

Der Hobby-Algorithmus

Philipp Stephani

14. April 2008

1 Einführung

In der zweidimensionalen Computergrafik stehen wir häufig vor dem Problem, durch eine gegebene Anzahl von Punkten eine glatte Kurve zu legen. Meist werden hierfür *Splines* verwendet. Dabei handelt es sich um miteinander verbundene Kurvenstücke. Besonders beliebt sind kubische Bézierkurven, da sie durch vier Kontrollpunkte mit einfacher geometrischer Bedeutung definiert sind und mit Hilfe des de-Casteljau-Algorithmus effizient und numerisch stabil berechnet werden können.

In [Abbildung 1](#) sehen wir zwei verbundene Bézierkurven, die beispielsweise Teil eines Béziersplines sein könnten. Die Bézierkurve auf der rechten Seite schneidet die beiden *Knotenpunkte* z_i und z_{i+1} . Die Tangenten an den Schnittpunkten schneiden die zugehörigen Kontrollpunkte p_i und q_i . Der Abstand von p_i und q_i von z_i bzw. z_{i+1} gibt grob gesprochen an, wie sehr sich die Kurve an die Tangenten anschmiegt.

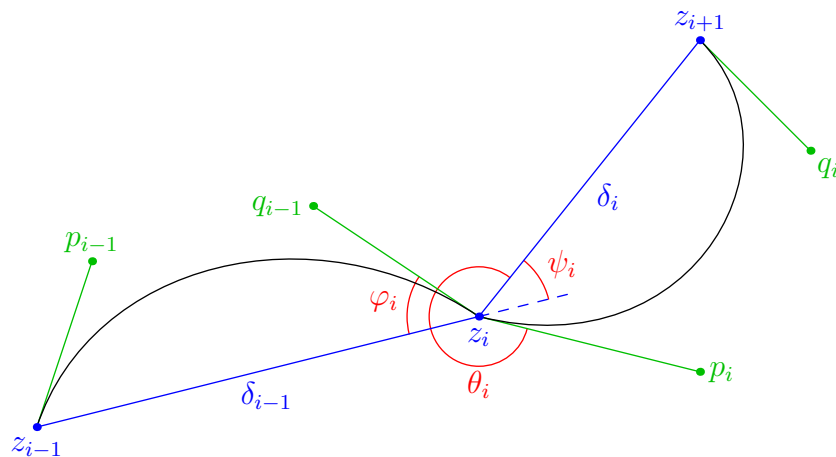


Abbildung 1: Winkel und Kontrollpunkte von Bézierkurven

Manchmal wollen wir Béziersplines durch eine Anzahl Knotenpunkte zeichnen, ohne die anderen Kontrollpunkte explizit gegeben zu haben. Vielmehr soll das resultierende

Spline die Knotenpunkte so natürlich wie möglich, d. h. mit möglichst geringer Gesamtkrümmung, interpolieren. Dies mathematisch zu quantifizieren, ist ziemlich kompliziert. Ich möchte hier lediglich als Ergebnis den *Hobby-Algorithmus* vorstellen. Dieser wurde von dem amerikanischen Mathematiker John D. HOBBY im Jahre 1985 entwickelt und von Donald E. KNUTH für sein Schriftdefinitionssystem METAFONT, welches in $\text{T}_\text{E}\text{X}$ und $\text{L}_\text{A}\text{T}_\text{E}\text{X}$ benutzt wird, eingesetzt. Verwendung findet der Algorithmus ebenfalls in den auf METAFONT basierenden Vektorgrafiksprachen METAPOST und Asymptote.

2 Definitionen

Sei $n \in \mathbb{N}$ mit $n > 1$. Gegeben seien $n + 1$ Knotenpunkte $(x_0, y_0), \dots, (x_n, y_n)$, durch die wir ein kubisches Bézierspline mit durch den Hobby-Algorithmus ermittelten Kontrollpunkten legen wollen. Zur notationellen Vereinfachung definieren wir:

$$\begin{aligned} \mathbf{z}_i &:= (x_i, y_i)^\top \text{ für } 0 \leq i \leq n \\ \delta x_i &:= x_{i+1} - x_i \text{ für } 0 \leq i < n \\ \delta y_i &:= y_{i+1} - y_i \text{ für } 0 \leq i < n \\ \boldsymbol{\delta}_i &:= \mathbf{z}_{i+1} - \mathbf{z}_i \text{ für } 0 \leq i < n \\ d_i &:= \|\boldsymbol{\delta}_i\| \text{ für } 0 \leq i < n \end{aligned}$$

Dabei bezeichnen \mathbf{v}^\top die Transposition des Vektors \mathbf{v} (sodass die \mathbf{z}_i Spaltenvektoren im \mathbb{R}^2 sind) und $\|\cdot\|$ die übliche euklidische Norm.

Grundsätzlich lassen sich zwei Typen von Splines unterscheiden: offene und geschlossene. Geschlossene Splines zeichnen sich gegenüber offenen dadurch aus, dass der letzte Punkt \mathbf{z}_n durch ein weiteres Kurvenstück wieder mit dem ersten Punkt \mathbf{z}_0 verbunden ist. Geschlossene Splines heißen auch *zyklisch*. Bezüglich des Hobby-Algorithmus unterscheiden sich offene und geschlossene Splines in einigen wichtigen Details. Das Verhalten an den *inneren* Knotenpunkten $\mathbf{z}_1, \dots, \mathbf{z}_{n-1}$ ist jedoch genau gleich, sodass wir zunächst diese Punkte betrachten und erst später zwischen den beiden Splinetypen differenzieren.

3 Die inneren Knotenpunkte

Zunächst führen wir für jeden inneren Knotenpunkt die in [Abbildung 1](#) eingezeichneten Winkel ein. ψ_i ist der Winkel zwischen der Verbindungsstrecke zwischen \mathbf{z}_i und \mathbf{z}_{i+1} sowie der Verbindungsstrecke zwischen \mathbf{z}_{i-1} und \mathbf{z}_i , d. h. zwischen den Vektoren $\boldsymbol{\delta}_i$ und $\boldsymbol{\delta}_{i-1}$. Die Winkel ψ_i werden dabei so gewählt, dass $-\pi < \psi_i \leq \pi$ gilt. θ_i ist der Winkel zwischen $\boldsymbol{\delta}_i$ und der Verbindungsstrecke von \mathbf{z}_i und \mathbf{p}_i . Schließlich ist φ_i der Winkel zwischen der Verbindungsstrecke von \mathbf{z}_{i-1} und \mathbf{z}_i und der Verbindungsstrecke von \mathbf{z}_i und \mathbf{q}_{i-1} . Zu beachten ist, dass auf diese Weise auch θ_0 und φ_n , nicht aber θ_n , φ_0 , ψ_0 und ψ_n definiert werden können.

Damit der Spline bei \mathbf{z}_i stetig differenzierbar ist, müssen \mathbf{q}_{i-1} , \mathbf{z}_i und \mathbf{p}_i auf einer Geraden liegen, und zwar derart, dass \mathbf{z}_i zwischen \mathbf{q}_{i-1} und \mathbf{p}_i liegt. Daraus folgt die

Bedingung, dass φ_i , θ_i und ψ_i sich zum Vollwinkel bzw. Nullwinkel ergänzen. Diese Bedingung werden wir später benötigen.

Durch die Winkel φ_i und θ_i ist zwar die Richtung der Kontrollpunkte \mathbf{p}_i und \mathbf{q}_{i-1} bezüglich \mathbf{z}_i gegeben, nicht aber ihr Abstand von \mathbf{z}_i . Wir benötigen also pro innerem Knotenpunkt noch zwei weitere Parameter. Diese sollen mit der lokalen Spannenergie des Splines an den Kontrollpunkten in Verbindung stehen und werden deshalb *Spannungsparameter* genannt. τ_i soll hierbei den Spannungsparameter des bei \mathbf{z}_i auslaufenden Kurvensegments bezeichnen, $\bar{\tau}_i$ entsprechend den des einlaufenden Segments. Im Hobby-Algorithmus werden die Spannungsparameter so definiert, dass sich minimale Gesamtspannungsenergie (und damit natürlichstes Aussehen) für $\tau_i = \bar{\tau}_i = 1$ ergibt.

Zunächst betrachten wir lediglich eine einzelne Bézierkurve zwischen den Punkten \mathbf{z}_i und \mathbf{z}_{i+1} . Wir können das Problem vereinfachen, indem wir ausnutzen, dass bei einer gegebenen affinen Transformation \mathcal{T} diejenige Bézierkurve, die durch die mit \mathcal{T} transformierten Kontrollpunkte definiert ist, identisch mit der durch \mathcal{T} transformierten Bézierkurve ist. Mathematisch formuliert bedeutet dies, dass für die Bézierkurve $\mathfrak{B}(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ mit den Kontrollpunkten \mathbf{a} , \mathbf{b} , \mathbf{c} und \mathbf{d} gilt:

$$\mathcal{T}(\mathfrak{B}(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})) = \mathfrak{B}(\mathcal{T}(\mathbf{a}), \mathcal{T}(\mathbf{b}), \mathcal{T}(\mathbf{c}), \mathcal{T}(\mathbf{d}))$$

Daraus folgern wir, dass wir eine Bézierkurve mit den Knotenpunkten $\tilde{\mathbf{z}}_i = (0, 0)^\top$ und $\tilde{\mathbf{z}}_{i+1} = (1, 0)^\top$ betrachten und die ermittelten Kontrollpunkte mit der affinen winkeltreuen Transformation \mathcal{T}_i , die $\tilde{\mathbf{z}}_i$ auf \mathbf{z}_i und $\tilde{\mathbf{z}}_{i+1}$ auf \mathbf{z}_{i+1} abbildet, transformieren können. \mathcal{T}_i ist eindeutig bestimmt durch

$$\mathcal{T}_i(\mathbf{v}) = \mathbf{z}_i + \begin{pmatrix} \delta x_i & -\delta y_i \\ \delta y_i & \delta x_i \end{pmatrix} \mathbf{v}.$$

Für die zwei untransformierten Kontrollpunkte $\tilde{\mathbf{p}}_i$ und $\tilde{\mathbf{q}}_i$ gilt dann:

$$\begin{aligned} \tilde{\mathbf{p}}_i &= \tilde{\mathbf{z}}_i + \xi_i \cdot \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix} = \begin{pmatrix} \xi_i \cos \theta_i \\ \xi_i \sin \theta_i \end{pmatrix}, \\ \tilde{\mathbf{q}}_i &= \tilde{\mathbf{z}}_{i+1} + \zeta_i \cdot \begin{pmatrix} -\cos \varphi_{i+1} \\ \sin \varphi_{i+1} \end{pmatrix} = \begin{pmatrix} 1 - \zeta_i \cos \varphi_{i+1} \\ \zeta_i \sin \varphi_{i+1} \end{pmatrix}. \end{aligned}$$

Dabei sind ξ_i der Abstand zwischen \mathbf{z}_i und \mathbf{p}_i sowie ζ_i der Abstand zwischen \mathbf{z}_{i+1} und \mathbf{q}_i . Hat man $\tilde{\mathbf{p}}_i$ und $\tilde{\mathbf{q}}_i$ erst einmal gefunden, ergeben sich die gesuchten Kontrollpunkte \mathbf{p}_i und \mathbf{q}_i einfach durch Anwendung von \mathcal{T}_i :

$$\mathbf{p}_i = \mathcal{T}_i(\tilde{\mathbf{p}}_i), \quad \mathbf{q}_i = \mathcal{T}_i(\tilde{\mathbf{q}}_i).$$

Die eigentliche Schwierigkeit am Hobby-Algorithmus ist die Ermittlung von ξ_i , ζ_i , φ_i und θ_i . HOBBY entwickelt in seiner Publikation [1] zunächst eine Methode, um ξ_i und ζ_i aus den Winkeln zu berechnen. Es gilt:

$$\xi_i := \frac{\rho(\theta_i, \varphi_{i+1})}{3\tau_i}, \quad \zeta_i := \frac{\sigma(\theta_i, \varphi_{i+1})}{3\bar{\tau}_{i+1}}.$$

Entscheidend für die Qualität des Ergebnisses ist die Auswahl der Funktionen ρ und σ . HOBBY wählt:

$$\rho(\theta, \varphi) := \frac{2 + \alpha(\theta, \varphi)}{1 + (1 - c) \cos \theta + c \cos \varphi},$$

$$\sigma(\theta, \varphi) := \frac{2 - \alpha(\theta, \varphi)}{1 + (1 - c) \cos \varphi + c \cos \theta}$$

mit

$$\alpha(\theta, \varphi) := a \cdot (\sin \theta - b \sin \varphi) \cdot (\sin \varphi - b \cos \theta) \cdot (\cos \theta - \cos \varphi).$$

Hierbei sind a , b und c experimentell zu ermittelnde Konstanten; METAFONT und Asymptote wählen:

$$a := \sqrt{2}, \quad b := \frac{1}{16}, \quad c := \frac{3 - \sqrt{5}}{2}.$$

Damit ist das Problem zumindest für die inneren Knotenpunkte bis auf die Bestimmung der Winkel φ_i und θ_i gelöst. Für letzteres führt HOBBY eine *Krümmungsfunktion* k ein:

$$k(\theta, \varphi, \tau, \bar{\tau}) := \frac{2\tau^2}{\rho(\theta, \varphi)^2} \left(\frac{\sigma(\theta, \varphi) \sin(\theta + \varphi)}{\bar{\tau}} - 3 \sin \theta \right)$$

Diese Funktion ist jedoch nichtlinear in den Variablen θ und φ , was auf ein schwierig zu lösendes nichtlineares Gleichungssystem für θ_i und φ_i hinausliefere. Um dies zu umgehen, wird k nach (θ, φ) entwickelt und nur der lineare Term $\hat{k}(\theta, \varphi, \tau, \bar{\tau})$ benutzt. Man erhält unter Benutzung der linearen Näherungen $\sin \theta \approx \theta$ und $\cos \theta \approx 1$:

$$\alpha(\theta, \varphi) \approx a \cdot (\theta - b\varphi) \cdot (\varphi - b) \cdot (1 - 1) = 0,$$

$$\rho(\theta, \varphi) \approx \frac{2}{1 + 1 - c + c} = 1,$$

$$\sigma(\theta, \varphi) \approx \frac{2}{1 + 1 - c + c} = 1,$$

$$\hat{k}(\theta, \varphi, \tau, \bar{\tau}) = 2\tau^2 \left(\frac{\theta + \varphi}{\bar{\tau}} - 3\theta \right).$$

Diese Linearisierung führt dazu, dass die Interpolation bei sehr großen Winkeln θ , φ nicht mehr optimal ist. Derartig große Winkel kommen jedoch in der Praxis relativ selten vor, sodass die linearisierte Krümmungsfunktion für unsere Zwecke eine mehr als ausreichende Näherung darstellt.

Damit die Krümmung an den Knotenpunkten übereinstimmt, muss

$$\frac{k(\theta_i, \varphi_{i+1}, \tau_i, \bar{\tau}_{i+1})}{d_i} = \frac{k(\varphi_i, \theta_{i-1}, \bar{\tau}_i, \tau_{i-1})}{d_{i-1}}$$

gelten. Aus der Stetigkeit der linearisierten Krümmungsfunktion ergibt sich wiederum:

$$\begin{aligned} \frac{\hat{k}(\theta_i, \varphi_{i+1}, \tau_i, \bar{\tau}_{i+1})}{d_i} &= \frac{\hat{k}(\varphi_i, \theta_{i-1}, \bar{\tau}_i, \tau_{i-1})}{d_{i-1}} \\ \frac{\tau_i^2}{d_i} \left(\frac{\theta_i + \varphi_{i+1}}{\bar{\tau}_{i+1}} - 3\theta_i \right) &= \frac{\bar{\tau}_i^2}{d_{i-1}} \left(\frac{\varphi_i + \theta_{i-1}}{\tau_{i-1}} - 3\varphi_i \right) \\ \frac{\tau_i^2}{d_i \bar{\tau}_{i+1}} \varphi_{i+1} + \frac{\tau_i^2}{d_i} \left(\frac{1}{\bar{\tau}_{i+1}} - 3 \right) \theta_i &= \frac{\bar{\tau}_i^2}{d_{i-1} \tau_{i-1}} \theta_{i-1} + \frac{\bar{\tau}_i^2}{d_{i-1}} \left(\frac{1}{\tau_{i-1}} - 3 \right) \varphi_i \\ d_{i-1} \tau_{i-1} \tau_i^2 \varphi_{i+1} + d_{i-1} \tau_{i-1} \tau_i^2 (1 - 3\bar{\tau}_{i+1}) \theta_i &= d_i \bar{\tau}_{i+1} \bar{\tau}_i^2 \theta_{i-1} + d_i \bar{\tau}_{i+1} \bar{\tau}_i^2 (1 - 3\tau_{i-1}) \varphi_i \end{aligned}$$

Mit den Abkürzungen

$$\begin{aligned} \lambda_i &:= 3\tau_i - 1, & \mu_i &:= 3\bar{\tau}_i - 1, \\ C_i &:= d_{i-1} \tau_i^2 \tau_{i-1}, & A_i &:= d_i \bar{\tau}_i^2 \bar{\tau}_{i+1} \end{aligned}$$

vereinfacht sich die letzte Gleichung zu:

$$A_i \theta_{i-1} + C_i \mu_i \theta_i - A_i \lambda_i \varphi_i - C_i \varphi_{i+1} = 0 \quad (1)$$

Weiterhin gilt, wie oben erwähnt, aufgrund der stetigen Differenzierbarkeit an den Knotenpunkten:

$$\varphi_i = -\psi_i - \theta_i \quad (2)$$

Dies alles gilt natürlich nur an den inneren Knotenpunkten. Wir haben aus der Stetigkeit der Krümmungsfunktion $n - 1$ und aus der stetigen Differenzierbarkeit des Splines ebenfalls $n - 1$, insgesamt also $2n - 2$ Bedingungen erhalten. Für die n Kurvensegmente benötigen wir je zwei zusätzliche Kontrollpunkte. Die Abstände ξ_i und ζ_i der Kontrollpunkte von den Knotenpunkten lässt sich wie oben gezeigt aus den Winkeln θ_i und φ_i berechnen, somit benötigen wir pro Kontrollpunkt nur eine Gleichung (für den jeweils anliegenden Winkel). Dennoch fehlen immer noch zwei Bedingungen, für die wir die äußeren Knotenpunkte \mathbf{z}_0 und \mathbf{z}_n betrachten und erstmals zwischen offenen und geschlossenen Splines unterscheiden müssen.

4 Offene Splines

Hier müssen wir zwei weitere Gleichungen künstlich einführen, da wird keine periodischen Randbedingungen haben. Analog zu den Spannungsparametern definiert HOBBY die *Wirbelparameter* χ_0 und χ_n derart, dass sich für $\chi_0 = \chi_n = 1$ ein natürliches Aussehen ergibt. Die Wirbelparameter beeinflussen das Aussehen des Splines an den Endpunkten. Für \mathbf{z}_0 gibt es darüberhinaus einen weiteren Ausgangsspannungsparameter τ_0 ; analog existiert für \mathbf{z}_n ein Eingangsspannungsparameter $\bar{\tau}_n$. Man wählt die folgenden Randbedingungen:

$$\begin{aligned} \hat{k}(\theta_0, \varphi_1, \tau_0, \bar{\tau}_1) = \chi_0 \hat{k}(\varphi_1, \theta_0, \bar{\tau}_1, \tau_0) &\iff \theta_0 = F_1 \varphi_1, \\ \hat{k}(\varphi_n, \theta_{n-1}, \bar{\tau}_n, \tau_{n-1}) = \chi_n \hat{k}(\theta_{n-1}, \varphi_n, \tau_{n-1}, \bar{\tau}_n) &\iff \varphi_n = G_{n-1} \theta_{n-1}. \end{aligned}$$

mit

$$F_1 := \frac{\tau_0^3 + \lambda_0 \chi_0 \bar{\tau}_1^3}{\mu_1 \tau_0^3 + \chi_0 \bar{\tau}_1^3}, \quad G_{n-1} := \frac{\bar{\tau}_n^3 + \mu_n \chi_n \tau_{n-1}^3}{\lambda_{n-1} \bar{\tau}_n^3 + \chi_n \tau_{n-1}^3}.$$

Damit lassen sich θ_0 und φ_n aus den anderen Gleichungen eliminieren:

$$\begin{aligned} 0 &= C_1 \mu_1 \theta_1 + A_1 (F_1 - \lambda_1) \varphi_1 - C_1 \varphi_2 \\ &= (A_1 (\lambda_1 - F_1) + C_1 \mu_1) \theta_1 - C_1 \varphi_2 + A_1 (\lambda_1 - F_1) \psi_1, \\ 0 &= A_{n-1} \theta_{n-2} + C_{n-1} (\mu_{n-1} - G_{n-1}) \theta_{n-1} - A_{n-1} \lambda_{n-1} \varphi_{n-1} \\ &= A_{n-1} \theta_{n-2} + (C_{n-1} (\mu_{n-1} - G_{n-1}) + A_{n-1} \lambda_{n-1}) \theta_{n-1} + A_{n-1} \lambda_{n-1} \psi_{n-1}. \end{aligned} \tag{3}$$

Für $n > 2$ lässt sich φ_2 in (3) ebenfalls durch θ_2 ersetzen:

$$(C_1 \mu_1 + A_1 (\lambda_1 - F_1)) \theta_1 + C_1 \theta_2 + A_1 (\lambda_1 - F_1) \psi_1 + C_1 \psi_2 = 0$$

Weiterhin gilt für $1 < i < n - 1$:

$$A_i \theta_{i-1} + (A_i \lambda_i + C_i \mu_i) \theta_i + C_i \theta_{i+1} + A_i \lambda_i \psi_i + C_i \psi_{i+1} = 0.$$

Definiert man nun

$$\begin{aligned} F_i &:= 0 & (1 < i < n), & & G_i &:= 0 & (0 < i < n - 1), \\ \psi_n &:= 0, \\ Q_i &:= A_i (\lambda_i - F_i), & & & R_i &:= C_i (\mu_i - G_i), \\ B_i &:= Q_i + R_i, & & & D_i &:= -Q_i \psi_i - C_i \psi_{i+1}, \end{aligned}$$

so gilt:

$$\begin{aligned} B_1 \theta_1 + C_1 \theta_2 &= D_1, \\ A_i \theta_{i-1} + B_i \theta_i + C_i \theta_{i+1} &= D_i & (i < 1 < n - 1), \\ A_n \theta_{n-1} + B_n \theta_n &= D_n. \end{aligned}$$

Dies ist ein lineares tridiagonales Gleichungssystem der Form

$$\begin{pmatrix} B_1 & C_1 & & & & & \\ A_2 & B_2 & C_2 & & & & \\ & & \ddots & & & & \\ & & & \cdots & A_{n-2} & B_{n-2} & C_{n-2} \\ & & & & \cdots & A_{n-1} & B_{n-1} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{n-2} \\ \theta_{n-1} \end{pmatrix} = \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_{n-2} \\ D_{n-1} \end{pmatrix}$$

mit Subdiagonale (A_i), Diagonale (B_i), Superdiagonale (C_i) und rechtsseitigem Vektor (D_i), welches mit dem Gaußalgorithmus in linearer Zeit gelöst werden kann [2, 3].

Sind alle τ_i und $\bar{\tau}_i$ größer als $\frac{2}{3}$, so gilt wegen $C_i > 0$, $A_i > 0$, $\lambda_i > 1$ und $\mu_i > 1$ an den inneren Knotenpunkten:

$$|B_i| = \lambda_i A_i + \mu_i C_i > A_i + C_i = |A_i| + |C_i|.$$

An den Randpunkten folgt analog, falls zusätzlich $\lambda_1 > F_1$ und $\mu_{n-1} > G_{n-1}$ gelten:

$$\begin{aligned} |B_1| &= A_1(\lambda_1 - F_1) + C_1\mu_1 > |C_1|, \\ |B_{n-1}| &= A_{n-1}\lambda_{n-1} + C_{n-1}(\mu_{n-1} - G_{n-1}) > |A_{n-1}|. \end{aligned}$$

Somit ist die Systemmatrix diagonaldominant. Das führt dazu, dass in diesem (in der Realität meistens angenommenen) Fall keine Pivotisierung nötig ist [2].

5 Geschlossene Splines

Hier haben wir periodische Randbedingungen vorliegen, weshalb es sich anbietet, die Indizes zyklisch zu definieren:

$$\begin{aligned} \boldsymbol{\delta}_n &:= z_0 - z_n, & d_n &:= \|\boldsymbol{\delta}_n\|, \\ d_{-1} &:= d_n, & d_{n+1} &:= d_0, \\ \theta_{-1} &:= \theta_n, & \theta_{n+1} &:= \theta_0, \\ \varphi_{-1} &:= \varphi_n, & \varphi_{n+1} &:= \varphi_0, \\ \tau_{-1} &:= \tau_n, & \bar{\tau}_{n+1} &:= \bar{\tau}_0. \end{aligned}$$

Weiterhin definieren wir ψ_0 und ψ_{n+1} als den Winkel zwischen $\boldsymbol{\delta}_n$ und $\boldsymbol{\delta}_0$ sowie ψ_n als den Winkel zwischen $\boldsymbol{\delta}_{n-1}$ und $\boldsymbol{\delta}_n$. Damit können wir die Gleichungen (1) und (2) auch für $i = 0$ und $i = n$ anwenden. Mit den Definitionen

$$\begin{aligned} Q_i &:= A_i\lambda_i, & R_i &:= C_i\mu_i, \\ B_i &:= Q_i + R_i, & D_i &:= -Q_i\psi_i - C_i\psi_{i+1} \end{aligned}$$

erhalten wir wieder

$$A_i\theta_{i-1} + B_i\theta_i + C_i\theta_{i+1} = D_i.$$

Für $i = 0$ und $i = n$ gelten mit obigen Definitionen die Spezialfälle:

$$\begin{aligned} B_0\theta_0 + C_0\theta_1 + A_0\theta_n &= D_0, \\ C_n\theta_0 + A_n\theta_{n-1} + B_n\theta_n &= D_n. \end{aligned}$$

Aufgrund der Periodizität des Problems ist die Gleichung für $i = n + 1$ identisch zu der Gleichung für $i = 0$. Damit haben wir also für die $n + 1$ Winkel $\theta_0, \dots, \theta_n$ (ein Winkel pro Knotenpunkt) $n + 1$ Gleichungen gefunden, die nun aber kein einfaches tridiagonales Gleichungssystem mehr, sondern ein *zyklisches* tridiagonales Gleichungssystem der Form

$$\begin{pmatrix} B_0 & C_0 & \cdots & & & & A_0 \\ A_1 & B_1 & C_1 & \cdots & & & \\ & & & \ddots & & & \\ & & & \cdots & A_{n-1} & B_{n-1} & C_{n-1} \\ C_n & & \cdots & & A_n & B_n & \end{pmatrix} \begin{pmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{n-1} \\ \theta_n \end{pmatrix} = \begin{pmatrix} D_0 \\ D_1 \\ \vdots \\ D_{n-1} \\ D_n \end{pmatrix}$$

bilden. Unter Zuhilfenahme der Sherman-Morrison-Formel [3] kann auch dieses System mit linearer Zeitkomplexität gelöst werden.

Literatur

- [1] John D. Hobby: Smooth, Easy to Compute Interpolating Splines, 1985
<ftp://db.stanford.edu/pub/cstr/reports/cs/tr/85/1047/>
- [2] Martin Kerscher, Joachim Puls, Sigmung Stintzing: Numerik für Physiker, 2007
<http://www.math.lmu.de/~kerscher/vorlesungen/numerikws0708/skript.html>
- [3] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery:
Numerical Recipes in C++, Second Edition, Cambridge University Press