TU Ilmenau
Institut für Mathematik
Prof. Dr. S. Vogel

Höhere Mathematik III für Wirtschaftsinformatiker

1 Funktionen von mehreren Variablen

1.1 Grenzwerte und Stetigkeit

Betrachtet werden Funktionen $f : D_f \subset R^n \to R^1$.

Beispiele:
a) $n = 2, z = f(x, y)$
In diesem Fall ist eine Veranschaulichung im R^3 (z.B. 3D-Plots mittels Computer) oder mit Hilfe von ebenen Schnitten (insbesondere Höhenlinien) möglich.

1.) $z = x^2 + y^2$

- α) Schnitte mit den Ebenen $z = c,\ c \in R$
  (Höhenlinien, Niveaulinien)
  liefern $x^2 + y^2 = c \quad (c \geq 0)$, also Kreise mit Radius $\sqrt{c}$.
- β) Der Schnitt z. B. mit der Ebene $x = 0$ ergibt $z = y^2$, der Schnitt mit $y = 0$ ergibt $z = x^2$.

Die Funktion beschreibt ein Rotationsparaboloid.

2.) $z = \sqrt{x^2 + y^2}$ beschreibt einen Kegel.

3.) $z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$ stellt ein elliptisches Paraboloid dar.

Die Höhenlinien $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = c \quad (c \geq 0)$ bzw. $\frac{x^2}{(a\sqrt{c})^2} + \frac{y^2}{(b\sqrt{c})^2} = 1$ sind Ellipsen.

4.) $z = x \cdot y$
Die Betrachtung der Höhenlinien zeigt, dass im Nullpunkt ein sogenannter Sattelpunkt vorliegt.

5.) Als Cobb-Douglas-Funktion (im R^2) bezeichnet man die Funktion $z = cx^{\alpha_1} \cdot y^{d_2}$,
wobei z den Output (Wertschöpfung), x den Kapitaleinsatz und y den Arbeitseinsatz bezeichnen.

b) Nun sei n beliebig und $\vec{x} = (x_1, \ldots, x_n)$. Wichtige Funktionstypen sind

1.) lineare Funktionen $y = f(\vec{x}) = f(x_1, \ldots, x_n) = a_o + a_1x_1 + \ldots, +a_nx_n$ und

2.) quadratische Funktionen

$$\begin{aligned} y = f(\vec{x}) &= f(x_1, \ldots, x_n) \\ &= a_o + a_1x_1 + \ldots a_nx_n + a_{11}x_1^2 + \cdots + a_{nn}x_n^2 \\ &\quad +a_{12}x_1x_2 + \cdots + a_{n-1,n}x_{n-1}x_n. \end{aligned}$$

Als Abstandsbegriff im R^n wird in der Regel die "**Euklidische Norm**" $||.||$ verwendet: Es seien $\vec{x} = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n)$. Dann

$$||\vec{x} - \vec{y}|| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Bemerkung: Es gibt auch andere Abstandsbegriffe, z. B. die sogenannte Maximumnorm $||.||_M$ mit
$||\vec{x} - \vec{y}||_M := \max\limits_{1 \leq i \leq n} |x_i - y_i|$.

Analog zum Eindimensionalen definiert man eine **ϵ-Umgebung** von $\vec{x}$:
$U_\epsilon\{\vec{x}\} := \{\vec{y} \mid ||\vec{x} - \vec{y}|| < \epsilon\}$.

Im Weiteren sei $\vec{x}_k = (x_1^{(k)}, \ldots, x_n^{(k)})$.

<u>Definition:</u> Eine Folge $(\vec{x}_k)_{k \in N}$ heißt **konvergent** gegen $\vec{x}_o$, falls
$\lim\limits_{k \to \infty} ||\vec{x}_k - \vec{x}_o|| = 0$ erfüllt ist.
(Schreibweise: $\lim\limits_{k \to \infty} \vec{x}_k = \vec{x}_o$)

$$\begin{aligned} \text{Es gilt: } \lim_{k \to \infty} \vec{x}_k = \vec{x}_o &\iff \lim_{k \to \infty} x_i^{(k)} = x_i^{(o)}\ \forall i \in \{1, \ldots, n\} \\ &\iff \forall \epsilon > 0\ \exists k_o\ \forall k \geq k_o :\ \vec{x}_k \in U_\epsilon\{\vec{x}_o\}. \end{aligned}$$

<u>Definition:</u> Die Zahl η heißt **Grenzwert** von f an der Stelle $\vec{x}_o \in D_f$, wenn $\lim\limits_{k \to \infty} f(\vec{x}_k) = \eta$ für alle Folgen $(\vec{x}_k)_{k \in N}$ mit $\vec{x}_k \in D_f \setminus \{\vec{x}_o\}$ und $\lim\limits_{k \to \infty} \vec{x}_k = \vec{x}_o$ gilt.

Schreibweise: $\lim\limits_{\vec{x} \to \vec{x}_o} f(\vec{x}) = \eta$ oder $\lim\limits_{x_1 \to x_1^{(o)}, \ldots, x_n \to x_n^{(o)}} f(x_1, \ldots, x_n) = \eta$

Definition: f heißt **stetig in $\vec{x}_o$** , wenn $\lim\limits_{\vec{x}\to\vec{x}_o} f(\vec{x}) = f(\vec{x}_o)$ gilt. f heißt **stetig**, wenn f in allen Punkten $\vec{x}_o \in D_f$ stetig ist.

Für zwei im Punkt $\vec{x}_o$ stetige Funktionen sind auch die Verknüpfungen $f+g$, αf $(\alpha \in R)$ und $f \cdot g$ stetig in $\vec{x}_o$. Die Funktion $\frac{f}{g}$ ist stetig in $\vec{x}_o$, falls zusätzlich $g(\vec{x}_o) \neq 0$ gilt.

Beispiele:

1.) Die Funktion f mit $f(x,y) = \dfrac{x^2 \cdot y^2}{x^2+y^2}$ ist stetig in allen Punkten $(x,y) \neq (0,0)$. Weiterhin gilt wegen $0 \leq \dfrac{x^2 \cdot y^2}{x^2+y^2} \leq y^2$ die Beziehung $\lim\limits_{x\to 0, y\to 0} \dfrac{x^2 \cdot y^2}{x^2+y^2} = 0$. Also kann die Funktion f im Nullpunkt "stetig fortgesetzt werden" zu der stetigen Funktion
$\tilde{f}$ mit $\tilde{f}(x,y) = \begin{cases} f(x,y) & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$.

2.) Die Funktion f mit $f(x,y) = \dfrac{x \cdot y}{x^2+y^2}$ ist stetig in allen Punkten $(x,y) \neq (0,0)$. Zur Untersuchung des Verhaltens in $(0,0)$ kann man verschiedene Folgen betrachten, die gegen $(0,0)$ konvergieren.

α) Für die Folge $(\vec{x}_k)_{k\in N}$ mit $\vec{x}_k = (x^{(k)}, y^{(k)}) = \left(\dfrac{1}{k}, 0\right)$ ergibt sich $\lim\limits_{k\to\infty} f(\vec{x}_k) = 0$.

β) Für die Folge $(\vec{x}_k)_{k\in N}$ mit $\vec{x}_k = \left(\dfrac{1}{k}, \dfrac{1}{k}\right)$ ergibt sich
$\lim\limits_{k\to\infty} f(\vec{x}_k) = \lim\limits_{k\to\infty} \dfrac{\frac{1}{k^2}}{\frac{2}{k^2}} = \dfrac{1}{2}$.
Also existiert der Limes $\lim\limits_{x\to 0, y\to 0} f(x,y)$ nicht, und die Funktion f kann nicht "stetig fortgesetzt werden".

1.2 Partielle Ableitungen

Es seien zunächst $n = 2$ und $z = f(x,y)$.
Für ein fest gewähltes y_o wird die Funktion $\varphi^{(y_o)} | \, R^1 \to R^1$ mit $\varphi^{(y_o)}(x) := f(x, y_o)$ betrachtet.
(Der Graph dieser Funktion kann als Schnittkurve der Fläche $z = f(x,y)$ mit der Ebene $y = y_o$ aufgefasst werden.)

Definition: Ist $\varphi^{(y_o)}$ in x_o differenzierbar (d. h., existiert
$\lim\limits_{h\to 0} \dfrac{\varphi^{(y_o)}(x_o+h) - \varphi^{(y_o)}(x_o)}{h} = \lim\limits_{h\to o} \dfrac{f(x_o+h, y_o) - f(x_o, y_o)}{h}$),

so heißt $(\varphi^{(y_o)})'(x_o)$ die **partielle Ableitung von f nach x im Punkt (x_o, y_o)**.

Anstelle von $\varphi^{(y_o)}(x_o)$ schreibt man $f_x(x_o, y_o)$ oder auch $\frac{\partial f}{\partial x}(x_o, y_o)$.

$f_x(x_o, y_o)$ beschreibt den Anstieg der Tangente an $\varphi^{(y_o)}$ im Punkt x_o oder, mit anderen Worten, den Anstieg der Tangente an f in (x_o, y_o) in Richtung der x-Achse.

Es sei nun $D_x \subset R^2$ die Menge aller Punkte (x_o, g_o), in denen f partiell nach x differenzierbar ist.
Die partielle Ableitung nach x kann dann ihrerseits als Funktion aufgefasst werden: $f_x|\ D_x \to R^1$.

Analog zur obigen Herleitung werden durch Vertauschung der Rolle von x und y die Ableitung $f_y(x_o, y_o)$ von f nach y im Punkt (x_o, y_o)
bzw. die partielle Ableitung $f_y|\ D_y \to R^1$ definiert.

Beispiel: Es sei $f(x, y) = x^2y^2 + x$. Dann ergeben sich $f_x(x, y) = 2xy^2 + 1$ und $f_y(x, y) = 2x^2y$ sowie (zum Beispiel) $f_x(1, 2) = 9,\ f_y(1, 2) = 4$.

Analog zum zweidimensionalen Fall wird für eine Funktion
$f|\ D_f \subset R^n \to R^1$ die **partielle Ableitung f_{x_i}** nach x_i in folgender Weise erklärt:
$f_{x_i}(\vec{x}_o) := \lim\limits_{h\to o} \dfrac{f(x_1^{(o)}, \ldots, x_{i-1}^{(o)}, x_i^{(o)} + h, x_{i+1}^{(o)}, \ldots, x_n^{(o)}) - f(x_1^{(o)}, \ldots, x_n^{(o)})}{h}$.
(Dabei seien $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{x}_o = (x_1^{(o)}, \ldots, x_n^{(o)})$.
Sind alle $f_{x_i}, i = 1 \ldots n$, stetige Funktionen, so heißt f **stetig partiell differenzierbar**.

Satz: Wenn eine $\epsilon-$Umgebung $U_\epsilon\{\vec{x}_o\}$ derart existiert, daß alle f_{x_i},
$i = 1, \ldots, n$, auf $U_\epsilon\{\vec{x}_o\}$ existieren und beschränkt sind, dann ist f stetig in $\vec{x}_o$.

Höhere partielle Ableitungen

Definition: Es sei $f|D_f \subset R^n \to R^1$ partiell differenzierbar nach allen Variablen $x_1, \ldots, x_n$. Sind die Ableitungen $f_{x_i}, i = 1, \ldots, n$, ihrerseits partiell differenzierbar, so heißt f **zweimal partiell** differenzierbar.

$\dfrac{\partial f_{x_i}}{\partial f_{x_j}}(\vec{x}_o)$ wird mit $\boldsymbol{f_{x_i,x_j}(x_1^{(o)},\ldots,x_n^{(o)})}$ bezeichnet, d.h., nach der näher an f stehenden Variablen wird zuerst abgeleitet. Üblich ist auch die Schreibweise $\dfrac{\partial^2 f}{\partial x_j \partial x_i} f(x_1^{(o)},\ldots,x_n^{(o)})$ für $f_{x_i,x_j}(x_1^{(o)},\ldots,x_n^{(o)})$.

Satz: (Vertauschungsregel) (für R^2, gilt im R^n analog)
Wenn f_x und f_y sowie f_{xy} auf $U_\epsilon\{(x_o,y_o)\}$ existieren und f_{xy} in (x_o,y_o) stetig ist, gilt $f_{xy}(x_o,y_o) = f_{yx}(x_o,y_o)$.

1.3 Ableitung mittelbarer Funktionen (verallgemeinerte Kettenregel)

Gegeben seien eine Funktion $f|D_f \subset R^n \to R^1$ sowie eine Parameterdarstellung für einen "Weg" im R^n in der Gestalt
$x_1 = g_1(t),\ldots,x_n = g_n(t),\ t \in D_g$.
Dabei wird vorausgesetzt, dass $(g_1(t),\ldots,g_n(t))$ für alle $t \in D_g$ zu D_f gehört. Durch
$F(t) := f(g_1(t), g_2(t),\ldots,g_n(t))$
ist dann eine Funktion $F|\ D_g \to R^1$ erklärt. Gesucht ist die Ableitung F'.

Satz (verallgemeinerte Kettenregel)
Die Funktion $f|D_f \subset R^n \to R^1$ sei stetig differenzierbar nach allen Variablen, und die Funktionen $g_i|D_g \to R^1,\ i = 1,\ldots,n$, seien differenzierbar. Dann ist die Funktion F differenzierbar, und es gilt
$F'(t_o) = f_{x_1}(g_1(t_o), g_2(t_o),\ldots,g_n(t_o)){g_1}'(t_o) + \ldots + f_{x_n}(g_1(t_o),\ldots,g_n(t_o)){g_n}'(t_o)$.

Beispiel: Es seien $f(x,y) = x^2y + y$, $x(t) = g_1(t) = t^2,\ y(t) = g_2(t) = 2t+1$. Dann kann die Ableitung von F einerseits nach Einsetzen der Parameterdarstellung gebildet werden:
$F'(t) = (\ (t^2)^2(2t+1) + (2t+1)\ )' = ... = 10t^4 + 4t^3 + 2$.

Unter Ausnutzung des obigen Satzes erhält man andererseits

$$\begin{aligned} F'(t) = 2t^2\cdot(2t+1)\cdot 2t + ((t^2)^2+1)\cdot 2 &= 4t^3(2t+1) + 2t^4 + 2 \\ &= 10t^4 + 4t^3 + 2. \end{aligned}$$

Es sei nun insbesondere $g_1(t) = x_1^{(o)} + ta_1,\ \ldots,\ g_n(t) = x_n^{(o)} + ta_n$,
wobei $\sqrt{\sum a^2} = 1$ angenommen wird. Weiterhin werden $\vec{a} = (a_1,\ldots,a_n)^T$, $\vec{x}_o = (x_1^{(o)},\ldots,x_n^{(o)})$ sowie $t_o = 0$ gesetzt.

Nach der verallgemeinerten Kettenregel erhält man

$$F'(0) = f_{x_1}(x_1^{(o)},\ldots,x_n^{(o)})\cdot a_1 + \ldots + f_{x_n}(x_1^{(o)},\ldots,x_n^{(o)})a_n.$$

Im R^2 gibt $F'(0)$ den Anstieg der Tangenten im Punkt $\vec{x}_o$ an die Fläche $z = f(x_1, x_2)$ in Richtung $\vec{a}$ an.

Definition: Die Zahl
$$\frac{\partial f}{\partial \vec{a}}(x_1^{(o)}, \ldots, x_n^{(o)}) \ := \ \sum_{i=1}^{n} f_{x_i}(x_1^{(o)}, \ldots, x_n^{(o)}) a_i$$
heißt **Richtungsableitung** von f an der Stelle $\vec{x}_o$ in Richtung $\vec{a}$.

Definition: Der Vektor $\nabla f(\vec{x}_o) = \nabla f(x_1^{(o)}, \ldots, x_n^{(o)}) := (f_{x_1}(\vec{x}_o), \ldots, f_{x_n}(\vec{x}_o))^T$ heißt **Gradient** von f im Punkt $\vec{x}_o$.

(Für den Gradienten ist auch die Schreibweise $grad\ f(\vec{x}_o)$ üblich.)

Für die Richtungsableitung gilt damit $\dfrac{\partial f}{\partial \vec{a}}(\vec{x}_o) = \vec{a}^T \nabla f(\vec{x}_o)$.

Das Skalarprodukt $\vec{a}^T \nabla f(\vec{x}_o)$ wird maximal für $\vec{a} = \frac{\nabla(\vec{x}_o)}{||\nabla(\vec{x}_o)||}$, d. h., $\nabla f(\vec{x}_o)$ gibt die **Richtung des steilsten Anstiegs** und entsprechend $-\nabla f(\vec{x}_o)$ die Richtung des steilsten Abstiegs einer Funktion an.

Beispiel: Es seien $z = x^2 + y^2, \quad \vec{a}_1 = \dfrac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\vec{a}_2 = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Dann erhält man $\nabla f(x, y) = \begin{pmatrix} 2\text{x} \\ 2\text{y} \end{pmatrix}$, $\nabla f(1, 1) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ sowie

$$\frac{\partial f}{\partial \vec{a}_1}(1,1) = \frac{1}{\sqrt{5}}(1\ 2) \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \frac{6}{\sqrt{5}}, \quad \frac{\partial f}{\partial \vec{a}_2}(1,1) = \frac{1}{\sqrt{2}}(1\ 1) \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\sqrt{2}.$$

Die durch den Gradienten gegebene Richtung des steilsten Anstiegs stimmt hier mit der durch $\vec{a}_2$ gegebenen Richtung überein, d.h., $\dfrac{\partial f}{\partial \vec{a}_2}(1,1)$ ist der steilste Anstieg von f im Punkt $(1, 1)$.

1.4 Das totale Differential

Betrachtet wird eine Funktion $f|D_f \subset R^2 \to R^1$, die in einer $\epsilon-$Umgebung von (x_o, y_o) stetig partiell differenzierbar ist. Der Wert $f(x_o, y_o)$ sei bekannt; gesucht ist (näherungsweise) der Wert $f(x_o + h, y_o + k)$ für kleine Werte h und k.

Bemerkung: Anstelle von h und k werden in der Regel die Schreibweisen dx und dy verwendet.

Es gilt
$$f(x_o + dx, y_o + dy) - f(x_o, y_o) = \underbrace{f_x(x_o, y_o)dx + f_y(x_o, y_o)dy}_{=:df(x_o,y_o)} + \eta(dx, dy).$$

Der Wert $\eta(dx, dy)$ gibt gerade die Differenz zwischen dem wahren Funktionswert an der Stelle $(x_o + dx, y + dy)$ und dem Funktionswert der Tangentialebene $f(x_o, y_o) + f_x(x_o, y_o)dx + f_y(x_o, y_o)dy$ an dieser Stelle an.

Gilt für η die Beziehung

$$\lim_{(dx,dy)\to(0,0)} \frac{|\eta(dx, dy)|}{||(dx, dy)||} = 0, \qquad (*)$$

so heißt f im Punkt (x_o, y_o) **total (oder vollständig) differenzierbar**, und
$df(x_o, y_o) := f_x(x_o, y_o)dx + f_y(x_o, y_o)dy$
heißt **totales (oder vollständiges) Differential** von f an der Stelle (x_o, y_o).

Bemerkungen:

1. Die vorausgesetzte Stetigkeit der partiellen Ableitungen sichert die Gültigkeit von (*).

2. Eine in (x_o, y_o) total differenzierbare Funktion f kann dort lokal durch die Tangentialebene angenähert werden.

Für eine Funktion $f|D_f \subset R^n \to R^1$ wird das totale Differential in analoger Weise erklärt:
$df(x_1^{(o)}, \ldots, x_n^{(o)}) = f_{x_1}(x_1^{(o)}, \ldots, x_n^{(o)})dx_1 + \cdots + f_{x_n}(x_1^{(o)}, \ldots, x_n^{(o)})dx_n.$

Anwendungsbeispiele für das totale Differential:

1.) Fehlerrechnung: Zwei Einflussgrößen können jeweils nur bis auf einen Fehler dx bzw. dy genau angegeben werden. Mit welchem (näherungsweisen) Fehler ist bei der Angabe von $z = f(x_o, y_o)$ zu rechnen?

Lösung: Es seien $x = x_o \pm \triangle x,\ y = y_o \pm \triangle y,\ z = f(x_o, y_o) \pm \triangle f$.
Man verwendet den Näherungswert
$|df(x_o, y_o)|$ für die wahre Abweichung $\triangle f$, wobei $\triangle x$ mit dx und $\triangle y$ mit dy identifiziert werden.

2.) Gegeben sei eine Outputfunktion der Gestalt
$f(x, y) = cx^{\alpha_1}y^{\alpha_2}, \quad c > 0,\ \alpha_1 > 0,\ \alpha_2 > 0.$

a) Wie verändert sich der Output, wenn sich die Produktionsfaktoren x und y um dx bzw. dy ändern (Ausgangspunkt (x_o, y_o))?
Lösung: Die Änderung ergibt sich näherungsweise zu
$df(x_o, y_o) = c\alpha_1 x_o^{\alpha_1 - 1} y_o^{\alpha_2} dx + c\alpha_2 x_o^{\alpha_1} y_o^{\alpha_2 - 1} dy.$

b) Der erste Produktionsfaktor ändert sich um dx. Um wieviel (dy) muss man den zweiten Produktionsfaktor ändern, damit das Ausgangsniveau des Outputs erhalten bleibt?

Lösung: Das Ausgangsniveau sei $f(x_o, y_o)$. Der Output soll sich nicht ändern, also muss $df(x_o, y_o) = 0$ gelten. Daraus folgt
$0 = f_x(x_o, y_o)dx + f_y(x_o, y_o)dy$ und schließlich
$dy = -\frac{f_x(x_o, y_o)}{f_y(x_o, y_o)}dx$.

Der Quotient $-\frac{f_x(x_o, y_o)}{f_y(x_o, y_o)}$ heißt **Substitutionsrate** oder auch Grenzwert der Faktorsubstitution.

1.5 Implizite Funktionen

Gegeben sei die Gleichung $F(x, y) = 0$. Wann beschreibt die Lösungsmenge dieser Gleichung (lokal) eine Funktion? (Erinnerung: Funktionen wurden mit ihrem Graphen identifiziert.)

Beispiele:

1.) Betrachtet wird die Gleichung $x^2 + y^2 - 1 = 0$.

- α) $\{(x, y) \in R^2 |\ x^2 + y^2 - 1 = 0\}$ stellt nicht den Graphen einer Funktion, sondern nur eine Relation dar.
- β) $\{(x, y) \in R^1 \times [0, \infty)|\ x^2 + y^2 - 1 = 0\}$ stellt eine Funktion f mit der expliziten Form $y = \sqrt{1 - x^2},\ |x| \leq 1$, dar.
- γ) $\{(x, y) \in [0, \infty) \times R^1|\ x^2 + y^2 - 1 = 0\}$ stellt eine Funktion g mit der expliziten Form $x = g(y) = \sqrt{1 - y^2},\ |y| \leq 1$, dar.

2.) Gegeben sei die Relation $x^3 + y^3 - 3xy = 0$ (Kartesisches Blatt).

In jeder Rechteckumgebung des Nullpunktes ist die Relation $x^3 + y^3 - 3xy = 0$ weder als Funktion von x noch als Funktion von y darstellbar.

Betrachtet wird eine “Rechteckumgebung“
$R_{\delta_1\delta_2}(x_o, y_o) := \{(x, y)|(x - x_o) < \delta_1,\ |y - y_o| < \delta_2\}$ eines Punktes (x_o, y_o).
Eine Funktion $y = f(x)$ bzw. $x = g(y)$) mit $(x, y) \in R_{\delta_1\delta_2}(x_o, y_o)$, die in der Form $F(x, y) = F(x, f(x)) = 0$ (bzw. $F(x, y) = F(g(y), y) = 0$) vorliegt, heißt **implizit gegeben** oder **implizite Funktion**.

Satz: Die Funktion F sei auf einer Rechteckumgebung $R_{\delta_1\delta_2}(x_o, y_o)$ definiert, und die folgenden Bedingungen seien erfüllt:

1.) $F(x_o, y_o) = 0$,

2.) F_y existiert auf $R_{\delta_1\delta_2}(x_o, y_o)$ und ist dort stetig, und es gilt $F_y(x_o, y_o) \neq 0$.

Dann existieren eine Rechteckumgebung $R_{\alpha_1\alpha_2}(x_o, y_o)$ und eine Funktion $y = f(x)$ derart, dass für alle $x \in (x_o - \alpha_1, x_o + \alpha_1)$ die Beziehungen $F(x, f(x)) = 0$ und $|f(x) - y_o| < \alpha_2$ gelten.
Sind darüber hinaus die Ableitungen F_x und F_y stetig auf $R_{\delta_1\delta_2}(x_o, y_o)$, so ist f in allen Punkten $x \in (x_o - \alpha_1, x_o + \alpha_1)$ differenzierbar, und es gilt

$$f'(x) = \frac{dy}{dx} = -\frac{F_x(x, f(x))}{F_y(x, f(x))}. \qquad (*)$$

Herleitung für (*): Aus $F(x, f(x)) = 0$ folgt
$F_x(x, f(x)) + F_y(x, f(x)) \cdot f'(x) = 0$.

Bemerkungen:

1.) Eine analoge Aussage gilt im Fall $F_x(x_o, y_o) \neq 0$:
Ist Bedingung 1.) erfüllt und gilt Bedingung 2.) für F_x, dann existieren eine Rechteckumgebung $R_{\alpha_1\alpha_2}(x_o, y_o)$ und eine Funktion $x = g(y)$ derart, dass für $y \in (y_o - \alpha_2, y_o + \alpha_2)$ die Beziehungen $F(g(y), y) = 0$ und $|g(y) - x_o| < \alpha_1$ gelten. Sind darüber hinaus die Ableitungen F_x und F_y stetig auf $R_{\delta_1\delta_2}(x_o, y_o)$, so ist g in allen Punkten $y \in (y_o - \alpha_2, y_o + \alpha_2)$ differenzierbar, und es gilt
$g'(y) = \frac{dx}{dy} = -\frac{F_y(g(y), y)}{F_x(g(y), y)}$.

2.) Sind die Voraussetzungen des Satzes erfüllt, muss es nicht notwendig möglich sein, die Gleichung $F(x, y) = 0$ rechnerisch (in geschlossener Form) nach einer Variablen aufzulösen.

Beispiel: Betrachtet wird das Kartesische Blatt in den Punkten
$(x_o, y_o) = (0, 0)$ und $(x_1, y_1) = (\sqrt[3]{2}, \sqrt[3]{4})$. Es ergibt sich
$F_x(x, y) = 3x^2 - 3y = 3(x^2 - y),\ F_x(0, 0) = 0,\ F_x(x_1, x_2) = 3(\sqrt[3]{4} - \sqrt[3]{4}) = 0$ und
$F_y(x, y) = 3y^2 - 3x = 3(y^2 - x),\ F_y(0, 0) = 0,\ F_y(x_1, x_2) = 3(\sqrt[3]{16} - \sqrt[3]{2}) > 0$.
Im Punkt (x_o, y_o) sind die Voraussetzungen des Satzes nicht erfüllt. Im Punkt (x_1, y_1) ist ein Auflösung in der Gestalt $y = f(x)$ möglich. Wegen $f'(x_1) = -\frac{F_x(x_1, y_1)}{F_y(x_1, y_1)} = 0$ besitzt die durch $x^3 + y^3 - 3xy = 0$ in einer Umgebung von (x_1, y_1) gegebene Kurve in (x_1, y_1) eine waagerechte Tangente.

1.6 Extremwerte für Funktionen von mehreren Veränderlichen

1.6.1 Definitionen

Betrachtet wird eine Funktion $f|D_f \subset R^n \to R^1$.

Definition: Ein Punkt $\vec{x}_o \in D_f$ heißt **lokale Maximalstelle** (bzw. **lokale Minimalstelle**) und $f(\vec{x}_o)$ **lokales Maximum** (bzw. **lokales Minimum**) von f, falls eine ϵ-Umgebung $U_\epsilon\{\vec{x}_o\}$ derart existiert, dass
$f(\vec{x}_o) \geq f(\vec{x}) \ \ \forall\, \vec{x} \in U_\epsilon\{\vec{x}_o\}$
$(f(\vec{x}_o) \leq f(\vec{x}) \ \ \forall\, \vec{x} \in U_\epsilon\{\vec{x}_o\})$
gilt.

Falls $f(\vec{x}_o) > f(\vec{x}) \ \ \forall\, \vec{x} \in U_\epsilon\{\vec{x}_o\} \setminus \{\vec{x}_o\}$ erfüllt ist, nennt man $f(\vec{x}_o)$ auch isoliertes (oder eigentliches) lokales Maximum.

Definition: Falls $f(\vec{x}_o) \geq f(\vec{x}) \ \ \forall\, \vec{x} \in M$ gilt, heißt $f(\vec{x}_o)$ **globales Maximum** von f über M. Entsprechend heißt $f(\vec{x}_o)$ **globales Minimum** von f über M, falls $f(\vec{x}_o) \leq f(\vec{x}) \ \ \forall\, \vec{x} \in M$ erfüllt ist.

Beispiele:

1.) $f(x, y) = x^2 + y^2$ besitzt auf $M = R^2$ ein globales Minimum im Punkt $(0, 0)$, aber kein lokales oder globales Maximum.

2.) $f(x, y) = x^2 + y^2$ besitzt auf $[-2, +2] \times [-2, +2]$ ein globales Minimum im Punkt $(0, 0)$ und globale Maxima in den Punkten $(+2, +2), (+2, -2), (-2, +2), (-2, -2)$.

Definition:

a) Eine Menge $M \subset R^n$ heißt **beschränkt**, wenn es eine Konstante $K \in R$ mit $||\vec{x}|| \leq K \ \ \forall \vec{x} \in M$ gibt.

b) Eine Menge $M \subset R^n$ heißt **abgeschlossen**, wenn für alle Folgen $(\vec{x}_n)_{n \in N}$ mit $\vec{x}_n \in M \quad \forall n \in N$ und $\lim_{n \to \infty} \vec{x}_n = \vec{x}_o$ die Beziehung $\vec{x}_o \in M$ gilt.

Beispiele: R^n ist abgeschlossen, Q ist nicht abgeschlossen.
Die Menge $[a, b] \ \times \ [c, d]$ ist abgeschlossen, $[a, b) \ \times \ [c, d]$ ist nicht abgeschlossen.

Bemerkung: Sind Teilmengen des R^n sowohl abgeschlossen als auch beschränkt, werden sie auch **kompakte Mengen** genannt.

Satz: Stetige Funktionen $f | D_f \subset R^n \to R^1$ nehmen auf einer beschränkten und abgeschlossenen Menge ihr globales Maximum und ihr globales Minimum an.

1.6.2 Lokale Extrema

Satz: $f|D_f \subset \mathbb{R}^n \to \mathbb{R}^1$ sei auf einer ϵ-Umgebung $U_\epsilon\{\vec{x}_o\}$ erklärt, und es existieren alle partiellen Ableitungen. Wenn f in $\vec{x}_o$ eine lokale Maximalstelle oder Minimalstelle besitzt, dann gilt
$f_{x_i}(\vec{x}_o) = 0 \; \forall i \in \{1, \ldots, n\}$.

Beweis: Für $i = 1, \ldots, n$ sei $g_i(t) := f(x_1^{(o)}, \ldots, x_{i-1}^{(o)}, t, x_{i+1}^{(o)}, \ldots, x_n^{(o)})$.
$g_i, i \in \{1 \ldots n\}$, ist jeweils eine Funktion einer Veränderlichen, die in $t = x_i^{(o)}$ differenzierbar ist und dort ein Extremum besitzt. Also gilt
$g_i'(x_i^{(o)}) = 0 = f_{x_i}(\vec{x}_o)$.

Punkte $(x_1^{(o)}, \ldots, x_n^{(o)})$ mit $f_{x_i}(x_1^{(o)}, \ldots, x_n^{(o)}) = 0 \quad \forall i \in \{1, \ldots, n\}$ heißen **stationäre Punkte**. Sie sind "extremwertverdächtig".

Beispiele:

1.) Betrachtet wird die Funktion $f(x, y) = x^2y - 2xy + \frac{3}{4}e^y$. Die partiellen Ableitungen haben die Gestalt
$f_x(x, y) = 2xy - 2y, \quad f_y(x, y) = x^2 - 2x + \frac{3}{4}e^y$.

Die notwendigen Bedingungen für das Vorliegen eines Extremums lauten damit:

$$f_x(x, y) = 2xy - 2y = 0 \qquad (1)$$
$$f_y(x, y) = x^2 - 2x + \tfrac{3}{4}e^y = 0 \qquad (2).$$

Aus (1) folgt $x = 1 \vee y = 0$. Für die Gleichung (2) ergeben sich folglich zwei Fälle:

a) Für $x = 1$ nimmt (2) die Gestalt $\frac{3}{4}e^y = 1$ an. Daraus erhält man $y = \ln \frac{4}{3}$.

b) Für $y = 0$ nimmt (2) die Gestalt $x^2 - 2x + \frac{3}{4} = 0$ an. Diese Gleichung besitzt die Lösungen $x_{1,2} = 1 \pm \sqrt{1 - \frac{3}{4}}$.

Damit sind alle Lösungen des Gleichungssystems (1)+(2) gefunden. Es gibt 3 stationäre Punkte: $P_1(\frac{1}{2}, 0), P_2(\frac{3}{2}, 0), P_3(1, \ln \frac{4}{3})$.
Diese stationären Punkte können z.B. unter Zuhilfenahme von Höhenlinien weiter untersucht werden. Hinreichende Bedingungen, die die zweiten partiellen Ableitungen nutzen, werden später angegeben.

2.) Für die Funktion f mit $f(x, y) = xy$ gilt $f_x(0, 0) = 0 = f_y(0, 0)$. Im Punkt $(0, 0)$ liegt aber kein Extremum, sondern ein Sattelpunkt vor.

Definition:

a) Eine Menge $M \subset R^n$ heißt **konvex**, wenn aus $\vec{x}_1 \in M$ und $\vec{x}_2 \in M$ auch $\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2 \in M \quad \forall \lambda \in [0,1]$ folgt.

b) $D \subset R^n$ sei eine konvexe Menge. Die Funktion $f|D \to R^1$ heißt **konvex**, (**streng konvex**) wenn
$f(\lambda\vec{x}_1+(1-\lambda)\vec{x}_2) \leq \lambda f(\vec{x}_1)+(1-\lambda)f(\vec{x}_2) \quad \forall \vec{x}_1 \in D,\ \vec{x}_2 \in D,\ \lambda \in [0,1]$
$(f(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2) < \lambda f(\vec{x}_1) + (1-\lambda)f(\vec{x}_2) \quad \forall \vec{x}_1 \in D,\ \vec{x}_2 \in D,\ \lambda \in (0,1))$
gilt.

c) Eine Funktion f heißt **konkav** (**streng konkav**), wenn $-f$ konvex (streng konvex) ist.

Beispiele:

a) ϵ-Umgebungen und Rechteckumgebungen im R^2 sind konvexe Mengen.

b) f mit $f(x,y) = x^2 + y^2$ ist eine (streng) konvexe Funktion.

c) f mit $f(x,y) = -(x^2 + y^2)$ ist eine (streng) konkave Funktion.

Satz: $f|D \subset R^n \to R^1$ sei konvex (konkav) und stetig partiell differenzierbar. Gilt in $\vec{x}_o \in D$ die Beziehung $\nabla f(\vec{x}_o) = \mathbf{0}$, dann besitzt f in $\vec{x}_o$ ein globales Minimum (Maximum).

Zur Untersuchung, ob eine gegebene Funktion lokal, d.h. in einer ϵ-Umgebung eines Punktes, konvexes oder konkaves Verhalten aufweist, kann die **Hessematrix $\boldsymbol{H(\vec{x}_o)}$** herangezogen werden:

$$H(\vec{x}_o) := \begin{pmatrix} fx_1x_1(\vec{x}_o) & fx_1x_2(\vec{x}_o) & \dots & fx_1x_n(\vec{x}_o) \\ \vdots & \vdots & & \vdots \\ fx_nx_1(\vec{x}_o) & fx_nx_2(\vec{x}_o) & \dots & fx_nx_n(\vec{x}_o) \end{pmatrix} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_o)\right)_{i,j}.$$

Falls f zweimal stetig differenzierbar ist, ist die Hessematrix symmetrisch.

Satz: Eine Funktion f sei auf einer ϵ-Umgebung $U_\epsilon\{\vec{x}_o\}$ von $\vec{x}_o$ zweimal stetig differenzierbar. Dann gilt

a) f konvex auf $U_\epsilon\{\vec{x}_o\} \Longleftrightarrow H(\vec{x}_o)$ positiv semidefinit $\forall \vec{x}_o \in U_\epsilon\{\vec{x}_o\}$.

b) f streng konvex auf $U_\epsilon\{\vec{x}_o\} \Leftarrow H(\vec{x}_o)$ positiv definit $\forall \vec{x} \in U_\epsilon\{\vec{x}_o\}$.

Es sei nun $f|D_f \subset R^2 \to R^1$ zweimal stetig differenzierbar.
Die Hessematrix im Punkt $\vec{x}_o$ besitzt dann die Gestalt $H_f(\vec{x}_o) = \begin{pmatrix} f_{xx}(\vec{x}_o) & f_{xy}(\vec{x}_o) \\ f_{xy}(\vec{x}_o) & f_{yy}(\vec{x}_o) \end{pmatrix}$.

Die Hauptminoren der Hessematrix sind also $f_{xx}(\vec{x}_o)$ und $\det H_f(\vec{x}_o)$. Aus dem Abschnitt über quadratische Formen ist bekannt, dass eine Matrix genau dann positiv definit (*negativ definit*) ist, wenn alle Hauptminoren positiv (*positiv für gerade Zeilenzahl und negativ für ungerade Zeilenzahl*) sind. Daraus ergibt sich, dass eine Hessematrix im vorliegenden zweidimensionalen Fall nur dann positiv definit oder negativ definit sein kann, wenn die sogenannte **Diskriminante**
$\delta(x,y) := \det(H_f(\vec{x}_o)) = f_{xx}(x,y) \cdot f_{yy}(x,y) - (f_{xy}(x,y))^2$
positiv ist. In Abhängigkeit vom Vorzeichen von $f_{xx}(\vec{x}_o)$ kann dann auf positive oder negative Definitheit geschlossen werden.

Satz: $f|D_f \subset R^2 \to R^1$ sei in einer ϵ-Umgebung $U_\epsilon\{\vec{x}_o\}$ zweimal stetig differenzierbar, und es gelte $f_x(x_o,y_o) = 0,\ f_y(x_o,y_o) = 0$.

a) Ist $\delta(x_o,y_o) > 0$, besitzt f an der Stelle (x_o,y_o) ein lokales Extremum. Es handelt sich dabei um ein lokales Minimum, falls $f_{xx}(x_o,y_o) > 0$ gilt, und um ein lokales Maximum, falls $f_{xx}(x_o,y_o) < 0$ erfüllt ist.

b) Gilt $\delta(x_o,y_o) < 0$, liegt an der Stelle $\vec{x}_o$ kein Extremwert vor.

c) Für $\delta(x_o,y_o) = 0$ ist eine Entscheidung ohne weitere Untersuchung nicht möglich.

Zum Beweis von a): Es seien $\delta(x_o,y_o) > 0$ und $f_{xx}(x_o,y_o) > 0$. Wegen der Stetigkeit der 2. Ableitung gibt es eine ϵ-Umgebung $U_\epsilon\{(x_o,y_o)\}$, so dass $f_{xx}(x,y) > 0$ und $\delta(x,y) > 0\ \forall (x,y) \in U_\epsilon\{(x_o,y_o)\}$ gilt. Also ist f dort (streng) konvex. Somit liegt in (x_o,y_o) ein lokales Minimum vor.
Gilt $\delta(x_o,y_o) > 0$ und $f_{xx}(x_o,y_o) < 0$, sind die obigen Überlegungen für $g(x,y) = -f(x,y)$ durchführbar. f ist dann (streng) konkav, und in (x_o,y_o) liegt ein lokales Maximum vor.

Allgemeiner gilt für Funktionen $f|D_f \subset R^n \to R^1$, die zweimal stetig differenzierbar sind, die folgende Aussage:
Es sei $\nabla f(\vec{x}_o) = \mathbf{0}$ erfüllt. Ist dann die Hessematrix $H_f(\vec{x}_o)$

- positiv definit, liegt in $\vec{x}_o$ ein lokales Minimum vor,
- negativ definit, liegt in $\vec{x}_o$ ein lokales Maximum vor,
- indefinit, liegt in $\vec{x}_o$ kein Extremun vor,
- positv semidefinit oder negativ semidefinit, ist eine Entscheidung ohne weitere Untersuchungen nicht möglich.

Fortsetzung des Beispiels: Betrachtet wird die Funktion f mit
$f(x,y) = x^2y - 2xy + \frac{3}{4}e^y$, für die die stationären Punkte $(\frac{1}{2}, 0), (\frac{3}{2}, 0), (1, \ln\frac{4}{3})$ ermittelt wurden.

Mit den zweiten partiellen Ableitungen
$f_{xx}(x,y) = 2y,\ f_{xy}(x,y) = 2x-2 = f_{yx}(x,y),\ f_{yy}(x,y) = \frac{3}{4}e^y$
erhält man
$\delta(x,y) = \begin{vmatrix} 2y & 2(x-1) \\ 2(x-1) & \frac{3}{4}e^y \end{vmatrix} = 2y \cdot \frac{3}{4}e^y - 4(x-1)^2.$

Einsetzen der stationären Punkte ergibt nun $\delta(\frac{1}{2},0) < 0$ und $\delta(\frac{3}{2},0) < 0$, d.h., in den Punkten $(\frac{1}{2},0)$ und $(\frac{3}{2},0)$ liegen keine Extremwerte vor.

Weiterhin ergibt sich für den dritten stationären Punkt
$\delta(1,\ln\frac{4}{3}) = 2\ln\frac{4}{3} > 0$, d.h., hier liegt ein Extremwert vor.
Wegen $f_{xx}(1,\ln\frac{4}{3}) = 2\ln\frac{4}{3} > 0$ handelt es sich dabei um ein Minimum.

1.6.3 Die Methode der kleinsten Quadrate

Aufgabenstellung: Gegeben seien Wertepaare $(x_i,y_i),\ i = 1\ldots,n$. Gesucht ist eine Funktion f, die "möglichst gut" den durch Wertpaare (x_i,y_i) gegebenen Zusammenhang widerspiegelt (Regressionsanalyse).

Vorgehen: Ausgangspunkt ist eine Vermutung über die Gestalt von f, die gewisse Parameter $a_1,\ldots,a_p$ enthält (z.B. $y = ax+b$ mit den Parametern $a,\ b$).
Die Parameter werden durch Lösung der Optimierungsaufgabe

$$\min_{a_1,\ldots,a_p \in R^p} \sum_{i=1}^{n} (y_i - f(x_i,a_1,\ldots,a_p))^2 \qquad (*)$$

ermittelt.
Die Berechnung der stationären Punkte der Extremwertaufgabe (*) führt im Allgemeinen auf ein nichtlineares Gleichungssystem.

Im Folgenden soll der Spezialfall der **einfachen linearen Regression** betrachtet werden. Der vermutete lineare Zusammenhang lautet dabei
$y = ax+b$.
Gesucht sind die Minimalstellen der Funktion φ mit
$\varphi(a,b) := \sum\limits_{i=1}^{n} (y_i - ax_i - b)^2,\ a \in R, b \in R.$

Das Gleichungssystem zur Bestimmung der stationären Punkte lautet:
$\varphi_a(a,b) = \sum\limits_{i=1}^{n} 2(y_i - ax_i - b)(-x_i) = 0, \qquad (1)$
$\varphi_b(a\ ,b) = \sum\limits_{i=1}^{n} 2(y_i - ax_i - b)(-1) = 0. \qquad (2)$

Aus (1) ergibt sich $\quad a \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i,$

aus (2) erhält man $\quad a \sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i.$

Damit liegt ein lineares Gleichungssystem zur Bestimmung von $a,\ b$ vor. Es kann in der Gestalt

$$A \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{n} x_i y_i \\ \sum\limits_{i=1}^{n} y_i \end{pmatrix} \qquad (*)$$

mit der Koeffizientenmatrix $A = \begin{pmatrix} \sum\limits_{i=1}^{n} x_i^2 & \sum\limits_{i=1}^{n} x_i \\ \sum\limits_{i=1}^{n} x_i & n \end{pmatrix}$ geschrieben werden.

Es wird nun vorausgesetzt, dass mindestens 2 der x_i-Werte verschieden sind (anderenfalls ist die Aufgabenstellung sinnlos). Dann folgt $\det A > 0$. Das ergibt sich aus der folgenden Rechnung:

$$\begin{aligned} \tfrac{1}{n} \det A &= \sum_{i=1}^{n} x_i^2 - \tfrac{1}{n} \left(\sum_{j=1}^{n} x_j \right)^2 \\ &= \sum_{i=1}^{n} x_i^2 - \tfrac{2}{n} \sum_{i=1}^{n} x_i \sum_{j=1}^{n} x_j + \tfrac{n}{n^2} \left(\sum_{i=1}^{n} x_j \right)^2 \\ &= \sum_{i=1}^{n} (x_i - \tfrac{1}{n} \sum_{j=1}^{n} x_j)^2 \ > \ 0. \end{aligned}$$

Somit ist das lineare Gleichungssystem (*) eindeutig lösbar. Die Lösung kann z.B. mittels Cramerscher Regel angegeben werden:

$$\hat{a} = \frac{\begin{vmatrix} \sum\limits_{i=1}^{n} x_i y_i & \sum\limits_{i=1}^{n} x_i \\ \sum\limits_{i=1}^{n} y_i & n \end{vmatrix}}{\det A} = \frac{n \sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n} x_i \sum\limits_{j=1}^{n} y_j}{n \sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{j=1}^{n} x_j)^2},$$

$$\hat{b} = \frac{\begin{vmatrix} \sum\limits_{i=1}^{n} x_i^2 & \sum x_i y_i \\ \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} y_i \end{vmatrix}}{\det A} = \frac{\sum\limits_{i=1}^{n} x_i^2 \sum\limits_{j=1}^{n} y_i - \sum\limits_{i=1}^{n} x_i y_i \sum\limits_{j=1}^{n} x_j}{n \sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}.$$

$(\hat{a}, \hat{b})$ ist der einzige stationäre Punkt. (Die Entscheidung, ob es sich dabei um ein Minimalstelle handelt, wäre an dieser Stelle auch durch Betrachtung des Verhaltens von φ an den "Rändern"des Definitionsgebietes möglich.)
Im Folgenden sollen die hinreichenden Bedingungen aus 1.6.2 angewandt werden. Die zweiten partiellen Ableitungen haben die Gestalt

$\varphi_{aa}(a,b) = 2\sum\limits_{i=1}^{n} x_i^2,\;\; \varphi_{ab}(a,b) = \varphi_{ba}(a,b) = 2\sum\limits_{i=1}^{n} x_i,\;\; \varphi_{bb}(a,b) = 2n.$

Damit ergibt sich $\delta(a,b) = \begin{vmatrix} 2\sum\limits_{i=1}^{n} x_i^2 & 2\sum\limits_{i=1}^{n} x_i \\ 2\sum\limits_{i=1}^{n} x_i & 2n \end{vmatrix} = 4\det A > 0,$

unabhängig von den Parameterwerten $a,\ b$. Somit ist die Funktion φ streng konvex. Außerdem gilt $2\sum\limits_{i=1}^{n} x_i^2 > 0$, also liegt im stationären Punkt eine globale Minimalstelle vor.
Als Ergebnis erhält man
$\min\limits_{(a,b)\in R^2} \varphi(a,b) = \sum\limits_{i=1}^{n}(y_i - \hat{a}x_i - \hat{b})^2$, und die gesuchte Funktion lautet
$y = \hat{a}x + \hat{b}$.

1.6.4 Extremwerte mit Nebenbedingungen

Es werden nur Punkte $\vec{x} \in R^n$, die den Bedingungen $g_i(\vec{x}) \le 0, i = 1,\dots,m$, genügen, bei der Suche nach Extremwerten zugelassen.

Es sei $M := \{\vec{x} \in R^n |\ g_i(\vec{x}) \le 0, i = 1,\dots,m\}$

Definition: $f|D_f \subset R^n \to R^1$ besitzt an der Stelle $\vec{x}_o \in D_f$ ein **lokales Maximum (lokales Minimum) unter den Nebenbedingungen $\boldsymbol{g_i(\vec{x}) \le 0, i = 1,\dots,m}$**, wenn $\vec{x}_o \in M$ und
$f(\vec{x}_o) \ge f(\vec{x}) \quad \forall\vec{x} \in D_f \cap U_\epsilon\{\vec{x}_o\} \cap M$
$(f(\vec{x}_o) \le f(\vec{x}) \quad \forall\vec{x} \in D_f \cap U_\epsilon\{\vec{x}_o\} \cap M)$
für eine geeignete ϵ-Umgebung $U_\epsilon\{\vec{x}_o\}$ gilt.

Im Folgenden wird nur der Spezialfall behandelt, dass **alle Nebenbedingungen in Form von Gleichungsrestriktionen** vorliegen.

Beispiel: Gesucht sind die Extremwerte der Funktion f mit $f(x,y) = x^2+2y^2$ unter der Nebenbedingung $\varphi(x,y) = y - x^2 + 1 = 0$.
Durch Betrachtung der Höhenlinien der Funktion f kann man erkennen, dass in den Punkten $(+\frac{1}{2}\sqrt{3}, -\frac{1}{4})$ und $(-\frac{1}{2}\sqrt{3}, -\frac{1}{4})$ jeweils lokale Minima und im Punkt $(0,-1)$ ein lokales Maximum vorliegt.

Die Lösung kann aber auch gesucht werden, indem die Nebenbedingung in die Gleichung der Zielfunktion f eingesetzt wird:

(1) Einsetzen von $y = x^2 - 1$ liefert die Funktion h mit
$h(x) := f(x, y(x)) = x^2 + 2(x^2-1)^2$.
Als notwendige Optimalitätsbedingung ergibt sich daher
$h'(x) = 2x + 4(x^2-1)2x = 2x(4x^2-3) = 0$.

"Extremwertverdächtig" sind somit die Werte $x_{E_1} = 0$ und $x_{E_{2,3}} = \frac{1}{2}\sqrt{3}$. Die zugehörigen y-Werte berechnet man aus $y = x^2 - 1$.
Aus $h''(x) = 2(4x^2 - 3) + 2x \cdot 8x$ erhält man $h''(0) < 0$ und $h''(+\frac{1}{2}\sqrt{3}) > 0$ sowie $h''(-\frac{1}{2}\sqrt{3}) > 0$, d.h., im Punkt $(0, -1)$ liegt ein lokales Maximum vor, und die Punkte $(+\frac{1}{2}\sqrt{3}, -\frac{1}{4})$ sowie $(-\frac{1}{2}\sqrt{3}, -\frac{1}{4})$ sind jeweils lokale Minimalstellen.

(2) Setzt man nun die Nebenbedingung in der Gestalt $x^2 = y + 1$ in die Zielfunktion ein, erhält man die Funktion $\tilde{h}$ mit $\tilde{h}(y) = 2y^2 + y + 1$. Aus der notwendigen Optimalitätsbedingung $\tilde{h}'(y) = 4y + 1 = 0$ ergibt sich nur die Lösung $y_E = -\frac{1}{4}$, die unter Berücksichtigung der Nebenbedingung schließlich die Punkte $(+\frac{1}{2}\sqrt{3}, -\frac{1}{4})$ und $(-\frac{1}{2}\sqrt{3}, -\frac{1}{4})$ liefert. Das lokale Maximum im Punkt $(0, -1)$ wird auf diesem Weg nicht gefunden. Der Grund dafür liegt darin, dass aus der Nebenbedingung $y \geq -1$ folgt und somit $y = -1$ wie ein Randpunkt gesondert betrachtet werden muss. Dieser Effekt wiederum beruht darauf, dass die durch die Nebenbedingung gegebene Kurve in keiner Rechteckumgebung des Punktes $(0, -1)$ als Funktion von y darstellbar ist.

Im Folgenden sollen die Extremwerte einer Funktion $f|D_f \subset R^2 \to R^1$ unter der Nebenbedingung $\varphi(x, y) = 0$ untersucht werden. Dabei wird vorausgesetzt, dass die Funktionen f und φ stetig partiell differenzierbar sind.

Angenommen, es gelte $\varphi_y(x_o, y_o) \neq 0$ für ein $(x_o, y_o) \in D_f$. Nach dem Hauptsatz über implizite Funktionen gibt es dann eine Rechteckumgebung $R_{\delta_1\delta_2}(x_o, y_o)$ und eine Funktion g derart, dass die Beziehung $\varphi(x, y) = 0$ auch in der Gestalt $y = g(x)\ \forall (x, y) \in R_{\delta_1\delta_2}(x_o, y_o)$ angegeben werden kann.

Es sei nun $h(x) := f(x, g(x))$. Wenn h ein Extremum bei x_o besitzt, muss (mit $y_o = g(x_o)$)
$h'(x_o) = f_x(x_o, y_o) + f_y(x_o, y_o) \cdot g'(x_o) = 0$ gelten.

Mit $g'(x_o) = \dfrac{-\varphi_x(x_o, y_o)}{\varphi_y(x_o, y_o)}$ ergibt sich $f_x(x_o, y_o) - \dfrac{f_y(x_o, y_o)}{\varphi_y(x_o, y_o)}\varphi_x(x_o, y_o) = 0$.
Setzt man nun $\lambda_o := \dfrac{-f_y(x_o, y_o)}{\varphi_y(x_o, y_o)}$, so erhält man einerseits
$f_x(x_o, y_o) + \lambda_o\varphi_x(x_o, y_o) = 0$
und andererseits durch Umstellen der Definitionsgleichung für λ_o
$f_y(x_o, y_o) + \lambda_o\varphi_y(x_o, y_o) = 0$.

Zusammenfassend kann man feststellen, dass (x_o, y_o) und λ_o die Bedingungen

$$\left.\begin{aligned} f_x(x_o, y_o) + \lambda_o\varphi_x(x_o, y_o) &= 0 \\ f_y(x_o, y_o) + \lambda_o\varphi_y(x_o, y_o) &= 0 \\ \varphi(x_o, y_o) &= 0 \end{aligned}\right\} (*)$$

erfüllen.

Ein analoges Vorgehen ist im Fall $\varphi_x(x_o, y_o) \neq 0$ möglich.

Unter Verwendung der sogenannten **Lagrangefunktion** L mit
$L(x, y, \lambda) := f(x, y) + \lambda(x, y)$
kann das System (*) in der folgenden Form geschrieben werden. λ heißt dabei Lagrangemultiplikator.

$$\left.\begin{array}{l} \frac{\partial L}{\partial x}(x_o, y_o, \lambda_o) = 0 \\ \frac{\partial L}{\partial y}(x_o, y_o, \lambda_o) = 0 \\ \frac{\partial L}{\partial \lambda}(x_o, y_o, \lambda_o) = 0 \end{array}\right\} (**)$$

Der folgende Satz fasst die obigen Überlegungen zusammen:

Satz (Multiplikatorenregel von Lagrange):
Die Funktion $f|D_f \subset R^2 \to R^1$ besitze an der Stelle $(x_o, y_o) \in D_f$ ein lokales Extremum unter der Nebenbedingung $\varphi(x, y) = 0$. Die Funktionen f und φ seien in allen Punkten der Menge $\{(x, y) \in R^2 | \ \varphi(x, y) = 0\}$ stetig partiell differenzierbar, und es gelte $\varphi_x(x_o, y_o) \neq 0 \vee \varphi_y(x_o, y_o) \neq 0$.
Dann existiert ein $\lambda_o \in R$ derart, dass das Gleichungssystem (**) erfüllt ist.

Beispiel: Für das oben betrachtete Beispiel hat die Lagrangefunktion die Gestalt $L(x, y, \lambda) := x^2 + 2y^2 + \lambda(y - x^2 + 1)$. Daraus ergibt sich das Gleichungssystem

$$\begin{array}{ll} L_x(x, y, \lambda) = 2x - 2x\lambda = 2x(1 - \lambda) = 0 & (1) \\ L_y(x, y, \lambda) = 4y + \lambda = 0 & (2) \\ F_\lambda(x, y, \lambda) = y - x^2 + 1 = 0 & (3)\ . \end{array}$$

Aus (2) erhält man $\lambda = -4y$, was durch Einsetzen in (1) zu $2x(1 + 4y) = 0$ führt. Diese Gleichung ist erfüllt für $x = 0$ (Fall 1) oder $y = -\frac{1}{4}$ (Fall 2).

Im Fall 1 ergibt sich mit unter Ausnutzung von (3) $y = -1$; im Fall 2 nimmt (3) die Gestalt $x^2 = \frac{3}{4}$ an. Somit erhält man $x_{1,2} = \pm\frac{1}{2}\sqrt{3}$.

Die Lagrangesche Multiplikatorenregel ist nur ein notwendiges Kriterium, die weitere Untersuchung der gefundenen "extremwertverdächtigen" Punkte ist z. B. durch Betrachtung der Höhenlinien möglich.

Für Funktionen $f|R^n \to R^1$, deren Extremwerte unter den m Gleichungsrestriktionen $\varphi_i(x_1, \ldots, x_n) = 0,\ i = 1, \ldots, m$ gesucht werden, kann man in ähnlicher Weise vorgehen.

Die **Lagrangefunktion** hat die folgende Gestalt:
$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) := f(x_1, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i \varphi_i(x_1, \ldots, x_n)$.

Wenn f und alle $\varphi_i,\ i = 1, \ldots, m$, stetig partiell differenzierbar sind und die Matrix $\begin{pmatrix} \frac{\delta\varphi_1}{\partial x_1}(\vec{x}_o) & \ldots & \frac{\partial\varphi_1}{\partial x_n}(\vec{x}_o) \\ \vdots & & \vdots \\ \frac{\partial\varphi_m}{\partial x_1}(\vec{x}_o) & \ldots & \frac{\partial\varphi_m}{\partial x_n}(\vec{x}_o) \end{pmatrix}$ den Rang m besitzt, dann ist die Existenz reeller Zahlen $\lambda_1^{(o)}, \ldots, \lambda_m^{(o)}$ mit

$$L_{x_i}(x_1^{(o)}, \ldots, x_n^{(o)}, \lambda_1^{(o)}, \ldots, \lambda_m^{(o)}) = 0 \quad i = 1, \ldots, n,$$
$$L_{\lambda_j}(x_1^{(o)}, \ldots, x_n^{(o)}, \lambda_1^{(o)}, \ldots, \lambda_m^{(o)}) = 0 \quad j = 1, \ldots, m,$$
$$(\Leftrightarrow \varphi_j(x_1^{(o)}, \ldots, x_n^{(o)}) = 0 \quad j = 1, \ldots, m)$$

notwendig dafür, dass an der Stelle $\vec{x}_o$ ein lokales Extremum der Funktion f unter den Nebenbedingungen $\varphi_i(\vec{x}) = 0,\ i = 1, \ldots m$, vorliegt.

2 Gewöhnliche Differenzengleichungen und Differentialgleichungen

2.1 Differentialgleichungen

2.1.1 Einführung

Gleichungen, in denen (eine oder mehrere) Ableitungen einer gesuchten Funktion einer Variablen enthalten sind, heißen **gewöhnliche Differentialgleichungen**:
$F(x, y, y', \ldots, y^{(n)}) = 0$
ist eine gewöhnliche Differentialgleichung der Ordnung n (in impliziter Form).

Gesucht sind alle Funktionen $y = \varphi(x)$, die der Differentialgleichung genügen (sogenannte allgemeine Lösung).

Im Gegensatz zu den gewöhnlichen Differentialgleichungen kommen bei **partiellen Differentialgleichungen** partielle Ableitungen von Funktionen mehrerer Veränderlichen vor.

Beispiel B1: Wachstumsmodell des Volkseinkommens (nach Boulding)

Es bezeichnen
$y(t)$ das Volkseinkommen zum Zeitpunkt t,
$c(t)$ den Konsum zum Zeitpunkt t,
$i(t)$ den Umfang der Investitionen zum Zeitpunkt t.

Das Modell geht von folgenden Annahmen aus:
$y(t) = c(t) + i(t)$,
$c(t) = \alpha + \beta y(t) \quad (\alpha \geq 0, \quad 0 < \beta < 1)$,

$y'(t) = \gamma i(t) \quad (\gamma > 0)$.

Daraus erhält man die Differentialgleichung
$y'(t) = \gamma(y(t) - c(t)) = \gamma(y(t) - \beta y(t) - \alpha) = \gamma(1-\beta)y(t) - \alpha\gamma$,
in Kurzform $y' = \gamma(1-\beta)y - \alpha\gamma$.

<u>Definition:</u> Eine Differentialgleichung der Gestalt

$$y^{(n)}(x) + p_{n-1}(x)y^{(n-1)}(x) + \cdots + p_1(x)y'(x) + p_o(x)y(x) = q(x)$$

heißt **lineare Differentialgleichung n-ter Ordnung**. Ist $q(x) \equiv 0$, heißt die lineare Differentialgleichung **homogen**, anderenfalls **inhomogen**. Gilt $p_i(x) = a_i, i = 0, \ldots, n-1$, spricht man von einer linearen Differentialgleichung mit konstanten Koeffizienten.

Beispiele:

1.) $y''' + x^2y'' - \ln xy = 0$ ist eine homogene lineare Differentialgleichung 3. Ordnung.

2.) Die Differentialgleichung aus Beispiel B1 $y' = \gamma(1-\beta)y - \alpha\gamma$ ist eine inhomogene lineare Differentialgleichung 1. Ordnung mit konstanten Koeffizienten.

2.2 Gewöhnliche Differentialgleichungen 1. Ordnung

Beispiel: Die Differentialgleichung $y' = x$ besitzt die allgemeine Lösung $y(x) = \frac{x^2}{2} + c$, d.h., die Lösung stellt eine Kurvenschar dar.

Will man aus der Kurvenschar einer allgemeinen Lösung eine Kurve auswählen, die durch einen vorgegebenen Punkt (x_o, y_o) verläuft, kann dies durch Einsetzen des Punktes in die allgemeine Lösung geschehen (allerdings nur unter gewissen Existenz- und Eindeutigkeitsbedingungen). Der Punkt (x_o, y_o) charakterisiert häufig einen "Anfangszustand" des jeweils beschriebenen Phänomens; entsprechend wird die Bedingung $y(x_o) = y_o$ dann **Anfangsbedingung** genannt.
Im betrachteten Beispiel $y' = x$ ergibt $y(x_o) = y_o$ die Beziehung $y_o = \frac{x_o^2}{2} + c$, damit $c = y_o - \frac{x_o^2}{2}$ und schließlich die **spezielle** (oder partikuläre) **Lösung** $y(x) = \frac{x^2}{2} + y_o - \frac{x_o^2}{2}$.

2.2.1 Die Methode der Trennung der Variablen

Ausgangspunkt ist eine Differentialgleichung der Gestalt $y' = f(x) \cdot g(y)$ (Differentialgleichung mit trennbaren Variablen).

Formal wird wie folgt vorgegangen:

$$y' = \frac{dy}{dx} = f(x) \cdot g(y)$$

$$\frac{dy}{g(y)} = f(x)dx$$
$$\int \frac{dy}{g(y)} = \int f(x)dx.$$

Beispiel: Für die Differentialgleichung $y' = \frac{y}{2x},\ x \neq 0$, ergibt sich mit dieser Regel (unter der Annahme, dass $y \not\equiv 0$) :
$\int \frac{1}{y} dy = \int \frac{1}{2x} dx$
$\ln|y| + c_1 = \frac{1}{2}\ln|x| + c_2$
$\ln|y| = \frac{1}{2}\ln|x| + c_3$
$e^{\ln|y|} = e^{\frac{1}{2}\ln|x| + c_3} = (e^{\ln|x|})^{\frac{1}{2}} \cdot \underbrace{e^{c_3}}_{=:c_4>0}$
$|y| = c_4\sqrt{|x|}$.
Für $y > 0$ folgt dann wegen $|y| = y$ die Beziehung $y = +c_4\sqrt{|x|}$ und für $y < 0$ erhält man $y = -c_4\sqrt{|x|}$.

Es bleibt noch die Frage zu klären, ob die beim obigen Vorgehen ausgeschlossene Funktion $y \equiv 0$ zur allgemeinen Lösung gehört. Einsetzen in die Differentialgleichung zeigt, daß $y \equiv 0$ ebenfalls Lösung ist.

In zusammenfassender Form kann die allgemeine Lösung in der Gestalt
$y = c\sqrt{|x|},\ c \in R$ beliebig, angegeben werden.

Bei Anwendungen in Ökonomie oder Technik sind zusätzliche Voraussetzungen über den Definitionsbereich, z. B. $x > 0$, häufig anzutreffen.

Es soll nun die Frage behandelt werden, unter welchen Voraussetzungen durch einen vorgegebenen Punkt (x_o, y_o) genau eine Kurve der Lösungskurvenschar verläuft.

Satz: Gegeben sei die Differentialgleichung $y' = f(x, y)$. Ist die Funktion f auf einer Rechteckumgebung $R_{ab}(x_o, y_o)$ eines Punktes (x_o, y_o) stetig und erfüllt die Bedingung
$|f(x, y_1) - f(x, y_2)| \leq k|y_1 - y_2|$ für alle $(x, y_1),\ (x, y_2) \in R_{a,b}(x_o, y_o)$ (*),
so hat die Differentialgleichung $y' = f(x, y)$ genau eine Lösung in $R_{a,b}(x_o, y_o)$, welche die Anfangsbedingung $y(x_o) = y_o$ erfüllt.

Bemerkung: Hinreichend für (*) ist die Stetigkeit von f_y auf $R_{a,b}(x_o, y_o)$

Fortsetzung des Beispiels B1: Die Differentialgleichung $y' = \gamma(1-\beta)y - \alpha\gamma$ geht mit den Abkürzungen $a := \gamma(1-\beta),\ b := \alpha\gamma$ in die Form $y' = ay - b$ über, die mit der Methode der Trennung der Variablen behandelt werden kann. Die Annahme, dass das Volkseinkommen zum Ausgangszeitpunkt $t = 0$ den Wert y_o besitzt, erfasst man mit der Anfangsbedingung $y(0) = y_o$.

Die Lösung kann wie folgt ermittelt werden:
$\int \frac{dy}{ay \cdot b} = \int 1dx \qquad (y \not\equiv \frac{b}{a})$

$$\frac{1}{a} \ln|y - \frac{b}{a}| = t + c_1$$

$$\ln|y - \tfrac{b}{a}| = ta + c_2$$

$$|y - \tfrac{b}{a}| = e^{at} \cdot c_3.$$

Schließlich ergibt sich $y - \frac{b}{a} = c_4 e^{at}$, wobei $c_4 > 0$ oder $c_4 < 0$ gilt. Da $y = \frac{b}{a}$ ebenfalls Lösung ist, lautet die allgemeine Lösung $y = ce^{at} + \dfrac{b}{a},\ c \in R$ beliebig.
Ersetzt man die Abkürzungen wieder durch die Originalsymbole, erhält man $y = ce^{\gamma(1-\beta)t} + \frac{\alpha}{1-\beta}, c \in R$.

Wegen $y(0) = c + \frac{\alpha}{1-\beta} = y_o$ folgt $c = y_o - \frac{\alpha}{1-\beta}$, also
$y(t) = (y_o - \frac{\alpha}{1-\beta})e^{\gamma(1-\beta)t} + \frac{\alpha}{1-\beta}$.

2.2.2 Lineare Differentialgleichungen 1. Ordnung

Gegeben sei eine Differentialgleichung der Form $y' + p(x)y = q(x)$.

Die zugehörige homogene Differentialgleichung $y' + p(x)y = 0$ kann stets durch Trennung der Variablen gelöst werden:

$y' = -p(x)y$
$\int \frac{dy}{y} = -\int p(x)dx \qquad y \not\equiv 0$
$\ln|y| = -\int p(x)dx + c_1$.

Die **allgemeine Lösung der homogenen linearen Differentialgleichung** 1. Ordnung lautet somit
$y_h = ce^{-\int p(x)dx} =: c\varphi(x),\ c \in R$, beliebig.

Die Struktur der allgemeinen Lösung der (inhomogenen) Ausgangsdifferentialgleichung beschreibt der folgende Satz:

Satz: Es seien y_s eine (spezielle) Lösung der inhomogenen linearen Differentialgleichung $y' + p(x)y = q(x)$ und y_h die allgemeine Lösung der zugehörigen homogenen Differentialgleichung. Dann ist $y_a := y_h + y_s$ die allgemeine Lösung der inhomogenen linearen Differentialgleichung.
Beweis:

1.) Es wird zunächst gezeigt, dass jede Funktion y_a der angegebenen Gestalt die Differentialgleichung erfüllt:

$y_a'(x) = y_h'(x) + y_s'(x) = -p(x)y_h(x) - p(x)y_s(x) + q(x)$
$= -p(x)(y_h(x) + y_s(x)) + q(x) = -p(x)y_a(x) + q(x).$

2.) Es sei $\tilde{y}$ eine spezielle Lösung der inhomogenen Differentialgleichung. Es wird gezeigt, dass dann die Funktion $\tilde{y} - y_s$ die homogene Differentialgleichung erfüllt.
$(\tilde{y}(x) - y_s(x))' = \tilde{y}'(x) - y_s'(x) = -p(x)\tilde{y}(x) + q(x) + p(x)y_s(x) - q(x)$
$= -p(x)(\tilde{y}(x) - y_s(x)).$

Damit werden die folgenden Schritte zur Lösung einer inhomogenen linearen Differentialgleichung nahe gelegt:

1.) Bestimmung der allgemeinen Lösung y_h der homogenen Differentialgleichung,

2.) Bestimmung einer speziellen Lösung y_s der inhomogenen Differentialgleichung,

3.) Angabe der allgemeinen Lösung gemäß $y_a := y_h + y_s$.

Bestimmung einer speziellen Lösung einer inhomogenen linearen Differentialgleichung

Ein Verfahren, das stets angewendet werden kann (sofern es möglich ist, die Integrale in geschlossener Form auszuwerten), ist die Methode der **Variation der Konstanten** (J. Bernoulli):

Dazu wird ein spezieller Ansatz für die Gestalt von y_s gemacht:

Ansatz: $y_s(x) = c(x) \cdot \varphi(x)$,
wobei φ eine Lösung der zugehörigen homogenen Differentialgleichung und c eine zu bestimmende Funktion von x bezeichnen.

Einsetzen des Ansatzes in die Differentialgleichung ergibt

$c'(x)\varphi(x) + c(x)\varphi'(x) + p(x)(c(x) \cdot \varphi(x)) = q(x)$
$c'(x)\varphi(x) + c(x)[\varphi'(x) + p(x)\varphi(x)] = q(x).$

Da φ Lösung der homogenen Differenentialgleichung ist, muss $\varphi'(x)+p(x)\varphi(x) = 0$ gelten. Somit erhält man $c'(x)\cdot\varphi(x) = q(x)$ und daraus $c'(x) = \dfrac{q(x)}{\varphi(x)}$, also eine Bestimmungsgleichung für $c(x)$.
Die allgemeine Lösung der inhomogenen Differentialgleichung lautet dann:

$y_a(x) = c \cdot \varphi(x) + c(x) \cdot \varphi(x).$

Beispiel: Gegeben sei die Differentialgleichung $y' + \frac{y}{x} = \sqrt{x^2+1},\ x > 0.$

1.) Im ersten Schritt wird die allgemeine Lösung der homogenen linearen Differentialgleichung bestimmt:

$y' = -\frac{y}{x}$
$\displaystyle\int \frac{dy}{y} = -\int \frac{dx}{x} \qquad y \not\equiv 0$
$\ln|y| = -\ln|x| + c_1 \qquad$ (Wegen $x > 0$ gilt $|x| = x$).
$y = e^{c_1} e^{-\ln x} = c_2 \dfrac{1}{e^{\ln x}} = \dfrac{c_2}{x}$
$y_h = \dfrac{c}{x},\ c \in R.$

2.) Im zweiten Schritt wird eine spezielle Lösung der inhomogenen linearen Differentialgleichung mittels Variation der Konstanten bestimmt.

Ansatz: $y_s = \frac{c(x)}{x}$
Einsetzen in die Differentialgleichung liefert:
$\dfrac{c'x - c}{x^2} + \dfrac{c}{x^2} = \sqrt{x^2+1}$
$c' = x\sqrt{x^2+1}$
$c(x) = \int x\sqrt{x^2+1}dx = \frac{1}{2}\int t^{\frac{1}{2}}dt = \frac{1}{2}\cdot\frac{2}{3}t^{\frac{3}{2}} = \frac{1}{3}(x^2+1)^{\frac{3}{2}} + K.$

Da nur eine spezielle Lösung benötigt wird, kann $K = 0$ gewählt werden. Somit ergibt sich $y_s = \dfrac{1}{x}\,\dfrac{1}{3}(x^2+1)^{\frac{3}{2}}$.

3.) Die allgemeine Lösung der Differentialgleichung lautet nun
$y_a = c\cdot\dfrac{1}{x} + \dfrac{1}{x}\cdot\dfrac{1}{3}(x^2+1)^{\frac{3}{2}} = \dfrac{1}{x}(c + \dfrac{1}{3}(x^2+1)^{\frac{3}{2}})$.
Eine Probe kann durch Einsetzen in die Differentialgleichung erfolgen.

2.3 Lineare Differentialgleichungen höherer Ordnung

2.3.1 Allgemeine Eigenschaften

Gegeben sei eine lineare Differentialgleichung n-ter Ordnung in der allgemeinen Gestalt :
$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_o(x)y = q(x). \qquad (1)$

Gesucht ist eine n-mal differenzierbare Funktion y, die die Differentialgleichung (1) und außerdem die Beziehungen

$y(x_o) = \eta_o,\ y'(x_o) = \eta_1,\ \ldots, y^{(n-1)}(x_o) = \eta_{n-1} \qquad (2)$

erfüllt (Anfangswertaufgabe (AWA).

Satz: Wenn die Funktionen $p_o, \ldots, p_{n-1}; q$ stetig auf dem Intervall $[a,b]$ mit $x_o \in (a,b)$ sind, dann besitzt die Differentialgleichung (1) genau eine Lösung y, welche die Bedingungen (2) erfüllt.

Wie für lineare Differentialgleichung 1. Ordnung kann man zeigen, dass $y_a = y_h + y_s$ gilt, wobei y_a die allgemeine Lösung der Differentialgleichung,

y_h die allgemeine Lösung der homogenen Differentialgleichung und y_s eine spezielle Lösung der inhomogenen Differentialgleichung bezeichnen.

Definition: Die Funktionen $\varphi_1, \varphi_2, \ldots, \varphi_n$ heißen im Intervall $[a, b]$ **linear unabhängig**, wenn aus der Beziehung
$c_1\varphi_1(x) + \ldots + c_n\varphi_n(x) = 0 \;\forall x \in [a, b]$
stets $c_1 = c_2 = \ldots = c_n = 0$ folgt.
Anderenfalls (d.h., wenn $c_1\varphi_1(x) + \cdots + c_n\varphi_n(x) = 0 \;\forall x \in [a, b]) \wedge \exists c_i \neq 0)$ heißen die Funktionen **linear abhängig**.

Beispiele:

1.) Die Funktionen $\varphi_1(x) = 1,\ \varphi_2(x) = x,\ \varphi_3(x) = x^2$ sind im Intervall $[a, b]$ linear unabhängig. Das ergibt sich aus den folgenden Überlegungen:
Es gelte $c_1 + c_2x + c_3x^2 = 0 \;\forall x \in [a, b]$. Einsetzen von drei x-Werten $x_1,\ x_2,\ x_3 \in [a, b],\ x_i \neq x_j, i \neq j$, führt auf eine homogenes lineares Gleichungssystem zur Bestimmung von $c_1,\ c_2,\ c_3$, das nur die triviale Lösung $c_1 = c_2 = c_3 = 0$ besitzt. (Die Koeffizientendeterminante ist eine sogenannte Vandermondesche Determinante, die unter den getroffenen Annahmen stets von Null verschieden ist.)

2.) Die Funktionen $\varphi_1(x) = 1,\ \varphi_2(x) = x,\ \varphi_3(x) = 5x + 6$ sind linear abhängig im Intervall $[a, b]$, denn es gilt
$-6\varphi_1(x) - 5\varphi_2(x) + \varphi_3(x) = 0 \;\forall x \in [a, b]$.

Definition: Ist jede der n Funktionen $\varphi_1, \ldots, \varphi_n$ mindestens $(n-1)$-mal differenzierbar, so heißt

$$W_{\varphi_1,\ldots,\varphi_n}(x) := \begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \ldots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \ldots & \varphi_n'(x) \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \ldots & \varphi_n^{(n-1)}(x) \end{vmatrix}$$

die **Wronskische Determinante** des Funktionensystems $\varphi_1, \ldots, \varphi_n$.

Satz: Sind die Funktionen $\varphi_1, \ldots, \varphi_n$ linear abhängig in $[a, b]$ und $(n-1)$-mal differenzierbar, so gilt $W_{\varphi_1,\ldots,\varphi_n}(x) = 0 \;\forall x \in [a, b]$.

Äquivalent dazu ist die folgende Aussage: Gilt $W_{\varphi_1,\ldots,\varphi_n}(x) \neq 0$ für mindestens ein $x \in [a, b]$, so sind die Funktionen $\varphi_1, \ldots, \varphi_n$ linear unabhängig in $[a, b]$.

Die Umkehrung gilt im Allgemeinen nicht.

Definition: Jedes System von n linear unabhängigen Lösungen einer homogenen linearen Differentialgleichung n-ter Ordnung heißt **Fundamentalsystem** (FS) von Lösungen der Differentialgleichung.

Satz: Die Lösungen $\varphi_1, \ldots, \varphi_n$ einer linearen homogenen Differentialgleichung n-ter Ordnung bilden genau dann ein Fundamentalsystem, wenn die Wronskische Determinante die Beziehung $W_{\varphi_1,\ldots,\varphi_n}(x) \neq 0$ erfüllt.

Bemerkung: Für Lösungen einer homogenen linearen Differentialgleichung n-ter Ordnung gilt (für ein Intervall $[a, b]$ aus dem gemeinsamen Definitionsbereich der Lösungen) entweder
$W_{\varphi_1,\ldots,\varphi_n}(x) \neq 0 \; \forall x \in [a, b]$ oder $W_{\varphi_1,\ldots,\varphi_n}(x) = 0 \; \forall x \in [a, b]$.

Satz: Gegeben sei die homogene lineare Differentialgleichung n-ter Ordnung $y^{(n)} + p_{n-1}(x)y^{(n-1)} + \ldots + p_o(x)y = 0$ mit auf $[a, b]$ stetigen Funktionen $p_k, \; k = 0, \ldots, n-1$. Dann besitzt die allgemeine Lösung dieser Differentialgleichung die Gestalt
$y(x) = c_1\varphi_1(x) + \ldots + c_n\varphi_n(x)$,
wobei $\{\varphi_1, \ldots, \varphi_n\}$ ein Fundamentalsystem von Lösungen der Differentialgleichung und $c_1, \ldots, c_n$ beliebige reelle Konstanten bezeichnen.

2.3.2 Fundamentalsysteme für homogene lineare Differentialgleichungen mit konstanten Koeffizienten

Gegeben sei die Differentialgleichung
$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + a_oy = 0$ (3)
mit den konstanten Koeffizienten $a_i, \; i = 0, \ldots, n-1$.

Die Funktionen des Fundamentalsystems werden mit dem

Ansatz: $\boldsymbol{y(x) = e^{\;\,x}}$

gesucht. Einsetzen des Ansatzes in die Differentialgleichung ergibt dann
$\lambda^n e^{\lambda x} + a_{n-1}\lambda^{n-1}e^{\lambda x} + \ldots + a_1\lambda e^{\lambda x} + a_o e^{\lambda x} = 0$ und weiter

$$\underbrace{e^{\lambda x}}_{\neq 0 \forall x \forall \lambda} (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_o) = 0.$$

Die enstandene Gleichung
$\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_o = 0$ (4)
heißt **charakteristische Gleichung der Differentialgleichung**.
$P_n(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_o$ wird auch charakteristisches Polynom genannt.

(4) besitzt n (ggf. komplexe) Lösungen $\lambda_i, \; i = 1, \ldots, n$.

Satz:

a) Besitzt die charakteristische Gleichung (4) n verschiedene reelle Lösungen $\lambda_1, \ldots, \lambda_n$ ($\lambda_i \in R,\ \lambda_i \neq \lambda_j,\ i \neq j$), dann ist
$\varphi_1(x) = e^{\lambda_1 x}, \varphi_2(x) = e^{\lambda_2 x}, \ldots, \varphi_n(x) = e^{\lambda_n x}$
ein Fundamentalsystem für (3).

b) Besitzt (4) nur reelle Lösungen, wobei gewisse Lösungen mehrfach auftreten, so werden jeder k-fachen Nullstelle λ_i die k Funktionen
$\varphi_{i1}(x) = e^{\lambda_i x},\ \varphi_{i2}(x) = xe^{\lambda_i x}, \ldots,\ \varphi_{i,k}(x) = x^{k-1}e^{\lambda_i x}$
zugeordnet.

Beispiel: Gegeben sei die Differentialgleichung $y''' - 3y' - 2y = 0$.

Die charakteristische Gleichung für diese Differentialgleichung lautet
$\lambda^3 - 3\lambda - 2 = 0$.
Sie besitzt die Nullstellen $\lambda_1 = 2, \lambda_{2,3} = -1$.
Damit ist e^{2x}, e^{-x}, xe^{-x} ein Fundamentalsystem von Lösungen für die gegebene Diffrentialgleichung und die allgemeine Lösung der Differentialgleichung lautet $y = c_1e^{2x} + c_2e^{-x} + c_3xe^{-x}$.

Um auch dann ein Fundamentalsystem von Lösungen angeben zu können, wenn nicht alle Nullstellen der charakteristischen Gleichung reell sind, werden Grundkenntnisse über komplexe Zahlen benötigt. Deshalb wird im Folgenden ein kurzer Abschnitt zu komplexen Zahlen eingefügt.

2.3.3 Komplexe Zahlen

Komplexe Zahlen wurden bereits zu Beginn des 16. Jahrhunderts eingeführt. Ausgangspunkt der Überlegungen war der Wunsch, den Zahlenbereich so zu erweitern, dass auch z.B. die Gleichung $x^2 + 1 = 0$ eine Lösung besitzt.

Mit der **imaginären Einheit** i, die durch $i^2 = -1$ definiert ist, lautet die **algebraische Form einer komplexen Zahl**: $z = \alpha + i\beta, \quad \alpha \in R,\ \beta \in R$.

α heißt dann **Realteil** von z ($\alpha = Re(z)$), und β heißt **Imaginärteil** von z ($\beta = Im(z)$).

Ist $\alpha = 0$, spricht man von einer imaginären Zahl.
Ist $\beta = 0$, stellt z eine reelle Zahl dar.(Mit der Einführung der komplexen Zahlen wurde also eine Erweiterung des Zahlenbereiches vorgenommen.)

Die komplexen Zahlen können in der sogenannten Gaußschen Zahlenebene dargestellt werden.

Die **trigonometrische Form** einer komplexen Zahl $z = \alpha + i\beta$ lautet
$z = r(\cos\varphi + i\sin\varphi),\ 0 \leq r \leq \infty,\ -\pi < \varphi \leq \pi.$

$r = \sqrt{\alpha^2 + \beta^2} =: |z|$ heißt **Betrag** von z;
$\varphi =: \arg z$ heißt **Argument** von z.
Aus den Beziehungen $\alpha = r\cos\varphi$ und $\beta = r\sin\varphi$ erhält man $\tan\varphi = \frac{\beta}{\alpha}$ und damit (unter Beachtung des Vorzeichens von α und β) eine Möglichkeit, bei Kenntnis von α und β den Wert φ zu bestimmen.

Es gilt die **Eulersche Formel**: $e^{\alpha+i\beta} = e^{\alpha}(\cos\beta + i\sin\beta)$.

Rechenregeln für komplexe Zahlen:

Es seien $z_1 = \alpha_1 + i\beta_1,\ z_2 = \alpha_2 + i\beta_2$.

(a) $z_1 = z_2 \Longleftrightarrow = \alpha_1 = \alpha_2 \wedge \beta_1 = \beta_2$

(b) $(\alpha_1 + \beta_1) \pm (\alpha_2 + i\beta_2) = \alpha_1 \pm \alpha_2 + i(\beta_1 \pm \beta_2)$

(c) $(\alpha_1 + i\beta_1) \cdot (\alpha_2 + i\beta_2) = (\alpha_1\alpha_2 - \beta_1\beta_2) + i(\alpha_1\beta_2 + \beta_1\alpha_2)$

(d) $\frac{\alpha_1+i\beta_1}{\alpha_2+i\beta_2} = \frac{(\alpha_1+i\beta_1)(\alpha_2-i\beta_2)}{(\alpha_2+i\beta_2)(\alpha_2-i\beta_2)} = \frac{\alpha_1\alpha_2+\beta_1\beta_2}{\alpha_2^2+\beta_2^2} + i\frac{\beta_1\alpha_2-\alpha_1\beta_1}{\alpha_2^2+\beta_2^2}$

Die Paare $z_1 = \alpha + i\beta,\ z_2 = \alpha - i\beta$ heißen (zueinander) **konjugiert komplexe Zahlen** (Schreibweise: $z_2 = \bar{z}_1$).

Für ein Paar zueinander konjugiert komplexer Zahlen $z = \alpha+i\beta,\ \bar{z} = \alpha-i\beta$ gilt $z + \bar{z} = 2\alpha,\ z \cdot \bar{z} = \alpha^2 + \beta^2$. Daraus ergibt sich
$(x-z)(x-\bar{z}) = x^2 - 2\alpha x + \alpha^2 + \beta^2$. Setzt man nun $-2\alpha =: p$ und $\alpha^2+\beta^2 =: q$ sowie (ohne Beschränkung der Allgemeinheit) $\beta > 0$, so erhält man
$\alpha = -\frac{p}{2}, \beta = \sqrt{q - \alpha^2} = \sqrt{q - \left(\frac{p}{2}\right)^2}$ mit $q - \left(\frac{p}{2}\right)^2 > 0$.
$z = \alpha + i\beta$ und $\bar{z}$ mit
$\alpha = -\frac{p}{2}$ und $\beta = \sqrt{q - \left(\frac{p}{2}\right)^2}$
sind daher **Lösungen der Gleichung $\boldsymbol{x^2 + px + q = 0}$,** $p \in R,\ q \in R$, falls die sogenannte Diskriminante $D := \frac{p^2}{2} - q$ negativ ist.

Beispiel: Gegeben sei die Gleichung $\lambda^2 + 2\lambda + 3 = 0$. Es gilt
$D = (-1)^2 - 3 = -2 < 0$. Also lauten die beiden Nullstellen
$\lambda_{1,2} = -1 \pm i\sqrt{-D} = -1 \pm i\sqrt{2}$.

Ein Polynom n-ten Grades $x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_o$ läßt sich stets in folgender Weise darstellen:

$$P_n(x) = \prod_{i=1}^{k}(x - x_i) \cdot \prod_{j=k+1}^{l}(x^2 + p_jx + q_j),$$

wobei $2(l - k) + k = n,\ x_i, p_j, q_j \in R$, und die quadratischen Ausdrücke

$x^2 + p_j x + q_j,\ j = k+1, \ldots, l$, keine reellen Nullstellen besitzen.

2.3.4 Fundamentalsystem einer homogenen linearen Differentialgleichung bei komplexen Nullstellen des charakteristischen Polynoms

Die Ableitung einer komplexwertigen Funktion einer reellen Variablen ist durch $(z(x))' = (u(x) + iv(x))' = u'(x) + iv'(x)$ definiert.

Es sei $\lambda = \alpha + i\beta$. Dann gilt

$$
\begin{aligned}
(e^{\lambda x})' = \tfrac{d}{dx}(e^{\lambda x}) &= [e^{\alpha x}(\cos(\beta x) + i\sin(\beta x))]' = (e^{\alpha x}\cos(\beta x))' + i(e^{\alpha x}\sin(\beta x))' \\
&= e^{\alpha x}(-\beta\sin(\beta x) + i\beta\cos(\beta x)) + \alpha e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)) \\
&= e^{\alpha x}[\sin(\beta x)\underbrace{(i\alpha - \beta)}_{i(\alpha+i\beta)} + \cos(\beta x)(\alpha + i\beta)] \\
&= (\alpha + i\beta)e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)) \\
&= \lambda e^{\lambda x},
\end{aligned}
$$

d.h., die Rechenregel gilt wie im Reellen.

Bemerkung: Für e^{λ} mit $\lambda = \alpha + i\beta$ gelten auch die aus dem Reellen bekannten Potenzgesetze; insbesondere ist $(e^{\lambda})^{\frac{p}{q}} = e^{\frac{p}{q}\lambda}$.

Ist $e^{\lambda x}$, λ komplex, Lösung einer homogenen linearen Differentialgleichung mit konstanten Koeffizienten , so sind sowohl Realteil als auch Imaginärteil Lösungen. Darüber hinaus sind Realteil und Imaginärteil linear unabhängig.

Satz:

a) Besitzt die charakteristische Gleichung (4) einer homogenen linearen Differentialgleichung mit konstanten Koeffizienten auch komplexe Lösungen $\alpha \pm i\beta$, so sind die beiden Funktionen
$\varphi_1(x) = e^{\alpha x}\cos(\beta x),\ \varphi_2(x) = e^{\alpha x}\sin(\beta x)$
linear unbhängige Lösungen der Differentialgleichung (1).

b) Tritt ein Paar $\alpha \pm i\beta$ k-fach auf, so werden ihm die Funktionen

$$
\begin{aligned}
&\varphi_1(x) = e^{\alpha x}\cos(\beta x),\ \varphi_2(x) = e^{\alpha x}\sin(\beta x), \\
&\varphi_3(x) = xe^{\alpha x}\cos(\beta x),\ \varphi_4(x) = xe^{\alpha x}\sin(\beta x), \\
&\vdots \\
&\varphi_{2k-1}(x) = x^{k-1}e^{\alpha x}\cos(\beta x),\ \varphi_{2k}(x) = x^{k-1}e^{\alpha x}\sin(\beta x)
\end{aligned}
$$

zugeordnet.

Beispiel: Die Nullstellen des charakteristischen Polynoms seien
$\lambda_{1,2} = 0,\ \lambda_3 = 4,\ \lambda_{4,5} = -1 + 3i,\ \lambda_{6,7} = -1 - 3i.$

Dann ergibt sich das Fundamentalsystem

$1,\ x,\ e^{4x},\ e^{-x}\cos(3x),\ e^{-x}\sin(3x);\ xe^{-x}\cos(3x),\ xe^{-x}\sin(3x).$

2.3.5 Bestimmung einer speziellen Lösung einer inhomogenen linearen Differentialgleichung mit konstanten Koeffizienten

Bei linearen Differentialgleichungen mit **konstanten Koeffizienten** und **spezieller Gestalt der Störfunktion** $q(x)$ kann man y_s in Form eines (an $q(x)$ orientierten) Ansatzes suchen:

$q(x)$ besitze die Gestalt

$\boldsymbol{q(x) = e^{\ x}[Q_m(x)\cos(\beta x) + \tilde{Q}_r(x)\sin(\beta x)]}, \qquad (5)$

wobei α und β reelle Parameter, Q_m ein Polynom vom Grad $m \geq 0$ und $\tilde{Q}_r$ ein Polynom vom Grad $r \geq 0$ bezeichnen.

Wichtige Spezialfälle von (5) sind die folgenden:

a) $\alpha = 0,\ \beta = 0:\ q(x) = Q_m(x)$,

b) $\beta = 0:\ q(x) = e^{\alpha x}Q_m(x)$,

c) $\alpha = 0,\ Q_m(x) = 0:\ q(x) = \tilde{Q}_r(x)\sin(\beta x)$.

Für die weiteren Betrachtungen seien α, β aus (5) festgelegt. Weiterhin sei $u := \max\{m, r\}$.

Fall 1: Ist $\alpha + i\beta$ keine Nullstelle des charakteristischen Polynoms der homogenen Differentialgleichung, wählt man den

Ansatz $\boldsymbol{y_s = e^{\ x}[R_u(x)\cos(\beta x) + \tilde{R}_u(x)\sin(\beta x)]}$,

wobei $R_u, \tilde{R}_u$ Polynome vom Grade u mit noch zu bestimmenden Koeffizienten bezeichnen.

Fall 2: Ist $\alpha + i\beta$ k-fache Nullstelle des charakteristischen Polynoms der homogenen Differentialgleichung (d.h., $q(x)$ ist Lösung der homogenen Differentialgleichung, Resonanzfall) wählt man den

Ansatz $\boldsymbol{y_s = x^k e^{\ x}[R_u(x)\cos(\beta x) + \tilde{R}_u(x)\sin(\beta x)]}$.

Bemerkung: Gilt $q(x) = q_1(x) + \ldots + q_l(x)$, wobei die q_i in der Form (5) darstellbar sind, so ist der Ansatz $y_s = y_s^{(1)} + \ldots + y_s^{(l)}$ möglich, wobei $y_s^{(i)}$ einen zu q_i gehörenden Ansatz bezeichnet.

Beispiele:

1.) Gegeben sei die Differentialgleichung $y''' + 2y'' = 24x^2 - 6$.

a) Bestimmung der allgemeinen Lösung y_h der homogenen Differentialgleichung:
Aus dem Ansatz $y = e^{\lambda x}$ ergibt sich die charakteristische Gleichung $\lambda^3 + 2\lambda^2 = \lambda^2(\lambda + 2) = 0$ mit den Lösungen
$\lambda_{1,2} = 0,\ \lambda_3 = -2$.
Daraus folgt $y_h = c_1 + c_2 x + c_3 e^{-2x}$.

b) Bestimmung einer speziellen Lösung der inhomogenen Differentialgleichung:
$q(x) = 24x^2 - 6$ lässt sich mit $\alpha = 0,\ \beta = 0,\ Q_m(x) = 24x^2 - 6$ in der Form (5) schreiben. Da 0 zweifache Nullstelle des charakteristischen Polynoms der homogenen Differentialgleichung ist, ergibt sich der folgende Ansatz:

$$y_s = x^2(Ax^2 + Bx + C) = Ax^4 + Bx^3 + Cx^2.$$

Einsetzen in die Differentialgleichung liefert

$$y' = 4Ax^3 + 3Bx^2 + 2Cx,$$
$$y'' = 12Ax^2 + 6Bx + 2C,$$
$$y''' = 24Ax + 6B,$$

$$24Ax + 6B + 2(12Ax^2 + 6Bx + 2C) = 24x^2 - 6.$$

Da die Funktionen $1,\ x,\ x^2$ linear unabhängig sind, kann man $A,\ B,\ C$ durch **Koeffizientenvergleich** ermitteln.
$x^2:\ 24A = 24 \Rightarrow A = 1,$
$x^1:\ 24A + 12B = 0 \Rightarrow B = -2,$
$x^o:\ 6B + 4C = -6 \Rightarrow C = \frac{3}{2}.$

Man erhält $y_s = x^4 - 2x^3 + \frac{3}{2}x^2$.

c) Die allgemeine Lösung der Differentialgleichung lautet nun
$y_a = c_1 + c_2 x + c_3 e^{-2x} + x^4 - 2x^3 + \frac{3}{2}x^2$.

2.) Gegeben sei die Differentialgleichung
$y'' + y = x^2 + 3e^{7x} + \cos x + \cos(4x) + (3x - 1)\sin(4x)$.
Die charakteristische Gleichung der homogenen Differentialgleichung lautet $\lambda^2 + 1 = 0$. Sie besitzt die Lösungen $\lambda_{1,2} = \pm i$.

Eine spezielle Lösung für die inhomogene Differentialgleichung kann folglich mit dem Ansatz

$$\begin{aligned} y_s = {} & (A_1x^2 + A_2x + A_3) + Be^{7x} + x(C_1\cos x + C_2\sin x) \\ & + [(D_1x + D_2)\cos(4x) + (D_3x + D_4)\sin(4x)] \end{aligned}$$

gesucht werden.

Bemerkung: Für lineare Differentialgleichungen höherer Ordnung mit beliebigen (nicht notwendigen konstanten) Koeffizienten des homogenen Teils und beliebigem q ist zur Bestimmung von y_s die Methode der **Variation der Konstanten** anwendbar.

y_s wird mit dem **Ansatz**

$y_s = C_1(x)\varphi_1(x) + \ldots + C_n(x)\varphi_n(x)$

gesucht, wobei $\varphi_1, \ldots, \varphi_n$ ein Fundamentalsystem der homogenen linearen Differentialgleichung bezeichnet.

Zur Bestimmung der Funktionen $C_i,\ i = 1, \ldots, n$, reicht die Differentialgleichung nicht aus, es können $n-1$ weitere Forderungen gestellt werden (so dass die Rechnung "durchführbar" wird).

Beispiel: Gegeben sei die Differentialgleichung $y'' + 4y = \frac{1}{\sin(2x)},\ 0 < x < \frac{\pi}{2}$.

a) Die homogene Differentialgleichung führt zu der charakteristischen Gleichung $\lambda^2 + 4 = 0$ mit den Lösungen $\lambda_{1,2} \pm 2i$. Daraus ergibt sich
$y_h = c_1 \cos(2x) + c_2 \sin(2x)$.

b) Eine spezielle Lösung der inhomogenen Differentialgleichung wird mit dem folgenden Ansatz gesucht:
$y_s = C_1(x)\cos(2x) + C_2(x)\sin(2x)$.
Die erste Ableitung liefert
$y_s' = \underbrace{C_1'(x)\cos(2x) + C_2'(x)\sin(2x)}_{Zusatzforderung: `=0'} + C_1(x)(-2\sin(2x)) + C_2(x)(2\cos(2x))$.

Unter Beachtung von $C_1'(x)\cos(2x) + C_2'(x)\sin(2x) = 0$ ergibt sich für die zweite Ableitung

$y_s'' = C_1'(x)(-2\sin(2x)) + C_2'(x)(2\cos(2x)) + C_1(x)(-4\cos(2x)) + C_2(x)(-4\sin(2x))$.

Die zweite Ableitung wird nun in die Differentialgleichung eingesetzt:
$C_1'(x)(-2\sin(2x)) + C_2'(x)(2\cos(2x)) + C_1(x)(-4\cos(2x)) + C_2(x)(-4\sin(2x)) +$
$4(C_1(x)\cos(2x) + C_2(x)\sin(2x))$
$= C_1'(x)(-2\sin(2x)) + C_2'(x)(2\cos(2x)) = \dfrac{1}{\sin(2x)}$.

Zur Bestimmung von C_1 und C_2 hat man zwei Gleichungen, nämlich
$C_1'(x)\cos(2x) + C_2'(x)\sin(2x) = 0$
$C_1'(x)(-2\sin(2x)) + C_2'(x)(2\cos(2x)) = \dfrac{1}{\sin(2x)}$.

Es handelt sich hierbei um ein lineares Gleichungssystem für C_1' und C_2', dessen Koeffizientendeterminante gerade die Wronskische Determinante des Fundamentalsystems der homogenen Differentialgleichung ist. Das Gleichungssystem kann daher stets mit der Cramerschen Regel gelöst werden:

$$C_1' = \frac{\begin{vmatrix} 0 & \sin(2x) \\ \frac{1}{\sin(2x)} & 2\cos(2x) \end{vmatrix}}{\begin{vmatrix} \cos(2x) & \sin(2x) \\ -2\sin(2x) & 2\cos(2x) \end{vmatrix}} = -\frac{1}{2}\ ,$$

$$C_2' = \frac{\begin{vmatrix} \cos(2x) & 0 \\ -2\sin(2x) & \frac{1}{\sin(2x)} \end{vmatrix}}{2} = \frac{1}{2}\cot(2x).$$

Durch Integration erhält man schließlich $C_1(x) = -\frac{x}{2};\ C_2(x) = +\frac{1}{4}\ln(\sin(2x))$.

Die allgemeine Lösung der Differentialgleichung lautet somit
$y = c_1\cos(2x) + c_2\sin(2x) - \frac{x}{2}\cos(2x) + \frac{1}{4}\ln(\sin(2x)) \cdot \sin(2x)$.

2.4 Lineare Differentialgleichungssysteme

Es werden nur lineare Systeme 1. Ordnung mit konstanten Koeffizienten betrachtet.

Ausgangspunkt der folgenden Betrachtungen ist die allgemeine Form:

$$y_1' = a_{11}y_1 + \ldots + a_{1n}y_n + q_1(x)$$
$$\vdots$$
$$y_n' = a_{n1}y_1 + \ldots + a_{nn}y_n + q_n(x).$$

In Vektorschreibweise hat das System die Gestalt

$$\vec{y}' = A\vec{y} + \vec{q} \text{ mit } \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \vec{y}' = \begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix}, A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}, \vec{q} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}.$$

a) Zunächst werden **homogene lineare Differentialgleichungssysteme** $\vec{y}' = A\vec{y}$ betrachtet.

Die allgemeine Gestalt der Lösung lautet

$$\vec{y}_h = c_1\vec{y}_1(x) + \ldots + c_n\vec{y}_n(x),$$

wobei $\vec{y}_1(x), \ldots, \vec{y}_n(x)$ linear unabhängige Lösungsvektoren bezeichnen.
Die Lösungsvektoren werden wieder mit einem Ansatz gesucht:
Der **Ansatz** zur Bestimmung der $\vec{y}_i$ lautet:

$$\vec{y} = \vec{a}e^{\lambda x},\ \vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \neq \vec{0},\ \lambda \in R^1.$$

Einsetzen des Ansatzes in das Differentialgleichungssystem ergibt

$$\lambda\vec{a}e^{\lambda x} - A\vec{a}e^{\lambda x} = 0 \text{ und weiter } \lambda\vec{a} - A\vec{a} = (A - \lambda E)\vec{a} = 0.$$

Damit ist die Aufgabe auf ein Eigenwertproblem zurückgeführt worden. Die gesuchten Werte λ sind gerade die Eigenwerte der Matrix A und die Koeffizientenvektoren $\vec{a}$ die zugehörigen Eigenvektoren. Allerdings müssen jetzt auch komplexwertige Eigenwerte und Eigenvektoren in die Betrachtungen einbezogen werden. Sie können analog zu dem in Teil I, 3.6.1., beschriebenen Vorgehen ermittelt werden.

Die Eigenwerte λ_i werden aus der Gleichung $\det(A - \lambda E) = 0$ ermittelt. Die zugehörigen Eigenvektoren $\vec{a}_i$ erhält man dann als Lösung der linearen Gleichungssysteme $(A - \lambda_i E)\vec{a}_i = 0$.

Beispiel: Gegeben sei das Differentialgleichungssystem
$$\begin{array}{l} y_1' = -3y_1 - 2y_2 \\ y_2' = 2y_1 + 2y_2 \end{array} .$$
Mit $\vec{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix}$ ergibt sich die Gestalt $\vec{y}' = \underbrace{\begin{pmatrix} -3 & -2 \\ 2 & 2 \end{pmatrix}}_{A} \vec{y}$.

Der Ansatz $\vec{y} = \vec{a}e^{\lambda x}$ mit $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ führt zu dem Eigenwertproblem $(A - \lambda E)\vec{a} = 0$. Die Eigenwerte ergeben sich aus

$$\begin{vmatrix} -3-\lambda & -2 \\ 2 & 2-\lambda \end{vmatrix} = \lambda^2 + \lambda - 2 = 0$$

zu $\lambda_1 = 1$ und $\lambda_2 = -2$.

Nun werden für die beiden Eigenwerte die zugehörigen Eigenvektoren $\vec{a}_i$ bestimmt:

Betrachtung von $\lambda_1 = 1$: Aus $(A - \lambda_1 E)\vec{a}_1 = \begin{pmatrix} -4 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ folgt $a_2 = -2a_1$. Eine Variable ist frei wählbar. Mit der speziellen Wahl $a_1 = 1$ ergibt sich $\vec{a}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

(Hier und im Folgenden wird auf eine zusätzliche Indizierung der Hilfsvariablen $a_1,\ a_2$ mit der Nummer des Eigenwertes verzichtet.)

Betrachtung von $\lambda_2 = -2$: Aus $(A - \lambda_2 E)\vec{a}_2 = \begin{pmatrix} -1 & -2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ folgt $a_1 = -2a_2$. Mit der speziellen Wahl $a_2 = 1$ ergibt sich $\vec{a}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

Die beiden Vektorfunktionen $y_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^x$ und $y_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{-2x}$ sind linear unabhängige Lösungen des Differentialgleichungssystems; sie bilden also ein Fundamentalsystem. Die allgemeine Lösung des homogenen Systems lautet somit

$$y_h = c_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^x + c_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{-2x}.$$

Allgemein können nach der Lösung des Eigenwertproblems folgende **Fälle** auftreten:

Fall 1: Alle Eigenwerte $\lambda_i,\ i = 1, \ldots, n$, sind reell und verschieden. Dann wird mit den zu $\lambda_i,\ i = 1, \ldots, n$, gehörigen Eigenvektoren $\vec{a}_i$ die allgemeine Lösung in der Gestalt

$y_h = c_1\vec{a}_1 e^{\lambda_1 x} + \ldots + c_n \vec{a}_n e^{\lambda_n x}$

aufgebaut.

Fall 2: Es treten Paare konjugiert komplexer Eigenwerte $\alpha \pm i\beta$ auf. Alle Paare sind verschieden.
Dann wird zunächst (für jedes konjugiert komplexe Paar) der (komplexe) Eigenvektor $\vec{a}$ zu $\alpha + i\beta$ bestimmt.
(Das reicht aus, denn ist $\vec{a} = \vec{b} + i\vec{c}$ komplexer Eigenvektor zu $\alpha + i\beta$, so ist $\vec{b} - i\vec{c}$ Eigenvektor zu $\alpha - i\beta$.)
Zu jedem Paar werden dann die beiden linear unabhängigen Lösungen

$y_1 := Re(\vec{a}[e^{\alpha x}\cos(\beta x) + ie^{\alpha x}\sin(\beta x)])$ und
$y_2 := Im(\vec{a}[e^{\alpha x}\cos(\beta x) + ie^{\alpha x}\sin(\beta x)])$

gebildet.

Fall 3: Es treten mehrfache Eigenwerte auf, d.h., die algebraische Vielfachheit (für mindestens einen Eigenvektor) ist größer als 1. (Man beachte, dass die geometrische Vielfachheit eines Eigenwertes, die der maximalen Anzahl der zum betrachteten Eigenwert gehörigen linear unabhängigen Eigenvektoren entspricht, kleiner als die algebraische Vielfachheit sein kann.)

Existieren zu einem k-fachen Eigenwert λ genau k linear unabhängige Eigenvektoren $\vec{a}_1, \ldots, \vec{a}_k$, werden die k Lösungen
$\vec{y}_1 = \vec{a}_1 e^{\lambda x},\ \ldots,\ \vec{y}_k = \vec{a}_k e^{\lambda x}$
zum Aufbau des Fundamentalsystems verwendet.
Anderenfalls müssen fehlende Fundamentallösungen in der Gestalt $p_v(x)e^{\lambda x}$ mit einem Vektorpolynom $p_v(x) = \vec{c}_o + \vec{c}_1 x + \ldots + \vec{c}_v t^v$ gesucht werden.

In jedem Fall können zu einem k-fachen Eigenwert λ k linear unabhängige Lösungen $p_o(x)e^{\lambda x},\ \ldots,\ p_{k-1}(x)e^{\lambda x}$ des Differentialgleichungssystems gefunden werden.

b) Nunmehr werden **inhomogene lineare Differentialgleichungssysteme** der Gestalt $\vec{y}' = A\vec{y} + \vec{q}$ betrachtet.

Die allgemeine Lösung hat die Form $\vec{y}_a = \vec{y}_h + \vec{y}_s$, wobei $\vec{y}_s$ eine spezielle

Lösung des inhomogenen Systems bezeichnet.

Zur Bestimmung von $\vec{y}_s$ sind die

• Variation der Konstanten oder

• spezielle Ansätze bei konstanten Koeffizienten und spezieller Gestalt von $\vec{q}$ möglich.

Beispiel: Gegeben sei das Differentialgleichungssystem
$\vec{y}' = \begin{pmatrix} -3 & -2 \\ 2 & 2 \end{pmatrix} \vec{y} + \begin{pmatrix} 0 \\ 2x \end{pmatrix}$.

α) Aus früheren Berechnungen ist die allgemeine Lösung des homogenen Systems $y_h = c_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^x + c_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{-2x}$ bekannt.

β) Eine spezielle Lösung des inhomogenen Systems wird mit dem Ansatz $\vec{y}_s = \begin{pmatrix} A_o + A_1 x \\ B_o + B_1 x \end{pmatrix} = \begin{pmatrix} y_{1,s}(x) \\ y_{2,s}(x) \end{pmatrix}$ gesucht.
(Da 0 kein Eigenwert von A ist, sind keine weiteren Zusätze nötig.)

Einsetzen in das Differentialgleichungssystem liefert die beiden Gleichungen
$A_1 + 3(A_o + A_1 x) + 2(B_o + B_1 x) = 0,$
$B_1 - 2(A_o + A_1 x) - 2(B_o + B_1 x) = 2x.$

Durch Koeffizientenvergleich erhält man dann folgende Bestimmungsgleichungen für die Konstanten:

$$\begin{array}{lll} \text{1. Gleichung: } x^o: & A_1 + 3A_o + 2B_o & = 0 \\ \phantom{\text{1. Gleichung: }} x^1: & 3A_1 + 2B_1 & = 0 \\ \text{2. Gleichung: } x^o: & B_1 - 2A_o - 2B_o & = 0 \\ \phantom{\text{2. Gleichung: }} x^1: & -2A_1 - 2B_1 & = 2. \end{array}$$

Das entstandene lineare Gleichunssystem besitzt die Lösungen
$A_o = 1,\ A_1 = 2,\ B_o = -\frac{5}{2},\ B_1 = -3.$

Daraus bildet man die speziellen Lösungen des Differentialgleichungssystems $y_{1,} = 1+2x,\ y_{2,s} = -\frac{5}{2}-3x$. In Vektorform lautet die allgemeine Lösung schließlich $\vec{y} = c_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^x + c_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{-2x} + \begin{pmatrix} 1+2x \\ -\frac{5}{2} - 3x \end{pmatrix}$.

Bemerkung: Ein lineares Differentialgleichungssystem 1. Ordnung kann im Allgemeinen auf eine lineare Differentialgleichung höherer Ordnung zurückgeführt werden. (Gelegentlich entstehen mehrere lineare Differentialgleichungen höherer Ordnung, die unabhängig voneinander gelöst werden können.)

Beispiel: Gegeben sei das bereits gelöste Differentialgleichungssystem
$y_1' = -3y_1 - 2y_2$
$y_2' = 2y_1 + 2y_2 + 2x.$

Durch Ableiten der ersten Gleichung und Ausnutzen der zweiten Gleichung erhält man $y_1'' = -3y_1' - 2y_2' = -3y_1' - 2(2y_1 + 2y_2 + 2x) = -3y_1' - 4y_1 - 4y_2 - 4x.$
y_2 wird mit Hilfe der ersten Gleichung ersetzt. Es ergibt sich die Differentialgleichung
$y_1'' = -3y_1' - 4y_1 + 2y_1' + 6y_1 - 4x = -y_1' + 2y_1 - 4x.$

Diese Differentialgleichung hat die Lösung
$y_1 = c_1e^x + c_2e^{-2x} + 1 + 2x.$

Aus der ersten Gleichung folgt nun
$y_2 = -\frac{3}{2}y_1 - \frac{1}{2}y_1' = -\frac{3}{2}(c_1e^x + c_2e^{-2x} + 1 + 2x) - \frac{1}{2}(c_1e^x - 2c_2e^{-2x} + 2)$
$= -2c_1e^x - \frac{1}{2}c_2e^{-2x} - \frac{5}{2} - 3x.$

Zusammengefasst lautet die Lösung
$$y = c_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^x + c_2 \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} e^{-2x} + \begin{pmatrix} 1 + 2x \\ -\frac{5}{2} - 3x \end{pmatrix}.$$

2.5 Differenzengleichungen

2.5.1 Lineare Differenzengleichungen 1. Ordnung

Differenzengleichungen treten auf in Modellen, in denen Prozesse zu diskreten, gleichabständigen Zeitpunkten, z. B. $n = 0, 1, 2 \ldots$, betrachtet werden.

Beispiel: Wachstumsmodell für das Volkseinkommen nach Boulding

Im Folgenden bezeichnen
$y(n)$ das Volkseinkommen in der Perode n,
$c(n)$ den Konsum in der Periode n,
$i(n)$ die Investitionen in der Periode n.

Es wird angenommen, dass die betrachteten Größen den folgenden Beziehungen genügen:
$y(n) = c(n) + i(n),$
$c(n) = \alpha + \beta y(n), \qquad \alpha \geq 0,\ 0 < \beta < 1,$
$y(n+1) - y(n) = \gamma i(n), \quad \gamma > 0.$

Aus diesen Modellannahmen ergibt sich
$\Delta y(n) := y(n+1) - y(n) = \gamma i(n) = \gamma(y(n) - c(n)) = \gamma(y(n) - \beta y(n) - \alpha)$
$= \gamma(1 - \beta)y(n) - \alpha\gamma.$

Die $y(n),\ n \in N$, können rekursiv aus den vorhergehenden Werten berechnet werden, wenn ein "Anfangswert" $y(0)$ bekannt ist.

Im Folgenden wird die Schreibweise y_n anstelle von $y(n)$ verwendet ($n \in N$).

Definition: Es seien $(a_n)_{n\in N_o}$ und $(b_n)_{n\in N_o}$ Folgen reeller Zahlen. Eine Vorschrift der Form

$$y_{n+1} = a_n y_n + b_n \qquad (1)$$

heißt **lineare Differenzengleichung 1. Ordnung**.

Satz: Die lineare Differenzengleichung (1) besitzt die allgemeine Lösung

$$y_n = c \prod_{k=0}^{n-1} a_k + \sum_{k=0}^{n-2} b_k \prod_{l=k+1}^{n-1} a_l + b_{n-1} \quad (n \geq 2) \qquad (2),$$

$y_1 = ca_o + b_o,$

$y_o = c,\ c \in R$ beliebig.

Beweis:

1.) Für $n = 2$ ist die Beziehung (2) richtig, denn es gilt

$$y_2 = a_1 y_1 + b_1 = a_1(ca_o + b_o) + b_1 = ca_o a_1 + b_o a_1 + b_1$$
$$= c \prod_{k=0}^{1} a_k + \sum_{k=0}^{0} b_k \prod_{l=k+1}^{1} a_l + b_1.$$

2.) Wir zeigen, dass aus der Gültigkeit von (2) für y_n auch die Gültigkeit für y_{n+1} folgt. Damit kann dann aus der Gültigkeit von (2) für $n = 2$ auf die Gültigkeit für $n = 3$ geschlossen werden, woraus wiederum die Gültigkeit für $n = 4$ folgt usw.

Nach Beziehung (1) gilt

$$y_{n+1} = a_n y_n + b_n = a_n \left(c \prod_{k=0}^{n-1} a_k + \sum_{k=0}^{n-2} b_k \prod_{l=k+1}^{n-1} a_l + b_{n-1} \right) + b_n$$
$$= c \prod_{k=0}^{n} a_l + \sum_{k=0}^{n-2} b_k \prod_{l=k+1}^{n} a_l + a_n b_{n-1} + b_n$$
$$= c \prod_{k=0}^{n} a_l + \sum_{k=0}^{n-1} b_k \prod_{l=k+1}^{n} a_l + b_n,$$

d.h., die gewünschte Beziehung ist erfüllt.

Bemerkung: Im Fall $a_n = a,\ b_n = b\ \forall n \in N_o$ ergibt sich

$$y_n = ca^n + \sum_{k=0}^{n-2} ba^{n-k-1} + b = ca^n + b \sum_{k=0}^{n-1} a^k = \begin{cases} ca^n + b\frac{a^n-1}{a-1}, & \text{falls } a \neq 1, \\ c + bn, & \text{falls } a = 1. \end{cases}$$

2.5.2 Der Differenzenoperator

$y(n)$ kann als Funktion $y|\ N_o \to R$ aufgefasst werden. Für diese Funktion wird dann durch $\Delta(y(n)) := y(n+1) - y(n)$ eine neue Funktion $\Delta y|\ N_o \to R$

gebildet. Sie wird (erste) Differenz von y genannt.

Die k-te Differenz kann dann rekursiv definiert werden:

$\Delta^k(y_n) := \Delta(\Delta^{k-1}(y(n)))$ mit $\Delta^1(y(n)) = \Delta(y(n))$.

Speziell für $k = 2$ und $k = 3$ ergibt sich

$\Delta^2(y(n)) = \Delta[y(n+1) - y(n)] = y(n+2) - y(n+1) - y(n+1) + y(n)$
$= y(n+2) - 2y(n+1) + y(n)$, sowie

$\Delta^3 y(n) = \Delta[y(n+2) - 2y(n+1) + y(n)] = y(n+3) - 2y(n+2) + y(n+1) -$
$y(n+2) + 2y(n+1) - y(n)$
$= y(n+3) - 3y(n+2) + 3y(n+1) - y(n)$.

Auch hier wird wieder die abkürzende Schreibweise $y_n = y(n)$ verwendet.

Eine Gleichung, in der neben n auch Differenzen bis zur Ordnung r vorkommen (Differenzengleichung r-ter Ordnung), lässt sich folglich auch als Gleichung darstellen, in der neben n die Größen $y_n,\ y_{n+1}, \ldots, y_{n+r}$ auftreten.

2.5.3 Lineare Differenzengleichungen höherer Ordnung

Im Folgenden werden nur lineare Differenzengleichungen mit konstanten Koeffizienten betrachtet.

Definition: Eine Vorschrift der Gestalt

$$y_{n+r} + a_{r-1}y_{n+r-1} + \ldots + a_1 y_{n+1} + a_o y_n = q_n \qquad (3)$$

mit $a_o \neq 0$ heißt **lineare Differenzengleichung der Ordnung r mit konstanten Koeffizienten**. Im Fall $q_n = 0\ \forall n \in N_o$ heißt die Differenzengleichung **homogen**, anderenfalls **inhomogen**.

Satz: Die Differenzengleichung (3) mit $a_o \neq 0$ hat eine eindeutige Lösung $y_n,\ n \in N_o$, wenn r aufeinander folgende (Anfangs-) Werte y_k vorgegeben sind.

In diesem Fall ist eine rekursive Berechnung, ausgehend von den Anfangswerten, möglich. (Beispiel: $y_{k+2} - y_k = 0, y_o = 2, y_1 = 4$)

Die allgemeine Lösung $y_n^{(a)}$ einer linearen Differenzengleichung r-ter Ordnung lässt sich stets in der folgenden Form schreiben:

$y_n^{(a)} = y_n^{(h)} + y_n^{(s)}$,
wobei $y_n^{(h)}$ ... die allgemeine Lösung der homogenen Differenzengleichung und
$y_n^{(S)}$ ... eine spezielle Lösung der inhomogenen Differenzengleichung
bezeichnen.

Weiter gilt
$y_n^{(h)} = c_1 y_n^{(1)} + \ldots + c_r y_n^{(r)}$, wobei $y_n^{(1)}, \ldots, y_n^{(r)}$ linear unabhängige Lösungen

der Differenzengleichung darstellen (Fundamentalsystem).

Die Funktionen $y_n^{(1)}, \ldots, y_n^{(r)}$ heißen **linear unabhängig**, wenn aus der Gültigkeit der Beziehung
$c_1 y_n^{(1)} + \ldots + c_r y_n^{(r)} = 0 \; \forall n \geq 0$ die Beziehung
$c_1 = \ldots = c_r = 0$ folgt.

Zur Überprüfung der linearen Unabhängigkeit kann die sogenannte Casorati-Determinante

$$W(y_n^{(1)}, \ldots, y_n^{(r)}) := \begin{vmatrix} y_o^{(1)} & y_o^{(2)} & \ldots & y_o^{(r)} \\ y_1^{(1)} & y_1^{(2)} & \ldots & y_1^{(r)} \\ \vdots & \vdots & & \vdots \\ y_{r-1}^{(1)} & y_{r-1}^{(2)} & \ldots & y_{r-1}^{(r)} \end{vmatrix}$$

herangezogen werden. Es gelten die folgenden Aussagen:

1.) Sind die Funktionen $y_n^{(1)}, \ldots, y_n^{(r)}$ linear abhängig, so ist $W(y_n^{(1)}, \ldots, y_n^{(r)}) = 0$.

2.) Für Lösungen einer homogenen linearen Differenzengleichung gilt auch die Umkehrung, d.h., die Funktionen $y_n^{(1)}, \ldots, y_n^{(r)}$ sind genau dann linear unabhängig, wenn $W(y_n^{(1)}, \ldots, y_n^{(r)}) \neq 0$ gilt.

Beispiel: Betrachtet werden die Funktionen
$y_n^{(1)} = 1, \; y_n^{(2)} = (-1)^n, \; y_n^{(3)} = 5^n$. Es ergibt sich

$$W(y_n^{(1)}, y_n^{(2)}, y_n^{(3)}) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 5 \\ 1 & 1 & 25 \end{vmatrix} \neq 0.$$

Die drei Funktionen sind also linear unabhängig.

Bestimmung der allgemeinen Lösung einer homogenen linearen Differenzengleichung r-ter Ordnung mit konstanten Koeffizienten

Die Funktionen für das Fundamentalsystem werden mit dem
Ansatz $y_n := \lambda^n, \quad \lambda \neq 0,$
gesucht.
Einsetzen in den homogenen Teil der Differenzengleichung (3) liefert
$\lambda^{n+r} + a_{r-1}\lambda^{n+r-1} + \ldots + a_1\lambda^{n+1} + a_o\lambda^n = 0.$

Da $\lambda \neq 0$ vorausgesetzt wurde, kann diese Gleichung durch λ^n dividiert werden, und man erhält die sogenannte **charakteristische Gleichung**
$\lambda^r + a_{r-1}\lambda^{r-1} + \ldots + a_1\lambda^1 + a_o = 0.$

Die Lösungen der charakteristischen Gleichung werden nun zur Zusammenstellung des Fundamentalsystems verwendet. Dabei geht man wie folgt vor:

(a) Für jede reelle Lösung λ_i der charakteristischen Gleichung mit Vielfachheit k_i bildet man die k_i Lösungen

$y_n^{(i,1)} = \lambda_i^n,\ y_n^{(i,2)} = n\lambda_i^n, \ldots,\ y_n^{(i,k_i)} = n^{k_i-1}\lambda_i^n.$

(b) Für jedes Paar konjugiert komplexer Lösungen $\lambda_{1,2}^{(j)} = \alpha \pm i\beta$ der Vielfachheit k_j bildet man die $2k_j$ Lösungen

$y_n^{(j,1)} = r^n \cos(n\varphi),\ y_n^{(j,2)} = r^n \sin(n\varphi),$
$y_n^{(j,3)} = nr^n \cos(n\varphi),\ y_n^{(j,4)} = nr^n \sin(n\varphi),$
$\vdots$
$y_n^{(j,2k_j-1)} = n^{k_j-1} r^n \cos(n\varphi),\ y_n^{(j,2k_j)} = n^{k_j-1} r^n \sin(n\varphi),$

wobei $r = \sqrt{\alpha^2 + \beta^2}$ gilt und φ durch $\cos\varphi = \frac{\alpha}{r},\ \sin\varphi = \frac{\beta}{r}$ (bzw. $\tan\varphi = \frac{\beta}{\alpha}$ bei vorheriger Bestimmung des Quadranten) definiert ist.

Die Gestalt der y_n im Fall komplexer Lösungen des charakteristischen Polynoms resultiert aus der Umformung
$(\alpha + i\beta)^n = (re^{i\varphi})^n = r^n e^{in\varphi} = r^n[\cos(n\varphi) + i\sin(n\varphi)]$

und der Überlegung, dass Realteil und Imaginärteil linear unabhängige Lösungen der homogenen linearen Differenzengleichung ergeben.

Beispiel: Gegeben sei die Differenzengleichung
$y_{n+4} + 4y_{n+3} + 7y_{n+2} + 6y_{n+1} + 2y_n = 0.$
Der Ansatz $y_n = \lambda^n$ liefert die charakteristische Gleichung
$\lambda^4 + 4\lambda^3 + 7\lambda^2 + 6\lambda + 2 = (\lambda+1)^2((\lambda+1)^2 + 1) = 0,$
die die Lösungen $\lambda_{1,2} = -1$ sowie $\lambda_{3,4} = -1 \pm i$ besitzt. Die komplexe Lösung $z = -1 + i$ kann in die trigonometrische Form $z = \sqrt{2} e^{i\frac{3}{4}\pi}$ umgewandelt werden. Die folgenden Funktionen bilden somit ein Fundamentalsystem:
$y_n^{(1)} = (-1)^n,\ y_n^{(2)} = n(-1)^n,\ y_n^{(3)} = 2^{\frac{n}{2}} \cos(n\frac{3}{4}\pi),\ y_n^{(4)} = 2^{\frac{n}{2}} \sin(n\frac{3}{4}\pi),$

und die allgemeine Lösung der Differenzengleichung ergibt sich als Linearkombination der Funktionen aus dem Fundamentalsystem.

Bestimmung einer speziellen Lösung der inhomogenen linearen Differenzengleichung

Zur Bestimmung einer speziellen Lösung gibt es (analog zu den Differentialgleichungen) im Wesentlichen zwei Möglichkeiten: die im Allgemeinen aufwändige Variation der Konstanten und - im Fall konstanter Koeffizienten - die Suche nach geeigneten Lösungen mit einem speziellen Ansatz, der sich nach der Gestalt des 'Störgliedes' q_n richtet.

Im Folgenden sollen nur die speziellen Ansätze näher betrachtet werden. Sie werden nach ähnlichen Prinzipien wie im Fall von Differentialgleichungen aufgestellt.

Wenn q_n die Gestalt $q_n = r^n(Q_m(n)\cos(n\varphi) + \tilde{Q}_{\tilde{m}}(n)\sin(n\varphi))$ (wobei $Q_m(n)$ ein Polynom des Grades m und $\tilde{Q}_{\tilde{m}}(n)$ ein Polynom des Grades $\tilde{m}$ bezeichnen) aufweist, wird der

Ansatz $y_n^{(s)} = r^n(R_u(n)\cos(n\varphi) + \tilde{R}_u(n)\sin(n\varphi))$

mit $u = \max\{m, \tilde{m}\}$ und unbestimmten Koeffizienten der Polynome $R_u(n)$ und $\tilde{R}_u(n)$ verwendet. Ist $re^{i\varphi}$ k-fache Nullstelle des charakteristischen Polynoms der homogenen Differenzengleichung, muss dieser Ansatz noch mit dem Faktor n^k multipliziert werden.

Beispiel: Gegeben sei die Differenzengleichung $y_{n+2} - 2y_{n+1} - 3y_n = n3^n$.

1.) Die charakteristische Gleichung der homogenen Differenzengleichung lautet $\lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1) = 0$. Daraus ergibt sich die allgemeine Lösung der homogenen Differenzengleichung $y_n^{(h)} = c_1(-1)^n + c_2 3^n$.

2.) Zur Bestimmung einer speziellen Lösung der inhomogenen Differenzengleichung wird der Ansatz

$y_n^{(s)} = n3^n(An + B) = 3^n(An^2 + Bn)$

gewählt. Dabei wurde berücksichtigt, dass $3 = 3e^{i0}$ einfache Nullstelle des charakteristischen Polynoms unter 1.) ist.

Einsetzen dieses Ansatzes in die Differenzengleichung liefert

$3^{n+2}(A(n+2)^2 + B(n+2)) - 2 \cdot 3^{n+1}(A(n+1)^2 + B(n+1)) - 3 \cdot 3^n(An^2 + Bn)$
$= n3^n$,

woraus sich nach Division durch 3^n und Koeffizientenvergleich bez. der Funktionen n^2, n und 1 die Lösung $y_n^{(s)} = 3^n(\frac{1}{24}n^2 - \frac{5}{48}n)$ ergibt.

Die folgende Tabelle enthält wichtige Spezialfälle der allgemeinen Formel für einen speziellen Ansatz:

q_n	$\tilde{\lambda}$	Ansatz
$b_0 + b_1 n^1 + \ldots + b_s n^s$	1	$B_0 + B_1 n^1 + \ldots + B_s n^s$
$a^n(b_0 + b_1 n^1 + \ldots + b_s n^s)$	a	$a^n(B_0 + B_1 n^1 + \ldots + B_s n^s)$
$\cos(\gamma n)(b_0 + \ldots + b_s n^s)$ oder $\sin(\gamma n)(b_0 + \ldots + b_s n^s)$	$\cos\gamma + i\sin\gamma$	$(A_0 + \ldots + A_s n^s)\cos(\gamma n)$ $+(B_0 + \ldots + B_s n^s)\sin(\gamma n)$
$a^n\cos(\gamma n)(b_0 + \ldots + b_s n^s)$ oder $a^n\sin(\gamma n)(b_0 + \ldots + b_s n^s)$	$a\cos\gamma + ia\sin\gamma$	$a^n(A_0 + \ldots + A_s n^s)\cos(\gamma n)$ $+a^n(B_0 + \ldots + B_s n^s)\sin(\gamma n)$

Spalte 2 dieser Tabelle gibt an, für welchen Wert $\tilde{\lambda}$ nachgeprüft werden muss, ob er unter den Nullstellen der charakteristischen Gleichung der homogenen Differenzengleichung vorkommt. Ist $\tilde{\lambda}$ k-fache Nullstelle, muss der Ansatz in Spalte 3 mit dem Faktor n^k multipliziert werden.

Bemerkung: Ist a negativ, kann $ae^{i\gamma}$ in der Form $(-a)e^{i(\gamma+\pi)}$ geschrieben werden und hat damit auch die Gestalt $re^{i\varphi}$ mit $r > 0$. Weiterhin gilt dann $y_n^{(s)} = (-a)^n\cos(n(\gamma + \pi)) + (-a)^n\sin(n(\gamma + \pi)) = (-a)^n\cos(n\gamma)\cos(n\pi) +$

$(-a)^n \sin(n\gamma)\cos(n\pi) = (-a)^n(-1)^n \cos(n\gamma) + (-a)^n(-1)^n \sin(n\gamma)$
$= a^n \cos(n\gamma) + a^n \sin(n\gamma)$.

Beispiel: Gegeben sei die Differenzengleichung
$y_{n+2} + 2y_{n+1} + 2y_n = (-1)^n \cos(\frac{3}{4}\pi n)$.

1.) Für die homogene Differenzengleichung erhält man mit dem Ansatz $y_n^{(h)} = \lambda^n$ die charakteristische Gleichung $\lambda^2+2\lambda+2 = 0$ mit den Nullstellen $\lambda_{1,2} = -1 \pm i$. Umwandlung der Lösung $\lambda_1 = -1+i$ in die trigonometrische Form $\sqrt{2}e^{i\frac{3}{4}\pi}$ führt schließlich zur allgemeinen Lösung
$y_n^{(h)} = c_1 2^{\frac{n}{2}} \cos(\frac{3}{4}\pi n) + c_2 2^{\frac{n}{2}} \sin(\frac{3}{4}\pi n)$.

2.) q_n lässt sich in der Gestalt $a^n \cos(\gamma n)(b_0+\ldots+b_s n^s)$ mit $a = -1,\ \gamma = \frac{3}{4}\pi$ und $s = 0$ darstellen.
Da $(-1) \cdot \cos(\frac{3}{4}\pi) + i(-1) \cdot \sin(\frac{3}{4}\pi) = \frac{1}{2}\sqrt{2} - i\frac{1}{2}\sqrt{2}$ keine Nullstelle der charakteristischen Gleichung unter 1.) ist, wird der Ansatz
$y_n^{(s)} = A(-1)^n \cos(\frac{3}{4}\pi n) + B(-1)^n \sin(\frac{3}{4}\pi n)$ verwendet.
Einsetzen in die Differenzengleichung liefert
$A(-1)^{n+2} \cos(\frac{3}{4}\pi(n+2)) + B(-1)^{n+2} \sin(\frac{3}{4}\pi(n+2))$
$+ 2A(-1)^{n+1} \cos(\frac{3}{4}\pi(n+1)) + 2B(-1)^{n+1} \sin(\frac{3}{4}\pi(n+1))$
$+ 2A(-1)^n \cos(\frac{3}{4}\pi n) + 2B(-1)^n \sin(\frac{3}{4}\pi n) = (-1)^n \cos(\frac{3}{4}\pi n)$.

Unter Berücksichtigung der Beziehungen
$\cos(\frac{3}{4}\pi(n+2)) = \cos(\frac{3}{4}\pi n)\cos(\frac{3\pi}{2}) - \sin(\frac{3}{4}\pi n)\sin(\frac{3\pi}{2}) = \sin(\frac{3}{4}\pi n)$ sowie
$\cos(\frac{3}{4}\pi(n+1)) = \frac{1}{2}\sqrt{2}(-\cos(\frac{3}{4}\pi n - \sin(\frac{3}{4}\pi n)$,
$\sin(\frac{3}{4}\pi(n+2)) = -\cos(\frac{3}{4}\pi n$,
$\sin(\frac{3}{4}\pi(n+1)) = \frac{1}{2}\sqrt{2}(-\sin(\frac{3}{4}\pi n) + \cos(\frac{3}{4}\pi n))$

ergibt sich

$A(-1)^n \sin(\frac{3}{4}\pi n) + B(-1)^n(-\cos(\frac{3}{4}\pi n))$
$-2A(-1)^n\frac{1}{2}\sqrt{2}(-\cos(\frac{3}{4}\pi n) - \sin(\frac{3}{4}\pi n)) - 2B(-1)^n\frac{1}{2}\sqrt{2}(-\sin(\frac{3}{4}\pi n) + \cos(\frac{3}{4}\pi n))$
$+ 2A(-1)^n \cos(\frac{3}{4}\pi n) + 2B(-1)^n \sin(\frac{3}{4}\pi n) = (-1)^n \cos(\frac{3}{4}\pi n)$.

Division durch $(-1)^n$ und Koeffizientenvergleich bez. der linear unabhängigen Funktionen $\cos(\frac{3}{4}\pi n)$ sowie $\sin(\frac{3}{4}\pi n)$ liefert das lineare Gleichungssystem

$(2+\sqrt{2})A + (-1-\sqrt{2})B = 1$,
$(1+\sqrt{2})A + (\quad 2+\sqrt{2})B = 0$.

Dieses Gleichungssystem besitzt eine eindeutige Lösung. Mit den ermittelten Werten für A und B kann dann $y_n^{(s)}$ und schließlich die allgemeine Lösung der Differenzengleichung angegeben werden.