Ein Algorithmus zur stückweise linearen Approximation implizit gegebener Mannigfaltigkeiten

Diplomarbeit von Frank Koster

Aufgabensteller: Prof. Dr. H.G Jeggle
Betreuer: Dr. M. Hinze

Inhaltsverzeichnis

Kapitel 1

Einleitung

Gegenstand der vorliegenden Arbeit soll die möglichst effiziente Lösung
der Gleichung

$$H : \bar{\Omega} \subset \mathbf{R}^{n+k} \to \mathbf{R}^n$$
$$H(x) = 0 \tag{1.1}$$

sein. Also die Bestimmung der Mannigfaltigkeit $M := H^{-1}(0)$.
Ω ist dabei ein beschränktes Gebiet und H eine zunächst noch beliebige Funktion. Eine erste Lösbarkeitsaussage bzgl. (1.1) gibt der Satz über implizite Funktionen (siehe [1] §8 Satz 1 und 2).

Theorem 1.1 *Sei H stetig differenzierbar und $x_0 \in \Omega$ mit $H(x_0) = 0$ sowie $rgDH(x_0) = n$, dann existiert $U \subset \mathbf{R}^k$, eine Umgebung der Null, und $c : U \to \mathbf{R}^{n+k}$ stetig differenzierbar mit $c(0) = x_0$ und $H(c(s)) = 0 \ \forall s \in U$.*

Für den Fall $k = 1$ führt eine solche Parametrisierung der Lösungsmenge auf die für 'glatte' Funktionen gebräuchlichsten Verfahren.

1.1 Prädiktor-Korrektor Methoden

Nach einer ursprünglichen Idee von Davidenko (1953) kann die Parametrisierung c als Lösung der Davidenko Differentialgleichung

$$\begin{aligned} DH(c(s)) \cdot \dot{c}(s) &= 0 \quad s \in U \quad \text{eine Nullumgebung} \\ \|\dot{c}\| &= 1 \\ c(0) &= x_0 \end{aligned} \tag{1.2}$$

erhalten werden. Ein Hilfsmittel zur Lösung dieser Dgl. ist der für Matrizen $A \in M(n+1, n)$ mit maximalem Rang durch

$$\begin{aligned} A \cdot t(A) &= 0 \\ \|t(A)\| &= 1 \\ det(A^T, t(A)) &> 0 \end{aligned} \tag{1.3}$$

eindeutig definierte Tangantialvektor. Es läßt sich zeigen, daß die Abbildung $A \mapsto t(A)$ auf der Menge $R = \{A \in M(n+1, n) \mid rgA = n\}$ stetig differenzierbar ist. (siehe [11, Lemma 2.18])
Prädiktor-Korrektor (PC) Methoden berechnen Approximationen $\{x_i\}_{0 \le i \le N}$ von Lösungspunkten im Wesentlichen durch folgendes Schema

Algorithmus 1.2 (elementares PC Schema)

$$\begin{array}{ccc} x_0 & & \textit{gegebener erster Lösungspunkt} \\ y_{i+1} & := & x_i + ht(DH(x_i)) \\ x_{i+1} & := & y_{i+1} - \begin{pmatrix} DH(x_i) \\ t(DH(x_i)) \end{pmatrix}^{-1} \cdot \begin{pmatrix} H(x_i) \\ 0 \end{pmatrix} \end{array}$$

Dabei ist h eine gegebene Schrittweite. Ein Konvergenzbeweis für dieses einfache Verfahren findet sich bei Allgower und Georg [11, Theorem 5.2.1.]
Dieses Verfahren kann modifiziert werden durch

- mehrere Iterationen im Korrektorschritt
- Verwendung von Approximationen von $DH(x_i)$ bzw. $t(DH(x_i))$
- Schrittweitensteuerungen
- Prädiktor höherer Ordnung

Alle diese Modifikationen lassen sich auch einfach implementieren.

Im Prinzip können PC Methoden auch auf den Fall $k \geq 2$ verallgemeinert werden. Das erste dieser Verfahren ist der Moving Frame Algorithmus von W.C. Rheinboldt [15].
Dabei wird versucht, schrittweise Punkten y_i einer Triangulierung $\mathcal{T}$ des $\mathbf{R}^k$ Lösungspunkte auf M zuzordnen. Dies geschieht mittels einer (hier vereinfachend) als Moving Frame bezeichneten QR-Zerlegung von $DH(x_i)$, welche zur Berechnung der Tangentialvektoren und zur Durchführung der Korrektorschritte benutzt wird.
Der Moving Frame wird für alle noch nicht mit einer Zuordnung versehenen Punkte $\{y_{j_1}, ..., y_{j_l} \in \mathcal{T} \mid \|y_i - y_{j_s}\| \leq \delta_{Frame}\}$ benutzt.
Problematisch ist dieses Vorgehen, wenn M diffeomorph zu einer Kugeloberfläche ist. Gegeben seien bspw. Lösungspunkte x_i, x_j, die auf M so dicht beieinander liegen, daß man sie als benachbart bezeichnen würde. Ihre 'Urbilder' $y_i, y_j \in \mathcal{T}$ können aber weit auseinaderliegen, so daß die Nachbarschaft nicht erkannt wird. Überlappungen sind also recht wahrscheinlich.

Bei den hier vorgestellten Methoden sind diese Probleme entweder nicht vorhanden, oder aber einfacher zu lösen.

1.2 Simpliziale Methoden

Hier wird $\bar{\Omega}$ durch $\mathcal{T} = \{\sigma \subset \bar{\Omega} \mid \sigma \quad n+k\text{-Simplex}\}$ trianguliert. Die älteste dieser Methoden geht auf Sperner [3] zurück. Eine genaue Analyse findet sich bei Allgower/Schmidt [6].
Für alle $x \in \Omega$ wird eine ganzzahlige Bewertungsfunktion definiert durch

$$l(x) = i + 1 :\Leftrightarrow (H(x))_1, .., (H(x))_i \geq 0 \text{ und sofern } i < n \text{ gilt } (H(x))_{i+1} < 0 \quad .$$

Ein Simplex σ mit Ecken $v_0, .., v_r$ heißt dann vollständig bewertet, falls

$$\{l(v_0), ..., l(v_r)\} = \{1, ..., n+1\} \quad .$$

Es kann gezeigt werden, daß jeder vollständig bewertete (v.b.) $n+k$-Simplex zwischen $k+1$ und $2k$ v.b Facetten hat. Außerdem ist ein $n+k$ Simplex genau dann v.b., wenn er über v.b. Facetten mit anderen v.b. Simplices benachbart ist. Auf diese Weise lassen sich schrittweise alle zusammenhängenden v.b. Simplices ermitteln. Die Kontrolle darüber, ob ein Simplex schon durchlaufen wurde oder nicht, ist recht einfach und wird im Kapitel 3 und 6 behandelt.
Als Approximation von $H^{-1}(0) \bigcap \sigma$ kann bspw. die konvexe Hülle der Schwerpunkte der v.b. $n+1$-Flächen $\tau_1, .., \tau_l \quad (k+1 \leq l \leq 2^k)$ benutzt werden.
Der Vorteil dieser Methode liegt darin, daß keine linearen Gleichungssysteme gelöst werden müssen

und daß an H keine Glattheitsvorausetzungen bestehen. Dieser Vorteil kann aber auch als Nachteil betrachtet werden. Die Glattheit der Lösung wird nicht ausgenutzt, was zu schlechten Fehlerschätzungen und zu einer sehr großen Anzahl von zu durchlaufenden Simplices führt.
Die in diesem Sinne günstigeren Methoden mit Vektor-Bewertung sind das zentrale Thema dieser Arbeit. Dazu werden im 2. Kapitel Triangulierungen und deren Eigenschaften vorgestellt. Im 3. Kapitel wird das Konzept der Vektor-Bewertung und bisherige Algorithmen dargelegt. In den folgenden Kapiteln wird eine neue Variante dieser Verfahren behandelt.

Kapitel 2

Triangulierungen

Zunächst zur Begriffsbildung, die weitgehend an [4] angelehnt ist.

Definition 2.1 *Eine Menge von Punkten* $\{v_i\}_{0\leq i\leq r} \subset \mathbf{R}^m$ *heißt affin unabhängig, wenn die Menge* $\{v_i - v_0\}_{1\leq i\leq r}$ *linear unabhängig ist.*
$\sigma := [v_0, ..., v_r]$ *- die konvexe Hülle der* v_i *- heißt* *r-Simplex* *mit den* *Ecken* $\{v_i\}_{0\leq i\leq r}$.
Wenn $\{w_l\}_{0\leq l\leq s} \subset \{v_i\}_{0\leq i\leq r}$ *dann heißt* $\tau := [w_0, ..., w_s]$ *s-Fläche* *von* σ.
$(r-1)$*-Flächen von* σ *heißen* *Facetten* *von* σ.
Wenn $\{v_{i_j}\}_{1\leq j\leq r-s}$ *die Ecken von* σ *sind, die nicht zu* τ *gehören, dann wird* τ *geschrieben als* $\tau = [\sigma; \hat{v}_{i_1}, .., \hat{v}_{i_{r-s}}]$. $\hat{}$ *bezeichnet das Weglassen der entsprechenden Ecke und* $\{v_{i_j}\}_{1\leq j\leq r-s}$ *sind die* *Fehlecken* *von* τ.
$b(\sigma) := \frac{1}{r+1}\sum_{i=0}^{r} v_i$ *ist der* *Schwerpunkt* *von* σ.
Eine *Triangulierung* $\mathcal{T}$ *von* $\mathbf{R}^m$ *ist eine Zerlegung von* $\mathbf{R}^m$ *in* m*-Simplices, so daß*
(i) $\sigma_1, \sigma_2 \in \mathcal{T}$ *dann ist* $\sigma_1 \bigcap \sigma_2 = \emptyset$ *oder eine gemeinsame Fläche von* σ_1 *und* σ_2
(ii) Jede beschränkte Menge $\mathbf{R}^m$ *schneidet nur endlich viele Simplices aus* $\mathcal{T}$

Da sich r-Flächen als Schnitt von m-Simplices darstellen lassen, folgt aus dieser Definition, daß auch der Schnitt zweier r-Flächen τ_1, τ_2 eine gemeinsame Fläche von τ_1, τ_2 oder $\emptyset$ sein muß.
Da der Schwerpunkt einer r-Fläche in keiner anderen r-Fläche enthalten ist, kann eine r-Fläche durch ihren Schwerpunkt eindeutig identifiziert werden.
Es gibt eine Reihe von Triangulierungen des $\mathbf{R}^n$, deren Ecken gerade im $\mathbf{Z}^n$ liegen. Bedeutsam sind vorallem Todd's Triangulierung $\mathcal{J}$ (siehe [9]) und die Freudenthal-Kuhn Triangulierung $\mathcal{K}$.
In der hier vorliegenden Arbeit wird aber ausschließlich ein affines Bild der Freudenthal-Kuhn Triangulierung benutzt.

Definition 2.2 (Die Freudenthal-Kuhn Triangulierung $\mathcal{K}$ des $\mathbf{R}^n$)
Ein m*-Simplex* $\sigma = [v_0, ..., v_m]$ *ist genau dann in* $\mathcal{K}$, *wenn* $v_0 \in \mathbf{Z}^m$ *ist und eine Permutation* $\pi \in S(m)$ *existiert, so daß für* $0 \leq i \leq m-1$ *gilt* $v_{i+1} = v_i + e_{\pi(i+1)}$.
Es wird die Schreibweise $\sigma =: [v_0, \pi]$ *benutzt.*

Wie bereits in der Einleitung bei der Beschreibung der ganzzahligen Bewertung angedeutet, wird bei simplizialen Methoden von einem m-Simplex über eine Facette zum nächsten Simplex gegangen. Dieser Schritt wird in dem folgenden Lemma analysiert. Gleichzeitig ergibt sich, daß sich mittels der darin angegeben Regeln ausgehend von einem Start-Simplex alle Simplices aus $\mathcal{K}$ durchlaufen lassen.

Lemma 2.3 ([4] Prop. 2 Pivot-Regeln für $\mathcal{K}$)
Sei $\sigma = [v_0, ..., v_m] = [v_0, \pi] \in \mathcal{K}$, $0 \leq j \leq m$ *,dann ist* $\sigma^* = [v_0, .., r(v_j), .., v_m]$ *auch aus* $\mathcal{K}$.
Hierbei bezeichnet $r(v_j) := v_{j-1} - v_j + v_{j+1}$ *die* *Reflexion* *von* v_j *über die Facette* j.
In dieser Definition muß $\{v_{-1} := v_m; v_{m+1} := v_0\}$ *gesetzt werden.*
Dieser sogenannte *Pivotschritt* *wird auch durch* $\sigma \xrightarrow{j} \sigma^*$ *dargestellt.*

Wenn $\sigma^* = [v_0^*, ..., v_m^*] = [v_0^*; \pi^*]$, $b = b(\sigma)$ *und* $b^* = b(\sigma^*)$ *ist, dann gilt:*

- $j = 0$: $v_0^* = v_1, \pi^*(i) = \begin{cases} \pi(i+1) & \textit{für } i < m \\ \pi(1) & \textit{für } i = m \end{cases}$
  also $v_0^* = v_1, ...v_{m-1}^* = v_m, v_m^* = r(v_0)$, $b^* = b + \mathbf{1}_m + e_{\pi(1)}$

- $0 < j < m$: $v_0^* = v_0, \pi^*(i) = \begin{cases} \pi(i) & \textit{für } i \neq j, j+1 \\ \pi(j+1) & \textit{für } i = j \\ \pi(j) & \textit{für } i = j+1 \end{cases}$
  also $r(v_j) = v_j^*$ *else* $v_i^* = v_i$, $b^* = b + e_{\pi(j+1)} - e_{\pi(j)}$

- $j = m$: $v_0^* = r(v_m), \pi^*(i) = \begin{cases} \pi(m) & \textit{für } i = 1 \\ \pi(i-1) & \textit{für } i > 1 \end{cases}$
  also $v_0^* = r(v_m), v_1^* = v_0, ..., v_m^* = v_{m-1}$, $b^* = b - \mathbf{1}_m - e_{\pi(m)}$

Wobei $\mathbf{1}_m := (1, ..., 1)^T \in \mathbf{R}^m$ ist und mit e_i der $i$te kanonische Einheitsvektor bezeichnet wurde. Die folgende Verallgemeinerung von [6, Lemma 5.6] zeigt, daß Simplices durch ihren Schwerpunkt eindeutig identifiziert werden können.

Lemma 2.4 (Rekonstruktion von v_0, π)
Sei $x \in \mathbf{R}^m$ *und* w *definiert durch* $w_i := \begin{cases} x_i - \lfloor x_i \rfloor & \textit{für } x_i > \lfloor x_i \rfloor \\ 1 \textit{ oder } 0 & \textit{für } x_i = \lfloor x_i \rfloor \end{cases}$.
π *sei eine Permutation , so daß* $1 \geq w_{\pi(1)} \geq ... \geq w_{\pi(m)} \geq 0$.
Dann gilt $x \in [x - w, \pi] =: \sigma$

Beweis Zunächst werden Zahlen λ_i definiert durch

$$\lambda_m := w_{\pi(m)} \quad ; \quad \lambda_i := w_{\pi(i)} - w_{\pi(i+1)} \geq 0 \quad \text{und} \quad \lambda_0 := 1 - \lambda_1 - ... - \lambda_m \quad .$$

Da $\lambda_m + ... + \lambda_k = w_{\pi(k)} \leq 1$, gilt $\lambda_0 \geq 0$ (setze $k = 1$). Außerdem gilt $\sum_{i=0} \lambda_i = 1$.
Die λ_i bilden gerade die baryzentrischen Koordinaten von x in $\sigma =: [v_0, ..., v_m]$, denn

$$x = x - w + w = x - w + \textstyle\sum_{k=1}^m e_{\pi(k)} w_{\pi(k)} =$$

$$x - w + \textstyle\sum_{k=1}^m e_{\pi(k)} \left(\sum_{i=k}^m \lambda_i\right) = x - w + \sum_{k=1}^m \lambda_k \left(\sum_{i=1}^k e_{\pi(i)}\right) = \sum_{k=0}^m \lambda_k v_k \quad ,$$

wenn $v_0 = x - w$ und $v_i = v_0 + \sum_{k=1}^i e_{\pi(k)}$ gesetzt wird.

Mit Hilfe obigen Schemas können v_0, π und so alle Ecken von $\sigma \in \mathcal{K}$ aus dem Schwerpunkt berechnet werden, da in diesem Fall w und π eindeutig sind. Für die Implementierung einer simplizialen Methode wird $\mathcal{K}$ auf $a(\mathcal{K})$, eine Triangulierung mit i.A. feineren Simplices, abgebildet. Dabei ist a eine affine Abbildung

$$a : \mathbf{R}^m \to \mathbf{R}^m; a(x) = W \cdot x + z \quad .$$

Für spätere Zwecke werden die äußeren Normalen-Einheitsvektoren (aNev) auf Simplices aus $a(\mathcal{K})$ benötigt.

Lemma 2.5 (äußere Normalenvektoren auf Facetten)
Sei $\sigma = [0, id] \in \mathcal{K}$ *und* $\tau = [v_{j_0}, ..., v_{j_{m-1}}] = [\sigma; \hat{v}_i]$ $(j_0 < ... < j_{m-1})$ *eine Facette, dann ist*

$$w_i := \begin{cases} e_1 & : \textit{für } i = 0 \\ (e_{i+1} - e_i)/\sqrt{2} & : \textit{für } 0 < i < m \quad \textit{der aNev auf } \tau \\ -e_m & : \textit{für } i = m \end{cases} .$$

Für die Normalenvektoren u_i *auf* $a([[0, id]; \hat{v}_i])$ *gilt die Transformationsformel* $u_i = W^{-T} \cdot w_i$

Der Beweis ist im Anhang zu finden.
Speziell für $W = diag(\delta_1, ..., \delta_m)$ sind die aNev von $a([v_0, \pi])$ durch

$$w_i := \begin{cases} e_{\pi(1)} & : \quad \text{für } i = 0 \\ (-\delta_{\pi(i+1)} e_{\pi(i)} + \delta_{\pi(i)} e_{\pi(i+1)}) \frac{1}{\sqrt{\delta^2_{\pi(i+1)} + \delta^2_{\pi(i)}}} & : \quad \text{für } 0 < i < m \\ -e_{\pi(m)} & : \quad \text{für } i = m \end{cases} \tag{2.1}$$

gegeben.

Abgeschlossen werden soll dieses Kapitel mit einer Betrachtung der prinzipiellen Effizienz von Triangulierungen und damit von Simplizialen Methoden.
Wie aus den Ausführungen des nächsten Kapitels ersichtlich, ist dabei von entscheidender Bedeutung die mittelere Anzahl aid_k von $m - k$ Flächen, die ein Einheitsstück einer k-dimensionalen Hyperebene schneiden.
J.C. Alexander [17] hat für diese Größe eine Formel angegeben, deren asymptotische Entwicklung für $m \to \infty$

$$aid_k \sim C(k) \cdot m^{3k/2} \tag{2.2}$$

lautet.
Ist bspw. die Lösungsmenge einer parameterabhängigen gewöhnlichen DGL gesucht, dann führt eine Halbierung des Diskretisierungsparameters h zu einer Verdopplung der Anz. d. Unbekannten und damit etwa zu einer 2.8-fachen Anzahl von zu durchlaufenden Simplices.

Kapitel 3

Vektor-Bewertung

Die Idee bei Methoden mit Vektor-Bewertung ist, auf jedem Simplex $\sigma = [v_0, ..., v_{n+k}]$ H zu approximieren durch die eindeutige affine Abb. H_σ, mit $H_\sigma(v_i) = H(v_i)$.
Dann wird $H^{-1}(0)$ durch $H_\sigma^{-1}(0)$ approximiert. Ist $\mathcal{T} = a(\mathcal{K})$ die benutzte Triangulierung, dann bezeichnet $H_\mathcal{T}$ die Abb. mit $H_\mathcal{T}\mid_\sigma = H_\sigma$.

Definition 3.1 (Vektor-Bewertung) *Sei $\tau = [v_0, ..., v_r]$ ein r-Simplex, dann ist*
$L(\tau) = \begin{pmatrix} 1 & ... & 1 \\ H(v_0) & ... & H(v_r) \end{pmatrix}$ *die Bewertungsmatrix von τ.*
Für den Fall $r = n$ heißt τ vollständig bewertet (v.b.), falls $L^{-1}(\tau)$ existiert und lexikographisch positiv ist. Das heißt, daß für jede Zeile $(c_0, ..., c_n)$ von $L^{-1}(\tau)$ der erste von Null verschiedene Eintrag positiv sein muß.
Ein Simplex σ , der mindestens eine v.b. n-Fläche enthält, wird transversal genannt.

Jetzt ist das Konzept zur Nullstellenberechnung von H_σ klar. Wenn $L^{-1}(\tau)$ lex. positiv ist, dann bildet die erste Spalte von $L^{-1}(\tau)$ (welche mit $(\lambda_0, ..., \lambda_n)^T = L^{-1}(\tau) \cdot e_1$ bezeichnet sei) die baryzentrischen Koordinaten von $H_\tau^{-1}(0)$, denn :

$$\mathbf{0} = \lambda_0 H(v_0) + ... + \lambda_0 H(v_0) = H_\sigma(\lambda_0 v_0 + ... + \lambda_n v_n)$$

Allgower und Gnutzmann [7, Theorem 3.3] haben gezeigt, daß $\{H_\tau^{-1}(0) \mid \tau \subset \sigma \text{ ist v.b.}\}$ die Menge aller Ecken von $H_\sigma^{-1}(0)$ ist.
Zur Vereinfachung werden ab jetzt $n+k$-Simplices einfach als Simplices, $n+k-1$-Simplices als Facetten und n-Simplices als n-Flächen bezeichnet.

Zur vollständigen Berechnung der stückweise linearen (p.l.) Mannigfaltigkeit sind demnach alle transversalen Simplices, die über transversale Facetten miteinander zusammenhängen und alle darin enthaltenen v.b. n-Flächen zu finden.

Im folgenden Alg., der dieses leistet, bezeichnet Σ eine Menge von transversalen Simplices. Für einen transv. Simplex σ ist $\sigma.\phi$ eine Menge der transv. Facetten und $\sigma.\tau$ die Menge der v.b. n-Flächen.

Algorithmus 3.2 (Hauptschleife)

1. $\Sigma = \emptyset$
2. *Finde einen ersten transv. Simplex* σ *und eine v.b.* n*-Fläche* τ
3. *Berechne alle weiteren v.b.* n*-Flächen* $\sigma.\tau$ *von* σ
4. *Berechne alle transv. Facetten* $\sigma.\phi$ *von* σ
5. *Wenn* $\sigma.\phi = \emptyset$ *dann gehe zu 9.*
   andernfalls wähle $\phi = [\sigma; \hat{v}_j] \in \sigma.\phi$
6. *Pivotschritt* $\sigma \xrightarrow{j} \sigma^*$
7. *Wenn* $\sigma^* \in \Sigma$, *dann* $\sigma.\phi = \sigma.\phi \backslash \phi$ *und* $\sigma^*.\phi = \sigma^*.\phi \backslash \phi$, *gehe zu 5.*
8. $\Sigma = \Sigma \bigcup \sigma^*$, $\sigma = \sigma^*$, *gehe zu 3.*
9. *Falls möglich wähle* $\sigma \in \Sigma$ *mit* $\sigma.\phi \neq \emptyset$ *und gehe zu 5.*
   Falls nicht möglich ENDE

Im nächsten Lemma wird ein Algorithmus angegeben, der die Grundlage für die Berechnung aller n-Flächen eines Simplex σ legt.

Lemma 3.3 ([5] Lemma 3.11 Algorithmus für l.p.-Step)
Sei $\tau = [v_{j_0}, ..., v_{j_n}] = [\sigma; \hat{v}_{i_1}, .., \hat{v}_{i_k}]$ *v.b. und* v *eine Fehlecke, dann existiert genau eine Ecke* v_{j_l} *so daß* $\tau^* = [v_{j_0}, .., v_{j_{l-1}}, v, v_{j_{l+1}}, .., v_{j_n}]$ *wieder v.b. ist. Dargestellt wird dieser sogenannte linear programming step (l.p.-Step) durch* $\tau^* = \tau \bigcup \{v\} \backslash \{v_{j_l}\}$.
Die neue Fehlecke v_{j_l} *wird durch folgenden Alg. gefunden:*

Algorithmus 3.4
Sei $L^{-1}(\tau) = (c_{r,t})_{0 \leq r,t \leq n}$.

1. *Berechne* $\beta = (\beta_0, ..., \beta_n)^T = L^{-1}(\tau) \cdot \begin{pmatrix} 1 \\ H(v) \end{pmatrix}$
2. $J_{-1} := \{j \mid \beta_j > 0\}$; $s := -1$
3. *Wenn* $|J_s| > 1$, *dann*
   - *Finde* $J_{s+1} := \left\{ j \in J_s \mid \frac{c_{j,s+1}}{\beta_j} = min\{\frac{c_{t,s+1}}{\beta_t} \mid t \in J_s\} \right\}$
   - $s = s + 1$, *gehe zu 3.*
4. *Wenn* $|J_s| = 1$, *dann ENDE und* $J_s = \{l\}$

In [5] haben Allgower/Georg gezeigt, daß ein $s_0 \in \{0, ..., n\}$ existiert, ab dem J_s nur noch den gesuchten Index l enthält.
Wegen seiner Ähnlichkeiten mit dem Austausch-Schritt des SIMPLEX-Verfahrens hat der Austausch von v und v_{i_l} den Namen 'linear programming step'.
Die Inverse der Bewertungsmatrix von τ^* kann mit der Sherman Morrison Formel durch einfache Rang 1-Modifikation aus $L^{-1}(\tau)$ berechnet werden. Dabei wird nur $L^{-1}(\tau)$ und β benötigt.

$$L^{-1}(\tau^*) = L^{-1}(\tau) - \frac{1}{\beta_l}((\beta - e_{l+1})c_{l,.})$$

Eine sehr wichtige Beobachtung aus der Praxis ist dabei, daß fast immer bereits J_0 nur noch den gesuchten Index enthält. D.h. es reicht i.d.R. aus ,'irgendwie' β und die erste Spalte von $L^{-1}(\tau)$ zu berechnen. Im nächsten Kapitel wird eine entsprechende Modifikation des Verfahrens angegeben. Noch bevor zwei weitere schon bekannte Modifikationen dieses Verfahrens erwähnt werden, soll auf das Problem der Berechnung aller v.b. n-Flächen eingegangen werden.

Der Algorithmus stammt von Allgower u. Gnutzmann [7]. In seiner Struktur ist er dem Alg. 3.2 sehr ähnlich.
Die v.b. n-Flächen werden als Knoten eines Graphen betrachtet. Die Kanten stellen gerade die l.p.-Steps dar, so daß zwei n-Flächen genau dann auf einer Kante liegen, falls die eine n-Fläche durch einen l.p.-Step aus der anderen hervorgeht.
Da jede v.b. n-Fläche k Fehlecken hat, ist sie mit genau k anderen v.b. n-Flächen verbunden.
Ausgehend von einer schon bekannten n-Fläche werden sukzessive alle anderen v.b. n-Flächen

berechnet.
Die Anzahl der deshalb auszuführenden l.p-Steps ist demnach abhängig von der Anzahl der v.b. n-Flächen in einem transversalen Simplex.

Lemma 3.5 (Anzahl von v.b. n-Flächen (Gnutzmann [8]))
Für die Anzahl S von v.b. n-Flächen in einem transversalen $n+k$-Simplex gelten folgende scharfe Abschätzungen

$$k+1 \leq S \leq \binom{n+k+1-\lfloor\frac{k+1}{2}\rfloor}{n+1} + \binom{n+k+1-\lfloor\frac{k+2}{2}\rfloor}{n+1} \tag{3.1}$$

Für $k=1$ ergibt sich damit $S=2$ und für $2 \leq k << n$ ergibt sich die asymptotische Abschätzung $S \leq O(n^{\lfloor k/2 \rfloor})$.

Bei der bisherigen Realisierung des l.p.-Steps mittels aufdatierter Inverser der Bewertungsmatrix wären demnach zur Berechnung aller v.b. n-Flächen $O(n^{2+\lfloor k/2 \rfloor})$ Operationen nötig.
Ein wichtiges Kriterium für eine Modifikation des l.p.-Steps ist deshalb, ob tatsächlich bei jedem l.p.-Step viele Rechenoperationen nötig sind.
Die folgenden beiden Modifikationen sind bisher nur für den Fall $k=1$ angegeben worden. Ich denke aber, daß es möglich ist, sie auf den Fall $k>1$ so zu verallgemeinern, daß nicht bei jedem l.p.-Step $O(n^2)$ Operationen anfallen.
Die erste Variante stammt von Todd [9, Abschnitt 4], ist aber auch bei Allgower/Georg [11, Abschnitt 12.4] zu finden. Dabei wird eine Rechtsinverse B von $L(\sigma)$ und ein Vektor γ, der $KerL(\sigma)$ aufgspannt, benutzt.
Eine n-Fläche $\tau = [\sigma; \hat{v_i}]$ ist genau dann v.b., wenn es einen Vektor z mit $z_i = 0$ gibt, so daß $B - \gamma \cdot z^T$ lex. positiv ist. Beim l.p-step wird die andere v.b. n-Fläche durch

$$l = arglexmin\{B_{s,.}/\gamma_s \mid s \neq i, \gamma_s > 0\} \quad \text{bestimmt.}$$

Da auch hier wieder die Minimierung über die erste Spalte von B fast immer ausreicht, kann eine Zerlegung von B benutzt werden.

Auch die zweite Variante [10] wurde von Todd entwickelt. Die Idee hierbei ist es, ausgehend von $x := H_\tau^{-1}(0)$ mittels ψ, einem Basisvektor von $KerDH_\sigma$, eine Liniensuche zu starten:
Finde ein maximales t_0 mit

$$\text{entweder} \quad \forall 0 \leq t \leq t_0 \quad x(t) := x + t\cdot\psi \in \sigma \quad \text{oder} \quad \forall 0 \geq t \geq -t_0 \quad x(t) \in \sigma$$

Zur Bestimmung von t_0 müssen folgende linearen Ungleichungen erfüllt sein

$$w_i^T \cdot (x(t) - b([\sigma; \hat{v_i}])) \leq 0 \quad \text{für} \quad 0 \leq i \leq n+1 \quad . \tag{3.2}$$

Dabei sind w_i gerade die aNev aus Lemma 2.5. Für t_0 muß eine der Ungleichungen scharf sein.
Die Besonderheit von Todd's Methode liegt nun aber darin, daß speziell bei Homothopie-Verfahren zur Lösung nichtlinerarer Gleichungen bestimmten Typs die affine Abb. $H_{\mathcal{T}}$ über mehrere Simplices $\sigma_1, ..\sigma_r$ unverändert ist. D.h. $H_{\sigma_1} = ... = H_{\sigma_r}$. Das kann zu großen Einsparungen genutzt werden.
Anwenden ließe sich das Verfahren bswp., wenn $H(u,\lambda)$ von der Diskretisierung einer parameterabhängigen (P)DGL stammt - und λ FESTGEHALTEN wird. Dann hat D_uH in der Regel eine Bandstruktur, die nach Todd [9] zur solchen größeren 'linearen Stücken' führt (siehe auch num. Beispiel 'Möbius-Band').
Für den Fall, daß die Lösung nicht nur für ein bestimmtes λ gesucht ist, läßt sich die Methode trotzdem anwenden.
Gesucht seien die Lösungen von $F(u_1,\lambda_1) = ... = F(u_r,\lambda_r) = \mathbf{0}$ für $\lambda_1 < ... < \lambda_r$.
Dann setzt man H bspw. als Homothopie zwischen $F(.,\lambda_i)$ und $F(.,\lambda_{i+1})$ also $H(u,t) := tF(u,\lambda_{i+1}) + (1-t)F(u,\lambda_i)$.
Sukzessive lassen sich so alle gewünschten Lösungspunkte finden.

Nicht nur für die Methode von Todd, sondern auch für die Methode aus dem nächsten Kapitel wird die Ableitung der affinen Approximation von H benötigt. Deshalb sollen jetzt einige elementare Eigenschaften hergeleitet werden.

Wie schon erwähnt wird als Triangulierung $a(\mathcal{K})$ verwendet, wobei $a(x) = W \cdot x + z$.

Lemma 3.6 (Darstellung von DH_σ)
Sei $\sigma = [v_0, ..., v_{n+k}] = a([\bar{v}_0, ..., \bar{v}_{n+k}]) = [a(\bar{v}_0), ..., a(\bar{v}_{n+k})] = a([\bar{v}_0, \pi])$, dann ist DH_σ gegeben durch
$DH_\sigma = (H(v_1) - H(v_0), ..., H(v_{n+k}) - H(v_0)) \cdot V^{-1} \quad ; \quad V := (v_1 - v_0, ..., v_{n+k} - v_0)$.

V ist von der Form $W \cdot P \cdot T$

$$T = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ & & & \\ 0 & 0 & \ddots & 1 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in M(n+k, n+k) \ , \ T^{-1} = \begin{pmatrix} 1 & -1 & 0 & \cdots \\ 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & -1 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

P ist die Permutationsmatrix von π , wenn $\sigma = a([v; \pi])$

Beweis Die Darstellung von DH_σ findet sich bei Allgower u. Gnutzmann [7]. Die Darstellung von V folgt aus

$$V \cdot e_i = v_i - v_0 = a(\bar{v}_i) - a(\bar{v}_0) = W \cdot (\bar{v}_i - \bar{v}_0) = W \cdot (\sum_{l=1}^{i} e_{\pi(l)}) = W \cdot P \cdot T \cdot e_i \quad .$$

Definition 3.7 *In den folgenden Kapiteln wird neben DH_σ vor allem folgende Modifikation von Bedeutung sein.*

$$\mathcal{D}H_\sigma := DH_\sigma \cdot W = (H(v_1) - H(v_0), ..., H(v_{n+k}) - H(v_0)) \cdot T^{-1} \cdot P^T \quad .$$

Analog zu V wird die Matrix

$$\mathcal{V} := P \cdot T = W^{-1} \cdot V \qquad \textit{definiert.}$$

Folgendes Lemma untersucht das Besetzungsmuster von DH_σ und $\mathcal{D}H_\sigma$ in Abhängigkeit von W.

Lemma 3.8 (Besetztheit von DH_σ und $\mathcal{D}H_\sigma$)
Mit den Bezeichnungen von Lemma 3.6 hat $\mathcal{D}H_\sigma$ das gleiche Besetzungsmuster wie $DH(x)$.
Ist $W = diag(\delta_1, ..., \delta_{n+k})$ eine Diagonalmatrix, dann hat demnach auch DH_σ das gleiche Besetzungsmuster wie $DH(x)$.

Beweis Die $i$te Komponente von $H(x)$ sei abhängig von $x_{r(i,1)}, ..., x_{r(i,n_i)}$. Dann hat $DH(x)$ in der $j$ten Spalte nur in den Zeilen $i_1, .., i_s$ Nicht-Null-Elemente , welche $j \in \{r(i_l, 1), ..., r(i_l, n_{i_l})\}$ erfüllen.

Aus der Darstellung $\mathcal{D}H_\sigma = (H(v_1) - H(v_0), H(v_2) - H(v_1), ..., H(v_{n+k}) - H(v_{n+k-1})) \cdot P^T$ folgt, daß die $\pi(j)$te Spalte von $\mathcal{D}H_\sigma$ gleich $H(v_j) - H(v_{j-1})$ ist.
Da $v_j - v_{j-1} = \delta_{\pi(j)} \cdot e_{\pi(j)}$, ist $H(v_j) - H(v_{j-1}) \neq 0$ nur für solche Komponenten $i_1, ..., i_s$, für die gilt $\pi(j) \in \{r(i_l, 1), ..., r(i_l, n_{i_l})\}$.

Daß die Einschränkung auf Diagonalmatrizen für W nicht zu restriktiv ist, zeigt das folgende

Beispiel 3.9 *Sei $B(n, m) = \{A \in M(n) \mid A_{ij} = 0 \text{ für } |i - j| > m\}$ die Menge der Bandmatrizen mit Bandbreite $2m + 1 < n$.*
Eine Matrix $W \in M(n)$ erfüllt $\forall A \in B(n, m) : A \cdot W \in B(n, m)$ genau dann, wenn $W = \begin{pmatrix} O & & \\ & D & \\ & & U \end{pmatrix}$, wobei O eine obere, U eine untere Dreiecksmatrix $\in M(m+1)$ und

$D \in M(n - 2m - 2)$ eine Diagonalmatrix ist.

BEWEIS Sei zunächst $l > m$ angenommen.
Für beliebiges $A \in B(n,m)$ soll gelten $0 = (A \cdot W)_{i,i+l} = \sum_{k=max(i-m,1)}^{min(i+m,n)} A_{ik} \cdot W_{k,i+l}$.
Durch geeignete Wahl von A erhält man die notwendige und hinreichende Bedingung
$W_{k,i+l} = 0$ für $max(i-m,1) \leq k \leq min(i+m,n)$.
Das ist äquivalent zu $0 = W_{i,i+l} = W_{i+1,i+1+(l-1)} = .. = W_{i+m,i+m+(l-m)}$, sofern die entsprechenden Indexkombinationen $(i+s, i+s+(l-s)) \in \{1,..,n\}^2$ sind.
Also $W_{i,j} = 0$ für $j > max(i, m+1)$.
Analog folgt für $l < -m$ die Bedingung $W_{i,j} = 0$ für $j < min(i, n-m-1)$.

Eine wichtige Eigenschaft von DH_σ ist, daß sich nach einem Pivot-Schritt $\sigma \xrightarrow{j} \sigma^*$ DH_σ und DH_{σ^*} nur um eine Rang-1 Modifikation unterscheiden. Derartige Formeln wurden offenbar bisher nur von Todd [10] für Todd's Triangulierung angegeben. Außerdem wurde auch nicht untersucht, ob es sich um Rang-1 Modifikationen handelt.

Lemma 3.10 (Update-Formeln für DH_σ und $\mathcal{D}H_\sigma$ beim Pivot-Schritt)
Mit den Bezeichnungen von 3.6, 3.8, 2.3 gelten die folgenden Rang-1 Update Formeln für $\mathcal{D}H_\sigma$:
Betrachte den Pivot-Schritt $\sigma \xrightarrow{j} \sigma^*$

- $0 < j < n+k$:
  $\mathcal{D}H_{\sigma^*} = \mathcal{D}H_\sigma + (H(v_{j+1}) - H(r(v_j)) + H(v_{j-1}) - H(v_j)) \cdot (e_{\pi(j)} - e_{\pi(j+1)})^T$
  die $\pi(j)$te Spalte wird ersetzt durch $H(v_{j+1}) - H(r(v_j))$
  und die $\pi(j+1)$te durch $H(v_{j-1}) - H(r(v_j))$

- $j = 0$:
  $\mathcal{D}H_{\sigma^*} = \mathcal{D}H_\sigma + (H(r(v_0)) - H(v_1) + H(v_0) - H(v_{n+k})) \cdot e_{\pi(1)}^T$
  die $\pi(1)$te Spalte wird ersetzt durch $H(r(v_0)) - H(v_{n+k})$

- $j = n+k$:
  $\mathcal{D}H_{\sigma^*} = \mathcal{D}H_\sigma + (-H(r(v_{n+k})) - H(v_{n+k}) + H(v_0) + H(v_{n+k-1})) \cdot e_{\pi(n+k)}^T$
  die $\pi(n+k)$te Spalte wird ersetzt durch $H(v_0) - H(r(v_{n+k}))$

Die Update-Formeln für DH_σ folgen aus $DH_{\sigma^*} = \mathcal{D}H_{\sigma^*} \cdot W^{-1}$

BEWEIS Da der Beweis sehr technisch ist, und für das weitere Verständnis nicht notwendig, wurde er in den Anhang verbannt.

Kapitel 4

Modifikation der Konzepts der Vektor-Bewertung

Im vorangegangenen Kapitel wurde bereits bemerkt, daß es nicht nötig ist, die komplette Inverse der Bewertungsmatrix zur Verfügung zu haben. Das führt unmittelbar auf die Idee, beim l.p.-Step β und die benötigten Spalten von $L^{-1}(\tau)$ durch die Lösung von linearen Gleichungsystemen (LGS) zu ermitteln. In diesem Abschnitt wird das Konzept der Vektor-Bewertung so modifiziert, daß diese LGS immer dann einfach zu lösen sind, wenn $DH(x)$ dünnbesetzt ist.

Definition 4.1 (Ergänzte Bewertungsmatrix)
Für $\tau = [\sigma; \hat{v}_{i_1}, .., \hat{v}_{i_k}]$ ist die <u>*ergänzte Bewertungsmatrix*</u> $\tilde{L}^T(\tau) := \left(L^T(\sigma), e_{i_1+1}, ..., e_{i_k+1}\right)$,
mit der Konvention $i_1 < ... < i_k$.

Die folgende Beziehung wird für den Nachweis benötigt, daß $\tilde{L}^{-1}$ genau dann lex. positiv ist, wenn L^{-1} lex. positiv ist .

Lemma 4.2 *Sei $\tau = [v_{j_0}, ..., v_{j_n}] = [\sigma; \hat{v}_{i_1}, .., \hat{v}_{i_k}]$ mit existierender Inverser* $L^{-1}(\tau) = \begin{pmatrix} c_0 \\ \vdots \\ c_n \end{pmatrix}$,

dann gilt mit $u_r := e_{i_r+1} - \sum_{l=0}^n e_{j_l+1} c_l \cdot \begin{pmatrix} 1 \\ H(v_{i_r}) \end{pmatrix}$ *die Darstellung*

$$\tilde{L}^{-1}(\tau) = \left(\sum_{l=0}^n e_{j_l+1} c_l \mid u_1, .., u_k \right) \tag{4.1}$$

Beweis

$$\begin{aligned} \left(\sum_{l=0}^n e_{j_l+1} c_l \mid u_1, .., u_k \right) \cdot \tilde{L}(\tau) &= \sum_{l=0}^n e_{j_l+1} c_l \cdot L(\sigma) + \sum_{r=1}^k u_r \cdot \begin{pmatrix} e_{i_1+1}^T \\ \vdots \\ e_{i_k+1}^T \end{pmatrix} \qquad (4.2) \\ &= \sum_{l=0}^n e_{j_l+1} \left(e_{j_l+1}^T + \sum_{r=1}^k e_{i_r+1}^T \cdot c_l \cdot \begin{pmatrix} 1 \\ H(v_{i_r}) \end{pmatrix} \right) \\ &\quad + \sum_{r=1}^k \left(e_{i_r+1}^T - \sum_{l=0}^n e_{j_l+1} \cdot c_l \cdot \begin{pmatrix} 1 \\ H(v_{i_r}) \end{pmatrix} \right) \cdot e_{i_r+1}^T \\ &= \sum_{l=0}^n e_{j_l+1} \cdot e_{j_l+1}^T + \sum_{r=1}^k e_{i_r+1} \cdot e_{i_r+1}^T \end{aligned}$$

$$
\begin{aligned}
&+\sum_{l=0}^{n} e_{j_l+1}\left(\sum_{r=1}^{k} e_{i_r+1}^T \cdot c_l \cdot \begin{pmatrix} 1 \\ H(v_{i_r}) \end{pmatrix}\right) \\
&-\sum_{r=1}^{k}\left(\sum_{l=0}^{n} e_{j_l+1} \cdot c_l \cdot \begin{pmatrix} 1 \\ H(v_{i_r}) \end{pmatrix}\right) \cdot e_{i_r+1}^T \\
= \quad & \sum_{l=0}^{n} e_{j_l+1} \cdot e_{j_l+1}^T + \sum_{r=1}^{k} e_{i_r+1} \cdot e_{i_r+1}^T + 0 = \\
= \quad & I_{n+k+1}
\end{aligned}
$$

Die Zeilen von $\tilde{L}^{-1}$ haben die Form

$$(c_j, *) \quad \text{oder} \quad (\mathbf{0}_{n+1}, e_r^T) \quad ; e_r \in \mathbf{R}^k \quad . \tag{4.3}$$

Es gibt also eine Permutationsmatrix S, so daß

$$S \cdot \tilde{L}^{-1}(\tau) = \begin{pmatrix} L^{-1}(\tau) & * \\ 0 & I_k \end{pmatrix} \quad \text{ist.}$$

Als direkte Konsequenz ergibt sich

Korollar 4.3 $\tilde{L}^{-1}(\tau) \geq_{lex} \mathbf{0} \Leftrightarrow L^{-1}(\tau) \geq_{lex} \mathbf{0}$

Analog zum bisherigen l.p.-Step gibt es auch bei der modifizierten Vektorbewertung eine Rang-1 Update-Formel für die Bewertungsmatrix. Eine Folgerung daraus ist dann die Adaption des Alg. für den l.p.-Step .

Lemma 4.4 (Update beim l.p.-Step)
Betrachte den l.p.-Step $[v_{j_0}, ..., v_{j_n}] = \tau \to \tau^* = \tau \bigcup \{v_{i_l}\} \backslash \{v_j\} = [v_{j_0^*}, ..., v_{j_n^*}]$.

Für τ *bzw.* τ^* *seien die Inversen der Bewertungsmatrizen mit* $\tilde{L}^{-1}(\tau) =: C =: \begin{pmatrix} c_0 \\ \vdots \\ c_{n+k} \end{pmatrix}$ *bzw.* C^* *bezeichnet.*

Weiterhin sei $C \cdot \begin{pmatrix} 1 \\ H(v_{i_l}) \\ \mathbf{0}_k \end{pmatrix} =: \beta =: (\beta_0, ..., \beta_{n+k})^T$. *Dann gilt*

$$C^* = (C + \frac{1}{\beta_j}(e_{i_l+1} - \beta) \cdot (e_{j+1}^T \cdot C - e_{n+1+l}^T)) \cdot O^T \tag{4.4}$$

wobei O *die Matrix der folgenden Permutation* ω *ist, die aus der Konvention* $i_1^* < ... < i_k^*$ *stammt. Sei* s*, so daß* $i_s < j < i_{s+1}$ *ist, dann ist*

- *für* $s < l$
  $\omega(t) = t$ *für* $0 \leq t \leq n+1+s$ *und* $n+1+l \leq t \leq n+1+k$
  $\omega(n+1+s+1) = n+1+l$
  $\omega(t) = t-1$ *für* $n+1+s+1 < t \leq n+1+l$

- *für* $s \geq l$
  $\omega(t) = t$ *für* $0 \leq t < n+1+i$ *und* $n+1+s+1 \leq t \leq n+1+k$
  $\omega(n+1+s) = n+1+l$
  $\omega(t) = t+1$ *für* $n+1+l \leq t < n+1+s$

Beweis Offensichtlich ist

$$\tilde{L}(\tau^*) = O \cdot (\tilde{L}(\tau) + e_{n+1+l} \cdot (e_{j+1} - e_{i_l+1}))^T \quad . \tag{4.5}$$

Die Sherman-Morrison Formel und die Identitäten

$$\begin{aligned} C \cdot e_{n+1+l} = e_{i_l+1} - \beta \quad &\Leftrightarrow \quad e_{n+1+l} = \tilde{L}(\tau) \cdot e_{i_l+1} - \begin{pmatrix} 1 \\ H(v_{i_l}) \\ \mathbf{0}_k \end{pmatrix} \\ e_{i_l+1}^T \cdot C = e_{n+1+l}^T \quad &\Leftrightarrow \quad e_{i_l+1}^T = e_{n+1+l}^T \cdot \tilde{L}(\tau) \end{aligned} \tag{4.6}$$

vervollständigen den Beweis.

Korollar 4.5 *Algorithmus 3.4 wobei alle Schleifen von* 0 *bis* $n+k$ *laufen, findet die neue Fehlecke* v_j.

BEWEIS Da eine Multiplikation mit O^T von rechts nur die letzten k Spalten umsortiert, was keinen Einfluß auf die lex. Positivheit hat (siehe 4.1), reicht es,

$$e_{i_l+1}^T \cdot (C + \frac{1}{\beta_j}(e_{i_l+1} - \beta) \cdot (e_{j+1}^T \cdot C - e_{n+1+l}^T)) = C_{i_l} + \frac{1}{\beta_j} C_j \overset{!}{\geq}_{lex} \mathbf{0}$$

zu betrachten. Aus der Darstellung von $\tilde{L}^{-1}$ in (4.3) folgt,

$$C_{i_l} \leq_{lex} e_{n+2}^T \quad \text{und} \quad C_j >_{lex} K \cdot e_{n+2}^T \text{ für } K \in \mathbf{R} \text{ beliebig}$$

Deshalb muß $\beta_j > 0$ sein.
Der Fall $r \neq i_l$ führt zu den Bedingungen

$$e_{r+1}^T \cdot (C + \frac{1}{\beta_j}(e_{i_l+1} - \beta) \cdot (e_{j+1}^T \cdot C - e_{n+1+l}^T)) = C_r - \frac{\beta_r}{\beta_j} C_j \overset{!}{\geq}_{lex} \mathbf{0} \quad . \tag{4.7}$$

Wenn $\beta_r \leq 0$ ist, ist (4.7) erfüllt. Also ist (4.7) äquivalent zu

$$j = argmin_{lex}\{\frac{1}{\beta_r} C_r \mid \beta_r > 0\}$$

In der numerischen Praxis ist $argmin\{\frac{1}{\beta_r} C_r^0 \mid \beta_r > 0\}$ i.d.R. schon eindeutig. Deshalb reicht es, die erste(n) Spalte(n) von C zu berechnen, um die neue Fehlecke zu ermitteln. Lemma 4.4 zeigte die Modifikation zwischen $\tilde{L}^{-1}(\tau)$ und $\tilde{L}^{-1}(\tau^*)$ nach einem l.p.-Step innerhalb eines $(n+k)$-Simplex. Diese Formel kann benutzt werden, um die erste(n) Spalte(n) von $\tilde{L}^{-1}(\tau^*)$ durch einfaches Aufdatieren zu erhalten. Jetzt ist noch zu überprüfen, wie sich die Spalte(n) bei einem Pivotschritt von einem Simplex zum nächsten verändert, wobei die n-Fläche aber beibehalten wird.

Lemma 4.6 (Update beim Pivot-Schritt)
Sei $C = (C^0, ..., C^{n+k})$ *die Inverse von* $\tilde{L}(\tau)$, *der ergänzten Bewertungsmatrix von* τ *bezüglich* σ, *und* $C^* = (C^{*0}, ..., C^{*n+k})$ *die Inverse von* $\tilde{L}^*(\tau)$ *bezüglich* σ^*, *dem neuen Simplex. Weiterhin sei* $\tau \subset [\sigma; \hat{v}_{i_l}]$ *die Pivot-Facette. Dann gilt für die ersten* n *Spalten von* C^*

$$(C^{*0}, ..., C^{*n}) = Q^T \cdot (C^0, ..., C^n) \quad . \tag{4.8}$$

Q ist die Matrix die zur Permutation q gehört, welche die Ecken wie folgt sortiert $v_0, .., r(v_{i_l}), .., v_{n+k} \mapsto v_0^*, ..., v_{n+k}^*$. *Also*

$$q = \begin{cases} id & \text{für} \quad 0 < i_l < n+k \\ (2, 3, .., n+k+1, 1) & \text{für} \quad i_l = 0 \\ (n+k+1, 1, 2, .., n+k-1) & \text{für} \quad i_l = n+k \end{cases}$$

BEWEIS Sei O die Permutationsmatrix induziert durch $i_1^* < .. < i_k^*$, dann ist

$$\begin{aligned}\tilde{L}^*(\tau) &= O\left(\tilde{L}(\tau) + \begin{pmatrix} 0 \\ H(r(v_{i_l})) - H(v_{i_l}) \\ \mathbf{0}_k \end{pmatrix} \cdot e_{i_l+1}^T\right) Q = \\ &= O\left(I + \begin{pmatrix} 0 \\ H(r(v_{i_l})) - H(v_{i_l}) \\ \mathbf{0}_k \end{pmatrix} \cdot e_{i_l+1}^T \cdot C\right) \tilde{L}(\tau) Q \quad .\end{aligned} \tag{4.9}$$

Demnach gilt mit der Identität $e_{i_l+1}^T \cdot C = e_{n+1+l}^T$

$$\tilde{L}^{*-1}(\tau) = C^* = Q^T C(I + \begin{pmatrix} 0 \\ H(r(v_{i_l})) - H(v_{i_l}) \\ \mathbf{0}_k \end{pmatrix} \cdot e_{n+l+1}^T) O^T \quad .$$

Da sich die Aussage nur auf die ersten n Spalten bezog und Multiplikation von rechts mit O^T nur die letzten k Spalten umsortiert, ist alles gezeigt.

Nach diesen Vorarbeiten kann das wichtigste Lemma dieses Abschnittes gegeben werden. Es stellt heraus, weshalb es günstiger ist, LGSe der Form

$$\tilde{L}(\tau) \cdot x = \left(y^T, \mathbf{0}_k^T\right)^T \qquad \text{wobei } y = e_i \text{ oder } y = \begin{pmatrix} 1 \\ H(v_i) \end{pmatrix} \tag{4.10}$$

anstatt von

$$L(\tau) \cdot x = y \tag{4.11}$$

zu lösen.

Lemma 4.7 (Faktorisierung von $\tilde{L}(\tau)$)
Mit den Bezeichnungen von Lemma 3.6 gilt folgende Beziehung

$$\tilde{L}(\tau) \quad = \quad J \cdot \tilde{M} \cdot \tilde{V} \quad := \quad \left[I_{n+k+1} + \begin{pmatrix} 0 \\ H(v_0) \\ \mathbf{0}_k \end{pmatrix} \cdot e_1^T \right] \cdot \tilde{M} \cdot \begin{pmatrix} 1 & \mathbf{0}_{n+k}^T \\ \mathbf{0}_{n+k} & \mathcal{V} \end{pmatrix} \quad . \tag{4.12}$$

J ist also eine Frobenius-Matrix und $\tilde{M}$ ist die folgende Matrix

$i_1 \neq 0$:

$$\tilde{M} = \begin{pmatrix} 1 & c^T \\ \mathbf{0}_n & \mathcal{D}H_\sigma \\ 0 & u_{i_1}^T \\ \vdots & \vdots \\ 0 & u_{i_k}^T \end{pmatrix} \quad \text{mit } c^T = \mathbf{1}_{n+k}^T \cdot \mathcal{V}^{-1} = e_{\pi(1)} \text{ und } u_j^T = e_j^T \cdot \mathcal{V}^{-1} = e_{\pi(j)}^T - e_{\pi(j-1)}^T \tag{4.13}$$

bzw.
$i_1 = 0$:

$$\tilde{M} = \begin{pmatrix} 1 & c^T \\ \mathbf{0}_n & \mathcal{D}H_\sigma \\ 1 & \mathbf{0}_{n+k}^T \\ 0 & u_{i_2}^T \\ \vdots & \vdots \\ 0 & u_{i_k}^T \end{pmatrix} \tag{4.14}$$

BEWEIS Die entscheidende Identität für den Beweis ist

$$L(\sigma) = \begin{pmatrix} 1 & \cdots & 1 \\ H(v_0) & \cdots & H(v_{n+k}) \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{1}_{n+k}^T \\ \mathbf{0}_n & \mathcal{D}H_\sigma \cdot \mathcal{V} \end{pmatrix} + \begin{pmatrix} 0 \\ H(v_0) \end{pmatrix} \cdot \mathbf{1}_{n+k+1}^T \quad .$$

Da die weiteren Rechnungen wiederum recht technisch sind, werden sie nur im Anhang dargelegt.

Eine unmittelbare Folgerung aus der Faktorisierung ist das folgende

Korollar 4.8 (Transformation der LGSe)
Mit den Bezeichnungen von 3.6; $y =: (y_0, \bar{y})$ und $x =: (x_0, ..., x_{n+k}) =: (x_0, \bar{x})$ ist die Lösung von (4.10) gegeben durch $(x_0, \bar{x}) = (z_0, \mathcal{V}^{-1} \cdot \bar{z})$, wobei

$i_1 \neq 0$:

$$\begin{pmatrix} \mathcal{D}H_\sigma \\ u_{i_1}^T \\ \vdots \\ u_{i_k}^T \end{pmatrix} \cdot \bar{z} = \begin{pmatrix} \bar{y} - y_0 H(v_0) \\ \mathbf{0}_k \end{pmatrix} \tag{4.15}$$

und $z_0 = y_0 - \cdot z_{\pi(1)}$ *beziehungsweise*

$i_1 = 0$:

$$\begin{pmatrix} \mathcal{D}H_\sigma \\ u_{i_1}^T \\ u_{i_2}^T \\ \vdots \\ u_{i_k}^T \end{pmatrix} \cdot \bar{z} = \begin{pmatrix} \bar{y} - y_0 H(v_0) \\ y_0 \\ \mathbf{0}_{k-1} \end{pmatrix} \tag{4.16}$$

und $z_0 = 0$

BEWEIS Die obigen Formeln folgen unmittelbar aus der Faktorisierung von $\tilde{L}$. Weiterhin wird

$$J^{-1} \cdot \begin{pmatrix} y \\ \mathbf{0}_k \end{pmatrix} = \begin{pmatrix} y_0 \\ \bar{y} - y_0 H(v_0) \\ \mathbf{0}_{k-1} \end{pmatrix}$$

ausgenutzt.
Im Fall $i_1 = 0$ ist allerdings noch ein Vertauschen der 1. und der $n + 2$.ten Zeile des LGSs nötig.

Bemerkung 4.9 *Da $DH_\sigma = \mathcal{D}H_\sigma \cdot W^{-1}$ ist, und $u_j^T \cdot W^{-1} = w_j^T$ ein aNev auf eine Facette von σ ist, ist die Koeffizientenmatrix des LGSs von der Form*

$$\begin{pmatrix} \mathcal{D}H_\sigma \\ u_{i_1}^T \\ \vdots \\ u_{i_k}^T \end{pmatrix} = \begin{pmatrix} DH_\sigma \\ w_{i_1}^T \\ \vdots \\ w_{i_k}^T \end{pmatrix} \cdot W \quad . \tag{4.17}$$

Nach all diesen Umformungen stellt sich die Frage, welchen Einfluß die Transformation auf die Kondition der zu lösenden LGSe hat.
Die Antwort basiert auf der Faktorisierung (4.12) von $\tilde{L}$.
Außerdem wird die vereinfachende Annahme $W = \delta I_{n+k}$ gemacht. Dann gilt für den Durchmesser eines Simplex $\sigma \in a(\mathcal{K})$

$$d := diam(\sigma) = \sqrt{n}\delta \quad .$$

Da die Simplices transversal sind, existiert $z \in \sigma$ mit $H_\sigma(z) = 0$, so daß sich für $H(v_0)$ die Abschätzung

$$\|H(v_0)\| = \|H_\sigma(z) + DH_\sigma \cdot (v_0 - z)\| = O(\sqrt{n}\delta) \quad \text{ergibt. Also} \quad \|J\| = O(max\{1, \sqrt{n}\delta\}) \quad .$$

Aus der einfachen Darstellung von J folgt, daß

$$J^{-1} = \left[I_{n+k+1} - \begin{pmatrix} 0 \\ H(v_0) \\ \mathbf{0}_k \end{pmatrix} \cdot e_1^T \right] \quad \text{ist} \ .$$

Damit ergibt sich für die Konditionszahl von J die Abschätzung $cond(J) = O(max\{1, n\delta^2\})$.
Zur Abschätzung der Kondition von $\tilde{V}$ wird

$$cond(\tilde{V}) \approx cond(\mathcal{V}) \leq cond(P)cond(T) = 1 \cdot 1 \cdot O(n) \quad \text{benutzt} \ .$$

Mit der Darstellung von T in Lemma 3.6 ergibt sich

$$n + k = \|T\|_{1,1} = \|T\|_{\infty,\infty} \quad \text{und} \quad 2 = \|T^{-1}\|_{1,1} = \|T^{-1}\|_{\infty,\infty} \quad .$$

Unter der Annahme $cond(\tilde{V}) = cond(\mathcal{V})$ folgt aus obigen Abschätzungen

$$cond(\tilde{L}) = cond(J \cdot \tilde{M} \cdot \tilde{V}) \leq cond(\tilde{M}) \cdot O(max\{n, n^2\delta^2\}) \quad .$$

Umgekehrt kann aber auch $cond(\tilde{M})$ durch

$$cond(\tilde{M}) = cond(J^{-1} \cdot \tilde{L} \cdot \tilde{V}^{-1}) \leq cond(\tilde{L}) \cdot O(max\{n, n^2\delta^2\})$$

abgeschätzt werden.

Bemerkung 4.10
1) In Lemma 7.1 wird unter der Annahme $\delta = O(n^{-3/2})$ die Existenz von transversalen Simplices in der Nähe von Nullstellen von H gezeigt. Damit hätte man in obiger Abschätzung $O(n)$ und nicht $O(n^2\delta^2)$.

2) Da $H_\tau^{-1}(0) = (v_0, ..., v_{n+k}) \cdot x = v_0 + V \cdot \bar{x} = v_0 + W \cdot \bar{z}$, führt die Berechnung von $\tilde{L}^{-1}(\tau)$ zu einem Newton-Schritt ausgehend von v_0.

4.1 Vergleich der verschiedenen l.p.-Step Varianten

In diesem Abschnitt soll die Komplexität der vier l.p-Step Varianten miteinander verglichen werden. Dazu möchte ich ein Resultat des nächsten Kaptiels vorwegnehmen.
Bei der Realisierung mittels modifizierter Vektorbewertung ist es i.d.R. möglich, die Lösungen aller LGSe, die für die Berechnung aller v.b. n-Flächen in einem Simplex notwendig sind, zurückzuführen auf die Lösungen von genau k LGSen.
Weiterhin sind einige Bemerkungen zur zweiten Methode von M.J. Todd angebracht. Wie bereits im dritten Kapitel erwähnt, ist die Berechnung von ψ mit $DH_\sigma \cdot \psi = \mathbf{0}$ nötig. Werden mit w_i die aNev auf Simplices aus $\mathcal{K}$ bezeichnet und mit $u_i = W^{-T} \cdot w_i$ die aNev auf Simplices aus $a(\mathcal{K})$, dann sind nach (3.2) Ungleichungen der Form

$$w_i^T \cdot W^{-1} \cdot t \cdot \psi =: t w_i^T \cdot \bar{\psi} \leq w_i^T \cdot W^{-1} \cdot (b([\sigma; \hat{v}_i) - x) \quad 0 \leq i \leq n+1 \quad \text{zu erfüllen} \ .$$

$\bar{\psi}$ muß nach (4.17) die Gleichung $\mathcal{D}H_\sigma \cdot \bar{\psi} = \mathbf{0}$ erfüllen, so daß $\bar{\psi}$ bei dünnbesetztem $DH(x)$ schnell berechnet werden kann.
Der Aufwand für die Berechnung der rechten Seite der Ungl. hängt aber stark von W ab. Da die Vektoren w_i maximal zwei Nicht-Null-Elemente (NNE) haben, liegt der Aufwand zwischen $O(n)$ für vollbesetztes W^{-1} und $O(1)$ für bspw. diagonales W^{-1}.

Aufwand für	Update von $L^{-1}(\tau)$	Rechtsinverse von $L(\sigma)$	Todd II $k=1$	modifizierte Vektorbewertung
l.p.-Step	$O(n^2)$	$O(n^2)$	$1\times$ LGS mit $\mathcal{D}H_\sigma$ + $O(n \cdot \mathrm{NNER}(W^{-1}))$	$k\times$ LGS mit $\mathcal{D}H_\sigma$ $+O(n)$
alle v.b. n-Flächen je Simplex	$O(n^{\lfloor k/2 \rfloor}) \cdot O(n^2)$	$\star$	$\star$	$k\times$ LGS mit $\mathcal{D}H_\sigma$ $+O(n^{\lfloor k/2 \rfloor}) \cdot O(n)$

$NNER(W^{-1})$ ist dabei die maximale Anzahl von NNE in einer Zeile von W^{-1}.

Abschließend läßt sich bemerken, daß durch die modifizierte Vektor-Bewertung das erste Verfahren angegeben worden ist, das unabhängig von der Triangulierung (und damit W) die Struktur von $DH(x)$ ausnutzen kann !

Kapitel 5

Lösungsmethoden für die Gleichungssysteme

In diesem Kapitel werden Lösungsmethoden für die LGSe (4.15),(4.16) betrachtet. Die LGSe sind von der Art:

$$\begin{pmatrix} A & B \\ C & F \end{pmatrix} \cdot \begin{pmatrix} o \\ p \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix} \text{ wobei } A \in M(n,n); B \in M(n,k) \tag{5.1}$$

LGSe dieser Art lassen sich mit direkten aber auch mit iterativen Verfahren lösen, sofern bspw. A dünnbesetzt ist. In den folgenden Abschnitten werden aber nur direkte Methoden ausführlich behandelt.
Besonderes Augenmerk liegt dabei auf dem Fall, daß H aus der Diskretisierung einer parameterabhängigen Differential-Gleichung (DGL) stammt. In diesem Fall wäre $x =: (u, \lambda)$, wobei λ den k-dimensionalen Parameter und u die Lösung bezeichnet.
Weiterhin wäre $D_u H$ dünnbesetzt, was ausgenutzt werden soll. In [13] hat Keller zwei Algorithmen angegeben, mit denen LGSe obiger Art effizient gelöst werden können.
Die beiden Alg. behandeln den Fall A regulär und A singulär. Bei der Lösung von parameterabhängigen DG hat man es in der überwiegenden Zahl mit dem ersten Fall zu tun.
Dazu wird die Vereinfachung $W = diag(\delta_1, .., \delta_{n+k})$ betrachtet. In den anschließenden Kapiteln wird ohnehin nur diese Wahl von W weiteruntersucht.
$(A \mid 0) \cdot W^{-1} = (D_u H_\sigma \mid 0)$ ist dann eine Approximation von $(D_u H \mid 0)$ [11, Prop. 15.5.4]. Da $D_u H$ nur in der Nähe von Bifurkations- oder Umkehrpunkten (nahezu) singulär wird, gilt dies im Prinzip auch für A.
Weiterhin sei darauf hingewiesen, daß bspw. in einfach-singulären Punkten zwar $rgDH(x) = n-1$ gilt (Fall $k = 1$ siehe H.B. Keller [13, Definition 5.3]), trotzdem aber DH_σ als Approximation für $DH(x)$ auf der p.l. Mannigfaltigkeit immer maximalen Rang hat. Dies folgt aus der Faktorisierung (4.12) und der Tatsache, daß $\tilde{L}(\tau)$ regulär ist.

Eine weitere Beobachtung ist, daß selbst in der Nähe von Umkehr- oder Bifurkationspunkten A häufig regulär bleibt, da A eben nur eine Approximation ist und solche Punkte i.d.R. nicht exakt durchlaufen werden (siehe drittes num. Bsp.).

5.1 A ist regulär

Eine einfache Blockelimination in (5.1) ergibt

$$p = (C \cdot A^{-1} \cdot B - F)^{-1} \cdot (C \cdot A^{-1} \cdot b - c); o = A^{-1} \cdot b - A^{-1} \cdot B \cdot p \tag{5.2}$$

Werden die rechten Seiten in (4.15),(4.16) betrachtet und wird berücksichtigt, daß die erste Spalte von $L^{-1}(\tau)$ nur sehr selten durch das Lösen eines LGS ermittelt wird, dann sind fast immer LGSe

der Form $A^{-1} \cdot B$ oder $A^{-1} \cdot (H(v_r) - H(v_0))$ zu lösen.
Im folgenden Lemma wird gezeigt, daß auch die Berechnung von $A^{-1} \cdot B$ auf LGSe mit rechter Seite vom Typ $H(v_r) - H(v_0)$ führt.

Lemma 5.1 *Mit den Bezeichnungen der Lemmata 3.3 und 4.8 sowie (5.1) gilt für B^i - {die $n+i$te Spalte von $\mathcal{D}H_\sigma$} - die Identität*

$$B^i = H(v_{q_i}) - H(v_0) - H(v_{q_i-1}) + H(v_0) \quad \textit{wobei} \quad q_i := \pi^{-1}(n+i) \quad \textit{ist.}$$

Beweis

$$\begin{aligned}
\mathcal{D}H_\sigma \cdot e_{n+i} &= (H(v_1) - H(v_0), ..., H(v_{n+k}) - H(v_0)) \cdot T^{-1} \cdot P^T \cdot e_{n+i} \\
&= (H(v_1) - H(v_0), ..., H(v_{n+k}) - H(v_0)) \cdot T^{-1} \cdot e_{q_i} \\
&= (H(v_1) - H(v_0), ..., H(v_{n+k}) - H(v_0)) \cdot (e_{q_i} - e_{q_i-1}) \quad \text{mit} \quad e_0 = \mathbf{0}
\end{aligned}$$

Diese spezielle Form der rechten Seiten wird benutzt, um die Lösungen der LGS auf die Lösung von k LGSen zurückzuführen, welche simultan gelöst werden können, falls z.B. LAPACK [20] zur Verfügung steht.

Lemma 5.2 *Mit obigen Bezeichnungen, sowie $0 \le j_0 < j \le n+k$ und $\pi(j_0+1), ..., \pi(j) \le n$. Dann ist die Lösung von $A \cdot x = H(v_j) - H(v_{j_0})$ durch $x = e_{\pi(j_0+1)} + ... + e_{\pi(j)}$ gegeben.*

Beweis $A \cdot x = (H(v_1) - H(v_0), ..., H(v_{n+k}) - H(v_0)) \cdot T^{-1} \cdot P^T \cdot \begin{pmatrix} I_n \\ \mathbf{0} \end{pmatrix} \cdot x$
Direktes Nachrechnen beendet den Beweis.

Die folgenden Algorithmen fassen die Aussagen dieses Abschnittes zusammen:

Algorithmus 5.3 (Vorbereitungen für einen neuen Simplex)

1. *für $1 \le i \le k$ Finde $q_i = \pi^{-1}(n+i)$*
2. *für $1 \le i \le k$ Löse $A \cdot s_i = H(v_{q_i}) - H(v_0)$*
3. *für $1 \le i \le k$ Löse $A \cdot x = H(v_{q_i-1}) - H(v_0)$ mittels Alg. 5.4*
   Setze dann die ite Spalte von $A^{-1} \cdot B = \hat{B}$ zu $\hat{B}^i := s_i - x$

Algorithmus 5.4 (Löse $A \cdot x = H(v_j) - H(v_0)$)

1. *Wenn $j = 0$ dann $x := \mathbf{0}_n$*
2. *Wenn $\pi(1), ..., \pi(j) \le n$, dann $x := e^n_{\pi(1)} + ... + e^n_{\pi(j)}$*
3. *Wenn $\pi(j) > n$ dann $x := s_{\pi(j)-n}$*
4. *Wenn $\pi(1), ..., \pi(j) \not\le n$ und $\pi(j) \le n$ dann finde das maximale q_i mit $q_i < j$*
   $x := s_i + e^n_{\pi(q_i+1)} + ... + e^n_{\pi(j)}$

Algorithmus 5.5 (Löse (4.10) wobei $A = D_u H_\sigma$ regulär ist)

1. *Wenn $y = \begin{pmatrix} 1 \\ H(v_i) \end{pmatrix}$ Berechne mittels Alg. 5.4 $\hat{b} = A^{-1} \cdot (H(v_i) - H(v_0))$*
   sonst ($y == e_i$) Löse $\hat{b} = A^{-1} \cdot (y - y_0 H(v_0))$
2. *Berechne $C, F, c, z_0, C \cdot \hat{B} - F$ wobei C, F, c, z_0 aus (5.1) mit (4.15),(4.16)*
3. *Löse $p = (C \cdot \hat{B} - F)^{-1} \cdot (C \cdot \hat{b} - c)$*
4. *$o = \hat{b} - \hat{B} \cdot p$; $\bar{z} := (o^T, p^T)$*
5. *Berechne $(x_1, ..., x_{n+k})^T = T^{-1} \cdot P^T \cdot \bar{z}$, $x_0 = z_0$*

5.2 A ist singulär

In diesem Abschnitt wird eine Verallgemeinerung von H.B. Keller's Bordering Algorithmus auf reguläre Systeme (5.1), mit ansonsten beliebigem Rangdefizit von A, angegeben. Insbesondere kann das Vorgehen im Fall $rgA = n$ natürlich in diesen Algorithmus eingebettet werden.

Der erste Schritt besteht in der Berechnung einer Matrix $\psi^T \in M(k, n+k)$ mit $rg(\psi^T) = k$ und $\psi^T \cdot \begin{pmatrix} A \\ C \end{pmatrix} = 0$. Multiplikation von (5.1) von links mit ψ^T ergibt das reguläre $k \times k$ System

$$\psi^T \cdot \begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} o \\ p \end{pmatrix} = \psi^T \cdot \begin{pmatrix} B \\ D \end{pmatrix} \cdot p = \psi^T \cdot \begin{pmatrix} b \\ c \end{pmatrix} \tag{5.3}$$

Für die Konstruktion von ψ^T sei $\phi = (\phi_1, .., \phi_l)$ eine Matrix aus Basisvektoren von $KerA$. Also insbesondere $dimKerA = l$, wobei $l = 0$ auch zulässig sein soll.
Analog sei $\bar{\psi} = (\bar{\psi}_1, .., \bar{\psi}_l)$ eine Matrix aus Basisvektoren von A^T.
Anschließend bestimmt man die $s \times k$ Matrix η^T mit

$$\eta^T \cdot C \cdot \phi = 0 \quad \text{und} \quad rg(\eta^T) = s \quad ,$$

also $\eta^T \cdot C \in (KerA)^\perp = ImA^T$.
Demnach können Partikulärlösungen $\mu^T \in M(s,n)$ mit $\mu^T \cdot A = -\eta^T \cdot C$ bestimmt werden.
Nunmehr kann

$$\psi^T = \begin{pmatrix} \bar{\psi}^T & 0 \\ \mu^T & \eta^T \end{pmatrix}$$

gesetzt werden.
Für den Fall $l = 0$ ergibt sich, daß $\eta^T = I_k$ und $\mu^T = -C \cdot A^{-1}$ geeignet sind. Das System (5.3) wird dann zu

$$(-C \cdot A^{-1} \cdot B + D) \cdot p = (-C \cdot A^{-1} \cdot b + c)$$

- der Schurkomplement-Gleichung.

Nachdem p bekannt ist, wird zunächst eine partikuläre Lösung $\bar{o}$ von $A \cdot \bar{o} = b - B \cdot p$ berechnet.
Wenn A regulär ist, ist $\bar{o} = o = A^{-1}(b - B \cdot p)$, so daß bereits die Lösung von (5.1) gefunden wurde.
Ansonsten gilt $o = \bar{o} + \phi \cdot w$ mit $w \in \mathbf{R}^l$.
Zur Bestimmung von w benutzt man das überbestimmte, aber eindeutig lösbare LGS

$$C \cdot (\bar{o} + \phi \cdot w) = c - D \cdot p \quad \Leftrightarrow \quad C \cdot \phi \cdot w = c - D \cdot p - C \cdot \bar{o}$$

Unklar ist noch, wie die Basisvektoren der Kerne oder die Partikulärlösungen bestimmt werden können. Dazu bietet sich eine singuläre LR-Zerlegung von A an.
Auf Grund der Seltenheit des Falls A singulär, wird man versucht sein, immer z.B. eine LR-Zerlegung von A zu benutzen.
In diesem Fall ist also R singulär, was bedeutet, daß es l Diagonalelemente $R_{s_1,s_1}, .., R_{s_l,s_l}$ gibt, die ungefähr 0 sind.
Mit der (beliebigen) Wahl $\bar{o}_{s_1} := ... := \bar{o}_{s_l} := 1$ kann bspw. $\bar{o}$ durch einfaches Rückwärts-Lösen bestimmt werden.

5.3 A ist nahezu singulär

Bei H.B. Keller wird neben den beiden obigen Algorithmen auch der Fall einer nahezu singulären Matrix A behandelt.
In diesem Fall gilt bei einer LU-Zerlegung mit geeigneter Pivot-Strategie

$$A = P \cdot L \cdot U \cdot Q \quad \text{wobei} \quad U = \begin{pmatrix} \bar{U} & u \\ 0 & \epsilon \end{pmatrix} \quad \text{mit} \quad \epsilon \in M(l)$$

Es gibt die Möglichkeit, entweder die fast immer gegebene Regularität von ϵ auszunutzen und analog zum ersten Fall 'A regulär' zu rechnen, oder $\epsilon = 0$ anzunehmen und mit dem Algorithmus für A singulär zu rechnen.

In [13] wird für den Fall $k = 1$ unter der Annahme, daß L und die Koeffizientenmatrix von (5.1) unabhängig von ϵ beschränkte Inverse haben, gezeigt, daß sich die Lösungen für beide Vorgehensweisen nur um Terme der Ordnung $O(\epsilon)$ voneinander unterscheiden.
Diese Annahmen werden dort allerdings nicht extra formuliert, und der Beweis ist auf den hier vorliegenden Fall nicht ohne weiteres übertragbar.
Trotzdem gilt auch hier die entsprechende Fehlerabschätzung, da

$$\|M - M_0\| := \left\| \begin{pmatrix} P \cdot L \cdot \begin{pmatrix} \bar{U} & u \\ 0 & \epsilon \end{pmatrix} \cdot Q & B \\ C & D \end{pmatrix} - \begin{pmatrix} P \cdot L \cdot \begin{pmatrix} \bar{U} & u \\ 0 & 0 \end{pmatrix} \cdot Q & B \\ C & D \end{pmatrix} \right\| = O(\epsilon)$$

Mit Hilfe der Neumann'schen Reihe kann gezeigt werden, daß für die Inversen gilt

$$\|M^{-1} - M_0^{-1}\| = O(\epsilon)$$

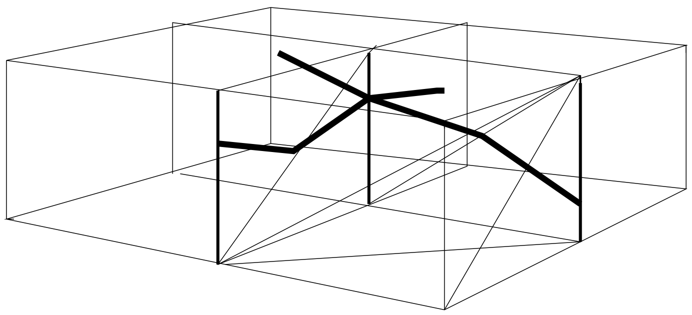
Also unterscheiden sich auch die Lösungen der LGSe nur um Terme $O(\epsilon)$.
Beim Durchführen der Algorithmen müssen aber im Fall 'A regulär' LGSe mit der schlecht konditionierten Matrix U gelöst werden, wogegen bei Gebrauch des Algorithmus für 'A singulär' nur LGSe mit den gut konditionierten Matrizen $L, \bar{U}$ nötig sind.

Kapitel 6

Der Gitter-Algorithmus für $k \geq 2$

Das nächste Ziel ist, die Anzahl von zu durchlaufenden Simplices und von zu berechnenden v.b. n-Flächen zu reduzieren. Diese Zahl ist speziell für $k \geq 2$ und große n sehr groß.
Die Idee des Gitter-Algorithmus ist die Folgende:
Der $\mathbf{R}^{n+k}$ wird in Quader mit den Seitenlängen $\delta_1, .., \delta_{n+k}$ so zerlegt, daß die Simplices der Triangulierung $a(\mathcal{K})$ genau in die Quader passen.
Anstatt alle v.b. n-Flächen der p.l. Mannigfaltigkeit zu bestimmen, beschränkt man sich auf solche v.b. n-Flächen, die im Durchschnitt von k Facetten von k Quadern liegen.
Für den Fall $n = 1$, $k = 2$ sind diese Durchschnitte die Kanten des Quaders. Aus diesem Grund werden die entsprechenden n-Flächen Kanten-n-Flächen (KNF) genannt. Jede KNF τ ist über einen Teil der p.l. Mannigfaltigkeit mit maximal $2k$ anderen KNF $\tau_1, ..., \tau_{2k}$ verbunden, wobei τ_i und τ in der gleichen Facette eines Quaders liegen. Die KNF τ_i werden Nachbarn von τ genannt.

In der nachfolgenden Skizze ist die Situation für $n = 1$ und $k = 2$ dargestellt.
Die dünnen Linien zeigen die Quader und einige Simplices. Die etwas dickeren Linien zeigen einige v.b. KNF.
Die dicken Linien zeigen den Schnitt der p.l. Mannigfaltigkeit mit den Seitenflächen der Quader. Die Lösungspunkte auf genau diesen Linien soll der Gitter-Alg. berechnen.



Zunächst soll die Hauptschleife des Algorithmus erläutert werden.

Für jede KNF τ wird Folgendes gespeichert

$\tau.b$: der Schwerpunkt; dient zur Identifikation in einem binären Suchbaum ;
$\tau.nfa[k][2]$: Zeiger auf benachbarte KNF ;
$\tau.nv[k]$: Normalen-Richtungen der Hyperebenen in der τ liegt: $\pi(n+1), .., \pi(n+k)$
$\tau.sol$: Lösungspunkt $H_\tau^{-1}(0)$

Algorithmus 6.1 (Hauptschleife)

1.	$\Sigma = \emptyset$
2.	*Finde eine erste KNF* τ *und einen Start-Simplex* σ
3.	$\tau.nfa := \emptyset$; $\tau.b := b(\tau)$
loop:	*Wenn möglich, wähle eine Normalen-Richtung, in der ein Nachbar liegt, der noch nicht durchlaufen wurde, also* s, t *mit* $\tau.nfa[s][t] = undefiniert$ *Wenn alle Nachbarn bekannt sind, gehe zu restart*
4.	*Verfolge mittels Alg. 6.4 den p.l. Pfad im Durchschnitt der übrigen Hyperebenen, bis der Nachbar* τ^* *gefunden wurde*
5.	$\tau.nfa[s][t] := \tau^*$
6.	*Wenn* $\tau^* \in \Sigma$ *dann* $\tau^*.nfa[s^*][t^*] := \tau$ *Sonst* $\tau^*.nfa[s^*][t^*] := \tau$; $\tau^*.b := b(\tau^*)$; $\Sigma := \Sigma \bigcup \tau^*$
7.	$\tau := \tau^*$, *gehe zu loop*
restart:	*Finde* $\tau \in \Sigma$ *mit unbekannten Nachbarn* *Wenn es keine solche KNF in* Σ *gibt, dann ENDE*
8.	*Rekonstruiere einen Simplex* σ , $\tau \subset \sigma$ *und alle übrigen Informationen, die für die Pfad-Verfolgung nötig sind* $DH_\sigma, \pi, v_0, ...$

Wie bereits in 2.4 erwähnt, ist eine KNF durch ihren Schwerpunkt eindeutig in der Menge der KNF repräsentiert. Deshalb werden alle KNF mit unbekannten Nachbarn in einem binären Suchbaum gespeichert, wobei $\tau_1 <_{lex} \tau_2$ bedeutet, daß $b(\tau_1) <_{lex} b(\tau_2)$ ist.
Diese Idee wurde von Gnutzmann eingeführt [8].
Wenn die p.l. Mannigfaltigkeit den Rand von $\bar{\Omega} \subset \mathbf{R}^{n+k}$ schneidet, wird der Algorithm 6.1 etwas komplizierter. In diesem Fall werden Dummy KNF für solche Nachbarn benutzt, die außerhalb von $\bar{\Omega}$ liegen würden. KNF mit Dummy Nachbarn werden nicht aus dem Suchbaum gestrichen. Vielmehr wird eine zusätzliche doppelt verkette Liste benutzt, in der alle KNF mit noch unbekannten Nachbarn gespeichert sind. Diese Liste kann einfach aufdatiert werden.
Es sind aber noch eine Reihe von Teilproblemen von 6.1 zu lösen.

Zunächst wird eine einfache Charakterisierung der KNF gegeben.

Lemma 6.2 *Sei* $\tau := [\sigma; \hat{v}_{i_1}, ..., \hat{v}_{i_k}]$. τ *ist genau dann eine KNF, wenn* $i_2 = i_1 \oplus 1, ..., i_k = i_{k-1} \oplus 1$ *und* $\{0, n+k\} \bigcap \{i_1, ..., i_k\} \neq \emptyset$ *wobei* $\oplus$ *die Addition modulo* $n+k$ *ist.*

Beweis Mit der Konvention $i_1 < .. < i_k$ und wegen der Pivot-Regeln (Lemma 2.3 Fälle $j = 0, n+k$) kann angenommen werden, daß $i_k = n + k$.
Aus der Darstellung der aNev (Lemma 2.5) folgt, daß für den Fall $i_1 = n + 1, .., i_k = n + k$ der Normalraum von τ durch

$$-\delta_{\pi(n+2)} e_{\pi(n+1)} + \delta_{\pi(n+1)} e_{\pi(n+2)} \quad , \quad -\delta_{\pi(n+3)} e_{\pi(n+2)} + \delta_{\pi(n+2)} e_{\pi(n+3)} \quad , ..., -e_{\pi(n+k)}$$

also durch

$$e_{\pi(n+1)}, .., e_{\pi(n+k)}$$

aufgespannt wird.
Analog zeigt sich, daß der Normalraum von τ NICHT von k kanonischen Einheitsvektoren aufgespannt wird, falls die Fehlecken NICHT $v_{n+1}, .., v_{n+k}$ sind.

Definition 6.3 *Eine Darstellung einer KNF* τ *als* $\tau = [\sigma, \hat{v}_{n+1}, .., \hat{v}_{n+k}]$ *heißt Standard-Darstellung.*

Betrachte nun den Fall, daß eine KNF $\tau = [\sigma, \hat{v}_{n+1}, .., \hat{v}_{n+k}] \subset \{x \mid x_{l_1} = c_1, ..., x_{l_k} = c_k\}$ gegeben ist. Gesucht ist die benachbarte KNF $\tau^* \subset \{x \mid x_{l_2} = c_{l_2}, ..., x_{l_k} = c_{l_k}, x_{l_1^*} = c^*_{l_1}, x_{l_1} > (<) c_{l_1}\}$.
Wie im Beweis von 6.2 folgt aus 2.5 daß die letzten $k - 1$ Fehlecken dafür verantwortlich sind, daß τ in $\{x \mid x_{\pi(n+2)} = c_{\pi(n+2)}, ..., x_{\pi(n+k)} = c_{\pi(n+k)}\}$ liegt. Deshalb findet der nächste Algorithmus τ^*

Algorithmus 6.4 (p.l.-Verfolgung im Durchschnitt von $k-1$ Hyperebenen)

1. *Finde einen Simplex σ so daß $\tau = [\sigma, \hat{v}_{n+2}, .., \hat{v}_{n+k}, \hat{v}_{i_1}]$ und* $\{\pi(n+2), ..., \pi(n+k)\} = \{l_2, ..., l_k\}$ *und* $\sigma \in \{x \mid x_{l_1} > (<) c_{l_1}\}$.
2. *l.p.-Step um v_{i_1} in die l.p.-Basis hineinzutauschen; v_l ist die neue Fehlecke*
3. *Wenn $l = n+1$ dann STOP: KNF τ^* in Standard Darstellung gefunden*
4. *Pivote über v_l*
5. *Wenn die Nummern der Fehlecken jetzt $\{n+1, .., n+k\}$ sind, dann STOP: KNF τ^* in Standard Darstellung gefunden*
6. *gehe zu 2.*

Die Beibehaltung der Fehlecken $v_{n+2}, ..., v_{n+k}$ erhält $\{\pi(n+2), ..\pi(n+k)\} = \{l_2, .., l_k\}$.
Weiterhin sind die durch diesen Alg. berechneten v.b. n-Flächen enthalten in $\{x \mid x_{\pi(n+2)} = c_{\pi(n+2)}, ..., x_{\pi(n+k)} = c_{\pi(n+k)}\}$.
Ausgenommen eines Pivot-Schritts über v_0, der unmittelbar zu der 'Wenn dann'- Bedingung in Schritt 5 führt, bleibt v_0 erhalten. Die einzige Veränderung gibt es bei den Permutationen.
Deshalb wird der Quader, in dem der Simplex von Schritt 1 liegt, nicht verlassen.
Also $\tau^* \in \{x \mid x_{l_1} > (<) c_{l_1}\}$.

Das nächste Ziel ist die Regeln für Schritt 1 des obigen Alg. anzugeben. Dazu wird eine KNF in Standard Darstellung betrachtet $\tau = [\sigma, \hat{v}_{n+1}, .., \hat{v}_{n+k}]$. Für die Koordinaten des Schwerpunktes gilt mit $n+1 \leq i \leq n+k$

$$b_{\pi(i)}(\sigma) = (v_0)_{\pi(i)} + \frac{n+k+1-i}{n+k+1} < (v_0)_{\pi(i)} + \frac{n+1}{n+1} = b_{\pi(i)}(\tau) \quad . \tag{6.1}$$

Deshalb ist für den Fall '<' in 6.4-1 der Simplex schon im richtigen Quader bzw. liegt auf der richtigen Seite der Hyperebene.
Um die π-Bedingung zu erfüllen, werden einige Pivot-Schritte über Fehlecken mit abfallenden Nummern ausgeführt, so daß für den resultierenden Simplex $\pi^*(n+1) = l_1$ gilt. Sei $\pi(l) = l_1$, $1 \leq l \leq k$, dann folgt aus Lemma 2.3 daß die folgende Sequenz von Pivotschritten den gewünschten Simplex σ^* findet: Wenn $l > n+1$, dann $\sigma \overset{l-1}{\to} .... \overset{n+1}{\to} \sigma^*$. Jetzt ist $\tau = [\sigma^*, \hat{v}^*_{n+1}, .., \hat{v}^*_{n+k}]$.

Für den Fall '>' in 6.4-1 erhält man den Simplex σ^* durch Pivot-Schritte über Facetten mit anwachsenden Nummern.
Sei $\pi(l) = l_1$, $n+1 \leq l \leq n+k$ dann folgt wiederum aus Lemma 2.3 daß die folgende Sequenz σ^* findet: $\sigma \overset{l}{\to} ... \overset{n+k-1}{\to} \bar{\sigma} \overset{n+k}{\to} \sigma^*$.
Dann ist $\tau = [\sigma^*, \hat{v}^*_{n+2}, .., \hat{v}^*_{n+k}, \hat{v}^*_0]$, und es gilt $\pi^*(n+2) = l_2, ..., \pi^*(n+k) = l_k, \pi^*(1) = l_1$.

Eine weitere nützliche Vereinfachung kann gemacht werden, falls die Lösungen von parameterabhängigen DGLen gesucht sind.
Wenn $x_{n+1}, .., x_{n+k}$ diese Parameter sind, macht es Sinn, nur solche KNF zu suchen, deren Normal-Raum durch $\{e_{n+1}, .., e_{n+k}\}$ aufgespannt wird. D.h. es sind nur die Lösungen bei bestimmten Parametern zu berechnen.

Kapitel 7

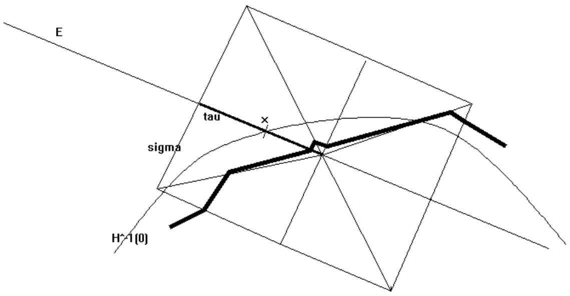
Finden einer ersten v.b. n-Fläche

Ein noch offenes Problem ist die Berechnung eines Start Simplex einer geeigneten Triangulierung und einer ersten v.b. n-Fäche. Die Triangulierung sollte das affine Bild von $\mathcal{K}$ sein, so daß letzlich eine geeignete affine Abbildung a gesucht ist.
In diesem Abschnit bezeichnet $\|.\|$ die euklidische Norm des $\mathbf{R}^{n+k}$ oder die zugehörige Matrix-Norm.
Folgendes Lemma gewährleistet die Existenz von v.b. n-Flächen in der Nähe von regulären Lösungspunkten für hinreichend feine Triangulierungen. Es handelt sich um eine Modifikation des unten angegeben Lemmas von Allgower/Gnutzmann/Georg

Lemma 7.1 ([11] Proposition 15.5.6 Existenz einer v.b. n-Fläche)
Sei $x \in H^{-1}(0)$ mit $rg(DH(x)) = n$ ein regulärer Lösungspunkt. Weiterhin sei $E := x + span\{w_1, ..., w_n\}$ ein n-dimensionaler affiner Unterraum mit $KerDH|x \bigcap span\{w_1, ..., w_n\} = \{0\}$. Der Normalraum von $span\{w_1, ..., w_n\}$ habe die Basis $\{u_1, ..., u_k\}$.
$\{\mathcal{T}_d\}$ sei eine Familie von uniformen Triangulierungen, wobei die Simplices den Durchmesser d haben. Es gibt also ein $\theta > 0$, so daß es für jede Triangulierung und jeden Simplex σ eine Kugel mit Radius $d \cdot \theta$ gibt, die ganz in σ liegt.
Außerdem soll jede Triangulierung $\mathcal{T}_d$ eine n-Fläche $\tau =: [v_0, ..., v_n]$ enthalten, für die $\tau \subset E$ und $b(\tau) = x$ gilt.

Dann ist τ v.b., sofern d hinreichend klein ist.

Folgende Skizze zeigt die Situation im Fall $n = k = 1$.



Bemerkung 7.2 ([11] Definition 15.5.3)
Für die Freudenthal-Kuhn-Triangulierung $\mathcal{K}$ und damit alle affinen Bilder unter der Abb. $a(x) = W \cdot x + w$ mit $W = \delta I_{n+k}$ gilt

$$\theta = 1/((n + k)\sqrt{2}) \tag{7.1}$$

Beweis von Lemma 7.1
σ (eigentlich σ_δ) sei für jede der Triangulierungen ein $n + k$-Simplex, der die oben bezeichnete n-Fläche τ enthält.

Seien U und B die folgenden Matrizen $U := (u_1, ..., u_k)$ bzw. $B := \begin{pmatrix} DH_\sigma \\ U^T \end{pmatrix}$. Auf Grund der Bedingung an $u_1, ..., u_k$ ist $\begin{pmatrix} DH(x) \\ U^T \end{pmatrix}$ invertierbar. Da DH_σ eine Approximation von $DH(x)$ ist [11, Prop. 15.5.4], ist für hinreichend kleines d auch B regulär.

Betrachte den Newton-Schritt $x_0 := x - y_0$ wobei $y_0 := B^{-1} \cdot \begin{pmatrix} H_\sigma(x) \\ \mathbf{0}_k \end{pmatrix}$.
Vorausgesetzt es ist $x_0 \in \sigma$, dann gilt

$$H_\sigma(x_0) = H_\sigma(x) - DH_\sigma \cdot y_0 = H_\sigma(x) - DH_\sigma \cdot \begin{pmatrix} DH_\sigma \\ U^T \end{pmatrix}^{-1} \cdot \begin{pmatrix} H_\sigma(x) \\ \mathbf{0}_k \end{pmatrix} = H_\sigma(x) - H_\sigma(x) = \mathbf{0}_n \quad .$$

Mit $U^T \cdot y_0 = \mathbf{0}_k$ ist $x_0 \subset E$. Es bleibt zu zeigen, daß für hinreichend kleines d x_0 in τ liegt. Das ist sicherlich erfüllt, falls

$$\|y_0\| \leq d\theta \tag{7.2}$$

Aus [11] Prop. 15.5.2 folgt $\| \begin{pmatrix} H_\sigma(x) \\ \mathbf{0}_k \end{pmatrix} \| \leq \alpha d^2/2$, also

$$\|y_0\| = \|B^{-1} \cdot \begin{pmatrix} H_\sigma(x) \\ \mathbf{0}_k \end{pmatrix} \| \leq \|B^{-1} |_{\mathbf{R}^n \times \mathbf{0}_k} \| \cdot \alpha d^2/2 \quad \text{wobei} \quad \alpha \geq \|H''(.)\|$$

Da $DH_\sigma \to DH(x)$ für $d \to 0$ gibt es ein von d unabhängiges C mit $\|B^{-1} |_{\mathbf{R}^n \times \mathbf{0}_k} \| \leq C$. Für hinreichend kleines d ist demnach Bedingung (7.2) erfüllt .

Bemerkung 7.3

Obige Abschätzung eingesetzt in (7.2) führt zur Bedingung

$$d \leq \frac{2\theta}{C\alpha} \quad .$$

Wird eine isotrope Triangulierung (also bspw. $W = \delta \cdot I$) benutzt, dann gilt für $a(\mathcal{K})$

$$d = \sqrt{n}\delta \quad .$$

Zusammen mit (7.1) folgt die für (7.2) hinreichende Bedingung

$$\delta \leq \frac{\sqrt{2}}{C\alpha n^{3/2}} \quad .$$

Als Folgerung aus Lemma 7.1 kann jetzt ein Algorithmus zum Finden einer ersten v.b. n-Fläche und eines Start-Simplex angegeben werden. Dabei ist das Problem der Wahl von E aber noch ausgespart.

Algorithmus 7.4 (Erste v.b. n-Fläche)

0. *Input: π ; x_0 ein regulärer Lösungspunkt und $W := diag(\delta_1, ..., \delta_{n+k})$ - mit W einer Vorgabe für die Seitenverhältnisse der Triangulierung und einer oberen Schranke für den Durchmesser der Simplices sowie π einer Permutation*
1. *Setze $a(z) := W \cdot z + w$, wobei $w := x_0 - W \cdot b([[0, \pi], \hat{v}_{n+1}, .., \hat{v}_{n+k}])$*
2. *Berechne $L(\tau)$ mit $\tau := a([[0, \pi], \hat{v}_{n+1}, .., \hat{v}_{n+k}])$*
3. *Wenn $L^{-1}(\tau)$ nicht existiert oder nicht lex. positiv ist, dann $W := \frac{1}{2}W$, gehe zu 1*

Nun werden verschiedene Wahlen von E und $(u_1, ..., u_k)$ betrachtet.
Es ist wünschenswert, $\|B^{-1} |_{\mathbf{R}^n \times \mathbf{0}_k} \|$ zu minimieren, so daß δ so groß wie möglich gewählt werden kann. Da für die Realisierung des Gitter-Alg. die Konstruktion einer v.b. KNF notwendig ist, ist

W als Diagonalmatrix anzusetzen, so daß $(u_1, ..., u_k)$ kanonische Einheitsvektoren sein müssen. Das nächste Ziel ist π bzw. $(u_1, ..., u_k) = (e_{\pi(n+1)}, ..., e_{\pi(n+k)})$ zu finden, so daß $\|B^{-1}|_{\mathbf{R}^n \times \mathbf{0}_k}\|$ minimiert wird über alle Vektoren $(u_1, ..., u_k)$, die zu einem 'achsenparallelen' affinen Unterraum führen. Die entscheidende Ungleichung lautet

$$\| \begin{pmatrix} DH_\sigma \\ U^T \end{pmatrix}^{-1} |_{\mathbf{R}^n \times \mathbf{0}_k} \| \leq \| \begin{pmatrix} DH_\sigma \\ KerDH_\sigma^T \end{pmatrix}^{-1} |_{\mathbf{R}^n \times \mathbf{0}_k} \| / \sqrt{1 - dist^2(KerDH_\sigma, span\{U\})} \quad .$$

Ist $k = 1$, u eine ONB von $span\{U\}$ und v eine ONB von $KerDH_\sigma$, dann ist der Ausdruck in der Wurzel gerade $(u^T \cdot v)^2$ (siehe nachfolgende Def. mit euklid. Norm) . Die Wurzel ist demnach der Betrag des Cosinus zwischen den beiden Basisvektoren.
In diesem Fall kann die Abschätzung 7.3 benutzt werden, um Rheinboldt's Matrix-Ergänzung ([14] statem. 4.4-4.6) für den Newton-Schritt bei Prädiktor-Korrektor Methoden zu rechtfertigen.
Dabei wird als Ergänzung für $DH(x)$ der Einheitsvektor vorgeschlagen, dessen Skalarprodukt mit dem Tangentialvektor maximal ist.

Obige Ungleichung kann auch allgemeiner im Kontext von Hilberträumen formuliert werden. Deshalb sei an dieser Stelle ein Abstecher in die Funktionalanalysis erlaubt.

Definition 7.5 ('Cosinus des Winkels' zwischen Unterräumen)
Seien X ein Hilbertraum und U, V abgeschlossene Unterräume von X, dann bezeichnet $P(x, U)$ die orthogonale Projektion von x auf U. Sei $\mu(x, U) := \|P(x, U)\|^2$, dann bezeichnet $\mu(V, U) := inf\{\mu(x, U) \mid x \in V, \|x\|_2 = 1\}$.

Bemerkung 7.6 *Mit der Definition des Abstandes zwischen Unterräumen nach Kato [19] IV.2.1*

$$dist(V, U) := sup\{dist(v, U) \mid v \in V, \|v\| = 1\} \quad gilt \quad dist^2(V, U) = 1 - \mu(V, U) \quad ,$$

weshalb aus [19, Theorem IV.2.9]

$$\mu(V, U) = \mu(U^\perp, V^\perp) \quad folgt.$$

Mit diesen Bezeichnungen gilt

Lemma 7.7 (Abschätzung für die Inverse)
Seien X, Y Hilberträume, $L : X \to Y$ linear und stetig, so daß $KerL$ abgeschlossen ist, sowie $X = KerL \oplus KerL^\perp$.
Außerdem sei die Einschränkung $L_{KerL^\perp} = L|_{KerL^\perp}$ ein Isomorphismus.
$U \subset X$ sei ein abgeschlossener Unterraum mit $\mu(KerL, U) > 0$.

Dann ist die Einschränkung $L_{U^\perp} = L|_{U^\perp}$ injektiv und erfüllt

$$\|L_{U^\perp}^{-1}\| \leq \|L_{KerL^\perp}^{-1}\| / \sqrt{\mu(KerL, U)} \tag{7.3}$$

Beweis Mit $R(L) = \{V \subset X$ abgeschlossen und $L_{V^\perp}$ ist injektiv$\}$ werde die Menge der L-regulären Unterräume bezeichnet.
U ist aus $R(L)$, denn sei $x \in U^\perp$, dann ist einerseits $P(x, U) = 0$. Gilt dann aber $Lx = 0$, dann ist x auch aus $KerL$, so daß $\mu(KerL, U) = 0$. Widerspruch zur Voraussetzung.

Seien $V_1, V_2 \in R(L)$ und $y \in L(V_1) \bigcap L(V_2)$. Weiterhin seien $L_{V_i^\perp}^{-1} y = x_0(V_i) \oplus x_1(V_i)$
mit $x_0(.) \in KerL, x_1(.) \in KerL^\perp$.
Offensichtlich ist $x_1(V_1) - x_1(V_2) \in KerL^\perp$.
Andererseits gilt

$$L(x_1(V_1) - x_1(V_2)) = L(x_1(V_1) + x_0(V_1) - x_1(V_2) - x_0(V_2)) = y - y = 0$$

Also ist auch $x_1(V_1) - x_1(V_2) \in KerL$, so daß $x_1(V_1) = x_1(V_2) =: x_1$.
Da $x_0(KerL) = 0$, ist

$$x_1 = L^{-1}_{KerL^\perp} y = P(L^{-1}_{V_i^\perp} y, KerL^\perp)$$

Für $V \in R(L)$, $y \in L(V)$ erhält man

$$\|L^{-1}_{KerL^\perp} y\|^2 = \|P(L^{-1}_{V^\perp} y, KerL^\perp)\|^2 \geq \|L^{-1}_{KerL^\perp} y\|^2 \mu(V^\perp, KerL^\perp)$$

Korollar 7.8 *Gilt zusätzlich $\mu(U, KerL) > 0$, dann ist $L_{U^\perp}$ auch surjektiv.*

Beweis Es ist zu zeigen, daß zu $x_1 \in KerL^\perp$ ein $u \in U^\perp$ existiert mit $P(u, KerL^\perp) = x_1$.
$P(U^\perp, KerL^\perp)$ ist abgeschlossen, denn:
Sei $\{v_i = P(u_i, KerL^\perp)\}$ mit $v_i \to v \in X$, dann ist wegen

$$\|v_i - v_j\| \geq \mu(U^\perp, KerL^\perp)\|u_i - u_j\|$$

auch $\{u_i\}$ eine C.F. in $U^\perp$.
Da $U^\perp$ abgeschlossen ist, konvergiert u_i in $U^\perp$ gegen u. Die Stetigkeit von $P(., KerL^\perp)$ zeigt $v = P(u, KerL^\perp)$.
Angenommen $KerL^\perp \ni x_1 \notin P(U^\perp, KerL^\perp)$, dann ließe sich $KerL^\perp$ darstellen als

$$KerL^\perp = P(U^\perp, KerL^\perp) \oplus W \quad \text{mit} \quad W \neq \{0\} \quad \text{und} \quad W \perp P(U^\perp, KerL^\perp)$$

Alle $u \in U^\perp$ haben eine Darstellung $u = u_1 + u_0$ mit $u_1 \in P(U^\perp, KerL^\perp)$ und $u_0 \in KerL$.
Also $U^\perp \perp W$, woraus $\mu(KerL^\perp, U^\perp) = \mu(U, KerL) = 0$ folgt.

7.3 läßt vermuten, daß es besonders günstig ist, $l_1, ..., l_k$ paarweise verschieden zu finden, so daß $\mu(KerDH_\sigma, span\{e_{l_1}, ..., e_{l_k}\})$ maximal ist. Natürlich ist die Triangulierung und damit σ , H_σ nicht bekannt, bevor Alg. 7.4 ausgeführt wurde. Deshalb wird $DH(x)$ als eine Approximation für das zukünftige DH_σ benutzt.
In diesem Sinne ist auch die Vorgehensweise von Allgower/Georg in [11] S. 242 zu verstehen.
Es wird eine $(L \mid 0) \cdot Q$-Zerlegung von $DH(x)$ berechnet. W wird dann gerade als Q^T genommen. Die letzten k Spalten von W spannen demnach $KerDH(x)$ auf. Weiterhin wird $\pi = id$ gesetzt und die erste n-Fläche ist $\tau = [\sigma; \hat{v}_{n+1}, ..., \hat{v}_{n+k}]$.
Mit der Transformationsformel für die Normalenvektoren auf Facetten aus Lemma 2.5 und $W^{-T} = W$ folgt, daß der Normalraum von τ gleich $KerDH(x)$ ist.

Wie die praktischen Erfahrungen bzgl. der Größe der Startsimplices und damit der Effizienz des Verfahrens zeigen, sollten zukünftigen Entwicklungen von p.l.-Methoden in der Lage sein, auch mit beliebigem W effizient umzugehen.
Die modifizierte Vektor-bewertung, durch welche die zu lösenden LGSe unabhängig von W auf gegebenenfalls dünnbesetzte LGSe zurückgeführt werden, leistet dabei den entscheidenden Beitrag !

Zur Bestimmung der optimalen Indizes $(l_1, ..., l_k) = (\pi(n+1), ..., \pi(n+k))$ ist folgendes Problem zu lösen.

$$max_{\{l_1 \neq .. \neq l_k\}} [\mu(KerDH(x), span\{e_{l_1}, ..., e_{l_k}\})] \tag{7.4}$$

Sei $\{u_1, ..., u_k\} =: U$ eine Orthonormalbasis (ONB) von $KerDH(x)$ die wie oben mittels einer QR-Zerlegung von $DH^T(x)$ (oder einer Approximation davon) berechnet werden kann.
Dann ist $\mu(w, span\{e_{l_1}, ..., e_{l_k}\}) = \sum_{i=1}^k < w, e_{l_i} >^2 = w^T \cdot E \cdot w$, wobei $E = \sum_{i=1}^k e_{l_i} \cdot e_{l_i}^T$ eine symmetrische, positiv semidefinte Matrix ist.
Das Minimum-Problem in (7.4) ist dann

$$min\{\alpha^T \cdot \bar{E} \cdot \alpha \mid \alpha \in \mathbf{R}^k \quad ; \quad \|a\| = 1\}$$

wobei $\bar{E} = U^T \cdot E \cdot U$ eine symmetrische, positiv definite $k \times k$-Matrix ist.
Das Minimum ist durch den kleinsten Eigenwert von $\bar{E}$ gegeben, dessen Berechnung kein Problem

ist.
Unglücklicherweise müssen aber $\binom{n+k}{k} \approx n^k/k!$ k-Tupel $(l_1, .., l_k)$ überprüft werden. Um die Suche abzukürzen, wird ausgenutzt, daß die Diagonalelemente von $\bar{E}$ $\quad \{\bar{E}_{s,s} = (e_{l_i}^T \cdot u_s)^2\}$ obere Schranken für den kleinsten Eigenwert sind.

Algorithmus 7.9 (Finde günstiges π für 7.4-0)
Berechne $(u_1, ..., u_k)$ *eine ONB von* $KerDH(x)$
Berechne $L := L[1], ..., L[n+k]$ *wobei* $L[l] := (L[l].val, L[l].index) := (min_{1 \leq s \leq k}(e_l^T \cdot u_s)^2; l)$
Sortiere L *so daß* $L[1].val \geq ... \geq L[n+k].val$
$max = -\infty$
Do label $l_1 = 1, n$
    Wenn $(max < L[l_1].val)$ *Do label* $l_2 = l_1 + 1, n+1$
        $\vdots$
        Wenn $(max < L[l_{k-1}].val)$ *Do label* $l_k = l_{k-1}, n+k$
            Berechne $\bar{E}$ *abhängig von* $(l_1, .., l_k)$
            Berechne den kleinsten EW von $\bar{E}$*:* λ
            Wenn $(\lambda > max)$ *dann* $max = \lambda$, $\mathbf{i} = (L[l_1].index, ..., L[l_k].index)$
label: *If* $(max < L[l_k].val)$ $l_k = n+k+1$
continue
Setze $\pi(n+1) = \mathbf{i}_1, ..., \pi(n+k) = \mathbf{i}_k$ *und den Rest so daß* π *eine Permutation ist*

Bemerkung 7.10
Die Vereinfachung am Ende des letzten Kapitels, bei der ja $\{\pi(n+1), ..., \pi(n+k)\} = \{n+1, .., n+k\}$ *gelten soll, macht obigen Alg. natülich hinfällig.*
In diesem Fall sollte der Startwert so gewählt werden, daß Abschätzung (7.3) nicht auf eine besonders ungünstige Situation hindeutet.

Kapitel 8

Beschleunigung durch Prädiktor-Schritte im Fall $k = 1$

In diesem Kapitel soll gezeigt werden, wie sich Prädiktor-Techniken nach D. Saupe [16] mit der modifizierten Vektor-Bewertung und den zur Verfügung stehenden LGS-Lösern kombiniert werden können.
Die Idee von [16] ist es, einen Prädiktor-Schritt zu y, einer Schätzung für den nächsten p.l. Lösungspunkt, auszuführen. Ein Simplex $\sigma \ni y$ ist i.A. nicht transversal bzgl. H.
Deshalb wird eine Pertubation $\tilde{H}$ von H definiert, so daß σ transversal bzgl. $\tilde{H}$ ist. Unter gewissen Voraussetzungen führt eine Verfolgung von $\tilde{H}^{-1}(0)$ auf einen Simplex σ^* der transversal bzgl. $\tilde{H}$ und H ist.
Eines der Probleme ist zu entscheiden, in welche Richtung des 1-dimensionalen p.l.-Pfades verfolgt werden soll.
Aus diesem Grund wird eine Invariante des p.l.-Pfades - die Orientierung - eingeführt. Es muß eine Definition für eine solche Orientation gegeben werden, die einerseits die gewünschte Invarianz-Eigenschaft hat und andererseits schnell zu berechnen ist.
Die in diesem Kapitel gemachte Einschränkung $k = 1$ ließe sich zumindest bei Verwendung des Gitter-Alg. auch vermeiden, denn der Alg. sucht ja letztlich doch nur 1-dimensionale Lösungsmannigfaltigkeiten.

Definition 8.1 (Index)
Sei $\tau = [\sigma, \hat{v}_{i_1}]$; $0 \le i_1 \le n+1$ eine v.b. n-Fläche, dann ist $ind(\tau, \sigma) := sign(det F(\sigma) \cdot det \tilde{L}(\tau))$ der Index von τ bzgl. σ. Wobei $F(\sigma) := \begin{pmatrix} 1 & \cdots & 1 \\ v_0 & \cdots & v_{n+1} \end{pmatrix}$ die Eckenmatrix von σ ist.

Obige Definition ist die kanonische Modifaktion der Definition von ind nach [16],[5].

Lemma 8.2 (Index und l.p.-Step)
Sei $\tau = [\sigma, \hat{v}_i]$ eine v.b. n-Fläche. Betrachte den l.p,-Step $\tau \to \tau^* := [\sigma; \hat{v}_j]$; $j \neq i$, dann ist $sign(det\tilde{L}(\tau)) = -sign(det\tilde{L}(\tau^*))$

Beweis Mit den Konventionen von Definition 4.1 ist nach (4.5)

$$\tilde{L}(\tau^*) = \tilde{L}(\tau) + e_{n+1+1} \cdot (e_{j+1} - e_{i+1})^T.$$

Mit Hilfe der nach (4.5) bereits benutzten Beziehung $\tilde{L}^{-1}(\tau) \cdot e_{n+1+1}^T = e_{i+1} - \beta$ und der Aussage von [11, Lemma 7.2.5] bezüglich der Determinate von Rang-1 Modifikationen folgt

$$\begin{aligned} det\tilde{L}(\tau^*) &= det\tilde{L}(\tau)(1 + (e_{j+1} - e_{i+1})^T \cdot \tilde{L}^{-1}(\tau) \cdot e_{n+1+1}) \\ &= det\tilde{L}(\tau)((e_{j+1} - e_{i+1})^T \cdot (e_{i+1} - \beta)) = det\tilde{L}(\tau)(-\beta_j). \end{aligned}$$

Da nach den Regeln des l.p.-Step $\beta_j > 0$ gilt, folgt die Behauptung.

Lemma 8.3 (Index und 1-dim. Pfadverfolgung)
Sei $\sigma^1, \sigma^2, ...$ die Kette der vom p.l. Algorithmus durchlaufenen Simplices und $\tau^s := \sigma^s \bigcap \sigma^{s+1}$ ist die v.b. Pivot-n-Fläche (Door-out Facette von σ^s). Dann gilt $ind(\tau^s, \sigma^s) = -ind(\tau^s, \sigma^{s+1}) = ind(\tau^{s+1}, \sigma^{s+1})$ für jedes $s \geq 1$.

Beweis Identifiziert man σ^s mit der Matrix $(v_0, ..., v_{n+k})$ und σ^{s+1} mit $(v_0^*, ..., v_{n+k}^*)$ läßt sich schreiben $\sigma^{s+1} = \sigma^s R_j \cdot Q$, wobei v_j die Fehlecke der Pivotfacette und
$R_{s-1} = I_{n+k+1} + (e_{s\ominus 1} - 2e_s + e_{s\oplus 1}) \cdot e_l^T$ die sogenannte Reflexionsmatrix ist.
Q ist wieder die Permutations-Matrix, die aus dem Umsortieren der Ecken-Nummern bei einem Pivotschritt über v_0 oder v_{n+k} resultiert (siehe Lemma 4.6).
Also gilt wiederum wegen [11, Lemma 7.2.5] $detR_l = -1$ und damit

$$detF(\sigma^{s+1}) = -detF(\sigma^s) \cdot detQ$$

Seien $\tilde{L}$ und $\tilde{L}^*$ die modifizierten Bewertungsmatrizen der Pivotfacette bzgl. σ^s und σ^{s+1}. Die Gleichung (4.9) impliziert

$$det\tilde{L}^* = \det \tilde{L} \cdot detQ \cdot (1 + e_{n+1+l}^T \cdot \begin{pmatrix} 0 \\ H(r(v_{i_l})) - H(v_{i_l}) \\ \mathbf{0}_k \end{pmatrix} = \det \tilde{L} \cdot detQ \cdot 1.$$

Damit folgt die Behauptung aus Lemma 8.2.

Nun kann die Orientierung erklärt werden.

Definition 8.4 (Orientierung)
Mit den Bezeichnungen von 8.2 ist $or := ind(\tau^s, \sigma^s) = ind(\tau^1, \sigma^1)$ die Orientierung der Kette σ_1, σ_2.

Da $ind(\tau^s, \sigma^s)$ entlang des Pfades invariant ist, ist die Entscheidung darüber, in welche Richtung nach einem Prädiktorschritt verfolgt werden soll, einfach zu treffen.
Ist $ind(\tau, \sigma) = or$, dann muß erst ein l.p.-Step ausgeführt werden und dann kann über die neue v.b. n-Fläche ein Pivotschritt erfolgen.
Im anderen Fall $ind(\tau, \sigma) = -or$ kann gleich über τ dieser Schritt erfolgen.
Es bleibt einen Algorithmus anzugeben, mit dem ind schnell berechnet werden kann.

Lemma 8.5 (erste Äquivalenz) *Sei σ ein transversaler Simplex und $\tau := [\sigma, \hat{v}_{i_1}]$ v.b. .
Dann gilt*

$$ind(\tau, \sigma) = sign(detW \cdot detM) = sign(det \begin{pmatrix} DH_\sigma \\ w_{i_1}^T \end{pmatrix}) \quad ,$$

wobei M die Koeffizienten-Matrix des LGSs (4.15) bzw. (4.16) ist und w_{i_1} der aNev auf $[\sigma; \hat{v}_{i_1}]$ ist.

Beweis Offensichtlich folgt mit (4.17) die zweite Identität aus der ersten.
Der Beweis folgt aus der Faktorisierung von $\tilde{L} = J \cdot \tilde{M} \cdot \tilde{V}$ nach Lemma 4.7. Weiterhin werden die Bezeichnungen von Lemma 3.6 benutzt.
Es gilt $detJ = 1$, da J eine untere Dreiecksmatrix mit Diagonalelementen 1 ist.
Mit der Definition

$$\bar{V} = \begin{pmatrix} 1 & \mathbf{1}_{n+k}^T \\ 0 & V \end{pmatrix} \quad \text{folgt} \quad \mathbf{1}_{n+k+1}^T \cdot \bar{V}^{-1} = e_1^T \quad \text{sowie} \quad det\bar{V} = detV = detW \cdot det\mathcal{V} = detW \cdot det\tilde{V} \quad .$$

Mit der Aussage von [11, Lemma 7.2.5] folgt daraus

$$detF(\sigma) = det\left(\bar{V} + \begin{pmatrix} 0 \\ v_0 \end{pmatrix} \cdot \mathbf{1}_{n+k+1}^T\right) = det(I + \begin{pmatrix} 0 \\ v_0 \end{pmatrix} \cdot e_1^T) \cdot det\bar{V} \overset{7.2.5}{=} (1+0) \det \bar{V} = det(W \cdot \tilde{V}) \quad .$$

Mit dieser Identität und der Faktorisierung von $\tilde{L}$ gilt

$$ind(\tau,\sigma) = sign(detJ \cdot det\tilde{M} \cdot \det\tilde{V} \cdot detW \cdot det\tilde{V}) = sign(detW \cdot det\tilde{M}) \quad .$$

Mit Hilfe des Laplaceschen Entwicklungssatzes folgt sowohl für $i_1 > 0$ also auch für $i_1 = 0$

$$det\tilde{M} = detM \quad .$$

Mit der Formel (4.17) ist die Behauptung bewiesen.

Korollar 8.6 *Es muß also* $d := det\begin{pmatrix} A & B \\ C & F \end{pmatrix}$ *berechnet werden. Da die Determinante unter einfachen Blockeliminationsoperationen invariant ist, gilt* $d = detA \cdot det(F - C \cdot A^{-1} \cdot B)$ *, sofern* A *regulär ist. In diesem Fall ist die Berechnung einfach.*

Für den Fall A singulär könnte ausgenutzt werden, daß A i.d.R. nur nahezu singulär ist.
Offen ist auch noch das Problem der Berechnung der Determinante (genauer: deren Vorzeichen), falls iterative Lösungsmethoden für das GLS (5.1) benutzt werden sollen.

Sei

$$x : I \to \mathcal{R}(H) := \{x \in D \mid rg(DH(x) = n\} \quad \text{die Lösung von} \quad H(x(s)) = \mathbf{0} \quad .$$

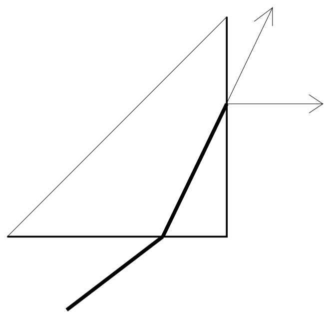
Dann ist

$$sign(det\begin{pmatrix} DH(x(s)) \\ \dot{x}^T(s) \end{pmatrix}) \tag{8.1}$$

invariant.
Analog zu $x(s)$ kann die p.l. Mannigfaltigkeit als nach s parametrisiert betrachtet werden: $x_{pl}(s)$. In diesem Sinne kann für einen transversalen Simplex eine door-in- und eine door-out-Facette definiert werden, wobei erstere die v.b. n-Fläche ist, über welche der Simplex zuerst erreicht wurde.
Für den aNev w_i einer door-out-Facette $\subset \sigma$ gilt dann

$$\dot{x}_{pl}^T(s) \cdot w_i > 0 \quad ,$$

wenn degenerierte Fälle vernachlässigt werden.



Das führt zu folgendem Lemma.

Lemma 8.7 (zweite Äquivalenz)
Sei σ *ein transv. Simplex ,* τ, τ^* *seien die door-in bzw. door-out* n*-Facetten von* σ*. Dann gilt*

$$sign(det\begin{pmatrix} DH_\sigma \\ (H_{\tau^*}^{-1}(\mathbf{0}) - H_\tau^{-1}(\mathbf{0}))^T \end{pmatrix}) = sign(det\begin{pmatrix} DH_\sigma \\ w^{*T} \end{pmatrix}) \overset{Lemma\ 8.5}{=} ind(\tau^*,\sigma) \quad ,$$

wobei w^* *der aNev auf* τ^* *ist.*

Beweis $t := (H_{\tau^*}^{-1}(\mathbf{0}) - H_\tau^{-1}(\mathbf{0}))$ ist eine Basis von $KerDH_\sigma$.
Die dafür eigentlich notwendige Annahme $t \neq \mathbf{0}$, kann dadurch umgangen werden, daß für einen

solchen Fall kurzzeitig $H_\sigma^{-1}(\vec{\epsilon})$ gesucht ist. (Siehe [8] S. 96)
Da τ^* die door-out-Facette ist, gilt

$$w^{*T} \cdot t > 0 \quad . \tag{8.2}$$

Weil $DH_\sigma \cdot t = \mathbf{0}$, ist $t \in (ImDH_\sigma^T)^\perp$. Wegen (8.2) ist deshalb

$$w^* \notin span\{DH_{\sigma 1,.}^T, ..., DH_{\sigma n,.}^T\} = ImDH_\sigma^T \quad .$$

Für die Homothopie

$$v(s) := s \cdot t + (1-s) \cdot w^* \quad \text{für} \quad s \in [0,1] \quad \text{gilt analog}$$

$$v(s)^T \cdot t = s + (1-s)w^{*T} \cdot t > 0 \quad \text{, woraus} \quad v(s) \notin span\{DH_{\sigma 1,.}^T, ..., DH_{\sigma n,.}^T\} \quad \text{folgt.}$$

Also ist für $s \in [0,1]$ die Matrix $\begin{pmatrix} DH_\sigma \\ v^T(s) \end{pmatrix}$ regulär.
Aus $sign\left(det\begin{pmatrix} DH_\sigma \\ v^T(s) \end{pmatrix}\right) = sign\left(det\begin{pmatrix} DH_\sigma \\ w^{*T} \end{pmatrix}\right)$ für $s \in [0,1]$ folgt die Behauptung.

Dieses Lemma besitzt eine nicht unwichtige Folgerung.
Gegeben sei eine Parametrisierung $x(s)$ des kontinuierlichen Problems. Um o.B.d.A $x(0)$ wird entsprechend des letzten Kaptitels ein transversaler Simplex konstruiert.
Ob gleich ein Pivotschritt über die erste v.b. n-Fläche τ ausgeführt wird, oder zuerst noch ein l.p.-Step und dann ein Pivotschritt über die zweite v.b. n-Fläche τ^* , wird danach entschieden, ob $ind(\tau, \sigma)$ mit der Invariante von (8.1) übereinstimmt oder nicht.
Bei diesem Vorgehen folgt aus dem Lemma, daß für hinreichend feine Triangulierungen der kontinuierliche und der p.l. Lösungspfad gleich orientiert sind ($\dot{x}_{pl}^T(s) \cdot \dot{x}(s) > 0$), denn DH_σ ist eine Approximation von $DH(x)$.
Sollte es nötig sein, den Lösungspfad auf verschiedenen Triangulierungen aber in der gleichen Richtung zu durchlaufen, kann die Unabhängigkeit der Orientierung von der Triangulierung angenommen werden.

Die letzte Aussage kann als Begründung des Vorgehens von Lippke[18] dienen, bei der nach jedem Prädiktor-Schritt eine neue Triangulierung berechnet wird !

Eine weitere nunmehr einfache Folgerung kann für die Definition des Index nach [16],[5] angegeben werden.

Lemma 8.8 *Es gilt für* $\tau = [\sigma; \hat{v}_i] = [v_{j_0}, ..., v_{j_n}]$ *die Beziehung*

$$sign\left[det\begin{pmatrix} 1 & \cdots & 1 \\ H(v_{j_0}) & \cdots & H(v_{j_n}) \end{pmatrix} \cdot det\begin{pmatrix} 1 & \cdots & 1 & 1 \\ v_{j_0} & \cdots & v_{j_n} & v_i \end{pmatrix}\right] = -ind(\tau, \sigma) \tag{8.3}$$

Beweis Offensichtlich ist die linke Seite unabhängig von der Reihenfolge der $j_0, .., j_n$, so daß $j_0 < \cdots < j_n$ angenommen werden kann.
Entwicklung der Determinanten nach der letzten Zeile zeigt

$$det\tilde{L}(\tau) = (-1)^{n+1+i+1} det\begin{pmatrix} 1 & \cdots & 1 \\ H(v_{j_0}) & \cdots & H(v_{j_n}) \end{pmatrix} \quad .$$

Weiterhin liefert wiederholtes Vertauschen von Spalten

$$det\begin{pmatrix} 1 & \cdots & 1 & 1 \\ v_{j_0} & \cdots & v_{j_n} & v_i \end{pmatrix} = (-1)^{n+1+i} detF(\sigma) \quad .$$

Diese Gleichungen ergeben die Behauptung.
Da die linke Seite in (8.3) gerade die Definition des Index (bzw. der Orientierung) nach [16],[5] darstellt, und die rechte Seite mittels Lemma 8.5 auf die Definition der Orientierung im Sinne von

[8] ,[11] zurückgeführt werden konnte, wurde mit obigem Lemma eine bislang noch unbekannte Beziehung zwischen diesen Begriffen gezeigt.

Nach diesen Vorarbeiten kann das eigentliche Ziel der Adaption der Prädiktor-methode in Angriff genommen werden. Wie bereits erwähnt, wird eine Schätzung y für den nächsten p.l. Lösungspunkt gemacht. Die eigentliche simpliziale Methode wird dann 'nur' noch für den Korrektor-Schritt benutzt.
Zur Konstruktion einer topologischen Pertubation $\tilde{H}$ von H wird zunächst eine Abbildung $\gamma \in C(\mathbf{R}^{n+1}, \mathbf{R})$ ausgewählt, so daß $v_r := argmin\{\gamma(v_j) \mid 0 \leq j \leq n\}$ eindeutig ist.
Dabei sind $v_0, .., v_n$ die Ecken der n-Fläche, in der y liegt.
Sei $s := min\{\gamma(v_j)\}$, dann ist $\tilde{H}$ gegeben durch

$$\tilde{H}(v) := \begin{cases} -d & \text{für } \gamma(v) \leq s \\ H(v) & \text{sonst} \end{cases} \quad .$$

Eine geeignete Wahl für d ist durch $d := \sum_{j \neq r} H(v_j)$ gegeben. In diesem Fall gilt

$$(H(v_0), \cdots, -d, \cdots, H(v_n)) \cdot \frac{1}{n+1} \mathbf{1}_{n+1} = \mathbf{0} \quad .$$

Insbesondere liegt der Lösunsgpunkt genau im Schwerpunkt der n-Fläche.
Für γ sind in [16] oder [18] verschieden Möglichkeiten angegeben und getestet worden.
Problematisch an der topologischer Pertubation ist, daß $D\tilde{H}$ i.A. nicht das gleiche Besetzungsmuster wie DH hat. Bandstrukturen werden i.A. zerstört, was sehr unangenehm sein kann, wenn gerade diese Struktur ausgenutzt werden soll.

Eine Möglichkeit dies zu umgehen ist die Folgende.
Zunächst fällt auf, daß $L(\sigma)$ maximal zwei künstlich bewertete Spalten haben kann.
Wären unter den $n + 2$ Spalten drei künstlich bewertete Spalten, so wäre es nicht möglich $n + 1$ lienar unabhängige Spalten auszuwählen. Dies überträgt sich natürlich auch auf $\tilde{L}$.
Aus diesem Grund ist $\mathcal{D}\tilde{H}_\sigma$ maximal eine Rang-2 Modifikation von $\mathcal{D}H_\sigma$. Mit der Darstellung von $\mathcal{D}H_\sigma$ aus Lemma 3.6 folgt

$$\begin{aligned} \mathcal{D}\tilde{H}_\sigma = \mathcal{D}H_\sigma + R \cdot S^T \cdot P^T =: \mathcal{D}H_\sigma + R \cdot Z^T \qquad \text{mit} \\ R = (-d - H(v_l), -d - H(v_s)) \quad , \quad S = (e_l - e_{l+1}, e_s - e_{s+1}) \end{aligned} \tag{8.4}$$

Dabei sind l, s die Nummern der künstlich bewerteten Ecken.
Mit den Bezeichnungen von (5.1) gilt dann i.A.

$$\mathcal{D}_u\tilde{H}_\sigma = \tilde{A} = A + R \cdot Z^T$$

Zur Lösung eines LGS mit $\tilde{A}$ als Koeffizienten-Matrix kann die Sherman/Morrison/Woodbury Formel [22, Lemma 3.4] benutzt werden:

$$\tilde{A}^{-1} = A - A^{-1}RU^{-1}Z^TA^{-1} \quad \text{mit } U = I_2 + Z^{-1}A^{-1}R \in M(2) \quad .$$

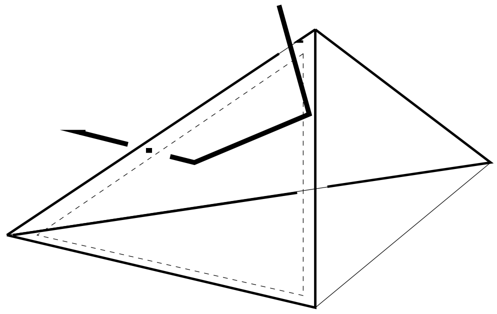
Es sind also zusätzlich maximal 2 LGSe zu lösen.

Kapitel 9

Numerische Stabilität

Es ist klar, daß die Lösungen der LGSe mit Fehlern behaftet sind.
Zunächst scheinen diese Fehler dann wichtig zu sein, wenn sie zu einer falschen Entscheidung darüber führen, welche n-Fläche die nächste v.b. n-Fläche ist.
Offensichtlich ist eine solche 'Verwechslung' dann wahrscheinlich, wenn der zugehörige Lösungspunkt am Rand der n-Fläche liegt, also wenn mindestens eine Komponente von $L^{-1}(\tau^*) \cdot e_1$ klein ist. Mit $\varepsilon > 0$ sei eine Schranke für 'klein' bezeichnet.

In nebenstehender Zeichnung für den Fall $n = 2$, $k = 1$ sind die etwas dickeren Linien jeweils der Rand einer v.b. n-Fläche.
Die dicken Linien stellen einen Teil der Lösungsmannigf. dar und die gestrichelten Linien zeigen den 'kritischen' Bereich.



Das nächste Lemma gibt Anlaß zu größtem Pessimismus, denn es wird gezeigt, daß für großes n immer mindestens eine Komponente $< \varepsilon$ ist.

Lemma 9.1 *Sei* $\varepsilon > 0$ *und* $\tilde{V}_1 := \{(x_0, ..., x_n) \in \mathbf{R}_+^{n+1} \mid \sum_{i=0}^n x_i = 1\}$ *die Menge der baryzentrischen Koodinaten einer n-Fläche sowie* $\tilde{W}_\varepsilon := \{(x_0, ..., x_n) \in \bar{V}_1 \mid \exists x_i < \varepsilon\}$ *die Menge der Koordinaten der kritischen Punkte.*
Dann $\exists N(\varepsilon) : \forall n \geq N \quad Vol(\tilde{W}_\varepsilon)/Vol(\tilde{V}_1) = 1$, *also* $\tilde{W}_\varepsilon = \tilde{V}_1$.

Bemerkung Obige Aussage ist eigentlich offensichtlich, denn das Minimum von $n + 1$ Zahlen $x_0, .., x_n$, deren Summe gleich 1 ist, muß kleiner gleich $1/(n+1)$ sein. Ansonsten wäre $x_0 + \cdots + x_n > (n + 1) \cdot 1/(n + 1) = 1$.
Im folgenden Beweis wird allerdings das Verhältnis der Volumina genau quantifiziert.

Beweis von Lemma 9.1. Zunächst wird

$$\tilde{V}_\alpha := \{(x_0, ..., x_n) \in \mathbf{R}_+^{n+1} \mid \sum_{i=0}^{n} x_i = \alpha\} \quad \text{und} \quad V_\alpha := \{(x_1, ..., x_n) \in \mathbf{R}_+^{n} \mid \sum_{i=1}^{n} x_i \leq \alpha\}$$

definiert. Weiterhin sei

$$W_\varepsilon := \{(x_1, ..., x_n) \in \bar{V}_1 \mid \exists x_i < \varepsilon \text{ oder } \sum_{i=0}^{n} x_i \leq 1 - \varepsilon\} \quad .$$

Da die Gram'sche Determinante g von $\phi : V_1 \mapsto \tilde{V}_1; \phi(x_1, ..., x_n) = (1 - x_1 - ... - x_n, x1, ..., x_n)$ konstant ist, gilt

$$Vol(\tilde{V}_\alpha) = \int_{V_\alpha} 1dS = g \int_{V_\alpha} 1dx = gVol(V_\alpha)$$

Es ist also $Vol(W_\varepsilon)/Vol(V_1) = 1$ für $n \geq N$ zu zeigen.
Mit der Translation $T_\varepsilon(x) := x - \varepsilon \cdot \mathbf{1}_n$ folgt

$$T_\varepsilon(V_1 \backslash W_\varepsilon) = V_{1-(n+1)\varepsilon}$$

Also $Vol(W_\varepsilon)/Vol(V_1) = 1 - \frac{|1-(n+1)\varepsilon|_+^n}{1}$.

Ein jetzt entstehendes Problem ist, daß bspw. nicht davon ausgegangen werden kann, daß der p.l.-Pfad für einen Kreis wieder genau auf den Start-Simplex stößt. Es ist vielmehr wahrscheinlich, daß der Pfad daran vorbeiläuft, bis irgendwann doch einmal ein schon berechneter Simplex gefunden wird.
Für die Approximation 1-dimensionaler Mannigfaltigkeiten (insbesondere wenn sie nicht diffeomorph zu einem Kreis sind) tritt das Problem selten (oder nicht) auf.
Bei der Approximation mehrdimensionaler Mannigfaltigkeiten müssen Lösungen für dieses Problem gefunden werden, denn hier sind die 1-dim. Pfade, die der Gitter-Alg. errechnet, sehr oft homöomorph zu einem Kreis (siehe Abb. zum zweiten num. Bsp.).
Eine Möglichkeit ist , im Gitter-Alg. die Suche '$\tau \in \Sigma$?' zu erweitern auf: Finde $\tau^1, .., \tau^r \subset \Sigma$ mit $b(\tau) - \varepsilon \leq_{lex} b(\tau^i) \leq_{lex} b(\tau) + \varepsilon$. $\varepsilon = (\varepsilon_1, ..., \varepsilon_{n+k})$ ist dabei eine erlaubte Abweichung. Natürlich sollte gelten $\varepsilon_{\pi(n+1)} = .. = \varepsilon_{\pi(n+k)} = 0$, wenn $\tau = [[v_0, \pi]; \hat{v}_{n+1}, .., \hat{v}_{n+k}]$ ist.

Mit diesen algorithmischen Korrekturen ist aber noch nicht geklärt, ob die Fehler beschränkt bleiben. Die Idee der folgenden Ausführungen ist zu zeigen, daß die vom Alg. erzeugten Lösungspunkte exakte Lösungspunkte von $\bar{H}_\mathcal{T}$ sind, wobei dann $\|H - \bar{H}\| \leq K$ mit möglichst kleinem K gelten soll.

Durch den Alg. des l.p.-Step in der ursprünglichen Form (Alg 3.4) wird jeweils $C \approx L^{-1}(\tau)$ aufdatiert. Wegen der Rundungsfehler ist C i.A. nur eine Approximation der Inversen von $L(\tau)$. Im Falle der modifizierten Vektorbewertung ist dann also C^0 nur eine Approximation der ersten Spalte von $\tilde{L}(\tau)$.
Aus diesem Grund werden die Rechnungen zunächst nur mit im Sinne der normalen Vektor-Bewertung durchgeführt. Außerdem wird nur der Fall $k = 1$ betrachtet, wo C im Prinzip immer durch Aufdatieren berechnet wird,da ein Restart maximal einmal ausgeführt wird.

Die erste Beobachtung ist, daß für C bis auf extrem selten vorkommende Fälle exakt $C^0 > 0$ gilt.
Beim $l.p. - Step$ wird

$$j = argmin\{\frac{C_r^0}{\beta_r} \mid \beta_r > 0\} \tag{9.1}$$

bestimmt. Dann wird

$$C_r^{*0} = C_r^0 - \frac{C_j^0}{\beta_j}\beta_r$$

aufdatiert.
Sollte dabei dabei ein r mit $C_r^{*0} < 0$ entstehen, kann $C_r^{*0} = 0$ gesetzt werden, da < 0 nur durch Rundungsfehler entstanden sein kann.
In diesem Fall oder für $C_r^{*0} = 0$ war also das Minimum (9.1) nicht oder nahezu nicht eindeutig, denn

$$0 \approx C_r^{*0} = C_r^0 - \frac{C_j^0}{\beta_j}\beta_r \quad \Leftrightarrow \quad \frac{C_r^0}{\beta_r} \approx \frac{C_j^0}{\beta_j} \quad .$$

Beim nächsten l.p.-Step für C^{**} können nun zwei Fälle eintreten.
1. Fall $\beta_r^* \leq 0$, dann gilt

$$C_r^{**0} = C_r^{*0} - \frac{C_{j^*}^{*0}}{\beta_{j^*}^*}\beta_r^* \geq 0$$

2. Fall $\beta_r^* > 0$, dann ist $j^* = r$ also

$$C^{**0} = C^{*0} \geq 0$$

Da aber C bzw. C^* sich trotzdem verändert, tritt in der Praxis irgendwann der erste Fall ein, so daß dann i.d.R. $C^0 > 0$ gilt.

Die nächste Beobachtung ist, daß C als die exakte Inverse von $\bar{B} = B + \Delta B$ mit $B = L(\tau)$ betrachtet werden kann.
Nach einer Idee von Georg/Hettich [21] kann untersucht werden, wie sich ΔB bei einem Sherman-Morrison Update zu ΔB^* verändert. Dabei werden lokale Fehler beim Update-Schritt selbst vernächlässigt.
Um sich spätere Diskussionsmöglichkeiten offenzuhalten, wird jetzt noch nicht spezifiziert, wie $\beta \approx L^{-1}(\tau) \cdot y$ ermittelt wird. Der Sherman-Morrison Update für C ist dann mit $(i = j + 1)$

$$C^* = C - \frac{1}{e_i^T \cdot \beta}(\beta - e_i) \cdot e_i^T \cdot C \tag{9.2}$$

Nochmalige Anwendung der Sherman-Morrison Formel ergibt dann für $\bar{B}^* = L(\tau^*) + \Delta B^*$

$$\begin{aligned} \bar{B}^* &= \bar{B} + \frac{\bar{B}(\beta - e_i) \cdot e_i^T \cdot C \cdot \bar{B}}{e_i^T \beta (1 - \frac{e_i^T \cdot C \cdot \bar{B}(\beta - e_i)}{e_i^T \cdot \beta}} \\ &= \bar{B} + \bar{B}(\beta - e_i) \cdot e_i^T \end{aligned} \tag{9.3}$$

Im Falle des bisherigen Konzepts der Vektor-Bewertung gilt (unter Vernachlässigung von Rundungsfehlern bei der Matrix Multiplikation) $\beta = C \cdot y$. Dabei ist y exakt die neue Spalte der Bewertungsmatrix $L(\tau^*)$. Damit erhält man aus (9.3)

$$\bar{B}^* = B^* + \Delta B^* = B + \Delta B + (y - B \cdot e_i) \cdot e_i^T - \Delta B \cdot e_i \cdot e_i^T \tag{9.4}$$

Mit $B^* = B + (y - B \cdot e_i) \cdot e_i^T$ folgt daraus

$$\Delta B^* = \Delta B(I - e_i \cdot e_i^T) \quad .$$

Dies ist die in [21] hergeleitete Formel, die zeigt, daß der Fehler ΔB bei jedem Spaltenaustausch gefiltert wird. Bei exakter Rechnung wäre ein Anfangsfehler nach $n + 1$ Spaltenaustauschen weggefiltert.
Wenn jetzt noch die erste Zeile von ΔB^* ungefähr 0 wäre, dann könnte C^* offensichtlich als lexik. positive Inverse einer Bewertungsmatrix von τ^* bzgl. $\bar{H}$ betrachtet werden.
Weiterhin wäre $\|H - \bar{H}\| = \|\Delta B\|$ immer in der Größenordnung von n-mal dem lokalen Fehler.
Leider gilt i.A. nicht $e_1^T \cdot \Delta B = 0$.
Es gibt aber eine recht einfache Möglichkeit die lokalen Fehler für die erste Zeile von ΔB sehr klein zu halten.
Dazu benutzt man die Kaczmarz-Projektion [23] für die erste Zeile von $L(\tau) = \mathbf{1}_{n+1}$:

$$f_x(z) = z + \frac{(x - \mathbf{1}^T \cdot z) \cdot \mathbf{1}}{n + 1}$$

Anstatt $C \cdot y$ wird dann $\beta = f_1(C \cdot y)$ benutzt. Bevor für diese Wahl von eine Analyse von ΔB^* im Sinne von (9.4) gegeben wird, soll gezeigt werden, daß durch die Wahl von β $\quad e_1^T \cdot \Delta B \approx 0$

erreicht wird.
Dazu wird (9.2) von links mit $\mathbf{1}^T$ multipliziert.

$$\mathbf{1}^T \cdot C^* = \mathbf{1}^T \cdot C - \frac{1}{e_i^T \cdot \beta}(\mathbf{1}^T \cdot \beta - 1) \cdot e_i^T \cdot C$$

Durch die Kaczmarz-Projektion gilt bis auf nur sehr kleine Rundungsfehler $\mathbf{1}^T \cdot \beta = 1$, so daß $\mathbf{1}^T \cdot C^* \approx \mathbf{1}^T \cdot C \approx e_1^T$ ist.
Also muß die Inverse von C^*, C in der ersten Zeile $\mathbf{1}^T$ haben !
Bei obiger Argumentation ist zu beachten, daß $e_i^T \cdot \beta = \beta_j$ der Nenner eines Bruches mit Minimum-Eigenschaften (9.1) ist, so daß β_j i.d.R. nicht sehr klein ist (siehe Tabelle weiter unten) .

Die Fehleranalyse im Sinne von (9.4) ergibt mit modifiziertem β

$$\bar{B}^* = B^* + \Delta B^* = B + \Delta B + (y - B \cdot e_i) \cdot e_i^T - \Delta B \cdot e_i \cdot e_i^T + \frac{1 - \mathbf{1}^T \cdot C \cdot y}{n+1} \bar{B} \cdot \mathbf{1} \cdot e_i^T \qquad (9.5)$$

Also $\Delta B^* = \Delta B(I - e_i \cdot e_i^T) + \frac{1-\mathbf{1}^T \cdot C \cdot y}{n+1} \bar{B} \cdot \mathbf{1} \cdot e_i^T$.
Mit der vorangegangenen Argumentation ($\mathbf{1}^T \cdot C \approx e_1^T$)und $y_0 = 1$ folgt, daß der gegenüber (9.4) zusätzliche Term klein ist, so daß wieder ein filterndes Fehlerverhalten zu erwarten ist.

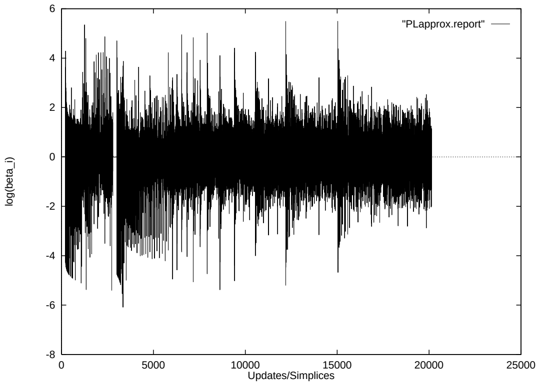
Bei der Analyse der modifizierten Vektor-Bewertung, wird davon ausgegangen, daß aus dem Vektor $\tilde{L}^{-1}(\tau) \cdot y$ die k Komponenten, die Null sein müssen einfach gestrichen werden, um β zu berechnen. Es wird also ein LGS der Form (4.11) gelöst. Dadurch kann die Formel (9.3) angewandt werden. Weiterhin wird die Annahme gemacht, daß bei der Berechnung von β nur lokalen aber NICHT mit akkumulierten Fehlern auftreten. Deshalb wird $\beta = B^{-1} \cdot y$ angenommen.
Diese Annahme ist insofern gerechtfertigt, da zumindest in den Beispielrechnungen die Lösung von parmeterabhängigen gewöhnlichen DG gesucht war. Es wurde daher auf eine Aufdatierung einer LR-Zerlegung (o.ä.) verzichtet.
Mit dieser Annahme folgt aus (9.3)

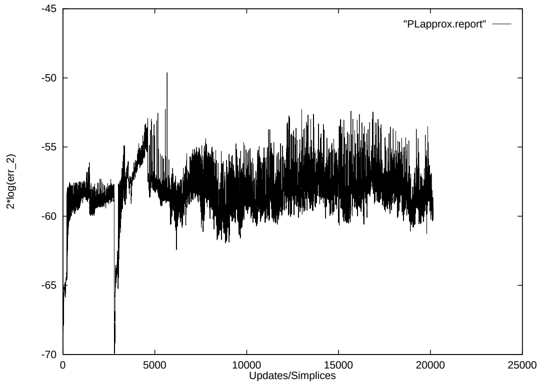
$$\Delta B^* = \Delta B(I - e_i \cdot e_i^T + B^{-1} \cdot y \cdot e_i^T) = \Delta B(I - (e_i - \beta) \cdot e_i^T) =: \Delta B \cdot T_{i,\beta}^T \quad . \qquad (9.6)$$

Für eine Fehleranalyse ist das (gegebenenfalls unendliche) Produkt von Operatoren der Form $T_{i,\beta} := I - e_i \cdot (e_i - \beta)^T$ zu untersuchen.
Diese Operatoren haben nur reelle Eigenwerte (EW). Offensichtlich sind alle Vektoren $v \in span\{e_i - \beta\}^\perp$ Eigenvektoren (EV) zum EW 1. Gilt zusätzlich $\beta_j \neq 1$, dann ist $e_i \notin span\{e_i - \beta\}^\perp$, so daß der letzte der möglichen EWe gleich β_j zum EV e_i ist.
Wegen der Regeln des l.p.-Steps gilt $\beta_j > 0$, aber die Eigenschaft $\beta_j \leq 1$, welche für eine Stabilitätsbedingung der Form $\|T_{i,\beta}\| \leq 1$ notwendig wäre, gilt i.A. nicht.
Trotzdem zeigt die modifizierte Vektorbewertung gute Stabiliätseigenschaften. Es wurde die durch Aufdatieren erhaltene erste Spalte von C mit einer in jedem Simplex neu berechneten Lösung $L^{-1}(\tau) \cdot e_1$ verglichen:

$$err_2 = \|C \cdot e_1 - L^{-1}(\tau) \cdot e_1\|^2 \quad .$$

Nachfolgende Tabellen zeigt die Verteilung von $ln(\beta_j)$ und $ln(err_2)$ während der Berechnung des dritten numerischen Beispiels.





Da zu Beginn der Rechnung und bei jedem Restart im p.l.-Algorithmus die erste Spalte nicht durch Aufdatieren, sondern durch Lösen des LGS (4.11) ermittelt wird, ist der Fehler err_2 am Anfang (Simplex 1 ,2 ,..) und nach dem Restart beim etwa 2500. Simplex besonders klein. Entsprechend dieser Tabellen haben die Operatoren $T_{i,\beta}$ recht oft eine Norm größer als 1.

Immerhin gilt aber mit großer Genauigkeit $\mathbf{1}^T \cdot y = 1$, denn sei $z \in \mathbf{R}^{n+k}$, so wird die Lösung $x = (x_0, \bar{x})$ von $\tilde{L}(\tau) \cdot x = y$ durch folgende Formeln bestimmt
1. Fall $i_1 = 0$ $x_0 = 1 - z_{\pi(1)}, \bar{x} = \mathcal{V}^{-1} \cdot z$
2. Fall $i_1 > 0$ hier gilt exakt $z_{\pi(1)} = \delta_{\pi(1)}$ und $x_0 = 0, \bar{x} = \mathcal{V}^{-1} \cdot z$

Da $\mathbf{1}^T \cdot \mathcal{V}^{-1} = e_{\pi(1)}$ folgt in beiden Fällen $\mathbf{1}^T \cdot y = 1$, im Wesentlichen unabhängig davon, wie genau die Lösung z des transformierten LGS ist.

Kapitel 10

Numerische Beispiele

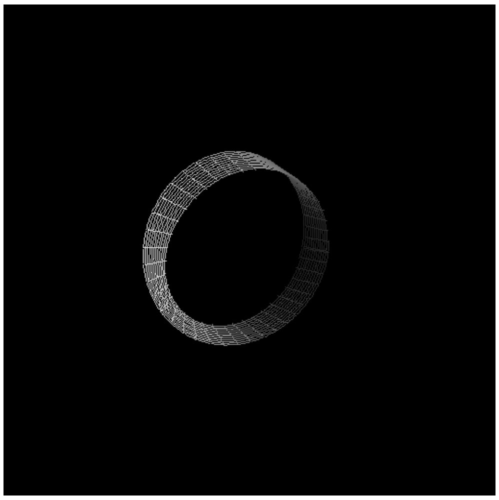
Im Folgenden seien drei numerische Beispiele dargestellt, die mit einem auf den bisherigen Ausarbeitungen basierenden Code erzeugt wurden. Da der Code selbst sehr lang ist (ca 2500 Zeilen) und andererseits alle wesentlichen Algorithmen bereits dargestellt wurden, erspare ich mir den Abdruck des Codes.
Zu bemerken ist, daß in dem Code kein Gebrauch von den beschleunigenden Techniken gemacht wurde, obgleich die angegebenen Anzahlen von traversierten Simplices die Notwendigkeit dieser Techniken deutlich machen.

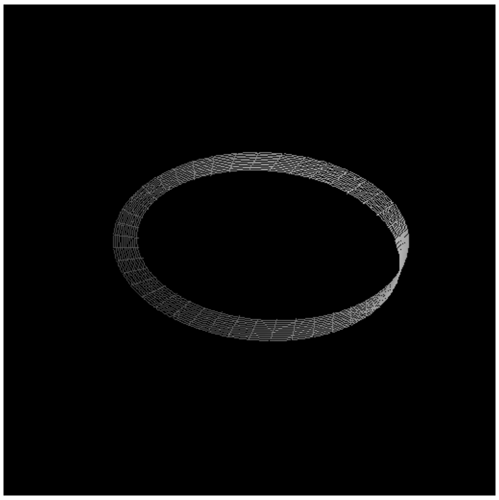
Beispiel 10.1 (Möbius-Band)

In diesem Fall wurde die Lösungsmenge von

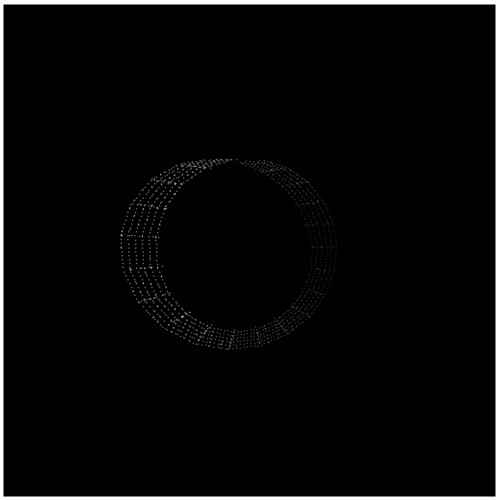
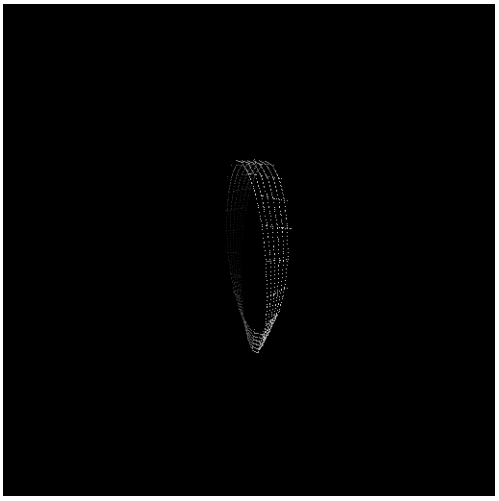
$$H : \mathbf{R}^3 \times [-1,1] \times [0,2\pi] \to \mathbf{R}^3; \qquad (x,r,t) \mapsto x - \begin{pmatrix} sin(t) & 0 & -cos(t) \\ 0 & 1 & 0 \\ -cos(t) & 0 & sin(t) \end{pmatrix} \cdot \begin{pmatrix} R + sin(t/2) \\ cos(t/2) \\ 0 \end{pmatrix}$$

berechnet. Das Matrix-Vektor-Produkt in der Definition von H stellt eine Parametrisierung des Möbiusbandes mit Windungszahl 2 dar.
Es gilt $D_x H = Id$, so daß das Problem gut in den Kontext der Dünnbesetztheit von DH_σ und der Methoden aus dem Kapitel 5 paßt.
Nachfolgend sind einige Abbildungen der ermittelten Lösungsmenge wiedergegeben.





Bei obigen Abb. wurden die Lösungspunkte mit Linien verbunden. Dadurch wird das Vorhandensein von Simplex-Clustern verdeckt, auf denen H_σ affin linear ist. In den beiden folgenden Abb. wurden nur die Lösungspunkte auf den KnF gezeichnet. Die Kantenlänge der Triangulierung ist doppelt so groß, wie in der Rechnung für obige Bilder.

Deutlich ist zu erkennen, daß jeweils die Lösungspunkte mehrerer KnF auf einer Geraden liegen. Dieser Effekt kann mittels der Update-Formeln aus Lemma 3.10 erklärt werden. Dazu wird der Einfachheit halber der Fall $0 < j < n+k$ und $\pi(j) \leq 3$ betrachtet. Aus der Definition von H folgt, daß H in den ersten 3 Argumenten $(x_1, .., x_3)$ linear ist.
Beim Aufdatieren ist folgender Term zu betrachten

$$\begin{aligned}
H(v_{j+1} - H(r(v_j)) + H(v_{j-1}) - H(v_j)) &= H(v_{j-1} + e_{\pi(j)} + e_{\pi(j+1)}) \\
&\quad -H(v_{j-1} - v_{j-1} - e_{\pi(j)} + v_{j-1} + e_{\pi(j)} + e_{\pi(j+1)}) \\
&\quad +H(v_{j-1}) - H(v_{j-1} + e_{\pi(j)}) \\
&= H(v_{j-1} + e_{\pi(j)} + e_{\pi(j+1)}) \\
&\quad -H(v_{j-1} + e_{\pi(j+1)}) \\
&\quad +H(v_{j-1}) - H(v_{j-1} + e_{\pi(j)}) \\
\text{Da } H \text{ linear in } x_1, .., x_3 & \\
&= H(v_{j-1} + e_{\pi(j+1)}) + c_{\pi(j)} \cdot e_{\pi(j)} - H(v_{j-1} + e_{\pi(j+1)}) \\
&\quad +H(v_{j-1}) - H(v_{j-1}) - c_{\pi(j)} \cdot e_{\pi(j)} \\
&= \mathbf{0}
\end{aligned}$$

Dabei wurde $\delta = 1$ gesetzt.
Analoge Rechnungen gelten für $j = 0, n+k$. D.h. bei jedem Pivotschritt um eine Facette $[\sigma; \hat{v}_j]$ mit $\pi(j) \leq 3$ verändert sich DH_σ nicht. Für solche Fälle ist die zweite Methode von M.J. Todd geeignet.

Beispiel 10.2 (Zylindrischer Reaktor)

Dieses Beispiel wurde auf Anregung von H. Buss [24] behandelt. In der Darstellung wird seiner Diplomarbeit gefolgt. Es handelt sich dabei um zwei gekoppelte gewöhnliche Differentialgleichungen, welche die stationären Konzentrations- und Temperaturverteilungen in einem zylindrischen Reaktor beschreiben. Dabei wird ein Stoff A in einen Stoff B umgewandelt.
Alle Zustandsgrössen seien nur vom Abstand zum Reaktoreinlaß s abhängig.
Die Differentialgleichungen und Randbedingungen lauten

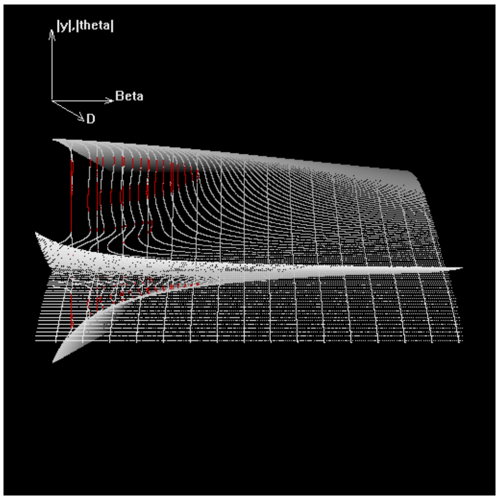
$$\begin{aligned}
0 &= \frac{1}{P_y} y'' - y' - Dy \exp(\gamma - \gamma/\theta) \qquad \text{für} \qquad s \in [0,1] \\
0 &= \frac{1}{P_\theta} \theta'' - \theta' - DB\theta \exp(\gamma - \gamma/\theta) - \beta(\theta - 1)
\end{aligned}$$

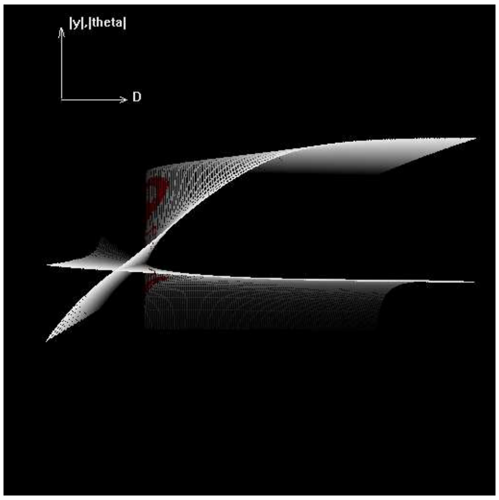
$$\begin{aligned} y'(0) &= P_y(y(0)-1) \quad ; \quad y'(1) = 0 \\ \theta'(0) &= P_\theta(\theta(0)-1) \quad ; \quad \theta'(1) = 0 \end{aligned}$$

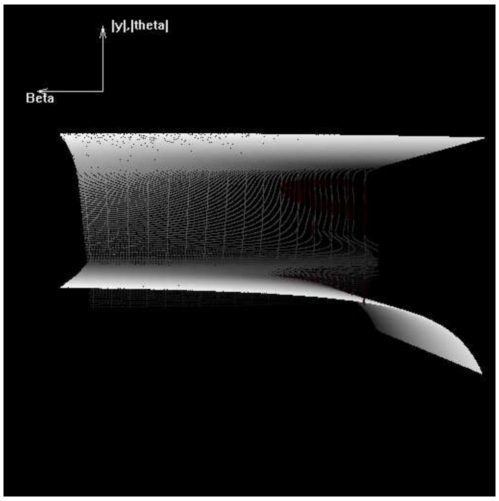
Folgende Tabelle gibt eine Kurzbeschreibung der auftauchenden Grössen. Für weitere Informationen sei auf die Diplomarbeit von H. Buss verwiesen.

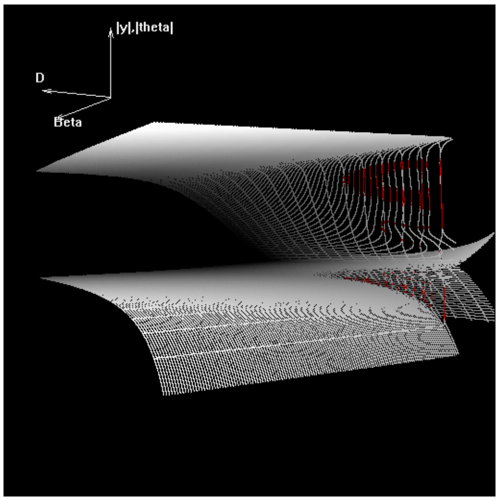
Variable	Wert	Beschreibung
y		normierte Konzentration von A
θ		normierte Temperatur
s	[0,1]	Abstand vom Reaktoreinlass
P_y	5.0	Peclet-Zahl für Massetransport
P_θ	5.0	Peclet-Zahl für Wärmeleitung
γ	25.0	reduzierte Aktivierungsenergie
B	0.5	reduzierte Reaktionswärme
β	-0.35...11.0	reduzierter Wärmetransportkoeffizient
D	-0.5...2.7	Damköhlersche Zahl

Das DGL-system wurde auf einem äquidistanten Gitter mit Maschenweite $h = 1/10$ diskretisiert, so daß $n = 20$ und $k = 2$ ist.
In den folgenden Bildern ist $\|y\|_2, \|\theta\|_2 \in [0.36, 4.8]$ über β und D aufgetragen.









In den Bildern wurden die Punkte, bei denen für $A = D_u H_\sigma$ ein Vorzeichenwechsel der Determinanten auftritt, ROT markiert. Offensichtlich ist dieses Kriterium zur Detektion von Umkehr- und Verzweigungspunkten bei p.l.-Methoden nicht geeignet.
Für obige Bilder wurden ca. 430.000 KnF und ca. 10.000.000 Simplices traversiert. In Anbetracht dieser großen Zahl ist die Rechenzeit von etwa 1.5 Stunden eher klein.

Beispiel 10.3 (nach A.Lippke[18])

In diesem Fall stammt H aus der Diskretisierung folgender gewöhnlicher DGL

$$u'' + f(x,u,\lambda) = 0 \quad \text{für} \quad x \in [0,1] \quad ; \quad \lambda \in [-0.8, 1.0] \tag{10.1}$$

mit den Randbedingungen

$$u(0) = u(1) = 0 \quad .$$

Dabei ist

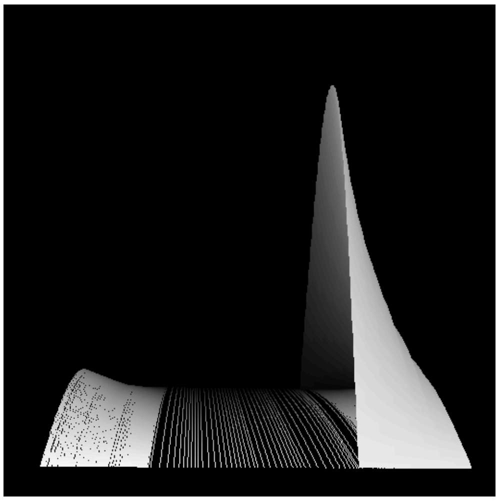
$$f(x,u,\lambda) = 2q(\lambda) + \pi^2 \lambda p(u - q(\lambda)x(1-x)) \quad \text{mit} \quad q(s) = s^2 e^{-s/2} \quad ; \quad p(s) = s + s^2 \quad .$$

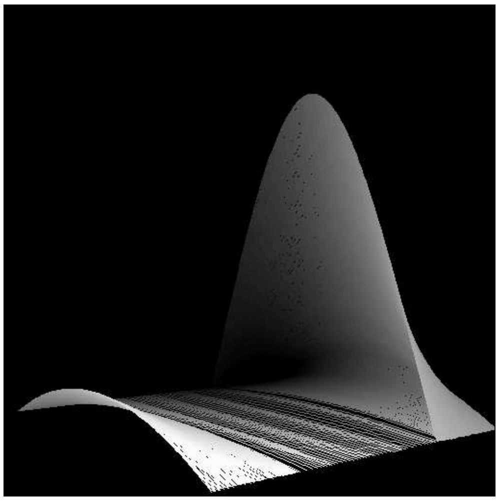
Offensichtlich ist

$$q(\lambda)x(1-x) \tag{10.2}$$

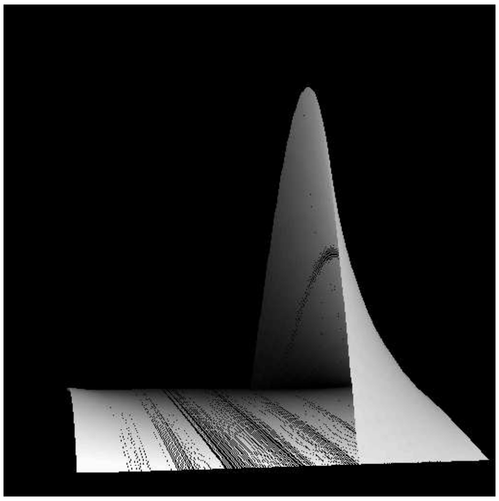
eine Lösung der DGL. Eine Linearisierung der DGL (10.1) in dieser Lösung zeigt, daß für $\lambda = k^2$ ($k \in \mathbf{N}$) von der Lösung andere Lösungen abzweigen können.

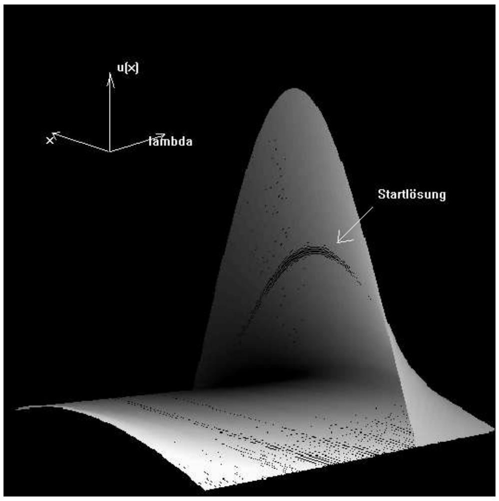
Die DGL wurde mittels zentraler Differenzen auf einem äquidistanten Gitter der Maschenweite $h = 1/100$ diskretisiert.
In diesem Fall ist also $n = 100$ und $k = 1$. Die Kantenlänge der Triangulierung δ wurde mit 0.2 recht groß gewählt.
Als Startwert wurde die offensichtliche Lösung $\lambda = 0$ und $u = 0$ benutzt.
Wird die Lösung u als $u(x,\lambda)$ dargestellt, läßt sich die Lösung als Fläche im $\mathbf{R}^3$ repräsentieren. Die folgenden Bilder zeigen diese Fläche. Dabei läuft im ersten Bild der Parameter λ von links nach rechts und x in die Tiefe.





Eine interessante Beobachtung ist nun, daß die Lösungen weg vom Startpunkt sehr dicht liegen, was zu einer scheinbar geschlossenen Fläche führt. In der Nähe des Startpunktes liegen die Lösungen aber recht weit auseinander.
Daß dieses Phänomen nicht allein an der Wahl des Startpunktes liegt, zeigen die folgenden Bilder, bei denen der Startwert $\lambda = 0.5$ gewählt wurde. Die Startlösung u wurde mittels Newton-Verfahren ausgehend von $v(x) = 4 \cdot 1.18x(1-x)$ berechnet. Diese Startlösung ist weniger günstig für die Konstruktion einer ersten v.b. n-Fläche, so daß hier $\delta = 0.17$ gesetzt werden mußte.





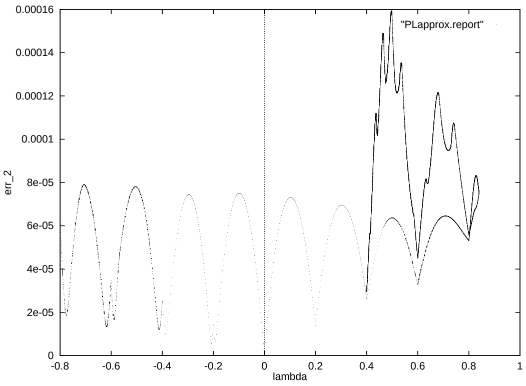
Etwa in der Mitte der Fläche, die durch den oberen Lösungsast gebildet wird, ist der Startpunkt. Dort liegen die Lösungspunkte weniger dicht beeinander.
Die Ursache für das Phänomen liegt meines Erachtens darin, daß die erste v.b. n-Fläche so konstruiert wird, daß ihr Schwerpunkt genau in einem Lösunsgpunkt des kontinuierlichen Problems liegt. Geht nun der p.l.-Pfad ebenfalls ungefähr durch den Schwerpunkt, kann davon ausgegangen werden, daß die Länge der Strecke $H_{\mathcal{T}}^{-1}(0) \bigcap \sigma$ wesentlich größer als der in Gleichung (2.2) vorhergesagte Wert von $\delta/(C(k) \cdot 100^{3/2})$.
Offenbar ist der Schnitt der p.l. Lösungsmenge dann auch für die nächsten Simplices wesentlich größer als die pessimistische Abschätzung. Dieser Effekt ist ausgesprochen günstig, falls bspw. die Prädiktor-Korrektor Methode von A. Lippke verwendet wird. Wenn um jede Schätzung y eines Prädiktor-Schrittes eine neue Triangulierung konstruiert wird, profitiert JEDE p.l. Korrektur zunächst von dieser günstigen Situation.

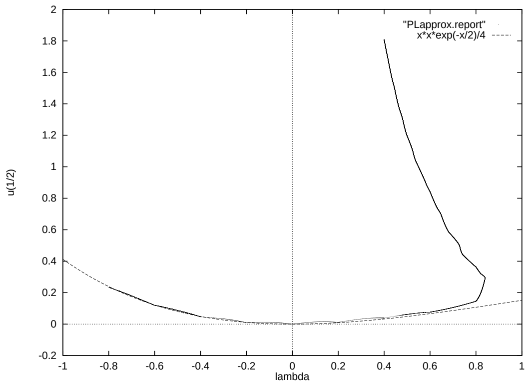
Abgeschlossen werden soll die Betrachtung dieses Beispiels mit einigen Werten bezüglich der Genauigkeit des Verfahrens. Dazu zeigen die nächsten Tabellen die Größe des Residuums

$$err_2 = \|H(H_\sigma^{-1}(0))\|_2$$

und den Vergleich der Scheitelpunkte von $q(\lambda)x(1-x)$ und $H_\sigma^{-1}(0)$ - also von

$$(H_\sigma^{-1}(0))(1/2, \lambda) \quad \text{mit} \quad q(\lambda)1/2(1-1/2) \quad .$$





Beim rechten Bild sei darauf hingewiesen, daß bei $\lambda = 1$ ein Bifurkationspunkt ist.

Dank

Herzlich bedanken möchte ich hiermit bei Herrn Prof. Hans-Georg Jeggle und Dr. Michael Hinze für ihre moralische und fachliche Unterstützung.

Anhang A

Rechnungen für die Beweise einiger Lemmata

A.1 Beweis für Lemma 2.5

Zunächst werden die äußeren Normalenvektoren auf die Facetten von $[0, id]$ berechnet. Anschließend wird gezeigt, wie sich die aNev bei affinen Transformationen verändern.

Die Vektoren $\{v_{j_1} - v_{j_0}, ..., v_{j_{m-1}} - v_{j_0}\}$ spannen den Tangentialraum von τ auf. Deshalb muß w_i zunächst normiert sein und

$$\mathbf{0}_{m-1} = \begin{pmatrix} v_{j_1}^T - v_{j_0}^T \\ \vdots \\ v_{j_{n+k-1}}^T - v_{j_0}^T \end{pmatrix} \cdot w_i =: A \cdot w_i$$

erfüllen.
Für den Fall $i = 0$ ist $j_0 = 1, ..., j_{m-1} = m$, so daß $A = (e_2, e_2 + e_3, ..., e_2 + .. + e_m)^T$ ist.
Es folgt unmittelbar $w_0 = \pm e_1$.
Analog ist für den Fall $i = m$ $\quad j_0 = 0, ..., j_{m-1} = m - 1$ also $A = (e_1, e_1 + e_2, ..., e_1 + .. + e_{m-1})^T$, woraus $w_m = \pm(-e_m)$ folgt.
Im Falle $0 < i < m$ gilt $j_0 = 0, .., j_{i-1} = i - 1, j_{i+1} = i + 2, .., j_{m-1} = m$ also
$A = (e_1, .., e_1 + .. + e_{i-1}, e_1 + .. + e_{i+1}, ..., e_1 + .. + e_m)^T$, woraus $w_i = \pm(e_{i+1} - e_i)/\sqrt{2}$ folgt.

Um zu zeigen, daß jeweils + das richtige Vorzeichen ist, sei x aus dem Inneren von $\tau = [\sigma; \hat{v}_i]$. Also $x = (v_0, ..., v_{n+k}) \cdot c$ mit $c_j > 0$ für $j \neq i$; $c_i = 0$ und $\sum_{j=0} c_j = 1$. Weiterhin sei $c = (c_0, \bar{c})$ gesetzt, so daß $x = v_0 + T \cdot \bar{c}$ mit der Matrix T aus Lemma 3.6 $(n + k = m)$.
Betrachte nun $\tilde{x} = x - \epsilon w_i = v_0 + T \cdot (\bar{c} - \epsilon \cdot T^{-1} \cdot w_i) =: v_0 + T \cdot \tilde{c} \overset{!}{\in} \sigma$ für ϵ hinreichend klein.
Es ist zu zeigen

$$\tilde{c}_j \geq 0 \quad (j \geq 1) \tag{A.1}$$

$$\sum_{j=1} \tilde{c}_j \leq 1 \tag{A.2}$$

1. Fall $i = 0$: Also $c_0 = 0$ und $c_1, ..., c_m > 0$. Da $T^{-1} \cdot e_1 = e_1$ folgt $\tilde{c}_1 = c_1 - \epsilon > 0$ für ϵ hinreichend klein und $\tilde{c}_j = c_j$ sonst. Bedingung (A.2) ist erfüllt, weil $\sum_{j=1} \tilde{c}_j = \sum_{j=1} c_j - \epsilon < 1$.

2. Fall $i = m$: Also $c_0, .., c_{m-1} > 0$ und $c_m = 0$. Da hier $T^{-1} \cdot (-e_m) = e_{m-1} - e_m$ ist, folgt für $j < m - 1$: $\tilde{c}_j = c_j$ sowie $\tilde{c}_{m-1} = c_{m-1} - \epsilon$ und $\tilde{c}_m = c_m + \epsilon = \epsilon$. Bedingung (A.2) ist also für jedes ϵ und Bed. (A.1) für hinreichend kleines ϵ erfüllt.

3. Fall $0 < i < m$: Also $c_i = 0$. Da $T^{-1} \cdot (e_{i+1} - e_i) = e_{i+1} - 2e_i + e_{i-1}$ (mit $e_0 = 0$) ist, folgt

$\tilde{c}_j = c_j$ $(|j-i|>1)$ und $\tilde{c}_{i+1} = c_{i+1} - \epsilon$, $\tilde{c}_i = 2\epsilon$ und für $i > 1$ $\tilde{c}_{i-1} = c_{i-1} - \epsilon$.
Für hinreichend kleines ϵ ist (A.1) offenbar erfüllt.
Da $c_0 > 0$ ist, gilt $\sum_{j=1} c_j < 1$, so daß $\sum_{j=1} \tilde{c}_j = \sum_{j=1} c_j - \epsilon + 2\epsilon \leq 1$ für ϵ hinreichend klein.

Gesucht sind jetzt die aNev $u_0, .., u_m$ auf $a([0, id])$ mit $a(x) = W \cdot x + w$. Sei $span\{q_1, ..., q_{m-1}\}$ der Tangentialraum von $\tau = [[0, id]; \hat{v}_i]$, dann gilt $w_i \perp q_1, .., q_{m-1}$ und es muß für u_i gelten $u_i \perp W \cdot q_1, .., W \cdot q_{m-1}$. Mit $(W \cdot q_l)^T \cdot u_i = q_l^T \cdot W^T \cdot q_i$ folgt daraus, daß $u_i = \pm v_i$ wobei $v_i = W^{-T} \cdot w_i / \|W^{-T} \cdot w_i\|$ sein muß.
Wieder ist $+$ das richtige Vorzeichen, denn für $t > 0$ klein, ist

$$b(a(\tau)) + tv_i = a(b(\tau)) + tv_i = W \cdot b(\tau) + w + tv_i = W \cdot (b(\tau) + tW^{-1} \cdot v_i) + w \notin a(\sigma) \Leftrightarrow$$

$$b(\tau) + tW^{-1} \cdot v_i \notin \sigma \quad \Leftrightarrow \quad 0 < tw_i^T \cdot W^{-1} \cdot v_i = tw_i^T \cdot W^{-1} \cdot W^{-T} \cdot w_i = t\|W^{-T} \cdot w_i\|^2$$

Sollen die aNev von $a([v_0, \pi])$ berechnet werden, so nutzt man $[v_0, \pi] = P \cdot [0, id] + v_0$, wobei P die Permutationsmatrix zu π ist.

A.2 Beweis für Lemma 3.10

Um die Rechnungen übersichtlicher zu gestalten, werden die Abkürzungen
$H_i := H(v_i)$ und $Hr := H(r(v_j))$ sowie $m := n + k$ benützt.
Weiterhin seien P, P^* die Permutationsmatrizen für π, π^*.
Aus der Darstellung für $\mathcal{D}H_\sigma$ in Lemma 3.6 folgt

$$\mathcal{D}H_\sigma = (H_1 - H_0, ..., H_m - H_{m-1}) \cdot P^T \tag{A.3}$$

Sei zunächst $j = 0$. Für die Rechnungen benötigt man die Permutation sl und deren Matrix SL. Dabei ist sl gegeben durch

$$sl(i) = \begin{cases} i+1 & \text{für} \quad i < m \\ 1 & \text{für} \quad i = m \end{cases}$$

Die Inverse von sl bzw. SL sei mit sr bzw. SR bezeichnet. Offensichtlich gilt $SL^{-1} = SR = SL^T$.
Mit dieser Definition folgt aus Lemma 2.3 $P^* = P \cdot SL$.
Außerdem gilt mit (A.3)

$$\mathcal{D}H_\sigma \cdot P \cdot SL = (H_2 - H_1, ..., H_m - H_{m-1}, H_1 - H_0)$$

Die Formel (A.3) ausgewertet für $\mathcal{D}H_{\sigma^*}$ ergibt

$$\begin{aligned} \mathcal{D}H_{\sigma^*} &= \left(\mathcal{D}H_\sigma \cdot P \cdot SL + (Hr - H_m - H_1 + H_0) \cdot e_m^T\right) \cdot SL^T \cdot P^T \\ &= \mathcal{D}H_\sigma + (Hr - H_m - H_1 + H_0) \cdot e_m^T \cdot SL^T \cdot P^T \end{aligned}$$

Da $e_m^T \cdot SL^T \cdot P^T = e_1^T \cdot P^T = e_{\pi(1)}$ ist, ist die Behauptung bewiesen.

Der Fall $0 < j < m$ wird ähnlich bewiesen, allerdings wird hier die Permutation ex benötigt

$$ex(i) = \begin{cases} j+1 & \text{für} \quad i = j \\ j & \text{für} \quad i = j+1 \\ i & \text{sonst} \end{cases}$$

Hier gilt nun $EX^{-1} = EX = EX^T$ und $P^* = P \cdot EX$.
Aus (A.3) folgt

$$\mathcal{D}H_\sigma \cdot P \cdot EX = (H_2 - H_1, .., H_{j+1} - H_j, H_j - H_{j-1}, .., H_m - H_{m-1})$$

Somit ergibt sich

$$\begin{aligned} \mathcal{D}H_{\sigma^*} &= \left(\mathcal{D}H_\sigma \cdot P \cdot SL + (Hr - H_{j-1} - H_{j+1} + H_j) \cdot (e_j - e_{j+1})^T\right) \cdot EX^T \cdot P^T \\ &= \mathcal{D}H_\sigma + (Hr - H_{j-1} - H_{j+1} + H_j) \cdot (e_j - e_{j+1})^T \cdot EX^T \cdot P^T \end{aligned}$$

Da
$(e_j - e_{j+1})^T \cdot EX^T \cdot P^T = (e_{j+1} - e_j)^T \cdot P^T = e^T_{\pi(j+1)} - e^T_{\pi(j)}$
gilt, ist die Behauptung bewiesen.

Im letzten Fall $j = m$ ergibt sich $P^* = P \cdot SR$. Also

$$\mathcal{D}H_\sigma \cdot P \cdot SR = (H_m - H_{m-1}, H_1 - H_0, ..., H_{m-1} - H_{m-2})$$

Die Formel (A.3) ausgewertet für $\mathcal{D}H_{\sigma^*}$ ergibt mit den Regeln des Pivotschrittes

$$\begin{aligned} \mathcal{D}H_{\sigma^*} &= \left(\mathcal{D}H_\sigma \cdot P \cdot SL + (H_0 - Hr - H_m - H_1 + H_{m-1}) \cdot e_1^T\right) \cdot SR^T \cdot P^T \\ &= \mathcal{D}H_\sigma + (H_0 - Hr - H_m - H_1 + H_{m-1}) \cdot e_1^T \cdot SR^T \cdot P^T \end{aligned}$$

Da $e_1^T \cdot SL \cdot P^T = e_m^T \cdot P^T = e_{\pi(m)}$ ist, ist die Behauptung bewiesen.

Die Aussagen bezüglich der ausgetauschten Spalten folgen aus

$$\begin{aligned} \mathcal{D}H_\sigma \cdot e_{\pi(j)} &= (H_1 - H_0, ..., H_{n+k} - H_{n+k-1}) \cdot P^T \cdot e_{\pi(j)} \\ &= (H_j - H_{j-1}) \end{aligned}$$

A.3 Beweis für Lemma 4.7

Es gilt die Identität

$$L(\sigma) = \begin{pmatrix} 1 & \cdots & 1 \\ H(v_0) & \cdots & H(v_{n+k}) \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{1}^T_{n+k} \\ \mathbf{0}_n & \mathcal{D}H_\sigma \cdot \mathcal{V} \end{pmatrix} + \begin{pmatrix} 0 \\ H(v_0) \end{pmatrix} \cdot \mathbf{1}^T_{n+k+1}$$

Damit ist

$$\tilde{L}(\tau) = \begin{pmatrix} \mathbf{1}^T_{n+k+1} \\ (0, \mathcal{D}H_\sigma) \cdot \tilde{V} + H(v_0) \cdot \mathbf{1}^T_{n+k+1} \\ e^T_{i_1+1} \\ \vdots \\ e^T_{i_k+1} \end{pmatrix} = \begin{pmatrix} 1, c^T \\ (0, \mathcal{D}H_\sigma) + H(v_0) \cdot (1, c^T) \\ e^T_{i_1+1} \cdot \tilde{V}^{-1} \\ \vdots \\ e^T_{i_k+1} \cdot \tilde{V}^{-1} \end{pmatrix} \cdot \tilde{V}$$

Da für $i_l > 0$: $e^T_{i_l+1} \cdot \tilde{V} = (0, u^T_{i_l})$ gilt, und für $i_l = 0$: $e_1^T \cdot \tilde{V} = (1, \mathbf{0}^T_{n+k})$, sowie $(1, c^T) \cdot \tilde{M}^{-1} = e_1^T$, läßt sich also $\tilde{L}$ darstellen als

$$\tilde{L}(\tau) = J \cdot \tilde{M} \cdot \tilde{V} := \left[I_{n+k+1} + \begin{pmatrix} 0 \\ H(v_0) \\ \mathbf{0}_k \end{pmatrix} \cdot e_1^T \right] \cdot \tilde{M} \cdot \begin{pmatrix} 1 & \mathbf{0}^T_{n+k} \\ \mathbf{0}_{n+k} & \mathcal{V} \end{pmatrix}$$

Literaturverzeichnis

[1] O. Forster, *Analysis, Band 2*
Verlag Vieweg 1991

[2] O. Forster, *Analysis, Band 3*
Verlag Vieweg 1990

[3] E. Sperner, *Neuer Beweis über die Invarianz der Dimensionszahl ...*
Abh. Math. Sem. Univ. Hamburg 6

[4] E.L. Allgower, K. Georg , *Generation of triangulations by reflections*
Utilitas Math., 16 (1979)

[5] E.L. Allgower, K. Georg , *Simplicial and continuation methods ...*
SIAM Rev., 22 (1980)

[6] E.L. Allgower, P. Schmidt *Piecewise-linear approximation ...*
SIAM J. num. Anal., 22 (1985)

[7] E.L. Allgower, S. Gnutzmann *Piecewiese linear approximation ...*
SIAM J. num. Anal., 24 (1987)

[8] S. Gnutzmann *Stückweise lineare Approximation implizit definierter*
Ph.D thesis, University of Hamburg

[9] M.J. Todd *Exploiting structure in pl homotopy algorithms for ...*
Math. programming 18, (1980)

[10] M.J. Todd *PL Homotopy algorithms for sparse systems of nonlinear ...*
SIAM J. Control and Optimization, 21 (1983)

[11] E.L. Allgower, K. Georg *Numerical Continuation Methods*
Springer Verlag (1990)

[12] M.J. Todd *On triangulations for computing fixed points*
Math. programming, 10 (1976)

[13] H.B. Keller *Lect. on num. meth. in bifurcation problems*
TATA Institute, Springer Verlag (1990)

[14] W.C. Rheinboldt *solution fields of nonlinear equations and ...*
SIAM J. num. Anal., 17 (1980)

[15] W.C. Rheinboldt *solution fields of nonl. equations and cont. methods*
SIAM J. num. Anal., 17 (1980)

[16] D. Saupe *On accelerating PL cont. methods by predictor-corrector methods*
Math. Programming 23, (1982)

[17] Alexander *Average intersection and pivoting densities* SIAM J. num. Anal., 24 (1987)

[18] A. Lippke *Fortsetzungsmethoden und schnelle elliptische Löser ...* Dissertation an der TU Berlin 1985

[19] T. Kato *Pertubation theory for linear Operators*
Springer Verlag (1976)

[20] E. Anderson,Z. Bai, ... *LAPACK User's Guide*
SIAM ,University City Science Center Philadelphia

[21] K. Georg,R. Hettich, *On the Numerical Stability of Simplex-Algorithms*
Optimization, 18 (1987)

[22] C. Großmann, J. Terno, *Numerik der Optimierung*
Teubner Verlag (1993)

[23] W. Hackbusch, *Iterative Lösung großer schwachbesetzter LGS*
Teubner Verlag (1993)

[24] H. Buss, *Diplomarbeit*
TU-Berlin (1996)