Rheinisch-Westfälische Technische Hochschule Aachen
Institut für Geometrie und Praktische Mathematik

Numerisches Rechnen — WS 2011/2012

Prof. Dr. M. Grepl — P. Esser, G. Welper, L. Zhang

Klausur Numerisches Rechnen (16.02.2012) (Musterlösung)

- Hilfsmittel: nur dokumentenechtes Schreibgerät (blau oder schwarz); genau ein Taschenrechner, der auf der Liste der erlaubten Taschenrechner steht; zwei beidseitig handbeschriebene Din-A4-Blätter
- kein eigenes Papier benutzen und nicht mit Blei-, Rot- oder Grünstift schreiben
- Bearbeitungszeit: 120 Minuten
- Deckblätter ausfüllen und unterschreiben
- Aufgabenblätter kontrollieren: insgesamt sechs Aufgaben
- jedes Blatt mit Namen und Matrikelnummer versehen
- Studenten- und Lichtbildausweis zur Kontrolle bereitlegen
- keine vorzeitige Abgabe während der letzten 15 Minuten

Zum Bestehen der Klausur sind 40 der insgesamt 80 erreichbaren Punkte erforderlich. Die Klausurergebnisse werden voraussichtlich ab Freitag, den 24. Februar 2012, auf der Webseite zur Veranstaltung bekanntgegeben. Die Klausureinsicht findet am Montag, den 27. Februar 2012, von 14:00 – 16:00 Uhr im Raum 149 Hauptgebäude statt. Danach sind keine Einsprüche gegen die Korrektur mehr möglich. Die Klausur kann nach einer Aufbewahrungsfrist von 5 Jahren innerhalb von 3 Wochen am Institut für Geometrie und Praktische Mathematik abgeholt werden.

Matrikelnummer: _____

Name: _____ Vorname: _____

Hiermit erkläre ich, dass ich keine anderen als die erlaubten Hilfsmittel benutze. Ferner nehme ich zur Kenntnis, dass bei Täuschungsversuchen, auch solchen zugunsten anderer, die Klausur als *nicht bestanden* bewertet wird.

Datum: _____ Unterschrift: _____

Korrekturvermerke

A 1	A 2	A 3	A 4	A 5	A 6	$\sum$

Aufgabe 1

a) Nennen Sie drei Probleme, die beim Rechnen mit Maschinenzahlen auftreten können.

b) Zeigen Sie, dass die Multiplikation zweier reeller Zahlen gut konditioniert ist.

c) Gegeben sei eine symmetrisch positiv definite Matrix $A \in \mathbb{R}^{n\times n}$ und ein Vektor $b \in \mathbb{R}^n$. Welches in der Vorlesung vorgestellte Verfahren ist am besten zur Lösung des Gleichungssystems $Ax = b$ geeignet? Geben Sie eine kurze Begründung für Ihre Antwort.

d) Betrachten Sie ein lineares Ausgleichsproblem $||Ax - b||_2 \to \min$ mit der Matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n, b \in \mathbb{R}^m$.

(i) Welche Bedingung muss A erfüllen, damit das Problem eindeutig lösbar ist?

(ii) Welche Bedingung muss b erfüllen, damit die Norm des Residuums gleich Null ist?

e) Seien $R \in \mathbb{R}^{n\times n}$, $b, x \in \mathbb{R}^n$ und $c \in \mathbb{R}^{m-n}$ mit $m \geq n$. Zeigen Sie:

$$\left\| \begin{pmatrix} R \\ 0 \end{pmatrix} x - \begin{pmatrix} b \\ c \end{pmatrix} \right\|_2^2 = \|Rx - b\|_2^2 + \|c\|_2^2.$$

f) Beweisen Sie: für alle $x \in \mathbb{R}^n$ gilt $\|x\|_\infty \leq \|x\|_2$.

g) Geben Sie drei wesentliche Unterschiede zwischen den Newton-Cotes-Formeln und der Gauß-Quadratur für die numerische Integration an.

h) Sei $f(x) = \frac{1}{2}x^T Ax - b^T x + c$ eine Funktion von $\mathbb{R}^n$ nach $\mathbb{R}$. $A \in \mathbb{R}^{n\times n}$ sei symmetrisch positiv definit, $b \in \mathbb{R}^n, c \in \mathbb{R}$. Zeigen Sie:

$$x^* = \arg\min_{x\in\mathbb{R}^n} f(x) \Rightarrow Ax^* = b.$$

i) Gegeben sei das lineare Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 6 & 2 & 3 \\ 2 & 8 & 2 \\ 3 & 2 & 6 \end{pmatrix} \quad \textit{und} \quad b = \begin{pmatrix} -2.5 \\ 1 \\ 3 \end{pmatrix}.$$

(i) Konvergiert das Gauß-Seidel-Verfahren eingesetzt zur Lösung dieses linearen Gleichungssystems?

(ii) Konvergiert das Jacobi-Verfahren eingesetzt zur Lösung dieses linearen Gleichungssystems?

3+3+2+2+3+3+3+3+2=24 Punkte

Musterlösung

a) (i) Auslöschung

(ii) relative Maschinengenauigkeit

(iii) overflow / underflow

b) Wir betrachten $f(x) = x_1 x_2$ mit $x = (x_1, x_2)^T$. Damit erhält man

$$\frac{\partial f(x)}{\partial x_1} = x_2, \quad \frac{\partial f(x)}{\partial x_2} = x_1, \quad \text{und} \quad \phi_j(x) = \frac{x_1 x_2}{f(x)} = 1, \quad j = 1, 2.$$

Mit der Definition $\kappa_{\text{rel}}(x) = \max_j |\phi_j(x)|$ ergibt sich damit $\kappa_{\text{rel}} = 1$.

c) Das Cholesky Verfahren, wegen geringem Aufwand und Stabilität.

d) (i) A muss vollen Rang haben, damit die Normalgleichungen eindeutig lösbar sind.

(ii) $b \in \text{Bild}(A)$. Dann gibt es ein $x \in \mathbb{R}$ mit $Ax = b$.

e) Für $z = (z^1, z^2)^T$ mit $z^1 \in \mathbb{R}^n$ und $z^2 \in \mathbb{R}^{m-n}$ gilt:

$$\|z\|_2^2 = \sum_{i=1}^{m} z_i^2 = \sum_{i=1}^{n} z_i^2 + \sum_{i=n+1}^{m} z_i^2 = \sum_{i=1}^{n} (z_i^1)^2 + \sum_{i=1}^{m-n} (z_i^2)^2 = \|z^1\|_2^2 + \|z^2\|_2^2.$$

Anwenden auf

$$z = \begin{pmatrix} R \\ 0 \end{pmatrix} x - \begin{pmatrix} b \\ c \end{pmatrix}$$

ergibt

$$\left\| \begin{pmatrix} R \\ 0 \end{pmatrix} x - \begin{pmatrix} b \\ c \end{pmatrix} \right\|_2^2 = \|Rx - b\|_2^2 + \|c\|_2^2.$$

f) Sei $|x_k| = \|x\|_\infty$. Dann gilt

$$\|x\|_\infty^2 = |x_k|^2 = x_k^2 \leq \sum_{l=1}^{n} x_l^2 = \|x\|_2^2$$

und somit $\|x\|_\infty \leq \|x\|_2$.

g) Unterschiede:

	Newton-Cotes	Gauß
Stützstellen	äquidistant	nicht äquidistant
Gewichte	wechselnde Vorzeichen	positiv
Exaktheitsgrad ($m+1$ Knoten)	m oder $m+1$	$2m+1$

h) Für das Minimum gilt $\nabla f(x^*) = 0$. Also ist

$$\nabla f(x^*) = Ax^* - b = 0$$

und somit $Ax^* = b$.

i) Untersuche A mit dem Zeilensummenkriterium:

$$\max_{i=1,2,3} \sum_{\substack{j=1 \\ j\neq i}}^{3} \frac{|a_{ij}|}{|a_{ii}|} = \max\left\{ \frac{2}{6} + \frac{3}{6}, \frac{2}{8} + \frac{2}{8}, \frac{3}{6} + \frac{2}{6} \right\} = \frac{5}{6} \leq 1.$$

Damit konvergieren sowohl das Jacobi- also auch das Gauss-Seidel-Verfahren.

Aufgabe 1
Gegeben sei das lineare Gleichungssystem $Ax = b$ mit

$$A := \begin{pmatrix} 0.2 & 0.5 \\ 0.8 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2} \quad \text{und} \quad b := \begin{pmatrix} 0.75 \\ -0.25 \end{pmatrix} \in \mathbb{R}^2.$$

Zur Lösung des Gleichungssystems stehen Ihnen Approximationen von A und b zur Verfügung, die mit einem absoluten Fehler von 0.05 in jedem Eintrag behaftet sind. Mit welchem relativen Fehler der Lösung des gestörten linearen Gleichungssystems - gemessen in der Norm $||\cdot||_1$ - müssen Sie rechnen? **Hinweis:** $||A^{-1}||_1 = 9$.

10 Punkte

Musterlösung Aus

$$\|b\|_1 = 1, \quad \|A\|_1 = \max\{1, \frac{3}{2}\} = \frac{3}{2} \quad \text{und} \quad \left\|A^{-1}\right\|_1 = 9$$

ergibt sich die Kondition von A zu

$$\kappa_1(A) = \|A\|_1 \cdot \left\|A^{-1}\right\|_1 = \frac{3}{2} \cdot 9 = \frac{27}{2}.$$

Der Fehler in jedem Eintrag beträgt maximal $\frac{1}{20}$, somit ist im schlimmsten Fall

$$\Delta A = \begin{pmatrix} \frac{1}{20} & \frac{1}{20} \\ \frac{1}{20} & \frac{1}{20} \end{pmatrix} \quad \text{und} \quad \Delta b = \begin{pmatrix} \frac{1}{20} \\ \frac{1}{20} \end{pmatrix}.$$

Es folgt

$$\|\Delta A\|_1 \leq \frac{1}{10} \quad \text{und} \quad \|\Delta b\|_1 \leq \frac{1}{10}$$

und daher auch

$$\frac{\|\Delta A\|_1}{\|A\|_1} \leq \frac{1}{10} \cdot \frac{2}{3} = \frac{1}{15} \quad \text{und} \quad \frac{\|\Delta b\|_1}{\|b\|_1} \leq \frac{1}{10}.$$

Der relative Fehler der Lösung lässt sich daher abschätzen durch

$$\begin{aligned} \frac{\|\Delta x\|_1}{\|x\|_1} &\leq \frac{\kappa_1(A)}{1 - \kappa_1(A) \cdot \frac{\|\Delta A\|_1}{\|A\|_1}} \cdot \left(\frac{\|\Delta A\|_1}{\|A\|_1} + \frac{\|\Delta b\|_1}{\|b\|_1} \right) \\ &\leq \frac{\frac{27}{2}}{\underbrace{1 - \frac{27}{2} \cdot \frac{1}{15}}_{=\frac{1}{10}>0}} \cdot \left(\frac{1}{15} + \frac{1}{10} \right) = 22.5. \end{aligned}$$

Aufgabe 2
Betrachten Sie

$$A = \begin{pmatrix} 2 & 7 & 17 \\ -8 & -30 & 37 \\ 6 & 21 & -12 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

a) Berechnen Sie die LR-Zerlegung von A mit Spaltenpivotisierung (ohne Äquilibrierung). Geben Sie L, R und P explizit an!

b) Bestimmen Sie $\det(A)$.

c) Lösen Sie das lineare Gleichungssystem $Ax = b$ mit $b = (7, -17, 0)^T \in \mathbb{R}^3$. Verwenden Sie die LR-Zerlegung aus Teil a).

7+1+5=13 Punkte

Musterlösung

a) Berechne LR-Zerlegung von A mit Spaltenpivotisierung:

$$A = \begin{pmatrix} 2 & 7 & 17 \\ -8 & -30 & 37 \\ 6 & 21 & -12 \end{pmatrix} \overset{\sigma_1=(12)}{\rightsquigarrow} \begin{pmatrix} -8 & -30 & 37 \\ 2 & 7 & 17 \\ 6 & 21 & -12 \end{pmatrix} \rightsquigarrow \left(\begin{array}{ccc} -8 & -30 & 37 \\ \hline -\frac{1}{4} & -\frac{1}{2} & \frac{105}{4} \\ -\frac{3}{4} & -\frac{3}{2} & \frac{63}{4} \end{array}\right)$$

$$\overset{\sigma_2=(23)}{\rightsquigarrow} \left(\begin{array}{ccc} -8 & -30 & 37 \\ \hline -\frac{3}{4} & -\frac{3}{2} & \frac{63}{4} \\ -\frac{1}{4} & -\frac{1}{2} & \frac{105}{4} \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc} -8 & -30 & 37 \\ \hline -\frac{3}{4} & -\frac{3}{2} & \frac{63}{4} \\ -\frac{1}{4} & \frac{1}{3} & 21 \end{array}\right)$$

$$\Rightarrow\ P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},\ L = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{4} & 1 & 0 \\ -\frac{1}{4} & \frac{1}{3} & 1 \end{pmatrix},\ R = \begin{pmatrix} -8 & -30 & 37 \\ 0 & -\frac{3}{2} & \frac{63}{4} \\ 0 & 0 & 21 \end{pmatrix}$$

b) Die Determinante läßt sich berechnen als

$$\begin{aligned} \det(A) &= \det(P^{-1}LR) \\ &= \det(P)^{-1} \cdot \underbrace{\det(L)}_{=1} \cdot \det(R) \\ &= (-1)^{\#Vertauschungen} \cdot 1 \cdot \det(R) \\ &= (-8) \cdot \left(-\frac{3}{2}\right) \cdot 21 = 252. \end{aligned}$$

c) Löse das lineare Gleichungssystem $Ax = b$:

$$Ax = b \quad \Leftrightarrow \quad PAx = Pb \quad \Leftrightarrow \quad L\underbrace{Rx}_{=:y} = Pb.$$

Die permutierte rechte Seite lautet

$$Pb = \begin{pmatrix} -17 \\ 0 \\ 7 \end{pmatrix}.$$

Vorwärtseinsetzen liefert: $Ly = Pb \quad \Leftrightarrow \quad \begin{cases} y_1 = -17, \\ y_2 = 0 + \frac{3}{4} \cdot (-17) = -\frac{51}{4}, \\ y_3 = 7 + \frac{1}{4}(-17) - (\frac{1}{3}) \cdot (-\frac{51}{4}) = 7. \end{cases}$

Rückwärtseinsetzen liefert: $Rx = y \quad \Leftrightarrow \quad \begin{cases} x_3 = \frac{1}{21} \cdot 7 = \frac{1}{3}, \\ x_2 = -\frac{2}{3} \cdot (-\frac{51}{4} - \frac{63}{4} \cdot \frac{1}{3}) = 12, \\ x_1 = -\frac{1}{8}(-17 + 30 \cdot 12 - 37 \cdot \frac{1}{3}.) = -\frac{124}{3}. \end{cases}$

Die Lösung des linearen Gleichungssystems ist $x = \begin{pmatrix} -\frac{124}{3} \\ 12 \\ \frac{1}{3} \end{pmatrix}$.

Aufgabe 3
Gegeben sei das nichtlineare Gleichungssystem

$$\begin{aligned} 4x &= \sin(x+y), \\ 4y &= \cos(x-y) \end{aligned} \qquad \text{auf } B = \{(x,y) : \ x^2 + y^2 \leq 1\}.$$

a) Skizzieren Sie B.

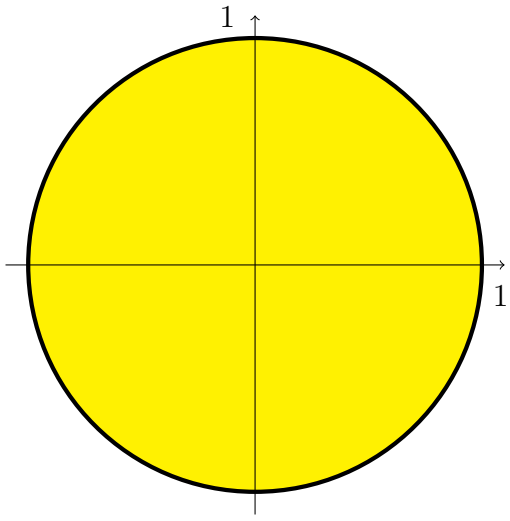
b) Zeigen Sie mit dem Banachschen Fixpunktsatz, dass dieses Gleichungssystem auf B genau eine Lösung besitzt. Verwenden Sie für den Kontraktivitätsbeweis die $||.||_\infty$-Norm.

c) Führen Sie einen Schritt der entsprechenden Fixpunktiteration mit dem Startwert $(0.7, -0.2)^T$ aus und geben Sie an, wie viele Schritte höchstens notwendig sind, um die Lösung mit der Genauigkeit $\epsilon = 10^{-4}$, gemessen in der $||.||_\infty$-Norm, zu approximieren.

1+8+4=13 Punkte

Musterlösung

a)



b) Eine Lösung des Gleichungssystems ist ein Fixpunkt der Funktion $F(x,y)$:

$$F(x,y) = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix} = \frac{1}{4} \begin{pmatrix} \sin(x+y) \\ \cos(x-y) \end{pmatrix}$$

Überprüfung der Voraussetzungen des Fixpunktsatzes:

(i) B ist offensichtlich abgeschlossen.

(ii) F ist selbstabbildend, da

$$\|F(x,y)\|_2^2 = \frac{1}{16}\sin(x+y)^2 + \frac{1}{16}\cos(x-y)^2 \leq \frac{1}{16} + \frac{1}{16} = \frac{1}{8} \leq 1$$

d.h. $F(B) \subset B$.

(iii) Kontraktivität von F. Da B Konvex folgt mit dem Mittelwertsatz

$$L := \sup_{(x,y)\in B} \|F'(x,y)\| < 1 \Rightarrow F \text{ ist kontrahierend.}$$

Es ist

$$F'(x,y) = \frac{1}{4}\begin{pmatrix} \cos(x+y) & \cos(x+y) \\ -\sin(x-y) & \sin(x-y) \end{pmatrix}.$$

Damit ist

$$\begin{aligned} \|F'(x,y)\|_\infty &= \frac{1}{4}\max\{|\cos(x+y)| + |\cos(x+y)|, |\sin(x-y)| + |\sin(x-y)|\} \\ &\leq \frac{1}{2} =: L \end{aligned}$$

(iv) Nach dem Banachschen Fixpunktsatz folgt, dass F auf B genau einen Fixpunkt besitzt.

c) Mit dem Fixpunkt $z^* := (x^*, y^*)^T$ und $z^n := (x^n, y^n)^T$ lautet die a-priori Abschätzung

$$\|z^* - z^n\|_\infty \leq \frac{L^n}{1-L}\|z^1 - z^0\|_\infty \leq \epsilon.$$

Es ist $F(0.7, -0.2) = (0.11985, 0.15540)^T$ und somit $\|z^1 - z^0\|_\infty = 0.58014$. Das gibt

$$n \geq \frac{\ln\left(\frac{\epsilon(1-L)}{\|z^1-z^0\|_\infty}\right)}{\ln L} \geq 13.51$$

Also genügen 14 Schritte.

Aufgabe 4
Die Funktion $f(x) = \cos x + e^{-x^2}$ soll interpoliert werden. Gegeben ist folgende Tabelle

x_i	0.5	1.0	1.5	2.0
$f(x_i)$	1.6564	0.9082	0.1761	−0.3978

a) Gesucht ist ein Näherungswert für $f(1.2)$ mit dem Neville-Aitken-Schema unter Benutzung aller Tabellenwerte. Ergänzen Sie dazu das folgende Tableau:

$$
\begin{array}{l|lllllll}
x_0 = 0.5 & 1.6564 & & & & & & \\
 & & \searrow & & & & & \\
x_1 = 1.0 & 0.9082 & \rightarrow & 0.60892 & & & & \\
 & & \searrow & & \searrow & & & \\
x_2 = 1.5 & 0.1761 & \rightarrow & P_{2,1} & \rightarrow & 0.613428 & & \\
 & & \searrow & & \searrow & & \searrow & \\
x_3 = 2.0 & P_{3,0} & \rightarrow & 0.52044 & \rightarrow & 0.596376 & \rightarrow & P_{3,3}
\end{array}
$$

b) Schätzen Sie die Interpolationsfehler $|P_{2,1} - f(1.2)|$ und $|P_{3,3} - f(1.2)|$ möglichst gut ab **ohne** $f(1.2)$ zu berechnen.

Hinweis:

$$\max_{x\in[0.5,2]} |f'(x)| < 4.1152, \qquad \max_{x\in[0.5,2]} |f''(x)| < 11.903, \qquad \max_{x\in[0.5,2]} |f'''(x)| < 32.152,$$
$$\max_{x\in[0.5,2]} \left|f^{(4)}(x)\right| < 60.189, \qquad \max_{x\in[0.5,2]} \left|f^{(5)}(x)\right| < 102.24.$$

c) Die Funktion $g(x)$ soll an den Stellen $x_0, x_1, x_2 ... x_n$ interpoliert werden. Dazu sei

$$P(x) = \frac{x - x_0}{x_n - x_0} P(g|x_1, ..., x_n)(x) + \frac{x_n - x}{x_n - x_0} P(g|x_0, ..., x_{n-1})(x),$$

wobei $P(g|x_1, ..., x_n)$ und $P(g|x_0, ..., x_{n-1})$ die Interpolationspolynome an den entsprechenden Stellen sind. Zeigen Sie, dass $P(x_i) = g(x_i), i \in 0, 1, ..., n$ gilt. Was können Sie daraus folgern?

5+4+5=14 Punkte

Musterlösung

a) Für $x = 1.2$,

$$P_{3,0} = -0.3978$$
$$P_{2,1} = 0.1761 + \frac{1.2 - 1.5}{1.5 - 1.0}(0.1761 - 0.9082) = 0.61536$$
$$P_{3,3} = 0.596376 + \frac{1.2 - 2.0}{2.0 - 0.5}(0.596376 - 0.613428) = 0.6054704$$

b) Benutzen

$$|f(x^*) - P(f|x_0, ..., x_n)(x^*)| \leq \left|\prod_{j=0}^{n}(x^* - x_j)\right| \max_{x\in[0.5,2]} \frac{\left|f^{(n+1)}(x)\right|}{(n+1)!}, \quad x^* \in [0.5, 2]$$

$$\begin{aligned}
|P_{2,1} - f(1.2)| &\leq |(1.2 - 1.0)(1.2 - 1.5)| \max_{x\in[0.5,2]} \frac{|f''(x)|}{2!} \\
&\leq 0.06 \times \frac{11.903}{2} = 0.35709 \\
|P_{3,3} - f(1.2)| &\leq |(1.2 - 0.5)(1.2 - 1.0)(1.2 - 1.5)(1.2 - 2.0)| \max_{x\in[0.5,2]} \frac{\left|f^{(4)}(x)\right|}{4!} \\
&\leq 0.0336 \times \frac{60.189}{24} = 0.084265
\end{aligned}$$

c) Beweis. Für $0 < i < n$ gilt

$$\begin{aligned}
P(x_i) &= \frac{x_i - x_0}{x_n - x_0} P(g|x_1, ..., x_n)(x_i) + \frac{x_n - x_i}{x_n - x_0} P(g|x_0, ..., x_{n-1})(x_i) \\
&= \frac{x_i - x_0}{x_n - x_0} g(x_i) + \frac{x_n - x_i}{x_n - x_0} g(x_i) = g(x_i)
\end{aligned}$$

Ferner erhält man

$$\begin{aligned}
P(x_0) &= \frac{x_0 - x_0}{x_n - x_0} P(g|x_1, ..., x_n)(x_i) + \frac{x_n - x_0}{x_n - x_0} P(g|x_0, ..., x_{n-1})(x_i) \\
&= 0 + \frac{x_n - x_0}{x_n - x_0} g(x_0) = g(x_0)
\end{aligned}$$

und ebenso

$$\begin{aligned}
P(x_n) &= \frac{x_n - x_0}{x_n - x_0} P(g|x_1, ..., x_n)(x_n) + \frac{x_n - x_n}{x_n - x_0} P(g|x_0, ..., x_{n-1})(x_n) \\
&= \frac{x_n - x_0}{x_n - x_0} g(x_n) + 0 = g(x_n)
\end{aligned}$$

$P(x)$ ist ein Polynom vom Grad höchstens n (als Linearkombimations zweier Polynome vom Grad höchstens n-1 mit x-Potenzen vom Grad 1 als Koeffizienten). Außerden interpoliert $P(x)$ die Funktion $g(x)$ an allen n+1 Stützstellen $x_0, ..., x_n$. Auf Grund der Eindeutigkeit der Polynominterpolation ist $P(x)$ das interpolationspolynom zu $g(x)$ in $x_0, ..., x_n$.