Universität Karlsruhe (T.H.)
Mathematisches Institut I
Institutsleiter: Prof. Dr. Michael Plum

Existenz- und Einschließung bei einem einfachen Modell zur Beschreibung von Hängebrücken

Diplomarbeit von
cand. math. tech. Holger Obermaier

betreut von
Prof. Dr. Michael Plum

Hiermit bestätige ich, Holger Obermaier, daß ich die hier vorliegende Diplomarbeit selbstständig ausgearbeitet habe. Alle dabei verwendeten mathematischen Artikel und Bücher sind dem Literaturverzeichnis zu entnehmen.

Karlsruhe, den 5. März 2003 Holger Obermaier

Vorwort:

Während meines Studiums habe ich mich hauptsächlich in der Richtung Analysis vertieft. Dabei mußte ich feststellen, daß nur für eine sehr beschränkte Klasse von Differentialgleichungen, egal ob gewöhnlich oder partielle, expliziete Lösungen existieren. Möchte man für die anderen iregendwelche Aussagen über die Lösung(en) machen, so ist man auf numerische Näherungen angewiesen. Für einen Mathematiker sind aber Näherungsaussagen, ohne den Beweis der Existenz oder Aussagen über die Entfernung zur exakten Lösung sehr unbefriedigend, da bei vorhandensein einer numerischen Lösung, selbst mit sehr kleinem Defekt, keinesfalls sicher ist, daß eine exakte Lösung existiert und wie weit diese Näherungslösung von einer eventuell existierenden exakten Lösung abweicht. Große Probleme hat man auch, wenn man bei einer Differentialgleichung zwei numerische Lösungen erhält, die sehr nah beieinander liegen und man nun wissen möchte, ob tatsächlich zwei verschiedene Lösungen vorliegen oder ob die Numerik für die leicht verschiedenen Lösungen verantwortlich ist. All diese Fragestellungen führten dazu, daß ich mir eine Vorlesung über Computerunterstütze Methoden bei Randwertproblemen bei Herrn Prof. M. Plum anhörte, in der er eben gerade auf die Frage nach dem Beweis der Existenz einer exakten Lösung in einer angebbaren Umgebung einer Näherungslösung für eine große Klasse von Differentialgleichungen einging. Ich war begeistert von den Möglichkeiten und sehr daran interessiert noch mehr zu erfahren und das Gelernte auch einmal praktisch umzusetzen.

Das Resultat dieser Neugierde ist die hier vorliegende Diplomarbeit, in der ich die Methoden zum Beweis der Existenz und Einschließungen von Lösungen bei gewöhnlichen Differentialgleichungen, die M. Plum schon ausgearbeitet hatte, auf Systeme von periodischen Randwertproblemen übertragen konnte (siehe Kapitel 1). Um die Möglichkeiten der Methoden zu demonstrieren wurden sie auf ein einfaches Modell für eine Hängebrücke angewandt (für Ergebnisse siehe Kapitel 3).

Unabhängig davon war ich fasziniert davon physikalische Objekte und ihre Bewegung mit Hilfe von unterschiedlichen Modellen zu beschreiben und deren unterschiedliches Verhalten zu analysieren. Da ich der Ansicht bin daß man sich als Mathematiker nicht scheuen sollte den Blick über den Tellerrand in Richtung Ingeneurwissenschaften zu wagen, sind in dieser Diplomarbeit in Kapitel 2 auch einige dieser einfachen Hängebrückemodelle vorgestellt.

Am Ende dieses Vorwortes bleibt mir nun nur noch die schöne Aufgabe all

jenen zu danken, die es überhaupt erst ermöglicht haben, daß diese Diplomarbeit entstehen konnte. An erster Stelle sind wohl meine Eltern zu nennen, die mich in jeglicher Form bei meiner Arbeit unterstützt haben. Aber auch meine ehemaligen Lehrer, die mein mathematisches Interesse erst geweckt haben sind zu nennen. Dann gebührt natürlich besonderer Dank meinem Betreuer Herrn Plum. Wie heißt es im Englischen so treffend "last but not least" schulde ich auch all jenen Dank, die meine Diplomarbeit Korektur gelesen haben und damit hoffentlich zu einer besseren Rechtschreibung beigetragen haben.

Inhaltsverzeichnis

Kapitel 1

Gesicherte Existenz- und Einschließungsergebnisse für periodische Systeme von Randwertproblemen:

1.1 Einleitung:

Ziel dieses Abschnitts wird es sein ausgehend von dem in [1] und [3] vorgestellten Algorithmus eine Verallgemeinerug, für nichtlineare periodische Systeme von Randwertaufgaben der Form

$$-\ddot{U} + B \cdot \dot{U} + F(t, U) = 0 \quad 0 \leq t \leq k\frac{2\pi}{\mu} \tag{1.1a}$$

mit den Randbedingungen

$$U(0) - U(k\tfrac{2\pi}{\mu}) \;=\; 0 \tag{1.1b}$$

$$\dot{U}(0) - \dot{U}(k\tfrac{2\pi}{\mu}) \;=\; 0 \tag{1.1c}$$

mit $U : [0, k\frac{2\pi}{\mu}] \to \mathbb{R}^n$, $B : [0, k\frac{2\pi}{\mu}] \to \mathbb{R}^{n\times n}$ stetig differenzierbar und $F : \mathbb{R}^{n+1} \to \mathbb{R}^n$ stetig differenzierbar (wobei B und F als bekannt vorausgesetzt seine) vorzustellen, mit deren Hilfe man die Existenz einer C_2-Lösung beweisen kann und eine gesicherte Einschließung für diese erhalten kann.

Existiert eine Lösung der Differentialgleichung, so ist die einzige Voraussetzung für die Durchführbarkeit des Algorithmus, daß man eine Näherungslösung

$$\omega \in R := \left\{ u \in H_2^n(0, k\tfrac{2\pi}{\mu}) : u(0) - u(k\tfrac{2\pi}{\mu}) = 0 \text{ und } \dot{u}(0) - \dot{u}(k\tfrac{2\pi}{\mu}) = 0 \right\}$$

finden kann, die den beiden folgenden Bedingungen genügt:

1. Definiert man den Defekt durch

$$d[\omega] := -\ddot{\omega} + B \cdot \dot{\omega} + F(., \omega)$$

und die L_2-Norm durch $||u||_2 := \sqrt{\int_0^{k\frac{2\pi}{\mu}} u_1^2 + \cdots + u_n^2 \, dt}$, so ist die erste zu erfüllende Bedingung, daß die L_2-Norm des Defektes hinreichend klein ist, mit anderen Worten, daß die Näherungslösung die Differentialgleichung hinreichend genau erfüllt. (Was mit den Worten hinreichend klein gemeint ist, wird bei der detailierten Vorstellung des Algorithmus noch klar spezifiziert).

2. Die zweite Bedingung ist, daß eine untere Grenze für den Abstand zwischen 0 und dem Spektrum des linearen Operators $L^*L : R \to L_2^n(0, k\frac{2\pi}{\mu})$ am Punkt ω, den man aus der Linearisierung des gegebenen Problems erhält, berechenbar sein muß. (Auch dies findet man formalisiert und näher beschrieben bei der Ausführung des Algorithmus).

Das erste Problem, das sich aus diesen zwei Bedingungen ergibt, ist daher natürlich zuallererst die Bestimmung einer Näherungslösung ω mit hinreichender Genauigkeit. Hat man diese nun bestimmt, so ergibt sich sogleich das Problem, eine gesicherte obere Schranke für $||d[\omega]||_2$ und eine gesicherte untere Schranke für den Abstand zwischen 0 und dem Spektrum von L^*L zu bestimmen. Das Interessante an dem Algorithmus, der im nächsten Abschnitt vorgestellt wird, ist die Möglichkeit, diese drei Unterprobleme völlig unabhängig voneinander mit jedem Verfahren, das zur Lösung des jeweiligen Teilproblems führt, ausführen zu können.

Zur Bestimmung der Näherungslösung ω kann es je nach Problem sinnvoll sein, ein Runge-Kutta-Verfahren oder aber ein Newtonverfahren zu benutzen. Für die Anwendung des Algorithmus macht es keinen Unterschied, auf welche Weise die Näherung gewonnnen wurde. Was ebenfalls das Spektrum der verwendbaren Näherungsalgorithmen erweitert, ist die Möglichkeit, die Näherungslösung mit der ganz normalen Zahlenarithmetik des Computers zu bestimmen, ohne auf Rundungsfehler achten zu müssen. Erst bei der Berechnung von Schranken für den Defekt $||d[\omega]||_2$ und des linearen Operators L^*L wird es dann notwendig, auf Intervallarithmetik, die z.B. Programmiersprachen wie Pascal XSC oder Programmbibliotheken wie Profil zur Verfügung stellen, zurückzugreifen. Das einzige, worauf man bei der Bestimmung der Näherungslösung achten muß, ist das Vorliegen der Näherung in einer Form, die geeignet ist, Intervallrechnungen durchführen zu können. (Ein Reihe von Punkten etwa dürfte nur schwer verwendbar sein, wohingegen eine Darstellung der Lösung als Fourierreihe leicht zu Einschließungen führen dürfte).

Es ist noch wichtig zu bemerken, daß das gesamte Vorgehen keine Monotonie- oder Inverspositivitätsannahmen benötigt und daher einen weiten Anwendungsbereich ermöglicht.

1.2 Das Algorithmengerüst:

In diesem Abschnitt wird ein Gerüst vorgegeben, das den Algorithmus zur gesicherten Existenz- und Einschließung für periodische Systeme von Randwertproblemen zwar vollständig, aber auf eine kurze und prägnante Art und Weise, beschreibt. Dieser Abschnitt ist eher als Übersicht, denn als vollständiger Algorithmus zu begreifen. Zu jedem der folgenden Punkte gibt es dann im nächsten Abschnitt eine ausführliche Beschreibung. Im übertragenen Sinne könnte man davon sprechen, daß in diesem Abschnitt das Skelett beschrieben wird, im nächsten Abschnitt dann das Fleisch, das der ganzen Sache Leben gibt.

Der Algorithmus kann in die folgenden Schritte gegliedert werden:

1. Berechne eine Näherungslösung

$$\omega \in R := \left\{ u \in H_2^n(0, k\tfrac{2\pi}{\mu}) : u(0) - u(k\tfrac{2\pi}{\mu}) = 0 \text{ und } \dot{u}(0) - \dot{u}(k\tfrac{2\pi}{\mu}) = 0 \right\} \tag{1.2}$$

   des Randwertproblems (1.1).

2. Berechne eine Konstante $\delta \geq 0$, für die mit der berechneten Näherungslösung ω gilt

$$||d[\omega]||_2 := || - \ddot{\omega} + B \cdot \dot{\omega} + F(., \omega)||_2 \leq \delta. \tag{1.3}$$

3. Berechne eine Konstante σ mit

$$0 < \sigma \leq \lambda \qquad \forall \lambda \text{ Eigenwert von } L^*L \text{ auf } R.$$

4. Berechne eine Konstante $K \geq 0$ mit

$$||u||_\infty \leq K||L[u]||_2 \quad \forall u \in R, \tag{1.4}$$

   wobei mit $L : R \to L_2^n(0, k\frac{2\pi}{\mu})$ der lineare Operator

$$L[u] := -\ddot{u} + B \cdot \dot{u} + C \cdot u \tag{1.5}$$

   bezeichnet sei, den man erhält, wenn man das Randwerproblem (1.1) linearisiert. C ist somit definiert durch $C(t) := \left(\frac{\partial F_i}{\partial y_j}(t, \omega(t))\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ $(0 \leq t \leq k\frac{2\pi}{\mu})$. (Insbesondere verlangt (1.4), daß der lineare Operator L auf R invertierbar ist).

5. Berechne eine nichtfallende Funktion $G : [0, \infty) \to [0, \infty)$, die der Bedingung

$$||F(., \omega + y) - F(., \omega) - C \cdot y||_2 \leq G(||y||_\infty) \quad (\forall y \in \mathbb{R}^n) \tag{1.6}$$

   genügt und für die ebenfalls $G(t) = o(t)$ für $t \to 0$ gilt.

6. Prüfe, ob eine Konstante α gefunden werden kann, die der Bedingung

$$\delta \leq \frac{\alpha}{K} - G(\alpha) \tag{1.7}$$

   genügt. Nach der Forderung an die Funktion G in (1.6) erfordert (1.7) nun in der Tat, daß δ hinreichend klein ist mit anderen Worten, daß die Näherungslösung ω hinreichend genau bestimmt wurde.

Theorem:

Wenn (1.7) erfüllt ist, so existiert eine Lösung $U \in C_2^n[0, k\frac{2\pi}{\mu}]$ des Randwertproblems (1.1), die der folgenden Abschätzung (Einschließung) genügt:

$$\boxed{||U - \omega||_\infty \leq \alpha}.$$

Beweis: Durch einfaches Umformen erhält man die Äquivalenz des Problems (1.1) mit den Randbedingungen zu einem System von Randwertaufgaben für den Fehler $u := U - \omega \in R$ (d.h. für das Randwertproblem für U existiert genau dann eine Lösung, wenn eine Lösung des Randwertproblems für u existiert).

$$\begin{aligned}
-\ddot{U} + B \cdot \dot{U} + F(.,U) &= 0 \\
-(\ddot{u} + \ddot{\omega}) + B \cdot (\dot{u} + \dot{\omega}) + F(.,u+\omega) &= 0 \\
-\ddot{u} + B \cdot \dot{u} + F(.,\omega + u) &= \ddot{\omega} - B \cdot \dot{\omega} \\
-\ddot{u} + B \cdot \dot{u} + C \cdot u + F(.,\omega + u) - F(.,\omega) - C \cdot u &= \ddot{\omega} - B \cdot \dot{\omega} - F(.,\omega) \\
L[u] + f(.,u) &= -d[\omega] \qquad (1.8)
\end{aligned}$$

In Differerentialgleichung (1.8), der Differentialgleichung für den Fehler u, seien $f(.,u) := F(.,\omega + u) - F(.,\omega) - C \cdot u$, $d[\omega] := -\ddot{\omega} + B \cdot \dot{\omega} + F(.,\omega)$ und $C := \left(\frac{\partial F_i}{\partial y_j}(.,\omega)\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ gesetzt. Wegen der Äquivalenz von (1.1) und der Differentialgleichung (1.8) mit den Randbedingugenügt $u(0) = u(k\frac{2\pi}{\mu})$ und $\dot{u}(0) = u(k\frac{2\pi}{\mu})$ genügt es also zu zeigen, daß die Differentialgleichung eine Lösung $u \in R$ besitzt, die die Abschätzung $||u||_\infty \leq \alpha$ erfüllt.

Nach (1.4) gilt: Der inverse Operator $L^{-1} : L_2^n(0, k\frac{2\pi}{\mu}) \to R \subseteq H_2^n(0, k\frac{2\pi}{\mu})$ existiert und ist stetig. Da die Einbettung $H_2^n(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1^n[0, k\frac{2\pi}{\mu}]$ nach dem Einbettungssatz von Sobolev, Kondrachov und Rellich kompakt ist, gilt auch für den Operator $T : C_1^n[0, k\frac{2\pi}{\mu}] \to C_1^n[0, k\frac{2\pi}{\mu}]$, definiert durch

$$Tu := -L^{-1}[d[\omega] + f(.,u)],$$

daß er stetig und kompakt ist.

Die Behauptung folgt nun nach dem Schauderschen Fixpunktsatz, der besagt, daß es mindestens einen Fixpunkt $Tu = u$ in D gibt, wenn man zeigen kann, daß

$$TD \subseteq D$$

gilt, für $D := \{u \in C_1[0, \frac{2\pi}{\mu}] : ||u||_\infty \leq \alpha\}$. Sei nun also $u \in D$

$$\begin{aligned}
\Rightarrow \quad ||Tu||_\infty &\overset{(1.4)}{\leq} K||L[T[u]]||_2 \\
&= K||d[\omega] + f(.,u)||_2 \\
&\overset{\triangle-\text{Ungl}}{\leq} K(||d[\omega]|| + ||f(.,u)||_2) \\
&\overset{(1.3)(1.6)}{\leq} K(\delta + G(||u||_\infty)) \\
&\overset{\text{Monotonie}}{\leq} K(\delta + G(\alpha))
\end{aligned}$$

$$\stackrel{(1.7)}{\leq} \quad \alpha$$

$\Rightarrow TD \subseteq D.$ $\square$

Bemerkung:
Für denjenigen, der mit der Sobolevtheorie nicht so vertraut ist, ist im Abschnitt 4.3.1 "Sobolev Räume" eine kurze Definition der Räume $H_m(\Omega)$ und der Einbettungssatz von Sobolev, Kondrachov und Rellich gegeben.

1.3 Die Ausfüllung des Algorithmengerüsts:

1.3.1 Zu Punkt 1: Bestimmung der Näherungslösung ω:

Wie schon zuvor erwähnt, kann man die Bestimmung der Näherungslösung ω mit jedem beliebigen Algorithmus durchführen, der das Ergebnis in einer Form liefert, die man dann bei der benötigten Intervallrechnung verwenden kann. Hier soll ein Algorithmus vorgestellt werden, der zwei Schritte kombiniert, um die Näherungslösung zu erhalten. Im ersten Schritt wird mit einem Runge-Kutta-Verfahren eine grobe Näherungslösung bestimmt, die dann im zweiten Schritt durch einen Newtonalgorithmus verbessert wird, dem die Näherung aus dem ersten Schritt als Startnäherung dient.

Bestimmung der groben Startnäherung für das Newtonverfahren mittels des Runge-Kutta-Verfahrens:

Hat man die Differentialgleichung in der Form

$$\dot{z} = \tilde{F}(t, z),$$

so liefert das Runge-Kutta-Verfahren, ausgehend von einem Anfangswert (t_0, z^0), Näherungswerte für die Zeitpunkte $t_{i+1} = t_i + h$ $(i = 0, \ldots)$ (mit $h := \frac{2\pi}{\mu m}$) durch das folgende Vorgehen für $i = 0, \ldots$

$$\begin{aligned} k_1 &:= \tilde{F}(t_i, y^i) \\ k_2 &:= \tilde{F}(t_i + h/2, 1/2hk_1) \\ k_3 &:= \tilde{F}(t_i + h/2, 1/2hk_2) \\ k_4 &:= \tilde{F}(t_i + h, hk_3) \\ z^{i+1} &:= z^i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4). \end{aligned}$$

Bemerkungen:

- Das Differentialgleichungssystem

  $$-\ddot{U} + B \cdot \dot{U} + F(t, U) = 0$$

  hat eine Form, die dem Runge-Kutta-Verfahren leider nicht direkt zugänglich ist, da es verlangt, daß die Differentialgleichung in der Form

$\dot{z} = \tilde{F}(t, z)$ geschrieben werden kann. Durch Aufblähen des Differentialgleichungssystems kann man diese Form aber dennoch erreichen, konkret heißt dies, setzt man $z_i := U_i$ für $i = 1, \ldots, n$ und $z_{n+i} := \dot{u}_i$ für $i = 1, \ldots, n$, so erhält man das äquivalente Differentialgleichungssystem

$$\begin{aligned} \dot{z}_i &= z_{n+1} & i &= 1, \ldots, n \\ \dot{z}_{n+i} &= -\sum_{j=1}^{n} b_{i,j} z_{n+j} - F(t, z_1, \ldots, z_n) & i &= 1, \ldots, n \ , \end{aligned}$$

das die verlangte Form aufweist.

- Das Runge-Kutta-Verfahren ist eigentlich zur Lösung von Anfangswertaufgaben gedacht und nicht zur Lösung periodischer Randwertprobleme, wie sie hier eigentlich gelöst werden sollen (siehe (1.1)). Dennoch liefert das Verfahren meist gute Lösungen. Das Vorgehen besteht darin, mit irgendeinem Anfangswert das Iterationsverfahren anlaufen zu lassen und dann zu hoffen, daß dieser Anfangswert irgendwann in einen periodischen Zyklus hineinläuft. In meinem Modellproblem (siehe Kapitel 2) zum Beispiel habe ich das Iterationsverfahren über die Zeit von 2000 Perioden der anregenden Kraft laufen lassen und dann meist ein periodische Näherungslösung vorgefunden. Dieses Vorgehen ist natürlich reine Heuristik und man kann nicht dafür garantieren, daß man auf diesem Weg zu Lösungen kommt. Instabile Lösungen können mit diesem Verfahren leider auch nicht gefunden werden, denn schon die kleinsten durch arithmetische Rundungen verursachten Fehler würden zum Verlassen dieses Lösungszweiges führen. Wenn man aber Parameterstudien betreibt, so ist dieses Verfahren recht gut geeignet aus bereits berechneten Lösungen weitere Lösungen für andere Parameter zu gewinnen. Variiert man den Parameter nur leicht, so ändert sich der Funktionswert der periodischen Lösung zum Zeitpunkt $t = 0$ meist nur sehr wenig (falls stetige Abhängigkeit von den Parametern vorliegt). Man kann also den Funktionswert einer bereits berechneten periodischen Lösung zum Zeitpunkt $t = 0$ als guten Anfangswert für das Runge-Kutta-Verfahren mit einem nur leicht variierten Parameter verwenden.

- Das Runge-Kutta Verfahren liefert nur eine Folge von Punkten. Wie läßt sich nun mit diesen feststellen, ob eine periodische Lösung vorliegt? Hat die Anregung wie beim Modellbeispiel (siehe Kapitel 2) z.B. eine Periode $\frac{2\pi}{\mu}$ und möchte man nach 2000 Periodendurchgängen der Anregung feststellen, ob man eine periodische Näherungslösung gefunden hat, so kann man $\eta^i := z^{i \cdot m} \approx z|_{t=i\frac{2\pi}{\mu}}$ $(i = 1, \ldots)$ setzen und nun prüfen, ob die euklidische Norm $\sqrt{\sum_{j=1}^{n} (\eta_j^{2000} - \eta_j^{2000+k})^2}$ für ein $k \geq 1 \in \mathbb{N}$ kleiner als eine vorgegebene Schranke ε wird. Das Problem dabei ist eine vernünftige Wahl für ε zu treffen. Wählt man die

Schranke ε, ab der Periodizität vorliegen soll, zu klein, so wird durch den Fehler, den das Runge-Kutta-Verfahren verursacht, nie eine Periodizität festgestellt. Wird hingegen die Schranke zu groß gewählt, so bescheinigt das Verfahren einer Lösung, bei der einfach nur die Näherungsfunktionswert für den Zeitpunkt η^{2000} und einen späteren Zeitpunkt η^{2000+k} nahe beieinander liegen, Periodizität, obwohl keine solche vorliegt. Der gesamte Einschließungsalgorithmus kann dann natürlich nicht funktionieren, da ja keine periodische Lösung vorliegt.

- Das Runge-Kutta-Verfahren hat als Ausgabe nur eine Menge von Näherungspunkten für die gesuchte Funktion U. Für den folgenden Newtonalgorithmus und auch für die Intervallarithmetik muß man die Näherung aber in einer diesen zugänglichen Darstellung bereitstellen. Hat die gefundene Näherungslösung eine Periode von $k\frac{2\pi}{\mu}$, so ist es geschickt die Näherungsfunktion ω darzustellen als

$$\omega(t) = \sum_{i=1}^{\text{anz}} a_i \Phi_i(t),$$

wobei die Funktionen $\Phi_i(t) : \mathbb{R} \to \mathbb{R}^n$ die Periode $k\frac{2\pi}{\mu}$ besitzen und orthonormal bezüglich des L_2-Innenprodukts sein sollen und die a_i Konstanten seien. Die Konstanten a_i erhält man dadurch, daß die Bestapproximation einer Funktion in den Raum, der durch die Funktionen Φ_i (i = 1, …, anz) aufgespannt wird, durch die orthogonale Projektion gegeben ist (für nähere Erläuterung dieses Sachverhaltes siehe 4.3.2 “Die Fourierentwicklung im Hilbertraum”). Somit ergibt sich für die Konstanten a_i:

$$\begin{aligned} a_i &:= <\omega, \Phi_i>_2 \\ &= \int_0^{k\frac{2\pi}{\mu}} \omega^t \Phi_i \, dt \\ &\approx h\left(\tfrac{1}{2} z^{2000m}\Phi_i(0) + \sum_{j=1}^{km-1} z^{2000m+j}\Phi_i(jh) + \tfrac{1}{2} z^{(2000+k)m}\Phi_i(k\tfrac{2\pi}{\mu}) \right) \\ &= h\left(\tfrac{1}{2}\Phi_i(0)(z^{2000m} + z^{(2000+k)m}) + \sum_{j=1}^{km-1} z^{2000m+j}\Phi_i(jh) \right) \end{aligned}$$

- Die Approximation ω aus dem Runge-Kutta-Verfahren kann schon gut genug sein, um damit direkt zu Punkt 2 überzugehen. Ist dies nicht der Fall, so verwendet man $\omega_0 := \omega$ als Startnäherung für den zweiten Schritt: das Newtonverfahren.

Die Verbesserung der Startnäherung mittels des Newtonverfahrens

In diesem Abschnitt werden zwei Methoden zur Verbesserung der Näherungslösung, ausgehend von der Startnäherung ω_0 vorgestellt, die beide auf einem Newtonverfahren basieren. Methode 1 beruht geometrisch gesehen auf der Linearisierung des Operators $Tu := -\ddot{u} + B\dot{u} + F(t,u)$, wohingegen Methode 2 auf der Minimierung von $d(\omega)$ beruht.

Methode 1: Setzt man

$$T(u) := -\ddot{u} + B \cdot \dot{u} + F(t,u),$$

so kann man die Bestimmung der Lösung des periodischen Randwertproblems (1.1) auch als $Tu = 0$ in R schreiben. Das Newtonverfahren für die Lösung dieser Operatorengleichung lautet dann

$$\underbrace{T(\omega_n)}_{d[\omega_n]} + T'(\omega_n) \cdot \underbrace{[\omega_{n+1} - \omega_n]}_{=:v_n} \stackrel{!}{=} 0 \quad n = 0,1,2,\ldots$$

wobei T' die Frchetableitung des Operators T bezeichnet. Anders geschrieben ergibt sich also das iterative Verfahren

$$\begin{array}{rcl} T'(\omega_n) \cdot v_n &=& -d[\omega_n] \\ \omega_{n+1} &=& \omega_n + v_n \end{array} \quad n = 0,1,2,\ldots$$

Dabei kann man als Startnäherung ω_0 jenes ω_0 verwenden, das man aus dem Runge-Kutta-Verfahren gewonnnen hat. Setzt man den den Term für die Frchetableitung ein, erhält man

$$\begin{array}{rcl} -\ddot{v}_n + B \cdot \dot{v}_n + \left(\frac{\partial F_i}{\partial u_j}(t,\omega_n)\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \cdot v_n &=& -d[\omega_n] \\ \omega_{n+1} &=& \omega_n + v_n \end{array} \quad n = 0,1,2,\ldots$$

Verwendet man nun den schon im Runge-Kutta-Verfahren gemachten Ansatz $v_n := \sum_{k=1}^{\text{anz}} a_k^{(v_n)} \Phi_k$ (die Bedingungen an die Φ_k waren, wie schon bei der Beschreibung des Runge-Kutta-Verfahrens erwähnt, daß sie jeweils die Randbedingungen erfüllen, also $k\frac{2\pi}{\mu}$ periodisch und in Intervallarithmetik auswertbar sind), so erhält man:

$$\sum_{k=1}^{\text{anz}} a_k^{(v_n)} \left(-\ddot{\Phi}_k + B \cdot \dot{\Phi}_k + \left(\frac{\partial F_i}{\partial u_j}(t,\omega_n)\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \cdot \Phi_k \right) = -d[\omega_n] \tag{1.9}$$

$$\omega_{n+1} = \sum_{k=1}^{\text{anz}} (a_k^{(\omega_n)} + a_k^{(v_n)})\, \Phi_k \tag{1.10}$$

(1.9) ist ein System von n Gleichungen für die unbekannten Konstanten $a_k^{(v_n)}$, das man natürlich nur in den allerseltensten Fällen exakt lösen kann. Um dieses Gleichungssystem dennoch weiter behandeln zu können, bieten sich zwei Wege an:

1. Der erste Weg ist die sogenannte Kollokationsmethode: Man fordert, daß Gleichung (1.9) an bestimmten Punkten ξ_i des Intervalls $[0, k\frac{2\pi}{\mu}]$ exakt erfüllt wird. Verwendet man $\frac{\text{anz}}{n}$ Punkte ξ_i $i = 1, \ldots, \frac{\text{anz}}{n}$, so erhält man ein lineares Gleichungssystem mit anz Gleichungen, für

die anz Unbekannten $a_k^{(v_n)}$. Meist verwendet man über das Intervall $[0, k\frac{2\pi}{\mu}]$ gleichverteilte Punkte ξ_i. Ist aber an einer bestimmten Stelle der Defekt $|d[\omega_n]|_{t=\xi_i}|$ besonders groß, so kann es sinnvoll sein, dort in einer Umgebung die Punkte dichter zu wählen.

2. Der zweite Weg ist das sogenannte Galerkin-Verfahren: Man fordert die sogenannte Fehlerorthogonalität bezüglich des Raumes, der durch die Ansatzfunktionen Φ_k $(k = 1, \ldots, \text{anz})$ aufgespannt wird. In Formeln ausgedrückt heißt dies

$$\left\langle \sum_{k=1}^{\text{anz}} a_k^{(v_n)} \left(-\ddot{\Phi}_k + B\dot{\Phi}_k + \left(\frac{\partial F_i}{\partial u_j}(t, \omega_n)\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \Phi_k \right) - d[\omega_n], \Phi_l \right\rangle \stackrel{!}{=} 0$$

$$\sum_{k=1}^{\text{anz}} a_k^{(v_n)} \left\langle -\ddot{\Phi}_k + B\dot{\Phi}_k + \left(\frac{\partial F_i}{\partial u_j}(t, \omega_n)\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \Phi_k, \Phi_l \right\rangle = - < d[\omega_n], \Phi_l >$$

für $l = 1, \ldots, \text{anz}$. Auch dies liefert also wieder ein lineares Gleichungssystem von anz Gleichungen für die anz Unbekannten $a_k^{(v_n)}$. Bei diesem Gleichungssystem ist allerdings jeder Matrix- oder Vektoreintrag ein L_2-Innenprodukt, das man im allgemeinen numerisch berechnen muß, was einen erhöhten Rechenaufwand gegenüber der Kollokationsmethode bedeutet.

Methode 2: Bei der zweiten Methode, die hier vorgestellt werden soll, wird das Ziel darin bestehen, die Funktion $d(\omega) = ||-\ddot{\omega} + B\dot{\omega} + F(., \omega)||_2$ in dem Raum, der durch die Funktionen $\{\Phi_1, \ldots, \Phi_{\text{anz}}\}$ aufgespannt wird zu minimieren. Da die Näherungslösung $\omega(a_1, \ldots, a_{\text{anz}}; t) = \sum_{i=1}^{\text{anz}} a_i \Phi_i(t)$ als Funktion in Abhängigkeit von den Variablen a_i aufgefaßt werden kann, ist es möglich auch die Funktion $d(\omega(a_1, \ldots, a_{\text{anz}}))$ in Abhängigkeit der Variablen a_i zu sehen. Das Bestreben, kann also so umformuliert werden: Suche das Minimum von $d(\omega(a_1, \ldots, a_{\text{anz}}))$ für $(a_i)_{i=1,\ldots,\text{anz}} \in \mathbb{R}^{\text{anz}}$. Ist nun $d(\omega)$ nach den Variablen a_i differenzierbar, so gilt für das gesuchte Minimum $(\tilde{a}_1, \ldots, \tilde{a}_{\text{anz}})$

$$\frac{\partial d(\omega)}{\partial a_i}(\tilde{a}_1, \ldots, \tilde{a}_{\text{anz}}) = 0 \quad i = 1, \ldots, \text{anz}$$

Setzt man die Funktion $g : \mathbb{R}^{\text{anz}} \to \mathbb{R}^{anz}$; $g_i(a_1, \ldots, a_{\text{anz}}) := \frac{\partial d(\omega)}{\partial a_i}$, so hat man das ursprüngliche Minimierungsproblem zurückgeführt auf das Problem der Nullstellenbestimmung einer vektorwertigen Funktion im $\mathbb{R}^{\text{anz}}$. Linearisiert man nun die Funktion g, so erhält man

$$0 = g(\tilde{a}_1, \ldots, \tilde{a}_{\text{anz}}) \approx g(a_1, \ldots, a_{\text{anz}}) + Dg(a_1, \ldots, a_{\text{anz}}) \cdot ((\tilde{a}_i)_i - (a_i)_i)$$

mit

$$Dg(a_1, \ldots, a_{\text{anz}}) = \begin{pmatrix} \frac{\partial g_1}{\partial a_1} & \cdots & \frac{\partial g_1}{\partial a_{\text{anz}}} \\ \vdots & & \vdots \\ \frac{\partial g_{\text{anz}}}{\partial a_1} & \cdots & \frac{\partial g_{\text{anz}}}{\partial a_{\text{anz}}} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 d(\omega)}{\partial a_1^2} & \cdots & \frac{\partial^2 d(\omega)}{\partial a_1 \partial a_{\text{anz}}} \\ \vdots & & \vdots \\ \frac{\partial^2 d(\omega)}{\partial a_{\text{anz}} \partial a_1} & \cdots & \frac{\partial^2 d(\omega)}{\partial a_{\text{anzt}} \partial a_{\text{anz}}} \end{pmatrix}$$

Falls nun die Funktionalmatrix $Dg(a_1 \ldots, a_{\text{anz}})$ nicht singulär ist, kann man das Gleichungssystem

$$0 = g(a_1, \ldots, a_{\text{anz}}) + Dg(a_1, \ldots, a_{\text{anz}}) \cdot ((\tilde{a}_i)_i - (a_i)_i)$$

nach $(\tilde{a}_i)_i$ auflösen und erhält dann

$$(\tilde{a}_i)_i = (a_i)_i - (Dg(a_1, \ldots, a_{\text{anz}}))^{-1} \cdot g(a_1, \ldots, a_{\text{anz}}).$$

Für die Berechnung der Koeffizienten der Näherungslösung ergibt sich somit die folgende Iterationsvorschrift

$$\begin{pmatrix} a_1^{\omega_{n+1}} \\ \vdots \\ a_{\text{anz}}^{\omega_{n+1}} \end{pmatrix} = \begin{pmatrix} a_1^{\omega_n} \\ \vdots \\ a_{\text{anz}}^{\omega_n} \end{pmatrix} - \begin{pmatrix} \frac{\partial^2 d(\omega_n)}{\partial a_1^2} & \cdots & \frac{\partial^2 d(\omega_n)}{\partial a_1 \partial a_{\text{anz}}} \\ \vdots & & \vdots \\ \frac{\partial^2 d(\omega_n)}{\partial a_{\text{anz}} \partial a_1} & \cdots & \frac{\partial^2 d(\omega_n)}{\partial a_{\text{anz}} \partial a_{\text{anz}}} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{\partial d(\omega_n)}{\partial a_i} \\ \vdots \\ \frac{\partial d(\omega_n)}{\partial a_{\text{anz}}} \end{pmatrix}$$

Um entscheiden zu können, wann die Newtoniteration abzubrechen ist, kann man den, näherungsweise etwa mit der Simpson Formel berechneten, Defekt $||d[\omega_{n+1}]||_2$ heranziehen. Abbrechen würde man dann etwa, wenn $||d[\omega_{n+1}]||_2$ gegenüber $||d[\omega_n]||_2$ nur noch geringfügig kleiner geworden ist. Man erhält dann aus dem Newtonverfahren die benötigte Näherungslösung

$$\omega = \sum_{k=1}^{\text{anz}} a_k \Phi^k := \sum_{k=1}^{\text{anz}} a_k^{\omega_{\text{Abbruch}}} \Phi^k = \omega_{\text{Abbruch}},$$

die dann zur Berechnung der Konstante δ in Punkt 2 und den weiteren Schritten verwendet wird.

1.3.2 Zu Punkt 2: Berechnung der Konstanten δ:

In diesem Abschnitt soll gezeigt werden, wie man eine Konstante δ finden kann, so daß der L_2-Defekt durch δ nach oben beschränkt wird:

$$||d[\omega]||_2 := || - \ddot{\omega} + B \cdot \dot{\omega} + F(., \omega)||_2 \leq \delta.$$

Um solch eine Abschätzung zu erhalten, gibt es, wie bei der Berechnung der Näherungslösung ω selbst, viele Wege, angefangen mit analytischen Methoden bis hin zu computerunterstützen Methoden. Hier werden zwei Möglichkeiten aufgezeigt, wie es mit Hilfe des Computers möglich ist, die gewünschte Abschätzung zu erhalten. Schreibt man den Ausdruck $||d[\omega]||_2$ einmal aus

$$||d[\omega]||_2 = \Big(\int_0^{k\frac{2\pi}{\mu}} \underbrace{\sum_{i=1}^{n} (-\ddot{\omega}_i(t) + (B_{ij}(t))_j \cdot \omega(t) + F_i(t, \omega(t)))^2 \, dt}_{=:h(t)} \Big)^{\frac{1}{2}},$$

so sieht man, daß es sich bei der Bestimmung der Konstanten δ nur darum handelt ein Integral nach oben zu beschränken. Bevor die Beschreibung der beiden Möglichkeiten beginnt, werden noch kurz ein paar Schreibweisen und Konventionen, die dabei verwendet werden, vorgestellt:

- $\mathrm{RND_M} h := \mathrm{RND_M}[h,h] \in \mathrm{I}(\mathbb{R}_\mathrm{M})$ sei das durch Maschinenintervallrundung erhaltene maschinenrepräsentierbare Intervall, das h enthält (siehe 4.1.4 für weitergehende Erläuterungen).

- $\overline{[a,b]} := a$ sei die Obergrenze und $\underline{[a,b]} := a$ sei die Untergrenze eines Intervalls.

- $\oplus_\mathrm{M}$, $\ominus_\mathrm{M}$, $\odot_\mathrm{M}$ und $\oslash_\mathrm{M}$ seien die Maschinenintervalloperationen, die erklärt seien durch

$$\forall A, B \in \mathrm{I}(\mathbb{R}_\mathrm{M}) : A \circledast_\mathrm{M} B := \mathrm{RND_M}(A \circledast B) \in \mathrm{I}(\mathbb{R}_\mathrm{M})$$

  mit $\mathrm{RND_M}$ einer Maschinenintervallrundung, die durch gerichtete Rundung der Intervallgrenzen gewonnen wird. Für die so gewonnenen Maschinenintervalloperationen $* \in \{+, -, \cdot, :\}$ gilt:

$$a \in A \in \mathrm{I}(\mathbb{R}_\mathrm{M}), b \in B \in \mathrm{I}(\mathbb{R}_\mathrm{M}) \Rightarrow \; a * b \in A \circledast_\mathrm{M} B \in \mathrm{I}(\mathbb{R}_\mathrm{M}).$$

- $\mathrm{SIN_M}$, $\mathrm{COS_M}$, $\mathrm{SQRT_M}$, ... seien die intervallmäßig ausgewerteten Pedanten zu den bekannten Standardfunktionen. Beispielhaft gelte für $\mathrm{SIN_M}$

$$\{\sin x : x \in X\} \subseteq \mathrm{SIN_M} X \in \mathrm{I}(\mathbb{R}_\mathrm{M}).$$

Nachdem nun die Schreibweisen, die bei der Verwendung von Intervallarithmetik benutzt werden, eingeführt wurden, folgen nun zwei Vorschläge, wie man die Konstanten δ bestimmen kann.

1. Möglichkeit: Man kann das Problem mit Riemanschen Ober- und Untersummen behandeln. Gemeint ist damit: Unterteilt man das Intervall $[0, k\frac{2\pi}{\mu}]$, über das integriert werden soll, in m gleichgroße Teilintervalle und wertet die Funktion $h(t)$ in jedem dieser Teilintervalle intervallmäßig aus, so kann man das Integral durch die Riemannschen Ober- und Untersummen abschätzen. Will man das soeben in Worte gefaßte ausführen, so muß man also für jedes Intervall $[\underline{t}, \overline{t}]$ die Funktion $h(t)$ durch ein Maschinenintervall $H_\mathrm{M}[\underline{t}, \overline{t}]$ einschränken, so daß gilt

$$\{h(t) : t \in [\underline{t}, \overline{t}]\} \subseteq H_\mathrm{M}([\underline{t}, \overline{t}])$$

   mit $H_\mathrm{M}([\underline{t}, \overline{t}]) \in \mathrm{I}(\mathbb{R}_\mathrm{M})$ Maschinenintervall. Die Bestimmung einer Maschinenintervallfunktion $H_\mathrm{M}([\underline{t}, \overline{t}])$, die diese Forderung erfüllt, ist nach der Wahl der Ansatzfunktionen und der Beschaffenheit der Funktion $h(t)$ aber nicht schwierig, wenn man einfach alle Additionen, Subtraktionen, Multiplikationen, Divisionen und Auswertungen von elementaren Funktionen, die in $h(t)$ aufteten durch Intervalladditionen ($\oplus_\mathrm{M}$), Intervallsubtraktionen ($\ominus_\mathrm{M}$), Intervallmultiplikationen ($\odot_\mathrm{M}$), Intervalldivisionen ($\oslash_\mathrm{M}$) und intervallmäßig ausgewertete Elementarfunktionen ersetzt und dabei ein definierter Intervallausdruck entsteht. Für $||d[\omega]||_2$ gilt dann nach Definition des Riemannintegrals

$$||d[\omega]||_2 \in \mathrm{SQRT_M}\left(\Delta \odot_\mathrm{M} \Big(\bigoplus_{i=1}^{m}{}_\mathrm{M} H_\mathrm{M}([t_{i-1}, t_i])\Big)\right) =: [\underline{d}, \overline{d}]$$

(mit $\Delta := \mathrm{RND_M} k \diamond_\mathrm{M} \mathrm{RND_M} 2 \diamond_\mathrm{M} \mathrm{RND_M} \pi \diamond_\mathrm{M} (\mathrm{RND_M} \mu \diamond_\mathrm{M} \mathrm{RND_M} m)$ und $[t_{i-1}, t_i] := [\underline{\mathrm{RND_M}(i-1) \diamond_\mathrm{M} \Delta}, \overline{\mathrm{RND_M} i \diamond_\mathrm{M} \Delta}])$. Man erhält also eine Einschließung für den Wert von $||d[\omega]||_2$. Die Untergrenze $\underline{d}$ braucht man für die Bestimmung der Konstante δ nicht, aber setzt man $\delta := \overline{d}$, so erhält man die gewünschte Abschätzung $||d[\omega]||_2 \leq \delta$.

2. Möglichkeit: Man kann das Problem mit der Trapezformel mit Fehlerabschätzung behandeln (siehe hierzu auch 4.2.3). Unterteilt man das Intervall $[0, k\frac{2\pi}{\mu}]$ äquidistant in m Teile, so ergibt sich mit Hilfe der Trapezformel

$$\int_0^{k\frac{2\pi}{\mu}} h(t)\,dt = h\left(\frac{h(t_0)}{2} + \sum_{i=1}^{n-1} h(t_i) + \frac{h(t_n)}{2}\right) + \frac{k\frac{2\pi}{\mu}h^2}{12}\ddot{h}(\xi)$$

mit $h := k\frac{2\pi}{\mu m}$, $\xi \in [0, k\frac{2\pi}{\mu}]$ und $t_i := ih$. Ersetzt man nun alle Operationen durch Intervalloperationen und die Funktionen h und $\ddot{h}$ durch ihre computerrepräsentierbare intervallmäßige Auswertungen H_M und $\ddot{H}_\mathrm{M}$ (diese Intervallmäßige Auswertung erhält man sehr einfach, wenn bei der Ersetzung aller Verknüpfungen durch Intervallverknüpfungen und aller Standartfunktionen durch Intervallstandartfunktionen in h und $\ddot{h}$ ein definierter Intervallausdruck entsteht), so erhält man:

$$\begin{aligned}
||d[\omega]||_2^2 \in\ & \mathrm{RND_M} h \diamond_\mathrm{M} \\
& \left((H_\mathrm{M}(T_0) \diamond_\mathrm{M} \mathrm{RND_M} 2) \mathbin{\diamond\!\!\!+}_\mathrm{M} (\mathop{\diamond\!\!\!+_\mathrm{M}}_{i=1}^{n-1} H_\mathrm{M}(T_i)) \mathbin{\diamond\!\!\!+}_\mathrm{M} (H_\mathrm{M}(T_n) \diamond_\mathrm{M} \mathrm{RND_M} 2)\right) \\
& \mathbin{\diamond\!\!\!+}_\mathrm{M} (\mathrm{RND_M}(k \cdot 2) \diamond_\mathrm{M} \mathrm{RND_M} \pi \diamond_\mathrm{M} \mathrm{RND_M} \mu \diamond_\mathrm{M} \mathrm{SQR_M}(\mathrm{RND_M} h) \diamond_\mathrm{M} \mathrm{RND_M} 12 \\
& \diamond_\mathrm{M} [\min_{\xi \in [0, k\frac{2\pi}{\mu}]} \ddot{h}(\xi), \max_{\xi \in [0, k\frac{2\pi}{\mu}]} \ddot{h}(\xi)] \\
=:\ & [\underline{d}, \overline{d}]
\end{aligned}$$

Setzt man $\delta := \overline{\mathrm{SQRT_M}[\underline{d}, \overline{d}]}$ oder $\delta := \overline{\mathrm{SQRT_M}[0, \overline{d}]}$ falls $\underline{d} \leq 0$, so hat man also die gewünschte Abschätzung $||d[\omega]||_2 \leq \delta$. Die Berechnung der Konstanten δ erfolgt somit durch einfache Intervalladditionen und Intervallmultiplikationen von intervallmäßigen Auswertungen von h an Punktintervallen und durch die Bestimmung eines Minimums und eines Maximums von $\ddot{h}$ in $[0, k\frac{2\pi}{\mu}]$. Die Abschätzung dieses Minimums und Maximums wiederum, kann auch mit Intervallarithmetik erfolgen, es gilt nämlich:

$$\begin{aligned}
\max_{\xi \in [0, k\frac{2\pi}{\mu}]} \ddot{h}(\xi) &\leq \max_{0 \leq i \leq m-1} \overline{\ddot{H}_\mathrm{M}(T_i)}, \\
\min_{\xi \in [0, k\frac{2\pi}{\mu}]} \ddot{h}(\xi) &\geq \min_{0 \leq i \leq m-1} \underline{\ddot{H}_\mathrm{M}(T_i)}
\end{aligned}$$

mit

$$\begin{aligned}
T_i := [&\underline{\mathrm{RND_M}(i \cdot k \cdot 2) \diamond_\mathrm{M} \mathrm{RND_M} \pi \diamond_\mathrm{M} (\mathrm{RND_M} \mu \diamond_\mathrm{M} \mathrm{RND_M} m)}, \\
&\overline{\mathrm{RND_M}((i+1) \cdot k \cdot 2) \diamond_\mathrm{M} \mathrm{RND_M} \pi \diamond_\mathrm{M} (\mathrm{RND_M} \mu \diamond_\mathrm{M} \mathrm{RND_M} m)}].
\end{aligned}$$

Bemerkungen:

- Die 2te Möglichkeit verspricht ein Verfahren, bei dem der Fehler des Integrals mit h^2 gegen null geht, für den Defekt hingegen, der ja die Wurzel aus diesem Integral ist bedeutet dies, das der Fehler nur mit h gegen null konvergiert. Es gibt zwei prinzipielle Gründe, die dagegen sprechen könnten, Variante 2 zu verwenden. Es wäre zum ersten denkbar, daß die Auswertung der zweiten Ableitung einen sehr hohen Aufwand erfordert. Zweitens wäre der Verzicht auf Variante 2 sinnvoll, wenn der Wert der zweiten Ableitung an einer Stelle sehr groß wird, wie dies etwa bei Funktionen, die beinahe Knicke in der ersten Ableitung haben, der Fall ist.

- Es ist durchaus denkbar, bei genügend glatter Funktion h, Integrationsverfahren mit Fehlerabschätzung noch höherer Ordnung zu verwenden, um eine schnellere Konvergenz zu erhalten.

1.3.3 Zu Punkt 3: Berechnung der Konstanten σ

Bevor mit dem Thema der Berechnung der Konstante σ begonnen wird, sei darauf hingewiesen, daß die Konstante ausschließlich zur Berechnung der Konstanten K benötigt wird. Möchte man also einen anderen, als den in "zu Punkt 4: Berechnung der Konstanten K", beschriebenen Algorithmus verwenden, kann man diesen Abschnitt getrost überspringen.

Ziel dieses Abschnittes wird es sein, eine untere Schranke σ für den kleinsten Eigenwerte des Operators $L^*L : D(L^*L) \subseteq H \to H \subseteq L_2^n(0, k\frac{2\pi}{\mu})$ zu finden. Dabei wird sich herausstellen, daß es nicht genügt, nur den kleinsten Eigenwert einzuschließen, sondern daß es nötig sein wird die ersten N Eigenwerte einzuschließen, mit einem gegebenen $N \in \mathbb{N}$.

Als Generalvoraussetzung für alle weiteren Untersuchungen in diesem Abschnitt gelte: Der pre-Hilbertraum H sei unendlichdimensional und besitze eine Orthonormalbasis $(\varphi_n)_{n\in\mathbb{N}}$ aus Eigenfunktionen des Operators $L^*L : D(L^*L) \subseteq H \to H \subseteq L_2^n(0, k\frac{2\pi}{\mu})$, wobei $D(L^*L)$ linearer dichter Unterraum von H sei. Die zugehörigen Eigenwerte $(\lambda_n)_{n\in\mathbb{N}}$ seien der Größe nach angeordnet und es gelte $\lim_{n\to\infty} \lambda_n = \infty$

Zu allererst soll in diesem Abschnitt das sogenannte Rayleigh-Ritz-Verfahren vorgestellt werden, das hier dazu dienen soll zu den ersten N Eigenwerten und Eigenfunktionen, Näherungseigenwerte und Näherungseigenfunktionen zu berechnen.

Nachdem man nun diese Näherungen berechnet hat, stehen einem zwei Wege für das weitere Vorgehen offen:

1. Weg: Man berechnet zu jedem Näherungseigenwerte und zu jeder Näherungseigenfunktione einen sogenannten Defekt und erhält dann durch die Aussage des Satzes von Weinstein ein Intervall, das mindestens einen Eigenwert des Operators L^*L enthält. Problematisch hierbei ist, daß man keine Information über den Index des eingeschlossenen

Eigenwertes erhält und auch nicht ob man in einem Intervall eventuell mehr als einen Eigenwert eingeschlossen hat. Um Informationen über den Index des eingeschlossenen Eigenwertes zu bekommen, wird es notwendig sein, vorab eine untere Schranke des Eigenwertes λ_{N+1} zu kennen. Dem ebenfalls auftretenden Problem von überlappenden Eigenwerteinschließungen kann mit einer Variante des Satzes von Weinstein, dem Satz von Weinstein für geclusterte Eigenwerte, begegnet werden. Die nur mäßige Qualität der Eigenwerteinschließungen des Weinsteinverfahrens läßt sich mit dem Satz von Kato in den Griff bekommen.

2. Weg: Man verwendet die Näherungseigenfunktionen als Ansatzfunktionen und erhält dann mit Hilfe des Rayleigh-Ritz-Verfahrens zur Bestimmung von oberen Eigenwertschranken ebensolche. Hat man nun als Vorabinformation eine untere Schranke des Eigenwertes λ_{N+1}, so kann man mit dem Verfahren von Lehmann untere Eigenwertschranken bestimmen. Nimmt man beide Verfahren zusammen erhält man sowohl obere, als auch untere Schranken für die Eigenwerte, also Einschließungen.

Beiden Wegen ist gemeinsam, daß man vorab eine untere Schranke für den Eigenwert λ_{N+1} benötigt. Mit der sogenannten Homothopiemethode wird nun als letztes ein Verfahren vorgestellt, mit dem man ausgehend von einem Vergleichsoperator, bei dem man Einschließungen der Eigenwerte und Eigenvektoren kennt (etwa weil er so einfach ist, daß man sie exakt berechnen kann), über eine ganze Familie von Operatoren, man schließlich zu einer unteren Schranke für den Eigenwert λ_{N+1} des betrachteten Operators L^*L gelangt. Bei dieser Homothopiemethode besteht jeder einzelne Schritt darin, für einen Operator dieser Familie von Operatoren, der zwischen dem Vergleichsoperator und L^*L liegt, mit den oben beschriebenen Verfahren Einschließungen mit Indexinformationen für dessen Eigenwerte zu erhalten.

Das Rayleigh-Ritz Verfahren zur Bestimmung von Näherungseigenwerten und Näherungseigenfunktionen:

Seien $v_1, \ldots, v_M \in D(L^*L)$ M linear unabhängige Ansatzfunktionen (z.B. trigonometrische Polynome). Bildet man dann mit diesen Ansatzfunktionen die $M \times M$-Matrizen

$$\begin{aligned} A_1 &:= (< L^*Lv_i, v_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}} = (< Lv_i, Lv_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}} \\ A_2 &:= (< v_i, v_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}} \end{aligned}$$

und löst damit, z.B. mit dem in 4.2.2 beschriebenem QR-Algorithmus, das symmetrische Eigenwertproblem

$$A_1 x = \tilde{\lambda} A_2 x \quad x \in \mathbb{R}^M \backslash \{0\}, \tag{1.11}$$

so approximieren, die dabei gefundenen Eigenwerte $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_N$ für ausreichend kleines N (etwa $N = \frac{M}{2}$) die Eigenwerte von L^*L meist recht gut.

Außerdem sind die Funktionen $\tilde{u}_i := \sum_{k=1}^{M} x_k^{(n)} v_k \; n = 1, \dots, N$ recht gute Näherungen für die Eigenfunktionen $\varphi_1, \dots, \varphi_N$. Wobei $x^{(1)}, \dots, x^{(M)} \in \mathbb{R}^M$ die Eigenvektoren des Problemes (1.11) seien, die

$$(x^{(i)})^t A_2 x^{(j)} = \delta_{ij}, \qquad (x^{(i)})^t A_1 x^{(j)} = \tilde{\lambda}_i \delta_{ij}$$

erfüllen.

Bemerkung:
Die Qualität der Approximation der Eigenwerte und Eigenfunktionen hängt stark von der Wahl der Ansatzfunktionen ab. Die wahren Eigenfunktionen $\varphi_1, \dots, \varphi_N$ sollten "nicht weit" von der linearen Hülle der Ansatzfunktionen $v_1, \dots, v_M$ entfernt sein.

Satz (Weinstein):

Sei $\tilde{u} \in D(L^*L) \backslash \{0\}$ und $\tilde{\lambda} \in \mathbb{R}$. Weiterhin sei der Defekt δ definiert durch:

$$\delta := \frac{||L^*L\tilde{u} - \tilde{\lambda}\tilde{u}||_2}{||\tilde{u}||_2}.$$

Dann enthält das Intervall $[\tilde{\lambda} - \delta, \tilde{\lambda} + \delta]$ mindestens einen Eigenwert des Problems $L^*Lu = \lambda u, \; u \in D(L^*L)$.

Beweis:

$$\begin{aligned}
||L^*L\tilde{u} - \tilde{\lambda}\tilde{u}||_2^2 &= \sum_{i=1}^{\infty} < L^*L\tilde{u} - \tilde{\lambda}\tilde{u}, \varphi_i >_2^2 \\
&= \sum_{i=1}^{\infty} (\lambda_n - \tilde{\lambda})^2 < \tilde{u}, \varphi_i >_2^2 \\
&\geq \min_{i \in \mathbb{N}} |\lambda_i - \tilde{\lambda}|^2 \sum_{i=1}^{\infty} < \tilde{u}, \varphi_i >_2^2 \\
&= \min_{i \in \mathbb{N}} |\lambda_i - \tilde{\lambda}|^2 ||\tilde{u}||_2^2.
\end{aligned}$$

Teilt man diese Ungleichung durch $||\tilde{u}||_2^2$ und zieht dann auf beiden Seiten die Wurzel, so erhält man:

$$\delta \geq \min_{i \in \mathbb{N}} |\lambda_i - \tilde{\lambda}| \Rightarrow \; \exists i \in \mathbb{N} : \lambda_i \in [\tilde{\lambda} - \delta, \tilde{\lambda} + \delta]$$

$\square$

Bemerkungen:

1. Üblicherweise wird man den Satz von Weinstein nicht mit irgendwelchen $\tilde{\lambda} \in \mathbb{R}$ und $\tilde{u} \in D(L^*L)$ anzuwenden versuchen, sondern vielmehr mit zuvor numerisch berechneten Näherungseigenwerten und Näherungseigenfunktionen. In diesem Fall ist dann auch der berechnete Defekt δ klein und man erhält gute Einschließungen, sprich ein kleines Intervall, in dem ein Eigenwert des Problems $L^*Lu = \lambda u, \; u \in D(L^*L)$ liegt.

2. Berechnet man nach dem Satz von Weinstein eine Einschließung für einen Eigenwerte von L^*L, so hat man leider keinerlei Informationen darüber, den wievielten Eigenwert man denn nun tatsächlich eingeschlossen hat. Kann man aber N nichtüberlappende Eigenwerteinschließungen berechnen, für die jeweils gilt $[\tilde{\lambda} - \delta, \tilde{\lambda} + \delta] \leq \lambda_{N+1}$, so kann man sicher sein, die ersten N Eigenwerte eingeschlossen zu haben. Man sieht aber, daß man um Informationen über den Index der ersten N Eigenwerte zu erhalten als Vorinformation eine untere Schranke des Eigenwertes λ_{N+1} benötigt, die ihrerseits auch noch größer sein muß, als alle Intervalleinschließungen.

3. Die Qualität der Weinsteineinschließungen ist nicht besonders gut, d.h. der Durchmesser der Einschließungen ist relativ groß. Diesem Problem kann man aber mit Verwendung des Satzes von Kato begegnen, der immerhin Grenzen in der Güte $O(\delta^2)$ liefert.

4. Berechnet man für mehrere Eigenwerte λ Einschließungsintervalle und überlappen sich etwa zwei dieser Intervalle, so ist im allgmeinen nicht sichergestellt, daß man zwei verschiedene Eigenwerte eingeschlossen hat. In diesem Fall kann man aber immer noch versuchen, den folgenden Satz (Weinstein für geclusterte Eigenwerte) anzuwenden.

Satz (Weinstein für geclusterte Eigenwerte):

Seien $\tilde{u}_1, \ldots, \tilde{u}_l \in D(L^*L)$ linear unabhängig. Seien weiterhin $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l \in \mathbb{R}$ der Größe nach geordnet. Für alle $i = 1, \ldots, l$ sei der Defekt δ_i definiert durch

$$\delta_i := \frac{||L^*L\tilde{u}_i - \tilde{\lambda}_i \tilde{u}_i||_2}{||\tilde{u}_i||_2} \quad i = 1, \ldots, l.$$

Wenn dann für den maximalen Eigenwert ρ der $l \times l$-Matrix $I - U$ mit $U_{ij} := \frac{<\tilde{u}_i, \tilde{u}_j>_2}{||\tilde{u}_i||_2 ||\tilde{u}_j||_2}$ $1 \leq i, j \leq l$ gilt: $\rho < 1$, so enthält das Intervall $[\tilde{\lambda}_1 - \delta, \tilde{\lambda}_l + \delta]$ mindestens l Eigenwerte des Problems $L^*Lu = \lambda u, \ u \in D(L^*L)$, wobei δ definiert sei durch

$$\delta := (1 - \rho)^{-\frac{1}{2}} \left(\sum_{i=1}^{l} \delta_i^2 \right)^{\frac{1}{2}}.$$

Bemerkungen:

1. Falls die Näherungseigenfunktionen $\tilde{u}_1, \ldots \tilde{u}_l$ näherungsweise orthogonal sind, wie sie eigentlich jeder numerische Algorithmus liefert, so ist die Matrix $I - U$ beinahe die Nullmatrix und auch der größte Eigenwert dieser Matrix ist nahezu null. Die Forderung $\rho < 1$ ist damit keine Zusatzforderung, sondern eigentlich immer erfüllt.

2. Man wird den Satz von Weinstein für geclusterte Eigenwerte im Allgemeinen nur dann anwenden, wenn beim Satz von Weinstein überlappende Eigenwerteinschließungen aufgetreten sind und man feststellen möchte, wieviele Eigenwerte denn nun mindestens eingeschlossen wurden.

3. Auch der Satz von Weinstein für geclusterte Eigenwerte liefert wieder keine Informationen über die Indizes der eingeschlossenen Eigenwerte. Dafür benötigt man abermals, möchte man N Eigenwerte einschließen, eine untere Schranke für den Eigenwerte λ_{N+1}.

Vorgehen zur Bestimmung von Einschließungen für die N ersten Eigenwerte:
Falls man als Vorabinformation eine untere Schranke für den $N+1$ Eigenwert von L^*L kennt, so bieten die beiden letzten Sätze eine Möglichkeit die ersten N Eigenwerte einzuschließen. Das Vorgehen hierfür ist wie folgt:

1. Man bestimme mit beliebigem numerischen Algorithmus N Näherungseigenwerte $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_N$ und zugehörige Näherungseigenfunktionen $\tilde{u}_1, \ldots, \tilde{u}_N$. Hierbei kann z.B. das zuvor beschriebene Rayleigh-Ritz Verfahren Verwendung finden.

2. Nun bestimme man für diese N Näherungseigenwerte jeweils eine obere Schranke für $\delta_1, \ldots, \delta_N$ und bilde damit die Eigenwertintervalleinschließungen. Falls sich diese Einschließungen nicht überlappen, so gehe man zu Punkt 4 ansonsten zu Punkt 3.

3. Auf Näherungseigenwerte, deren Eigenwertintervalleinschließungen sich überlappen, versuche man Satz (Weinstein für geclusterte Eigenwerte) anzuwenden. Kommt es nach Anwendung des Satzes zu neuen Intervallüberlappungen, so wende man Punkt 3, mit den zusätzlichen überlappenden Eigenwerten, erneut an. Auf diese Weise produziert man keine Endlosschleife, denn das Verfahren bricht spätestens dann ab, wenn alle Eigenwerteinschließungen überlappen, also spätestens nach N-facher Ausführung von Punkt 3.

4. Gilt für die größte Intervallgrenze der Eigenwerteinschließungen, daß sie kleiner ist als die als Vorabinformation bekannte untere Schranke von λ_{N+1}, so hat man genau die ersten N Eigenwerte des Operators L^*L eingeschlossen und kennt somit den Index des Eigenwertes der jeweiligen Eigenwerteinschließung.

Satz (Kato):

Sei $\tilde{u} \in D(L^*L)\backslash\{0\}$ und $\tilde{\lambda}$ definiert durch

$$\tilde{\lambda} := \frac{< L^*L\tilde{u}, \tilde{u} >_2}{< \tilde{u}, \tilde{u} >_2} = \frac{||L[\tilde{u}]||_2^2}{||\tilde{u}||_2^2}.$$

Außerdem sei der Defekt δ, wie schon beim Satz von Weinstein, definiert durch:

$$\delta := \frac{||L^*L\tilde{u} - \tilde{\lambda}\tilde{u}||_2}{||\tilde{u}||_2} = \frac{\left[||L^*L\tilde{u}||_2^2||\tilde{u}||_2^2 \cdot - < L^*L\tilde{u}, \tilde{u} >_2^2\right]^{\frac{1}{2}}}{||\tilde{u}||_2^2}.$$

Schließlich seien noch für ein $n \in \mathbb{N}$ eine obere Schranke μ für λ_{n-1} und eine untere Schranke ν für λ_{n+1} gegeben, für die gilt:

$$\lambda_{n-1} \leq \mu < \tilde{\lambda} < \nu \leq \lambda_{n+1}.$$

Dann gilt (mit $\lambda_0 := \mu := -\infty$, falls $n = 1$):

$$\boxed{\tilde{\lambda} - \frac{\delta^2}{\nu - \tilde{\lambda}} \leq \lambda_n \leq \tilde{\lambda} + \frac{\delta^2}{\tilde{\lambda} - \mu}}.$$

Beweis: Zuerst ist noch zu zeigen, daß die angegebene Umformung für δ korrekt ist. Es gilt:

$$\begin{aligned}
\delta \quad &:= \quad \frac{||L^*L\tilde{u} - \tilde{\lambda}\tilde{u}||_2}{||\tilde{u}||_2} = \frac{(||L^*L\tilde{u} - \tilde{\lambda}\tilde{u}||_2^2)^{\frac{1}{2}}}{||\tilde{u}||_2} \\
&= \quad \frac{(||L^*L\tilde{u}||_2^2 - 2\tilde{\lambda} < L^*L\tilde{u}, u >_2 + \tilde{\lambda}^2||\tilde{u}||_2^2)^{\frac{1}{2}}}{||\tilde{u}||_2} \\
&\overset{\text{Def.}\tilde{\lambda}}{=} \quad \frac{\left(||L^*L\tilde{u}||_2^2 - 2\frac{<L^*L\tilde{u},u>_2^2}{||\tilde{u}||_2^2} + \frac{<L^*L\tilde{u},u>_2^2}{||\tilde{u}||_2^2}\right)^{\frac{1}{2}}}{||\tilde{u}||_2} \\
&= \quad \frac{(||L^*L\tilde{u}||_2^2 \cdot ||\tilde{u}||_2^2 - < L^*L\tilde{u}, \tilde{u} >_2^2)^{\frac{1}{2}}}{||\tilde{u}||_2^2}. \qquad (1.12)
\end{aligned}$$

Für $\rho = \mu$ oder $\rho = \nu$ und alle $i \in \mathbb{N}$ gilt

$$\underbrace{(\lambda_i - \lambda_n)(\lambda_i - \rho)}_{\geq 0} \underbrace{< \tilde{u}, \varphi_i >_2^2}_{\geq 0} \geq 0.$$

Summiert man all diese Ungleichungen über i, so erhält man:

$$\begin{aligned}
\sum_{i=1}^{\infty}(\lambda_i - \lambda_n)(\lambda_i - \rho) < \tilde{u}, \varphi_i >_2^2 &\geq 0 \\
\sum_{i=1}^{\infty}[\lambda_i^2 < \tilde{u}, \varphi_i >_2^2 -(\lambda_n + \rho)\lambda_i < \tilde{u}, \varphi_i >_2^2 +\lambda_n\rho < \tilde{u}, \varphi_i >_2^2] &\geq 0 \\
||L^*L\tilde{u}||_2^2 - (\lambda_n + \rho) < L^*L\tilde{u}, \tilde{u} >_2 +\lambda_n\rho||\tilde{u}||_2^2 &\geq 0 \\
\frac{||L^*L\tilde{u}||_2^2}{||u||_2^2} - (\lambda_n + \rho) \underbrace{\frac{< L^*L\tilde{u}, \tilde{u} >_2}{||\tilde{u}||_2^2}}_{=\tilde{\lambda}} +\lambda_n\rho &\geq 0.
\end{aligned}$$

Für den ersten Term $\frac{||L^*L\tilde{u}||_2^2}{||u||_2^2}$ folgt nach (1.12) und weiterem umformen:

$$\begin{aligned}
\delta||\tilde{u}||_2^2 &= \left[||L^*L\tilde{u}||_2^2||\tilde{u}||_2^2 - < L^*L\tilde{u}, \tilde{u} >_2^2\right]^{\frac{1}{2}} \\
\delta^2||\tilde{u}||_2^4 &= ||L^*L\tilde{u}||_2^2||\tilde{u}||_2^2 - < L^*L\tilde{u}, \tilde{u} >_2^2 \\
\delta^2 &= \frac{||L^*L\tilde{u}||_2^2}{||\tilde{u}||_2^2} - \underbrace{\frac{< L^*L\tilde{u}, \tilde{u} >_2^2}{||\tilde{u}||_2^4}}_{\tilde{\lambda}^2} \\
\delta^2 + \tilde{\lambda}^2 &= \frac{||L^*L\tilde{u}||_2^2}{||\tilde{u}||_2^2}.
\end{aligned}$$

Setzt man diese gefundene Umformung in die obige Ungleichung ein, so erhält man insgesamt also:

$$\begin{aligned} &\delta^2 + \tilde{\lambda}^2 - (\lambda_n + \rho)\tilde{\lambda} + \lambda_n \rho \geq 0 \\ \Rightarrow\ &\lambda_n(\rho - \tilde{\lambda}) \geq \tilde{\lambda}(\rho - \tilde{\lambda}) - \delta^2. \end{aligned}$$

Für $\rho = \mu$ ergibt sich der erste Fall:

$$\lambda_n \underbrace{(\mu - \tilde{\lambda})}_{<0} \geq \tilde{\lambda}(\mu - \tilde{\lambda}) - \delta^2$$

$$\lambda_n \leq \tilde{\lambda} + \frac{\delta^2}{\tilde{\lambda} - \mu}.$$

Für $\rho = \nu$ der zweite:

$$\lambda_n \underbrace{(\nu - \tilde{\lambda})}_{>0} \geq \tilde{\lambda}(\nu - \tilde{\lambda}) - \delta^2$$

$$\lambda_n \geq \tilde{\lambda} - \frac{\delta^2}{\nu - \tilde{\lambda}}.$$

$\square$

Bemerkung:
Wenn die Eigenwerte gut voneinander getrennt sind, also nicht clustern, so kann man die a priori Information einer oberen Schranke von λ_{n-1} und einer unteren Schranke von λ_{n+1} aus dem Verfahren von Weinstein gewinnen. Die dort erhaltenen Grenzen der Eigenwerteinschließungen kann man als Werte für μ und ν im Satz von Kato verwenden. Wichtig hierbei ist wiederum, daß man auch dafür erneut das Wissen über den Index des mit dem Weinsteinverfahren eingeschlossenen Eigenwertes benötigt, was die Kenntnis einer unteren Schranke des Eigenwertes λ_{N+1} voraussetzt. Mit dem Satz von Kato ist also eine Verbesserung der Qualität der durch den Satz von Weinstein gewonnenen Eigenwerteinschließungen möglich.

Satz (Rayleigh-Ritz Verfahren zur Bestimmung von oberen Eigenwertschranken):

Für die aus dem Rayleigh-Ritz Verfahren gewonnnenen Näherungseigenwerte $\tilde{\lambda}_i$ gilt:

$$\boxed{\lambda_i \leq \tilde{\lambda}_i \quad i = 1, \ldots, M.}$$

Bemerkungen:

- Das Rayleigh-Ritz Verfahren läßt sich also zur Gewinnung oberer Eigenwertschranken nutzen. Dazu müssen aber die Integrale aus denen die Matrizen

$$\begin{aligned} A_1 &:= (< Lv_i, Lv_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}} \\ A_2 &:= (< v_i, v_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}} \end{aligned}$$

aufgebaut sind in Intervalle eingeschlossen werden. Hierzu kann man wie schon in 1.3.2 "Zu Punkt 2: Berechnung der Konstanten δ:" beschrieben Riemannsche Ober- und Untersummen oder aber Integrationsformeln mit Fehlerabschätzung verwenden. Die Integrationen, die für die Matrix A_2 benötigt werden können eventuell so einfach sein, daß sie analytisch durchgeführt werden können, aber denoch wird auch hier die Intervallrechnung benötigt um Rundungsfehler mit zu erfassen.

- Wenn man statt irgendwelcher Ansatzfunktionen v_i $i = 1, \ldots, M$ normierte Näherungseigenfunktionen $\tilde{u}_i$ $i = 1, \ldots, M$ verwendet, um die Matrizen

$$A_1 := (< L\tilde{u}_i, L\tilde{u}_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}} \text{ und } A_2 := (< \tilde{u}_i, \tilde{u}_j >_2)_{\substack{i=1,\ldots,M \\ j=1,\ldots,M}}$$

  zu bilden, so ist die Matrix A_1 nahezu eine Diagonalmatrix mit den Eigenwerten auf der Diagonale und die Matrix A_2 nahezu die Einheitsmatrix. Solche normierten Näherungseigenfunktionen $\tilde{u}_i$ kann man leicht aus einem zuvor durchgeführten Rayleigh-Ritz Verfahren ohne Berücksichtigung von Rundungsfehlern und nicht exakt ausgeführten Integrationen gewinnen.

- Verwendet man nun normierte Näherungseigenfunktionen $\tilde{u}_i$ zur Bildung der Matrizen A_1 und A_2 und berücksichtigt Integrations- und Rundungsfehler (d.h. A_1 und A_2 sind Intervallmatrizen), so kann man leicht den Satz von Gerschgorin zur Berechnung von Einschließung der Eigenwerte $\tilde{\lambda}$ des Problems $A_1 x = \tilde{\lambda} A_2 x$ verwenden. Bildet man die Matix $\triangle := I - A_2$, so ist diese beinahe die Nullmatrix und die Forderung $||\triangle||_\infty < 1$ sicherlich erfüllt. Die Matrix A_2^{-1} läßt sich dann als Neumannsche Reihe schreiben:

$$\begin{aligned} A_2^{-1} &= \sum_{k=0}^{\infty} \triangle^k = \sum_{k=0}^{m-1} \triangle^k + \sum_{k=m}^{\infty} \triangle^k \\ &= \sum_{k=0}^{m-1} \triangle^k + \triangle^m \sum_{k=0}^{\infty} \triangle^k = \sum_{k=0}^{m-1} \triangle^k + \triangle^m A_2^{-1} \\ &= \sum_{k=0}^{m-1} \triangle^k + \triangle^m (I - \triangle)^{-1} \end{aligned}$$

  was durch

$$\left|\left| A_2^{-1} - \sum_{k=0}^{m-1} \triangle^k \right|\right|_\infty \leq \frac{||\triangle^m||_\infty}{1 - ||\triangle||_\infty} \qquad \text{für beliebiges } m \in \mathbb{N}$$

  eine hinreichend genaue Einschließung für die Matrix A_2^{-1} liefert. Führt man dann mit dieser Einschließung das Matixprodukt $A_2^{-1} A_1$ in Intervallarithmetik aus, so erhält man eine immer noch nahezu diagonale Einschließnung für $A_2^{-1} A_1$. Der Satz von Gerschgorin, der besagt, daß die Vereinigung aller Kreisscheiben

$$K_i := \left\{ \mu \in \mathbb{C} : |\mu - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^{M} |a_{ik}| \right\} \qquad 1 \leq i \leq M$$

allle Eigenwerte der $m \times m$-Matrix A enthält liefert, unter Berücksichtigung der Tatsache das alle Eigenwerte $\tilde{\lambda}_i$ reell sind, Einschließungen für die Rayleigh-Ritz-Eigenwerte.

Für den Beweis dieses Satzes wird die Variationsformulierung der Eigenwerte benötigt, die deswegen hier zwischen Satz und dessen Beweis eingeschoben wird.

Satz (Variationsformulierung der Eigenwerte):

Sind die Eigenwerte des Operators L^*L in aufsteigender Reihenfolge angeordnet, so gilt für alle $n \in \mathbb{N}$:

$$\boxed{\lambda_n = \min_{\substack{U \subseteq D(L^*L) \text{ Unterraum} \\ \dim U = n}} \max_{\substack{u \in U \\ u \neq 0}} \frac{< L^*Lu, u >_2}{< u, u >_2} = \min_{\substack{U \subseteq D(L^*L) \text{ Unterraum} \\ \dim U = n}} \max_{\substack{u \in U \\ u \neq 0}} \frac{||Lu||_2^2}{||u||_2^2}}.$$

Insbesondere gilt für den Spezialfall $n = 1$:

$$\boxed{\lambda_1 = \min_{u \in D(L^*L) \setminus \{0\}} \frac{||Lu||_2^2}{||u||_2^2}}.$$

Beweis:

"$\leq$" Sei $U \subseteq D(L^*L)$ beliebiger Unterraum mit $\dim U = n$. Wähle ein $u \in U$ mit $< u, \varphi_i >= 0$ $i = 1, \ldots, n-1$. Dann gilt:

$$\begin{aligned}
\frac{< L^*Lu, u >_2}{< u, u >_2} &= \frac{< L^*L \sum_{i=1}^{\infty} < u, \varphi_i >_2 \varphi_i, \sum_{j=1}^{\infty} < u, \varphi_j >_2 \varphi_j >_2}{< \sum_{i=1}^{\infty} < u, \varphi_i >_2 \varphi_i, \sum_{j=1}^{\infty} < u, \varphi_j >_2 \varphi_j >_2} \\
&= \frac{\sum_{i=n}^{\infty} \lambda_i < u, \varphi_i >_2^2}{\sum_{i=n}^{\infty} < u, \varphi_i >_2^2} \\
&\geq \lambda_n \frac{\sum_{i=n}^{\infty} < u, \varphi_i >_2^2}{\sum_{i=n}^{\infty} < u, \varphi_i >_2^2} = \lambda_n.
\end{aligned}$$

$$\Rightarrow \max_{u \in U \setminus \{0\}} \frac{< L^*Lu, u >_2}{< u, u >_2} \geq \lambda_n \quad \forall U \subseteq D(L^*L) \text{ Unterraum mit } \dim U = n$$

$$\Rightarrow \inf_{\substack{U \subseteq D(L^*L) \text{ Unterraum} \\ \dim U = n}} \max_{u \in U \setminus \{0\}} \frac{< L^*Lu, u >_2}{< u, u >_2} \geq \lambda_n.$$

"$\geq$" Sei $U := \text{span}\{\varphi_1, \ldots, \varphi_n\}$. Dann gilt für alle $u \in U \setminus \{0\}$:

$$\frac{< L^*Lu, u >_2}{< u, u >_2} = \frac{\sum_{i=1}^{\infty} \lambda_i < u, \varphi_i >_2^2}{\sum_{i=1}^{\infty} < u, \varphi_i >_2^2}$$

$$
\begin{aligned}
&= \frac{\sum_{i=1}^{n} \lambda_i < u, \varphi_i >_2^2}{\sum_{i=1}^{n} < u, \varphi_i >_2^2} \\
&\leq \lambda_n \frac{\sum_{i=1}^{n} < u, \varphi_i >_2^2}{\sum_{i=1}^{n} < u, \varphi_i >_2^2} = \lambda_n.
\end{aligned}
$$

$$
\Rightarrow \ \exists U \subseteq D(L^*L) \text{ Unterraum mit } \dim U = n: \ \max_{u \in U \setminus \{0\}} \frac{< L^*Lu, u >_2}{< u, u >_2} \leq \lambda_n
$$

$$
\Rightarrow \inf_{\substack{U \subseteq D(L^*L) \text{ Unterraum} \\ \dim U = n}} \max_{u \in U \setminus \{0\}} \frac{< L^*Lu, u >_2}{< u, u >_2} \leq \lambda_n.
$$

Da λ_n kleiner- und größer-gleich diesem Infimum ist wird das Minimum angenommen, man kann also min dafür schreiben. □

Beweis zu Satz (Rayleigh-Ritz Verfahren zur Bestimmung von oberen Eigenwertschranken): Seien $\tilde{u}_n := \sum_{k=1}^{M} x_k^{(n)} v_k$ die Näherungseigenvektoren zu L^*L, die man aus dem Rayleigh-Ritz Verfahren erhält, wobei $x^{(n)}$ die Eigenvektoren des Problems

$$
A_1 x = \tilde{\lambda} A_2 x
$$

mit

$$
(x^i)^t A_2 x^{(j)} = \delta_{ij}, \qquad (x^{(i)})^t A_1 x^{(j)} = \tilde{\lambda}_i \delta ij
$$

seien. Für diese Näherungseigenvektoren gilt:

$$
\begin{aligned}
< \tilde{u}_i, \tilde{u}_j >_2 &= \left\langle \sum_{k=1}^{M} x_k^{(i)} v_k, \sum_{l=1}^{M} x_l^{(j)} v_l \right\rangle_2 \\
&= \sum_{k=1}^{M} \sum_{l=1}^{M} x_k^{(i)} x_l^{(j)} < v_k, v_l >_2 \\
&= (x^{(i)})^t A_2 x^{(j)} = \delta_{ij} \\
< L^*L\tilde{u}_i, \tilde{u}_j >_2 &= \left\langle L^*L \sum_{k=1}^{M} x_k^{(i)} v_k, \sum_{l=1}^{M} x_l^{(j)} v_l \right\rangle_2 \\
&= \sum_{k=1}^{M} \sum_{l=1}^{M} x_k^{(i)} x_l^{(j)} < L^*Lv_k, v_l >_2 \\
&= (x^{(i)})^t A_1 x^{(j)} = \tilde{\lambda}_i \delta_{ij}
\end{aligned}
$$

Man definiert nun für jedes $n \in \{1, \ldots, M\}$: $U := \operatorname{span}\{\tilde{u}_1, \ldots, \tilde{u}_n\}$. Nach Satz (Variationsformulierung der Eigenwerte) gilt:

$$
\begin{aligned}
\lambda_n &\leq \max_{u \in U \setminus \{0\}} \frac{< L^*Lu, u >_2}{< u, u >_2} \\
&= \max_{c \in \mathbb{R}^n \setminus \{0\}} \frac{< L^*L \sum_{k=1}^{M} c_k \tilde{u}_k, \sum_{l=1}^{M} c_l \tilde{u}_l >_2}{< \sum_{k=1}^{M} c_k \tilde{u}_k, \sum_{l=1}^{M} c_l \tilde{u}_l >_2} \\
&= \max_{c \in \mathbb{R}^n \setminus \{0\}} \frac{\sum_{k=1}^{M} \sum_{l=1}^{M} c_k c_l < L^*L\tilde{u}_k, \tilde{u}_l >_2}{\sum_{k=1}^{M} \sum_{l=1}^{M} c_k c_l < \tilde{u}_k, \tilde{u}_l >_2}
\end{aligned}
$$

$$= \max_{c\in\mathbb{R}^n\setminus\{0\}} \frac{\sum_{i=1}^M \tilde{\lambda}_i c_i^2}{\sum_{i=1}^M c_i^2} \le \tilde{\lambda}_n.$$

□

Satz (Lehmann Verfahren zur Bestimmung unterer Eigenwertschranken)

Seien $\tilde{u}_1, \ldots, \tilde{u}_n \in D(L^*L)$ N linear unabhängige Ansatzfunktionen. (Man denke sich dabei die $\tilde{u}_i$ ruhig als Approximationen der exakten Eigenwerte von L^*L). Es existiere nun weiterhin ein $\nu \in \mathbb{R}$, das

$$\tilde{\lambda}_N < \nu \le \lambda_{N+1}$$

erfüllt. Hierbei sei $\tilde{\lambda}_N$ die aus dem Rayleigh-Ritz-Verfahren erhaltene obere Schranke für den N-ten Eigenwert und λ_{N+1} der $N+1$-te exakte Eigenwert von L^*L. Bildet man mit diesen Ansatzfunktionen und der Konstanten ν wieder die bereits im Rayleigh-Ritz-Verfahren aufgetretenen Matrizen

$$\begin{aligned}
A_1 &:= (< L^*L\tilde{u}_i, \tilde{u}_j >_2)_{\substack{i=1,\ldots,N\\ j=1,\ldots,N}} = (< LL\tilde{u}_i, L\tilde{u}_j >_2)_{\substack{i=1,\ldots,N\\ j=1,\ldots,N}},\\
A_2 &:= (< \tilde{u}_i, \tilde{u}_j >_2)_{\substack{i=1,\ldots,N\\ j=1,\ldots,N}}
\end{aligned}$$

und weiterhin die Matrizen

$$\begin{aligned}
A_3 &:= (< L^*L\tilde{u}_i, L^*L\tilde{u}_j >_2)_{\substack{i=1,\ldots,N\\ j=1,\ldots,N}},\\
B_1 &:= A_1 - \nu A_2 = (< (L^*L-\nu)\tilde{u}_i, \tilde{u}_j >_2)_{\substack{i=1,\ldots,N\\ j=1,\ldots,N}},\\
B_2 &:= A_3 - 2\nu A_1 + \nu^2 A_2 = (< (L^*L-\nu)\tilde{u}_i, (L^*L-\nu)\tilde{u}_j >_2)_{\substack{i=1,\ldots,N\\ j=1,\ldots,N}},
\end{aligned}$$

so kann man das Eigenwertproblem

$$B_1 x = \mu B_2 x$$

formulieren. Für dessen N negative Eigenwerte (siehe Bemerkungen) μ_i gilt:

$$\boxed{\lambda_{N+1-n} \ge \nu + \frac{1}{\mu_n} \qquad (n = 1, \ldots, N).}$$

Beweis: Seien $x^{(1)}, \ldots, x^{(N)} \in \mathbb{R}^N$ die Eigenvektoren des Problems $B_1 x = \mu B_2 x$, die zusätzlich

$$(x^{(i)})^t B_2 x^{(j)} = \delta_{ij}, \qquad (x^{(i)})^t B_1 x^{(j)} = \mu_1 \delta_{ij}$$

erfüllen. Definiert man damit

$$u_i := \sum_{k=1}^N x_k^{(i)} \tilde{u}_k \quad (i = 1, \ldots, N),$$

so gilt

$$< (L^*L-\nu)u_i, (L^*L-\nu)u_j >_2 = \left\langle (L^*L-\nu)\sum_{k=1}^N x_k^{(i)}\tilde{u}_k, (L^*L-\nu)\sum_{l=1}^N x_l^{(j)}\tilde{u}_l \right\rangle_2$$

$$
\begin{aligned}
&= \sum_{k=1}^{N}\sum_{l=1}^{N} x_k^{(i)} x_l^{(j)} < (L^*L-\nu)\tilde{u}_k, (L^*L-\nu)\tilde{u}_l >_2 \\
&= x^{(i)} B_2 x^{(j)} = \delta_{ij}, \\
< (L^*L-\nu)u_i, u_j >_2 &= < (L^*L-\nu) \sum_{k=1}^{N} x_k^{(i)}\tilde{u}_k, \sum_{l=1}^{N} x_l^{(j)}\tilde{u}_l >_2 \\
&= \sum_{k=1}^{N}\sum_{l=1}^{N} x_k^{(i)} x_l^{(j)} < (L^*L-\nu)\tilde{u}_k, \tilde{u}_l >_2 \\
&= x^{(i)} B_1 x^{(j)} = \mu_i \delta_{ij}.
\end{aligned}
$$

Sei nun $n \in \{1, \ldots, N\}$ beliebig aber fest. Wähle nun $(c_1, \ldots, c_n)^t \in \mathbb{R}^n \backslash \{0\}$ so, daß $u := \sum_{i=1}^{n} c_i u_i$ den $n-1$ linearen Gleichungen $< u, \varphi_k >_2 = 0$ $(k = N+2-n, \ldots, N)$ genügt. Nach der vorausgesetzten Ungleichung $\tilde{\lambda}_N < \nu \leq \lambda_{N+1}$ gilt:

$$
\begin{aligned}
&\lambda_i - \nu \geq \ \lambda_i - \lambda_{N+1} \geq 0 && \text{für } i \geq N+1, \\
&\lambda_i - \nu < \lambda_i - \tilde{\lambda}_N \overset{\substack{\text{Rayleigh-}\\ \text{Ritz}}}{\leq} 0 && \text{für } i \leq N.
\end{aligned}
$$

Damit erhält man dann:

$$
\begin{aligned}
< (L^*L-\nu)u, u >_2 &= \left\langle (L^*L-\nu) \sum_{i=1}^{\infty} < u, \varphi_i >_2 \varphi_i, u \right\rangle_2 \\
&= \sum_{i=1}^{\infty} < u, \varphi_i >_2 < (L^*L-\nu)\varphi_i, u >_2 \\
&= \sum_{i=1}^{\infty} < u, \varphi_i >_2 < (\lambda_i-\nu)\varphi_i, u >_2 \\
&= \sum_{i=1}^{\infty} (\lambda_i-\nu) < u, \varphi_i >_2^2 \\
&= \sum_{i=1}^{N+2-n-1} (\lambda_i-\nu) < u, \varphi_i >_2^2 + \sum_{i=N+1}^{\infty} \underbrace{(\lambda_i-\nu)}_{\geq 0} \underbrace{< u, \varphi_i >_2^2}_{\geq 0} \\
&\geq \sum_{i=1}^{N+1-n} \frac{(\lambda_i-\nu)^2}{\lambda_i-\nu} < u, \varphi_i >_2^2 \\
&\geq \frac{1}{\lambda_{N+1-n}-\nu} \sum_{i=1}^{N+1-n} (\lambda_i-\nu)^2 < u, \varphi_i >_2^2 \\
&\geq \frac{1}{\lambda_{N+1-n}-\nu} \sum_{i=1}^{\infty} ((\lambda_i-\nu) < u, \varphi_i >_2)^2 \\
&= \frac{1}{\lambda_{N+1-n}-\nu} \sum_{i=1}^{\infty} < u, (L^*L-\nu)\varphi_i >_2^2 \\
&= \frac{1}{\lambda_{N+1-n}-\nu} \sum_{i=1}^{\infty} < (L^*L-\nu)u, \varphi_i >_2^2
\end{aligned}
$$

$$= \frac{1}{\lambda_{N+1-n} - \nu} < (L^*L - \nu)u, (L^*L - \nu)u) >_2 .$$

Andererseits gilt noch:

$$\begin{aligned}
< (L^*L - \nu)u, u >_2 &= \left\langle (L^*L - \nu) \sum_{i=1}^{n} c_i u_i, \sum_{j=1}^{n} c_j u_j \right\rangle_2 \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j < (L^*L - \nu)u_i, u_j >_2 \\
&= \sum_{i=1}^{n} \mu_i c_i^2 \\
&\leq \mu_n \sum_{i=1}^{n} c_i^2 = \mu_n \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \delta_{ij} = \mu_n \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j (x^{(i)})^t B_2 x^{(j)} \\
&= \mu_n \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j < (L^*L - \nu)u_i, (L^*L - \nu)u_j > \\
&= \mu_n \left\langle (L^*L - \nu) \sum_{i=1}^{n} c_i u_i, (L^*L - \nu) \sum_{j=1}^{n} c_j u_j \right\rangle_2 \\
&= \mu_n < (L^*L - \nu)u, (L^*L - \nu)u >_2 .
\end{aligned}$$

Da $(L^*L - \nu)u \neq 0$ ist folgt somit aus den beiden letzten Ungleichungen:

$$\begin{aligned}
\frac{1}{\lambda_{N+1-n} - \nu} < (L^*L - \nu)u, (L^*L - \nu)u) >_2 &\leq < (L^*L - \nu)u, u >_2 \\
&\leq \mu_n < (L^*L - \nu)u, (L^*L - \nu)u >_2
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{\lambda_{N+1-n} - \nu} \leq \mu_n \\
\overset{\text{beide Seiten negativ}}{\Rightarrow} \quad & \lambda_{N+1-n} \geq \nu + \frac{1}{\mu_n}.
\end{aligned}$$

□

Bemerkungen:

1. Die Negativheit der Eigenwerte des Eigenwertproblems $B_1 x = \mu B_2 x$ resultiert aus der negativen Definitheit der Matrix B_1 und der positiven Definitheit der Matrix B_2.

   - Zur negativen Definitheit der Matrix B_1:
     Seien $x^{(i)} \in \mathbb{R}^N$ $(i = 1, \ldots, N)$ die im Rayleigh-Ritz-Verfahren berechneten Eigenvektoren mit

     $$(x^{(i)})^t A_1 x^{(j)} = \tilde{\lambda}_i \delta_{ij}, \qquad (x^{(i)})^t A_2 x^{(j)} = \delta_{ij},$$

     dann gilt

     $$(x^{(i)})^t B_1 x^{(j)} = (x^{(i)})^t A_1 x^{(j)} - \nu (x^{(i)})^t A_2 x^{(j)} = (\tilde{\lambda}_i - \nu)\delta_{ij}.$$

Die Eigenvektoren $x^{(i)}$ bilden eine Basis des $\mathbb{R}^N$. Somit läßt sich jeder Vektor des $\mathbb{R}^N$ darstellen als $v = \sum_{i=1}^N c_i x^{(i)}$, womit man dann erhält:

$$\begin{aligned} vB_1v &= (\sum_{i=1}^N c_i x^{(i)})^t (A_1 - \nu A_2) \sum_{j=1}^N c_j x^{(j)} \\ &= \sum_{i=1}^N \sum_{j=1}^N c_i c_j (\tilde{\lambda}_i - \nu)\delta_{ij} \\ &= \sum_{i=1}^N c_i^2 (\tilde{\lambda}_i - \nu) < 0. \end{aligned}$$

B_1 ist somit negativ definit.

- Zur positiven Definitheit der Matrix B_2:
  Für jedes $x \in \mathbb{R}^N$ gilt

$$\begin{aligned} x^t B_2 x &= \left\langle (L^*L - \nu) \sum_{i=1}^N x_i \tilde{u}_i, (L^*L - \nu) \sum_{j=1}^N x_j \tilde{u}_j \right\rangle_2 \\ &= \left\| (L^*L - \nu) \sum_{i=1}^N x_i \tilde{u}_i \right\|_2^2 \geq 0. \end{aligned}$$

  B_2 ist somit also positiv semidefinit. Nimmt man an es gibt ein $x \in \mathbb{R}^N$ für das $x^t B_2 x = 0$ gilt, so folgt sofort daraus $(L^*L - \nu)(\sum_{i=1}^N x_i \tilde{u}_i) = 0$, was wiederum im Widerspruch zur negativ Definitheit von B_1 steht. B_2 ist somit positiv definit.

2. Wie schon das Rayleigh-Ritz-Verfahren benötigt auch das Verfahren von Lehmann als a priorie Information eine untere Schranke des $N+1$-ten Eigenwertes des Operators L^*L.

3. Die Schranken, die man durch das Lehmann-Verfahren erhält, sind in dem Sinne optimal, daß es bei alleiniger Kenntnis von $\tilde{u}_1, \ldots, \tilde{u}_N$ und $L^*L\tilde{u}_1, \ldots, L^*L\tilde{u}_N$ keine Möglichkeit gibt, bessere untere Schranken zu erhalten, als die durch das Lehmann-Verfahren gegebenen (siehe N.J. Lehmann *Optimale Eigenwerteinschließungen* Numerische Mathematik 5 (1963), 246 - 272).

Homothopie-Methode:

Als Problem bei der Eigenwerteinschließung, egal ob nach der Methode von Weinstein, Kato oder Lehmann, ist nun nur noch die benötigte vorab Kenntnis einer unteren Schanke von λ_{N+1} des Operators L^*L verblieben. Hier soll gezeigt werden, wie man solch eine a priorie Information, ausgehend von einem Vergleichsoperator $L_0^*L_0$, bei dem man Einschließungen für die Eigenwerte kennt, über eine ganze Familie von Operatoren, erhält. Dazu wird auf die sogenannte *Homothopie-Methode*, die von M. Plum und F. Goerisch unabhängig voneinander entwickelt wurde, zurückgegriffen.

Der folgende Satz demonstriert, welche Bedingung benötigt wird, um eine Verbindung zwischen den Eigenwerten der Operatoren $L_s^* L_s$ und $L_t^* L_t$ mit $0 \leq s \leq t \leq 1$ der Operatorenfamilie herstellen zu können.

Satz:

Sei $((H_s, < .,. >_s))_{s \in [0,1]}$ eine Familie von pre-Hilberträumen mit $\dim H_s = \infty$, $(D(L_s^* L_s))_{s \in [0,1]}$ eine Familie von Unterräumen mit $D(L_s^* L_s) \subseteq H_s$ dicht und $(L_s^* L_s)_{s \in [0,1]}$ eine Familie linearer symmetrischer Operatoren $L_s^* L_s : D(L_s^* L_s) \to H_s$. Für jedes $s \in [0,1]$ sollen die Eigenelemente von $L_s^* L_s$ eine Orthonormalbasis von H_s bilden und die zugehörige Eigenwertsequenz $(\lambda_n^s)_{n \in \mathbb{N}}$ gegen ∞ konvergieren. Genügt diese Familie von Hilberträumen und Operatoren den Bedingungen

$$H_s \supseteq H_t, \quad D(L_s^* L_s) \supseteq D(L_t^* L_t),$$
$$\frac{< L_s u, L_s u >_s}{< u, u >_s} \leq \frac{< L_t u, L_t u >_t}{< u, u >_t} \quad \forall u \in D(L_t^* L_t),\ 0 \leq s \leq t \leq 1,$$

so gilt

$$\boxed{\lambda_n^s \leq \lambda_n^t \quad \forall n \in \mathbb{N} \quad 0 \leq s \leq t \leq 1}.$$

Beweis: Man benutzt die Variationsformulierung der Eigenwerte, um λ_n^s darzustellen, und schätzt dann ab:

$$\begin{aligned}
\lambda_n^s &= \min_{\substack{U \subseteq D(L_s^* L_s) \\ \dim U = n}} \max_{u \in U \setminus \{0\}} \frac{< L_s u, L_s u >_s}{< u, u >_s} \\
&\leq \min_{\substack{U \subseteq D(L_t^* L_t) \\ \dim U = n}} \max_{u \in U \setminus \{0\}} \frac{< L_s u, L_s u >_s}{< u, u >_s} \\
&\leq \min_{\substack{U \subseteq D(L_t^* L_t) \\ \dim U = n}} \max_{u \in U \setminus \{0\}} \frac{< L_t u, L_t u >_t}{< u, u >_t} \\
&= \lambda_n^t.
\end{aligned}$$

□

Dieser Satz zeigt nun, wie man ausgehend von einem Vergleichsoperator $L_0^* L_0$, dessen Eigenwerteinschließungen man kennt, über einige Schritte mit Operatoren $L_{s_i}^* L_{s_i}$ man sich schließlich bis zum Operator $L_1^* L_1 := L^* L$ durchhangeln kann und dabei jeweils Einschließungen für die Eigenwerte der Zwischenschrittoperatoren berechnen kann, da man ja die untere Schranke von $\lambda_n^{s_i}$ auch als a priorie benötigte untere Schranke für $\lambda_n^{s_{i+1}}$ verwenden kann. Der folgende Algorithmus zeigt detailliert, wie man dazu vorzugehen hat:

Der Homothopie-Algorithmus:

1. Man bestimme eine Familie von Vergleichsräumen $((H_s, < .,. >_s))_{s \in [0,1]}$ und Vergleichsoperatoren $(L_s^* L_s)_{s \in [0,1]}$ mit $L_s^* L_s : D(L_s^* L_s) \subseteq H_s \to H_s$, die den Voraussetzungen des letzten Satzes genügen. Dabei

sei der Operator $L_0^* L_0$ so gewählt, daß Einschließungen für seine Eigenwerte bekannt sind ($\lambda_n^0 \in [\underline{\lambda}_n^0, \overline{\lambda}_n^0]$) und $L_1^* L_1 := L^* L$ gesetzt. Man setze $i = 0$ und $s_i = 0$. Man starte jetzt den Algorithmus mit N_i Eigenwerteinschließungen für den Operator $L_{s_i}^* L_{s_i}$, die ja als bekannt vorausgesetzt wurden. Dabei ist N_i so zu wählen, daß erstens $\lambda_{N_i}^{s_i}$ und $\lambda_{N_i-1}^{s_i}$ gut separiert sind und zweitens $\lambda_N < \lambda_{N_i}^{s_i}$ mit nicht zu kleiner Differenz gilt (Die Bedingung $\lambda_N < \lambda_{N_i}^{s_0}$ ist auf jeden Fall erfüllt, wenn statt ihrer die Bedingung $\tilde{\lambda}_N < \underline{\lambda}_{N_i}^{s_0}$, für die man nur eine bekannte untere Grenze von $\lambda_{N_i}^{s_0}$ und eine mit dem Rayleigh-Ritz Verfahren leicht beschaffte Obergrenze für λ_N kennen muß, verwendet wird. Solch ein N_i zu finden ist immer möglich, da nach den Voraussetzungen die Eigenwertsequenz gegen Unendlich konvergiert).

2. Übergang von Eigenwerteinschließungen für den Operator $L_{s_i}^* L_{s_i}$ zu Eigenwerteinschließungen für den Operator $L_{s_{i+1}}^* L_{s_{i+1}}$ mit $s_{i+1} > s_i$: Man suche $s_{i+1} \in [0, 1]$, so daß gilt:

$$\lambda_{N_i-1}^{s_{i+1}} < \lambda_{N_i}^{s_i}.$$

   Dabei wähle man s_{i+1} möglichst nahe an $\sup\{s \in [0,1] : \lambda_{N_i-1}^{s} < \lambda_{N_i}^{s_i}\}$ (solch ein s_{i+1} existiert, wenn die Eigenwerte in s stetig sind, oder nur kleine Sprünge machen, da N_i gerade so gewählt war, daß $\lambda_{N_i}^{s_i}$ und $\lambda_{N_i-1}^{s_i}$ gut separiert sind. Das Erfülltsein der Bedingung $\lambda_{N_i-1}^{s_{i+1}} < \lambda_{N_i}^{s_i}$ kann man, wie schon unter 1. erwähnt, leicht überprüfen, indem man stattdesen die Bedingung $\tilde{\lambda}_{N_i-1}^{s_{i+1}} < \underline{\lambda}_{N_i}^{s_i}$ verwendet, wobei $\underline{\lambda}_{N_i}^{s_i}$ eine bekannt untere Schranke von $\lambda_{N_i}^{s_i}$ ist und $\tilde{\lambda}_{N_i-1}^{s_{i+1}}$ eine nach dem Rayleigh-Ritz-Verfahren berechnete obere Schranke für $\lambda_{N_i-1}^{s_{i+1}}$ darstellt). Da alle Voraussetzungen des vorherigen Satzes erfüllt sind, gilt dann:

$$\lambda_{N_i}^{s_i} \leq \lambda_{N_i}^{s_{i+1}}.$$

   Man kann also, indem man $\nu = \underline{\lambda}_{N_i}^{s_i}$ setzt, eine untere Schranke für den N_i-ten Eigenwert des Operators $L_{s_{i+1}}^* L_{s_{i+1}}$ erhalten. Dieses ν kann man dann dazu verwenden, um mit dem Verfahren von Weinstein, Kato oder Lehmann Einschließungen für die ersten $N_i - 1$ Eigenwerte von $L_{s_{i+1}}^* L_{s_{i+1}}$ zu gewinnen. Man suche nun $N_{i+1} \leq N_i - 1$ möglichst groß, so daß die Eigenwerte $\lambda_{N_{i+1}}^{s_{i+1}}$ und $\lambda_{N_{i+1}-1}^{s_{i+1}}$ gut separiert sind.

3. Wenn $s_{i+1} = 1$, dann ist der Algorithmus zu Ende und man hat Einschließungen für $L_1^* L_1 = L^* L$ gewonnen. (Hat man dabei aber nur $\tilde{N} < N$ Eigenwerte eingeschlossen, so war das gewählte N_0 zu klein, und man muß den Algorithmus noch einmal mit einem $\tilde{N}_0 > N_0$ von vorne starten). Ist dies nicht der Fall, so erhöhe man i um eins und wiederhole Schritt 2. des Algorithmus.

Bemerkungen:

- Die verlangte Kenntnis von Eigenwerteinschließungen für $L_0^* L_0$ kann z.B. dadurch erreicht werden, daß der Operator so einfach ist, daß seine Eigenwerte in geschlossener Form bestimmt werden können.

- Bei der praktischen Durchführung des Algorithmus ist es oft möglich, $(H_0, <.,.>_0) := (H, <.,.>_2)$ und $D(L_0^*L_0) := D(L^*L)$ zu wählen. In diesem Fall kann die Konstruktion der Familie von Vergleichsoperatoren und Vergleichsräumen auf die folgende einfache Weise durchgeführt werden:

$$(H_s, <.,.>_s) := (H, <.,.>_2), \quad D(L_s^*L_s) := D(L^*L),$$
$$L_s^*L_s := (1-s)\cdot L_0^*L_0 + s\cdot L^*L \quad \forall s \in [0,1].$$

Konstruktion einer Familie von Vergleichsoperatoren:

Im letzten Abschnitt wurde gezeigt wie man, wenn man Einschließungen für die Eigenwerte eines Vergleichsoperators kennt, über eine Familie von Operatoren schließlich zu den gesuchten Einschließungen der Eigenwerte des Operators L^*L kommt. Hier nun soll gezeigt werden, wie man solch einen Vergleichsoperator zu konstruieren hat, das er einerseits die für die Monotonie der Eigenwerte notwendige Bedingung

$$\frac{< L^*Lu, u >_2}{< u, u >_2} \geq \frac{< L_0^*L_0u, u >_2}{< u, u >_2} := \frac{B_0[u,u]}{< u, u >_2} \quad \forall u \in D(L^*L)$$

erfüllt und andererseits so einfach ist, daß seine Eigenwerte expliziet berechenebar sind. Dazu wird zuerst das Innenprodukt $< L^*Lu, u >_2$ solange nach unten abgeschätzt, bis man zu einer Bilinearform $B_0[u,u]$ mit konstanten Koeffizienten gelangt. Dann betrachtet man die zu dieser Bilinearform gehörige starke Eigenwertformulierung, die wegen der konstanten Koeffizienten expliziet lösbar ist. Man kann somit Einschließungen für die Eigenwerte des Operators angeben, der zu dieser Bilinearform gehört. Auf die gleiche Art und Weise wie in den Bemerkungen zum letzten Satz, nämlich durch setzen von $B_s[u,v] := (1-s)B_0[u,v] + s < Lu, Lv >_2$, erhält man eine ganze Familie von Bilinearformen mit assozierten Operatoren, die jeweils paarweise der Monotoniebedingung genügen. Man hat also, das Problem der Konstruktion einer Familie von Vergleichsoperatoren mit einfachem Operator $L_0^*L_0$ gelöst.

Wendet man sich nun dem Problem des Abschätzens der Bilinearform $< L^*Lu, u >_2$ zu so findet man durch einfaches ausmultiplizieren:

$$\begin{aligned}
< L^*Lu, u >_2 &=< Lu, Lu >_2 \\
&=< -\ddot{u} + B\dot{u} + Cu, -\ddot{u} + B\dot{u} + Cu >_2 \\
&=< \ddot{u}, \ddot{u} >_2 \underbrace{- < \ddot{u}, B\dot{u} >_2}_{(A1)} \underbrace{- < \ddot{u}, Cu >_2}_{(B_1)} \\
&\quad \underbrace{- < B\dot{u}, \ddot{u} >_2}_{(A_2)} + < B\dot{u}, B\dot{u} >_2 \underbrace{+ < B\dot{u}, Cu >_2}_{(C_1)} \\
&\quad \underbrace{- < Cu, \ddot{u} >_2}_{(B_2)} \underbrace{+ < Cu, B\dot{u} >_2}_{(C_2)} + < Cu, Cu >_2 .
\end{aligned}$$

Das Innenprodukt zerfällt also in 9 einzelne Innenprodukte, von denen die unterklammerten noch weiter untersucht werden sollen. Zuerst wird hier die

Summe aus $(A1)$ und $(A2)$ näheren Betrachtungen unterzogen:

$$\begin{aligned}(A_1)+(A_2) &= -\int_0^{k\frac{2\pi}{\mu}} (\ddot{u}^t B\dot{u} + \dot{u}^t B^t \ddot{u})\,dt \\ &= -2\int_0^{k\frac{2\pi}{\mu}} \ddot{u}^t B\dot{u}\,dt \\ &\geq -2\int_0^{k\frac{2\pi}{\mu}} |\ddot{u}|_e |B|_e |\dot{u}|_e\,dt \\ &\geq -\int_0^{k\frac{2\pi}{\mu}} |B|_e(\varepsilon_1 \ddot{u}^t\ddot{u} + \tfrac{1}{\varepsilon_1}\dot{u}^t\dot{u})\,dt.\end{aligned}$$

Auch die Termen (B_1) und (B_2) werden zuerst zusammengefaßt, danach wendet man aber am besten noch partielle Integration an, um einen weiteren Term bei $\ddot{u}$ zu verhindern:

$$\begin{aligned}(B_1)+(B_2) &= -\int_0^{k\frac{2\pi}{\mu}} (\ddot{u}^t C u + u^t C^t \ddot{u})\,dt \\ &= -2\int_0^{k\frac{2\pi}{\mu}} \ddot{u}^t C u\,dt \\ &= -2\left(\underbrace{\dot{u}^t C u|_0^{k\frac{2\pi}{\mu}}}_{=0} - \int_0^{k\frac{2\pi}{\mu}} \dot{u}^t \dot{C} u\,dt - \int_0^{k\frac{2\pi}{\mu}} \dot{u}^t C\dot{u}\,dt\right) \\ &\geq -2\int_0^{k\frac{2\pi}{\mu}} |\dot{u}|_e|\dot{C}|_e|u|_e\,dt + 2\int_0^{k\frac{2\pi}{\mu}} \dot{u}^t C\dot{u}\,dt \\ &\geq -\int_0^{k\frac{2\pi}{\mu}} |\dot{C}|_e(\varepsilon_2\dot{u}^t\dot{u} + \tfrac{1}{\varepsilon_2}u^t u)\,dt + 2\int_0^{k\frac{2\pi}{\mu}} \dot{u}^t C\dot{u}\,dt.\end{aligned}$$

Bei der Addition der Term (C_1) und (C_2) kann man wieder wie bei den Termen $(A1)$ und $(A2)$ vorgehen und erhält dann:

$$\begin{aligned}(C_1)+(C_2) &= \int_0^{k\frac{2\pi}{\mu}} (\dot{u}^t B^t C u + u^t C^t B\dot{u})\,dt \\ &= 2\int_0^{k\frac{2\pi}{\mu}} \dot{u}^t B^t C u\,dt \\ &\geq -2\int_0^{k\frac{2\pi}{\mu}} |\dot{u}|_e|B^t C|_e|u|_e\,dt \\ &\geq -\int_0^{k\frac{2\pi}{\mu}} |B^t C|_e(\varepsilon_3\dot{u}^t\dot{u} + \tfrac{1}{\varepsilon_3}u^t u)\,dt.\end{aligned}$$

Nimmt man all diese Abschätzungen zusammen und schätzt weiter ab, so erhält man:

$$< L^*Lu, u >_2 \geq \int_0^{k\frac{2\pi}{\mu}} \ddot{u}^t \underbrace{((1-\varepsilon_1|B|_e)I)}_{=:M_2} \ddot{u}\,dt$$

$$
\begin{aligned}
&+\int_0^{k\frac{2\pi}{\mu}} \dot{u}^t \underbrace{(B^tB + 2C - (\tfrac{1}{\varepsilon_1}|B|_e + \varepsilon_2|\dot{C}|_e + \varepsilon_3|B^tC|_e)I)}_{=:M_1} \dot{u}\, dt \\
&+\int_0^{k\frac{2\pi}{\mu}} u^t \underbrace{(C^tC - (\tfrac{1}{\varepsilon_2}|\dot{C}|_e + \tfrac{1}{\varepsilon_3}|B^tC|_e)I)}_{=:M_0} u\, dt \\
=&\int_0^{k\frac{2\pi}{\mu}} (\ddot{u}^t \tfrac{1}{2}(M_2 + M_2^t)\ddot{u} + \dot{u}^t \tfrac{1}{2}(M_1 + M_1^t)\dot{u} + u^t \tfrac{1}{2}(M_2 + M_2^t)u)\, dt \\
\geq&\int_0^{k\frac{2\pi}{\mu}} \tfrac{1}{2}\lambda_{\min}(M_2 + M_2^t)\ddot{u}^t\ddot{u}\, dt \\
&+\int_0^{k\frac{2\pi}{\mu}} \tfrac{1}{2}\lambda_{\min}(M_1 + M_1^t)\dot{u}^t\dot{u}\, dt + \int_0^{k\frac{2\pi}{\mu}} \tfrac{1}{2}\lambda_{\min}(M_2 + M_2^t)u^t u\, dt \\
\geq&\int_0^{k\frac{2\pi}{\mu}} \underbrace{\tfrac{1}{2}\min_{t\in[0,k\frac{2\pi}{\mu}]} \lambda_{\min}(M_2 + M_2^t)}_{=:\alpha}\, \ddot{u}^t\ddot{u}\, dt \\
&+\int_0^{k\frac{2\pi}{\mu}} \underbrace{\tfrac{1}{2}\min_{t\in[0,k\frac{2\pi}{\mu}]} \lambda_{\min}(M_1 + M_1^t)}_{=:\beta}\, \dot{u}^t\dot{u}\, dt \\
&+\int_0^{k\frac{2\pi}{\mu}} \underbrace{\tfrac{1}{2}\min_{t\in[0,k\frac{2\pi}{\mu}]} \lambda_{\min}(M_2 + M_2^t)}_{=:\gamma}\, u^t u\, dt \\
=&\int_0^{k\frac{2\pi}{\mu}} \alpha\ddot{u}^t\ddot{u} + \beta\dot{u}^t\dot{u} + \gamma u^t u\, dt
\end{aligned}
$$

Setzt man also

$$B_0[u,v] := \int_0^{k\frac{2\pi}{\mu}} \alpha\ddot{u}^t\ddot{v} + \beta\dot{u}^t\dot{v} + \gamma u^t v\, dt,$$

so gilt für diese Bilinearform die für ein Vergleichsproblem benötigte Abschätzung

$$\frac{< L^*Lu, u >_2}{< u, u >_2} \geq \frac{B_0[u,u]}{< u, u >_2} \quad \forall u \in D(L^*L)$$

Für die Eigenwerte des schwach formulierten Eigenwertproblems

$$< L_0^*L_0u, \varphi >_2 := B_0[u,\varphi] = \lambda^0 < u, \varphi >_2 \quad \forall \varphi \in D(L^*L) \tag{Sch}$$

erhält man somit $\lambda_i^0 \leq \lambda_i \ \forall i \in \mathbb{N}$. Durch partielle Integration kann man nun die zu dieser schwachen Formulierung der Eigenwertaugabe gehörende starke gewinnen:

$$< L_0^*L_0u, \varphi >_2 := B_0[u,\varphi] = \lambda^0 < u, \varphi >_2$$

$$\int_0^{k\frac{2\pi}{\mu}} \alpha\ddot{u}^t\ddot{\varphi} + \beta\dot{u}^t\dot{\varphi} + \gamma u^t\varphi\, dt = \lambda^0 \int_0^{k\frac{2\pi}{\mu}} u^t\varphi\, dt$$

$$\int_0^{k\frac{2\pi}{\mu}} -\alpha \ddot{u}^t \dot{\varphi} - \beta \ddot{u}^t \varphi + \gamma u^t \varphi \, dt = \lambda^0 \int_0^{k\frac{2\pi}{\mu}} u^t \varphi \, dt$$

$$\int_0^{k\frac{2\pi}{\mu}} (\alpha \dddot{u}^t - \beta \ddot{u}^t + \gamma u^t) \varphi \, dt = \lambda^0 \int_0^{k\frac{2\pi}{\mu}} u^t \varphi \, dt$$

$$L_0^* L_0 u := \alpha \dddot{u}^t - \beta \ddot{u}^t + \gamma u^t = \lambda^0 u. \tag{St}$$

Die Eigenwerte der stark formulierten Eigenwertaufgabe (St) lassen sich aber, da die Koeffizienten α, β und γ konstante reelle Zahlen sind, auf einfache Art expliziet berechnen. Um nun eine ganze Familie von Bilinearformen $B_s[u,v]$ mit zugehörigen Operatoren $L_s^* L_s$ zu erhalten muß man nun nur noch, wie in der Einleitung dieses Abschnittes bereites erwähnt

$$< L_s^* L_s u, v >:= B_s[u,v] := (1-s) B_0[u,v] + s < Lu, Lv >_2 \qquad \forall s \in [0,1]$$

setzen. Für $s = 0$ erhält man dann das expliziet berechenbare Eigenwertproblem und für $s = 1$ das Eigenwertproblem, für das Einschließungen gesucht werden. Außerdem ist die Monotonie der Eigenwerte im Parameter s garantiert, da $B_s[u,u] \geq B_t[u,u]$ für alle $1 \geq s \geq t \geq 0$ gilt. Die Konstruktion einer Familie von Operatoren bzw. Bilinearformen, die von einem einfachen Vergleichsproblem zu $L^* L$ führen ist somit vollzogen und das Problem eine Konstante σ als untere Schranke des kleinsten Eigenwerts von $L^* L$ zu finden vollständig gelöst.

1.3.4 Zu Punkt 4: Berechnung der Konstanten K

Auch für die Berechnung der Konstanten K gibt es, wie schon bei den vorherigen Punkten, mehrere Möglichkeiten der Bestimmung. Hier soll nur eine Möglichkeit vorgestellt werden, die man leicht auch auf elliptische Randwertprobleme übertragen kann. Die numerische Hauptaufgabe bei dieser Möglichkeit besteht darin, eine Konstante σ zu finden mit

$$0 < \sigma \leq \lambda \qquad \forall \lambda \text{ Eigenwert von } L^* L \text{ auf } R.$$

Dieser Punkt wurde bereits im letzten Abschnitt besprochen. Es sei noch einmal erwähnt, verwendet man ein anderes als das hier beschriebene Verfahren um die Konstante K zu berechnen, so kann man Punkt 3 überspringen, da die dort berechnete Konstante σ nur in diesem Punkt Verwendung findet. Hat man nun eine solche Konstante σ gefunden, so helfen die folgenden drei Sätze Konstanten K_0, K_1, C_0 und C_1 zu finden, die für alle $u \in R$ die folgenden Bedingungen erfüllen:

$$||u||_2 \leq K_0 ||L[u]||_2,$$
$$||\dot{u}||_2 \leq K_1 ||L[u]||_2 \text{ und}$$
$$||u||_\infty \leq C_0 ||u||_2 + C_1 ||\dot{u}||_2.$$

Die Abschätzung $||u||_\infty \leq K ||L[u]||_2 \; \forall u \in R$ (siehe (1.4)) erhält man dann offensichtlich in dem man

$$\boxed{K := K_0 C_0 + K_1 C_1}$$

setzt.

Satz (Bestimmung der Konstanten K_0):

Unter der Voraussetzung, daß es eine Konstante σ gibt, für die gilt $0 < \sigma \leq \lambda$ für alle Eigenwerte λ von L^*L auf R gilt

$$\boxed{||u||_2 \leq K_0 ||L[u]||_2 \ \forall u \in R \quad \text{mit } K_0 := \frac{1}{\sqrt{\sigma}}}.$$

Beweis: Nach dem Satz (Variationsformulierung der Eigenwerte) (siehe Seite 21) gilt:

$$\begin{aligned} 0 < \sigma \leq \lambda_1 = \min_{\substack{u \in R \\ u \neq 0}} \frac{< L[u], L[u] >_2}{< u, u >_2} & \\ \Rightarrow \ 0 < \sigma \leq \frac{< L[u], L[u] >_2}{< u, u >_2} = \frac{||L[u]||_2^2}{||u||_2^2} & \qquad \forall u \in R \end{aligned}$$

$$\Rightarrow \ ||u||_2 \leq \underbrace{\frac{1}{\sqrt{\sigma}}}_{=:K_0} ||L[u]||_2 \quad \forall u \in R.$$

□

Satz (Bestimmung der Konstanten K_1):

Unter den Voraussetzungen $B^t = B$ und $B(0) = B(k\frac{2\pi}{\mu})$ an den linearisierten Operator

$$L[u] = -\ddot{u} + B\dot{u} + \left(\frac{\partial F_i}{\partial y_j}\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} u$$

(siehe (1.5)) des nichtlinearen Problems erhält man mit

$$\gamma := \frac{1}{2} \min_{t \in [0, k\frac{2\pi}{\mu}]} \lambda_{\min} \left[\left(\frac{\partial F_i}{\partial y_j}\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} + \left(\frac{\partial F_i}{\partial y_j}\right)^t_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} - \dot{B} \right]$$

und

$$\nu := \begin{cases} K_0 - \gamma K_0^2 & \text{falls } \gamma K_0 < \frac{1}{2} \\ \frac{1}{4\gamma} & \text{falls } \gamma K_0 \geq \frac{1}{2} \end{cases}$$

die Abschätzung

$$\boxed{||\dot{u}||_2 \leq \underbrace{\sqrt{\nu}}_{=:K_1} ||L[u]||_2}.$$

Beweis: Zur Verkürzung der Schreibweise setze man $C := \left(\frac{\partial F_i}{\partial y_j}\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ wie in (1.5). Dann läßt sich der lineare Operator $L[u]$ schreiben als

$$L[u] = -\ddot{u} + B\dot{u} + Cu,$$

wobei B und C jeweils $n \times n$-Matrizen und $\ddot{u}$, $\dot{u}$ und u Vektoren aus $\mathbb{R}^n$ sind. Bildet man das Innenprodukt zwischen u und $L[u]$, so läßt sich dieses in drei Terme aufspalten, von denen die ersten beiden weiter untersucht werden.

$$\Rightarrow < u, L[u] >_2 = \underbrace{< u, -\ddot{u} >_2}_{(A)} + \underbrace{< u, B\dot{u} >_2}_{(B)} + < u, Cu >_2 \tag{1.13}$$

Die Terme (A) und (B) lassen sich umformen zu

$$\begin{aligned}
(A) &= -\int_0^{k\frac{2\pi}{\mu}} u^t \ddot{u}\, dt \\
&= \underbrace{-u^t \dot{u}\Big|_0^{k\frac{2\pi}{\mu}}}_{=0} + \int_0^{k\frac{2\pi}{\mu}} \dot{u}^t \dot{u}\, dt \\
&= ||\dot{u}||_2^2,
\end{aligned}$$

$$\begin{aligned}
(B) &= \int_0^{k\frac{2\pi}{\mu}} u^t B \dot{u}\, dt \\
&= \underbrace{u^t B u\Big|_0^{k\frac{2\pi}{\mu}}}_{=0} - \int_0^{k\frac{2\pi}{\mu}} \underbrace{\dot{u}^t B u}_{\substack{=u^t B^t \dot{u} \\ =u^t B \dot{u}}}\ dt - \int_0^{k\frac{2\pi}{\mu}} u^t \dot{B} u\, dt \\
&= -\int_0^{k\frac{2\pi}{\mu}} u^t B \dot{u}\, dt - \int_0^{k\frac{2\pi}{\mu}} u^t \dot{B} u\, dt \\
&= -(B) - < u, \dot{B}u >_2
\end{aligned}$$

$$\Rightarrow\ (B) = -\tfrac{1}{2} < u, \dot{B}u >_2 .$$

Setzt man Ausdruck (A) und (B) in Gleichung (1.13) ein, so erhält man:

$$\begin{aligned}
< u, L[u] >_2 &= ||\dot{u}||_2^2 - \tfrac{1}{2} < u, \dot{B}u >_2 + < u, Cu >_2 \\
&= ||\dot{u}||_2^2 + \underbrace{< u, \underbrace{(C - \tfrac{1}{2}\dot{B})}_{=:M} u >_2}_{(C)} .
\end{aligned} \tag{1.14}$$

Ausdruck (C) kann man abschätzen durch:

$$\begin{aligned}
(C) &= < u, Mu >_2 \overset{< u, Mu >_2 \in \mathbb{R}}{=} < Mu, u >_2 = < u, M^t u >_2 \\
\Rightarrow\ 2(C) &= < u, Mu >_2 + < u, M^t u >_2 \\
\Rightarrow\ (C) &= < u, \tfrac{1}{2}(M + M^t) u >_2 \\
&\geq \min_{t \in [0, k\frac{2\pi}{\mu}]} \lambda_{\min}[\tfrac{1}{2}(M + M^t)] < u, u >_2 \\
&= \underbrace{\min_{t \in [0, k\frac{2\pi}{\mu}]} \lambda_{\min}[\tfrac{1}{2}(C + C^t - \dot{B})]}_{=\gamma} ||u||_2^2.
\end{aligned}$$

Setzt man (C) in (1.14) ein, so erhält man:

$$< u, L[u] >_2 \geq ||\dot{u}||_2^2 + \gamma ||u||_2^2.$$

Nach der Cauchy-Schwarzschen-Ungleichung gilt aber auch:

$$< u, L[u] >_2 \leq ||u||_2 ||L[u]||_2.$$

Verwendet man jeweils die rechte Seite der letzten beiden Ungleichungen, so ergibt sich:

$$||u||_2 ||L[u]||_2 \geq ||\dot{u}||_2^2 + \gamma ||u||_2^2 \quad \Leftrightarrow \quad ||\dot{u}||_2^2 \leq ||u||_2 ||L[u]||_2 - \gamma ||u||_2^2.$$

Setzt man $\delta := \frac{||u||_2}{||L[u]||_2}$, so gilt nach der Wahl von K_0 ($||u||_2 \leq K_0 ||L[u]||_2$) $0 \leq \delta \leq K_0$, und man kann die obige Ungleichung umschreiben zu

$$||\dot{u}||_2^2 \leq ||L[u]||_2^2 (\delta - \gamma\delta^2) \leq ||L[u]||_2^2 \underbrace{\max_{0 \leq \delta \leq K_0} (\delta - \gamma\delta^2)}_{=:\nu}$$

$$\Rightarrow \; ||\dot{u}||_2 \leq \sqrt{\nu} ||L[u]||_2.$$

Bestimmung von $\nu := \max_{0 \leq \delta \leq K_0} (\delta - \gamma\delta^2)$:

- Fall 1: $\gamma \leq 0 \quad \Rightarrow \; \nu = K_0 - \gamma K_0^2$.
- Fall 2: $\gamma > 0$
  - Fall 2.1: $K_0 < \frac{1}{2\gamma} \quad \Rightarrow \; \nu = K_0 - \gamma K_0^2$.
  - Fall 2.2: $K_0 \geq \frac{1}{2\gamma} \quad \Rightarrow \; \nu = \frac{1}{4\gamma}$.

Fall 1 und Fall 2.1 liefern den gleichen Wert für ν und man kann sie daher zu einem einzigen Fall zusammenfassen. Tut man dies, so erhält man genau das Resultat des Satzes. □

Satz (Bestimmung der Konstanten C_0, C_1):

Für eine $k\frac{2\pi}{\mu}$ periodische Funktion $u \in H_1^n(0, k\frac{2\pi}{\mu})$ gibt es Konstanten C_0 und C_1, so daß gilt:

$$||u||_\infty \leq C_0 ||u||_2 + C_1 ||\dot{u}||_2.$$

Dabei ist $C_0 \geq \sqrt{\frac{\mu}{2\pi k}}$ beliebig wählbar und C_1 so zu wählen, daß $C_1 \geq \sqrt{\frac{1}{12}\frac{1}{C_0}}$ gilt.

Beweis: Zunächst wird die Aussage dieses Satzes für Funktionen $v \in H_1(0, k\frac{2\pi}{\mu})$, die $k\frac{2\pi}{\mu}$ periodisch sind, bewiesen und danach gezeigt, wie man dieses Resultat auf $k\frac{2\pi}{\mu}$ periodische Funktionen $u \in H_1^n(0, k\frac{2\pi}{\mu})$ verallgemeinern kann. Wird in diesem Beweis die Funktion v außerhalb des Intervalls $[0, k\frac{2\pi}{\mu}]$ ausgewertet (wo sie ja eigentlich nicht definiert ist), so verwende man die Funktionswerte, die man erhält, wenn man v über das Intervall $[0, k\frac{2\pi}{\mu}]$ hinaus periodisch fortsetzt. Es sei nun also zunächst $v \in H_1(0, k\frac{2\pi}{\mu})$ beliebige $k\frac{2\pi}{\mu}$ periodische Funktion und $C_0 \geq \sqrt{\frac{\mu}{2\pi k}}$ beliebig aber fest. Dann wähle man für jedes beliebige aber feste $t \in [0, k\frac{2\pi}{\mu}]$ zwei reelle Zahlen α und β, so daß einerseits

$$0 < \beta - \alpha = \frac{1}{C_0^2} \leq k\tfrac{2\pi}{\mu}$$

und andererseits auch

$$t := \frac{1}{2}(\alpha + \beta)$$

gilt. Dies ist immer möglich, da $\det \begin{pmatrix} -1 & 1 \\ 0.5 & 0.5 \end{pmatrix} \neq 0$ ist. Durch einfaches Differenzieren und darauf angewandte Integration erhält man:

$$\begin{aligned}
\underbrace{(t-\alpha)}_{=\frac{1}{2}\frac{1}{C_0^2}>0} |v(t)| &= \left| \int_\alpha^t \frac{\partial}{\partial \bar{t}} [(\bar{t}-\alpha)v(\bar{t})] \, d\bar{t} \right| \\
&= \left| \int_\alpha^t v(\bar{t}) \, d\bar{t} + \int_\alpha^t (\bar{t}-\alpha)\dot{v}(\bar{t}) \, d\bar{t} \right| \\
&\leq \left| \int_\alpha^t v(\bar{t}) \, d\bar{t} \right| + \left| \int_\alpha^t (\bar{t}-\alpha)\dot{v}(\bar{t}) \, d\bar{t} \right| \\
&\leq \int_\alpha^t |v(\bar{t})| \, d\bar{t} + \int_\alpha^t \underbrace{(\bar{t}-\alpha)}_{\geq 0} |\dot{v}(\bar{t})| \, d\bar{t}. \qquad (1.15)
\end{aligned}$$

Auf die selbe Art und Weise erhält man:

$$\underbrace{(\beta - t)}_{=\frac{1}{2}\frac{1}{C_0^2}>0} |u(t)| \leq \int_t^\beta |v(\bar{t})| \, d\bar{t} + \int_t^\beta (\beta - \bar{t})|\dot{v}(\bar{t})| \, d\bar{t}. \qquad (1.16)$$

Addiert man (1.15) und (1.16), so ergibt sich:

$$\begin{aligned}
(\beta-\alpha)|v(t)| &\leq \int_\alpha^\beta 1 \cdot |v(\bar{t})| \, d\bar{t} + \int_\alpha^t (\bar{t}-\alpha)|\dot{v}(\bar{t})| \, d\bar{t} + \int_t^\beta (\beta-\bar{t})|\dot{v}(\bar{t})| \, d\bar{t} \\
&\overset{\text{CSU}}{\leq} \sqrt{\int_\alpha^\beta 1^2 \, d\bar{t}} \sqrt{\int_\alpha^\beta v^2(\bar{t}) \, d\bar{t}} \\
&\quad + \sqrt{\int_\alpha^t (\bar{t}-\alpha)^2 \, d\bar{t}} \sqrt{\int_\alpha^t \dot{v}^2(\bar{t}) \, d\bar{t}} \\
&\quad + \sqrt{\int_t^\beta (\beta-\bar{t})^2 \, d\bar{t}} \sqrt{\int_t^\beta \dot{v}^2(\bar{t}) \, d\bar{t}} \\
&= \sqrt{\beta-\alpha} \sqrt{\int_\alpha^\beta v^2(\bar{t}) \, d\bar{t}} \\
&\quad + \sqrt{\tfrac{(t-\alpha)^3}{3}} \sqrt{\int_\alpha^t \dot{v}^2(\bar{t}) \, d\bar{t}} + \sqrt{\tfrac{(\beta-t)^3}{3}} \sqrt{\int_t^\beta \dot{v}^2(\bar{t}) \, d\bar{t}}
\end{aligned}$$

$$
\begin{aligned}
&\stackrel{\beta-\alpha\le k\frac{2\pi}{\mu}}{\le} && \sqrt{\beta-\alpha}||v||_2 + \begin{pmatrix} \sqrt{\frac{(t-\alpha)^3}{3}} \\ \sqrt{\frac{(\beta-t)^3}{3}} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{\int_\alpha^t \dot v^2(\bar t)\, d\bar t} \\ \sqrt{\int_t^\beta \dot v^2(\bar t)\, d\bar t} \end{pmatrix} \\
&\stackrel{\text{CSU}}{\le} && \sqrt{\beta-\alpha}||v||_2 + \sqrt{\frac{(t-\alpha)^3+(\beta-t)^3}{3}} \sqrt{\int_\alpha^t \dot v^2(\bar t)\, d\bar t + \int_t^\beta \dot v^2(\bar t)\, d\bar t} \\
&\stackrel{t=\frac12(\alpha+\beta)}{=} && \sqrt{\beta-\alpha}||v||_2 + \sqrt{\tfrac{1}{12}(\beta-\alpha)^3} \sqrt{\int_\alpha^\beta \dot v^2(\bar t)\, d\bar t} \\
&\stackrel{\beta-\alpha\le k\frac{2\pi}{\mu}}{\le} && \sqrt{\beta-\alpha}||v||_2 + \sqrt{\tfrac{1}{12}(\beta-\alpha)^3}||\dot v||_2
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \ \forall t \in [0, k\tfrac{2\pi}{\mu}] : \ |v(t)| &\le \frac{1}{\sqrt{\beta-\alpha}}||v||_2 + \sqrt{\tfrac{1}{12}}\sqrt{\beta-\alpha}||\dot v||_2 \\
&= C_0||v||_2 + \sqrt{\tfrac{1}{12}\tfrac{1}{C_0}}||\dot v||_2 \\
\Rightarrow \ ||v||_\infty &\le C_0||v||_2 + \sqrt{\tfrac{1}{12}\tfrac{1}{C_0}}||\dot v||_2 .
\end{aligned}
$$

Sei nun $u \in H_1^n(0, k\frac{2\pi}{\mu})$ eine beliebige $k\frac{2\pi}{\mu}$ periodische Funktion, so gilt natürlich nach obiger Überlegung für jede Komponente von u:

$$
||u_i||_\infty \le C_0||u_i||_2 + \sqrt{\tfrac{1}{12}\tfrac{1}{C_0}}||\dot u_i||_2 \quad 1 \le i \le n
$$

$$
\begin{aligned}
\Rightarrow \ ||u||_\infty &= \max_{1\le i\le n} ||u_i||_\infty \\
&\le \max_{1\le i\le n} (C_0||u_i||_2 + \sqrt{\tfrac{1}{12}\tfrac{1}{C_0}}||\dot u_i||_2) \\
&\le \max_{1\le i\le n} (C_0||u||_2 + \sqrt{\tfrac{1}{12}\tfrac{1}{C_0}}||\dot u||_2) \\
&= C_0||u||_2 + \sqrt{\tfrac{1}{12}\tfrac{1}{C_0}}||\dot u||_2 .
\end{aligned}
$$

□

1.3.5 Zu Punkt 5: Finden einer monoton wachsenden Funktion G

In diesem Abschnitt werden zwei Sätze vorgestellt, die beide eine monoton wachsende Funktion G liefern, die die Abschätzung

$$
||F(.,\omega+y) - F(.,\omega) - Cy||_2 \le G(||y||_\infty) \quad (\forall y \in \mathbb{R}^n)
$$

gewährleistet (siehe Seite 3). Der Unterschied zwischen beiden Sätzen besteht darin, daß der erste Satz nur erste Ableitungen von F benötigt, wohingegen der zweite auch zweite Ableitungen von F heranzieht. Der Vorteil des zweiten Satzes besteht darin, daß dadurch daß $\alpha^2 := ||y||_\infty^2$ ausgeklammert werden kann, für kleines α der Rest des Funktionsausdruckes für G in Intervallarithmetik nur sehr grob berechnet werden muß. Nachteilig wirkt auf den zweiten Satz, wenn etwa die zweiten Ableitungen von F sehr große Werte annehmen (etwa bei beinahe Sprüngen in der ersten Ableitung). Manchmal existiert auch gar keine zweite Ableitung von F. In diesen beiden Fällen ist man dann genötigt, Satz eins anzuwenden.

Satz:

Setzt man

$$G(\alpha) := \alpha \sqrt{\int_0^{k\frac{2\pi}{\mu}} \sum_{i=1}^n \left(\max_{||y||_\infty \leq \alpha} \sum_{j=1}^n |D^j F_i(t,\omega(t)+y) - D^j F_i(t,\omega(t)|\right)^2 dt},$$

so gilt die gewünschte Abschätzung (1.6)

$$||F(.,\omega+y) - F(.,\omega) - Cy||_2 \leq G(||y||_\infty) \quad (\forall y \in \mathbb{R}^n)$$

und weiterhin $G(\alpha) = o(\alpha)$ für $\alpha \to 0$. (Mit D^j sei dabei die erste Ableitung nach der j-ten Raumkomponente bezeichnet).

Beweis: Schreibt man den Term, den man durch $G(\alpha)$ abschätzen möchte ausführlich, so erhält man

$$\begin{aligned}
&||F(.,\omega+y) - F(.,\omega) - Cy||_2 \\
&= \Big(\int_0^{k\frac{2\pi}{\mu}} \sum_{i=1}^n \underbrace{\Big(F_i(t,\omega(t)+y) - F_i(t,\omega(t)) - \sum_{j=1}^n D^j F_i(t,\omega(t)) \cdot y_j\Big)^2}_{(a)} dt\Big)^{\frac{1}{2}}.
\end{aligned}$$

Durch einfaches Umformen des Termes (a) erhält man

$$\begin{aligned}
(a) &= F_i(t,\omega(t)+y) - F_i(t,\omega(t)) - \sum_{j=1}^n D^j F_i(t,\omega(t)) \cdot y_j \\
&= F_i(t,\omega(t)+s\cdot y)|_{s=0}^{s=1} - \sum_{j=1}^n D^j F_i(t,\omega(t)) \cdot y_j \\
&= \int_0^1 \sum_{j=1}^n D^j F_i(t,\omega(t)+s\cdot y) \cdot y_j \, ds - \sum_{j=1}^n D^j F_i(t,\omega(t)) \cdot y_j \\
&= \int_0^1 \sum_{j=1}^n \big(D^j F_i(t,\omega(t)+s\cdot y) - D^j F_i(t,\omega(t))\big) \cdot y_j \, ds.
\end{aligned}$$

Versucht man nun den Term $|(a)|$ für $||y||_\infty \leq \alpha$ abzuschätzen, so erhält man

$$\begin{aligned}
|(a)| &= \left| \int_0^1 \sum_{j=1}^n \big(D^j F_i(t,\omega(t)+s\cdot y) - D^j F_i(t,\omega(t))\big) \cdot y_j \, ds \right| \\
&\leq \int_0^1 \left| \sum_{j=1}^n \big(D^j F_i(t,\omega(t)+s\cdot y) - D^j F_i(t,\omega(t))\big) \cdot y_j \right| ds \\
&\leq \int_0^1 \sum_{j=1}^n \big|D^j F_i(t,\omega(t)+s\cdot y) - D^j F_i(t,\omega(t))\big| \cdot \underbrace{|y_j|}_{\leq \alpha} \, ds
\end{aligned}$$

$$\leq \alpha \int_0^1 \sum_{j=1}^{n} \left| D^j F_i(t,\omega(t)+s\cdot y) - D^j F_i(t,\omega(t)) \right| \, ds$$

$$\leq \alpha \max_{0\leq s\leq 1} \sum_{j=1}^{n} \left| D^j F_i(t,\omega(t)+s\cdot y) - D^j F_i(t,\omega(t)) \right|$$

$$\leq \alpha \max_{||y||_\infty \leq \alpha} \sum_{j=1}^{n} \left| D^j F_i(t,\omega(t)+y) - D^j F_i(t,\omega(t)) \right| .$$

Setzt man dies in unsere erste Gleichung ein, so erhält man die gesuchte Beziehung:

$$||F(.,\omega+y) - F(.,\omega) - Cy||_2$$

$$\leq \alpha \sqrt{\int_0^{k\frac{2\pi}{\mu}} \sum_{i=1}^{n} \left(\max_{||y||_\infty\leq\alpha} \sum_{j=1}^{n} |D^j F_i(t,\omega(t)+y) - D^j F_i(t,\omega(t)| \right)^2 dt} =: G(\alpha)$$

für alle y mit $||y||_\infty \leq \alpha$. $\square$

Bemerkungen:

- Wie schon bei der Bestimmung der Konstante δ (siehe 1.3.2 auf Seite 10) muß auch hier der Wert des Integrals in gesicherten Grenzen bestimmt werden. Von den dort beschriebenen zwei Möglichkeiten der exakten Integration bietet sich hier das Verfahren der Riemanschen Ober- und Untersummen an, da die für Integrationsformeln mit Fehlerabschätzung mindestens benötigte 2te Ableitung des Integranten, wenn überhaupt vorhanden nur sehr schwer berechenbar sein dürfte.

- Zur Bechnung des im Integranten auftretenden Maximums über den n-dimensionalen Quader $||y||_\infty \leq \alpha$. kann man wiederum Intervallarithmetik verwenden. Dazu unterteilt man den Quader in mehrere möglichst gleich große Quader Y_l und berechnet dann mit Intervallrechnung für jedes l eine Obergrenze für den Intervallausdruck

  $$\bigoplus_{j=1}^{n}{}_{\mathrm{M}} \mathrm{ABS_M} \left((D^j F_i)_{\mathrm{M}}(T,\omega_{\mathrm{M}}(T)+Y_l) - (D^j F_i)_{\mathrm{M}}(T,\omega_{\mathrm{M}}(T)) \right)$$

  mit $(D^j F_i)_{\mathrm{M}}$ als intervallmäßige Auswertung von $D^j F_i$ und $\omega_{\mathrm{M}}(T)$ als intervallmäßige Auswertung von $\omega(t)$. Danach muß man nur noch das Maximum über diese Maxima bilden.

Satz:

Setzt man

$$G(\alpha) := \alpha^2 \sqrt{\int_0^{k\frac{2\pi}{\mu}} \sum_{i=1}^{n} \left(\max_{||y||_\infty\leq\alpha} \sum_{j=1}^{n} \sum_{k=1}^{n} |D^{jk} F_i(t,\omega(t)+y)| \right)^2 dt},$$

so gilt die gewünschte Abschätzung (1.6):

$$||F(.,\omega+y)-F(.,\omega)-Cy||_2 \le G(||y||_\infty) \quad (\forall y \in \mathbb{R}^n)$$

und weiterhin $G(\alpha)=o(\alpha)$ für $\alpha \to 0$. (Mit D^{jk} sei dabei die zweite Ableitung nach den Raumkomponenten j und k bezeichnet).

Beweis: Wie im Beweis des vorherigen Satzes, kann man durch Umformen mit Hilfe des Mittelwertsatzes der Integralrechnung erhalten, daß

$$(a)=\int_0^1 \sum_{j=1}^n \left(D^j F_i(t,\omega(t)+s\cdot y)-D^j F_i(t,\omega(t))\right)\cdot y_j\, ds$$

gilt. Erneutes Anwenden des Mittelwertsatzes der Integralrechung liefert

$$\begin{aligned}
(a) &= \int_0^1 \sum_{j=1}^n \left(D^j F_i(t,\omega(t)+s\cdot y)-D^j F_i(t,\omega(t))\right)\cdot y_j\, ds\\
&= \int_0^1 \sum_{j=1}^n D^j F_i(t,\omega(t)+u\cdot s\cdot y)\Big|_{u=0}^{u=1}\cdot y_j\, ds\\
&= \int_0^1 \sum_{j=1}^n \int_0^1 \sum_{k=1}^n D^{jk} F_i(t,\omega(t)+u\cdot s\cdot y)\cdot y_k\, du\cdot y_j\, ds\\
&= \int_0^1 \int_0^1 \sum_{j=1}^n \sum_{k=1}^n D^{jk} F_i(t,\omega(t)+u\cdot s\cdot y)\cdot y_k\cdot y_j\, du\, ds.
\end{aligned}$$

Versucht man nun den Term $|(a)|$ für $||y||_\infty \le \alpha$ abzuschätzen, so erhält man

$$\begin{aligned}
|(a)| &= \left|\int_0^1 \int_0^1 \sum_{j=1}^n \sum_{k=1}^n D^{jk} F_i(t,\omega(t)+u\cdot s\cdot y)\cdot y_k\cdot y_j\, du\, ds\right|\\
&\le \int_0^1 \int_0^1 \sum_{j=1}^n \sum_{k=1}^n \left|D^{jk} F_i(t,\omega(t)+u\cdot s\cdot y)\right|\cdot \underbrace{|y_k|}_{\le\alpha}\cdot \underbrace{|y_j|}_{\le\alpha}\, du\, ds\\
&\le \alpha^2 \max_{||y||_\infty\le\alpha} \sum_{j=1}^n \sum_{k=1}^n \left|D^{jk} F_i(t,\omega(t)+y)\right|.
\end{aligned}$$

Setzt man die gefundene Beziehung in die erste Gleichung aus dem letzten Beweis ein, so sieht man:

$$\begin{aligned}
&||F(.,\omega+y)-F(.,\omega)-Cy||_2\\
&\le \alpha^2 \sqrt{\int_0^{k\frac{2\pi}{\mu}} \sum_{i=1}^n \left(\max_{||y||_\infty\le\alpha} \sum_{j=1}^n \sum_{k=1}^n |D^{jk} F_i(t,\omega(t)+y)|\right)^2 dt} =: G(\alpha)
\end{aligned}$$

für alle y mit $||y||_\infty \le \alpha$. □

Bemerkung:
Für diesen Satz gelten die schon zum letzten Satz gemachten Bemerkungen ebenfalls, man muß ausschließlich den dort angegebenen Intervallausdruck durch

$$\mathop{\Phi_{\mathrm{M}}}_{j=1}^{n} \mathop{\Phi_{\mathrm{M}}}_{k=1}^{n} \mathrm{ABS}_{\mathrm{M}} \Big((D^{jk} F_i)_{\mathrm{M}}(T, \omega_{\mathrm{M}}(T) + Y_l) \Big)$$

ersetzen.

1.3.6 Zu Punkt 6: Bestimmung der Konstanten α

Im letzten Punkt verbleibt noch die Aufgabe eine Konstante α zu bestimmen, die der Bedingung

$$\delta \leq \frac{\alpha}{K} - G(\alpha),$$

mit der Konstanten K aus Punkt 4 und der monoton wachsenden Funktion G aus Punkt 5, genügt. Diese Bedingung kann durch Ausklammern von α umgeformt werden zu

$$\delta \leq \alpha \Big(\frac{1}{K} - \underbrace{\frac{G(\alpha)}{\alpha}}_{\substack{\alpha \to 0 \\ \to 0}} \Big).$$

(Daß $\frac{G(\alpha)}{\alpha} \stackrel{\alpha \to 0}{\to} 0$ gilt, ergibt sich gerade aus der Eigenschaft $G(\alpha) = o(\alpha)$ für $\alpha \to 0$). Aus der Umschreibung der ursprünglichen Bedingung sieht man aber deutlich, daß das Geforderte für hinreichend kleines δ, also hinreichend genau berechnete Näherungslösung ω, immer erfüllbar ist. Wobei das Wort hinreichend klein konkret bedeutet, daß der Defekt δ gerade kleiner als $\sup(\frac{\alpha}{K} - G(\alpha))$ zu sein hat. Punkt 6 liefert somit keine weitere Forderungen, sondern ist unter der natürlich erscheinenden Voraussetzung δ klein, immer erfüllbar. Abbildung 1.1 verdeutlicht die Situation noch einmal graphisch. Hat man einen Defekt (A), so konnte die Näherungslösung nicht ausreichend genau bestimmt werden und man erhält keine Einschließung, wohingegen bei einem Vorliegen von Defekt (B) eine ganze Reihe von Einschließungen möglich sind. Die beste Einschließung der exakten Lösung erhält man gerade dadurch, daß man $\min\{\alpha : \delta_{(B)} \leq \frac{\alpha}{K} - G(\alpha)\}$ bildet.

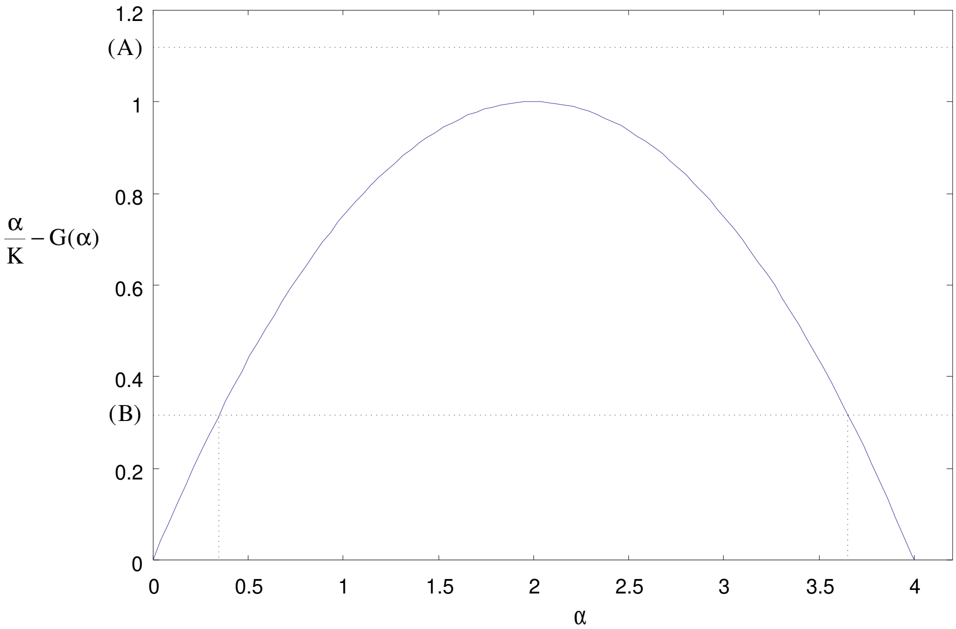

Abbildung 1.1: Beispielverlauf der Funktion $\frac{\alpha}{K} - G(\alpha)$ zur Verdeutlichung der Bedingung $\delta \leq \frac{\alpha}{K} - G(\alpha)$

Kapitel 2

Modell einer Hängebrücke:

In diesem Abschnitt wird ein möglichst einfaches Modell für eine Hängebrücke entworfen, um daran das Schwingungsverhalten bei periodischer Anregung, z.B. durch Wind, untersuchen zu können. Ein einfaches Modell bedeutet, daß möglichst viele Vereinfachungen vorgenommen werden sollen, die Charakteristiken der realen schwingenden Brücke sich in dem Modell aber dennoch wiederspiegeln sollen.

Dieses Modell wird dann auf vielfälltige Weise in Kapitel 3 weiteruntersucht. Zum einen werden für die verschiedenen Parameter des Modelles Parameterstudien in Form von Verzweigungsdiagrammen vorgestellt. Einzelne Bewegungsbilder von Lösungen des Modells sind, um deren Verschiedenheit zu demonstrieren ebenfalls vorhanden. Zum anderen ist aber auch der in Kapitel 1 vorgestellten Algorithmus zur gesicherten Existenz und Einschließung von periodischen Systemen von Randwertproblemen beispielhaft auf zwei Lösungen dieses Modells angewandt, dort zu finden.

2.1 Annahmen und Vereinfachungen:

Als erste Vereinfachung gegenüber einer realen Hängebrücke, denke man sich deren gesammte Massen vereinigt in einem Punkte. Zur weiteren Simplifizierung bestehe in diesem Modell die Aufhängung der Brückenmasse nur aus zwei Brückenträgern, die mit der Masse jeweils über ein Kabel verbunden seien. Die Kabel verhalten sich unter Ausdehnung wie einfache Hooksche-Federn (die Zugkraft ist der Kabelverlängerung proportional). Bei Zusammenziehen üben die Kabel aber, anders als Federn, auf die Masse keinerlei Kraft aus. Weiterhin seien die beiden Kabel in Bezug auf ihre Federkonstanten als identisch vorausgesetzt. Die Brückenseile weisen also ein asymmetrisches Steifigkeitsverhalten auf.

Auf die Masse wirkt die Schwerkraft, die sie nach unten zu einem Gleichgewichtspunkt zieht, bei dem sich Schwerkraft und die Rückstellkräfte der gedehnten Kabel das Gleichgewicht halten. Es soll dann, mit Hilfe des Modells, die Antwort des Systems auf vertikale periodische Kraftanregung, wie sie etwa durch Wind oder durch eine marschierende Soldatenkolone gegeben ist, untersucht werden. Dabei wirke weiterhin eine der Bewegung entgegengesetzte Dämpfung. Abbildung 2.1 soll helfen das in Worte gefaßte Modell

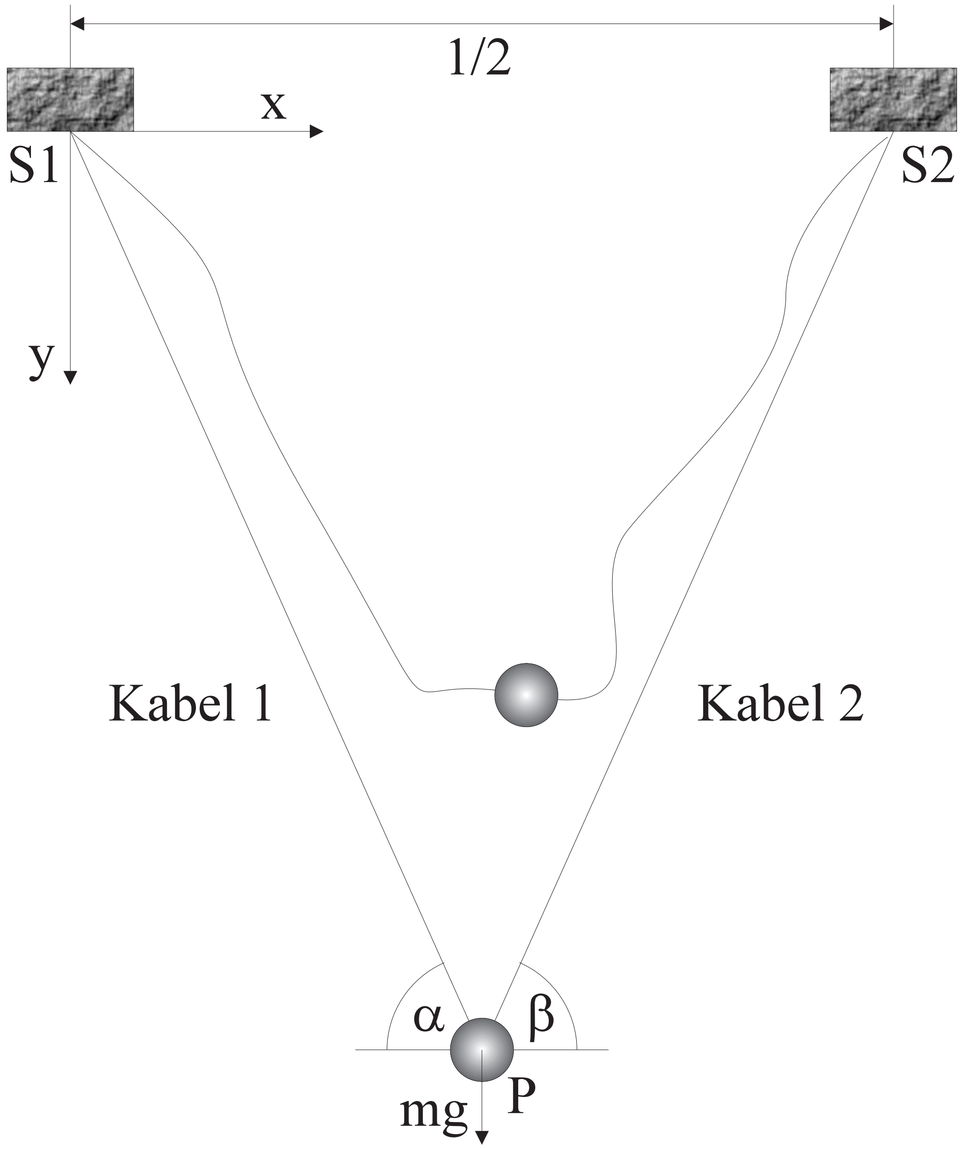

Abbildung 2.1: Einfaches Modell einer Hängebrücke: Die Punktmasse wird von zwei Kabeln, die an Brückenträgern (S1 und S2) befestigt sind, gehalten. Die Abbildung zeigt zwei mögliche Zustände des Systems bei Anregung durch eine vertikale Kraft ($\lambda \sin(\mu t)$). Der erste Zustand zeigt die Masse an einem tiefen Punkt, so daß die Kabel gespannt sind und sich wie Hooksche-Federn verhalten. Beim zweiten Zustand hängen die Kabel lose durch und üben keine Kräfte auf die Punktmasse aus.

zu verdeutlichen.

In der Nähe des Gleichgewichtspunktes, bei nur schwacher vertikaler Anregung und in Bezug zu den beiden Brückenträgern symmetrischen Anfangsbedingungen (Bewegung der Masse exakt in der Mitte zwischen den Aufhängungspunkten, keine horizontale Bewegung) verhält sich die Bewegung annähernd linear und besteht aus der erwarteten Auf- und Abbewegung. Die Lösung wird nur sehr schwache vertikale Auslenkungen haben und auch weiterhin symmetrisch zu den Brückenträgern bleiben.

Eine Frage, die sich dann natürlich sofort stellt: Wie verhalten sich Lösungen, wenn man ihnen keine symmetrischen Anfangsbedingungen mitgibt, sondern ihnen die Bewegung in der ganzen zweidimensionalen Brückenträgerebene erlaubt? Existieren nur vertikale Schwingungen oder gibt es bei gleicher Kraftanregung auch zweidimensionale periodische Bewegungen?

2.2 Die beschreibenden Gleichungen:

Hier werden nun die Gleichungen, die die Bewegung der Punktmasse (als Modell für die Hängebrücke) beschreiben, aufgestellt. Dazu werden ausschließlich zwei Anwendungen des Newtonschen-Gesetzes benötigt:

$$\frac{d}{dt}(m\vec{v}) = \vec{F},$$

wobei m die Masse des Körpers, $\vec{v}$ dessen Geschwindigkeit und $\vec{F}$ die an ihm angreifenden Kräfte bezeichnet. Da bei dem betrachteten Modell die Masse über die Bewegung hinweg konstant ist, wird das Newtonsche-Gesetz in der Form $m\dot{v} = m\ddot{x} = F$ für jede Koordinatenrichtung angewendet.

Wie Abbildung 2.1 zeigt, ist die Punktmasse P über zwei identische masselose Kabel mit den unbeweglichen Brückenträgern S_1 und S_2 verbunden. Die beiden Kabel verhalten sich, wie schon zuvor beschrieben: Bei Ausdehnung wie lineare Hooksche-Federn, ihre Zugkraft ist also proportional zur Ausdehnung. Im unverlängerten Zustand (also etwa bei Durchhängen der Kabel) üben diese keine Kraft auf die Punktmasse aus. Bei Verwendung des in Abbildung 2.1 dargestellten Koordinatensystems, hat der Punkt S_1 die Koordinaten $(0,0)$ und Punkt S_2 die Koordinaten $(\frac{1}{2},0)$. An der Punktmasse greifen die vertikale periodische Erregerkraft, die Gewichtskraft und eine der Bewegung entgegengesetzte Dämpfungskraft, die sich der Geschwindigkeit proportional verhalten soll, an. Die Position der Punktmasse wird in Abhängigkeit der Zeit t durch $(x(t), y(t))$ beschrieben. Es ergeben sich dann die beiden folgenden Gleichungen:

$$\begin{aligned}
m\ddot{x} &= -k(\overline{S_1P} - l_0)^+ \cos\alpha + k(\overline{S_2P} - l_0)^+ \cos\beta - c\dot{x}, \\
m\ddot{y} &= -k(\overline{S_1P} - l_0)^+ \sin\alpha - k(\overline{S_2P} - l_0)^+ \sin\beta - c\dot{y} + mg + \lambda\sin(\mu t),
\end{aligned}$$

wobei $\overline{S_1P}$ und $\overline{S_2P}$ die Distanzen der Punktmasse zum jeweiligen Träger, l_0 die Länge des ungedehnten Kabels, k die Federkonstante des Kabels, α und β die in der Zeichnung angegebenen Winkel jeweils zwischen dem ersten und zweiten Kabel und der x-Achse, m die Masse, g die Erdbeschleunigung, c den Dämpfungskoeffizienten, λ ein Parameter für die Stärke der vertikalen Erregerkraft und μ die Frequenz der Erregerkraft bezeichnet. Weiterhin sei die Notation $(.)^+$ definiert durch:

$$u^+ := \begin{cases} u & u > 0 \\ 0 & u \leq 0 \end{cases}$$

Solche Nichtlinearität taucht in den Gleichungen wegen des unterschiedlichen Verhaltens der Trägerkabel bei Auseinanderziehen und Zusammenstauchen auf. Ersetzt man die Längen und Winkelbeziehungen in den obigen beiden Gleichungen durch Terme in x und y erhält man schließlich:

$$\begin{aligned}
m\ddot{x} = &-k\left(\sqrt{x^2+y^2}-l_0\right)^+ \frac{x}{\sqrt{x^2+y^2}} \\
&+k\left(\sqrt{(\tfrac{1}{2}-x)^2+y^2}-l_0\right)^+ \frac{\tfrac{1}{2}-x}{\sqrt{(\tfrac{1}{2}-x)^2+y^2}} \\
&-c\dot{x}, \\
m\ddot{y} = &-k\left(\sqrt{x^2+y^2}-l_0\right)^+ \frac{y}{\sqrt{x^2+y^2}} \\
&-k\left(\sqrt{(\tfrac{1}{2}-x)^2+y^2}-l_0\right)^+ \frac{y}{\sqrt{(\tfrac{1}{2}-x)^2+y^2}} \\
&-c\dot{y}+mg+\lambda\sin(\mu t).
\end{aligned}$$

Dies sind zwei gekoppelte, nicht autonome, nichtlineare Differentialgleichungen zweiter Ordnung, die die Bewegung der Punktmasse in zwei Dimensionen beschreiben. Setzt man für $x(t) = \frac{1}{4}$, so ist die erste Differentialgleichung identisch erfüllt und man erhält nur noch eine Differentialgleichung für $y(t)$. Diese Lösung steht in Verbindung mit der zu beiden Brückenträgern symmetrischen Lösung. Wählt man jedoch unsymmetrische Anfangswerte, so kann die Lösung durchaus auch unsymmetrisch sein.

Da es in Bezug auf die Untersuchungen des Modelles mit Hilfe des in Kapitel 1 gegebenen Algorithmus von Vorteil sein wird, eine stetig differenzierbare Nichtlinearität vorliegen zu haben, wird fortan aber das Problem betrachtet, das man erhält, falls man statt der oben angegebenen Funktion $(.)^+$, die folgende stetig differenzierbare Definition verwendet.

$$u^+ := \frac{u}{2}+\sqrt{\left(\frac{u}{2}\right)^2+\varepsilon}.$$

Man sieht, für $\varepsilon = 0$ stimmt das zweite Modell mit dem ersten überein. Für $\varepsilon > 0$ klein erhält man aber eine gewisse Glättung, was nichts anderes heißen soll, als daß die Ableitung von $(.)^+$ nun keinen Sprung mehr aufweist, sondern stetig ist. Es mag auf den ersten Blick relativ willkürlich erscheinen, die Funktion $(.)^+$ auf diese Art zu definieren, aber die zuvor vorgestellte Funktion $(.)^+$ ist ebenso wie die neue nur eine Näherung der Wirklichkeit. Kein Seil dieser Welt verhält sich tatsächlich wie eine ideale Hooksche Feder und übt bei zusammengezogenem Zustand keinerlei Kraft auf die Masse aus. Allein dadurch, daß nicht alle Fasern eines Tragseiles exakt gleich lang sind, werden immer zuerst nur einige Fasern Zug ausüben, bis dann schließlich alle Fasern an der Brückenmasse ziehen. Man hat also einen kontinuierlichen Übergang zwischen keiner Ausübung eines Zugs und voller Ausübung des Zugs nach dem Hookschen Gesetz, also keinen Knick. Insofern ist das gewählte Modell vielleicht sogar ein Stück näher an der Realität, da ein glatterer Übergang von Kraftausübung zu keiner Kraftausübung auf die Masse gegeben ist.

2.3 Eigenschaften der Lösung:

Hat man Lösungen, etwa auf numerischem Wege, für das Modell der Hängebrücke gefunden, so ist es möglich, aus diesen auf rein analytischem Wege weitere Lösungen zu gewinnen und damit ein weitaus umfangreicheres Verzweigungsdiagramm zu erhalten, als nur mit Hilfe der Numerik. Die Vorgehensweise beschreiben die beiden folgenden Lemmata:

2.3.1 Lemma (Zeitsymmetrie):

Ist $\begin{pmatrix} x \\ y \end{pmatrix}$ eine Lösung des Modellproblems, so ist auch $\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}(t) := \begin{pmatrix} x \\ y \end{pmatrix}(t + i\frac{2\pi}{\mu})$ mit $i \in \mathbb{Z}$ eine Lösung.

Beweis: Der Beweis erfolgt durch einfaches einsetzen:

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{\tilde{y}} \end{pmatrix}(t) = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}(t + i\tfrac{2\pi}{\mu}) \qquad \begin{pmatrix} \ddot{\tilde{x}} \\ \ddot{\tilde{y}} \end{pmatrix}(t) = \begin{pmatrix} \ddot{x} \\ \ddot{y} \end{pmatrix}(t + i\tfrac{2\pi}{\mu})$$

Setzt man $s := t + i\frac{2\pi}{\mu}$, so erhält man:

$$\begin{aligned}
m\ddot{\tilde{x}} &= -k(\sqrt{\tilde{x}^2+\tilde{y}^2} - l_0)^+ \frac{\tilde{x}}{\sqrt{\tilde{x}^2+\tilde{y}^2}} \\
&\quad +k(\sqrt{(\tfrac{1}{2}-\tilde{x})^2+\tilde{y}^2} - l_0)^+ \frac{\frac{1}{2}-\tilde{x}}{\sqrt{(\frac{1}{2}-\tilde{x})^2+\tilde{y}^2}} - c\dot{\tilde{x}}, \\
m\ddot{x}(s) &= -k(\sqrt{x^2(s)+y^2(s)} - l_0)^+ \frac{x(s)}{\sqrt{x^2(s)+y^2(s)}} \\
&\quad +k(\sqrt{(\tfrac{1}{2}-x(s))^2+y^2(s)} - l_0)^+ \frac{\frac{1}{2}-x(s)}{\sqrt{(\frac{1}{2}-x(s))^2+y^2(s)}} - c\dot{x}(s),
\end{aligned}$$

$$\begin{aligned}
m\ddot{\tilde{y}} &= -k(\sqrt{\tilde{x}^2+\tilde{y}^2} - l_0)^+ \frac{\tilde{y}}{\sqrt{\tilde{x}^2+\tilde{y}^2}} \\
&\quad -k(\sqrt{(\tfrac{1}{2}-\tilde{x})^2+\tilde{y}^2} - l_0)^+ \frac{\tilde{y}}{\sqrt{(\frac{1}{2}-\tilde{x})^2+\tilde{y}^2}} - c\dot{\tilde{y}} + mg + \lambda\sin(\mu t) \\
m\ddot{y}(s) &= -k(\sqrt{x^2(s)+y^2(s)} - l_0)^+ \frac{y(s)}{\sqrt{x^2(s)+y^2(s)}} \\
&\quad -k(\sqrt{(\tfrac{1}{2}-x(s))^2+y^2(s)} - l_0)^+ \frac{y(s)}{\sqrt{(\frac{1}{2}-x(s))^2+y^2(s)}} \\
&\quad -c\dot{y}(s) + mg + \lambda\sin(\mu s).
\end{aligned}$$

$\square$

2.3.2 Lemma (Raumsymmetrie):

Ist $\begin{pmatrix} x \\ y \end{pmatrix}$ eine Lösung des Modellproblems, so ist auch $\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}(t) := \begin{pmatrix} \frac{1}{2} - x \\ y \end{pmatrix}(t)$ eine Lösung.

Beweis: Der Beweis erfolgt durch einfaches einsetzen:

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{\tilde{y}} \end{pmatrix}(t) = \begin{pmatrix} -\dot{x} \\ \dot{y} \end{pmatrix}(t) \qquad \begin{pmatrix} \ddot{\tilde{x}} \\ \ddot{\tilde{y}} \end{pmatrix}(t) = \begin{pmatrix} -\ddot{x} \\ \ddot{y} \end{pmatrix}(t)$$

$$\begin{aligned}
m\ddot{\tilde{x}} &= -k(\sqrt{\tilde{x}^2 + \tilde{y}^2} - l_0)^+ \frac{\tilde{x}}{\sqrt{\tilde{x}^2 + \tilde{y}^2}} \\
&\quad + k(\sqrt{(\tfrac{1}{2} - \tilde{x})^2 + \tilde{y}^2} - l_0)^+ \frac{\frac{1}{2} - \tilde{x}}{\sqrt{(\frac{1}{2} - \tilde{x})^2 + \tilde{y}^2}} - c\dot{\tilde{x}}, \\
-m\ddot{x} &= -k(\sqrt{(\tfrac{1}{2} - x)^2 + y^2} - l_0)^+ \frac{\frac{1}{2} - x}{\sqrt{(\frac{1}{2} - x)^2 + y^2}} \\
&\quad + k(\sqrt{(\tfrac{1}{2} - (\tfrac{1}{2} - x))^2 + y^2} - l_0)^+ \frac{\frac{1}{2} - (\frac{1}{2} - x)}{\sqrt{(\frac{1}{2} - (\frac{1}{2} - x))^2 + y^2}} + c\dot{x},
\end{aligned}$$

$$\begin{aligned}
m\ddot{\tilde{y}} &= -k(\sqrt{\tilde{x}^2 + \tilde{y}^2} - l_0)^+ \frac{\tilde{y}}{\sqrt{\tilde{x}^2 + \tilde{y}^2}} \\
&\quad - k(\sqrt{(\tfrac{1}{2} - \tilde{x})^2 + \tilde{y}^2} - l_0)^+ \frac{\tilde{y}}{\sqrt{(\frac{1}{2} - \tilde{x})^2 + \tilde{y}^2}} - c\dot{\tilde{y}} + mg + \lambda \sin(\mu t) \\
m\ddot{y} &= -k(\sqrt{(\tfrac{1}{2} - x)^2 + y^2} - l_0)^+ \frac{y}{\sqrt{(\frac{1}{2} - x)^2 + y^2}} \\
&\quad - k(\sqrt{(\tfrac{1}{2} - (\tfrac{1}{2} - x))^2 + y^2} - l_0)^+ \frac{y}{\sqrt{(\frac{1}{2} - (\frac{1}{2} - x))^2 + y^2}} \\
&\quad - c\dot{y} + mg + \lambda \sin(\mu t).
\end{aligned}$$

Wenn man nun die erste Gleichung mit -1 multipliziert und die Ausdrücke $\frac{1}{2} - (\frac{1}{2} - x)$ zu x vereinfacht und einzelne Terme umstellt, erhält man die Behauptung. □

2.4 Verfeinerungsvorschläge zum einfachen Modell der Hängebrücke:

Beim Studium der Literatur, die sich mit dem Modellieren von schwingenden Hängebrücken beschäftigt, fällt auf, daß es dafür ganz unterschiedliche Ideen und Ansatzpunkte gibt. In den folgenden Abschnitten wurde nun versucht einige diese Ideen so umzuformen, daß sie auf das im letzten Abschnitt vorgestellte Modell angewandt werden können. Der nächste Abschnitt z.B. sucht nach einer Antwort auf die Frage, wie man den Energieverlust beim Straffen der Seile beschreiben kann. Der darauf folgende Abschnitt untersucht wie nicht mehr statische, sondern selbst schwingende Brückenaufhängungen, Einfluß auf das Modell nehmen.

2.4.1 Hinzufügen einer asymmetrischen Dissipation:

Annahmen und Voraussetzungen:

In diesem Abschnitt soll das einfache Modell einer Hängebrücke, das ja asymmetrische Steifigkeitseigenschaften aufweist (Tragseile verhalten sich unter Zug wie Hooksche Federn, bei Durchhängen üben sie hingegen keinerlei Kraft aus), noch um eine asymmetrische Dissipation ergänzt werden. Dabei soll dem von P. Vielsack und H. Wei in [6] vorgeschlagenen Weg gefolgt werden.

Das ursprüngliche Modell besitzt ja bereits eine kleine viskose Dämpfung, die aber unabhängig von der Größe und der Schwingweite der Bewegungsform ist. Die Erweiterung soll nun darin bestehen für größere Schwingweiten zusätzliche Dissipation einzuführen. Dies ist notwendig, da Untersuchungen an seilabgespannten hohen Masten gezeigt haben, daß für solche Bewegungen die dann auftretenden Eigenbewegungen der Seile infolge stehender und laufender Wellen zu erheblicher Energieabstrahlung führen, die nicht mehr durch die lineare viskose Dämpfung adäquat beschrieben werden kann. Diese Energieverluste, die auftreten, wenn eines der Seile von Zustand durchhängend in den Zustand gespannt übergeht, sollen durch einen Stoß modelliert werden. Die asymmetrische Seileigenschaft wird somit ergänzt durch einen asymmetrischen Energieverlust.

Das mechanische Modell:

Sei m die Brückenmasse, l_0 die ungedehnte Seillänge, k die Federkonstante des Tragseils, c die Dämpfungskonstante der viskosen Dämpfung, g die Erdbeschleunigung und λ die Stärke der Kraftanregung, so gelten in dem Zeitbereich, in dem kein Übergang von ungespanntem zu gespanntem Seil vorkommt, weiterhin die Bewegungsgleichungen für die Punktmasse in der Ebene, die im einfachen Modell einer Hängebrücke gefunden wurden:

$$\begin{aligned}
m\ddot{x} &= -k\left(\sqrt{x^2+y^2}-l_0\right)^+ \frac{x}{\sqrt{x^2+y^2}} \\
&\quad + k\left(\sqrt{(\tfrac{1}{2}-x)^2+y^2}-l_0\right)^+ \frac{\tfrac{1}{2}-x}{\sqrt{(\tfrac{1}{2}-x)^2+y^2}} \\
&\quad - c\dot{x}, \\
m\ddot{y} &= -k\left(\sqrt{x^2+y^2}-l_0\right)^+ \frac{y}{\sqrt{x^2+y^2}}
\end{aligned}$$

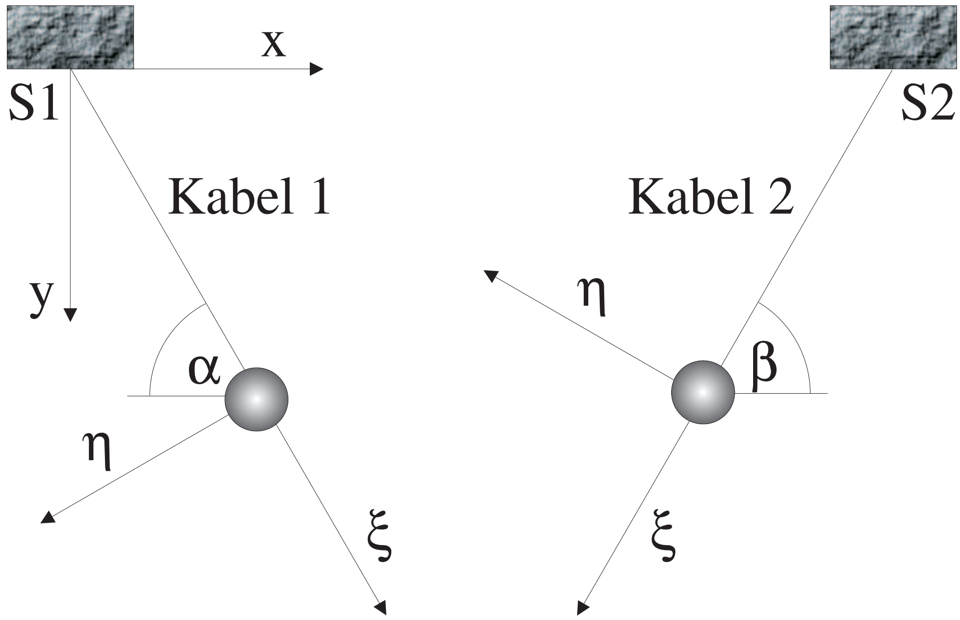

Abbildung 2.2: Verdeutlichung der lokalen Koordinatensysteme ξ-η in Kabel 1 bzw. Kabel 2 jeweils gerade beim Übergang zwischen durchhängendem und gestrafftem Seil.

$$
\begin{aligned}
&- k\left(\sqrt{(\tfrac{1}{2}-x)^2+y^2}-l_0\right)^+ \frac{y}{\sqrt{(\frac{1}{2}-x)^2+y^2}} \\
&- c\dot{y} + mg + \lambda\sin(\mu t).
\end{aligned}
$$

Sei nun t_1 der erste Zeitpunkt, für den eines der beiden Seile vom durchhängenden, in den gespannten Zustand übergeht, d.h. für alle $\epsilon > 0$ gelte

$$
\sqrt{x^2(t_1-\epsilon)+y^2(t_1-\epsilon)} < l_0 \text{ und } \sqrt{x^2(t_1+\epsilon)+y^2(t_1+\epsilon)} > l_0
$$

oder

$$
\sqrt{(\tfrac{1}{2}-x(t_1-\epsilon))^2+y^2(t_1-\epsilon)} < l_0 \text{ und } \sqrt{(\tfrac{1}{2}-x(t_1+\epsilon))^2+y^2(t_1+\epsilon)} > l_0.
$$

Zu diesem Zeitpunkt findet also, aufgrund der oben gemachten Vorüberlegungen eine Energieabstrahlung über die Eigenbewegung des plötzlich gespannten Seiles statt, die hier durch einen Stoß nachmodelliert werden soll. Dabei soll die Stoßwirkung aber nur in Seilrichtung, die wie in Abbildung 2.2 mit ξ bezeichnet sei, und nicht in der senkrechten dazu stehenden η Komponente spürbar werden. Man hat also in Seilrichtung

$$
v_{\xi,\text{neu}} := \lim_{t\to t_1+0} v_\xi(t) = \nu \lim_{t\to t_1-0} v_\xi(t) =: \nu v_{\xi,\text{alt}}
$$

und senkrecht dazu

$$
v_{\eta,\text{neu}} := \lim_{t\to t_1+0} v_\eta(t) = \lim_{t\to t_1-0} v_\eta(t) =: v_{\eta,\text{alt}},
$$

mit einer Stoßzahl ν zwischen null und eins. Für das x-y-Koordinatensystem liest sich das dann als

$$\begin{pmatrix}\dot{x}_{\text{neu}}\\ \dot{y}_{\text{neu}}\end{pmatrix} = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\nu & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\dot{x}_{\text{alt}}\\ \dot{y}_{\text{alt}}\end{pmatrix}$$
$$= \begin{pmatrix}\nu\cos^2\alpha\sin^2\alpha & \nu\cos\alpha\sin\alpha - \cos\alpha\sin\alpha\\ \nu\cos\alpha\sin\alpha - \cos\alpha\sin\alpha & \nu\sin^2\alpha + \cos^2\alpha\end{pmatrix}\begin{pmatrix}\dot{x}_{\text{alt}}\\ \dot{y}_{\text{alt}}\end{pmatrix}$$

für einen Stoß im linke Seil und als

$$\begin{pmatrix}\dot{x}_{\text{neu}}\\ \dot{y}_{\text{neu}}\end{pmatrix} = \begin{pmatrix}-\cos\beta & -\sin\beta\\ \sin\beta & -\cos\beta\end{pmatrix}\begin{pmatrix}\nu & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}-\cos\beta & \sin\beta\\ -\sin\beta & -\cos\beta\end{pmatrix}\begin{pmatrix}\dot{x}_{\text{alt}}\\ \dot{y}_{\text{alt}}\end{pmatrix}$$
$$= \begin{pmatrix}\nu\cos^2\beta\sin^2\beta & -\nu\cos\beta\sin\beta + \cos\beta\sin\beta\\ -\nu\cos\beta\sin\beta + \cos\beta\sin\beta & \nu\sin^2\beta + \cos^2\beta\end{pmatrix}\begin{pmatrix}\dot{x}_{\text{alt}}\\ \dot{y}_{\text{alt}}\end{pmatrix}$$

für einen Stoß im rechten Seil. Für den nächsten Zeitpunkt t_2, für den wieder ein Übergang von unbelastetem Seil zu gespanntem Seil vorliegt, geht man dann analog vor.

Ergebnisse:

Um zu verdeutlichen, wie die Bewegung einer Punktmasse für dieses Modell aussehen kann, wurde in Abbildung 2.3 für einen bestimmten Parametersatz eine periodische Näherungslösung in der x-y-Ebene dargestellt. Man kann die Knicke in der Bewegung, verursacht durch den Stoß beim Spannen des Seiles, deutlich erkennen. Bei einer realen Bewegung sind die Knicke allerdings nicht so offensichtlich, da sie durch die unterschiedliche Wahl der Skalierung auf der x- und y-Achse in der Abbildung stark überbetont werden. In den beiden Abbildungen in 2.4, sind zu der in Abbildung 2.3 gezeigten Näherungslösung, einmal $\dot{x}$ über x und zum anderen $\dot{y}$ über y aufgetragen. Man kann deutlich die vertikalen Sprünge in den beiden Phasenportraits erkennen, die durch den Stoß erzeugt werden.

2.4.2 In Phase und 180 Grad in Gegenphase schwingende Brückenträger

In diesem Abschnitt soll die Bewegung des Brückenmassenpunktes bei periodischer Anregung der Brückenträger untersucht werden. Die Oszilation der Türme, die die Brücke tragen, kann durch verschiedene Dinge hervorgerufen sein. Zum einen ist etwa denkbar, daß die seitlichen Abspannseile, die dem Wind ausgesetzt sind, ihre periodische Schwingung in die Brückentürme induzieren. Zum anderen können aber auch Windeffekte verantwortlich sein, die unmittelbar auf die Türme oder das Kabelsystem der Brücke wirken. An der "berühmten" Takoma Narrows Bridge wurden z.B. auch solche Turmoszilationen festgestellt. Dabei lief die Bewegung der beiden Türme entweder in gleicher Phase oder aber die Brückenträger schwangen 180 Grad in Gegenphase.

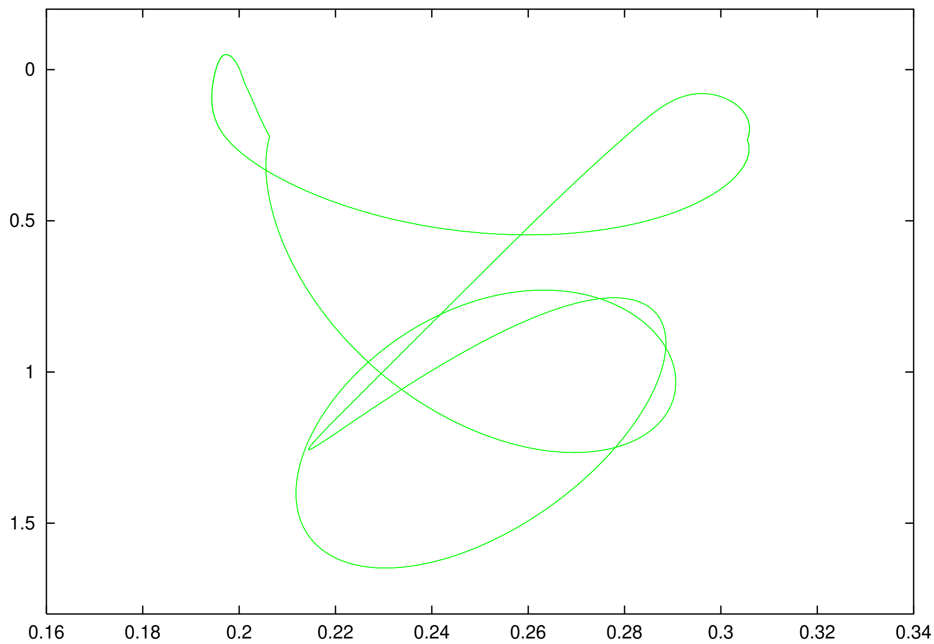

Abbildung 2.3: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\epsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$ und $\mu = 1$. Als Dämpfung beim Stoß wurde $\nu = 0.9$ verwendet. Die dargestellte Lösung hat die gleiche Periode, wie die Anregung. Man kann im Bild sehr deutlich den Knick erkennen, den die Bahn beim Stoß, also beim Übergang zwischen durchhängendem und gestrafftem Seil macht. Bei Verwendung gleicher Maßstäbe auf der x- und y-Achse, fällt der Knick dann aber nicht mehr so stark ins Auge.

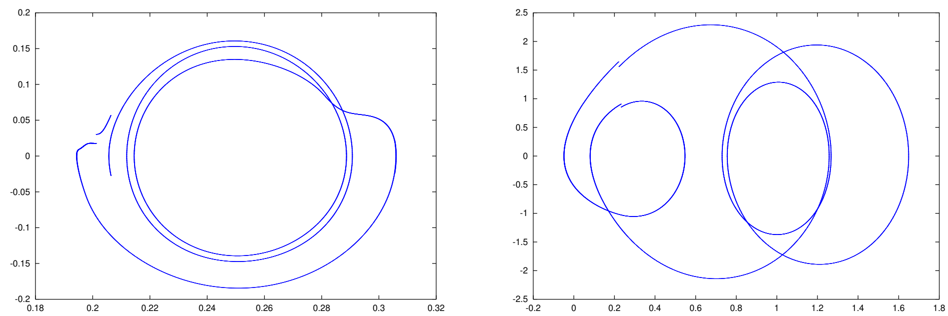

Abbildung 2.4: Auf dem linken Bild sieht man, für die in Abbildung 2.3 gezeigte Lösung, $\dot{x}$ gegen x aufgetragen. Rechts wurde $\dot{y}$ gegen y aufgetragen. In beiden Bildern sind deutlich die vertikalen Sprünge, d.h. die unstetigen Übergänge der Geschwindigkeit, verursacht durch den Stoß beim Übergang des Kabels vom durchhängendem zum gestrafften Zustand zu erkennen.

Das mechanische Modell:

Sei m die Brückenmasse, l_0 die ungespannte Seillänge, k die Federkonstante des Tragseils, c die Dämpfungskonstante der viskosen Dämpfung, g die Erdbeschleunigung und λ die Stärke der Kraftanregung. Da nur das ursprünglich betrachtete einfache Brückenmodell erweitert werden soll, gilt weiterhin die in 2.2 aufgestellte Grundgleichung:

$$
\begin{aligned}
m\frac{d^2x}{dt^2} &= -k(\overline{S_1P} - l_0)^+ \cos\alpha + k(\overline{S_2P} - l_0)^+ \cos\beta - c\frac{dx}{dt}, \\
m\frac{d^2y}{dt^2} &= -k(\overline{S_1P} - l_0)^+ \sin\alpha - k(\overline{S_2P} - l_0)^+ \sin\beta - c\frac{dy}{dt} + mg + \lambda\sin(\mu t),
\end{aligned}
$$

wobei nun allerdings die Aufhängepunkte S_1 und S_2 der Brücke nicht mehr als ortsfest angesehen werden. Sei die Ortskoordinate des ersten Aufhängepunktes mit $(x_1(t), y_1(t))$ und die des zweiten Aufhängepunktes mit $(x_2(t), y_2(t))$ beschrieben. Es ergeben sich somit die Beziehungen:

$$
\begin{aligned}
\overline{S_1P} &= \sqrt{(x-x_1)^2 + (y-y_1)^2} \\
\overline{S_2P} &= \sqrt{(x_2-x)^2 + (y-y_2)^2}
\end{aligned}
$$

$$
\begin{aligned}
\cos\alpha &= \frac{x-x_1}{\sqrt{(x-x_1)^2 + (y-y_1)^2}} & \sin\alpha &= \frac{y-y_1}{\sqrt{(x-x_1)^2 + (y-y_1)^2}} \\
\cos\beta &= \frac{x_2-x}{\sqrt{(x_2-x)^2 + (y-y_2)^2}} & \sin\beta &= \frac{y-y_2}{\sqrt{(x_2-x)^2 + (y-y_2)^2}}
\end{aligned}
$$

Da die vertikale Verschiebung des Aufhängepunktes gegenüber der horizontalen sehr klein ist, soll diese hier vernachlässigt werden. Es werden die zwei Fälle der in Phase und der um 180 Grad in Gegenphase schwingenden Brückenträger untersucht.

Fall 1: Bei den um 180 Grad aus der Phase schwingenden Brückenträgern unter Vernachlässigung der vertikalen Bewegung, ergeben sich die Ortskoordinaten der Aufhängepunkte zu

$$
S_1 = \begin{pmatrix} -\delta\sin\mu t \\ 0 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} \frac{1}{2} + \delta\sin\mu t \\ 0 \end{pmatrix},
$$

mit der Aufhängungsschwingweite δ. Setzt man alle gefundenen Beziehungen in die Grundgleichung ein so erhält man:

$$
\begin{aligned}
m\ddot{x} = &-k(\sqrt{(x+\delta\sin\mu t)^2 + y^2} - l_0)^+ \frac{x+\delta\sin\mu t}{\sqrt{(x+\delta\sin\mu t)^2 + y^2}} \\
&+ k(\sqrt{(\tfrac{1}{2} + \delta\sin\mu t - x)^2 + y^2} - l_0)^+ \frac{\frac{1}{2} + \delta\sin\mu t - x}{\sqrt{(\frac{1}{2} + \delta\sin\mu t - x)^2 + y^2}} \\
&- c\dot{x}, \\
m\ddot{y} = &-k(\sqrt{(x+\delta\sin\mu t)^2 + y^2} - l_0)^+ \frac{y}{\sqrt{(x+\delta\sin\mu t)^2 + y^2}}
\end{aligned}
$$

$$
\begin{aligned}
&- k(\sqrt{(\tfrac{1}{2} + \delta \sin \mu t - x)^2 + y^2} - l_0)^+ \frac{y}{\sqrt{(\tfrac{1}{2} + \delta \sin \mu t - x)^2 + y^2}} \\
&- c\dot{y} + mg + \lambda \sin(\mu t).
\end{aligned}
$$

Fall 2: Bei den in Phase schwingenden Brückenträgern unter Vernachlässigung der vertikalen Bewegung ergeben sich die Ortskoordinaten der Aufhängepunkte zu

$$
S_1 = \begin{pmatrix} -\delta \sin \mu t \\ 0 \end{pmatrix} \qquad S_2 = \begin{pmatrix} \frac{1}{2} - \delta \sin \mu t \\ 0 \end{pmatrix}
$$

mit der Aufhängungsschwingweite δ. Setzt man alle gefundenen Beziehungen in die Grundgleichung ein so erhält man:

$$
\begin{aligned}
m\ddot{x} = &-k(\sqrt{(x + \delta \sin \mu t)^2 + y^2} - l_0)^+ \frac{x + \delta \sin \mu t}{\sqrt{(x + \delta \sin \mu t)^2 + y^2}} \\
&+ k(\sqrt{(\tfrac{1}{2} - \delta \sin \mu t - x)^2 + y^2} - l_0)^+ \frac{\frac{1}{2} - \delta \sin \mu t - x}{\sqrt{(\tfrac{1}{2} - \delta \sin \mu t - x)^2 + y^2}} \\
&- c\dot{x}, \\
m\ddot{y} = &-k(\sqrt{(x + \delta \sin \mu t)^2 + y^2} - l_0)^+ \frac{y}{\sqrt{(x + \delta \sin \mu t)^2 + y^2}} \\
&- k(\sqrt{(\tfrac{1}{2} - \delta \sin \mu t - x)^2 + y^2} - l_0)^+ \frac{y}{\sqrt{(\tfrac{1}{2} - \delta \sin \mu t - x)^2 + y^2}} \\
&- c\dot{y} + mg + \lambda \sin(\mu t).
\end{aligned}
$$

Ergebnisse:

In Abbildung 2.5 sind zwei Näherungslösungen für das Problem der um 180 Grad in Gegenphase schwingenden Brückenaufhängungen zu sehen. Ihre Ähnlichkeit mit in Abbildung 3.2 und Abbildung 3.3 gezeigten Näherungslösungen für unbewegliche Brückenaufhängungen ist unübersehbar. Wohingegen die in Abbildung 2.6 gezeigte Näherungslösung für in Phase schwingende Brückenaufhängungen keine Ähnlichkeiten mit diesen Lösungen zeigt.

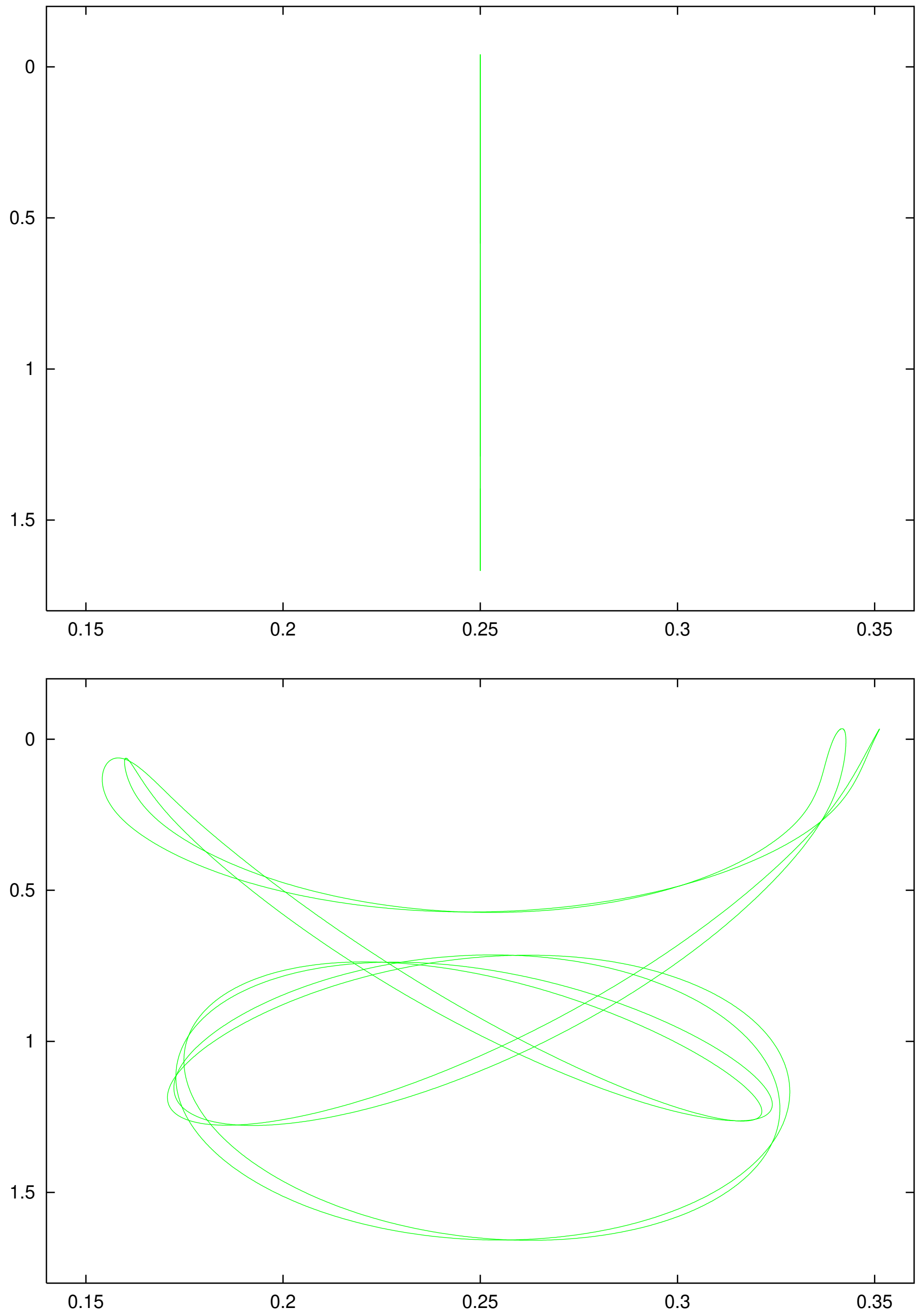

Abbildung 2.5: Bahn der Brückenpunktmasse in der x-y-Ebene bei um 180 Grad aus der Phase schwingenden Brückenaufhängungen unter Verwendung der Parameter $\epsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$, $\mu = 1$ und Aufhängungsschwingweite $\delta = 0.05$. Die zuoberst dargestellte Lösung hat die gleiche Periode wie die Anregung. Sie zeigt eine leicht erkennbare Verwandtschaft, mit der in Abbildung 3.2 dargestellten Lösungen für unbewegliche Brückenaufhängungen. Die unten dargestellte Lösung hat die doppelte Periode der Anregung. Diese Lösung zeigt wiederum eine leicht erkennbare Verwandtschaft, mit der in Abbildung 3.3 dargestellten Lösungen für unbewegliche Brückenaufhängungen.

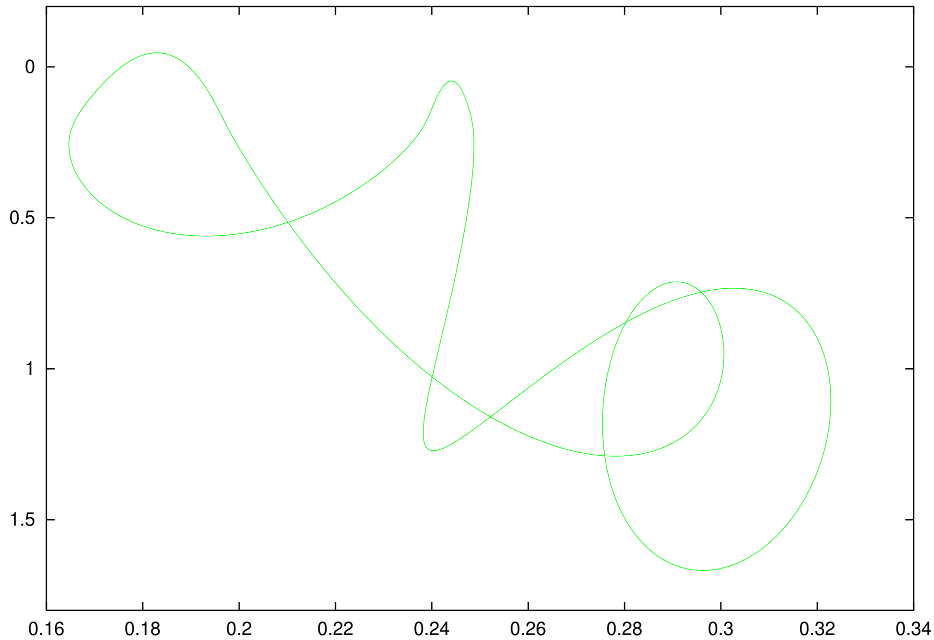

Abbildung 2.6: Bahn der Brückenpunktmasse in der x-y-Ebene bei in gleicher Phase schwingenden Brückenaufhängungen, bei Verwendung der Parameter $\epsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$, $\mu = 1$ und Aufhängungsschwingweite $\delta = 0.05$. Die dargestellte Lösung hat die gleiche Periode, wie die Anregung. Die gefundene Lösung hat im Gegensatz zu den Lösungen für die um 180 Grad aus der Phase schwingenden Brückenaufhängungen in Abbildung 2.5 keine offensichtlichen Ähnlichkeiten mit den in Abbildung 3.2, Abbildung 3.3 und Abbildung 3.4 dargestellten Lösungen für unbewegliche Brückenaufhängungen.

Kapitel 3

Numerische Ergebnisse:

Im Kapitel 3 werden die Ergebnisse der Diplomarbeit wiedergegeben. Das Kapitel läßt sich in zwei große Teile aufgliedern. Im erste Teil 3.1 wird das in Kapitel 2 vorgestellte einfache Modell der Hängebrücke anhand von Verzweigungsdiagrammen für unterschiedliche Parameter und ausgewählten Darstellungen der Bahnkurven für interessant erscheinende Parameter untersucht. Der zweite Teil enthält die Ergebnisse, die man erhält, wenn man den in Kapitel 1 vorgestellten Algorithmus auf zwei unterschiedliche Näherungslösungen des Modells der Hängebrücke anwendet, um für diese die Existenz einer exakten Lösung in einer Umgebung der Näherungslösung zu beweisen.

3.1 Parameterstudien für das einfache Modell der Hängebrücke:

In diesem ersten großen Abschnitt des Kapitel 3 werden, wie schon oben erwähnt, die Eigenschaften des Hängebrückemodells in Form von Verzweigungsdiagrammen und einzelnen Bahnkurven dargestellt. Zur Berechnung der einzelnen Lösungen dieser Verzweigungsdiagramme wurde meist ein Runge-Kutta Verfahren verwendet, das mit einem vernünftig gewählten Startwert über 2000 Periodendurchgänge der anregenden Kraft laufen gelassen wurde und danach ein Newtonalgorithmus auf diese Lösungen angewandt. (Für genauere Informationen über das Verfahren, das verwendet wurde siehe 1.3.1). Lösungen, die instabil waren wurden nur mit dem Newtonalgorithmus berechnet, wobei eine Näherungslösung eines benachbarten Parametersatzes als Startlösung der Iteration diente.

In den nun folgenden Unterabschnitten findet man einige Erleuterungen zu den Verzweigungsdiagrammen, die sich am Schluß von 3.1 befinden. Die Werte die in den einzelnen Verzweigungsdiagrammen aufgetragen wurden, sind so gewählt, daß die Verzweigungen möglichst gut hervor treten und keine gegenseitigen Überdeckungen der Lösungszweige vorliegen. Dies bedingt nun aber leider, daß die einzelnen Diagramme nicht direkt (an Hand der aufgetragenen Zahlenwerte) vergleichbar sind.

3.1.1 Studie für den Parameter λ:

In Abbildung 3.1 sieht man ein Verzweigungsdiagramm für den Parameter λ zwischen 8.85 und 9.2. Der in der Mitte ziemlich gerade verlaufende Zweig aus Pluszeichen gehört zu einer einfachen Auf- und Abschwingung, die in Abbildung 3.2 für $\lambda = 8.9$ dargestellt ist. Man kann wunderbar erkennen, wie es ungefähr beim Parameter $\lambda = 9$ zu einer Verzweigung mit Periodenverdoppelung kommt. Für den weiterhin mit einfacher Periode verlaufende Zweig, zwischen den Zweigen mit der Peridenverdoppelung werden die Lösungen instabil und können dadurch nur noch mit dem Newtonalgorithmus, aber nicht mehr mit dem Runge-Kutta Verfahren, bestimmt werden. Auch bei den Lösungszweigen, deren Lösungen eine vierfachen Periode in Bezug zur Anregung besitzen kann man den Effekt des Verzweigens und der Periodenverdoppelung ganz am Ende der vier Zweige erkennen. Abbildung 3.4 zeigt, ebenfalls für den Parameter $\lambda = 8.9$, eine Lösung aus diesem Lösungszweig mit vierfacher Periode der Anregung. Die beiden Zweige mit einfacher Periode, die neben der einfachen Auf- und Abschwingung liegen, gehören zu einer unsymmetrischen Schwingung, die in Abbildung 3.3, auch für den gleichen λ-Wert wie die letzten Bahnkurven, dargestellt ist. Die Parameter der Lösungen, die man in den Abbildungen 3.2, 3.3 und 3.4 sieht wurden absichtlich exakt gleich gewählt um dem Leser die Vielfalt und Unterschiedlichkeit der Lösungen bei gleichen Parametern zu zeigen.

Abbildung 3.5 zeigt ebenso wie Abbildung 3.1 ein Verzweigungsdiagramm für den Parameter λ, nun aber mit einem anderen Wert für die ungestreckte Länge der Tragseile l_0. Man kann deutlich die Ähnlichkeit zwischen den beiden Verzweigungsdiagrammen erkennen, mit dem einen Unterschied, daß diesmal die unsymmetrische Lösung, die mit der gleichen Periode wie die anregende Kraft schwingt im hinteren Bereich, also erst nach Periodenverdoppelung der einfachen Auf- und Abschwingung, auftritt.

Abbildung 3.6 zeigt dann für den gleichen Parametersatz ein Verzweigungsdiagramm in λ, das nur die Lösungszweige enthält, deren Lösungen mit der drei- oder sechsfachen Periode der Anregung schwingen. Man kann erkennen, wie zuerst eine Periodenhalbierung und später wieder eine Periodenverdopplung stattfindet. Die beiden Abbildungen 3.7 und 3.8 zeigen, wie solch ein periodenhalbierender Übergang bei den Bahnkurven der Brückenmasse aussehen kann. Die Grundcharakteristika sind sowohl bei der Lösung mit der dreifachen als auch bei der Lösung mit der sechsfachen Periode ähnlich.

3.1.2 Studie für den Parameter ε:

Abbildung 3.9 zeigt ein Verzweigungsdiagramm für den Parameter ε für den $x^2 + y^2$ zum Zeitpunkt $t = 0$ aufgetragen wurde. Alle darin dargestellten Lösungen haben die gleiche Periode wie die anregende Kraft. Man sieht deutlich wie die beiden äußeren Zweige, die zu einer unsymmetrischen (in Bezug zu $x = 0.25$) Lösung gehören in den mittleren Zweig, der zu einer einfachen eindimensionalen Auf- und Abschwingung gehört, einmünden. Abbildung 3.10 zeigt diesen Vorgang des Einmündens direkt an einzelnen Bahnkurven für bestimmte Parameter ε. Das gleiche Verhalten konnte auch an einer Näherungslösung beobachtet werden, die die vierfache Periode der Anregung besaß und die für ε im Bereich von 10e-4 auch in eine einfache Auf- und Abbewegung einmündete.

Diese Beobachtungen an sich stellen aber schon eine starke Kritik am Modell selbst dar. Offensichtlich existieren viele Lösungen nur für den Fall eines Knickes in der $(.)^+$ Funktion, also im alternativen Modell für $\varepsilon \to 0$ und verschwinden für glättere Übergänge zwischen Ausübung einer Kraft bei gespanntem Seil und keiner Ausübung einer Kraft bei durchhängendem Seil. Diesen harten Knick beim Übergang wird man aber in der Natur vergeblich suchen, da schon allein dadurch, daß ein Stahlseil aus mehreren Trossen aufgebaut ist, die nicht alle exakt gleichlang sind, also einzelne Stränge etwas früher Kraft ausüben als andere ein kontinuierlicherer Übergang gegeben ist. Es ist daher höchst fraglich wie und ob überhaupt Ergebnisse aus der Untersuchung an diesem Modell auf real existierende Brücken übertragen werden können.

3.1.3 Studie für den Parameter k:

Für den Parameter k findet man zwei Verzweigungsdiagramme. Im ersten Diagramm (Abbildung 3.11) sind ausschließlich Lösungen dargestellt, die nur eine einfache eindimensionale Auf- und Abschwingung durchführen, während im zweiten (Abbildung 3.12) nur Zweige mit Lösungen zu finden sind, die eine tatsächlich zweidimensionale Bewegung durchführen. Auch hier kann man, wie später auch bei der Untersuchung des Parameters μ, die Periodenverdopplung an mehreren Stellen beobachten. Zum einen sieht man wie eine Lösung mit der dreifachen Periode der Anregung übergeht in eine Lösung mit der sechsfachen Periode, zum anderen kann man auch einen Übergang von der gleichen Periode wie die Anregung zu der doppelten und wieder zurück zur einfachen beobachten. In Abbildung 3.12 sind solche Übergänge leider nicht zu finden, obwohl auch dort Lösungszweige nur über einen bestimmten Bereich hin verfolgt werden konnten. Warum die Lösungen außerhalb dieses Bereichs nicht mehr mit dem Runge-Kutta Verfahren zu finden sind ist mir nicht klar, da eigentlich alle Lösungszweige sehr glatt sind, auch an den Stellen, an denen sie verloren gehen.

Obwohl Abbildung 3.13 und Abbildung 3.14, wenn man sich Abbildung 3.13 in y-Richtung etwas auseinandergezogen vorstellt, sehr ähnlich sind entstammen sie doch ganz unterschiedlichen Lösungszweigen. Abbildung 3.14 stammt dabei aus dem oberen und Abbildung 3.13 aus dem unteren Lösungszweig mit zweifacher Periode der Anregung im Verzweigungsdiagramm 3.12.

In Abbildung 3.15 ist dann eine Näherungslösung dargestellt, die obwohl sie so chaotisch und an manchen Stellen sogar kantig wirkt, doch nur eine sechsfache Periode der Anregung besitzt und somit keinesfalls chaotisch ist.

3.1.4 Studie für Parameter μ:

Im Verzweigungsdiagramm in Abbildung 3.16, in dem über dem Parameter μ der Wert $y^2 + \dot{y}^2$ aufgetragen wurde, sind ausschließlich Lösungen dargestellt, die nur eindimensionale Schwingungen ausführen, also nur einfache auf und ab Bewegungen. Verblüffend ist hierbei die Menge unterschiedlicher Lösungszweige. Dies Fülle unterschiedlicher Lösungen wäre noch größer, wären auch die Lösungen, die eine dreifache Periode in Bezug zur Anregung besitzen mit in das Diagramm aufgenommen worden. Diese wurden aber, da nur sehr viel, oftmals nur über kurze μ-Bereiche bestehende Zweige gefunden wurden der Übersichtlichkeit zu liebe, hier außen vor gelassen. Interessante Bereiche sind der Bereich um $\mu = 1.2$ in dem zwei Zweige mit der gleichen Periode sehr nah beieinander vorbeilaufen, und doch beide stabil weiter existieren. Auch sehr schön beobachten kann man eine Periodenverdoppelung, die dann wieder in eine einfache Periode übergeht. Es ist zu vermuten, daß bei dieser Periodenverdoppelung zwischen den beiden Zweigen mit der doppelten Periode weiterhin ein Zweig mit der gleichen Periode, wie die Anregung existiert, der aber da er vermutlich instabil ist, mit dem hier verwendeten Runge-Kutta Verfahren, nicht berechnet werden konnte.

In Abbildung 3.17 sind dann die Lösungszweige zu sehen, bei denen eine tatsächliche zweidimensionale Bewegung stattfindet. Der Lösungszweig mit den gelben Pluszeichen gehört zu einer unsymmetrischen Bewegung, die die gleiche Periode wie die Anregung besitzt, weswegen man durch Spiegelung an der $x = 0.25$ Achse Lösungen eines zweiten Lösungszweiges erhält. Bei den Lösungszweigen mit der vierfachen Periode der Anregung kann man deutlich erkennen, wie diese aus einer Lösung mit zweifacher Periode hervorgehen, die aber leider nicht weiter verfolgt werden konnte.

3.1.5 Zusammenfassung:

Die einzelnen Verzweigungsdiagramme und auch die einzelnen exemplarisch herausgegriffenen Bewegungsverläufe demonstrieren, wie unermeßlich groß das Spektrum verschiedenartiger Lösungen ist. Die Bandbreite reicht dabei von einfachen Auf- und Abbewegungen, die selbst schon in ungeheuerer Zahl vorkommen, über zu $x = 0.25$ symmetrischen hübsch anzuschauenden Lösungen, hin zu unsymmetrische geradezu chaotisch wirkenden Lösungen, oder gar zu Lösungen deren Periode entweder so groß war, daß sie nicht gefunden wurde, oder, die tatsächlich unperiodisch waren.

In wie weit diese Ergebnisse nun helfen, bei realen Hängebrücken auftretende Phänomene zu verstehen ist aus mehreren Gründen nicht ganz klar:

1. Die Untersuchung des Parameters ε legt nahe, daß viele im Modell auftretende Phänomene, durch den Knick in der Funktion $(.)^+$ bedingt sind, der aber in der Natur nicht vorkommen wird.

2. Viele der gefundenen Näherungslösungen sind von rein mathematischer Natur, da sie so große Schwingweiten aufweisen, daß sie in der Natur nie vorkommen würden, da schon weit vor diesen Schwingweiten die Brücke in sich zusammenbrechen würde.

3. Es gibt unterschiedliche Auffassungen darüber, was geschieht, wenn man von dem einfachen Modell übergeht zu einer realen Brücke. Die eine besagt, daß man die prinzipiellen Ergebnisse auch auf die viel komplexere Brückenstruktur anwenden kann. Die andere geht aber davon aus, daß dadurch, daß bei einer realen Brücke viel mehr Komponente miteinander in Wechselwirkung treten und dadurch die einzelnen Bauteile nur geringere Bewegungsmöglichkeiten haben, es eher zu einer Versteifung der Konstruktion kommt und damit ein ganz anderes Verhalten vorliegt, als jenes, das durch das Modell beschrieben wird.

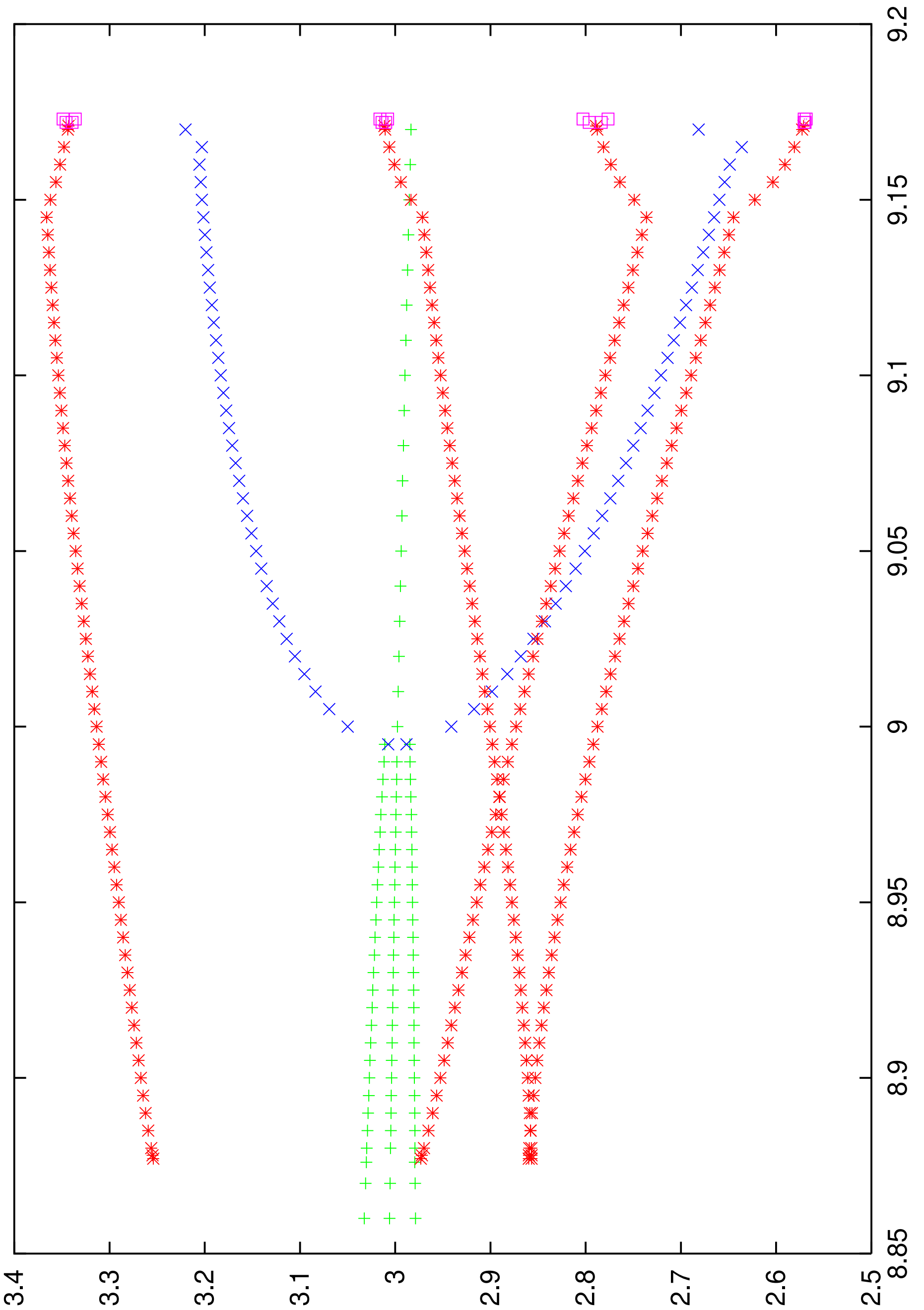

Abbildung 3.1: Parameterstudie für den Parameter λ zwischen 8.86 und 9.173. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E - 10$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$ und $\mu = 1$. Die gelben Pluszeichen gehören zu Lösungen, die mit der gleichen, die blauen x-Symbole zu Lösungen, die mit der doppelten, die roten Sterne zu Lösungen, die mit der vierfachen, und die blauen Quadrate gehören zu Lösungen, die mit der achtfachen Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung $x + \dot{x} + y + \dot{y}$ zum Zeitpunk $t = 0$.

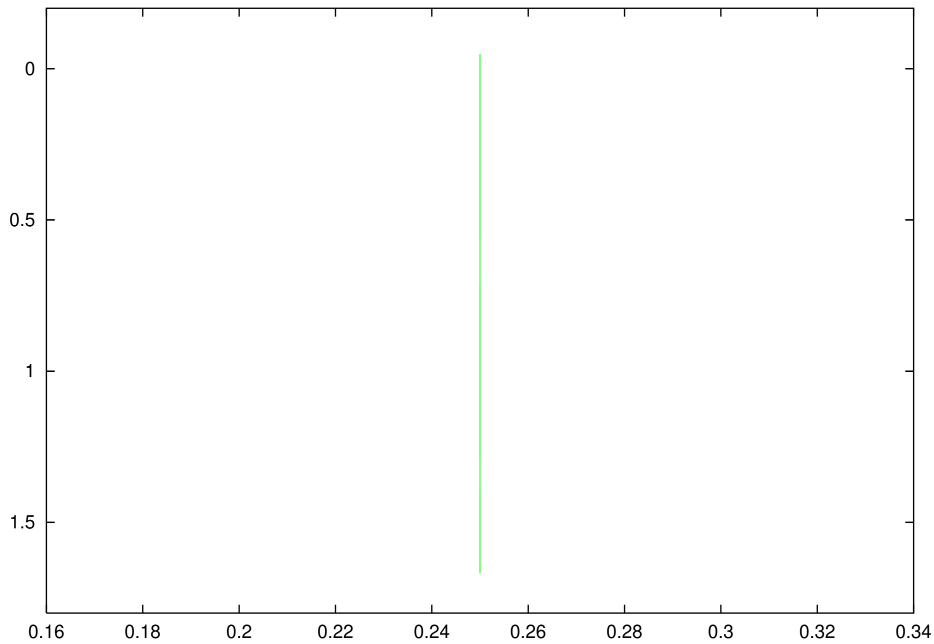

Abbildung 3.2: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$ und $\mu = 1$. Die dargestellte Lösung hat die gleiche Periode wie die Anregung

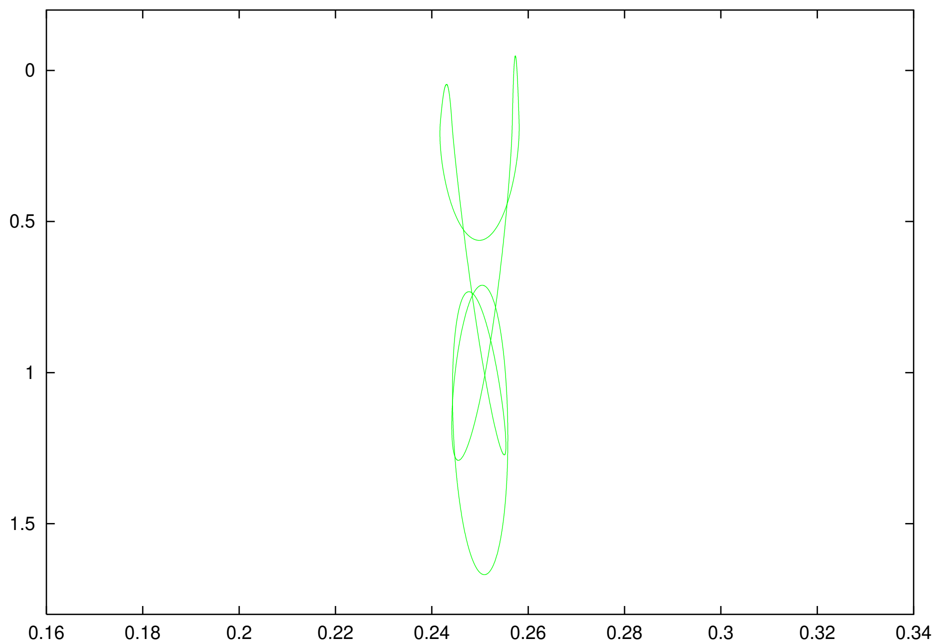

Abbildung 3.3: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$ und $\mu = 1$. Die dargestellte Lösung hat die gleiche Periode wie die Anregung

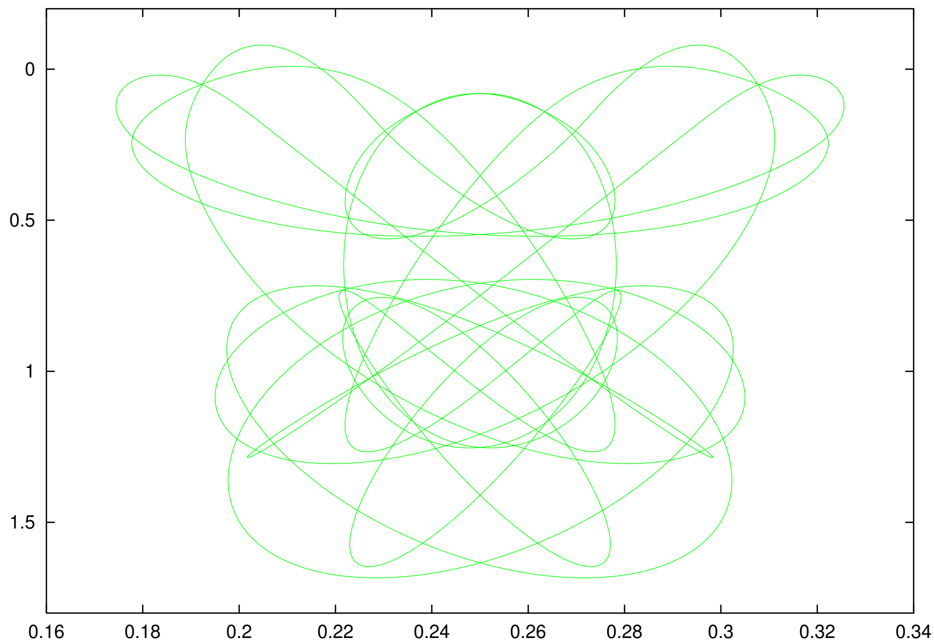

Abbildung 3.4: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$ und $\mu = 1$. Die dargestellte Lösung hat die vierfache Periode der Anregung

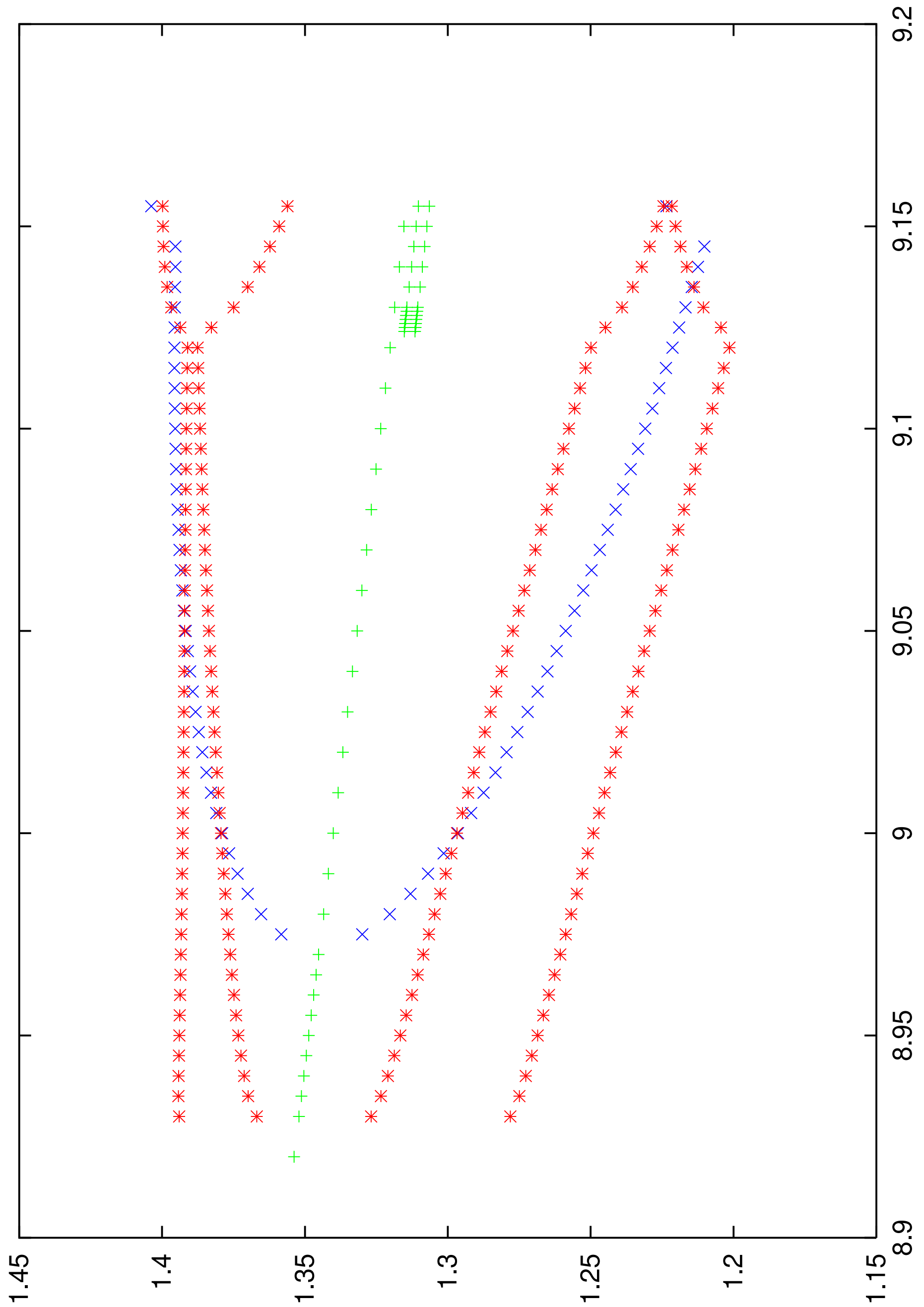

Abbildung 3.5: Parameterstudie für den Parameter λ zwischen 8.92 und 9.155. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E - 10$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.32$ und $\mu = 1$. Die gelben Pluszeichen gehören zu Lösungen, die mit der gleichen, die blauen x-Symbole zu Lösungen, die mit der doppelten, die roten Sterne zu Lösungen, die mit der vierfachen Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung x^2+y^2 zum Zeitpunk $t = 0$. Die Lösungen mit drei- und sechsfacher Periode der Anregung finden sich in der Abbildung 3.6

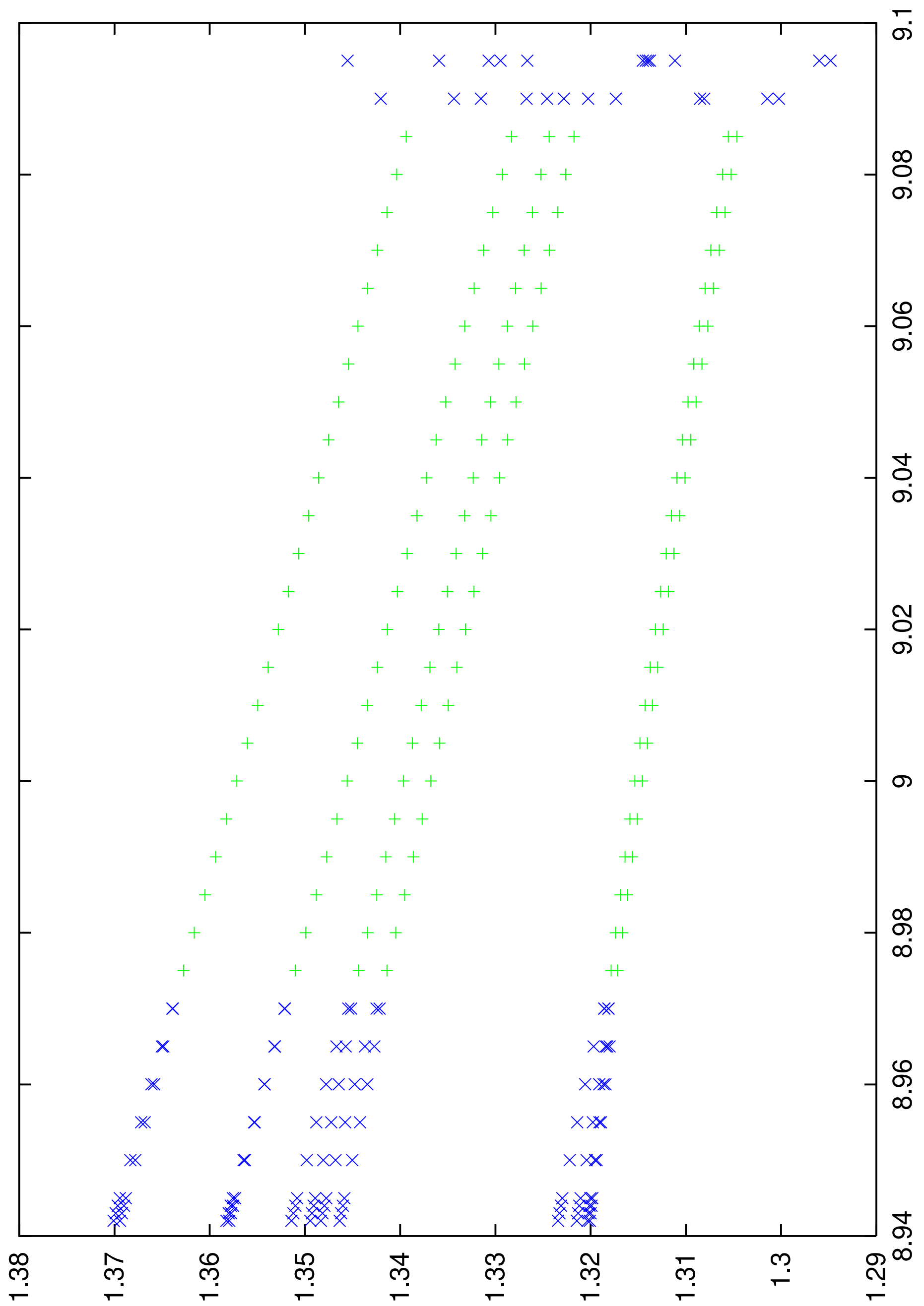

Abbildung 3.6: Parameterstudie für den Parameter λ zwischen 8.942 und 9.095. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E-10$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.32$ und $\mu = 1$. Die gelben Pluszeichen gehören zu Lösungen, die mit der dreifachen, die blauen x-Symbole zu Lösungen, die mit der sechsfachen Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung $x^2 + y^2$ zum Zeitpunk $t = 0$. Die Lösungen mit ein-, zwei- und vierfacher Periode der Anregung finden sich in der Abbildung 3.5

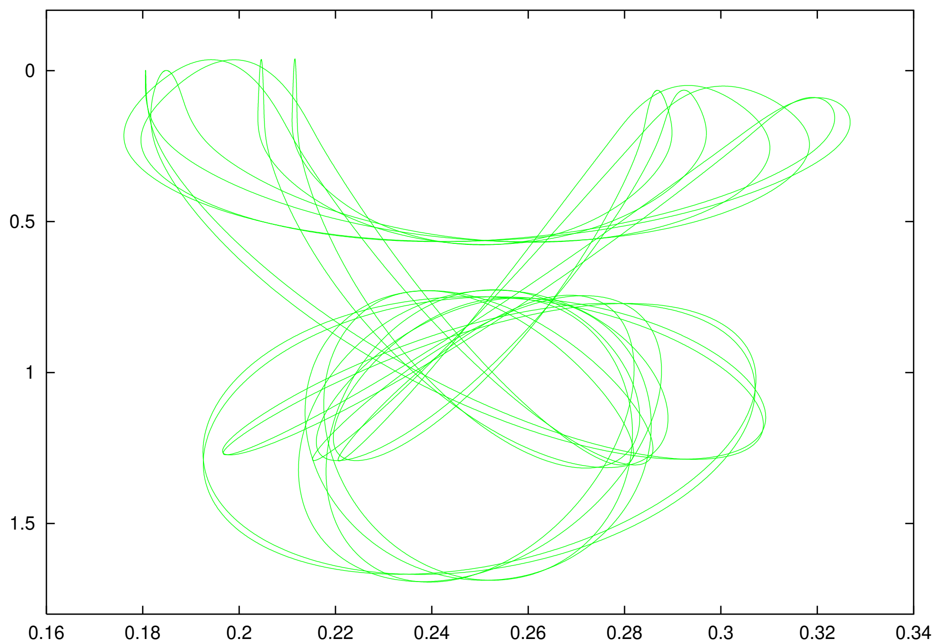

Abbildung 3.7: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E-9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.32$, $\lambda = 8.942$ und $\mu = 1$. Die dargestellte Lösung hat die sechsfache Periode der Anregung

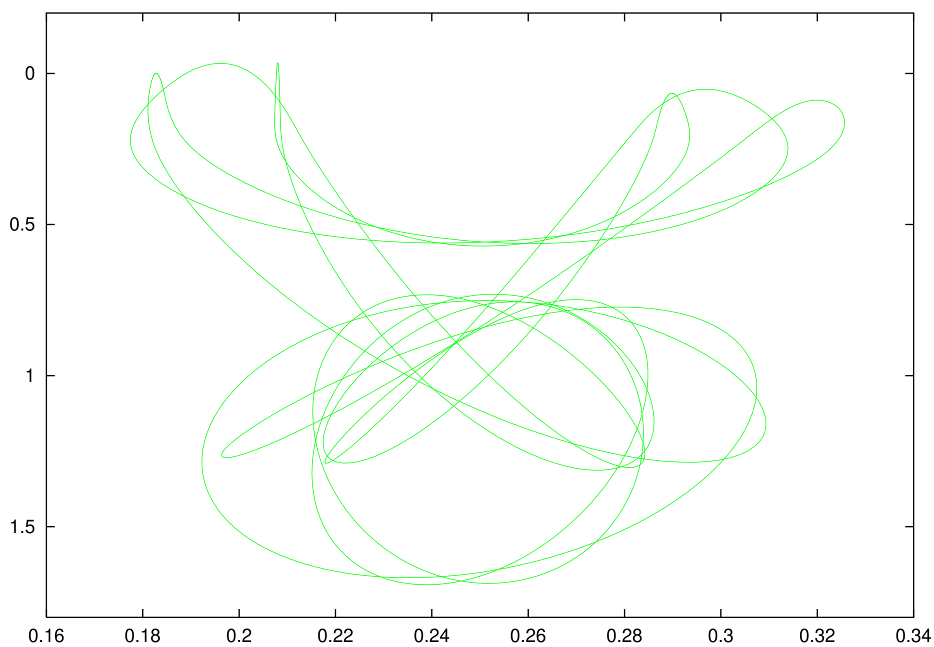

Abbildung 3.8: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E-9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.32$, $\lambda = 8.975$ und $\mu = 1$. Die dargestellte Lösung hat die dreifache Periode der Anregung

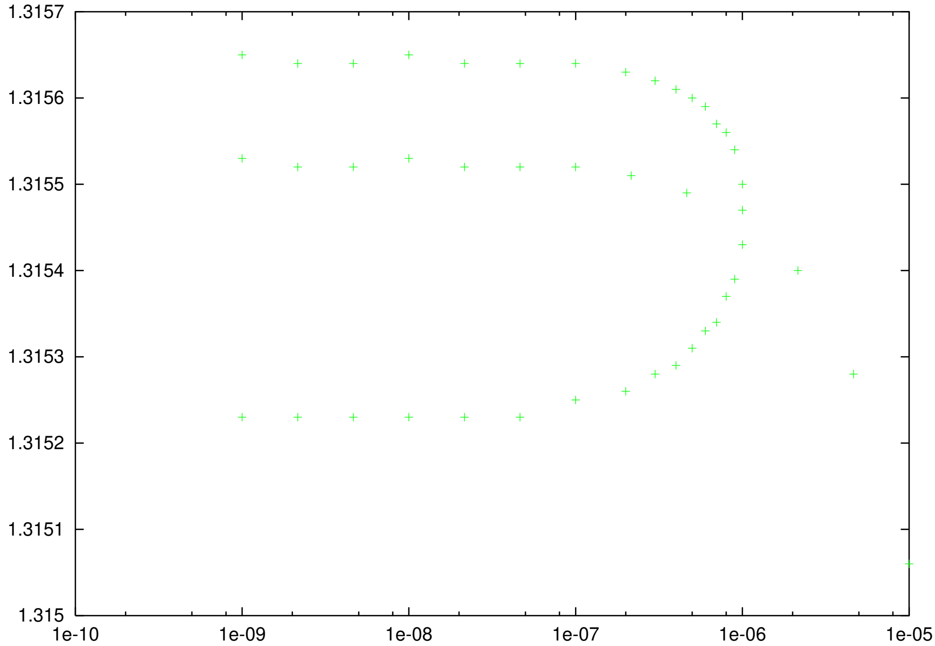

Abbildung 3.9: Parameterstudie für den Parameter ε zwischen 1e-9 und 1e-5. Die übrigen Parameter sind festgelegt auf: $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$ und $\mu = 1$. Alle berechneten Näherungslösungen haben die gleiche Periode wie die anregende Kraft.

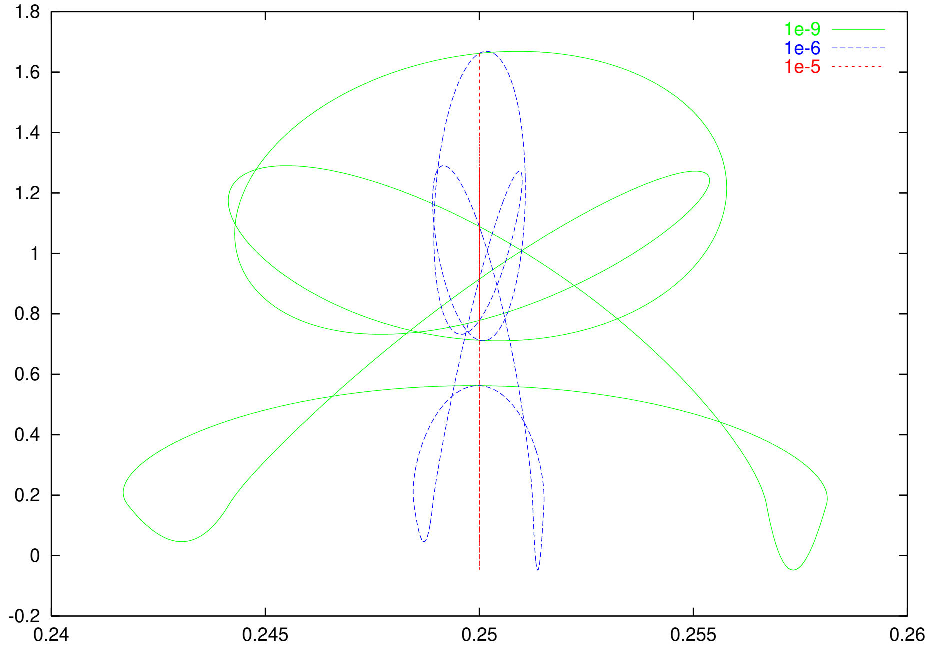

Abbildung 3.10: Bahnen der Brückenpunktmassen in der x-y-Ebene bei Verwendung der bei allen drei Näherungslösungen gleichen Parameter $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.9$ und $\mu = 1$. Einzig der Parameter ε hat bei der ersten Bahn den Wert $\varepsilon = 1e-9$, bei der nächsten den Wert $\varepsilon = 1e-6$ und bei der letzten den Wert $\varepsilon = 1e-5$. Alle dargestellten Näherungslösungen haben die gleiche Periode wie die anregende Kraft.

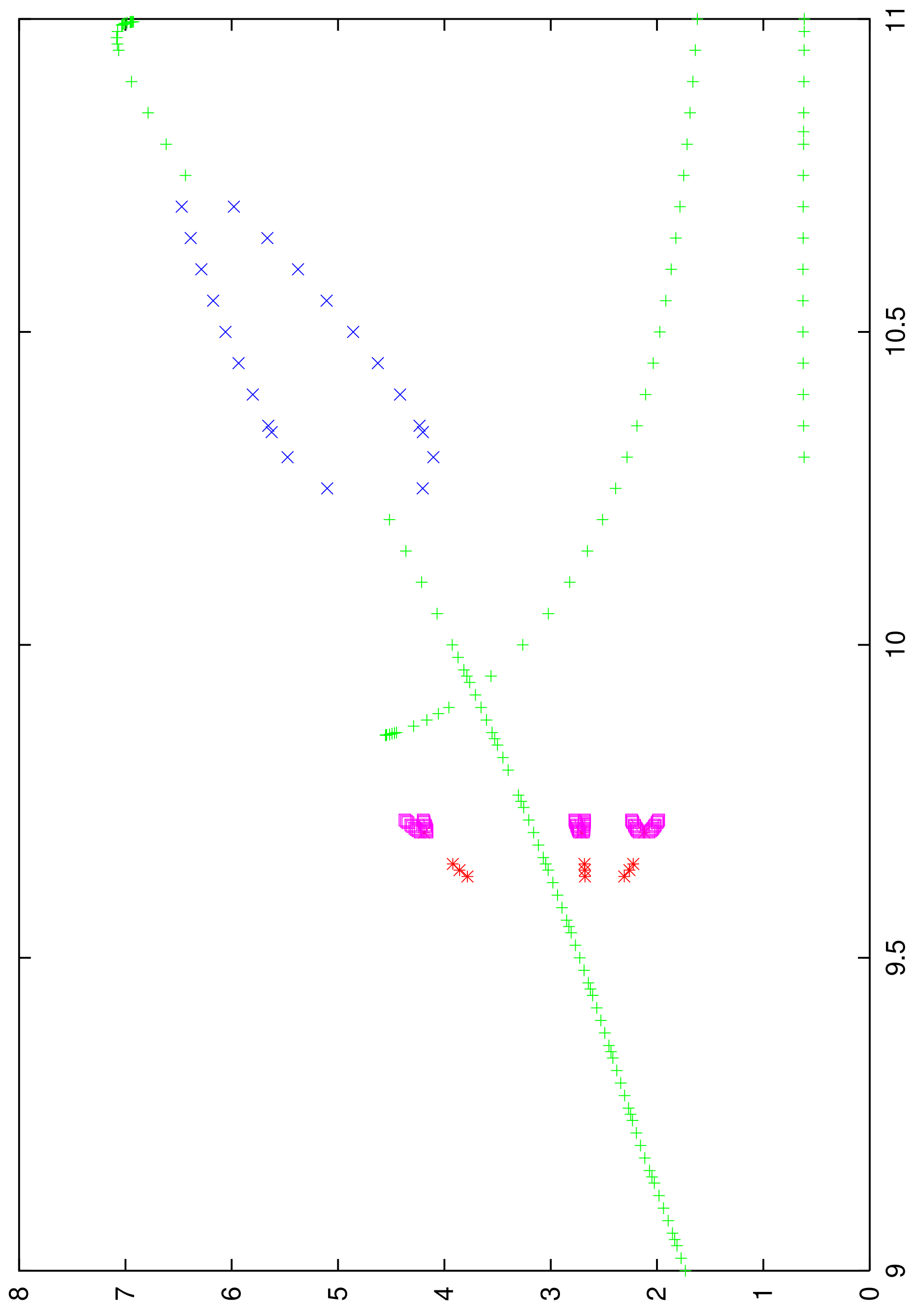

Abbildung 3.11: Parameterstudie für den Parameter k zwischen 9 und 11. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E-10$, $m = 1$, $\lambda = 8.875$, $c = 0.05$, $l_0 = 0.3025$ und $\mu = 1$. Die gelben Pluszeichen gehören zu Lösungen, die mit der gleichen, die blauen x-Symbole zu Lösungen, die mit der doppelten, die roten Sterne zu Lösungen, die mit der dreifachen, und die blauen Quadrate gehören zu Lösungen, die mit der sechsfachen Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung $y^2 + \dot{y}^2$ zum Zeitpunk $t = 0$.

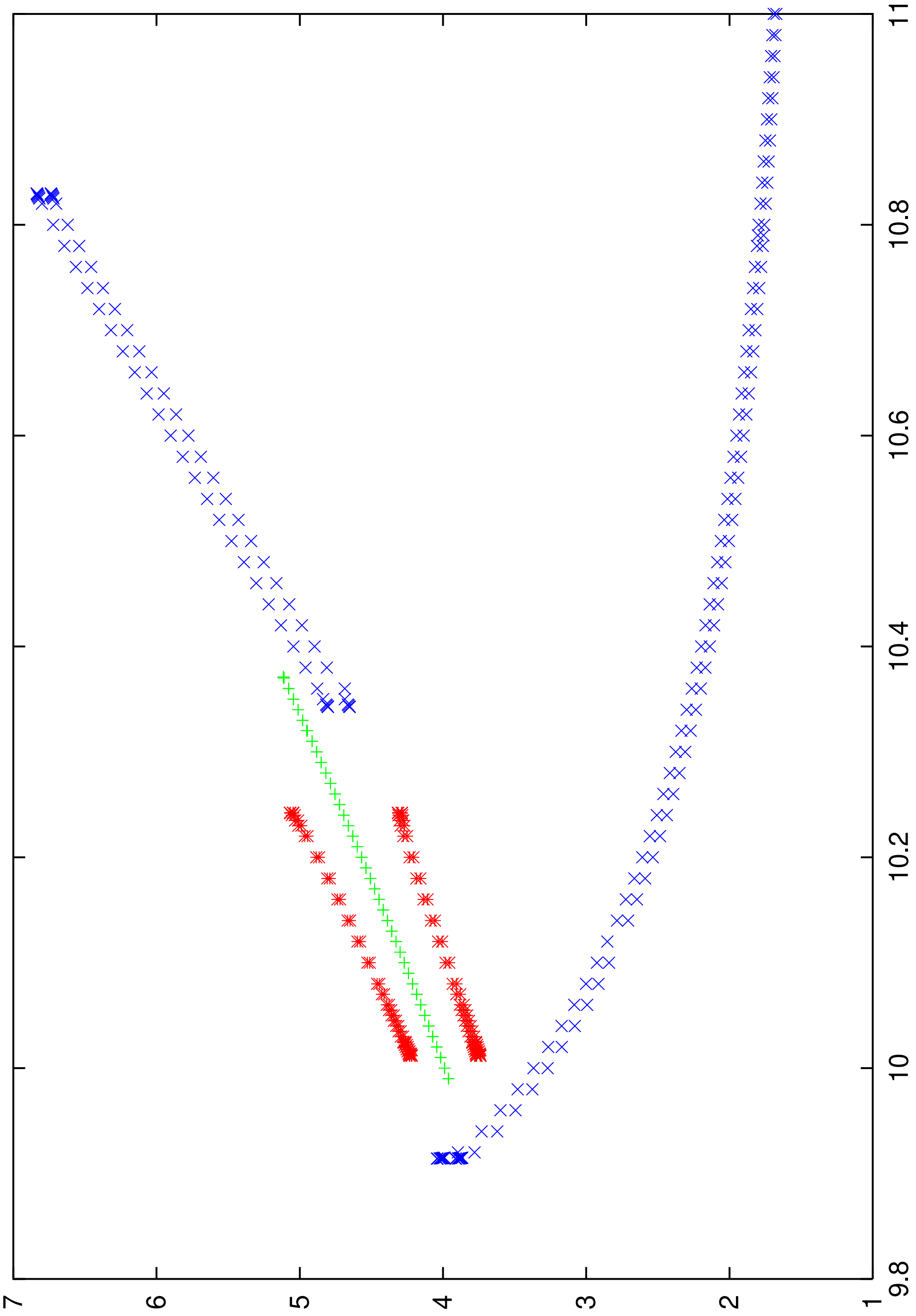

Abbildung 3.12: Parameterstudie für den Parameter k zwischen 9.8 und 11. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E - 10$, $m = 1$, $\lambda = 8.875$, $c = 0.05$, $l_0 = 0.3025$ und $\mu = 1$. Die gelben Pluszeichen gehören zu Lösungen, die mit der gleichen, die blauen x-Symbole zu Lösungen, die mit der doppelten, und die roten Sterne zu Lösungen, die mit der vierfachen Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung $x^2 + \dot{x}^2 + y^2 + \dot{y}^2$ zum Zeitpunk $t = 0$.

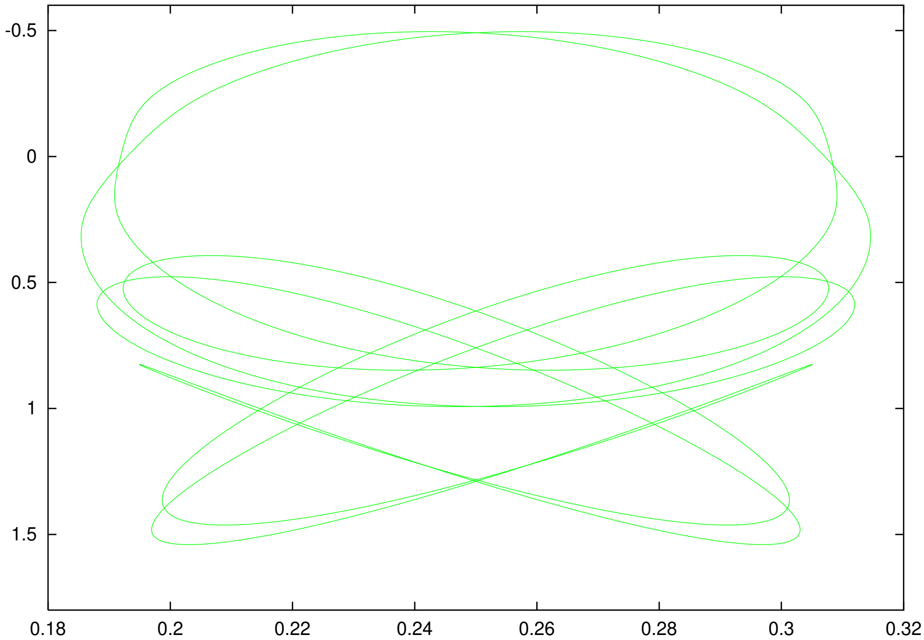

Abbildung 3.13: Bahnen der Brückenpunktmassen in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 10E - 10$, $m = 1$, $k = 10.4$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.875$ und $\mu = 1$. Die dargestellte Lösung hat die zweifache Periode der Anregung.

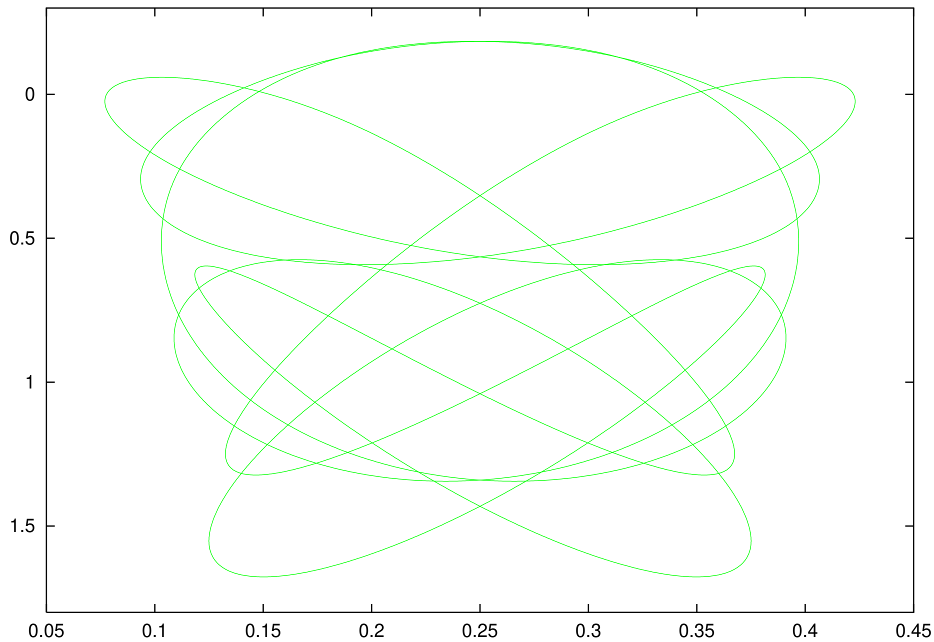

Abbildung 3.14: Bahnen der Brückenpunktmassen in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 10E - 10$, $m = 1$, $k = 10.76f$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.875$ und $\mu = 1$. Die dargestellte Lösung hat die zweifache Periode der Anregung.

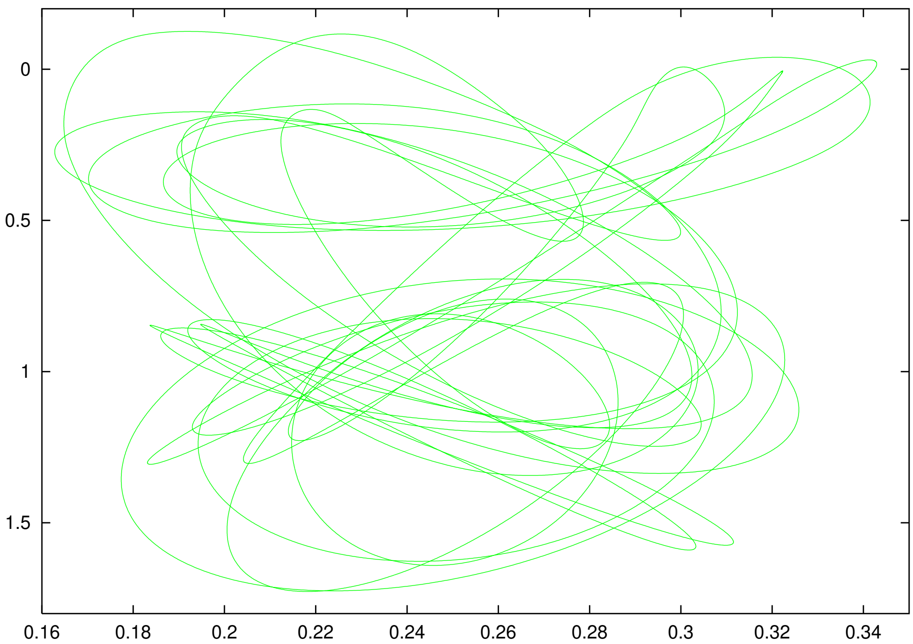

Abbildung 3.15: Bahnen der Brückenpunktmassen in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 10E - 10$, $m = 1$, $k = 9.78f$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.875$ und $\mu = 1$. Die dargestellte Lösung hat die sechsfache Periode der Anregung.

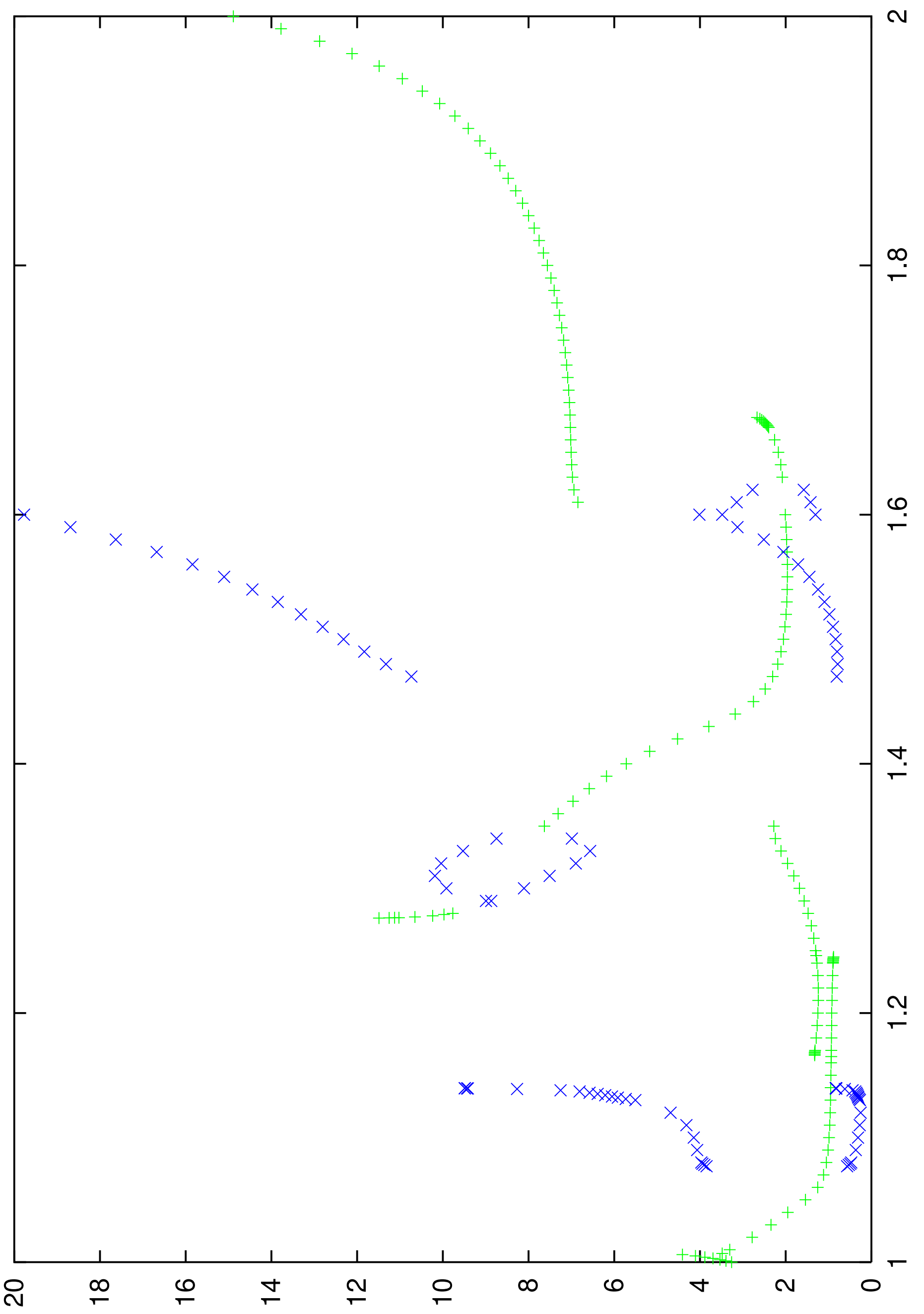

Abbildung 3.16: Parameterstudie für den Parameter μ zwischen 1 und 2. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E - 10$, $m = 1$, $k = 10$, $\lambda = 8.875$, $c = 0.05$ und $l_0 = 0.3025$ Die gelben Pluszeichen gehören zu Lösungen, die mit der gleichen und die blauen x-Symbole zu Lösungen, die mit der doppelten Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung $y^2 + \dot{y}^2$ zum Zeitpunk $t = 0$.

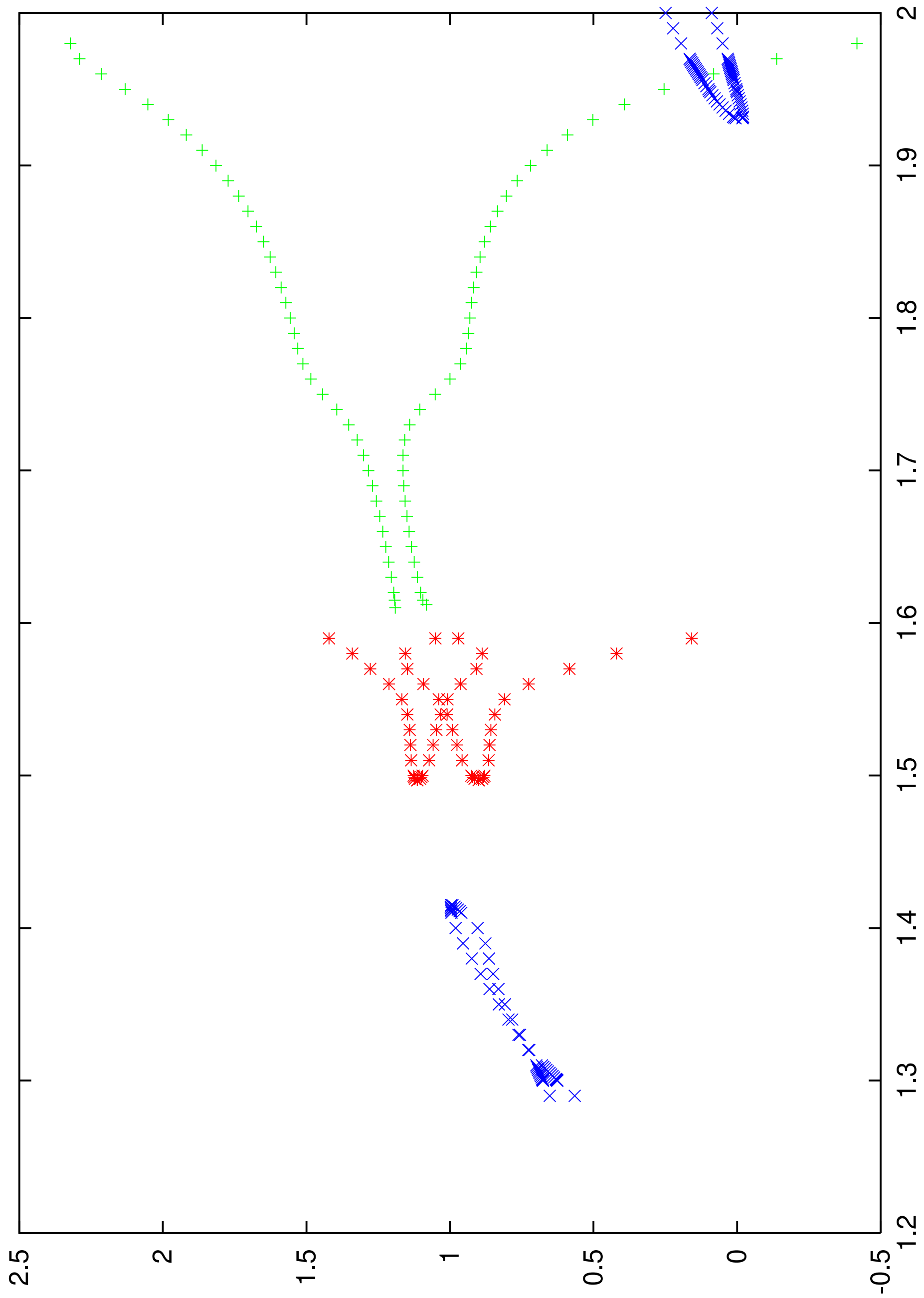

Abbildung 3.17: Parameterstudie für den Parameter μ zwischen 1.2 und 2. Die übrigen Parameter sind festgelegt auf: $\varepsilon = 10E - 10$, $m = 1$, $k = 10$, $\lambda = 8.875$, $c = 0.05$ und $l_0 = 0.3025$ Die gelben Pluszeichen gehören zu Lösungen, die mit der gleichen, die blauen x-Symbole zu Lösungen, die mit der doppelten und die roten Sterne zu Lösungen, die mit der vierfachen Periode der Anregung schwingen. Aufgetragen wurde für jede Lösung $x + y$ zum Zeitpunk $t = 0$.

3.2 Anwendung des Einschließungsalgorithmus auf das Modellproblem

In den folgenden Unterabschnitten dieses Abschnittes werden die einzelnen Punkte des Algorithmus aus Kapitel 1 angewandt auf die Gleichungen des Modellproblems in 2.2, um die Verwendbarkeit dieses Algorithmus zu demonstrieren. Es wurden zwei einfache Lösungen, an denen die Prinzipien und Vorgehensweisen sehr gut hervortreten, ausgewählt. Der Algorithmus ist auch auf kompliziertere Lösungen anwendbar. Solche Lösungen werden hier aber Aufgrund des höheren Rechenaufwands und der geringeren Übersichtlichkeit nicht behandelt.

Die einzelnen Unterabschnitte haben die gleichen Überschriften, wie die Unterabschnitte in Kapitel 1, sodaß man leicht zwischen Kapitel 1 und Kapitel 3 hin und herspringen kann. Dies erlaubt einem dann z.B. bei den Ergebnissen noch einmal schnell einen Blick auf die theoretischen Grundlagen zu werfen, oder umgekehrt bei der mathematischen Theorie in Kapitel 1 schon einmal einen Blick auf die Ergebnisse zu werfen, um sich alles bildlicher vorstellen zu können.

3.2.1 Zu Punkt 1: Bestimmung der Näherungslösung ω:

In Abbildung 3.18 ist die Näherungslösung zu sehen, die fortan mit Lösung *(Lsg1)* bezeichnet wird. Abbildung 3.19 zeigte die zweite Näherungslösung, die hier näher untersucht werden soll. Diese bekomme zur Kennzeichnung den Namen: Lösung *(Lsg2)*. Die verwendeten Parameter für die beiden Lösungen sind bei der jeweiligen Abbildung zu finden. Sowohl Lösung *(Lsg1)*, als auch Lösung *(Lsg2)* haben die gleiche Periode wie die anregende Kraft. Bei Lösung *(Lsg1)* handelt es sich genau genommen nicht um eine Lösung eines Differentialgleichungssystems, sondern da $x = 0.25 = \text{const}$ und somit die erste Gleichung des Differentialgleichungssystems exakt erfüllt ist, nur um die Lösung einer gewöhnlichen Differentialgleichung, die aber dennoch nicht trivial ist, da es sich um eine nichtlineare Differentialgleichung handelt, die man nicht in geschlossener Form lösen kann. Lösung *(Lsg2)* hingegen zeigt eine wirklich zweidimensionale Bewegung.

Beide Lösungen wurden aus einem zweistufigen Verfahren gewonnen. In der ersten Stufe wurde ein Runge-Kutta-Verfahren verwendet, um ausgehend von einem Startpunkt eine Folge von Punkten zu gewinnen. Dabei wurden pro Periodendurchgang der Anregung 4096 Zeitschritte verwendet. Die Berechnungen des Runge-Kutta-Verfahrens erstreckten sich auf 2000 Periodendurchläufe der anregenden Kraft. Die Punkte ab der zweitausendsten Periode wurden dann genutzt, um eine Fourierreihenapproximation mit insgesammt 402 Koeffizienten zu erhalten. Diese Näherung

$$\begin{aligned}\omega = &\left(xa_0 + \sum_{i=1}^{100}[xa_i \cos(ik\tfrac{2\pi}{\mu}t) + xb_i \sin(ik\tfrac{2\pi}{\mu}t)]\right)\vec{e}_x \\ &+ \left(ya_0 + \sum_{i=1}^{100}[ya_i \cos(ik\tfrac{2\pi}{\mu}t) + yb_i \sin(ik\tfrac{2\pi}{\mu}t)]\right)\vec{e}_y\end{aligned}$$

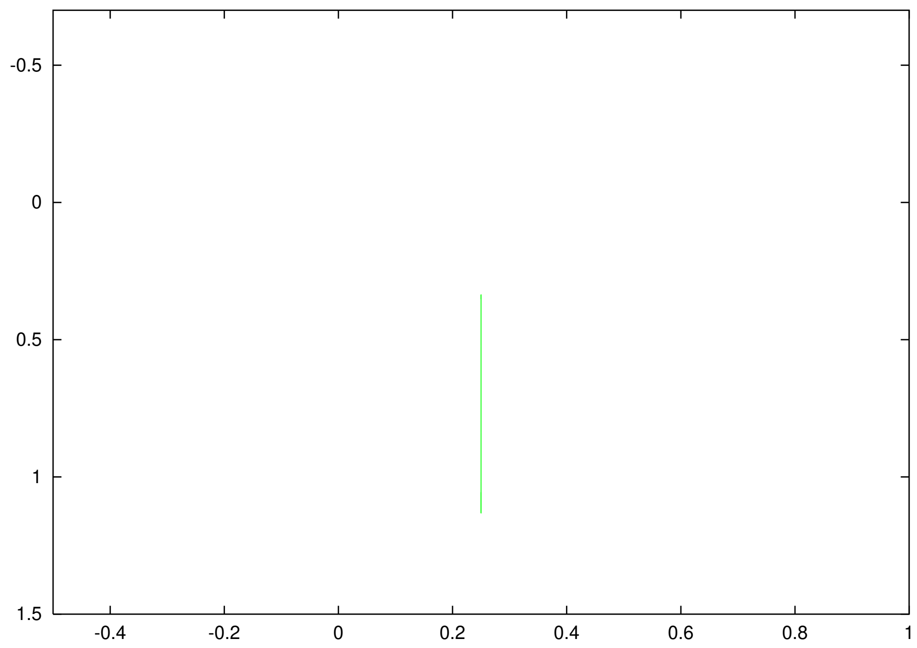

Abbildung 3.18: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.32$, $\lambda = 6.0$ und $\mu = 2.0$. Die dargestellte Lösung hat die gleiche Periode wie die Anregung.

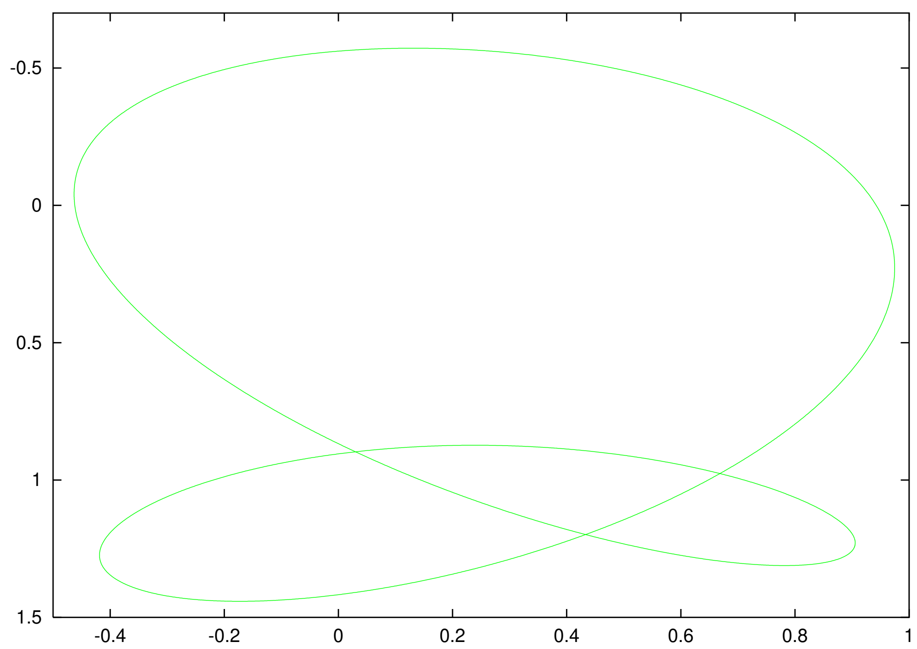

Abbildung 3.19: Bahn der Brückenpunktmasse in der x-y-Ebene bei Verwendung der Parameter $\varepsilon = 1E - 9$, $m = 1$, $k = 10$, $c = 0.05$, $l_0 = 0.3025$, $\lambda = 8.875$ und $\mu = 1.85$. Die dargestellte Lösung hat die gleiche Periode wie die Anregung.

diente in der zweiten Stufe, dem Newtonverfahren, als Startnäherung. Beim Newtonverfahren kam die sogenannte Kollokationsmethode zum Einsatz. In Tabelle 3.1 findet man die Entwicklung des Defektes (nur näherungsweise bestimmt) für den nullten bis vierten Schritt wiedergegeben. Es fällt auf, daß selbst die Näherung, die aus dem Runge-Kutta-Verfahren gewonnen wurde schon einen sehr geringen Defekt besitzt. Das Newtonverfahren schafft dann beide Male im ersten Schritt eine enorme Verringerung des Defektes und verbessert diesen dann in den folgenden Schritten nur noch geringfügig. Für die Bestimmung einer Näherungslösung mit geringem Defekt reicht also die Verwendung des Runge-Kutta-Verfahrens mit nur einem nachgeschalteten Newtonschritt.

Defektentwicklung beim Newtonverfahren		
Schritt	Lösung *(Lsg1)*	Lösung *(Lsg2)*
0	7.5501925e-11	9.0790226e-11
1	1.6784519e-14	2.5101062e-14
2	1.0806724e-14	2.2934523e-14
3	1.0488676e-14	2.2457330e-14
4	1.0580239e-14	2.2439403e-14

Tabelle 3.1: Entwicklung des Defektes (nur näherungsweise berechnet) beim Newtonverfahren.

Den beiden Näherungslösungen ist gemeinsam, daß die Funktion $(.)^+$ bei ihnen keinen Nulldurchgang besitzt, sondern stattdessen ständig echt größer null ist, die Seile also permanent gespannt sind. Dies ist kein Zufall, denn es ist leider so, daß bei Näherungslösungen, bei denen die Funktion $(.)^+$ einen Nulldurchgang hat, die Fourierentwicklung nur schlechte Approximationsergebnisse im Bereich von 10e-3 geliefert hat, was für die folgenden Punkte einen zu großen Wert darstellt. Auch Versuche die Näherungslösungen mit anderen Ansatzfunktionen, etwa Polynome oder Polynome mit trigonometrischen Funktionen kombiniert, zu approximieren führten nicht zum Erfolg. Die schlechte Approximationseigenschaft bei Näherungsfunktionen mit Nulldurchgang der Plusfunktion dürfte wohl durch den Knick in der zweiten Ableitung bedingt sein.

3.2.2 Zu Punkt 2: Berechnung der Konstanten δ:

In den Tabellen 3.2 und 3.3 findet man die Defekteeinschließungen $||d(\omega)||_2$ für die zu untersuchenden Lösungen *(Lsg1)* und *(Lsg2)*.

Zur Schreibweise ist zu sagen, daß z.B. 0.0[1,2] abkürzend für das Intervall [0.01, 0.02] stehen soll. Der Strich in einigen Feldern bedeutet, daß diese Berechnung nicht durchgeführt wurde, da das Verhalten aus schon berechneten Ergebnissen klar wird und das andere Einschließungsverfahren sich als besser herausgestellt hat. Er soll nicht andeuten, daß diese Berechnung nicht durchführbar wäre.

Beide Verfahren zeigen das erwartete Verhalten, daß mit Verdoppelung

Näherungslösung *Lsg1*		
Punkte	Riemanverfahren	Trapezformel
4096	0.0[000000,307426]	0.0[000000,953118]
8192	0.0[000000,153689]	0.0[000000,475728]
16384	0.00[00000,768410]	0.0[000000,237656]
32768	0.00[00000,384221]	0.0[000000,118776]
65536	0.00[00000,192139]	0.00[00000,593752]
131072	0.000[0000,961022]	0.00[00000,296844]
262144	0.000[0000,480842]	0.00[00000,148413]
524288	0.000[0000,240754]	-
1048576	0.000[0000,120711]	-
2097152	0.0000[000,606901]	-
4194304	0.0000[000,306801]	-

Tabelle 3.2: Intervalleinschließungen des Defekts für Näherungslösung *(Lsg1)* berechnet, zum einen durch Riemannsche Ober- und Untersummen und zum anderen durch die Trapezformel mit Fehlerabschätzung.

der Integrationspunkte, also Halbierung der eingesetzten Zeitintervalle, sich auch der Durchmesser der Defekteinschließung halbiert. Sowohl bei Lösung *(Lsg1)* als auch Lösung *(Lsg2)* liefert das Verfahren, das auf Verwendung der Riemannschen Ober- und Untersummen beruht bessere Ergebnisse als die Trapezformel mit Fehlerabschätzung. Dies liegt bei diesen beiden Beispielen daran, daß der Wert der zweiten Ableitung sehr groß ist. Im allgemeinen ist aber das Riemannverfahren nicht zwingendermaßen die bessere Wahl.

Von der Verwendung einer Integrationsformel mit höherer Konvergenzordnung wurde bei diesem Beispiel abgesehen da, wie man der Differentialgleichung des Modellproblems entnehmen kann (beachte die Definition von $(.)^+$) die zweite Ableitung beinahe einen Knick besitzen kann (für den Fall, daß die Funktion $(.)^+$ überhaupt zum Tragen kommt, die Seile könnten natürlich auch permanent gespannt sein). Falls also dieser beinahe Knick (beinahe nur deswegen, da die ursprüngliche Definition der Funktion $(.)^+$ durch die ε Definition ersetzt wurde) auftritt, haben die Ableitungen der zweiten Ableitung, die man ja für dieses Verfahren benötigt, zwingendermaßen einen sehr großen Wert. Als weiteres Problem steht dem Wunsch nach einem Verfahren höherer Konvergenzordnung, in diesem Fall die enorm umfangreichen Ableitungen, die damit auch rechenintensiv auszuwerten sind, im Wege.

Der ohne Intervalleinschließung berechnete Defekt liegt für Lösung *(Lsg1)* bei 1.0441719e-14 und für Lösung *(Lsg2)* bei 2.2941469e-14, sodaß zu vermuten ist, daß noch wesentlich bessere Einschließungen, als die hier gegebenen, berechenbar gewesen wären. Da sich aber mit jeder Verdoppelung der Integrationspunkte auch der Rechenaufwand verdoppelt, müssen die Einschließungen als Kompromiß zwischen Wartezeit und gewünschter Qualität des Defektes gewertet werden.

Näherungslösung *Lsg2*		
Punkte	Riemanverfahren	Trapezformel
4096	0.[0000000,1764955]	0.[0000000,8295783]
8192	0.0[000000,8818403]	0.[0000000,4128401]
16384	0.0[000000,4407624]	0.[0000000,2059323]
32768	0.0[000000,2203432]	0.[0000000,1028441]
65536	0.0[000000,1101637]	0.0[000000,5139157]
131072	0.00[00000,5508271]	0.0[000000,2568815]
262144	0.00[00000,2754386]	0.0[000000,1284217]
524288	0.00[00000,1377480]	0.00[00000,6420609]
1048576	0.000[0000,6890390]	0.00[00000,3210185]
2097152	0.000[0000,3448199]	0.00[00000,1605063]
4194304	0.000[0000,1727108]	-
8388608	0.0000[0000,866569]	-
16777216	0.0000[0000,436309]	-
33554432	0.0000[0000,221196]	-

Tabelle 3.3: Intervalleinschließungen des Defekts für Näherungslösung *(Lsg2)* berechnet, zum einen durch Riemannsche Ober- und Untersummen und zum anderen durch die Trapezformel mit Fehlerabschätzung.

Als Resultat für die eigentlich gesuchte Konstante δ kann man den Tabellen als besten Wert

$$\boxed{\delta_{(Lsg1)} := 0.0000306801} \qquad \boxed{\delta_{(Lsg2)} := 0.0000221196}$$

entnehmen.

3.2.3 Zu Punkt 3: Berechnung der Konstanten σ

Dadurch, daß der Operator L die einfachere Gestallt

$$L\left[\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right](t) = -m\begin{pmatrix} \ddot{u}_1 \\ \ddot{u}_2 \end{pmatrix}(t) - c\begin{pmatrix} \dot{u}_1 \\ \dot{u}_2 \end{pmatrix}(t) + \begin{pmatrix} D^1F_1(t) & D^2F_1(t) \\ D^1F_2(t) & D^2F_2(t) \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(t)$$

als die allgemeine in 1.3.3 verwendete besitzt, ergeben sich bei der Berechnung eines Vergleichsproblems einige Vereinfachungen, die hier noch vorgestellt werden sollen. Die prinzipielle Vorgehensweise, also das Abschätzen von $< L[u], L[u] >_2$ nach unten, bleibt aber bestehen. Die Aufsplittung des Innnenproduktes sieht in diesem Fall wie folgt aus:

$$\begin{aligned} < L[u], L[u] >_2 &=< -m\ddot{u} - c\dot{u} + DFu, -m\ddot{u} - c\dot{u} + DFu >_2 \\ &= m^2 < \ddot{u}, \ddot{u} >_2 \underbrace{+mc < \ddot{u}, \dot{u} >_2}_{(A1)} \underbrace{-m < \ddot{u}, DFu >_2}_{(B1)} \\ &\quad \underbrace{+mc < \dot{u}, \ddot{u} >_2}_{(A2)} + c^2 < \dot{u}, \dot{u} >_2 \underbrace{-c < \dot{u}, DFu >_2}_{(C1)} \end{aligned}$$

$$\underbrace{-m < DFu, \ddot{u} >_2}_{\text{(B2)}} \underbrace{-c < DFu, \dot{u} >_2}_{\text{(C2)}} + < DFu, DFu >_2 .$$

Es bleiben sechs unterklammerte Ausdrücke übrig, die der Reihe nach untersucht werden sollen. Begonnen wird mit den beiden A-Termen:

$$\begin{aligned}
\text{(A1)} + \text{(A2)} &= mc \int_0^{\frac{2\pi}{\mu}} \ddot{u}^t \dot{u}\, dt + mc \int_0^{\frac{2\pi}{\mu}} \dot{u}^t \ddot{u}\, dt \\
&= \underbrace{mc\, \dot{u}^t \dot{u}|_0^{\frac{2\pi}{\mu}}}_{=0} - mc \int_0^{\frac{2\pi}{\mu}} \dot{u}^t \ddot{u}\, dt + mc \int_0^{\frac{2\pi}{\mu}} \dot{u}^t \ddot{u}\, dt \\
&= 0.
\end{aligned}$$

Die beiden A-Terme heben sich also, da m und c reelle Konstanten sind und daher aus dem Innenprodukt herausgezogen werden können, gegenseitig auf, was beim allgemeinen Operator L nicht ging, da dort vor dem Vektor mit den ersten Ableitungen eine eventuell zeitabhängige Matrix stand. Bei der Untersuchung der beiden B-Terme ergeben sich keine großen Änderungen, abgesehen davon, daß man wieder die reellen Konstanten aus dem Innenprodukt herausziehen kann:

$$\begin{aligned}
\text{(B1)} + \text{(B2)} &= -m \int_0^{\frac{2\pi}{\mu}} (\ddot{u}^t DFu + u^t DF^t \ddot{u})\, dt \\
&= -2m \int_0^{\frac{2\pi}{\mu}} \ddot{u}^t DFu\, dt \\
&= -2m \left(\underbrace{\dot{u}^t DFu|_0^{\frac{2\pi}{\mu}}}_{=0} - \int_0^{\frac{2\pi}{\mu}} \dot{u}^t \dot{DF} u\, dt - \int_0^{\frac{2\pi}{\mu}} \dot{u}^t DF \dot{u}\, dt \right) \\
&\geq 2m \left(- \int_0^{\frac{2\pi}{\mu}} |\dot{u}|_e |\dot{DF}|_e |u|_e\, dt + \int_0^{\frac{2\pi}{\mu}} \dot{u}^t DF \dot{u}\, dt \right) \\
&\geq -m |\dot{DF}|_e \int_0^{\frac{2\pi}{\mu}} \dot{u}^t \dot{u} + u^t u\, dt + 2m \int_0^{\frac{2\pi}{\mu}} \dot{u}^t DF \dot{u}\, dt.
\end{aligned}$$

Auch bei der Abschätzung der C-Terme sind die Änderungen gegenüber dem früheren Ergebnis nicht gravierend:

$$\begin{aligned}
\text{(C1)} + \text{(C2)} &= -c \int_0^{\frac{2\pi}{\mu}} \dot{u}^t DFu + u^t DF^t \dot{u}\, dt \\
&= -2c \int_0^{\frac{2\pi}{\mu}} \dot{u}^t DFu\, dt \\
&\geq -2c \int_0^{\frac{2\pi}{\mu}} |\dot{u}|_e |DF|_e |u|_2\, dt \qquad \geq -c |DF|_e \int_0^{\frac{2\pi}{\mu}} \dot{u}^t \dot{u} + u^t u\, dt.
\end{aligned}$$

Die obigen Abschätzungen kann man, wenn man die Matrizen

$$M_1 = \begin{pmatrix} c^2 - c|DF|_e - m|\dot{DF}|_e - 2mD^1F_1 & 2mD^2F_1 \\ 2mD^1F_2 & c^2 - c|DF|_e - m|\dot{DF}|_e + 2mD^2F_2 \end{pmatrix},$$

$$M_0 = \begin{pmatrix} (D^1F_1)^2 + (D^1F_2)^2 - c|DF|_e - m|\dot{DF}|_e & D^1F_1 \cdot D^2F_1 + D^1F_2 \cdot D^2F_2 \\ D^1F_1 \cdot D^2F_1 + D^1F_2 \cdot D^2F_2 & (D^2F_1)^2 + (D^2F_2)^2 - c|DF|_e - m|\dot{DF}|_e \end{pmatrix}$$

bildet zusammenfassend schreiben und weiter abschätzen durch:

$$\begin{aligned} < L[u], L[u] >_2 &\geq \int_0^{\frac{2\pi}{\mu}} m^2 \ddot{u}^t \ddot{u} + \dot{u} M_1 \dot{u} + u^t M_0 u \, dt \\ &= \int_0^{\frac{2\pi}{\mu}} m^2 \ddot{u}^t \ddot{u} + \dot{u} \tfrac{1}{2}(M_1 + M_1^t)\dot{u} + u^t \tfrac{1}{2}(M_0 + M_0^t) u \, dt \\ &\geq \int_0^{\frac{2\pi}{\mu}} m^2 \ddot{u}^t \ddot{u} + \tfrac{1}{2}\lambda_{\min}(M_1 + M_1^t)\dot{u}^t\dot{u} + \tfrac{1}{2}\lambda_{\min}(M_0 + M_0^t) u^t u \, dt \\ &\geq \int_0^{\frac{2\pi}{\mu}} \underbrace{m^2}_{\alpha} \ddot{u}^t \ddot{u} + \underbrace{\tfrac{1}{2} \min_{t \in [0, \frac{2\pi}{\mu}]} \lambda_{\min}(M_1 + M_1^t)}_{\beta} \dot{u}^t \dot{u} \\ &\quad + \underbrace{\tfrac{1}{2} \min_{t \in [0, \frac{2\pi}{\mu}]} \lambda_{\min}(M_0 + M_0^t)}_{\gamma} u^t u \, dt \end{aligned}$$

Diese Vorgehensweise wurde auch bei der Berechnung der Vergleichsprobleme für die Homothopieen, die in Abbildung 3.20 und in Abbildung 3.21 zu finden sind verwendet. Dabei wurde das Vergleichsproblem allerdings nicht exakt mit Intervallarithmetik bestimmmt, sondern alle Berechnungen nur näherungsweise mit normaler Arithmetik durchgeführt. Man kann aber dennoch sehr gut die Monotonie der Eigenwerte im Homothopieparameter erkennen.

In jedem Homothopieschritt müßte man eigentlich die Eigenwerte mit Hilfe des Satzes von Weinstein, Kato, Rayleigh-Ritz oder Lehmann einschließen. Dies ist hier nicht geschehen. Statt dessen wurde in jedem der 200 Homothopieschritte (zum Einschließen von Eigenwerten sind nicht soviele Schritte nötig, die hohe Schrittanzahl wurde nur gewählt um ein zusammenhängendes Bild zu bekommen) das Rayleigh-Ritz Eigenwertproblem, das nur Näherungseigenwerte und Näherungseigenfunktionen liefert, mit Hilfe des QR-Verfahrens gelöst. Der Raum der Ansatzfunktionen für das Rayleigh-Ritz Verfahren bestand dabei aus 402 trigonometrischen Polynomen.

Es ist klar, daß dadurch, daß weder das Vergleichsproblem noch die einzelnen Homothopieschritte exakt berechnet wurden, man nicht mehr vom Beweis der Existenz von Lösungen sprechen kann. Die beiden Abbildungen legen meiner Ansicht nach aber nahe, daß einer exakten Berechenung der Konstanten σ nichts im Wege stehen würde.

Es fällt auf, daß beim Vergleichsproblem einige Eigenwerte mehrfach auftreten und auch im Verlauf der Homothopie einige Eigenwerte clustern. Dies behindert die Durchführbarkeit des Homothopiealgorithmus aber nicht, man verliert dann zwar in jedem Homothopieschritt mehr als nur einen Eigenwert, was aber nicht stört, wenn man mit genügenden gestartet ist.

Als Wert für die Konstante σ kann man den Abbildungen 3.20 und 3.21

$$\boxed{\sigma_{(Lsg1)} = 5.16329258} \qquad \boxed{\sigma_{(Lsg2)} = 0.313831534}$$

entnehmen.

3.2.4 Zu Punkt 4: Berechnung der Konstanten K

Ziel dieses Abschnittes ist es eine Konstante $K > 0$ zu finden, die die Bedingung

$$||u||_\infty \leq K||L[u]||_2 \quad \forall u \in R$$

erfüllt. Dabei wird das Problem zerlegt in die Teilschritte, Konstanten K_0, K_1, C_0 und C_1 zu finden, die den Bedingungen

$$\begin{aligned} ||u||_2 &\leq K_0||L[u]||_2, \\ ||\dot{u}||_2 &\leq K_1||L[u]||_2 \text{ und} \\ ||u||_\infty &\leq C_0||u||_2 + C_1||\dot{u}||_2. \end{aligned}$$

für alle $u \in R$ genügen.

Die einzelnen Teilschritte zur Berechnung der Konstanten K wurden, wie schon die Berechnung der Konstanten σ im letzten Abschnitt, nur näherungsweise, also unter Verwendung der normalen Arithmetik des Computers berechnet. Dies erschien als sinnvoll, da die Berechnung der Konstanten K direkt auf der Berechnung der Konstanten σ aufsetzt, die ja schon nicht exakt bestimmt wurde. Die exakte Durchführung dieses Abschnittes stellt aber kein wirkliches Problem dar, da es sich nur um elementare Berechungen handelt. Selbst die einschließende Berechnung der Eigenwerte zur Bestimmung der Konstanten γ ist bei einer 2×2-Matrix sehr einfach auszuführen.

Als Ergebnisse dieses Schrittes erhält man die Konstanten

$K_{0,(Lsg1)} = 0.440085067$	$K_{0,(Lsg2)} = 1.78505545$
$K_{1,(Lsg1)} = 2.06492145$	$K_{1,(Lsg2)} = 8.09404182$
$C_{0,(Lsg1)} = 1.16382623$	$C_{0,(Lsg2)} = 1.1440936$
$C_{1,(Lsg1)} = 0.248039723$	$C_{1,(Lsg2)} = 0.252317761$
$K_{(Lsg1)} = 1.024365089$	$K_{(Lsg2)} = 4.084541025$

für die beiden Lösungen *(Lsg1)* und *(Lsg2)*.

3.2.5 Zu Punkt 5: Finden einer monoton wachsenden Funktion G

Die Abbildungen 3.22 bis 3.25 zeigen die Ergebnisse der Berechnung der monoton wachsenden Funktion G, die die Abschätzung

$$||F(.,\omega + y) - F(.,\omega) - C \cdot y||_2 \leq G(||y||_\infty) \quad (\forall y \in \mathbb{R}^n)$$

erfüllt. Die Zahlennummern der einzelnen Kurven G geben an, mit wievielen Integrationspunkten diese berechnet wurden.

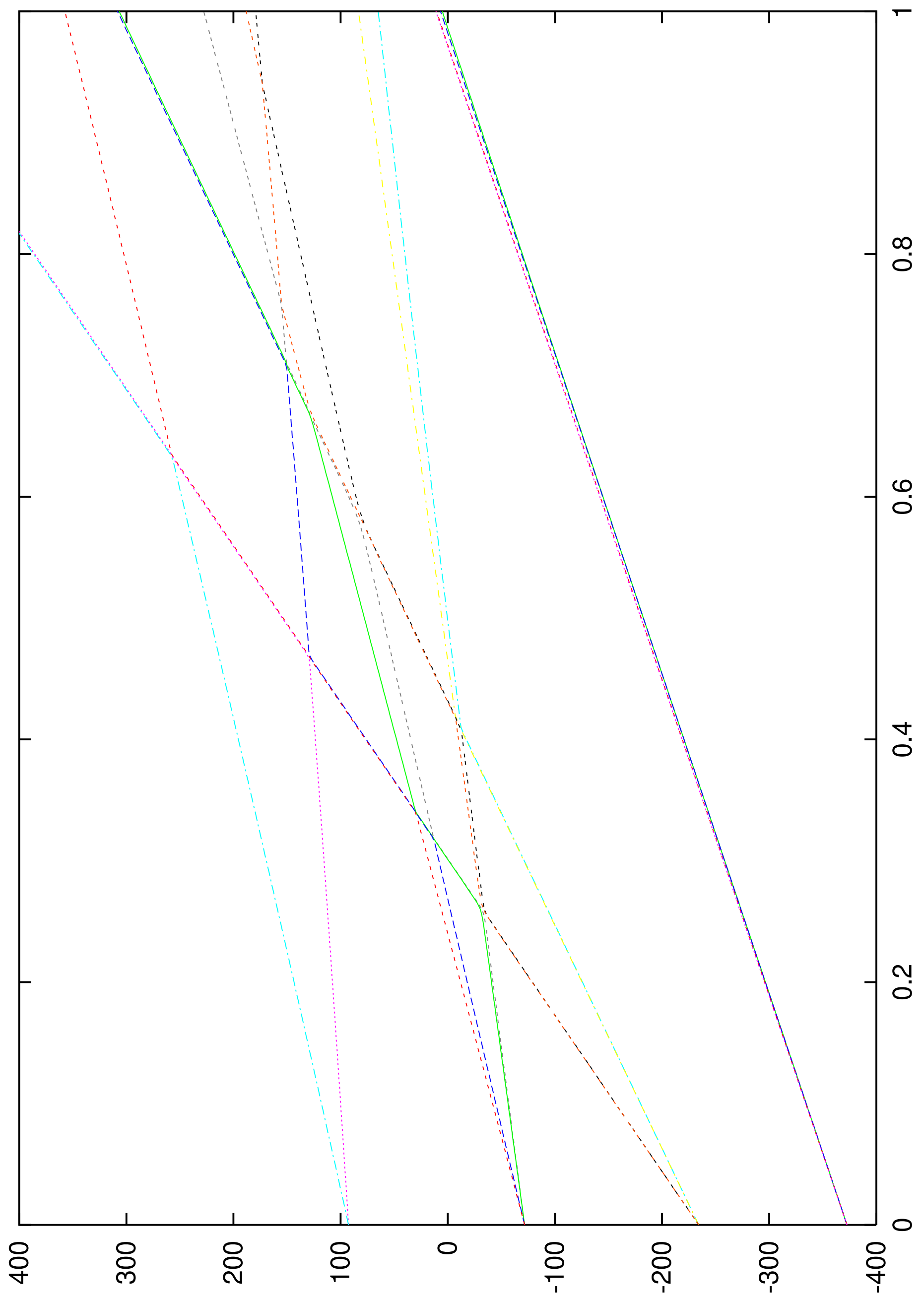

Abbildung 3.20: Ergebnis des Homothopie-Algorithmus zur Verbindung der Eigenwerte eines einfachen Vergleichsproblems mit den Eigenwerten von L^*L für Lösung *(Lsg1)*.

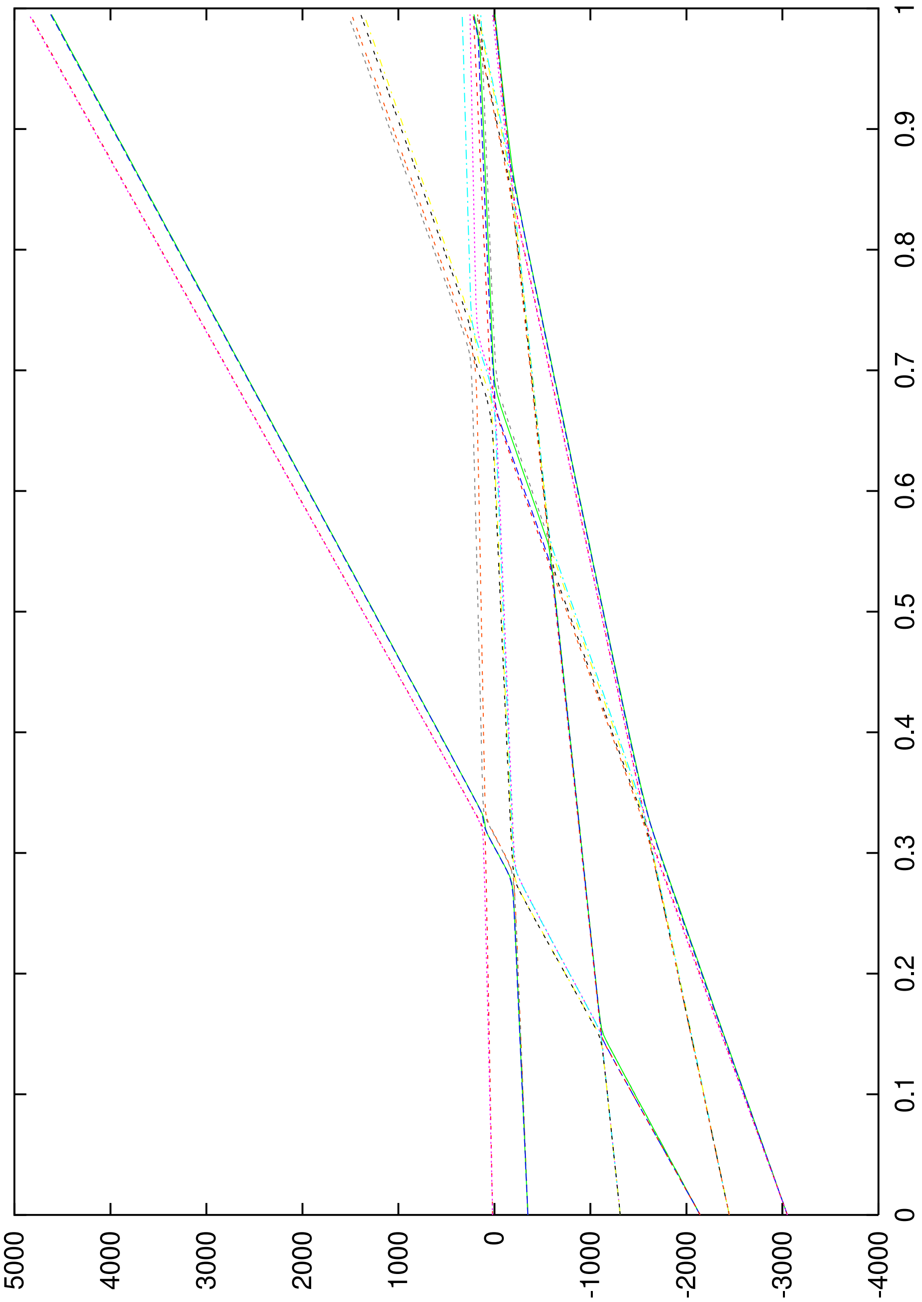

Abbildung 3.21: Ergebnis des Homothopie-Algorithmus zur Verbindung der Eigenwerte eines einfachen Vergleichsproblems mit den Eigenwerten von L^*L für Lösung *(Lsg2)*.

Die Berechnung erfolge gemäß den zwei unterschiedlichen Sätzen, die in 1.3.5 vorgestellt wurden. Abbildung 3.22 und 3.24 wurden mit Hilfe des ersten Satzes, also ausschließlich unter Verwendung höchstens erster Ableitungen von F berechnet, wohingegen in Abbildungen 3.23 und 3.25 die Ergebnisse zu finden sind, die man erhält, wenn man Satz zwei, der auch zweite Ableitungen von F verwendet, zur Berechnung der gesuchten Funktion G anwendet. In den Bemerkungen zum zweiten Satz in 1.3.5 war die Rede davon, daß dadurch, daß in diesem Satz α^2 ausgeklammert wird, die Berechnung des Restes relativ grob durchgeführt werden kann. Diese Behauptung findet man durch die Abbildungen recht gut bestätigt, denn während bei den Abbildungen die nach Satz eins berechnet wurden eine Verdoppelung der Integrationspunkte, also eine Halbierung der Durchmesser der Integrationsintervalle, noch eine erhebliche Verringerung der Funktionswerte von G nach sich zieht, ist bei den Abbildungen, die nach Satz zwei berechnet wurden bei einer Integrationspunkteerhöhung keine wesentliche Veränderung mehr festzustellen.

Es mag so scheinen, daß Satz zwei die bessere Wahl darstellt, was bei den hier vorgestellten Lösungen auch tatsächlich so war. Dabei muß man aber auch sehen, das die Berechnung der zweiten Ableitung für Satz zwei einen wesentlich höheren Rechenaufwand benötigt und damit ein mit 8192 Integrationspunkten nach Satz eins durchgeführte Berechnung eventuell immer noch weniger Zeit benötigt als die Berechnung mit 4096 Integrationspunkten nach Satz zwei. Außerdem sind mir bei Berechnungen mit verschiedenen Parametern auch solche Lösungen begegnet, deren zweite Ableitung einen so großen Wert aufwies, daß der Effekt, der durch das Ausklammern von α^2 zustande kommt erst für eine große Anzahl von Integrationspunkten zum tragen kam.

Die hier vorgestellten Ergebnisse sind, anders als die Ergebnisse bei der Berechnung der Konstanten σ und K, exakt, das soll heißen, daß sie in Intervallarithmetik berechnet wurden. Zur Berechnung wurde dabei die $x-y-$Ebene in $200 \cdot 200$ Quadrate mit einer Kantenlänge von 0.0001 aufgeteilt. Außerdem wurde noch eine Unterteilung des Integrationsintervalls zur einschließenden Durchführung der Integration nötig. Dadurch wurde dieser Punkt zu einem der rechenintensivsten Aufgaben des gesammten Algorithmus.

Da das Ziel der Einschließung der exakten Lösung mit einem Fehler von 0.0001 gelingen soll, werden hier die beiden besten Funktionswerte von G für diesen Wert noch einmal getrennt aufgeführt:

$$\boxed{G_{(Lsg1)}(0.0001) = 5.54588e - 07} \quad \boxed{G_{(Lsg2)}(0.0001) = 3.40391e - 07}.$$

3.2.6 Zu Punkt 6: Bestimmung der Konstanten α

In den Abbildungen 3.26 und 3.27 werden die Berechnungen der letzten Punkte zusammengeführt, um ein α zu bestimmen, das der Bedingung

$$\delta \leq \frac{\alpha}{K} - G(\alpha)$$

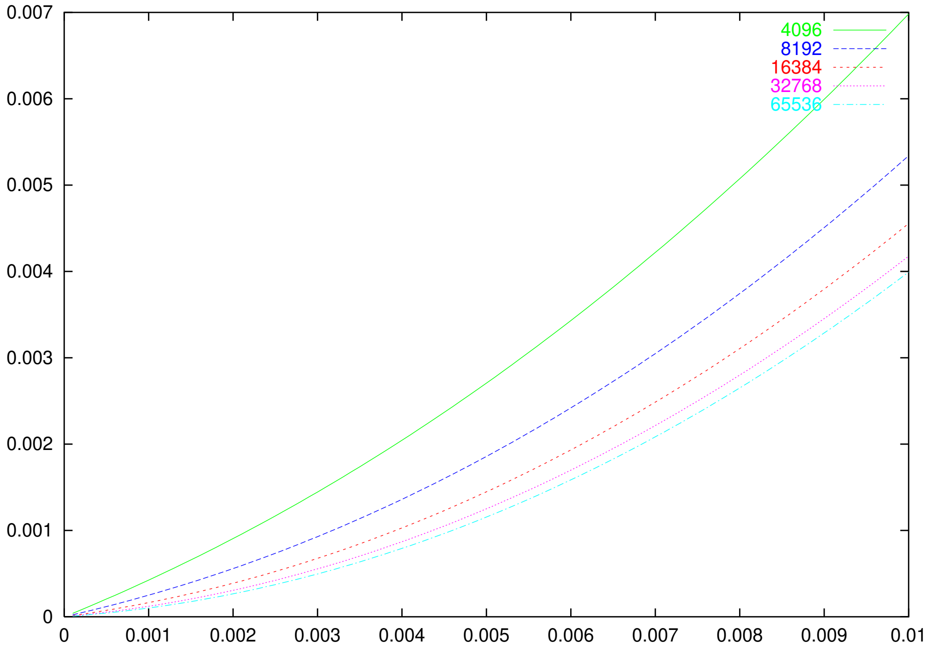

Abbildung 3.22: Monoton wachsenden Funktion G zu Lösung *(Lsg1)* berechnet unter Verwendung ausschließlich erster Ableitungen.

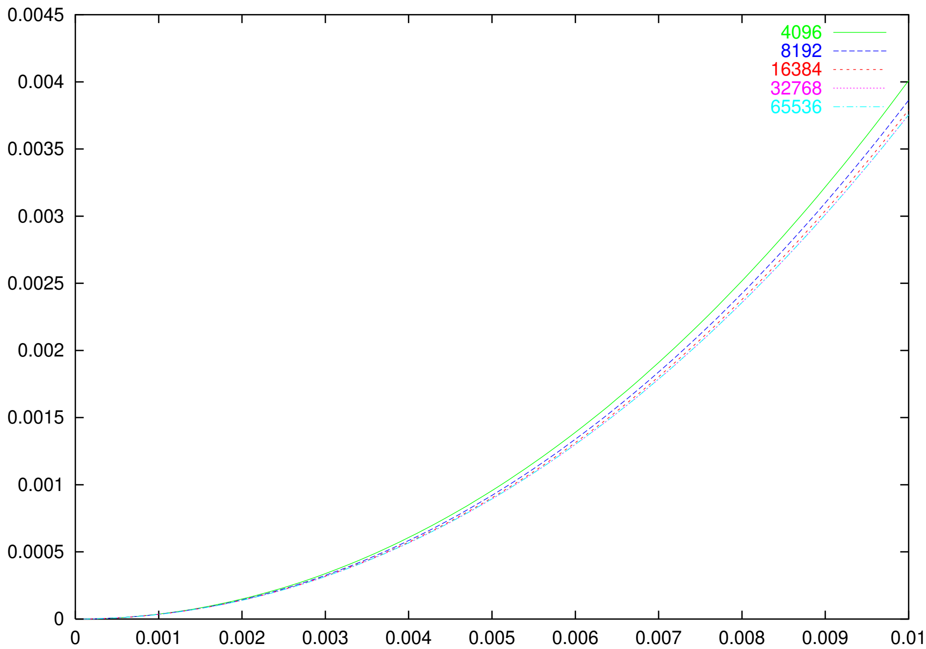

Abbildung 3.23: Monoton wachsenden Funktion G zu Lösung *(Lsg1)* berechnet unter Verwendung von zweiten Ableitungen.

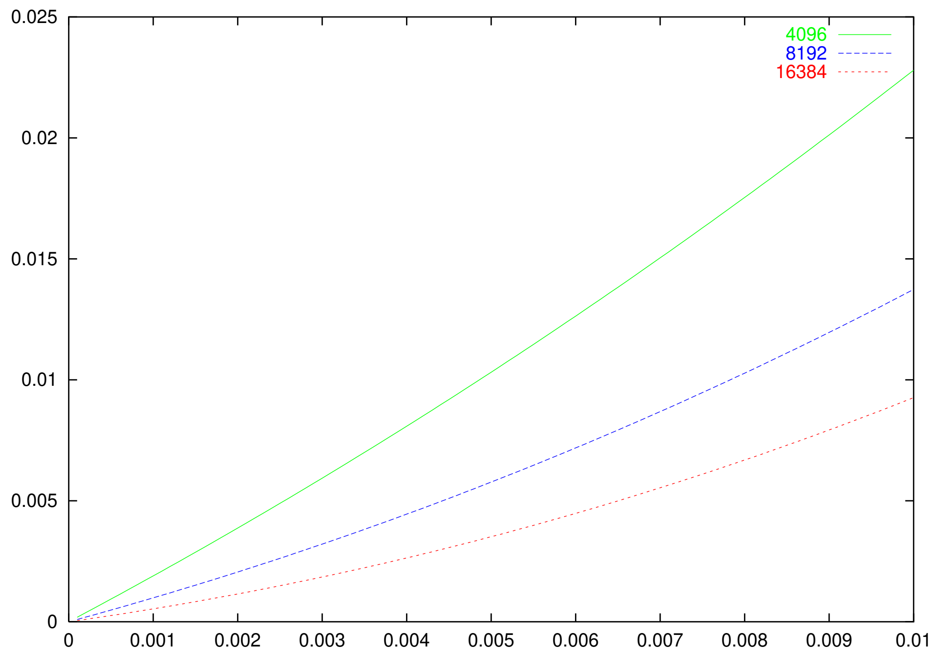

Abbildung 3.24: Monoton wachsenden Funktion G zu Lösung *(Lsg2)* berechnet unter Verwendung ausschließlich erster Ableitungen.

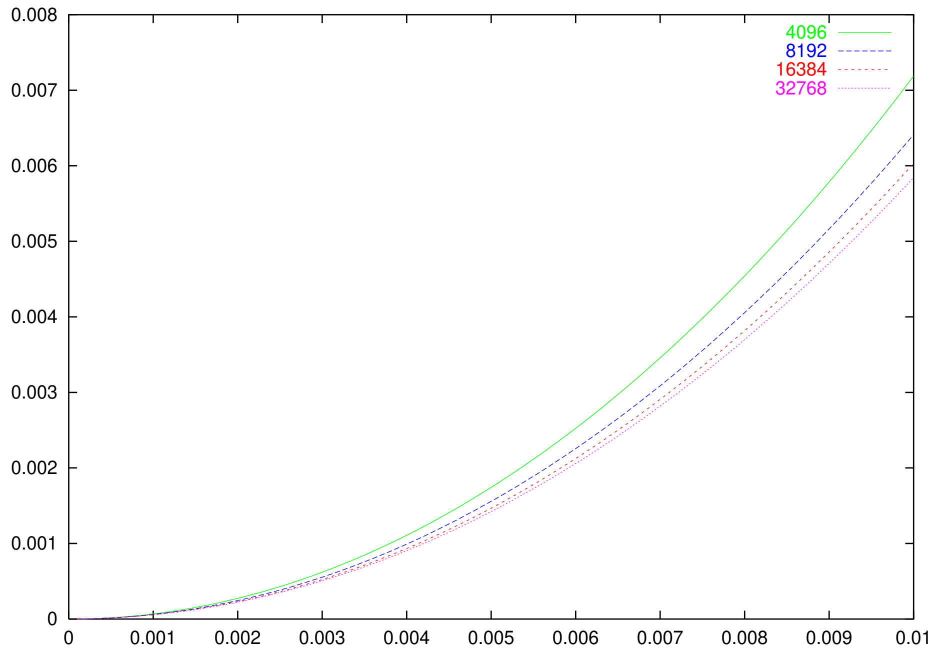

Abbildung 3.25: Monoton wachsenden Funktion G zu Lösung *(Lsg1)* berechnet unter Verwendung von zweiten Ableitungen.

genügt. Man kann deutlich sehen, daß dies in beiden Fällen gelingt (beachte: durchgezogene Linie gibt den gesichert berechneten Defekt δ wieder). Bei der Erzeugung der Bilder wurden selbstverständlich immer die besten Resultate der letzten Punkte verwendet, also beispielsweise der kleinste gesichert berechnete Defekt δ, die niedrigste monotone Funktion G und so weiter. Für den Wert $\alpha = 0.0001$ ergibt sich für die beiden Näherungslösungen:

$$\begin{aligned}
\frac{\alpha}{K_{(Lsg1)}} - G_{(Lsg1)}(\alpha) &= \frac{0.0001}{1.024365089} - 5.54588e-07 \\
&\geq 0.00009 \geq 0.0000306801 = \delta_{(Lsg1)} \quad \text{ok}, \\
\frac{\alpha}{K_{(Lsg1)}} - G_{(Lsg1)}(\alpha) &= \frac{0.0001}{4.084541025} - 3.40391e-07 \\
&\geq 0.00002414 \geq 0.0000221196 = \delta_{(Lsg2)} \quad \text{ok}.
\end{aligned}$$

Somit gilt also sowohl für Lösung *(Lsg1)* als auch für Lösung *(Lsg2)*, daß jeweils ein $U \in C_2^2(0, \frac{2\pi}{\mu})$ existiert mit der gleichen Periode wie die Näherungslösung, für das gilt:

$$\boxed{||U_{(Lsg1)} - \omega_{(Lsg1)}||_\infty \leq 0.0001} \quad \boxed{||U_{(Lsg2)} - \omega_{(Lsg2)}||_\infty \leq 0.0001}.$$

3.2.7 Zusammenfassung:

Der exemplarisch auf zwei Beispielen angewandte Algorithmus zum Beweis der Existenz und Einschließung von Lösungen von periodischen Systemen von Randwertproblemen hat wohl in den letzten Abschnitten seine Verwendbarkeit demonstriert, auch wenn die Berechnung der Konstanten σ und K nicht exakt d.h. mit Intervallrechnung durchgeführt wurde. Die nicht exakte Durchführung der Bestimmung der beiden Konstanten mag aber als verzeilich hingenommen werden, da einem die Abbildungen 3.20 und 3.21 nahe legen, daß es sich bei dem näherungsweise berechneten Eigenwert tatsächlich um den kleinsten Eigenwert von L^*L handelt.

Es mag vielleicht so scheinen, daß der Algorithmus doch nicht so allgemein angewandt werden kann, wie behauptet, da er an keiner Lösung mit einem Nulldurchgang der Funktion $(.)^+$ durchgeführt wurde. Dem ist aber nicht so, denn es handelt sich hierbei eher um ein Problem der Rechenleistung, als um ein Problem des Algorithmus. Eine Voraussetzung des Algorithmus ist, daß die Lösung einen hinreichend kleinen Defekt besitzt. Dies läßt sich aber bei diesen Lösungen, da sie sich schlecht approximieren lassen, nur mit umfangreichen Fourierentwicklungen erreichen, was aber zur Folge hat, daß alle folgenden Berechnungen mit dieser großen Entwicklung durchgeführt werden müssen. Mit der heutigen Rechenleistung stößt man dann aber sehr schnell an Grenzen, wenn man nicht gerade bereit ist einige Monate auf das Ergebnis zu warten. Wenn die derzeitige Entwicklung der Computertechnik (jedes Jahr beinahe eine Verdopplung der Rechengeschwindigkeit) auch in Zukunft anhält ist es aber nur noch eine Frage der Zeit, bis auch solche Lösungen mit dem Algorithmus in akzeptabler Zeit untersucht werden können.

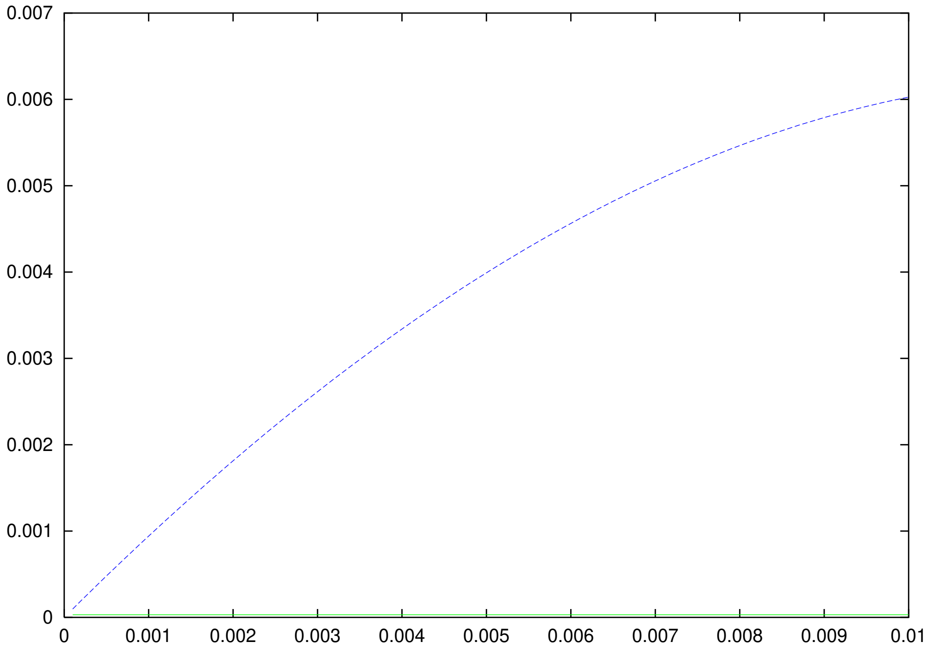

Abbildung 3.26: Funktion $\frac{\alpha}{K_{(Lsg1)}} - G_{(Lsg1)}(\alpha)$ aufgetragen über α für die Lösung *(Lsg1)*. Die gerade Linie ist der für diese Lösung gesichert berechnete Defekt.

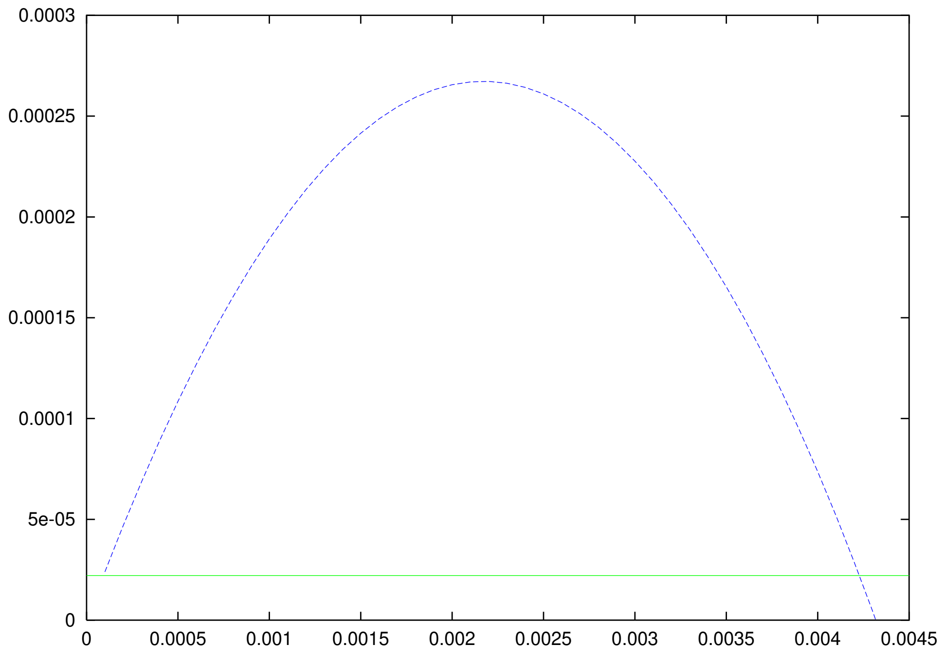

Abbildung 3.27: Funktion $\frac{\alpha}{K_{(Lsg2)}} - G_{Lsg2}(\alpha)$ aufgetragen über α für die Lösung *(Lsg2)*. Die gerade Linie ist der für diese Lösung gesichert berechnete Defekt.

Die Rechenzeit stellt aber nicht nur für schlecht zu approximierende Lösungen ein Problem da, sondern auch für gut zu approximierende (wenn hier auch ein lösbares). Durch auf das Problem angepaßte Teile des Algorithmus kann man hier aber Verbesserungen erzielen. Dadurch, daß der Algorithmus aus verschiedenen Bausteinen aufgebaut ist, die unabhängig voneinander durchgeführt werden können bietet es sich geradezu an einige dieser Bausteine auf das betrachtete Problem zuzuschneiden. Auch die Idee den Algorithmus zu parallelisieren ist naheliegend. Die einzelnen Teile des Algorithmus können sowieso unabhängig voneinander auf verschiedenen Rechnern gleichzeitig durchgeführt werden. Aber auch nahezu jeder Teil des Algorithmus bietet Möglichkeiten der Parallelisierung. Beispielhaft sei hier die Berechnung des Defektes δ genannt. Hier kann man eine Verteilung auf verschiedene Prozessoren einfach dadurch erreichen, daß man das Integrationsgebiet in soviele Teilgebiete zerlegt, wie Prozessoren vorhanden sind und jeden Prozessor genau eines der Teilgebiete bearbeiten läßt.

Abschließend läßt sich sagen, daß der Algorithmus zwar sehr rechenintensiv ist, daß man dies aber durch bestimmte Ideen in den Griff bekommen kann. Dadurch, daß er aber zu Existenz- und Einschließungsbeweisen führt, die man auf andere Art und Weise kaum erhalten kann behält er auch mit diesem Manko seine große Bedeutung.

Kapitel 4

Ergänzungen:

Im Kapitel 4 Ergänzungen, sind die Hilbertraumtheorie, die numerischen Methoden und die Einführung in die Intervallarithmetik zu finden, kurzum all die Dinge, auf denen der Algorithmus in Kapitel 1 fußt. Es gibt zweierlei Gründe, weswegen diese Dinge in einem extra Kapitel zu finden sind und nicht einfach im ersten Kapitel aufgenommen wurden. Der Erste ist, daß sonst Kapitel 1 zu umfangreich geworden wäre und seine Übersichtlichkeit verloren hätte. Zum Zweiten ist es möglich, daß man Teile dieser Grundlagen schon kennt und diese deswegen nicht noch einmal lesen möchte, was aber schwierig wäre, falls sie in Kapitel 1 mit eingewoben wären.

4.1 Eine kurze Einführung in die Intervallrechnung:

An vielen Stellen des Algorithmus für gesicherte Exisztenz- und Einschließungsergebnisse für periodische Systeme von Randwertproblemen (etwa bei der Berechnung des Defektes δ, dem finden einer monotonen Funktion G und der Einschließung der Eigenwerte des linearisierten Operators) werden Sachverhalte aus der Intervallrechnung verwendet, weshalb es nötig erscheint, eine kurze Einführung und einen groben Überblick über diesen Bereich zu geben. Dieser Abschnitt hat aus diesem Grunde keinesfalls den Anspruch der Vollständigkeit, sondern enthält nur die Bereiche der Intervallrechnung, die in Hinsicht auf Kapitel 1 von Bedeutung sind. Auch die Beweise einiger Sätze wurden in Hinsicht auf den gewünschten Umfang dieses Abschnittes teilweise weggelassen.

4.1.1 Reelle Intervallrechnung

Schreibkonventionen:

In diesem Abschnitt seien Elemente aus $\mathbb{R}$ mit kleinen Buchstaben $a, b, c, \ldots$ bezeichnet. Die Intervalle hingegen, also Teilmengen der reellen Zahlen vom folgenden Typ

$$A = [\underline{a}, \overline{a}] = \{t : \underline{a} \leq t \leq \overline{a} \text{ mit } \underline{a}, \overline{a} \in \mathbb{R}\}$$

seien mit Großbuchstaben bezeichnet.

Definition ($\mathbf{I}(\mathbb{R})$):

Die Menge aller abgeschlossenen reellen Intervalle $A = [\underline{a}, \overline{a}]$ sei bezeichnet durch $\mathrm{I}(\mathbb{R})$.

Definition (Intervalloperationen):

Sei $* \in \{+, -, \cdot, :\}$ die bekannte Operation für reelle Zahlen. Dann gelte für $A, B \in \mathrm{I}(\mathbb{R})$:

$$A \mathbin{\Diamond\!\!\!\!\!*} B = \{z = a * b : a \in A, b \in B\}.$$

Bei der Division wird zusätzlich $0 \notin B$ vorausgesetzt, was im folgenden nicht mehr jeweils erwähnt wird. Verknüpfungen in $\mathrm{I}(\mathbb{R})$ werden mit den gleichen Symbolen, wie die reellen Operationen, umrahmt von einer Raute, bezeichnet, um kenntlich zu machen, daß es sich um eine Intervalloperation handelt. Für die explizite Durchführung der Verknüpfungen ergeben sich mit $A = [\underline{a}, \overline{a}]$ und $B = [\underline{b}, \overline{b}]$ die folgenden Rechenregeln:

$$\begin{aligned}
A \mathbin{\Diamond\!\!\!\!\!+} B &= [\underline{a} + \underline{b}, \overline{a} + \overline{b}],\\
A \mathbin{\Diamond\!\!\!\!\!-} B &= [\underline{a} - \underline{b}, \overline{a} - \overline{b}] = A \mathbin{\Diamond\!\!\!\!\!+} [-1, -1] \mathbin{\Diamond\!\!\!\!\!\cdot} B,\\
A \mathbin{\Diamond\!\!\!\!\!\cdot} B &= [\min\{\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}\}, \max\{\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}\}],\\
A \mathbin{\Diamond\!\!\!\!\!:} B &= [\underline{a}, \overline{a}] \mathbin{\Diamond\!\!\!\!\!\cdot} \left[\frac{1}{\overline{b}}, \frac{1}{\underline{b}}\right].
\end{aligned}$$

Denn $z = f(x, y)$ mit $f(x, y) = x * y$, $* \in \{+, -, \cdot, :\}$, ist auf der kompakten Menge $A \times B$ eine stetig Funktion und nimmt daher einen größten und kleinsten Wert an, und außerdem wird jeder Wert zwischen diesen beiden Extrema angenommen. $A \mathbin{\Diamond\!\!\!\!\!*} B$ ist somit wieder ein abgeschlossenes reelles Intervall. Daß die angegebenen Grenzen jeweils das Minimum und Maximum von $f(x, y)$ darstellen, ist durch leichtes nachrechnen ersichtlich.

Die Menge $\mathrm{I}(\mathbb{R})$ ist also bezüglich der eingeführten Verknüpfungen abgeschlossen.

Definition (einstellige Operationen in $\mathbf{I}(\mathbb{R})$):

Bezeichnet $r(x)$ eine stetige einstellige Operation in $\mathbb{R}$, dann ist durch

$$r(X) = [\min_{x \in X} r(x), \max_{x \in X} r(x)]$$

eine (zugehörige) einstellige Operation in $\mathrm{I}(\mathbb{R})$ erklärt (z.B. lassen sich so die einstelligen Operationen $\sin X$, $X^k, \ldots$ definieren).

Satz (elementare Eigenschaften von Intervalloperationen):

Seien A, B und C Intervalle aus $\mathrm{I}(\mathbb{R})$, dann gilt:

1. $A \mathbin{\Diamond\!\!\!\!\!+} B = B \mathbin{\Diamond\!\!\!\!\!+} A$, $\quad A \mathbin{\Diamond\!\!\!\!\!\cdot} B = B \mathbin{\Diamond\!\!\!\!\!\cdot} A$ (Kommutativität).
2. $(A \mathbin{\Diamond\!\!\!\!\!+} B) \mathbin{\Diamond\!\!\!\!\!+} C = A \mathbin{\Diamond\!\!\!\!\!+} (B \mathbin{\Diamond\!\!\!\!\!+} C)$, $\quad (A \mathbin{\Diamond\!\!\!\!\!\cdot} B) \mathbin{\Diamond\!\!\!\!\!\cdot} C = A \mathbin{\Diamond\!\!\!\!\!\cdot} (B \mathbin{\Diamond\!\!\!\!\!\cdot} C)$ (Assoziativität).

3. Das Punktintervall $[0,0]$ ist das eindeutig bestimmte neutrale Element der Addition d.h.

$$A = X \oplus A = A \oplus X \quad \forall A \in \mathrm{I}(\mathbb{R}) \Leftrightarrow X = [0,0].$$

   Genauso gilt: Das Punktintervall $[1,1]$ ist das eindeutig bestimmte neutale Element der Multiplikation; in Symbolsprache:

$$A = X \odot A = A \odot X \quad \forall A \in \mathrm{I}(\mathbb{R}) \Leftrightarrow X = [1,1].$$

4. ein beliebiges Element $A = [\underline{a}, \overline{a}] \in \mathrm{I}(\mathbb{R})$ mit $\underline{a} \neq \overline{a}$ besitzt keine inversen Elemente bezüglich der Addition und Multiplikation. Es gilt jedoch:

$$0 \in A \ominus A \quad \text{und} \quad 1 \in A \oslash A.$$

5. $A \odot (B \oplus C) \subseteq (A \odot B) \oplus (A \odot C)$ (Subdistributivität),
   $[a,a] \odot (B \oplus C) = ([a,a] \odot B) \oplus ([a,a] \odot C) \quad \forall a \in \mathbb{R}$.

Beweis

1. Sei $* \in \{+, \cdot\}$, dann gilt:

$$\begin{aligned} A \circledast B &= \{z = a * b : \ a \in A, \ b \in B\} \\ &= \{z = b * a : \ a \in A, \ b \in B\} \\ &= B \circledast A. \end{aligned}$$

2. Sei $* \in \{+, \cdot\}$. Dann gilt

$$\begin{aligned} (A \circledast B) \circledast C &= \{z = y * c : \ y \in A \circledast B, \ c \in C\} \\ &= \{z = (a * b) * c : \ a \in A, \ b \in B, \ c \in C\} \\ &= \{z = a * (b * c) : \ a \in A, \ b \in B, \ c \in C\} \\ &= \{z = a * x : \ a \in A, \ x \in B \circledast C\} \\ &= A \circledast (B \circledast C) \end{aligned}$$

3. "$\Leftarrow$":
   Sei $X = [0,0]$ so folgt durch einfaches Nachrechnen $A = A \circledast X = X \circledast A$. Für $Y = [1,1]$ folgt ebenso einfach $A = A \odot Y = Y \odot A$.
   "$\Rightarrow$":
   Sind N und $\overline{N}$ zwei neutrale Elemente bezüglich der Addition, so gilt

$$N = \overline{N} \oplus N = N \oplus \overline{N} = \overline{N},$$

   womit die Eindeutigkeit des neutralen Elements bewiesen wäre. Für den Beweis der Eindeutigkeit des neutralen Elementes bezüglich der Multiplikation, geht man ebenso vor.

4. Die beiden Aussagen sind äquivalent zu

$$A \ominus B = [0,0] \qquad \Rightarrow \qquad A = B = [a,a]$$

$$A \odot B = [1,1] \qquad \Rightarrow \qquad A = [a,a],\ B = [\tfrac{1}{a}, \tfrac{1}{a}].$$

Sei $A \ominus B = \{z = a - b:\ a \in A,\ b \in B\} = [0,0]$, dann gilt $z = a - b = 0$ für alle $a \in A$ und alle $b \in B$. Für festes $b \in B$ folgt somit $a = b$ für alle $a \in A$, mit anderen Worten: A ist das Punktintervall $[b,b]$. Umgekehrt sieht man, wenn man ein $a \in A$ festhält, daß gilt $B = [a,a]$ und $a = b$, also die gewünschte Aussage. Für den Beweis der Aussage mit der Multiplikation kann man analog vogehen.

Wegen $0 = a - a \in \{z = x - y:\ x \in A,\ y \in A\}$ für alle $a \in A$ gilt $0 \in A \ominus A$. Ebenso folgt mit $1 = a : a \in \{z = x : y:\ x \in A,\ y \in A\}$ für alle $a \in A$, daß $1 \in A \oslash A$ gilt.

5.

$$\begin{aligned} A \odot (B \oplus C) &= \{z = a \cdot (b + c):\ a \in A,\ b \in B,\ c \in C\} \\ &\subseteq \{z = a \cdot b + \tilde{a} \cdot b:\ a, \tilde{a} \in A,\ b \in B,\ c \in C\} \\ &= (A \odot B) \oplus (A \odot C) \end{aligned}$$

□

Satz:

Seien A_1, A_2, B_1, B_2 Intervalle in $\mathrm{I}(\mathbb{R})$ und es gelte $A_1 \subseteq B_1$ und $A_2 \subseteq B_2$. Dann gilt für die Verknüpfungen $* \in \{+, -, \cdot, :\}$ zwischen Intervallen:

$$A_1 \circledast A_2 \subseteq B_1 \circledast B_2.$$

Insbesondere gilt für reelle Zahlen $a_1 \in B_1$ und $a_2 \in B_2$

$$a_1 * a_2 \in B_1 \circledast B_2.$$

Beweis: Der zweite Teil der Aussage ist nur ein Spezialfall des ersten Teiles und muß somit nicht bewiesen werden. Deshalb folgt hier der erste Teil des Beweises. Nach Voraussetzung gilt:

$$\begin{aligned} A_1 \circledast A_2 &= \{z = x * y:\ x \in A_1,\ y \in A_2\} \\ &\subseteq \{z = x * y:\ x \in B_1,\ y \in B_2\} \\ &= B_1 \circledast B_2 \end{aligned}$$

□

4.1.2 Abstand, Betrag und Durchmesser von reellen Intervallen und deren Eigenschaften

Definition (Abstand):

Für zwei Intervalle $A = [\underline{a}, \overline{a}]$, $B = [\underline{b}, \overline{b}]$ ist der *Abstand* definiert durch:

$$q(A,B) = \max\{|\underline{a} - \underline{b}|, |\overline{a} - \overline{b}|\}.$$

Durch den Abstand q wird in $\mathrm{I}(\mathbb{R})$ eine Metrik eingeführt, da q die notwendigen Eigenschaften

$$
\begin{aligned}
&q(A,B) \geq 0 \quad \forall A,B \in \mathrm{I}(\mathbb{R}), \quad q(A,B)=0 \Leftrightarrow A=B, && \text{(Definitheit)}\\
&q(A,B) = q(B,A), && \text{(Symmetrie)}\\
&q(A,B) \leq q(A,C)+q(B,C) && \text{(Dreiecksungleichung)}
\end{aligned}
$$

besitzt.

Bemerkung:
Für Punktintervalle geht der obige Abstandsbegriff in den für reelle Zahlen üblichen Abstandsbegriff $q(a,b)=|a-b|$ über.

Durch Einführung der Metrik q wird $\mathrm{I}(\mathbb{R})$ zu einem topologischen Raum. Damit lassen sich die Begriffe Konvergenz und Stetigkeit genauso, wie in anderen topologischen Räumen erklären. Es gilt z.B. für die Konvergenz:

$$A_k \overset{k\to\infty}{\to} A \Leftrightarrow q(A_k,A) \overset{k\to\infty}{\to} 0 \Leftrightarrow \underline{a}_k \overset{k\to\infty}{\to} \underline{a} \text{ und } \overline{a}_k \overset{k\to\infty}{\to} \overline{a}.$$

Satz (Vollständigkeit von $(\mathbf{I}(\mathbb{R}),q)$):

Der metrische Raum $(\mathrm{I}(\mathbb{R}),q)$ ist ein vollständiger metrischer Raum (d.h. daß jede Cauchyfolge von Intervallen ein Intervall als Grenzwert besitzt).

Satz (Stetigkeit der Intervalloperationen):

Die Intervalloperationen $\oplus$, $\ominus$, $\odot$ und $\oslash$ sind stetig.

Beweis: Seien (A_k) und (B_k) zwei beliebige Folgen mit $A_k \overset{k\to\infty}{\to} A$ und $B_k \overset{k\to\infty}{\to} B$. Für die Folge der Verknüpfungsergebnisse $(A_k \circledast B_k)$ mit $\star \in \{+,-,\cdot,:\}$ gilt dann:

$$
\begin{aligned}
A_k \oplus B_k &= \left[\underline{a}_k+\underline{b}_k, \overline{a}_k+\overline{b}_k\right] \overset{k\to\infty}{\to} \left[\underline{a}+\underline{b}, \overline{a}+\overline{b}\right] = A \oplus B,\\
A_k \ominus B_k &= \left[\underline{a}_k-\underline{b}_k, \overline{a}_k-\overline{b}_k\right] \overset{k\to\infty}{\to} \left[\underline{a}-\underline{b}, \overline{a}-\overline{b}\right] = A \ominus B,\\
A_k \odot B_k &= \left[\min\{\underline{a}_k\cdot\underline{b}_k, \underline{a}_k\cdot\overline{b}_k, \overline{a}_k\cdot\underline{b}_k, \overline{a}_k\cdot\overline{b}_k\}, \max\{\underline{a}_k\cdot\underline{b}_k, \underline{a}_k\cdot\overline{b}_k, \overline{a}_k\cdot\underline{b}_k, \overline{a}_k\cdot\overline{b}_k\}\right]\\
&\overset{k\to\infty}{\to} \left[\min\{\underline{a}\cdot\underline{b}, \underline{a}\cdot\overline{b}, \overline{a}\cdot\underline{b}, \overline{a}\cdot\overline{b}\}, \max\{\underline{a}\cdot\underline{b}, \underline{a}\cdot\overline{b}, \overline{a}\cdot\underline{b}, \overline{a}\cdot\overline{b}\}\right] = A \odot B,\\
A_k \oslash B_k &= \left[\min\{\underline{a}_k\cdot\tfrac{1}{\underline{b}_k}, \underline{a}_k\cdot\tfrac{1}{\overline{b}_k}, \overline{a}_k\cdot\tfrac{1}{\underline{b}_k}, \overline{a}_k\cdot\tfrac{1}{\overline{b}_k}\}, \max\{\underline{a}_k\cdot\tfrac{1}{\underline{b}_k}, \underline{a}_k\cdot\tfrac{1}{\overline{b}_k}, \overline{a}_k\cdot\tfrac{1}{\underline{b}_k}, \overline{a}_k\cdot\tfrac{1}{\overline{b}_k}\}\right]\\
&\overset{k\to\infty}{\to} \left[\min\{\underline{a}\cdot\tfrac{1}{\underline{b}}, \underline{a}\cdot\tfrac{1}{\overline{b}}, \overline{a}\cdot\tfrac{1}{\underline{b}}, \overline{a}\cdot\tfrac{1}{\overline{b}}\}, \max\{\underline{a}\cdot\tfrac{1}{\underline{b}}, \underline{a}\cdot\tfrac{1}{\overline{b}}, \overline{a}\cdot\tfrac{1}{\underline{b}}, \overline{a}\cdot\tfrac{1}{\overline{b}}\}\right] = A \oslash B.
\end{aligned}
$$

□

Korollar (Stetigkeit von einstelligen Operationen in $\mathbf{I}(\mathbb{R})$):

Ist $r(x)$ eine stetige Funktion, so ist $r(X) := [\min_{x\in X}, r(x), \max_{x\in X} r(x)]$ eine stetige einstellige Operatione in $\mathrm{I}(\mathbb{R})$.

Definition (Betrag):

Für ein Intervall $A = [\underline{a}, \overline{a}] \in \mathrm{I}(\mathbb{R})$ heißt

$$|A| := q(A, [0,0]) = \max\{|\underline{a}|, |\overline{a}|\} = \max_{a \in A} |a|$$

der *Betrag* von A.

Definition (Durchmesser):

Für ein Intervall $A = [\underline{a}, \overline{a}] \in \mathrm{I}(\mathbb{R})$ heißt

$$d(A) := \overline{a} - \underline{a} \geq 0$$

der *Durchmesser* von A.

4.1.3 Intervallmäßige Auswertung und Wertebereiche von reellen Funktionen:

Im ganzen folgenden Abschnitt werden, da es sich nur um eine Übersicht über und eine Einführung in die Intervallrechnung handeln soll, die Beweise der Sätze weggelassen und stattdessen nur deren Resultate wiedergegeben. Wer an den Beweisen Interesse hat, findet diese in [12].

Als Generalvoraussetzung für den gesamten Abschnitt gelte, daß alle betrachteten Funktionen f reell und stetig seien.

Defininition (Funktionsausdruck):

Ein zu f gehörender *Funktionsausdruck* sei eine Rechenvorschrift, mit der zu jedem x, nach endlich vielen Operationen, zu denen auch die entsprechenden Intervalloperationen definiert seien, der jeweilige Funktionswert bestimmt werden kann.

Enthält der zu f gehörende Funktionsausdruck auch noch die Konstanten $a_1, \ldots, a_m$ so wird zur Verdeutlichung dafür $f(x; a_0, \ldots, a_m)$ geschrieben. O.B.d.A. wird vorausgesetzt, daß jede dieser Konstanten im Funktionsausdruck nur einmal vorkommt, denn käme eine Konstante mehrfach vor, so vergebe man für sie einfach mehrere Namen a_i.

Definition (Wertebereich):

Das Intervall das alle Funktionswerte von f enthält, wenn unabhängig voneinander $x \in X$, $a_k \in A_k$ $(k = 1, \ldots, m)$ genommen werden

$$W(f, X; A_0, \ldots A_m) := \{f(x, a_0, \ldots, a_m) : \ x \in X, \ a_1 \in A_1, \ \ldots, \ a_m \in A_m\},$$

heißt *Wertebereich* der Funktion f. Wie man anhand der Definition sehen kann, ist der Wertebereich der Funktion f unabhängig vom jeweiligen Funktionsausdruck für f.

Hat man zu einer reellen Funktion einen Funktionsausdruck gegeben, so ist klar, wie man diesen auszuwerten hat; will man die Funktion hingegen intervallmäßig auswerten, so muß man erst definieren, wie dies zu geschehen hat. Die folgende Definition soll erklären, was man unter einer intervallmäßigen Auswertung zu verstehen hat.

Definition (intervallmäßige Auswertung):

Ist eine Funktion f und deren Funktionsausdruck gegeben, so bezeichnet man $f(X, A_0, \ldots, A_m)$ als *intervallmäßige oder intervallarithmetische Auswertung* von f, falls bei der Ersetzung aller Operanden durch Intervalle, die im Definitionsbereich enthalten sind, und aller Operationen durch Intervalloperationen stets ein definierter Ausdruck entsteht.

Bemerkungen:

- Die intervallmäßige Auswertung ist, wie man an dem Beispiel

$$\begin{aligned} g_1(x,a) &= \frac{ax}{1-x} && \Rightarrow\ g_1([2,3],[0,1]) = [-3,0] \\ g_2(x,a) &= \frac{a}{\frac{1}{x}-1} && \Rightarrow\ g_2([2,3],[0,1]) = [-2,0] \end{aligned}$$

  sehen kann, von der Wahl des Funktionsausdruckes abhängig.

- Nicht jeder reelle Funktionsausdruck führt zu einem definierten Intervallausdruck. Hat man etwa die reelle Funktion definiert durch den Funktionsausdruck $f(x) = \frac{1}{x \cdot x + 0.5}$ und wertet f intervallmäßig an der im Definitionsbereich enthaltenen Stelle $X = [-1,1]$ aus, so erhält man den undefininierten Ausdruck:

$$\begin{aligned} f([-1,1]) &= [1,1] \mathbin{\diamond\!\!\!\!/} ([-1,1] \mathbin{\diamond\!\!\!\!\cdot} [-1,1] \mathbin{\diamond\!\!\!\!+} [0.5,0.5]) \\ &= [1,1] \mathbin{\diamond\!\!\!\!/} ([-1,1] \mathbin{\diamond\!\!\!\!+} [0.5,0.5]) \\ &= [1,1] \mathbin{\diamond\!\!\!\!/} [-0.5,1.5]. \end{aligned}$$

Aus der Definition der intervallmäßigen Auswertung von Funktionen ergeben sich sofort ein paar grundlegende Eigenschaften, die im nächsten Satz zusammengestellt sind.

Satz:

Sei f stetige Funktion der reellen Veränderlichen $x_1, \ldots x_n$ mit zugehörigem Funktionsausdruck $f(x_1, \ldots, x_n, a_1, \ldots, a_m)$. Ferner sei die intervallmäßige Auswertung $f(\tilde{X}_1, \ldots, \tilde{X}_n, \tilde{A}_1, \ldots, \tilde{A}_m)$ für die Intervalle $\tilde{X}_1, \ldots, \tilde{X}_n$ und $\tilde{A}_1, \ldots \tilde{A}_m$ definiert. Dann gilt:

1. Für alle $X_k \subseteq \tilde{X}_k \quad k = 1, \ldots, n$ und alle $A_k \subseteq \tilde{A}_k \quad k = 1, \ldots, m$ ist

$$W(f, X_1, \ldots, X_n, A_1, \ldots, A_m) \subseteq f(X_1, \ldots X_n, A_1 \ldots, A_m)$$

   (Einschließungseigenschaft).

2. Für alle $X_k \subseteq \hat{X}_k \subseteq \tilde{X}_k \quad k = 1, \ldots, n$ und alle $A_k \subseteq \hat{A}_k \subseteq \tilde{A}_k \quad k = 1, \ldots, m$ ist

$$f(X_1, \ldots, X_n, A_1, \ldots, A_m) \subseteq f(\hat{X}_1, \ldots, \hat{X}_n, \hat{A}_1, \ldots, \hat{A}_m)$$

(Teilmengeneigenschaft).

Die nun folgenden beiden Sätze beschäftigen sich mit der Güte, mit der der Wertebereich einer Funktion f durch die intervallmäßige Auswertung der Funktion approximiert werden kann.

Satz:

Sei f eine reelle Funktion einer Veränderlichen x mit zugehörigem Funktionsausdruck $f(x, a_1, \ldots, a_m)$. Die intervallmäßige Auswertung des Ausdruckes für f existiere für $\tilde{X}, A_1, \ldots, A_m \in \mathrm{I}(\mathbb{R})$. Der Funktionsausdruck, den man erhält, wenn man im Funktionsausdruck für f bei jedem Auftreten von x eine neue Variable x_i einführt, genüge in jeder dieser Variablen im Intervall $\tilde{X}$ einer Lipschitzbedingung. Dann gilt für $X \subseteq \tilde{X}$:

$$q(W(f, X; A_1 \ldots, A_m), f(X, A_1 \ldots, A_m)) \leq \gamma \cdot d(X) \quad \gamma \geq 0,$$

für $d(X) \to 0$ folgt also $f(X, A_1, \ldots, A_m) \to W(f, X; A_1, \ldots A_m)$.

Bemerkungen:

- Die geforderte Lipschitzbedingung in jeder der neuen Variablen läßt sich bei Funktionen, die aus intervallmaßig ausführbaren Operationen zusammengesetzt sind, oft erfüllen.
- Die Approximation des Wertebereichs ist abhängig von der Wahl des Funktionsausdruckes.

Satz:

Sei f eine reelle Funktion mit zugehörigem Funktionsausdruck $f(x)$. Seien weiterhin die Voraussetzungen des vorherigen Satzes erfüllt, dann gilt für $X \subseteq \tilde{X}$:

$$d(f(X)) \leq c \cdot d(X) \quad c \geq 0.$$

4.1.4 Maschinenintervallarithmetik:

In den letzten Abschnitten über die Intervallrechnung wurde beschrieben, wie man mit Intervallen rechnen kann und welche Gesetze dabei gelten. Hierbei wurde immer von den reellen Zahlen als zugrundeliegendem Körper für die Intervalle ausgegangen. Will man nun aber die Intervallrechnung auf Computern verwirklichen, so stehen einem dort nicht alle reellen Zahlen, sondern nur eine endliche Anzahl von *Maschinenzahlen* zur Verfügung. Dabei seien die Maschinenzahlen definiert durch:

Definition (Maschinenzahlen):

Zahlen, die sich in der Form

$$x = mb^e$$

darstellen lassen, mit $m =$ Mantisse (generalisiert durch $\frac{1}{2} \leq |m| < 1$), $b =$ Basis (auf Computern meist 2) und $e =$ Exponent (mit $\underline{e} \leq e \leq \overline{e}$), wobei für m und e nur eine feste Anzahl von endlich viele Stellen zur Verfügung stehen, seien als *Maschinenzahlen* bezeichnet. Die Menge aller Maschinenzahlen eines Rechners sei bezeichnet mit $\mathbb{R}_\mathrm{M}$, wobei noch die für spätere Überlegungen nützliche Forderung $\mathbb{R}_\mathrm{M} = -\mathbb{R}_\mathrm{M}$ erfüllt sei.

Dadurch daß aber nur eine endliche Anzahl von Maschinenzahlen zur Verfügung steht, ergibt sich nun das Problem, wie man $z \in [\min_{y\in\mathbb{R}_\mathrm{M}}, \max_{y\in\mathbb{R}_\mathrm{M}}]$ durch $\tilde{z} \in \mathbb{R}_\mathrm{M}$ repräsentieren kann. Ganz allgemein ist die Lösung dieses Problems möglich, durch die sogenannte *Rundung*.

Definition (Rundung):

Die Abbildung $\mathrm{rnd}_\mathrm{M} : x \in \mathbb{R} \to \mathrm{rnd}_\mathrm{M} x \in \mathbb{R}_\mathrm{M}$ bezeichnet man als *Rundung*, wenn aus $x \leq y$ folgt $\mathrm{rnd}_\mathrm{M} x \leq \mathrm{rnd}_\mathrm{M} y$ (Monotonie). Eine Abbildung $\mathrm{rnd}_\mathrm{M} : x \in \mathbb{R} \to \mathrm{rnd}_\mathrm{M} x \in \mathbb{R}_\mathrm{M}$, für die weiterhin gilt $x = \mathrm{rnd}_\mathrm{M} x \ \forall x \in \mathbb{R}_\mathrm{M}$, heißt optimale Rundung.

Um später reelle Zahlen in Intervalle auf dem Computer einschließen zu können, wird eine Rundung gebraucht, die stets nach unten oder stets nach oben rundet. Man definiert deshalb:

Definition (gerichtete Rundung):

Eine Rundung $\underline{\mathrm{rnd}}_\mathrm{M}$, für die stets gilt

$$\forall x \in \mathbb{R} \Rightarrow \ \underline{\mathrm{rnd}}_\mathrm{M} x \in \mathbb{R}_\mathrm{M} \leq x$$

heißt nach unten gerichtete Rundung.

Bemerkung:
Aus jeder nach unten gerichteten Rundung kann man durch die einfache Definition $\overline{\mathrm{rnd}}_\mathrm{M} x := -\underline{\mathrm{rnd}}_\mathrm{M}(-x)$ eine nach oben gerichtete Rundung erhalten.

Mit den Definitionen dieses Abschnittes wurde gezeigt, wie man reelle Zahlen auf dem Computer repräsentieren kann. Es bleibt jetzt noch übrig, auch für reelle Intervalle $X \in \mathrm{I}(\mathbb{R})$ Repräsentanten auf dem Rechner in geeigneter Weise zu finden. Man definiert ähnlich wie bei den Maschinenzahlen:

Definition (Maschinenintervalle):

Ein Maschinenintervall X ist gegeben durch $X = [\underline{x}, \overline{x}]$ mit $\underline{x}, \overline{x} \in \mathbb{R}_\mathrm{M}$ und $\underline{x} \leq \overline{x}$. Die Menge aller Maschinenintervalle $\mathrm{I}(\mathbb{R}_\mathrm{M})$ ist somit gegeben durch

$$\mathrm{I}(\mathbb{R}_\mathrm{M}) := \{[\underline{x}, \overline{x}] : \underline{x}, \overline{x} \in \mathbb{R}_\mathrm{M} \text{ und } \underline{x} \leq \overline{x}\} \subseteq \mathrm{I}(\mathbb{R}).$$

Auch für die Menge der Maschinenintervalle muß man sich wieder eine geeignete Rundung überlegen, um die Intervalle auf die Maschinenintervalle abzubilden. Die folgende Definition wird dieses Problem lösen:

Definition (Intervallrundung):

Die Abbildung $\mathrm{RND_M} : X \in \mathrm{I}(\mathbb{R}) \to \mathrm{RND_M}X \in \mathrm{I}(\mathbb{R}_\mathrm{M})$ heißt Rundung, falls gilt:

$$X \in \mathrm{I}(\mathbb{R}) \Rightarrow \ X \subseteq \mathrm{RND_M}X$$

und

$$X, Y \in \mathrm{I}(\mathbb{R}) \text{ mit } X \subseteq Y \Rightarrow \ \mathrm{RND_M}X \subseteq \mathrm{RND_M}Y.$$

Bemerkungen:

- Bedingung zwei ist ähnlich der Forderung der Monotonie bei der Rundung reeller Zahlen.
- Bedingung eins wird notwendig sein, um für Operationen zwischen Maschinenintervallen ähnliche Resultate zu erhalten, wie man sie bei Operationen zwischen Intervallen erhält.
- Jede solche Intervallrundung läßt sich darstellen durch

$$\mathrm{RND_M}X = \mathrm{RND_M}([\underline{x}, \overline{x}]) = [\underline{\mathrm{rnd}}_\mathrm{M}\underline{x}, \overline{\mathrm{rnd}}_\mathrm{M}\overline{x}].$$

Über Verknüpfungen von Maschinenzahlen wurde zwar zuvor nicht gesprochen, aber es ist klar, daß auch das Resultat einer Verknüpfung, abgesehen von Über- oder Unterläufen, selbst wieder eine Maschinenzahl sein muß. Dies erreicht man, indem man zuerst die Verknüpfung exakt durchführt und danach rundet. Auf ähnliche Weise kann man dies auch für Intervalle definieren.

Definition (Maschinenintervall-Operationen):

Für $A, B \in \mathrm{I}(\mathbb{R}_\mathrm{M})$ wird die Operation $\circledast_\mathrm{M}$ mit $* \in \{+, -, \cdot, :\}$ erklärt durch:

$$A \circledast_\mathrm{M} B := \mathrm{RND_M}(A \circledast B) \in \mathrm{I}(\mathbb{R}_\mathrm{M}),$$

wobei $\mathrm{RND_M}$ eine vorgegebene Intervallrundung sei.

Definiert man die Verknüpfung zwischen Maschinenintervallen auf diese Weise, so sagen die folgenden Sätze, daß die grundlegenden Eigenschaften, die man bei der Intervallrechnung gefunden hat, auch bei Maschinenintervallen gelten.

Satz (Maschinenintervall-Operationen):

Für die Operationen $\circledast_{\mathrm{M}}$ mit $* \in \{+,-,\cdot,:\}$ gilt nach obiger Definition für $A, B, \tilde{A}, \tilde{B} \in \mathrm{I}(\mathbb{R}_{\mathrm{M}})$ mit $A \subseteq \tilde{A}$ und $B \subseteq \tilde{B}$:

$$A \circledast_{\mathrm{M}} B \subseteq \tilde{A} \circledast_{\mathrm{M}} \tilde{B}.$$

Mit anderen Worten, es gilt auch für die Maschinenintervalle, die von den Intervallen bekannte Teilmengeneigenschaft.

Für die Einschließung von Rundungsfehlern oder die Einschließung des Wertebereichs einer Funktion liefert der folgende Satz die entscheidenden Aussagen:

Satz:

Für die Maschinenintervallrundung $\mathrm{RND}_{\mathrm{M}}$, die durch gerichtete Rundungen ($\underline{\mathrm{rnd}}_{\mathrm{M}}$, $\overline{\mathrm{rnd}}_{\mathrm{M}}$) der Grenzen gewonnnen wird und $* \in \{+,-,\cdot,:\}$ gilt:

$$\begin{aligned} A, B \in \mathrm{I}(\mathbb{R}_{\mathrm{M}}) \Rightarrow &\; A \circledast B \subseteq A \circledast_{\mathrm{M}} B \in \mathrm{I}(\mathbb{R}_{\mathrm{M}}), \\ a \in A \in \mathrm{I}(\mathbb{R}_{\mathrm{M}}), b \in B \in \mathrm{I}(\mathbb{R}_{\mathrm{M}}) \Rightarrow &\; a * b \in A \circledast_{\mathrm{M}} B \in \mathrm{I}(\mathbb{R}_{\mathrm{M}}). \end{aligned}$$

Insbesondere folgt damit auch:

$$a \in \mathbb{R}, b \in \mathbb{R} \Rightarrow \; a * b \in \mathrm{RND}_{\mathrm{M}}[a,a] \circledast_{\mathrm{M}} \mathrm{RND}_{\mathrm{M}}[b,b].$$

4.2 Numerische Verfahren:

In diese Abschnitt sollen solche numerischen Verfahren vorgestellt werden, die für die Durchführung des Algorithmus in Kapitel 1, nach der dort beschriebenen Vorgehensweise, verwendet werden können. Da der ganze Algorithmus so konzipiert war, daß er leicht auf Computer umgesetzt werden kann, sind auch die hier vorgestellten Verfahren für die Verwendung eines Computers ausgelegt. An entsprechender Stelle in Kapitel 1 findet sich ein Verweis auf das entsprechende numerische Verfahren, damit man nicht immer im Hinterkopf behalten muß, welche Methoden denn nun hier beschrieben werden und für welchen Zweck sie gedacht sind. Wer mehr über die entsprechenden Themen lesen möchte findet in [11] die Informationen über die numerische Integration und in [10] die Informationen über das QR-Verfahren.

4.2.1 Die QR-Zerlegung:

Das erste numerische Verfahren, das hier vorgestellt werden soll ist die QR-Zerlegung. Es gibt zwei Gründe, weswegen die QR-Zerlegung beschrieben wird. Zum einen bildet die QR-Zerlegung die Grundlage des danach beschriebenen QR-Verfahrens. Zum anderen läßt sich die QR-Zerlegung aber auch nutzen um das lineare Gleichungssystem, das man beim Newtonalgorithmus in 1.3.1 “Zu Punkt 1: Bestimmung der Näherungslösung ω” erhält zu lösen.

Satz:

Zu jeder Matrix $A \in \mathbb{R}^{n\times n}$ existiert eine orthogonale Matrix Q (d.h. $Q^tQ = I$) und eine rechte obere Dreiecksmatrix R, so daß sich A zerlegen läßt in

$$\boxed{A = QR}.$$

Wie man mit der QR-Zerlegung ein lineares Gleichungssystem lösen kann sieht man sehr einfach an den folgenden Äquivalenzen:

$$Ax = b \Leftrightarrow QRx = b \Leftrightarrow Rx = Q^tb.$$

Aus der letzte Äquivalenz läßt sich, dadurch daß R eine obere Dreiecksmatrix ist, x sehr einfach bestimmen.

Für die Berechnung der QR-Zerlegung gibt es mehrere Möglichkeiten. Eine beruht darauf, daß ausgehend von der Matrix $A_0 := A$ Matrizen $A_i := P_iA_{i-1}$ durch Multiplikation der vorherigen Iterationsmatrix mit einer sogenannten *Housholdermatrix* P_i berechnet werden. Dabei wird in jedem Schritt von i nach $i+1$ die Spalte $i+1$ von A_i auf die Form von R gebracht.

Die andere Möglichkeit, die hier näher beschrieben werden soll, beruht auf dem *Orthogonalisierungsverfahren von Schmidt.* Betrachtet man die Zerlegung $A = QR$ als zusammengefaßte Schreibweise für mehrere Vektorgleichungen, so erkennt man, daß jede Spalte a_k von A eine Linearkombination der ersten k orthonormalen Vektoren q_i (Spalten von Q) ist:

$$a_k = \sum_{i=1}^{k} r_{ik}q_i \quad k = 1,\ldots,n.$$

Diese Erkenntnis legt einem die folgende Vorgehensweise nahe:

- Für den Start ergeben sich die beiden Gleichungen $a_1 = r_{11}q_1$ und $q_1^tq_1 = 1$, die man, wenn man

  $$r_{11} := ||a_1|| \quad \text{und} \quad q_1 := \frac{a_1}{r_{11}}$$

  setzt erfüllen kann.

- Für den Schritt von $k-1$ nach k ergeben sich die Gleichungen $a_k = \sum_{i=1}^{k} r_{ik}q_i$ und $q_i^tq_j = \delta_{ij}$ für alle $i,j \leq k$. Multipliziert man die erste Gleichung für $j = 1,\ldots,k-1$ mit q_j und verwendet dann noch die Orthonormalitätsgleichungen so ergibt sich

  $$q_j^ta_k = \sum_{i=1}^{k} r_{ik}q_j^tq_i = \sum_{i=1}^{k} r_{ik}\delta_{ji} = r_{jk}$$

  Aus der Gleichung $a_k - \sum_{i=1}^{k-1} r_{ik}q_i = q_k$ und der Gleichung $q_k^tq_k = 1$ ergibt sich

  $$r_{kk} := ||a_k - \sum_{i=1}^{k-1} q_i|| \text{ und damit } q_k := \frac{a_k - \sum_{i=1}^{k-1} q_i}{r_{kk}}.$$

  Somit sind also alle Unbekannten dieses Schrittes bestimmt.

Der Nachteil des Berechnens der QR-Zerlegung nach der Schmidtschen Methode gegenüber der Vorgehensweise mit den Housholdermatrizen ist, daß das Verfahren nicht gutartig ist d.h. das kleinst Rundungsfehler die für das Verfahren wichtige Eigenschaft der Orthogonalität der q_i vernichten können. Wenn die q_i nicht mehr orthogonal zueinander sind, kann aber eine sogenannte Nachorthogonalisierung, um den Preis des doppelten Rechenaufwandes, helfen.

4.2.2 Das QR-Verfahren:

Bei der Berechnung von Näherungseigenwerten für den Operator L^*L nach dem Rayleigh-Ritz Verfahren hat man wie in 1.3.3 "Zu Punkt 3: Berechnung der Konstanten σ" beschrieben, das Eigenwertproblem

$$A_1 x = \tilde{\lambda} A_2 x \quad x \in \mathbb{R}^M \backslash \{0\}$$

mit den Matrizen

$$\begin{aligned} A_1 &:= (< Lv_i, Lv_j >_2)_{\substack{i=1,\dots,M \\ j=1,\dots,M}} \\ A_2 &:= (< v_i, v_j >_2)_{\substack{i=1,\dots,M \\ j=1,\dots,M}} \end{aligned}$$

zu lösen. Für diese Aufgabe bietet sich das QR-Verfahren durch seine einfache Umsetzbarkeit und seine Stabilität geradezu an. Der erste Satz beschreibt die grundlegenden Eigenschaften der QR-Iteration. Danach wird in getrennten Abschnitten auf die Konvergenz und die praktische Umsetzung des Verfahrens eingegangen.

Satz (QR-Iteration):

Bildet man ausgehend von der $n \times n$-Matrix $A_1 := A$ die QR-Zerlegungen

$$A_i := Q_i R_i$$

(wobei Q_i unitäre Matrix sei (d.h. es gelte $Q_i^t Q_i = I$) und R_i obere Dreiecksmatrix sei) und damit die nächste Iterationsmatrix

$$A_{i+1} := R_i Q_i = Q_{i+1} R_{i+1},$$

so gilt:

1. A_{i+1} ist ähnlich zu A_i d.h. A_{i+1} hat die gleichen Eigenwerte wie A_i und A_{i+1} läßt sich darstellen durch $A_{i+1} = Q_i^t A_i Q_i$.
2. $A_{i+1} = (Q_1 Q_2 \cdots Q_n)^t A_1 Q_1 Q_2 \cdots Q_n$

Beweis.

1. Durch einfaches Rechnen sieht man

$$A_{i+1} = R_i Q_i = Q_i^t Q_i R_i Q_i = Q_i^t A_i Q_i,$$

also den zweiten Teil der Aussage. Sei nun λ_j beliebiger Eigenwert zur $n \times n$-Matix A_i mit Eigenvektor $\varphi_j^{(i)}$, dann gilt:

$$
\begin{aligned}
A_i \varphi_j^{(i)} &= \lambda_j \varphi_j^{(i)} \\
A_i \underbrace{Q_i Q_i^t}_{=I} \varphi_j^{(i)} &= \lambda_j \varphi_j^{(i)} \\
\underbrace{Q_i^t A_i Q_i}_{=A_{i+1}} Q_i^t \varphi_j^{(i)} &= \lambda_j Q_i^t \varphi_j^{(i)} \\
A_{i+1} \underbrace{Q_i^t \varphi_j^{(i)}}_{=:\varphi_j^{(i+1)}} &= \lambda_j \underbrace{Q_i^t \varphi_j^{(i)}}_{=:\varphi_j^{(i+1)}} .
\end{aligned}
$$

Somit ist also λ_j Eigenwert von A_{i+1} um Eigenvektor $\varphi_j^{(i+1)} := Q_i^t \varphi_j^{(i)}$. Sei nun umgekehrt λ_j beliebiger Eigenwert von A_{i+1} zum Eigenvektor $\varphi_j^{(i+1)}$, dann gilt:

$$
\begin{aligned}
A_{i+1} \varphi_j^{(i+1)} &= \varphi_j^{(i+1)} \\
Q_i^t A_i Q_i \varphi_j^{(i+1)} &= \varphi_j^{(i+1)} \\
A_i \underbrace{Q_i \varphi_j^{(i+1)}}_{=:\varphi_j^{(i)}} &= \underbrace{Q_i \varphi_j^{(i+1)}}_{=:\varphi_j^{(i)}} .
\end{aligned}
$$

Somit ist also λ_j Eigenwert von A_i zum Eigenvektor $\varphi_j^{(i)} := Q_i \varphi_j^{(i+1)}$.

2. Durch wiederholtes Anwenden von Aussage 1. erhält man

$$
\begin{aligned}
A_{i+1} &\overset{1.}{=} Q_i^t A_i Q_i \overset{1.}{=} Q_i^t Q_{i-1}^t A_{i-1} Q_{i-1} Q_i \overset{1.}{=} \cdots \\
&\overset{1.}{=} Q_i^t Q_{i-1}^t \cdots Q_1^t A_1 Q_1 \cdots Q_{i-1} Q_i \\
&= (Q_1 Q_2 \cdots Q_i)^t A_1 (Q_1 Q_2 \cdots Q_i).
\end{aligned}
$$

□

Bemerkungen:

1. Die QR-Zerlegung der Matrizen A_i existiert stets und kann mit den Orthogonalisierungsverfahren von Housholder oder Schmidt numerisch stabil berechnet werden.

2. Die QR-Zerlegung einer Matrix ist nicht eindeutig, denn ist z.B. $A_i = Q_i R_i$ eine QR-Zerlegung der Matrix A_i, so ist auch $(Q_i S)(S^h R_i) = \tilde{Q}_i \tilde{R}_i = A_i$ mit $S = \operatorname{diag}\left(e^{i\Phi_1}, e^{i\Phi_2}, \ldots, e^{i\Phi_n}\right)$ ebenfalls eine QR-Zerlegung von A_i.

Satz (Konvergenzeigenschaft des QR-Verfahrens):

Sind die Eigenwerte λ_i des Eigenwertproblems $Ax = \lambda x$ betragsmäßig verschieden

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

und gelten gewisse Zusatzvoraussetzungen (siehe [10]) so gilt:

1. $\lim\limits_{i \to \infty} a_{jk}^{(i)} = 0$ für alle j, k mit $j > k$,

2. $\lim\limits_{i \to \infty} a_{jj}^{(i)} = \lambda_j \quad j = 1, \ldots, n.$

Sind die Eigenwerte hingegen nicht alle betragsmäßig verschieden, sondern gilt etwa, wie bei konjungiert komplexen Eigenwerten:

$$|\lambda_1| > \cdots > |\lambda_r| = |\lambda_{r+1}| > \cdots > |\lambda_n|,$$

so gilt unter den gleichen Voraussetzungen (siehe [10]):

1. $\lim\limits_{i \to \infty} a_{jk}^{(i)} = 0$ für alle $j, k \neq r, r+1$ mit $j > k$.

2. $\lim\limits_{i \to \infty} a_{jj}^{(i)} = \lambda_j$ für $j \neq r, r+1$.

3. Obwohl die 2×2 Matrizen

$$\begin{bmatrix} a_{r,r}^{(i)} & a_{r,r+1}^{(i)} \\ a_{r+1,r}^{(i)} & a_{r+1,r+1}^{(i)} \end{bmatrix}$$

   im allgemeinen für $i \to \infty$ divergieren, konvergieren ihre Eigenwerte für $i \to \infty$ gegen die Eigenwerte λ_r und λ_{r+1}.

Wegen des Umfanges wird der Beweis des Satzes hier nicht aufgeführt, sondern auf das Buch [10]: Stoer und Bulisch *Numerische Mathematik 2* verwiesen, in dem man ihn finden kann.

Praktische Durchführung des QR-Verfahrens:

1. Das QR-Verfahren ist sehr aufwendig. Pro Schritt von A_i nach A_{i+1} werden für vollbesetzte Matrizen $O(n^3)$ Rechenoperationen benötigt. Diesen Aufwand kann man jedoch verhindern, indem man das QR-Verfahren nicht auf die vollbesetzten Matrizen, sondern auf reduzierte Matrizen von Hessebergform, oder bei symmetrischen Matrizen auf Tridiagonalmatrizen anwendet. Dies funktioniert, da diese Matrixformen bezüglich der QR-Transformationen invariant sind. Bei Hessenbergmatrizen sinkt der Aufwand pro Schritt damit auf $O(n^2)$, bei Tridiagonalmatizen gar auf $O(n)$.

2. Die Konvergenz des QR-Verfahrens ist sehr langsam, wenn die Eigenwerte λ_j von A betragsmäßig nur schlecht getrennt sind d.h. $|\frac{\lambda_j}{\lambda_k}| \approx 1$ für $j > k$. Diesem Problem des QR-Verfahrens kann man begegnen,

indem man statt des QR-Verfahrens, das QR-Verfahren mit Shifts verwendet:

$$\begin{aligned} A_1 &:= A \\ A_i - k_i I &=: Q_i R_i \quad \text{(QR-Zerlegung)} \\ A_{i+1} &:= R_i Q_i + k_i I, \end{aligned}$$

wobei k_i eine möglichst gute Näherung für einen der Eigenwerte von A sei.

4.2.3 Numerische Integration mit Fehlerabschätzung:

Bei der Berechnung des Defektes δ, der monotonen Funktion G und an vielen anderen Stellen des Algorithmus in Kaptel 1 traten Integrale auf, die exakt ausgerechnet werden mußten, d.h. deren Wert in einem festen Intervall angegeben werden mußte. Eine Möglichkeit solche Intervalleinschließungen zu bekommen ist die Berechnung Riemanscher Ober- und Untersummen. Eine andere Möglichkeit beruht darauf, daß man bei Integrationsformeln den Fehler abschätzen kann. Im folgenden Satz wird genau dies für die Trapezformel getan. Der Beweis dieses Satzes ist so gehalten, daß er durch leichte Modifikatione auch für Verfahren mit höherer Fehlerordnung verwendet werden kann.

Satz (Trapezformel mit Fehlerabschätzung):

Berechnet man den Wert des Integrales $\int_a^b f(x)\,dx$ näherungsweise mit der Trapezformel

$$\int_a^b f(x)\,dx \approx h\left(\frac{f(x_0)}{2} + \sum_{i=1}^{n-1} f(x_i) + \frac{f(x_n)}{2}\right)$$

mit $h = \frac{b-a}{n}$ und $x_i = a + i \cdot h$, so gilt für den dabei gemachten Fehler

$$Rf(x) := \int_a^b f(x)\,dx - h\left(\frac{f(x_0)}{2} + \sum_{i=1}^{n-1} f(x_i) + \frac{f(x_n)}{2}\right)$$

die Abschätzung

$$\boxed{|Rf(x)| \leq \frac{h^2}{12}(b-a) \max_{\xi\in[a,b]} |f''(\xi)|}.$$

Beweis: Die Funktion $f(x)$ läßt sich schreiben als

$$f(x) = f(x_i) + f'(x_i)\cdot(x - x_i) + \int_{x_i}^x f''(t)\cdot(x-t)\,dt$$

Da $R_n f(x) := \int_{x_i}^{x_{i+1}} f(x)\,dt - \frac{h}{2}(f(x_i) + f(x_{i+1}))$ ein lineares Funktional ist, das für Polynome bis zum Grad eins den Wert 0 annimmt erhält man mit der obigen Darstellung von $f(x)$:

$$R_n f(x) = R\left[\int_{x_i}^x f''(t)\cdot(x-t)\,dt\right]$$

$$
\begin{aligned}
&= \int_{x_i}^{x_{i+1}} \int_{x_i}^{x} f''(t) \cdot (x-t)\, dt\, dx \\
&\quad - \frac{h}{2} \left(\int_{x_i}^{x_i} f''(t) \cdot (x_i - t)\, dt + \int_{x_i}^{x_{i+1}} f''(t) \cdot (x_{i+1} - t)\, dt \right) \\
&= \int_{x_i}^{x_{i+1}} \int_{t}^{x_{i+1}} f''(t) \cdot (x-t)\, dx\, dt - \frac{h}{2} \int_{x_i}^{x_{i+1}} f''(t) \cdot (x_{i+1} - t)\, dt \\
&= \int_{x_i}^{x_{i+1}} f''(t) \left(\int_{t}^{x_{i+1}} (x-t)\, dx - \frac{h}{2}(x_{i+1} - t) \right) dt \\
&= \int_{x_i}^{x_{i+1}} f''(t) \left(\frac{1}{2}(x_{i+1} - t)^2 - \frac{h}{2}(x_{i+1} - t) \right) dt \\
&= f''(\xi_n) \int_{x_i}^{x_{i+1}} \left(\frac{1}{2}(x_{i+1} - t)^2 - \frac{h}{2}(x_{i+1} - t) \right) dt \\
&= f''(\xi_n) \left(\frac{1}{6}(x_{i+1} - t)^3 - \frac{h}{4}(x_{i+1} - t)^2 \right) \Bigg|_{t=x_i}^{t=x_{i+1}} \\
&= f''(\xi_n) \frac{h^3}{12}.
\end{aligned}
$$

Für das Funktional R erhält man somit:

$$
\begin{aligned}
|Rf(x)| &= \left| \sum_{i=0}^{n-1} R_i f(x) \right| \\
&= \frac{h^3}{12} \left| \sum_{i=0}^{n-1} f''(\xi_i) \right| \\
&\leq \frac{h^3}{12} n \max_{\xi \in [a,b]} |f(\xi)| \\
&= \frac{h^2}{12} (b-a) \max_{\xi \in [a,b]} |f(\xi)|.
\end{aligned}
$$

□

4.3 Hilbertraumtheorie:

In den beiden Unterabschnitten dieses Abschnittes werden die theoretischen Grundlagen der Hilbertraumtheorie, die in Kapitel 1 benötigt werden, dargestellt. Der erste Teil beschäftigt sich mit den Voraussetzungen und den Räumen, die für den Schauderschen Fixpunktsatz benötigt werden. Der zweite Teil vermittelt Kenntnisse, wie man mit Orthonormalfolgen Funktionen approximieren kann, diese also als Linearkombination anderer Funktionen dartstellen kann.

4.3.1 Sobolev Räume:

Im Beweis zum Theorem in Kapitel 1 “Gesicherte Existenz- und Einschließungsergebnisse für periodische Systeme von Randwertproblemen” wurde der Schaudersche Fixpunktsatz und der Einbettungssatz von Sobolev, Kondrachew und Rellich verwendet, ohne dessen Voraussetzungen und dessen

Aussage anzugeben. Dies wird nun hier nachgeholt. Zunächst soll aber der Sobolevraum $H_m(\Omega)$, über den der Einbettungssatz ja Aussagen macht, definiert werden und einige seiner Eigenschaften beschrieben werden. Sei $\Omega \subseteq \mathbb{R}^d$ ein beschränktes Gebiet mit stückweise glattem Rand $\Gamma = \partial\Omega$. Dann seien die folgenden Räume definiert durch:

Definition ($L_2(\Omega), C_\infty(\Omega), C_{\infty,0}(\Omega)$):

$$\begin{aligned}
L_2(\Omega) &:= \Big\{u : \Omega \to \mathbb{R} \text{ mit } ||u||_2 \leq \infty \text{ mit durch Innenprodukt} \\
&\qquad < u, v >_2 := \int_\Omega u(t)\bar{v}(t)\,dt \text{ induzierter Norm}\Big\} \\
C_\infty(\Omega) &:= \{u \in L_2(\Omega) : u \text{ beliebig oft differenzierbar in } \Omega\} \\
C_{\infty,0}(\Omega) &:= \{u \in C_\infty(\Omega) : u = 0 \text{ außerhalb einer kompakten Teilmenge von } \Omega\}
\end{aligned}$$

Um im weiteren den Schreibaufwand zu verringern und die Übersichtlichkeit zu erhöhen, sei hier der Begriff des *Multi-Indizes* für ein Tupel $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$ eingeführt. Der Begriff *Betrag von* α bezeichnet den Wert $|\alpha| := \sum_{i=1}^d \alpha_i$. Mit diesen Vereinbarungen lassen sich dann die partiellen Ableitungen einer Funktion $u \in C_\infty(\Omega)$ auf einfache Art und Weise schreiben als:

$$D^\alpha := \frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial t_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial t_d^{\alpha_d}}.$$

In dieser Schreibweise definiert man:

Definition (schwache Ableitung):

$\partial^\alpha u \in L_2(\Omega)$ heißt *α-te schwache Ableitung* von $u \in L_2(\Omega)$ falls

$$< \Phi, \partial^\alpha u >_2 = (-1)^{|\alpha|} < D^\alpha \Phi, u >_2$$

für alle $\Phi \in C_{\infty,0}(\Omega)$ gilt.

Bemerkung:
Wie der Begriff der schwachen Ableitung motiviert ist, sieht man am besten, wenn man die geforderte Bedingung einmal für ein tatsächlich differenzierbares $u \in C_1(\Omega)$ betrachtet. Nach dem Gaußschen Integralsatz gilt dann nämlich:

$$\left\langle \Phi, \frac{\partial}{\partial t_i} u \right\rangle_2 + \left\langle \frac{\partial}{\partial t_i} \Phi, u \right\rangle_2 = \int_\Omega \frac{\partial}{\partial t_i}(\Phi u)\,dt = \int_\Gamma \Phi \cdot u \cdot n_i\,ds = 0.$$

In der obigen Notation gilt also für alle $|\alpha| = 1$

$$< \Phi, \partial^\alpha u >_2 = (-1)^{|\alpha|} < D^\alpha \Phi, u >_2 \quad \forall \Phi \in C_{\infty,0}(\Omega).$$

Die schwache Ableitung ist somit nichts anderes als eine Verallgemeinerung der Ableitung.

Mit dieser schwachen Ableitung ist es nun möglich Sobolev-Räume zu definieren:

Definition ($H_m(\Omega)$):

Der *Sobolev-Raum* $H_m(\Omega)$ mit $m = 0, 1, \ldots$ ist gegeben durch:

$$H_m(\Omega) := \{u \in L_2(\Omega) : \partial^\alpha u \text{ existiert und } \partial^\alpha u \in L_2(\Omega)\ \forall |\alpha| \leq m\}$$

Satz (Vollständigkeit von H_m bezüglich $||.||_{H_m}$)

Der Sobolevraum $H_m(\Omega)$ ist bezüglich des Skalarproduktes

$$< u, v >_{H_m} := \sum_{|\alpha| \leq m} < \partial^\alpha u, \partial^\alpha v >_2$$

vollständig, also Hilbertraum. Die dazugehörige Sobolevnorm $||.||_{H_m}$ hat die Gestallt:

$$||u||_{H_m} := \sqrt{< u, u >_{H_m}} = \left(\sum_{|\alpha| \leq m} ||\partial^\alpha u||_2^2 \right)^{\frac{1}{2}}.$$

Beweis: Sei $(u_i)_{i=1,\ldots,\infty}$ Cauchy-Folge bezüglich der Sobolevnorm $||.||_{H_m}$, dann ist für alle $\alpha : |\alpha| \leq m$ die Folge $(\partial^\alpha u_i)_{i=1,\ldots,\infty}$ Cauchy Folge in $L_2(\Omega)$. Aus der Vollständigkeit von $L_2(\Omega)$ erhält man nun $||\partial^\alpha u_i - u^\alpha||_2 \to 0$ für ein $u^\alpha \in L_2(\Omega)$. Es bleibt also nur noch zu zeigen: $\partial^\alpha u = u^\alpha$. Aus dem oben gezeigten $\partial^\alpha u_i \to u^\alpha$ in $L_2(\Omega)$ folgt dann nach der Cauchy-Schwarzschen-Ungleichung:

$$< \partial^\alpha u_i - u^\alpha, \Phi >_2 \leq ||\partial^\alpha u_i - u^\alpha||_2\, ||\Phi||_2 \overset{i \to \infty}{\to} 0 \quad \forall \Phi \in C_{\infty,0}(\Omega),$$

also $< \partial^\alpha u_i, \Phi >_2 \to < u^\alpha, \Phi >$. Somit gilt dann

$$\begin{aligned} < u^\alpha, \Phi >_2 &= \lim_{i \to \infty} < \partial^\alpha u_i, \Phi >_2 = \lim_{i \to \infty} (-1)^{|\alpha|} < u_i, D^\alpha \Phi >_2 \\ &= (-1)^{|\alpha|} < u, D^\alpha \Phi >_2, \end{aligned}$$

woraus folgt $u^\alpha = \partial^\alpha u$. □

Der nun folgende Satz zeigt, daß man Sobolevräume nicht nur als Räume mit schwachen Ableitungen sehen darf, sondern daß sie auch Verfollständigungen von $C_\infty(\Omega)$ sind:

Satz (Meyers/Serring 1964, "$H = W$"):

Es gilt

$$C_\infty \cap H_m(\Omega) \text{ liegt dicht in } H_m(\Omega).$$

Somit lassen sich die Sobolevräume $H_m(\Omega)$ auch als Verfollständigung von $C_\infty(\Omega)$ bezüglich der Sobolevnorm $||.||_{H_m}$ auffassen.

Beweis: Ein Beweis dieses Satzes ist nachzulesen in Joseph Wloka; *Partielle Differentialgleichungen*; Teubner Stuttgart 1982. □

Nun nachdem die Sobolevräume $H_m(\Omega)$ definiert wurden und einige grundlegende Eigenschaften gezeigt wurden, wird es Zeit sich den Voraussetzungen- und schließlich dem Einbettungssatz von Sobolev, Kondrachov und Rellich zuzuwenden. Sei dazu von nun an $(R, ||.||)$ ein Banachraum über $\mathbb{R}$:

Definition (Stetigkeit):

Sei $D(F) \subseteq R$. Eine Abbilung $F : D(F) \to R$ heißt *stetig*, wenn für jede Folge $(x_i)_{i\in\mathbb{N}}$ mit $x_i \in D(F)$ und $\lim_{i\to\infty} x_i = x \in D(F)$ gilt:

$$\lim_{i\to\infty} F(x_i) = F(x)$$

Definition (Kompaktheit):

Eine Teilmenge K von R heißt *kompakt*, wenn jede Folge in K eine konvergente Teilfolge mit Grenzwert in K enthält.

Achtung: Ist $\dim(R) = \infty$, so ist zwar jede kompakte Menge beschränkt und abgeschlossen, die Umkehrung, daß jede beschränkte und abgeschlossene Menge auch kompakt ist gilt aber nur für $\dim(R) < \infty$. (Beispiel: R Hilbertraum, K sei die Einheitskugel, $(x_i)_{i\in\mathbb{N}}$ sei ein Orthonormalsystem $\Rightarrow$ $||x_i - x_j||^2 = 2$ für $i \neq j$, es gibt somit keine konvergente Teilfolge, obwohl die Einheitskugel beschränkt und abgeschlossen ist).

Definition (relativer Kompaktheit):

Eine Teilmenge $K \subseteq R$ heißt *relativ kompakt* oder *totalbeschränkt*, wenn zu jeder Folge $(x_i)_{i\in\mathbb{N}}$ in K eine konvergente Teilfolge (mit Grenzwert nicht notwendigerweise in K) existiert ($\Leftrightarrow \bar{K}$ ist kompakt)

Definition (Kompaktheit eines Operators):

Ein Operator $T : D \to R$ mit $D \subseteq R$ heißt *kompakt* in D, wenn $T(D)$ relativ kompakt ist.

Definition (Konvexheit):

Eine Menge $D \subseteq R$ heißt *konvex*, wenn für beliebige $x, y \in D$ die Verbindungsstrecke $\overline{xy} = \{\lambda x + (1-\lambda)y : 0 \leq \lambda \leq 1\}$ zu D gehört.

Schauderscher Fixpunktsatz:

Sei $D \subseteq R$ abgeschlossen und konvex, $T : D \to R$ sei stetig und kompakt. Außerdem gelte $T(D) \subseteq D$. Dann existiert ein Fixpunkt von T in D.

Beweis: Zum Beweis sei auf [7] verwiesen, wo der Beweis des Schauderschen Fixpunktsatzes auf den des Brouwerschen- und eines weiteren Fixpunktsatzes zurückgeführt wird. $\square$

Einbettungssatz von Sobolev, Kondrachov und Rellich:

Sei $\Omega \subseteq \mathbb{R}^d$ nun ein beschränktes Gebiet mit Lipschitzrand $\partial\Omega$, dann ist für $m, j \in \mathbb{N}$, die Einbettung $H_m(\Omega) \hookrightarrow C_j(\bar{\Omega})$ kompakt, falls $m > j + \frac{d}{2}$.

Beweis: Zum Beweis sei auf [8] verwiesen, in dem alle möglichen Arten von Einbettungen zwischen $H_{m,p}(\Omega)$ und $H_{n,p}(\Omega)$ sowie auch zwischen $H_{m,p}(\Omega)$ und $C_j(\Omega)$ behandelt werden. (Dabei drückt der Index $p \in [0, \infty]$ aus, das allgemeiner als hier nicht nur Sobolevräume betrachtet werden, die aus L_2-Räumen konstruiert wurden, sondern allgemeiner solche, die aus L_p-Räumen hervorgegangen sind). Die Orginalbeweise finden sich in F. Rellich; *Ein Satz über mittlere Konvergenz*; Göttinger Nachr. 1933 und B. I. Kondrachov; *Certain properties of functions in the space L^p*; Dokl. Akad. Nauk. SSSR 48 1945. $\square$

Bemerkungen:
Aus dem Einbettungssatz von Sobolev, Kondrachov und Rellich folgt sofort, daß die Einbettung $H_2(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1[0, k\frac{2\pi}{\mu}]$ kompakt ist. Es bleibt aber noch zu zeigen, daß auch die Einbettung $H_2^n(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1^n[0, k\frac{2\pi}{\mu}]$, die im Beweis des Theorems verwendet wurde, kompakt ist. Sei als (x_i) $i \in \mathbb{N}$ eine Folge in $H_2^n(0, k\frac{2\pi}{\mu})$. Da $H_2(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1[0, k\frac{2\pi}{\mu}]$ kompakt ist, kann man eine Teilfolge (x_{i_k}) von (x_i) auswählen, so daß $(Ex_{i_k})_1$ (wobei E die Einbettung von $H_2^n(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1^n[0, k\frac{2\pi}{\mu}]$ bezeichne) konvergent ist. Sei diese Teilfolge, die unter E in der ersten Komponente konvergent ist, zur Vereinfachung wiederum mit (x_i) bezeichnet. Dann gilt erneut, da $H_2(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1[0, k\frac{2\pi}{\mu}]$ kompakt ist, daß man eine Teilfolge (x_{i_k}) auswählen kann, für die gilt: $(Ex_{i_k})_2$ ist konvergent. Da die Folge (x_i) aber schon so gewählt war, daß die erste Komponente unter E konvergent ist, hat man nun die Konvergenz von (x_i) unter E in der ersten und zweiten Komponente. Setzt man diesen Prozess nun n-mal fort, so erhält man also ausgehend von der Folge (x_i) eine Teilfolge, die unter der Abbildung E in allen Komponenten konvergent ist. Da $||.||_{\infty, C_1^n[0,k\frac{2\pi}{\mu}]}$ als das Maximum von $||.||_{\infty, C_1[0,\frac{2\pi}{\mu}]}$ jeder Komponente definiert wurde, ist somit die Teilfolge auch in $C_1^n[0, k\frac{2\pi}{\mu}]$ versehen mit der Maximumnorm konvergent und somit $H_2^n(0, k\frac{2\pi}{\mu}) \hookrightarrow C_1^n[0, k\frac{2\pi}{\mu}]$ kompakt.

4.3.2 Die Fourierentwicklung im Hilbertraum:

Im Abschnitt 1.3.1 "Zu Punkt 1: Bestimmung der Näherungslösung ω" wurde die Fourierentwicklung zur Gewinnung einer Darstellung der Näherungslösung ω verwendet, ohne dabei auf Konvergenz- und Approximationseigenschaften einzugehen. Dies soll hier nun in allgemeinerem Rahmen (nicht nur für L_2) nachgeholt werden. Dabei wird der Vorgehensweise in [7] gefolgt.

Im ganzen folgenden Abschnitt sei H einer reeller oder komplexer Hilbertraum mit zugehörigem Innenprodukt $< u, v >$ und dadurch induzierter Norm $||u|| := \sqrt{< u, u >}$.

Für die Betrachtung der Fourierentwicklung wird sich der Begriff der Orthonormalfolge als elementar herausstellen, weshalb hier die folgende Definition folgt.

Definition (Orthonormalfolge):

- Zwei Elemente $u, v \in H$ heißen orthogonal, wenn $< u, v >= 0$ gilt.
- Eine Folge $(u_i)_{i=0,\ldots,\infty}$ heißt Orthonormalfolge, wenn

$$\begin{aligned} < u_i, u_j > &= 0 \qquad && \forall i, j \in \mathbb{N} \text{ mit } i \neq j, \\ < u_i, u_i > &= 1 && \forall i \in \mathbb{N} \end{aligned}$$

  gilt. Mit anderen Worten, wenn alle Folgenglieder paarweise orthogonal sind und die Norm eins haben.

Mit den so definierten Orthonormalfolgen ist es nun möglich, ganz allgemein Fourierentwicklungen zu gewinnen und für diese eine Approximationsformel abzuleiten.

Satz (Approximationsformel):

Definiert man für ein Element $u \in H$ die sogenannten *Fourierkoeffizienten* durch

$$\gamma_i :=< u, u_i >$$

mit $(u_i)_{i=0,\ldots,\infty}$ Orthonormalfolge, so gilt für endliche Indexmengen A die folgende Approximationsformel:

$$\left|\left| u - \sum_{i\in A} \alpha_i u_i \right|\right| = ||u||^2 - \sum_{i\in A} |\gamma_i|^2 + \sum_{i\in A} |\alpha_i - \gamma_i|^2.$$

Beweis:

$$\begin{aligned} \left|\left| u - \sum_{i\in A} \alpha_i u_i \right|\right| &= \left\langle u - \sum_{i\in A} \alpha_i u_i, u - \sum_{i\in A} \alpha_i u_i \right\rangle \\ &= ||u||^2 - \left\langle u, \sum_{i\in A} \alpha_i u_i \right\rangle - \left\langle \sum_{i\in A} \alpha_i u_i, u \right\rangle + \left|\left| \sum_{i\in A} \alpha_i u_i \right|\right|^2 \\ &= ||u||^2 - \sum_{i\in A} \bar{\alpha}_i \underbrace{< u, u_i >}_{=\gamma_i} - \sum_{i\in A} \alpha_i \underbrace{< u_i, u >}_{=\bar{\gamma}_i} + \sum_{i\in A} \alpha_i \bar{\alpha}_i \\ &= ||u||^2 + \sum_{i\in A} \underbrace{(\alpha_i \bar{\alpha}_i - \bar{\alpha}_i \gamma_i - \alpha_i \bar{\gamma}_i)}_{=|\alpha_i - \gamma_i|^2 - |\alpha_i|^2} \end{aligned}$$

□

Mit der soeben bewiesenen Approximationsformel läßt sich nun leicht der folgende Approximationssatz gewinnen, der zeigt, daß die Fourierreihe bezüglich der $||.||$ Norm eine optimale Näherung darstellt.

Approximationssatz:

Sei $(u_i)_{i=0,\ldots,\infty}$ beliebige Orthonormalfolge.

Endlicher Fall: Sei A eine beliebige endliche Indexmenge, so stellt die endliche Fourierreihe bezüglich der Orthonormalfolge $(u_i)_{i=0,\ldots,\infty}$

$$u_A := \sum_{i\in A} < u, u_i > u_i = \sum_{i\in A} \gamma_i u_i$$

unter allen Linearkombinationen $\sum_{i\in A} \alpha_i u_i$ die beste Approximation von u dar. Außerdem gibt es keine weitere beste Approximation. Es gilt die Approximationsformel:

$$||u - u_A||^2 = ||u|| - \sum_{i\in A} |\gamma_i|^2.$$

Unendlicher Fall: Die unendliche Reihe

$$u^* := \sum_{i=0}^{\infty} < u, u_i > u_i = \sum_{i=0}^{\infty} \gamma_i u_i$$

ist konvergent und wird *Fourierreihe* von u bezüglich der Orthonormalfolge $(u_i)_{i=0,\ldots,\infty}$ genannt. Für die so definierte Fourierreihe u^* gilt, wie im endlichen Fall

$$||u - u^*||^2 = ||u|| - \sum_{i=0}^{\infty} |\gamma_i|^2,$$

sowie

$$< u - u^*, u_i >=< u, u_i > - < u^*, u_i >= \gamma_i - \sum_{j=1}^{\infty} \gamma_j < u_j, u_i >= 0 \quad \forall i.$$

Beweis:

Endlicher Fall: Für die Norm der Differenz zwischen dem zu approximierenden Element u und der endlichen Fourierentwicklung u_A gilt

$$\begin{aligned} ||u - u_A||^2 &\overset{\substack{\text{Approximations-}\\ \text{formel}}}{=} ||u||^2 - \sum_{i\in A} |\gamma_i|^2 \\ &< ||u||^2 - \sum_{i\in A} |\gamma_i|^2 + \sum_{i\in A} |\alpha_i - \gamma_i|^2 \text{ falls } \alpha_i \neq \gamma_i \text{ für ein } i \in A \\ &\overset{\substack{\text{Approximations-}\\ \text{formel}}}{=} \left|\left| u - \sum_{i\in A} \alpha_i u_i \right|\right|^2 . \end{aligned}$$

Unendlicher Fall: Da $0 \leq ||u - u_A||^2 = ||u||^2 - \sum_{i\in A} |\gamma_i|^2$ gilt, haben alle endlichen Summen $\sum_{i\in A} |\gamma_i|^2$ die obere Schranke $||u||^2$. Die Reihe $\sum_{i=0}^{\infty} |\gamma_i|^2$ ist somit konvergent. Für die Differenz der q-ten und p-ten

Teilsumme der Fourierreihe erhält man somit $||s_q - s_p||^2 = \sum_{p+1}^{p} |\gamma_i|^2$. Da $\sum_{i=0}^{\infty} |\gamma_i|^2$ konvergiert gibt es zu jedem $\varepsilon > 0$ einen Index N, derart daß $\sum_{p+1}^{p} |\gamma_i|^2 < \varepsilon \ \forall p, q \geq N$ gilt. Es gilt somit $\forall \varepsilon > 0 \ \exists N$: $||s_q - s_p|| \leq \varepsilon \ \forall q, p \geq N$. Die Teilsummen der Fourierreihe bilden somit eine Cauchyfolge. Da H vollständig ist, konvergiert die Fourierreihe dann gegen ein Element $u^* \in H$. Die Approximationsformel für den unendlichen Fall erhält man durch Grenzübergang in der Approximationsformel für den endlichen Fall.

□

Bemerkung:
Die Koeffizienten der Fourierreihe sind völlig unabhänig von der Indexmenge A.

Bisher wurde also gezeigt, wie man, wenn man eine Orthonormalfolge gegeben hat, die Koeffizienten der Reihe zu wählen hat, um eine Bestapproximation bezüglich ——.—— zu erhalten. Als nächstes stellt sich nun auf natürliche Weise die Frage, wann nun gar $u = u^*$ gilt. Um diese Fragestellung beleuchten zu können, wird sich der Begriff der vollständigen Orthonormalfolge, der hier als nächstes definiert werden soll, als notwendig erweisen.

Definition (vollständige Orthonormalfolge)

Eine Orthogonalfolge $(u_i)_{i=0,\ldots,\infty}$ heißt *vollständige orthogonale Folge*, wenn aus $< u, u_i >= 0 \ \forall i = 0, \ldots, \infty$ folgt $u = 0$.

Der Bergriff der vollständigen Orthogonalfolge stellt ein einfach zu handhabendes Kriterium dar, das sagt, wann die Fourierreihe gegen das zu approximierende Element konvergiert. Es gilt der

Darstellungssatz:

Ist $(u_i)_{i=1,\ldots,\infty}$ vollständige Orthogonalfolge, so läßt sich jedes Element $u \in H$ durch seine Fourierreihe darstellen:

$$u = \sum_{i=0}^{\infty} < u, u_i > u_i.$$

Beweis: Aus dem unendlichen Fall des Approximationssatz ersieht man, daß $< u - u^*, u_i >= 0 \ \forall i = 0, \ldots, \infty$ gilt. Da $(u_i)_{i=0,\ldots,\infty}$ nach Voraussetzung vollständige Orthogonalfolge ist folgt nach Definition $u - u^* = 0$, also die Behauptung. □

Bemerkung:
Im Hilbertraum $L_2(0, k\frac{2\pi}{\mu})$ mit dem Innenprodukt $< u, v >_2 := \int_0^{k\frac{2\pi}{\mu}} u(t)\bar{v}(t)\, dt$ ist $\sqrt{\frac{\mu}{k\pi}}(\frac{1}{\sqrt{2}}, \cos t, \sin t, \cos 2t, \sin 2t, \ldots)$ ein vollständiges Orthogonalsystem. Somit läßt sich jedes Element $u \in L_2(0, k\frac{2\pi}{\mu})$ durch seine Fourierreihe darstellen.

Literaturverzeichnis

[1] M. Plum; *Verified existence and inclusion results for two-point value problems* ; Contributions to Computer Arithmetic and Self-Vaidating numerical Methods, Scientific Publishing Co.

[2] M. Plum; *Guaranted Numerical Bounds for Eigenvalues*; Mathematisches Institut I, Universität Karlsruhe, D-76128 Karlsruhe, Germany.

[3] M. Plum; *Computer-assisted existence proofs for two-point boundary value problems*; Computing 46, 19 - 34 (1991).

[4] P. J. McKenna, Y. S. Choi, Hyeyoung Oh Her; *Periodic Oscillations of a Suspension Bridge Model.*

[5] Y.S.Choi, Hyeyoud Oh Her, P.J. McKenna; *Galloping: nonlinear oscillation in a periodically forced loaded hanging cable*; Journal of Computational and Applied Mathematics 52 (1994) 23 - 34.

[6] P. Vielsack, H. Wei; *Sensitivität der harmonischen Schwingung eines Hängebrückenmodells mit asymmetrischer Dissipation*; Archive of Applied Mechanics 64 (1994) 408 - 416 Springer Verlag.

[7] W. Walter; *Gewöhnliche Differentialgleichungen, 3. Aufl.*; Heidelberger Taschenbücher 110, Springer 1986.

[8] R. A. Adams; *Sobolev Spaces*; New York: Acadamic Press 1975.

[9] J. Stoer; *Numerische Mathematik 1, 6. Aufl.*; Springer-Verlag Berlin Heidelberg 1993.

[10] J. Stoer, R. Bulirsch; *Numerische Mathmatik 2, 3. Aufl.*; Springer-Verlag Berlin Heidelberg 1990.

[11] G. Hämmerlin K.-H. Hoffmann; *Numerische Mathematik*; Springer Verlag.

[12] G. Alefeld, J. Herzberger; *Einführung in die Intervallrechnung*; B.I. Wissenschaftsverlag 1974.

Die Bücher und Artikel des Literaturverzeichnis sind nicht nach den Verfassern alphabetisch geordnet sondern nach Themengebieten: [1] bis [3] beschäftigen sich mit dem Einschließungsalgorithmus; [4] bis [6] enthält die Punkte, die das Modellieren einer Hängebrücke betreffen; aus [7] bis [8] stammen die funktionalanalytischen Grundlagen, und [9] bis [12] enthält die numerischen Algorithem und Verfahren um den Existenzalgorithmus auf dem Computer. umsetzen zu können. Ich hoffe, daß sich keiner der Verfasser durch meine Ordnung bevorzugt oder benachteiligt fühlt, denn das entspricht nicht meiner Absicht.

Index