

### Agenda

1. Rückblick Theorie „fehlerbasierter Abstieg“
2. Konstruktion des globalen Newton Algorithmus NLEQ-ERR mittels adaptiver „trust region“ Strategie
3. Zusammenhang zwischen der gedämpften Newton Iteration und dem Newtonpfad
4. Anhang

1.

Im letzten Vortrag haben wir uns mit der Konstruktion einer affin kovarianten Variante des gedämpften Newtonverfahrens beschäftigt.

- $\|F'(x)^{-1}(F'(\bar{x}) - F'(x))(\bar{x} - x)\| \leq \omega \|\bar{x} - x\|^2, \quad x, \bar{x} \in D$
- $F'(x^k) \Delta x^k = -F(x^k), \quad x^{k+1} = x^k + \lambda_k \Delta x^k, \quad \lambda_k \in ]0,1], \quad k=0,1, \dots$

Im Verlauf des Vortrags sind wir im Rahmen der Globalisierung von Theorem 3.12 (siehe Anhang) auf einen Schwachpunkt von gedämpften Newtonverfahren, die mit einem Test auf kontrahierendes Residuum

$$T(x^{k+1} | I) \leq T(x^k | I) \quad \Rightarrow \quad \|F(x^{k+1})\| < \|F(x^k)\|$$

ausgestattet sind gestoßen:

Bei schlecht konditionierten Jacobi Matrizen verkürzt sich die optimale Schrittweite

$$\bar{\lambda}_k(I) = (h_k \text{cond}_2(F'(x^k)))^{-1}$$

unter Umständen trotz globaler Konvergenzresultate so weit, dass es zu einem Stillstand des Verfahrens kommen kann, d.h.

$$x^{k+1} \approx x^k.$$

⇒

„Schwach nichtlineare“ Probleme, d.h. solche bei denen man in „guter“ Entfernung vom Lösungspunkt ist (im Sinne von lokalen Konvergenztheoremen) können wie „hochgradig nichtlineare“ Probleme erscheinen.

Aus diesen Überlegungen heraus haben wir die natürliche Niveaufunktion

$$T(x^k | A_k) \stackrel{\text{def}}{=} \frac{1}{2} \|A_k F(x^k)\|_2^2, \quad A_k \stackrel{\text{def}}{=} F'(x^k)^{-1}$$

und den assoziierten natürlichen Monotonietest

$$T(x^{k+1} | F'(x^k)^{-1}) \leq T(x^k | F'(x^k)^{-1}) \quad \Rightarrow \quad \|\bar{\Delta} x^{k+1}\| < \|\Delta x^k\|$$

eingeführt.

An dieser Stelle tritt das vereinfachte Newtonverfahren als theoretische Größe auf, d.h.

$$F'(x^k) \bar{\Delta} x^{k+1} = -F(x^{k+1})$$

Aus der Einführung der natürlichen Niveaufunktion ergaben sich einige interessante Eigenschaften:

- Extremaleigenschaft für Reduktionsfaktoren  $t_k$  und Schrittweiten  $\bar{\lambda}_k$  :

$$\begin{aligned} t_k(\lambda|A_k) &\leq t_k(\lambda|A), \\ \bar{\lambda}_k(A_k) &\geq \bar{\lambda}_k(A) \end{aligned}$$

- Verschmelzungseigenschaft:

$$h_k \leq 1 \Rightarrow \bar{\lambda}_k(F'(x^k)^{-1}) = 1$$

$\Rightarrow$  quadratische Konvergenz wird asymptotisch erreicht

- Asymptotische Abstandsfunktion:

$$\text{Für } F \in C^2(D) \text{ gilt } T(x|F'(x^*)^{-1}) = \frac{1}{2}\|x - x^*\|^2 + O(\|x - x^*\|^3)$$

$\Rightarrow$  Für  $x^k \rightarrow x^*$  geht das natürliche Monotoniekriterium in ein Abstandskriterium der Form  $\|x^{k+1} - x^*\| \leq \|x^k - x^*\|$  über. (D.h.: Globale geht in lokale Form der Lösungsannäherung über.)

Das wesentliche Ergebnis zur Konstruktion einer adaptiven „trust region“ Strategie, insbesondere zur Wahl von theoretisch optimalen Dämpfungsfaktoren  $\bar{\lambda}_k$  ist Korollar 3.15 (siehe Anhang). Das Korollar erlaubt uns die Substitution der globalen, affin kovarianten Lipschitzkonstante  $\omega$  durch seine lokale Variante  $\omega(x^k) = \omega_k$ , definiert durch

$$\|F'(x^k)^{-1}(F'(x) - F'(x^k))(x - x^k)\| \leq \omega_k \|x - x^k\|^2, \quad \text{für } x, x^k \in D_0 \quad (3.42).$$

Dies führt uns zu folgender theoretisch optimalen Dämpfungsstrategie für das exakte Newtonverfahren:

$$x^{k+1} = x^k + \bar{\lambda}_k \Delta x^k, \quad \bar{\lambda}_k \stackrel{\text{def}}{=} \min \left\{ 1, \frac{1}{h_k} \right\}, \quad h_k = \omega_k \|\Delta x^k\| \quad (3.43).$$

Es bleibt anzumerken, dass für diese Strategie kein globales Konvergenzresultat vorliegt!

2.

Problem: Kantorovichgrößen  $h_k$  sind über die nicht berechenbaren Lipschitzkonstanten  $\omega_k$  definiert.

Lösung: Einführung von berechenbaren Größen  $[\omega_k]$  und  $[h_k] = [\omega_k] \|\Delta x^k\|$ .

Dadurch können wir die affin kovariante Struktur erhalten. Aus Korollar 3.15 (siehe Anhang) folgt:

$$\left. \begin{aligned} [\omega_k] \leq \omega_k \leq \omega_k(z), \\ [h_k] \leq h_k \leq h_k(z). \end{aligned} \right\} (3.44)$$

Die Effizienz einer solchen Strategie hängt im Wesentlichen von der benötigten Genauigkeit der berechenbaren Größe  $[h_k] = [\omega_k] \|\Delta x^k\|$  ab. Allerdings lässt sich zeigen, dass es reicht wenn  $h_k$  und  $[h_k]$  im führenden Bit übereinstimmen, d.h.:

Lemma 3.16:

Es sei das gedämpfte Newtonverfahren mit Dämpfungsstrategie (3.43) und berechenbarer Größe  $[h_k]$  realisiert. Für die Genauigkeit der berechenbaren Größe gelte

$$0 \leq h_k - [h_k] < \sigma \max \{1, [h_k]\}, \quad \sigma < 1.$$

Dann ergibt der natürliche Monotonietest

$$\|\bar{\Delta} x^{k+1}\| \leq (1 - \frac{1}{2}(1 - \sigma)\lambda) \|\Delta x^k\|.$$

□

Aus (3.43) und (3.44) erhalten wir:

$$\bar{\lambda}_k \leq [\bar{\lambda}_k] \stackrel{\text{def}}{=} \min \left\{ 1, \frac{1}{[h_k]} \right\}$$

Das heißt  $[\bar{\lambda}_k]$  könnte zu groß sein. Eine zu implementierende Strategie zur Wahl von optimalen Dämpfungsfaktoren muss diesem Umstand Rechnung tragen. Eine Aufteilung in Prädiktor- und Korrektorstrategie erscheint angebracht.

Die Korrektorstrategie:

Stellt man sich das gedämpfte Newtonverfahren als Fortsetzung der Tangentenrichtung entlang des Newtonpfads vor ergibt sich folgende obere Schranke:

$$\|\bar{\Delta} x^{k+1} - (1 - \lambda)\Delta x^k\| \leq \frac{1}{2} \lambda^2 \omega_k \|\Delta x^k\|^2 = \frac{1}{2} \lambda^2 h_k \|\Delta x^k\|$$

$$\Rightarrow [h_k] = [\omega_k] \|\Delta x^k\| \stackrel{\text{def}}{=} \frac{2 \|\bar{\Delta} x^{k+1} - (1 - \lambda)\Delta x^k\|}{\lambda^2 \|\Delta x^k\|} \leq h_k$$

$$\Rightarrow \lambda_k^{j+1} \stackrel{\text{def}}{=} \min \left\{ \frac{1}{2} \lambda_k^j, \frac{1}{[h_k^j]} \right\}, \quad j = 0, 1, \dots$$

Die Prädiktorstrategie:

Wir benötigen nun eine effiziente Initialschätzung für  $\lambda_k^0$ . Dazu müssen wir allerdings ein wenig von der Definition der Lipschitzkonstante  $\omega_k$  abweichen, d.h.:

$$\|F'(x^k)^{-1}(F'(x) - F'(x^k))v\| \leq \bar{\omega}_k \|x - x^k\| \|v\|.$$

Mit dieser modifizierten Lipschitzbedingung erhalten wir auf folgende Weise eine geeignete Schranke:

$$\begin{aligned} & \|\bar{\Delta} x^k - \Delta x^k\| \\ &= \|-F'(x^{k-1})^{-1}F(x^k) + F'(x^k)^{-1}F(x^k)\| \\ &= \|F'(x^k)^{-1}F'(x^k)(-F'(x^{k-1})^{-1}F(x^k)) + F'(x^k)^{-1}(-F'(x^{k-1})(-F'(x^{k-1})^{-1}))F(x^k)\| \\ &= \|F'(x^k)^{-1}(F'(x^k) - F'(x^{k-1}))(-F'(x^{k-1})^{-1}F(x^k))\| \\ &= \|F'(x^k)^{-1}(F'(x^k) - F'(x^{k-1}))\bar{\Delta} x^k\| \leq \bar{\omega}_k \|x^k - x^{k-1}\| \|\bar{\Delta} x^k\| = \bar{\omega}_k \lambda_{k-1} \|\Delta x^{k-1}\| \|\bar{\Delta} x^k\|. \end{aligned}$$

$$\Rightarrow [\bar{h}_k] \stackrel{\text{def}}{=} [\bar{\omega}_k] \|\Delta x^k\| \stackrel{\text{def}}{=} \frac{\|\bar{\Delta} x^k - \Delta x^k\| \|\Delta x^k\|}{\lambda_{k-1} \|\Delta x^{k-1}\| \|\bar{\Delta} x^k\|} \leq \bar{\omega}_k \|\Delta x^k\|$$

$$\Rightarrow \lambda_k^0 \stackrel{\text{def}}{=} \min \left\{ 1, \frac{1}{[\bar{h}_k]} \right\}, \quad k > 0.$$

Es muss allerdings darauf geachtet werden, dass die Richtung  $\bar{\Delta} x^k$  „nicht zu weit“ von der Richtung  $\Delta x^{k-1}$  entfernt ist.

Desweiteren verbleibt die Wahl des Startwerts  $\lambda_0^0$ . Stuft der Benutzer das Problem als „schwach nichtlinear“ ein, dann sollte  $\lambda_0^0=1$  und andernfalls  $\lambda_0^0=\lambda_{\min} \ll 1$  gewählt werden.

Quasi-Newton Zwischenschritte:

Wenn  $\lambda_k=1$  und der natürliche Monotonietest

$$\Theta_k = \frac{\|\bar{\Delta} x^{k+1}\|}{\|\Delta x^k\|} < \frac{1}{2},$$

liefert, dann kann auf den fehlerbasierten Quasi-Newton Algorithmus QNERR umgeschaltet werden.

Die beschriebene adaptive „trust region“ Strategie führt zum globalen Newton Algorithmus NLEQ-ERR (siehe Anhang), der eine leichte Modifikation des Algorithmus NLEQ1 darstellt.

3.

Wie bereits erwähnt fehlt dem hier besprochenen affin kovarianten Ansatz zur Entwicklung eines globalen fehlerbasierten Newtonverfahrens ein nachweisliches Konvergenzresultat.

Die Konstruktion des Algorithmus erfolgt anhand einer zugrundeliegenden geometrischen Idee. Man versucht den Newtonpfad, der in Theorem 3.6 (siehe Anhang) definiert ist zu imitieren.

Interessante Aussagen bezüglich des Newtonpfads sind:

- Alle (!) Niveaufunktionen  $T(x|A)$  nehmen entlang dem Newtonpfad  $\bar{x}$  ab.
- Der Newtonpfad  $\bar{x}$  endet entweder im Lösungspunkt, in einem kritischen Punkt mit singulärer Jacobi Matrix oder am Rand des Gebiets auf dem die Funktion  $F$  definiert ist.

Man könnte also auch versuchen den Newtonpfad numerisch zu berechnen und so die Lösung ausfindig machen. Andererseits liefert uns Theorem 3.6 in Aussage (3.25) (siehe Anhang) folgende lokale Information:

Die Tangentenrichtung des Newtonpfads entspricht der Newtonrichtung, d.h.:

Sogar „weit weg“ von der Lösung ist die Newtonrichtung eine beachtenswerte Größe. Man muss sich allerdings um die Länge dieser Richtung Gedanken machen, wenn man „hochgradig nichtlineare“ Probleme betrachtet.

Diese Überlegungen führen geradewegs zu dem Konzept des gedämpften Newtonpfads. Wobei die Dämpfungsfaktoren uns sagen wie weit wir der Linearisierung unseres Problems vertrauen können.

Veranschaulicht wird dies in Fig. 3.10. (siehe Anhang).

Beispielhaft wird die Lösung folgender Gleichungen gesucht:

$$\begin{aligned} \exp(x^2 + y^2) - 3 &= 0, \\ x + y - \sin(3(x + y)) &= 0. \end{aligned}$$

Auf der Abbildung sieht man von den kritischen Interfaces  $x=y$  und

$$y = -x \pm \arccos\left(\frac{1}{3}\right) \pm \frac{2}{3} \pi j, \quad j = 1, 2, 3, \dots \text{ fünf Stück und dazu sechs der Lösungspunkte.}$$

Abgebildet wird ein quadratischer Teilausschnitt des Gebiets auf dem die Funktion definiert ist. Links oben sieht man die Newtonpfade und rechts unten die Newtoniterierten. Klar wird, dass unsere Konstruktion den Newtonpfad nachahmt und wir in diesem Sinne guter Hoffnung sein können, dass mit unserem Vorgehen tatsächlich Lösungspunkte gefunden werden.

Eine weitere nützliche Eigenschaft der Folge von Newtoniterierten, die durch den fehlerbasierten, adaptiven „trust region“ Ansatz erzeugt werden ist in Fig. 3.11. (siehe Anhang) illustriert. Dort sieht man klar, dass die Folge von Newtoniterierten im Vergleich zu anderen Verfahren ihr „attraction basin“ nicht verlässt. Insbesondere bedeutet dies, dass ein unerwünschter Übergang über die kritischen Interfaces hinweg nicht stattfindet.

4.

Theorem 3.12:

Sei  $F \in C^1(D)$  mit  $D \subset \mathbb{R}^n$  konvex und  $F'(x)$  invertierbar  
 $\forall x \in D$ . Für eine gegebene Iterierte  $x^k \in D$  sei  $G(x^k | A)$   
 $\subset D$  für beliebiges  $A \in GL(n)$ .

Es gelte die affin kovariante Lipschitzbedingung für  $x, y \in D$

$$\|F'(x)^{-1}(F'(y) - F'(x))(y - x)\| \leq \omega \|y - x\|^2.$$

Dann erhält man mit der Notation

$$h_k := \|\Delta x^k\| \omega, \quad \bar{h}_k := h_k \operatorname{cond}(AF'(x^k))$$

für  $\lambda \in [0, \min(1, \frac{2}{\bar{h}_k})]$ :

$$\|AF(x^k + \lambda \Delta x^k)\| \leq t_k(\lambda | A) \|AF(x^k)\|, \quad (3.37)$$

wobei

$$t_k(\lambda | A) := 1 - \lambda + \frac{1}{2} \lambda^2 \bar{h}_k. \quad (3.38)$$

Die optimale Wahl der Dämpfungsfaktor ausgedrückt in diesen  
 lokalen Größen ist

$$\bar{\lambda}_k(A) := \min\left(1, \frac{1}{\bar{h}_k}\right)$$

Korollar 3.15:

Sei  $F: \mathcal{D} \rightarrow \mathbb{R}^n$  eine stetig differenzierbare Abbildung mit  $\mathcal{D} \subset \mathbb{R}^n$  offen und konvex. Seien weiterhin  $x^0, z \in \mathcal{D}$ , die Jacobimatrix  $F'(x)$  invertierbar  $\forall x \in \mathcal{D}$  und sei die wegzusammenhängende Komponente  $\mathcal{D}_0$  von  $G(x^0 | F'(z)^{-1})$  in  $x_0$  kompakt und enthalten in  $\mathcal{D}$ . Gilt zusätzlich folgende affin kovariante Lipschitzbedingung

$$\|F'(z)^{-1} (F'(y) - F'(x)) (y-x)\| \leq \omega(z) \|y-x\|^2 \quad \text{für } x, y, z \in \mathcal{D}_0.$$

Dann kann gezeigt werden, dass folgende lokale Reduktionseigenschaft der Niveaufunktion

$$T(x^k + \lambda \Delta x^k | F'(z)^{-1}) \leq (1 - \lambda + \frac{1}{2} \lambda^2 h_k(z))^2 T(x^k | F'(z)^{-1})$$

mit  $h_k(z) := \|\Delta x^k\| \omega(z) \|F'(x^k)^{-1} F'(z)\|$  vorliegt und ein lokal optimaler Dämpfungsfaktor

$$\bar{\lambda}_k(z) := \min\left(1, \frac{1}{h_k(z)}\right)$$

gilt.

Ist die benutzte Matrixnorm zudem submultiplikativ, dann erhält man für die best möglichen Schätzungen  $\omega(z)$  die Extremaleigenschaften

$$\|F'(x^k)^{-1} F'(z)\| \omega(z) \geq \omega(x^k)$$

und  $h_k(x^k) \leq h_k(z)$ ,  $\bar{\lambda}_k \geq \bar{\lambda}_k(z)$ ,  $z \in \mathcal{D}_0$ .

Algorithmus < NLEQ-ERR >

- Setze Fehlergenauigkeit  $\varepsilon$
- Wähle Startiterierte  $x^0$ , Berechne  $F(x^0)$
- Setze Dämpfungsfaktor  $\lambda_0 := 1$  oder  $\lambda_0 \ll 1$ .

Für  $k = 0, 1, \dots$  do:

1. Schritt  $k$ : Berechne Jacobi  $F'(x^k)$ . Löse LGS

$$F'(x^k) \Delta x^k = -F(x^k)$$

Konvergenztest: If  $\|\Delta x^k\| \leq \varepsilon$ : STOP. Lösung  $x^* := x^k + \Delta x^k$ .

Für  $k > 0$ : Berechne Prädiktor

$$\lambda_k := \min(1, \mu_k), \quad \mu_k := \frac{\|\Delta x^{k-1}\| \cdot \|\bar{\Delta x}^k\|}{\|\bar{\Delta x}^k - \Delta x^k\| \cdot \|\Delta x^k\|} \cdot \lambda_{k-1}$$

Regularitätstest: If  $\lambda_k < \lambda_{\min}$ : STOP. Konvergenzfehler

2. Else: berechne Versuchsiterierte  $x^{k+1} := x^k + \lambda_k \Delta x^k$  und werte  $F(x^{k+1})$  aus. Löse LGS

$$F'(x^k) \bar{\Delta x}^{k+1} = -F(x^{k+1})$$

3. Berechne die Monitorgrößen

$$\theta_k := \frac{\|\bar{\Delta x}^{k+1}\|}{\|\Delta x^k\|}, \quad \mu'_k := \frac{\frac{1}{2} \|\Delta x^k\| \cdot \lambda_k^2}{\|\bar{\Delta x}^{k+1} - (1 - \lambda_k) \Delta x^k\|}$$

If  $\theta_k \geq 1$

then replace  $\lambda_k$  by  $\lambda'_k := \min(\mu'_k, \frac{1}{2} \lambda_k)$ . Go to Regularitätstest.

Else let  $\lambda'_k := \min(1, \mu'_k)$ .

If  $\lambda'_k = \lambda_k = 1$ :

If  $\|\bar{\Delta x}^{k+1}\| \leq \varepsilon$ : STOP Lösung  $x^* := x^{k+1} + \bar{\Delta x}^{k+1}$

If  $\theta_k < \frac{1}{2}$ : SWITCH zu QNERR

Else: If  $\frac{1}{4} \lambda'_k \geq \lambda_k$ : ersetze  $\lambda_k$  durch  $\lambda'_k$  und goto 2.

Else: nimm  $x^{k+1}$  als neue Iterierte an.  
Goto 1 mit  $k \rightarrow k+1$ .



Der Newton Pfad

Theorem 3.6:

Sei  $F \in C^1(D)$ ,  $D \subseteq \mathbb{R}^n$ ,  $F'(x)$  invertierbar  $\forall x \in D$ . Für beliebiges  $\hat{A} \in GL(n)$ , sei die weg-zusammenhängende Komponente von  $G(x^0 | \hat{A})$  in  $x^0$  kompakt und enthalten in  $D$ .

Dann ist die weg-zusammenhängende Komponente von  $\bar{G}(x^0) := \bigcap_{A \in GL(n)} G(x^0 | A)$  ein topologischer Pfad  $\bar{x}: [0, 2] \rightarrow \mathbb{R}^n$ ,

der sogenannte Newtonpfad mit folgenden Eigenschaften

$$F(\bar{x}(\lambda)) = (1-\lambda) F(x^0), \quad (3.22)$$

$$T(\bar{x}(\lambda) | A) = (1-\lambda)^2 T(x^0 | A), \quad (3.23)$$

$$\frac{d\bar{x}}{d\lambda} = -F'(\bar{x})^{-1} F(x^0), \quad (3.24)$$

$$\bar{x}(0) = x^0, \quad \bar{x}(1) = x^*,$$

$$\left. \frac{d\bar{x}}{d\lambda} \right|_{\lambda=0} = -F'(x^0)^{-1} F(x^0) \equiv \Delta x^0, \quad (3.25)$$

wobei  $\Delta x^0$  die gewöhnliche Newtonrichtung ist.

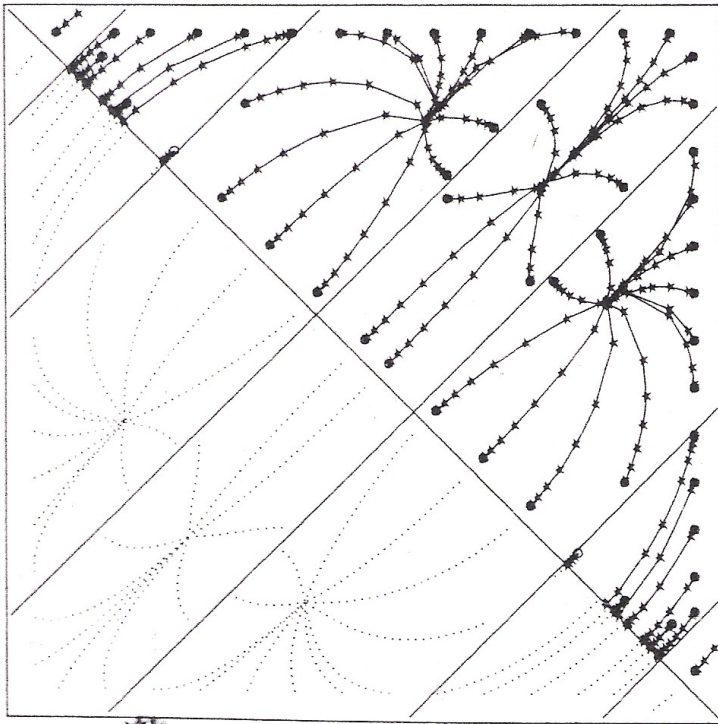


Fig. 3.10. Example 3.2: Newton paths ( $\dots$ ) versus Newton sequences ( $---$ )

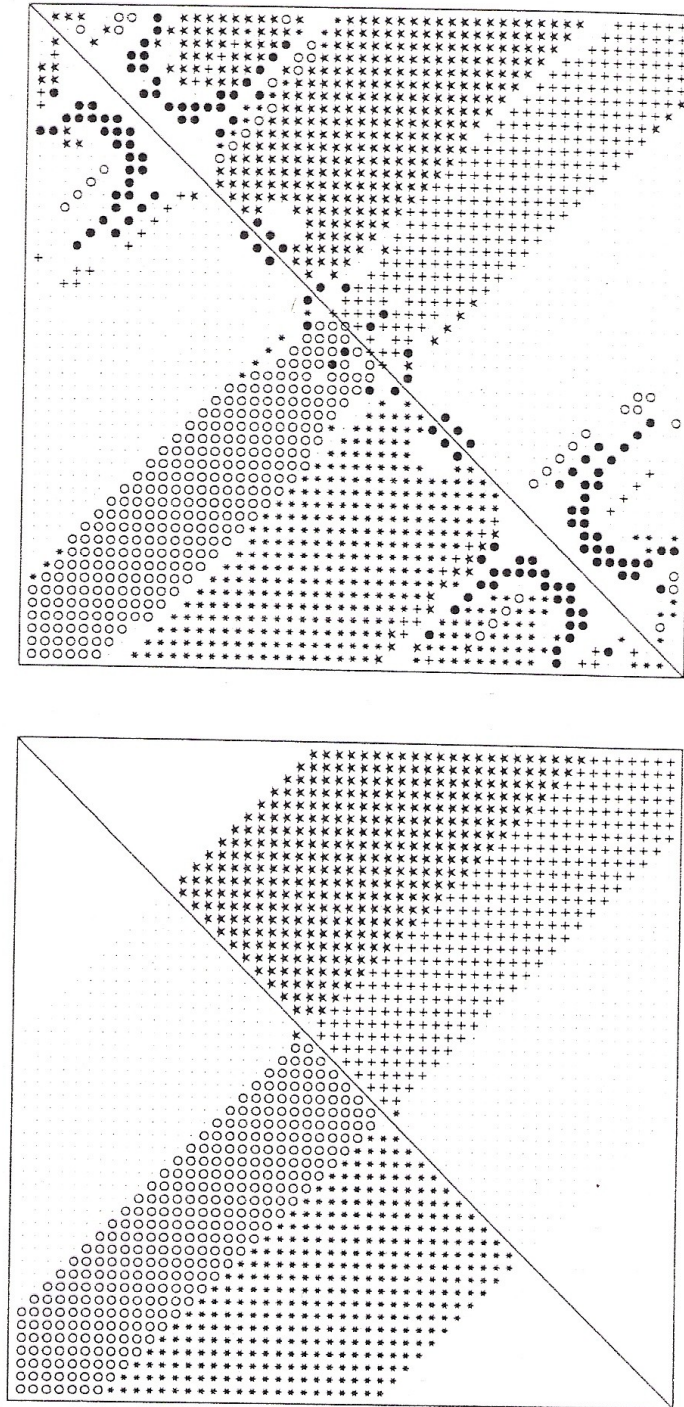


Fig. 3.11. Example 3.2: attraction basins. *Left:* Global Newton method, code NLEQ1 [161]. *Right:* hybrid method, code HYBRJ1 [153]. Outliers are indicated as bullets (●).