

Aber Achtung: Wie üblich beim Problemlösen durch "Losrechnen" ist auch hier nicht klar, ob man wirklich die gesuchte Lösung erhält. Tatsächlich liefern zum Beispiel die Polynome $P_\lambda(s) = (s^2-1)(s^2-\lambda)(s^2+\frac{1}{\lambda}) + s(s^2-\lambda)(s^2+\frac{1}{\lambda})$ für alle $\lambda \neq 0$ den selben Quotienten $\frac{B(s)}{A(s)}$ und deshalb dieselbe LC-Kette. Aber welches $P_\lambda(s)$ ist dann die Übertragungsfunktion?

Mit anderen Worten: Wie sieht man einem vorgegebenen $p(s)$ an, ob man es durch eine LC-Kette in der oben angegebenen Weise realisieren kann? Nun, leider nicht auf den ersten Blick, sondern nur im Verlauf der Kettenbruchentwicklung. $p(s)$ ist tatsächlich die Übertragungsfunktion, wenn folgende drei Bedingungen erfüllt sind:

- 1) Der konstante Term von $p(s)$ ist 1. (Das ist sicher notwendig, vgl. Bedingung (ii) auf Seite 4.1.33.
- 2) Die Kettenbruchentwicklung hat die oben angegebene Form mit positiven $L_1, \dots, L_m, C_1, \dots, C_m$.
- 3) Der Grad von $p(s)$ ist

$$\begin{aligned} & 2m+1, \quad \text{falls } L_{m+1} \neq 0 \\ & 2m, \quad \text{falls } L_{m+1} = 0. \end{aligned}$$

Konkrete Beispiele:

a) Gegeben sei das Polynom

$$p(s) = s^4 + 2s^3 + 3s^2 + 2s + 1.$$

Dann ist

$$A(s) = s^4 + 3s^2 + 1$$

$$B(s) = 2s^3 + 2s$$

$$\begin{aligned} \frac{B(s)}{A(s)} &= \frac{2s^3 + 2s}{s^4 + 3s^2 + 1} \\ &= \frac{1}{(s^4 + 3s^2 + 1) : (2s^3 + 2s)} \end{aligned}$$

Die Division mit Rest liefert

$$(s^4 + 3s^2 + 1) : (2s^3 + 2s) = \frac{1}{2}s + \frac{2s^2 + 1}{2s^3 + 2s}.$$

Also

$$\frac{B(s)}{A(s)} = \frac{1}{\frac{1}{2}s + \frac{1}{(2s^3 + 2s) : (2s^2 + 1)}}$$

und wie eben ergibt sich weiter

$$\frac{B(s)}{A(s)} = \frac{1}{\frac{1}{2}s + \frac{1}{s + \frac{1}{2s+1}}}$$

Die Bedingung 1) - 3) sind erfüllt, und deshalb hat



die gesuchte Übertragungsfunktion

$$\frac{1}{s^4 + 2s^3 + 3s^2 + 2s + 1}.$$

b) BUTTERWORTHfilter 3. Ordnung.
Die Nullstellen von $1-z^6$ sind

$$\pm 1, \quad \frac{1}{2} \pm \frac{j}{2}\sqrt{3}, \quad -\frac{1}{2} \pm \frac{j}{2}\sqrt{3}.$$

Die Übertragungsfunktion für das BUTTERWORTHfilter 3. Ordnung ist daher

$$\begin{aligned} & 1/(s+1) \left(s + \frac{1}{2} - \frac{j}{2}\sqrt{3}\right) \left(s + \frac{1}{2} + \frac{j}{2}\sqrt{3}\right) \\ &= 1/(s^3 + 2s^2 + 2s + 1). \end{aligned}$$

Kettenbruchentwicklung:

$$\frac{s^3 + 2s}{2s^2 + 1} = \frac{1}{2} s + \frac{\frac{4}{3} s + \frac{1}{2}}{\frac{3}{2} s + \frac{1}{2}}$$

und weil die Bedingungen 1), 2), 3) offenbar erfüllt sind, ist



ein BUTTERWORTHfilter 3. Ordnung.

Schlussbemerkung: Die Berechnung der L_j und C_j aus dem Polynom mittels Kettenbruchverfahren ist so einfach, daß man dafür ein wunderbares Taschenrechnerprogramm schreiben kann.

Versuchen Sie's mal!

Literatur zur LAPLACETRansformation

Als verlässliches Handbuch für Ingenieure empfiehlt sich

G. DOETSCH, Anleitung zum praktischen Gebrauch der Laplace-Transformation und der Z-Transformation, Wien 1967 (3. Auflage)

Speziell für Ingenieure geschrieben sind auch

O. FÖLLINGER, Laplace- und Fouriertransformation, Berlin 1977 und

I. HOLBROOK, Laplace-Transformationen, Braunschweig 1970.

Im letzten Buch sind ausführlich Fragen der Netzwerkanalyse und -synthese sowie elektronischer Filter behandelt.

Ich habe mich im ersten Teil dieses Abschnitts teilweise eng gehalten an die sehr elementare Darstellung in

KREIDER, KULLER, OSTBERG, Elementary Differential Equations, Massachusetts 1968.

Kleine Tabelle zur LAPLACETRansformation
(nach KREIDER, KULLER, OSTBERG)

$f(t)$	$\mathcal{L}\{f\} = \int_0^{\infty} e^{-st} f(t) dt$
1	$\frac{1}{s}$
e^{at}	$\frac{1}{s-a}$
t^n	$\frac{n!}{s^{n+2}}$
$\sin at$	$\frac{a}{s^2+a^2}$
$\cos at$	$\frac{s}{s^2+a^2}$
$\sinh at$	$\frac{a}{s^2-a^2}$
$\cosh at$	$\frac{s}{s^2-a^2}$
$\delta(t)$	1
$\delta(t-a)$	e^{-as}
$\frac{t^{n-1} e^{at}}{(n-1)!}$	$\frac{1}{(s-a)^n} \quad (n \geq 1)$
$\frac{1}{2a^3} (\sin at - at \cos at)$	$\frac{1}{(s^2+a^2)^2}$
$\frac{t}{2a} \sin at$	$\frac{s}{(s^2+a^2)^2}$
$\frac{t}{2n} \mathcal{L}^{-1} \left[\frac{1}{(s^2+a^2)^n} \right] dt$	$\frac{1}{(s^2+a^2)^{n+1}}$
$\frac{t}{2n} \mathcal{L}^{-1} \left[\frac{1}{(s^2+a^2)^n} \right] dt$	$\frac{s}{(s^2+a^2)^{n+1}}$

14.2 Die FOURIERtransformation

Vorbemerkung. Während die LAPLACetransformation speziell auf Anfangswertprobleme zugeschnitten ist, weil sie nur die Werte von $f(t)$ für $t > 0$ berücksichtigt, tritt z.B. bei der Behandlung von nicht-kausalen Systemen etwa in der Nachrichtentechnik häufig der Fall auf, das der Verlauf über die ganze reelle Achse wesentlich ist. Die FOURIERtransformation nimmt darauf Rücksicht. Sie ist in vielem der LAPLACetransformation ähnlich, im mathematischen Detail allerdings schwieriger: durch ein schlechteres Konvergenzverhalten ist man viel häufiger zur Verwendung verallgemeinerter Funktionen gezwungen. Wir beschränken uns hier aber auf einige grundlegende Tatsachen. Die Anwendungsweisen der FOURIERtransformation sind ähnlich wie für die LAPLACetransformation. Wir verweisen insbesondere auf das für Elektroingenieure gedachte Buch von PAPOULIS (s. Literatur am Ende des Abschnitts).

Definition der FOURIERtransformation. Die FOURIERtransformation ordnet komplexen Funktionen eines reellen Arguments $t \in (-\infty, \infty)$ wieder solche Funktionen zu durch

$$\mathcal{F}[f](\omega) := \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt.$$

Dabei ist das rechte Integral definiert als

$$\lim_{\alpha \rightarrow -\infty} \int_{\alpha}^{\beta} f(t)e^{-i\omega t} dt.$$

Der Unterschied zur LAPLACetransformation besteht im veränderten Integrationsbereich und in der Beschränkung auf die imaginäre Achse (iω statt z).

Die Definition von $\mathcal{F}[f]$ ist in der Literatur nicht einheitlich; man findet vor dem Integral auch Faktoren $1/2\pi$ oder $1/\sqrt{2\pi}$.

$\alpha f(t) + \beta g(t)$	$\alpha \mathcal{L}[f] + \beta \mathcal{L}[g]$
$f'(t)$	$s \mathcal{L}[f] - f(0^+)$
$f''(t)$	$s^2 \mathcal{L}[f] - sf(0^+) - f'(0^+)$
$f^{(n)}(t)$	$s^n \mathcal{L}[f] - s^{n-1}f(0^+) - s^{n-2}f'(0^+) - \dots - f^{(n-1)}(0^+)$
$\int_a^t f(t) dt$	$\frac{1}{s} \mathcal{L}[f]$
$\int_a^t \int_a^t f(t) dt$	$\frac{1}{s} \mathcal{L}[f] - \frac{1}{s} \int_a^a f(t) dt$
$\int_a^t \dots \int_a^t f(t) dt \dots dt$ n times	$\frac{1}{s^n} \mathcal{L}[f]$
$\int_a^t \dots \int_a^t \int_a^t f(t) dt \dots dt$ n times	$\frac{1}{s^n} \mathcal{L}[f] - \frac{1}{s^{n-1}} \int_a^a f(t) dt$ $- \frac{1}{s^{n-1}} \int_a^a \int_a^t f(t) dt dt$ $- \dots - \frac{1}{s} \int_a^a \int_a^t \dots \int_a^t f(t) dt \dots dt$ n times
$e^{at} f(t)$	$\bar{f}(s-a)$, where $\bar{f}(s) = \mathcal{L}[f]$
$t^n f(t)$	$(-1)^n \frac{d^n}{ds^n} \mathcal{L}[f]$
$u_a(t)g(t) = \begin{cases} 0, & t < a \\ g(t), & t > a \end{cases}$	$e^{-as} \mathcal{L}[g(t+a)]$
$u_a(t)g(t-a) = \begin{cases} 0, & t < a \\ g(t-a), & t > a \end{cases}$	$e^{-as} \mathcal{L}[g]$
$\int_a^t f(t-\xi)g(\xi) d\xi$	$\mathcal{L}[f] \mathcal{L}[g]$
$f(t)$ periodic with period p ($p > 0$)	$\int_0^p \frac{e^{-st} f(t) dt}{1 - e^{-ps}}$
$\frac{f(t)}{t}$ if $\lim_{t \rightarrow 0^+} \frac{f(t)}{t}$ exists	$\int_0^{\infty} \mathcal{L}[f] ds$

Eine Klasse von FOURIERtransformierbaren Funktionen. Wir definieren nun eine Klasse von Funktionen, die sogenannten SCHWARTZschen Funktionen, für die sich die Theorie der FOURIERtransformierung ohne Konvergenzprobleme entwickeln läßt. Damit haben Sie dann die wesentlichen Eigenschaften entwickelt vor Augen. Es ist allerdings zu betonen, daß die meisten in der Praxis auftretenden Funktionen sich zwar sinnvoll durch SCHWARTZsche Funktionen approximieren lassen, selbst aber keine sind. Eine praxisgerechte und mathematisch einwandfreie Theorie kommt wohl ohne Distributionen nicht aus.

Wir nennen eine komplexwertige Funktion $f(t)$ eines reellen Arguments eine SCHWARTZsche Funktion, wenn sie

- i) auf ganz \mathbb{R} definiert ist und dort beliebig oft differenzierbar ist;
- ii) $f(t)$ und alle seine Ableitungen im Unendlichen schnell fallen.

Das bedeutet: für jede natürliche Zahl m und alle k ist

$$f^{(k)}(t) \cdot t^m$$

beschränkt.

Wir nennen SCHWARTZsche Funktionen kurz S-Funktionen.

Beispiele: Die Funktion $f(t) = e^{-t^2}$ ist eine SCHWARTZsche Funktion. Ebenso jede beliebig oft differenzierbare Funktion, die außerhalb eines beschränkten Intervalls verschwindet. Man sieht leicht aus der Definition, daß Summen und Produkte von S-Funktionen wieder S-Funktionen sind. Auch Produkte von Polynomen mit S-Funktionen sind S-Funktionen. Insbesondere gibt es deshalb zu jeder S-Funktion $f(t)$ eine Konstante $K > 0$ mit

$$(1+t^2) |f(t)| \leq K.$$

Dann ist aber

$$\int_{\beta_1}^{\beta_2} |f(t)e^{-i\omega t}| dt \leq K \int_{\beta_1}^{\beta_2} \frac{1}{1+t^2} dt$$

$$= K \cdot (\arctan \beta_2 - \arctan \beta_1).$$

Läßt man β_1 und β_2 gegen $-\infty$ bzw. $+\infty$ gehen, so sieht man, daß $\int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt$ existiert.

Satz: Die FOURIERtransformierte existiert für jede S-Funktion.

Die obige Abschätzung zeigt überdies, daß

$$|\mathcal{F}\{f\}(\omega)| \leq K \cdot \pi,$$

also beschränkt ist.

Wir untersuchen nun die Eigenschaften von \mathcal{F} in der Klasse der S-Funktionen.

Differentiation. Ist $f(t)$ eine S-Funktion, so auch $f'(t)$, und wir haben den

Satz:

$$\mathcal{F}\{f'\}(\omega) = i\omega \mathcal{F}\{f\}(\omega)$$

$$\mathcal{F}\{f\}'(\omega) = \mathcal{F}\{-itf(t)\}(\omega)$$

Beweis:

$$\mathcal{F}\{f'\}(\omega) = \int_{-\infty}^{\infty} f'(t)e^{-i\omega t} dt$$

$$= f(t)e^{-i\omega t} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(t)(-i\omega)e^{-i\omega t} dt.$$

Da $tf(t)$ beschränkt ist, ist $\lim_{t \rightarrow \pm\infty} f(t) = 0$ für $t \rightarrow \pm\infty$.

Daher verschwindet der erste Term und es folgt die erste Behauptung.

Durch (erlaubte) Differentiation unter dem Integral folgt ebenso

$$\begin{aligned}
\mathcal{F}[f]'(\omega) &= \frac{d}{d\omega} \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \\
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial \omega} f(t)e^{-i\omega t} dt \\
&= \int_{-\infty}^{\infty} -itf(t)e^{-i\omega t} dt \\
&= \mathcal{F}[-itf(t)](\omega).
\end{aligned}$$

Aus dem damit bewiesenen Satz folgt

$$|\omega^m \mathcal{F}[f(t)]^{(k)}(\omega)| = |\mathcal{F}[(t^k f(t))^{(m)}](\omega)|.$$

Aber $(t^k f(t))^{(m)}$ ist eine S-Funktion, und die FOURIERtransformierte deshalb beschränkt, vgl. oben. Also erhalten wir

Satz: Die FOURIERtransformierten von S-Funktionen sind S-Funktionen.

Die Faltung. Für zwei S-Funktionen f, g definieren wir die Faltung durch

$$f * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau.$$

Beachten Sie, daß diese Definition mit der bei der LAPLACETRansformation gegebenen übereinstimmt, wenn $f(t) = g(t) = 0$ für alle $t < 0$.

Ähnlich wie bei der LAPLACETRansformation beweist man den

Faltungssatz. Sind $f(t)$ und $g(t)$ S-Funktionen, so ist auch $f * g(t)$ eine S-Funktion und

$$\mathcal{F}[f * g] = \mathcal{F}[f]\mathcal{F}[g].$$

Die Umkehrung der FOURIERtransformation. Wir beweisen den folgenden

Satz: Für eine S-Funktion $f(t)$ gilt

$$\mathcal{F}[\mathcal{F}[f]](t) = 2\pi f(-t).$$

Insbesondere ist deshalb \mathcal{F} umkehrbar und

$$\mathcal{F}^{-1}[\mathcal{F}(\omega)](t) = \frac{1}{2\pi} \mathcal{F}[F(\omega)](-t).$$

Bemerkung: Auch die Umkehrtransformation der LAPLACETRansformation läßt sich durch ein - allerdings kompliziertes - Integral beschreiben.

Der Beweis des Satzes ist etwas trickreich, aber weil Zwischenresultate auch von selbständigem Interesse sind, führe ich ihn vor. Wir berechnen für eine feste S-Funktion $h(\omega)$ und $a > 0$

$$\begin{aligned}
\mathcal{F}[h(a\omega)]\mathcal{F}[f](\omega)(\theta) &= \\
&= \int_{-\infty}^{\infty} h(a\omega) \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt e^{-i\theta\omega} d\omega \\
&= \int_{-\infty}^{\infty} f(t) \mathcal{F}[h(a\omega)](t+\theta) dt
\end{aligned}$$

Nun ist aber

$$\begin{aligned} \mathcal{F}[h(a\omega)](t) &= \frac{1}{a} \int_{-\infty}^{\infty} h(a\omega) e^{-iat} d(a\omega) \\ &= \frac{1}{a} \mathcal{F}[h(\omega)]\left(\frac{t}{a}\right), \end{aligned}$$

und wir können die vorangegangene Rechnung fortsetzen:

$$\begin{aligned} &= \int_{-\infty}^{\infty} f(t) \mathcal{F}\left[h(\omega)\right]\left(\frac{t+\theta}{a}\right) d\left(\frac{t}{a}\right) \\ &= \int_{-\infty}^{\infty} f(a\tau - \theta) \mathcal{F}[h](\tau) d\tau, \quad \tau = \frac{t+\theta}{a}. \end{aligned}$$

Bei genauem Hinsehen hängt diese Gleichung stetig von a ab, und für $a \rightarrow 0$ folgt

$$\begin{aligned} h(0) \mathcal{F}[\mathcal{F}[f]](0) &= \int_{-\infty}^{\infty} f(-\theta) \mathcal{F}[h](\tau) d\tau \\ &= f(-\theta) \mathcal{F}[\mathcal{F}[h]](0). \end{aligned}$$

Diese Formel liefert uns

$$\mathcal{F}[\mathcal{F}[f]](t) = C f(-t)$$

mit einer von h unabhängigen Konstanten $C = \mathcal{F}[\mathcal{F}[h]](0)/h(0)$. Zur Bestimmung von C genügt also die Berechnung von $\mathcal{F}[\mathcal{F}[h]](0)$ für eine S -Funktion $h(t)$ mit $h(0) \neq 0$. Wir betrachten $h(t) = e^{-t^2/2}$. Die direkte Auswertung der FOURIERtransformation macht Schwierigkeiten, wir brauchen wiederum einen Trick:

$$\begin{aligned} \mathcal{F}[h]'(\omega) &= \mathcal{F}[-it h(t)](\omega) \\ &= \int_{-\infty}^{\infty} -it h(t) e^{-i\omega t} dt \\ &= \int_{-\infty}^{\infty} -ite^{-t^2/2} e^{-i\omega t} dt \\ &= ie^{-t^2/2} e^{-i\omega t} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} ie^{-t^2/2} (-i\omega) e^{-i\omega t} dt \\ &= -\omega \mathcal{F}[h](\omega). \end{aligned}$$

Also ist $\mathcal{F}[h]$ eine Lösung der Differentialgleichung

$$y'(\omega) + \omega y(\omega) = 0,$$

und es folgt

$$\mathcal{F}[h](\omega) = C_0 e^{-\omega^2/2}$$

mit einer Konstanten $C_0 = \mathcal{F}[h](0)$.

Aber (vgl. HM III)

$$\mathcal{F}[h](0) = \int_{-\infty}^{\infty} e^{-t^2/2} dt = \sqrt{2\pi}.$$

Wir erhalten

$$\mathcal{F}[h] = \sqrt{2\pi} h$$

und

$$\mathcal{F}[\mathcal{F}[h]](0) = 2\pi h(0).$$

Damit folgt

$$\mathcal{F}[\mathcal{F}[f]](t) = 2\pi f(-t)$$

und der Satz ist bewiesen.

Physikalische Interpretation. FOURIERtransformation und FOURIERreihen.

Die komplexe FOURIERreihe einer T -periodischen Funktion f_T war

$$f_T(t) \sim \sum_{k=-\infty}^{\infty} c_k e^{ik\omega t}$$

mit $\omega = \frac{2\pi}{T}$ und

$$c_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(\tau) e^{-ik\omega \tau} d\tau.$$

Wir schreiben $\Delta\omega$ statt ω und ω_k statt $k\omega$.

Dann erhalten wir

$$f_T(t) \sim \sum_{k=-\infty}^{\infty} \frac{1}{2\pi} \left\{ \int_{-\frac{T}{2}}^{\frac{T}{2}} f(\tau) e^{-i\omega_k \tau} d\tau e^{i\omega_k t} \right\} \Delta\omega.$$

Lassen wir T gegen Unendlich und damit $\Delta\omega \rightarrow 0$ gehen, so erhalten wir (wenn alles gut geht) für die nicht mehr periodische Funktion $f(t) = \lim_{T \rightarrow \infty} f_T(t)$ die Darstellung

$$\begin{aligned} f(t) &\sim \int_{-\infty}^{\infty} \frac{1}{2\pi} \mathcal{F}[f](\omega) e^{i\omega t} d\omega \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi} F(\omega) e^{i\omega t} d\omega \end{aligned} \quad (*)$$

mit

$$\frac{1}{2\pi} F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt. \quad (**)$$

Wie man periodische Schwingungen durch Überlagerung von harmonischen Schwingungen mit diskretem Frequenzspektrum $k\omega, k \in \mathbb{Z}$, erhalten kann, so kann man also aperiodische Schwingungen (wie sie z.B. bei unbeschränkten Bereichen typischer Weise auftreten) erhalten, indem man harmonische Schwingungen $e^{i\omega t}$ mit kontinuierlichem Frequenzspektrum $-\infty < \omega < \infty$ und der Spektraldichte $\frac{1}{2\pi} F(\omega)$ überlagert. An die Stelle der FOURIERreihe tritt dann das FOURIERintegral (*), und die FOURIERtransformation (**) liefert zu gegebener Zeitfunktion $f(t)$ deren Spektraldichte.

Der Produktsatz. Als Konsequenz aus dem Umkehrsatz erhält man - anders als bei der LAPLACETRANSFORMATION - auch eine einfache Formel für die Transformation eines Produkts:

Satz: Für S-Funktionen $f(t), g(t)$ gilt:

$$\mathcal{F}[fg] = \frac{1}{2\pi} \mathcal{F}[f] * \mathcal{F}[g]$$

Beweis: Da \mathcal{F} eindeutig, folgt die Behauptung aus

$$\begin{aligned} \mathcal{F}\left[\frac{1}{2\pi} \mathcal{F}[f] * \mathcal{F}[g]\right](t) &= \frac{1}{2\pi} \mathcal{F}[\mathcal{F}[f]] \mathcal{F}[\mathcal{F}[g]](t) \\ &= \frac{1}{2\pi} \cdot 2\pi f(-t) \cdot 2\pi g(-t) \\ &= \mathcal{F}[\mathcal{F}[fg]](t). \end{aligned}$$

Die PARSEVALSche Gleichung. Wir bezeichnen die FOURIERtransformierten der beiden S-Funktionen $f(t)$ und $g(t)$ mit $F(\omega)$ und $G(\omega)$. Dann gilt nach dem vorstehenden Satz

$$\int_{-\infty}^{\infty} f(t)g(t) e^{-i\omega t} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\tau)G(\omega-\tau) d\tau.$$

Wir setzen $\omega = 0$ und erhalten

$$\int_{-\infty}^{\infty} f(t)g(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\tau)G(-\tau) d\tau.$$

Wir überlegen zwischendurch schnell, wie die FOURIERTRANSFORMATION sich ändert, wenn wir $g(t)$ durch die konjugiert-komplexe Funktion $\bar{g}(t)$ ersetzen:

$$\begin{aligned} \mathcal{F}[\bar{g}](\omega) &= \int_{-\infty}^{\infty} \bar{g}(t) e^{-i\omega t} dt \\ &= \int_{-\infty}^{\infty} \overline{g(t) e^{i\omega t}} dt \\ &= \int_{-\infty}^{\infty} g(t) e^{-i(-\omega)t} dt \\ &= \mathcal{F}[g](-\omega). \end{aligned}$$

Ersetzen wir also in obiger Formel g durch \bar{g} , so folgt

$$\int_{-\infty}^{\infty} f(t)\bar{g}(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\bar{G}(\omega) d\omega.$$

Dies ist die sogenannte PARSEVALSche Gleichung. Sie impliziert insbesondere für $f(t) = g(t)$:

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega$$

Das läßt sich so interpretieren: Ist $f(t)$ der Spannungsverlauf an einem Widerstand von 1Ω , so ist $|f(t)|^2$ die abgegebene Leistung und das linke Integral stellt die gesamte Energie dar. $|F(\omega)|^2$ nennt man wegen der PARSEVALSchen Gleichung die spektrale Energiedichte.

Beispiele zur FOURIERtransformation. Beim Beweis der Umkehrformel haben wir schon gezeigt, daß

$$\mathcal{F}\{e^{-t^2}/2\}(\omega) = \sqrt{2\pi} e^{-\omega^2}/2$$

Die Funktion $h(t) = e^{-t^2}/2$ ist eine S-Funktion. Weitere Beispiele von S-Funktionen sind etwa die Produkte von $h(t)$ mit Polynomen oder trigonometrischen Funktionen. Wir geben stattdessen noch ein Beispiel für die transformierte einer "Nicht-S-Funktion":

Sei

$$r_T(t) = \begin{cases} 1 & \text{für } |t| < T \\ 0 & \text{sonst.} \end{cases}$$

$r_T(t)$ ist also ein Rechteckimpuls der Breite $2T$. Dafür erhält man

$$\begin{aligned} \mathcal{F}\{r_T\}(\omega) &= \int_{-\infty}^{\infty} r_T(t) e^{-i\omega t} dt \\ &= \int_{-T}^T e^{-i\omega t} dt \\ &= \frac{-1}{i\omega} e^{-i\omega t} \Big|_{-T}^T \\ &= + \frac{1}{i\omega} (e^{i\omega T} - e^{-i\omega T}) \\ &= 2 \frac{\sin \omega T}{\omega} \end{aligned}$$

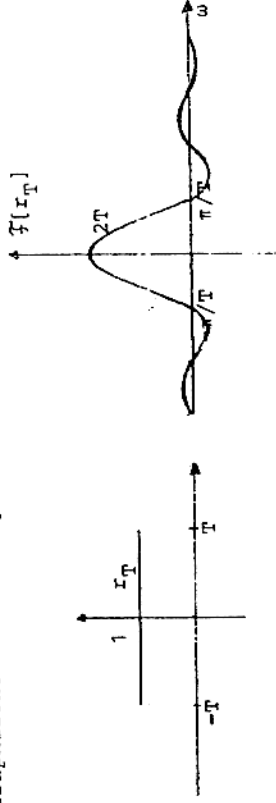
Also

$$\mathcal{F}\{r_T\}(\omega) = 2 \frac{\sin \omega T}{\omega}$$

und mit der Umkehrformel (die außerhalb der Sprungstellen auch hier anwendbar ist)

$$\mathcal{F}\left\{2 \frac{\sin \omega T}{\omega}\right\}(t) = 2\pi r_T(-t) \quad t \neq \pm T$$

Graphische Darstellung:



Es zeigt sich folgendes: In der FOURIERanalyse von r_T überwiegen die Frequenzen im Bereich $|\omega| < \pi/T$. Ein Sandpaßfilter für diesen Frequenzbereich verzerrt Rechteckimpulse r_T mit $T_1 > T$ nicht sehr stark. Da aber die Breite π/T des wesentlichen Frequenzbandes der Impulsdauer umgekehrt proportional ist, ist für $T_1 \ll T$ eine starke Verzerrung zu erwarten.

Aus dem letzten Beispiel folgt

$$\mathcal{F}\left\{\frac{1}{\sqrt{2\pi}} r_T\right\}(\omega) = \frac{\sin \omega T}{\omega T}$$

und durch "Grenzübergang" für $T \rightarrow 0$ ergibt sich die FOURIERtransformierte der δ -Funktion:

$$\mathcal{F}[\delta] = 1$$

und (falls die Umkehrformel richtig bleibt)

$$\mathcal{F}[1] = 2\pi\delta.$$

Ich will aber auf verallgemeinerte Funktionen hier nicht näher eingehen. Vergleichen Sie dazu die Literatur. Zwei wichtige und leicht zu verifizierende Identitäten seien noch angegeben:

$$\mathcal{F}[f(r-a)](\omega) = e^{-i\omega a} \mathcal{F}[f(t)](\omega)$$

$$\mathcal{F}[f(\alpha t)](\omega) = \frac{1}{|\alpha|} \mathcal{F}[f(t)](\omega/\alpha), \alpha \neq 0.$$

Der Abtastatz von SHANNON. Gegeben sei eine S-Funktion $f(t)$ mit der Spektraldichte $F(\omega)$. Dann gilt also

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega.$$

Wir setzen nun voraus, daß $f(t)$ vor endlicher Bandbreite ω_0 ist, d.h. daß

$$F(\omega) = 0 \text{ für } |\omega| \geq \omega_0 > 0.$$

Der SHANNONSche Abtastatz besagt, daß man unter diesen Voraussetzungen den ganzen Funktionsverlauf von $f(t)$ rekonstruieren kann, wenn man die Funktion nur an äquidistanten Stellen vom Abstand π/ω_0 "abtastet", d.h. wenn man die Werte $f(k\pi/\omega_0)$ für alle ganzzahligen Werte k kennt:

Satz (KOTELNIKOV-SHANNON) Seien $f(t)$ und $\tilde{f}(t)$ S-Funktionen von endlicher Bandbreite ω_0 . Gilt dann $f(k\pi/\omega_0) = \tilde{f}(k\pi/\omega_0)$ für alle ganzen Zahlen k , so gilt $f(t) = \tilde{f}(t)$ für alle t .

Im nun zu erbringenden Beweis werden wir eine explizite Formel herleiten, wie sich $f(t)$ aus den $f(k\pi/\omega_0)$ berechnet.

Wir haben

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} F(\omega) e^{i\omega t} d\omega. \end{aligned}$$

Wir setzen $F(\omega)$ mit der Periode $2\omega_0$ auf \mathbb{R} fort und entwickeln in eine FOURIERREihe, die offenbar gleichmäßig konvergent ist, weil $F(\omega)$ beliebig oft differenzierbar ist. Beachten Sie: Bei diesem Schritt nutzen wir die endliche Bandbreite aus!

$$F(\omega) = \sum_{k=-\infty}^{\infty} c_k e^{ik\pi\omega/\omega_0}$$

mit

$$c_k = \frac{1}{2\omega_0} \int_{-\omega_0}^{\omega_0} F(\omega) e^{-ik\pi\omega/\omega_0} d\omega.$$

Einsetzen in den Ausdruck für $f(t)$ liefert nach Vertauschen von Integral und Summe

$$\begin{aligned} f(t) &= \sum_{k=-\infty}^{\infty} \frac{c_k}{2\pi} \int_{-\omega_0}^{\omega_0} e^{ik\pi\omega/\omega_0 + i\omega t} d\omega \\ &= \sum_{k=-\infty}^{\infty} \frac{c_k}{2\pi} \frac{1}{i(k\pi/\omega_0 + t)} e^{i(k\pi/\omega_0 + t)\omega} \Big|_{\omega=-\omega_0}^{\omega=\omega_0} \\ &= \sum_{k=-\infty}^{\infty} \frac{c_k}{\pi} \frac{\omega_0}{k\pi + t\omega_0} \sin(k\pi + t\omega_0). \end{aligned}$$

Damit ist $f(t)$ durch die FOURIERKoeffizienten c_k eindeutig bestimmt. Für diese gilt aber

$$c_k = \frac{\pi}{\omega_0} \cdot \frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} F(\omega) e^{i\omega(-k\pi/\omega_0)} d\omega$$

$$= \frac{\pi}{\omega_0} f(-k\pi/\omega_0).$$

Es folgt

$$f(t) = \sum_{k=-\infty}^{\infty} f(-k\pi/\omega_0) \frac{\sin(k\pi + t\omega_0)}{k\pi + t\omega_0}$$

und damit der behauptete Satz, ja ein stärkeres Resultat, eben weil wir eine explizite Formel für $f(t)$ gefunden haben.

Literatur zur FOURIERtransformation

Die elementaren mathematischen Grundlagen der F.T. für S-Funktionen findet man in

S. LANG, Analysis. Inter European Editions, Amsterdam 1977 (oder engl. Ausgabe "Analysis I", Addison-Wesley 1968).

Ausführlichere Darstellung der mathematischen Grundlagen:

G. DOETSCH, Funktionaltransformationen. Abschnitt C in "Mathematische Hilfsmittel des Ingenieurs, Teil I" Hrsg. Sauer, Szabo, Springer 1967.

Weitere Bücher speziell für (Elektro-) Ingenieure:
O. FÖLLINGER, Laplace- und Fouriertransformationen, Elitera-Verlag, Berlin 1977

A. PAPOULIS, The Fourier Integral and its Applications. McGraw Hill 1962.

14.3 Die Z-Transformation

Vorbemerkung. Die Z-Transformation ist nicht eigentlich eine Integraltransformation, obwohl man sie "mit etwas Geschick" auch als solche interpretieren kann. Sie wird (wenigstens primär) nicht zur Behandlung von kontinuierlichen Prozessen benutzt, sondern bei diskreten Problemen, wie sie zum Beispiel bei digitalen Prozessen auftreten. Häufig sind die diskreten Vorgaben allerdings das Resultat eines "Abtastvorgangs" an kontinuierlichen Funktionen, so in der Theorie der Impulselemente. Aus mathematischer Sicht ist die Z-Transformation für Differenzgleichungen, was die LAPLACETRansformation für Differentialgleichungen ist.

Definition der Z-Transformation. Sei $(f_n) = (f_0, f_1, f_2, \dots)$ eine Folge reeller oder komplexer Zahlen. Dann ist die Z-Transformation dieser Folge eine komplexe Funktion, nämlich

$$Z\{f_n\}(z) = \sum_{n=0}^{\infty} f_n z^{-n}.$$

Diese Funktion ist also durch eine LAURENTreihe um den Ursprung gegeben, die nach der Funktionentheorie außerhalb eines Kreises vom Radius R ($0 < R \leq \infty$) konvergiert. Beachten Sie, daß Sie die rechte Seite auch als Potenzreihe in $1/z$ betrachten können. Die geschlossene Berechnung von $Z\{f_n\}$ ist nichts anderes als die geschlossene Darstellung der Reihensumme.

Die Umkehrung der Z-Transformation. Vorgegeben sei eine Funktion $F^*(z)$; gesucht ist eine Folge (f_n) mit

$$F^*(z) = \sum_{n=0}^{\infty} f_n z^{-n} = Z\{f_n\}.$$

Dies Problem kann man z.B. lösen, indem man $F^*(\frac{1}{z})$ in eine TAYLORreihe um 0 entwickelt:

$$f_n = \frac{1}{n!} \frac{d^n}{dz^n} F^*\left(\frac{1}{z}\right).$$

Ist $F^*(z) = \frac{P(z)}{Q(z)}$ rational, so führen auch Partialbruchzerlegung und die Formel für die geometrische Reihe zum Ziel.

Fundamentale Eigenschaften. Die folgenden Rechenregeln sind ganz leicht zu beweisen.

Für $k=0,1,2,\dots$ fest sei $f_{n-k} := 0$, falls $n-k < 0$. Dann gilt

$$\begin{aligned} \mathcal{Z}\{f_{n-k}\} &= z^{-k} \mathcal{Z}\{f_n\} \\ \mathcal{Z}\{f_{n+k}\} &= z^k \mathcal{Z}\{f_n\} - \sum_{j=0}^{k-1} f_j z^{k-j} \\ \mathcal{Z}\left\{\sum_{j=0}^{n-1} f_j\right\} &= \frac{1}{z-1} \mathcal{Z}\{f_n\} \\ \mathcal{Z}\left\{\sum_{j=0}^n f_j\right\} &= \frac{z}{z-1} \mathcal{Z}\{f_n\} \\ \mathcal{Z}\{\alpha^n f_n\} &= \mathcal{Z}\{f_n\}(az) \\ \mathcal{Z}\{nf_n\} &= -z \frac{d}{dz} \mathcal{Z}\{f_n\} \end{aligned}$$

Etwas schwieriger zu zeigen ist die Faltungsgleichung

$$\mathcal{Z}\left[\sum_{j=0}^n f_j g_{n-j}\right] = \mathcal{Z}\{f_n\} \mathcal{Z}\{g_n\}.$$

Bei der Anwendung auf Differenzgleichungen wichtig ist folgender

Satz: Setzt man $\Delta f_n := f_{n+1} - f_n$
 $\Delta^m f_n := \Delta(\Delta^{m-1} f_n)$
 und $\Delta^0 f_n := f_n$,
 so folgt

$$\mathcal{Z}\{\Delta f_n\} = (z-1) \mathcal{Z}\{f_n\} - f_0 z$$

$$\mathcal{Z}\{\Delta^2 f_n\} = (z-1)^2 \mathcal{Z}\{f_n\} - f_0 z(z-1) - \Delta f_0 z$$

etc.

Beachten Sie die Ähnlichkeit mit der Formel für $\mathcal{L}\{f\}$.

Beispiele für die \mathcal{Z} -Transformation

$$\begin{aligned} \mathcal{Z}\{a^n\} &= \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n = \frac{z}{z-a} \\ \mathcal{Z}\{na^n\} &= -z \frac{d}{dz} \frac{z}{z-a} = \frac{az}{(z-a)^2} \\ \mathcal{Z}\{a^{n/n!}\} &= e^{a/z}. \end{aligned}$$

Die beiden folgenden Formeln geben wir im Hinblick auf ein späteres Beispiel:

$$\begin{aligned}
 \mathcal{Z}[\cosh nx] &= \mathcal{Z}\left[\frac{1}{2}(e^{nx} + e^{-nx})\right] \\
 &= \frac{1}{2} \mathcal{Z}[(e^x)^n] + \frac{1}{2} \mathcal{Z}[(e^{-x})^n] \\
 &= \frac{1}{2} \frac{z}{z-e^x} + \frac{1}{2} \frac{z}{z-e^{-x}} = \frac{z}{2} \cdot \frac{2z - (e^x + e^{-x})}{z^2 - z \cdot (e^x + e^{-x}) + 1} \\
 &= \frac{z - z \cosh x}{z^2 - 2z \cosh x + 1}
 \end{aligned}$$

Analog ergibt sich

$$\mathcal{Z}[\sinh nx] = \frac{z \sinh x}{z^2 - 2z \cosh x + 1}$$

Anwendungen: Differenzengleichungen. Eine lineare Differenzengleichung mit konstanten Koeffizienten ist eine Gleichung für eine gesuchte Folge (x_n) von der Form

$$a_0 \Delta^r x_n + a_1 \Delta^{r-1} x_n + \dots + a_r x_n = b_n \quad (n = 0, 1, 2, \dots)$$

Durch Ausschreiben der Differenzen wird daraus eine Gleichung der Form

$$c_0 x_{n+r} + c_1 x_{n+r-1} + \dots + c_r x_n = b_n.$$

Gibt man Anfangswerte x_0, \dots, x_{r-1} vor, so kann man die Folge (x_n) rekursiv daraus berechnen. In der Theorie der Differenzengleichungen sucht man aber einen geschlossenen Ausdruck für die Lösung. Ähnlich wie die LAPLACETRANSFORMATION bei Differenzialgleichungen kann man hierzu die \mathcal{Z} -TRANSFORMATION benutzen. Ich will das aber nur an einem konkreten Beispiel aus der E-Technik illustrieren:

Gegeben sei ein Vierpol mit der Schaltmatrix

$$\mathcal{M}(s) = \begin{pmatrix} A(s) & B(s) \\ C(s) & A(s) \end{pmatrix}$$

mit $\det \mathcal{M}(s) = A(s)^2 - B(s)C(s) = 1$. Vergleichen Sie dazu die letzten Seiten über die LAPLACETRANSFORMATION. Gesucht ist die Übertragungsmatrix für die Kettenschaltung einer großen Anzahl von Exemplaren dieses Vierpols:



Wir halten s fest und schreiben A, B , usw. für $A(s), B(s)$, usw.. Dann erhalten wir die Übertragungsgleichungen

$$\begin{aligned}
 E_n &= A E_{n+1} + B I_{n+1} \\
 I_n &= C E_{n+1} + A I_{n+1}
 \end{aligned}$$

Wendet man darauf die \mathcal{Z} -TRANSFORMATION an und setzt $E^* := \mathcal{Z}\{E_n\}, I^* := \mathcal{Z}\{I_n\}$, so folgt

$$\begin{aligned}
 E^* &= A(zE^* - zE_0) + B(zI^* - zI_0) \\
 I^* &= C(zE^* - zE_0) + A(zI^* - zI_0),
 \end{aligned}$$

und durch Elimination von I^* nach einfacher Rechnung wegen $A^2 - BC = 1$

$$E^* = \frac{E_0(z^2 - zA) - I_0 z B}{z^2 - 2zA + 1}$$

Wir setzen $A \geq 1$ voraus und wählen nun τ mit $A = \cosh \tau$. Wegen $A^2 - BC = 1$ ist dann $BC = \sinh^2 \tau$, und wir haben für ein geeignetes $\alpha > 0$

$$\begin{aligned}
 B &= \alpha \sinh \tau \\
 C &= 1/\alpha \sinh \tau
 \end{aligned}$$

Indem man gegebenenfalls τ durch $-\tau$ ersetzt, kann man $\alpha > 0$ erreichen.

Damit wird

$$E^* = E_0 \frac{z^2 - z \cosh \tau}{z^2 - 2z \cosh \tau + 1} - I_0 \frac{az \sinh \tau}{z^2 - 2z \cosh \tau + 1},$$

und es folgt nach den vorstehenden Beispielen

$$\begin{aligned} E_n &= E_0 \cdot \cosh n\tau - \alpha I_0 \sinh n\tau \\ &= E_0 \cdot \cosh n\tau - \sqrt{B/C} I_0 \sinh n\tau. \end{aligned}$$

Ebenso ergibt sich

$$I_n = -E_0 \sqrt{C/B} \sinh n\tau + I_0 \cosh n\tau.$$

Die Übertragungsmatrix ist also die Inverse von

$$\begin{pmatrix} \cosh n\tau & -\sqrt{B/C} \sinh n\tau \\ -\sqrt{C/B} \sinh n\tau & \cosh n\tau \end{pmatrix}$$

d.h. die Matrix

$$\begin{pmatrix} \cosh n\tau & \sqrt{B/C} \sinh n\tau \\ \sqrt{C/B} \sinh n\tau & \cosh n\tau \end{pmatrix}.$$

Bemerkung: Dies Beispiel wird in der Ingenieurmathematik gern angeführt, vgl. DOETSCH und JURY. Ich bin nicht ganz überzeugt, daß es wirklich die Qualitäten der λ -Transformation sehr gut demonstriert, weil man die n-te Potenz der Matrix

$$\begin{pmatrix} \cosh \tau & \alpha \sinh \tau \\ \frac{1}{\alpha} \sinh \tau & \cosh \tau \end{pmatrix}$$

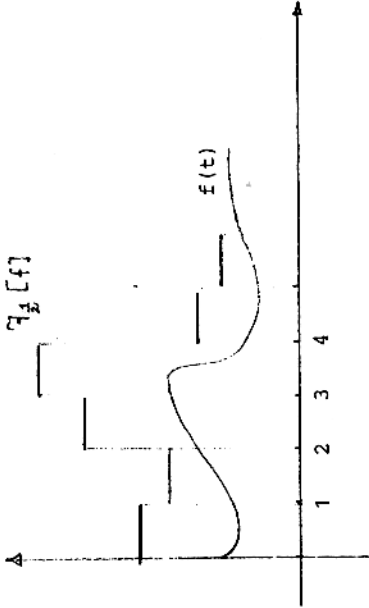
auch leicht direkt ausrechnen kann.

λ -Transformation und LAPLACE-Transformation

Ein Impulselement (ein Taster) verwandelt eine kontinuierliche Eingangsfunktion $f(t)$ in eine "Treppenfunktion", und zwar auf

folgende Weise: Für die Ausgangsfunktion $\gamma_\epsilon[f]$ gilt mit festem $\epsilon, 0 < \epsilon \leq 1$:

$$\gamma_\epsilon[f](t) = \begin{cases} \frac{1}{\epsilon} f(n) & \text{für } n \leq t < n+\epsilon \\ 0 & \text{sonst.} \end{cases}$$



Wir berechnen die LAPLACE-Transformation von $\gamma_\epsilon[f]$:

$$\begin{aligned} \mathcal{L}[\gamma_\epsilon[f]](s) &= \int_0^\infty \gamma_\epsilon[f](t) e^{-st} dt \\ &= \sum_{n=0}^{\infty} \frac{1}{\epsilon} \int_n^{n+\epsilon} f(n) e^{-st} dt \\ &= \sum_{n=0}^{\infty} \frac{1}{\epsilon} f(n) \frac{e^{-ns}}{s} (1 - e^{-\epsilon s}) \\ &= \frac{1 - e^{-\epsilon s}}{s\epsilon} \sum_{n=0}^{\infty} f(n) (e^{-s})^n \\ &= \frac{1 - e^{-\epsilon s}}{s\epsilon} \mathcal{Z}[f(n)](e^{+s}) \end{aligned}$$

Wir sehen: Bis auf den Faktor $\frac{1 - e^{-\epsilon s}}{\epsilon s}$, der für $\epsilon \rightarrow 0$ gegen 1 konvergiert, ist die LAPLACE-Transformation von $\gamma_\epsilon[f]$ gerade die λ -Transformation der Wertefolge $(f(n))$ angewendet auf $z = e^{+s}$. In der digitalen Nachrichtentechnik treten sehr häufig

Impulselemente obiger Art auf, und deren LAPLACE-Transformation, also die \mathcal{L} -Transformation, ist deshalb wichtig. Wie schon früher bemerkt, ist es üblich, den Grenzübergang $\epsilon \rightarrow 0$ "durchzuführen". Dann erhält man mit der Deltafunktion

$$f^*(t) := \lim_{\epsilon \rightarrow 0} \mathcal{L}_{\epsilon}[f](t) = \sum_{n=0}^{\infty} f(n) \delta(n-t)$$

und

$$\mathcal{L}[f^*(t)](s) = \mathcal{L}[f(n)](e^{+s}) .$$

für weitere Einzelheiten vergleiche das 8. Kapitel im angegebenen Buch von DOETSCH.

Literatur zur \mathcal{L} -Transformation

- G. DOETSCH, Anleitung zum praktischen Gebrauch der LAPLACE-Transformation und der \mathcal{L} -Transformation. 3. Auflage Oldenbourg Verlag München, Wien 1967
- E. JURY, Theory and Applications of the \mathcal{L} -Transformation Method. Wiley, New York 1964

KAPITEL 15

WAHRSCHEINLICHKEITSRECHNUNG

- 15.1 WAHRSCHEINLICHKEIT
- 15.2 BEDINGTE WAHRSCHEINLICHKEIT
- 15.3 ZUFALLSVARIABLE UND ERWARTUNGSWERT
- 15.4 UNABHÄNGIGE ZUFALLSVARIABLE
- 15.5 DER ZENTRALE GRENZWERTSATZ
- 15.6 STATISTISCHE METHODEN

Kap. 15: Wahrscheinlichkeitsrechnung

Wahrscheinlichkeitsrechnung und statistische Methoden spielen überall da eine Rolle, wo man auf unvollständige oder fehlerhafte Informationen angewiesen ist, z.B. in der Nachrichtentechnik, bei sehr komplexen Systemen, bei Fertigungsprozessen usw.

15.1 Wahrscheinlichkeit

Der Begriff der Wahrscheinlichkeit.

Man geht aus von einer Menge Ω , dem *Wahrscheinlichkeitsraum*. Seine Teilmengen A , also die Mengen $A \subset \Omega$, sind die möglichen *Ereignisse*.

Bei überabzählbaren Mengen Ω darf man nicht alle Teilmengen betrachten; diese Einschränkung ist aber in der Praxis irrelevant und soll im folgenden übergangen werden.

Jedem Ereignis $A \subset \Omega$ wird eine Zahl $P(A)$ mit

$$(1) \quad 0 \leq P(A) \leq 1$$

zugeordnet, die *Wahrscheinlichkeit* von A , wobei

$$(2) \quad P(\Omega) = 1$$

ist. Wenn die abzählbar vielen Mengen A_1, A_2, \dots disjunkt sind, so soll

$$(3) \quad P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

sein; die Summe kann endlich oder unendlich sein.

Die Wahrscheinlichkeit $P(A)$ des Ereignisses A ist so zu interpretieren:

$P(A) = 0$: A ist praktisch unmöglich;

$P(A) = \frac{1}{2}$: A hat 50% Chance vorzukommen;

$P(A) = 1$: A ist praktisch sicher, man sagt auch *fast sicher*.

Jedoch besagt $P(A) = 0$ nicht immer, daß A logisch ausgeschlossen ist.

Ein besonders wichtiger Fall ist der *diskrete Wahrscheinlichkeitsraum*. Dann ist entweder Ω endlich, d.h.

$$(4) \quad \Omega = \{\omega_k : k=1, \dots, m\} \quad \text{mit} \quad m \in \mathbb{N}$$

oder abzählbar unendlich, d.h.

$$(5) \quad \Omega = \{\omega_k : k=1, 2, \dots\}.$$

Die einpunktigen Teilmengen $\{\omega_k\}$ nennt man die *Elementarereignisse*. Wir schreiben ω_k statt $\{\omega_k\}$, identifizieren also Elemente und einpunktige Mengen.

Nach (1) ist

$$(6) \quad 0 \leq P(\omega_k) \leq 1.$$

Da ja eine Menge A die disjunkte Vereinigung ihrer Elemente ist, so folgt aus (3)

$$(7) \quad P(A) = \sum_{\omega_k \in A} P(\omega_k).$$

Wegen $P(\Omega) = 1$ ist daher

$$(8) \quad \sum_{k=1}^m P(\omega_k) = 1 \quad \text{falls} \quad \Omega = \{\omega_1, \dots, \omega_m\} \text{ endlich ist,}$$

$$(9) \quad \sum_{k=1}^{\infty} P(\omega_k) = 1 \quad \text{falls} \quad \Omega \text{ abzählbar unendlich ist.}$$

Dies gilt also für diskrete Wahrscheinlichkeitsräume.

[Für überabzählbares Ω ist $\sum_{\omega \in A} P(\omega) = 1$ sinnlos!]

Beispiele.

1. Ein Würfel werde zweimal nacheinander geworfen. Das erste Mal ergebe sich die Augenzahl x , das zweite Mal y . Die 36 Elementarereignisse sind

$$(x, y) : x=1, \dots, 6; y=1, \dots, 6.$$

Es ist

$$\Omega = \{(1,1), (1,2), \dots, (6,5), (6,6)\}.$$

Wenn es sich um gute Würfel handelt, so sind alle 36 Elementarereignisse gleich wahrscheinlich. Aus (8) mit $m=36$ folgt also

$$P(x, y) = \frac{1}{36} \quad \text{für} \quad x=1, \dots, 6; y=1, \dots, 6.$$

Wir betrachten einige spezielle Ereignisse:

$A =$ (beide Male hat man dasselbe Ergebnis),

$B =$ (die Augenzahl zusammen ist = 6).

Man hat

$$A = \{(1,1), (2,2), \dots, (6,6)\},$$

also nach (7)

$$P(A) = \frac{6}{36} = \frac{1}{6} \approx 16.7\%.$$

Weiter ist

$$B = \{(1,5), (2,4), (3,3), (4,2), (5,1)\},$$

also nach (7)

$$P(B) = \frac{5}{36} \approx 13.9\%.$$

2. Wir betrachten die Lebensdauer eines Gerätes in Monaten. Als Wahrscheinlichkeitsraum betrachten wir

$$\Omega = \{0, 1, 2, \dots\} = \mathbb{N} \cup \{0\};$$

zwar hält nichts mehr als 10^{10} Monate, aber die Beschreibung wird so einfacher. Wir setzen

$P(0) =$ Wahrscheinlichkeit für von Anfang an defektes Gerät,

$P(k) =$ Wahrscheinlichkeit dafür, daß das Gerät im k -ten Monat defekt wird.

Es muß

$$\sum_{k=0}^{\infty} P(k) = 1$$

sein. Weiter ist

$$1 - P(0) = \sum_{k=1}^{\infty} P(k) = P(\text{Gerät funktioniert bei Lieferung}),$$

$$\sum_{k=0}^{24} P(k) = P(\text{Gerät hält maximal zwei Jahre}).$$

Rechenregeln. Wir betrachten einen beliebigen Wahrscheinlichkeitsraum. Wenn A ein Ereignis ist, d.h. $A \subset \Omega$, so beschreibt

$$(10) \quad \bar{A} = \Omega \setminus A,$$

daß das Ereignis nicht eintritt. Da A und \bar{A} disjunkt sind, folgt aus (2) und (3)

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}),$$

also

$$(11) \quad P(\bar{A}) = 1 - P(A).$$

Weiter seien A, B beliebig. Wir setzen $C = B \setminus (A \cap B)$. Dann ist

$$A \cap C = \emptyset, \quad A \cup C = A \cup B.$$

Aus (3) folgt also

$$(12) \quad P(A \cup B) = P(A \cup C) = P(A) + P(C).$$

Weiter ist B die disjunkte Vereinigung von $A \cap B$ und C, also nach (3)

$$P(B) = P(A \cap B) + P(C).$$

Zusammen mit (12) ergibt dies

$$(13) \quad P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Wegen $P(A \cap B) \geq 0$ folgt hieraus

$$(14) \quad P(A \cup B) \leq P(A) + P(B).$$

Allgemeiner ist immer

$$(15) \quad P(U \cap A_n) \leq \sum \frac{P(A_n)}{n};$$

man vergleiche (3) für den disjunkten Fall.

Beispiel. Von einem Spiel mit 32 Karten werden zwei Karten ausgeteilt ("Skat"). Der Wahrscheinlichkeitsraum besteht aus den $31 \cdot 32 = 992$ sich ergebenden Elementarereignissen wie z.B.

(Pik A, Karo 3), (Pik A, Herz 8), (Karo 3, Pik A).

Sie haben alle die Wahrscheinlichkeit $\frac{1}{992}$.

Wir betrachten die Ereignisse

B = "erste Karte ist ein As" (und die zweite beliebig),

C = "zweite Karte ist ein As" (und die erste beliebig),

A_j = "insgesamt sind es j = 0, 1, 2 Asse".

Es ist

$$P(B) = \frac{4}{32} = \frac{1}{8}, \quad P(C) = \frac{1}{8}.$$

Insgesamt gibt es 12 Möglichkeiten, zwei Asse zu erhalten. Also ist

$$(16) \quad P(A_2) = \frac{12}{992} = \frac{3}{248} \approx 1,2\%.$$

Weiter ist $A_2 = B \cap C$, also nach (13)

$$P(B \cup C) = P(B) + P(C) - P(B \cap C) = \frac{1}{8} + \frac{1}{8} - \frac{3}{248} = \frac{59}{248}$$

Nun ist $B \cup C =$ "mindestens ein As" = \bar{A}_0 , also nach (11)

$$P(A_0) = 1 - P(\bar{A}_0) = 1 - \frac{59}{248} = \frac{189}{248} \approx 76,2\%.$$

Da Ω die disjunkte Vereinigung von A_0, A_1 und A_2 ist, verbleibt

$$P(A_1) = 1 - P(A_0) - P(A_2) = \frac{56}{248} \approx 22,6\%.$$

15.2.1. Bedingte Wahrscheinlichkeit

1. Definition. Wir betrachten einen Wahrscheinlichkeitsraum Ω . B ein Ereignis mit $P(B) > 0$. Man definiert die bedingte Wahrscheinlichkeit durch

$$(1) \quad P(A | B) = \frac{P(A \cap B)}{P(B)}.$$

Dies ist die Wahrscheinlichkeit für das Eintreten von A , vorausgesetzt daß B eintritt.

Beispiel (Fortsetzung von S. 15.1.6). Sei wieder

$B =$ "erste Karte ist ein As",
 $C =$ "zweite Karte ist ein As".

Wenn B eintritt, so kann $A_0 =$ "insgesamt kein As" nicht eintreten, d.h.

$$P(A_0 | B) = 0.$$

Weiter enthalten die übrigen 31 Karten genau 3 Asse. Also ist

$$(2) \quad P(C | B) = \frac{3}{31}.$$

In (16) auf S. 15.1.6 hatten wir gesehen, daß $P(B \cap C) = P(A_2) = \frac{3}{248}$ ist. Also ist

$$(3) \quad \frac{P(B \cap C)}{P(B)} = \frac{3/248}{3/31} = \frac{3}{178}.$$

Nach (1) müssen ja (2) und (3) übereinstimmen. Weiter ist $A_2 \cap B = B \cap C$, also nach (3)

$$P(A_2 | B) = \frac{P(B \cap C)}{P(B)} = \frac{3}{31} \approx 9.7\%.$$

Es folgt noch

$$P(A_1 | B) = 1 - P(A_2 | B) = \frac{28}{31} \approx 90.3\%.$$

Rechenregeln. Man rechnet mit $P(A | B)$ genau wie mit $P(A)$. So ist z.B.

$$(4) \quad P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B) - P(A_1 \cap A_2 | B).$$

Sei Ω zerlegt in abzählbar viele Ereignisse B_n , die sich gegenseitig ausschließen und jeweils positive Wahrscheinlichkeit haben, d.h.

$$(5) \quad \Omega = \cup_n B_n, \quad B_m \cap B_n = \emptyset \quad \text{für } m \neq n, \quad P(B_n) > 0.$$

Dann ist für $A \subset \Omega$

$$\begin{aligned} A &= A \cap \Omega = A \cap (B_1 \cup B_2 \cup \dots) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup \dots \end{aligned}$$

Wegen $(A \cap B_m) \cap (A \cap B_n) = A \cap B_m \cap B_n = \emptyset$ für $m \neq n$ ist nach (3) auf S. 15.1.1 daher

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots$$

und nach der Definition der bedingten Wahrscheinlichkeit ist dies

$$= P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots$$

Es ist also unter der Voraussetzung (4)

$$(6) \quad P(A) = \sum_n P(A | B_n)P(B_n).$$

Beispiel. Bei einer Krankheit sei die Todesrate 18% bei Männern und 7% bei Frauen. Die Bevölkerung setze sich zu 48% aus Männern und zu 52% aus Frauen zusammen. Wie groß ist die gesamte Todesrate der Krankheit? Offenbar wird stillschweigend vorausgesetzt, daß Männer und Frauen gleichoft daran erkranken. Wir wählen $\Omega = \{\text{Erkrankte}\}$, $M = \{\text{Männer}\}$, $F = \{\text{Frauen}\}$, $T = \{\text{Todesfälle}\}$ und erhalten aus (9)

$$P(T) = P(T | M)P(M) + P(T | F)P(F) \\ = \frac{18}{100} \cdot \frac{48}{100} + \frac{7}{100} \cdot \frac{52}{100} = 0.1228 \approx 12.3\%$$

Unabhängige Ereignisse. Die Ereignisse A und B heißen unabhängig (genauer stochastisch oder statistisch unabhängig), wenn

$$(7) \quad P(A \cap B) = P(A)P(B)$$

ist. Falls $P(B) > 0$ ist, kann man (7) wegen (1) als

$$(8) \quad P(A | B) = P(A)$$

schreiben. Das heißt, daß die Wahrscheinlichkeit von A unabhängig davon ist, ob B eintritt oder nicht. Falls $P(A) > 0$ ist, kann man (7) schreiben als

$$(9) \quad P(B | A) = P(B),$$

d.h. die Wahrscheinlichkeit von B hängt nicht vom Eintreten von A ab.

Beispiel. Wir nehmen ein Kartenspiel mit 32 Karten und ziehen daraus eine Karte. Dies definiert einen Wahrscheinlichkeitsraum mit 32 Elementarereignissen der Wahrscheinlichkeit $\frac{1}{32}$. Wir betrachten die Ereignisse

$$A = \text{"ein As"}, \quad H = \text{"eine Herzkarte"}, \quad K = \text{"ein König"}.$$

Es ist

$$P(A) = \frac{4}{32} = \frac{1}{8}, \quad P(H) = \frac{8}{32} = \frac{1}{4}, \quad P(K) = \frac{1}{8}, \quad P(\bar{K}) = 1 - \frac{1}{8} = \frac{7}{8}.$$

Weiter ist

$$P(A \cap \bar{K}) = P(A) = \frac{1}{8}.$$

Also ist

$$P(A | \bar{K}) = \frac{P(A \cap \bar{K})}{P(\bar{K})} = \frac{1}{8} \cdot \frac{8}{7} = \frac{1}{7} \neq P(A) = \frac{1}{8}.$$

Die Ereignisse "As" und "kein König" sind also abhängig. Das ist ja auch klar, da es zwar 32 Karten, aber nur 28 Nichtkönige gibt. Dagegen ist

$$P(A \cap H) = \frac{1}{32} = \frac{1}{8} \cdot \frac{1}{4} = P(A)P(H),$$

d.h. A und H sind unabhängig.

Unabhängigkeit bei mehreren Ereignissen. Die Ereignisse

$$A_1, A_2, \dots$$

heißen unabhängig, wenn

$$(10) \quad P(A_m \cap A_n) = P(A_m)P(A_n) \quad (m \neq n)$$

$$(11) \quad P(A_m \cap A_n \cap A_k) = P(A_m)P(A_n)P(A_k) \quad (m \neq n, n \neq k, k \neq m)$$

usw. ist. Es genügt also nicht nur, die wechselseitige Unabhängigkeit zu fordern, die in (10) formuliert ist.

Wiederholte Ereignisse. Sei Ω ein Wahrscheinlichkeitsraum.

Man bildet

$$(12) \quad \Omega \times \Omega = \{(\omega_1, \omega_2) : \omega_1 \in \Omega, \omega_2 \in \Omega\}.$$

Wenn z.B. durch Ω ein Experiment beschrieben wird, so beschreibt $\Omega \times \Omega$ die zweimalige Durchführung. Die n-malige Durchführung wird beschrieben durch

$$(13) \quad \Omega^n = \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i \in \Omega, i=1, \dots, n\}$$

Wir nennen Ω^n den durch n-fache Wiederholung erhaltenen Raum.

Besonders wichtig ist die unabhängige Wiederholung. Durch

$$(A_1, A_2, \dots, A_n) \subset \Omega^n$$

wird das folgende Ereignis beschrieben: Beim ersten Experiment A_1 , beim zweiten A_2 , usw. Bei unabhängiger Wiederholung ist

$$(14) \quad P(A_1, A_2, \dots, A_n) = P(A_1)P(A_2) \dots P(A_n) \text{ für } A_1, \dots, A_n \subset \Omega.$$

Beispiel 1. Wir betrachten ein Würfelspiel. Durch $\Omega = \{1, 2, \dots, 6\}$ wird ein Wurf beschrieben, durch

$$\Omega^2 = \{(1, 1), (1, 2), \dots, (6, 6)\} \quad (\text{dazu S. 15.1.3})$$

zwei Würfe usw. Dies ist eine unabhängige Wiederholung. Nehmen wir etwa das Ereignis $A_1 =$ "es kommt eine Eins". Dann ist

$$(A_1, \dots, A_1) = \text{"es kommt n-mal eine Eins"}.$$

Wegen $P(A_1) = \frac{1}{6}$ hat dies nach (14) die Wahrscheinlichkeit $\frac{1}{6^n}$.

Ähnlich ist

$$P(\text{"gerade, ungerade, gerade, usw."}) = \frac{1}{2^n};$$

$$P(\text{"n-mal keine Sechsen"}) = \left(\frac{5}{6}\right)^n,$$

da ja $P(\text{"keine Sechsen"}) = \frac{5}{6}$ ist. Für $n=4$ ist dies $\approx 0.482 \approx 48.2\%$.

Beispiel 2. Sei P die Wahrscheinlichkeit für das Auftreten eines Buchstabens in deutschen Texten. Bekanntlich ist

$$P(E) > P(N) > \dots,$$

Die Zahlenwerte weiß ich nicht.

Die unabhängige Wiederholung sieht so aus: Man schlägt irgend ein Buch auf und wählt etwa den 177. Buchstaben. Dann nimmt man ein anderes Buch und tut dasselbe, usw.

Dagegen ist es keine unabhängige Wiederholung, wenn man ein Buch aufschlägt und dort n aufeinanderfolgende Buchstaben nimmt. Offenbar ist

$$P(EE) < P(E)^2 \quad (\text{da "EE" sehr selten ist})$$

und da "EAXUVXQQFAXYZ" in der deutschen Literatur eventuell nur einmal vorkommt (\hat{A}), ist die Wahrscheinlichkeit extrem gering, jedenfalls kleiner als der Wert

$$P(A)^2 P(E) P(F) P(Q)^2 P(U) P(V) P(X)^3 P(Y) P(Z),$$

der sich bei unabhängiger Wiederholung zu ergeben hätte.

Wir werden Zufallsvariable mit großen Buchstaben bezeichnen. Wir betrachten nur reelle eindimensionale Zufallsvariable. Eine zweidimensionale Zufallsvariable ist z.B. der Ort (X, Y) , wo ein Würfel liegen bleibt.

Diskrete Zufallsvariable. Die Zufallsvariable X heißt diskret wenn sie nur endlich oder abzählbar unendlich viele mögliche Werte x_k hat. Die Zufallsvariable X wird vollständig beschrieben durch die Ereignisse

$$(1) \quad A_k = \{\omega \in \Omega : X(\omega) = x_k\}, \text{ kurz } A_k = \{X = x_k\},$$

wobei $k = 1, \dots, m$ (endlicher Fall) oder $k = 1, 2, \dots$ (abzählbar unendlicher Fall) ist.

Sei X eine Zufallsvariable mit den möglichen Werten x_k . Der Erwartungswert von X ist definiert durch

$$(2) \quad E(X) = \sum_k x_k P\{X = x_k\}.$$

Im endlichen Fall ist also

$$(3) \quad E(X) = \sum_{k=1}^m x_k P\{X = x_k\}.$$

Im unendlichen Fall ist

$$(4) \quad E(X) = \sum_{k=1}^{\infty} x_k P\{X = x_k\}. \quad (\text{oder auch z.B. } \sum_{k=0}^{\infty} \dots)$$

Man verlangt, daß die Reihe absolut konvergiert, da es bei absolut konvergenten Reihen auf die Reihenfolge der Summanden nicht ankommt.

Beispiele. 1. (Fortsetzung von S 15.3.1, Beispiel 1). Es ist $x_1 = 2, \dots, x_{11} = 12$. Weiter kann der Fall $X = 3$ durch (1,2) und (2,1) realisiert werden. Da jeder Fall die Wahrscheinlichkeit $\frac{1}{36}$ hat, ist $P\{X = 3\} = \frac{2}{36} = \frac{1}{18}$. Man erhält so als Erwartungswert

15.3. Zufallsvariable und Erwartungswert

Zufallsvariable. Wir gehen aus von einem Wahrscheinlichkeitsraum Ω , für dessen Ereignisse $A \subset \Omega$ Wahrscheinlichkeiten $P(A)$ gegeben sind.

Eine reelle Zufallsvariable X ist (mathematisch gesehen) eine Abbildung von Ω in \mathbb{R} , d.h.

$$\omega \in \Omega \rightarrow X(\omega) \in \mathbb{R}.$$

Also ist X eine Größe, deren Wert vom Zufall abhängt.

Beispiele 1. (Fortsetzung von S.15.1.3) Die Summe X der Augenzahlen ist eine Zufallsvariable. Als Werte kommen die Zahlen 2,3,...,12 in Frage. Eine andere Zufallsvariable Y ist wie folgt definiert

$$Y = 1 \text{ falls beide Augenzahlen gleich,} \\ Y = 0 \text{ sonst.}$$

2. Von einer Fabrik werde serienmäßig ein Gerät ausgeliefert. Die Lebensdauer L ist dann eine Zufallsvariable. Man beachte, daß der zugrunde gelegte Wahrscheinlichkeitsraum gar nicht klar definiert ist, aber trotzdem die Zufallsvariable L eindeutig bestimmt ist.

3. (Fortsetzung von S.15.1.6) Die Anzahl X der Asse bei den beiden ausgeteilten Karten ist eine Zufallsvariable. Es ist

$$X(\omega) = \begin{cases} 0 & \text{für } \omega \in A_0, \\ 1 & \text{für } \omega \in A_1, \\ 2 & \text{für } \omega \in A_2. \end{cases}$$

Umgekehrt hätten wir definieren können

$$A_j = \{\omega \in \Omega : X(\omega) = j\}, \quad j = 0, 1, 2.$$

für die Augenzahl zweier Würfel

$$\begin{aligned}
E(X) &= 2P\{X=2\} + 3P\{X=3\} + \dots + 12P\{X=12\} \\
(5) \quad &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} \\
&\quad + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{3}{36} + 12 \cdot \frac{1}{36} = \frac{252}{36} = 7.
\end{aligned}$$

Weiter ist für Y wie auf S. 15.3.1, Mitte:

$$E(Y) = 0 \cdot P\{\text{verschieden}\} + 1 \cdot P\{\text{gleich}\} = 1 \cdot \frac{1}{6} = \frac{1}{6}.$$

2. Sei X eine Zufallsvariable mit den möglichen Werten 1, 2, ... und zwar sei

$$P\{X = k\} = \frac{1}{2^k} \quad (k = 1, 2, \dots).$$

Dann ist

$$(6) \quad E(X) = \sum_{k=1}^{\infty} k \cdot \frac{1}{2^k} = 2;$$

es ist nämlich

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k, \quad \frac{1}{(1-x)^2} = \frac{d}{dx} \sum_{k=0}^{\infty} x^k = \sum_{k=1}^{\infty} k x^{k-1}$$

für $|x| < 1$, woraus (6) für $x = \frac{1}{2}$ folgt.

Rechenregeln für den Erwartungswert. Bei der Menge der möglichen Werte x_k der Zufallsvariablen X darf man auch Werte zulassen, die nie vorkommen. Diese haben dann die Wahrscheinlichkeit 0, d.h. es ist $P\{X = x_k\} = 0$, falls X immer $\neq x_k$ ist. In der Summe

$$E(X) = \sum_k x_k P\{X = x_k\}$$

liefern diese x_k dann keinen Beitrag.

Weiter ist die disjunkte Vereinigung der $\{\omega \in \Omega : X = x_k\} = \{X = x_k\}$. Da es sich um höchstens abzählbar viele (und disjunkte) Mengen handelt, folgt aus (3) und (2) auf S.15.1.1

$$(7) \quad \sum_k P\{X = x_k\} = P(\Omega) = 1.$$

Sei a eine Konstante (kleines a), d.h. a hängt nicht vom Zufall ab. Dann können wir a natürlich als Zufallsvariable betrachten, für die a der einzig mögliche Wert ist. Also ist

$$(8) \quad E(a) = a.$$

Der Erwartungswert verhält sich linear, d.h. es ist

$$(9) \quad E(aX + bY) = aE(X) + bE(Y) \quad (a, b \text{ Konstante}).$$

Zum Beweis seien x_k und y_l die möglichen Werte von X bzw. Y. Dann hat $aX + bY$ die möglichen Werte

$$(10) \quad ax_k + by_l \quad (k, l \text{ variieren unabhängig}),$$

wobei verschiedene (k, l)-Kombinationen eventuell denselben Wert ergeben. Aus (2) und 15.1.1 (3) folgt

$$\begin{aligned}
(11) \quad E(aX + bY) &= \sum_k \sum_l (ax_k + by_l) P\{X = x_k, Y = y_l\} \\
&= a \sum_k x_k \sum_l P\{X = x_k, Y = y_l\} + b \sum_l y_l \sum_k P\{\dots\}.
\end{aligned}$$

Nun ist $\{X = x_k\}$ die disjunkte Vereinigung der Mengen $\{X = x_k, Y = y_l\}$. Somit ist

$$P\{X = x_k\} = \sum_l P\{X = x_k, Y = y_l\}.$$

Aus (11) folgt also (zus. mit einer analogen Identität $P\{Y = y_l\} = \dots$)

$$E(ax+by) = a \sum_k x_k P\{X=x_k\} + b \sum_l y_l P\{Y=y_l\},$$

und das ist unsere Behauptung (9).

Wenn $f(t)$ ($t \in \mathbb{R}$) eine reelle Funktion ist, so ist $f(X)$ eine Zufallsvariable mit den möglichen Werten $f(x_k)$. Also ist

$$(12) \quad E(f(X)) = \sum_k f(x_k) P\{X=x_k\}.$$

Für $f(t) = t^2$ ergibt sich so z.B.

$$(13) \quad E(X^2) = \sum_k x_k^2 P\{X=x_k\}.$$

Die Varianz. Wir definieren die *Varianz* der Zufallsvariablen X durch

$$(14) \quad V(X) = E([X-E(X)]^2).$$

Wenn wir $c = E(X)$ schreiben, so ist also nach (9) und (8)

$$\begin{aligned} V(X) &= E([X-c]^2) = E(X^2 - 2cX + c^2) \\ &= E(X^2) - 2cE(X) + c^2 \\ &= E(X^2) - 2c^2 + c^2 = E(X^2) - c^2. \end{aligned}$$

Also ist

$$(15) \quad V(X) = E(X^2) - E(X)^2.$$

Ausgeschrieben ist nach (14) mit $c = E(X)$

$$(16) \quad V(X) = \sum_k (x_k - c)^2 P\{X=x_k\}.$$

Beispiel. Wir betrachten wieder die zwei Würfel und die Zufallsvariable $X =$ Summe der Augenzahlen. Dann ist (man vergl. (5) auf S.15.3.3)

$$E(X^2) = 4 \cdot \frac{1}{36} + 9 \cdot \frac{2}{36} + \dots + 144 \cdot \frac{1}{36} = \frac{1974}{36} = \frac{329}{6}.$$

Wegen $E(X) = 7$ folgt aus (15)

$$V(X) = \frac{329}{6} - 49 = \frac{35}{6} \approx 5.83.$$

Direkt aus (16) folgt mit $c = 7$

$$V(X) = (-5)^2 \cdot \frac{1}{36} + (-4)^2 \cdot \frac{2}{36} + \dots + 5^2 \cdot \frac{1}{36} = \frac{35}{6}.$$

Für die Zufallsvariable Y mit $Y = 1$ falls Augenzahlen gleich, $Y = 0$ sonst, ist $E(Y) = \frac{1}{6}$, also

$$V(Y) = E(Y^2) - E(Y)^2 = 0 \cdot \frac{5}{6} + 1 \cdot \frac{1}{6} - \frac{1}{36} = \frac{5}{36}.$$

Die TSCHEBYSCHESCHE UNGLEICHUNG. Wir kommen nun zur Bedeutung der Varianz. Sei $t > 0$ gegeben und sei $c = E(X)$. Es ist

$$(16) \quad V(X) = \sum_k (x_k - c)^2 P\{X=x_k\}.$$

Wir summieren nun nur noch über alle k mit

$$(17) \quad |x_k - c| \geq t\sqrt{V(X)}.$$

Da alle Summanden in (16) nichtnegativ sind, folgt

$$\begin{aligned} V(X) &\geq \sum_{k \text{ mit (17)}} (x_k - c)^2 P\{X=x_k\} \\ &\geq t^2 V(X) \sum_{k \text{ mit (17)}} P\{X=x_k\}. \end{aligned}$$

Nach (7) auf S.15.1.2 ist die letzte Summe $= P\{|X-c| \geq t\sqrt{V(X)}\}$. Also ist

$$V(X) \geq t^2 V(X) P\{|X-c| \geq t\sqrt{V(X)}\},$$

also

$$(18) \quad P\{|X-c| \geq t\sqrt{V(X)}\} \leq \frac{1}{t^2} \quad (0 < t < \infty)$$

mit $c = E(X)$. Das ist die TSCHEBYSCHESCHE Ungleichung. Für $t = 100$ erhält man z.B.

$$(19) \quad P(|X-c| \geq 100\sqrt{V(X)}) \leq \frac{1}{10000} \approx 0.01\%$$

Es ist also sehr unwahrscheinlich, daß X vom Erwartungswert um mehr als $100\sqrt{V(X)}$ abweicht. Man kann (19) auch so ausdrücken: Es ist ziemlich sicher, daß

$$E - 100\sqrt{V(X)} < X < E(X) + 100\sqrt{V(X)}$$

ist. Man nennt $\sqrt{V(X)}$ daher die Streuung von X und bezeichnet sie mit σ , also

$$(20) \quad \sigma = \sqrt{V(X)}.$$

Kontinuierliche Zufallsvariable. Nun betrachten wir eine Zufallsvariable X , die kontinuierliche Werte in $I \subset \mathbb{R}$ annimmt. Wir definieren die Verteilungsfunktion durch

$$(21) \quad F(x) = P\{X \leq x\} \quad \text{für} \quad -\infty < x < +\infty.$$

Wir setzen künftig voraus, daß $F(x)$ im Wertebereich I stetig differenzierbar ist, also dort eine stetige Dichte

$$(22) \quad f(x) = F'(x) \quad (x \in I)$$

besitzt. Aus der Stetigkeit von F folgt $P\{X < x\} = P\{X \leq x\} = F(x)$.

Für $h > 0$ ist

$$\begin{aligned} F(x+h) - F(x) &= P\{x < X < x+h\} = P\{X < x\} + P\{x \leq X \leq x+h\} \\ &= F(x) + P\{x \leq X \leq x+h\}, \end{aligned}$$

also nach (22) für $x \in I$

$$(23) \quad f(x) = \lim_{h \rightarrow +0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow +0} \frac{P\{x \leq X \leq x+h\}}{h}.$$

Somit liegt X mit der Wahrscheinlichkeit $\approx f(x)h$ im Intervall $[x, x+h]$.

Beispiel. Es werde eine reelle Zahl Z (rational oder irrational) zufällig herausgegriffen. Ihr nichtganzzahliger Anteil sei X , also

$$Z = \text{int } Z + X, \quad 0 \leq X < 1.$$

Dann kommt jeder Wert x in $I = [0, 1)$ mit gleicher Wahrscheinlichkeit vor. Es ist

$$F(x) = x \quad \text{für} \quad 0 \leq x < 1,$$

außerdem $F(x) = 0$ für $x < 0$ und $F(x) = 1$ für $1 \leq x < \infty$. Weiter ist

$$f(x) = F'(x) = 1 \quad \text{für} \quad 0 \leq x < 1.$$

Man beachte, daß $F(x)$ im Punkt 0 von rechts differenzierbar ist, also vom Intervall I aus kommend.

Erwartungswert und Varianz für kontinuierliche Zufallsvariable.

Sei nun X eine Zufallsvariable mit der Dichte $f(x)$ ($x \in I$). Nach (23) ist

$$(24) \quad f(x) \geq 0 \quad \text{für} \quad x \in I.$$

Weiter ist

$$(25) \quad \int_I f(x) dx = P\{X \in I\} = 1.$$

Wir definieren den Erwartungswert durch

$$(26) \quad E(X) = \int_I x f(x) dx$$

und die Varianz durch

$$(27) \quad V(X) = \int_I (x-c)^2 f(x) dx, \quad c = E(X).$$

Wie in (15) ist

$$(28) \quad V(X) = E(X^2) - E(X)^2.$$

Die Streuung ist definiert als $\sigma = \sqrt{V(X)}$. Die TSCHEBYSCHESCHE Ungleichung

$$(29) \quad P\{|X-c| \geq t\} \leq \frac{1}{t^2} \quad (0 < t < \infty), \quad c = E(X), \quad \sigma = \sqrt{V(X)}$$

gilt genau wie im diskreten Fall.

Beispiele.

1. (Fortsetzung von S.15.3.8) Sei $f(x) = 1$ für $0 \leq x < 1$. Dann ist

$$E(X) = \int_0^1 x dx = \frac{1}{2},$$

$$V(X) = \int_0^1 x^2 dx - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

2. Sei $f(x) = e^{-x}$ für $0 \leq x < \infty$. Es ist

$$E(X) = \int_0^{\infty} x e^{-x} dx = [-x e^{-x}]_0^{\infty} + \int_0^{\infty} e^{-x} dx = 1,$$

$$V(X) = \int_0^{\infty} x^2 e^{-x} dx - 1 = 2 - 1 = 1.$$

15.4. Unabhängige Zufallsvariable

Definition. Seien X und Y diskrete Zufallsvariable mit den möglichen Werten x_k, y_k ($k = 1, 2, \dots$). Wir sagen, daß X und Y unabhängig sind, wenn die Ereignisse

$$(1) \quad \{X = x_k\}, \{Y = y_l\} \quad (k, l = 1, 2, \dots)$$

unabhängig sind. Nach S.15.2.3 heißt das also

$$(2) \quad P\{X = x_k, Y = y_l\} = P\{X = x_k\}P\{Y = y_l\}$$

oder auch (mit der bedingten Wahrscheinlichkeit)

$$(3) \quad P\{Y = y_l | X = x_k\} = P\{Y = y_l\}.$$

Die Kenntnis, daß X den Wert x_k annimmt, ändert also nichts an der Wahrscheinlichkeit dafür, daß Y den Wert y_l annimmt.

Für mehrere diskrete Zufallsvariable X, Y, Z, \dots mit den möglichen Werten $x_k^{(n)}$ fordert man für die Unabhängigkeit, daß alle Ereignisse

$$(4) \quad \{X = x_k\}, \{Y = y_l\}, \{Z = z_j\}, \dots$$

unabhängig sind.

Für kontinuierliche Zufallsvariable X, Y ersetzt man in der Definition Werte durch Intervalle: Man fordert, daß die Ereignisse

$$(5) \quad \{X \leq x\}, \{Y \leq y\} \quad (x \in \mathbb{R}, y \in \mathbb{R})$$

unabhängig sind. Es soll also (analog zu (3)) gelten

$$(6) \quad P\{X \leq x | Y \leq y\} = P\{X \leq x\}.$$

Entsprechendes gilt für mehr als zwei unabhängige Zufallsvariable.

Beispiele.

1. Aus einem Kartenspiel werde eine Karte gezogen. Wir geben ihr den Wert der darauf stehenden Zahl (mit $B = 11$, $D = 12$, $K = 13$, $A = 14$). Das ist eine Zufallsvariable X mit den möglichen Werten $2, 3, \dots, 14$.

Nun ziehen wir eine zweite Karte. Das definiert eine weitere Zufallsvariable Y . Ob X und Y unabhängig sind, hängt von den Details ab. Wir haben zwei Fälle:

(i) Die erste Karte wird zurückgelegt und dann wird gründlich durchgemischt. Dann sind X und Y unabhängige Zufallsvariable.

(ii) Die erste Karte wird nicht zurückgelegt. Dann sind X und Y nicht unabhängig. So ist z.B.

$$P\{Y = 10 \mid X = 10\} = \frac{3}{51}, \text{ aber } P\{X = 10\} = \frac{4}{52} > \frac{3}{51}.$$

2. Bei einem Experiment ergebe sich ein Wert X . Das ist natürlich eine Zufallsvariable, da das Ergebnis von Zufallsschwankungen der Apparatur usw. abhängt.

Nun werde das Experiment unabhängig mehrfach wiederholt, wobei "unabhängig" heißen soll, daß das Ergebnis eines Versuches keinen Einfluß auf die Ergebnisse der anderen Versuche haben soll. Das Ergebnis des n -ten Versuchs bezeichnen wir mit X_n . Dann haben wir die Zufallsvariablen

$$X_1, X_2, \dots$$

und wir dürfen annehmen, daß diese unabhängig sind.

In gewisser Weise sind X_1, X_2, \dots alle "dieselben" Zufallsvariablen X , wir müssen sie aber bezeichnungstechnisch unterscheiden, da sie natürlich verschiedene Werte annehmen können. Jedoch sind sie *identisch verteilt*, d.h. es sind

$$(7) \quad P\{X_n = x_k\}, E\{X_n\}, V\{X_n\} \text{ unabhängig von } n.$$

Eigenschaften. Seien X und Y unabhängige Zufallsvariable. Dann ist

$$(8) \quad E\{XY\} = E\{X\}E\{Y\}.$$

Das wollen wir für diskrete Zufallsvariable beweisen. Nach (2) auf S. 15.3.2 ist

$$\begin{aligned} E\{X\}E\{Y\} &= \sum_k x_k P\{X = x_k\} \cdot \sum_l y_l P\{Y = y_l\} \\ &= \sum_k \sum_l x_k y_l P\{X = x_k\} P\{Y = y_l\}. \end{aligned}$$

Nach (2) ist daher

$$(9) \quad E\{X\}E\{Y\} = \sum_k \sum_l x_k y_l P\{X = x_k, Y = y_l\}.$$

Nun hat die Zufallsvariable XY als mögliche Werte $x_k y_l$. Wenn alle diese Produkte verschieden sind, so steht rechts in (9) der Erwartungswert $E\{XY\}$. Wenn einige Produkte zusammenfallen, so muß man die entsprechenden Werte noch zusammenfassen.

Dies wenden wir auf die Varianz an. Seien X und Y unabhängig. Aus der Linearität folgt

$$\begin{aligned} E\{(X+Y)^2\} &= E\{X^2 + 2XY + Y^2\} \\ &= E\{X^2\} + 2E\{XY\} + E\{Y^2\}. \end{aligned}$$

Aus (8) erhält man also

$$E\{(X+Y)^2\} = E\{X^2\} + 2E\{X\}E\{Y\} + E\{Y^2\}.$$

Nach (15) auf S. 15.3.5 ist daher

$$\begin{aligned} V\{(X+Y)\} &= [E\{X^2\} + 2E\{X\}E\{Y\} + E\{Y^2\}] - [E\{X\} + E\{Y\}]^2 \\ &= E\{X^2\} - E\{X\}^2 + E\{Y^2\} - E\{Y\}^2. \end{aligned}$$

Wir erhalten also nach (15) auf S. 15.3.5

(11) $V(X+Y) = V(X) + V(Y)$ (X, Y unabhängig)

Wenn X_1, X_2, \dots, X_n unabhängige Zufallsvariable sind, so gilt allgemeiner

(12) $E(X_1 X_2 \dots X_n) = E(X_1) E(X_2) \dots E(X_n)$

(13) $V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n)$

Das Gesetz der großen Zahl. Seien X, X_1, X_2, \dots unabhängige, identisch verteilte Zufallsvariablen. Dann gilt fast sicher

(14) $\frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow E(X)$ für $n \rightarrow \infty$,

falls der (von n unabhängige) Erwartungswert $E(X)$ existiert.

Die Zufallsvariablen X_n sollen also alle dieselbe Verteilung wie X besitzen, d.h. es soll (z.B. für diskrete Zufallsvariable)

(15) $P\{X_n = x_k\} = P\{X = x_k\}$ ($n = 1, 2, \dots$)

für alle k gelten. Es folgt $E(X_n) = E(X)$, $V(X_n) = V(X)$ ($n = 1, 2, \dots$).

Man kann das obige (sog. starke) Gesetz der großen Zahl auch so ausdrücken: Seien X_n unabhängige identisch verteilte Zufallsvariable. Dann hat das Ereignis

(16) $\lim_{n \rightarrow \infty} \frac{1}{n} (X_1 + \dots + X_n) = E(X)$

die Wahrscheinlichkeit 1, wobei die Existenz des Erwartungswerts vorausgesetzt wird. Für großes n ist es also sehr wahrscheinlich, daß

$\frac{1}{n} (X_1 + \dots + X_n) \approx E(X)$

ist. Im Mittel ergibt sich also der Erwartungswert.

In dieser allgemeinen Form ist der Beweis recht schwierig. Wir beweisen nur etwas weniger, nämlich das sog. schwache Gesetz der großen Zahl unter der zusätzlichen Annahme, daß $V(X) < \infty$ ist.

Sei $S_n = X_1 + \dots + X_n$. Dann ist nach (13)

$E(S_n) = E(X_1) + \dots + E(X_n) = nE(X)$,

$V(S_n) = V(X_1) + \dots + V(X_n) = nV(X)$.

Die TSCHEBYSCHESCHE Ungleichung angewandt auf S_n ergibt daher

$P\{|S_n - nE(X)| \geq t\sqrt{nV(X)}\} \leq \frac{1}{t^2}$ für $0 < t < \infty$.

Wir wählen $t = \epsilon\sqrt{n}$, $\epsilon > 0$ und erhalten

(17) $P\left\{\left|\frac{S_n}{n} - E(X)\right| \geq \epsilon\sqrt{V(X)}\right\} \leq \frac{1}{\epsilon^2 n} \rightarrow 0$ für $n \rightarrow \infty$.

D.h. daß

$\left|\frac{S_n}{n} - E(X)\right| \geq \epsilon\sigma$, $\sigma = \sqrt{V(X)}$ Streuung

ist, hat eine Wahrscheinlichkeit $\leq \frac{1}{\epsilon^2 n}$, anders ausgedrückt

(18) $\left|\frac{S_n}{n} - E(X)\right| < \epsilon\sigma$

mit einer Wahrscheinlichkeit $\geq 1 - \frac{1}{\epsilon^2 n}$ ($\rightarrow 1$ für $n \rightarrow \infty$).

Das ist zwar mathematisch nicht genau unsere Behauptung, kommt ihr aber doch recht nahe.

Beispiel. Ein Würfel werde unabhängig n-mal geworfen. Die Augenzahlen seien X_1, \dots, X_n . Sie sind identisch verteilt mit

$E(X) = \frac{7}{2}$, $V(X) = E(X^2) - \frac{49}{4} = \frac{35}{12}$, $\sigma < 1.71$.

Also ist nach (17) mit $\epsilon = 0.1$, $n = 1000$

$P\left\{\left|\frac{S_{1000}}{1000} - \frac{7}{2}\right| \geq 0.171\right\} \leq \frac{1}{10}$.

Später werden wir ein besseres Ergebnis erhalten.

Die Binomialverteilung. Gegeben sei ein Experiment mit den möglichen Ereignissen 0 und 1 mit $P\{X=1\} = p$ und $P\{X=0\} = 1-p$. Dieses Experiment wird nun n-mal unabhängig durchgeführt und die Zufallsvariablen X_v seien

$$(19) \quad X_v := \begin{cases} 1 & \text{falls } 1 \text{ im } v\text{-ten Versuch} \\ 0 & \text{falls } 0 \text{ im } v\text{-ten Versuch.} \end{cases}$$

Die Zufallsvariable

$$(20) \quad S_n := X_1 + \dots + X_n$$

gibt dann die Anzahl der Einsen bei n Versuchen an. Es gilt

$$(21) \quad P\{S_n = k\} = \binom{n}{k} p^k (1-p)^{n-k} =: b(k;n,p) \quad (k = 0, 1, \dots, n),$$

d.h. die Wahrscheinlichkeit bei n Versuchen genau k Einsen zu erhalten ist $\binom{n}{k} p^k (1-p)^{n-k}$. Man sagt die Zufallsvariable S_n ist binomial verteilt mit den Parametern n und p.

Man zeigt (vgl. Übungen)

$$(22) \quad E(S_n) = np, \quad V(S_n) = np(1-p).$$

Beispiel. Ein "gutes" Einmarkstück wird 4 mal geworfen. Dann ist $p = \frac{1}{2}$ und die Wahrscheinlichkeit, daß sich $k = 0, 1, \dots, 4$ Einsen ergeben ist

$$P\{S_4 = k\} = \binom{4}{k} \frac{1}{2^4}, \quad \text{z.B. } P\{S_4 = 2\} = \frac{4 \cdot 3}{1 \cdot 2} \frac{1}{2^4} = \frac{3}{8}.$$

Approximation der Binomialverteilung durch die POISSONverteilung. Da für großes n die Berechnung von $\binom{n}{k} p^k (1-p)^{n-k}$ mühsam ist, ist man an Näherungsformeln interessiert.

Setzt man $\lambda := E(S_n) = np$, so ergibt sich

$$\begin{aligned} b(k;n,p) &= \binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1) \dots (n-k+1)}{k!} p^k (1-p)^{n-k} \\ &= \frac{n(n-1) \dots (n-k+1)}{n^k} (1-p)^{-k} (1-p)^n \frac{(np)^k}{k!} \\ &= \underbrace{1 \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)}_{\approx 1} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\approx 1} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\approx e^{-\lambda}} \frac{\lambda^k}{k!} \\ &\approx e^{-\lambda} \frac{\lambda^k}{k!} =: p(k;\lambda) \end{aligned}$$

für großes n, kleines k und nicht zu großes λ , also kleines p.

Satz. Sei (S_n) eine Folge von binomialverteilten Zufallsvariablen mit gleichem Erwartungswert $\lambda := E(S_n)$. Dann gilt

$$\lim_{n \rightarrow \infty} P\{S_n = k\} = p(k;\lambda) \quad (k = 0, 1, \dots).$$

Eine Zufallsvariable Y, die nur die Werte $k = 0, 1, \dots$ annimmt und zwar mit den Wahrscheinlichkeiten

$$(23) \quad P\{Y=k\} = e^{-\lambda} \frac{\lambda^k}{k!} \quad (k = 0, 1, \dots)$$

nennt man Poissonverteilt mit dem Parameter λ .

Beispiel. In einer Stadt ereignen sich V Verkehrsunfälle pro Tag, durchschnittlich 3. Wenn wir davon ausgehen, daß diese Unfälle unabhängig voneinander stattfinden, so sind die Voraussetzungen für die Poissonverteilung erfüllt, und zwar ist $\lambda = E(V) = 3$. Nach (23) ist

k	$P\{V=k\}$	$P\{V \leq k\}$
0	e^{-3}	0.0497
1	$3e^{-3}$	0.1494
2	$\frac{9}{2}e^{-3}$	0.2240
3	$\frac{9}{2}e^{-3}$	0.2240
4	$\frac{27}{8}e^{-3}$	0.1680
5	$\frac{81}{40}e^{-3}$	0.1008
6	...	0.0804
7	...	0.0216
8	...	0.0081

Mit der Substitution $X = \frac{k-\lambda}{\sigma}$ folgt weiter

$$\int_0^m \frac{1}{\sigma} \varphi\left(\frac{k-\lambda}{\sigma}\right) dk = \frac{m-\lambda}{\sigma} \int_{-\frac{\lambda}{\sigma}}^{\frac{m-\lambda}{\sigma}} \varphi(x) dx \approx \int_0^m \varphi(x) dx$$

da das Integral $\int_{-\infty}^{\infty} \varphi(x) dx = 1$ mit der GAUSSschen Funktion praktisch keinen Beitrag liefert. Definiert man jetzt

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\xi^2} d\xi,$$

so ist

$$P\{S_n \leq m\} \approx \phi\left(\frac{m-\lambda}{\sigma}\right).$$

Approximation der Binomialverteilung durch die GAUSSsche Funktion. Mit etwas größerem mathematischen Aufwand erhält man eine sehr gute Approximation der Binomialverteilung.

Satz. Sei $\lambda = E(S_n) = np$ und $\sigma = \sqrt{np(1-p)}$. Dann ist

$$b(k;n,p) \approx \frac{1}{\sigma} \varphi\left(\frac{k-\lambda}{\sigma}\right),$$

wobei $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ die GAUSSsche Funktion ist.

Faustregel. Für $\sigma > 3$ ist (24) praktisch anwendbar!

Die Wahrscheinlichkeit für genau k Einsen bei n Versuchen ist also

$$P\{S_n = k\} = b(k;n,p) \approx \frac{1}{\sigma} \varphi\left(\frac{k-\lambda}{\sigma}\right)$$

und die Wahrscheinlichkeit für höchstens m Einsen ist

$$P\{S_n \leq m\} = \sum_{k=0}^m b(k;n,p) \approx \int_0^m \frac{1}{\sigma} \varphi\left(\frac{k-\lambda}{\sigma}\right) dk.$$

d.h. die Zufallsvariable S_n^* entsteht aus der Zufallsvariablen S_n durch Normierung auf Erwartungswert 0 und Varianz 1.

Mit dieser Zufallsvariablen schreibt sich das Ergebnis als

$$P\{S_n^* \leq x\} \approx \phi(x).$$

15.5 Der zentrale Grenzwertsatz für identisch verteilte Zufallsvariable

Zentraler Grenzwertsatz. Sei (X_n) eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit existierender Varianz. Für die Zufallsvariablen

$$S_n := X_1 + \dots + X_n \quad \text{bzw.} \quad S_n^* := \frac{S_n - E(S_n)}{\sqrt{V(S_n)}}$$

und für $-\infty < x < +\infty$ gilt dann

(1) $P\{S_n^* \leq x\} \rightarrow \phi(x)$ für $n \rightarrow \infty$.

Der Beweis dieses Satzes liegt jenseits der Möglichkeiten dieser Vorlesung. Er benutzt die FOURIERtransformation.

Der Satz soll nun analysiert werden:

Man geht aus von unabhängigen Zufallsvariablen X_n ; das ist eine wesentliche Voraussetzung. Die Voraussetzung der identischen Verteilung dagegen könnte durch eine schwächere Forderung an die Varianzen ersetzt werden, ist aber für die Praxis ausreichend.

Bildet man nun die Summen S_n und normiert diese auf Erwartungswert 0 und Varianz 1, so gilt für diese normierten Zufallsvariablen S_n^*

(2) $P\{S_n^* \leq x\} \approx \phi(x)$.

Diese Aussage gilt, egal mit welcher Folge unabhängiger, identisch verteilter Zufallsvariablen man startet! Das macht diesen Satz so zentral.

Für die Praxis wichtig ist

(3) $P\{a \leq S_n^* \leq b\} = P\{S_n^* \leq b\} - P\{S_n^* < a\} \approx \phi(b) - \phi(a)$,

also

(4) $P\left\{a \leq \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq b\right\} \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\xi^2} d\xi$.

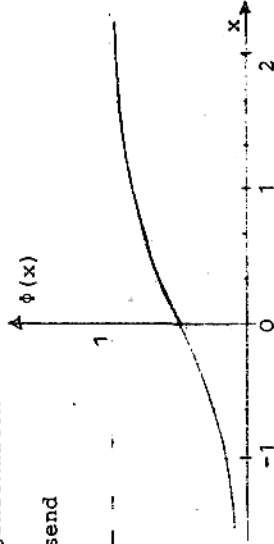
Die Funktion ϕ hat die Eigenschaften

(5) ϕ streng monoton wachsend

(6) $\lim_{x \rightarrow \infty} \phi(x) = 1$

$\lim_{x \rightarrow -\infty} \phi(x) = 0$

(7) $\phi(-x) = 1 - \phi(x)$



und liegt ansonsten tabelliert vor (siehe S.15.5.3).

Beispiele:

1. Sei (X_n) eine Folge unabhängiger, identisch verteilter Zufallsvariablen und

$$S_n := X_1 + \dots + X_n$$

Es soll die Wahrscheinlichkeit

$$P\left\{\left| \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \right| \leq x\sqrt{V(S_n)}\right\}$$

berechnet werden; also, daß die Zufallsvariable S_n weniger als das x -fache der Streuung vom Erwartungswert abweicht.

(Weiter Seite 15.5.4)

Normierung von S_n und der zentrale Grenzwertsatz liefern

$$P\{|S_n - E(S_n)| \leq x\sqrt{V(S_n)}\} = P\{-x\sqrt{V(S_n)} \leq S_n - E(S_n) \leq x\sqrt{V(S_n)}\}$$

$$= P\{-x \leq \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq x\} = P\{-x \leq S_n^* \leq x\}$$

$$\approx \Phi(x) - \Phi(-x) = \Phi(x) - (1 - \Phi(x)) = 2\Phi(x) - 1.$$

Speziell für $x = 1, 2, 3$ ergibt sich mit $\lambda := E(S_n)$ und $\sigma := \sqrt{V(S_n)}$

$$P\{\lambda - \sigma \leq S_n \leq \lambda + \sigma\} \approx 2\Phi(1) - 1 \approx 0.683$$

$$P\{\lambda - 2\sigma \leq S_n \leq \lambda + 2\sigma\} \approx 2\Phi(2) - 1 \approx 0.955$$

$$P\{\lambda - 3\sigma \leq S_n \leq \lambda + 3\sigma\} \approx 2\Phi(3) - 1 \approx 0.997.$$

2. Wie groß ist die Wahrscheinlichkeit, daß bei 1000 unabhängigen Würfeln eines Würfels die Summe der Augenzahlen vom Erwartungswert 3500 um mehr als 100 abweicht?

Man betrachtet die Zufallsvariablen

$$X_n := \text{"Augensumme beim } n\text{-ten Wurf"}.$$

Sie sind nach Voraussetzung unabhängig und identisch verteilt mit

$$E(X_n) = \frac{7}{2} \text{ und } V(X_n) = \frac{35}{12}.$$

Die Zufallsvariable

$$S_{1000} := X_1 + \dots + X_{1000}$$

gibt also die Summe der Augenzahlen bei 1000 Würfeln an und mit der Unabhängigkeit der X_n folgt

$$E(S_{1000}) = \frac{7}{2} \cdot 1000 = 3500, \quad V(S_n) = \frac{35}{12} \cdot 1000.$$

Berechnet werden soll also

$$P\{|S_{1000} - 3500| > 100\}.$$

Tabelle für $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\xi^2}{2}} d\xi$

		NORMAL DISTRIBUTION FUNCTION									
		0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359	0.5399
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	0.5793
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	0.6179
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	0.6554
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	0.6915
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	0.7258
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549	0.7581
0.7	0.7581	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	0.7881
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	0.8160
0.9	0.8160	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8390	0.8415
1.0	0.8415	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621	0.8643
1.1	0.8643	0.8665	0.8686	0.8718	0.8729	0.8749	0.8770	0.8790	0.8810	0.8836	0.8859
1.2	0.8859	0.8888	0.8917	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015	0.9031	0.9049
1.3	0.9049	0.9066	0.9083	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177	0.9192	0.9207
1.4	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319	0.9332	0.9345
1.5	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441	0.9452	0.9463
1.6	0.9463	0.9474	0.9485	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545	0.9555	0.9565
1.7	0.9565	0.9574	0.9583	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633	0.9641	0.9649
1.8	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706	0.9713	0.9720
1.9	0.9720	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767	0.9773	0.9778
2.0	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817	0.9821	0.9826
2.1	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9858	0.9862	0.9866
2.2	0.9866	0.9869	0.9873	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890	0.9893	0.9896
2.3	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916	0.9918	0.9920
2.4	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936	0.9938	0.9940
2.5	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952	0.9954	0.9955
2.6	0.9955	0.9956	0.9957	0.9959	0.9960	0.9962	0.9963	0.9964	0.9965	0.9966	0.9967
2.7	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974	0.9975	0.9976	0.9977
2.8	0.9977	0.9978	0.9979	0.9980	0.9981	0.9982	0.9983	0.9984	0.9985	0.9986	0.9987
2.9	0.9987	0.9988	0.9989	0.9990	0.9991	0.9992	0.9993	0.9994	0.9995	0.9996	0.9997
3.0	0.9997	0.9998	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Beispiel: $\Phi(1.52) \approx 0.9357$.

Zur Berechnung für $x < 0$ verwende man $\Phi(-x) = 1 - \Phi(x)$.

Normierung von S_{1000} und der zentrale Grenzwertsatz liefern wegen $\frac{100}{\sqrt{V(S_{1000})}} \approx 1.85$

$$\begin{aligned} P\{|S_{1000} - 3500| > 100\} &= 1 - P\{|S_{1000} - 3500| \leq 100\} \\ &= 1 - P\{-100 \leq S_{1000} - 3500 \leq 100\} \\ &= 1 - P\left\{-\frac{100}{\sqrt{V(S_{1000})}} \leq \frac{S_{1000} - E(S_{1000})}{\sqrt{V(S_{1000})}} \leq \frac{100}{\sqrt{V(S_{1000})}}\right\} \\ &= 1 - P\{-1.85 \leq S_{1000}^* \leq 1.85\} \\ &\approx 1 - [\Phi(1.85) - \Phi(-1.85)] = 2 - 2\Phi(1.85) \approx 0.064, \end{aligned}$$

d.h. die Wahrscheinlichkeit ist ungefähr 6 %.

3. Bei der Abfüllung von Zucker in 1000 g - Tüten schwanke die wirklich eingefüllte Menge mit der Streuung $\sigma = 5g$. Die durchschnittlich eingefüllte Menge λ sei am Abfüllgerät einstellbar. Wie hoch muß λ eingestellt werden, um mit 99,9 % Wahrscheinlichkeit eine eingefüllte Menge von 1000 g zu erzielen? Die (zufällige) Abfüllmenge werde durch die Zufallsvariable M beschrieben. Nach Voraussetzung ist

$$E(M) = \lambda \quad \text{und} \quad \sqrt{V(M)} = 5.$$

Gelten soll $P(1000 \leq M) = 0.999$.

Normierung von M und der "zentrale Grenzwertsatz" liefern

$$\begin{aligned} P(1000 \leq M) &= P\left\{\frac{1000 - \lambda}{5} \leq \frac{M - E(M)}{\sqrt{V(M)}}\right\} \\ &= P\left\{\frac{1000 - \lambda}{5} \leq M^*\right\} \approx 1 - \Phi\left(\frac{1000 - \lambda}{5}\right) \stackrel{!}{=} 0.999. \end{aligned}$$

Also muß

$$\Phi\left(\frac{1000 - \lambda}{5}\right) = 0.001 \quad \text{bzw.} \quad \Phi\left(\frac{\lambda - 1000}{5}\right) = 0.999$$

gelten. Wegen $\Phi(3.09) \approx 0.999$ ist somit

$$\frac{\lambda - 1000}{5} = 3.09 \quad \text{bzw.} \quad \lambda \approx 1015.45,$$

d.h. ungefähr 1016 g.

15.6. Statistische Methoden

Wegen der Kürze der Zeit kann von diesem wichtigen Gebiet nur ein ganz kleiner Ausschnitt gebracht werden. Zwei Fragestellungen sollen exemplarisch behandelt werden:

- (i) Schätzen von Parametern;
- (ii) Testen von Hypothesen.

Schätzen von Parametern. Eine Zufallsvariable X mit unbekannter Wahrscheinlichkeitsverteilung soll untersucht werden. Man möchte Parameter wie den Erwartungswert und die Varianz möglichst gut schätzen.

Das "Experiment", das auf die Zufallsvariable X führt, werde n-mal unabhängig wiederholt. Die Ergebnisse seien

(1) X_1, X_2, \dots, X_n .

Wir haben hier also n unabhängige identisch verteilte Zufallsvariable. Wir nennen (1) eine Stichprobe vom Umfang n.

Eine Schätzung T für den Parameter θ ist eine Funktion, die von X_1, \dots, X_n abhängt, also wieder eine Zufallsvariable. Die Schätzung T für θ heißt erwartungstreu, wenn

(2) $E(T) = \theta$

ist. Eine solche erwartungstreue Schätzung enthält also keinen systematischen Fehler. Man nennt die Schätzung konsistent, wenn T für $n \rightarrow \infty$ gegen θ konvergiert, d.h. genauer: Für jedes $\epsilon > 0$ gilt

(3) $P\{|T - \theta| > \epsilon\} \rightarrow 0$ für $n \rightarrow \infty$.

Die Schätzung des Erwartungswertes. Als Parameter θ betrachten wir speziell den Erwartungswert $E(X)$. Als Schätzung definieren wir das arithmetische Mittel

(4) $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n) = \frac{1}{n} \sum_{v=1}^n X_v$

Da die Zufallsvariablen X_1, \dots, X_n alle dieselbe Verteilung wie X haben, ist $E(X_1) = \dots = E(X_n) = E(X)$, also nach (4)

(5) $E(\bar{X}) = \frac{1}{n} \cdot n E(X) = E(X)$.

Daher wird durch (4) eine erwartungstreue Schätzung von $E(X)$ gegeben. Nach dem Gesetz der großen Zahl (S.15.4.4) ist

(6) $\frac{X_1 + \dots + X_n}{n} \rightarrow E(X)$ ($n \rightarrow \infty$).

Daraus kann man folgern, daß (4) eine konsistente Schätzung ist.

Die Schätzung der Varianz. Als Parameter wählen wir nun die Varianz $V(X)$. Wir versuchen

(7) $T = \frac{1}{n} \sum_{v=1}^n (X_v - \bar{X})^2$ mit $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$

als Schätzung. Mit $c = E(X)$ ist nach (5)

$$\begin{aligned}
 E(T) &= \frac{1}{n} E\left(\sum_{v=1}^n [(X_v - c) - (\bar{X} - c)]^2\right) = E([(X - c)^2] - E[(\bar{X} - c)^2]), \\
 &= \frac{1}{n} E\left(\sum_{v=1}^n (X_v - c)^2 - 2 \sum_{v=1}^n (X_v - c)(\bar{X} - c) + n(\bar{X} - c)^2\right) \\
 &= \frac{1}{n} E\left(\sum_{v=1}^n (X_v - c)^2 - n(\bar{X} - c)^2\right) \\
 &= \frac{1}{n} n E((X - c)^2) - E((\bar{X} - c)^2)
 \end{aligned}$$

(8)

da alle X_v dieselbe Verteilung wie X haben. Es ist weiter nach (4)

$$E((\bar{X}-c)^2) = E\left(\left[\frac{1}{n} \sum_{v=1}^n (X_v - c)\right]^2\right) = \frac{1}{n^2} nV(X) = \frac{1}{n} V(X);$$

die Varianzen addieren sich, da die Zufallsvariablen X_v unabhängig sind. Also ist nach (8)

$$(9) \quad E(T) = (1 - \frac{1}{n})V(X) = \frac{n-1}{n} V(X) < V(X).$$

Daher ist (7) keine erwartungstreue Schätzung für die Varianz. Aus (9) folgt aber, daß $E(\frac{n}{n-1}T) = V(X)$ ist. Somit ist

$$(10) \quad \frac{1}{n-1} \sum_{v=1}^n (X_v - \bar{X})^2$$

eine erwartungstreue Schätzung für $V(X)$. Man kann weiter zeigen, daß dies eine konsistente Schätzung ist.

Zusammenfassend haben wir also

$$\text{Schätzung für } E(X): \quad \bar{X} = \frac{1}{n} \sum_{v=1}^n X_v,$$

(11)

$$\text{Schätzung für } V(X): \quad \frac{1}{n-1} \sum_{v=1}^n (X_v - \bar{X})^2,$$

und diese Schätzungen sind erwartungstreu und konsistent.

Konfidenzintervalle. Wir wollen uns auf ein Beispiel beschränken. Sei X eine normalverteilte Zufallsvariable mit

$$\mu = E(X) \text{ unbekannt,} \quad \sigma = \sqrt{V(\bar{X})} \text{ bekannt.}$$

Wir betrachten eine Stichprobe des Umfangs n (d.h. n -fache unabhängige Wiederholung des Experiments), also X_1, \dots, X_n . Nach S. 15.6.2 ist \bar{X} eine erwartungstreue konsistente Schätzung für c . Aber wie gut ist diese Schätzung?

Man kann zeigen, daß \bar{X} wieder normalverteilt ist, und zwar ist

$$E(\bar{X}) = \mu, \quad \sqrt{V(\bar{X})} = \frac{\sigma}{\sqrt{n}}.$$

Also ist

$$(12) \quad Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

normalverteilt mit $E(Z) = 0, V(Z) = 1$. Somit ist für $0 < z < \infty$

$$\begin{aligned} P\{-z \leq Z \leq z\} &= P\{Z \leq z\} - P\{Z < -z\} \\ &= \Phi(z) - \Phi(-z) = 2\Phi(z) - 1. \end{aligned}$$

Nach (12) gilt daher

$$(13) \quad P\left\{\mu - z \frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu + z \frac{\sigma}{\sqrt{n}}\right\} = 2\Phi(z) - 1.$$

Das können wir schreiben als

$$(14) \quad P\left\{\mu \in \left[\bar{X} - z \frac{\sigma}{\sqrt{n}}, \bar{X} + z \frac{\sigma}{\sqrt{n}}\right]\right\} = 2\Phi(z) - 1,$$

in Worten: Mit Wahrscheinlichkeit $2\Phi(z) - 1$ liegt der unbekanntes Erwartungswert im Konfidenzintervall

$$(15) \quad \left[\bar{X} - z \frac{\sigma}{\sqrt{n}}, \bar{X} + z \frac{\sigma}{\sqrt{n}}\right].$$

Wenn wir $z = 3$ wählen, so wird $2\Phi(z) - 1 \approx 0.9972$, d.h. mit Wahrscheinlichkeit 99.72% liegt μ im Konfidenzintervall (15).

Das Testen von Hypothesen. Wir betrachten eine Zufallsvariable, deren Wahrscheinlichkeitsverteilung von einer Größe μ abhängt, die unbekannt ist, z.B. $\mu = E(X)$. Man testet eine statistische Hypothese, die sog. *Nullhypothese* die dann aufgrund eines Tests akzeptiert oder abgelehnt werden kann, und zwar auf einem gewissen Signifikanzniveau α , d.h. Irrtumswahrscheinlichkeit.

Als Beispiel betrachten wir wieder eine normalverteilte Zufallsvariable mit unbekanntem Erwartungswert μ und bekannter Streuung σ . Es soll die Nullhypothese

$$(16) \quad H_0: \mu = \mu_0$$

auf dem Signifikanzniveau α untersucht werden. Sei X_1, \dots, X_n eine Stichprobe vom Umfang n .

Wir wählen nun z , daß $2\Phi(z) - 1 = 1 - \alpha$, d.h. $\Phi(z) = 1 - \frac{\alpha}{2}$ ist. Nach (13) ist

$$\bar{X} \in \left[\mu - z \frac{\sigma}{\sqrt{n}}, \mu + z \frac{\sigma}{\sqrt{n}} \right]$$

mit Wahrscheinlichkeit $1 - \alpha$. Wenn dies nun erfüllt ist, so akzeptieren wir H_0 (zumindest provisorisch); wenn dies nicht erfüllt ist, so lehnen wir H_0 ab.

Beispiel. Sei X normalverteilt mit $\sigma = 10$. Man vermutet $E(X) = 800$, hat also als Nullhypothese

$$H_0: \mu = E(X) = 800.$$

Als Signifikanzniveau soll $\alpha = 0.05 \approx 5\%$ zugrundegelegt werden. Aus $\Phi(z) = 1 - \frac{\alpha}{2} = 0.975$ folgt nach Tabelle $z \approx 2$.

Unser Intervall ist

$$\left[\mu - \frac{2\sigma}{\sqrt{n}}, \mu + \frac{2\sigma}{\sqrt{n}} \right] = \left[800 - \frac{20}{\sqrt{n}}, 800 + \frac{20}{\sqrt{n}} \right].$$

Bei einer Stichprobe von $n = 100$ ergebe sich $\bar{X} = 805$. Unser Intervall ist [798, 802]. Da \bar{X} nicht darin liegt, müssen wir die Hypothese $\mu = 800$ ablehnen. Natürlich kann man trotzdem nicht ausschließen, daß sie doch zutrifft. Man beachte: Wenn wir $n = 10$ gewählt hätten, so wäre unser Intervall [793.68, 806.32]. Da 805 in diesem Intervall liegt, hätten wir H_0 akzeptieren müssen, sozusagen provisorisch. Das Ziel ist also eher, die Nullhypothese zu falsifizieren und nicht zu verifizieren.

Vergleich zweier Erwartungswerte. Gegeben seien zwei normalverteilte Zufallsvariable X und Y mit

$$E(X) = \mu, \quad E(Y) = \nu, \quad \sqrt{V(X)} = \sigma, \quad \sqrt{V(Y)} = \tau.$$

Dabei sollen μ und ν unbekannt, jedoch σ und τ bekannt sein. Die zu testende Hypothese sei

$$H_0: \mu \leq \nu, \quad \text{d.h.} \quad E(X) \leq E(Y).$$

So können z.B. X und Y die Leistungen zweier Motorentypen sein, die unter den verschiedenartigsten Bedingungen geprüft werden.

Gegeben seien Stichproben X_1, \dots, X_m bzw. Y_1, \dots, Y_n . Dann sind die Stichprobenmittel \bar{X} und \bar{Y} unabhängige Zufallsvariable. Also ist

$$V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(-\bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \frac{\sigma^2}{m} + \frac{\tau^2}{n}.$$

Weiter ist $E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu - \nu$. Sei

$$(17) \quad Z = \frac{(\bar{X} - \bar{Y}) - (\mu - \nu)}{\sqrt{V(\bar{X} - \bar{Y})}} = \frac{(\bar{X} - \bar{Y}) - (\mu - \nu)}{\sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}}$$

Man kann zeigen, daß Z wieder normalverteilt ist; aus dem zentralen Grenzwertsatz (vergleiche S. 15.5.1) folgt das jedenfalls für großes m und n . Weiter ist

$$E(Z) = 0, \quad V(Z) = 1.$$

Also ist

$$P(Z \leq z) = \Phi(z),$$

d.h. mit Wahrscheinlichkeit $\Phi(z)$ gilt $Z \leq z$, also gemäß (17)

$$(18) \quad \bar{X} - \bar{Y} \leq \mu - \nu + z \sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}} \leq z \sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}.$$

Bei gegebenem Signifikanzniveau α wählt man z so, daß $\Phi(z) = 1 - \alpha$ ist. Nun soll $\mu \leq \nu$ sein. Also gilt

$$(19) \quad \bar{X} - \bar{Y} \leq z \sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}.$$

Wenn (19) erfüllt ist, so akzeptiert man $H_0: \mu \leq \nu$, sonst lehnt man H_0 ab (auf dem Signifikanzniveau α).

Beispiel. Sei $\sigma = 20$, $\tau = 10$, $m = 200$, $n = 100$. Dann ist

$$\sqrt{\dots} = \sqrt{\frac{400 \cdot 100}{200} + \frac{100}{100}} = \sqrt{3}. \text{ Wir betrachten zwei verschiedene Signifikanzniveaus } \alpha_1 = 0.05 \text{ und } \alpha_2 = 0.01. \text{ Dazu gehören } z_1 = 1.65 \text{ und } z_2 = 2.33. \text{ Zu untersuchen ist also}$$

$$(20) \quad \bar{X} - \bar{Y} \leq 2.86 \quad \text{bzw.} \quad \bar{X} - \bar{Y} \leq 4.04.$$

Die Mittelwerte der Meßwerte seien $\bar{X} = 805$, $\bar{Y} = 802$. Dann ist die erste Bedingung (20) erfüllt, aber nicht die zweite. Daher haben wir: Auf dem Signifikanzniveau $\alpha = 0.05$ muß man $H_0: \mu \leq \nu$ ablehnen, auf dem Signifikanzniveau $\alpha = 0.01$ muß man $H_0: \mu \leq \nu$ noch provisorisch akzeptieren.