Taylorentwicklung

Manfred Hörz

Die Linearkombination von Potenzfunktionen x^k nennt man Polynomfunktionen oder ganzrationale Funktionen $P(x)$:

$P(x)=a_0 x^0+a_1 x^1+a_2 x^2+...+a_n x^n=\sum_{k=0}^{n} a_k x^k$, n heißt der Grad der Polynomfunktion, a_k die Koeffizienten der Polynomfunktion.

Beispiel 1: $P(x)=-1+2\mathrm{x}+4\mathrm{x}^2-\frac{1}{2}x^3$ ist eine Polynomfunktion 3. Grades mit rationalen Koeffizienten.

Beispiel 2: $s(t)=\frac{1}{2}a_0 t^2+v_0 t+s_0$ ist eine Polynomfunktion 2. Grades mit reellen Koeffizienten, die die Bewegungsgleichung (Weg-Zeit-Funktion) für gleichmäßige Bewegungsabläufe (etwa freier Fall in der Nähe der Erdoberfläche) angibt.

Polynomfunktionen sind sehr leicht zu differenzieren und zu integrieren und sind wieder Polynomfunktionen:

$$P'(x)=a_1+2a_2 x+...+n a_{n-1} x^{n-1}=\sum_{k=0}^{n-1} k\, a_k x^{k-1}$$

$$\int P(x)=c+a_0 x+\frac{1}{2}a_1 x^2+...+\frac{1}{n+1}a_n x^{n+1}=\sum_{k=0}^{n}\frac{1}{k+1}a_k x^{k+1}+c$$

Eine Polynomfunktion n-ten Grades ist durch ihre Koeffizienten eindeutig festgelegt. Man könnte sie also auch als n+1-dimensionalen Vektor bzgl. der Basis $\{1, x, x^2, ..., x^n\}$ darstellen:

$P(x)=\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$. Die Koeffizienten sind wiederum durch die Ableitungen $P^{(n)}$ an der Stelle 0 und gewissen (blau geschriebenen) Zahlen gegeben, die durch das Ableiten entstehen.

Beispiel 1: $P(x)=-1+2\mathrm{x}+4\mathrm{x}^2-\frac{1}{2}x^3 \qquad P(0)=P^{(0)}(0)=-1$

$P'(x)=2\cdot 1+4\cdot 2x+(-\frac{1}{2})\cdot 3x^2 \qquad \frac{P'(0)}{1}=2$

$P''(x)=4\cdot 2\cdot 1+(-\frac{1}{2})\cdot 3\cdot 2x \qquad \frac{P''(0)}{2\cdot 1}=\frac{8}{2\cdot 1}=4$

$P'''(x)=(-\frac{1}{2})\ \cdot 3\ \cdot 2\ \cdot 1 \qquad \frac{P'''(0)}{3\ \cdot 2\ \cdot 1}=-\frac{1}{2}$. Man erkennt hier im Nenner die Fakultät.

Insgesamt hat man $-1=\frac{P(0)}{0!} \quad 2=\frac{P'(0)}{1!} \quad 4=\frac{P''(0)}{2!} \quad -\frac{1}{2}=\frac{P'''(0)}{3!}$, sodass das

Polynom auch in der Form $P(x)=\frac{P(0)}{0!}+\frac{P'(0)}{1!}x+\frac{P''(0)}{2!}x^2+\frac{P'''(0)}{3!}x^3=\sum_{k=0}^{3}\frac{P^{(k)}(0)}{k!}x^k$

geschrieben werden kann.

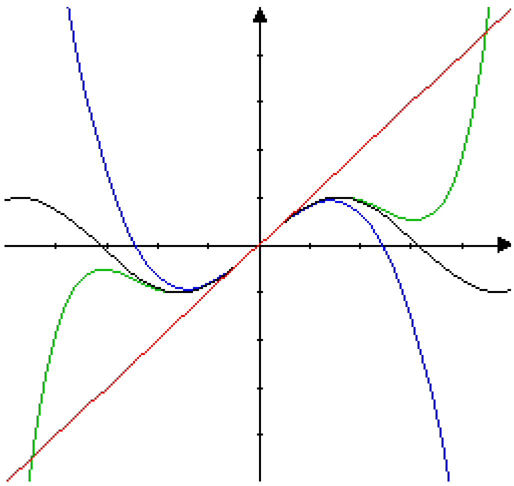
Das Erstaunliche ist nun, dass man *beliebige* Funktionen mit genügend hohen Ableitungen durch solche einfachen Polynome annähern kann.

Beispiel: $f(x)=\sin(x)$ ist unendlich oft differenzierbar.
$f'(x)=\cos(x) \quad f''(x)=-\sin(x) \quad f'''(x)=-\cos(x)\, f^{(4)}(x)=\sin(x),\ldots$

Ich bilde nun das Polynom 1. Grades und das Polynom 3. Grades als Näherungen der Sinusfunktion um die Stelle 0:

$$\sum_{k=0}^{1}\frac{f^{(k)}(0)}{k!}x^k=\sin(0)+\cos(0)x=x$$

$$\sum_{k=0}^{3}\frac{f^{(k)}(0)}{k!}x^k=\sin(0)+\cos(0)x-\frac{\sin(0)}{2}x^2-\frac{\cos(0)}{6}x^3=x-\frac{1}{6}x^3$$



Das Polynom 5.Grades schmiegt sich noch besser an:

$$\sum_{k=0}^{3}\frac{f^{(k)}(0)}{k!}x^k=\sin(0)+\cos(0)x-\frac{\sin(0)}{2}x^2-\frac{\cos(0)}{6}x^3+\frac{\sin(0)}{24}x^4+\frac{\cos(0)}{120}x^5=x-\frac{1}{6}x^3+\frac{1}{120}x^5$$

Sei $f: I \subseteq \mathbb{R} \to \mathbb{R}$ eine Funktion aus C^{n+1} (d.h. n+1 mal stetig differenzierbar), dann heißt das Polynom

$$T_n f_{x_0} = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

das **n-te Taylorpolynom** von f an der Entwicklungsstelle x_0 .

Man sagt, dass die Funktion f bis zur n-ten Ordnung durch das Taylorpolynom entwickelt (expanded) wurde.

Dabei gilt für alle $x \in I$ und jede beliebige Entwicklungsstelle $x_0 \in I$

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-z)^n}{n!} f^{(n+1)}(z)\,dz \quad (*)$$

Dabei heißt das Integral das **integrale Restglied** $R_n f_{x_0}$, das angibt wie sehr sich die Funktion f von dem Polynom unterscheidet.

Ist $f \in C^{\infty}$ und konvergiert die Folge von Taylorpolynomen $\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ für $n \to \infty$ gegen $f(x)$, so heißt die Reihe

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

die **Taylorreihe von f**. Ist $x_0 = 0$ so nennt man die Reihe auch **Maclaurin-Reihe**.

Beweis von (*) über vollst. Induktion:

IA: $n=0$ $\quad f(x) = \frac{f^{(0)}(x_0)}{0!}(x-x_0)^0 + \int_{x_0}^{x} \frac{(x-z)^0}{0!} f'(z)\,dz \Leftrightarrow f(x) = f(x_0) + \int_{x_0}^{x} f'(z)\,dz$.

Das ist der Hauptsatz der Differenzial-und Integralrechnung.

IS: $n \to n+1$: IV: $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-z)^n}{n!} f^{(n+1)}(z)\,dz$

IB: $f(x) = \sum_{k=0}^{n+1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-z)^{n+1}}{(n+1)!} f^{(n+2)}(z)\,dz$

Beweis: Das Integral der IB wird über die partielle Integration umgeschrieben:

$$\int_{x_0}^{x} \frac{(x-z)^{n+1}}{(n+1)!} f^{(n+1)}{}'(z)\,dz = \left[\frac{(x-z)^{n+1}}{(n+1)} ! \, f^{(n+1)}(z)\right]_{x_0}^{x} - \int_{x_0}^{x} \frac{-(n+1)(x-z)^n}{(n+1)!} f^{(n+1)}(z)\,dz =$$

$$=\int_{x_0}^{x}\frac{(x-z)^n}{n!}f^{(n+1)}(z)\,dz-\frac{(x-x_0)^{n+1}}{(n+1)!}f^{(n+1)}(x_0)$$ und auf das Integral die IV angewandt.

Also gilt $\sum_{k=0}^{n+1}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\int_{x_0}^{x}\frac{(x-z)^{n+1}}{(n+1)!}f^{(n+2)}(z)\,dz=$

$$=\sum_{k=0}^{n+1}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k-\frac{(x-x_0)^{n+1}}{(n+1)!}f^{(n+1)}(x_0)+\int_{x_0}^{x}\frac{(x-z)^n}{(n)!}f^{(n+1)}(z)\,dz\overset{IV}{=}$$

$$=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\int_{x_0}^{x}\frac{(x-z)^n}{(n)!}f^{(n+1)}(z)\,dz\overset{IV}{=}f(x)$$ wenn die beiden ersten Terme (schwarz

und blau) zur kleineren Summe zusammengefasst werden.

Bemerkung: Die Taylorreihe muss erstens nicht für jede Funktion konvergieren, (bspw. nicht für $\ln x$ an der Stelle 1) und zweitens, falls sie das tut, nicht gegen f konvergieren (bspw. $f(x)=\begin{cases} e^{-\frac{1}{x^2}} & x\neq 0 \\ 0 & x=0\end{cases}$ an der Entwicklungsstelle 0). Für die „üblichen“ Funktionen der Physik tut sie das aber im Allgemeinen.

Philosophische Bemerkung: Es besteht eine gewissen Ähnlichkeit zwischen dem Verhältnis von rationalen Zahlen zu irrationalen auf der einen Seite und dem Verhältnis von Taylorpolynomen zu den transzendenten Funktionen wie bspw. die trigonometrischen oder die Exponentialfunktion. Die irrationalen Zahlen sind im Allgemeinen geometrisch definiert so ähnlich wie auch die trigonometrischen Funktionen. Man nähert die irrationalen Zahlen über rationale an, indem nach einer Dezimalstelle abgebrochen wird ebenso wie man transzendente Funktionen über Polynomfunktionen, also rationale, annähert.

Es sollen noch die wichtigsten und einfachsten transzendente Funktionen mit ihren Taylorpolynomen an der Entwicklungsstelle $x_0=0$ angegeben werden (MacLaurin-Reihen):

$$\sin x=x-\frac{x^3}{6}+\frac{x^5}{120}-...=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}$$ für alle x

$$\cos x=1-\frac{x^2}{2}+\frac{x^4}{24}-...=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!}$$ für alle x

$$e^x=1+x+\frac{x^2}{2}+...=\sum_{k=0}^{\infty}\frac{x^k}{k!}$$ für alle x

Und für den Logarithmus an der Entwicklungsstelle $x_0=1$:

$$\ln x=(x-1)-\frac{(x-1)^2}{2}+\frac{(x-1)^3}{3}-...=\sum_{k=1}^{\infty}(-1)^{k+1}\frac{(x-1)^k}{k}$$ für alle $x\in]0,2]$

Besonders schön ist, dass die Entwicklungen auch für Funktionen mit mehrerer Variablen möglich sind.

Es sollen Funktionen mit zwei unabhängigen Variablen x , y betrachtet werden.

Sei $f: G \subseteq \mathbb{R}^2 \to \mathbb{R}$, wobei G ein Gebiet sein soll. Die Entwicklung von f bis zum Taylorpolynom 2. Grades an der Stelle (x_0, y_0) :

$$f(x,y) \approx f(x_0,y_0) + \nabla f(x_0,y_0) \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \cdot \left(H_f(x_0,y_0) \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \right)$$, wobei

∇ der Gradient und H_f die Hessematrix ist (also das Analogon des Gradienten für die 2. Ableitungen) oder ausgeschrieben:

$$f(x,y) \approx f(x_0,y_0) + \begin{pmatrix} \frac{\partial f}{\partial x}(x_0,y_0) \\ \frac{\partial f}{\partial y}(x_0,y_0) \end{pmatrix} \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \cdot \left(\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0,y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(x_0,y_0) & \frac{\partial^2 f}{\partial y^2}(x_0,y_0) \end{pmatrix} \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \right)$$

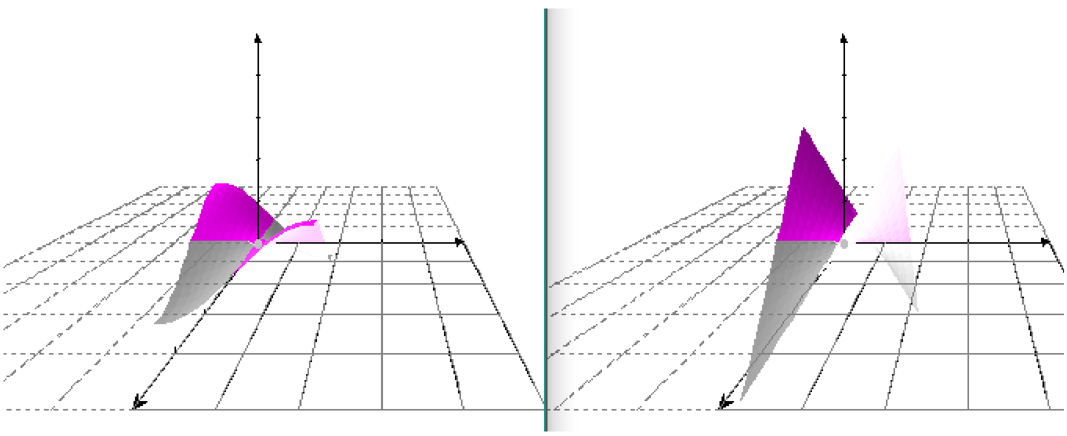
Beispiel: $f(x,y) = \sin(x)\sin(y)$ an der Entwicklungsstelle $(0,0)$ und $(\pi/2, \pi/2)$.

Zunächst die Ableitungen: $\frac{\partial f}{\partial x} = \cos(x)\sin(y)$ $\frac{\partial f}{\partial y} = \sin(x)\cos(y)$

$\frac{\partial^2 f}{\partial x^2} = -\sin(x)\sin(y)$ $\frac{\partial^2 f}{\partial y^2} = -\sin(x)\sin(y)$ $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \cos(x)\cos(y)$, sodass der

Gradient von f ist $\nabla f = \begin{pmatrix} \cos(x)\sin(y) \\ \sin(x)\cos(y) \end{pmatrix}$ und $H_f = \begin{pmatrix} -\sin(x)\sin(y) & \cos(x)\cos(y) \\ \cos(x)\cos(y) & -\sin(x)\sin(y) \end{pmatrix}$

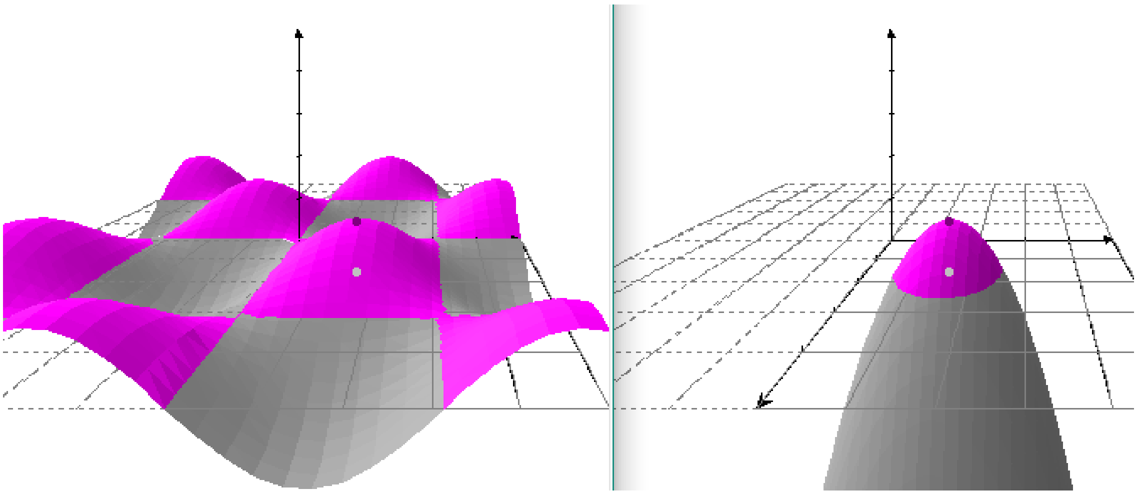
a) $(0,0)$: $f(x,y) \approx 0 + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x \\ y \end{pmatrix} \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right) = \frac{1}{2} \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} = xy = T_2 f_{(0,0)}$



Gradient ist Null. Hessedeterminante ist negativ (-1). Hier liegt also ein Sattelpunkt vor.

b) $(\pi/2,\pi/2)$: $f(x,y)\approx 1+\begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}x-\pi/2\\y-\pi/2\end{pmatrix}+\frac{1}{2}\begin{pmatrix}x-\pi/2\\y-\pi/2\end{pmatrix}\left(\begin{pmatrix}-1 & 0\\0 & -1\end{pmatrix}\right)\begin{pmatrix}x-\pi/2\\y-\pi/2\end{pmatrix}=$

$=1-(x-\pi/2)^2-(y-\pi/2)^2=T_2 f_{(\pi/2,\pi/2)}$



Gradient ist Null, Hessedeterminante ist positiv (1) , also Extrempunkt. Weiter ist $\frac{\partial f^2}{\partial x^2}|_{(\pi/2,\pi/2)}=-1<0$. Hier liegt also ein lokales Maximum.

Bemerkung: Approximiert man die Funktion f durch das Taylorpolynom 2. Grades, so nennt man dieses Taylorpolynom auch die Schmiegequadrik von f.

Ist die Funktion f eine Funktion mit mehreren Variablen, so ist ihre Schmiegequadrik folgende:

$$T_2 f_{\boldsymbol{x_0}}=f(\boldsymbol{x_0})+\nabla f(\boldsymbol{x_0})^T\cdot(\boldsymbol{x}-\boldsymbol{x_0})+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x_0})^T\cdot\left(H_f(\boldsymbol{x_0})\cdot(\boldsymbol{x}-\boldsymbol{x_0})\right)$$

Dabei ist $\nabla f(\boldsymbol{x_0})=\begin{pmatrix}\frac{\partial f}{\partial x_1}(\boldsymbol{x_0})\\ \vdots\\ \frac{\partial f}{\partial x_n}(\boldsymbol{x_0})\end{pmatrix}$ und $H_f(\boldsymbol{x_0})=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1\partial x_1}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}(\boldsymbol{x_0})\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_n\partial x_n}(\boldsymbol{x_0})\end{pmatrix}$ oder

die Matrizenmultiplikationen ausgerechnet

$$f(\boldsymbol{x})\approx f(\boldsymbol{x_0})+\sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(\boldsymbol{x_0})\cdot(x_k-x_{0k})+\frac{1}{2}\sum_{i,k=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_k}(\boldsymbol{x_0})(x_i-x_{0i})(x_k-x_{0k})$$

Beispiel: $f(x,y,z)=x+y e^z \quad \frac{\partial f}{\partial x}=1 \quad \frac{\partial f}{\partial y}=e^z \quad \frac{\partial f}{\partial z}=y e^z \quad \nabla f^T=(1, e^z, ye^z)$

$$\frac{\partial^2 f}{\partial x^2}=0 \quad \frac{\partial^2 f}{\partial x \partial y}=0 \quad \frac{\partial^2 f}{\partial x \partial z}=0 \quad \frac{\partial^2 f}{\partial y \partial x}=0 \quad \frac{\partial^2 f}{\partial y^2}=0 \quad \frac{\partial^2 f}{\partial y \partial z}=e^z$$

$$\frac{\partial^2 f}{\partial z \partial x}=0 \quad \frac{\partial^2 f}{\partial z \partial y}=e^z \quad \frac{\partial^2 f}{\partial z^2}=y e^z \quad H_f=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & e^z \\ 0 & e^z & ye^z \end{pmatrix}$$

$$f(x,y,z)\approx f(x_0,y_0,z_0)+1 \cdot (x-x_0)+e^{z_0}(y-y_0)+y_0 e^{z_0}(z-z_0)+$$
$$+\frac{1}{2}\Big(0(x-x_0)^2+0(x-x_0)(y-y_0)+0(x-x_0)(z-z_0)+0(y-y_0)(x-x_0)+0(y-y_0)^2+e^{z_0}(y-y_0)(z-z_0)+$$
$$+0(z-z_0)(x-x_0)+e^{z_0}(z-z_0)(y-y_0)+y_0 e^{z_0}(z-z_0)^2\Big)$$

Speziell für die Entwicklungsstelle $(x_0, y_0, z_0)=(1,2,1)$:

$$f(x,y,z)\approx 1+2e+(x-1)+e(y-2)+2e(z-1)+\frac{1}{2}(e(y-2)(z-1)+e(z-1)(y-2)+2e(z-1)^2)$$

$$f(x,y,z)\approx x+ez(y+z-2)+e \quad .$$

oder für den Ursprung: $f(x,y,z)\approx x+y+2y z$

Das n-te Taylorpolynom von $f=f(x_1,\ldots,x_n)$ an der Entwicklungsstelle x_0 unter den entsprechenden Voraussetzungen ist:

$$T_n f_{x_0}=f(\mathbf{x_0})+\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\mathbf{x_0}) \cdot (x_i-x_{0i})+\frac{1}{2!}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x_0})(x_i-x_{0i})(x_j-x_{0j})+$$
$$+\frac{1}{3!}\sum_{i,j,k=1}^{n}\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x_0})(x_i-x_{0i})(x_j-x_{0j})(x_k-x_{0k})+\ldots+$$
$$+\frac{1}{n!}\sum_{i_1,\ldots,i_n=1}^{n}\frac{\partial^n f}{\partial x_{i_1}\ldots\partial x_{i_n}}(\mathbf{x_0})(x_{i_1}-x_{0i_1})\ldots(x_{i_n}-x_{0i_n})$$

Für die Variationsrechnung:

Bezeichnet das Differenzial dx die infinitesimale Änderung des Wertes x, so bezeichnet df die damit einhergehende infinitesimale Änderung des Funktionswertes $f(x)$:
$df=f(x+dx)-f(x)$.

Ist f eine Funktion in mehreren Veränderlichen $f(x_1,\ldots,x_n,t)$ und erfahren die Koordinaten eine infinitesimale Änderung $dx_1,\ldots,dx_n, dt$, so ist die entsprechende Änderung der Funktionswertes

$$df=f(x_1+dx_1,\ldots,x_n+dx_n,t+dt)-f(x_1,\ldots,x_n,t)=\frac{\partial f}{\partial x_1}dx_1+\ldots+\frac{\partial f}{\partial x_n}dx_n+\frac{\partial f}{\partial t}dt \quad .$$

Im Unterschied zu einer wirklichen, physikalischen (zeitlichen) Änderung, versteht man unter einer Änderung, die die Zeit fest lässt, eine *virtuelle* Änderung. Dadurch ist $dt=0$ und die Änderung des Funktionswertes schreibt man nach Lagrange mit δf sowie die entsprechenden Koordinatenänderungen mit $\delta x_1, \ldots, \delta x_n$. Somit ergibt sich entsprechend:

$$\delta f = f(x_1+\delta x_1, \ldots x_n+\delta x_n, t) - f(x_1, \ldots, x_n, t) = \frac{\partial f}{\partial x_1}\delta x_1 + \ldots + \frac{\partial f}{\partial x_n}\delta x_n \quad .$$

Man nennt den Ausdruck $\delta f = \frac{\partial f}{\partial x_1}\delta x_1 + \ldots + \frac{\partial f}{\partial x_n}\delta x_n$ die **erste Variation** der Funktion.

Man kann auch die infinitesimalen Elemente δx_i durch endliche ausdrücken: $\delta x_i = \epsilon a_i$, wobei a_i die Richtungskosinus sind und ϵ eine positive Zahl, die man gegen Null streben lässt.

Man hat dann $\delta f = \frac{\partial f}{\partial x_1}\epsilon a_1 + \ldots + \frac{\partial f}{\partial x_n}\epsilon a_n = \epsilon\left(\frac{\partial f}{\partial x_1}a_1 + \ldots + \frac{\partial f}{\partial x_n}a_n\right)$ oder

$$\frac{\delta f}{\epsilon} = \frac{\partial f}{\partial x_1}a_1 + \ldots + \frac{\partial f}{\partial x_n}a_n$$

die Änderungsrate der Funktion in die gegebene Richtung, durch die Richtungskosinus ausgedrückt.

Wenn die Funktion einen stationären Wert hat, so muss die Änderungsrate in jeder Richtung Null sein, also muss gelten

$$\frac{\partial f}{\partial x_1}a_1 + \ldots + \frac{\partial f}{\partial x_n}a_n = 0$$

für alle Richtungen $(a_1, \ldots, a_n)$ (und also auch $\delta f = 0$). Dann müssen die Koeffizienten auch alle Null sein: $\bigwedge_k \frac{\partial f}{\partial x_k} = 0$.

Umgekehrt ist $\bigwedge_k \frac{\partial f}{\partial x_k} = 0$, dann ist $\delta f = 0$ und auch die Änderungsrate in jede Richtung Null.

Also ist $\bigwedge_k \frac{\partial f}{\partial x_k} = 0$ oder $\delta f = 0$ hinreichend und notwendig für einen stationären Wert von f.

Die **zweite Variation** der Funktion ist $\delta^2 f := \epsilon^2 \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k} a_i a_k$. Man erhält dies auf folgende Weise:

$$f(\boldsymbol{y}) = f(y_1, \ldots, y_n) = f(x_1+\epsilon a_1, \ldots, x_n+\epsilon a_n) \qquad \boldsymbol{y_0} = \boldsymbol{x} = (x_1, \ldots, x_n) \text{ mit } \epsilon = 0 \quad .$$

$f(\boldsymbol{y})$ werde an der Stelle $\boldsymbol{y_0}$ durch das Taylorpolynom 2. Grades approximiert:

$$f(\boldsymbol{y}) \approx f(\boldsymbol{y_0}) + \sum_{k=1}^{n} \frac{\partial f}{\partial y_k}(\boldsymbol{y_0}) \cdot (y_k - y_{0k}) + \frac{1}{2}\sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial y_i \partial y_k}(\boldsymbol{y_0})(y_i - y_{0i})(y_k - y_{0k}) \quad .$$

Aufgrund der Kettenregel gilt

$$\frac{\partial f}{\partial x_k} = \frac{\partial f}{\partial y_k} \cdot \frac{d(x_k + \epsilon a_1)}{d x_k} = \frac{\partial f}{\partial y_k} \quad \text{und} \quad \frac{\partial^2 f}{\partial y_i \partial y_k} = \frac{\partial f}{\partial x_i \partial x_k} \quad \text{, sodass wegen}$$

$$y_k - y_{0k} = x_k + \epsilon a_k - x_k = \epsilon a_k \quad \text{gilt}$$

$$f(\boldsymbol{y}) - f(\boldsymbol{y_0}) \approx \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\boldsymbol{x}) \cdot \epsilon a_k + \frac{1}{2} \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}(\boldsymbol{x}) \epsilon a_i \epsilon a_k \quad \text{oder}$$

$$f(\boldsymbol{y}) - f(\boldsymbol{y_0}) \approx \epsilon \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\boldsymbol{x}) \cdot a_k + \frac{1}{2} \epsilon^2 \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}(\boldsymbol{x}) a_i a_k \quad \text{und da}$$

$$f(\boldsymbol{y}) - f(\boldsymbol{y_0}) = f(x_1 + \epsilon a_1, \ldots, x_n + \epsilon a_n) - f(x_1, \ldots, x_n) = \delta f \quad \text{gilt weiter}$$

$$\delta f \approx \epsilon \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\boldsymbol{x}) \cdot a_k + \frac{1}{2} \epsilon^2 \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}(\boldsymbol{x}) a_i a_k$$

Wenn nun $\boldsymbol{x}$ eine stationäre Stelle ist, dann ist $\bigwedge_k \frac{\partial f}{\partial x_k} = 0$,also fällt der erste Summand weg und es bleibt auf der rechten Seite $\frac{1}{2} \epsilon^2 \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}(\boldsymbol{x}) a_i a_k$ übrig. Für sehr kleine ϵ können die höheren als die 2. Glieder (die ja in 3. und höherer Potenz von ϵ vorkommen) ohnehin vernachlässigt werden.

Es bleibt also die Variation in zweiter Potenz erhalten und die wird mit $\delta^2 f$ bezeichnet:

$$\delta^2 f = \epsilon^2 \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}(\boldsymbol{x}) a_i a_k$$

Ist nun zusätzlich $\delta^2 f > 0$ für alle Richtungen $\boldsymbol{a} = (a_1, \ldots, a_n)$, dann ist

$f(x_1 + \epsilon a_1, \ldots, x_n + \epsilon a_n) - f(x_1, \ldots, x_n) > 0$ und $\boldsymbol{x}$ muss mithin eine lokale Minimumstelle sein.

Ist zusätzlich $\delta^2 f < 0$ für alle Richtungen $\boldsymbol{a} = (a_1, \ldots, a_n)$, dann ist

$f(x_1 + \epsilon a_1, \ldots, x_n + \epsilon a_n) - f(x_1, \ldots, x_n) < 0$ und $\boldsymbol{x}$ muss eine lokale Maximumstelle sein.

Das Vorzeichen der zweiten Variation liefert also ein Kriterium für die Art des Extremum, vorausgesetzt, dass $\delta^2 f \neq 0$ in jeder Richtung.

Es ist jedoch nicht immer notwendig, die zweite Variation zu bemühen, falls aufgrund des Problems klar ist, dass etwa ein Minimum vorliegen muss. Dann reicht die stationäre Stelle aus.