

PHYSIKALISCHE MODELLIERUNG UND EFFIZIENTE NUMERISCHE SIMULATION VON LIION-HOCHLEISTUNGSAKKUMULATOREN MITTELS SPEKTRAL- UND LINIENMETHODEN

JULIUS ZWIRNER

ZUSAMMENFASSUNG. **Zwischenstand bitte nicht weitergeben.**

INHALTSVERZEICHNIS

1. Einleitung	1
2. Physikalische Modellierung einer Lithium-Ionen-Zelle	2
2.1. Funktionsprinzip	2
2.2. Newman-Modell	3
2.3. Gleichungs-Übersicht	9
3. Numerische Lösung	14
3.1. Linienmethode und räumliche Diskretisierung	14
3.2. Differential-Algebraische Gleichungen	23
3.3. Zeitliche Integration der DAE	32
3.4. Analytischen Teillösung	41
3.5. Aspekte der Implementierung	48
4. Ergebnisse	50
4.1. Konvergenz	50
4.2. Vergleich mit COMSOL	50
5. Zusammenfassung	51
Anhang A. Parameter	51
Literatur	53

1. EINLEITUNG

Seit ihrer Entdeckung in den späten 1970er und 1980er Jahren haben Lithium-Ionen-Akkumulatoren sich nicht nur als Energieträger für mobile Geräte etabliert, sondern dank steigender Energiedichten und fallender Preise auch eine neue Phase in der Entwicklung der Elektromobilität eingeleitet [26, 28, 34, 46]. Im Zuge des Ausbaus erneuerbarer, fluktuierender Energien gewinnen auch Heimspeicher und Speicherkraftwerke auf Lithium-Ionen-Basis weiter an Bedeutung [28, 36]. Zugleich wächst das Interesse an Fragen der maximalen Ausnutzung und

der Schnellladung bei gleichzeitiger Wahrung der Sicherheits- und Alterungsbeschränkungen. Solche Fragen erfordern die Entwicklung fortschrittlicher Algorithmen innerhalb des für die Überwachung und Steuerung zuständigen Batteriemanagementsystems (BMS).

Gegenwärtige Algorithmen basieren im Allgemeinen auf empirischen Ersatzschaltbildmodellen [8, 23]. Diese sind zwar mit niedrigem Rechenaufwand zu lösen, aber nur im Bereich der bei der Parametrisierung vorliegenden Betriebsbedingungen gültig.

Eine Alternative stellen physikalische Zellmodelle dar, die die Reaktionskinetik der Elektroden, den Ionen-Transport im Elektrolyten und die Thermodynamik in der Zelle beschreiben. Diese Modelle führen zu gekoppelten Systemen gewöhnlicher und partieller Differentialgleichungen, deren Lösung zwar einen deutlich höheren Rechenaufwand erfordert, aber auch in einem größeren Bereich gültig bleibt und einen Einblick in die in der Zelle ablaufenden Prozesse gewährt.

Das erste und bekannteste Modell dieser Art ist das Newman-Modell [10, 15, 33]. Ein alternatives Modell wurde in [30] vorgestellt. Der Trend geht aber in Richtung Weiterentwicklungen des Newman-Modells hinsichtlich detaillierter Modellierung der Mikrostruktur oder Berücksichtigung der Elektroden-Elektrolyt-Grenzschicht (SEI) [11, 1] auf der einen Seite und in Richtung Integration des Newman-Modell in mehrdimensionale Thermo- und CFD-Modelle auf Zell- oder Packebene auf der anderen Seite [17, 14, 24, 20].

Gleichzeitig hat man verschiedene Vereinfachungen des Newman-Modells entwickelt mit dem Ziel, diese in kompliziertere Modelle zu integrieren oder auf Mikrocontrollern zu implementieren. Eine bekannte Vereinfachung ist das „Single Particle Model“, in dem räumliche Variationen in Elektroden und Elektrolyt vernachlässigt werden und nur die Variation in der Partikeldimension betrachtet werden (siehe Beschreibung des Newman-Modells) [37, 31]. Diese Approximation ist allerdings nur für kleine Lade- oder Entladeströme zulässig. Andere Verfahren zur Ordnungsreduktion des Newman-Modells wurden in [7, 9, 43] betrachtet. Diese Modelle müssen aber i. A. *ad hoc* Approximationen zur Ordnungsreduktion treffen, was die Interpretierbarkeit der Ergebnisse erschwert.

In dieser Arbeit haben wir uns dagegen bemüht, dass vollständige Newman-Modell möglichst effizient zu lösen.

2. PHYSIKALISCHE MODELLIERUNG EINER LITHIUM-IONEN-ZELLE

2.1. Funktionsprinzip. Eine Lithium-Ionen-Zelle besteht aus zwei Elektroden, in denen Lithium eingespeichert (interkaliert) werden kann, und einem nichtleitenden Separator, der einen Kurzschluss der Elektroden verhindert. Elektroden und Separator sind in einen Elektrolyten getränkt, der den Ionentransport ermöglicht. Als Elektroden-Materialien

werden in der Regel Übergangsmetallverbindungen wie LiCoO_2 , LiMn_2O_4 oder LiFePO_4 für die positive Elektrode und Graphit für die negative Elektrode verwendet; für den Elektrolyten nutzt man nicht-wässrige Lösungsmittel wie etwa Ethylencarbonat und Leitsalz wie LiPF_6 [26, 28]. Im geladenen Zustand ist Lithium vorwiegend in der Graphit-

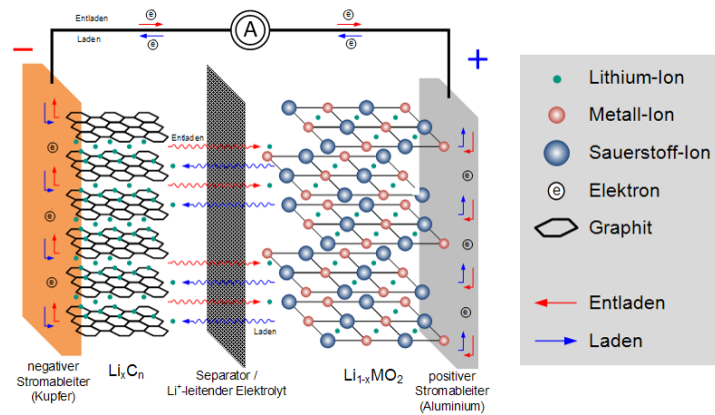


ABBILDUNG 1. Funktionsprinzip einer Li-Ionen-Zelle,
KIT Institute of Thermal Process Engineering [?]

struktur der negativen Elektrode interkaliert. Werden die Stromableiter beider Elektroden nun über einen externen Stromkreis miteinander verbunden, so fließen - aufgrund des niedrigeren Redoxpotentials von Lithium - Elektronen von der negativen zur positiven Elektrode. Gleichzeitig werden an der negativen Elektrode (Anode) Lithium-Ionen abgeschieden, um, angetrieben von einem Potential- und Konzentrationsgefälle, durch den Elektrolyten zur positiven Elektrode (Kathode) zu wandern, wo sie wieder in die feste Struktur interkalieren. Die Zelle wird entladen und die Potentialdifferenz zwischen den Elektroden nimmt ab. Wird umgekehrt eine Spannung angelegt, so kehren sich die Vorgänge um und die Zelle wird wieder geladen.

Bemerkung. Nach der üblichen Konvention werden die Begriffe Anode und Kathode am Entladefall festgemacht. Unabhängig von der Richtung des Stromflusses macht es daher Sinn, von einer ‘Graphit-Anode’ oder einer ‘LFP-Kathode’ zu sprechen. Der Satz ‘Kationen wandern zur Kathode’ gilt dagegen nur für den Entladenfall.

2.2. Newman-Modell. Die in Abbildung 1 dargestellten Elektroden- und Separatorschichten werden in einer realen Zelle als dünne Lagen gefertigt und spiralförmig aufgerollt. Da die wesentlichen elektrochemischen Prozessen senkrecht zu den Elektroden-Schichten stattfinden, werden wir in unserem Modell die Effekte der Spiral-Geometrie vernachlässigen, die Lagen also wieder als „abgerollt“ betrachten, und darüber hinaus nur die Variation senkrecht zu den Lagen betrachten.

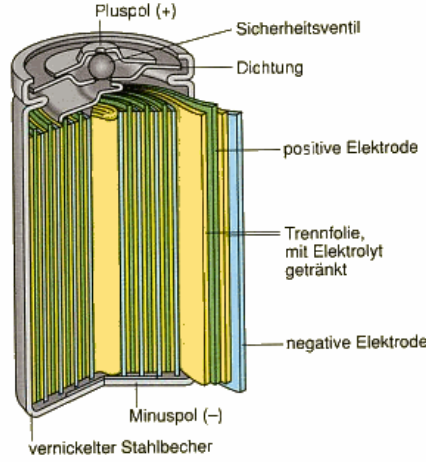


ABBILDUNG 2. Rundzelle mit aufgerollten Lagen

Diese Vereinfachung ist in dem Maße zulässig, wie die Lagen keine großen Inhomogenitäten aufweisen und die Dimensionen entlang der Lagen ihre Dicken übertreffen, sodass Randeffekte vernachlässigt werden können. Bei einer typischen 18650-Rundzelle beträgt die kombinierte Dicke der Elektroden- und Separatorlagen ca. 0.2 mm, ihre Höhe dagegen 65 mm und ihre Länge etwa 1.5 m, sodass diese Bedingung hinreichend erfüllt ist.

Wir erhalten somit ein räumliches 1D-Makro-Modell. Hinzu kommt eine weitere Mikro- oder Pseudo-Dimension r , in welcher der Interkationsprozess in den Elektrodenpartikeln modelliert wird. Aus x - und r -Dimension resultiert die gebräuchliche Bezeichnung P2D-Modell.

Als Zustandsgrößen betrachten wir: Die Konzentrationen c_l und c_s von Lithium-Ionen im Elektrolyten und im Aktivmaterial der Elektroden (in mol/m^3), die Ströme i_l und i_s in Elektrolyt und Elektroden (in A/m^2), der Reaktionsstrom i_r an den Partikel-Elektrolyt-Grenzflächen der Elektroden (in A/m^3), die elektrischen Elektrolyt- und Elektroden-Potentiale ϕ_l und ϕ_s , die Durchschnittstemperatur T sowie als Eingangsgröße den Klemmenstrom i . Alle diese Größen sind natürlich Funktionen der Zeit t . Bis auf T und i sind sie alle außerdem Funktionen von x . c_s ist schließlich auch noch eine Funktion von r .

Aufgrund der Ladungserhaltung gilt dabei

$$i_l + i_s = i. \quad (2.1)$$

Der Elektrolyt-Strom i_l wird von Ionen, der Elektroden-Strom i_s von Elektronen getragen. An den Stromableitern (d.h. für $x = 0$ und $x = l$) ist $i_l = 0$ und $i_s = i$, im Separator ($x_{neg} \leq x \leq x_{pos}$) gilt umgekehrt $i_l = i$ und $i_s = 0$. Die weiteren Beziehungen zwischen diesen Größen inklusive des gekoppelten thermischen Modells werden wir in den

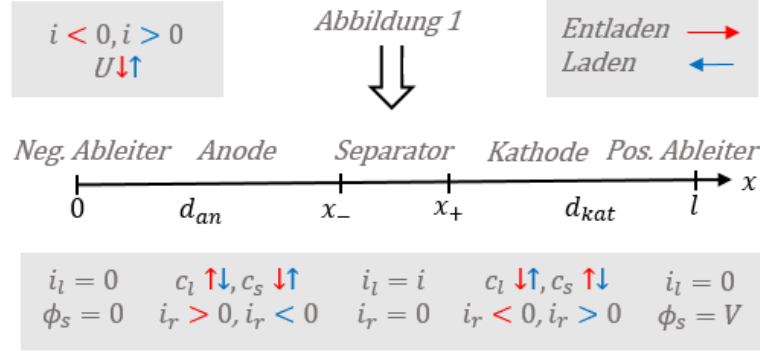


ABBILDUNG 3. Vom Zellschema zum mathematischen Modell.

nächsten fünf Unterabschnitten beschreiben und am Schluss noch mal alle Modellgleichungen in einer Übersicht zusammenstellen.

2.2.1. Poröse Struktur und Festkörperdiffusion. Um die Rate der Inter- bzw. Deinterkalationsreaktionen (und somit die Leistung der Zelle) zu steigern, werden die Elektroden kommerzieller Zellen in aller Regel aus porösen Strukturen gefertigt, deren Oberflächen mehrere Größenordnungen über denen solider Strukturen gleicher Volumina liegen [26]. Da eine mikroskopische Modellierung der porösen Struktur nicht in Frage kommt, tragen wir ihr in unserem Modell durch vier Parameter Rechnung: den Volumenanteilen ε_s und ε_l von Festkörper und Elektrolyt, der spezifischen Oberfläche a_s (in m^2/m^3), sowie dem typischen Partikelradius r_p . Wenn wir uns die poröse Struktur als Anzahl kleiner Kugeln des Radius r_p vorstellen, dann gilt $\varepsilon_s V = \frac{4}{3}\pi r_p^3 N$ und $a_s V = 4\pi r_p^2 N$, wobei N die Anzahl der Kugeln in einem Volumen V sind. Daher ist

$$a_s = \frac{3\varepsilon_s}{r_p}. \quad (2.2)$$

Die Interkalation von Lithium in den Partikeln modellieren wir als Diffusionsprozess in einer solchen Kugeln des Radius r_p . $c_s = c_s(x, r, t)$ bezeichnet somit die Lithium-Konzentration im Festkörper zur Zeit t an der Stelle x entlang der Dimension unseres Modells und eine Distanz r vom Partikelzentrum entfernt. (Es ist hilfreicher, sich die r -Dimension nicht strikt räumlich vorzustellen, sondern als Maß für die Verfügbarkeit von Lithium bzw. von Freistellen im Festkörper für die Deinterkalations bzw. Interkalation zu betrachten. Nur die Oberflächen-Konzentration $\bar{c}_s(x, t) = c_s(x, r = r_p, t)$ kann in den Elektrolyten übergehen; bei schneller Ladung oder Entladung kann es also zu „Stauungen“ bzw. „Engpässen“ kommen, wenn die Reaktion an der Grenzschicht schneller abläuft als die Diffusion von der bzw. an die Oberfläche.)

Es gilt also

$$\frac{\partial c_s}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 D_s \frac{\partial c_s}{\partial r} \right) \quad (2.3)$$

mit der Randbedingung

$$D_s \frac{\partial c_s}{\partial r}(r = r_p) = -\frac{i_r}{a_s F}. \quad (2.4)$$

Diese beschreibt die Reaktion an der Partikel-Oberfläche. Das Vorzeichen ist so gewählt, dass ein positiver Reaktionsstrom eine Deinterkation von Ionen aus dem Festkörper heraus anzeigt.

2.2.2. Butler-Volmer Kinetik. Der Reaktionsstrom i_r ergibt sich nach der Butler-Volmer-Gleichung zu

$$i_r = a_s i_0 \left(e^{\frac{\alpha_a F}{RT} \eta} - e^{-\frac{\alpha_c F}{RT} \eta} \right), \quad (2.5)$$

wobei F die Faraday-, R die allgemeine Gaskonstante und α_a und α_c sogenannte Symmetriekonstanten sind. Die Überspannung η ist definiert durch

$$\eta = \phi_s - \phi_l - \phi_0. \quad (2.6)$$

Die Potentiale ϕ_s und ϕ_l werden im nächsten Abschnitt eingeführt; $\phi_0 = \phi_0(\bar{c}_s, T)$ ist die experimentell bestimmte Gleichgewichtsspannung der jeweiligen Elektrode in Abhängigkeit der Oberflächenkonzentration $\bar{c}_s = c_s(r = r_p)$ und der Temperatur T . Die Austauschstromdichte i_0 ist schließlich erklärt durch

$$i_0 = F k_c^{\alpha_a} k_a^{\alpha_c} (c_s^{max} - \bar{c}_s)^{\alpha_a} \bar{c}_s^{\alpha_c} (c_l / c_{l,ref})^{\alpha_a} \quad (2.7)$$

für Reaktionsraten k_a , k_c sowie eine maximale Festkörperkonzentration c_s^{max} und eine Referenz-Elektrolytkonzentration $c_{l,ref}$.

Bemerkung. Die Verwendung von $c_{l,ref}$ hat den Vorteil, dass für den meistens betrachteten Fall $\alpha_a + \alpha_c = 1$ k_a und k_b die natürliche Dimension Länge/Zeit haben.

2.2.3. Potentiale in Elektroden und Elektrolyten. Im nächsten Schritt beschreiben wir die Beziehungen zwischen Strom und Spannung in fester und flüssiger Phase. In der festen Phase ist dies einfach das ohmsche Gesetz in differenzieller Form

$$\sigma_{eff} \frac{\partial \phi_s}{\partial x} + i_s = 0. \quad (2.8)$$

Aufgrund der porösen Struktur wird für die elektrische Leitfähigkeit die sogenannte Bruggeman-Korrektur $\sigma_{eff} = \sigma \varepsilon_s^{1.5}$ verwendet. In der flüssigen Phase wird diese Beziehung durch die Abhängigkeit von der Konzentration c_l kompliziert:

$$\kappa_{eff} \frac{\partial \phi_l}{\partial x} - \frac{2\kappa_{eff} RT(1 - t_+)}{F} \left(1 + \frac{\partial \ln f_{\pm}}{\partial c_l} \right) \frac{\partial \ln c_l}{\partial x} + i_l = 0. \quad (2.9)$$

Hierbei ist $\kappa_{eff} = \kappa \varepsilon_l^{1.5}$ die effektive elektrolytische Leitfähigkeit und f_{\pm} die mittlere molare Aktivität. Die Überföhrungszahl t_+ bezeichnet den Anteil des Elektrolyt-Stromes i_l , der von Li^+ -Ionen getragen wird. Man erhält Gleichung (refeq: Elektrolyt-Potential), wenn man Gleichung 2.32 in [32] nach x ableitet:

$$\frac{\partial \mu_e}{\partial x} = 2RT \left(\frac{\partial \ln c_l}{\partial x} + \frac{\partial \ln f_{\pm}}{\partial x} \right) = 2RT \left(1 + \frac{\partial \ln f_{\pm}}{\partial \ln c_l} \right) \frac{\partial \ln c_l}{\partial x} \quad (2.10)$$

und in Gleichung 12.27 für einen binären Elektrolyten in 1D einsetzt. Wir weisen darauf hin, da der entsprechende Ausdruck meist in allgemeinerer Form [33], ohne Herleitung [10, 15] oder in nicht offensichtlich äquivalenter Form verwendet wird [3, 13]. Siehe dazu auch [4].

Dass allgemein eine solche Konzentrations-Abhängigkeit existieren muss, kann man sich am folgenden Beispiel einer Konzentrationszelle klar machen. Dort werden zwei identische Kupfer-Elektroden in CuSO_4 -Lösungen unterschiedlicher Konzentration c_1, c_2 getaucht. Die Lösungen sind durch eine poröse Glasscheibe verbunden, die eine schnelle Durchmischung der Lösungen verhindert, aber einen Ionen-Strom und langsame Diffusion zulässt. Zwischen den Elektroden lässt sich dann eine Potentialdifferenz von etwa

$$\phi_1 - \phi_2 = (1 - t_+) \frac{RT}{F} \ln \frac{c_1}{c_2} \quad (2.11)$$

messen. Werden die beiden Elektroden über einen externen Stromkreis miteinander verbunden, fließt Strom bis sich die Konzentrationen angeglichen haben [32, S. 18-19].

2.2.4. Transport im Elektrolyten. Als nächstes wenden wir uns der Evolution der Li^+ -Konzentration im Elektrolyten zu. Der Gaußsche Satz, angewendet auf ein Volumen bestehend aus porösem Material und Elektrolyt, liefert die differentielle Bilanzgleichung

$$\varepsilon_l \frac{\partial c_l}{\partial t} = -\frac{\partial N_l}{\partial x} + R_l, \quad (2.12)$$

wobei N_l den Li^+ -Ionen-Fluss und R_l den Quellterm bezeichnet. Der Flussterm setzt sich aus einem diffusiven und einem migrativen Anteil zusammen:

$$N_l = -D_{l,eff} \frac{\partial c_l}{\partial x} + t_+ i_l / F. \quad (2.13)$$

Für den Quellterm gilt

$$R_l = i_r / F. \quad (2.14)$$

Außerdem gilt die Ladungsbilanz

$$\frac{\partial i_l}{\partial x} = i_r \quad (2.15)$$

Einsetzen von (2.13) und (2.14) in (2.12) ergibt unter der Annahme eines konzentrationsunabhängigen Diffusionskoeffizienten, einer konstanten Überföhrungszahl und Verwendung von (2.15)

$$\varepsilon_l \frac{\partial c_l}{\partial t} = D_{l,eff} \frac{\partial^2 c_l}{\partial x^2} - t_+/F \frac{\partial i_l}{\partial x} + a_s i_n = D_{l,eff} \frac{\partial^2 c_l}{\partial x^2} + (1 - t_+) i_r / F. \quad (2.16)$$

2.2.5. Thermisches Modell. Schließlich beschreiben wir das gekoppelte Thermomodell. Aufgrund der geringen Lagendicke und der guten Wärmeleitfähigkeit der Ableiter können wir Temperaturgradienten in der Zelle vernachlässigen und vereinfachend eine homogene Zelltemperatur annehmen. Somit haben wir nur die Wärmequellen und -senken in der Bilanzgleichung

$$\rho c_p \frac{dT}{dt} = q_r + q_{rev} + q_R + q_{ab} + q_{konv} \quad (2.17)$$

im Einheitsvolumen zu beschreiben. Wie [3] verwenden wir die Quellterme

$$q_r = \frac{1}{l} \int_0^l i_r (\phi_s - \phi_l - \phi_0) dx \quad (2.18)$$

$$q_{rev} = \frac{1}{l} \int_0^l i_r T \frac{\partial \phi_0}{\partial T} dx \quad (2.19)$$

$$q_{ab} = Ri^2 A^2 / V \quad (2.20)$$

$$q_{konv} = -hA(T - T_\infty) / V \quad (2.21)$$

für die Reaktionswärme, den reversiblen Wärmeeffekt, ohmsche Verluste an den Ableitern und den konvektiven Austausch mit der Umgebung, wobei R den Kontaktierungswiderstand, h den Wärmeaustauschkoeffizienten (Einheiten W/m^2K) und A und V die Zelloberfläche und das Zellvolumen bezeichnen.

Wir gehen von einer zylindrische 18650-Zelle ($r = 18/2 mm$, $h = 65 mm$) aus, sodass

$$A = 2\pi r^2 + 2\pi r h = 41.85 cm^2$$

$$V = \pi r^2 h = 16.54 cm^3.$$

Bei dem ohmschen Wärmeterm q_R ist - unter Berücksichtigung der jeweiligen Formulierung des ohmschen Gesetzes im Elektrolyten - ein Vorzeichenunterschied in den beiden von [3] hierzu zitierten Quellen [29, 49] festzustellen. Um diese Diskrepanz zu untersuchen, gehen wir auf die allgemeine von Newman aufgestellte Energieerhaltung in [32,

Kap. 13] zurück. Von dort führt die Herleitung in [19] zu dem Term

$$q_R = -\frac{1}{l} \int_0^l i_s \frac{\partial \phi_s}{\partial x} + i_l \frac{\partial \phi_l}{\partial x} dx \quad (2.22)$$

$$= \frac{1}{l} \int_0^l \left[\sigma_{eff} \left(\frac{\partial \phi_s}{\partial x} \right)^2 + \kappa_{eff} \left(\frac{\partial \phi_l}{\partial x} \right)^2 - \kappa_{eff} \frac{2aRT}{F} \frac{\partial \ln c_l}{\partial x} \frac{\partial \phi_l}{\partial x} \right] dx \quad (2.23)$$

mit einem numerisch motivierten „Aktivitätsfaktor“

$$a := (1 - t_+) \left(1 + \frac{\partial \ln f_{\pm}}{\partial c_l} \right). \quad (2.24)$$

Partielle Integration von (2.22) ergibt

$$lq_R = -[i_s \phi_s]_0^l - \int_0^l i_r \phi_s dx - [i_l \phi_l]_0^l + \int_0^l i_r \phi_l dx = -iU - \int_0^l i_r (\phi_s - \phi_l) dx$$

und somit

$$q_r + q_{rev} + q_R = \frac{1}{l} \int_0^l i_r \left(T \frac{\partial \phi_0}{\partial T} - \phi_0 \right) dx - iU/l. \quad (2.25)$$

Nimmt man konstante Gleichgewichtspotentiale ϕ_0^+ , ϕ_0^- in beiden Elektroden an und setzt $U_0 = \phi_0^+ - \phi_0^-$ ergibt sich

$$\begin{aligned} q_r + q_{rev} + q_R &= \frac{1}{l} \left(T \frac{\partial \phi_0^-}{\partial T} - \phi_0^- \right) \int_0^{x_-} i_r dx + \frac{1}{l} \left(T \frac{\partial \phi_0^+}{\partial T} - \phi_0^+ \right) \int_{x_+}^l i_r dx - iU \\ &= \frac{i}{l} (U_0 - U - T \frac{\partial U_0}{\partial T}). \end{aligned}$$

Dies ist gerade die bekannte Bernadi-Gleichung [2]. Wir verwenden sie in der räumlich aufgelösten Form (2.25). Wie dargestellt, enthält sie Reaktionswärme, reversible und ohmsche Wärmeverluste. Wärme durch Mischung oder Änderung des Aggregatzustandes werden dagegen vernachlässigt.

2.3. Gleichungs-Übersicht. Das gesamte Gleichungssystem des betrachteten Newman-Modells ist in Tabelle 1 zusammengestellt. Wir fanden es hilfreich, die Gleichungen den einzelnen Variablen zuzuordnen, auch wenn sie natürlich Beziehungen mehrerer Variablen zueinander ausdrücken. Heuristisch gesprochen haben wir so „sieben Gleichungen für sieben Unbekannte“. Wir haben in der Darstellung der Gleichungen die übliche Annahme $\alpha_a = \alpha_c = \frac{1}{2}$ getroffen, wodurch sich Gleichungen (2.5) und (2.7) vereinfachen. Die Spalte ‚Ableiter‘ enthält die Randbedingungen bei $x = 0$ und $x = l$. „ $i_l = 0$ “ etwa ist zu verstehen als $i_l(x = 0) = i_l(x = l)$, „ $\phi_s = 0$, U “ als $\phi_s(x = 0) = 0$, $\phi_s(x = l) = U$.

Tabelle 2 enthält die verwendeten Modellparameter, die wir [?] entnommen haben. Die Koeffizienten der Funktionen κ , a und D_l sowie

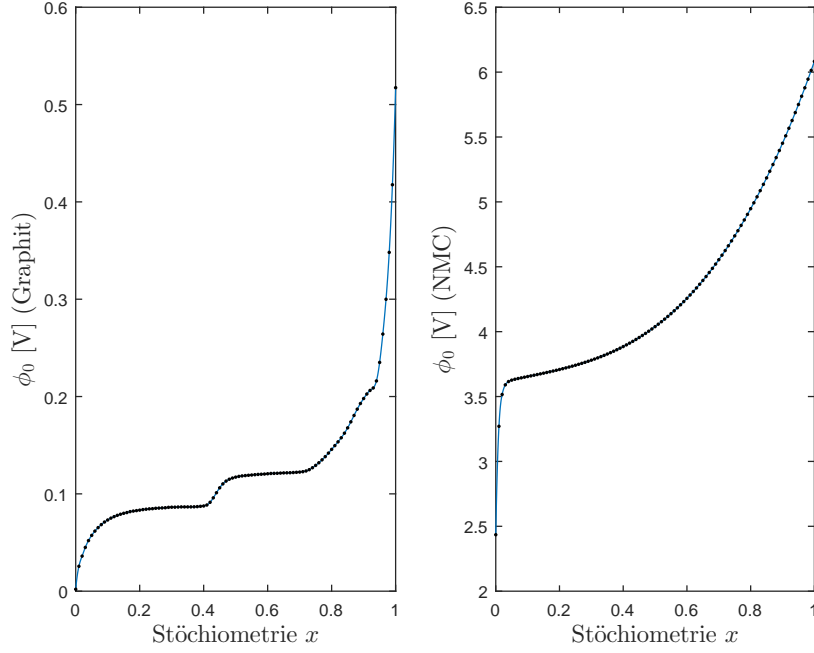


ABBILDUNG 4. Gleichgewichtsspannungen Graphit und NMC

die Messwerte der Funktionen ϕ_0 , $\phi_{0,T}$ sind in Appendix A aufgeführt. Die Messwerte für ϕ_0 wurden mit kubischen Splines interpoliert (siehe Abbildung 2.3).

Zu Messdaten $(x_i, y_i)_{i=0}^n$ berechnet man dazu bekanntlich einen C^2 -glatten Polynomzug P derart, dass die $p_i = P|_{[x_{i-1}, x_i]}$ kubische Polynome sind. Die insgesamt $4n$ Koeffizienten zur Spezifizierung von P werden durch die $2n$ Interpolationsbedingungen

$$p_i(x_{i-1}) = y_{i-1}, \quad p_i(x_i) = y_i, \quad i = 1, \dots, n, \quad (2.26)$$

die $2n - 2$ Glattheitsbedingungen

$$p'_i(x_i) = p'_{i+1}(x_i), \quad p''_i(x_i) = p''_{i+1}(x_i) \quad i = 1, \dots, n - 1 \quad (2.27)$$

sowie zwei Randbedingungen bestimmt. Für $\phi_{0,T}$ hätte Spline-Interpolation zu einem Overfitting und ungewünschten Oszillationen geführt (siehe Abbildung 2.3), was besonders bei der Berechnung der Ableitungen problematisch gewesen wäre. Wir haben deshalb die Funktion ‚Splinefit‘ aus dem MATLAB-Austausch verwendet [25]. Diese berechnet einen C^2 -glatten Polynomzug \tilde{P} mit einer kleineren Anzahl von Splines $k < n$ und bestimmt deren Koeffizienten mit einer kleinsten-Quadrate Methode. Es werden dadurch zwar nicht mehr alle Interpolationswerte exakt getroffen, dafür aber unrealistische Oszillationen vermieden.

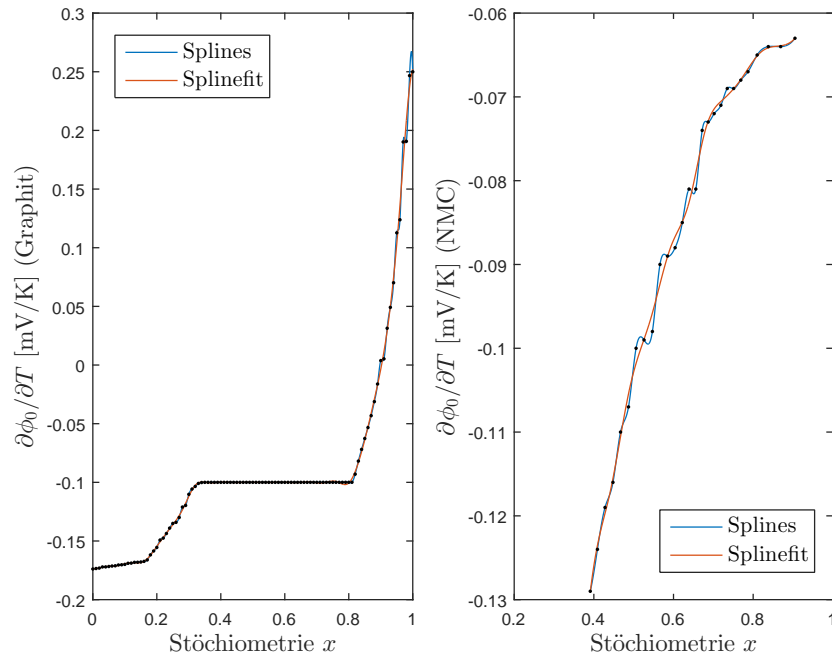


ABBILDUNG 5. Entropiekoeffizienten Graphit und NMC

TABELLE 1. Gleichungsübersicht

Variable	Ableiter	Separator	Elektroden	Nr
i_l	$i_l = 0$	$i_l = i$	$\frac{\partial i_l}{\partial x} = i_r$	(2.15)
ϕ_s	$\phi_s = 0, U$	-	$\sigma_{eff} \frac{\partial \phi_s}{\partial x} + i - i_l = 0$	(2.8)
ϕ_l	-	$\kappa_{eff} \frac{\partial \phi_l}{\partial x} - \kappa_{eff} \frac{2aRT}{F} \frac{\partial \ln c_l}{\partial x} + i_l = 0$		(2.9)
c_l	$\frac{\partial c_l}{\partial x} = 0$	$\varepsilon_l \frac{\partial c_l}{\partial t} = \frac{\partial}{\partial x} (D_{l,eff} \frac{\partial c_l}{\partial x}) + (1 - t_+) i_r / F$		(2.16)
c_s	-	-	$\frac{\partial c_s}{\partial t} - \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 D_s \frac{\partial c_s}{\partial r}) = 0$	(2.3)
	-	-	$D_s \frac{\partial c_s}{\partial r} _{r=r_p} = -\frac{i_r}{a_s F}$	(2.4)
i_r	-	$i_r = 0$	$i_r - 2a_s i_0 \sinh\left(\frac{F\eta}{RT}\right) = 0$	(2.5)
	-	-	$\eta = \phi_s - \phi_l - \phi_0(\bar{c}_s, T)$	(2.6)
	-	-	$i_0 = F \sqrt{k_a k_c (c_s^{max} - \bar{c}_s) \bar{c}_s c_l / c_{l,ref}}$	(2.7)
T	-	$\rho c_p \frac{dT}{dt} = q_r + q_{rev} + q_R + q_{ab} + q_{konv}$		(2.17)
	-	$q_r + q_{rev} + q_R = \frac{1}{l} \int_0^l i_r (T \frac{\partial \phi_0}{\partial T} - \phi_0) dx - iU/l$		(2.25)
	-	$q_{ab} = Ri^2 A^2 / V$		(2.20)
	-	$q_{konv} = -hA(T - T_\infty) / V$		(2.21)

TABELLE 2. Modellparameter

Parameter	Anode	Separator	Kathode	Einheiten
d	50	20	45	μm
D_s	0.03	-	0.001	$\mu m^2/s$
r_p	10	-	5	μm
ϵ_s	0.6	-	0.55	1
ϵ_l	0.35	0.5	0.35	1
c_s^{max}	28000	-	50000	mol/m^3
σ	100	-	1	S/m
k	$2 \cdot 10^{-5}$	-	2×10^{-5}	$\mu m/s$
α	0.5	-	0.5	1
$c_{l,ref}$		1		mol/m^3
t_+		0.38		1
ϕ_0	pp.	-	pp.	V
$\phi_{i0,T}$	pp.	-	pp.	mV/K
κ	$c_l \left(\sum_{i,j=0}^2 a_{i,j} c_l^i T^j \right)^2$			S/m
a	$b_0 + \sqrt{c_l}(b_1 + b_2 c_l + b_3 c_l T)$			1
D_l	$\exp(d_0 + d_1/(T + d_2 + d_3 c_l) + d_4 c_l)$			mm^2/s
h		10		W/m^2K
ρ		2		g/cm^3
c_p		1		kJ/kgK
R		253		Ω
A		41.85		cm^2
V		16.54		cm^3

3. NUMERISCHE LÖSUNG

Als gekoppeltes System gewöhnlicher und partieller Differentialgleichungen lässt sich das Newman-Modell nur numerisch lösen. In diesem Abschnitt sollen nun die numerische Lösung des Newman-Modells und ihre theoretischen Grundlagen betrachtet werden. Generell fällt unser Ansatz in die Klasse der Linienmethoden, auf die wir zunächst eingehen werden. Deren Grundidee ist es, alle bis auf eine der Modelldimensionen (meist die Zeitdimension) zu diskretisieren, um so zu einem System gewöhnlicher Differentialgleichungen (engl. *ordinary differential equation*, kurz ODE) bzw. differential-algebraischer Gleichungen (DAE) zu gelangen, dass dann mit einem etablierten ODE/DAE-Integrator gelöst werden kann. Entsprechend gehen wir als nächstes auf die räumliche Diskretisierung des Modells ein. In den folgenden Abschnitten gehen wir näher auf die Theorie von DAE-Systemen und ihre Integration mittels linearen Mehrschrittverfahren ein. Des Weiteren stellen wir ein analytisches Verfahren zur Lösung der Festkörper-Diffusionsgleichung (2.3) vor und untersuchen dessen numerische Eignung im Gesamtsystem. Schließlich gehen wir noch auf einige untergeordnete, aber für die Rechenzeit wichtige Details der Implementierung ein.

3.1. Linienmethode und räumliche Diskretisierung. Die *vertikale* Linienmethode (engl. *method of lines*, MOL) bezeichnet eine Technik zur Lösung partieller Differentialgleichungen wie (3.40), bei der alle bis auf eine der Dimensionen, üblicherweise die Zeitdimension, diskretisiert werden. Ist das so diskretisierte ODE- oder DAE System ein System erster Ordnung (keine höheren Ableitungen), so kann es in vielen Fällen mittels Ein- oder Mehrschrittverfahren numerisch gelöst werden (Wir werden diese Aussage im Abschnitt zur zeitlichen Integration von DAEs präzisieren). In diesem Kontext wird die Linienmethode mindestens seit den frühen sechziger Jahren verwendet [38]. Seitdem sind viele Publikationen zur Theorie und Praxis der vertikalen Linienmethode erschienen [40, 39, 47, 50].

Dem gegenüber steht die auch als Rothe-Methode bekannte *horizontale* Linienmethode, bei der zuerst die zeitliche Dimension, meist mit einem impliziten Verfahren, diskretisiert wird. Anschließend muss in jedem Zeitschritt das resultierende Ortsproblem gelöst werden. Auch diese Methode wurde ausgiebig untersucht [6, 35], soll aber hier nicht weiter betrachtet werden.

3.1.1. Spektralmethoden. Die Wahl des räumlichen Diskretisierungsverfahrens (Finite Differenzen, Elemente, Volumen oder Spektralmethoden) stets uns dabei frei. Wie auch Bizeray [3] haben wir uns für eine

(Pseudo-)Spektral- oder Kollokationsmethode entschieden. Diese zeichnet sich zum einen durch eine einfache Implementierung mittels Differentiationsmatrizen und andererseits durch das Phänomen der „spektralen Genauigkeit“ aus, nachdem für glatte Lösungen exponentielle Konvergenzraten von $O(e^{-cn})$ oder $O(e^{-c\sqrt{n}})$ möglich sind [5, 45, 48] - im Gegensatz zu den algebraischen Konvergenzraten finiter Differenzen oder finiter Elemente (typischerweise $O(n^{-2})$ oder $O(n^{-4})$). Wir wollen die Idee der Spektralmethoden und der spektralen Genauigkeit an einem einfachen Beispiel illustrieren.

Beispiel 1 (Spektrale Genauigkeit). Wir betrachten das Diffusionsproblem mit periodischer Randbedingung

$$\begin{cases} \frac{\partial y}{\partial t} = \frac{\partial^2 y}{\partial x^2}, & 0 < x < 2\pi, \\ y(x, 0) = f(x), & 0 \leq x \leq 2\pi, \\ y(0, t) = y(2\pi, t). \end{cases} \quad (3.1)$$

Mittels Fourier-Transformation findet man die analytische Lösung

$$y(x, t) = \sum_{k=-\infty}^{\infty} \hat{f}_k e^{ikx - k^2 t}, \quad (3.2)$$

mit den Fourier-Koeffizienten

$$\hat{f}_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx. \quad (3.3)$$

Wir bestimmen eine diskrete Lösung $y_n(x, t)$ durch Abschneiden nach Term n und betrachten den Fehler

$$\varepsilon_n = y - y_n = \sum_{|k| > n} \hat{f}_k e^{ikx - k^2 t}. \quad (3.4)$$

Ist nun $f \in C^m$, so gilt wegen $\widehat{f'_k} = ik\hat{f}_k$

$$|\hat{f}_k| = k^{-m} |\widehat{f_k^{(m)}}| \leq C k^{-m} \quad (3.5)$$

und daher $\varepsilon_n = \mathcal{O}(n^{-m})$. Für $f \in C^\infty$ klingt der Fehler daher schneller ab als jede rationale Funktion. \square

Wenn man eine nicht-periodische Funktion periodisch fortsetzt, erhält man Unstetigkeiten, die gerade diese gute Konvergenz zunichte macht. Daher verwendet man für nicht periodische Probleme anstelle von Fourier-Reihen polynomiale Basisfunktionen. Auch hier lässt sich für glatte Funktion spektrale Genauigkeit zeigen bzw. beobachten [5, 27]. Grundidee der Spektralmethoden ist es in jedem Fall, die gesuchte Lösungsfunktion einer Differentialgleichung durch eine Linearkombination von *globalen* Ansatzfunktionen zu approximieren [5, 18, 16].

Anstatt nur die Funktionswerte an benachbarten Stützstellen zur Berechnung einer Ableitung zu verwenden, wie in Finite-Differenzen-Methoden, werden die Funktionswerte an *allen* Stützstellen verwendet.

Man erkaufte sich die verbesserte Konvergenz der Spektralmethoden also mit vollbesetzten Matrizen anstelle der dünnbesetzten Matrizen, die bei Diskretisierungen mit Finiten Differenzen oder Finiten Elementen auftreten. Auch die Implementierung auf unregelmäßigen Gebieten ist bei Spektralmethoden schwieriger.

Die einzelnen Verfahren unterscheiden sich dabei sowohl hinsichtlich der Wahl der Ansatzfunktionen als auch hinsichtlich der Methode zur Bestimmung der Koeffizienten der Linearkombination.

Sei $y = y(x)$ die eindeutige Lösung einer Operatorgleichung $\mathcal{A}y = 0$ in einem Funktionenraum $V \subset \mathcal{L}^2(X)$ mit innerem Produkt $\langle \cdot, \cdot \rangle$ und

$$y_n(x) = \sum_{i=0}^n c_i \phi_i(x) \quad (3.6)$$

eine Linearkombination von Ansatzfunktionen $\phi_i(x) \in V$ mit $\text{supp } \phi_i = X$, die y approximiere. Dann können die Koeffizienten c_i bestimmt werden, indem man fordert, dass die schwache Formulierung

$$\langle \mathcal{A}y, \psi \rangle = 0 \text{ für alle } \psi \in V \quad (3.7)$$

für Testfunktionen $\psi = \psi_j$, $j = 0, \dots, n$ erfüllt sei. Wählt man dabei $\psi_j = \phi_j$, erhält man ein *Galerkin-Verfahren*. Alternativ kann man verlangen, dass die starke Formulierung an den Stützstellen x_j exakt erfüllt sei, dass also

$$\mathcal{A}y(x_j) = 0 \text{ für } j = 0, \dots, n. \quad (3.8)$$

In diesem Fall erhält man ein *Kollokationsverfahren*. Dabei bietet es sich an, die diskrete Lösung $y_n(x)$ nicht durch die modalen Koeffizienten c_i , sondern die nodalen Werte $y_i = y_n(x_i)$ auszudrücken. Gegenüber Galerkin-Spektralmethoden haben die Kollokationsverfahren vor allem den Vorteil, dass sie leichter zu implementieren sind [27].

3.1.2. Tschebyschow-Kollokation. Da die guten Konvergenzeigenschaften der Spektralmethoden nur für glatte Lösungen realisiert werden, haben wir die Modell-Domänen Anode, Separator und Kathode einzeln diskretisiert. Da unser Problem nicht periodisch ist, verwenden wir polynomiale Basisfunktionen. Als fundamentale Unbekannte betrachten wir die nodalen Werten $y_i = y_n(x_i)$, sodass $y_n(x)$ das eindeutig bestimmte Polynom von Grad n mit $y_i = y_n(x_i)$, $i = 0, \dots, n$ ist. Zur Wahl der x_i transformieren wir die drei Domänen zunächst auf $I = [-1, 1]$.

Es zeigt sich nun heraus, dass der naive Ansatz äquidistant verteilter Punkte in I spektakulär scheitern kann. Nicht nur, dass y_n für $n \rightarrow \infty$ nicht gegen y konvergiert, der maximale Fehler kann sogar exponentiell wachsen! Dieses als Runge-Phänomen bekannte Problem ist im oberen

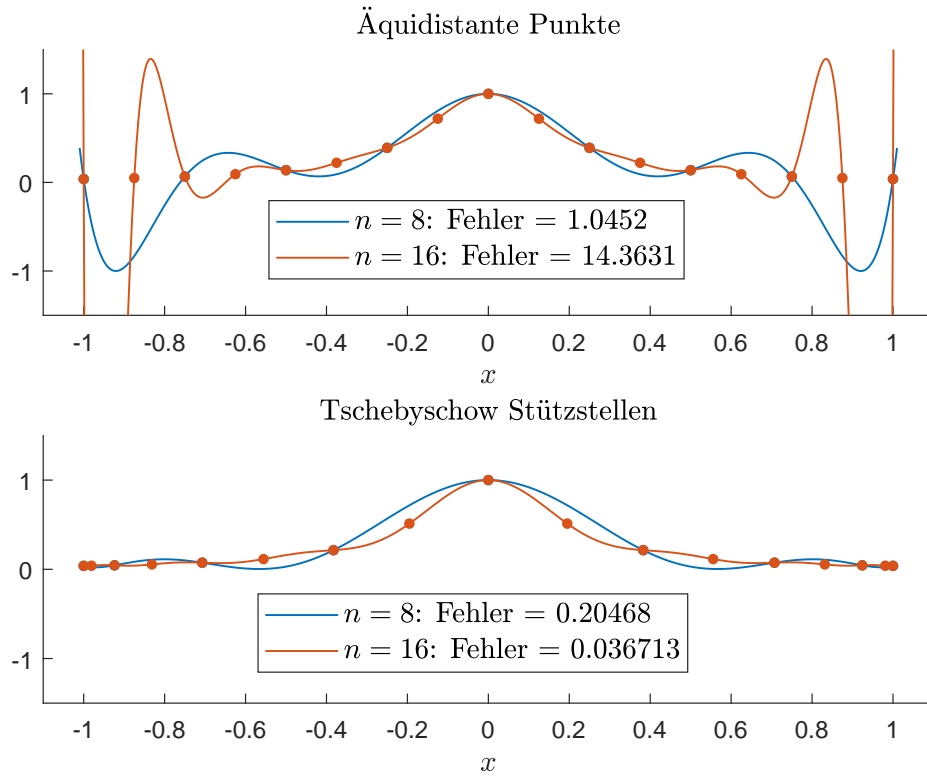


ABBILDUNG 6. Polynominterpolation der Runge-Funktion $y(x) = 1/(1 + 25x^2)$ an äquidistanten und an Tschebyschow-Stützstellen.

Graph der Abbildung 3.1.2 zu sehen. Dargestellt sind die Interpolationspolynome $y_n(x)$ für $n = 8$ und $n = 16$ an die Runge-Funktion $y(x) = 1/(1 + 25x^2)$ mit ihren maximalen Fehlern

$$e_n = \max_{x \in I} |y(x) - y_n(x)|. \quad (3.9)$$

Man erkennt, dass mit wachsendem Polynomgrad n zunehmend Schwingungen an der Enden auftreten, die die Güte der Interpolation zerstören.

Diesen Schwingungen kann man Herr werden, indem man an den Enden die Anzahl der Stützstellen erhöht. Unterschiedliche Punkte können dazu verwendet werden, es zeigt sich aber, dass sie alle die asymptotische Dichte

$$\rho(x) \sim \frac{n}{\pi \sqrt{1 - x^2}}, \quad n \rightarrow \infty. \quad (3.10)$$

haben [45]. Das einfachste Beispiel so verteilter Punkte sind die *Tschebyschow-Punkte* (auch bekannt als Tschebyschow-Lobatto- oder Gauss-Tschebyschow-Lobatto-Punkte)

$$x_j = \cos(\pi j/n), \quad j = 0, \dots, n. \quad (3.11)$$

Anschaulich lassen sie sich als Projektionen von gleichmäßig auf dem Einheitskreis verteilten Punkten auf die x -Achse interpretieren. Im unteren Graphen von Abbildung 3.1.2 sieht man, dass der Interpolationsfehler an die Runge-Funktion mit zunehmendem Polynomgrad nun rasch abklingt. Eine interessante Erklärung dieses Phänomens ist in [45] zu finden. Zu dem Interpolations-Polynom

$$y_n(z) = \prod_{k=1}^n (z - z_k) \quad (3.12)$$

mit Nullstellen z_i (gezählt mit Vielfachheit) betrachtet er das „Potential“

$$\phi_n(z) = \frac{1}{n} \sum_{k=1}^n \log |z - z_k|. \quad (3.13)$$

Nach Konstruktion ist dann

$$|y_n(z)| = e^{n\phi_n(z)}. \quad (3.14)$$

Man erkennt bereits die Idee: Ist $\phi_n(z)$ in etwa konstant auf I , so auch $|y_n(z)|$. Variiert $\phi_n(z)$ dagegen auf I , so führt dies zu Variationen in $|y_n(z)|$, die exponentiell in n wachsen. Analog führt er auch zur asymptotischen Verteilung $\rho(x)$ das Potential

$$\phi(z) = \frac{1}{n} \int_{-1}^1 \rho(x) \log |z - x| dx \quad (3.15)$$

und gelangt damit zu folgendem

Satz 1. Für Interpolationspunkte $(x_j)_{j=0}^n$ in $I = [-1, 1]$, die für $n \rightarrow \infty$ gegen eine Dichte $\rho(x)$ konvergieren mit korrespondierendem Potential $\phi(z)$ gegeben durch (3.15), sei

$$\phi_I = \sup_{x \in I} \phi(x). \quad (3.16)$$

Sei y_n das Polynom von Grad $\leq n$, dass y an den Stellen x_i interpoliert. Existiert eine Konstante $\phi_y > \phi_{[-1,1]}$, sodass y analytisch auf dem abgeschlossenen Gebiet

$$\{z \in \mathbb{C} : \phi(z) \leq \phi_y\} \quad (3.17)$$

ist, so ist für jedes $\nu \geq 0$ und $c = \phi_y - \phi_I$

$$\max_{x \in I} |y(x) - y_n(x)| = \mathcal{O}(e^{-cn}). \quad (3.18)$$

Weiterhin ist für jedes $\nu > 0$

$$\max_{x \in I} |y^{(\nu)}(x) - y_n^{(\nu)}(x)| = \mathcal{O}(e^{-cn}). \quad (3.19)$$

Wir können nun mit (3.15) für die konstante Dichte $\rho(x) = n/2$ und die Tschebyschow-Dichte $\rho(x) = n/(\pi\sqrt{1-x^2})$ die zugehörigen Potentiale berechnen. Mit etwas Arbeit kann man zeigen, dass für $\rho(x) = n/2$

die Äquipotentialkurven in etwa die Gestalt konvexer Linsen haben. Für die Tschebyschow-Dichte erhält man dagegen

$$\phi(z) = \log \frac{|z - \sqrt{z^2 - 1}|}{2}. \quad (3.20)$$

Die Äquipotentialkurven sind Ellipsen mit Foki ± 1 , und der Wert entlang einer solchen Ellipse ist der Logarithmus der halben Summe aus halben Haupt- und Nebenachsen. Insbesondere ist die degenerierte Ellipse $[-1, 1]$ Äquipotentialkurve mit Wert $\phi(z) = -\log 2$. Der springende Punkt ist nun, dass die Rungefunktion die Pole $\pm i/5$ besitzt. Für $\rho(x) = n/2$ umschließen alle Äquipotentialkurven, die I enthalten, auch die Pole. Für die Tschebyschow-Dichte schmiegen sich dagegen die Äquipotentialkurven an I an, sodass sich eine Konstante ϕ_y aus Satz 1 für jede Funktion, die in einer offenen Umgebung von I analytisch ist, finden lässt.

Was passiert nun, wenn man statt $y(x) = 1/(1 + 25x^2)$ $y(x) = 1/(1 + 4x^2)$ wählt? In diesem Fall sind die Pole $\pm i/2$ so gewählt, dass auch die konstanten Dichte Äquipotentialkurven hat, die I nicht aber die Pole umschließt. In diesem Fall konvergieren auch die Interpolationspolynom mit gleichverteilten Stützstellen - und zwar ebenfalls exponentiell in n , aber mit kleinerem c .

Satz 1 sagt aus, dass wir nicht nur $y(x)$ gut durch das Interpolationspolynom $y_n(x)$, sondern auch $y'(x)$ durch $y'_n(x)$ approximieren können. Es bleibt noch zu klären, wie wir praktisch $y'_n(x)$ bzw. die Werte $y'_j = y'_n(x_j)$ aus den Werten $y_j = y_n(x_j)$ berechnen können. Da sowohl die Interpolation als auch die Differentiation lineare Operationen sind, gibt es eine Matrix $D \in \mathbb{R}^{n+1, n+1}$ derart, dass für $\mathbf{y}_n = (y_j)_{j=0}^n$ und $\mathbf{y}'_n = (y'_j)_{j=0}^n$

$$\mathbf{y}'_n = D\mathbf{y}_n \quad (3.21)$$

gilt. Für höheren Ableitungen enthält man dann entsprechend

$$\mathbf{y}_n^{(k)} = D^k \mathbf{y}_n \quad (3.22)$$

Es folgt unmittelbar, dass $D^{n+1} = 0$. Die Matrizen D^k berechnen wir durch Matrixmultiplikation. Zwar gibt es auch Iterationsverfahren, die eine weitere Ableitung in $\mathcal{O}(n^2)$ anstatt $\mathcal{O}(n^3)$ berechnen, aber für unsere Zwecke ($n < 50$, nur erste und zweite Ableitungen, einmaliges Berechnen der Differentiationsmatrizen) reicht Matrixmultiplikation vollkommen aus. Die Koeffizienten dieser Matrix sind in folgendem Satz beschrieben [45].

Satz 2 (Differentiationsmatrix). *Die spektrale Tschebyschow-Differentiationsmatrix $D \in \mathbb{R}^{n+1, n+1}$ hat die Einträge*

$$D_{i,j} = \begin{cases} \frac{c_i(-1)^{i+j}}{c_j(x_i - x_j)} & i \neq j \\ -\frac{x_j}{2(1-x_i^2)} & i = j \neq 1, n+1 \\ \frac{2n^2+1}{6} & i = j = 1 \\ -\frac{2n^2+1}{6} & i = j = n+1 \end{cases} \quad (3.23)$$

Hierbei ist $c_i = 1$ für $i \neq 1, n+1$ und $c_1 = c_{n+1} = 2$.

Aus Gründen der numerischen Stabilität berechnet man in der Praxis die Diagonalterme nicht mit dem Satz 2 [48], sondern durch

$$D_{ii} = - \sum_{j \neq i} D_{i,j}. \quad (3.24)$$

Diese Identität folgt aus der Tatsache, dass konstante Vektoren im Kern von D liegen. Ist nämlich $y_i = y_j$ für alle $0 \leq i, j \leq n$, so ist $y_n(x)$ konstant und daher $y'_n(x) = 0$ und $\mathbf{y}' = 0$. Da wir keine Publikation finden konnten, in denen dieser Satz bewiesen wird (in [45] wird er als Übung aufgegeben), geben wir hier den

Beweis. Wir leiten zunächst die Formeln für beliebige Werte x_i her. Sei dazu

$$p_j(x) = \frac{1}{a_j} \prod_{k \neq j} (x - x_k), \quad a_j = \prod_{k \neq j} (x_j - x_k) \quad (3.25)$$

das eindeutige Polynom von Grad n mit $p_j(x_i) = \delta_{ij}$. Dann ist

$$y_n(x) = \sum_j y_j p_j(x), \quad y'_n(x) = \sum_j y_j p'_j(x) \quad (3.26)$$

und somit $D_{i,j} = p'_j(x_i)$. Nehmen des Logarithmus und Ableiten gibt

$$p'_j(x) = p_j(x) \sum_{k \neq j} (x - x_k)^{-1}. \quad (3.27)$$

Für $i = j$ erhalten wir somit

$$D_{i,i} = p'_i(x_i) = \sum_{k \neq i} (x_i - x_k)^{-1}. \quad (3.28)$$

Für $i \neq j$ schreiben wir die Ableitung in der Form

$$p'_j(x) = \frac{1}{a_j} \sum_{k \neq j} \prod_{l \neq k, j} (x - x_l). \quad (3.29)$$

Nach Einsetzen von x_i bleibt nur ein Term der Summe übrig, sodass

$$p'_j(x_i) = \frac{1}{a_j} \prod_{k \neq i,j} (x_i - x_k) = \frac{a_i}{a_j(x_i - x_j)}. \quad (3.30)$$

Dies sind die allgemeinen Formel, gültig für beliebige voneinander verschiedene x_j . Wir kommen nun zu den speziellen Einträgen für die Tschebyschow-Punkte $x_j = \cos(j\pi/n)$, $j = 0, \dots, n$. Wir beweisen hier nur die Formel für $i \neq j$. Die anderen Werte berechnen sich auf ähnliche Weise. Vergleichen wir (3.30) und (3.23), so bleibt

$$\frac{a_i}{a_j} = \frac{c_i}{c_j} (-1)^{i+j} \quad (3.31)$$

zu zeigen. Aufgrund der Identität

$$\cos a - \cos b = -2 \sin \frac{a+b}{2} \sin \frac{a-b}{2} \quad (3.32)$$

ist

$$a_i = (-2)^n \prod_{k \neq i} \sin \frac{\pi(i+k)}{2n} \sin \frac{\pi(i-k)}{2n}. \quad (3.33)$$

In dem man nun für $k > i$ im zweiten Faktor die Identität

$$\sin a = -\sin(\pi - a) \quad (3.34)$$

und für $i \neq 1, n$ die Identität

$$\sin 2a = 2 \sin a \cos a = 2 \sin a \sin(a + \frac{\pi}{2}) \quad (3.35)$$

verwendet, erhält man

$$a_i = 2^{n-1} (-1)^i c_i \prod_{k=1}^{2n-1} \sin \frac{\pi k}{2n}. \quad (3.36)$$

Es folgt die Behauptung. \square

Die Behandlung der räumlichen Diskretisierung mittels Tschebyschow-Kollokation ist damit zu einem gewissen Abschluss abgebracht. Wir sind dabei nicht auf die natürliche Frage eingegangen, was die Tschebyschow-Punkte mit den Tschebyschow-Polynomen zu tun hat. Die Antwort hängt mit der diskreten Tschebyschow-Transformation zusammen, die für uns zwar keine praktische Bedeutung hat, zum allgemeinen Verständnis aber wichtig ist. Wir skizzieren eine Antwort daher in der folgenden Exkurs und verweisen noch einmal auf die ausführlichere Darstellung in [45].

Exkurs. Die Tschebyschow-Polynomen T_n (erster Art) lassen sich für $z = x + iy = e^{i\theta}$ auf dem Einheitskreis durch die Identität

$$T_n(\operatorname{Re} z) = \operatorname{Re} z^n \quad (3.37)$$

definieren. Dies definiert in der Tat ein Polynom, da

$$\begin{aligned} T_n(x) &= \operatorname{Re}(x + iy)^n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} x^{n-2k} (-1)^k y^{2k} \\ &= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} x^{n-2k} (x^2 - 1)^k. \end{aligned}$$

Sei nun $p(x)$ das Interpolationspolynom einer Funktion $f(x)$ an den Tschebyschow-Punkten $x_j = \cos(\theta_j)$ für $\theta_j = \pi j/n$ und f_j die zugehörigen Werte. Wir wollen den Index j nun von 1 bis $2n$ laufen lassen, sodass die θ_j gleichverteilt in $[0, 2\pi]$ sind, und setzen entsprechend die f_j mit $f_{2n-j} = f_j$ fort. Wir können nun mit der Diskreten Fourier Transformation (DFT) die Fourier-Koeffizienten von f berechnen. Diese sind

$$\hat{f}_k = \frac{\pi}{n} \sum_{j=1}^{2n} e^{-ik\theta_j} f_j, \quad k = -n+1, \dots, n \quad (3.38)$$

Da die f_j reell und $f_j = f_{2n-j}$ ist, gilt, wie man schnell zeigt, $\hat{f}_k = \hat{f}_{-k}$. Mit der inversen DFT ist dann

$$p(x) = p(\cos \theta) = \frac{1}{2\pi} \sum_{k=-n+1}^n e^{ik\theta} \hat{f}_k = \frac{1}{\pi} \sum_{m=0}^n \hat{f}_m T_m(x). \quad (3.39)$$

Wir sehen also: Die nodale Darstellung des Polynoms $p(x)$ über den Vektor $(f_j)_{j=0}^n$ an den Tschebyschow-Punkten entspricht der modalen Darstellung $p(x) = \frac{1}{\pi} \sum_{m=0}^n \hat{f}_m T_m(x)$ mittels der Tschebyschow-Polynome.

Dieser Zusammenhang gibt uns nun ein alternatives Verfahren zur Berechnung der Ableitung. Beginnend mit $(f_j)_{j=0}^n$ berechnen wir zunächst $(\hat{f}_k)_{k=-n+1}^n$. Die Fourier-Koeffizienten der Ableitung sind dann einfach $\hat{f}'_k = ik\hat{f}_k$. Rücktransformation gibt dann bis auf eine Skalierung mit $-1/\sqrt{1-x_j^2}$ die gesuchten Werte f'_j für $j = 1, \dots, n-1$. Für die Randwerte müssen spezielle Formeln verwendet werden.

Da man die DFT mit der schnellen Fourier Transformation in $\mathcal{O}(n \log n)$ Operationen berechnen kann, lässt sich auf diese Weise auch die Ableitung in $\mathcal{O}(n \log n)$ Operationen berechnen. Für große n ist dies günstiger als die $\mathcal{O}(n^2)$ Operationen, die für die Berechnung der Ableitung mittels Matrixmultiplikation erforderlich sind.

Allerdings interessieren wir uns in unserem Anwendungsfall für kleine n ($n < 50$). In diesem Fall ist die Matrixmultiplikation jedenfalls nicht langsamer als die Tschebyschow-Transformation und deutlich einfacher zu implementieren [27]. Wir bemerken noch, dass unsere Beschreibung

der Tschebyschow-Transformation mittels der DFT nicht ganz effizient ist. In der Praxis würde man stattdessen eine diskrete Kosinus Transformation verwendet und damit etwa einen Faktor 4 gewinnen.

3.2. Differential-Algebraische Gleichungen. Wir wenden nun das beschriebene Diskretisierungsverfahren auf das Gleichungssystem des Newman-Modells an. Setzen wir $y = (c_s, c_l, \phi_s, \phi_l, i_r, i_l, T)^\top$, so können wir dieses zusammenfassen als

$$\begin{aligned} \mathcal{M} \frac{\partial y}{\partial t} &= \mathcal{F}(t, y) && \text{in } [0, t_f] \times (0, l) \times B(r_p), \\ \mathcal{G}(y) &= 0 && \text{in } [0, t_f] \times \partial(0, l) \times B(r_p), \\ \mathcal{H}(y) &= 0 && \text{in } [0, t_f] \times (0, l) \times \partial B(r_p). \end{aligned} \quad (3.40)$$

Hierbei ist \mathcal{M} die Diagonalmatrix $\text{diag}(1, 1, 0, 0, 0, 0)$ und \mathcal{F} , \mathcal{G} und \mathcal{H} sind Operatoren, die die x - und r -Ableitungen enthalten und aus Tabelle 1 abgelesen werden können. Nach einer räumlichen Diskretisierung mit k bzw. l Stützstellen in der x - bzw. r -Dimension erhalten wir

$$M \dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}) \quad \text{in } [0, t_f], \quad (3.41)$$

mit

$$\begin{aligned} \mathbf{y} &= (c_s, c_l, \phi_s, \phi_l, i_r, i_l, T)^\top \in \mathbb{R}^n = \mathbb{R}^{(l+5)k+1}, \\ \mathbf{f} : [0, t_f] \times \mathbb{R}^n &\rightarrow \mathbb{R}^n \text{ nicht-linear}, \\ M &= M_{ii} \delta_{ij} \in \mathbb{R}^{n,n}, \quad M_{ii} \in \{0, 1\}. \end{aligned} \quad (3.42)$$

Die Randbedingungen $\mathcal{G}(y) = 0$ und $\mathcal{H}(y) = 0$ wurden dabei bereits in die Funktion \mathbf{f} integriert. Hinzu kommen natürlich jeweils noch Anfangsbedingungen. Die Matrix M nennen wir Massematrix.

Die Darstellung (3.2) enthält nur noch erste Zeitableitungen. Allerdings ist (3.2) keine ODE und kann auch nicht ohne Weiteres in eine ODE transformiert werden, da M nicht invertierbar ist. Es fehlen also zeitliche Ableitungen für einige der Komponenten von \mathbf{y} . Man spricht in diesem Fall von einer *differential-algebraischen Gleichung*, kurz DAE. Es ist zunächst nicht klar, wie sich ODEs und DAEs zueinander verhalten. Können DAEs mit ODE-Methoden gelöst werden oder erfordern sie eine eigenständige (numerische) Behandlung? Zur Beantwortung dieser Frage wollen wir uns zunächst etwas allgemeiner mit DAEs und der instruktiven Spezialklasse der linearen DAEs beschäftigen, bevor wir zu unserem konkreten Anwendungsfall zurückkehren. Unsere Darstellung orientiert sich dabei an [44] und [21]. In der allgemeinsten Form können wir also die implizite Differentialgleichung

$$F(y, y') = 0. \quad (3.43)$$

betrachten. Hierbei seien y und $F(y, y')$ von der selben Dimension n und F glatt. Ein nicht-autonomes System kann in die Form (3.43) gebracht werden, indem wir den Vektor y um die Variable t erweitern und die Gleichung $t' = 1$ hinzufügen.

Ist $\partial F/\partial y'$ invertierbar, können wir (3.43) formal nach y' auflösen und erhalten eine gewöhnliche Differentialgleichung. Wir interessieren uns in diesem Abschnitt für den Fall, in dem $\partial F/\partial y'$ singulär ist. In diesem Fall sprechen wir von einer allgemeinen DAE.

Bemerkung. Fortan benutzen wir auch für Vektoren einfache Lettern y . Wir hatten \mathbf{y} nur verwendet, um die räumliche Diskretisierung $y = y(x, r, t) \rightarrow \mathbf{y}(t)$ deutlich zu machen.

3.2.1. *Lineare DAEs mit konstanten Koeffizienten.* Wir beginnen mit dem einfachsten Fall einer linearen DAE mit konstanten Koeffizienten

$$Ay'(t) + By(t) = f(t), \quad (3.44)$$

mit A singulär. Daraus folgt sofort, dass wir die Anfangsbedingung $y(0) = y^0$ nicht beliebig wählen können, vielmehr muss

$$By^0 - f(0) \in \text{im } A \quad (3.45)$$

gelten. In diesem Fall sagen wir, die Anfangsbedingungen seien *konsistent*. Auch dann ist noch keinesfalls Existenz oder Eindeutigkeit einer Lösung garantiert, wie das Beispiel

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (3.46)$$

bzw. in Komponenten-Schreibweise

$$y_1' = f_1, \quad y_1 = f_2 \quad (3.47)$$

deutlich macht: Es existiert eine Lösung nur dann, wenn $f_2' = f_1$ gilt, und in diesem Fall kann y_2 frei gewählt werden. Wählen wir dagegen

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.48)$$

bzw. in Komponenten-Schreibweise

$$y_1' = f_1, \quad y_2 = f_2, \quad (3.49)$$

so hat das System die eindeutige Lösung

$$y_1 = y_1^0 + \int_0^t f_1(s)ds, \quad y_2 = f_2 \quad (3.50)$$

(vorausgesetzt die Konsistenzbedingung $y_2^0 = f_2(0)$ ist erfüllt). Offenbar liegt der Unterschied zwischen beiden Fällen darin, dass die Matrizen $\lambda A + B$ im ersten Fall singulär und im zweiten (für $\lambda \neq 0$) regulär sind. Das führt zur

Definition 1. Das *Matrixbüschel* $(A, B) := \{\lambda A + B : \lambda \in \mathbb{R}\}$ heißt *regulär*, wenn ein $c \in \mathbb{R}$ existiert, sodass $cA + B$ regulär (d.h. invertierbar) ist.

Diese Bedingung notwendig und hinreichend für die Existenz einer eindeutigen Lösung von (3.44). Notwendigkeit zeigen wir in der folgenden Proposition. Die Gegenrichtung wird etwas mehr Arbeit erfordern.

Proposition 1. *Hat (3.44) eine eindeutige Lösung, so ist (A, B) regulär.*

Beweis. Ist (A, B) singular, so existieren

$$c_1, \dots, c_{n+1} \in \mathbb{R}, v_1, \dots, v_{n+1} \in \mathbb{R}^n : (c_i A + B)v_i = 0, v_i \neq 0. \quad (3.51)$$

Seien $\alpha_1, \dots, \alpha_{n+1} \in \mathbb{R}$ nicht alle null, sodass $\sum \alpha_i v_i = 0$. Dann ist

$$y(t) = \sum \alpha_i v_i e^{c_i t} \neq 0 \quad (3.52)$$

eine Lösung der DAE $Ay' + By = 0$, $y(0) = 0$. Die Lösung lässt sich somit zu jeder Lösung von (3.44) hinzuaddieren, womit diese nicht eindeutig sein kann. \square

Während im ersten Beispiel Existenz bzw. Eindeutigkeit verletzt waren, so waren im zweiten Fall beide nahezu automatisch, weil die differentiellen und die algebraischen Gleichungen entkoppelt waren. Mittels des folgenden Satzes lässt sich dies auch im allgemeinen Fall erreichen.

Satz 3 (Weierstrass-Kronecker Normalform). *Sei (A, B) ein reguläres Matrixbüschel. Dann existieren reguläre Matrizen $P, Q \in \mathbb{C}^{n,n}$, sodass*

$$PAQ = \begin{pmatrix} I_d & 0 \\ 0 & N \end{pmatrix}, PBQ = \begin{pmatrix} R & 0 \\ 0 & I_{n-d} \end{pmatrix} \quad (3.53)$$

gilt, wobei R und N in Jordanscher Normalform sind und N nilpotent ist.

Beweis. Da (A, B) regulär ist, existiert eine reelle Zahl c , sodass $P_0 := cA + B$ regulär ist. Wir setzen

$$A_1 = P_0^{-1}A, \quad B_1 = P_0^{-1}B. \quad (3.54)$$

Dann gilt die Beziehung

$$I - cA_1 = P_0^{-1}(cA + B - cA) = B_1. \quad (3.55)$$

Transformation von A_1 auf Jordan-Normalform gibt

$$A_2 := P_1^{-1}A_1P_1 = \begin{pmatrix} M & 0 \\ 0 & N_1 \end{pmatrix}, \quad (3.56)$$

wobei M die Jordan-Blöcke mit Eigenwerten ungleich 0 und N_1 die Jordan-Blöcke zum Eigenwert null enthält. Mit (3.55) folgt dann

$$B_2 := P_1^{-1}B_1P_1 = I - cP_1^{-1}A_1P_1 = \begin{pmatrix} I - cM & 0 \\ 0 & I - cN_1 \end{pmatrix}. \quad (3.57)$$

Nach Konstruktion ist N_1 nilpotent, d.h. es existiert ein $k \in \mathbb{N}$ mit $N_1^k = 0$, aber $N_1^{k-1} \neq 0$). Somit ist $I - cN_1$ regulär mit inverser Matrix

$$(I - cN_1)^{-1} = I + cN_1 + \dots + c^{k-1}N_1^{k-1}. \quad (3.58)$$

Folglich ist auch

$$P_2 := \begin{pmatrix} M & 0 \\ 0 & I - cN_1 \end{pmatrix} \quad (3.59)$$

regulär. Linksmultiplikation mit P_2^{-1} gibt

$$A_3 := P_2^{-1}A_2 = \begin{pmatrix} I & 0 \\ 0 & (I - cN_1)^{-1}N_1 \end{pmatrix}, \quad B_3 := P_2^{-1}B_2 = \begin{pmatrix} M^{-1} - cI & 0 \\ 0 & I \end{pmatrix}.$$

Damit sind wir beinahe am Ziel; es müssen nur noch der erste Block von A_3 und der letzte Block von B_3 auf Jordansche Normalform gebracht werden:

$$N := V^{-1}(I - cN_1)^{-1}N_1V, \quad R := U^{-1}(M^{-1} - cI)U. \quad (3.60)$$

Aufgrund von (3.58) kommutieren die Matrizen $(I - cN_1)^{-1}$ und N_1 , sodass mit N_1 auch $(I - cN_1)^{-1}N_1$ und also auch N nilpotent mit selbem Nilpotenzindex k ist. Mit

$$P_3 := \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}, \quad (3.61)$$

erhalten wir also

$$P_3^{-1}A_3P_3 = \begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix}, \quad P_3^{-1}B_3P_3 = \begin{pmatrix} R & 0 \\ 0 & I \end{pmatrix}, \quad (3.62)$$

wie gefordert. Für die Matrizen P und Q des Satzes gilt dann

$$P = P_3^{-1}P_2^{-1}P_2^{-1}P_0^{-1}, \quad Q = P_1P_3. \quad (3.63)$$

□

Definition 2. Das Matrixbüschel $P(A, B)Q$ mit P und Q aus Satz 3 heißt *Weierstrass-Kronecker Normalform* des regulären Matrixbüschels (A, B) . Der Nilpotenzindex k von N heißt *Index* von (A, B) .

Die Definition setzt voraus, dass $P(A, B)Q$ eindeutig durch (A, B) bestimmt ist. Dies ist bis auf Ähnlichkeit tatsächlich der Fall, wie wir im folgenden Lemma zeigen.

Lemma 1. *Die Weierstrass-Kronecker Normalform des regulären Matrixbüschels (A, B) ist bis auf Ähnlichkeit eindeutig bestimmt. Für zwei Transformationen*

$$(A_1, B_1) = P_1(A, B)Q_1 \text{ und } (A_2, B_2) = P_2(A, B)Q_2$$

von (A, B) auf Weierstrass-Kronecker Normalform existiert also ein reguläres P mit $(A_1, B_1) = P^{-1}(A_2, B_2)P$. Die Matrizen R_1, R_2 bzw. N_1, N_2 sind somit eindeutig bestimmt bis auf Vertauschen der Jordan-Blöcke. Insbesondere ist also der Index k von (A, B) als Größe des größten Jordanblocks zum Eigenwert 0 wohldefiniert.

Beweis. Nach Voraussetzung ist $(A_1, B_1) = P_1 P_2^{-1} (A_2, B_2) Q_2^{-1} Q_1$. Wir setzen daher $P = P_2 P_1^{-1}$ und $Q = Q_2^{-1} Q_1$, sodass $P = Q$ zu zeigen ist. Seien

$$A_1 = \begin{pmatrix} I_{d_1} & 0 \\ 0 & N_1 \end{pmatrix}, B_1 = \begin{pmatrix} R_1 & 0 \\ 0 & I_{n-d_1} \end{pmatrix} \quad (3.64)$$

und

$$A_2 = \begin{pmatrix} I_{d_2} & 0 \\ 0 & N_2 \end{pmatrix}, B_2 = \begin{pmatrix} R_2 & 0 \\ 0 & I_{n-d_2} \end{pmatrix}. \quad (3.65)$$

Wir betrachten das Polynom

$$\begin{aligned} p(\lambda) &= \det(\lambda A + B) = \det(P_1^{-1}) \det(\lambda A_1 + B_1) \det(Q_1) \\ &= \det(P_1^{-1}) \det(\lambda I_{d_1} + R_1) \det(Q_1) \\ &= \det(P_2^{-1}) \det(\lambda I_{d_2} + R_2) \det(Q_2). \end{aligned}$$

Daher ist $d_1 = d_2$, wir können die Indizes d_1, d_2 daher weglassen. Wir schreiben die Gleichungen $PA_1 = A_2Q$ und $PB_1 = B_2Q$ aus:

$$\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & N_1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & N_2 \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \quad (3.66)$$

$$\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} R_1 & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} R_2 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}. \quad (3.67)$$

Es ist $P_{11} = Q_{11}$ und $P_{22} = Q_{22}$. Weiterhin gilt $P_{12} = R_2 Q_{12} = R_2 P_{12} N_2 = \dots = R_2^k P_{12} N_2^k = 0$ für hinreichend großes k . Mit einem ähnlichen Argument kann man zeigen, dass alle Gegendiagonalblöcke null sind. Damit folgt $P = Q$ und die Behauptung. \square

Satz 4. *Ist (A, B) regulär, so hat (3.44) eine eindeutige Lösung.*

Beweis. Seien P und Q wie in Satz 3. Wir multiplizieren die DAE (3.44) von links mit P und erhalten

$$PAy' + PBy = Pf \quad (3.68)$$

Indem wir

$$y = Q \begin{pmatrix} u \\ v \end{pmatrix}, \quad Pf = \begin{pmatrix} g \\ h \end{pmatrix}, \quad y, g \in \mathbb{R}^d \quad (3.69)$$

überführen wir die DAE (3.44) in die Form

$$u' + Ru = g(t) \quad (3.70)$$

$$Nv' + v = h(t). \quad (3.71)$$

(3.70) stellt eine gewöhnliche lineare Differentialgleichung dar und besitzt folglich eine eindeutige Lösung für beliebige Anfangsbedingung $u^0 \in \mathbb{R}^d$. Die DAE (3.70) können wir mittels $k - 1$ -fachen Differenzieren nach v auflösen, wobei k der Index von (A, B) ist:

$$v = h - Nv' = h - N(h' - Nv'') = \dots = \sum_{l=0}^{k-1} (-1)^l N^l h^{(l)}. \quad (3.72)$$

Diese entspricht formal der „Neumannschen Reihe“

$$Nv' + v = (N\frac{d}{dt} + I)v = h \implies v = \left(\sum_{l=0}^{k-1} (-1)^l N^l \frac{d^l}{dt^l} \right) h. \quad (3.73)$$

$v(t)$ ist somit ebenfalls eindeutig bestimmt, sofern die Konsistenzbedingung

$$v^0 = \sum_{l=0}^{k-1} (-1)^l N^l h^{(l)}(0) \quad (3.74)$$

erfüllt ist. (Die allgemeine Konsistenzbedingung (3.45) transformiert sich zur Bedingung $v^0 - h(0) \in \text{im } N$. Aufgrund der Differenziationen kommen jetzt aber noch Bedingungen an die Ableitungen hinzu.) \square

Wir sehen, dass der Index k misst, wie weit die DAE von einer ODE ist. k Differenziationen sind nötig, um die DAE (3.44) in eine ODE zu transformieren. Mit dieser Einsicht kehren wir nun zur allgemeinen (nicht-linearen) Theorie zurück.

3.2.2. Allgemeine DAEs. Wir beginnen mit der Verallgemeinerung des wichtigen Indexbegriffs für beliebige DAEs. Dazu geben wir

Definition 3 (Differentiationsindex). Sei die allgemeine differential-algebraische Gleichung $F(t, y, \dot{y}) = 0$ lokal eindeutig lösbar und die Funktion $F : [0, t_f] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ hinreichend oft stetig differenzierbar. Die kleinste natürliche Zahl m für die man aus den Gleichungen

$$F(t, y, \dot{y}) = 0, \quad \frac{dF(t, y, \dot{y})}{dt} = 0, \quad \dots, \quad \frac{d^m F(t, y, \dot{y})}{dt^m} = 0 \quad (3.75)$$

$\dot{y}(t)$ eindeutig bestimmen kann, heißt *Differentiationsindex* oder einfach nur *Index* der differential-algebraischen Gleichung.

Bemerkung 1. Der Beweis von Satz 4 zeigt, dass der Differentiationsindex der linearen DAE

$$Ay'(t) + By(t) = f(t) \quad (3.76)$$

gerade der Index des Matrixbüschels (A, B) ist.

Bemerkung 2. Wir betrachten die DAE

$$My'(t) = f(t, y) \quad (3.77)$$

mit M singulär (Von dieser Form war unsere DAE). O.B.d.A. können wir davon ausgehen, dass M diagonal ist, da wir andernfalls die DAE durch Singulärwertzerlegung von M in diese Form bringen. Fassen wir dann die Komponenten y_i mit $M_{ii} \neq 0$ zu einem differentiellen Anteil u und die restlichen Komponenten zu einem algebraischen Anteil v

zusammen und teilen entsprechend die Komponenten von f in g und h auf, zerfällt (3.2) in

$$u' = g(t, u, v) \quad (3.78)$$

$$0 = h(t, u, v). \quad (3.79)$$

Ableiten der algebraischen Gleichungen (3.79) gibt

$$0 = \frac{\partial h}{\partial t} + \frac{\partial h}{\partial u}g + \frac{\partial h}{\partial v}v', \quad (3.80)$$

sodass die DAE Index 1 hat, genau dann, wenn die Jacobi-Matrix der algebraischen Gleichungen

$$\partial h / \partial v \text{ nichtsingulär} \quad (3.81)$$

in einer Umgebung der Lösung ist. Wir bezeichnen dies als Index-1-Bedingung. Die Anfangsbedingungen sind konsistent, falls $0 = h(t, u_0, v_0)$.

Bemerkung 3. Ist die Index 1 Bedingung nicht erfüllt, können wir versuchen die DAE in sogenannte Hessenbergform zu bringen. Wir betrachten dazu die DAE

$$u' = g(t, u, v) \quad (3.82)$$

$$0 = h(t, u). \quad (3.83)$$

Ableiten der Zwangsbedingung liefert die „versteckte“ Zwangsbedingung

$$0 = h_t + h_u u' = h_t + h_u g \quad (3.84)$$

Nochmaliges Ableiten liefert

$$0 = h_{tt} + 2h_{tu}ug + h_{uu}(g, g) + h_u g_u g + h_u g_v v' \quad (3.85)$$

Ist die Index-2-Bedingung

$$h_u(t, u)g_v(t, u, v) \text{ nichtsingulär} \quad (3.86)$$

erfüllt, so hat die DAE Index 2. Die Anfangsbedingungen sind konsistent, sofern die Zwangsbedingung *und* die versteckte Zwangsbedingung erfüllt sind.

Ist auch diese Bedingung nicht erfüllt, so kann man versuchen die DAE in eine höhere Hessenbergform zu bringen. Die Index 3 Form lautet

$$u' = f(t, u, v, w) \quad (3.87)$$

$$v' = g(t, u, v) \quad (3.88)$$

$$0 = h(t, u). \quad (3.89)$$

Analog kann man zeigen, dass diese DAE Index 3 hat, sofern die Index-3-Bedingung

$$h_u(t, u)g_v(t, u, v)f_w(t, u, v, w) \text{ nichtsingulär} \quad (3.90)$$

erfüllt ist. Es treten in dem Fall zwei versteckte Zwangsbedingungen auf. Die Verallgemeinerung für höhere Indexe als 3 ist nun offensichtlich. Für allgemeine DAEs kann der Index schwierig zu bestimmen sein. Wir wollen nun noch ein einfaches Anwendungsbeispiel besprechen, bevor wir auf die Konsequenzen der Theorie für unseren Anwendungsfall eingehen.

Beispiel 2. Wir betrachten dazu das idealisierte Pendel in Abbildung 7, bestehend aus einem leichten, inelastischen Stab der Länge l und einer Masse m im Punkt (x, y) . Die Bewegungsgleichungen ergeben sich nach dem zweiten newtonschen Gesetz zu

$$\begin{aligned}\ddot{x} &= -\frac{1}{m}T \sin \theta = -\frac{Tl}{m}x =: -\lambda x \\ \ddot{y} &= \frac{1}{m}(mg - T \cos \theta) = g - \lambda y\end{aligned}\tag{3.91}$$

Gemeinsam mit der geometrischen Zwangsbedingung

$$x^2 + y^2 = l^2\tag{3.92}$$

erhalten wir ein DAE-System in den Variablen x , y und λ . (Wie gewohnt, kann man $u = \dot{x}$, $v = \dot{y}$ setzen, um entsprechend Definition 3 ein DAE-System in x , y , u , v und λ zu erhalten, das nur erste Ableitungen enthält.) Zweimaliges Ableiten der Zwangsbedingung liefert unter Verwendung der Bewegungsgleichungen (3.91)

$$2x\dot{x} + 2y\dot{y} = 0\tag{3.93}$$

$$2(\dot{x}^2 + \dot{y}^2) + 2gy - 2\lambda l^2 = 0.\tag{3.94}$$

Die dritte Ableitung enthält somit $\dot{\lambda}$. Da offensichtlich das Ableiten der Bewegungsgleichungen nicht weiterführt (mit den Ableitungen von λ entstehen auch höhere als zweite Ableitungen von x und y), hat das DAE-System (3.91) -(3.92) also Differentiationsindex 3.

Bekanntlich lässt sich das Pendel über den Drehimpulssatz

$$I\dot{\omega} = ml^2\ddot{\theta} = -l \cdot mg \sin \theta\tag{3.95}$$

auch durch eine gewöhnliche Differentialgleichung in θ beschreiben, aber für kompliziertere Systeme ist solch eine Darstellung weniger offensichtlich, wenn nicht sogar unmöglich.

3.2.3. Konsequenzen für den Anwendungsfall. Unsere DAE hat die Gestalt

$$My'(t) = f(t, y)\tag{3.96}$$

für M diagonal und singular mit Einträgen 1 und 0 auf der Diagonale. Nach Trennung der differentiellen und algebraischen Variablen schreiben wir dies in der Form

$$y_d' = f_d(t, y_d, y_a)\tag{3.97}$$

$$0 = f_a(t, y_d, y_a).\tag{3.98}$$

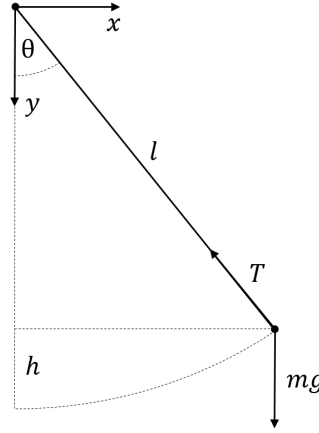


ABBILDUNG 7. Mathematisches Pendel

Nach Bemerkung 3.3.3 hat diese Index 1 genau dann, wenn die Index-1-Bedingung

$$\partial f_a / \partial y_a \text{ nichtsingulär} \quad (3.99)$$

erfüllt ist. Die Jacobi-Matrix $\partial f / \partial y$ wird, wie in Abschnitt 3.5 besprochen, analytisch berechnet. Somit kann die Index-1-Bedingung für beliebige Diskretisierung und jeden Wert y überprüft werden. Es zeigt sich, dass diese für physikalisch sinnvolle y und typische Diskretisierungen $k, l \leq 100$ im Allgemeinen erfüllt sind. Aufgrund der komplizierten, nicht-linearen Struktur der Jacobi-Matrix ist es allerdings schwierig, dies zu beweisen.

Wie in Abschnitt 3.3.3 beschrieben, überprüft *ode15s* die Index-1-Bedingung bei der Berechnung konsistenter Startwerte (d.h. Werte $y_0 = [y_{0,d}, y_{0,a}]$ mit $f_a(0, y_{0,d}, y_{0,a}) = 0$). Die Nutzer vorgegeben Startwerte \tilde{y}_0 werden dabei als Schätzwerte behandelt. Dies ist zwar nur erforderlich, falls die vom Nutzer vorgegeben Startwerte nicht konsistent sind. Allerdings ist es in der Praxis schwierig, konsistente Startwerte außerhalb des Gleichgewichts vorzugeben, da dafür gerade das nichtlineare Gleichungssystem $f_a(0, y_d, y_a) = 0$ zu lösen ist. Sollte in dieser Situation die Index-1-Bedingung verletzt sein, bricht das Programm mit der Fehlermeldung „This DAE appears to be of index greater than 1“ ab. Diese (theoretische) Limitation des Programms ist vernünftig, da nach dem Satz von der impliziten Funktion die Index-1-Bedingung gerade aussagt, dass $y_{0,a}$ in der Nähe \tilde{y}_0 nicht eindeutig bestimmt ist.

Wir wollen aber betonen, dass es sich nach unserer Erfahrung hierbei nur um eine theoretische und keine praktische Limitation handelt. Normalerweise weist die Fehlermeldung entweder auf eine vergessene Randbedingung oder einen Fehler in der Berechnung der Jacobi-Matrix hin!

3.3. Zeitliche Integration der DAE. Wir kommen nun zur numerischen Integration des nach der räumlichen Diskretisierung entstandenen DAE-Systems. In MATLAB stehen für die Lösung von Index 1 DAE-Problemen die Integratoren *ode15s* und *ode23t* zur Verfügung. *ode15s* ist ein implizites lineares Mehrschritt-Verfahren mit Ordnungssteuerung ($1 \leq p \leq 5$). *ode23t* dagegen basiert auf dem impliziten Trapezverfahren und ist vor allem für geringe Genauigkeit und Probleme mittlerer Steifigkeit geeignet [41, 42]. Da bei niedrigen Genauigkeitsanforderungen der Einsatz eines Ersatzschaltbild-Modells anstelle des rechenaufwendigen Newman-Modell zweckmäßiger sein kann, sollte der Integrator auch für hohe Genauigkeitsanforderungen effizient sein. Darüber hinaus haben auch numerische Tests ergeben, dass *ode15s* in dem hier betrachteten Problem mit einer typischen räumlichen Diskretisierung mit 30 Stützstellen in der x - und 25 Stützstellen in der r -Dimension (auf dem verwendeten Rechner) schon bei kleinen Genauigkeitsanforderungen weniger Rechenzeit erfordert. Aus diesem Grund haben wir uns für die Verwendung von *ode15s* entschieden.

In diesem Abschnitt stellen wir erst die wichtigsten Definitionen und Resultate zur Konvergenz und Stabilität linearer Mehrschrittverfahren zusammen, um diese dann zur Charakterisierung der *ode15s* zugrunde liegenden sogenannten BDF- bzw. NDF-Verfahren zu benutzen. Dabei können wir der Einfachheit halber zunächst ein ODE-System betrachten. Im Abschluss beschreiben wir dann, wie mit den betrachteten Verfahren auch Index 1 DAEs gelöst werden können.

3.3.1. Lineare Mehrschrittverfahren. Wir beginnen mit dem allgemeinen Anfangswert-Problem

$$\begin{aligned} y' &= f(t, y), \quad t \in [0, t_f], \\ y(0) &= y_0. \end{aligned} \tag{3.100}$$

Dabei erfülle $f : [0, t_f] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine globale Lipschitzbedingung bezüglich der zweiten Variable, sodass (3.100) nach dem Satz von Picard-Lindelöf eine eindeutige Lösung $y : [0, t_f] \rightarrow \mathbb{R}^n$ besitzt.

Der Grundgedanke der Mehrschrittverfahren ist es, mehrere bereits berechnete Näherungswerte y_m, \dots, y_{m+k-1} zur Berechnung einer neuen Näherung y_{m+k} heranzuziehen. Bei *linearen* Mehrschrittverfahren der Schrittlänge k (LMV $_k$) geht man von der allgemeinen Verfahrensvorschrift

$$\sum_{l=0}^k \alpha_l y_{m+l} = h \sum_{l=0}^k \beta_l f(t_{m+l}, y_{m+l}) \tag{3.101}$$

mit festen reellen Koeffizienten $\alpha_l, \beta_l, l = 0, \dots, k$ aus. Wir setzen dabei voraus, dass $\alpha_k \neq 0, \alpha_0 \beta_0 \neq 0$, sodass tatsächlich die letzten k Werte in die Berechnung eines neuen Wertes eingehen. Ist $\beta_k = 0$ sprechen wir von einem *expliziten*, anderenfalls von einem *impliziten* Verfahren. Die Bedingung $\alpha_k \neq 0$ garantiert nach dem Satz von der impliziten

Funktion, dass (3.101) auch im impliziten Fall für hinreichend kleine h eindeutig lösbar ist.

Wir möchten zunächst einige grundlegende Definitionen und Sätze zu dieser Verfahrensklasse zusammenstellen. Unsere Darstellung folgt dabei den Büchern von Hairer und Wanner [22, 21] und Strehmel, Weiner und Podhaisky [44]. Um die Notation einfach zu halten, gehen wir von einer Konstanten Schrittweite h aus.

Definition 4 (Konsistenz). Das LMV (3.101) heißt *konsistent von Ordnung p* bzw. hat *Konsistenzordnung p* , falls für den linearen Differenzenoperator

$$L[y(t), h] = \sum_{l=0}^k (\alpha_l y(t + lh) - \beta_l y'(t + lh))$$

für alle $y(t) \in C^{p+1}[0, t_f]$

$$L[y(t), h] = \mathcal{O}(h^{p+1}) \quad \text{für } h \rightarrow 0.$$

Proposition 2. Ein LMV besitzt die Konsistenzordnung p genau dann, wenn die Konsistenzbedingungen

$$\sum_{l=0}^k \alpha_l = 0, \quad \sum_{l=0}^k \alpha_l l^q = \sum_{l=0}^k q \beta_l l^{q-1}, \quad q = 1, \dots, p. \quad (3.102)$$

gelten.

Beweis. Für $y \in C^{p+1}[0, t_f]$ entwickeln wir y und y' als Taylorreihen um den Punkt t und setzen in den linearen Differenzenoperator ein:

$$\begin{aligned} L[y(t), h] &= \sum_{l=0}^k \left(\sum_{q=0}^p \alpha_l \frac{(lh)^q}{q!} y^{(q)}(t) - \sum_{q=1}^p h \beta_l \frac{(lh)^{q-1}}{(q-1)!} y^{(q)}(t) \right) + \mathcal{O}(h^{p+1}) \\ &= y(t) \sum_{l=0}^k \alpha_l + \sum_{q=1}^p \frac{h^q y^{(q)}(t)}{q!} \sum_{l=0}^k (\alpha_l l^q - q \beta_l l^{q-1}) + \mathcal{O}(h^{p+1}). \end{aligned}$$

$L[y(t), h] = \mathcal{O}(h^{p+1})$ für beliebige $y \in C^{p+1}$ somit genau dann, wenn die Konsistenzbedingungen erfüllt sind. \square

Definition 5 (Fehlerkonstante). Sei $y \in C^{p+2}[0, t_f]$. Für ein LMV der Konsistenzordnung p gilt dann

$$L[y(t), h] = C_{p+1} h^{p+1} y^{(p+1)}(t) + \mathcal{O}(h^{p+2}). \quad (3.103)$$

Die Konstante $C_{p+1}^* = C_{p+1} / \sum_{l=0}^k \beta_l$ heißt *Fehlerkonstante* des Verfahrens.

Die Skalierung mit $\sum_{l=0}^k \beta_l$ garantiert, dass die Fehlerkonstante invariant unter der Skalierung $\alpha'_l = c \alpha_l$, $\beta'_l = c \beta_l$ ist.

Im Gegensatz zu Einschrittverfahren muss ein Mehrschrittverfahren hoher Konsistenzordnung und kleiner Fehlerkonstante noch nicht zu

brauchbaren Lösungen führen. Grund dafür ist die Instabilität der numerischen Lösung, wie wir zuerst an einem Beispiel deutlich machen wollen.

Beispiel 3. Wir betrachten das implizite Verfahren

$$y_{m+2} + 4y_{m+1} - 5y_m = h(f_{m+2} + 2y_{m+1} + 3f_m) \quad (3.104)$$

der Konsistenzordnung 2, wie wir mit Satz 2 leicht überprüfen können. Wenden wir es nun auf die Testgleichung

$$y'(t) = 0, \quad y(0) = 1 \quad (3.105)$$

an, so erhalten wir die Differenzengleichung

$$y_{m+2} + 4y_{m+1} - 5y_m = 0. \quad (3.106)$$

mit allgemeiner Lösung

$$y_m = a + b(-5)^m. \quad (3.107)$$

Einsetzen der exakten Anfangsbedingung $y_0 = y_1 = 1$ gibt $a = 1, b = 0$. Wir erhalten also die erwartete Lösung $y_m = 1$. Aber diese Lösung ist instabil: Leicht gestörte Anfangsdaten $y_0 = 1, y_1 = 1 + \varepsilon$ führen zu der unbeschränkten Lösung $y_m = 1 + \frac{1}{6}\varepsilon[1 - \frac{1}{6}(-5)^m]$. \square

Zur Definition eines Stabilitätsbegriffs wollen wir diese Überlegungen nun für ein allgemeines LMV k und die allgemeinere Klasse von Testgleichungen

$$y' = \lambda y, \quad y(0) = y_0 \quad (3.108)$$

fortsetzen. Durch Einsetzen in (3.101) erhalten wir die Differenzengleichung

$$\sum_{l=0}^k \alpha_l y_{m+l} = z \sum_{l=0}^k y_{m+l}, \quad z = h\lambda. \quad (3.109)$$

Da die Lösungen der Testgleichung (3.108) für beliebige Anfangswerte beschränkt bleiben, solange $\operatorname{Re}(\lambda) \leq 0$, könnte man verlangen, dass Gleiches für Lösungen der Differenzengleichung (3.109) für beliebige Startwerte y_0, \dots, y_{k-1} gelten soll. Dies führt zum Begriff der A-Stabilität. Wie wir noch sehen werden, ist dies aber bereits eine starke Forderung. Wir schwächen den Begriff daher gleich wieder etwas ab und geben

Definition 6 ($A(\alpha)$ -Stabilität). Das lineare Mehrschrittverfahren (3.101) heißt $A(\alpha)$ -stabil mit $\alpha \in [0, \frac{\pi}{2})$, falls die Menge S aller $z \in \mathbb{C}$, für die alle Lösungen der Differenzengleichung (3.109) beschränkt bleiben, *Stabilitätsgebiet* genannt, die Menge

$$S(\alpha) = \{z \in \mathbb{C}^- : |\arg(z) - \pi| \leq \alpha\}$$

enthält. Kann $\alpha = 0$ gewählt werden, so ist das Verfahren A -stabil. Ist $0 \in S$, so heißt das Verfahren *nullstabil*.

Lemma 2. Definieren wir das charakterische Polynom zu (3.101) durch

$$\phi_z(\lambda) = (\alpha_k - z\beta_k)\lambda^k + (\alpha_{k-1} - z\beta_{k-1})\lambda^{k-1} + \dots + (\alpha_0 - z\beta_0),$$

so lässt sich das Stabilitätsgebiet S charakterisieren als Menge aller $z \in \mathbb{C}$, sodass $\phi_z(x)$ hat keine Nullstellen außerhalb des Einheitskreises und nur einfache Nullstellen auf dessen Rand besitzt.

Beweis. Dies folgt unmittelbar aus der Theorie linearer Differenzengleichungen: Sind $\lambda_1, \dots, \lambda_n$ die Nullstellen von $\phi_z(\lambda)$ der Vielfachheit m_1, \dots, m_n , so ist die allgemeine Lösung von (3.109) gegeben durch

$$y_m = \sum_{i=1}^n p_i(m) \lambda_i^{m_i} \quad (3.110)$$

für Polynome $p_i(m)$ von Grad $m_i - 1$. y_m ist also für beliebige Anfangswerte genau dann beschränkt, wenn $|\lambda_i| \leq 1$ und $m_i = 1$, falls $|\lambda_i| = 1$. \square

Daraus folgt, dass ein explizites Verfahren kein unbeschränktes Stabilitätsgebiet und damit nicht $A(\alpha)$ -stabil sein kann. Denn für explizite Verfahren ist $\beta_k = 0$ und mindestens ein $\beta_l \neq 0$, $l = 0, \dots, k-1$. Für $|z| \rightarrow \infty$ strebt dann der Koeffizient von x^l dem Betrag nach gegen unendlich und nach dem Satz von Vieta mit ihm auch mindestens eine der Nullstellen von ϕ_z . Für große λ muss also eine kleine Schrittweite h gewählt werden, sodass $z = \lambda h$ im Stabilitätsgebiet liegt.

Wir können nun zur Frage der Konvergenz zurückkehren. Sei $y_h(mh) = y_m$ die auf dem Gitter $h\mathbb{N} \cap [0, t_f]$ definierte und mit dem k -Schritt Verfahren (3.101) berechnete numerische Lösung des Anfangswertproblems (3.100), wobei die Werte $y_h(mh)$ für $m < k$ mit Verfahren niedriger Schrittlänge berechnet wurden.

Definition 7 (Konvergenz). Das LMV k heißt *konvergent von Ordnung p* bzw. hat *Konvergenzordnung p* , falls für alle Anfangswertprobleme (3.100) mit

$$f(t, y) \in C^p([0, t_f] \times \mathbb{R}^n)$$

und für alle Startwerte mit

$$\|y_0(lh) - y_h(lh)\| = \mathcal{O}(h^p), \quad \text{für } h \rightarrow 0, \quad l = 0, \dots, k,$$

folgende Bedingung erfüllt ist:

$$\|y(t^*) - y_h(t^*)\| = \mathcal{O}(h^p), \quad \text{für } h \rightarrow 0, \quad t^* = mh \in [0, t_f] \text{ fest.}$$

Mit dieser Definition können wir folgenden Satz zur Konvergenz linearer Mehrschrittverfahren aufstellen.

Satz 5. Ist ein LMV k konsistent von Ordnung p und nullstabil, so ist es auch konvergent von Ordnung p . Umgekehrt, ist es konvergent von Ordnung p , so ist es konsistent und nullstabil.

Beweis. Siehe [21]. \square

Wir hätten gerne Verfahren mit hoher Konsistenzordnung und guten Stabilitätseigenschaften. Da ein LMV k $2k$ unbestimmte Koeffizienten besitzt, lässt sich vermuten, dass diese so gewählt werden können, dass die Konsistenzbedingungen für $q = 1$ bis $q = 2k$ erfüllt sind, das Verfahren also Ordnung $2k$ besitzt. Dies ist in der Tat möglich, allerdings sind diese Verfahren nicht nullstabil. Welche Konsistenzordnung noch möglich ist für nullstabile bzw. A-stabiles LMV k ist Inhalt der beiden bekannten Dahlquist-Schranken.

Satz 6 (Erste und zweite Dahlquist-Schranke). *Ein nullstabiles k -Schrittverfahren hat die maximale Konsistenzordnung*

$$p_{\max} = \begin{cases} k+2, & k \equiv 0 \pmod{2} \\ k+1, & k \equiv 1 \pmod{2} \\ k, & \beta_k/\alpha_k \leq 0 \end{cases} \quad (3.111)$$

Ein A-stabiles LMV k hat die maximale Konsistenzordnung $p_{\max} = 2$.

Beweis. Siehe [21]. □

3.3.2. BDF-Verfahren. Nach dieser allgemeinen Betrachtung kommen wir jetzt zu einer bestimmten Klasse von Mehrschritt-Verfahren, den BDF-Verfahren (von engl. „backward differentiation formula“). Wie die zweite Dahlquist-Barriere zeigt, müssen wir uns für Verfahren der Ordnung $p > 2$ mit A(α)-stabilen Verfahren für möglichst große Öffnungswinkel α zufrieden geben. Nach Lemma 2 erfordert dies, dass die Nullstellen des charakteristischen Polynoms für $|z| \rightarrow \infty$ beschränkt bleiben. Es macht aber Sinn, darüber hinaus zu fordern, dass die Nullstellen für $|z| \rightarrow \infty$ gegen 0 gehen. Dies entspricht der Forderung, dass numerische Lösungen der Testgleichung $y' = \lambda y$ nicht nur beschränkt bleiben, sondern für hinreichend große λ und nicht zu kleine h gegen null gehen, und so zwar je größer λ , desto schneller.

Nach dem Satz von Vieta müssen dann auch die Koeffizienten

$$(\alpha_l - z\beta_l)/(\alpha_k - z\beta_k)$$

von $\phi(x)/(\alpha_k - z\beta_k)$ für $l < k$ nach 0 gehen. Dies erfordert offensichtlich $\beta_l = 0$ für $l < k$. Setzen wir $\beta_k = 1$, hat unser Verfahren also die Gestalt

$$\alpha_k y_{m+k} + \dots + \alpha_0 y_m = hf(t_{m+k}, y_{m+k}). \quad (3.112)$$

Die α_l bestimmen wir wie folgt: Wir legen zunächst ein äquidistantes Gitter $h\mathbb{N} \cap [0, t_f]$ zugrunde und nehmen an, dass bereits Näherungswerte $y_m, y_{m+1}, \dots, y_{m+k-1}$ an den Stützstellen $t_m, t_{m+1}, \dots, t_{m+k-1}$ berechnet worden sind. Wir legen dann gedanklich ein Interpolationspolynom $p(t)$ durch die Stützpunkte (t_{m+l}, y_{m+l}) für $l < k$ und den noch unbekannten Punkt (t_{m+k}, y_{m+k}) und bestimmen y_{m+k} durch die Forderung

$$p'(t_{m+k}) = f(t_{m+k}, y_{m+k}). \quad (3.113)$$

Das Interpolationspolynom p können wir in der Newton-Form

$$p(t) = p(t_{m+k} - sh) = \sum_{l=0}^k (-1)^l \binom{s}{l} \nabla^l y_{m+k} \quad (3.114)$$

schreiben, wobei die *Rückwärtsdifferenzen* rekursiv definiert sind durch

$$\nabla^0 u_m = u_m, \quad \nabla^l u_m = \nabla^{l-1} u_m - \nabla^{l-1} u_{m-1}. \quad (3.115)$$

Für Leser, die mit dieser Darstellung des Interpolationspolynoms nicht vertraut sind, weisen wir nach, dass p tatsächlich die Werte (t_{m+l}, y_{m+l}) interpoliert.

Proposition 3. $p(t_{m+l}) = y_{m+l}$ für $l = 0, \dots, k$.

Beweis. Aus der Definition der Rückwärtsdifferenzen folgt

$$\nabla^l y_{m+k} = \sum_{i=0}^l (-1)^i \binom{l}{i} y_{m+k-i}. \quad (3.116)$$

Für $j = k - l$ ist daher

$$\begin{aligned} p(t_{m+l}) &= \sum_{l=0}^k (-1)^l \binom{j}{l} \nabla^l y_{m+k} = \sum_{l=0}^j (-1)^l \binom{j}{l} \nabla^l y_{m+k} \\ &= \sum_{l=0}^j (-1)^l \binom{j}{l} \sum_{i=0}^l (-1)^i \binom{l}{i} y_{m+k-i} \\ &= \sum_{i=0}^l (-1)^i y_{m+k-i} \sum_{l=i}^j (-1)^l \binom{j}{l} \binom{l}{i} \\ &= \sum_{i=0}^l (-1)^i y_{m+k-i} \binom{j}{i} \sum_{l=0}^{j-i} (-1)^l \binom{j-i}{l} \\ &= \sum_{i=0}^l (-1)^i y_{m+k-i} \binom{j}{i} (1-1)^{j-i} = y_{m+k-j} = y_{m+l}. \quad \square \end{aligned}$$

Einsetzen der Newton-Form der Interpolierenden (3.114) in die Forderung (3.113) liefert

$$\begin{aligned} p'(t_{m+k}) &= \left. \frac{dp(t_{m+k} - sh)}{h ds} \right|_{s=0} \\ &= \frac{1}{h} \sum_{l=0}^k (-1)^l \left. \frac{d}{ds} \binom{s}{l} \right|_{s=0} \nabla^l y_{m+k} \\ &= \frac{1}{h} \sum_{l=0}^k \left(\frac{d}{ds} \prod_{i=0}^{l-1} \frac{i-s}{i+1} \right) \bigg|_{s=0} \nabla^l y_{m+k} = \frac{1}{h} \sum_{l=1}^k \frac{1}{l} \nabla^l y_{m+k} \end{aligned}$$

Wir erhalten also die Verfahrensvorschrift des BDF-Verfahrens der Schrittlänge k (BDF k)

$$\sum_{l=1}^k \frac{1}{l} \nabla^l y_{m+k} = hf(t_{m+k}, y_{m+k}) \quad (3.117)$$

In der praktischen Implementierung merken wir uns anstelle der Werte y_{m+l} die Werte $\nabla^l y_{m+k}$, die wir iterativ berechnen. Die Darstellung hat den Vorteil, dass die Schrittlänge einfach erhöht werden kann, indem ein weiterer Term zur Summe (3.117) hinzugefügt wird. Bei variabler Schrittweite treten anstelle der Rückwärtsdifferenzen die dividierten Differenzen $y[t_{m+k}, \dots, t_{m+k-l}]$, die iterativ definiert sind mit $y[t_{m+k}] = y_{m+k}$ und

$$y[t_{m+k}, \dots, t_{m+k-l}] = \frac{y[t_{m+k}, \dots, t_{m+k-l+1}] - y[t_{m+k-1}, \dots, t_{m+k-l}]}{t_{m+k} - t_{m+k-l}}.$$

Im folgenden Satz wollen wir die Konsistenz und Stabilität der BDF-Verfahren untersuchen.

Proposition 4. *Die Konsistenzordnung von BDF k ist k mit Fehlerkonstante $C_{p+1}^* = -\frac{1}{k+1}$. Es ist A -stabil für $k \leq 2$ und $A(\alpha)$ -stabil für $k \leq 6$ mit den maximalen α aus Abbildung 3.3.2.*

Beweis. Nach Konstruktion mittels Interpolationspolynomen integriert BDF k Polynome bis zum Grad k theoretisch (d.h. ohne Rundungsfehler) exakt. Daher ist $L[y(t), h] = 0$, falls $y(t)$ ein Polynom von Grad höchstens k ist.

Nach dem Beweis von Proposition 2 sind dann die Konsistenzbedingungen für $q \geq k$ erfüllt, d.h. $p \geq k$. Nun wählen wir $f(t) = (k+1)t^k$, so dass $y = y_0 + t^{k+1}$. Das Verfahren der Schrittweite $k+1$ integriert die Gleichung exakt, sodass

$$\begin{aligned} 0 &= L_{k+1}[y(t), h] = L_k[y(t), h] + \frac{1}{k+1} \nabla^{(k+1)} y(t) \\ &= L_k[y(t), h] + \frac{1}{k+1} h^{k+1} y^{(k+1)}(t). \end{aligned}$$

Die letzte Gleichung lässt sich leicht mit (3.116) nachrechnen. Vergleich mit (3.103) gibt dann $C_{p+1} = -\frac{1}{k+1}$. Es folgt, dass $p = k$ und mit $\sum_{l=0}^k \beta_l = 1$ auch die genannte Fehlerkonstante C_{p+1}^* . Mit (3.116) gilt für das charakteristische Polynom

$$\phi_z(\lambda) = \sum_{l=1}^k \frac{1}{l} \sum_{i=1}^l (-1)^i \binom{l}{i} \lambda^{k-i} - z\lambda^k = \sum_{l=1}^k \frac{1}{l} \lambda^k (1 - \lambda^{-1})^l - z\lambda^k. \quad (3.118)$$

Für den Rand des Stabilitätsgebiet S gilt daher

$$\partial S \subset \{z \in \mathbb{C} | z = \sum_{l=1}^k \frac{1}{l} (1 - w)^l, |w| = 1\}. \quad (3.119)$$

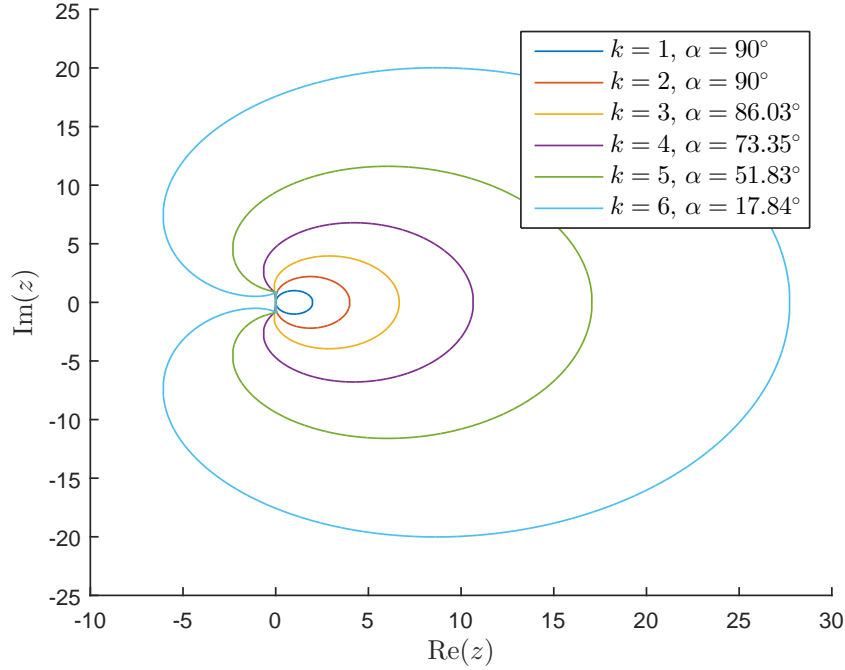


ABBILDUNG 8. Stabilitätsgebiete der BDF-Verfahren der Ordnung 1-6. Die Stabilitätsgebiete sind die unbeschränkten Gebiete außerhalb der dargestellten Ränder.

Wie in Abbildung 3.3.2 dargestellt, lassen sich daraus die Stabilitätsgebiete und die Winkel α berechnen. \square

Abbildung 3.3.2 zeigt, dass die Stabilität der BDF-Verfahren mit wachsendem k abnimmt. Tatsächlich ist BDF k für $k > 6$ nicht mehr $A(\alpha)$ -stabil für irgendein α . Einen Beweis dieser Tatsache findet man in [21]. In der Praxis wird auch schon BDF6 aufgrund des kleinen α in Mehrschritt-Codes nicht mehr verwendet [44].

Wir schließen diesen Abschnitt mit einer kurzen Beschreibung der von Klopfenstein und Reiher eingeführten NDF-Verfahren ab [?], die in *ode15s* von Shapine und Reichelt implementiert wurden [41]. Sie stellen eine Variation der Verfahren BDF1-5 dar und haben die Form

$$\sum_{l=1}^k \frac{1}{l} \nabla^l y_{m+k} - \kappa \gamma_k (y_{m+k} - y_{m+k}^{(0)}) = h f(t_{m+k}, y_{m+k}), \quad (3.120)$$

wobei der Ausdruck

$$y_{m+k}^{(0)} = \sum_{l=0}^k \nabla^l y_{m+k-1} \quad (3.121)$$

als Prädiktor (Startwert) zur Lösung der nichtlinearen Gleichungssystems (3.120) verwendet wird. Hierbei ist $\gamma_k = \sum_{i=1}^k \frac{1}{i}$ und κ ein frei zu

TABELLE 3. Vergleich der BDF- und NDF-Verfahren

k	κ	η	α_{BDF}	α_{NDF}
1	-0.1850	1.26	90°	90°
2	-1/9	1.26	90°	90°
3	-0.0823	1.26	86°	80°
4	-0.0415	1.12	73°	66°
5	0	1.00	51°	90°

wählender Parameter. Der Prädiktor ist so gewählt, dass

$$\begin{aligned}
y_{m+k} - y_{m+k}^{(0)} &= y_{m+k} + \sum_{l=0}^k (\nabla^{l+1} y_{m+k} - \nabla^l y_{m+k}) \\
&= y_{m+k} + \nabla^{k+1} y_{m+k} - \nabla^0 y_{m+k} \\
&= \nabla^{k+1} y_{m+k} \approx h^{k+1} y^{(k+1)}|_{(m+k)h}.
\end{aligned}$$

Es folgt, dass das NDFk-Verfahren (3.120) ebenfalls Konsistenzordnung k hat mit Fehlerkonstante

$$C_{p+1}^* = - \left(\frac{1}{k+1} + \kappa \gamma_k \right). \quad (3.122)$$

Der Parameter κ wird nun von Shampine/Reichelt genutzt, um die Fehlerkonstante zu verkleinern, ohne dabei die Stabilität stark zu beeinträchtigen. Die gewählten κ -Werte zusammen mit den Änderungen der Stabilitätswinkel α sind in Tabelle 3.3.2 aufgeführt. Die η -Werte geben an, um welchen Faktor die Schrittweite bei gleicher Genauigkeit aufgrund der kleineren Fehlerkonstante erhöht werden kann. Aufgrund des schon relativ kleinen Stabilitätswinkel für $k = 5$ wurde hier $\kappa = 0$ gewählt.

3.3.3. *Integration von Index 1 DAEs mit ode15s.* Wir wenden uns nun der numerischen Lösung von Problemen der Form

$$My' = f(t, y), \quad (3.123)$$

mittels *ode15s* zu, wobei wir zunächst annehmen, dass M regulär ist, auch wenn wir uns natürlich letztlich gerade für den entgegengesetzten Fall interessieren. Durch Multiplikation mit M^{-1} von links können wir das Problem (3.123) dann auf den bisher betrachteten Fall zurückführen. Aufgrund der großen Lipschitz-Konstanten, die in steifen Problemen auftreten, wird das nichtlineare Gleichungssystem (3.120) nicht durch Funktionaliteration, sondern durch ein vereinfachtes Newton-Verfahren gelöst. Wie in [41] dargelegt, zeigt die Identität

$$\sum_{l=1}^k \frac{1}{l} \nabla^l y_{n+1} = \gamma_k (y_{n+1} - y_{n+1}^{(0)}) + \sum_{l=1}^k \gamma_l \nabla^l y_n, \quad (3.124)$$

dass (3.120) äquivalent zu

$$(1 - \kappa)\gamma_k(y_{n+1} - y_{n+1}^{(0)}) + \sum_{l=1}^k \gamma_l \nabla^l y_n - hF(t_{n+1}, y_{n+1}) = 0 \quad (3.125)$$

ist, wobei wir $F = M^{-1}f$ gesetzt haben. Die Korrektur

$$\Delta^{(i)} = y_{n+1}^{(i+1)} - y_{n+1}^{(i)} \quad (3.126)$$

des vereinfachten Newtonverfahrens wird dann mittels

$$\left(I - \frac{h}{(1 - \kappa)\gamma_k} J_F \right) \Delta^{(i)} = \frac{h}{(1 - \kappa)\gamma_k} F(t_{n+1}, y_{n+1}^{(i)}) - \Psi - (y_{n+1}^{(i)} - y_{n+1}^{(0)})$$

bestimmt. Hierbei ist J_F eine Approximation der Jacobi-Matrix von F und

$$\Psi = \frac{1}{(1 - \kappa)\gamma_k} \sum_{l=1}^k \gamma_l \nabla^l y_n \quad (3.127)$$

eine Größe, die während der Berechnung von y_{n+1} konstant ist. Als Iterationsmatrix erhalten wir somit den Ausdruck

$$I - \frac{h}{(1 - \kappa)\gamma_k} J_F = I - \frac{h}{(1 - \kappa)\gamma_k} M^{-1} J_f, \quad (3.128)$$

Multiplizieren wir nun wieder mit M , so erhalten wir die Iterationsmatrix

$$M - \frac{h}{(1 - \kappa)\gamma_k} J_f, \quad (3.129)$$

für das neue Problem (3.123). Die einfache, aber wichtige Beobachtung ist nun die, dass dieser Ausdruck auch für *singuläre* Matrizen M definiert ist. Für $h \rightarrow 0$ geht die Kondition der Iterationsmatrix zwar nach unendlich, aber für nicht zu kleine h lässt sich die Newton-Iteration dennoch erfolgreich verwenden. Strategien zur Präkonditionierung der Iterationsmatrizen, zur Steuerung der Schrittweite, zu Konvergenzkriterien für die Folge $y_{n+1}^{(i)}$ und zur Berechnung neuer Jacobi-Matrizen werden in [42] beschrieben. Bisher hat der Index der DAE (3.123) keine Rolle gespielt. Tatsächlich wird die Index-1-Bedingung nur zur Berechnung konsistenter Startwerte benötigt, wie in Abschnitt 3.2.3 erklärt. In [42] wird gezeigt, dass konsistente Anfangsbedingungen eindeutig bestimmt werden können, falls das Problem $M(t)y' = f(t, y)$ von Index 1 ist und eine Lösung $y(t)$ besitzt. Der Beweis verläuft wie die Argumentation in Bemerkung und wendet dann den Satz von der impliziten Funktion an.

3.4. Analytischen Teillösung. Die Gleichungsübersicht in Tabelle 1 macht deutlich, dass die Festkörper-Konzentration $c_s(r, x, t)$ nur über die Werte an den Partikeloberflächen $\bar{c}_s(x, t) = c_s(r_p, x, t)$ in das Gesamtmodell eingeht. Da die Partikel-Diffusionsgleichung (2.3) linear ist, lässt sich eine analytische Lösung mit Hilfe einer Greenschen Funktion angeben. Auf diese Weise kann man $\bar{c}_s(x, t)$ im Prinzip nur aus den

Anfangswerten und den Randflüssen i_r berechnen. Dies wiederum wäre auch aus Sicht der numerischen Berechnung des Gesamtsystems sehr attraktiv, da so auf eine Diskretisierung der r -Dimension verzichtet werden kann und somit ein wesentlich kleineres DAE-System integriert werden kann. Ein ähnlicher Ansatz wurde bereits von der Newman-Gruppe benutzt und ist in [10, 15] beschrieben.

Allerdings hat die Elimination der r -Dimension einen Preis. Anstelle an jedem x -Knoten eine Diffusionsgleichung zu lösen, muss dort ein Integral über die vergangenen Neumann-Randflüsse ausgewertet werden. Wie wir sehen werden, erfordert die numerische Berechnung des Integrals bei längeren Simulationszeiten eine große Anzahl von Stützstellen, die zu einem großen Speicher- und Rechenaufwand führen. Hinzu kommt, dass die Methode nicht für einen temperaturabhängigen Diffusionskoeffizienten D_s geeignet ist, da in diesem Fall die Greensche Funktion nicht mehr im Voraus berechnet werden kann. Aus diesen Gründen haben wir uns - im Gegensatz zu [10, 15] - gegen eine Implementierung dieses Ansatzes entschieden.

Dennoch soll er hier vorgestellt werden. Zunächst führen wir die allgemeine Greensche Funktion der Diffusionsgleichung ein. Dann berechnen wir die Greensche Funktion für die Diffusion in einer Kugel in \mathbb{R}^3 mit Neumann-Randbedingung. Anschließend untersuchen wir die numerische Implementierung der resultierenden analytischen Lösungsformel. Schließlich beschreiben wir eine andere Methode, die uns erlaubt die Linearität der Partikeldiffusionsgleichung auch bei temperaturabhängigen Diffusionskoeffizienten auszunutzen und so die Berechnung in der r -Dimension zu beschleunigen.

3.4.1. Greensche Funktion der allgemeinen Diffusionsgleichung. Wir betrachten zunächst die inhomogene, lineare Diffusionsgleichung in einem Gebiet Ω

$$\begin{cases} \mathcal{L}_{t,x}y := y_t - D\Delta_x y &= f(\mathbf{x}, t) & \mathbf{x} \in \Omega, 0 < t \leq t_f \\ y \text{ bzw. } \mathbf{n} \cdot \nabla_x y &= g(\mathbf{x}, t) & \mathbf{x} \in \partial\Omega, 0 < t \leq t_f \\ y &= y_0(\mathbf{x}) & \mathbf{x} \in \Omega, t = 0 \end{cases} \quad (3.130)$$

mit Anfangs- und *Dirichlet* bzw. *Neumann*-Randbedingungen und geben folgende

Definition 8 (Greensche Funktion). Die Lösung $G(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ des assoziierten Problems

$$\begin{cases} \mathcal{L}_{t,x}^* G := -G_t - D\Delta_x G &= \delta(\mathbf{x} - \boldsymbol{\xi})\delta(t - \tau) \quad (\text{Dist.}) \\ G \text{ bzw. } \mathbf{n} \cdot \nabla_x G &= 0 & \mathbf{x} \in \partial\Omega, 0 \leq t \leq t_f \\ G &= 0 & t < \tau \quad (\text{„Kausalität“}) \end{cases} \quad (3.131)$$

heißt *Greensche Funktion* des Problems (3.130). Die erste Gleichung ist als Gleichung von Distributionen zu verstehen.

Satz 7. Seien y und G Lösungen von (3.130) bzw. (3.131). Dann ist

$$\begin{aligned} y(\mathbf{x}, t) &= \int_{\Omega} G(\mathbf{x}, t; \boldsymbol{\xi}, 0) y_0(\boldsymbol{\xi}) d\boldsymbol{\xi} + \int_0^t \int_{\Omega} G(\mathbf{x}, t; \boldsymbol{\xi}, \tau) f(\boldsymbol{\xi}, \tau) d\boldsymbol{\xi} d\tau \\ &\quad + D \int_0^t \int_{\partial\Omega} (-\nabla G \text{ bzw. } G)(\mathbf{x}, t; \boldsymbol{\xi}, \tau) g(\boldsymbol{\xi}, \tau) d\boldsymbol{\xi} d\tau. \end{aligned} \quad (3.132)$$

Der Ausdruck $(-\nabla G \text{ bzw. } G)$ ist dabei als $-\nabla G$ im Fall einer Dirichlet- und als G im Fall einer Neumann-Randbedingung zu lesen.

Beweis. Für feste \mathbf{x} , $t < t_f$ gilt mittels partieller Integration bezüglich τ und Greenscher Formel bezüglich $\boldsymbol{\xi}$

$$\begin{aligned} \int_0^t \int_{\Omega} G f(\boldsymbol{\xi}, \tau) d\boldsymbol{\xi} d\tau &= \int_0^{t_f} \int_{\Omega} G \mathcal{L}_{\tau, \boldsymbol{\xi}} y d\boldsymbol{\xi} d\tau = \int_0^{t_f} \int_{\Omega} y \mathcal{L}_{\tau, \boldsymbol{\xi}}^* G d\boldsymbol{\xi} d\tau \\ &\quad - D \int_0^{t_f} \int_{\partial\Omega} (G \nabla_{\boldsymbol{\xi}} y - y \nabla_{\boldsymbol{\xi}} G) \cdot \mathbf{d}\boldsymbol{\xi} d\tau - \int_{\Omega} ([Gy]_{\tau=t_f} - [Gy]_{\tau=0}) d\boldsymbol{\xi} \\ &= y(\mathbf{x}, t) - D \int_0^t \int_{\partial\Omega} (-\nabla G \text{ bzw. } G)(\mathbf{x}, t; \boldsymbol{\xi}, \tau) g(\boldsymbol{\xi}, \tau) d\boldsymbol{\xi} d\tau \\ &\quad - \int_{\Omega} G(\mathbf{x}, t; \boldsymbol{\xi}, 0) y_0(\boldsymbol{\xi}) d\boldsymbol{\xi}. \end{aligned}$$

Umstellen nach y liefert die Behauptung. \square

Bemerkung 4. Satz 7 setzt die Existenz von Lösungen (3.130) und (3.131) bereits vorausgesetzt. Mit mehr Arbeit lässt sich (3.132) nutzen, um unter geeigneten Annahmen an die Glattheit von f, g , und y_0 die Existenz einer Lösung zu beweisen. Wir verweisen dafür auf [12].

Bemerkung 5. Mit $y = y_0 \equiv 1$, $f = g \equiv 0$ folgt aus (3.138) die Normalisierung

$$\int_{\Omega} G d\mathbf{x} = 1. \quad (3.133)$$

3.4.2. Greensche Funktion für die Diffusion in einer Kugel. Wir beschränken uns nun auf ein Kugel-Gebiet, Neumann-Randbedingung und radialsymmetrischen Daten f, g, y_0

$$\begin{aligned} \Omega &= B = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| \leq 1\} \\ f &= f(r, t), g = g(t), y_0 = y_0(r) \end{aligned}$$

und leiten in diesem Fall eine explizite Formel für die Greensche Funktion her. Aus der Symmetrie des Problems folgern wir $y = y(r, t)$ und $G = G(r, t; \rho, \tau)$. Mit

$$\mathcal{L}_{t,r}^* = -\frac{\partial}{\partial t} - D\Delta_r = -\frac{\partial}{\partial t} - D\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right)$$

erscheint das Problem (3.131) in Kugelkoordinaten in der Form

$$\begin{cases} \mathcal{L}_{t,r}^* G &= \frac{1}{4\pi r^2} \delta(r - \rho) \delta(t - \tau) \quad (\text{als Distributionen}) \\ \frac{\partial G}{\partial r} &= 0 & r = 1 \\ G &= 0 & t < \tau \text{ („Kausalität“)} \end{cases} \quad (3.134)$$

Wie man leicht nachrechnet, sind die Eigenfunktionen des Laplace-Operators in der Kugelkoordinate ρ mit Neumann-Randbedingung

$$y_n = \begin{cases} 1 & n = 0 \\ \frac{1}{\rho} \sin(\lambda_n \rho) & n > 0, \lambda_n \text{ nte positive Nullstelle von } x = \tan x \end{cases}$$

mit zugehörigen Eigenwerten $-\lambda_n^2$. Die Eigenfunktionen sind orthogonal bezüglich des inneren Produkts

$$\langle f, g \rangle = \int_B f g dV.$$

Wir suchen nach einer Greenschen Funktion in der Form

$$G(r, t, \rho, \tau) = \sum_{n=0}^{\infty} a_n(t, r, \tau) y_n(\rho). \quad (3.135)$$

Aufgrund der Definitionen von G in (3.134) und $\delta(r - \rho)$ ist einerseits

$$\langle \mathcal{L}_{\tau,\rho}^* G, y_m \rangle = \langle \frac{1}{4\pi r^2} \delta(r - \rho) \delta(t - \tau), y_m \rangle = y_m(r) \delta(t - \tau),$$

andererseits aufgrund der Orthogonalität der Eigenfunktionen

$$\begin{aligned} \langle \mathcal{L}_{t,r}^* G, y_m \rangle &= \sum_{n=0}^{\infty} \left(-\frac{\partial a_n}{\partial \tau} + D\lambda_n^2 a_n \right) \langle y_n, y_m \rangle \\ &= \left(-\frac{\partial a_m}{\partial \tau} + D\lambda_m^2 a_m \right) \langle y_m, y_m \rangle. \end{aligned}$$

Für $t < \tau$ folgt daraus $-\frac{\partial a_m}{\partial \tau} + D\lambda_m^2 a_m = 0$ und daher

$$a_m = C_m(r, t) e^{D\lambda_m^2 \tau}.$$

Die Koeffizienten $C_m(r, t)$ ergeben sich aus der Sprungbedingung bei $\tau = t$:

$$\begin{aligned} 1 = \int_{t^-}^{t^+} \delta(t - \tau) d\tau &= \frac{\langle y_m, y_m \rangle}{y_m(r)} \int_{t^-}^{t^+} \left(-\frac{\partial a_m}{\partial \tau} + D\lambda_m^2 a_m \right) d\tau \\ &= \frac{\langle y_m, y_m \rangle}{y_m(r)} a_m|_{t^-}^{t^+} \\ &= \frac{\langle y_m, y_m \rangle}{y_m(r)} C_m e^{D\lambda_m^2 t}, \end{aligned}$$

wobei wir im letzten Schritt $G = 0$ für $\tau > t$ genutzt haben. Es folgt

$$G(r, t; \rho, \tau) = \sum_{m=0}^{\infty} \frac{y_m(r)y_m(\rho)}{\langle y_m, y_m \rangle} e^{-D\lambda_m^2(t-\tau)}.$$

Schließlich berechnen sich die Normalisierungskonstanten für $m > 0$ zu

$$\begin{aligned} \langle y_m, y_m \rangle &= 4\pi \int_0^1 \sin^2(\lambda_m r) dr = 2\pi \int_0^1 1 - \cos(2\lambda_m r) dr \\ &= 2\pi R(1 - \sin 2\lambda_m / 2\lambda_m) = 2\pi(1 - \cos^2 \lambda_m) = 2\pi \sin^2 \lambda_m \\ &= 2\pi R \frac{\sin^2 \lambda_m}{\lambda_m^2 \cos^2 \lambda_m + \cos^2 \lambda_m} = 2\pi \frac{\lambda_m^2}{\lambda_m^2 + 1}. \end{aligned}$$

Wir erhalten also

$$G(r, t; \rho, \tau) = \frac{3}{4\pi} + \sum_{m=1}^{\infty} \frac{\lambda_m^2 + 1}{2\pi r \rho \lambda_m^2} \sin(\lambda_m r) \sin(\lambda_m \rho) e^{-D\lambda_m^2(t-\tau)}. \quad (3.136)$$

Kehren wir nun zur speziellen Diffusionsgleichung in den Partikeln des Aktivmaterials in der Form

$$\begin{cases} \frac{\partial c_s}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 D \frac{\partial c_s}{\partial r} \right) & \mathbf{r} \in B = \{\mathbf{r} \in \mathbb{R}^3 : r = \|\mathbf{r}\| \leq 1\} \\ \frac{\partial c_s}{\partial r} = g & \mathbf{r} \in \partial B. \end{cases} \quad (3.137)$$

zurück, wobei wir x -Abhängigkeit von c_s und g zunächst außer Acht lassen dürfen. Unsere Lösungsformel (3.132) lautet dann

$$c_s(r, t) = \int_0^1 G(r, t; \rho, 0) c_{s,0}(\rho) 4\pi \rho^2 d\rho + D \int_0^t G(r, t; 1, \tau) g(\tau) 4\pi d\tau. \quad (3.138)$$

mit G aus (3.136). Wie anfangs festgestellt, interessiert uns vor allem der Wert von c_s bei $r = 1$. Hier lässt sich Formel (3.138) weiter vereinfachen. Sei zunächst

$$c_{s,av} = \oint_B c_s dV = \frac{3}{4\pi} \int_0^1 4\pi r^2 c_s dr. \quad (3.139)$$

die Durchschnittskonzentration in B . Aufgrund der Normalisierung (3.133) gilt dann

$$c_{s,av} = c_{s,av}^0 + 3D \int_0^t g(\tau) d\tau, \quad (3.140)$$

wobei $c_{s,av}^0$ die anfängliche Durchschnittskonzentration ist. Dies ist auch physikalisch klar, da die Durchschnittskonzentration gleich der durchschnittlichen Anfangskonzentration plus dem Netto-Durchtritt an der Oberfläche mal der spezifischen Oberfläche ist. Aus (3.133) folgt, dass,

falls $c_{s,0}(\rho) = c_{s,av}^0$ keine ρ -Abhängigkeit aufweist, das erste Integral in (3.138) den konstanten Wert $c_{s,av}^0$ annimmt. Da weiter

$$G(r, t, \rho, 0) \rightarrow \frac{3}{4\pi} \quad \text{für } t \rightarrow \infty \quad (3.141)$$

gilt selbiges auch für hinreichend große Zeiten t , wenn $c_{s,0}(\rho) \neq \text{konst.}$ Für einen hinreichend großen Zeithorizont oder Equilibriums-Anfangsbedingungen können wir uns somit auf die Auswertung des zweiten Integrals beschränken. Aufgrund der Identität

$$(\lambda_m^2 + 1) \sin^2 \lambda_m = \lambda_m^2 \quad (3.142)$$

ergibt sich für $r = 1$ somit

$$\bar{c}_s(t) = c_{s,av} + 2D \int_0^t \sum_{m=1}^{\infty} e^{-D\lambda_m^2(t-\tau)} g(\tau) d\tau. \quad (3.143)$$

Für fixes D lässt sich die Funktion

$$w(t) = D \sum_{m=1}^{\infty} e^{-D\lambda_m^2 t}. \quad (3.144)$$

an gewählten Punkten im Voraus berechnen. Man sieht, dass w eine Singularität in $t = 0$ besitzt, aber so, dass

$$\int_0^{\infty} w(t) dt = \sum_{m=1}^{\infty} \frac{1}{\lambda_m^2} < \frac{1}{\pi^2} \sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{1}{6} < \infty. \quad (3.145)$$

Bemerkung 6. Es gilt die Identität

$$\sum_{m=1}^{\infty} \frac{1}{\lambda_m^2} = \frac{1}{10}. \quad (3.146)$$

Natürlich ist der exakte Werte der Summer für die Numerik unbedeutend. Da er sich aber über eine instruktive und so weit uns bekannt neue Anwendung der Greenschen Funktion berechnen lässt, geben wir den

Beweis. Wir betrachten die Diffusionsgleichung (3.137) erweitert um einen Quellterm f :

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 D \frac{\partial u}{\partial r} \right) + f & \mathbf{r} \in B = \{\mathbf{r} \in \mathbb{R}^3 : r = \|\mathbf{r}\| \leq 1\} \\ \frac{\partial u}{\partial r} = g & \mathbf{r} \in \partial B. \end{cases} \quad (3.147)$$

(Wir wählen u statt c_s , da im Newman-Modell kein Quellterm auftritt.) Die Lösungsformel (3.138) ist dann nach (3.132) um einen Term für den

mit einem Quellterm zu erweitern:

$$\begin{aligned} u(r, t) = & \int_B G(r, t; \rho, 0) u_0(\rho) d\boldsymbol{\rho} + \int_0^t \int_B G(\rho, t; \rho, \tau) f(\rho, \tau) d\boldsymbol{\rho} d\tau \\ & + D \int_0^t \int_{\partial B} G(r, t; 1, \tau) g(\rho, \tau) d\boldsymbol{\rho} d\tau =: I_1 + I_2 + I_3. \end{aligned} \quad (3.148)$$

Wir setzen nun eine passend gewählte Lösung ein. Um den Ausdruck $\sum \lambda_m^{-2}$ zu erhalten, bietet es sich an $t \rightarrow \infty$ gehen lassen und g konstant zu wählen. Dies legt die stationäre Funktion $u = \frac{1}{2}r^2$ nahe. Diese ist eine Lösung der Diffusionsgleichung für den Randterm $g = 1$ und den Quellterm $f = -D\Delta_r u = -3D$. Die drei Integrale berechnen sich dann zu:

$$I_1 = \int_0^1 G(1, t; \rho, 0) \frac{1}{2} \rho^2 4\pi \rho^2 d\rho \rightarrow \int_0^1 \frac{3}{4\pi} 2\pi \rho^4 d\rho = \int_0^1 \frac{3}{2} \rho^4 d\rho = \frac{3}{10},$$

für $t \rightarrow \infty$ (da $G(r, t, \rho, 0) \rightarrow \frac{3}{4\pi}$),

$$I_2 = f \int_0^t \int_B G d\boldsymbol{\rho} d\tau = f \int_0^t d\tau = ft = -3Dt$$

aufgrund der Normalisierung (3.133) und

$$I_3 = D \int_0^t G(1, t; 1, \tau) 4\pi d\tau = 3Dt + D \int_0^t \sum_{m=1}^{\infty} G_m(r, t, 1, \tau) 4\pi d\tau$$

mit G_m den Summanden der Reihenentwicklung (3.136). Für $r = 1$ ist dann

$$I_2 + I_3 = 2D \int_0^t \sum_{m=1}^{\infty} e^{-D\lambda_m^2(t-\tau)} d\tau \rightarrow 2 \sum_{m=1}^{\infty} \frac{1}{\lambda_m^2} \quad \text{für } t \rightarrow \infty.$$

Das Vertauschen von Summation und Integration ist zulässig aufgrund des Satzes von Fubini. Einsetzen der Ergebnisse in die Lösungsformel (3.148) gibt

$$\frac{1}{2} = \frac{3}{10} + 2 \sum_{m=1}^{\infty} \frac{1}{\lambda_m^2}. \quad (3.149)$$

Die Behauptung folgt. \square

Nach diesem Exkurs kommen wir nun zu der Frage, wie wir das Integral

$$\int_0^t w(t-\tau)g(\tau)d\tau. \quad (3.150)$$

numerisch berechnen können. Da $w \rightarrow \infty$ für $\tau \rightarrow t$, ist einerseits für $\tau \approx t$ eine sehr feine Diskretisierung erforderlich. Für große Werte von $t - \tau$ ist zwar $w(t - \tau)$ nahezu konstant, g (der Reaktionsstrom) kann aber durchaus volatil sein, jedenfalls, wenn nicht nur der konstante

Lade- oder Entladevorgang, sondern auch dynamisches Verhalten modelliert werden soll.

Andererseits geht w für ein D der typischen Größenordnung 10^{-5} 1/s nur langsam gegen 0:

$$\int_t^\infty w(\tau) d\tau < \frac{1}{100} \int_0^\infty w(\tau) d\tau \quad (3.151)$$

erst ab $t > 19330 \text{ s}$. Insofern dürfen wir auch nicht zu früh abschneiden. Um beide Bedingungen zu erfüllen, wählen wir Stützstellen $(t_k)_{k=0}^n$ derart, dass

$$\int_{t_k}^{t_{k+1}} w(\tau) d\tau = \frac{1}{n} \int_0^\infty w(\tau) d\tau = \frac{1}{10n} D \quad (3.152)$$

3.5. Aspekte der Implementierung.

3.5.1. *Diskretisierung der Partikel-Diffusionsgleichung.* Wir führen hier die räumliche Diskretisierung der Diffusionsgleichung in der Partikeldimension r durch, einerseits um die allgemeine Methode zu illustrieren, andererseits weil hier einige Besonderheiten zu beachten sind. Wir betrachten also

$$\frac{\partial c_s}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 D_s \frac{\partial c_s}{\partial r} \right) = D_s \left(\frac{2}{r} \frac{\partial c_s}{\partial r} + \frac{\partial^2 c_s}{\partial r^2} \right) \quad (3.153)$$

mit der Randbedingung

$$D_s \frac{\partial c_s}{\partial r} \Big|_{r=r_p} = -\frac{i_r}{a_s F} =: g \quad (3.154)$$

Um Probleme mit dem bei $r = 0$ undefinierten Ausdruck $2c_{s,r}/r$ zu vermeiden, setzen wir zunächst $u = rc_s$. Zusätzlich ändern wir noch die Skalierung indem wir r durch r/r_p ersetzen. Die Diffusionsgleichung erscheint dann in die Form

$$\frac{\partial u}{\partial t} - \hat{D}_s \frac{\partial^2 u}{\partial r^2} = 0 \quad (3.155)$$

bringen, und als Randbedingung

$$\hat{D}_s \left(\frac{\partial u}{\partial r} \Big|_{r=1} - u(1) \right) = g \quad (3.156)$$

mit $\hat{D}_s = D_s/r_p^2$. Zur Diskretisierung dieser Gleichung wollen wir wieder Tschebyschow-Punkte verwenden. Während es allerdings vorteilhaft ist, eine höhere Dichte an Punkte in Nähe des Rande zu verwenden, scheint dies im Zentrum überflüssig, da dort die niedrigsten Gradienten auftreten. Wir verwenden daher folgenden Trick: Wir diskretisieren einen ganzen Durchmesser von $r = -1$ bis $r = 1$ und benutzen dann die Symmetrie $u(r) = -u(-r)$.

Seien also x_i , $i = 1, \dots, 2m+1$ die Tschebyschow-Punkte, D die zugehörige $(2m+1) \times (2m+1)$ Differentiationsmatrix und $u_i = u(r = x_i)$

mit x_i . Für die Implementierung ist es einfacher die Reihenfolge der x_i umzudrehen, sodass $x_i < x_{i+1}$, die x -Achse also in vertrauter Weise nach rechts zeigt. Entsprechend muss D an der horizontalen und vertikalen Mittellinie gespiegelt werden, was einer Rotation um 180° entspricht. Es zeigt sich, dass die resultierende Matrix gerade $-D$ ist. Nach Satz 2 ist der zentrale Koeffizient

$$D_{m+1,m+1} = -\frac{x_{m+1}}{2(1 - x_{m+1}^2)} = 0 \quad (3.157)$$

Dann ist

$$u'_i / \hat{D}_s = \sum_{j=1}^m D_{i,j}^2 u_j + \sum_{j=m+2}^{2m+1} D_{i,j}^2 u_j = \sum_{j=m+2}^{2m+1} (D_{i,j}^2 - D_{i,2m+1-j}^2) u_j. \quad (3.158)$$

Wir können also die $m \times m$ Matrix $\tilde{D} = (D_{i,j}^2 - D_{i,2m+1-j}^2)_{i,j}$ mit $m+2 \leq i, j \leq 2m+1$ verwenden. Schließlich ersetzen wir noch die letzte Reihe von \tilde{D} durch die diskretisierte Randbedingung

$$\hat{D}_s(\tilde{D}_{2m+1,j} u_j - u_{2m+1}) = g. \quad (3.159)$$

Wir nennen die entstandene Matrix A . Außerdem sei M die $m \times m$ Diagonalmatrix mit $M_{ii} = 1$ für $i < m$ und $M_{mm} = 0$ und b der m -Vektor mit $b_i = 1$ für $i < m$ und $b_m = 1$. Die räumlich Diskretisierung lautet dann

$$M\mathbf{u}' = A\mathbf{u} + gb. \quad (3.160)$$

3.5.2. Die Jacobi-Matrix. In Abschnitt 3.3.3 wurde bereits deutlich, dass die Jacobi-Matrix $\partial f / \partial y$ in *ode15s* eine entscheidende Rolle spielt. Wenn dies vom Nutzer nicht vorgegeben wird, müssen *ode15s* diese mittels finiten Differenzen berechnen werden. Dies erfordert N Funktionsauswertungen pro Berechnung der Jacobi-Matrix, wobei N die Anzahl der Freiheitsgrade des Systems ist.

Alternativ kann die Jacobi-Matrix analytisch berechnet werden. Dies erfordert die Ableitung der Gleichungen des Newman-Systems „von Hand“, bringt aber eine signifikante Beschleunigung: der Rechenaufwand zur Berechnung einer Jacobi-Matrix entspricht etwa einer bis maximal zwei Funktionsauswertungen. Bei einer typischen Diskretisierung mit insgesamt $N \approx 600$ Freiheitsgraden führt dies zu einer Reduktion der Rechenzeit um einen Faktor von 20.

In Abbildung 3.5.2 ist die Besetzungsstruktur der Jacobi-Matrix für eine Diskretisierung mit jeweils 10 Stützstellen in Anode, Separator und Kathode und 30 Stützstellen in der r -Dimension zu sehen. Insgesamt ist $y(t)$ für jeden Zeitpunkt t ein Vektor in \mathbb{R}^N mit $N = 717$ und $\partial f / \partial y$ entsprechend eine $N \times N$ Matrix. Die 20 großen Blöcke sind gerade Kopien der Matrix A aus dem letzten Abschnitt - einer für jede Stützstelle in Anode und Kathode (im Separator findet ja keine Reaktion statt). Da wir Spektralverfahren verwenden sind diese Blöcke wie auch die anderen kleineren Böcke vollbesetzt (linke Abbildung). Das war, wie

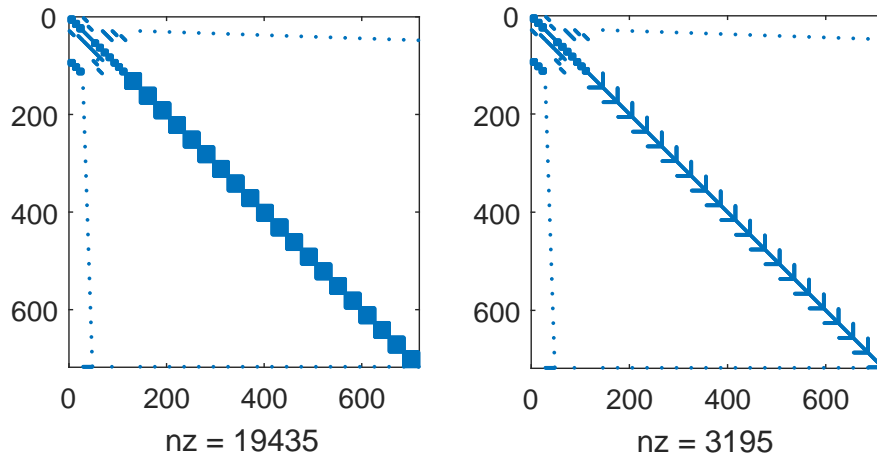


ABBILDUNG 9. Besetzungsstruktur der $N \times N$ Jacobi-Matrix ohne bzw. mit Diagonalisierung für $N = 717$. nz bezeichnet jeweils die Anzahl der Einträge ungleich null.

gesagt, der Preis den wir für die Verwendung von Spektralverfahren zahlen. Dennoch ist die Matrix als Ganze dünnbesetzt mit unter 4% der Einträge ungleich null.

Wir können dies ausnutzen, indem wir Massematrix M und Jacobi-Matrix $\partial f / \partial y$ an *ode15s* als dünnbesetzte Matrizen übergeben. *ode15s* wählt dann automatisch geeignete Verfahren für das Lösen linearer Systeme. Natürlich lohnt sich dies umso mehr, je größer N ist. Für $N = 717$ verkürzte sich die Rechenzeit auf dem verwendeten PC um einen Faktor von ca. 3,5. Man kann diesen Effekt noch verstärken, indem man die Untermatrix der ersten $m - 1$ Reihen und Spalten von A diagonalisiert. Interessanterweise scheinen die Eigenwerte dieser Matrizen alle reell zu sein. Wir haben diese Behauptung für $m \leq 1000$ numerisch überprüft, aber bisher noch keinen Beweis gefunden.

3.5.3. Dünnbesetzte Matrizen.

3.5.4. Vorfaktorisierung.

3.5.5. Singulär gestörte Probleme. Dünnbesetzte Matrizen - Dennoch LU-Zerlegung. Könnte man durch Krylow-Verfahren ersetzen?

- (Schätzung der Anfangsbedingungen) Substitution $u = rc_s$ führt zu

4. ERGEBNISSE

4.1. Konvergenz.

4.2. Vergleich mit COMSOL.

TABELLE 4. Koeffizienten $a_{i,j}, b_i, d_i$

i/j	0	1								2	
0	-105	0.74	$-6.96 \cdot 10^{-4}$	1	6680	178	$2.8 \cdot 10^{-4}$	2	3	4	

5. ZUSAMMENFASSUNG

ANHANG A. PARAMETER

TABELLE 5. Gleichgewichtsspannungen und Entropiekoeffizienten

x	ϕ_0^-	ϕ_0^+	$\phi_{0,T}^-$	$\phi_{0,T}^+$	x	ϕ_0^-	ϕ_0^+	$\phi_{0,T}^-$	$\phi_{0,T}^+$
0.00	0.51734	6.08260	0.00527	0.00000	0.50	0.11719	4.03871	-0.11022	-0.08668
0.01	0.41761	6.01339	0.04916	0.00000	0.51	0.11626	4.02071	-0.10349	-0.08323
0.02	0.34811	5.94558	0.11269	0.00000	0.52	0.11502	4.00334	-0.10093	-0.07741
0.03	0.29993	5.87916	0.19070	0.00000	0.53	0.11312	3.98657	-0.10016	-0.07216
0.04	0.26418	5.81412	0.24996	0.00000	0.54	0.11025	3.97041	-0.09997	-0.06937
0.05	0.23506	5.75043	0.24670	0.00000	0.55	0.10620	3.95482	-0.09994	-0.06924
0.06	0.21602	5.68809	0.19018	0.00000	0.56	0.10116	3.93980	-0.09995	-0.07067
0.07	0.20879	5.62708	0.12383	0.00000	0.57	0.09589	3.92533	-0.09997	-0.07064
0.08	0.20626	5.56738	0.07015	0.00000	0.58	0.09132	3.91139	-0.09999	-0.06870
0.09	0.20260	5.50897	0.03134	0.00000	0.59	0.08855	3.89797	-0.10000	-0.06792
0.10	0.19795	5.45185	0.00379	0.00000	0.60	0.08755	3.88506	-0.10000	-0.06874
0.11	0.19283	5.39599	-0.01617	0.00000	0.61	0.08718	3.87263	-0.10000	-0.06891
0.12	0.18702	5.34138	-0.03121	0.00000	0.62	0.08689	3.86067	-0.10000	-0.06747
0.13	0.18056	5.28801	-0.04315	0.00000	0.63	0.08670	3.84917	-0.10000	-0.06480
0.14	0.17399	5.23585	-0.05330	0.00000	0.64	0.08662	3.83811	-0.10000	-0.06293
0.15	0.16782	5.18490	-0.06263	0.00000	0.65	0.08660	3.82747	-0.10000	-0.06339
0.16	0.16234	5.13514	-0.07192	0.00000	0.66	0.08660	3.81724	-0.10000	-0.06422
0.17	0.15763	5.08655	-0.08188	0.00000	0.67	0.08660	3.80741	-0.10000	-0.06411
0.18	0.15345	5.03911	-0.09305	0.00000	0.68	0.08647	3.79795	-0.10000	-0.06375
0.19	0.14957	4.99282	-0.10572	0.00000	0.69	0.08631	3.78886	-0.10000	-0.06628
0.20	0.14571	4.94765	-0.11969	0.00000	0.70	0.08629	3.78012	-0.10000	-0.07137
0.21	0.14188	4.90359	-0.13411	0.00000	0.71	0.08612	3.77170	-0.10000	-0.07242
0.22	0.13826	4.86063	-0.14755	0.00000	0.72	0.08574	3.76360	-0.10000	-0.07334
0.23	0.13496	4.81874	-0.15853	0.00000	0.73	0.08555	3.75581	-0.10000	-0.07884
0.24	0.13196	4.77792	-0.16615	0.00000	0.74	0.08536	3.74829	-0.10000	-0.08147
0.25	0.12926	4.73815	-0.17045	0.00000	0.75	0.08517	3.74105	-0.10000	-0.08089
0.26	0.12684	4.69940	-0.17212	0.00000	0.76	0.08478	3.73406	-0.10000	-0.08422
0.27	0.12485	4.66168	-0.17207	0.00000	0.77	0.08459	3.72731	-0.10000	-0.08824
0.28	0.12352	4.62495	-0.17114	0.00000	0.78	0.08421	3.72078	-0.10000	-0.08878
0.29	0.12278	4.58921	-0.16993	0.00000	0.79	0.08383	3.71445	-0.10000	-0.08954
0.30	0.12238	4.55444	-0.16885	0.00000	0.80	0.08326	3.70832	-0.10000	-0.09315
0.31	0.12213	4.52063	-0.16813	0.00000	0.81	0.08278	3.70237	-0.10000	-0.09800
0.32	0.12191	4.48775	-0.16790	0.00000	0.82	0.08231	3.69657	-0.10000	-0.10246
0.33	0.12173	4.45579	-0.16820	0.00000	0.83	0.08174	3.69092	-0.10000	-0.10660
0.34	0.12163	4.42475	-0.16902	0.00000	0.84	0.08078	3.68540	-0.10000	-0.11089
0.35	0.12153	4.39459	-0.17027	0.00000	0.85	0.08002	3.67999	-0.10000	-0.11555
0.36	0.12134	4.36531	-0.17174	0.00000	0.86	0.07907	3.67468	-0.10000	-0.12019
0.37	0.12114	4.33690	-0.17310	0.00000	0.87	0.07792	3.66945	-0.10000	-0.12439
0.38	0.12105	4.30932	-0.17386	0.00000	0.88	0.07659	3.66429	-0.10000	-0.12768
0.39	0.12094	4.28258	-0.17345	0.00000	0.89	0.07506	3.65918	-0.10000	-0.12961
0.40	0.12066	4.25665	-0.17135	-0.11844	0.90	0.07316	3.65410	-0.09999	-0.12975
0.41	0.12036	4.23152	-0.16737	-0.11564	0.91	0.07106	3.64904	-0.09999	0.00000
0.42	0.12016	4.20717	-0.16184	-0.11122	0.92	0.06839	3.64395	-0.09999	0.00000
0.43	0.11997	4.18359	-0.15549	-0.10590	0.93	0.06532	3.63875	-0.09999	0.00000
0.44	0.11969	4.16077	-0.14919	-0.10095	0.94	0.06173	3.63318	-0.09999	0.00000
0.45	0.11937	4.13867	-0.14361	-0.09896	0.95	0.05734	3.62638	-0.09999	0.00000
0.46	0.11908	4.11730	-0.13901	-0.10016	0.96	0.05199	3.61543	-0.09999	0.00000
0.47	0.11878	4.09664	-0.13501	-0.09911	0.97	0.04496	3.59019	-0.09999	0.00000
0.48	0.11844	4.07666	-0.12994	-0.09374	0.98	0.03604	3.51566	-0.09999	0.00000
0.49	0.11793	4.05736	-0.12106	-0.08891	0.99	0.02558	3.27059	-0.09999	0.00000
0.50	0.11719	4.03871	-0.11022	-0.08668	1.00	0.00200	2.43507	-0.09998	0.00000

LITERATUR

- [1] T. R. Ashwin, A. McGordon, and P. A. Jennings. A mass transfer based variable porosity model with particle radius change for a lithium-ion battery. *Electrochimica Acta*, 232:203–214, 2017.
- [2] D. Bernardi. A general energy balance for battery systems. *Journal of The Electrochemical Society*, 132(1):5, 1985.
- [3] A. M. Bizeray, S. Zhao, S. R. Duncan, and D. A. Howey. Lithium-ion battery thermal-electrochemical model-based state estimation using orthogonal collocation and a modified extended kalman filter. *Journal of Power Sources*, 296:400–412, 2015.
- [4] Adrien M. Bizeray, David A. Howey, and Charles W. Monroe. Resolving a discrepancy in diffusion potentials, with a case study for li-ion batteries. *Journal of The Electrochemical Society*, 163(8):E223–E229, 2016.
- [5] Claudio Canuto. *Spectral methods in fluid dynamics*. Springer series in computational physics. Springer, Berlin u.a., corr. 3. print edition, 1988.
- [6] Roman Chapko and Rainer Kress. Rothe’s method for the heat equation and boundary integral equations. *Journal of Integral Equations and Applications*, 9(1):47–69, 1997.
- [7] Thanh-Son Dao, Chandrika P. Vyasarayani, and John McPhee. Simplification and order reduction of lithium-ion battery model based on porous-electrode theory. *Journal of Power Sources*, 198:329–337, 2012.
- [8] D. Di Domenico, Y. Creff, E. Prada, P. Duchêne, J. Bernard, and V. Sauvant-Moynot. A review of approaches for the design of li-ion bms estimation functions. *Oil & Gas Science and Technology – Revue d’IFP Energies nouvelles*, 68(1):127–135, 2013.
- [9] Domenico Di Domenico, Anna Stefanopoulou, and Giovanni Fiengo. Lithium-ion battery state of charge and critical surface charge estimation using an electrochemical model-based extended kalman filter. *Journal of Dynamic Systems, Measurement, and Control*, 132(6):061302, 2010.
- [10] Marc Doyle. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *Journal of The Electrochemical Society*, 140(6):1526, 1993.
- [11] Moses Ender. *Mikrostrukturelle Charakterisierung, Modellentwicklung und Simulation poröser Elektroden für Lithiumionenzellen*. Dissertation, 2014.
- [12] Lawrence C. Evans. *Partial differential equations*, volume Volume 19 of *Graduate studies in mathematics*. American Mathematical Society, Providence, Rhode Island, second edition, reprinted with corrections edition, 2015.
- [13] Guodong Fan, Ke Pan, Marcello Canova, James Marcicki, and Xiao Guang Yang. Modeling of li-ion cells for fast simulation of high c-rate and low temperature operations. *Journal of The Electrochemical Society*, 163(5):A666–A676, 2016.
- [14] Mohammed Farag, Haitham Sweity, Matthias Fleckenstein, and Saeid Habibi. Combined electrochemical, heat generation, and thermal model for large prismatic lithium-ion batteries in real-time applications. *Journal of Power Sources*, 360:618–633, 2017.
- [15] Thomas F. Fuller. Simulation and optimization of the dual lithium ion insertion cell. *Journal of The Electrochemical Society*, 141(1):1, 1994.
- [16] D. Funaro. *Polynomial Approximation of Differential Equations*, volume 8. Springer Berlin Heidelberg, Berlin, Heidelberg, 1992.
- [17] Maryam Ghalkhani, Farid Bahiraei, Gholam-Abbas Nazri, and Mehrdad Saif. Electrochemical? thermal model of pouch-type lithium-ion batteries. *Electrochimica Acta*, 2017.

- [18] David Gottlieb and Steven A. Orszag. *Numerical Analysis of Spectral Methods*. Society for Industrial and Applied Mathematics, 1977.
- [19] W. B. Gu and C. Y. Wang. Thermal-electrochemical modeling of battery systems. *Journal of The Electrochemical Society*, 147(8):2910, 2000.
- [20] Emre Gümüřsu, Özgür Ekici, and Murat Köksal. 3-d cfd modeling and experimental testing of thermal behavior of a li-ion battery. *Applied Thermal Engineering*, 120:484–495, 2017.
- [21] Ernst Hairer and Gerhard Wanner. *Stiff and differential-algebraic problems*, volume 14 of *Springer series in computational mathematics*. Springer, Berlin, 2., rev. ed., corrected printing, 1. softcover printing edition, 2010.
- [22] Ernst Hairer, Gerhard Wanner, and Syvert P. Nørsett. *Solving Ordinary Differential Equations I: Nonstiff Problems*, volume 8 of *Springer series in computational mathematics*. Springer Berlin Heidelberg, Berlin Heidelberg, second revised edition edition, 1993.
- [23] Xiaosong Hu, Shengbo Li, and Huei Peng. A comparative study of equivalent circuit models for li-ion batteries. *Journal of Power Sources*, 198:359–367, 2012.
- [24] Hsuan-Han Huang, Hsun-Yi Chen, Kuo-Chi Liao, Hong-Tsu Young, Ching-Fei Lee, and Jhen-Yang Tien. Thermal-electrochemical coupled simulations for cell-to-cell imbalances in lithium-iron-phosphate based battery packs. *Applied Thermal Engineering*, 123:584–591, 2017.
- [25] Jonas Lundgren. Splinefit - file exchange - matlab central, 2014.
- [26] Andreas Jossen and Wolfgang Weydanz. *Moderne Akkumulatoren richtig einsetzen: 36 Tabellen*. Ubooks-Verl., Neusäß, 2006.
- [27] David A. Kopriva. *Implementing Spectral Methods for Partial Differential Equations: Algorithms for Scientists and Engineers*. Scientific computation. Springer Netherlands, s.l., 1. aufl. edition, 2009.
- [28] Reiner Korthauer. *Handbuch Lithium-Ionen-Batterien*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [29] Karthikeyan Kumaresan, Godfrey Sikha, and Ralph E. White. Thermal model for a li-ion cell. *Journal of The Electrochemical Society*, 155(2):A164, 2008.
- [30] Wei Lai and Francesco Ciucci. Mathematical modeling of porous battery electrodes? revisit of newman’s model. *Electrochimica Acta*, 56(11):4369–4377, 2011.
- [31] Scott J. Moura, Nalin A. Chaturvedi, and Miroslav Krsti. Adaptive partial differential equation observer for battery state-of-charge/state-of-health estimation via an electrochemical model. *Journal of Dynamic Systems, Measurement, and Control*, 136(1):011015, 2014.
- [32] John Newman and Karen E. Thomas-Alyea. *Electrochemical Systems*. Wiley-Interscience, s.l., 3. aufl. edition, 2012.
- [33] John Newman and William Tiedemann. Porous-electrode theory with battery applications. *AIChE Journal*, 21(1):25–41, 1975.
- [34] Björn Nykvist and Måns Nilsson. Rapidly falling costs of battery packs for electric vehicles. *Nature Climate Change*, 5(4):329–332, 2015.
- [35] Volker Pluschke. Anwendung der rothe-methode auf eine quasilineare parabolische differentialgleichung. *Mathematische Nachrichten*, 114(1):105–121, 1983.
- [36] Matt Rennie and Ian Wood. Leading the charge: Battery storage is set to jump-start the clean energy transition while at the same time disrupting the traditional utility business model. *recai*, (49), 2017.
- [37] Shriram Santhanagopalan and Ralph E. White. Online estimation of the state of charge of a lithium ion cell. *Journal of Power Sources*, 161(2):1346–1355, 2006.

- [38] E. N. Sarmin and L. A. Chudov. On the stability of the numerical integration of systems of ordinary differential equations arising in the use of the straight line method. *USSR Computational Mathematics and Mathematical Physics*, 3(6):1537–1543, 1963.
- [39] William E. Schiesser. *The Numerical Method of Lines: Integration of Partial Differential Equations*. Elsevier Science, s.l., 2012.
- [40] William E. Schiesser and Graham W. Griffiths. *A compendium of partial differential equation models: Method of lines analysis with MATLAB*. Cambridge Univ. Press, Cambridge, 1. publ edition, 2009.
- [41] Lawrence F. Shampine and Mark W. Reichelt. The matlab ode suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, 1997.
- [42] Lawrence F. Shampine, Mark W. Reichelt, and Jacek A. Kierzenka. Solving index-1 daes in matlab and simulink. *SIAM Review*, 41(3):538–552, 1999.
- [43] Kandler A. Smith, Christopher D. Rahn, and Chao-Yang Wang. Model-based electrochemical estimation of lithium-ion batteries. In *2008 IEEE International Conference on Control Applications*, pages 714–719, Piscataway, NJ, 2008. IEEE Service Center.
- [44] Karl Strehmel, Rüdiger Weiner, and Helmut Podhaisky. *Numerik gewöhnlicher Differentialgleichungen: Nichtsteife, steife und differential-algebraische Gleichungen*. Vieweg & Teubner, Wiesbaden, 2., überarb. und erw. aufl. edition, 2012.
- [45] Lloyd N. Trefethen. *Spectral Methods in MATLAB*. Society for Industrial and Applied Mathematics, 2000.
- [46] Richard van Noorden. The rechargeable revolution: A better battery. *Nature*, 507(7490):26–28, 2014.
- [47] J. G. Verwer and J. M. Sanz-Serna. Convergence of method of lines approximations to partial differential equations. *Computing*, 33(3-4):297–313, 1984.
- [48] J. A. C. Weideman and S. C. Reddy. A matlab differentiation matrix suite. *ACM Transactions on Mathematical Software*, 26(4):465–519, 2000.
- [49] Billy Wu, Vladimir Yufit, Monica Marinescu, Gregory J. Offer, Ricardo F. Martinez-Botas, and Nigel P. Brandon. Coupled thermal–electrochemical modelling of uneven heat generation in lithium-ion battery packs. *Journal of Power Sources*, 243:544–554, 2013.
- [50] A. Zafarullah. Application of the method of lines to parabolic partial differential equations with error estimates. *Journal of the ACM*, 17(2):294–302, 1970.