

Optimale Steuerung und Modellreduktion  
von Maxwell-Gleichungen zur  
mikrowellenunterstützten Erwärmung

**Dissertation**

zur Erlangung des akademischen Grades

Dr. rer. nat.

eingereicht an der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

der Universität Augsburg

von

Christian Andreas Lochner

Augsburg, Juli 2020



---

Erstgutachterin: Prof. Dr. Tatjana Stykel  
Zweitgutachterin: Prof. Dr. Malte Peter  
Tag der mündlichen Prüfung: 13 Juli 2020

## Abstrakt

In der vorliegenden Dissertation zur mikrowellenunterstützten Erwärmung werden Optimalsteuerungsprobleme von Maxwell-Gleichungen in Bezug auf reduzierte Modelle mehrerer mit finiten Elementen diskretisierten Modellen betrachtet. Hierzu werden die zu betrachtenden partiellen Differentialgleichungen auf Basis der Maxwell-Gleichungen hergeleitet. Die Existenz und Eindeutigkeit der Lösung der hergeleiteten Gleichungen wird sowohl für die kontinuierliche als auch für die diskretisierte Gleichung bewiesen. Bevor die Optimalsteuerung zur mikrowellenunterstützten Erwärmung durchgeführt werden kann, benötigt es aufgrund der hohen Rechenlaufzeiten der finiten Elemente Simulation die Verwendung von Modellreduktionsalgorithmen. Hierzu wird die Reduzierte-Basis-Methode verwendet, um die Theorie der Lagrange-Systeme erweitert und dann anhand von drei ausgewählten Modellen auf ihre Funktionalität hin getestet. Daraufhin werden Steuerungsprobleme hergeleitet, die Existenz von Lösungen dieser gezeigt und sowohl die Konvergenztheorie als auch die Fehlertheorie für den diskreten und den reduzierten Fall durchgeführt. Zudem werden für die hergeleiteten Steuerungsprobleme die Existenz von Ableitungen bewiesen und die zu den Optimierungsproblemen zugehörigen adjungierten Probleme hergeleitet. Die Steuerungsprobleme werden dann wiederum auf ihre Funktionalität zur mikrowellenunterstützten Erwärmung, anhand der drei ausgewählten Modelle, bezüglich der Benchmark-Kriterien der Gleichmäßigkeit, des Energieeintrags und der gezielten Erwärmung getestet.

## Danksagung

Ich möchte mich recht herzlich bei meinen Betreuern Frau Prof. Tatjana Stykel und Herrn Prof. Malte Peter für die Möglichkeit der Promotion im Bereich Mathematik bedanken. In den vielen gemeinsamen und sehr lehrreichen Gesprächen konnten wir über einige sehr interessante Fragestellungen diskutieren. Insbesondere bedanke ich mich für die große Hilfe bei der Beantwortung komplexer Fragestellungen in der Analysis und Numerik.

Ein weiterer Dank gilt meiner Frau Tanja Lochner, die zum einen als Gesprächspartnerin bei Fragen zur Seite stand und mich zum anderen stets zu der Erarbeitung der Dissertation motivierte. Des Weiteren möchte ich mich bei ihr für die detaillierten Anmerkungen bei der Korrektur dieser Arbeit bedanken.

Ich danke meiner Mutter Kornelia und meinem Vater Peter für den steten Rückhalt in meinem Leben, sowie meinen Schwestern Sabine und Anna für die schönen gemeinsamen Stunden, sodass ich immer wieder neue Motivation schöpfen konnte. Ich danke meinen Schwiegereltern Elisabeth und Roland für die private Unterstützung und Hilfe während des Studiums und der Promotion. Zudem möchte ich mich bei meinen Freunden Michael und Jannik bedanken, die mich seit meiner Kindheit begleiten und in jeder Lebenslage an meiner Seite stehen.

Zuletzt möchte ich noch bei allen weiteren Freunden und Verwandten für die gemeinsamen Momente und deren Unterstützung bedanken.

## Inhaltsverzeichnis

<b>1. Einführung</b>	<b>1</b>
1.1. Motivation . . . . .	1
1.2. Vorgegangene Arbeiten . . . . .	2
1.3. Dissertationsarbeit . . . . .	3
1.4. Aufbau der Arbeit . . . . .	4
<b>2. Elektromagnetische Simulation</b>	<b>5</b>
2.1. Maxwell-Gleichungen . . . . .	5
2.1.1. Integraldarstellung . . . . .	6
2.1.2. Differentialdarstellung . . . . .	7
2.1.3. Zeitharmonische Darstellung . . . . .	8
2.1.4. Materialgleichungen . . . . .	10
2.1.5. Zeitharmonische Wellengleichung . . . . .	12
2.2. Funktionenräume . . . . .	14
2.3. Satz von Poynting . . . . .	17
2.4. Randbedingungen . . . . .	19
2.5. PDE elektromagnetischer Probleme mit Koaxial-Portanregung . . . . .	26
<b>3. Kontinuierliche und diskrete variationelle Formulierung</b>	<b>30</b>
3.1. Existenztheorie der kontinuierlichen Formulierung . . . . .	31
3.1.1. Kompakter Operator . . . . .	31
3.1.2. Existenz der Lösung . . . . .	45
3.2. Finite-Elemente-Methode . . . . .	53
3.3. Existenztheorie der diskreten Formulierung . . . . .	59
3.3.1. Interpolationsoperator . . . . .	60
3.3.2. Existenz der Lösung . . . . .	61
3.4. Finite Elemente Assemblierung . . . . .	63
<b>4. Reduzierte-Basis-Methode</b>	<b>67</b>
4.1. Grundlagen der RBM . . . . .	67
4.1.1. Parameterische Differentialgleichungen . . . . .	67
4.1.2. Offline-Online-Zerlegung . . . . .	73
4.1.3. Basisgenerierung des reduzierten Raumes . . . . .	74
4.1.4. Fehlerdarstellung . . . . .	75
4.1.5. Berechnung des Riesz-Repräsentanten . . . . .	77
4.1.6. Berechnung der inf-sup-Konstante . . . . .	78
4.2. RB-Lagrange-Systeme . . . . .	89
4.2.1. RB-Reduktion der Lagrange-Systeme . . . . .	90
4.2.2. Fehlerdarstellung . . . . .	99

4.2.3.	Berechnung des Riesz-Repräsentanten . . . . .	100
4.2.4.	Berechnung der inf-sup-Konstante . . . . .	103
4.2.5.	Getrennte RB-Systeme . . . . .	107
4.3.	Numerische Ergebnisse . . . . .	110
4.3.1.	Modell I: Zwei Antennen, ein Objekt . . . . .	113
4.3.2.	Modell II: Vier Antennen, ein Objekt . . . . .	137
4.3.3.	Modell III: Vier Antennen, zwei Objekte . . . . .	149
4.3.4.	Zusammenfassung . . . . .	156
<b>5.</b>	<b>Optimalsteuerung elektromagnetischer Felder</b>	<b>157</b>
5.1.	Kontinuierliche Optimalsteuerung . . . . .	157
5.1.1.	Stetige Abhängigkeit der Lösung bezüglich der Steuerparameter . . . . .	158
5.1.2.	Optimalsteuerungsprobleme . . . . .	164
5.1.3.	Erste Variation und Gâteaux-Differenzierbarkeit . . . . .	168
5.1.4.	Adjungierte Probleme . . . . .	179
5.1.5.	Getrennte Systeme . . . . .	192
5.2.	Diskrete Optimalsteuerungsprobleme . . . . .	201
5.2.1.	Fehlertheorie . . . . .	202
5.2.2.	$L^2$ -Feldoptimierung . . . . .	206
5.3.	RB Optimalsteuerungsprobleme . . . . .	220
5.4.	Steuerungsprobleme für den Erwärmungsprozess . . . . .	230
5.5.	Numerische Ergebnisse . . . . .	240
5.5.1.	Modell I: Zwei Antennen, ein Objekt . . . . .	242
5.5.2.	Modell II: Vier Antennen, ein Objekt . . . . .	274
5.5.3.	Modell III: Vier Antennen, zwei Objekte . . . . .	283
5.5.4.	Zusammenfassung . . . . .	295
<b>6.</b>	<b>Zusammenfassung und Ausblick</b>	<b>296</b>
6.1.	Zusammenfassung . . . . .	296
6.2.	Ausblick . . . . .	297
<b>A.</b>	<b>Anhang</b>	<b>300</b>
A.1.	Funktionenräume und Operatoren . . . . .	300
A.2.	Einbettungssätze . . . . .	303
A.3.	Existenztheorie . . . . .	305
	<b>Literaturverzeichnis</b>	<b>308</b>

## Symbolverzeichnis

$\partial_x f$	partielle Ableitung von $f$ nach $x$
$\nabla f$	Gradient von $f$
$\nabla \cdot f$	Divergenz-Operator angewendet auf $f$
$\nabla \times f$	Curl-Operator angewendet auf $f$
$\ \cdot\ _X$	Norm des Raumes $X$
$(\cdot, \cdot)_X$	$X$ -Skalarprodukt
$a \cdot b$	komplexes Standardskalarprodukt zweier Vektoren $a, b$
$ a $	Betragsnorm eines komplexen Vektors $a$
$ A $	Determinante der Matrix $A$
$ K $	Fläche oder Volumen des Gebietes $K$
$X_e, Y_e$	Funktionsraum
$X_N, Y_N$	Finite Elemente Raum
$X_N, Y_N$	Reduzierter Raum
$u_e$	Elemente der Funktionenräume $X_e$ oder $Y_e$
$u_N$	Elemente der Finiten Elementen Räume $X_N$ oder $Y_N$
$u_N$	Elemente des reduzierten Räume $X_N$ oder $Y_N$
$u^N$	Komplexer Koeffizientenvektor der Finiten Elemente Darstellung
$u^N$	Komplexer Koeffizientenvektor der reduzierten Basis-Darstellung
$[V_K], [V_{D,N}]$	Kontinuierliches und diskretes Variationsproblem auf den Seiten 30 und 59
$[V_{K,Q}], [V_{D,Q,N_L}]$	Kontinuierliches und diskretes $L^2$ -Variationsproblem auf den Seiten 209 und 217
$[V_{K,Div_l}]$	Kontinuierliches getrenntes Variationsproblem auf der Seite 192
$([O_k]_K)_{k=1}^3$	Kontinuierliche Optimierungsprobleme auf den Seiten 164, 165 und 209

$([O_k]_{K, \text{Div}})_{k=1}^3$	Kontinuierliche getrennte Optimierungsprobleme auf den Seiten 193 und 219
$([D_k]_{K, \text{Div}_i})_{k=1}^2$	Kontinuierliche adjungierte getrennte Probleme auf den Seiten 197 und 198
$([D_k]_K)_{k=1}^3$	Kontinuierliche adjungierte Probleme auf den Seiten 184, 185 und 213
$([O_k]_{D, \mathcal{N}})_{k=1}^2$	Diskrete Optimierungsprobleme auf der Seite 201
$[O]_{D, \mathcal{N}, \text{Div}}$	Allgemeines diskretes Optimierungsproblem für affine abhängige Probleme auf der Seite 229
$[O_3]_{D, \mathcal{N}_L}$	Diskretes Optimierungsproblem auf der Seite 218
$[O]_{D, \mathcal{N}}^{RG}, [O]_{D, \mathcal{N}}^{PG}$	Reduzierte Optimierungsprobleme auf der Seite 221
$[D]_{D, \mathcal{N}}^{RG}, [D]_{D, \mathcal{N}}^{PG}$	Reduzierte adjungierte Probleme auf der Seite 227

## Liste an Acronymen

<b>FEM</b>	Finite Elemente Methode
<b>PDE</b>	Partielle Differentialgleichung (partial differential equation)
<b>ODE</b>	Gewöhnliche Differentialgleichung (ordinary differential equation)
<b>RBM</b>	Reduced Basis Method
<b>PEC</b>	Perfekter elektrischer Leiter (perfect electric conducting)
<b>PMC</b>	Perfekter magnetischer Leiter (perfect magnetic conducting)
<b>IMP</b>	Impedanz
<b>SCM</b>	Successive Constraint Method

## Abbildungsverzeichnis

2.1. Querschnitt eines Koaxialkabels. . . . .	22
2.2. Modenfeld im Koaxialkabel. . . . .	22
3.1. Skizze der Modellmengen. . . . .	31
3.2. Skizze der Randmengen und Indexmengen mit $I_{\Sigma,1} = \{1, 4, 8, 9, 10\}$ und $I_{\Sigma,2} = \{2, 3, 5, 6, 7\}$ . . . . .	32
3.3. Skizze der sphärischen Umrandung der Modellmengen. . . . .	33
3.4. Skizze der Gebiete für die rechtsinversen Operatoren. . . . .	35
4.1. Quaderförmige Kavität. . . . .	110
4.2. Monopol-Antenne. . . . .	110
4.3. Darstellung des $S_{11}$ -Parameters in dB. . . . .	112
4.4. Geometrie des ersten Modells. . . . .	113
4.5. Geometrie des ersten Modells: zwei Antennen (links), ein zu erwärmendes Objekt (mitte), eine Schale (rechts). . . . .	113
4.6. Maße des Erwärmungsobjektes. . . . .	114
4.7. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	119
4.8. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase ohne Kontrolle der Kondition des reduzierten Systems. . . . .	120
4.9. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	121
4.10. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase . . . . .	122
4.11. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	122
4.12. Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline- Phase. . . . .	123
4.13. Maximaler relativer Fehler bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p)$ der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	124
4.14. Durchschnittlicher relativer Fehler bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p)$ der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	125
4.15. Minimaler relativer Fehler bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p)$ der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	125
4.16. Relativer Abstand bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p_{64})$ mit $p_{64} \in$ $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	126
4.17. Relativer Abstand bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p_{284})$ mit $p_{284} \in$ $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	127
4.18. Maximaler relativer Fehler bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p)$ der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	128
4.19. Maximaler relativer Fehler bezüglich der inf-sup-Konstanten $\beta_{N,c}^{approx}(p)$ der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	128

4.20. Maximaler relativer Fehler bezüglich der inf-sup-Konstanten $\beta_{\mathcal{N},c}^{approx}(p)$ der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	129
4.21. Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge $\mathbb{P}_{300}^{train}$ während der Offline-Phase. . . . .	130
4.22. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	131
4.23. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	132
4.24. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	133
4.25. Komplexer Betrag der ortsabhängigen elektromagnetischen Feldvektoren zu $p_{\max} \in \mathbb{P}_{100}^{test}$ . . . . .	135
4.26. Absoluter und relativer Fehler der ortsabhängigen elektromagnetischen Feldvektoren zu $p_{\max} \in \mathbb{P}_{100}^{test}$ . . . . .	136
4.27. Geometrie des zweiten Modells. . . . .	137
4.28. Geometrie des zweiten Modells: vier Antennen (links), ein zu erwärmendes Objekt (mitte), eine Schale (rechts). . . . .	138
4.29. Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge $\mathbb{P}_{400}^{train}$ während der Offline-Phase. . . . .	141
4.30. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	141
4.31. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	142
4.32. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	142
4.33. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	143
4.34. Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge $\mathbb{P}_{400}^{train}$ während der Offline-Phase. . . . .	143
4.35. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	144
4.36. Position und Nummerierung der Antennen des ersten (links) und zweiten (rechts) Modells. . . . .	145
4.37. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	146
4.38. Komplexer Betrag der ortsabhängigen elektromagnetischen Feldvektoren zu $p_{\max} \in \mathbb{P}_{100}^{test}$ . . . . .	147
4.39. Absoluter und relativer Fehler der ortsabhängigen elektromagnetischen Feldvektoren zu $p_{\max} \in \mathbb{P}_{100}^{test}$ . . . . .	148
4.40. Geometrie des dritten Modells. . . . .	149
4.41. Geometrie des dritten Modells: vier Antennen (links), zwei zu erwärmende Objekte (mitte), zwei Schalen (rechts). . . . .	149
4.42. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	153
4.43. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	153
4.44. Maximale Werte des relativen Fehlerschätzers während der Offline-Phase. . . . .	154
4.45. Maximaler relativer Fehler zu der Testparametermenge $\mathbb{P}_{100}^{test}$ . . . . .	155

5.1. Stehende Welle $\tilde{y}(t, x)$ zu verschiedenen Phasenzeitpunkten. . . . .	232
5.2. Kumulativ-iterative Rechenzeit der lokalen und globalen Optimierungsalgorithmen unter Verwendung der Probleme $[I_3]_1^l$ für $K_{opt} = 11$ und für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . . . . .	245
5.3. Kumulativ-iterative RSDHR- und MP-Werte der lokalen Optimierungsalgorithmen unter Verwendung der Probleme $[I_3]_1^l$ für $K_{opt} = 11$ und für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . . . . .	247
5.4. Kumulativ-iterative RSDHR- und MP-Werte der globalen Optimierungsalgorithmen unter Verwendung der Probleme $[I_3]_1^l$ für $K_{opt} = 11$ und für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . . . . .	248
5.5. Kumulativ-iterative Vergleichswerte der lokalen und globalen Optimierungsalgorithmen der elften Iteration für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . . . . .	249
5.6. RTM-Werte der lokalen und globalen Optimierungsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . . . . .	250
5.7. Kumulativ-iterative Rechenzeit der lokalen und globalen Optimierungsalgorithmen unter Verwendung der Probleme $[I_3]_1^l$ für $K_{opt} = 11$ und für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . . . . .	251
5.8. Kumulativ-iterative RSDHR- und MP-Werte der lokalen Optimierungsalgorithmen unter Verwendung der Probleme $[I_3]_1^l$ für $K_{opt} = 11$ und für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . . . . .	252
5.9. Kumulativ-iterative RSDHR- und MP-Werte der globalen Optimierungsalgorithmen unter Verwendung der Probleme $[I_3]_1^l$ für $K_{opt} = 11$ und für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . . . . .	253
5.10. Kumulativ-iterative Vergleichswerte der lokalen und globalen Optimierungsalgorithmen der elften Iteration für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . . . . .	254
5.11. RTM-Werte der lokalen und globalen Optimierungsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . . . . .	255
5.12. Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	256
5.13. Darstellung der S-Parametermatrix in dB. (Modell I) . . . . .	257
5.14. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_1^{rel}]_k$ , $[I_2^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	258
5.15. Kumulativ-iterative RSDHR- und MP-Werte der Steuerungsprobleme $[I_1^{rel}]_1$ und $[I_2^{rel}]_1$ für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	259
5.16. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_3]_k$ , $[I_3^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	262
5.17. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_4]_k$ , $[I_4^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	263

5.18. Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems $[I_3]_1$ und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	264
5.19. Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	264
5.20. RSDHR- und MP-Werte des Steuerparametervektors $p_S \in \mathbb{P}_S$ für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	265
5.21. Hyperfläche des steuerungsabhängigen kumulativ-iterativen RSDHR-Wertes für den Vergleichsalgorithmus RSDHR-Optimierung. (Modell I) . . . . .	266
5.22. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_1]_k, [I_1^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	268
5.23. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_2]_k, [I_2^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	269
5.24. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_3]_k, [I_3^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	270
5.25. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_4]_k, [I_4^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	271
5.26. Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems $[I_3]_1$ und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	272
5.27. Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $1 \leq m \leq 4$ . (Modell II) . . . . .	274
5.28. Darstellung der S-Parametermatrix in dB. (Modell II) . . . . .	275
5.29. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_3]_k, [I_3^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $1 \leq m \leq 4$ . (Modell II) . . . . .	276
5.30. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_4]_k, [I_4^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $1 \leq m \leq 4$ . (Modell II) . . . . .	277
5.31. Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems $[I_3]_1$ und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1]$ , $1 \leq m \leq 4$ . (Modell II) . . . . .	278
5.32. Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $1 \leq m \leq 4$ . (Modell II) . . . . .	279
5.33. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_3]_k, [I_3^{rel}]_k$ , $k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1$ , $1 \leq m \leq 4$ . (Modell II) . . . . .	280

5.34. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II) . . . . .	281
5.35. Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems $[I_3]_1$ und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II) . . . . .	282
5.36. Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III) . . . . .	283
5.37. Darstellung der S-Parametermatrix in dB. (Modell III) . . . . .	284
5.38. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III) . . . . .	285
5.39. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III) . . . . .	286
5.40. Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems $[I_3]_1$ und der Vergleichsalgorithmen für den Fall a) $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III) . . . . .	287
5.41. Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III) . . . . .	287
5.42. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III) . . . . .	288
5.43. Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III) . . . . .	289
5.44. Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems $[I_3]_1$ und der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III) . . . . .	290
5.45. Kumulativ-iterative RSDHR- und MPDR-Werte der Vergleichsalgorithmen für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . . . . .	292
5.46. Kumulativ-iterative RSDHR-, HTQ und MP-Werte der Vergleichsalgorithmen im warmen und kalten Gebiet für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . . . . .	293
5.47. Kumulativ-iterative RSDHR- und MPDR-Werte des Steuerungsproblems $[I_{3,mpdr}]_1$ für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . . . . .	294
5.48. Kumulativ-iterative RSDHR-, HTQ und MP-Werte des Steuerungsproblems $[I_{3,mpdr}]_1$ im warmen und kalten Gebiet für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . . . . .	294

## Tabellenverzeichnis

1.	RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der Fehlerschätzer-Schranken zu Abbildung 4.23. . . . .	132
2.	RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der maximalen relativen Fehler-Schranken zu Abbildung 4.24. . . . .	133
3.	RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der Fehlerschätzer-Schranken zu Abbildung 4.35. . . . .	145
4.	RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der echten maximalen relativen Fehler-Schranken zu Abbildung 4.37. . . . .	146
5.	RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der Fehlerschätzer-Schranken zu Abbildung 4.44 . . . . .	154
6.	RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der echten maximalen relativen Fehler-Schranken zu Abbildung 4.45. . . . .	155
7.	RB-Dimensionen des vollen und der getrennten Systeme für die Verwendung in der Optimalsteuerung. . . . .	242
8.	Durchschnittliche Rechenzeit der einzelnen Bausteine für die Berechnung der Kostenfunktionale $Z_k$ , $k = 1, 2, 3$ , in Sekunden. . . . .	242
9.	Durchschnittliche Rechenzeit der einzelnen Bausteine für die Berechnung der Ableitung der Kostenfunktionale $\nabla Z_k$ , $k = 1, 2, 3$ , bei Verwendung der Ableitungen $\nabla \mathcal{E}_N$ und $\nabla \mathcal{P}_{N_L}$ in Sekunden. . . . .	242
10.	Durchschnittliche Rechenzeit der einzelnen Bausteine für die Berechnung der Ableitung der Kostenfunktionale $\nabla Z_k$ , $k = 1, 2, 3$ , bei Verwendung der adjungierten Lösungen $\lambda$ in Sekunden. . . . .	243
11.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>1</sub> ] für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	258
12.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>2</sub> ] für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	259
13.	Optimale Steuerparameter des iterativen Steuerungsproblems (SP) $[I_1^{rel}]_1^l$ , $1 \leq l \leq 11$ , für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	260
14.	Optimale Steuerparameter des iterativen Steuerungsproblems (SP) $[I_2^{rel}]_1^l$ , $1 \leq l \leq 11$ , für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	260
15.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>3</sub> ] für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	262
16.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} \in [0, 1]$ , $m = 1, 2$ . (Modell I) . . . . .	263
17.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>1</sub> ] für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	268
18.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>2</sub> ] für den Fall $P_{in,m} = 1$ , $m = 1, 2$ . (Modell I) . . . . .	269

---

19.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>3</sub> ] für den Fall $P_{in,m} = 1, m = 1, 2$ . (Modell I) . . . . .	270
20.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} = 1, m = 1, 2$ . (Modell I) . . . . .	271
21.	Optimale Steuerparameter der iterativen Steuerungsprobleme (SP) bei Verwendung von Referenzfeld $E_{V,\mathcal{N}} = \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]$ , bzw. $Q_{V,\mathcal{N}_L} = \mathcal{P}_{Q,\mathcal{N}_L}[\tilde{p}_S]$ , mit $\tilde{p}_S \in \mathbb{P}_S$ , für den Fall $P_{in,m} \in [0, 1], m = 1, 2$ (Modell I) . . . . .	273
22.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell II) . . .	276
23.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell II) . . .	277
24.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>3</sub> ] für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II) . . . . .	280
25.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II) . . . . .	281
26.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III) . . .	285
27.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III) . . .	286
28.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>3</sub> ] für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III) . . . . .	288
29.	Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP) [I <sub>4</sub> ] für den Fall $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III) . . . . .	289



---

# 1. Einführung

## 1.1. Motivation

Mit der Erforschung neuer Halbleitertechniken zur Erzeugung elektromagnetischer Strahlung und der einhergehenden Möglichkeit deren Phase, Frequenz und Amplitude zu steuern, hat in den letzten Jahren die Steuerung von elektromagnetischen Systemen für die Industrie an Bedeutung gewonnen. Ein großer Forschungsbereich ist die mikrowellenunterstützte Erwärmung von Objekten, wie beispielsweise von Harzsystemen oder zum Garen von Lebensmitteln. In der Arbeit [1] wird beispielsweise der Aushärtungsprozess von Epoxidharzsystemen unter Mikrowelleneinfluss betrachtet. Dabei konnte festgestellt werden, dass eine nicht optimal gesteuerte Mikrowellenerwärmung eine schlechte Auswirkung auf die ausgehärteten Materialeigenschaften besitzt. Jedoch bietet die mikrowellenunterstützte Erwärmung einige Vorteile in diesem Bereich, denn im Gegensatz zur konventionellen Erwärmung mit Heißluft, welche das Objekt von Außen nach Innen erwärmt, ist der Erwärmungsprozess bei Mikrowellen von Innen nach Außen. Dies ermöglicht eine deutlich schnellere Erwärmung und beschleunigt dabei den Aushärtungsprozess des Epoxidharzes. Um mehr Wissen über den Erwärmungsprozess mit Mikrowellen zu erlangen, werden in dieser Arbeit elektromagnetische Modelle mit komplexer Geometrie betrachtet, welche auf den partiellen Differentialgleichungen (PDEs) der Maxwell-Theorie basieren. Diese PDEs können verwendet werden, um Optimierungsprobleme zur mikrowellenunterstützten Erwärmung in Abhängigkeit der elektromagnetischen Felder zu definieren. Die dafür notwendigen Berechnungen sind aufgrund der Komplexität der Modelle sehr rechenintensiv und eignen sich nicht für eine in Echtzeit funktionierende optimale Steuerung. Eine Abhilfe dafür bietet die Modellreduktion, welche die Rechenzeiten durch Vernachlässigung redundanter Informationen verkürzen kann. Dementsprechend ist der Bereich der Modellreduktion innerhalb der letzten Jahre zu einem wichtigen Forschungsbereich innerhalb der Mathematik geworden. Zur Steuerung elektromagnetischer Felder wird dabei die Frequenz, die Phase und die Leistung verwendet, weshalb sich parametrische Modellreduktionsalgorithmen für die Generierung der reduzierten Modellen eignen. Eine der am häufigsten verwendeten Methoden zur parametrischen Modellreduktion stellt die Reduzierte-Basis-Methode (RBM) dar, die entweder auf das direkte Eingang-Ausgang-Verhalten der Parameter oder auf das Eingang-Zustand-Verhalten reduziert. Wenn das reduzierte Modell das gewünschte Verhalten gut approximiert, dann lässt sich mit diesem das Optimalsteuerungsproblem effektiv und schnell berechnen.

Um das entstandene Optimalsteuerungsproblem in der Realität verwenden zu können, muss das mathematische Modell den realen Prozess gut approximieren. Hierfür kann die erweiterte Reduzierte-Basis-Element-Methode (RBEM) verwendet werden, denn mit affinen Transformationen lassen sich geometrische Parameter in das Modell einarbeiten. Diese geometrischen Parameter können dann als Regelgröße dienen, um Simulationsdaten an gemessene Daten anzugleichen.

### 1.2. Vorangegangene Arbeiten

Im Bereich der Modellreduktion für elektromagnetische Probleme gibt es neben der RBM noch weitere Methoden, wie Modale Approximation, Asymptotische Wellenauswertung, Balanciertes Abschneiden und Krylov-Unterraum Methoden für lineare Probleme und Proper Orthogonal Decomposition für nichtlineare Systeme. Eine Übersicht der verschiedenen Methoden gibt es in [2, 3] und auf der Internetseite MOR-Wiki<sup>1</sup>.

Die in dieser Arbeit verwendete RBM wurde erstmals am Anfang der achtziger Jahre im Bereich der nichtlinearen Strukturanalyse in den Arbeiten [4, 5, 6] beschrieben. Nach und nach wurde die Methode dann auf weitere Bereiche, wie beispielsweise der statischen Wärmeleitungsgleichung in [7] oder nicht komprimierbarer Fluide [8], erweitert. Zeitgleich wurde auch intensiv an einer effektiven Fehlertheorie der RBM [9] geforscht, denn zu diesem Zeitpunkt war nicht klar, welche Approximationsgüte das generierte RB Modell besitzt. Diese Forschung resultierte in zuverlässigen Schranken an den Ausgang [10] um die Jahrtausendwende und wurde in den darauffolgenden Jahren [11] erweitert. Zusätzlich dazu wurden Fehlerschranken auch für nichtkoerzive Probleme ausgearbeitet [12, 13]. Eine ausführlichere Zusammenfassung der Historie ist dem Buch [14] zu entnehmen, welches sich gut als einführende Lektüre in die Thematik der RBM eignet. Die Entwicklung von Fehlerschranken für nichtkoerzive partielle Differentialgleichungen ermöglichte die Verwendung der RBM im Bereich der zeitharmonischen Maxwell-Gleichungen [15, 16, 17, 18, 19]. Die existierenden Fehlerschranken und der damit einhergehenden effektiven Berechnung der reduzierten Modelle führte zur Verwendung der RBM im Bereich der optimalen Steuerungstheorie [20, 21, 22, 23].

Die optimale Steuerung besitzt eine weitaus längere Historie, die bis in die Antike zurückreicht. Für einige zählt das Jahr 1697 als Johann Bernoulli seine Lösung des Brachystochronen Problems veröffentlichte als Anfang der optimalen Steuerung, denn die Arbeit von Bernoulli war der Anreiz für einige berühmte Mathematiker sich mit ähnlichen Problemen zu beschäftigen. Hervorzuheben sind dabei Leonhard Euler als Schüler Bernoullis und Joseph-Louis Lagrange, die aufgrund der Arbeiten von Bernoulli einige generelle Techniken zur Betrachtung von Optimierungsproblemen entwickelten. Bis ins zwanzigste Jahrhundert hinein arbeiteten viele Mathematiker im Bereich der Optimierung und der Variationsrechnung. Für viele ist hingegen die Arbeit [24] von Lew Semjonowitsch Pontryagin der eigentliche Beginn der heutigen optimalen Steuerungstheorie, denn in dieser Arbeit wurden erstmals dynamische Bedingungen an die Lösungsmenge gestellt. Als Standardwerk für optimale Steuerung von linearen Gleichungen und konvexen Zielfunktionalen unter der Nebenbedingung von partiellen Differentialgleichungen dient das Werk [25] von Joseph-Louis Lions. Eine allgemeine Einführung in die Theorie von optimaler Steuerung bieten die Bücher [26] und [27] mehrerer Autoren.

---

<sup>1</sup>[https://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Main\\_Page](https://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Main_Page)

### 1.3. Dissertationsarbeit

Der Hauptschwerpunkt dieser Arbeit liegt in der Theorie von optimalen Steuerungssystemen elektromagnetischer Prozesse. Hierzu wird eine partielle Differentialgleichung zur Beschreibung eines geschlossenen Systems hergeleitet und die Existenz- und Eindeutigkeitsbeweise auf Basis der Arbeit [28] von Salim Meddahi auf die hergeleiteten Gleichungen erweitert. Die partiellen Differentialgleichungen werden dann in Kombination mit der Simulationssoftware COMSOL Multiphysics<sup>2</sup> verwendet, um komplexe geometrische Probleme zur mikrowellenunterstützten Erwärmung von Epoxidharzen aufzubauen. Dafür bietet COMSOL eine Vielzahl an Möglichkeiten zur einfachen Behandlung komplexer CAD-Modelle mithilfe einer strukturierten GUI. Neben der einfachen Erstellung von Finite-Elemente-Methode (FEM) basierenden elektromagnetischen Simulationsmodellen mit vordefinierten Randbedingungen lassen sich auch die Variationsformulierungen der verschiedenen Bausteine an die eigenen Anforderungen anpassen. Mit der von COMSOL bereitgestellten Schnittstelle zu MATLAB<sup>3</sup> wird eine Schnittstellenkommunikation für die Assemblierung der FEM Matrizen aufgebaut. Damit gibt es die Möglichkeit der Kombination von RBM mit der Simulationssoftware COMSOL. Für die Anwendung der RBM wird als Grundgerüst die Programmbibliothek RBmatlab<sup>4</sup> verwendet. Aufgrund der betrachteten PDE, der Datendarstellung in COMSOL, MATLAB spezifischen Eigenschaften und der Dimensionsgröße der Probleme muss der Programmcode der Bibliothek RBmatlab erweitert, ergänzt oder zum Teil neu geschrieben werden. Wegen der von COMSOL sehr allgemein aufgestellten Theorie zum Lösen von PDEs wird die RB Theorie um die Theorie der Lagrange-Systeme erweitert, mit der auch allgemeiner definierte PDEs mit COMSOL betrachtet werden können. Des Weiteren wird die Theorie von optimalen Steuerungen auf die Steuerungen von elektromagnetischen Feldern und im Besonderen auf die mikrowellenunterstützte Erwärmung angewandt und die entsprechende Theorie dazu entwickelt. Darunter fällt die Definition von Optimalsteuerungsproblemen, die Herleitung von Ableitungen und den adjungierten Problemen sowie die Fehlertheorie zwischen den für FEM diskretisierten und den mit der RBM reduzierten Optimalsteuerungsproblemen. Diese Theorie wird dann auf drei ausgewählte Probleme angewandt und auf entsprechende Eigenschaften untersucht.

---

<sup>2</sup><https://www.comsol.de/>

<sup>3</sup><https://de.mathworks.com>

<sup>4</sup><https://www.morepas.org/software/rbmatlab/>

#### 1.4. Aufbau der Arbeit

In Kapitel 2 wird eine allgemeine Einführung in die Maxwell-Theorie gegeben, die notwendigen Funktionenräume eingeführt, die Wohldefiniertheit von Operatoren auf dem Rand des Gebietes diskutiert und daraus die PDE zur Betrachtung elektromagnetischer geschlossener Systeme hergeleitet. Für diese PDE betrachten wir dann in Kapitel 3 die Existenz und Eindeutigkeit des kontinuierlichen Variationsproblems und des FEM diskretisierten Variationsproblems. Kapitel 4 widmet sich der Betrachtung der RBM, wofür wir in Abschnitt 4.1 die allgemeine Theorie für inf-sup stabile Sesquilinearformen angeben und in Abschnitt 4.2 um die Theorie der Lagrange-Systeme erweitern. Die verschiedenen Varianten der Generierung der reduzierten Modelle testen wir dann in Abschnitt 4.3 auf deren Funktionalität und Eigenschaften. Abschließend betrachten wir die Theorie der Optimalsteuerung elektromagnetischer Felder und im Besonderen in Bezug auf die mikrowellenunterstützte Erwärmung in Kapitel 5. Dabei betrachten wir in Abschnitt 5.1 die kontinuierliche und in Abschnitt 5.2 die diskrete Optimalsteuerungstheorie. Die Optimalsteuerung erweitern wir daraufhin mit der in Abschnitt 4.2 hergeleiteten RB-Theorie in Abschnitt 5.3. Unter Verwendung der RB-Modelle aus Abschnitt 4.3 untersuchen wir dann in Abschnitt 5.4 eine Reihe von verschiedenen Optimierungsalgorithmen auf deren Verhalten und deren Effektivität zur Bestimmung von optimalen Steuerungen. Numerische Ergebnisse werden in Abschnitt 5.5 vorgestellt.

---

## 2. Elektromagnetische Simulation

Als Grundlage für elektromagnetische Modelle dienen die Maxwell-Gleichungen, die James Clerk Maxwell erstmals in seinem 1864 veröffentlichten Paper „A Dynamical Theory of the Electromagnetic Field“ [29] beschrieb. In diesem Kapitel wird neben einer allgemeinen Einführung in die Theorie elektromagnetischer Felder ein Modell zur Approximation geschlossener Systeme mit koaxialer Anregung hergeleitet.

### 2.1. Maxwell-Gleichungen

Die Maxwell-Gleichungen werden aus der Betrachtung der Wechselwirkungen von Punktladungen mithilfe des Satzes von Gauß und des Satzes von Stokes [30] hergeleitet. Um diese Gleichungen anzugeben, definieren wir die dazu notwendigen Vektorfelder.

**Definition 2.1.** Sei  $\Omega \subset \mathbb{R}^3$  offen. Gegeben seien

- das elektrische Feld  $E : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{V}{m}]$ ,
- die elektrische Flussdichte  $D : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{As}{m^2}]$ ,
- das magnetische Feld  $H : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{A}{m}]$ ,
- die magnetische Flussdichte  $B : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{Vs}{m^2}]$ ,
- die Ladungsdichte  $\rho : [0, T] \times \Omega \rightarrow \mathbb{R}$  mit der physikalischen Einheit  $[\frac{As}{m^3}]$ ,
- die elektrische Stromdichte  $J : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{A}{m^2}]$ ,
- die Leitungsstromdichte  $J_l : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{A}{m^2}]$ ,
- die Verschiebungsstromdichte  $J_v : [0, T] \times \Omega \rightarrow \mathbb{R}^3$  mit der physikalischen Einheit  $[\frac{A}{m^2}]$ .

Die elektrische Stromdichte  $J = J_l + J_v$  ist über die Leitungsstromdichte  $J_l$  und die Verschiebungsstromdichte  $J_v$  gegeben.

Eine Erläuterung zur Herleitung der Maxwell-Gleichungen durch die Betrachtung von Punktladungen kann in [31] nachgelesen werden.

### 2.1.1. Integraldarstellung

Unter den Maxwell-Gleichungen werden vier Gleichungen verstanden, die elektromagnetische Kopplungen beschreiben. Die erste Gleichung heißt *Gaußsches Gesetz* und beschreibt die Gleichheit des elektrischen Flusses  $D$  an der Oberfläche  $\partial V$  eines Volumens  $V$  und der elektrischen Ladung  $\rho$  innerhalb des Volumens

$$\int_{\partial V} D \cdot n \, d\Gamma = \int_V \rho \, d\Omega \quad (2.1)$$

für ein beliebiges messbares Volumen  $V$  mit stückweise stetigen Rand  $\partial V$  und zugehörigen äußeren Normalenvektor  $n$ . Die Gleichung

$$\int_{\partial V} B \cdot n \, d\Gamma = 0 \quad (2.2)$$

wird *Gaußsches Gesetz des Magnetismus* genannt und gibt die Quellenfreiheit des magnetischen Flusses  $B$  an. Im Gegensatz zu (2.1) ist die rechte Seite von (2.2) gleich Null, da keine magnetischen Monopole existieren. Ein magnetischer Monopol ist ein theoretischer Magnet mit nur einem Pol. Während (2.1) keine magnetische und (2.2) keine elektrische Einflüsse beinhalten, beschreiben die nächsten beiden Gleichungen die Abhängigkeiten voneinander. Der Zusammenhang von elektrischen und magnetischen Kräften ist die Lorentz-Kraft, die im Bereich der Induktion eine Rolle spielt. Dementsprechend wird für eine hinreichend reguläre zweidimensionale Fläche  $S \subset V$  mit Normalenvektor  $n$  und Rand  $\partial S$  mit Tangentialvektor  $\tau$  die Integralgleichung

$$\int_S \partial_t B \cdot n \, d\Gamma + \int_{\partial S} E \cdot \tau \, ds = 0 \quad (2.3)$$

auch *Faraday'sches Induktionsgesetz* genannt. Es beschreibt die Zirkulation des elektrischen Feldes  $E$  bezüglich des Randes  $\partial S$  aufgrund der zeitlichen Änderung des magnetischen Flusses  $B$  durch die Fläche  $S$ . Die letzte Gleichung

$$\int_{\partial S} H \cdot \tau \, ds = \int_S \partial_t D \cdot n \, d\Gamma + \int_S J_l \cdot n \, d\Gamma \quad (2.4)$$

ist das *Ampère'sche Durchflutungsgesetz*. Diese beschreibt den Zusammenhang der magnetischen Zirkulation auf  $\partial S$  und der Summe aus der Leitungsstromdichte  $J_l$  und der zeitlichen Änderung des elektrischen Flusses  $D$ .

### 2.1.2. Differentialdarstellung

Um die differentielle Darstellung anzugeben, führen wir zunächst den Divergenz- und Rotations-Operator und deren Notation ein.

**Definition 2.2.** Der *Divergenzoperator*  $\operatorname{div} : C^1(\mathbb{C}^3, \mathbb{C}^3) \rightarrow C^0(\mathbb{C}^3, \mathbb{C})$  ist gegeben durch

$$E = \begin{pmatrix} E_{x_1} \\ E_{x_2} \\ E_{x_3} \end{pmatrix} \mapsto \operatorname{div}[E] = \nabla \cdot E = \frac{\partial E_{x_1}}{\partial x_1} + \frac{\partial E_{x_2}}{\partial x_2} + \frac{\partial E_{x_3}}{\partial x_3}.$$

Der *Rotationsoperator*  $\operatorname{curl} : C^1(\mathbb{C}^3, \mathbb{C}^3) \rightarrow C^0(\mathbb{C}^3, \mathbb{C}^3)$  ist gegeben durch

$$E = \begin{pmatrix} E_{x_1} \\ E_{x_2} \\ E_{x_3} \end{pmatrix} \mapsto \operatorname{curl}[E] = \nabla \times E = \begin{pmatrix} \frac{\partial E_{x_3}}{\partial x_2} - \frac{\partial E_{x_2}}{\partial x_3} \\ \frac{\partial E_{x_1}}{\partial x_3} - \frac{\partial E_{x_3}}{\partial x_1} \\ \frac{\partial E_{x_2}}{\partial x_1} - \frac{\partial E_{x_1}}{\partial x_2} \end{pmatrix}.$$

Durch Anwendung des Satzes von Gauß auf (2.1) und (2.2) und des Satzes von Stokes auf (2.3) und (2.4) ergeben sich die Maxwell-Gleichungen in differentieller Form:

$$\nabla \cdot D = \rho, \tag{2.5}$$

$$\nabla \cdot B = 0, \tag{2.6}$$

$$\partial_t B + \nabla \times E = 0, \tag{2.7}$$

$$-\partial_t D + \nabla \times H = J_l. \tag{2.8}$$

Wenden wir den Divergenzoperator auf (2.8) an und verwenden sowohl (2.5) als auch die Vektoridentität  $\nabla \cdot (\nabla \times H) = 0$ , dann ergibt sich die Ladungserhaltungsgleichung

$$0 = \nabla \cdot J_l + \partial_t \rho. \tag{2.9}$$

Mit dieser Gleichung lassen sich entweder (2.7), (2.8) und (2.5) oder (2.7), (2.8) und (2.9) unabhängig von den jeweiligen anderen Gleichungen betrachten. Innerhalb dieser Arbeit betrachten wir die zeitharmonische Abhängigkeit der einzelnen Vektorfelder in den Gleichungen (2.5)-(2.8).

### 2.1.3. Zeitharmonische Darstellung

Für die Zeitharmonizität gibt es unterschiedliche Darstellungen in der Literatur. In den Ingenieurwissenschaften wird die Zeigerform

$$E(t, x) = \operatorname{Re} \left( \hat{E}(\omega, x) e^{j\omega t} \right) \quad (2.10)$$

für die Zeitharmonizität eines Vektorfeldes  $E$  mit imaginärer Einheit  $j$  verwendet, während in der Mathematik und der Physik die Fouriertransformation angewandt wird, um die Vektorfelder vom Zeitbereich in den Frequenzbereich zu transformieren. Für die Definition der Fouriertransformation existieren viele verschiedene Möglichkeiten und die in dieser Arbeit verwendete Form wird aufgrund der beiden am weitesten verbreiteten zeitharmonischen Darstellungen der Maxwell-Gleichungen verwendet. Wir definieren die Fouriertransformation deshalb als

$$\hat{E}(\omega, x) = \int_{\mathbb{R}} E(t, x) e^{i\omega t} dt \quad (2.11)$$

mit der imaginären Einheit  $i$ . Hierbei ist

$$\omega = 2\pi f,$$

wobei  $f$  die Frequenz der zeitlich veränderlichen Vektorfelder darstellt. Je nachdem ob wir die Zeigerform oder die Fouriertransformation verwenden, verwenden wir  $j$  oder  $i$  als Indikator und imaginäre Einheit. Im Frequenzbereich lässt sich ein Vektorfeld relativ einfach in der Zeit verschieben. Wollen wir ein zum Zeitpunkt  $t + \tilde{t}$  betrachtetes Vektorfeld  $E(t + \tilde{t}, x)$  durch das Vektorfeld  $E(t, x)$  zum Zeitpunkt  $t$  im Frequenzbereich beschreiben und bezeichnen mit  $\hat{E}(\omega, x)$  die Fouriertransformierte zu  $E(t, x)$  und  $\tilde{E}(\omega, x)$  als die Fouriertransformierte zu  $E(t + \tilde{t}, x)$ , dann ergibt sich

$$\tilde{E}(\omega, x) = \hat{E}(\omega, x) e^{-i\omega \tilde{t}} \quad (2.12)$$

bei der Verwendung der Fouriertransformation und unter Verwendung der Zeigerform

$$\tilde{E}(\omega, x) = \hat{E}(\omega, x) e^{j\omega \tilde{t}}. \quad (2.13)$$

Definieren wir nun  $\tilde{E}(\omega, x)$  als die Fouriertransformierte der Zeitableitung  $\partial_t E(t, x)$ , dann folgt mit partieller Integration

$$\tilde{E}(\omega, x) = -i\omega \hat{E}(\omega, x) \quad (2.14)$$

bei der Verwendung der Fouriertransformation und unter Verwendung der Zeigerform

$$\tilde{E}(\omega, x) = j\omega \hat{E}(\omega, x). \quad (2.15)$$

Bei dem Vergleich der Zeitverschiebungen (2.12), (2.13) und der Transformation der Zeitableitungen (2.14), (2.15) fällt auf, dass die Gleichungen der Fouriertransformation (2.11) denen der Zeigerform (2.10) ähneln. Wir erhalten demnach die selben Gleichungen (2.12) und (2.14) der Fouriertransformation (2.11) bei Verwendung der Zeigerform

$$E(t, x) = \operatorname{Re} \left( \hat{E}(\omega, x) e^{-i\omega t} \right), \quad (2.16)$$

welche vor allem in der Physik verwendet wird. Aufgrund der  $2\pi$ -Periodizität der komplexen Exponentialfunktion  $e$ , betrachten wir nur Zeitverschiebungen  $\tilde{t} \in \left[0, \frac{1}{f}\right]$  und definieren die Phase

$$\theta = \omega \tilde{t},$$

für die gerade  $\theta \in [0, 2\pi]$  gilt. Damit erhalten wir die zeitharmonischen Maxwell-Gleichungen im Frequenzbereich

$$\nabla \cdot \hat{D} = \hat{\rho}, \quad (2.17)$$

$$\nabla \cdot \hat{B} = 0, \quad (2.18)$$

$$-i\omega \hat{B} + \nabla \times \hat{E} = 0, \quad (2.19)$$

$$i\omega \hat{D} + \nabla \times \hat{H} = \hat{J}_t \quad (2.20)$$

zu einem bestimmten Zeitpunkt  $t$  unter Verwendung der Fouriertransformation, beziehungsweise

$$\nabla \cdot \hat{D} = \hat{\rho}, \quad (2.21)$$

$$\nabla \cdot \hat{B} = 0, \quad (2.22)$$

$$j\omega \hat{B} + \nabla \times \hat{E} = 0, \quad (2.23)$$

$$-j\omega \hat{D} + \nabla \times \hat{H} = \hat{J}_t, \quad (2.24)$$

durch die Zeigerform.

### 2.1.4. Materialgleichungen

Die Einflüsse elektromagnetischer Effekte auf Materialien lassen sich durch die folgenden Materialparameter beschreiben:

- Permeabilität  $\mu$ , Maß für die Durchlässigkeit des magnetischen Feldes,
- Permittivität  $\varepsilon$ , Maß für die Durchlässigkeit des elektrischen Feldes,
- Konduktivität  $\sigma$ , Maß für die elektrische Leitfähigkeit.

Während die Konduktivität im Vakuum Null ist, sind die Permeabilität und Permittivität im Vakuum

$$\mu_0 = 4\pi \cdot 10^{-7} \frac{Vs}{Am}, \quad \varepsilon_0 = 8.854187817 \cdot 10^{-12} \frac{As}{Vm}$$

Naturkonstanten und dienen als Referenzwerte für jegliche Parameterwerte anderer Materialien. Um die Vektorfelder  $\hat{E}$  und  $\hat{H}$  mit den Flussdichten  $\hat{D}$  und  $\hat{B}$  zu verknüpfen führen wir zwei weitere Vektorfelder ein, wobei diese Transformation der Vektorfelder nur bei der Betrachtung nichtleitender Materialien gültig ist.

**Definition 2.3.** Sei  $\Omega \subset \mathbb{R}^3$  offen. Gegeben seien

- die elektrische Polarisation  $P_e : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{C}^3$ ,  $(\omega, x) \mapsto P_e(\omega, x, \hat{E}(\omega, x))$  mit der physikalischen Einheit  $[\frac{As}{m^2}]$ ,
- die magnetische Polarisation  $P_m : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{C}^3$ ,  $(\omega, x) \mapsto P_m(\omega, x, \hat{H}(\omega, x))$  mit der physikalischen Einheit  $[\frac{Vs}{m^2}]$ .

Die Definition 2.3 ist für die in dieser Arbeit betrachteten Modelle ausreichend, denn alle leitenden Materialien werden im späteren Verlauf durch Randbedingungen approximiert. Es ergeben sich die Beziehungen

$$\hat{D} = \varepsilon_0 \hat{E} + P_e, \tag{2.25}$$

$$\hat{B} = \mu_0 \hat{H} + P_m \tag{2.26}$$

für beliebige nichtleitende Materialien. Wir treffen nun eine erste Unterteilung in lineare und nichtlineare Materialien. Dabei heißt linear, dass  $P_e$  bzw.  $P_m$  linear von den jeweiligen Feldern  $\hat{E}$  bzw.  $\hat{H}$  abhängig sind. Wir erhalten für lineare Materialien

$$P_e = \varepsilon_0 \chi_e \hat{E}, \quad P_m = \mu_0 \chi_m \hat{H},$$

mit der elektrischen und magnetischen Suszeptibilität  $\chi_e$  und  $\chi_m$ . Dabei können  $\chi_e$  und  $\chi_m$  komplexwertig sein und hängen sowohl vom Ort  $x$  als auch der Frequenz  $\omega$

der elektromagnetischen Welle ab. Eine genaue Darstellung der elektrischen und magnetischen Suszeptibilität bezüglich der Frequenz und des Ortes hängt deshalb stark von dem betrachteten Material ab und wird hier nicht näher erläutert. Wir können somit den relativen Permittivitätsfaktor  $\varepsilon_r$  und Permeabilitätsfaktor  $\mu_r$  für lineare Materialien definieren.

**Definition 2.4.** Sei  $\Omega \subset \mathbb{R}^3$  offen und  $\mathbb{1}$  der  $3 \times 3$  Einheitsensor. Gegeben seien

- die relative Permittivität  $\varepsilon_r : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{C}^{3 \times 3}$ ,

$$(\omega, x) \mapsto \varepsilon_r(\omega, x) = \operatorname{Re}(\varepsilon_r(\omega, x)) + i \operatorname{Im}(\varepsilon_r(\omega, x)) = \mathbb{1} + \chi_e(\omega, x),$$

mit  $\operatorname{Re}(\varepsilon_r) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+^{3 \times 3}$  und  $\operatorname{Im}(\varepsilon_r) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_{\geq 0}^{3 \times 3}$ .

- die relative Permeabilität  $\mu_r : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{C}^{3 \times 3}$ ,

$$(\omega, x) \mapsto \mu_r(\omega, x) = \operatorname{Re}(\mu_r(\omega, x)) + i \operatorname{Im}(\mu_r(\omega, x)) = \mathbb{1} + \chi_m(\omega, x),$$

mit  $\operatorname{Re}(\mu_r) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+^{3 \times 3}$  und  $\operatorname{Im}(\mu_r) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_{\geq 0}^{3 \times 3}$ .

Dabei ergeben sich die Eigenschaften in Definition 2.4 aus der Darstellung der elektrischen und magnetischen Suszeptibilität, wenn für die Zeitharmonizität die Fouriertransformation angewandt wird. Im Falle der Zeigerform erhält man dieselbe Definition bis auf ein Minus im Imaginärteil, also

$$\begin{aligned} \varepsilon_r(\omega, x) &= \operatorname{Re}(\varepsilon_r(\omega, x)) - j \operatorname{Im}(\varepsilon_r(\omega, x)), \\ \mu_r(\omega, x) &= \operatorname{Re}(\mu_r(\omega, x)) - j \operatorname{Im}(\mu_r(\omega, x)). \end{aligned}$$

Wir können damit für lineare Materialien die Relationen

$$\hat{D} = \varepsilon_0 \varepsilon_r(\omega, x) \hat{E} = \varepsilon(\omega, x) \hat{E}, \quad (2.27)$$

$$\hat{B} = \mu_0 \mu_r(\omega, x) \hat{H} = \mu(\omega, x) \hat{H} \quad (2.28)$$

definieren und erhalten daraus für die Polarisierungen

$$P_e = \varepsilon_0(\varepsilon_r(\omega, x) - \mathbb{1})\hat{E}, \quad P_m = \mu_0(\mu_r(\omega, x) - \mathbb{1})\hat{H},$$

wobei  $\mathbb{1}$  der Einheitsensor bezeichnet. Des Weiteren können wir die Leitungsstromdichte  $\hat{J}_l = \hat{J}_c + \hat{J}_i$  als Summe konduktiver Stromdichte  $\hat{J}_c$  und Quellenstromdichte  $\hat{J}_i$  schreiben, wobei für isotrope Materialien das Ohm'sche Gesetz

$$\hat{J}_c = \sigma(x) \hat{E} \quad (2.29)$$

gilt. Für isotrope Materialien können wir zudem die tensorwertigen Funktionen  $\varepsilon_r$ ,  $\mu_r$  und  $\sigma$  als skalare Funktionen beschreiben. Des Weiteren gilt für inhomogene Materiali-

## 2. Elektromagnetische Simulation

---

en die örtliche Abhängigkeit der drei Parameter innerhalb des Materials, während diese für homogene Materialien im Ort konstant sind. Damit ergeben sich die Maxwell-Gleichungen bezüglich  $\hat{E}$  und  $\hat{H}$  für lineare, isotrope und homogene Materialien

$$\varepsilon_0 \nabla \cdot (\varepsilon_r(\omega) \hat{E}) = \hat{\rho}, \quad (2.30)$$

$$\mu_0 \nabla \cdot (\mu_r(\omega) \hat{H}) = 0, \quad (2.31)$$

$$-i\omega\mu_0\mu_r(\omega)\hat{H} + \nabla \times \hat{E} = 0, \quad (2.32)$$

$$i\omega\varepsilon_0\varepsilon_r(\omega)\hat{E} - \sigma\hat{E} + \nabla \times \hat{H} = \hat{J}_i \quad (2.33)$$

unter Verwendung der Fouriertransformation im Frequenzbereich, bzw.

$$\varepsilon_0 \nabla \cdot (\varepsilon_r(\omega) \hat{E}) = \hat{\rho}, \quad (2.34)$$

$$\mu_0 \nabla \cdot (\mu_r(\omega) \hat{H}) = 0, \quad (2.35)$$

$$j\omega\mu_0\mu_r(\omega)\hat{H} + \nabla \times \hat{E} = 0, \quad (2.36)$$

$$-j\omega\varepsilon_0\varepsilon_r(\omega)\hat{E} - \sigma\hat{E} + \nabla \times \hat{H} = \hat{J}_i, \quad (2.37)$$

mit der Zeigerform. Aufgrund der komplexeren Notation bei Betrachtung von anisotropen Materialien, werden wir im weiteren Verlauf nur isotrope, sowie lineare und homogene, Materialien betrachten und verwenden hierzu die Gleichungen (2.30)-(2.33). Die Erweiterung auf anisotrope Materialien folgt mit kleinen Anpassungen an die nachfolgende Theorie analog. Im weiteren Verlauf der Arbeit schreiben wir  $E$  und  $H$  für die Vektorfelder  $\hat{E}$  und  $\hat{H}$  im Frequenzbereich, um die Notation zu vereinfachen.

### 2.1.5. Zeitharmonische Wellengleichung

Die Gleichungen (2.33) und (2.37) lassen sich mit der Einführung effektiver komplexer relativer Permittivitäten

$$\tilde{\varepsilon}_r(\omega) = \operatorname{Re}(\varepsilon_r(\omega)) + i \left( \operatorname{Im}(\varepsilon_r(\omega)) + \frac{\sigma}{\omega\varepsilon_0} \right) \quad (2.38)$$

bzw.

$$\tilde{\varepsilon}_r(\omega) = \operatorname{Re}(\varepsilon_r(\omega)) - j \left( \operatorname{Im}(\varepsilon_r(\omega)) + \frac{\sigma}{\omega\varepsilon_0} \right) \quad (2.39)$$

umformen in

$$\nabla \times H = -i\omega\varepsilon_0\tilde{\varepsilon}_r(\omega)E + J_i$$

bzw.

$$\nabla \times H = j\omega\varepsilon_0\tilde{\varepsilon}_r(\omega)E + J_i.$$

Im weiteren Verlauf schreiben wir anstatt  $\tilde{\varepsilon}_r(\omega)$  wiederum  $\varepsilon_r(\omega)$ , da diese dieselben Eigenschaften wie in Definition 2.4 besitzt. Die Einführung der effektiven Permittivitäten (2.38) bzw. (2.39) hat dabei auch einen physikalischen Hintergrund, denn bei der Bestimmung des Imaginärteils der Permittivität vieler Materialien kann nicht zwischen dem Anteil der Polarisierung  $\text{Im}(\varepsilon_r(\omega))$  und der Leitfähigkeit  $\sigma$  unterschieden werden. Wir möchten nachfolgend eine Formulierung herleiten, sodass die elektrischen und magnetischen Felder unabhängig voneinander betrachtet werden können, was vor allem in der numerischen Betrachtung Vorteile mit sich bringt. Hierzu multiplizieren wir (2.32) bzw. (2.36) mit  $\frac{1}{\mu_0\mu_r(\omega)}$ , wenden den Rotationsoperator auf diese Gleichung an und multiplizieren (2.33) mit  $i\omega$ . Kombination der dadurch erhaltenen Gleichungen ergibt

$$\nabla \times \left( \frac{1}{\mu_0\mu_r(\omega)} \nabla \times E \right) - \omega^2 \varepsilon_0 \varepsilon_r(\omega) E = i\omega J_i, \quad (2.40)$$

bzw. unter (2.37) und (2.39)

$$\nabla \times \left( \frac{1}{\mu_0\mu_r(\omega)} \nabla \times E \right) - \omega^2 \varepsilon_0 \varepsilon_r(\omega) E = -j\omega J_i, \quad (2.41)$$

für das elektrische Feld  $E$ . Für das magnetische Feld  $H$  ergibt sich analog

$$\nabla \times \left( \frac{1}{i\omega \varepsilon_0 \varepsilon_r(\omega)} \nabla \times H \right) + i\omega \mu_0 \mu_r(\omega) H = \nabla \times \left( \frac{1}{i\omega \varepsilon_0 \varepsilon_r(\omega)} J_i \right), \quad (2.42)$$

bzw.

$$\nabla \times \left( \frac{1}{j\omega \varepsilon_0 \varepsilon_r(\omega)} \nabla \times H \right) + j\omega \mu_0 \mu_r(\omega) H = -\nabla \times \left( \frac{1}{j\omega \varepsilon_0 \varepsilon_r(\omega)} J_i \right), \quad (2.43)$$

je nach Definition der Zeitharmonizität. Im weiteren Verlauf der Arbeit wollen wir Optimallösungen von (2.40) mit  $J_i = 0$  zu gegebenen Zielfunktionalen bestimmen. Dabei bedeutet  $J_i = 0$ , dass keine Stromquelle mit Quellendichte  $J_i$  in unserem Modell existiert, es sich demnach um ein geschlossenes System handelt. Um das System anzuregen und die Existenz und Eindeutigkeit von Lösungen zu erhalten, benötigen wir noch verschiedene Randbedingungen. Doch zuvor führen wir die benötigten Funktionenräume und Spurooperatoren ein, um die Wohldefiniertheit der Randterme zu gewährleisten.

## 2.2. Funktionenräume

Es sei  $\Omega \subset \mathbb{R}^3$  offenes polyedrisches Gebiet mit Rand  $\Gamma$  und  $\bar{\Omega}$  seine abgeschlossene Hülle. Dabei kann der Rand  $\Gamma$  in  $N$  offene Flächen  $(\Gamma_j)_{1 \leq j \leq N}$  aufgesplittet werden. Es gilt demnach  $\Gamma = \bigcup_j \bar{\Gamma}_j$ . Für die in Anhang A.1 definierten Räumen, definieren wir zunächst die vektorwertigen Funktionenräume.

**Definition 2.5.** Sei  $(X, \|\cdot\|_X)$  ein normierter Raum, dann bezeichne  $[X]^3$  den vektorwertigen normierten Raum

$$[X]^3 = \{v = [v_1, v_2, v_3]^T : v_k \in X, k = 1, 2, 3\}$$

mit der Norm

$$\|v\|_{[X]^3} = (\|v_1\|_X^2 + \|v_2\|_X^2 + \|v_3\|_X^2)^{\frac{1}{2}}.$$

Damit definieren wir nun den natürlichen Raum für Lösungen  $E$  der Maxwell-Gleichungen (2.17)-(2.20).

**Definition 2.6** ( $H(\text{curl}; \Omega)$ -Räume). Es sei

$$H(\text{curl}; \Omega) = \left\{ u \in [L^2(\Omega)]^3 : \nabla \times u \in [L^2(\Omega)]^3 \right\}$$

der vektorwertige Raum mit Skalarprodukt

$$(u, v)_{H(\text{curl}; \Omega)} = (u, v)_{L^2(\Omega)} + \int_{\Omega} (\nabla \times u) \cdot (\nabla \times v) \, d\Omega \quad (2.44)$$

und

$$H(\text{div}; \Omega) = \left\{ u \in [L^2(\Omega)]^3 : \nabla \cdot u \in [L^2(\Omega)]^3 \right\}$$

der vektorwertige Raum mit Skalarprodukt

$$(u, v)_{H(\text{div}; \Omega)} = (u, v)_{L^2(\Omega)} + \int_{\Omega} (\nabla \cdot u) \cdot (\nabla \cdot v) \, d\Omega. \quad (2.45)$$

Die Räume  $H(\text{curl}; \Omega)$  bzw.  $H(\text{div}; \Omega)$  sind mit der durch die Skalarprodukte (2.44) bzw. (2.45) induzierten Normen Hilberträume, siehe [32]. Ebenso gilt nach [32], dass  $[L^2(\Gamma)]^3$  mit dem Skalarprodukt

$$(u, v)_{[L^2(\Gamma)]^3} = \int_{\Gamma} u \cdot v \, d\Gamma$$

ein Hilbertraum ist. Des Weiteren definieren wir zu dem vektorwertigen Tangentialraum

$$L_t^2(\Gamma) = \left\{ u \in [L^2(\Gamma)]^3 : u \cdot n = 0 \right\}, \quad (2.46)$$

wobei  $n$  der äußere Normalenvektor zu dem Rand  $\Gamma$  ist, zwei tangentielle Spurooperatoren.

**Definition 2.7.** Der Operator  $\pi_t : [\mathcal{D}(\overline{\Omega})]^3 \rightarrow L_t^2(\Gamma)$ , gegeben durch

$$u \mapsto n \times (u \times n)|_{\Gamma},$$

wird *tangentiale Komponente* genannt. Der Operator  $\gamma_t : [\mathcal{D}(\overline{\Omega})]^3 \rightarrow L_t^2(\Gamma)$ , gegeben durch

$$u \mapsto (u \times n)|_{\Gamma},$$

wird *tangentiale Spur* genannt.

Mit dem Spuroperator  $\gamma$  aus Satz A.10 und der Dichtheit von  $[\mathcal{D}(\overline{\Omega})]^3|_{\Gamma}$  in  $[L^2(\Gamma)]^3$  können die beiden Operatoren  $\pi_t$  und  $\gamma_t$  zu linearen stetigen Operatoren

$$\pi_t : [H^1(\Omega)]^3 \rightarrow H_{-}^{\frac{1}{2}}(\Gamma) = \left\{ u \in L_t^2(\Gamma) : u|_{\Gamma_l} \in H^{\frac{1}{2}}(\Gamma_l), l = 1, \dots, N \right\}$$

und

$$\gamma_t : [H^1(\Omega)]^3 \rightarrow H_{-}^{\frac{1}{2}}(\Gamma)$$

fortgesetzt werden. Für eine genauere Betrachtung des Raumes  $H_{-}^{\frac{1}{2}}(\Gamma)$  sei hier auf [33] verwiesen. Wir bezeichnen die Bildräume dieser Operatoren mit

$$H_{\perp}^{\frac{1}{2}}(\Gamma) = \gamma_t \left( [H^1(\Omega)]^3 \right) \quad (2.47)$$

und

$$H_{\parallel}^{\frac{1}{2}}(\Gamma) = \pi_t \left( [H^1(\Omega)]^3 \right). \quad (2.48)$$

Der Beweis der Hilbertraumeigenschaft der beiden Räume, sowie die Beschränktheit und die Surjektivität der Operatoren  $\gamma_t : [H^1(\Omega)]^3 \rightarrow H_{\perp}^{\frac{1}{2}}(\Gamma)$  und  $\pi_t : [H^1(\Omega)]^3 \rightarrow H_{\parallel}^{\frac{1}{2}}(\Gamma)$  kann in [33] nachgelesen werden. Des Weiteren sei

$$H_0(\text{curl}; \Omega) = \{ u \in H(\text{curl}; \Omega) : \gamma_t[u] = 0 \text{ auf } \Gamma \}.$$

Für  $\delta \in (0, 1)$  ist

$$H_{\parallel}^{\delta}(\Gamma) = \pi_t \left( \left[ H^{\delta + \frac{1}{2}}(\Omega) \right]^3 \right) \quad (2.49)$$

ein stetiger Operator, siehe [33]. Um tangentielle Operatoren auf  $H(\text{curl}; \Omega)$  und deren Bildräumen zu definieren, benötigen wir zum einen die Definition der Dualräume und zum anderen eine Definition des Divergenz- und Curl-Operators auf  $\Gamma$ . Zunächst widmen wir uns der Definition der Dualräume  $H_{\perp}^{-\frac{1}{2}}(\Gamma)$  und  $H_{\parallel}^{-\frac{1}{2}}(\Gamma)$  zu  $H_{\perp}^{\frac{1}{2}}(\Gamma)$  und  $H_{\parallel}^{\frac{1}{2}}(\Gamma)$

## 2. Elektromagnetische Simulation

---

hinsichtlich des Pivotraumes  $L_t^2(\Gamma)$ . Mit der Green'schen Formel

$$\int_{\Omega} (\nabla \times u) \cdot v - u \cdot (\nabla \times v) \, d\Omega = - \int_{\Gamma} \gamma_t(u) \cdot \pi_t(v) \, d\Gamma \quad (2.50)$$

für alle  $u, v \in [C^\infty(\bar{\Omega})]^3$ , und der Dichtheit von  $[C^\infty(\bar{\Omega})]^3$  in  $H(\text{curl}; \Omega)$  können  $\gamma_t$  und  $\pi_t$  zu beschränkten tangentialen Operatoren von  $H(\text{curl}; \Omega)$  auf  $H_{\perp}^{-\frac{1}{2}}(\Gamma)$  bzw.  $H_{\parallel}^{-\frac{1}{2}}(\Gamma)$  fortgesetzt werden. Auf die Definition des Divergenzoperators  $\text{div}_{\Gamma}$  und des Rotationsoperators  $\text{curl}_{\Gamma}$ , welche über lokale Lipschitzdarstellungen definiert sind, wird hier nicht genauer eingegangen und es wird hierfür auf [34] verwiesen. Aus der Kombination von Divergenz- und Rotationsoperator mit den tangentialen Operatoren ergibt sich

$$\begin{aligned} \text{div}_{\Gamma} [\gamma_t [u]] &= (\nabla \times u) \cdot n \in H^{-\frac{1}{2}}(\Gamma), \\ \text{curl}_{\Gamma} [\pi_t [u]] &= (\nabla \times u) \cdot n \in H^{-\frac{1}{2}}(\Gamma), \end{aligned}$$

und somit  $\text{div}_{\Gamma} (\gamma_t [u]) = \text{curl}_{\Gamma} (\pi_t [u])$  für  $u \in H(\text{curl}; \Omega)$ . Damit können wir ein wichtiges Resultat für tangentialen Operatoren von  $H(\text{curl}; \Omega)$  aus [34] angeben.

**Satz 2.8.** *Seien*

$$H^{-\frac{1}{2}}(\text{div}_{\Gamma}; \Gamma) = \left\{ u \in H_{\parallel}^{-\frac{1}{2}}(\Gamma) : \text{div}_{\Gamma}[u] \in H^{-\frac{1}{2}}(\Gamma) \right\} \quad (2.51)$$

und

$$H^{-\frac{1}{2}}(\text{curl}_{\Gamma}; \Gamma) = \left\{ u \in H_{\perp}^{-\frac{1}{2}}(\Gamma) : \text{curl}_{\Gamma}[u] \in H^{-\frac{1}{2}}(\Gamma) \right\}. \quad (2.52)$$

Dann sind

$$\gamma_t : H(\text{curl}; \Omega) \rightarrow H^{-\frac{1}{2}}(\text{div}_{\Gamma}; \Gamma) \quad (2.53)$$

und

$$\pi_t : H(\text{curl}; \Omega) \rightarrow H^{-\frac{1}{2}}(\text{curl}_{\Gamma}; \Gamma) \quad (2.54)$$

lineare, beschränkte und surjektive Operatoren und besitzen stetige Rechtsinverse. Zudem kann das  $[L^2(\Gamma)]^3$ -Skalarprodukt zum Dualitätsprodukt  $\langle \cdot, \cdot \rangle_{t, \Gamma}$  zwischen den Räumen  $H^{-\frac{1}{2}}(\text{div}_{\Gamma}; \Gamma)$  und  $H^{-\frac{1}{2}}(\text{curl}_{\Gamma}; \Gamma)$  erweitert werden.

*Bemerkung 2.9.* Auch die Tangentialoperatoren  $\gamma_t|_{\hat{\Gamma}}$  und  $\pi_t|_{\hat{\Gamma}}$ , die nur auf eine zusammenhängende Teilmenge  $\hat{\Gamma} \subset \Gamma$  abbilden, sind lineare stetige Operatoren.

Der Beweis von Satz 2.8 ist in [34, Theorem 4.1] gegeben und Bemerkung 2.9 wird in [33, Theoreme 3.15 und 3.16] bewiesen. Wir können nun (2.50) zu einer Formel für

$H(\text{curl}; \Omega)$  erweitern

$$\int_{\Omega} (\nabla \times u) \cdot v - u \cdot (\nabla \times v) \, d\Omega = - \int_{\Gamma} \gamma_t [u] \cdot \pi_t [v] \, d\Gamma, \quad u, v \in H(\text{curl}; \Omega). \quad (2.55)$$

### 2.3. Satz von Poynting

Bevor wir Randbedingungen für elektromagnetische Probleme herleiten, leiten wir noch eine Erhaltungsgleichung für elektromagnetische Leistungen her. Dazu betrachten wir das komplexe Skalarprodukt aus magnetischem Feld  $H$  und der Gleichung (2.32), also

$$(-i\omega\mu_0\mu_r(\omega)H + \nabla \times E) \cdot H = 0.$$

Dies ist äquivalent zu der Gleichung

$$(\nabla \times E) \cdot H = i\omega\mu_0\mu_r(\omega)|H|^2, \quad (2.56)$$

wobei  $|H|$  der komplexe Betrag ist. Im Folgenden bezeichnen wir mit  $|\cdot|$  verschiedene mathematische Operationen, wobei dies aus dem jeweiligen Kontext heraus klar wird. Analog erhalten wir die Gleichung

$$E \cdot (\nabla \times H) = E \cdot J_i + i\omega\varepsilon_0\overline{\varepsilon_r(\omega)}|E|^2, \quad (2.57)$$

wobei  $\overline{\varepsilon_r(\omega)}$  das komplex Konjugierte zu  $\varepsilon_r(\omega)$ , gegeben durch (2.38), darstellt. Mit der Vektoridentität

$$\nabla \cdot (a \times \bar{b}) = (\nabla \times a) \cdot b - a \cdot (\nabla \times b),$$

ergibt sich mit (2.56) und (2.57) die Gleichung

$$\nabla \cdot (E \times \bar{H}) = i\omega\mu_0\mu_r(\omega)|H|^2 - E \cdot J_i - i\omega\varepsilon_0\overline{\varepsilon_r(\omega)}|E|^2. \quad (2.58)$$

Wir integrieren (2.58) über ein zusammenhängendes Gebiet  $\Omega \subset \mathbb{R}^3$  mit glattem Rand  $\Gamma = \partial\Omega$  und wenden den Satz von Gauß an, womit sich

$$\int_{\Gamma} (E \times \bar{H}) \cdot n \, d\Gamma = i\omega\mu_0 \int_{\Omega} \mu_r(\omega)|H|^2 \, d\Omega - \int_{\Omega} E \cdot J_i \, d\Omega - i\omega\varepsilon_0 \int_{\Omega} \overline{\varepsilon_r(\omega)}|E|^2 \, d\Omega$$

ergibt, wobei  $n$  der äußere Normalenvektor ist. Umstrukturieren dieser Gleichung und Aufspalten von  $\mu_r(\omega)$ ,  $\varepsilon_r(\omega)$  in Real- und Imaginärteil nach Definition 2.4 ergibt dann

die Erhaltungsgleichung

$$\begin{aligned} \frac{1}{2} \int_{\Omega} E \cdot J_i \, d\Omega &= \frac{1}{2} \int_{\Gamma} (E \times \overline{H}) \cdot n \, d\Gamma + \frac{\omega}{2} \int_{\Omega} \mu_0 \operatorname{Im}(\mu_r(\omega)) |H|^2 + \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E|^2 \, d\Omega \\ &+ i \frac{\omega}{2} \int_{\Omega} \varepsilon_0 \operatorname{Re}(\varepsilon_r(\omega)) |E|^2 - \mu_0 \operatorname{Re}(\mu_r(\omega)) |H|^2 \, d\Omega. \end{aligned} \quad (2.59)$$

Die Erhaltungsgleichung (2.59) ist das Resultat des Satzes von Poynting, nachzulesen in [35]. Hierbei beschreibt das Integral der linken Seite die komplette Leistung  $P_S(\Omega)$ , die durch eine elektrische Quelle  $J_i$  erzeugt wird. Das erste Integral der rechten Seite repräsentiert den komplexen Leistungsfluss durch die Oberfläche  $\Gamma$  und ergibt den Leistungsfaktor  $P_O(\Omega)$ . Dabei wird der Leistungsfluss

$$L_P = E \times \overline{H}$$

als Poynting-Vektor bezeichnet. Das zweite Integral

$$P_Q(\Omega) = \frac{\omega}{2} \int_{\Omega} \mu_0 \operatorname{Im}(\mu_r(\omega)) |H|^2 + \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E|^2 \, d\Omega \in \mathbb{R} \quad (2.60)$$

beschreibt den zeitlich gemittelten Wärmeverlust  $Q$  im Gebiet  $\Omega$ . Der erste Summand in diesem Term sind die magnetischen und der zweite Term die elektrischen Wärmeverluste, die durch Reibung der Dipole aufgrund des stetig wechselnden elektromagnetischen Feldes entstehen. Der letzte Term in (2.59) kann als gespeicherte Leistung  $P_B(V)$  im elektromagnetischen Feld aufgefasst werden und wird in der Literatur auch als Blindleistung bezeichnet. Wir erhalten somit die Leistungserhaltungsgleichung

$$P_S(\Omega) = P_O(\Omega) + P_Q(\Omega) + P_B(\Omega) \quad (2.61)$$

für beliebige reguläre Gebiete  $\Omega$ .

## 2.4. Randbedingungen

Wegen der Wohldefiniertheit der Spur- und Tangentialoperatoren aus Abschnitt 2.2 können wir verschiedene Randbedingungen elektromagnetischer Probleme betrachten. Sei  $n$  der Einheitsnormalenvektor, der von Material zwei nach Material eins zeigt. Des Weiteren seien  $E_1, E_2$  die elektrischen Vektorfelder und  $H_1, H_2$  die magnetischen Vektorfelder in den Materialien eins und zwei. Analog dazu sind  $\varepsilon_{r,1}(\omega)$  und  $\varepsilon_{r,2}(\omega)$ , bzw.  $\mu_{r,1}(\omega)$  und  $\mu_{r,2}(\omega)$ , die Permittivitäten, bzw. die Permeabilitäten, der beiden Materialien. Betrachten wir für die Vektorfelder den Grenzwert an der Schnittfläche beider Materialien, so ergeben sich die natürlichen Randbedingungen

$$\varepsilon_0 \varepsilon_{r,1}(\omega) E_1 \cdot n - \varepsilon_0 \varepsilon_{r,2}(\omega) E_2 \cdot n = \rho_S, \quad (2.62)$$

$$\mu_0 \mu_{r,1}(\omega) H_1 \cdot n - \mu_0 \mu_{r,2}(\omega) H_2 \cdot n = 0, \quad (2.63)$$

$$\gamma_t [E_1] - \gamma_t [E_2] = 0, \quad (2.64)$$

$$\gamma_t [H_1] - \gamma_t [H_2] = J_{l,S}, \quad (2.65)$$

wobei  $\rho_S$  die Flächenladungsdichte und  $J_{l,S}$  die Flächenstromdichte darstellt. Die Herleitung der Randgleichungen kann in [31] nachgelesen werden. Aus diesen Gleichungen lassen sich weitere natürliche Randbedingungen für elektromagnetische Probleme herleiten.

### Perfekter elektrischer Leiter (perfect electric conducting=PEC)

Ein perfektes elektrisch leitendes Material besitzt eine theoretische Konduktivität von unendlich. Dies wird approximativ verwendet, insofern der elektrische Widerstand vernachlässigt werden kann. In perfekten elektrischen Leitern gilt

$$E_2 = 0, \quad H_2 = 0,$$

da das Innere des Materials keine Propagation des elektrischen und magnetischen Feldes ermöglicht. Damit erhalten wir für perfekte elektrische Leiter die Randbedingungen

$$\varepsilon_0 \varepsilon_{r,1}(\omega) E_1 \cdot n = \rho_S, \quad (2.66)$$

$$\mu_0 \mu_{r,1}(\omega) H_1 \cdot n = 0, \quad (2.67)$$

$$\gamma_t [E_1] = 0, \quad (2.68)$$

$$\gamma_t [H_1] = J_{l,S}, \quad (2.69)$$

wobei der Normalenvektor  $n$  in Richtung des zu approximierenden PEC-Materials zeigt. Wird (2.40) oder (2.41) betrachtet, dann wird für die Charakterisierung eines PEC-

Materials die Dirichlet-Randbedingung (2.68) verwendet.

### Perfekter magnetischer Leiter (perfect magnetic conducting=PMC)

Während perfekte elektrische Leiter approximativ sehr gut leitende Materialien beschreiben, existieren keine perfekt magnetisch leitende Materialien. Jedoch lässt sich aus theoretischen Überlegungen zu perfekt leitenden Materialien aus (2.65) die Gleichung

$$\gamma_t [H_2] = 0$$

herleiten, da für perfekt magnetisch leitende Materialien  $J_{l,S} = 0$  gelten muss. Für zeitharmonische Probleme ergibt dies auch

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_{r,2}(\omega)} \nabla \times E_2 \right] = 0. \quad (2.70)$$

Die Neumann-Randbedingung (2.70) wird in Simulationen auf Symmetrieflächen angewendet, um das zu diskretisierende Gebiet zu verkleinern.

Als nächstes widmen wir uns den approximativen Randbedingungen.

### Impedanz-Randbedingung (IMP)

Die Impedanzbedingung

$$\pi_t [E_2] = Z(\omega) \gamma_t [H_2] \quad (2.71)$$

mit dem Feldwellenwiderstand

$$Z(\omega) = \sqrt{\frac{\mu_0 \mu_{r,2}(\omega)}{\varepsilon_0 \varepsilon_{r,2}(\omega)}} \quad (2.72)$$

wird aus der Betrachtung der Maxwell-Gleichungen im Frequenzbereich hergeleitet und entspricht einer Approximation ersten Grades des Feldes in dem Material. Eine Herleitung der approximativen Gleichung (2.71) findet sich in [36]. Mit (2.32) ergibt sich für (2.71) die Impedanzbedingung

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_{r,2}(\omega)} \nabla \times E_2 \right] = i\omega Y(\omega) \pi_t [E_2] \quad (2.73)$$

mit

$$Y(\omega) = \frac{1}{Z(\omega)} = \sqrt{\frac{\varepsilon_0 \varepsilon_{r,2}(\omega)}{\mu_0 \mu_{r,2}(\omega)}}. \quad (2.74)$$

Falls auf der zu approximierenden Fläche ein elektrisches Quellenfeld  $E_S$  existiert, dann haben wir die veränderte Gleichung

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_{r,2}(\omega)} \nabla \times E_2 \right] = i\omega Y(\omega) \pi_t [E_S] + i\omega Y(\omega) \pi_t [E_2].$$

Die Verwendung der Impedanzrandbedingung innerhalb einer Simulation wird oft von der Eindringtiefe

$$\delta = \sqrt{\frac{2}{\sigma_2 \omega \mu_0 \mu_{r,2}(\omega)}} \sqrt{\sqrt{1 + \left( \frac{\omega \varepsilon_0 \varepsilon_{r,2}(\omega)}{\sigma_2} \right)^2} + \frac{\omega \varepsilon_0 \varepsilon_{r,2}(\omega)}{\sigma_2}},$$

wobei hier  $\varepsilon_{r,2}(\omega)$  der relativen Permittivität aus (2.27) entspricht, abhängig gemacht. Nur wenn  $\delta \ll L$ , wobei  $L$  Dicke des zu approximierenden Materials ist, und

$$\left| \sqrt{\frac{\mu_{r,2}(\omega) \varepsilon_{r,2}(\omega)}{\mu_{r,1}(\omega) \varepsilon_{r,1}(\omega)}} \right| \gg 1,$$

wobei  $\mu_{r,1}(\omega)$  und  $\varepsilon_{r,1}(\omega)$  die effektiven relativen Materialparameter des nicht zu approximierenden Materials sind, gilt, stellt (2.73) eine gute Approximation des Materials dar.

Mit den hergeleiteten Randbedingungen an den Grenzflächen zweier Materialien haben wir die Möglichkeit die zeitharmonische Maxwell Gleichung (2.40) zu lösen. Da wir jedoch  $J_i = 0$  annehmen, ist die Lösung von (2.40) mit bisherigen Randbedingungen PEC, PMC und IMP die triviale Lösung. Deswegen leiten wir noch eine Randbedingung her, mit der wir die Anregung des elektromagnetischen Feldes modellieren können. Wir beschränken uns hierbei auf die Modellierung der Anregung mithilfe eines Koaxial-Wellenleiters, der oft beim Bau von Antennen verwendet wird und somit eine Möglichkeit für die Behandlung vieler realer Probleme bietet.

### Port-Bedingung für koaxiale Wellenleiter (PORT)

Um die Randbedingung für die Anregung eines Systems mit koaxialem Wellenleiter zu bestimmen, betrachten wir dessen Querschnitt, siehe Abbildung 2.1. Innerhalb eines koaxialen Wellenleiters breiten sich lediglich transversal elektromagnetische Wellen, genannt TEM-Wellen, aus. Dabei bedeutet transversal, dass das elektrische und magneti-

## 2. Elektromagnetische Simulation

---

sche Feld senkrecht zur Ausbreitungsrichtung steht. Der Innenleiter und der Außenleiter agieren hierbei wie ein Kondensator, weswegen die Feldrichtungen des elektrischen Feldes von Innen nach Außen zeigen, während das magnetische Feld, welches senkrecht zur Ausbreitungsrichtung und dem elektrischen Feld ist, eine kreisförmige Feldausrichtung besitzt.

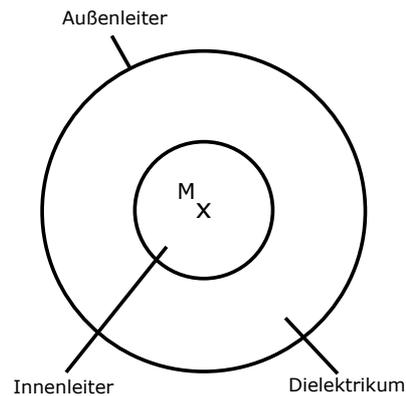


Abbildung 2.1: Querschnitt eines Koaxialkabels.

Die beiden Felder sind in Abbildung 2.2 für den Querschnitt visuell dargestellt.

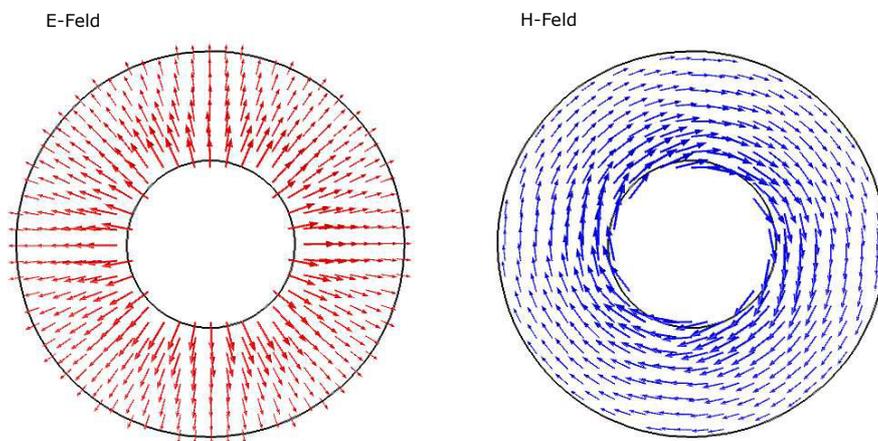


Abbildung 2.2: Modenfeld im Koaxialkabel.

Wie bereits erwähnt sind die beiden Felder von der Ausbreitungsrichtung  $k$  und der

Ausbreitungskonstante  $\alpha$  abhängig. Es ist dabei

$$k = -\alpha n$$

mit  $n$  äußerer Normalenvektor des Simulationsgebietes an der Portoberfläche und

$$\alpha = \sqrt{i\omega\mu_0\mu_r(\omega)(i\omega\varepsilon_0\varepsilon_r(\omega))} = i\mathbf{k}_0\sqrt{\mu_r(\omega)\varepsilon_r(\omega)}$$

mit der Wellenzahl  $\mathbf{k}_0 = \omega\sqrt{\mu_0\varepsilon_0}$  im Vakuum. Da wir je nach eingespeister Leistung unterschiedliche Feldstärken im Wellenleiter erhalten, leiten wir ein auf 1 Watt normiertes Referenzfeld her. Wie wir bereits wissen, ist die Richtung des elektrischen Feldes von innen nach außen gerichtet und nimmt nach außen hin in der Intensität ab, weswegen wir das Referenzfeld

$$E_0(X) = \frac{1}{\|X - M\|^2} (X - M)$$

für den Mittelpunkt  $M$  der Koaxial-Oberfläche formulieren. Das eben definierte Referenzfeld wird aus der Theorie von koaxialen Kondensatorstrukturen in der Hochfrequenztechnik abgeleitet, wie in [37] nachzulesen. Dabei ist  $E_0$  gerade so definiert, dass auf der Oberfläche  $\hat{\Gamma}$  die Gleichung  $\pi_t[E_0] = E_0$  gilt. Wegen der Orthogonalität des magnetischen Feldes zum elektrischen Feld und der Ausbreitungsrichtung  $k$  folgern wir

$$H_0 = \lambda k \times \pi_t[E_0] = -\lambda\alpha n \times \pi_t[E_0] = \lambda\alpha\gamma_t[\pi_t[E_0]], \quad (2.75)$$

wobei  $\lambda$  einen Streckungsfaktor darstellt, da wir die genaue Skalierung noch nicht kennen. Auch hier gilt wiederum  $\pi_t[H_0] = H_0$  aufgrund der Orthogonalität zum Normalenvektor  $n$ . Um nun das normierte Referenzfeld zu bestimmen, verwenden wir den Normierungsfaktor

$$\frac{1}{\sqrt{|P_m|}},$$

wobei  $P_m$  die Modenleistung des Referenzfeldes  $E_0$  auf der Koaxialoberfläche darstellt und im Folgenden hergeleitet wird. Hierzu betrachten wir die Erhaltungsgleichung (2.59). Dazu nehmen wir an, dass an die Randfläche  $\hat{\Gamma}$  ein beschränktes reguläres Gebiet  $\Upsilon \subset \mathbb{R}^3$ , mit  $\hat{\Gamma} = \overline{\Upsilon} \cap \overline{\Omega}$ , angefügt wird. Dieses Gebiet besitzt eine Quelle  $J_i$ , jedoch soll es keine elektromagnetischen Verluste und Blindleistungen aufweisen, also  $P_Q = P_B = 0$ . Dann ergibt sich daraus die Leistungserhaltungsgleichung

$$P_S(\Upsilon) = P_O(\Upsilon).$$

Nehmen wir dann noch an, dass ein Leistungsfluss  $L_P$  nur durch die Randfläche  $\hat{\Gamma} \subset \partial\Upsilon$  möglich ist und wir den Normalenvektor  $n$ , welcher von  $\Omega$  in Richtung  $\Upsilon$  zeigt, verwenden, dann ergibt sich die komplette erzeugte Modenleistung zu dem Phasenzeitpunkt

## 2. Elektromagnetische Simulation

---

$\theta = 0$  durch

$$P_O = -\frac{1}{2} \int_{\hat{\Gamma}} \left( \pi_t [E_0] \times \overline{\pi_t [H_0]} \right) \cdot n \, d\Gamma.$$

Betrachten wir  $P_O$  für ein um die Phase  $\theta$  verschobenes Feld, dann lässt sich mit wenigen Schritten die Zeitunabhängigkeit des komplexen Leistungsfaktors  $P_O$ , also

$$-\frac{1}{2} \int_{\hat{\Gamma}} \left( \pi_t [E_0 e^{-i\theta}] \times \overline{\pi_t [H_0 e^{-i\theta}]} \right) \cdot n \, dS = -\frac{1}{2} \int_{\hat{\Gamma}} \left( \pi_t [E_0] \times \overline{\pi_t [H_0]} \right) \cdot n \, d\Gamma,$$

zeigen. Nach dem Satz von Poynting entspricht dabei der Realteil von  $P_O$  der zeitlich gemittelten eingespeisten Leistung. Daher ergibt sich die reelle Modenleistung

$$P_m = -\frac{1}{2} \int_{\hat{\Gamma}} \operatorname{Re} \left( \pi_t [E_0] \times \overline{\pi_t [H_0]} \right) \cdot n \, d\Gamma.$$

Wir können damit die auf ein Watt normierten Modenfelder

$$\pi_t [E_{m,0}] = \frac{1}{\sqrt{|P_m|}} \pi_t [E_0], \quad \pi_t [H_{m,0}] = \frac{1}{\sqrt{|P_m|}} \pi_t [H_0]$$

zur Phase Null bestimmen, indem wir den Betrag des Leistungsfaktors  $P_m$  verwenden. Möchten wir das System nun mit einer bestimmten Leistung  $P_{in}$  und zu einer bestimmten Phase  $\theta$  bestimmen, dann ergeben sich die Modenfelder

$$\begin{aligned} \pi_t [E_m(P_{in}, \theta)] &= \sqrt{P_{in}} e^{-i\theta} \pi_t [E_{m,0}], \\ \pi_t [H_m(P_{in}, \theta)] &= \sqrt{P_{in}} e^{-i\theta} \pi_t [H_{m,0}]. \end{aligned}$$

Da das durch (2.40) bestimmte elektromagnetische Feld an der betrachteten koaxialen Oberfläche gerade dem elektrischen und magnetischen Modenfeld entsprechen soll, wird sowohl für die tangentielle Komponente des elektrischen Feldes

$$\pi_t [E] = \pi_t [E_m] \tag{2.76}$$

oder äquivalent

$$\gamma_t \left[ \frac{\alpha}{\mu_0 \mu_r(\omega)} \gamma_t [\pi_t [E]] \right] = \gamma_t \left[ \frac{\alpha}{\mu_0 \mu_r(\omega)} \gamma_t [\pi_t [E_m]] \right] \tag{2.77}$$

als auch für die tangentielle Komponente des magnetischen Feldes

$$\frac{1}{\lambda} \frac{1}{\mu_0 \mu_r(\omega)} \pi_t [H] = \frac{1}{\lambda} \frac{1}{\mu_0 \mu_r(\omega)} \pi_t [H_m] = \frac{\alpha}{\mu_0 \mu_r(\omega)} \gamma_t [\pi_t [E_m]] \tag{2.78}$$

gefordert. Wir bestimmen nun den Faktor  $\lambda$ , indem wir (2.32) für (2.78) benutzen wollen. Dazu verwenden wir, dass für ein Vektorfeld  $E$

$$\gamma_t [E] = \gamma_t [\pi_t [E]]$$

gilt, womit

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right] = \gamma_t \left[ \frac{i\omega \mu_0 \mu_r(\omega)}{\mu_0 \mu_r(\omega)} \pi_t [H] \right] = \gamma_t \left[ \frac{1}{\lambda} \frac{1}{\mu_0 \mu_r(\omega)} \pi_t [H] \right]$$

gelten sollte, und damit  $\lambda = \frac{1}{i\omega \mu_0 \mu_r(\omega)}$  folgt. Mit (2.78) erhalten wir die Bedingung

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right] = \gamma_t \left[ \frac{\alpha}{\mu_0 \mu_r(\omega)} \gamma_t [\pi_t [E_m]] \right]. \quad (2.79)$$

Eine Kombination aus (2.77) und (2.79) liefert die Robin-Randbedingung

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right] = 2\gamma_t \left[ \frac{\alpha}{\mu_0 \mu_r(\omega)} \gamma_t [\pi_t [E_m]] \right] - \gamma_t \left[ \frac{\alpha}{\mu_0 \mu_r(\omega)} \gamma_t [\pi_t [E]] \right].$$

Werden andere als koaxiale Oberflächen betrachtet, dann ergibt sich ein veränderter Term im ersten Term der rechten Seite. Wegen (2.78) und der Frequenzabhängigkeit des Faktors  $\lambda$  ist auch der Normierungsfaktor  $P_m$  frequenzabhängig und wir schreiben deshalb  $P_m(\omega)$ . Wegen der Betrachtung linearer isotroper Materialien und der Gleichung

$$\frac{\alpha}{\mu_0 \mu_r(\omega)} = \frac{\sqrt{i\omega \mu_0 \mu_r(\omega) (i\omega \varepsilon_0 \varepsilon_r(\omega))}}{\mu_0 \mu_r(\omega)} = i\omega \sqrt{\frac{\varepsilon_0 \varepsilon_r(\omega)}{\mu_0 \mu_r(\omega)}} = i\omega Y(\omega),$$

wobei  $Y(\omega)$  der Feldwellenwiderstand aus (2.74) ist, lässt sich die abschließende Robin-Randbedingung

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right] = i\omega Y(\omega) \pi_t [E] - 2i\omega Y(\omega) \frac{\sqrt{P_m}}{\sqrt{|P_m(\omega)|}} e^{-i\theta} \pi_t [E_0] \quad (2.80)$$

aufgrund der Vektoridentität

$$\gamma_t [\gamma_t [E]] = -\pi_t [E]$$

folgern. Die eben hergeleitete Randbedingung (2.80) entspricht für lineare isotrope Materialien gerade der Impedanzbedingung (2.73) mit einem entsprechenden Quellenfeld. Wir werden die eben hergeleitete Bedingung (2.80) als Approximation auch auf annähernd koaxiale Oberflächen anwenden.

## 2.5. PDE elektromagnetischer Probleme mit Koaxial-Portanregung

Wir haben damit alle Gleichungen hergeleitet, um elektromagnetische Probleme mit Koaxial-Portanregung zu betrachten. Zusammenfassend haben wir

$$\nabla \times \left( \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right) - \omega^2 \varepsilon_0 \varepsilon_r(\omega) E = 0 \quad \text{in } \Omega \quad (2.81)$$

für ein messbares polyedrisches Gebiet  $\Omega \subset \mathbb{R}^3$  mit stückweise stetigem Rand

$$\Gamma_T = \Gamma_{PEC} \cup \Gamma_{IMP} \cup \Gamma_{PORT}.$$

Wir führen  $\Gamma_R = \Gamma_{IMP} \cup \Gamma_{PORT}$  als die Menge der Ränder mit Robin-Randbedingung ein. Des Weiteren definieren wir  $K_A$  als die Anzahl an Port-Oberflächen, die auch als Anzahl der Antennen des Systems verstanden werden kann. Demnach ist

$$\Gamma_{PORT} = \bigcup_{1 \leq l \leq K_A} \Gamma_{P,l},$$

mit  $\Gamma_{P,l}$  die  $l$ -te Port-Oberfläche. Für  $1 \leq l \leq K_A$  sind die zugehörigen Randterme gegeben durch

$$\gamma_t [E] = 0 \quad \text{auf } \Gamma_{PEC}, \quad (2.82)$$

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right] = i\omega Y(\omega) \pi_t [E] \quad \text{auf } \Gamma_{IMP}, \quad (2.83)$$

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r(\omega)} \nabla \times E \right] = i\omega Y(\omega) \pi_t [E] - 2i\omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t [E_{0,l}] \quad \text{auf } \Gamma_{P,l}. \quad (2.84)$$

Wir treffen nun Voraussetzungen an die Parameterfunktionen bezüglich der in der Arbeit betrachteten Probleme. Hierzu sei  $C_S(\bar{\Omega})$  die Menge aller stückweise stetiger Funktionen in  $\bar{\Omega}$ .

**Voraussetzung 2.10.** Seien  $Q_\Omega \subset \bar{\Omega}$  und  $Q_\Gamma \subset \Gamma_R$  beliebige aber feste offene Teilmengen.

1. Für die Frequenz  $f$  gilt

$$0 < f_{\min} \leq f \leq f_{\max} < \infty$$

mit Konstanten  $f_{\min}, f_{\max} \in \mathbb{R}_+$ . Damit ist  $\omega \in [\omega_{\min}, \omega_{\max}]$  mit  $\omega_{\min} = 2\pi f_{\min}$  und  $\omega_{\max} = 2\pi f_{\max}$ .

2. Es gilt  $\mu_r : \bar{\Omega} \rightarrow \mathbb{R}$  mit  $\mu_r \in C_S(\bar{\Omega})$  und

$$0 < \mu_{\min} \leq \mu_r(x) \leq \mu_{\max}, \quad x \in \bar{\Omega},$$

und Konstanten  $\mu_{\min}, \mu_{\max} \in \mathbb{R}_+$ .

3. Es gilt  $\sigma : \bar{\Omega} \rightarrow \mathbb{R}$  mit  $\sigma \in C_S(\bar{\Omega})$ . Für Konstanten  $\sigma_{\min}, \sigma_{\max} \in \mathbb{R}_+$  gilt

$$0 < \sigma_{\min} \leq \sigma(x) \leq \sigma_{\max}, \quad x \in \bar{\Omega} \setminus Q_\Omega,$$

und

$$\sigma(x) = 0, \quad x \in Q_\Omega.$$

4. Es gilt  $\varepsilon_r : \bar{\Omega} \times [0, \sigma_{\max}] \times [\omega_{\min}, \omega_{\max}] \rightarrow \mathbb{C}$  mit  $\varepsilon_r(\cdot, \sigma, \omega) \in C_S(\bar{\Omega})$  und die Abbildungen  $\omega \mapsto \varepsilon_r(x, \sigma, \omega)$  und  $\sigma \mapsto \varepsilon_r(x, \sigma, \omega)$  sind stetig. Für Konstanten

$$\varepsilon_{\min, \text{Re}}, \varepsilon_{\max, \text{Re}}, \varepsilon_{\min, \text{Im}}, \varepsilon_{\max, \text{Im}} \in \mathbb{R}_+$$

und alle Frequenzen  $\omega \in [\omega_{\min}, \omega_{\max}]$  gilt

$$0 < \varepsilon_{\min, \text{Re}} \leq \text{Re}(\varepsilon_r(x, \sigma, \omega)) \leq \varepsilon_{\max, \text{Re}}, \quad x \in \bar{\Omega}.$$

Des Weiteren gilt

$$0 < \varepsilon_{\min, \text{Im}} \leq \text{Im}(\varepsilon_r(x, \sigma, \omega)) \leq \varepsilon_{\max, \text{Im}}, \quad x \in \bar{\Omega} \setminus Q_\Omega,$$

und

$$\text{Im}(\varepsilon_r(x, \sigma, \omega)) = 0, \quad x \in Q_\Omega.$$

5. Es gilt  $Y : \Gamma_R \times [0, \sigma_{\max}] \times [\omega_{\min}, \omega_{\max}] \rightarrow \mathbb{C}$ , mit  $Y(\cdot, \sigma, \omega) \in C_S(\Gamma_R)$ . Für Konstanten

$$Y_{\min, \text{Re}}, Y_{\max, \text{Re}}, Y_{\min, \text{Im}}, Y_{\max, \text{Im}} \in \mathbb{R}_+$$

und alle Frequenzen  $\omega \in [\omega_{\min}, \omega_{\max}]$  gilt

$$0 < Y_{\min, \text{Re}} \leq \text{Re}(Y(x, \sigma, \omega)) \leq Y_{\max, \text{Re}}, \quad x \in \Gamma_R.$$

Des Weiteren gilt

$$0 < Y_{\min, \text{Im}} \leq \text{Im}(Y(x, \sigma, \omega)) \leq Y_{\max, \text{Im}}, \quad x \in \Gamma_R \setminus Q_\Gamma,$$

und

$$\text{Im}(Y(x, \sigma, \omega)(x)) = 0, \quad x \in Q_\Gamma.$$

Hierbei gilt jedoch  $\text{Im}(Y(x, \sigma, \omega)) = 0$  genau dann, wenn  $\text{Im}(\varepsilon_r(x, \sigma, \omega)) = 0$  gilt.

## 2. Elektromagnetische Simulation

---

6. Es gilt  $P_{in} : \Gamma_{PORT} \rightarrow \mathbb{R}$  mit

$$P_{in} |_{\Gamma_{P,l}} = P_{in,l},$$

wobei  $P_{in,l} \in [0, P_{\max}]$  und  $P_{\max} \in \mathbb{R}_+$  für  $1 \leq l \leq K_A$ .

7. Es gilt  $\theta : \Gamma_{PORT} \rightarrow \mathbb{R}$  mit

$$\theta |_{\Gamma_{P,l}} = \theta_l,$$

wobei  $\theta_l \in [0, 2\pi]$  für  $1 \leq l \leq K_A$ .

Die Bedingung  $0 < \text{Im}(\varepsilon_r(x, \sigma, \omega))$  für  $x \in \bar{\Omega} \setminus Q_\Omega$  in Punkt 3 der Voraussetzung 2.10 wird explizit für eine wohldefinierte Betrachtung der Optimalsteuerung in Kapitel 5 benötigt. Die Abhängigkeit zwischen  $\text{Im}(Y(x, \sigma, \omega))$  und  $\text{Im}(\varepsilon_r(x, \sigma, \omega))$  in Punkt 5, ergibt sich wie wir noch später sehen werden bei Betrachtung der Definition (2.74) und der Voraussetzung der Reellwertigkeit der Permeabilität  $\mu_r$ . Im weiteren Verlauf werden wir die Ortsabhängigkeit der Parameter, aufgrund der Verwendung verschiedener Materialien, und die Abhängigkeiten bezüglich  $\sigma$  nicht mehr explizit angeben. Ein wichtiges Resultat, welches in Kapitel 5 benötigt wird, ist die stetige Abhängigkeit der frequenzabhängigen Parameterfunktionen.

**Satz 2.11.** *Ist die Abbildung  $\omega \mapsto \varepsilon_r(\omega)$  stetig, dann sind die frequenzabhängigen Abbildungen  $\omega \mapsto Y(\omega)$  und  $\omega \mapsto P_m(\omega)$  auch stetig.*

*Beweis.* Wir betrachten die Definition (2.74) der Inversen der Impedanz

$$Y(\omega) = \sqrt{\frac{\varepsilon_0 \varepsilon_r(\omega)}{\mu_0 \mu_r}} = \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \sqrt{\varepsilon_r(\omega)},$$

welche aufgrund der Stetigkeit der Wurzelfunktion und der Abbildung  $\omega \mapsto \varepsilon_r(\omega)$  selbst stetig ist. Um die Stetigkeit in der Frequenz für die Normierungsfaktoren  $P_{m,l}(\omega)$  zu beweisen, betrachten wir deren Definition

$$P_{m,l}(\omega) = -\frac{1}{2} \int_{\Gamma_{P,l}} \text{Re} \left( \pi_t [E_{0,l}] \times \overline{\pi_t [H_{0,l}]} \right) \cdot n \, dS.$$

Setzen wir nun (2.75) in diese Definition ein, dann ergibt sich

$$P_{m,l}(\omega) = -\frac{1}{2} \int_{\Gamma_{P,l}} \text{Re} \left( \frac{\bar{\alpha}}{-i\omega \mu_0 \mu_r} \right) (\pi_t [E_{0,l}] \times \gamma_t [\pi_t [E_{0,l}]]) \cdot n \, dS,$$

da  $E_0$  nach Definition ein reelles Vektorfeld darstellt. Mit der Definition

$$\alpha = i\omega \sqrt{\mu_0 \varepsilon_0} \sqrt{\mu_r} \sqrt{\varepsilon_r(\omega)}$$

folgt

$$P_{m,l}(\omega) = -\frac{1}{2} \int_{\Gamma_{P,l}} \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \operatorname{Re} \left( \sqrt{\varepsilon_r(\omega)} \right) (\pi_t [E_{0,l}] \times \gamma_t [\pi_t [E_{0,l}]] \cdot n \, dS$$

und damit die Stetigkeit von  $P_{m,l}(\omega)$  analog zur Stetigkeit von  $Y(\omega)$ .  $\square$

Für die Voraussetzung von Satz 2.11 betrachten wir die Definition der reformulierten komplexen Permittivität (2.38). Dabei ist ersichtlich, dass der hintere Term  $\frac{\sigma}{\omega \varepsilon_0}$  stetig in  $\omega$  ist und somit ist die Stetigkeit der Abbildung  $\omega \mapsto \varepsilon_r(\omega)$  direkt abhängig von der Stetigkeit des Polarisationsfaktors  $\chi_e(\omega)$  aus Abschnitt 2.1.4. Wir wollen die Korrektheit der Aussage 5 der Voraussetzungen 2.10 näher begründen. Dazu betrachten wir

$$Y(\omega) = \sqrt{\frac{\varepsilon_0 \varepsilon_r(\omega)}{\mu_0 \mu_r}} = \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \sqrt{\varepsilon_r(\omega)}.$$

Der erste Faktor ist nach Voraussetzung 2.10 eine positive reelle Zahl, weswegen wir lediglich den zweiten Faktor beachten müssen. Aus der Wurzel einer komplexen Zahl, deren Real- und Imaginärteil stets größer oder gleich Null sind, siehe Definition 2.4, lassen sich die Bedingungen

$$\operatorname{Re}(Y(\omega)), \operatorname{Im}(Y(\omega)) \geq 0, \quad \operatorname{Re}(Y(\omega)) > \operatorname{Im}(Y(\omega))$$

oder

$$\operatorname{Re}(Y(\omega)), \operatorname{Im}(Y(\omega)) \leq 0 \quad \operatorname{Re}(Y(\omega)) < \operatorname{Im}(Y(\omega))$$

an die komplexe Zahl  $Y(\omega)$  ableiten. Dabei macht jedoch nur der erste Fall aus physikalischer Sicht Sinn, da der Realteil des Feldwellenwiderstands  $Y(\omega)$  stets positiv sein muss. Für die numerische Betrachtung der Gleichungen (2.81)-(2.84) leiten wir in Kapitel 3 die kontinuierliche und diskrete variationelle Formulierung her und beweisen die Existenz der Lösungen unter Voraussetzung 2.10.

### 3. Kontinuierliche und diskrete variationelle Formulierung

Um die Existenz und Eindeutigkeit der variationellen Formulierung der Maxwell-Gleichungen (2.81)-(2.84) zu beweisen, benötigen wir zunächst den Lösungsraum

$$X_e = \{ E \in H(\text{curl}; \Omega) : \pi_t[E] \in L_t^2(\Gamma_R), \gamma_t[E] = 0 \text{ auf } \Gamma_{PEC} \} \quad (3.1)$$

für das komplexe Vektorfeld  $E$ , wobei  $\pi_t$  den tangentialen Projektionsoperator aus Satz 2.8 darstellt. Mit dem Skalarprodukt

$$(E, \varphi)_{X_e} = (E, \varphi)_{H(\text{curl}; \Omega)} + (\pi_{t, \Gamma_R}[E], \pi_{t, \Gamma_R}[\varphi])_{[L^2(\Gamma_R)]^3} \quad (3.2)$$

und der durch (3.2) induzierten Norm wird  $X_e$  zu einem Hilbertraum, siehe [32]. Wir betrachten (2.81) mit Voraussetzung 2.10 in der schwachen Formulierung

$$\int_{\Omega} \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times E \right) \cdot \varphi - (\omega^2 \varepsilon_0 \varepsilon_r(\omega)) E \cdot \varphi \, d\Omega = 0$$

mit  $\varphi \in X_e$ . Mit der Green'schen Formel (2.55) erhalten wir

$$\int_{\Omega} \frac{1}{\mu_0 \mu_r} (\nabla \times E) \cdot (\nabla \times \varphi) - (\omega^2 \varepsilon_0 \varepsilon_r(\omega)) E \cdot \varphi \, d\Omega - \int_{\Gamma_T} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] \cdot \pi_t[\varphi] \, d\Gamma = 0.$$

Nutzen wir aus, dass  $\varphi \in X_e$  und damit Nullrandwerte auf  $\Gamma_{PEC}$  besitzt, setzen die unterschiedlichen Randbedingungen (2.82)-(2.84) ein und multiplizieren die Gleichung mit  $\mu_0$ , so ergibt sich das kontinuierliche variationelle Problem:

Bestimme  $E \in X_e$ , sodass

$$[\text{VK}] \quad a(E, \varphi) = f(\varphi), \quad \varphi \in X_e,$$

gilt mit

$$a(E, \varphi) = \int_{\Omega} \frac{1}{\mu_r} (\nabla \times E) \cdot (\nabla \times \varphi) \, d\Omega - \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 E \cdot \varphi \, d\Omega \quad (3.3)$$

$$- \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 \pi_t[E] \cdot \pi_t[\varphi] \, d\Gamma - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \pi_t[E] \cdot \pi_t[\varphi] \, d\Gamma,$$

$$f(\varphi) = -2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma. \quad (3.4)$$

### 3.1. Existenztheorie der kontinuierlichen Formulierung

Um die Existenz einer Lösung von  $[V_K]$  zu beweisen, folgen wir dem Beweis aus [28], indem wir kompakte Operatoren herleiten und die Sesquilinearform  $a$  entsprechend umformen, um mit der Fredholm'schen Alternative die Existenz zu erhalten. Den Satz der Fredholm'schen Alternative entnehmen wir [32].

**Satz 3.1** (Fredholm'sche Alternative). *Sei  $B : X \rightarrow X$  ein beschränkter linearer Operator und  $X$  ein Hilbertraum. Ist  $B = I + A$ , wobei  $A$  kompakter Operator und  $I$  der Identitätsoperator ist, dann gilt genau eine der folgenden Aussagen:*

1. *Die homogene Gleichung  $B[u] = 0$  besitzt nur die triviale Lösung  $u = 0$  in  $X$ . In diesem Fall existiert für jedes  $f \in X$ , eine eindeutige Lösung  $u \in X$  mit  $B[u] = f$  und die Lösung hängt stetig von der rechten Seite  $f$  ab.*
2. *Die homogene Gleichung  $B[u] = 0$  hat genau  $p$  linear unabhängige Lösungen für ein festes  $p \in \mathbb{N}$ .*

Wir leiten die benötigten kompakten Operatoren in Abschnitt 3.1.1 her und wenden diese dann in Abschnitt 3.1.2 auf unser Problem an.

#### 3.1.1. Kompakter Operator

Sei  $\Omega_G \subset \mathbb{R}^3$  ein offenes polyedrisches Gebiet mit dem Rand  $\Gamma_G = \partial\Omega_G$  und  $\Omega_{\Sigma_k} \subset \Omega_G$ , wobei hier  $k \in I_\Sigma = \{1, \dots, K_\Sigma\}$  mit  $K_\Sigma$  Anzahl der offenen polyedrischen Teilgebiete, für die  $\Omega_{\Sigma_k} \cap \Omega_{\Sigma_l} = \emptyset$  für alle  $k, l \in I_\Sigma$  mit  $k \neq l$  gilt, siehe Abbildung 3.1.

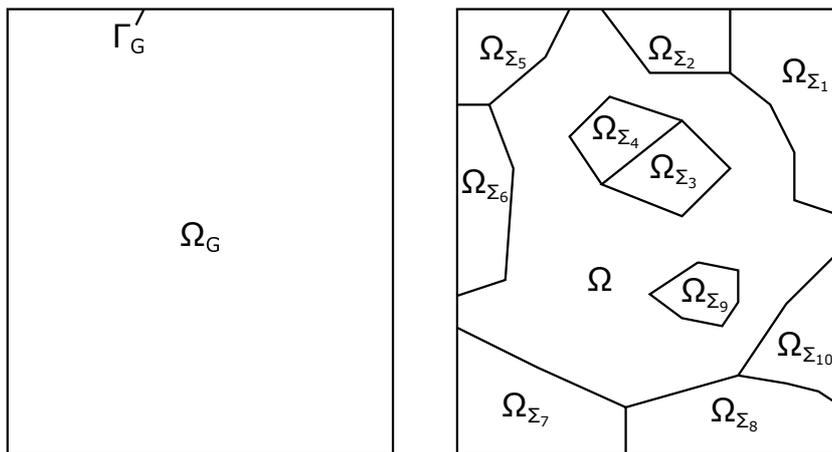


Abbildung 3.1: Skizze der Modellmengen.

### 3. Kontinuierliche und diskrete variationelle Formulierung

Mit  $\Sigma_k$  bezeichnen wir den Rand der Teilmenge  $\Omega_{\Sigma_k}$  und definieren das offene polyedrische Gebiet  $\Omega = \Omega_G \setminus \left\{ \bigcup_{k \in I_\Sigma} \overline{\Omega_{\Sigma_k}} \right\}$  mit dem Rand  $\Gamma = \partial\Omega$ , in welchem unsere Lösung existieren soll. Dabei können  $k, l \in I_\Sigma$  existieren, sodass  $\Omega_{\Sigma_k} \cap \Omega_{\Sigma_l} = \emptyset$ , aber  $\Sigma_k \cap \Sigma_l \neq \emptyset$ , weswegen wir die Menge  $II_\Sigma = \{(k, l) : \Sigma_k \cap \Sigma_l \neq \emptyset\}$  an Tupeln definieren, deren Ränder einen nichtleeren Schnitt aufweisen, siehe Abbildung 3.2.

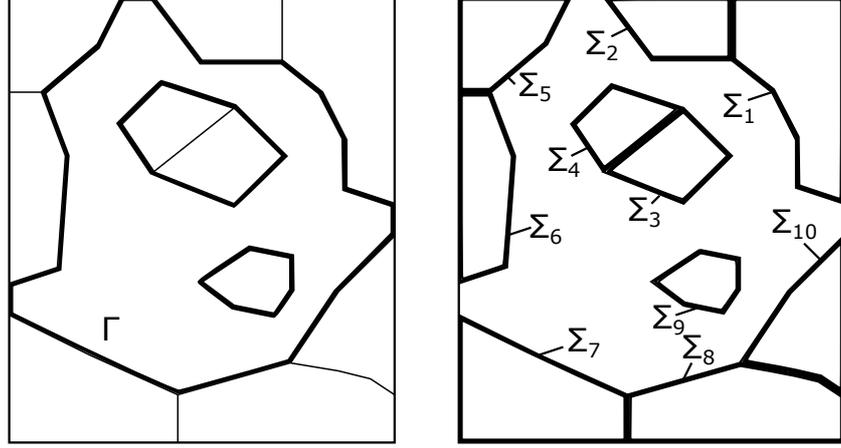


Abbildung 3.2: Skizze der Randmengen und Indextupeln mit  $I_{\Sigma,1} = \{1, 4, 8, 9, 10\}$  und  $I_{\Sigma,2} = \{2, 3, 5, 6, 7\}$ .

Des Weiteren führen wir  $\tilde{\Sigma}_k = \Gamma \cap \Sigma_k$  für alle  $k \in I_\Sigma$  und  $\Gamma_P = \Gamma \setminus \left\{ \bigcup_{k \in I_\Sigma} \tilde{\Sigma}_k \right\}$  ein. Wir teilen  $I_\Sigma$  in die beiden Indextupeln  $I_{\Sigma,1}$  und  $I_{\Sigma,2}$  auf, wobei  $I_\Sigma = I_{\Sigma,1} \dot{\cup} I_{\Sigma,2}$ , und definieren die Indextupeln

$$\begin{aligned} II_{\Sigma,1}(k) &= \{l \in I_\Sigma : (k, l) \in II_\Sigma \text{ und } l \in I_{\Sigma,1}\}, \\ II_{\Sigma,2}(k) &= \{l \in I_\Sigma : (k, l) \in II_\Sigma \text{ und } l \in I_{\Sigma,2}\}. \end{aligned}$$

In der variationellen Formulierung  $[V_K]$  entsprechen die durch  $\Gamma_{IMP}$  definierten Gebiete im Inneren des Modellgebietes den Gebieten zur Indextupel  $I_{\Sigma,1}$ , während zu den durch  $\Gamma_{PEC}$  definierten Gebiete  $I_{\Sigma,2}$  gehört. Der Rand  $\Gamma_{PORT}$  wird hierbei durch den Rand  $\Gamma_P$  dargestellt. Wir definieren den Lösungsraum

$$\begin{aligned} X_e = \left\{ v \in H(\text{curl}; \Omega) : \pi_{t, \tilde{\Sigma}_k} [v] \in L_t^2(\tilde{\Sigma}_k), k \in I_{\Sigma,1}; \pi_{t, \Gamma} [v] \in L_t^2(\Gamma_P); \right. \\ \left. \gamma_{t, \tilde{\Sigma}_k} [v] = 0, k \in I_{\Sigma,2} \right\}, \end{aligned} \quad (3.5)$$

wobei dieser die gleiche Struktur wie (3.1) aufweist. Die Norm in diesem Raum ist

entsprechend gegeben durch

$$\|v\|_{X_e}^2 = \|v\|_{H(\text{curl}; \Omega)}^2 + \|\pi_{t,\Gamma}[v]\|_{[L^2(\Gamma_P)]^3}^2 + \sum_{k \in I_{\Sigma,1}} \|\pi_{t,\tilde{\Sigma}_k}[v]\|_{[L^2(\tilde{\Sigma}_k)]^3}^2,$$

wodurch  $X_e$  mit dem durch die einzelnen Räume gegebenen Skalarprodukt zu einem Hilbertraum wird, siehe [32]. Weiterhin führen wir eine Sphäre  $\Gamma_0$  ein, die  $\bar{\Omega}_G$  vollständig umschließt und definieren die dazugehörige offene Kugel  $\Omega_{\Gamma_0}$ . Der Raum zwischen  $\Gamma_G$  und der Sphäre  $\Gamma_0$  wird definiert über  $\Omega_{\Gamma_G, \Gamma_0} = \Omega_{\Gamma_0} \setminus \{\bar{\Omega}_G\}$ , siehe Abbildung 3.3.

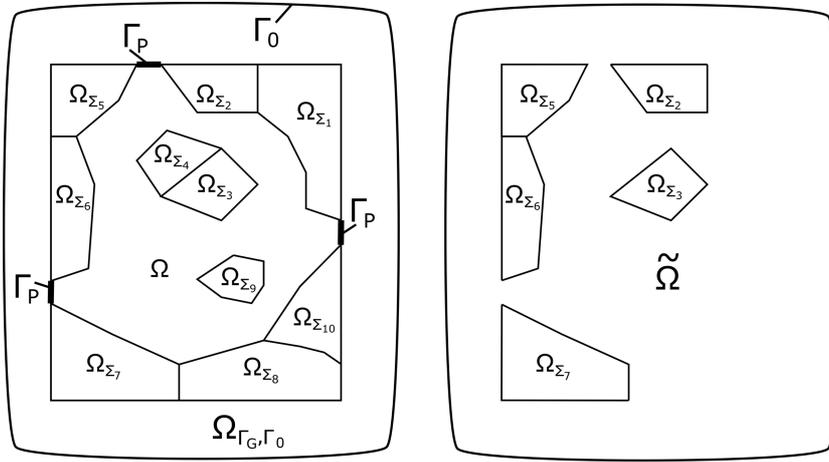


Abbildung 3.3: Skizze der sphärischen Umrandung der Modellmengen.

Als letztes definieren wir noch das Gebiet

$$\tilde{\Omega} = \Omega_{\Gamma_0} \setminus \left\{ \bigcup_{k \in I_{\Sigma,2}} \bar{\Omega}_{\Sigma_k} \right\}$$

und fordern zusätzlich für Funktionen  $v \in H(\text{curl}; \tilde{\Omega})$  die PEC-Eigenschaft

$$\gamma_{t,\Gamma_0}[v] = 0.$$

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

Für den nächsten Schritt definieren wir die folgenden Funktionenräume

$$V(\tilde{\Omega}) = \left\{ v \in H_0(\text{curl}; \tilde{\Omega}) : \text{div}[v] = 0 \text{ in } \Omega_{\Gamma_G, \Gamma_0}; \text{div}[v] = 0 \text{ in } \Omega_{\Sigma_l}, l \in I_{\Sigma,1} \right\}, \quad (3.6)$$

$$V_0(\Omega_{\Gamma_G, \Gamma_0}) = \{v \in H_0(\text{curl}; \Omega_{\Gamma_G, \Gamma_0}) : \text{div}[v] = 0 \text{ in } \Omega_{\Gamma_G, \Gamma_0}\}, \quad (3.7)$$

$$V_0(\Omega_{\Sigma_k}) = \{v \in H_0(\text{curl}; \Omega_{\Sigma_k}) : \text{div}[v] = 0 \text{ in } \Omega_{\Sigma_k}\} \quad (3.8)$$

für alle  $k \in I_{\Sigma,1}$ , und geben ein nützliches Lemma an, dessen Beweis in [38] zu finden ist.

**Lemma 3.2.** *Die Seminorm  $v \mapsto \|\text{curl}[v]\|_{[L^2(\Omega_{\Gamma_G, \Gamma_0})]^3}$  ist eine zu  $H(\text{curl}; \Omega_{\Gamma_G, \Gamma_0})$  äquivalente Norm auf  $V_0(\Omega_{\Gamma_G, \Gamma_0})$ . Sei  $k \in I_{\Sigma,1}$ , dann ist die Seminorm gegeben durch  $v \mapsto \|\text{curl}[v]\|_{[L^2(\Omega_{\Sigma_k})]^3}$  eine zu  $H(\text{curl}; \Omega_{\Sigma_k})$  äquivalente Norm auf  $V_0(\Omega_{\Sigma_k})$ .*

Um Satz 3.1 anwenden zu können leiten wir im Folgenden einen kompakten Operator her, der wiederum aus einzelnen Operatoren besteht. Der erste der benötigten Operatoren, ist ein Fortsetzungsoperator von  $X_e$  nach  $V(\tilde{\Omega})$ .

**Lemma 3.3.** *Es existiert ein linearer und beschränkter Fortsetzungsoperator*

$$\begin{aligned} \mathcal{E} : X_e &\rightarrow V(\tilde{\Omega}), \\ v &\mapsto \mathcal{E}[v], \end{aligned}$$

mit  $\mathcal{E}[v]|_{\Omega} = v$ , welcher

$$\int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times \mathcal{E}[v]) \cdot (\nabla \times q) \, d\Omega = 0, \quad q \in V_0(\Omega_{\Gamma_G, \Gamma_0}), \quad (3.9)$$

$$\int_{\Omega_{\Sigma_k}} (\nabla \times \mathcal{E}[v]) \cdot (\nabla \times q) \, d\Omega = 0, \quad q \in V_0(\Omega_{\Sigma_k}), \quad k \in I_{\Sigma,1}, \quad (3.10)$$

erfüllt.

*Beweis.* Nach Satz 2.8 und Bemerkung 2.9 existieren die tangentialen Spurooperatoren  $\gamma_{t, \Sigma_k}^+$  von  $\Omega_{\Sigma_k}$  nach  $\Sigma_k$  und  $\gamma_{t, \tilde{\Sigma}_k}^-$  von  $\Omega$  nach  $\tilde{\Sigma}_k$ . Wir leiten zunächst rechtsinverse Operatoren von  $\tilde{\Sigma}_k$  nach  $\Omega_{\Sigma_k}$ ,  $k \in I_{\Sigma,1}$ , her, wobei hier drei Fälle auftreten, siehe Abbildung 3.4.

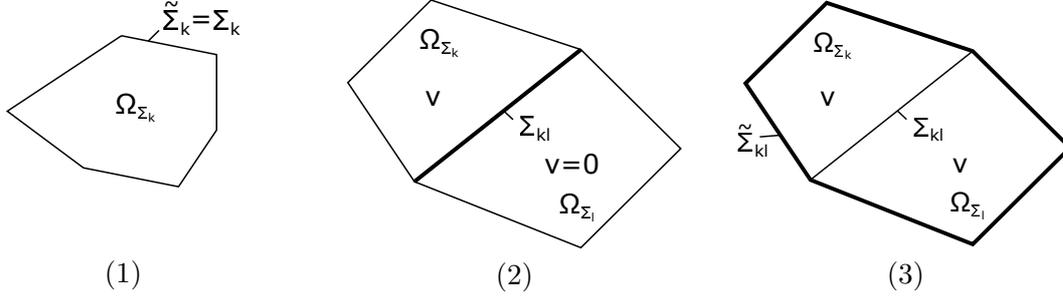


Abbildung 3.4: Skizze der Gebiete für die rechtsinversen Operatoren.

- (1)  $\Sigma_k \cap \Sigma_l = \emptyset$ ,  $l \in I_\Sigma$ ,  $l \neq k$ :

Es gilt  $\tilde{\Sigma}_k = \Sigma_k$  und damit  $\gamma_{t, \tilde{\Sigma}_k}^+ = \gamma_{t, \Sigma_k}^+$ , womit nach Satz 2.8 ein linearer stetiger rechtsinverser Operator

$$(\gamma_{t, \tilde{\Sigma}_k}^+)^{-1} : H^{-\frac{1}{2}}(\text{div}_{\Sigma_k}; \Sigma_k) \rightarrow H(\text{curl}; \Omega_{\Sigma_k})$$

existiert.

- (2)  $\Sigma_k \cap \Sigma_l \neq \emptyset$ ,  $k \in I_{\Sigma,1}$ ,  $l \in II_{\Sigma,2}(k)$ ,  $l \neq k$ :

Nach Voraussetzung existiert  $\Sigma_{kl} = \Sigma_k \cap \Sigma_l \neq \emptyset$  und es gilt

$$\gamma_{t, \tilde{\Sigma}_l}^- [v] = 0$$

für alle  $v \in X_e$ . Nach Herleitung der PEC-Randbedingung in Abschnitt 2.3 fordern wir  $v = 0$  in  $\Omega_{\Sigma_l}$ , womit

$$\gamma_{t, \Sigma_l}^+ [v] \Big|_{\Sigma_{kl}} = \gamma_{t, \Sigma_k}^+ [v] \Big|_{\Sigma_{kl}} = 0$$

für die Einschränkung auf deren Schnitt folgen muss. Als nächstes definieren wir den Fortsetzungsoperator

$$\mathcal{E}_{\Sigma_{kl}} : H^{-\frac{1}{2}}(\text{div}_{\tilde{\Sigma}_k}; \tilde{\Sigma}_k) \rightarrow H_{\Sigma_{kl}}^{-\frac{1}{2}}(\text{div}_{\Sigma_k}; \Sigma_k) = \left\{ u \in H^{-\frac{1}{2}}(\text{div}_{\Sigma_k}; \Sigma_k) : u \Big|_{\Sigma_{kl}} = 0 \right\},$$

der gegeben ist durch

$$\mathcal{E}_{\Sigma_{kl}} [u] = \begin{cases} u & \text{in } \tilde{\Sigma}_k, \\ 0 & \text{in } \Sigma_{kl}. \end{cases}$$

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

Da  $\tilde{\Sigma}_k \subset \Sigma_k$ , folgt die Linearität und Stetigkeit des Fortsetzungsoperators  $\mathcal{E}_{\Sigma_{kl}}$  aufgrund seiner Definition. Wiederum mit Satz 2.8 erhalten wir somit die Existenz des linearen stetigen rechtsinversen Operators

$$(\tilde{\gamma}_{t, \tilde{\Sigma}_k}^+)^{-1} : H^{-\frac{1}{2}}(\operatorname{div}_{\tilde{\Sigma}_k}; \tilde{\Sigma}_k) \rightarrow H(\operatorname{curl}; \Omega_{\Sigma_k}),$$

der gerade durch

$$(\tilde{\gamma}_{t, \tilde{\Sigma}_k}^+)^{-1} = (\gamma_{t, \Sigma_k}^+)^{-1} \circ \mathcal{E}_{\Sigma_{kl}}$$

gegeben ist.

$$(3) \quad \Sigma_k \cap \Sigma_l \neq \emptyset, \quad k \in I_{\Sigma,1}, \quad l \in II_{\Sigma,1}(k), \quad k \neq l:$$

Es sei wiederum  $\Sigma_{kl} = \Sigma_k \cap \Sigma_l$ . Nach Satz 2.8 existiert der lineare stetige rechtsinverse Operator

$$(\tilde{\gamma}_{t, \tilde{\Sigma}_{kl}}^+)^{-1} : H^{-\frac{1}{2}}(\operatorname{div}_{\tilde{\Sigma}_{kl}}; \tilde{\Sigma}_{kl}) \rightarrow H(\operatorname{curl}; \Omega_{\Sigma_{kl}}),$$

wobei  $\tilde{\Sigma}_{kl} = (\Sigma_k \cup \Sigma_l) \setminus (\Sigma_{kl})$  und  $\Omega_{\Sigma_{kl}} = \Omega_{\Sigma_k} \cup \Omega_{\Sigma_l} \cup \Sigma_{kl}$ . Mit dem linearen und stetigen Einschränkungoperator erhalten wir die Existenz des linearen stetigen rechtsinversen Operators

$$(\tilde{\gamma}_{t, \tilde{\Sigma}_k}^+)^{-1} : H^{-\frac{1}{2}}(\operatorname{div}_{\tilde{\Sigma}_k}; \tilde{\Sigma}_k) \rightarrow H(\operatorname{curl}; \Omega_{\Sigma_k})$$

gegeben durch

$$(\tilde{\gamma}_{t, \tilde{\Sigma}_k}^+)^{-1} = (\tilde{\gamma}_{t, \tilde{\Sigma}_{kl}}^+)^{-1} \Big|_{\Omega_{\Sigma_k}}.$$

Mit diesen Fällen wird auch eine Kombination der letzten beiden Fälle abgedeckt, indem die entsprechenden Operatoren der einzelnen Fälle verwendet werden. Zudem setzen wir für Gebiete  $\Omega_{\Sigma_k}$ ,  $k \in I_{\Sigma,1}$ , mit  $\Sigma_k \cap \Gamma_G \neq \emptyset$  für die Wohldefiniertheit und Eindeutigkeit des inversen Operators die entsprechenden Ränder mit Null fort. Dies entspricht so auch der physikalischen Interpretation der Impedanz-Bedingung. Damit lassen sich nun die Operatoren

$$\begin{aligned} \mathcal{L}_{\Sigma_k} : X_e &\rightarrow H(\operatorname{curl}; \Omega_{\Sigma_k}), \\ v &\mapsto \mathcal{L}_{\Sigma_k}[v] = (\tilde{\gamma}_{t, \tilde{\Sigma}_k}^+)^{-1}[\gamma_{t, \tilde{\Sigma}_k}^-[v]] \end{aligned} \tag{3.11}$$

für  $k \in I_{\Sigma,1}$  mit  $II_{\Sigma,2}(k) = \emptyset$  und

$$\begin{aligned} \mathcal{L}_{\Sigma_k} : X_e &\rightarrow H_{II_{\Sigma,2}(k)}(\operatorname{curl}; \Omega_{\Sigma_k}), \\ v &\mapsto \mathcal{L}_{\Sigma_k}[v] = (\tilde{\gamma}_{t, \tilde{\Sigma}_k}^+)^{-1}[\gamma_{t, \tilde{\Sigma}_k}^-[v]] \end{aligned} \tag{3.12}$$

mit

$$H_{II_{\Sigma,2}(k)}(\text{curl}; \Omega_{\Sigma_k}) = \{v \in H(\text{curl}; \Omega_{\Sigma_k}) : \gamma_{t, \Sigma_{kl}}[v] = 0, \quad l \in II_{\Sigma,2}(k)\}$$

für  $k \in I_{\Sigma,1}$  mit  $II_{\Sigma,2}(k) \neq \emptyset$  definieren. Die soeben definierten Operatoren  $\mathcal{L}_{\Sigma_k}$ ,  $k \in I_{\Sigma,1}$ , sind, wegen der Eigenschaften tangentialer Spuroperatoren und den entsprechend definierten rechtsinversen Operatoren, lineare und stetige Operatoren, weshalb Konstanten  $C_{\Sigma_k}$  existieren, sodass

$$\|\mathcal{L}_{\Sigma_k}[v]\|_{H(\text{curl}; \Omega_{\Sigma_k})} \leq C_{\Sigma_k} \|v\|_{X_e}, \quad v \in X_e, \quad k \in I_{\Sigma,1}. \quad (3.13)$$

Zu (3.11) und (3.12) definieren wir für  $k \in I_{\Sigma,1}$  die folgenden Variationsprobleme:

Bestimme die Funktionen  $z_{v, \Sigma_k} \in \mathcal{L}_{\Sigma_k}[v] + H_0(\text{curl}; \Omega_{\Sigma_k})$  und  $\chi_{\Sigma_k} \in H_0^1(\Omega_{\Sigma_k})$ , die

$$[P_{\Sigma_k}] \quad \begin{aligned} \int_{\Omega_{\Sigma_k}} (\nabla \times z_{v, \Sigma_k}) \cdot (\nabla \times q) \, d\Omega + \int_{\Omega_{\Sigma_k}} q \cdot \nabla \chi_{\Sigma_k} \, d\Omega &= 0, \quad q \in H_0(\text{curl}; \Omega_{\Sigma_k}), \\ \int_{\Omega_{\Sigma_k}} z_{v, \Sigma_k} \cdot \nabla \theta \, d\Omega &= 0, \quad \theta \in H_0^1(\Omega_{\Sigma_k}), \end{aligned}$$

lösen.

Jegliches dieser Probleme ist mit Satz A.21 lösbar. Hierzu schreiben wir  $[P_{\Sigma_k}]$  mit der Zerlegung  $z_{v, \Sigma_k} = \mathcal{L}_{\Sigma_k}[v] + z_{0, \Sigma_k}$  um und erhalten

$$\begin{aligned} a(z_{0, \Sigma_k}, q) + b(q, \chi_{\Sigma_k}) &= f(q), \quad q \in H_0(\text{curl}; \Omega_{\Sigma_k}), \\ b(z_{0, \Sigma_k}, \theta) &= g(\theta), \quad \theta \in H_0^1(\Omega_{\Sigma_k}), \end{aligned}$$

wobei

$$\begin{aligned} a(u, \xi) &= \int_{\Omega_{\Sigma_k}} (\nabla \times u) \cdot (\nabla \times \xi) \, d\Omega, & b(u, \xi) &= \int_{\Omega_{\Sigma_k}} u \cdot \nabla \xi \, d\Omega, \\ f(\xi) &= - \int_{\Omega_{\Sigma_k}} (\nabla \times \mathcal{L}_{\Sigma_k}[v]) \cdot (\nabla \times \xi) \, d\Omega, & g(\xi) &= - \int_{\Omega_{\Sigma_k}} \mathcal{L}_{\Sigma_k}[v] \cdot \nabla \xi \, d\Omega. \end{aligned}$$

Damit wir Satz A.21 anwenden können, müssen wir die  $V_0(\Omega_{\Sigma_k})$ -Koerzivität von  $a(\cdot, \cdot)$

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

nach Definition A.19 zeigen. Mit Lemma 3.2 folgt für eine Konstante  $C_\alpha > 0$

$$a(u, u) = \int_{\Omega_{\Sigma_k}} (\nabla \times u) \cdot (\nabla \times u) \, d\Omega = \|\operatorname{curl}[u]\|_{[L^2(\Omega_{\Sigma_k})]^3}^2 \geq C_\alpha \|u\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})}^2, \quad u \in V_0(\Omega_{\Sigma_k}).$$

Für die Vollständigkeit der Definition der  $V_0(\Omega_{\Sigma_k})$ -Koerzivität nach Definition A.19, müssen die Funktionen  $u \in V_0(\Omega_{\Sigma_k})$  noch die Bedingung  $b(u, \xi) = 0$  erfüllen. Wir erhalten mit partieller Integration und der Bedingung  $\operatorname{div}[u] = 0$  des Raumes  $V_0(\Omega_{\Sigma_k})$  die Gleichung

$$b(u, \xi) = \int_{\Omega_{\Sigma_k}} u \cdot \nabla \xi \, d\Omega = 0, \quad \xi \in H_0^1(\Omega_{\Sigma_k}).$$

Damit haben wir nun die  $V_0(\Omega_{\Sigma_k})$ -Koerzivität der Sesquilinearform  $a$  gezeigt. Wir benötigen noch, dass  $b(\cdot, \cdot)$  die Babuška-Brezzi-Bedingung, nach Definition A.20, erfüllt. Mit der natürlichen Einbettung  $\nabla(H_0^1(\Omega_{\Sigma_k})) \hookrightarrow H_0(\operatorname{curl}; \Omega_{\Sigma_k})$ , wegen

$$\nabla \times \nabla u = 0 \in [L^2(\Omega_{\Sigma_k})]^3, \quad \nabla u \in [L^2(\Omega_{\Sigma_k})]^3,$$

und der Poincaré-Ungleichung, siehe [39], folgt

$$\sup_{u \in H_0(\operatorname{curl}; \Omega_{\Sigma_k})} \frac{|\int_{\Omega_{\Sigma_k}} u \cdot \nabla \theta \, d\Omega|}{\|u\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})}} \geq \frac{\int_{\Omega_{\Sigma_k}} |\nabla \theta|^2 \, d\Omega}{\|\nabla \theta\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})}} = \|\nabla \theta\|_{[L^2(\Omega_{\Sigma_k})]^3} \geq \beta \|\theta\|_{H^1(\Omega_{\Sigma_k})}^2.$$

Somit sind die Voraussetzungen von Satz A.21 erfüllt und wir erhalten die Existenz einer Lösung  $z_{0, \Sigma_k}$  mit

$$\|z_{0, \Sigma_k}\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})} \leq C_{\Sigma_k, 1} \|\mathcal{L}_{\Sigma_k}[v]\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})},$$

woraus sich nach Definition von  $z_{v, \Sigma_k}$ , mit zusätzlicher Anwendung der Dreiecksungleichung und (3.13) die Abschätzung

$$\|z_{v, \Sigma_k}\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})} \leq (1 + C_{\Sigma_k, 1}) \|\mathcal{L}_{\Sigma_k}[v]\|_{H(\operatorname{curl}; \Omega_{\Sigma_k})} \leq C_{\Sigma_k, 2} \|v\|_{X_e} \quad (3.14)$$

mit  $C_{\Sigma_k, 2} = C_{\Sigma_k}(1 + C_{\Sigma_k, 1})$  ergibt. Bevor wir den Operator  $\mathcal{L}_\Gamma : X_e \rightarrow H_0(\operatorname{curl}; \Omega_{\Gamma_G, \Gamma_0})$  herleiten, definieren wir den Fortsetzungsoperator  $\mathcal{E}_{\Omega_G} : X_e \rightarrow H(\operatorname{curl}; \Omega_G)$  gegeben durch

$$\mathcal{E}_{\Omega_G}[v] = \begin{cases} v & \text{in } \Omega, \\ z_{v, \Sigma_k} & \text{in } \Omega_{\Sigma_k}, \quad k \in I_{\Sigma, 1}, \\ 0 & \text{in } \Omega_{\Sigma_k}, \quad k \in I_{\Sigma, 2}, \end{cases}$$

für den

$$\|\mathcal{E}_{\Omega_G}[v]\|_{H(\text{curl}; \Omega_G)} \leq \sqrt{1 + \sum_{k \in I_{\Sigma,1}} C_{\Sigma_k,2}^2} \|v\|_{X_e} \quad (3.15)$$

wegen (3.14) folgt. Mit Satz 2.8 erhalten wir wieder die Existenz der tangentialen Spuroperatoren  $\gamma_{t,\Gamma}^+$  von  $\Omega_{\Gamma_G, \Gamma_0}$  nach  $\Gamma$  und  $\gamma_{t,\Gamma}^-$  von  $\Omega_G$  nach  $\Gamma$ , sowie die Existenz des rechtsinversen Operators

$$(\gamma_{t,\Gamma}^+)^{-1} : H^{-\frac{1}{2}}(\text{div}_{\Gamma}; \Gamma) \rightarrow H_0(\text{curl}; \Omega_{\Gamma_G, \Gamma_0}) = \{v \in H(\text{curl}; \Omega_{\Gamma_G, \Gamma_0}) : \gamma_{t,\Gamma_0}[v] = 0\}.$$

Damit können wir den Fortsetzungsoperator folgendermaßen definieren

$$\begin{aligned} \mathcal{L}_{\Gamma} : X_e &\rightarrow H_0(\text{curl}; \Omega_{\Gamma_G, \Gamma_0}) \\ v &\mapsto \mathcal{L}_{\Gamma}[v] = (\gamma_{t,\Gamma}^+)^{-1}[\gamma_{t,\Gamma}^-[\mathcal{E}_{\Omega_G}[v]]]. \end{aligned} \quad (3.16)$$

Wir erhalten wiederum die Linearität und Stetigkeit mit (3.15) aus den Eigenschaften der Tangentialoperatoren und der Definition von  $\mathcal{E}_{\Omega_G}$ . Wie zuvor leiten wir eine Funktion  $z_{v,\Gamma} \in \mathcal{L}_{\Gamma}[v] + H_0(\text{curl}; \Omega_{\Gamma_G, \Gamma_0})$  über das folgende Problem her:

Bestimme  $z_{v,\Gamma} \in \mathcal{L}_{\Gamma}[v] + H_0(\text{curl}; \Omega_{\Gamma_G, \Gamma_0})$  und  $\chi_{\Gamma} \in H_0^1(\Omega_{\Gamma_G, \Gamma_0})$ , die

$$\begin{aligned} [P_{\Gamma}]_{\Omega_{\Gamma_G, \Gamma_0}} \int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times z_{v,\Gamma}) \cdot (\nabla \times q) \, d\Omega + \int_{\Omega_{\Gamma_G, \Gamma_0}} q \cdot \nabla \chi_{\Gamma} \, d\Omega &= 0, \quad q \in H_0(\text{curl}; \Omega_{\Gamma_G, \Gamma_0}), \\ \int_{\Omega_{\Gamma_G, \Gamma_0}} z_{v,\Gamma} \cdot \nabla \theta \, d\Omega &= 0, \quad \theta \in H_0^1(\Omega_{\Gamma_G, \Gamma_0}), \end{aligned}$$

lösen.

Mit denselben Schritten wie zuvor erhalten wir dann

$$\|z_{v,\Gamma}\|_{H(\text{curl}; \Omega_{\Gamma_G, \Gamma_0})} \leq (1 + C_{\Gamma,1}) \|\mathcal{L}_{\Gamma}[v]\|_{H(\text{curl}; \Omega_{\Gamma_G, \Gamma_0})} \leq C_{\Gamma,2} \|v\|_{X_e}. \quad (3.17)$$

Sei nun

$$\mathcal{E}[v] = \begin{cases} v & \text{in } \Omega, \\ z_{v,\Gamma} & \text{in } \Omega_{\Gamma_G, \Gamma_0}, \\ z_{v,\Sigma_k} & \text{in } \Omega_{\Sigma_k}, \quad k \in I_{\Sigma,1}, \end{cases}$$

dann erfüllt  $\mathcal{E}[v]$  die Bedingungen (3.9) und (3.10), wegen der Lösungseigenschaften von  $z_{v,\Gamma}$  zu  $[P_{\Gamma}]$  beziehungsweise von  $z_{v,\Sigma_k}$  zu  $[P_{\Sigma_k}]$ . Mit den Abschätzungen (3.14) und (3.17)

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

erhalten wir dann zum Abschluss

$$\|\mathcal{E}[v]\|_{H(\text{curl}; \tilde{\Omega})} \leq \sqrt{1 + C_{\Gamma,2}^2 + \sum_{k \in I_{\Sigma,1}} C_{\Sigma_k,2}^2} \|v\|_{X_e}. \quad (3.18)$$

□

Damit ist Lemma 3.3 bewiesen. Wir definieren nun den linearen Projektionsoperator

$$\begin{aligned} \mathcal{R} : H_0(\text{curl}; \tilde{\Omega}) &\rightarrow H_0(\text{curl}; \tilde{\Omega}), \\ v &\mapsto \mathcal{R}[v] = v - \nabla \varphi_v, \end{aligned} \quad (3.19)$$

wobei  $\varphi_v \in H_0^1(\tilde{\Omega})$  die eindeutige Lösung von

$$\int_{\tilde{\Omega}} \nabla \varphi_v \cdot \nabla \psi \, d\Omega = \int_{\tilde{\Omega}} v \cdot \nabla \psi \, d\Omega, \quad \psi \in H_0^1(\tilde{\Omega}), \quad (3.20)$$

zu einem festen  $v \in H_0(\text{curl}; \tilde{\Omega})$  ist. Hierbei erhalten wir die Existenz der eindeutigen Lösung  $\varphi_v$  durch Satz A.17. Die Projektionseigenschaften des Operators  $\mathcal{R}$  beweisen wir mit dem nächsten Lemma.

**Lemma 3.4.** *Der Operator  $\mathcal{R}$ , gegeben durch (3.19), ist linear, beschränkt und ein Projektionsoperator mit den Eigenschaften*

$$\nabla \times \mathcal{R}[v] = \nabla \times v, \quad \nabla \cdot \mathcal{R}[v] = 0 \text{ in } \tilde{\Omega} \quad (3.21)$$

für alle  $v \in H_0(\text{curl}; \tilde{\Omega})$ .

*Beweis.* Die Linearität ergibt sich direkt aufgrund der Definition des Operators. Um die Projektionseigenschaft  $(\mathcal{R}\mathcal{R})[v] = \mathcal{R}[v]$  zu zeigen, sei  $v \in H_0(\text{curl}; \tilde{\Omega})$ , denn damit erhalten wir

$$\int_{\tilde{\Omega}} (\nabla \varphi_v - v) \cdot \nabla \psi \, d\Omega = 0, \quad \psi \in H_0^1(\tilde{\Omega}), \quad (3.22)$$

für die eindeutige Lösung  $\varphi_v$ . Wir betrachten nun die Lösungseigenschaft (3.20) von  $\varphi_{(v-\nabla\varphi_v)}$  zu der Funktion  $\mathcal{R}[v] = v - \nabla\varphi_v$ , für die mit (3.22)

$$\int_{\tilde{\Omega}} \nabla \varphi_{(v-\nabla\varphi_v)} \cdot \nabla \psi \, d\Omega = \int_{\tilde{\Omega}} (v - \nabla\varphi_v) \cdot \nabla \psi \, d\Omega = 0, \quad \psi \in H_0^1(\tilde{\Omega}),$$

gilt. Dies ist die schwache Formulierung der Laplace-Gleichung mit homogenen Dirichlet-Randwerten  $\varphi_{(v-\nabla\varphi_v)} = 0$  auf dem Rand  $\partial\tilde{\Omega}$ . Mit der Theorie der Fundamentallösung der

Poisson-Gleichung ergibt sich damit  $\varphi_{(v-\nabla\varphi_v)} = 0$  auf  $\tilde{\Omega}$ . Somit gilt auch  $\nabla\varphi_{(v-\nabla\varphi_v)} = 0$  und damit die Projektionseigenschaft

$$(\mathcal{R}\mathcal{R})[v] = \mathcal{R}[v - \nabla\varphi_v] = v - \nabla\varphi_v - \nabla\varphi_{(v-\nabla\varphi_v)} = v - \nabla\varphi_v = \mathcal{R}[v].$$

Wir betrachten nun (3.22), führen eine partielle Integration durch und erhalten

$$\int_{\tilde{\Omega}} \nabla \cdot (\nabla\varphi_v - v)\psi \, d\Omega = \int_{\tilde{\Omega}} \nabla \cdot \mathcal{R}[v]\psi \, d\Omega = 0, \quad \psi \in H_0^1(\tilde{\Omega}),$$

und somit

$$\nabla \cdot \mathcal{R}[v] = 0$$

mit dem Fundamentallemma der Variationsrechnung. Die Wohldefiniertheit der partiellen Integration und damit auch der eben gezeigten Aussage ergibt sich aus der noch zu zeigenden Beschränktheit des Operators  $\mathcal{R} \in H(\text{curl}; \tilde{\Omega})$ . Aus der Linearität des curl-Operators ergibt sich die zweite Eigenschaft

$$\nabla \times \mathcal{R}[v] = \nabla \times (v - \nabla\varphi_v) = \nabla \times v - \nabla \times \nabla\varphi_v = \nabla \times v.$$

Die noch fehlende Beschränktheit folgt durch Anwenden von (3.20), (3.21) und den zuvor gezeigten Eigenschaften

$$\begin{aligned} \|\mathcal{R}[v]\|_{H(\text{curl}; \tilde{\Omega})}^2 &= \int_{\tilde{\Omega}} |\mathcal{R}[v]|^2 \, d\Omega + \int_{\tilde{\Omega}} |\nabla \times \mathcal{R}[v]|^2 \, d\Omega \\ &= \int_{\tilde{\Omega}} |v|^2 \, d\Omega - \int_{\tilde{\Omega}} |\nabla\varphi_v|^2 \, d\Omega + \int_{\tilde{\Omega}} |\nabla \times v|^2 \, d\Omega \\ &= \|v\|_{H(\text{curl}; \tilde{\Omega})}^2 - \|\nabla\varphi_v\|_{[L^2(Q)]^3}^2 \\ &\leq \|v\|_{H(\text{curl}; \tilde{\Omega})}^2, \end{aligned}$$

wobei in der zweiten Gleichheit zusätzlich zur Definition des Projektionsoperators  $\mathcal{R}$  auch die Bedingung (3.20) verwendet wird.  $\square$

Als nächstes betrachten wir den Einschränkungoperator  $\mathcal{T}$  von  $H_0(\text{curl}; \tilde{\Omega})$  nach  $X_e$ .

**Lemma 3.5.** *Der lineare Operator*

$$\begin{aligned} \mathcal{T} : H_0(\text{curl}; \tilde{\Omega}) &\rightarrow X_e, \\ v &\mapsto \mathcal{T}[v] = (\mathcal{R}[v])|_{\Omega}, \end{aligned} \tag{3.23}$$

*ist beschränkt.*

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

*Beweis.* Mit Lemma 3.4, Lemma A.15 und der Stetigkeit der tangentialen Operatoren aus (2.49) folgt

$$\begin{aligned} \|\mathcal{T}[v]\|_{X_e}^2 &= \|\mathcal{T}[v]\|_{H(\text{curl}; \Omega)}^2 + \|\pi_{t, \Gamma_P}[\mathcal{T}[v]]\|_{[L^2(\Gamma_P)]^3}^2 + \sum_{k \in I_{\Sigma, 1}} \|\pi_{t, \tilde{\Sigma}_k}[\mathcal{T}[v]]\|_{[L^2(\tilde{\Sigma}_k)]^3}^2 \\ &\leq C_1 \|v\|_{H(\text{curl}; \tilde{\Omega})}^2 + \|\pi_{t, \Gamma}[\mathcal{R}[v]]\|_{H_{\parallel}^{s_{\Gamma} - \frac{1}{2}}(\Gamma)}^2 + \sum_{k \in I_{\Sigma, 1}} \|\pi_{t, \Sigma_k}[\mathcal{R}[v]]\|_{H_{\parallel}^{s_{\Sigma_k} - \frac{1}{2}}(\Sigma_k)}^2 \\ &\leq C_1 \|v\|_{H(\text{curl}; \tilde{\Omega})}^2 + C_{\Gamma, 1} \|\mathcal{R}[v]\|_{[H^{s_{\Gamma}}(\Omega_{\Gamma_G, \Gamma_0})]^3}^2 + \sum_{k \in I_{\Sigma, 1}} C_{\Sigma_k, 1} \|\mathcal{R}[v]\|_{[H^{s_{\Sigma_k}}(\Omega_{\Sigma_k})]^3}^2 \end{aligned}$$

für  $s_{\Gamma}, s_{\Sigma_k} \in (\frac{1}{2}, 1)$  und für alle  $k \in I_{\Sigma, 1}$ . Daraus ergibt sich nun mit Lemma A.15 und (3.21) gerade

$$\|\mathcal{T}[v]\|_{X_e}^2 \leq C_1 \|v\|_{H(\text{curl}; \tilde{\Omega})}^2 + C_{\Gamma, 2} \|\mathcal{R}[v]\|_{H(\text{curl}; \tilde{\Omega})}^2 + \sum_{k \in I_{\Sigma, 1}} C_{\Sigma_k, 2} \|\mathcal{R}[v]\|_{H(\text{curl}; \tilde{\Omega})}^2$$

und abschließend mit der Beschränktheit von  $\mathcal{R}$  aus Lemma 3.4

$$\|\mathcal{T}[v]\|_{X_e}^2 \leq C \|v\|_{H(\text{curl}; \tilde{\Omega})}^2.$$

□

Damit erhalten wir also die Stetigkeit des Operators  $\mathcal{T}$ , wodurch wir nun den Projektionsoperator

$$\begin{aligned} \mathcal{P} : X_e &\rightarrow X_e, \\ v &\mapsto \mathcal{P}[v] = (\mathcal{T}\mathcal{E})[v] = ((\mathcal{R}\mathcal{E})[v])|_{\Omega}, \end{aligned} \tag{3.24}$$

als Komposition der zuvor eingeführten Operatoren definieren und folgendes Lemma zeigen können.

**Lemma 3.6.** *Der lineare Projektionsoperator  $\mathcal{P}$ , gegeben durch (3.24), besitzt die Eigenschaften*

$$\nabla \times \mathcal{P}[v] = \nabla \times v, \quad \nabla \cdot \mathcal{P}[v] = 0 \text{ in } \Omega \tag{3.25}$$

für alle  $v \in X_e$  und ist beschränkt.

*Beweis.* Die curl-Eigenschaft in (3.25) folgt aus der curl-Eigenschaft des Operators  $\mathcal{R}$ , aus Lemma 3.4, durch

$$\nabla \times \mathcal{P}[v] = (\nabla \times (\mathcal{R}\mathcal{E})[v])|_{\Omega} = (\nabla \times \mathcal{E}[v])|_{\Omega} = \nabla \times ((\mathcal{E}[v])|_{\Omega}) = \nabla \times v, \quad v \in X_e.$$

Auch die div-Eigenschaft des Operators  $\mathcal{P}$  in (3.25) folgt aus Lemma 3.4. Wir betrachten

nun

$$(\mathcal{P}\mathcal{P})[v] = ((\mathcal{R}\mathcal{E})[((\mathcal{R}\mathcal{E})[v])|_{\Omega}])|_{\Omega}$$

und zeigen dass  $\mathcal{E}[((\mathcal{R}\mathcal{E})[v])|_{\Omega}] = (\mathcal{R}\mathcal{E})[v]$ , denn dann ergibt sich

$$(\mathcal{P}\mathcal{P})[v] = ((\mathcal{R}\mathcal{R}\mathcal{E})[v])|_{\Omega} = ((\mathcal{R}\mathcal{E})[v])|_{\Omega} = \mathcal{P}[v],$$

mit der Projektionseigenschaft von  $\mathcal{R}$ . Wir definieren

$$z = \mathcal{E}[((\mathcal{R}\mathcal{E})[v])|_{\Omega}] - (\mathcal{R}\mathcal{E})[v],$$

denn dabei gilt

$$z|_{\Omega} = (\mathcal{E}[((\mathcal{R}\mathcal{E})[v])|_{\Omega}])|_{\Omega} - ((\mathcal{R}\mathcal{E})[v])|_{\Omega} = ((\mathcal{R}\mathcal{E})[v])|_{\Omega} - ((\mathcal{R}\mathcal{E})[v])|_{\Omega} = 0.$$

Des Weiteren ergibt sich aus (3.9) und Lemma 3.4

$$\begin{aligned} \int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times z) \cdot (\nabla \times z) \, d\Omega &= \int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times (\mathcal{E}[((\mathcal{R}\mathcal{E})[v])|_{\Omega}])) \cdot (\nabla \times z) \, d\Omega \\ &\quad - \int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times ((\mathcal{R}\mathcal{E})[v])) \cdot (\nabla \times z) \, d\Omega \\ &= - \int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times ((\mathcal{R}\mathcal{E})[v])) \cdot (\nabla \times z) \, d\Omega \\ &= - \int_{\Omega_{\Gamma_G, \Gamma_0}} (\nabla \times (\mathcal{E}[v])) \cdot (\nabla \times z) \, d\Omega \\ &= 0, \end{aligned}$$

also  $\nabla \times z = 0$  in  $\Omega_{\Gamma_G, \Gamma_0}$ . Wegen Lemma 3.2 ergibt sich somit  $z = 0$  in  $\Omega_{\Gamma_G, \Gamma_0}$ . Analog folgt auch  $z = 0$  in  $\Omega_{\Sigma_k}$ ,  $k \in I_{\Sigma, 1}$ , und somit  $z = 0$  in  $\tilde{\Omega}$ , was gerade zu zeigen war. Die Beschränktheit des Operators  $\mathcal{P}$  folgt aus der Beschränktheit der Operatoren  $\mathcal{T}$  und  $\mathcal{E}$ .  $\square$

Wir erhalten somit die Helmholtz-Zerlegung des Lösungsraums

$$X_e = \mathcal{P}[X_e] \oplus (\mathcal{I} - \mathcal{P})[X_e], \quad (3.26)$$

wobei  $\mathcal{I}$  der Identitätsoperator ist, in den curl-freien Teilraum  $(\mathcal{I} - \mathcal{P})[X_e]$  und den

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

div-freien Teilraum  $\mathcal{P}[X_e]$ , denn für eine Funktion  $u \in (\mathcal{I} - \mathcal{P})[X_e]$  gilt

$$\nabla \times u = \nabla \times (\mathcal{I} - \mathcal{P})[v] = \nabla \times \mathcal{I}[v] - \nabla \times \mathcal{P}[v] = \nabla \times v - \nabla \times v = 0, \quad (3.27)$$

bzw. für  $u \in \mathcal{P}[X_e]$  folgt

$$\nabla \cdot u = \nabla \cdot \mathcal{P}[v] = 0, \quad (3.28)$$

für  $v \in X_e$ . Damit erhalten wir für eine Funktion  $v \in X_e$  die eindeutige Zerlegung

$$v = \mathcal{P}[v] + (\mathcal{I} - \mathcal{P})[v]$$

und führen die Notation  $v_0 = \mathcal{P}[v]$  und  $v_\perp = (\mathcal{I} - \mathcal{P})[v]$  ein. Zusätzlich dazu definieren wir die Norm

$$\|v\|_{X_e} = (\|v_0\|_{X_e}^2 + \|v_\perp\|_{X_e}^2)^{\frac{1}{2}}, \quad (3.29)$$

die äquivalent zu  $\|v\|_{X_e}$  auf  $X_e$  ist. Nach [40, Lemma 5] gilt für die Operatornorm

$$\|\mathcal{P}\|_{\mathcal{L}(X_e, X_e)} = \|\mathcal{I} - \mathcal{P}\|_{\mathcal{L}(X_e, X_e)}$$

und damit

$$\frac{1}{\sqrt{2}\|\mathcal{P}\|_{\mathcal{L}(X_e, X_e)}} \|v\|_{X_e} \leq \|v\|_{X_e} \leq \sqrt{2} \|v\|_{X_e}, \quad v \in X_e. \quad (3.30)$$

Um Satz 3.1 anwenden zu können, benötigen wir kompakte Operatoren.

**Lemma 3.7.** *Die Einbettungsabbildungen  $\mathcal{P} : X_e \hookrightarrow [L^2(\Omega)]^3$ ,  $\pi_{t, \Gamma_P} \circ \mathcal{P} : X_e \hookrightarrow L_t^2(\Gamma_P)$  und  $\pi_{t, \tilde{\Sigma}_k} \circ \mathcal{P} : X_e \hookrightarrow L_t^2(\tilde{\Sigma}_k)$ ,  $k \in I_{\Sigma, 1}$ , sind kompakt.*

*Beweis.* Sei  $s \in (\frac{1}{2}, 1]$ . Die erste Kompaktheitsaussage folgt aus Lemma A.15 und der kompakten Einbettung

$$H^s(\Omega) \hookrightarrow_C L^2(\Omega)$$

nach Satz A.13. Nach [41, Lemma 3.2] ist

$$H_{\parallel}^{s-\frac{1}{2}}(\Gamma) \hookrightarrow_C L_t^2(\Gamma)$$

und damit ist auch  $\pi_{t, \Gamma_P} \circ \mathcal{P} : X_e \rightarrow L_t^2(\Gamma_P)$  kompakt. Analog folgt die Aussage auch für  $\tilde{\Sigma}_k$ .  $\square$

### 3.1.2. Existenz der Lösung

Wir definieren die Sesquilinearform

$$\mathbf{a}(E, \varphi) = \mathbf{a}_1(E, \varphi) + \mathbf{a}_2(E, \varphi)$$

mit

$$\begin{aligned} \mathbf{a}_1(E, \varphi) &= \int_{\Omega} \frac{1}{\mu_r} (\nabla \times E) \cdot (\nabla \times \varphi) \, d\Omega + \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 E \cdot \varphi \, d\Omega, \\ \mathbf{a}_2(E, \varphi) &= \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 \pi_t[E] \cdot \pi_t[\varphi] \, d\Gamma + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \pi_t[E] \cdot \pi_t[\varphi] \, d\Gamma. \end{aligned}$$

Um die Notation zu vereinfachen, schreiben wir  $\pi_t$  für jegliche Projektionsoperatoren auf den verschiedenen Oberflächen. Dabei entsprechen diese gerade den Operatoren  $\pi_{t,\Gamma}$  und  $\pi_{t,\tilde{\Sigma}_k}$  aus Abschnitt 3.1.1 der jeweiligen Randbedingung. Mithilfe des Projektionsoperators  $\mathcal{P}$ , der Helmholtz-Zerlegung (3.26), sowie der Linearität des curl-Operators und der Projektionsoperatoren  $\pi_t$  können wir die Sesquilinearform  $a(E, \varphi)$  in (3.3) umformen zu

$$\begin{aligned} a(E, \varphi) &= \mathbf{a}(E_0, \varphi_0) + \mathbf{a}(E_0, \varphi_{\perp}) + \mathbf{a}(E_{\perp}, \varphi_0) + \mathbf{a}(E_{\perp}, \varphi_{\perp}) \\ &\quad - 2 \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 (I_{E,\varphi}) \, d\Omega - 2 \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 (II_{E,\varphi}) \, d\Gamma \\ &\quad - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 (II_{E,\varphi}) \, d\Gamma \end{aligned}$$

mit den Termbezeichnungen

$$I_{E,\varphi} = E_0 \cdot \varphi_0 + E_0 \cdot \varphi_{\perp} + E_{\perp} \cdot \varphi_0 + E_{\perp} \cdot \varphi_{\perp}$$

und

$$II_{E,\varphi} = \pi_t[E_0] \cdot \pi_t[\varphi_0] + \pi_t[E_0] \cdot \pi_t[\varphi_{\perp}] + \pi_t[E_{\perp}] \cdot \pi_t[\varphi_0] + \pi_t[E_{\perp}] \cdot \pi_t[\varphi_{\perp}].$$

Mit der Definition von  $\mathbf{a}(\cdot, \cdot)$ , sowie (3.27) folgt

$$a(E, \varphi) = A(E, \varphi) + K(E, \varphi),$$

mit den Sesquilinearformen

$$A(E, \varphi) = \mathbf{a}(E_0, \varphi_0) - \mathbf{a}(E_{\perp}, \varphi_{\perp}) \tag{3.31}$$

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

und

$$\begin{aligned}
 K(E, \varphi) = & a(E_0, \varphi_\perp) + a(E_\perp, \varphi_0) - 2 \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 E_0 \cdot \varphi_0 \, d\Omega \\
 & - 2 \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 \pi_t[E_0] \cdot \pi_t[\varphi_0] \, d\Gamma - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \pi_t[E_0] \cdot \pi_t[\varphi_0] \, d\Gamma.
 \end{aligned} \tag{3.32}$$

Wir definieren die Operatoren

$$\begin{aligned}
 \mathcal{A} : X_e & \rightarrow (X_e)', \\
 E & \mapsto \mathcal{A}(E) = A(E, \cdot),
 \end{aligned} \tag{3.33}$$

und

$$\begin{aligned}
 \mathcal{K} : X_e & \rightarrow (X_e)', \\
 E & \mapsto \mathcal{K}(E) = K(E, \cdot),
 \end{aligned} \tag{3.34}$$

und das Dualraumelement

$$\mathcal{F} \in (X_e)', \quad \mathcal{F} = f(\cdot). \tag{3.35}$$

Damit ergibt sich gerade die zu  $[V_K]$  äquivalente Problemformulierung:

Bestimme  $E \in X_e$ , sodass

$$(\mathcal{A} + \mathcal{K})[E] = \mathcal{F}. \tag{3.36}$$

Um Satz 3.1 anwenden zu können, beweisen wir, dass  $A(\cdot, \cdot)$  die inf-sup-Bedingung erfüllt, womit wir mit dem Nečas-Satz, siehe [42, Satz 3.2.3], die Isomorphismus-Eigenschaft des Operators  $\mathcal{A}$  erhalten.

**Lemma 3.8.** *Es existiert  $\beta > 0$  sodass*

$$\sup_{\varphi \in X_e \setminus \{0\}} \frac{|A(E, \varphi)|}{\|\varphi\|_{X_e}} \geq \beta \|E\|_{X_e}, \quad E \in X_e,$$

und es gilt

$$\sup_{E \in X_e} |A(E, \varphi)| > 0, \quad \varphi \in X_e \setminus \{0\}.$$

*Beweis.* Sei  $\Theta : X_e \rightarrow X_e$  mit

$$\Theta[E] = \begin{cases} (1+i) (2\mathcal{P} - \mathcal{I})[E], & x \in \Omega, \\ (-1+i) (2\mathcal{P} - \mathcal{I})[E], & x \in \Gamma_R = \Gamma_{IMP} \cup \Gamma_{PORT}. \end{cases}$$

Wegen der Stetigkeit und Linearität von  $\mathcal{P}$  ist auch  $\Theta$  linear und stetig, also

$$\|\Theta[E]\|_{X_e} \leq C_\Theta \|E\|_{X_e}.$$

Es gilt somit

$$\sup_{\varphi \in X_e \setminus \{0\}} \frac{|A(E, \varphi)|}{\|\varphi\|_{X_e}} \geq \frac{|A(E, \Theta[E])|}{\|\Theta[E]\|_{X_e}} \geq \frac{\operatorname{Re}\{A(E, \Theta[E])\}}{\|\Theta[E]\|_{X_e}} \quad (3.37)$$

für  $E \in X_e$ . Wegen der Projektionseigenschaft von  $\mathcal{P}$  gilt

$$\mathcal{P}(2\mathcal{P} - \mathcal{I}) = \mathcal{P}, \quad (\mathcal{I} - \mathcal{P})(2\mathcal{P} - \mathcal{I}) = -(\mathcal{I} - \mathcal{P}),$$

und mit der Linearität der Sesquilinearform  $A$  folgt

$$\begin{aligned} A(E, \Theta[E]) &= (1 - i)\mathfrak{a}_1(E_0, E_0) + (1 - i)\mathfrak{a}_1(E_\perp, E_\perp) \\ &\quad - (1 + i)\mathfrak{a}_2(E_0, E_0) - (1 + i)\mathfrak{a}_2(E_\perp, E_\perp) \\ &= (1 - i)\mathfrak{a}_1(E, E) - (1 + i)\mathfrak{a}_2(E, E). \end{aligned} \quad (3.38)$$

Betrachten wir zunächst den Realteil für den ersten Term

$$\begin{aligned} \operatorname{Re}\{(1 - i)\mathfrak{a}_1(E, E)\} &= \int_{\Omega} \omega^2 \varepsilon_0 (\operatorname{Re}(\varepsilon_r(\omega)) + \operatorname{Im}(\varepsilon_r(\omega))) \mu_0 |E|^2 d\Omega \\ &\quad + \int_{\Omega} \frac{1}{\mu_r} |\nabla \times E|^2 d\Omega, \end{aligned}$$

dann ergibt sich

$$\operatorname{Re}\{(1 - i)\mathfrak{a}_1(E, E)\} \geq C_{E,1} \|E\|_{H(\operatorname{curl}; \Omega)}^2 \quad (3.39)$$

aus Voraussetzung 2.10 mit

$$C_{E,1} = \min \left\{ \frac{1}{\mu_{\max}}, \omega_{\min}^2 \varepsilon_0 (\varepsilon_{\min, \operatorname{Re}} + \varepsilon_{\min, \operatorname{Im}}) \mu_0 \right\}.$$

Des Weiteren erhalten wir ebenso

$$\begin{aligned} \operatorname{Re}\{-(1 + i)\mathfrak{a}_2(E, E)\} &= \int_{\Gamma_{IMP}} \omega (\operatorname{Re}(Y(\omega)) + \operatorname{Im}(Y(\omega))) \mu_0 |\pi_t[E]|^2 d\Gamma \\ &\quad + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \omega (\operatorname{Re}(Y(\omega)) + \operatorname{Im}(Y(\omega))) \mu_0 |\pi_t[E]|^2 d\Gamma, \end{aligned}$$

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

wodurch wir

$$\operatorname{Re} \{-(1+i) \mathbf{a}_2(E, E)\} \geq C_{E,2} \left( \|\pi_t[E]\|_{L^2(\Gamma_{IMP})}^2 + \sum_{l=1}^{K_A} \|\pi_t[E]\|_{L^2(\Gamma_{P,l})}^2 \right) \quad (3.40)$$

mit

$$C_{E,2} = \omega_{\min}(Y_{\min, \operatorname{Re}} + Y_{\min, \operatorname{Im}}) \mu_0$$

folgern können. Setzen wir (3.38) in (3.37) ein und verwenden (3.39) und (3.40), dann ergibt sich

$$\sup_{\varphi \in X_e \setminus \{0\}} \frac{|A(E, \varphi)|}{\|\varphi\|_{X_e}} \geq \min \{C_{E,1}, C_{E,2}\} \frac{\|E\|_{X_e}^2}{\|\Theta[E]\|_{X_e}}$$

für alle  $E \in X_e$ . Wenden wir abschließend die Normäquivalenz (3.30) und die Beschränktheit des Operators  $\Theta$  an, so folgt

$$\sup_{\varphi \in X_e \setminus \{0\}} \frac{|A(E, \varphi)|}{\|\varphi\|_{X_e}} \geq \beta \|E\|_{X_e}$$

mit

$$\beta = \frac{\min \{C_{E,1}, C_{E,2}\}}{2C_\Theta}$$

und damit die erste Aussage des Lemmas. Aufgrund der Symmetrie von  $A$  und der eben gezeigten inf-sup-Bedingung folgt auch die zweite Aussage.  $\square$

Als nächstes zeigen wir noch die Kompaktheit des Operators  $\mathcal{K}$ .

**Lemma 3.9.** *Der Operator  $\mathcal{K}$ , gegeben durch (3.34), ist kompakt.*

*Beweis.* Durch Verwendung von (3.27) erhalten wir

$$\begin{aligned} K(E, \varphi) &= - \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 (E_0 \cdot \varphi_\perp + E_\perp \cdot \varphi_0) \, d\Omega \quad (3.41) \\ &\quad - \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 (\pi_t[E_0] \cdot \pi_t[\varphi_\perp] + \pi_t[E_\perp] \cdot \pi_t[\varphi_0]) \, d\Gamma \\ &\quad - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 (\pi_t[E_0] \cdot \pi_t[\varphi_\perp] + \pi_t[E_\perp] \cdot \pi_t[\varphi_0]) \, d\Gamma \\ &\quad - 2 \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 E_0 \cdot \varphi_0 \, d\Omega - 2 \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 \pi_t[E_0] \cdot \pi_t[\varphi_0] \, d\Gamma \\ &\quad - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \pi_t[E_0] \cdot \pi_t[\varphi_0] \, d\Gamma. \end{aligned}$$

Die Abbildungen auf die Integranden in den letzten drei Integraltermen sind aufgrund von Lemma 3.7 kompakt. Wir müssen für die Kompaktheit des Operators  $\mathcal{K}$  noch die Kompaktheit in den vorderen Integraltermen zeigen. Hierzu betrachten wir den Operator  $\mathcal{G}_E$  gegeben durch

$$\begin{aligned}\mathcal{G}_E : X_e &\rightarrow \mathbb{C}, \\ v &\mapsto \mathcal{G}_E[v] = \int_{\Omega} p E_0 \cdot v_{\perp} d\Omega,\end{aligned}$$

mit dem Parameter

$$p = -\omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0.$$

Die Linearität erhalten wir aus der Linearität des Projektionsoperators  $\mathcal{P}$  und die Beschränktheit folgt aus der Hölder-Ungleichung und der Beschränktheit von  $\mathcal{P}$ . Damit gilt  $\mathcal{G}_E \in (X_e)'$ . Analog zeigen wir, dass der Operator  $\mathcal{G}_v$  gegeben durch

$$\begin{aligned}\mathcal{G}_v : X_e &\rightarrow \mathbb{C}, \\ E &\mapsto \mathcal{G}_v[E] = \int_{\Omega} p E_0 \cdot v_{\perp} d\Omega,\end{aligned}$$

Element des Dualraums  $(X_e)'$  ist. Wir betrachten den Operator

$$\begin{aligned}\mathcal{G} : X_e &\rightarrow (X_e)', \\ E &\mapsto \mathcal{G}[E] = \mathcal{G}_E,\end{aligned}$$

für den wir die Kompaktheit zeigen müssen. Sei  $M \subset X_e$  beschränkt. Zeigen wir, dass jede Folge in  $\overline{\mathcal{G}[M]} = \overline{\mathcal{G}[\overline{M}]}$  eine konvergente Teilfolge besitzt, dann ist  $\mathcal{G}$  kompakt. Sei dazu  $(\mathcal{G}_E)_k$  Folge in  $\overline{\mathcal{G}[\overline{M}]}$ , dann wird diese durch eine Folge  $(E_k)_k$  in  $\overline{M}$  erzeugt, also

$$(\mathcal{G}_E)_k = \mathcal{G}[E_k]$$

für alle  $k \in \mathbb{N}$ . Da  $M$  beschränkt im Hilbertraum  $X_e$  ist, ist die Folge  $(E_k)_k$  in  $X_e$  beschränkt. Damit existiert eine Teilfolge  $(E_{k_l})_l$  von  $(E_k)_k$ , die fortan wiederum mit  $(E_k)_k$  bezeichnet wird, die schwach gegen  $E \in \overline{M}$  konvergiert. Wir zeigen nun die schwach-\* Konvergenz von  $(\mathcal{G}[E_k])_k$  und die Normkonvergenz, was im Hilbertraum gerade der starken Konvergenz entspricht. Aus der schwachen Konvergenz der Folge  $(E_k)_k$  erhalten wir die schwach-\* Konvergenz

$$|\mathcal{G}_{E_k}[v] - \mathcal{G}_E[v]| \leq C \left| \int_{\Omega} ((E_k)_0 - E_0) \cdot v_{\perp} d\Omega \right| \xrightarrow{k \rightarrow \infty} 0$$

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

für alle  $v \in X_e$ , denn

$$\mathcal{T}_v[E] = \int_{\Omega} E_0 \cdot v_{\perp} \, d\Omega$$

ist Element des Dualraums  $(X_e)'$ . Wegen Satz A.16 können wir jedes Folgenglied  $\mathcal{G}[E_k]$  und auch den Grenzwert  $\mathcal{G}[E]$  mit einem Element  $v_k$  beziehungsweise  $v$  in  $X_e$  identifizieren. Das heißt es gilt zum einen die Gleichheit der Normen

$$\|\mathcal{G}[E_k]\|_{(X_e)'} = \|v_k\|_{X_e}, \quad \|\mathcal{G}[E]\|_{(X_e)'} = \|v\|_{X_e}$$

und die Abbildungseigenschaft

$$\mathcal{G}_{E_k}[w] = \langle v_k, w \rangle_{X_e}, \quad \mathcal{G}_E[w] = \langle v, w \rangle_{X_e}$$

für alle  $w \in X_e$ . Zudem erhalten wir aus der schwach-\* Konvergenz von  $(\mathcal{G}[E_k])_k$  direkt die schwache Konvergenz der Folge  $(v_k)_k$  gegen  $v$  und damit auch die Beschränktheit der Folge  $(v_k)_k$  in  $X_e$ . Aus der Gleichheit der Normen ergibt sich somit

$$\begin{aligned} | \|\mathcal{G}[E_k]\|_{(X_e)'}^2 - \|\mathcal{G}[E]\|_{(X_e)'}^2 | &= | \|v_k\|_{X_e}^2 - \|v\|_{X_e}^2 | = | \langle v_k, v_k \rangle_{X_e} - \langle v, v \rangle_{X_e} | \\ &= | \mathcal{G}_{E_k}[v_k] - \mathcal{G}_E[v] | \\ &\leq C \left| \int_{\Omega} (E_k)_0 \cdot (v_k)_{\perp} \, d\Omega - \int_{\Omega} E_0 \cdot v_{\perp} \, d\Omega \right| \\ &\leq C \| (E_k)_0 - E_0 \|_{[L^2(\Omega)]^3} \| (v_k)_{\perp} \|_{[L^2(\Omega)]^3} + C \left| \int_{\Omega} E_0 \cdot (v_k - v)_{\perp} \, d\Omega \right|. \end{aligned}$$

Der erste Term auf der rechten Seite konvergiert gegen Null, denn  $(v_k)_{\perp}$  ist wegen  $v_k$  beschränkt und es gilt

$$\| (E_k)_0 - E_0 \|_{[L^2(\Omega)]^3} = \| \mathcal{P}[E_k - E] \|_{[L^2(\Omega)]^3} \xrightarrow{k \rightarrow \infty} 0$$

aufgrund der kompakten Abbildung  $\mathcal{P} : X_e \rightarrow [L^2(\Omega)]^3$  nach Lemma 3.7. Der zweite Term auf der rechten Seite konvergiert aufgrund der schwachen Konvergenz der Folge  $v_k$  ebenso gegen Null, denn

$$\mathcal{T}_E[v] = \int_{\Omega} E_0 \cdot v_{\perp} \, d\Omega$$

ist wiederum ein Element des Dualraums  $(X_e)'$ . Mit den selben Argumenten folgt analog auch die Kompaktheit in den restlichen Integralen in (3.41), denn auch  $L_t^2(\Gamma)$  ist ein Hilbertraum.  $\square$

Wegen Lemma 3.8 und dem Nečas-Satz existiert damit die Inverse  $\mathcal{A}^{-1}$  des Operators  $\mathcal{A}$ , weswegen wir (3.36) umformen können zu

$$(\mathcal{I} + \mathcal{A}^{-1}\mathcal{K})[E] = \mathcal{A}^{-1}\mathcal{F}.$$

Mit der Kompaktheit von  $\mathcal{K}$  ist auch  $\mathcal{A}^{-1}\mathcal{K}$  kompakt. Unter der Annahme, dass das homogene Problem

$$(\mathcal{I} + \mathcal{A}^{-1}\mathcal{K})[E] = 0,$$

mit  $\mathcal{F} = 0$ , nur die triviale Lösung besitzt, haben wir mit Satz 3.1 den folgenden Satz bewiesen.

**Satz 3.10.** *Besitzt die Gleichung*

$$A(E, \varphi) + K(E, \varphi) = 0$$

*nur die triviale Lösung, dann existiert zu gegebenen Modenfeldern  $[E_{0,l}]_{l=1}^{K_A}$  eine eindeutige Lösung  $E \in X_e$  zu*

$$a(E, \varphi) = A(E, \varphi) + K(E, \varphi) = f(\varphi, [E_{0,l}]_{l=1}^{K_A}), \quad \varphi \in X_e,$$

*und es existiert eine Konstante  $C > 0$ , sodass*

$$\|E\|_{X_e} \leq C \|\mathcal{F}\|_{(X_e)'}. \quad (3.42)$$

*Beweis.* Betrachten wir die Voraussetzung

$$A(E, \varphi) + K(E, \varphi) = a(E, \varphi) = 0,$$

dann verschwindet die rechte Seite. Dies wiederum bedeutet  $P_{in,l} = 0$ ,  $1 \leq l \leq K_A$ , und damit wird kein Feld an den PORT-Oberflächen angeregt. Da wir zudem keine Quellenterme im Inneren definiert haben folgt damit  $E = 0$ . Mit den Definitionen (3.33), (3.34), sowie Lemma 3.8 und Lemma 3.9, erhalten wir die Operatorgleichung

$$(\mathcal{I} + \mathcal{A}^{-1}\mathcal{K})[E] = \mathcal{A}^{-1}\mathcal{F},$$

mit

$$\mathcal{F} \in (X_e)', \quad \mathcal{F} = f(\cdot).$$

Mit der eben hergeleiteten Operatorgleichung und  $E = 0$  bei der Betrachtung der Voraussetzung folgt die Aussage durch anwenden von Satz 3.1.  $\square$

Eine direkte Folgerung von Satz 3.10 ist die Existenz einer globalen inf-sup-Konstante  $\beta$ .

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

**Korollar 3.11.** *Zur Sesquilinearform (3.3) existiert die globale inf-sup-Konstante*

$$\beta = \inf_{u \in X_e \setminus \{0\}} \sup_{v \in X_e \setminus \{0\}} \frac{|a(u, v)|}{\|u\|_{X_e} \|v\|_{X_e}} > 0. \quad (3.43)$$

*Beweis.* Die Existenz folgt dabei direkt aus der Anwendung von Satz 3.1 und der stetigen Abhängigkeit von der rechten Seite. Sei dazu  $\tilde{E} \in X_e$ ,  $\tilde{E} \neq 0$ , beliebig aber fest. Wir definieren dann die neue rechte Seite

$$\tilde{f}(\varphi) = a(\tilde{E}, \varphi)$$

und betrachten Lösungen  $E \in X_e$  die das folgende Variationsproblem

$$a(E, \varphi) = \tilde{f}(\varphi), \quad \varphi \in X_e, \quad (3.44)$$

lösen. Wie in Satz 3.10 erhalten wir wieder

$$(\mathcal{I} + \mathcal{A}^{-1}\mathcal{K})[E] = \mathcal{A}^{-1}\mathcal{F},$$

mit

$$\mathcal{F} \in (X_e)', \quad \mathcal{F} = \tilde{f}(\cdot).$$

Dabei folgt  $\mathcal{F} \in (X_e)'$  aus der Linearität und Stetigkeit der Sesquilinearform  $a$ . Wenden wir nun Satz 3.1 an, so erhalten wir die Existenz einer eindeutigen Lösung von (3.44) und die stetige Abhängigkeit der Lösung von der rechten Seite

$$\|E\|_{X_e} \leq C \|\mathcal{F}\|_{(X_e)'}$$

mit einer Konstanten  $C > 0$  unabhängig von  $E$  und  $\mathcal{F}$ . Aus dieser Ungleichung ergibt sich

$$\frac{1}{C} \leq \frac{\|\mathcal{F}\|_{(X_e)'}}{\|E\|_{X_e}} = \sup_{v \in X_e \setminus \{0\}} \frac{|\tilde{f}(v)|}{\|E\|_{X_e} \|v\|_{X_e}}.$$

Da  $E \in X_e$  Lösung von (3.44) ergibt sich

$$\frac{1}{C} \leq \sup_{v \in X_e \setminus \{0\}} \frac{|a(E, v)|}{\|E\|_{X_e} \|v\|_{X_e}}.$$

Betrachten wir nun die Definition unserer rechten Seite  $\tilde{f}$  von (3.44), dann ergibt sich gerade

$$a(E, \varphi) = a(\tilde{E}, \varphi), \quad \varphi \in X_e.$$

Da  $\tilde{E}$  diese Gleichung immer erfüllt und die Lösung eindeutig ist gilt  $E = \tilde{E}$ . Wir haben damit gezeigt, dass für ein beliebiges  $\tilde{E} \in X_e$ ,  $\tilde{E} \neq 0$ , eine von  $\tilde{E}$  unabhängige Konstante

$C > 0$  existiert, sodass

$$0 < \frac{1}{C} \leq \sup_{v \in X_e \setminus \{0\}} \frac{|a(\tilde{E}, v)|}{\|\tilde{E}\|_{X_e} \|v\|_{X_e}}$$

gilt. Damit existiert auch ein  $\beta > 0$  mit

$$\beta \geq \frac{1}{C} > 0.$$

□

Wir leiten nun in dem nächsten Abschnitt das diskrete variationelle Problem zu  $[V_K]$  her.

### 3.2. Finite-Elemente-Methode

Um  $[V_K]$  zu diskretisieren gibt es unterschiedliche Methoden, wie beispielsweise die Finiten-Differenzen, die Finiten-Volumen oder die hier verwendete Finite-Elemente-Methode (FEM). In der FEM wird das Problem  $[V_K]$  durch eine Approximation des Raumes  $X_e$  gelöst. Hierzu wird ein Teilraum  $X_N \subset X_e$  mit endlicher Basis  $B = \{\varphi_1, \varphi_2, \dots, \varphi_N\}$  betrachtet. Wir approximieren eine Lösung  $E$  des Variationsproblems  $[V_K]$  durch die FEM-Funktion  $E_N \in X_N$ . Dazu nutzen wir die Basisdarstellung

$$E_N = \sum_{k=1}^N E_k^N \varphi_k,$$

und die Linearität der Sesquilinearform  $a(\cdot, \cdot)$  aus, verwenden die Basisfunktionen  $\varphi_k$  als Testfunktionen und gewährleisten die Einhaltung der Randbedingung

$$\gamma_t[E_N] = 0$$

auf  $\Gamma_{PEC}$ . Dann ergibt sich daraus ein lineares Gleichungssystem

$$A E^N = f, \tag{3.45}$$

wobei die Matrix  $A \in \mathbb{C}^{N \times N}$  und der Vektor  $f \in \mathbb{C}^N$  die Einträge  $A_{l,k} = a(\varphi_k, \varphi_l)$ ,  $f_l = f(\varphi_l)$  besitzen und  $E^N = (E_1^N, E_2^N, \dots, E_N^N)^T$  der komplexe Koeffizientenvektor ist. Bei der Herleitung des finiten Elements folgen wir [32, Kapitel 5]. Um  $X_N \subset X_e$  zu bestimmen, führen wir zunächst eine reguläre Zerlegung des Gebietes  $\Omega$  ein.

**Definition 3.12.** Sei  $M \in \mathbb{N}$ . Die Menge

$$\mathcal{T}_h(\Omega) = \{K_l : 1 \leq l \leq M\}$$

gegeben durch eine Anzahl an Teilmengen  $K_l \subset \Omega$ ,  $1 \leq l \leq M$ , eines polyedrischen Gebietes  $\Omega \subset \mathbb{R}^3$  heißt *Zerlegung* oder auch *Gitter*, wenn die folgenden Bedingungen gelten:

1.  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h(\Omega)} \bar{K}$ , wobei  $\bar{\Omega}$  der Abschluss von  $\Omega$ .
2. Jedes  $K_l \in \mathcal{T}_h(\Omega)$ ,  $1 \leq l \leq M$ , ist ein offenes Lipschitz-Gebiet mit positivem Volumen.
3. Für  $K_1, K_2 \in \mathcal{T}_h(\Omega)$  mit  $K_1 \neq K_2$  gilt  $K_1 \cap K_2 = \emptyset$ .

Zu dieser Definition erhalten wir für jedes Element  $K \in \mathcal{T}_h(\Omega)$  die Parameter

$$h_K = \text{diam}(K) = \sup \{d(x, y) = |x - y| : x, y \in K\},$$

$$\rho_K = \sup \left\{ d(x, y) = |x - y| : x, y \in K; \exists z \in K \text{ mit } S_{d(x,y)}^2(z) \subset \bar{K} \right\},$$

wobei

$$S_d^2(z) = \{x \in \mathbb{R}^3 : \|x - z\|_{\mathbb{R}^3} = d\}.$$

Dahingehend wird der Index  $h$ , der auch als Schrittweite des Gitters  $\mathcal{T}_h(\Omega)$  bezeichnet wird, definiert durch

$$h = \max_{K \in \mathcal{T}_h(\Omega)} h_K. \quad (3.46)$$

Für die weitere Betrachtung verwenden wir ein aus Tetraedern bestehendes Gitter. Dabei ist ein Tetraeder  $K \subset \mathbb{R}^3$  die konvexe Hülle von drei linear unabhängigen Vektoren  $v_k \in \mathbb{R}^3$ ,  $k = 1, 2, 3$ . Dieses Tetraeder besteht daher aus vier Eckpunkten  $P_k \in \mathbb{R}^3$ ,  $1 \leq k \leq 4$ , sowie den Kanten  $(e_k)_{k=1}^6$  und den Flächen  $(f_k)_{k=1}^4$ . Für diese Art an Gittern führen wir die Begriffe der Konformität und Regularität ein.

**Definition 3.13.** Das Gitter  $\mathcal{T}_h(\Omega)$  aus Tetraedern nach Definition 3.12 wird *konform* genannt, wenn es die folgende geometrische Bedingung erfüllt:

Für  $K_1, K_2 \in \mathcal{T}_h(\Omega)$  mit  $K_1 \neq K_2$  und  $\bar{K}_1 \cap \bar{K}_2 \neq \emptyset$  gilt genau eine der folgenden Bedingungen:

1. Die Schnittmenge ist ein gemeinsamer Eckpunkt  $P$  beider Elemente  $K_1, K_2$ .
2. Die Schnittmenge ist eine gemeinsame Kante  $e$  und die Endpunkte  $P_1, P_2$  sind jeweils Eckpunkte beider Elemente  $K_1, K_2$ .
3. Die Schnittmenge ist eine gemeinsame Fläche  $f$  und die Eckpunkte der Fläche  $P_1, P_2, P_3$  sind jeweils Eckpunkte beider Elemente  $K_1, K_2$ .

**Definition 3.14.** Sei

$$\sigma_K = \frac{h_K}{\rho_K}, \quad \sigma_h = \max_{K \in \mathcal{T}_h(\Omega)} \sigma_K,$$

für jedes Element  $K \in \mathcal{T}_h(\Omega)$ . Eine Familie von Gittern  $(\mathcal{T}_{h_l}(\Omega))$  nach Definition 3.12, mit  $h_l$  gegeben durch (3.46), wird *regulär* genannt, wenn Konstanten  $\sigma_{\max} > 0$  und  $h_0 > 0$  existieren sodass

$$\sigma_{h_l} \leq \sigma_{\max}$$

für alle  $h_l$  mit  $0 < h_l \leq h_0$ .

Wir betrachten im Folgenden lediglich konforme Gitter und im Falle einer Familie an Gittern soll diese Menge regulär sein. Des Weiteren definieren wir das Referenztetraeder  $\hat{K}$  als das Einheitssimplex mit den Eckpunkten

$$\xi_1 = (0, 0, 0)^T, \quad \xi_2 = (1, 0, 0)^T, \quad \xi_3 = (0, 1, 0)^T, \quad \xi_4 = (0, 0, 1)^T.$$

Unter Verwendung des Referenzelements  $\hat{K}$  kann jedes Tetraeder  $K \in \mathcal{T}_h(\Omega)$  durch eine affine Koordinatentransformation von  $\hat{K}$  beschrieben werden. Dazu definieren wir die Indexmenge  $I_{\mathcal{T}_h(\Omega)}$  zu  $K_m \in \mathcal{T}_h(\Omega)$  und beschreiben im Folgenden die Koordinaten im Referenzgebiet  $\hat{K}$  mit  $x_{\hat{K}}$  und die Koordinaten in  $K_m$  mit  $x_{K_m}$ . Die affine Koordinatentransformation ist gegeben durch die Abbildung

$$\begin{aligned} \phi_{K_m} : \hat{K} &\rightarrow K_m, \\ x_{\hat{K}} &\mapsto x_{K_m} = B_{K_m} x_{\hat{K}} + b_{K_m}, \end{aligned} \quad (3.47)$$

mit  $B_{K_m} \in \mathbb{R}^{3 \times 3}$  und  $b_{K_m} \in \mathbb{R}^3$ . Wollen wir nun Vektorfunktionen  $V_{K_m} \in H(\text{curl}; K_m)$  durch Vektorfunktionen  $V_{\hat{K}} \in H(\text{curl}; \hat{K})$  beschreiben, dann benötigen wir eine Abbildung

$$\Phi_{K_m} : H(\text{curl}; \hat{K}) \rightarrow H(\text{curl}; K_m).$$

Um eine korrekte Abbildung der Vektorfelder zwischen den Räumen  $H(\text{curl}; \hat{K})$  und  $H(\text{curl}; K_m)$  zu gewährleisten, muss die Piola-Transformation

$$V_{K_m} = \Phi_{K_m}[V_{\hat{K}}] = J_{K_m}^{-T} (V_{\hat{K}} \circ \phi_{K_m}^{-1}) \quad (3.48)$$

verwendet werden, wobei

$$J_{K_m}(x_{\hat{K}}) = \nabla \phi_{K_m}(x_{\hat{K}}) = \left( \frac{\partial(\phi_{K_m})_i}{\partial(x_{\hat{K}})_j}(x_{\hat{K}}) \right)_{1 \leq i, j \leq 3} \quad (3.49)$$

die Jacobimatrix der Koordinatentransformation  $\phi_{K_m}$  darstellt. Für eine detaillierte Beschreibung sei hier auf [43] verwiesen. Für eine affine Transformation (3.47) ergibt sich

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

$J_{K_m} = B_{K_m}$  und mit der Piola-Transformation folgt zudem

$$\nabla V_{K_m} = B_{K_m}^{-T} (\hat{\nabla} V_{\hat{K}} \circ \phi_{K_m}^{-1}) B_{K_m}^{-1}$$

und

$$\nabla \times V_{K_m} = \det(B_{K_m}^{-1}) B_{K_m} (\hat{\nabla} \times V_{\hat{K}} \circ \phi_{K_m}^{-1}),$$

wobei  $\hat{\nabla}$  den Gradientenoperator auf  $\hat{K}$  zu  $x_{\hat{K}}$  darstellt. Um nun das Nédelec-Element  $(K, P_K, \Sigma_K)$  einzuführen, wobei  $K \subset \Omega$ ,  $P_K$  Funktionenmenge und  $\Sigma_K$  die Freiheitsgrade definiert, sei hier  $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T \in \mathbb{Z}_+^3$  ein Multiindex mit

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}, \quad |\alpha| = \alpha_1 + \alpha_2 + \alpha_3.$$

Damit erhalten wir die beiden Polynomräume

$$P_k = \left\{ p(x) = \sum_{|\alpha| \leq k} a_\alpha x^\alpha \text{ mit Koeffizienten } a_\alpha \in \mathbb{C} \right\}, \quad (3.50)$$

$$\tilde{P}_k = \left\{ p(x) = \sum_{|\alpha|=k} a_\alpha x^\alpha \text{ mit Koeffizienten } a_\alpha \in \mathbb{C} \right\}. \quad (3.51)$$

Des Weiteren definieren  $P_k(e)$ ,  $P_k(f)$  die auf eine Kante  $e$  und auf eine Fläche  $f$  eingeschränkten Polynome aus  $P_k$  und analog gilt dies auch für  $\tilde{P}_k$ . Sei nun

$$\mathcal{S}_k = \left\{ P \in [\tilde{P}_k]^3 : x \cdot P = 0 \right\}, \quad (3.52)$$

dann definieren wir den vektorwertigen Polynomraum

$$R_k = [P_{k-1}]^3 \oplus \mathcal{S}_k. \quad (3.53)$$

Nach [32, Lemma 5.32] ist der Polynomraum  $R_k$  invariant unter der Piola-Transformation (3.48), womit wir das konforme und unisolvente Nédelec-Element definieren können, siehe [32, Theorem 5.37]. Die Freiheitsgrade dieser Elemente sind unabhängig von der affinen Koordinatentransformation von  $\hat{K}$  zu  $K$ , was wiederum in [32, Lemma 5.34] bewiesen wird.

**Definition 3.15.** Das Nédelec-Element  $(K, P_K, \Sigma_K)$  zu dem Polynomraum  $P_K$  auf einem Element  $K$  mit Referenzelement  $\hat{K}$  ist definiert durch:

- $K$  ist ein Tetraeder;
- $P_K = R_k$ ;

- Die Freiheitsgrade  $\Sigma_K$  sind von den Kanten  $e$ , den Seitenflächen  $f$  von  $K$  und  $K$  selbst abhängig. Sei  $\tau$  der Einheitstangentenvektor der Kante  $e$ , dessen Richtung von  $\xi_n$  nach  $\xi_m$  mit  $n < m$  gerichtet ist, und  $n$  der Einheitsnormalenvektor der Fläche  $f$ , dessen Richtung nach Außen gerichtet ist. Die drei verschiedenen Freiheitsgrade sind:

1. Kanten-Freiheitsgrade:

$$M_e(V) = \left\{ \int_e V \cdot \tau q \, ds \text{ für } q \in P_{k-1}(e) \text{ und Kanten } e \text{ von } K \right\},$$

2. Flächen-Freiheitsgrade:

$$M_f(V) = \left\{ \frac{1}{|f|} \int_f V \cdot q \, d\Gamma, \, q = B_K \hat{q}, \, \hat{q} \in [P_{k-2}(\hat{f})]^3, \, \hat{q} \cdot \hat{n} = 0 \right. \\ \left. \text{und für Flächen } f \text{ von } K \right\},$$

3. Volumen-Freiheitsgrade:

$$M_K(V) = \left\{ \int_K V \cdot q \, d\Omega, \, q = \det(B_K^{-1}) B_K (\hat{q} \circ \phi_K^{-1}) \text{ mit } \hat{q} \in [P_{k-3}]^3 \right\}$$

Es gilt damit  $\Sigma_K = M_e(V) \cup M_f(V) \cup M_K(V)$ .

Damit ist der globale Finite-Elemente-Raum

$$Y_{\mathcal{N}} = \{V \in H(\text{curl}; \Omega) : V|_K \in R_k, \, K \in \mathcal{T}_h(\Omega)\}, \quad (3.54)$$

definiert als Unterraum

$$Y_{\mathcal{N}} \subset Y_e = \{V \in H(\text{curl}; \Omega) : \pi_t[V] \in L_t^2(\partial\Omega)\} \subset H(\text{curl}; \Omega).$$

Die Norm in  $Y_e$  ist gegeben durch

$$\|V\|_{Y_e}^2 = \|V\|_{H(\text{curl}; \Omega)}^2 + \|\pi_t[V]\|_{[L^2(\partial\Omega)]^3}^2,$$

wobei  $\pi_t$  den entsprechenden stückweise definierten Spuoperator aus Abschnitt 2.2 auf dem Rand  $\partial\Omega$  darstellt. Durch die Indexierung  $I_{\mathcal{T}_h}$ , mit Kardinalität  $|I_{\mathcal{T}_h}| = \mathcal{N}$ , der globalen Ansatzfunktionen  $(\varphi_k)_{k=1}^{\mathcal{N}}$ , welche durch Einschränkung auf  $K_m \in \mathcal{T}_h(\Omega)$  mit

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

einer der lokalen Ansatzfunktionen  $\varphi_l^{K_m}$  übereinstimmt, erhalten wir

$$Y_{\mathcal{N}}(\Omega) = \text{span} \{\varphi_1, \dots, \varphi_{\mathcal{N}}\}, \quad (3.55)$$

bezüglich des konformen Gitters  $\mathcal{T}_h(\Omega)$ . Somit haben wir für  $V_{\mathcal{N}} \in Y_{\mathcal{N}}$  die Basisdarstellung

$$V_{\mathcal{N}} = \sum_{k=1}^{\mathcal{N}} V_k^{\mathcal{N}} \varphi_k$$

mit  $V_k^{\mathcal{N}} \in \mathbb{C}$  für alle  $k \in \{1, \dots, \mathcal{N}\}$ . Eine wichtige Eigenschaft der so definierten Ansatzfunktionen  $\varphi_k$  ist die tangentielle Stetigkeit (2.64) an der Grenzfläche zweier benachbarter Teilgebiete. Weitere Eigenschaften und Beweise zu den hier dargestellten Nédélec Elementen finden sich in [32, 43, 44]. Da die Nédélec-Basisfunktionen lokal auf den einzelnen Tetraeder  $K_m$  des regulären Gitters  $\mathcal{T}_h(\Omega)$  definiert werden, können wir die lokalen Nédélec-Basisfunktionen  $\varphi_n^{K_m}$  aus den Nédélec-Ansatzfunktionen  $\varphi_n^{\hat{K}}$  des Referenzelements herleiten, indem wir

$$\varphi_n^{K_m}(x_{K_m}) = \Phi(\varphi_n^{\hat{K}}(x_{\hat{K}})) = B_{K_m}^{-T} \varphi_n^{\hat{K}}(x_{\hat{K}}) \circ \phi_{K_m}^{-1}$$

mit der Affintransformation (3.47) anwenden.

### 3.3. Existenztheorie der diskreten Formulierung

Wir definieren den Lösungsraum als Unterraum

$$X_{\mathcal{N}} = \{V_{\mathcal{N}} \in Y_{\mathcal{N}} : \gamma_t[V_{\mathcal{N}}] = 0 \text{ auf } \Gamma_{PEC}\} \subset Y_{\mathcal{N}}. \quad (3.56)$$

Dabei folgt aus  $\pi_t[\varphi_k] \in L_t^2(\Gamma_R)$  für alle  $1 \leq k \leq \mathcal{N}$ , dass auch für  $E_{\mathcal{N}} \in X_{\mathcal{N}}$  die erste Bedingung

$$\pi_t[E_{\mathcal{N}}] \in L_t^2(\Gamma_R)$$

in (3.1) erfüllt ist und deshalb folgt

$$X_{\mathcal{N}} \subset X_e.$$

Wir betrachten das diskrete variationelle Problem:

Bestimme  $E_{\mathcal{N}} \in X_{\mathcal{N}}$ , sodass

$$[\text{VD}_{,\mathcal{N}}] \quad a(E_{\mathcal{N}}, \varphi_{\mathcal{N}}) = f(\varphi_{\mathcal{N}}), \quad \varphi_{\mathcal{N}} \in X_{\mathcal{N}},$$

gilt mit

$$a(E_{\mathcal{N}}, \varphi_{\mathcal{N}}) = \int_{\Omega} \frac{1}{\mu_r} (\nabla \times E_{\mathcal{N}}) \cdot (\nabla \times \varphi_{\mathcal{N}}) d\Omega - \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) \mu_0 E_{\mathcal{N}} \cdot \varphi_{\mathcal{N}} d\Omega \quad (3.57)$$

$$- \int_{\Gamma_{IMP}} i\omega Y(\omega) \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi_{\mathcal{N}}] d\Gamma$$

$$- \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi_{\mathcal{N}}] d\Gamma,$$

$$f(\varphi_{\mathcal{N}}) = -2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi_{\mathcal{N}}] d\Gamma. \quad (3.58)$$

### 3.3.1. Interpolationsoperator

Analog zu Definition A.9 definieren wir

$$H^\delta(\text{curl}; \Omega) = \left\{ u \in [H^\delta(\Omega)]^3 \mid \nabla \times u \in [H^\delta(\Omega)]^3 \right\} \quad (3.59)$$

mit der entsprechenden Norm

$$\|u\|_{H^\delta(\Omega)}^2 = \|u\|_{[H^\delta(\Omega)]^3}^2 + \|\nabla \times u\|_{[H^\delta(\Omega)]^3}^2.$$

Nach [38, Lemma 4.7] sind die Freiheitsgrade der Definition 3.15 für  $u \in H^\delta(\text{curl}; \Omega)$ ,  $\delta > \frac{1}{2}$ , wohldefiniert und beschränkt, weswegen der Interpolationsoperator

$$\Pi_{\mathcal{N}} : H^\delta(\text{curl}; \Omega) \rightarrow Y_{\mathcal{N}} \quad (3.60)$$

über die Bedingung

$$M_e(u - \Pi_{\mathcal{N}}[u]) = M_f(u - \Pi_{\mathcal{N}}[u]) = M_K(u - \Pi_{\mathcal{N}}[u]) = 0 \quad (3.61)$$

an die Freiheitsgrade charakterisiert wird.

**Lemma 3.16.** *Für alle  $\delta \in (\frac{1}{2}, 1]$  existiert eine von der Schrittweite  $h$  des Gitters  $\mathcal{T}_h(\Omega)$  unabhängige Konstante  $C > 0$ , sodass*

$$\|u - \Pi_{\mathcal{N}}[u]\|_{H(\text{curl}; \Omega)} \leq Ch^\delta \|u\|_{H^\delta(\text{curl}; \Omega)}, \quad u \in H^\delta(\text{curl}; \Omega), \quad (3.62)$$

*gilt.*

*Beweis.* Siehe [45, Proposition 5.6]. □

Mit der durch  $\mathcal{T}_h(\Omega)$  induzierten Triangulierung  $\mathcal{T}_h(\Gamma)$  auf den Rändern des Gebietes  $\Omega$  definieren wir den Interpolationsoperator für  $\delta \in (0, 1]$  auf dem Rand

$$\Pi_{\mathcal{N}, \Gamma} : H_{||}^\delta(\Gamma) \cap H(\text{curl}_\Gamma; \Gamma) \rightarrow (Y_{\mathcal{N}})_\Gamma,$$

wobei  $(Y_{\mathcal{N}})_\Gamma$  dem zweidimensionalen FEM-Raum auf einer Dreieckstriangulierung entspricht, siehe [46]. Für den Randinterpolationsoperator ergibt sich gerade folgendes Lemma.

**Lemma 3.17.** Für jedes  $\delta \in (0, 1]$  existiert eine Konstante  $C > 0$ , unabhängig von der Schrittweite  $h$  der Triangulierung  $\mathcal{T}_h(\Gamma)$ , sodass

$$\|u - \Pi_{\mathcal{N},\Gamma}[u]\|_{[L^2(\Gamma)]^3} \leq Ch^\delta \left( \|u\|_{H_{\parallel}^\delta(\Gamma)} + \|\nabla_\Gamma \times u\|_{L^2(\Gamma)} \right), \quad u \in H_{\parallel}^\delta(\Gamma) \cap H(\text{curl}_\Gamma; \Gamma). \quad (3.63)$$

*Beweis.* Siehe [47]. □

Aus (3.62) und Lemma 3.17 erhalten wir für die Norm von  $X_e$  die Existenz einer von der Schrittweite  $h$  der Triangulierung  $\mathcal{T}_h(\Omega)$  unabhängigen Konstante  $C > 0$ , sodass für alle Funktionen  $u \in H^\delta(\text{curl}; \Omega)$  mit  $\pi_t[u] \in H_{\parallel}^\delta(\Gamma) \cap H(\text{curl}_\Gamma; \Gamma)$  die Ungleichung

$$\|u - \Pi_{\mathcal{N}}[u]\|_{X_e} \leq Ch^\delta \left( \|u\|_{H^\delta(\text{curl}; \Omega)} + \|\pi_t[u]\|_{H_{\parallel}^\delta(\Gamma_R)} + \|\nabla_\Gamma \times \pi_t[u]\|_{L^2(\Gamma_R)} \right) \quad (3.64)$$

gilt. Dabei ist (3.64) gerade die Approximationseigenschaft des Raumes  $X_{\mathcal{N}}$  bei entsprechender Regularität der zu approximierenden Funktion.

### 3.3.2. Existenz der Lösung

Der Beweis läuft in großen Teilen analog zum Beweis der Existenz der Lösung des kontinuierlichen variationellen Problems. Die Beweise der Aussagen innerhalb dieses Kapitels können dem Abschnitt 7.2 in [28] entnommen werden. Zunächst definieren wir den diskreten Projektionsoperator

$$\begin{aligned} \mathcal{P}_{\mathcal{N}} : X_{\mathcal{N}} &\rightarrow X_{\mathcal{N}}, \\ v &\mapsto \mathcal{P}_{\mathcal{N}}[v] = (\Pi_{\mathcal{N}}\mathcal{P})[v] \end{aligned} \quad (3.65)$$

und verwenden das folgende Lemma, dessen Beweis auf der curl-Eigenschaft des Operators  $\mathcal{P}$  und entsprechenden Pendeldiagrammen für den Projektionsoperator  $\Pi_{\mathcal{N}}$  in [32] beruht.

**Lemma 3.18.** Es existiert eine Konstante  $C > 0$ , unabhängig von der Schrittweite  $h$  der Triangulierung  $\mathcal{T}_h(\Omega)$ , sodass

$$\|\mathcal{P}[u] - \mathcal{P}_{\mathcal{N}}[u]\|_{X_e} \leq Ch^{\delta - \frac{1}{2}} \|u\|_{X_e}, \quad u \in X_e,$$

mit  $\delta \in (\frac{1}{2}, 1]$ .

Mit Lemma 3.18 ergibt sich das Analogon zu Lemma 3.8 für das diskrete variationelle Problem mit den entsprechenden Definitionen aus Abschnitt 3.1.2.

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

**Lemma 3.19.** *Es existieren Konstanten  $\beta_{\mathcal{N}} > 0$ ,  $h_0 > 0$ , unabhängig von der Schrittweite  $h$  des Gitters  $\mathcal{T}_h(\Omega)$ , sodass für alle  $h < h_0$*

$$\sup_{\varphi_{\mathcal{N}} \in X_{\mathcal{N}} \setminus \{0\}} \frac{|A(E_{\mathcal{N}}, \varphi_{\mathcal{N}})|}{\|\varphi_{\mathcal{N}}\|_{X_e}} \geq \beta_{\mathcal{N}} \|E_{\mathcal{N}}\|_{X_e}, \quad E_{\mathcal{N}} \in X_{\mathcal{N}}.$$

Der Beweis von Lemma 3.19 wird analog zum Beweis von Lemma 3.8 geführt, nur das hierbei noch die Interpolationseigenschaft, Lemma 3.18, des Projektors  $\mathcal{P}_{\mathcal{N}}$  verwendet wird. Damit erhalten wir den Existenzsatz für das diskrete variationelle Problem  $[V_{D,\mathcal{N}}]$ .

**Satz 3.20.** *Angenommen es existiert eine Lösung zum kontinuierlichen Variationsproblem  $[V_K]$ . Dann existiert  $h_0 > 0$ , sodass für alle  $h \leq h_0$  das diskrete Variationsproblem  $[V_{D,\mathcal{N}}]$  eine eindeutige Lösung  $E_{\mathcal{N}} \in X_{\mathcal{N}}$  besitzt. Zudem existieren Konstanten  $C_1, C_2 > 0$ , unabhängig von der Schrittweite  $h$  des Gitters  $\mathcal{T}_h(\Omega)$ , sodass*

$$\|E_{\mathcal{N}}\|_{X_e} \leq C_1 \|\mathcal{F}\|_{(X_e)'} \quad (3.66)$$

und

$$\|E - E_{\mathcal{N}}\|_{X_e} \leq C_2 \inf_{V_{\mathcal{N}} \in X_{\mathcal{N}}} \|E - V_{\mathcal{N}}\|_{X_e}. \quad (3.67)$$

Falls zusätzlich ein  $\delta \in (\frac{1}{2}, 1]$  existiert, sodass  $E \in H^\delta(\text{curl}; \Omega)$  und  $\pi_t[E] \in H^\delta(\text{curl}_\Gamma; \Gamma)$ , so existiert eine von der Schrittweite  $h$  des Gitters  $\mathcal{T}_h(\Omega)$  unabhängige konstante  $C_3$ , sodass

$$\|E - E_{\mathcal{N}}\|_{X_e} \leq C_3 h^\delta (\|E\|_{H^\delta(\text{curl}; \Omega)} + \|\pi_t[E]\|_{H^\delta(\text{curl}_\Gamma; \Gamma)}). \quad (3.68)$$

Der Beweis beruht auf der Existenz der globalen inf-sup-Konstante nach Korollar 3.11 und dem Operator  $\mathcal{R}_{\mathcal{N}} : X_e \rightarrow X_{\mathcal{N}}$ , charakterisiert durch

$$A(E_{\mathcal{N}}, \mathcal{R}_{\mathcal{N}}[v]) = A(E_{\mathcal{N}}, v_{\mathcal{N}}), \quad E_{\mathcal{N}} \in X_{\mathcal{N}},$$

woraus sich die entsprechenden Abschätzungen mit dem Lemma von C ea f ur inf-sup stabile Sesquilinearformen herleiten lassen.

### 3.4. Finite Elemente Assemblierung

Um ein lineares Gleichungssystem zu erhalten, wollen wir den Nédelec-Ansatzraum (3.55) für die Gleichungen (3.57) und (3.58) verwenden, weshalb wir die diskretisierte Gleichung  $[V_{D,\mathcal{N}}]$  umformulieren müssen. Wir betrachten dazu die Bedingung

$$\gamma_t[V] = 0 \text{ auf } \Gamma_{PEC} \quad (3.69)$$

des Lösungsraums (3.56) und fügen diese Nebenbedingung der Formulierung  $[V_{D,\mathcal{N}}]$  mithilfe von Lagrange-Multiplikatoren hinzu. Dazu betrachten wir die tangentialen Operatoren  $\gamma_t$  und  $\pi_t$  für den Unterraum  $Y_{\mathcal{N}} \subset Y_e$ . Nach [32, Lemma 5.35] gilt  $\gamma_t[V_{\mathcal{N}}] = 0$  auf den diskretisierten Teilflächen von  $\Gamma_{PEC}$ , genau dann wenn  $M_f(V_{\mathcal{N}}) = 0$  mit  $f \subset \Gamma_{PEC}$  und  $M_e(V_{\mathcal{N}}) = 0$  für die Kanten  $e$  der Fläche  $f$  gilt. Damit sind die noch zu definierenden Tangentialräume von  $\Gamma_{PEC}$  nur von den Freiheitsgraden auf  $\Gamma_{PEC}$  abhängig und wir identifizieren alle Ansatzfunktionen  $\varphi_k$  für die ein Freiheitsgrad auf  $\Gamma_{PEC}$  existiert, der nicht Null wird. Wir bezeichnen die Indexmenge dieser Ansatzfunktionen mit  $I_{PEC} \subset I_{\mathcal{T}_h}$ . Des Weiteren sei  $\mathcal{N}_{PEC} = |I_{PEC}|$  die Kardinalität der Indexmenge. Daraus folgern wir die Basisdarstellungen der tangentialen Räume

$$Y_{\gamma,\mathcal{N}_{PEC}} = \bigoplus_{k \in I_{PEC}} \text{span} \{ \gamma_t[\varphi_k] \}$$

und

$$Y_{\pi,\mathcal{N}_{PEC}} = \bigoplus_{k \in I_{PEC}} \text{span} \{ \pi_t[\varphi_k] \}.$$

Aufgrund der Linearität der tangentialen Operatoren und der Basisdarstellung (3.55) von  $Y_{\mathcal{N}}$  ergibt sich daher

$$\begin{aligned} \gamma_t : Y_{\mathcal{N}} &\rightarrow Y_{\gamma,\mathcal{N}_{PEC}}, \\ V_{\mathcal{N}} = \sum_{k=1}^{\mathcal{N}} V_k^{\mathcal{N}} \varphi_k &\mapsto V_{\gamma,\mathcal{N}_{PEC}} = \sum_{k \in I_{PEC}} V_k^{\mathcal{N}} \gamma_t[\varphi_k], \\ \pi_t : Y_{\mathcal{N}} &\rightarrow Y_{\pi,\mathcal{N}_{PEC}}, \\ V_{\mathcal{N}} = \sum_{k=1}^{\mathcal{N}} V_k^{\mathcal{N}} \varphi_k &\mapsto V_{\pi,\mathcal{N}_{PEC}} = \sum_{k \in I_{PEC}} V_k^{\mathcal{N}} \pi_t[\varphi_k]. \end{aligned}$$

Analog lassen sich auch die Tangentialräume auf den anderen Teilflächen  $\Gamma_{IMP}$  und  $\Gamma_{P,l}$ ,  $1 \leq l \leq K_A$ , ableiten. Nach Satz 2.8 kann das  $[L^2(\Gamma_{PEC})]^3$ -Skalarprodukt als Dualitätsprodukt zwischen den Räumen  $H^{-\frac{1}{2}}(\text{div}_{\Gamma_{PEC}}, \Gamma_{PEC})$  und  $H^{-\frac{1}{2}}(\text{curl}_{\Gamma_{PEC}}, \Gamma_{PEC})$  aufgefasst werden. Deshalb verwenden wir  $Y_{\gamma,\mathcal{N}_{PEC}} \subset H^{-\frac{1}{2}}(\text{div}_{\mathcal{N}_{PEC}}, \Gamma_{PEC})$  als Dualraum von  $Y_{\pi,\Gamma_{PEC}} \subset H^{-\frac{1}{2}}(\text{curl}_{\Gamma_{PEC}}, \Gamma_{PEC})$  und der Lagrange-Faktor  $\lambda \in Y_{\pi,\mathcal{N}_{PEC}}$  zur

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

Nebenbedingung (3.69) ist gegeben durch

$$d(E_{\mathcal{N}}, \lambda) = \int_{\Gamma_{PEC}} \gamma_t[E_{\mathcal{N}}] \cdot \lambda \, d\Gamma. \quad (3.70)$$

Da auch die Testfunktionen Elemente des Raumes (3.56) sind, benötigen wir zu diesen das Dualitätsprodukt

$$d_f(\lambda_\varphi, \varphi_{\mathcal{N}}) = \int_{\Gamma_{PEC}} \lambda_\varphi \cdot \gamma_t[\varphi_{\mathcal{N}}] \, d\Gamma, \quad (3.71)$$

mit dem Lagrange-Faktor  $\lambda_\varphi \in Y_{\pi, \mathcal{N}_{PEC}}$ . Wir erhalten damit die zu  $[V_{D, \mathcal{N}}]$  äquivalente Formulierung:

Bestimme  $E_{\mathcal{N}} \in Y_{\mathcal{N}}$  und  $\lambda_\varphi \in Y_{\pi, \mathcal{N}_{PEC}}$ , sodass

$$\begin{aligned} a(E_{\mathcal{N}}, \varphi_{\mathcal{N}}) + d_f(\lambda_\varphi, \varphi_{\mathcal{N}}) &= f(\varphi_{\mathcal{N}}), \quad \varphi_{\mathcal{N}} \in Y_{\mathcal{N}}, \\ d(E_{\mathcal{N}}, \lambda) &= g(\lambda), \quad \lambda \in Y_{\pi, \mathcal{N}_{PEC}}, \end{aligned} \quad (3.72)$$

gilt, mit (3.57), (3.58), (3.70), (3.71) und  $g(\lambda) = 0$ .

Werden die Basisdarstellungen der Lösungsräume  $Y_{\mathcal{N}}$  und  $Y_{\pi, \mathcal{N}_{PEC}}$  ausgenutzt und die Gleichungen (3.72) für alle Basisfunktionen der Räume  $Y_{\mathcal{N}}$  und  $Y_{\pi, \mathcal{N}_{PEC}}$  getestet, dann ergibt sich das lineare Gleichungssystem

$$\begin{pmatrix} A & D_f \\ D & 0 \end{pmatrix} \begin{pmatrix} E_{\mathcal{N}} \\ \lambda^{\mathcal{N}_{PEC}} \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \quad (3.73)$$

mit

$$\begin{aligned} A &\in \mathbb{C}^{\mathcal{N} \times \mathcal{N}}, \quad D_f \in \mathbb{C}^{\mathcal{N} \times \mathcal{N}_{PEC}}, \quad D \in \mathbb{C}^{\mathcal{N}_{PEC} \times \mathcal{N}}, \\ f &\in \mathbb{C}^{\mathcal{N}}, \quad g \in \mathbb{C}^{\mathcal{N}_{PEC}}. \end{aligned}$$

Im Falle der hier betrachteten Modelle ergibt sich gerade  $g = 0$ , aufgrund der betrachteten PEC-Randbedingung, und die einzelnen Spalten der Matrizen  $D$  und  $D_f^H$  sind ausgewählte Einheitsvektoren. Die Indizes der Einträge entsprechen dabei den Indizes der auf  $\Gamma_{PEC}$  liegenden FEM-Basisfunktionen. Um die Gleichungen (3.72) voneinander zu entkoppeln, betrachten wir den Operator

$$\begin{aligned} \mathcal{D} : Y_{\mathcal{N}} &\rightarrow (Y_{\pi, \mathcal{N}_{PEC}})', \\ E_{\mathcal{N}} &\mapsto d(E_{\mathcal{N}}, \cdot), \end{aligned}$$

und Zerlegen  $Y_{\mathcal{N}}$  bezüglich  $\mathcal{D}$  in

$$Y_{\mathcal{N}} = \text{Kern}(\mathcal{D}) \oplus \text{Kern}(\mathcal{D})^{\perp},$$

dabei bezeichnet  $\text{Kern}(\mathcal{D})^{\perp}$  das orthogonale Komplement von  $\text{Kern}(\mathcal{D})$  nach Definition A.2. Damit erhalten wir die eindeutige Zerlegung für den Lösungsraum

$$E_{\mathcal{N}} = E_{\mathcal{N},c} + E_{\mathcal{N},d} = \sum_{k=1}^{\mathcal{N}} (E_c^{\mathcal{N}})_k \varphi_k + \sum_{k=1}^{\mathcal{N}} (E_d^{\mathcal{N}})_k \varphi_k$$

mit  $E_{\mathcal{N},c} \in \text{Kern}(\mathcal{D})$ ,  $E_{\mathcal{N},d} \in \text{Kern}(\mathcal{D})^{\perp}$ ,  $E_c^{\mathcal{N}} \in \mathbb{C}^{\mathcal{N}}$  und  $E_d^{\mathcal{N}} \in \mathbb{C}^{\mathcal{N}}$ . Des Weiteren betrachten wir die Zerlegung

$$Y_{\mathcal{N}} = \text{Kern}(\mathcal{D}_f) \oplus \text{Kern}(\mathcal{D}_f)^{\perp}$$

bezüglich

$$\begin{aligned} \mathcal{D}_f : Y_{\mathcal{N}} &\rightarrow (Y_{\pi, \mathcal{N}_{PEC}})', \\ \varphi_{\mathcal{N}} &\mapsto d_f(\cdot, \varphi_{\mathcal{N}}), \end{aligned}$$

und erhalten die eindeutige Zerlegung für den Testraum

$$\varphi_{\mathcal{N}} = \varphi_{\mathcal{N},c} + \varphi_{\mathcal{N},d}$$

mit  $\varphi_{\mathcal{N},c} \in \text{Kern}(\mathcal{D}_f)$  und  $\varphi_{\mathcal{N},d} \in \text{Kern}(\mathcal{D}_f)^{\perp}$ . Setzen wir die Zerlegung des Lösungsraums in die Gleichungen (3.72) ein, so ergibt sich aufgrund der Linearität und der Kern-Eigenschaft

$$\begin{aligned} a(E_{\mathcal{N},c}, \varphi_{\mathcal{N}}) + d_f(\lambda_{\varphi}, \varphi_{\mathcal{N}}) &= f(\varphi_{\mathcal{N}}) - a(E_{\mathcal{N},d}, \varphi_{\mathcal{N}}), \quad \varphi_{\mathcal{N}} \in Y_{\mathcal{N}}, \\ d(E_{\mathcal{N},d}, \lambda) &= g(\lambda), \quad \lambda \in Y_{\pi, \mathcal{N}_{PEC}}. \end{aligned}$$

Testen wir diese beiden Gleichungen nun anstatt der Funktionen aus  $Y_{\mathcal{N}}$  mit den Funktionen aus  $\text{Kern}(\mathcal{D}_f)$ , dann ergibt sich

$$\begin{aligned} a(E_{\mathcal{N},c}, \varphi_{\mathcal{N},c}) &= f(\varphi_{\mathcal{N},c}) - a(E_{\mathcal{N},d}, \varphi_{\mathcal{N},c}), \quad \varphi_{\mathcal{N},c} \in \text{Kern}(\mathcal{D}_f), \\ d(E_{\mathcal{N},d}, \lambda) &= g(\lambda), \quad \lambda \in Y_{\pi, \mathcal{N}_{PEC}}. \end{aligned} \tag{3.74}$$

Damit erhalten wir die Lösung  $E_{\mathcal{N}}$  von  $[V_{\mathcal{D}, \mathcal{N}}]$  indem wir zunächst die zweite Gleichung in (3.74) lösen, um  $E_{\mathcal{N},d}$  zu erhalten, und dies zur Lösung der ersten Gleichung in (3.74) verwenden, um  $E_{\mathcal{N},c}$  zu bestimmen. Mit derselben Herangehensweise erhalten wir eine Lösung  $E^{\mathcal{N}}$  des Lagrange-Gleichungssystems (3.73), indem wir zunächst die bei-

### 3. Kontinuierliche und diskrete variationelle Formulierung

---

den rechtsseitigen Kernbasismatrizen  $B_D$  und  $B_{D_f}$  zu den beiden Matrizen  $D$  und  $D_f^H$  der Diskretisierung (3.73) bestimmen. Aufgrund der speziellen Struktur der Matrizen  $D$  und  $D_f^H$ , sind die Spalten der Kernbasismatrizen  $B_D$  und  $B_{D_f}$  wiederum durch Einheitsvektoren gegeben, deren Indizes der Einträge nun mit den nicht auf  $\Gamma_{PEC}$  liegenden Basisfunktionen übereinstimmen. Multiplizieren wir daraufhin die Matrix

$$\begin{pmatrix} B_{D_f}^H & 0 \\ 0 & I \end{pmatrix}$$

von links an die lineare Gleichung (3.73), so ergibt sich mit

$$E^{\mathcal{N}} = E_c^{\mathcal{N}} + E_d^{\mathcal{N}} = B_D E_c^{\mathcal{N}_c} + E_d^{\mathcal{N}},$$

wobei  $\mathcal{N}_c = \mathcal{N} - \mathcal{N}_{PEC}$ , das zu (3.73) äquivalente Gleichungssystem

$$B_{D_f}^H A B_D E_c^{\mathcal{N}_c} = B_{D_f}^H f - B_{D_f}^H A E_d^{\mathcal{N}}, \quad (3.75)$$

$$D E_d^{\mathcal{N}} = g \quad (3.76)$$

mit

$$B_{D_f}^H A B_D \in \mathbb{C}^{\mathcal{N}_c \times \mathcal{N}_c}, \quad B_{D_f}^H f \in \mathbb{C}^{\mathcal{N}_c}, \quad B_{D_f}^H A E_d^{\mathcal{N}} \in \mathbb{C}^{\mathcal{N}_c}.$$

Die dünnbesetzte Struktur der Matrix  $A$  erhalten wir aufgrund der Darstellung der Kernbasismatrizen auch für die Matrix  $B_{D_f}^H A B_D$ . Das heißt, wir bestimmen zunächst eine Lösung  $E_d^{\mathcal{N}}$  des Gleichungssystems (3.76) und verwenden diese wiederum in (3.75), um eine Lösung  $E_c^{\mathcal{N}_c}$  zu erhalten. Auch hier vereinfacht sich auf Basis der hier betrachteten Modelle die Bestimmung der Lösung  $E^{\mathcal{N}}$ , denn aufgrund der speziellen Struktur von  $D$  und  $g = 0$  ergibt sich direkt  $E_d^{\mathcal{N}} = 0$ . Da die Matrix  $B_{D_f}^H A B_D$  in den hier betrachteten Fällen quadratisch und dünnbesetzt sind, kann zur Lösung der Gleichung (3.75) bei Dimensionen bis zu  $10^6$  in MATLAB der Backslash-Operator verwendet werden. Bei höheren Dimensionen sollte aufgrund der Implementierung des Backslash-Operators hingegen der MUMPS-Algorithmus<sup>5</sup> verwendet werden. Aufgrund der in dieser Arbeit betrachteten Modelle vereinfachen sich die Berechnung der Kernbasismatrizen  $B_D$  und  $B_{D_f}$  oder auch  $E_d^{\mathcal{N}}$ , aber die Theorie der Lagrange-Systeme kann auch auf viele andere Probleme angewandt werden und besitzt dennoch die gleiche Struktur. Auch wenn die Gleichungen und Bedingungen an den Funktionenraum komplexer sein können, ergeben sich in sehr vielen Problemen dünnbesetzte und einfach strukturierte Matrizen  $D$  und  $D_f^H$ , sowie die daraus resultierenden Kernbasismatrizen  $B_D$  und  $B_{D_f}$  und die relativ einfache Bestimmung einer Lösung  $E_d^{\mathcal{N}}$ .

---

<sup>5</sup><http://mumps.enseeiht.fr/>

---

## 4. Reduzierte-Basis-Methode

Für die optimale Steuerung von PDEs müssen wir viele Gleichungssysteme (3.73) hoher Dimension lösen. Da die Berechnung solcher Lösungen sehr rechenintensiv ist, möchten wir die in der Arbeit betrachteten Modelle mithilfe der Modellreduktion in ihrer Dimension reduzieren. Hierzu betrachten wir den parametrischen Modellreduktionsalgorithmus Reduzierte-Basis-Methode (RBM), welche wir in Abschnitt 4.1 allgemein einführen, wobei Details in [14, 48] nachgelesen werden können. In Abschnitt 4.2 wenden wir die RBM auf die in Abschnitt 3.4 bereits erwähnten Lagrange-Probleme an und erweitern damit die RBM um eine bisher nicht betrachtete Theorie. In Abschnitt 4.3 betrachten wir die zu Abschnitt 4.2 zugehörigen numerischen Ergebnisse anhand ausgewählter Modelle der Optimalsteuerung.

### 4.1. Grundlagen der RBM

Dieser Abschnitt stellt eine kurze Zusammenfassung der RBM dar. Für eine detailliertere Betrachtung wird auf [14, 48, 49] verwiesen.

#### 4.1.1. Parameterische Differentialgleichungen

In der RBM betrachten wir parametrische PDE  $k$ -ter Ordnung

$$F(D^k u(p), \dots, D^2 u(p), Du(p), u(p); p) = 0 \quad (4.1)$$

mit  $k$ -ter örtlicher Ableitung  $D^k u(p)$  der gesuchten Funktion  $u(p) \in X_e$  und Parametern  $p \in \mathbb{P} \subset \mathbb{R}^n$ . Zu der Funktion  $u(p)$  kann zusätzlich auch noch ein möglicher Ausgang

$$s(p) = l(u(p); p)$$

betrachtet werden. Dies kann beispielsweise eine spezifische Temperatur, die Ausbreitungsrichtung einer Welle oder im Falle der in dieser Arbeit betrachteten Modelle die elektromagnetische Energie in einem spezifischen Ort sein. Der Ausgang  $s(p)$  kann demnach beliebig definiert werden, wird aber in dieser Arbeit nicht näher betrachtet. Die Theorie zu ausgangsbezogenen Problemen kann dabei wieder in [14, 48] nachgelesen werden. Bei der Betrachtung komplexer parametrischer PDEs werden komplexe Parameter  $\hat{p} \in \mathbb{C}$  als zwei reelle Parameter  $\text{Re}(\hat{p}), \text{Im}(\hat{p}) \in \mathbb{R}$  aufgefasst. Die zur parametrischen PDE zugehörige variationelle Formulierung lautet:

Bestimme eine Lösung  $u(p) \in X_e$ , sodass

$$a(u(p), \varphi; p) = f(\varphi; p), \quad \varphi \in X_e, \quad (4.2)$$

#### 4. Reduzierte-Basis-Methode

---

mit  $a : X_e \times X_e \times \mathbb{P} \rightarrow \mathbb{C}$  und  $f : X_e \times \mathbb{P} \rightarrow \mathbb{C}$  gilt.

Wird die Sesquilinearform  $a$  und die rechte Seite  $f$  diskretisiert, beispielsweise durch die FEM, dann ergibt sich die Approximation der variationellen Formulierung (4.2):

Bestimme  $u_{\mathcal{N}} \in X_{\mathcal{N}}$ , sodass

$$a(u_{\mathcal{N}}(p), \varphi; p) = f(\varphi; p), \quad \varphi \in X_{\mathcal{N}}, \quad (4.3)$$

gilt, wobei  $X_{\mathcal{N}} \subset X_e$  mit  $\mathcal{N} = \dim(X_{\mathcal{N}}) < \infty$ .

Dieses Problem wird in der Literatur als „truth problem“ bezeichnet. Aufgrund der Definition des FE-Raumes  $X_{\mathcal{N}}$  können wir  $u_{\mathcal{N}}(p)$  als Linearkombination einer Basis von  $X_{\mathcal{N}} = \text{span}\{\varphi_1, \dots, \varphi_{\mathcal{N}}\}$  mit Koeffizientenvektor  $u^{\mathcal{N}}(p) = (u_k^{\mathcal{N}}(p))_{1 \leq k \leq \mathcal{N}}$  schreiben als

$$u_{\mathcal{N}}(p) = \sum_{k=1}^{\mathcal{N}} u_k^{\mathcal{N}}(p) \varphi_k.$$

Wird die Linearkombination in (4.3) eingesetzt, so ergibt sich aus der Linearität der Sesquilinearform und dem Testen mit den Basisfunktionen  $\varphi_l$ ,  $l = 1, \dots, \mathcal{N}$ , das Gleichungssystem

$$A_p^{\mathcal{N}} u^{\mathcal{N}} = f_p^{\mathcal{N}} \quad (4.4)$$

mit  $(A_p^{\mathcal{N}})_{m,n} = a(\varphi_n, \varphi_m; p)$  und  $(f_p^{\mathcal{N}})_m = f(\varphi_m; p)$ . Der Index  $p$  soll dabei die Parameterabhängigkeit der Matrix  $A_p^{\mathcal{N}}$  und des Vektors  $f_p^{\mathcal{N}}$  verdeutlichen. Da wir eine approximative Lösung über die FEM bestimmen, betrachten wir im Folgenden den Fehler  $\|u(p) - u_{\mathcal{N}}(p)\|_{X_e}$  der exakten Lösung  $u(p)$  und der FE-Lösung  $u_{\mathcal{N}}(p)$ . Ist die Sesquilinearform  $a$  stetig und koerziv, d.h. existieren Konstanten  $\alpha(p), \gamma(p) > 0$ , für die

$$\begin{aligned} |a(u(p), v(p); p)| &\leq \gamma(p) \|u(p)\|_{X_e} \|v(p)\|_{X_e}, \quad u, v \in X_e, \\ \alpha(p) \|u(p)\|_{X_e}^2 &\leq |a(u(p), u(p); p)|, \quad u \in X_e, \end{aligned}$$

gilt, dann folgt aufgrund der Konformität des FE-Approximationsraums  $X_{\mathcal{N}} \subset X_e$ , mit der Galerkin-Orthogonalität

$$a(u(p) - u_{\mathcal{N}}(p), \varphi; p) = 0, \quad \varphi \in X_{\mathcal{N}},$$

nach Céa's Lemma

$$\|u(p) - u_{\mathcal{N}}(p)\|_{X_e} \leq \left(1 + \frac{\gamma(p)}{\alpha(p)}\right) \inf_{\varphi \in X_{\mathcal{N}}} \|u(p) - \varphi\|_{X_e}. \quad (4.5)$$

Auch in dem Fall einer inf-sup-stabilen Sesquilinearform, bestimmt durch die Bedingung

$$0 < \beta(p) = \inf_{u \in X_e \setminus \{0\}} \sup_{v \in X_e \setminus \{0\}} \frac{|a(u, v; p)|}{\|u\|_{X_e} \|v\|_{X_e}}, \quad (4.6)$$

ergibt sich gerade

$$\|u(p) - u_{\mathcal{N}}(p)\|_{X_e} \leq \left(1 + \frac{\gamma(p)}{\beta(p)}\right) \inf_{\varphi \in X_{\mathcal{N}}} \|u(p) - \varphi\|_{X_e}. \quad (4.7)$$

Das bedeutet, dass die Qualität der Approximation von der Stetigkeitskonstanten  $\gamma(p)$  und der Koerzivitatskonstanten  $\alpha(p)$ , bzw. der inf-sup-Konstanten  $\beta(p)$ , abhangt. Die Idee der RBM ist es, eine weitere Projektion auf  $X_N \subset X_{\mathcal{N}}$  durchzufuhren. Dazu wird die Losungsmannigfaltigkeit

$$\mathcal{M} = \{u(p) \mid u(p) \text{ ist Losung von (4.2) und } p \in \mathbb{P}\} \subset X_e$$

der variationellen Formulierung bezuglich des Parameterraums betrachtet. Die diskretisierte Variante lautet

$$\mathcal{M}_{\mathcal{N}} = \{u_{\mathcal{N}}(p) \mid u_{\mathcal{N}}(p) \text{ ist Losung von (4.3) und } p \in \mathbb{P}\} \subset X_{\mathcal{N}}.$$

Die Voraussetzung ist nun, dass  $\mathcal{M}_{\mathcal{N}}$  durch einen Raum niedriger Dimension approximiert werden kann. In diesem Fall werden weniger Basisfunktionen  $\xi_1, \dots, \xi_N \in \mathcal{M}_{\mathcal{N}}$  benotigt, um  $u_{\mathcal{N}}(p) \in \mathcal{M}_{\mathcal{N}}$  zu beschreiben. Wir geben nachfolgend zwei verschiedene Varianten der Projektion des FE-Raumes  $X_{\mathcal{N}}$  auf den RB-Raum

$$X_N = \text{span} \{\xi_1, \dots, \xi_N\} \subset \mathcal{M}_{\mathcal{N}}$$

an. Die erste der beiden Projektionen ist die Ritz-Galerkin-Projektion und das zugehorige Problem ist gegeben durch:

Bestimme  $u_N^{RG}(p) \in X_N \subset X_{\mathcal{N}}$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u_N^{RG}(p), \xi; p) = f(\xi; p), \quad \xi \in X_N, \quad (4.8)$$

gilt.

Im Fall koerziver Sesquilinearformen, kann damit Stabilitat garantiert werden, siehe [14]. Dabei bedeutet Stabilitat, dass die Koerzivitatskonstante  $\alpha_N(p)$  des reduzierten Systems

die Abschätzung

$$\alpha_N(p) = \inf_{u_N \in X_N \setminus \{0\}} \frac{|a(u_N, u_N; p)|}{\|u_N\|_{X_e}^2} > \alpha_{\mathcal{N}}(p) = \inf_{u_{\mathcal{N}} \in X_{\mathcal{N}} \setminus \{0\}} \frac{|a(u_{\mathcal{N}}, u_{\mathcal{N}}; p)|}{\|u_{\mathcal{N}}\|_{X_e}^2}$$

bezüglich der Koerzitivitätskonstante  $\alpha_{\mathcal{N}}(p)$  des diskreten Systems erfüllt. Mit der Basisdarstellung im Unterraum  $X_N$  lässt sich jede Funktion  $u_N(p) \in X_N$  als Linearkombination

$$u_N(p) = \sum_{k=1}^N u_k^N(p) \xi_k \quad (4.9)$$

der Basisfunktionen  $\xi_k$  mit Koeffizientenvektor  $(u_k^N(p))_{1 \leq k \leq N}$  schreiben. Die Basisfunktionen  $\xi_k \in \mathcal{M}_N$  des reduzierten Problems lassen sich wiederum durch die FEM-Basisdarstellung des Raumes  $X_N$  als

$$\xi_l = \sum_{k=1}^{\mathcal{N}} \xi_{k,l}^{\mathcal{N}} \varphi_k, \quad l = 1, \dots, N,$$

darstellen. Setzen wir diese Darstellungen in (4.8) ein, so ergibt sich das reduzierte Gleichungssystem

$$A_p^{RG} u^{N, RG}(p) = f_p^{RG} \quad (4.10)$$

mit

$$\begin{aligned} (A_p^{RG})_{m,n} &= a(\xi_n, \xi_m; p) \quad \text{mit } m, n = 1, \dots, N, \\ (f_p^{RG})_m &= f(\xi_m; p) \quad \text{mit } m = 1, \dots, N, \\ (u^{N, RG}(p))_m &= u_m^N(p) \quad \text{mit } m = 1, \dots, N. \end{aligned}$$

Dabei sind die Matrix und der Vektor des Gleichungssystems gegeben durch

$$A_p^{RG} = \Xi^H A_p \Xi, \quad f_p^{RG} = \Xi^H f_p,$$

mit  $(\Xi)_{m,n} = \xi_{m,n}^{\mathcal{N}}$ . Um die Stabilität bei inf-sup-stabilen Sesquilinearformen zu garantieren, wird eine Petrov-Galerkin-Projektion benötigt, siehe [19]. Das zur Petrov-Galerkin-Projektion zugehörige Problem ist gegeben durch:

Bestimme  $u_N^{PG}(p) \in X_N$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u_N^{PG}(p), \psi; p) = f(\psi; p), \quad \psi \in Y_N, \quad (4.11)$$

gilt.

Um einen geeigneten Raum  $Y_N \subset X_N$  für die Stabilität zu bestimmen, führen wir den Supremumsoperator  $\mathcal{T}(p) : X_e \rightarrow X_e$  ein, welcher durch die Eigenschaft

$$(\mathcal{T}(p)[u], v)_{X_e} = a(u, v; p), \quad v \in X_e, \quad (4.12)$$

charakterisiert ist. Aus dieser Charakterisierungseigenschaft lässt sich zeigen, dass der Operator  $\mathcal{T}(p)$  gerade

$$\sup_{v \in X_e \setminus \{0\}} \frac{|a(u, v; p)|}{\|v\|_{X_e}} = \frac{|a(u, \mathcal{T}(p)[u]; p)|}{\|\mathcal{T}(p)[u]\|_{X_e}} \quad (4.13)$$

mit  $u \in X_e$  erfüllt, woraus sich auch der Name Supremumsoperator ableitet. Der Beweis der Eigenschaft (4.13) und weitere Eigenschaften zu den Supremumsoperatoren können in [14, 19] nachgelesen werden. Mit der Definition des Supremumsoperators ergibt sich für die inf-sup-Konstante  $\beta(p)$  aus (4.6) die Darstellung

$$\beta(p) = \inf_{u \in X_e \setminus \{0\}} \frac{\|\mathcal{T}(p)[u]\|_{X_e}^2}{\|u\|_{X_e} \|\mathcal{T}(p)[u]\|_{X_e}} = \inf_{u \in X_e \setminus \{0\}} \frac{\|\mathcal{T}(p)[u]\|_{X_e}}{\|u\|_{X_e}}. \quad (4.14)$$

Dieselben Aussagen gelten auch für das diskretisierte Problem (4.3) mit zugehöriger inf-sup-Konstante

$$\beta_N(p) = \inf_{u_N \in X_N \setminus \{0\}} \sup_{v_N \in X_N \setminus \{0\}} \frac{|a(u_N, v_N; p)|}{\|u_N\|_{X_e} \|v_N\|_{X_e}} = \inf_{u_N \in X_N \setminus \{0\}} \frac{\|\mathcal{T}_N(p)[u_N]\|_{X_e}}{\|u_N\|_{X_e}}, \quad (4.15)$$

indem wir den diskreten Supremumsoperator  $\mathcal{T}_N(p) : X_N \rightarrow X_N$  definieren, der wiederum durch die Eigenschaft

$$(\mathcal{T}_N(p)[u_N], v_N)_{X_e} = a(u_N, v_N; p), \quad v_N \in X_N, \quad (4.16)$$

charakterisiert ist. Definieren wir den Testraum als

$$Y_N = \text{span} \{ \mathcal{T}_N(p)[\xi_1], \dots, \mathcal{T}_N(p)[\xi_N] \},$$

dann erhalten wir für die Petrov-Galerkin-Projektion (4.11) und die zugehörigen inf-sup-Konstanten

$$\beta_N^{PG}(p) = \inf_{u_N \in X_N \setminus \{0\}} \sup_{v_N \in Y_N \setminus \{0\}} \frac{|a(u_N, v_N; p)|}{\|u_N\|_{X_e} \|v_N\|_{X_e}} \quad (4.17)$$

die gewünschte Stabilität.

**Lemma 4.1.** Für die Petrov-Galerkin-Projektion (4.11) gilt

$$\beta_N^{PG}(p) \geq \beta_N(p).$$

*Beweis.* Für die inf-sup-Konstanten  $\beta_N^{PG}(p)$  in (4.17) und  $\beta_N(p)$  in (4.15) erhalten wir

$$\begin{aligned} \beta_N^{PG}(p) &= \inf_{u_N \in \tilde{X}_N} \sup_{v_N \in \tilde{Y}_N} \frac{|a(u_N, v_N; p)|}{\|u_N\|_{X_e} \|v_N\|_{X_e}} = \inf_{u_N \in \tilde{X}_N} \frac{|a(u_N, \mathcal{T}_N(p)[u_N]; p)|}{\|u_N\|_{X_e} \|\mathcal{T}_N(p)[u_N]\|_{X_e}} \\ &= \inf_{u_N \in \tilde{X}_N} \frac{\|\mathcal{T}_N(p)[u_N]\|_{X_e}}{\|u_N\|_{X_e}} \geq \inf_{u_N \in \tilde{X}_N} \frac{\|\mathcal{T}_N(p)[u_N]\|_{X_e}}{\|u_N\|_{X_e}} \\ &= \beta_N(p), \end{aligned}$$

wobei  $\tilde{X}_N = X_N \setminus \{0\}$ ,  $\tilde{X}_N = X_N \setminus \{0\}$  und  $\tilde{Y}_N = Y_N \setminus \{0\}$  bezeichnet.  $\square$

Mit den Basisdarstellungen für  $u_N(p)$  und  $\xi_k$  ergibt sich aus (4.11) das reduzierte Gleichungssystem

$$A_p^{PG} u^{N,PG}(p) = f_p^{PG} \quad (4.18)$$

mit

$$\begin{aligned} (A_p^{PG})_{m,n} &= a(\xi_n, \mathcal{T}_N(p)[\xi_m]; p) \quad \text{mit } m, n = 1, \dots, N, \\ (f_p^{PG})_m &= f(\mathcal{T}_N(p)[\xi_m]; p) \quad \text{mit } m = 1, \dots, N, \\ (u^{N,PG}(p))_m &= u_m^N(p) \quad \text{mit } m = 1, \dots, N. \end{aligned}$$

Die Matrix und der Vektor in (4.18) sind dabei gegeben durch

$$A_p^{PG} = (\mathcal{T}_p^N(\Xi))^H A_p \Xi, \quad f_p^{PG} = (\mathcal{T}_p^N(\Xi))^H f_p,$$

mit  $(\Xi)_{m,n} = \xi_{m,n}^N$  und  $(\mathcal{T}_p^N(\Xi))_{m,n} = (\mathcal{T}_N(p)[\xi_m])_n^N$  berechnen. Dabei bezeichnen wir mit  $(\mathcal{T}_N(p)[\xi_m])_n^N$  den  $n$ -ten Koeffizient in der FEM-Basisdarstellung

$$\mathcal{T}_N(p)[\xi_m] = \sum_{n=1}^N (\mathcal{T}_N(p)[\xi_m])_n^N \varphi_n.$$

Eine genauere Betrachtung des Operators  $\mathcal{T}_N(p)$  ist in Abschnitt 4.1.4 gegeben. Die folgenden Abschnitte differenzieren nicht zwischen der Ritz-Galerkin-Projektion und der Petrov-Galerkin-Projektion, weswegen wir  $A_p^N$  und  $f_p^N$  als Verallgemeinerung schreiben, wenn wir keine spezifische Projektion betrachten. Die Matrix  $A_p^N$  und der Vektor  $f_p^N$  hängen dabei vom Parameter  $p$  ab und müssen für jeden neuen Parameter berechnet werden. Für die effiziente Berechnung des RB-Systems für beliebige Parameter benötigen wir eine spezielle Darstellung der einzelnen Komponenten, welche wir nun einführen.

### 4.1.2. Offline-Online-Zerlegung

Eine Offline-Online-Zerlegung besteht daraus, dass ein Teil der Methode in der Offline-Phase berechnet und in der Online-Phase auf die berechneten Daten aus der Offline-Phase zurückgegriffen wird. Dabei soll die Erstellung der reduzierten Matrix  $A_p^N$  und des Vektors  $f_p^N$  in der Offline-Phase erfolgen, um damit die Lösung von (4.10) in der Online-Phase für beliebige Parameter  $p \in \mathbb{P}$  schnell zu berechnen. Diese Zerlegung wird durch eine affine Darstellung ermöglicht.

**Definition 4.2** (Affine Darstellung). Für das parameterabhängige Funktional  $\mathcal{F}[u; p]$  wird

$$\mathcal{F}[u; p] = \sum_{k=1}^{Q_F} \Theta_k(p) \mathcal{F}_k[u], \quad (4.19)$$

mit  $\Theta_k : \mathbb{P} \mapsto \mathbb{K}$ ,  $\mathcal{F}_k : X_e \mapsto \mathbb{K}$ , wobei  $\mathbb{K}$  der zugrundeliegende Zahlenkörper, als *affine Zerlegung* oder *affine Darstellung* bezeichnet.

Wir fordern, dass die Sesquilinearform  $a$  und die Antilinearform  $f$  die affinen Darstellungen

$$a(u, v; p) = \sum_{k=1}^{Q_a} \Theta_k^a(p) a_k(u, v), \quad (4.20)$$

$$f(v; p) = \sum_{k=1}^{Q_f} \Theta_k^f(p) f_k(v) \quad (4.21)$$

besitzen. Denn damit ergibt sich bei der Diskretisierung von (4.2) das lineare Gleichungssystem (4.4) mit

$$A_p^N = \sum_{k=1}^{Q_a} \Theta_k^a(p) A_k^N, \quad f_p^N = \sum_{k=1}^{Q_f} \Theta_k^f(p) f_k^N \quad (4.22)$$

und

$$\begin{aligned} (A_k^N)_{m,n} &= a_k(\varphi_n, \varphi_m) \quad \text{mit } m, n = 1, \dots, \mathcal{N}, \\ (f_k^N)_m &= f_k(\varphi_m) \quad \text{mit } m = 1, \dots, \mathcal{N}, \end{aligned}$$

wobei  $\varphi_n, \varphi_m$  die FEM-Basisfunktionen darstellen. Eine äquivalente Darstellung wie (4.22) gilt auch für die reduzierte Matrix  $A_p^N$  und den Vektor  $f_p^N$ . Mit diesen affinen Darstellungen wird die Offline-Online-Zerlegung ermöglicht, indem die Matrizen  $A_k^N$  und die Vektoren  $f_k^N$  in der Offline-Phase assembliert, reduziert und dann abgespeichert werden. Bei der numerischen Reduktion wird dazu die Teilmenge  $\mathbb{P}_{NRB} \subset \mathbb{P}$  verwendet, auf dem die Basisvektoren  $\xi_k$  des reduzierten Systems durch die Snapshot-Lösungen  $u_N(p)$ ,

$p \in \mathbb{P}_{N_{RB}}$ , des FE-Problems bestimmt werden. In der Online-Phase werden dann die reduzierten Matrizen verwendet, um für beliebige Parameter  $p \in \mathbb{P}$  eines der reduzierten Probleme (4.10) oder (4.18) zu lösen. Als nächstes betrachten wir die optimale Wahl der Parameter  $p$  für die Bestimmung der reduzierten Basis.

### 4.1.3. Basisgenerierung des reduzierten Raumes

Um die bestmögliche Wahl der Parameter zu treffen, die für die Berechnung der Snapshot-Lösungen verwendet werden, gibt es verschiedene Methoden. Die am häufigsten verwendeten Methoden sind die Proper Orthogonal Decomposition (POD) und der Greedy-Algorithmus. Im POD-Ansatz wird der  $N$ -dimensionale Raum  $X_N \subset X_{\mathcal{N}}$  bestimmt, sodass der gemittelte Approximationsfehler

$$\sqrt{\frac{1}{N_{RB}} \sum_{p \in \mathbb{P}_{N_{RB}}} \inf_{v_N \in X_N} \|u_{\mathcal{N}}(p) - v_N\|_{X_e}^2}$$

mit diskretisierter Menge  $\mathbb{P}_{N_{RB}} \subset \mathbb{P}$  und  $N_{RB} = |\mathbb{P}_{N_{RB}}|$  minimal wird. Dazu wird für alle  $p \in \mathbb{P}_{N_{RB}}$  die Snapshot-Lösung  $u_{\mathcal{N}}(p)$  berechnet und mit diesen ein linearer Operator definiert, dessen  $N$  größten Eigenwerte  $\lambda_k$  das obige Minimierungsproblem lösen. Dies setzt  $N_{RB}$  Lösungen des FE-Problems voraus, welche bei komplexen großdimensionalen Systemen einen hohen Rechenaufwand besitzen. Ein weiterer Nachteil ist die mögliche Berechnung redundanter Informationen, das heißt, dass die Berechnung der Snapshot-Lösung für ein  $\tilde{p}$  nicht notwendig ist, da eine Basis kleinerer Dimension die Lösung  $u(\tilde{p})$  bereits gut genug approximiert. Auch wenn wir die Zerlegung in Offline-Online-Phasen haben und deswegen der Zeitaufwand in der Offline-Phase vernachlässigbar ist, möchten wir eine nicht so rechenaufwändige Methode für die Generierung der reduzierten Basis verwenden. Eine genauere Beschreibung der POD-Methode findet sich in [14]. Die andere Möglichkeit für eine gute Wahl der Parameter ist ein Greedy-Algorithmus, siehe Algorithmus 1, welcher gute Ergebnisse liefert und deshalb sehr häufig verwendet wird. Näheres zu den Konvergenzraten des Greedy-Algorithmus kann in [50] nachgelesen werden. Die Idee des Algorithmus 1 ist es den Parameter zum größtmöglichen Fehler zu verwenden, um die nächste Snapshot-Lösung  $u_{\mathcal{N}}(p)$  für die reduzierte Basis zu berechnen. Die Konvergenz des Algorithmus 1 wird in Abhängigkeit von der Kolmogorov- $N$ -Breite

$$d_N(\mathcal{M}_{\mathcal{N}}) = \inf_{X_N} \sup_{u_{\mathcal{N}} \in \mathcal{M}_{\mathcal{N}}} \inf_{v_N \in X_N} \|u_{\mathcal{N}} - v_N\|_{X_e} \quad (4.23)$$

in [50] bewiesen. Eine Verbesserung innerhalb des Greedy-Algorithmus 1 kann dadurch erreicht werden, dass die Basismatrix  $\Xi$  zusätzlich orthonormalisiert wird. Dadurch soll das Auftreten von Singularitäten in der Matrix  $A_p^N$ , die durch fast linear abhängige Vektoren  $u^{\mathcal{N}}$  entstehen, vermieden werden. Dabei muss jedoch gewährleistet werden, dass

---

**Algorithm 1** Generiere die Menge der Snapshot-Lösungen und die resultierende reduzierte Basis  $\Xi$  des Projektionsraums  $X_N$

---

**Require:** Diskrete Parametermenge  $\mathbb{P}_{N_{RB}}$ , Toleranz  $tol$ , Fehlerschätzer  $\Delta_N$

**Ensure:**  $\mathbb{P}_\Xi, \Xi$

$p_1 \leftarrow p \in \mathbb{P}_{N_{RB}}$ .

Löse (4.3) mit  $p = p_1$  for  $u^{\mathcal{N}}(p_1)$ .

$\mathbb{P}_\Xi \leftarrow \{p_1\}$ .

$\Xi \leftarrow [u^{\mathcal{N}}(p_1)]$ .

**while**  $\max_{p \in \mathbb{P}_{N_{RB}} \setminus \mathbb{P}_\Xi} \Delta_N(p) \geq tol$  **do**

$p_{N+1} \leftarrow \arg \max_{p \in \mathbb{P}_{N_{RB}} \setminus \mathbb{P}_\Xi} \Delta_N(p)$ .

Löse (4.3) mit  $p = p_{N+1}$  für  $u^{\mathcal{N}}(p_{N+1})$ .

$\mathbb{P}_\Xi \leftarrow \mathbb{P}_\Xi \cup \{p_{N+1}\}$ .

$\Xi \leftarrow [\Xi \ u^{\mathcal{N}}(p_{N+1})]$ .

$N \leftarrow N + 1$ .

**end while**

---

nach Orthonormalisierung die zu den einzelnen Spalten gehörige FE-Funktion  $\xi$  wiederum im Lösungsraum  $X_{\mathcal{N}}$  liegt. Bei der Wahl des Fehlerschätzers  $\Delta_N(p)$  gibt es unterschiedliche Ansätze, wobei bei allen die Evaluierung unabhängig von  $\mathcal{N}$  sein soll, damit eine effiziente Berechnung ermöglicht wird. Hierzu betrachten wir innerhalb der nächsten Abschnitte die Fehlertheorie der RBM und leiten daraus geeignete Fehlerschätzer her.

#### 4.1.4. Fehlerdarstellung

Um geeignete Fehlerschranken herzuleiten, um daraus einen Fehlerschätzer  $\Delta_N(p)$  abzuleiten, betrachten wir das Residuum

$$r(\varphi; p) = a(e(p), \varphi; p) = f(\varphi; p) - a(u_N(p), \varphi; p), \quad \varphi \in X_{\mathcal{N}}, \quad (4.24)$$

für den Fehler  $e(p) = u_{\mathcal{N}}(p) - u_N(p) \in X_{\mathcal{N}}$ , wobei wir die Linearität von  $a$  im ersten Argument ausnutzen. Dabei betrachten wir den Fehler der reduzierten Lösung  $u_N(p)$  nicht in Bezug auf die exakte Lösung  $u(p) \in X_e$ , sondern bezüglich der FE-Lösung  $u_{\mathcal{N}}(p)$ , da der Fehler

$$\|u(p) - u_{\mathcal{N}}(p)\|_{X_e}$$

bereits durch eine feine Diskretisierung klein ist. Da  $a$  in den zugrundeliegenden Zahlenkörper  $\mathbb{C}$  abbildet, können wir  $r(\cdot; p) : X_{\mathcal{N}} \rightarrow \mathbb{C}$  als Element des Dualraums  $(X_{\mathcal{N}})'$  auffassen. Nach dem Rieszschen Darstellungssatz A.16 existiert ein Element in  $X_{\mathcal{N}}$ , wel-

#### 4. Reduzierte-Basis-Methode

---

ches wir als  $\hat{r}(p)$  bezeichnen, sodass gilt

$$(\hat{r}(p), \varphi)_{X_e} = r(\varphi; p), \quad \varphi \in X_{\mathcal{N}}. \quad (4.25)$$

Somit gilt

$$\|\hat{r}(p)\|_{X_e} = \sup_{\varphi \in X_{\mathcal{N}} \setminus \{0\}} \frac{|r(\varphi; p)|}{\|\varphi\|_{X_e}} \quad (4.26)$$

und

$$(\hat{r}(p), \varphi)_{X_e} = a(e(p), \varphi; p). \quad (4.27)$$

Mithilfe dieser Darstellung lassen sich rigoros berechenbare Schranken herleiten, für die wir im koerziven Fall die Koerzivitätskonstante  $\alpha(p)$  und im inf-sup-stabilem Fall die inf-sup-Konstante  $\beta(p)$ , wie in Abschnitt 4.1.1 definiert, benötigen. Da unsere Probleme lediglich inf-sup-stabil sind, nehmen wir im Folgenden an, dass eine Konstante  $\beta_{LB} > 0$  existiert, für die gilt

$$\beta_{\mathcal{N}}(p) = \inf_{u_{\mathcal{N}} \in X_{\mathcal{N}} \setminus \{0\}} \sup_{v_{\mathcal{N}} \in X_{\mathcal{N}} \setminus \{0\}} \frac{|a(u_{\mathcal{N}}, v_{\mathcal{N}}; p)|}{\|u_{\mathcal{N}}\|_{X_e} \|v_{\mathcal{N}}\|_{X_e}} \geq \beta_{LB}, \quad p \in \mathbb{P}. \quad (4.28)$$

Des Weiteren verwenden wir den Supremumsoperator  $\mathcal{T}_{\mathcal{N}}(p)$  aus (4.16). Wir erhalten damit die folgende Fehlerschranke.

**Proposition 4.3.** *Für den Fehler  $e(p) = u_{\mathcal{N}}(p) - u_N(p)$  gilt*

$$\|u_{\mathcal{N}}(p) - u_N(p)\|_{X_e} \leq \frac{\|\hat{r}(p)\|_{X_e}}{\beta_{LB}} = \Delta_{en}(p), \quad (4.29)$$

wobei  $\beta_{LB}$  gegeben ist durch (4.28) und  $\hat{r}(p)$  der Riesz-Repräsentant (4.25) des Residuums in (4.24).

*Beweis.* Wegen (4.13) und (4.16) folgt

$$\begin{aligned} \|\hat{r}(p)\|_{X_e} &= \sup_{\varphi \in X_{\mathcal{N}} \setminus \{0\}} \frac{|r(\varphi; p)|}{\|\varphi\|_{X_e}} = \sup_{\varphi \in X_{\mathcal{N}} \setminus \{0\}} \frac{|a(e(p), \varphi; p)|}{\|\varphi\|_{X_e}} \\ &= \frac{|a(e(p), \mathcal{T}_{\mathcal{N}}(p)[e(p)]; p)|}{\|\mathcal{T}_{\mathcal{N}}(p)[e(p)]\|_{X_e}} = \frac{|(\mathcal{T}_{\mathcal{N}}(p)[e(p)], \mathcal{T}_{\mathcal{N}}(p)[e(p)])_{X_e}|}{\|\mathcal{T}_{\mathcal{N}}(p)[e(p)]\|_{X_e}} = \|\mathcal{T}_{\mathcal{N}}(p)[e(p)]\|_{X_e} \end{aligned}$$

und damit

$$\|e(p)\|_{X_e} = \|e(p)\|_{X_e} \frac{\|\hat{r}(p)\|_{X_e}}{\|\mathcal{T}_{\mathcal{N}}(p)[e(p)]\|_{X_e}} \leq \frac{\|\hat{r}(p)\|_{X_e}}{\inf_{\varphi \in X_{\mathcal{N}} \setminus \{0\}} \frac{\|\mathcal{T}_{\mathcal{N}}(p)[\varphi]\|_{X_e}}{\|\varphi\|_{X_e}}} = \frac{\|\hat{r}(p)\|_{X_e}}{\beta_{\mathcal{N}}(p)} \leq \frac{\|\hat{r}(p)\|_{X_e}}{\beta_{LB}}.$$

□

Während für den relativen Fehler bezüglich der reduzierten Lösung

$$\frac{\|u_{\mathcal{N}}(p) - u_N(p)\|_{X_e}}{\|u_N(p)\|_{X_e}} \leq \frac{\Delta_{en}(p)}{\|u_N(p)\|_{X_e}} \quad (4.30)$$

gilt, haben wir für den relativen Fehler bezüglich der FEM-Lösung

$$\frac{\|u_{\mathcal{N}}(p) - u_N(p)\|_{X_e}}{\|u_{\mathcal{N}}(p)\|_{X_e}} \leq \frac{\frac{\Delta_{en}(p)}{\|u_N(p)\|_{X_e}}}{1 - \frac{\Delta_{en}(p)}{\|u_N(p)\|_{X_e}}}, \quad (4.31)$$

falls  $\frac{\Delta_{en}(p)}{\|u_N(p)\|_{X_e}} < 1$ . Um die Schranken verwenden zu können, müssen wir sowohl  $\|\hat{r}(p)\|_{X_e}$  als auch die Schranke  $\beta_{LB}$  berechnen.

#### 4.1.5. Berechnung des Riesz-Repräsentanten

Um den Riesz-Repräsentanten  $\hat{r}(p)$  des Residuums zu bestimmen, betrachten wir (4.9) und (4.25) in Verbindung mit (4.20) und (4.21)

$$(\hat{r}(p), \varphi)_{X_e} = \sum_{k=1}^{Q_f} \Theta_k^f(p) f_k(\varphi) - \sum_{k=1}^{Q_a} \sum_{l=1}^N \Theta_k^a(p) u_l^N a_k(\xi_l, \varphi), \quad \varphi \in X_{\mathcal{N}}. \quad (4.32)$$

Aufgrund der Linearität der Sesquilinearform  $a$  und der Antilinearform  $f$  erhalten wir  $\hat{r}(p)$  durch

$$\hat{r}(p) = \sum_{k=1}^{Q_f} \Theta_k^f(p) \mathcal{C}_k + \sum_{k=1}^{Q_a} \sum_{l=1}^N \Theta_k^a(p) u_l^N \mathcal{L}_k^l, \quad (4.33)$$

wobei  $\mathcal{C}_k$  und  $\mathcal{L}_k^l$  die Riesz-Repräsentanten zu den Gleichungen

$$(\mathcal{C}_k, \varphi)_{X_e} = f_k(\varphi), \quad k = 1, \dots, Q_f, \quad (4.34)$$

$$(\mathcal{L}_k^l, \varphi)_{X_e} = -a_k(\xi_l, \varphi), \quad k = 1, \dots, Q_a, \quad l = 1, \dots, N, \quad (4.35)$$

mit  $\varphi \in X_{\mathcal{N}}$  sind. Die Gleichungen (4.34) und (4.35) werden innerhalb der Literatur als *FE-Poisson*-Probleme bezeichnet. Um  $\mathcal{C}_k$  und  $\mathcal{L}_k^l$  zu bestimmen, fordern wir  $\mathcal{C}_k, \mathcal{L}_k^l \in X_{\mathcal{N}}$  und verwenden so die Basisdarstellung in  $X_{\mathcal{N}}$ , um Koeffizientenvektoren zu erhalten. Diese hängen nicht von den Parametern ab, aber  $\mathcal{L}_k^l$  hängt von den Basisfunktionen  $\xi_l$  des Raumes  $X_{\mathcal{N}}$  ab, weswegen die Berechnung von  $\mathcal{C}_k$  nur einmal durchgeführt werden muss, während für jede neue Basisfunktion  $\xi_l$  genau  $Q_a$  Lösungen  $\mathcal{L}_k^l$  bestimmt werden müssen. Nach der Bestimmung der FE-Poisson-Lösungen lässt sich der Riesz-Repräsentant  $\hat{r}(p)$  dann allgemein für alle Parameter mithilfe der affinen Darstellung (4.33) bestimmen.

#### 4.1.6. Berechnung der inf-sup-Konstante

Um die Fehlerschranken aus Abschnitt 4.1.4 verwenden zu können, benötigen wir noch die Schranke  $\beta_{LB}$  an die inf-sup-Konstanten  $\beta_{\mathcal{N}}(p)$ . Dazu betrachten wir das zu (4.15) zugehörige Minimierungsproblem

$$\beta_{\mathcal{N}}(p) = \min_{u_{\mathcal{N}} \in X_{\mathcal{N}} \setminus \{0\}} \frac{\|\mathcal{T}_{\mathcal{N}}(p)[u_{\mathcal{N}}]\|_{X_e}}{\|u_{\mathcal{N}}\|_{X_e}}.$$

Um dieses Problem zu lösen, betrachten wir zunächst die Darstellung zweier Funktionen  $u_{\mathcal{N}}, v_{\mathcal{N}} \in X_{\mathcal{N}}$  im Skalarprodukt

$$(u_{\mathcal{N}}, v_{\mathcal{N}})_{X_e} = (v^{\mathcal{N}})^H H_{X_{\mathcal{N}}} u^{\mathcal{N}},$$

wobei hier  $(H_{X_{\mathcal{N}}})_{m,n} = (\varphi_n, \varphi_m)_{X_e}$  und  $u^{\mathcal{N}}, v^{\mathcal{N}}$  die FEM-Koeffizientenvektoren der FEM-Darstellung sind. Aufgrund der FEM-Darstellung der Funktionen  $u_{\mathcal{N}} \in X_{\mathcal{N}}$  und der Abbildungseigenschaft  $\mathcal{T}_{\mathcal{N}}(p)[u_{\mathcal{N}}] \in X_{\mathcal{N}}$  definieren wir den zu  $\mathcal{T}_{\mathcal{N}}(p)$  zugehörigen Operator  $\mathcal{T}_p^{\mathcal{N}} : \mathbb{C}^{\mathcal{N}} \rightarrow \mathbb{C}^{\mathcal{N}}$ , der den FEM-Koeffizientenvektor  $u^{\mathcal{N}}$  der Funktion  $u_{\mathcal{N}}$  auf den FEM-Koeffizientenvektor der Funktion  $\mathcal{T}_{\mathcal{N}}(p)[u_{\mathcal{N}}]$  abbildet. Betrachten wir (4.16) mit der Darstellung des Skalarproduktes für finite Elemente, dann ergibt sich

$$(v^{\mathcal{N}})^H H_{X_{\mathcal{N}}} (\mathcal{T}_p^{\mathcal{N}}[u^{\mathcal{N}}]) = (v^{\mathcal{N}})^H A_p^{\mathcal{N}} u^{\mathcal{N}}$$

mit  $u_{\mathcal{N}}, v_{\mathcal{N}} \in X_{\mathcal{N}}$ . Damit ist  $\mathcal{T}_p^{\mathcal{N}} = (H_{X_{\mathcal{N}}})^{-1} A_p^{\mathcal{N}}$  und so folgt

$$\begin{aligned} (\beta_{\mathcal{N}}(p))^2 &= \min_{u_{\mathcal{N}} \in X_{\mathcal{N}} \setminus \{0\}} \frac{\|\mathcal{T}_{\mathcal{N}}(p)[u_{\mathcal{N}}]\|_{X_e}^2}{\|u_{\mathcal{N}}\|_{X_e}^2} \\ &= \min_{u^{\mathcal{N}} \in \mathbb{C}^{\mathcal{N}} \setminus \{0\}} \frac{(u^{\mathcal{N}})^H (A_p^{\mathcal{N}})^H H_{X_{\mathcal{N}}}^{-1} A_p^{\mathcal{N}} u^{\mathcal{N}}}{(u^{\mathcal{N}})^H H_{X_{\mathcal{N}}} u^{\mathcal{N}}}. \end{aligned}$$

Um nun dieses Minimierungsproblem zu lösen geben wir die folgende Definition an und beweisen die folgenden Sätze.

**Definition 4.4.** Für die komplexe Matrix  $B \in \mathbb{C}^{\mathcal{N}_1 \times \mathcal{N}_1}$  definiere die reellen Matrizen

$$M[B] = \begin{pmatrix} \operatorname{Re}(B) & -\operatorname{Im}(B) \\ \operatorname{Im}(B) & \operatorname{Re}(B) \end{pmatrix}, \quad \hat{M}[B] = \begin{pmatrix} \operatorname{Re}(B) & -\operatorname{Im}(B) \\ -\operatorname{Im}(B) & -\operatorname{Re}(B) \end{pmatrix}. \quad (4.36)$$

**Satz 4.5.** Sei  $\mathcal{N}_K \in \mathbb{N}$ ,  $v \in \mathbb{C}^{\mathcal{N}_K}$ ,  $B_k \in \mathbb{C}^{\mathcal{N}_{k,1} \times \mathcal{N}_{k,2}}$ ,  $\mathcal{N}_{k,1}, \mathcal{N}_{k,2} \in \mathbb{N}$  und  $1 \leq k \leq K$ , mit

$$\mathcal{N}_{l,2} = \mathcal{N}_{l+1,1}, \quad \mathcal{N}_{K,2} = \mathcal{N}_K,$$

wobei  $1 \leq l \leq K-1$ , und  $H \in \mathbb{R}^{\mathcal{N}_{1,1} \times \mathcal{N}_{1,1}}$  eine reelle symmetrische Matrix, dann ist das komplexe Produkt

$$v^H \left( \prod_{l=2}^K B_l \right)^H B_1^H H B_1 \left( \prod_{l=2}^K B_l \right) v$$

äquivalent zum reellen Produkt

$$v_R^T \left( \prod_{l=2}^K M[B_l] \right)^T \hat{M}[B_1]^T \mathcal{H} \hat{M}[B_1] \left( \prod_{l=2}^K M[B_l] \right) v_R$$

mit

$$v_R = \begin{pmatrix} \operatorname{Re}(v) \\ \operatorname{Im}(v) \end{pmatrix}, \quad \mathcal{H} = \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}.$$

*Beweis.* Wir führen den Beweis durch die vollständige Induktion über  $K$ :

IA: Wir schreiben zunächst für den komplexen Vektor und Matrix

$$\tilde{v}_r = \operatorname{Re}(v), \quad \tilde{v}_i = \operatorname{Im}(v), \quad \tilde{B}_r = \operatorname{Re}(B_1), \quad \tilde{B}_i = \operatorname{Im}(B_1)$$

und erhalten

$$\begin{aligned} & v^H B_1^H H B_1 v \\ &= (\tilde{v}_r + i\tilde{v}_i)^H (\tilde{B}_r + i\tilde{B}_i)^H H (\tilde{B}_r + i\tilde{B}_i) (\tilde{v}_r + i\tilde{v}_i) \\ &= (\tilde{v}_r^T \tilde{B}_r^T - \tilde{v}_i^T \tilde{B}_i^T - i\tilde{v}_i^T \tilde{B}_r^T - i\tilde{v}_r^T \tilde{B}_i^T) H (\tilde{B}_r \tilde{v}_r - \tilde{B}_i \tilde{v}_i + i\tilde{B}_r \tilde{v}_i + i\tilde{B}_i \tilde{v}_r) \\ &= \tilde{v}_r^T \tilde{B}_r^T H \tilde{B}_r \tilde{v}_r - \tilde{v}_r^T \tilde{B}_r^T H \tilde{B}_i \tilde{v}_i - \tilde{v}_i^T \tilde{B}_i^T H \tilde{B}_r \tilde{v}_r + \tilde{v}_i^T \tilde{B}_i^T H \tilde{B}_i \tilde{v}_i \\ &\quad + \tilde{v}_i^T \tilde{B}_r^T H \tilde{B}_r \tilde{v}_i + \tilde{v}_i^T \tilde{B}_r^T H \tilde{B}_i \tilde{v}_r + \tilde{v}_r^T \tilde{B}_i^T H \tilde{B}_r \tilde{v}_i + \tilde{v}_r^T \tilde{B}_i^T H \tilde{B}_i \tilde{v}_r \\ &= \begin{pmatrix} \tilde{v}_r \\ \tilde{v}_i \end{pmatrix}^T \begin{pmatrix} \tilde{B}_r^T H \tilde{B}_r + \tilde{B}_i^T H \tilde{B}_i & \tilde{B}_i^T H \tilde{B}_r - \tilde{B}_r^T H \tilde{B}_i \\ \tilde{B}_r^T H \tilde{B}_i - \tilde{B}_i^T H \tilde{B}_r & \tilde{B}_r^T H \tilde{B}_r + \tilde{B}_i^T H \tilde{B}_i \end{pmatrix} \begin{pmatrix} \tilde{v}_r \\ \tilde{v}_i \end{pmatrix} \\ &= \begin{pmatrix} \tilde{v}_r \\ \tilde{v}_i \end{pmatrix}^T \begin{pmatrix} \tilde{B}_r & -\tilde{B}_i \\ -\tilde{B}_i & -\tilde{B}_r \end{pmatrix}^T \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix} \begin{pmatrix} \tilde{B}_r & -\tilde{B}_i \\ -\tilde{B}_i & -\tilde{B}_r \end{pmatrix} \begin{pmatrix} \tilde{v}_r \\ \tilde{v}_i \end{pmatrix} \\ &= v_R^T \hat{M}[B_1]^T \mathcal{H} \hat{M}[B_1] v_R \end{aligned}$$

und somit die Aussage für  $K = 1$ . Wir zeigen die Aussage auch noch für  $K = 2$ .

#### 4. Reduzierte-Basis-Methode

---

Mit der Verwendung der Aussage für  $K = 1$  erhalten wir

$$\begin{aligned} v^H B_2^H B_1^H H B_1 B_2 v &= v^H \tilde{B}_1^H H \tilde{B}_1 v \\ &= v_R^T \hat{M}[\tilde{B}_1]^T \mathcal{H} \hat{M}[\tilde{B}_1] v_R \\ &= v_R^T \hat{M}[B_1 B_2]^T \mathcal{H} \hat{M}[B_1 B_2] v_R \end{aligned}$$

mit  $\tilde{B}_1 = B_1 B_2$ . Hierzu betrachten wir die Definition (4.36) und das komplexe Produkt

$$\begin{aligned} B_1 B_2 &= (\operatorname{Re}(B_1) + i \operatorname{Im}(B_1)) (\operatorname{Re}(B_2) + i \operatorname{Im}(B_2)) \\ &= \operatorname{Re}(B_1) \operatorname{Re}(B_2) - \operatorname{Im}(B_1) \operatorname{Im}(B_2) \\ &\quad + i \operatorname{Im}(B_1) \operatorname{Re}(B_2) + i \operatorname{Re}(B_1) \operatorname{Im}(B_2) \end{aligned}$$

und erhalten somit

$$\begin{aligned} \hat{M}[B_1 B_2] &= \begin{pmatrix} \operatorname{Re}(B_1 B_2) & -\operatorname{Im}(B_1 B_2) \\ -\operatorname{Im}(B_1 B_2) & -\operatorname{Re}(B_1 B_2) \end{pmatrix} \\ &= \begin{pmatrix} \operatorname{Re}(B_1) & -\operatorname{Im}(B_1) \\ -\operatorname{Im}(B_1) & -\operatorname{Re}(B_1) \end{pmatrix} \begin{pmatrix} \operatorname{Re}(B_2) & -\operatorname{Im}(B_2) \\ \operatorname{Im}(B_2) & \operatorname{Re}(B_2) \end{pmatrix} \\ &= \hat{M}[B_1] M[B_2] \end{aligned}$$

und damit auch die Aussage für  $K = 2$ .

IV: Für  $K \in \mathbb{N}$  beliebig aber fest gilt die Äquivalenz des komplexen Produkts

$$v^H \left( \prod_{l=2}^K B_l \right)^H B_1^H H B_1 \left( \prod_{l=2}^K B_l \right) v$$

und des reellen Produkts

$$v_R^T \left( \prod_{l=2}^K M[B_l] \right)^T \hat{M}[B_1]^T \mathcal{H} \hat{M}[B_1] \left( \prod_{l=2}^K M[B_l] \right) v_R.$$

IS: Wir definieren

$$\tilde{B}_k = B_k, \quad 2 \leq k \leq K-1,$$

und

$$\tilde{B}_K = B_K B_{K+1}$$

und erhalten für das komplexe Produkt

$$\begin{aligned}
 & v^H \left( \prod_{l=2}^{K+1} B_l \right)^H B_1^H H B_1 \left( \prod_{l=2}^{K+1} B_l \right) v \\
 &= v^H \left( \prod_{l=2}^K \tilde{B}_l \right)^H B_1^H H B_1 \left( \prod_{l=2}^K \tilde{B}_l \right) v \\
 &\stackrel{IV}{=} v_R^T \left( \prod_{l=2}^K M[\tilde{B}_l] \right)^T \hat{M}[B_1]^T \mathcal{H} \hat{M}[B_1] \left( \prod_{l=2}^K M[\tilde{B}_l] \right) v_R \\
 &= v_R^T M[B_K B_{K+1}]^T \left( \prod_{l=2}^{K-1} M[B_l] \right)^T \hat{M}[B_1]^T \mathcal{H} \hat{M}[B_1] \left( \prod_{l=2}^{K-1} M[B_l] \right) M[B_K B_{K+1}] v_R.
 \end{aligned}$$

Um den Satz nun zu beweisen, müssen wir nur noch

$$M[B_K B_{K+1}] = M[B_K] M[B_{K+1}]$$

zeigen. Hierzu betrachten wir die Definition (4.36) und das komplexe Produkt

$$\begin{aligned}
 B_K B_{K+1} &= (\operatorname{Re}(B_K) + i \operatorname{Im}(B_K)) (\operatorname{Re}(B_{K+1}) + i \operatorname{Im}(B_{K+1})) \\
 &= \operatorname{Re}(B_K) \operatorname{Re}(B_{K+1}) - \operatorname{Im}(B_K) \operatorname{Im}(B_{K+1}) \\
 &\quad + i \operatorname{Im}(B_K) \operatorname{Re}(B_{K+1}) + i \operatorname{Re}(B_K) \operatorname{Im}(B_{K+1})
 \end{aligned}$$

und erhalten somit

$$\begin{aligned}
 M[B_K B_{K+1}] &= \begin{pmatrix} \operatorname{Re}(B_K B_{K+1}) & -\operatorname{Im}(B_K B_{K+1}) \\ \operatorname{Im}(B_K B_{K+1}) & \operatorname{Re}(B_K B_{K+1}) \end{pmatrix} \\
 &= \begin{pmatrix} \operatorname{Re}(B_K) & -\operatorname{Im}(B_K) \\ \operatorname{Im}(B_K) & \operatorname{Re}(B_K) \end{pmatrix} \begin{pmatrix} \operatorname{Re}(B_{K+1}) & -\operatorname{Im}(B_{K+1}) \\ \operatorname{Im}(B_{K+1}) & \operatorname{Re}(B_{K+1}) \end{pmatrix} \\
 &= M[B_K] M[B_{K+1}].
 \end{aligned}$$

□

**Lemma 4.6.** Sei  $A \in \mathbb{C}^{\mathcal{N} \times \mathcal{N}}$  mit

$$\operatorname{Re}(A) = \operatorname{Re}(A)^T, \quad \operatorname{Im}(A) = \operatorname{Im}(A)^T \quad (4.37)$$

und  $H \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$  eine symmetrische positiv definite Matrix, dann gilt

$$\min_{u \in \mathbb{C}^{\mathcal{N}} \setminus \{0\}} \frac{u^H A^H H^{-1} A u}{u^H H u} = |\lambda_{\min}(\mathcal{H}, \hat{M}[A])|^2, \quad (4.38)$$

#### 4. Reduzierte-Basis-Methode

---

wobei  $\lambda_{\min}(\mathcal{H}, \hat{M}[A])$  der betragskleinste Eigenwert zu dem verallgemeinerten Eigenwertproblem

$$\hat{M}[A]v = \lambda_{\min}(\mathcal{H}, \hat{M}[A]) \mathcal{H}v, \quad (4.39)$$

mit  $v \in \mathbb{R}^{2\mathcal{N}}$  ist.

*Beweis.* Mit Satz 4.5 folgt

$$u^H A^H H^{-1} A u = u_R^T \hat{M}[A]^T \mathcal{H}^{-1} \hat{M}[A] u_R$$

mit

$$u_R = \begin{pmatrix} \operatorname{Re}(u) \\ \operatorname{Im}(u) \end{pmatrix}, \quad \mathcal{H} = \begin{pmatrix} H^{-1} & 0 \\ 0 & H^{-1} \end{pmatrix}.$$

Analog folgt damit auch

$$u^H H u = u_R^T \mathcal{H} u_R$$

mit

$$\mathcal{H}^{-1} = \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}$$

und somit

$$\min_{u \in \mathbb{C}^{\mathcal{N}} \setminus \{0\}} \frac{u^H A^H H^{-1} A u}{u^H H u} = \min_{u_R \in \mathbb{R}^{2\mathcal{N}} \setminus \{0\}} \frac{u_R^T \hat{M}[A]^T \mathcal{H}^{-1} \hat{M}[A] u_R}{u_R^T \mathcal{H} u_R}.$$

Da nach Voraussetzung  $H$  und somit auch  $\mathcal{H}$  eine symmetrische positiv definite Matrix ist, existiert die Quadratwurzel der Matrix  $\mathcal{H}^{\frac{1}{2}}$  und wir erhalten

$$\begin{aligned} & \min_{u_R \in \mathbb{R}^{2\mathcal{N}} \setminus \{0\}} \frac{u_R^T \hat{M}[A]^T \mathcal{H}^{-1} \hat{M}[A] u_R}{u_R^T \mathcal{H} u_R} \\ &= \min_{u_R \in \mathbb{R}^{2\mathcal{N}} \setminus \{0\}} \frac{u_R^T \left(\mathcal{H}^{\frac{1}{2}}\right)^T \left(\mathcal{H}^{-\frac{1}{2}}\right)^T \hat{M}[A]^T \mathcal{H}^{-\frac{1}{2}} \mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}} \mathcal{H}^{\frac{1}{2}} u_R}{u_R^T \left(\mathcal{H}^{\frac{1}{2}}\right)^T \mathcal{H}^{\frac{1}{2}} u_R} \\ &= \min_{v_R \in \mathbb{R}^{2\mathcal{N}} \setminus \{0\}} \frac{v_R^T \left(\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}}\right)^T \left(\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}}\right) v_R}{v_R^T v_R} \\ &= \min_{v_R \in \mathbb{R}^{2\mathcal{N}} \setminus \{0\}} \frac{\|\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}} v_R\|_{\mathbb{R}^{2\mathcal{N}}}^2}{\|v_R\|_{\mathbb{R}^{2\mathcal{N}}}^2}. \end{aligned}$$

Aufgrund der Symmetrieeigenschaften (4.37) ist auch die Matrix  $\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}}$  sym-

metrisch, weshalb das Minimierungsproblem gelöst wird durch

$$\min_{v_R \in \mathbb{R}^{2N} \setminus \{0\}} \frac{\|\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}} v_R\|_{\mathbb{R}^{2N}}^2}{\|v_R\|_{\mathbb{R}^{2N}}^2} = |\lambda_{\min}(\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}})|^2$$

und  $\lambda_{\min}(\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}})$  ist der betragsmäßig kleinste Eigenwert bezüglich des Eigenwertproblems

$$\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}} v_R = \lambda_{\min}(\mathcal{H}^{-\frac{1}{2}} \hat{M}[A] \mathcal{H}^{-\frac{1}{2}}) v_R.$$

Dieses Eigenwertproblem ist mit  $v_R = \mathcal{H}^{\frac{1}{2}} v$  äquivalent zum verallgemeinerten Eigenwertproblem (4.39), womit wir das Lemma bewiesen haben.  $\square$

Mit Lemma 4.6 ergibt sich nun

$$\begin{aligned} (\beta_{\mathcal{N}}(p))^2 &= \min_{u^{\mathcal{N}} \in \mathbb{C}^{\mathcal{N}} \setminus \{0\}} \frac{(u^{\mathcal{N}})^H (A_p^{\mathcal{N}})^H H_{X_{\mathcal{N}}}^{-1} A_p^{\mathcal{N}} u^{\mathcal{N}}}{(u^{\mathcal{N}})^H H_{X_{\mathcal{N}}} u^{\mathcal{N}}} \\ &= |\lambda_{\min}(p)|^2, \end{aligned}$$

wobei  $\lambda_{\min}$  der betragsmäßig kleinste Eigenwert zum reellen Eigenwertproblem

$$\begin{pmatrix} \operatorname{Re}(A_p^{\mathcal{N}}) & -\operatorname{Im}(A_p^{\mathcal{N}}) \\ -\operatorname{Im}(A_p^{\mathcal{N}}) & -\operatorname{Re}(A_p^{\mathcal{N}}) \end{pmatrix} \begin{pmatrix} \operatorname{Re}(v^{\mathcal{N}}) \\ \operatorname{Im}(v^{\mathcal{N}}) \end{pmatrix} = \lambda_{\min}(p) \begin{pmatrix} H_{X_{\mathcal{N}}} & 0 \\ 0 & H_{X_{\mathcal{N}}} \end{pmatrix} \begin{pmatrix} \operatorname{Re}(v^{\mathcal{N}}) \\ \operatorname{Im}(v^{\mathcal{N}}) \end{pmatrix}. \quad (4.40)$$

*Bemerkung 4.7.* Anstelle des reellen Eigenwertproblems (4.40) kann approximativ auch das komplexe Eigenwertproblem

$$A_p^{\mathcal{N}} v^{\mathcal{N}} = \tilde{\lambda}_{\min}(p) H_{X_{\mathcal{N}}} v^{\mathcal{N}}, \quad (4.41)$$

betrachtet werden, denn mit dem Eigenvektor  $v^{\mathcal{N}}$  zum betragsmäßig kleinsten Eigenwert  $\tilde{\lambda}_{\min}(p)$  gilt

$$\begin{aligned} (\beta_{\mathcal{N}}(p))^2 &\leq \frac{(v^{\mathcal{N}})^H (A_p^{\mathcal{N}})^H H_{X_{\mathcal{N}}}^{-1} A_p^{\mathcal{N}} v^{\mathcal{N}}}{(v^{\mathcal{N}})^H H_{X_{\mathcal{N}}} v^{\mathcal{N}}} \\ &= \frac{(\tilde{\lambda}_{\min}(p) H_{X_{\mathcal{N}}} v^{\mathcal{N}})^H H_{X_{\mathcal{N}}}^{-1} (\tilde{\lambda}_{\min}(p) H_{X_{\mathcal{N}}} v^{\mathcal{N}})}{(v^{\mathcal{N}})^H H_{X_{\mathcal{N}}} v^{\mathcal{N}}} \\ &= |\tilde{\lambda}_{\min}(p)|^2. \end{aligned}$$

#### 4. Reduzierte-Basis-Methode

---

Damit folgt mit  $\beta_{\mathcal{N}}^{approx}(p) = |\tilde{\lambda}_{\min}(p)|$  für den Fehlerschätzer

$$\Delta_{en}^{approx}(p) = \frac{\|\hat{r}(p)\|_{X_e}}{\beta_{\mathcal{N}}^{approx}(p)} \leq \frac{\|\hat{r}(p)\|_{X_e}}{\beta_{LB}} = \Delta_{en}(p)$$

und wir erhalten einen approximativen Fehlerschätzer  $\Delta_{en}^{approx}(p)$  auf Basis des komplexen Eigenwertproblems (4.41).

Damit haben wir die Möglichkeit die inf-sup-Konstante für einen bestimmten Parameter  $p \in \mathbb{P}$  zu bestimmen. Während des Greedy-Algorithmus wird eine diskrete Menge von Parametern  $\mathbb{P}_{N_{RB}}$  betrachtet und für jedes  $p \in \mathbb{P}_{N_{RB}}$  muss die Konstante  $\beta_{\mathcal{N}}(p)$  berechnet werden. Damit ergeben sich  $N_{RB}$  zu berechnende verallgemeinerte Eigenwertprobleme, die wiederum von  $\mathcal{N}$  abhängen und damit zu sehr hohen Offline-Rechenzeiten führen. Um die Bestimmung der inf-sup-Konstante zur Berechnung des Fehlerschätzers zu beschleunigen, gibt es verschiedene Möglichkeiten. Doch zunächst gehen wir noch auf die Besonderheit des Supremumsoperators (4.16) in Zusammenhang mit der affinen Darstellung (4.20) ein, denn wir erhalten damit

$$\mathcal{T}_{\mathcal{N}}(p)[u_{\mathcal{N}}] = \sum_{k=1}^{Q_a} \Theta_k^a(p) \mathcal{T}_{\mathcal{N}}^k[u_{\mathcal{N}}], \quad (4.42)$$

für die Supremumsoperatoren  $\mathcal{T}_{\mathcal{N}}^k$ , die gerade

$$(\mathcal{T}_{\mathcal{N}}^k[u_{\mathcal{N}}], v_{\mathcal{N}})_{X_e} = a_k(u_{\mathcal{N}}, v_{\mathcal{N}}), \quad v_{\mathcal{N}} \in X^{\mathcal{N}}, \quad (4.43)$$

erfüllen. Diese Darstellung (4.42) wird im Zusammenhang mit der Petrov-Galerkin-Projektion verwendet, um eine affine Darstellung bezüglich parameterunabhängiger Operatoren  $\mathcal{T}_{\mathcal{N}}^k$  zu erhalten. Zudem kann diese auch bei der Berechnung des Fehlers verwendet werden, um dessen Berechnung noch weiter zu vereinfachen. Wir folgen bei der Approximation der inf-sup-Konstante den Erkenntnissen aus [19], wobei wir hier die weit verbreitete Successive-Constraint-Methode (SCM), siehe [14], zur Berechnung von inf-sup-Konstanten außer Acht lassen, da diese für großdimensionale Probleme aufgrund der Lösung großdimensionaler Eigenwertprobleme nicht immer realisierbar ist und in dieser Arbeit auch nicht verwendet wird.

### Galerkin-Schätzer

Beim Galerkin-Schätzer, nachzulesen in [19], schränken wir (4.28) auf den Unterraum  $X_N$  ein, also

$$\beta_N^{GA}(p) = \inf_{u_N \in X_N \setminus \{0\}} \sup_{v_N \in X_N \setminus \{0\}} \frac{|a(u_N, v_N; p)|}{\|u_N\|_{X_e} \|v_N\|_{X_e}}, \quad p \in \mathbb{P}.$$

Bevor wir uns die Darstellung genauer betrachten, beweisen wir noch den folgenden Satz.

**Satz 4.8.** *Sei  $B \in \mathbb{C}^{N \times N}$  hermitesch und  $H \in \mathbb{R}^{N \times N}$  eine hermitesche positiv definite Matrix, dann gilt*

$$\min_{v \in \mathbb{R}^N \setminus \{0\}} \frac{v^T B v}{v^T H v} = \lambda_{\min}$$

und  $\lambda_{\min}$  ist der kleinste Eigenwert zum verallgemeinerten Eigenwertproblem

$$Bv = \lambda_{\min} H v. \quad (4.44)$$

*Beweis.* Da  $H$  hermitesch positiv definit ist, existiert die Cholesky-Zerlegung  $H = L^T L$  und wir erhalten

$$\min_{v \in \mathbb{R}^N \setminus \{0\}} \frac{v^T B v}{v^T H v} = \min_{v \in \mathbb{R}^N \setminus \{0\}} \frac{v^T L^T L^{-T} B L^{-1} L v}{v^T L^T L v} = \min_{w \in \mathbb{R}^N \setminus \{0\}} \frac{w^T L^{-T} B L^{-1} w}{w^T w}.$$

Wegen der hermiteschen Eigenschaft von  $B$  ist auch  $L^{-T} B L^{-1}$  hermitesch. Analog zum Satz von Courant-Fischer [51, Satz 15.3] gilt nun

$$\min_{w \in \mathbb{R}^N \setminus \{0\}} \frac{w^T L^{-T} B L^{-1} w}{w^T w} = \lambda_{\min}$$

und  $\lambda_{\min}$  ist der kleinste Eigenwert zum Eigenwertproblem

$$L^{-T} B L^{-1} w = \lambda_{\min} w,$$

welches wiederum äquivalent ist zum verallgemeinerten Eigenwertproblem (4.44).  $\square$

Wir definieren den reduzierten Supremumsoperator  $\mathcal{T}_N(p) : X_N \rightarrow X_N$ , der wiederum durch die Eigenschaft

$$(\mathcal{T}_N(p)[u_N], v_N)_{X_e} = a(u_N, v_N; p), \quad v_N \in X_N,$$

#### 4. Reduzierte-Basis-Methode

---

gegeben ist. Zu dem Operator  $\mathcal{T}_N(p)$  definieren wir den zugehörigen Operator

$$\mathcal{T}_p^N : \mathbb{C}^N \rightarrow \mathbb{C}^N,$$

der den RB-Koeffizientenvektor  $u^N$  der Funktion  $u_N$  auf den RB-Koeffizientenvektor der Funktion  $\mathcal{T}_N(p)[u_N]$  abbildet. Es gilt somit die FEM-Darstellung

$$\mathcal{T}_N(p)[u_N] = \sum_{k=1}^N (\Xi \mathcal{T}_p^N [u^N])_k \varphi_k.$$

Wie zuvor können wir aus den Eigenschaften des Operators  $\mathcal{T}_N(p)$  die Darstellung

$$\mathcal{T}_p^N = (\Xi^H H_{X_N} \Xi)^{-1} \Xi^H A_p^N \Xi$$

herleiten und erhalten für die inf-sup-Konstante mit  $\tilde{A}_p^N = \Xi^H A_p^N \Xi$  gerade

$$(\beta_N^{GA}(p))^2 = \min_{u^N \in \mathbb{C}^N \setminus \{0\}} \frac{(u^N)^H (\tilde{A}_p^N)^H (\Xi^H H_{X_N} \Xi)^{-H} \Xi^H H_{X_N} \Xi (\Xi^H H_{X_N} \Xi)^{-1} \tilde{A}_p^N u^N}{(u^N)^H \Xi^H H_{X_N} \Xi u^N}.$$

Mit Satz 4.5 ist dieses Minimierungsproblem äquivalent zum reellen Problem

$$(\beta_N^{GA}(p))^2 = \min_{v^N \in \mathbb{R}^{2N} \setminus \{0\}} \frac{(v^N)^T \tilde{M}^T \hat{M}[\Xi]^T \mathcal{H} \hat{M}[\Xi] \tilde{M} v^N}{(v^N)^T \hat{M}[\Xi]^T \mathcal{H} \hat{M}[\Xi] v^N}$$

mit

$$\tilde{M} = M [(\Xi^H H_{X_N} \Xi)^{-1} \Xi^H A_p^N \Xi], \quad \mathcal{H} = \begin{pmatrix} H_{X_N} & 0 \\ 0 & H_{X_N} \end{pmatrix}.$$

Da die Matrix  $H$  symmetrisch positiv definit ist, ist  $\hat{M}[\Xi]^T \mathcal{H} \hat{M}[\Xi]$  eine hermitesche positiv definite Matrix. Demnach folgt mit Satz 4.8 gerade

$$(\beta_N^{GA}(p))^2 = \lambda_{\min}^{GA}(p)$$

mit  $\lambda_{\min}^{GA}(p)$  kleinster Eigenwert zu dem verallgemeinerten Eigenwertproblem

$$\tilde{M}^T \hat{M}[\Xi]^T \mathcal{H} \hat{M}[\Xi] \tilde{M} v^N = \lambda_{\min}^{GA}(p) \hat{M}[\Xi]^T \mathcal{H} \hat{M}[\Xi] v^N.$$

Es gibt für die inf-sup-Konstante  $\beta_N^{GA}(p)$  keinen Beweis, dass diese eine obere oder untere Schranke an (4.15) darstellt.

*Bemerkung 4.9.* Wie wir bereits wissen, können wir anstatt der inf-sup-Konstante (4.39) auch das komplexe Eigenwertproblem (4.41) lösen und erhalten einen approximativen

Fehlerschätzer. Analog dazu folgt

$$\beta_N^{GA}(p) \leq |\tilde{\lambda}_{\min}^{GA}(p)| = \beta_N^{\text{approx},GA}(p),$$

wobei  $\tilde{\lambda}_{\min}^{GA}(p)$  der betragsmäßig kleinste Eigenwert zu dem verallgemeinerten komplexen Eigenwertproblem

$$\Xi^H A_p^N \Xi u^N = \tilde{\lambda}_{\min}^{GA}(p) \Xi^H H_{X_N} \Xi u^N, \quad (4.45)$$

ist. Eine Verbesserung der approximativen Eigenschaft bezüglich  $\tilde{\lambda}_{\min}(p)$  aus Bemerkung 4.7 wird laut [19] dadurch erreicht, dass anstatt (4.45) das verallgemeinerte Eigenwertproblem

$$X^H A_p^N X u^N = \hat{\lambda}_{\min}^{GA}(p) X^H H_{X_N} X u^N,$$

gelöst wird, wobei  $X = [\Xi, v[\Xi]]$  und

$$v[\Xi] = (v[\xi_1], v[\xi_2], \dots, v[\xi_N]).$$

Dabei ist  $v[\xi_m]$  der Eigenvektor zu dem Eigenwertproblem (4.41) für den zur Bestimmung von  $\xi_m$  verwendeten Parameter  $p_m \in \mathbb{P}$ . Dabei wird davon ausgegangen, dass durch Hinzunahme der Eigenvektoren  $v[\xi_m]$  das Eigenwertproblem (4.41) besser approximiert wird.

### MinRes-Schätzer

Beim MinRes-Schätzer betrachten wir die inf-sup-Konstante  $\beta_N^{PG}(p)$  aus Lemma 4.1 für Petrov-Galerkin-Projektionen

$$\beta_N^{MR}(p) = \inf_{u_N \in X_N \setminus \{0\}} \frac{\|\mathcal{T}_N(p)[u_N]\|_{X_e}}{\|u_N\|_{X_e}} = \beta_N^{PG}(p),$$

für die nach Lemma 4.1

$$\beta_N^{MR}(p) \geq \beta_N(p)$$

gilt und damit eine obere Schranke an  $\beta_N(p)$  darstellt. Wird für die Reduktion des FEM-Problems (4.3) eine Petrov-Galerkin-Projektion (4.11) verwendet, dann stimmt die zu (4.11) zugehörige inf-sup-Konstante  $\beta_N^{PG}(p)$  mit  $\beta_N^{MR}(p)$  überein. Mit der Definition des Supremumsoperators  $\mathcal{T}_p^N = (H_{X_N})^{-1} A_p$  aus Abschnitt 4.1.6 und der Basisdarstellung

$$u_N(p) = \sum_{l=1}^N u_l^N \xi_l = \sum_{l=1}^N u_l^N \sum_{k=1}^N \xi_{k,l}^N \varphi_k = \sum_{k=1}^N (\Xi u^N)_k \varphi_k,$$

ergibt sich

$$\begin{aligned} (\beta_N^{MR}(p))^2 &= \inf_{u_N \in X_N \setminus \{0\}} \frac{(\mathcal{T}_N(p)[u_N], \mathcal{T}_N(p)[u_N])_{X_e}}{(u_N, u_N)_{X_e}} \\ &= \min_{u^N \in \mathbb{C}^N \setminus \{0\}} \frac{(u^N)^H \Xi^H (A_p^N)^H (H_{X_N})^{-1} A_p^N \Xi u^N}{(u^N)^H \Xi^H H_{X_N} \Xi u^N}. \end{aligned}$$

Nach den Sätzen 4.5 und 4.8 erhalten wir

$$(\beta_N^{MR}(p))^2 = \lambda_{\min}^{MR}(p),$$

wobei  $\lambda_{\min}^{MR}(p)$  der kleinste Eigenwert zu dem verallgemeinerten reellen Eigenwertproblem

$$M [\Xi]^T \hat{M} [A_p^N]^T \mathcal{H}^{-1} \hat{M} [A_p^N] M [\Xi] v^N = \lambda_{\min}^{MR}(p) \hat{M} [\Xi]^T \mathcal{H} \hat{M} [\Xi] v^N$$

ist.

*Bemerkung 4.10.* Anstelle des reellen Eigenwertproblems, können wir wegen Satz 4.8 auch

$$\beta_N^{\text{approx}, MR}(p) = \sqrt{\tilde{\lambda}_{\min}^{MR}(p)} = \beta_N^{MR}(p)$$

betrachten, wobei  $\tilde{\lambda}_{\min}^{MR}(p)$  der kleinste Eigenwert zum verallgemeinerten komplexen Eigenwertproblem

$$\Xi^H (A_p^N)^H (H_{X_N})^{-1} A_p^N \Xi u^N = \tilde{\lambda}_{\min}^{MR}(p) \Xi^H H_{X_N} \Xi u^N \quad (4.46)$$

ist. Wie für den Galerkin-Schätzer können wir auch eine Verbesserung der approximativen Eigenschaft bezüglich  $\tilde{\lambda}_{\min}(p)$  aus Bemerkung 4.7 erreichen, indem wir anstatt (4.46) das verallgemeinerte Eigenwertproblem

$$X^H (A_p^N)^H (H_{X_N})^{-1} A_p^N X u^N = \hat{\lambda}_{\min}^{MR}(p) X^H H_{X_N} X u^N,$$

betrachten, wobei wieder  $X = [\Xi, v[\Xi]]$  und  $v[\xi_m]$  die Eigenvektoren zu den Eigenwertproblemen (4.39) für die zur Bestimmung von  $\xi_m$  verwendeten Parameter  $p_m \in \mathbb{P}$  sind.

*Bemerkung 4.11.* Die Eigenvektorenerweiterung  $X = [\Xi, v[\Xi]]$  bei der Approximation der inf-sup-Konstante ist nur sinnvoll bei Problemen deren Dimension der FEM-Diskretisierung nicht zu groß sind, denn für jeden RB-Basisvektor  $\xi_m$  und entsprechenden Parameter  $p_m \in \mathbb{P}$  muss das großdimensionale und damit rechenintensive Eigenwertproblem (4.39) gelöst werden um den Eigenvektor  $v[\xi_m]$  zu bestimmen. Dies kann je nach Dimension der Diskretisierung zu deutlich längeren Rechenzeiten während der Offline-Phase führen, weswegen für Probleme mit Diskretisierungsdimension  $\mathcal{N} > 10^6$  die Eigenvektorenerweiterung nicht mehr verwendet werden sollte.

## 4.2. RB-Lagrange-Systeme

In diesem Abschnitt möchten wir die in Abschnitt 3.4 hergeleiteten Lagrange-Systeme in ihrer Allgemeinheit betrachten, das heißt unabhängig von der Darstellung der hier betrachteten Probleme aus Kapitel 3. Wir betrachten dazu das parameterabhängige allgemeine Problem:

Bestimme  $u(p) \in X_e$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u(p), \varphi; p) = f(\varphi; p), \quad \varphi \in Y_e, \quad (4.47)$$

mit  $a : X_e \times Y_e \times \mathbb{P} \rightarrow \mathbb{C}$  und  $f : Y_e \times \mathbb{P} \rightarrow \mathbb{C}$  gilt.

Hierbei seien

$$X_e = \{u \in V_e : c[u] = 0, c[u] \in W_e\} \subset V_e,$$

$$Y_e = \{u \in Z_e : c_f[u] = 0, c_f[u] \in U_e\} \subset Z_e,$$

mit  $U_e, V_e, W_e, Z_e$  Hilberträume, und  $c : V_e \rightarrow W_e$  und  $c_f : Z_e \rightarrow U_e$  affin lineare Operatoren. Das zu dem Problem (4.47) zugehörige Lagrange-Problem lautet:

Bestimme  $u(p) \in V_e$  und  $\lambda(p) \in (U_e)'$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u(p), \varphi; p) + d_f(\lambda(p), \varphi; p) = f(\varphi; p), \quad \varphi \in Z_e, \quad (4.48)$$

$$d(u(p), \phi; p) = g(\phi; p), \quad \phi \in (W_e)', \quad (4.49)$$

mit  $a : V_e \times Z_e \times \mathbb{P} \rightarrow \mathbb{C}$ ,  $d_f : (U_e)' \times Z_e \times \mathbb{P} \rightarrow \mathbb{C}$ ,  $d : V_e \times (W_e)' \times \mathbb{P} \rightarrow \mathbb{C}$  und  $f : Y_e \times \mathbb{P} \rightarrow \mathbb{C}$ ,  $g : (W_e)' \times \mathbb{P} \rightarrow \mathbb{C}$  gilt.

Dabei werden die Sesquilinearformen  $d$  und  $d_f$  aus den Dualitätsprodukten von  $\phi$  mit  $c[u]$  und  $\lambda(p)$  mit  $c_f[\varphi]$  in den Hilberträumen  $U_e, W_e$  abgeleitet. Für eine Lösung  $u(p)$  des Systems (4.48), (4.49) gilt  $u(p) \in X_e$  und  $u(p)$  ist Lösung von (4.47). Bei der Betrachtung vieler variationeller Probleme gilt, dass der Testraum  $Y_e$  im selben Funktionenraum  $V_e$ , also  $Z_e = V_e$ , wie der Ansatzraum  $X_e$  liegt und wir deshalb im weiteren Verlauf dieses Abschnittes die Einschränkung

$$Y_e = \{u \in V_e : c_f[u] = 0, c_f[u] \in U_e\} \subset V_e$$

verwenden. Wir betrachten demnach das veränderte Lagrange-System:

Bestimme  $u(p) \in V_e$  und  $\lambda(p) \in (U_e)'$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u(p), \varphi; p) + d_f(\lambda(p), \varphi; p) = f(\varphi; p), \quad \varphi \in V_e, \quad (4.50)$$

$$d(u(p), \phi; p) = g(\phi; p), \quad \phi \in (W_e)'. \quad (4.51)$$

*Bemerkung 4.12.* Die Theorie der RB-Lagrange-Systeme (4.50) und (4.51) in diesem Abschnitt kann mit Anpassungen an die Theorie und erweitertem Notationsaufwand auf die Systeme (4.48) und (4.49) erweitert werden.

*Bemerkung 4.13.* Die Verwendung der Dualräume in den Gleichungen (4.50) und (4.51) ergibt sich aus der Betrachtung der variationellen Form partieller Differentialgleichungen, bei deren Herleitung die PDE in einer schwachen Formulierung hinsichtlich des Dualraums betrachtet wird.

*Bemerkung 4.14.* Ein Beispiel für die Gleichungen (4.50) und (4.51) stellt das System (3.72) dar, bei dem die Operatoren  $c$  und  $c_f$  gerade dem Randoperator  $\gamma_t$  eingeschränkt auf den Rand  $\Gamma_{PEC}$  entsprechen.

Wegen der deutlich komplexeren Notation für die mögliche Erweiterung von  $Y_e$  in Bemerkung 4.12 hinsichtlich der diskretisierten Räume, betrachten wir in dieser Arbeit die Theorie der RB-Lagrange-Systeme (4.50) und (4.51) bezüglich des Spezialfalls  $Y_e \subset V_e$ . Mit einer entsprechenden Anpassung der Funktionenräume folgt die Erweiterung in Bemerkung 4.12 vollkommen analog zu der hier vorgestellten Theorie. Aufgrund der Theorie der RBM für Lagrange-Probleme geben wir hier noch eine Bemerkung in Bezug auf die Linearität des Lösungsraumes an. Beispielsweise erhalten wir bei nicht homogenen Dirichlet-Randwerten affin lineare Lösungsräume  $X_e$ . Werden zum Beispiel Wärmeleitungsprobleme betrachtet, dann führt die konstante Temperatur an einer Oberfläche zu einem affin linearen Lösungsraum.

*Bemerkung 4.15.* Der Lösungsraum  $M_e \subset X_e$  kann affin linear sein, das heißt für Funktionen  $v, w \in M_e$  und  $x_1, x_2 \in \mathbb{C}$  folgt dann  $u = x_1v + x_2w \notin M_e$ .

Dies hat zur Folge, dass bei einfacher Anwendung der RBM die Basisdarstellung wie in Abschnitt 4.1 nicht korrekt ist und die Methode angepasst werden muss. Das Lagrange-System entsteht dabei aus einer Galerkin-Projektion einer PDE unter Nebenbedingungen und wird so im Simulationstool COMSOL Multiphysics für jegliche PDEs verwendet, weswegen wir die zu diesem System zugehörige Theorie der RBM allgemein darstellen wollen. Wir befassen uns in Abschnitt 4.2.1 mit der Darstellung und Herleitung des RB-Systems für Lagrange-Systeme und geben in Abschnitt 4.2.2 und Abschnitt 4.2.3 die entsprechenden Berechnungen an, welche für eine Fehlertheorie wie in Abschnitt 4.1.4 benötigt werden.

#### 4.2.1. RB-Reduktion der Lagrange-Systeme

Wie in Abschnitt 4.1.1 betrachten wir das zu (4.47) diskretisierte Problem:

Bestimme  $u_{\mathcal{N}}(p) \in X_{\mathcal{N}}$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u_{\mathcal{N}}(p), \varphi; p) = f(\varphi; p), \quad \varphi \in Y_{\mathcal{N}}. \quad (4.52)$$

Analog zu (4.50), (4.51) erhalten wir auch das zu (4.52) parameterabhängige diskretisierte Lagrange-System:

Bestimme  $u_{\mathcal{N}}(p) \in V_{\mathcal{N}}$  und  $\lambda_{\mathcal{N}_{d,f}}(p) \in (U_{\mathcal{N}_{d,f}})'$  zum Parameter  $p \in \mathbb{P}$ , sodass

$$a(u_{\mathcal{N}}(p), \varphi; p) + d_f(\lambda_{\mathcal{N}_{d,f}}(p), \varphi; p) = f(\varphi; p), \quad \varphi \in V_{\mathcal{N}}, \quad (4.53)$$

$$d(u_{\mathcal{N}}(p), \phi; p) = g(\phi; p), \quad \phi \in (W_{\mathcal{N}_d})'. \quad (4.54)$$

Das zu dem Lagrange-Problem (4.53), (4.54) äquivalente Gleichungssystem hat die Darstellung

$$\begin{pmatrix} A_p^{\mathcal{N}} & (D_f)_p \\ D_p & 0 \end{pmatrix} \begin{pmatrix} u_{\mathcal{N}} \\ \lambda_{\mathcal{N}_{d,f}} \end{pmatrix} = \begin{pmatrix} f_p^{\mathcal{N}} \\ g_p \end{pmatrix} \quad (4.55)$$

mit

$$\begin{aligned} A_p^{\mathcal{N}} &\in \mathbb{C}^{\mathcal{N} \times \mathcal{N}}, & (D_f)_p &\in \mathbb{C}^{\mathcal{N} \times \mathcal{N}_{d,f}}, & D_p &\in \mathbb{C}^{\mathcal{N}_d \times \mathcal{N}}, \\ f_p^{\mathcal{N}} &\in \mathbb{C}^{\mathcal{N}}, & g_p &\in \mathbb{C}^{\mathcal{N}_d}. \end{aligned}$$

Um wie in Abschnitt 4.1.2 eine effiziente RBM anwenden zu können, benötigen wir eine affine Darstellung des Gleichungssystems.

**Voraussetzung 4.16.** Die Funktionale  $a, d_f, d, f$  und  $g$  besitzen die affinen Darstellungen

$$\begin{aligned} a(u, v; p) &= \sum_{k=1}^{Q_a} \Theta_k^a(p) a_k(u, v), & d_f(u, v; p) &= \sum_{k=1}^{Q_{d_f}} \Theta_k^{d_f}(p) d_{f,k}(u, v), \\ d(u, v; p) &= \sum_{k=1}^{Q_d} \Theta_k^d(p) d_k(u, v), & f(v; p) &= \sum_{k=1}^{Q_f} \Theta_k^f(p) f_k(v), \\ g(v; p) &= \sum_{k=1}^{Q_g} \Theta_k^g(p) g_k(v). \end{aligned}$$

Aus Voraussetzung 4.16 folgt damit auch die Existenz einer affinen Darstellung der entsprechenden Matrizen  $A_p^{\mathcal{N}}$ ,  $(D_f)_p$ ,  $D_p$  und Vektoren  $f_p^{\mathcal{N}}$ ,  $g_p$ . Um die RBM anwenden zu können, arbeiten wir in einem linearen Unterraum der Lösungsmannigfaltigkeit

$$\mathcal{M}_{\mathcal{N},u} = \left\{ u_{\mathcal{N}} \in V_{\mathcal{N}} : \exists \lambda_{\mathcal{N}_{d,f}} \in (U_{\mathcal{N}_{d,f}})' \text{ mit } (u_{\mathcal{N}}(p), \lambda_{\mathcal{N}_{d,f}}(p)) \text{ löst} \right. \\ \left. (4.53), (4.54) \text{ für } p \in \mathbb{P} \right\}.$$

Hierfür benötigen wir noch zusätzliche Voraussetzungen an das System. Dabei befolgen

#### 4. Reduzierte-Basis-Methode

---

wir die in Abschnitt 3.4 angegebene Vorgehensweise zur Berechnung einer Lösung des Lagrange-Systems (4.53), (4.54), indem wir die Zerlegung von  $V_{\mathcal{N}}$  bezüglich des Kernes und des orthogonalen Komplements der Operatoren

$$\begin{aligned}\mathcal{D}(p) : V_{\mathcal{N}} &\rightarrow (W_{\mathcal{N}_d})' ', \\ u_{\mathcal{N}} &\mapsto d(u_{\mathcal{N}}, \cdot; p),\end{aligned}$$

und

$$\begin{aligned}\mathcal{D}_f(p) : V_{\mathcal{N}} &\rightarrow (U_{\mathcal{N}_{d,f}})' ', \\ \varphi &\mapsto d_f(\cdot, \varphi; p),\end{aligned}$$

betrachten. Dabei gilt nun mit

$$\begin{aligned}\dim(V_{\mathcal{N}}) &= \mathcal{N}, \\ \dim(\text{Kern}(\mathcal{D}(p))) &= \mathcal{N}_c, \\ \dim(\text{Kern}(\mathcal{D}(p))^\perp) &= \mathcal{N}_r\end{aligned}$$

die Dimensionsgleichung

$$\mathcal{N} = \mathcal{N}_c + \mathcal{N}_r,$$

im Falle der Zerlegung

$$V_{\mathcal{N}} = \text{Kern}(\mathcal{D}(p)) \oplus \text{Kern}(\mathcal{D}(p))^\perp,$$

und mit

$$\begin{aligned}\dim(\text{Kern}(\mathcal{D}_f(p))) &= \mathcal{N}_{d,c}, \\ \dim(\text{Kern}(\mathcal{D}_f(p))^\perp) &= \mathcal{N}_{d,r}\end{aligned}$$

die Dimensionsgleichung

$$\mathcal{N} = \mathcal{N}_{d,c} + \mathcal{N}_{d,r},$$

für die Zerlegung

$$V_{\mathcal{N}} = \text{Kern}(\mathcal{D}_f(p)) \oplus \text{Kern}(\mathcal{D}_f(p))^\perp.$$

Mit der Vorgehensweise in Abschnitt 3.4 erhalten wir damit das zu (4.55) äquivalente Gleichungssystem

$$\begin{aligned}(B_{D_f})_p^H A_p^{\mathcal{N}} (B_D)_p u_c^{\mathcal{N}} &= (B_{D_f})_p^H f_p^{\mathcal{N}} - (B_{D_f})_p^H A_p^{\mathcal{N}} u_r^{\mathcal{N}}, \\ D_p u_r^{\mathcal{N}} &= g_p,\end{aligned}$$

wobei  $(B_D)_p \in \mathbb{C}^{\mathcal{N} \times \mathcal{N}_c}$  und  $(B_{D_f})_p \in \mathbb{C}^{\mathcal{N} \times \mathcal{N}_{d,c}}$  die rechtsseitigen Kernbasismatrizen der

Matrizen  $D_p$  und  $(D_f)_p^H$  sind. Der Koeffizientenvektor  $u^{\mathcal{N}}$  der Lösung  $u_{\mathcal{N}}(p)$  zu (4.53), (4.54) ist dabei gegeben durch

$$u^{\mathcal{N}} = u_c^{\mathcal{N}} + u_r^{\mathcal{N}} = (B_D)_p u_c^{\mathcal{N}_c} + u_r^{\mathcal{N}}.$$

Da die beiden Operatoren  $\mathcal{D}(p)$  und  $\mathcal{D}_f(p)$  parameterabhängig sind und damit sowohl der Kern als auch das orthogonale Komplement von den Parametern abhängig sind, benötigen wir für den noch zu kreierenden parameterunabhängigen linearen Unterraum folgende Voraussetzung.

**Voraussetzung 4.17.** Die Operatoren  $d$  und  $d_f$  sind unabhängig vom Parameter  $p \in \mathbb{P}$ , womit auch  $\mathcal{D}$  und  $\mathcal{D}_f$  parameterunabhängig sind. Dies ist äquivalent dazu, dass die Kernbasismatrizen  $B_D$  und  $B_{D_f}$  der Matrizen  $D$  und  $D_f^H$  parameterunabhängig sind.

**Voraussetzung 4.18.** Es gilt  $U_e = W_e$  und

$$d(u, v) = \overline{d_f(v, u)}.$$

Falls die Unterräume  $\text{Kern}(\mathcal{D}(p))$  und  $\text{Kern}(\mathcal{D}_f(p))$  abhängig vom Parameter  $p \in \mathbb{P}$  sind und damit Voraussetzung 4.17 nicht gilt, dann muss für die Erhaltung der affinen Darstellung der Gleichungen eine affine Darstellung der Basismatrizen  $(B_D)_p$  und  $(B_{D_f})_p$  bestimmt werden. Die Herleitung einer solchen Darstellung kann nicht allgemein angegeben werden und wird deshalb in dieser Arbeit nicht weiter betrachtet. Eine direkte Folgerung aus Voraussetzung 4.17 sind die parameterunabhängigen Zerlegungen

$$V_{\mathcal{N}} = \text{Kern}(\mathcal{D}) \oplus \text{Kern}(\mathcal{D})^{\perp} = \text{Kern}(\mathcal{D}_f) \oplus \text{Kern}(\mathcal{D}_f)^{\perp}, \quad (4.56)$$

sowohl im Test- als auch im Ansatzraum. Anstatt eine Basis in der Untermannigfaltigkeit  $\mathcal{M}_{\mathcal{N},u} \subset X_{\mathcal{N}} \subset V_{\mathcal{N}}$  zu generieren, betrachten wir die durch die Zerlegung von  $V_{\mathcal{N}}$  einhergehende Zerlegung des Lösungsraums

$$\mathcal{M}_{\mathcal{N},u} = \mathcal{M}_{\mathcal{N},u_c} \oplus \mathcal{M}_{\mathcal{N},u_r},$$

wobei  $\mathcal{M}_{\mathcal{N},u_c} \subset \text{Kern}(\mathcal{D})$  und  $\mathcal{M}_{\mathcal{N},u_r} \subset \text{Kern}(\mathcal{D})^{\perp}$ .

**Lemma 4.19.** Gelten die Voraussetzungen 4.16 und 4.17, dann besitzt  $\mathcal{M}_{\mathcal{N},u_r}$  eine affine Zerlegung, d.h.  $u_{\mathcal{N},r}(p) \in \mathcal{M}_{\mathcal{N},u_r}$  besitzt die Darstellung

$$u_{\mathcal{N},r}(p) = \sum_{k=1}^{Q_g} \Theta_k^g(p) u_{\mathcal{N},r,k},$$

wobei  $u_{\mathcal{N},r,k}$  die parameterunabhängige Lösung zu

$$d(u_{\mathcal{N},r,k}, \phi) = g_k(\phi), \quad \phi \in (W_{N_d})', \quad (4.57)$$

#### 4. Reduzierte-Basis-Methode

---

ist. Die zu  $u_{\mathcal{N},r,k}$  gehörige FEM-Darstellung ist gegeben durch

$$u_{\mathcal{N},r,k} = \sum_{l=1}^{\mathcal{N}} (u_{r,k}^{\mathcal{N}})_l \varphi_l.$$

Gilt zudem Voraussetzung 4.18, dann folgt

$$\text{Kern}(\mathcal{D}) = \text{Kern}(\mathcal{D}_f) \text{ und } \text{Kern}(\mathcal{D})^\perp = \text{Kern}(\mathcal{D}_f)^\perp,$$

was wiederum äquivalent ist zu

$$B_D = B_{D_f}.$$

*Beweis.* Die erste Aussage des Lemmas ist eine direkte Folgerung aus der Linearität der Sesquilinearform  $d(\cdot, \cdot)$ , der Zerlegung (4.56) des Raumes  $V_{\mathcal{N}}$  bezüglich des Operators  $\mathcal{D}$  und den genannten Voraussetzungen 4.16 und 4.17. Die zweite Aussage folgt direkt aus Voraussetzung 4.18.  $\square$

Wir definieren

$$A_{p,c} = \sum_{k=1}^{Q_a} \Theta_k^a(p) B_{D_f}^H A_k^{\mathcal{N}} B_D,$$

sowie

$$f_{p,c} = \sum_{k=1}^{Q_f} \Theta_k^f(p) B_{D_f}^H f_k^{\mathcal{N}} - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_g} \Theta_k^a(p) \Theta_l^g(p) B_{D_f}^H A_k^{\mathcal{N}} u_{r,l}^{\mathcal{N}},$$

wobei

$$\begin{aligned} (A_k^{\mathcal{N}})_{m,n} &= a_k(\varphi_n, \varphi_m) \quad \text{mit } m, n = 1, \dots, \mathcal{N}, \\ (f_k^{\mathcal{N}})_m &= f_k(\varphi_m) \quad \text{mit } m = 1, \dots, \mathcal{N} \end{aligned}$$

und  $\varphi_n$  die FEM-Basisfunktionen darstellt. Mit den Voraussetzungen 4.16 und 4.17, sowie Lemma 4.19, erhalten wir die zu (4.53), (4.54) äquivalente Gleichung

$$A_{p,c} u_c^{\mathcal{N}_c} = f_{p,c} \tag{4.58}$$

mit

$$A_{p,c} \in \mathbb{C}^{\mathcal{N}_{d,c} \times \mathcal{N}_c}, \quad f_{p,c} \in \mathbb{C}^{\mathcal{N}_{d,c}},$$

wobei der zur Lösung  $u_{\mathcal{N}}(p)$  zugehörige Koeffizientenvektor  $u^{\mathcal{N}}$  die folgende Darstellung besitzt

$$u^{\mathcal{N}}(p) = B_D u_c^{\mathcal{N}_c} + \sum_{k=1}^{Q_g} \Theta_k^g(p) u_{r,k}^{\mathcal{N}}.$$

Die zu (4.58) zugehörige Problembeschreibung ist dabei gegeben durch:

Sei  $p \in \mathbb{P}$  gegeben. Bestimme  $u_{\mathcal{N},c}(p) \in \text{Kern}(\mathcal{D})$ , sodass

$$a(u_{\mathcal{N},c}(p), \varphi; p) = f(\varphi; p) - \sum_{k=1}^{Q_g} \Theta_k^g(p) a(u_{\mathcal{N},r,k}, \varphi; p), \quad \varphi \in \text{Kern}(\mathcal{D}_f), \quad (4.59)$$

wobei  $u_{\mathcal{N},r,k}$  die Gleichung (4.57) löst.

Wegen der FEM-Darstellung in  $V_{\mathcal{N}}$  und der Kernbasismatrix  $B_D$  besitzt jede Kernlösung  $u_{\mathcal{N},c}(p) \in \text{Kern}(\mathcal{D})$  die FEM-Darstellung

$$u_{\mathcal{N},c}(p) = \sum_{l=1}^{\mathcal{N}} (u_c^{\mathcal{N}})_l \varphi_l = \sum_{l=1}^{\mathcal{N}} (B_D u_c^{\mathcal{N},c})_l \varphi_l, \quad (4.60)$$

wobei  $\varphi_l$  die Basisfunktionen der FEM darstellt. Aufgrund der Äquivalenz der Probleme (4.47) und (4.59) erhalten wir die Existenz einer Lösung und definieren die entsprechende inf-sup-Konstante über die Kerne

$$\beta_{\mathcal{N},c}(p) = \inf_{u_{\mathcal{N},c} \in \text{Kern}(\mathcal{D}) \setminus \{0\}} \sup_{v_{\mathcal{N},c} \in \text{Kern}(\mathcal{D}_f) \setminus \{0\}} \frac{|a(u_{\mathcal{N},c}, v_{\mathcal{N},c}; p)|}{\|u_{\mathcal{N},c}\|_{V_e} \|v_{\mathcal{N},c}\|_{V_e}}, \quad p \in \mathbb{P}. \quad (4.61)$$

Wir definieren nun den linearen Unterraum  $X_{\mathcal{N},c} \subset \mathcal{M}_{\mathcal{N},u_c}$  durch

$$X_{\mathcal{N},c} = \text{span} \{(\xi_1)_{\mathcal{N},c}, \dots, (\xi_N)_{\mathcal{N},c}\}$$

mit  $(\xi_n)_{\mathcal{N},c} \in \mathcal{M}_{\mathcal{N},u_c}$  für  $1 \leq n \leq N$  und daraus den Unterraum

$$X_N = X_{\mathcal{N},c} \oplus \mathcal{M}_{\mathcal{N},u_r} \subset \mathcal{M}_{\mathcal{N},u} \subset X_e.$$

Da die Funktionen  $(\xi_n)_{\mathcal{N},c}$  für  $1 \leq n \leq N$  durch Lösungen von (4.59) definiert sind, besitzen sie die FEM-Darstellungen

$$(\xi_n)_{\mathcal{N},c} = \sum_{l=1}^{\mathcal{N}} (B_D (\xi_n)_c^{\mathcal{N}})_l \varphi_l.$$

Damit einhergehend besitzt  $u_N \in X_N$  die Darstellung

$$\begin{aligned} u_N(p) &= u_{\mathcal{N},c}(p) + u_{\mathcal{N},r}(p) = \sum_{n=1}^N (u_c^{\mathcal{N}})_n (\xi_n)_{\mathcal{N},c} + u_{\mathcal{N},r}(p) \\ &= \sum_{n=1}^N (u_c^{\mathcal{N}})_n (\xi_n)_{\mathcal{N},c} + \sum_{l=1}^{Q_g} \Theta_l^g(p) u_{\mathcal{N},r,l} \end{aligned} \quad (4.62)$$

#### 4. Reduzierte-Basis-Methode

---

und daraus ergibt sich die FEM-Darstellung

$$u_N(p) = \sum_{k=1}^N \left( B_D \Xi_c u_c^N + \sum_{l=1}^{Q_g} \Theta_l^g(p) u_{r,l}^N \right)_k \varphi_k, \quad (4.63)$$

mit  $(\Xi_c)_{m,n} = ((\xi_n)_c^N)_m$ . Zudem definieren wir  $(\Xi)_{m,n} = ((\xi_n)_c^N)_m = (B_D(\xi_n)_c^N)_m$  als FEM-Koeffizientenmatrix der Lösungen im Kern von  $\mathcal{D}$ . Analog zu Abschnitt 4.1.1 können wir die stabile Petrov-Galerkin-Projektion im Unterraum definieren, indem wir den Supremumsoperator  $\mathcal{T}_{N,d,c}(p) : \text{Kern}(\mathcal{D}) \rightarrow \text{Kern}(\mathcal{D}_f)$  einführen, welcher durch die Eigenschaft

$$(\mathcal{T}_{N,d,c}(p)[u_{N,c}], v_{N,c})_{V_e} = a(u_{N,c}, v_{N,c}; p), \quad v_{N,c} \in \text{Kern}(\mathcal{D}_f), \quad (4.64)$$

charakterisiert ist. Wir erhalten wiederum die Darstellung

$$\beta_{N,c}(p) = \inf_{u_{N,c} \in \text{Kern}(\mathcal{D}) \setminus \{0\}} \frac{\|\mathcal{T}_{N,d,c}(p)[u_{N,c}]\|_{V_e}}{\|u_{N,c}\|_{V_e}} \quad (4.65)$$

der inf-sup-Konstanten (4.61). Für den eben eingeführten Operator  $\mathcal{T}_{N,d,c}(p)$  definieren wir nun zwei unterschiedliche Vektor-Operatoren

$$\mathcal{T}_{p,d,c}^N : \mathbb{C}^N \rightarrow \mathbb{C}^N, \quad \mathcal{T}_p^{N,d,c} : \mathbb{C}^{N_c} \rightarrow \mathbb{C}^{N_{d,c}}.$$

Der erste Operator  $\mathcal{T}_{p,d,c}^N$  bildet dabei den FEM-Koeffizientenvektor der FEM-Funktion  $u_{N,c} \in \text{Kern}(\mathcal{D})$  auf den FEM-Koeffizientenvektor der Funktion

$$\mathcal{T}_{N,d,c}(p)[u_{N,c}] \in \text{Kern}(\mathcal{D}_f)$$

ab, während der zweite Operator  $\mathcal{T}_p^{N,d,c}$  den Koeffizientenvektor der Funktion  $u_{N,c}$  bezüglich der Basismatrix  $B_D$  auf den Koeffizientenvektor der Funktion  $\mathcal{T}_{N,d,c}(p)[u_{N,c}]$  bezüglich der Basismatrix  $B_{D_f}$  abbildet. Sei nun wiederum der Testraum gegeben durch

$$Y_{N,c} = \text{span} \{ \mathcal{T}_{N,d,c}(p)[(\xi_1)_{N,c}], \dots, \mathcal{T}_{N,d,c}(p)[(\xi_N)_{N,c}] \}.$$

Das zur Petrov-Galerkin-Projektion zugehörige Problem lautet somit:

Sei  $p \in \mathbb{P}$  gegeben. Bestimme  $u_{N,c}^{PG}(p) \in X_{N,c}$  sodass

$$a(u_{N,c}^{PG}(p), \psi; p) = f(\psi; p) - \sum_{k=1}^{Q_g} \Theta_k^g(p) a(u_{N,r,k}, \psi; p), \quad \psi \in Y_{N,c}, \quad (4.66)$$

wobei  $u_{N,r,k}$  die Gleichung (4.57) löst.

Wir erhalten wiederum die Darstellung der inf-sup-Konstanten

$$\beta_{N,c}^{PG}(p) = \inf_{u_{N,c} \in X_{N,c} \setminus \{0\}} \sup_{v_{N,c} \in Y_{N,c} \setminus \{0\}} \frac{a(u_{N,c}, v_{N,c}; p)}{\|u_{N,c}\|_{V_e} \|v_{N,c}\|_{V_e}}$$

der mit der Petrov-Galerkin-Projektion durchgeführten Reduktion und die äquivalente Aussage zu Lemma 4.1.

**Lemma 4.20.** *Für die Petrov-Galerkin-Projektion (4.66) gilt*

$$\beta_{N,c}^{PG}(p) = \inf_{u_{N,c} \in X_{N,c} \setminus \{0\}} \sup_{v_{N,c} \in Y_{N,c} \setminus \{0\}} \frac{a(u_{N,c}, v_{N,c}; p)}{\|u_{N,c}\|_{V_e} \|v_{N,c}\|_{V_e}} \geq \beta_{N,c}(p).$$

Nach Voraussetzung 4.16 und (4.64) erhalten wir, wie bereits in Abschnitt 4.1.6 erwähnt, die affine Darstellung des Supremumsoperators

$$\mathcal{T}_{N_{d,c}}(p) = \sum_{k=1}^{Q_a} \Theta_k^a(p) \mathcal{T}_{N_{d,c}}^k, \quad (4.67)$$

womit wir das zu dem Problem (4.66) äquivalente Gleichungssystem

$$A_p^{PG} u_c^{N,PG} = f_p^{PG}, \quad (4.68)$$

erhalten, wobei

$$A_p^{PG} = \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \overline{\Theta_l^a(p)} \Theta_k^a(p) \left( \mathcal{T}_l^{N_{d,c}}[\Xi_c] \right)^H B_{D_f}^H A_k^N B_D \Xi_c$$

und

$$\begin{aligned} f_p^{PG} &= \sum_{k=1}^{Q_f} \sum_{l=1}^{Q_a} \overline{\Theta_l^a(p)} \Theta_k^f(p) \left( \mathcal{T}_l^{N_{d,c}}[\Xi_c] \right)^H B_{D_f}^H f_k^N \\ &\quad - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_g} \sum_{m=1}^{Q_a} \overline{\Theta_m^a(p)} \Theta_k^a(p) \Theta_l^g(p) \left( \mathcal{T}_m^{N_{d,c}}[\Xi_c] \right)^H B_{D_f}^H A_k^N u_{r,l}^N \end{aligned}$$

mit

$$A_p^{PG} \in \mathbb{C}^{N \times N}, \quad f_p^{PG} \in \mathbb{C}^N.$$

Aufgrund der affinen Darstellung (4.67) des Supremumsoperators  $\mathcal{T}_{N_{d,c}}(p)$  können wir den Operator für jeden Parameter  $p$  und jede reduzierte Lösung  $u_N$  auswerten, insofern

#### 4. Reduzierte-Basis-Methode

---

wir die Matrizen  $\mathcal{T}_l^{\mathcal{N}_{d,c}}[\Xi_c]$  bestimmen und abspeichern, denn dann erhalten wir gerade die FEM-Darstellung

$$\mathcal{T}_{\mathcal{N}_{d,c}}(p)[u_{N,c}] = \sum_{k=1}^{\mathcal{N}} \left( \sum_{l=1}^{Q_a} \Theta_l^a(p) B_{D_f} \mathcal{T}_l^{\mathcal{N}_{d,c}}[\Xi_c] u_c^N \right)_k \varphi_k.$$

Das Abspeichern der Matrizen  $\mathcal{T}_l^{\mathcal{N}_{d,c}}[\Xi_c]$  liefert noch weitere Vorteile bei der Berechnung der inf-sup-Konstanten, welche in Abschnitt 4.2.4 noch näher erläutert werden. Gilt zudem Voraussetzung 4.18, dann können wir für (4.59), wegen der eindeutigen Zerlegung von  $V_{\mathcal{N}}$  im Test- und Ansatzraum, auch eine Ritz-Galerkin-Projektion innerhalb des Kerns von  $\mathcal{D}$  auf den zugehörigen Unterraum  $X_{N,c}$  durchführen:

Sei  $p \in \mathbb{P}$  gegeben. Bestimme  $u_{N,c}^{RG}(p) \in X_{N,c}$ , sodass

$$a(u_{N,c}^{RG}(p), \varphi; p) = f(\varphi; p) - \sum_{k=1}^{Q_g} \Theta_k^g(p) a(u_{\mathcal{N},r,k}, \varphi; p), \quad \varphi \in X_{N,c}, \quad (4.69)$$

wobei  $u_{\mathcal{N},r,k}$  die Gleichung (4.57) löst.

Mit der Darstellung (4.63) ergibt sich dann das zu (4.69) äquivalente Gleichungssystem

$$A_p^{RG} u_c^{N, RG} = f_p^{RG}, \quad (4.70)$$

wobei

$$A_p^{RG} = \sum_{k=1}^{Q_a} \Theta_k^a(p) \Xi_c^H B_D^H A_k^N B_D \Xi_c$$

und

$$f_p^{RG} = \sum_{k=1}^{Q_f} \Theta_k^f(p) \Xi_c^H B_D^H f_k^N - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_g} \Theta_k^a(p) \Theta_l^g(p) \Xi_c^H B_D^H A_k^N u_{r,l}^N$$

mit

$$A_p^{RG} \in \mathbb{C}^{N \times N}, \quad f_p^{RG} \in \mathbb{C}^N.$$

Auch wenn Ritz-Galerkin-Projektionen von inf-sup-stabilen Problemen keine Stabilität aufweisen, gibt es dennoch Vorteile bei der Verwendung. Falls  $A_p^N$  Symmetrie-Eigenschaften besitzt, dann bleiben diese nach der Reduktion erhalten und zudem wird deutlich weniger Speicher sowohl in der Offline- als auch in der Online-Phase benötigt. Wie bereits in Abschnitt 4.1.3 generieren wir die Basis für den Raum  $X_{N,c}$  mithilfe des Greedy-Algorithmus 1. Hierzu leiten wir im folgenden Kapitel den Fehlerschätzer für die innerhalb dieses Kapitel hergeleiteten Projektionen her.

### 4.2.2. Fehlerdarstellung

Wir wollen die gleiche Aussage wie in Proposition 4.3 für die in Abschnitt 4.2.1 hergeleiteten RB-Lagrange-Systeme herleiten. Hierzu betrachten wir zunächst den Fehler zwischen der FEM-Lösung  $u_N(p) \in V_N$  des Problems (4.53), (4.54) und der RB-Lösung  $u_N(p) \in X_N \subset V_N$ . Aufgrund der Lösungstheorie, die wir im vorherigen Abschnitt diskutiert haben, erhalten wir die Lösung  $u_N$  zum Lagrange-System über die Zerlegung

$$V_N = \text{Kern}(\mathcal{D}) \oplus \text{Kern}(\mathcal{D})^\perp,$$

also

$$u_N(p) = u_{N,c}(p) + \sum_{k=1}^{Q_g} \Theta_k^g(p) u_{N,r,k}.$$

Die RB-Lösung  $u_N(p)$  besitzt nach (4.62) die Darstellung

$$u_N(p) = u_{N,c}(p) + \sum_{k=1}^{Q_g} \Theta_k^g(p) u_{N,r,k},$$

woraus direkt

$$u_N(p) - u_N(p) = u_{N,c}(p) - u_{N,c}(p)$$

folgt, d.h. wir betrachten lediglich den RB-Fehler

$$e_c(p) = u_{N,c}(p) - u_{N,c}(p)$$

zu dem Problem (4.59) im Kern( $\mathcal{D}$ ). Das dazugehörige Residuum kann geschrieben werden als

$$\begin{aligned} r_c(\varphi; p) &= a(e_c(p), \varphi; p) \\ &= f(\varphi; p) - \sum_{k=1}^{Q_g} \Theta_k^g(p) a(u_{N,r,k}, \varphi; p) - a(u_{N,c}(p), \varphi; p) \end{aligned} \quad (4.71)$$

mit  $\varphi \in \text{Kern}(\mathcal{D}_f)$ . Nun lässt sich  $r_c(\cdot; p) : \text{Kern}(\mathcal{D}_f) \rightarrow \mathbb{C}$  wiederum als Dualraumelement  $r_c \in (\text{Kern}(\mathcal{D}_f))'$  auffassen, womit nach dem Riesz'schen Darstellungssatz A.16 ein Element  $\hat{r}_c(p) \in \text{Kern}(\mathcal{D}_f)$  existiert, sodass

$$(\hat{r}_c(p), \varphi)_{V_e} = r_c(\varphi; p), \quad \varphi \in \text{Kern}(\mathcal{D}_f), \quad (4.72)$$

gilt und damit

$$(\hat{r}_c(p), \varphi)_{V_e} = a(e_c(p), \varphi; p).$$

Mit einer zu Proposition 4.3 vollkommen analogen Beweisführung erhalten wir die folgende Fehlerschranke.

**Proposition 4.21.** *Für den Fehler  $e(p) = u_{\mathcal{N}}(p) - u_N(p)$  gilt*

$$\|u_{\mathcal{N}}(p) - u_N(p)\|_{V_e} \leq \frac{\|\hat{r}_c(p)\|_{V_e}}{\beta_{\mathcal{N},c}(p)} = \Delta_{en,c}(p), \quad (4.73)$$

wobei  $\hat{r}_c(p)$  der Riesz-Repräsentant (4.72) des Residuums ist.

Wie bereits in Abschnitt 4.1.4 erläutert, ergeben sich damit auch entsprechende relative Fehlerschranken. In den beiden folgenden Abschnitten geben wir die für die Fehlerschranke notwendigen Konstanten an und gehen dabei auch sogleich auf die numerische Berechnung ein.

### 4.2.3. Berechnung des Riesz-Repräsentanten

Wie zuvor in Abschnitt 4.1.5 erhalten wir den Riesz-Repräsentanten

$$\hat{r}_c(p) = \sum_{k=1}^{Q_f} \Theta_k^f(p) \mathcal{C}_k - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_g} \Theta_k^a(p) \Theta_l^g(p) \mathcal{R}_{k,l} - \sum_{k=1}^{Q_a} \sum_{m=1}^N \Theta_k^a(p) (u_c^N)_m \mathcal{L}_{k,m},$$

des Residuums  $r_c(\varphi)$  über den Zusammenhang (4.72) und den affinen Darstellungen der Sesquilinearform  $a$  und der Antilinearform  $f$ , indem die dazugehörigen FE-Poisson-Probleme

$$(\mathcal{C}_k, \varphi)_{V_e} = f_k(\varphi), \quad \varphi \in \text{Kern}(\mathcal{D}_f), \quad k = 1, \dots, Q_f, \quad (4.74)$$

$$(\mathcal{R}_{k,l}, \varphi)_{V_e} = a_k(u_{\mathcal{N},r,l}, \varphi), \quad \varphi \in \text{Kern}(\mathcal{D}_f), \quad k = 1, \dots, Q_a, \quad l = 1, \dots, Q_g, \quad (4.75)$$

$$(\mathcal{L}_{k,m}, \varphi)_{V_e} = a_k((\xi_m)_{\mathcal{N},c}, \varphi), \quad \varphi \in \text{Kern}(\mathcal{D}_f), \quad k = 1, \dots, Q_a, \quad m = 1, \dots, N, \quad (4.76)$$

für  $\mathcal{C}_k, \mathcal{R}_{k,l}, \mathcal{L}_{k,m} \in \text{Kern}(\mathcal{D}_f)$  gelöst werden. Die zu (4.74), (4.75) und (4.76) äquivalenten Gleichungssysteme sind

$$\begin{aligned} B_{D_f}^H H_{V_{\mathcal{N}}} B_{D_f} \mathcal{C}_k^{N_{d,c}} &= B_{D_f}^H f_k, & k &= 1, \dots, Q_f, \\ B_{D_f}^H H_{V_{\mathcal{N}}} B_{D_f} \mathcal{R}_{k,l}^{N_{d,c}} &= B_{D_f}^H A_k u_{r,l}^{\mathcal{N}}, & k &= 1, \dots, Q_a, \quad l = 1, \dots, Q_g, \\ B_{D_f}^H H_{V_{\mathcal{N}}} B_{D_f} \mathcal{L}_{k,m}^{N_{d,c}} &= B_{D_f}^H A_k B_D (\xi_m)_c^{\mathcal{N}}, & k &= 1, \dots, Q_a, \quad m = 1, \dots, N, \end{aligned}$$

wobei  $(H_{V_{\mathcal{N}}})_{i,j} = (\varphi_j, \varphi_i)_{V_e}$  für  $i, j = 1, \dots, \mathcal{N}$  und  $\varphi_i$  die FEM-Basisfunktionen darstellt, und  $\mathcal{C}_k^{N_{d,c}}, \mathcal{R}_{k,l}^{N_{d,c}}, \mathcal{L}_{k,m}^{N_{d,c}}$  die zu  $\mathcal{C}_k, \mathcal{R}_{k,l}, \mathcal{L}_{k,m}$  zugehörigen Vektoren bezüglich der Darstellung (4.60) sind. Um die Notation in der Darstellung  $\|\hat{r}_c(p)\|_{V_e}$  zu vereinfachen,

führen wir zum einen die folgende Indexierung

$$\sum_{k=1}^{Q_f} \Theta_k^f(p) \mathcal{C}_k - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_g} \Theta_k^a(p) \Theta_l^g(p) \mathcal{R}_{k,l} = \sum_{m=1}^{Q_{f,a,g}} \Theta_m^{f,a,g}(p) \hat{\mathcal{C}}_m, \quad (4.77)$$

wobei  $Q_{f,a,g} = Q_f + Q_a Q_g$ , und zum anderen die Matrizen

$$\mathcal{L}_k^N = \left[ \mathcal{L}_{k,1}^{N_{d,c}}, \dots, \mathcal{L}_{k,N}^{N_{d,c}} \right] \quad (4.78)$$

ein, wobei der  $m$ -te Spaltenvektor der Matrix  $\mathcal{L}_k^N$  dem Vektor  $\mathcal{L}_{k,m}^{N_{d,c}}$  entspricht. Der Grund dieser Vereinfachung ist, dass die Berechnung von  $\hat{\mathcal{C}}_m$  für  $1 \leq m \leq Q_{f,a,g}$  unabhängig von der reduzierten Lösung  $u_N(p)$  ist und damit nur einmal durchgeführt werden muss, wohingegen für jeden neu generierten Basisvektor  $(\xi_m)_c^{N_c}$  zur Erweiterung des Unterraums  $X_N$  genau  $Q_a$  Berechnungen von (4.76) benötigt werden. Des Weiteren definieren wir noch die Vektoren  $\Theta^{f,a,g}(p) \in \mathbb{C}^{Q_{f,a,g}}$  und  $\Theta^a(p) \in \mathbb{C}^{Q_a}$ , deren  $k$ -ter Eintrag die entsprechende Parameterfunktion  $\Theta_k^{f,a,g}(p)$  bzw.  $\Theta_k^a(p)$  darstellt, und den Vektor

$$u_{\Theta_a}^{N, Q_a}(p) = ((\Theta_1^a(p) u_c^N)^T, \dots, (\Theta_{Q_a}^a(p) u_c^N)^T)^T. \quad (4.79)$$

Wir erhalten nun mit (4.77) und der Linearität des komplexen Skalarproduktes

$$\begin{aligned} \|\hat{r}_c(p)\|_{V_e}^2 &= \left( \sum_{k=1}^{Q_{f,a,g}} \Theta_k^{f,a,g}(p) \hat{\mathcal{C}}_k, \sum_{k=1}^{Q_{f,a,g}} \Theta_k^{f,a,g}(p) \hat{\mathcal{C}}_k \right)_{V_e} \\ &\quad - \left( \sum_{k=1}^{Q_{f,a,g}} \Theta_k^{f,a,g}(p) \hat{\mathcal{C}}_k, \sum_{k=1}^{Q_a} \sum_{m=1}^N \Theta_k^a(p) (u_c^N)_m \mathcal{L}_{k,m} \right)_{V_e} \\ &\quad - \left( \sum_{k=1}^{Q_a} \sum_{m=1}^N \Theta_k^a(p) (u_c^N)_m \mathcal{L}_{k,m}, \sum_{k=1}^{Q_{f,a,g}} \Theta_k^{f,a,g}(p) \hat{\mathcal{C}}_k \right)_{V_e} \\ &\quad + \left( \sum_{k=1}^{Q_a} \sum_{m=1}^N \Theta_k^a(p) (u_c^N)_m \mathcal{L}_{k,m}, \sum_{k=1}^{Q_a} \sum_{m=1}^N \Theta_k^a(p) (u_c^N)_m \mathcal{L}_{k,m} \right)_{V_e}. \end{aligned}$$

#### 4. Reduzierte-Basis-Methode

---

Mit der Verwendung von (4.78) erhalten wir wiederum mit der Linearität des komplexen Skalarproduktes

$$\begin{aligned}
\|\hat{r}_c(p)\|_{V_e}^2 &= \sum_{k=1}^{Q_{f,a,g}} \sum_{l=1}^{Q_{f,a,g}} \Theta_k^{f,a,g}(p) \overline{\Theta_l^{f,a,g}(p)} (\hat{\mathcal{C}}_l^{\mathcal{N}_{d,c}})^H B_{D_f}^H H_{V_N} B_{D_f} \hat{\mathcal{C}}_k^{\mathcal{N}_{d,c}} \\
&\quad - \sum_{k=1}^{Q_{f,a,g}} \sum_{l=1}^{Q_a} \Theta_k^{f,a,g}(p) \overline{\Theta_l^a(p)} (\mathcal{L}_l^N u_c^N)^H B_{D_f}^H H_{V_N} B_{D_f} \hat{\mathcal{C}}_k^{\mathcal{N}_{d,c}} \\
&\quad - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_{f,a,g}} \Theta_k^a(p) \overline{\Theta_l^{f,a,g}(p)} (\hat{\mathcal{C}}_l^{\mathcal{N}_c})^H B_{D_f}^H H_{V_N} B_{D_f} \mathcal{L}_k^N u_c^N \\
&\quad + \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p) \overline{\Theta_l^a(p)} (\mathcal{L}_l^N u_c^N)^H B_{D_f}^H H_{V_N} B_{D_f} \mathcal{L}_k^N u_c^N.
\end{aligned}$$

Verwenden wir nun noch die Notationen

$$\begin{aligned}
(\mathfrak{C})_{k,l} &= (\hat{\mathcal{C}}_l^{\mathcal{N}_{d,c}})^H B_{D_f}^H H_{V_N} B_{D_f} \hat{\mathcal{C}}_k^{\mathcal{N}_{d,c}}, & 1 \leq k, l \leq Q_{f,a,g}, \\
(\mathfrak{N}^k)_{l,m} &= (\hat{\mathcal{C}}_l^{\mathcal{N}_{d,c}})^H B_{D_f}^H H_{V_N} B_{D_f} \mathcal{L}_{k,m}^{\mathcal{N}_{d,c}}, & 1 \leq k \leq Q_a, 1 \leq l \leq Q_{f,a,g}, 1 \leq m \leq N, \\
(\mathfrak{L}^{l,k})_{m,n} &= (\mathcal{L}_{l,m}^{\mathcal{N}_{d,c}})^H B_{D_f}^H H_{V_N} B_{D_f} \mathcal{L}_{k,n}^{\mathcal{N}_{d,c}}, & 1 \leq k, l \leq Q_a, 1 \leq m, n \leq N.
\end{aligned}$$

und (4.79) ergibt sich somit

$$\|\hat{r}_c(p)\|_{V_e}^2 = \begin{pmatrix} \Theta^{f,a,g}(p) \\ -u_{\Theta^a}^{N,Q_a}(p) \end{pmatrix}^H \begin{pmatrix} \mathfrak{C} & \mathfrak{N}^1 & \dots & \mathfrak{N}^{Q_a} \\ (\mathfrak{N}^1)^H & \mathfrak{L}^{1,1} & \dots & \mathfrak{L}^{1,Q_a} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathfrak{N}^{Q_a})^H & \mathfrak{L}^{Q_a,1} & \dots & \mathfrak{L}^{Q_a,Q_a} \end{pmatrix} \begin{pmatrix} \Theta^{f,a,g}(p) \\ -u_{\Theta^a}^{N,Q_a}(p) \end{pmatrix},$$

Damit haben wir nun eine Darstellung von  $\|\hat{r}_c(p)\|_{V_e}^2$ , für die wir in jedem  $m$ -ten Iterationsschritt des Greedy-Algorithmus 1 für den neu generierten Basisvektor  $(\xi_m)_c^{\mathcal{N}_c}$  genau  $Q_a$  Lösungen  $\mathcal{L}_{k,m}^{\mathcal{N}_{d,c}}$ ,  $k = 1, \dots, Q_a$ , berechnen. Wir erhalten damit die von  $u_c^N$  und den Parametern unabhängige Matrix

$$\mathfrak{G} = \begin{pmatrix} \mathfrak{C} & \mathfrak{N}^1 & \dots & \mathfrak{N}^{Q_a} \\ (\mathfrak{N}^1)^H & \mathfrak{L}^{1,1} & \dots & \mathfrak{L}^{1,Q_a} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathfrak{N}^{Q_a})^H & \mathfrak{L}^{Q_a,1} & \dots & \mathfrak{L}^{Q_a,Q_a} \end{pmatrix}.$$

Hiermit lässt sich  $\|\hat{r}_c(p)\|_{V_e}^2$  für beliebige Vektoren  $u_c^N$  und Parameter  $p \in \mathbb{P}$  für den RB-Raum  $X_{N,c}$  bestimmen, indem wir

$$\|\hat{r}_c(p)\|_{V_e}^2 = \begin{pmatrix} (\Theta^{f,a,g}(p))^H & (-u_{\Theta^a}^{N,Q^a}(p))^H \end{pmatrix} \mathfrak{G} \begin{pmatrix} \Theta^{f,a,g}(p) \\ -u_{\Theta^a}^{N,Q^a}(p) \end{pmatrix} \quad (4.80)$$

auswerten.

#### 4.2.4. Berechnung der inf-sup-Konstante

Bevor wir uns der Berechnung der inf-sup-Konstanten widmen, können wir analog zu Abschnitt 4.1.6 die Koeffizientendarstellung des Supremumsoperators  $\mathcal{T}_p^{N_{d,c}}[u_c^{N_c}]$  bezüglich des Kerns von  $\mathcal{D}_f$  über die Lösung des Gleichungssystems

$$B_{D_f}^H H_{V_N} B_{D_f} \mathcal{T}_p^{N_{d,c}}[u_c^{N_c}] = B_{D_f}^H A_p^N B_D u_c^{N_c}$$

herleiten. Mit den Matrizen

$$\begin{aligned} A_p^{N_{d,c},N_c} &= B_{D_f}^H A_p^N B_D, \\ H_{V_N}^{N_{d,c}} &= B_{D_f}^H H_{V_N} B_{D_f} \end{aligned}$$

erhalten wir damit

$$\mathcal{T}_p^{N_{d,c}} = \left( H_{V_N}^{N_{d,c}} \right)^{-1} A_p^{N_{d,c},N_c}. \quad (4.81)$$

Der dazugehörige FEM-Koeffizientenvektor ist dann gegeben durch

$$\mathcal{T}_{p,d,c}^N[u_c^N] = B_{D_f} \mathcal{T}_p^{N_{d,c}}[u_c^{N_c}].$$

Wie bereits bei dem MinRes-Schätzer in Abschnitt 4.1.6 diskutiert, betrachten wir nun das quadrierte Minimierungsproblem (4.65) zur Bestimmung der inf-sup-Konstante. Mit den FEM-Darstellungen und (4.81) erhalten wir

$$\begin{aligned} (\beta_{N,c}(p))^2 &= \inf_{u_{N,c} \in \text{Kern}(\mathcal{D}) \setminus \{0\}} \frac{\|\mathcal{T}_{N_{d,c}}(p)[u_{N,c}]\|_{V_e}^2}{\|u_{N,c}\|_{V_e}^2} \\ &= \min_{u_c^N \in \mathbb{C}^{N_c} \setminus \{0\}} \frac{\left( \mathcal{T}_{p,d,c}^N[u_c^N] \right)^H H_{V_N} \mathcal{T}_{p,d,c}^N[u_c^N]}{(u_c^N)^H H_{V_N} u_c^N} \\ &= \min_{u_c^{N_c} \in \mathbb{C}^{N_c} \setminus \{0\}} \frac{(u_c^{N_c})^H \left( A_p^{N_{d,c},N_c} \right)^H \hat{H}_{V_N}^{N_{d,c}} A_p^{N_{d,c},N_c} u_c^{N_c}}{(u_c^{N_c})^H H_{V_N}^{N_c} u_c^{N_c}}, \end{aligned}$$

#### 4. Reduzierte-Basis-Methode

---

mit  $H_{V_N}^{\mathcal{N}_c} = B_D^H H_{V_N} B_D$  und  $\hat{H}_{V_N}^{\mathcal{N}_{d,c}} = \left(H_{V_N}^{\mathcal{N}_{d,c}}\right)^{-H} H_{V_N}^{\mathcal{N}_{d,c}} \left(H_{V_N}^{\mathcal{N}_{d,c}}\right)^{-1}$ . Nach den Sätzen 4.5 und 4.8 erhalten wir

$$(\beta_{\mathcal{N},c}(p))^2 = \lambda_{\min}(p),$$

wobei  $\lambda_{\min}(p)$  der kleinste Eigenwert zu dem verallgemeinerten Eigenwertproblem

$$\tilde{M} v^N = \lambda_{\min}(p) \hat{M}[B_D]^T \mathcal{H} \hat{M}[B_D] v^N \quad (4.82)$$

mit

$$\tilde{M} = M \left[ \left(H_{V_N}^{\mathcal{N}_{d,c}}\right)^{-1} A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \right]^T \hat{M}[B_{D_f}]^T \mathcal{H} \hat{M}[B_{D_f}] M \left[ \left(H_{V_N}^{\mathcal{N}_{d,c}}\right)^{-1} A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \right]$$

und

$$\mathcal{H} = \begin{pmatrix} H_{V_N} & 0 \\ 0 & H_{V_N} \end{pmatrix}$$

ist. Wie bereits in Abschnitt 4.1.6 genauer beschrieben, können wir auch anstatt (4.82) eine entsprechende Approximationen wie die SCM oder den MinRes-Schätzer anwenden. Die SCM, siehe [14], kann ohne weitere Anpassung auch für die Lagrange-Systeme angewandt werden.

*Bemerkung 4.22.* Wie in Abschnitt 4.1.6 können wir eine Approximation

$$\beta_{\mathcal{N},c}(p) \leq |\tilde{\lambda}_{\min}(p)| = \beta_{\mathcal{N},c}^{\text{approx}}(p)$$

verwenden, wobei  $\tilde{\lambda}_{\min}(p)$  der betragsmäßig kleinste Eigenwert zu dem verallgemeinerten komplexen Eigenwertproblem

$$\left(A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c}\right)^H \left(H_{V_N}^{\mathcal{N}_{d,c}}\right)^{-1} A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} u_c^{\mathcal{N}_c} = \tilde{\lambda}_{\min}(p) H_{V_N}^{\mathcal{N}_c} u_c^{\mathcal{N}_c} \quad (4.83)$$

ist. Gilt Voraussetzung 4.18, dann können wir anstatt (4.83) auch das Eigenwertproblem

$$A_p^{\mathcal{N}_c, \mathcal{N}_c} u_c^{\mathcal{N}_c} = \tilde{\lambda}_{\min}(p) H_{V_N}^{\mathcal{N}_c} u_c^{\mathcal{N}_c}$$

betrachten.

Für den MinRes-Schätzer für Lagrange-Systeme geben wir nachfolgend die dazugehörigen Eigenwertprobleme an und gehen dabei kurz auf die numerischen Besonderheiten ein.

### MinRes-Schätzer

Beim MinRes-Schätzer betrachten wir die inf-sup-Konstante  $\beta_N(p)$  aus Lemma 4.1 für Petrov-Galerkin-Projektionen

$$\beta_{N,c}^{MR}(p) = \inf_{u_{N,c} \in X_{N,c} \setminus \{0\}} \frac{\|\mathcal{T}_{N,c}(p)[u_{N,c}]\|_{V_e}}{\|u_{N,c}\|_{V_e}},$$

für den nach Lemma 4.20

$$\beta_{N,c}^{MR}(p) \geq \beta_{N,c}(p)$$

gilt und damit eine obere Schranke an  $\beta_{N,c}(p)$  darstellt. Wird bei der Reduktion die Petrov-Galerkin-Projektion verwendet, dann stimmt  $\beta_{N,c}^{MR}(p)$  mit der inf-sup-Konstante  $\beta_{N,c}^{PG}(p)$  überein. Mit der Darstellung des Supremumsoperators (4.81) erhalten wir das Minimierungsproblem

$$(\beta_{N,c}^{MR}(p))^2 = \min_{u_c^N \in \mathbb{C}^N \setminus \{0\}} \frac{(u_c^N)^H \Xi_c^H \left( A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \right)^H \hat{H}_{V_N}^{\mathcal{N}_{d,c}} A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \Xi_c u_c^N}{(u_c^N)^H \Xi_c^H H_{V_N}^{\mathcal{N}_c} \Xi_c u_c^N}$$

mit  $\hat{H}_{V_N}^{\mathcal{N}_{d,c}} = \left( H_{V_N}^{\mathcal{N}_{d,c}} \right)^{-H} H_{V_N}^{\mathcal{N}_{d,c}} \left( H_{V_N}^{\mathcal{N}_{d,c}} \right)^{-1}$ , welches wiederum durch das reelle Eigenwertproblem

$$\tilde{M}^T \mathcal{H} \tilde{M} v_c^N = \lambda_{\min}^{MR}(p) M [\Xi_c]^T \hat{M} [B_D]^T \mathcal{H} \hat{M} [B_D] M [\Xi_c] v_c^N \quad (4.84)$$

mit

$$\tilde{M} = \hat{M} [B_{D_f}] M \left[ \left( H_{V_N}^{\mathcal{N}_{d,c}} \right)^{-1} A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \Xi_c \right]$$

und

$$\mathcal{H} = \begin{pmatrix} H_{V_N} & 0 \\ 0 & H_{V_N} \end{pmatrix}$$

gelöst wird.

*Bemerkung 4.23.* Anstelle des reellen Eigenwertproblems, können wir wegen Satz 4.8 die approximative inf-sup-Konstante

$$\beta_{N,c}^{approx,MR}(p) = \sqrt{\tilde{\lambda}_{\min}^{MR}(p)} = \beta_{N,c}^{MR}(p)$$

bestimmen, indem wir das verallgemeinerte komplexe Eigenwertproblem

$$\Xi_c^H \left( A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \right)^H \left( H_{V_N}^{\mathcal{N}_{d,c}} \right)^{-1} A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} \Xi_c u_c^N = \tilde{\lambda}_{\min}^{MR}(p) \Xi_c^H H_{V_N}^{\mathcal{N}_c} \Xi_c u_c^N \quad (4.85)$$

#### 4. Reduzierte-Basis-Methode

---

lösen. Bei der Berechnung von (4.85) benötigen wir  $N$  Lösungen des linearen Gleichungssystems

$$H_{V_N}^{\mathcal{N}_{d,c}} \hat{u}_p^{\mathcal{N}_{d,c}} = A_p^{\mathcal{N}_{d,c}, \mathcal{N}_c} (\xi_m)_c^{\mathcal{N}_c}, \quad 1 \leq m \leq N,$$

für jeden beliebigen Parameter  $p \in \mathbb{P}$ . Eine deutlich effizientere Berechnung ergibt sich durch die affine Darstellung (4.67) und das Speichern der Matrizen  $\mathcal{T}_k^{\mathcal{N}_{d,c}}[\Xi_c]$ , denn wir erhalten damit aus der Darstellung (4.81) die zu (4.85) äquivalente Formulierung

$$T_c^N(p) u_c^N = \tilde{\lambda}_{\min}^{MR}(p) \Xi_c^H H_{V_N}^{\mathcal{N}_c} \Xi_c u_c^N \quad (4.86)$$

mit

$$T_c^N(p) = \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p) \overline{\Theta_l^a(p)} \left( \mathcal{T}_l^{\mathcal{N}_{d,c}}[\Xi_c] \right)^H H_{V_N}^{\mathcal{N}_{d,c}} \mathcal{T}_k^{\mathcal{N}_{d,c}}[\Xi_c].$$

Wie in Abschnitt 4.1 können wir eine Verbesserung der approximativen Eigenschaft der Konstante  $\beta_{N,c}^{\text{approx}, MR}(p)$  dadurch erreichen, dass anstatt  $\Xi_c$  die Matrix  $X_c = [\Xi_c, v[\Xi_c]]$  verwendet wird, wobei  $v[(\xi_m)_c^{\mathcal{N}_c}]$  der Eigenvektor zum Eigenwertproblem (4.83) für den zur Bestimmung von  $(\xi_m)_{N,c}$  verwendeten Parameter  $p \in \mathbb{P}$  ist. Das heißt, es müssen zusätzlich  $Q_a$  Matrizen der Art  $\mathcal{T}_l^{\mathcal{N}_{d,c}}[v[\Xi_c]]$  abgespeichert werden.

Falls Voraussetzung 4.18 gilt, dann erleichtert sich zum einen die Berechnung der bisherigen Darstellung der inf-sup-Konstante und zum anderen können wir dann einen Galerkin-Schätzer verwenden. Dieser benötigt im Vergleich zum MinRes-Schätzer deutlich weniger Speicherplatz und auch dessen Rechenaufwand ist aufgrund einer möglichen Symmetrieerhaltung wesentlich geringer.

#### Galerkin-Schätzer für Lagrange-Systeme

Es gilt Voraussetzung 4.18. Beim Galerkin-Schätzer schränken wir (4.61) auf den Unterraum  $X_{N,c}$  ein, also

$$\beta_{N,c}^{GA}(p) = \inf_{u_{N,c} \in X_{N,c} \setminus \{0\}} \sup_{v_{N,c} \in X_{N,c} \setminus \{0\}} \frac{|a(u_{N,c}, v_{N,c}; p)|}{\|u_{N,c}\|_{V_e} \|v_{N,c}\|_{V_e}}, \quad p \in \mathbb{P}.$$

Wie bereits in Abschnitt 4.1.6 erläutert, erhalten wir

$$(\beta_{N,c}^{GA}(p))^2 = \lambda_{\min}^{GA}(p),$$

wobei  $\lambda_{\min}^{GA}(p)$  der kleinste Eigenwert zum reellen Eigenwertproblem

$$\tilde{M}^T \hat{M}[B_D]^T \mathcal{H} \hat{M}[B_D] \tilde{M} v^N = \lambda_{\min}^{GA}(p) M[\Xi_c]^T \hat{M}[B_D]^T \mathcal{H} \hat{M}[B_D] M[\Xi_c] v^N$$

mit

$$\tilde{M} = M [(\Xi_c^H B_D^H H_{V_N} B_D \Xi_c)^{-1} (\Xi_c^H B_D^H A_p^N B_D \Xi_c)].$$

*Bemerkung 4.24.* Die approximative inf-sup-Konstante

$$\beta_{N,c}^{GA}(p) \leq \tilde{\lambda}_{\min}^{GA}(p) = \beta_{N,c}^{\text{approx},GA}(p)$$

erhalten wir wieder durch das Lösen des verallgemeinerten komplexen Eigenwertproblems

$$\Xi_c^H B_D^H A_p^N B_D \Xi_c u_c^N = \tilde{\lambda}_{\min}^{GA}(p) \Xi_c^H B_D^H H_{V_N} B_D \Xi_c u_c^N \quad (4.87)$$

zum kleinsten Eigenwert  $\tilde{\lambda}_{\min}^{GA}(p)$ . Eine Verbesserung der approximativen Eigenschaft der Konstante  $\beta_{N,c}^{\text{approx},GA}(p)$  kann wiederum durch das Lösen des folgenden verallgemeinerten Eigenwertproblems

$$X_c^H B_D^H A_p^N B_D X_c u^N = \hat{\lambda}_{\min}^{GA}(p) X_c^H B_D^H H_{V_N} B_D X_c u^N$$

erreicht werden, wobei  $X_c = [\Xi_c, v[\Xi_c]]$  und  $v[(\xi_m)_c^N]$  der Eigenvektor zum Eigenwertproblem (4.82) für den zur Bestimmung von  $(\xi_m)_{N,c}$  verwendeten Parameter  $p \in \mathbb{P}$ . Bei der Betrachtung der Eigenvektorerweiterung soll wiederum Bemerkung 4.11 bedacht werden.

#### 4.2.5. Getrennte RB-Systeme

Um die Theorie der getrennten RB-Systeme näher zu erläutern betrachten wir das Problem (4.59) mit den entsprechenden affinen Darstellungen aus Voraussetzung 4.16 und der Gültigkeit von Voraussetzung 4.17:

Sei  $p \in \mathbb{P}$  gegeben. Bestimme  $u_{N,c}(p) \in \text{Kern}(\mathcal{D})$ , sodass

$$\sum_{k=1}^{Q_a} \Theta_k^a(p) a_k(u_{N,c}(p), \varphi) = \sum_{k=1}^{Q_f} \Theta_k^f(p) f_k(\varphi) - \sum_{k=1}^{Q_a} \sum_{k=1}^{Q_g} \Theta_k^a(p) \Theta_k^g(p) a_k(u_{N,r,k}, \varphi) \quad (4.88)$$

mit  $\varphi \in \text{Kern}(\mathcal{D}_f)$  gilt.

Unter gewissen Voraussetzungen an die Parametermenge  $\mathbb{P}$  und die Parameterfunktionen  $\Theta^a$ ,  $\Theta^f$ ,  $\Theta^g$  können wir anstatt (4.88) eine Reihe von Problemen betrachten und diese unabhängig voneinander lösen und reduzieren. Hierzu werden die folgenden Bedingungen benötigt.

**Voraussetzung 4.25.** Sei  $\mathbb{P} \subset \mathbb{R}^m$  gegeben durch  $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$  mit  $\mathbb{P}_1 \subset \mathbb{R}^{m_1}$  und  $\mathbb{P}_2 \subset \mathbb{R}^{m_2}$ , also  $p = (p_1^T, p_2^T)^T \in \mathbb{P}$  und  $p_1 \in \mathbb{P}_1$  sowie  $p_2 \in \mathbb{P}_2$ . Die Parameterfunktionen

haben bezüglich  $\mathbb{P}$  die folgenden Abhängigkeiten

$$\begin{aligned}\Theta_k^a(p) &= \Theta_k^a(p_1), & 1 \leq k \leq Q_a, \\ \Theta_k^f(p) &= \Theta_k^{f,1}(p_1)\Theta_k^{f,2}(p_2), & 1 \leq k \leq Q_f, \\ \Theta_k^g(p) &= \Theta_k^{g,1}(p_1)\Theta_k^{g,2}(p_2), & 1 \leq k \leq Q_g.\end{aligned}$$

Diese Voraussetzung bedeutet, dass die linke Seite des Problems (4.88) unabhängig vom Parameter  $p_2$  darstellbar ist und die Parameterfunktionen der rechten Seite eine entsprechende multiplikative Darstellung bezüglich  $p_2$  besitzen. Anstatt nun (4.88) für die komplette Parametermenge  $\mathbb{P}$  zu betrachten, können wir das folgende getrennte Problem betrachten:

Sei  $p_1 \in \mathbb{P}_1$  sowie  $p_2 \in \mathbb{P}_2$  gegeben. Bestimme  $u_{\mathcal{N},c}(p) \in \text{Kern}(\mathcal{D})$  über die Darstellung

$$u_{\mathcal{N},c}(p) = \sum_{k=1}^{Q_f} \Theta_k^{f,2}(p_2)u_{\mathcal{N},c}^{f,k}(p_1) - \sum_{k=1}^{Q_g} \Theta_k^{g,2}(p_2)u_{\mathcal{N},c}^{g,k}(p_1), \quad (4.89)$$

wobei  $u_{\mathcal{N},c}^{f,k}(p_1) \in \text{Kern}(\mathcal{D})$ ,  $1 \leq k \leq Q_f$ , die Lösung zu

$$\sum_{l=1}^{Q_a} \Theta_l^a(p_1)a_l(u_{\mathcal{N},c}^{f,k}(p_1), \varphi) = \Theta_k^{f,1}(p_1)f_k(\varphi) \quad (4.90)$$

mit  $\varphi \in \text{Kern}(\mathcal{D}_f)$  ist, sowie  $u_{\mathcal{N},c}^{g,k}(p_1) \in \text{Kern}(\mathcal{D})$ ,  $1 \leq k \leq Q_g$ , die Lösung zu

$$\sum_{l=1}^{Q_a} \Theta_l^a(p_1)a_l(u_{\mathcal{N},c}^{g,k}(p_1), \varphi) = \sum_{l=1}^{Q_a} \Theta_l^a(p_1)\Theta_k^{g,1}(p_1)a_l(u_{\mathcal{N},r,k}, \varphi) \quad (4.91)$$

mit  $\varphi \in \text{Kern}(\mathcal{D}_f)$  ist.

Der Vorteil der Betrachtung der getrennten Probleme ist die Darstellung (4.89), da wir hier zwar zunächst die Lösungen  $u_{\mathcal{N},c}^{f,k}(p_1)$  und  $u_{\mathcal{N},c}^{g,k}(p_1)$  für einen bestimmten Parameter  $p_1$  bestimmen müssen, aber danach für die Lösung  $u_{\mathcal{N},c}(p)$  des vollen Problems (4.88) den Parameter  $p_2 \in \mathbb{P}_2$  vollkommen beliebig wählen können. Im Gegensatz dazu muss bei der Betrachtung des vollen Problems (4.88) und direktem Lösen, lediglich eine Lösung bestimmt werden. Das getrennte RB-System erhalten wir nun, indem die entsprechenden Petrov-Galerkin- oder Ritz-Galerkin-Projektionen auf die einzelnen Problemformulierungen (4.90) und (4.91) angewandt werden. Da wir dabei die Parametermenge trennen und nun eine kleinere Parametermenge für jedes einzelne Problem betrachten, werden

auch weniger Basisfunktionen benötigt, um eine gute Approximation der einzelnen RB-Systeme zu erhalten. Die Anwendung der Theorie des getrennten Systems bringt noch weitere Vorteile mit sich. Zum einen können die einzelnen Probleme getrennt voneinander gelöst werden und lassen sich demnach gut parallelisieren und zum anderen entspricht der Parametervektor  $p_2$  in vielen Fällen den Anregungsparametern eines Systems, was wiederum Vorteile in der Optimalsteuerung mit sich bringt.

### 4.3. Numerische Ergebnisse

Um die verschiedenen Varianten der RBM nach Abschnitt 4.2 für die Lagrange-Systeme aus Kapitel 3 zu testen, beschreiben wir innerhalb dieses Kapitels drei verschiedene Modelle. Diese Modelle werden wiederum zum Testen der in Kapitel 5 hergeleiteten Optimalsteuerungsprobleme verwendet. In allen drei Modellen betrachten wir die zu erwärmenden Epoxidharzobjekte innerhalb der gleichen quaderförmigen Kavität, siehe Abbildung 4.1, mit den Maßen:

Länge ( $x$ -Achse): 68.61 cm; Breite ( $y$ -Achse): 56.55 cm; Höhe ( $z$ -Achse): 27.53 cm.

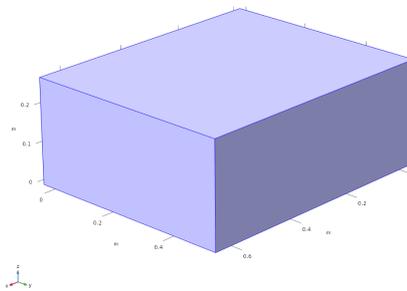


Abbildung 4.1: Quaderförmige Kavität.

Um die Randbedingung (2.80) für die Einspeisung der elektromagnetischen Wellen zu verwenden, benötigen wir eine Antenne mit koaxialer Anregung. Wir verwenden dabei in allen Modellen einfache Monopol-Antennen, dargestellt in Abbildung 4.2.

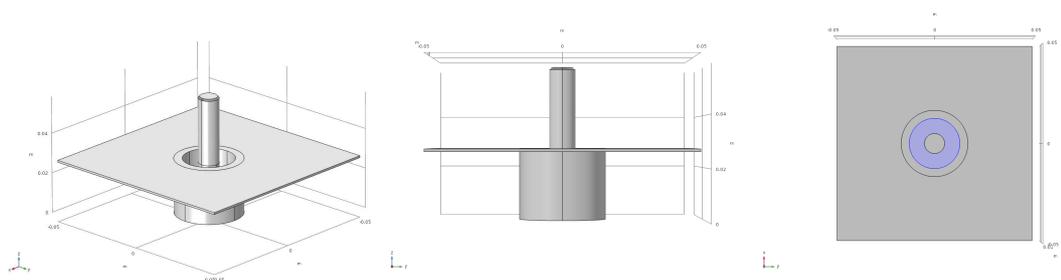


Abbildung 4.2: Monopol-Antenne.

Um das Verhalten der Antenne im Frequenzbereich von 2.4 bis 2.5 GHz bewerten zu können, werden in der Hochfrequenztechnik die Antennen im Freiraum vermessen und

das Verhältnis aus reflektierter Leistung und eingespeister Leistung betrachtet. In der Hochfrequenztechnik wird dieser Wert Streuparameter oder auch  $S$ -Parameter genannt. Wird zudem ein System mit  $N$  Eingängen betrachtet, wobei jeder Eingang auch als Ausgang aufgefasst werden kann, dann wird dieses System als  $N$ -Port-Netzwerk bezeichnet. Aus den Verhältnissen der reflektierten und transmittierenden Leistungen zu der Eingangsleistung an jedem Port, kann daraus die  $S$ -Parametermatrix, die das Eingangs-Ausgangsverhalten des Systems beschreibt, abgeleitet werden. Um den  $S_{ij}$ -Wert der Matrix  $S$  zu evaluieren, wird nur der  $i$ -te Port mit einem Watt Leistung angeregt und am  $j$ -ten Port die transmittierende Leistung, bzw. bei  $j = i$  die reflektierte Leistung, gemessen und das Verhältnis bestimmt. Auch die Modelle in dieser Arbeit stellen ein solches  $N$ -Port-Netzwerk dar, wobei  $N$  gleich der Anzahl der Antennen  $K_A$  ist, wodurch wir auch für die betrachteten Modelle eine  $S$ -Parametermatrix für das Eingangs-Ausgangsverhalten bestimmen können. Eine detaillierte Beschreibung zu  $S$ -Parametern und  $N$ -Port-Netzwerken kann in [37] nachgelesen werden. Die  $S_{11}$ -Parameterkurve der verwendeten Monopolantenne im Freiraum wird in Abbildung 4.3 dargestellt, wobei die in der Elektrotechnik weitverbreitete Dezibell-Darstellung verwendet wird. Diese wird mit der folgenden Formel bestimmt:

$$S_{11} \text{ in dB-Darstellung: } 20 \log_{10}(|S_{11}|). \quad (4.92)$$

Im betrachteten Frequenzbereich liegt die Anpassung bei unter  $-19$  dB, d.h. die Antenne überträgt innerhalb des betrachteten Frequenzbereiches mehr als 88% der Eingangsleistung. Für die Definition der Geometrie und die Diskretisierung der elektromagnetischen PDE (2.81)-(2.84) wird die FEM-Simulationssoftware COMSOL Multiphysics<sup>6</sup> verwendet. Da eine Schnittstelle zwischen der Simulationssoftware COMSOL Multiphysics und der Programmiersoftware MATLAB<sup>7</sup> existiert, wurde die Softwarebibliothek RBMatlab<sup>8</sup> als Grundgerüst für die RB-Lagrange-Systeme verwendet. Hierfür wurde als erstes eine Schnittstelle zu COMSOL implementiert, um die Daten der affinen Darstellungen mit COMSOL zu assemblieren. Für die RB-Reduktion der aus COMSOL assemblierten Daten, welche als Lagrange-System vorliegen, und damit der Theorie in Abschnitt 4.2 entsprechen, musste ein neuer Reduktionsalgorithmus für die Softwarebibliothek RBMatlab geschrieben werden. Hierfür wurde neben der Verwendung von Lagrange-Systemen auch die Möglichkeit der Petrov-Galerkin-Projektion und die Berechnung des Fehlerschätzers entsprechend der Theorie in Abschnitt 4.2 implementiert.

<sup>6</sup><https://www.comsol.de/>

<sup>7</sup><https://de.mathworks.com>

<sup>8</sup><https://www.morepas.org/software/rbmatlab/>

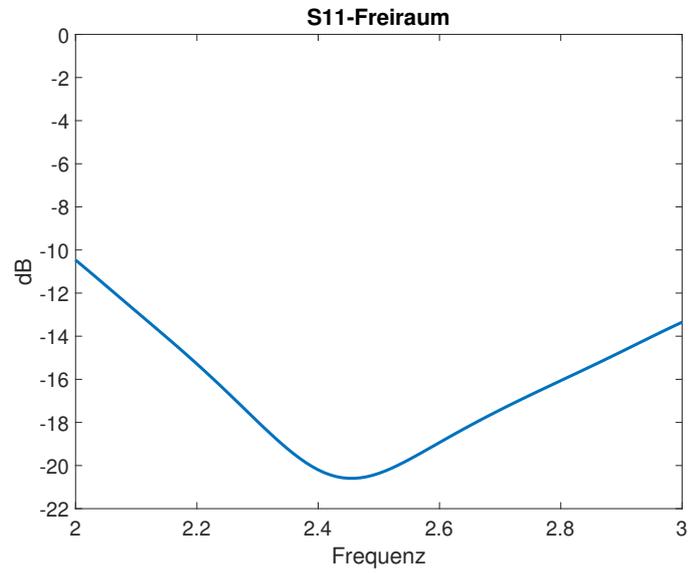


Abbildung 4.3: Darstellung des  $S_{11}$ -Parameters in dB.

### 4.3.1. Modell I: Zwei Antennen, ein Objekt

#### Modellbeschreibung

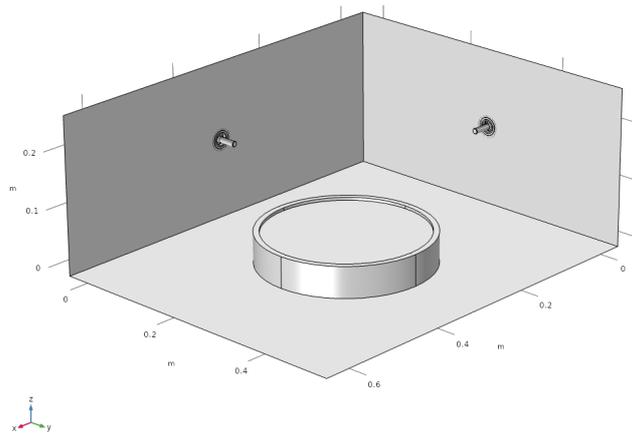


Abbildung 4.4: Geometrie des ersten Modells.

Das erste der drei Modelle besteht aus zwei Monopol-Antennen und einem zu erwärmenden Objekt in einer Schale, dargestellt in den Abbildungen 4.4 und 4.5. Dabei sind die beiden Monopol-Antennen mittig an den entsprechenden Randflächen der Kavität angebracht.

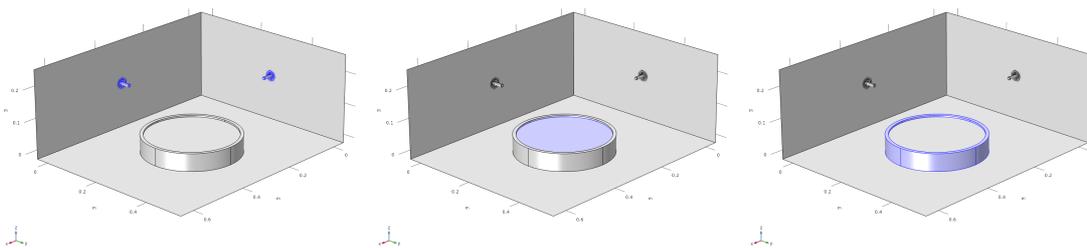


Abbildung 4.5: Geometrie des ersten Modells: zwei Antennen (links), ein zu erwärmendes Objekt (mitte), eine Schale (rechts).

#### 4. Reduzierte-Basis-Methode

---

Das zu erwärmende Objekt und die Schale haben die nach Abbildung 4.6 definierten Maße:

$$h_a = 5.5\text{cm}; \quad h_i = 4\text{cm}; \quad b = 1\text{cm}; \quad r = 14\text{cm}. \quad (4.93)$$

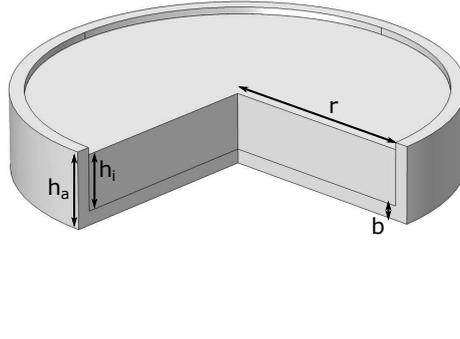


Abbildung 4.6: Maße des Erwärmungsobjektes.

Damit ergibt sich für das zu erwärmende Objekt das auf eine Nachkommastelle gerundete Volumen  $V = 2463.0\text{cm}^3$ . Bevor wir die affine Form des Modells angeben, müssen wir jedem Objekt ein Material zuordnen. Für die äußeren Wände der Kavität und der Monopol-Antenne verwenden wir Edelstahl, dessen Materialparameter gegeben sind durch:

$$\varepsilon_{r,S}(\omega) = 1 + \frac{1041700}{\omega \varepsilon_0} i; \quad \mu_{r,S} = 1; \quad \mu_0 = 4\pi \cdot 10^{-7} \frac{Vs}{Am}; \quad \varepsilon_0 = 8.854187817 \cdot 10^{-12} \frac{As}{Vm}.$$

Um ein Lagrange-System zu erhalten, wollen wir den Einfluss der PEC-Randbedingung nach Abschnitt 3.4 in dem Modell abbilden. Deshalb definieren wir den Innenleiter der Monopol-Antenne als PEC-Material. Die Kavität und die Antenne sind mit Luft gefüllt:

$$\varepsilon_{r,L}(\omega) = \varepsilon_{r,L} = 1; \quad \mu_{r,L} = 1.$$

Für die Schale verwenden wir Gummi mit den Materialparametern:

$$\varepsilon_{r,G}(\omega) = \varepsilon_{r,G} = 3 + 0.01i; \quad \mu_{r,G} = 1.$$

Als Material des zu erwärmenden Objektes verwenden wir das Epoxidharz Rapid aus [1], welches je nach Aushärtungsgrad innerhalb der folgenden Intervallgrenzen liegt

$$\text{Re}(\varepsilon_{r,E}(\omega)) \in [3, 3.7]; \quad \text{Im}(\varepsilon_{r,E}(\omega)) \in [2, 3.5]; \quad \mu_{r,E} = 1.$$

Aufgrund der Betrachtung des Frequenzbereiches von 2.4 bis 2.5 GHz nehmen wir die Frequenzunabhängigkeit der Materialparameter des Epoxidharzes Rapid an. Die Materialparameter hängen dabei nur vom Aushärtungsgrad ab, weswegen wir im Folgenden  $\varepsilon_{r,E}$  anstatt  $\varepsilon_{r,E}(\omega)$  schreiben werden. Mit den Anregungsparametern  $P_{in,1}$ ,  $P_{in,2}$ ,  $\theta_1$  und  $\theta_2$  der beiden Antennen ergibt sich für die RBM der Parametervektor

$$p = (p_1, p_2, p_3, p_4, p_5, p_6, p_7)^T = (\omega, \operatorname{Re}(\varepsilon_{r,E}), \operatorname{Im}(\varepsilon_{r,E}), P_{in,1}, P_{in,2}, \theta_1, \theta_2)^T$$

und die zugehörige Parametermenge

$$\mathbb{P} = [4.8\pi, 5\pi] \times [3, 3.7] \times [2, 3.5] \times [0, 1] \times [0, 1] \times [0, 2\pi] \times [0, 2\pi].$$

Als nächstes definieren wir das Gebiet  $\Omega$  und die Ränder  $\Gamma$  für das Variationsproblem  $[V_K]$  aus Kapitel 3, gegeben durch die Gleichungen (3.3) und (3.4). Das Gebiet  $\Omega$  ist definiert als

$$\Omega = \Omega_{Luft} \cup \Omega_{Gummi} \cup \Omega_{Epoxid} = \Omega_L \cup \Omega_G \cup \Omega_E.$$

Den Innenleiter der Monopol-Antennen approximieren wir mit der PEC-Randbedingung und die Edelstahlgebiete approximieren wir mit der Impedanzbedingung. Damit ist  $\Gamma_{IMP}$  nur durch die Ränder mit Edelstahl definiert. Als letztes bezeichnen wir mit  $\Gamma_{P,1}$  und  $\Gamma_{P,2}$  die beiden Portoberflächen der beiden Antennen. Wir erhalten damit die folgende affine Darstellung der Sesquilinearform  $a$  des Variationsproblems  $[V_{D,N}]$  aus Abschnitt 3.3 für das Modell eins:

$$\begin{aligned} \Theta_1^a(p) &= 1, & a_1(E_N, \varphi) &= \int_{\Omega_L} \frac{1}{\mu_{r,L}} (\nabla \times E_N) \cdot (\nabla \times \varphi) \, d\Omega \\ & & &+ \int_{\Omega_G} \frac{1}{\mu_{r,G}} (\nabla \times E_N) \cdot (\nabla \times \varphi) \, d\Omega \\ & & &+ \int_{\Omega_E} \frac{1}{\mu_{r,E}} (\nabla \times E_N) \cdot (\nabla \times \varphi) \, d\Omega, \\ \Theta_2^a(p) &= p_1^2, & a_2(E_N, \varphi) &= - \int_{\Omega_L} (10^9)^2 \mu_0 \varepsilon_0 \varepsilon_{r,L} E_N \cdot \varphi \, d\Omega \\ & & &- \int_{\Omega_G} (10^9)^2 \mu_0 \varepsilon_0 \varepsilon_{r,G} E_N \cdot \varphi \, d\Omega, \\ \Theta_3^a(p) &= p_1^2 (p_2 + ip_3), & a_3(E_N, \varphi) &= - \int_{\Omega_E} (10^9)^2 \mu_0 \varepsilon_0 E_N \cdot \varphi \, d\Omega, \end{aligned}$$

$$\begin{aligned}\Theta_4^a(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,S}(p_1)}{\mu_0 \mu_{r,S}}}, & a_4(E_{\mathcal{N}}, \varphi) &= - \int_{\Gamma_{IMP}} 10^9 \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi] d\Gamma, \\ \Theta_5^a(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}}, & a_5(E_{\mathcal{N}}, \varphi) &= - \sum_{k=1}^2 \int_{\Gamma_{P,k}} 10^9 \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi] d\Gamma.\end{aligned}$$

Den Faktor  $10^9$  erhalten wir dabei aus der Frequenz. Innerhalb der Assemblierung kompensiert der Faktor damit die Zehnerpotenzen der Naturkonstanten  $\varepsilon_0$  von  $10^{-12}$  und  $\mu_0$  von  $10^{-7}$ . Um eine in MATLAB berechenbare affine Darstellung der rechten Seite  $f$  der Variationsformulierung herzuleiten, müssen wir den Leistungsfaktor

$$P_{m,l}(\omega) = P_{m,l} = -\frac{1}{2} \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_{r,L}}} \int_{\Gamma_{P,l}} \operatorname{Re}(\sqrt{\varepsilon_{r,L}}) (\pi_t[E_{0,l}] \times \gamma_t[\pi_t[E_{0,l}]]) \cdot n d\Gamma$$

der einzelnen Ports genauer betrachten. Mit der Vektoridentität

$$(\pi_t[E_{0,l}] \times \gamma_t[\pi_t[E_{0,l}]]) \cdot n = -|\gamma_t[E_{0,l}]|^2 \quad (4.94)$$

erhalten wir somit

$$P_{m,l}(\omega) = \frac{1}{2} \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \int_{\Gamma_{P,l}} |\gamma_t[E_{0,l}]|^2 d\Gamma = \frac{1}{2} \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} P_{0,l}[E_{0,l}].$$

Damit die Lösung, die wir in MATLAB mit der affinen Darstellung erhalten, gleich der Lösung in COMSOL ist, muss der konstante Integralwert  $P_{0,l}[E_{0,l}]$  für jeden Port mithilfe von COMSOL bestimmt werden. Dies liegt daran, dass die Berechnung des Normierungsfaktors  $P_{m,l}(\omega)$ ,  $1 \leq l \leq K_A$ , in COMSOL von der Diskretisierung  $\mathcal{T}_h(\Omega)$  abhängt. Wir erhalten somit die folgende affine Darstellung der rechten Seite  $f$  der Variationsproblemformulierung [V<sub>D</sub>] für das Modell I:

$$\begin{aligned}\Theta_1^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_4}}{\sqrt{P_{m,1}}} e^{-ip_6}, & f_1(\varphi) &= -2 \int_{\Gamma_{P,1}} 10^9 \mu_0 \pi_t[E_{0,1}] \cdot \pi_t[\varphi] d\Gamma, \\ \Theta_2^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_5}}{\sqrt{P_{m,2}}} e^{-ip_7}, & f_2(\varphi) &= -2 \int_{\Gamma_{P,2}} 10^9 \mu_0 \pi_t[E_{0,2}] \cdot \pi_t[\varphi] d\Gamma.\end{aligned}$$

Aufgrund der Materialwahl Luft an der Antenne sind  $P_{m,1}$  und  $P_{m,2}$  von der Frequenz unabhängig. Mit der Lagrange-Formulierung der diskretisierten und evaluierten FEM-Probleme in COMSOL in Abschnitt 3.4 erhalten wir mit der eben hergeleiteten affinen

Darstellung des Lagrange-Systems

$$\begin{pmatrix} \sum_{k=1}^5 \Theta_k^a(p) A_k^{\mathcal{N}} & D^H \\ D & 0 \end{pmatrix} \begin{pmatrix} E^{\mathcal{N}} \\ \lambda^{\mathcal{N}_{PEC}} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^2 \Theta_k^f(p) f_k^{\mathcal{N}} \\ 0 \end{pmatrix}, \quad (4.95)$$

mit  $\mathcal{N} = 309218$  die Anzahl der Freiheitsgrade. Dabei ergibt sich die Dimensionszahl  $\mathcal{N}$  aus der Diskretisierung durch COMSOL. Das System (4.95) erfüllt die Voraussetzungen 4.16, 4.17 und 4.18, weswegen wir die in Abschnitt 4.2.1 hergeleitete RB-Lagrange-Lösungsmethode verwenden können, um ein reduziertes Modell herzuleiten. Der FEM-Koeffizientenvektor  $E^{\mathcal{N}}$  der Lösung  $E_{\mathcal{N}}$  ist gegeben durch

$$E^{\mathcal{N}} = E_c^{\mathcal{N}} + E_d^{\mathcal{N}} = B_D E_c^{\mathcal{N}_c}, \quad (4.96)$$

wobei  $\mathcal{N}_c = \mathcal{N} - \mathcal{N}_{PEC} = 306208$  und  $E_d^{\mathcal{N}} = 0$ , aufgrund der Randbedingung  $\Gamma_{PEC}$ . Das zu dem System (4.95) zugehörige affine Gleichungssystem, welches wir reduzieren wollen, ist somit gegeben durch

$$\begin{aligned} \left( \sum_{k=1}^5 \Theta_k^a(p) B_D^H A_k^{\mathcal{N}} B_D \right) E_c^{\mathcal{N}_c} &= \sum_{k=1}^2 \Theta_k^f(p) B_D^H f_k^{\mathcal{N}} - \sum_{k=1}^5 \Theta_k^a(p) B_D^H A_k^{\mathcal{N}} E_d^{\mathcal{N}} \\ &= \sum_{k=1}^2 \Theta_k^f(p) B_D^H f_k^{\mathcal{N}}. \end{aligned} \quad (4.97)$$

Die zu diesem Gleichungssystem zugehörige Kondition der FEM-Matrix für Parameter  $p \in \mathbb{P}$  liegt in der Größenordnung  $10^9$ . Betrachten wir die Parameterfunktionen  $\Theta_k^a(p)$ ,  $1 \leq k \leq 5$ , und  $\Theta_k^f(p)$ ,  $1 \leq k \leq 2$ , hinsichtlich der Theorie der getrennten Systeme in Abschnitt 4.2.5, dann erfüllen die Parameterfunktionen gerade die Voraussetzung 4.25 und wir können für die RBM anstatt des Gleichungssystems (4.97) die beiden getrennten Systeme

$$\left( \sum_{k=1}^5 \Theta_k^a(\tilde{p}) B_D^H A_k^{\mathcal{N}} B_D \right) E_c^{\mathcal{N}_c} = \tilde{\Theta}_l^f(\tilde{p}) B_D^H f_l^{\mathcal{N}}, \quad 1 \leq l \leq 2,$$

betrachten, wobei

$$\tilde{p} = (p_1, p_2, p_3)^T = (\omega, \operatorname{Re}(\varepsilon_{r,E}), \operatorname{Im}(\varepsilon_{r,E}))^T$$

und

$$\Theta_1^f(p) = \tilde{\Theta}_1^f(\tilde{p}) \sqrt{p_4} e^{-ip_6}, \quad \Theta_2^f(p) = \tilde{\Theta}_2^f(\tilde{p}) \sqrt{p_5} e^{-ip_7}$$

gilt. Dabei beschreibt das  $l$ -te System den Effekt der  $l$ -ten Antenne im ersten Modell.

### Ergebnisse der RBM

Um die RB-Lagrange Methode aus Abschnitt 4.2 anhand des ersten Modells mit Gleichungssystem (4.97) in vollem Umfang zu testen, wurden alle Kombinationen der nachfolgenden Stellschrauben für die RBM durchgeführt. Die Stellschrauben der RBM sind:

- Ritz-Galerkin-Projektion (RG)  $\leftrightarrow$  Petrov-Galerkin-Projektion (PG)
- Orthonormalisierung (ONJ)  $\leftrightarrow$  Normalisierung (ONN)
- Galerkin-Schätzer (GA)  $\leftrightarrow$  MinRes-Schätzer (MR)
- Eigenvektorenerweiterung (EVJ)  $\leftrightarrow$  keine Eigenvektorenerweiterung (EVN)
- Getrennte Systeme (GS-#)  $\leftrightarrow$  Volles System (VS)

Die verschiedenen RB-Systeme werden im Folgenden mit den in den Klammern angegebenen Bezeichnungen der Stellschrauben gekennzeichnet. Beispielsweise steht die Bezeichnung PG-ONJ-MR-EVN-GS-1 für die Verwendung einer Petrov-Galerkin-Projektion mit Orthonormalisierung, der Verwendung des MinRes-Schätzers, keiner Eigenvektorenerweiterung bei der Berechnung der inf-sup-Konstanten und der Verwendung von getrennten Systemen, wobei die 1 für Antenne eins steht. Bei dem zu reduzierenden Gleichungssystem (4.97) entspricht der Testraum dem Ansatzraum, weswegen wir eine Ritz-Galerkin-Projektion durchführen können. Auch wenn aus mathematischer Sicht keine Stabilität der RBM mit Ritz-Galerkin-Projektion für die in der Arbeit betrachteten elektromagnetischen Modelle garantiert werden kann, hat die Verwendung der Ritz-Galerkin-Projektion auch einige Vorteile gegenüber der Petrov-Galerkin-Projektion, weswegen wir die beiden Varianten im späteren Verlauf miteinander vergleichen wollen. Bei der Frage nach einer Orthonormalisierung der RB-Basismatrix gehen wir kurz auf die Notwendigkeit der Orthonormalisierung in Bezug auf die Approximationsgüte des reduzierten Systems und der Vermeidung annähernd singulärer reduzierter Matrizen ein. Unter der Eigenvektorenerweiterung versteht sich die Hinzunahme von Eigenvektoren bei der Verwendung des Galerkin- oder MinRes-Schätzers zur Bestimmung der inf-sup-Konstante. Dabei gehen wir auf die Approximationseigenschaften der inf-sup-Konstanten ein und vergleichen zudem die Wirkung bezüglich der generierten RB-Basismatrix. Neben der Verwendung eines Fehlerschätzers in Kombination mit dem Greedy-Algorithmus zur Bestimmung einer RB-Basis, wird auch die Kondition der reduzierten Systemmatrizen der RBM als Abbruchkriterium verwendet. Die Verschlechterung der Kondition der reduzierten Matrix führt mit zunehmender RB-Dimension zu einer Instabilität des reduzierten Systems und einer schlechteren Approximationsgüte. Bei der Erstellung der RB-Basismatrix wurde als obere Schranke an die Kondition der reduzierten Systemmatrizen die Toleranz  $tol_{cond} = 10^{12}$  verwendet. Damit wir die Varianten besser miteinander

vergleichen können, wurde für alle Varianten die gleiche diskrete Parametermenge  $\mathbb{P}_{300}^{train}$ , gegeben durch dreihundert zufällig gewählte und verschiedene Parametervektoren, für den Greedy-Algorithmus verwendet. Der einzige Unterschied in den Parametermengen ergibt sich dabei bei der Betrachtung der getrennten Systeme, bei denen die Eingangsleistungen  $P_{in,l}$  und Phasen  $\theta_l$  für  $1 \leq l \leq 2$  aus der diskreten Parametermenge eliminiert wurden. Wir widmen uns als erstes den verschiedenen Reduktionsmöglichkeiten und betrachten hierfür die Auswirkung der Orthogonalisierung der normierten Basisvektoren.

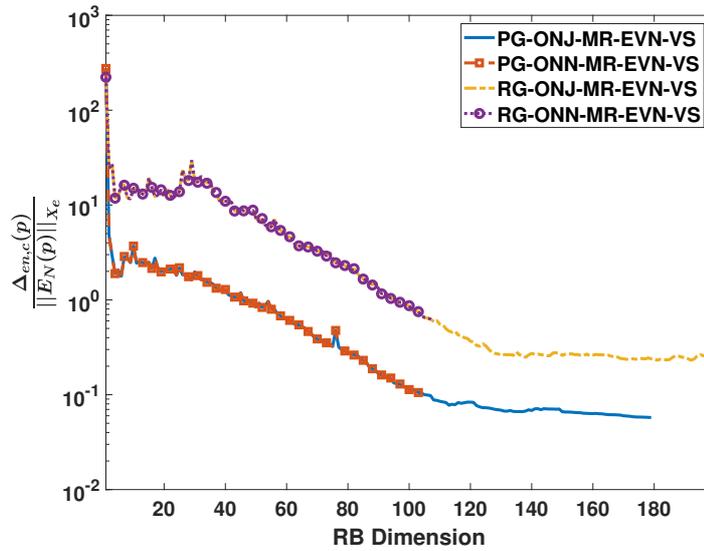


Abbildung 4.7: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

Nach Abbildung 4.7 sind die Werte des Fehlerschätzers bei Normalisierung und Orthonormalisierung gleich und der Fehlerschätzer für die Petrov-Galerkin-Projektion prognostiziert eine bessere Approximationsgüte im Vergleich zur Ritz-Galerkin-Projektion. Der ähnliche Verlauf liegt daran, dass wir bei beiden Varianten auf dieselben Unterräume projizieren. Wie wir anhand der Abbildung 4.7 sehen, bricht der Reduktionsprozess bei keiner Orthonormalisierung in der Offline-Phase früher ab. Dies liegt an dem Abbruchkriterium der Kondition der reduzierten Matrizen, welche wiederum von der Kondition der RB-Basismatrix abhängt. Die Konditionen der RB-Basismatrizen der vier Varianten aus Abbildung 4.7 bei der RB-Dimension von 104 sind:

$$\begin{aligned} \text{PG-ONJ-MR-EVN-VS: } &\sim 2.66, & \text{PG-ONN-MR-EVN-VS: } &\sim 5.57 \cdot 10^5, \\ \text{RG-ONJ-MR-EVN-VS: } &\sim 2.65, & \text{RG-ONN-MR-EVN-VS: } &\sim 2.24 \cdot 10^5. \end{aligned}$$

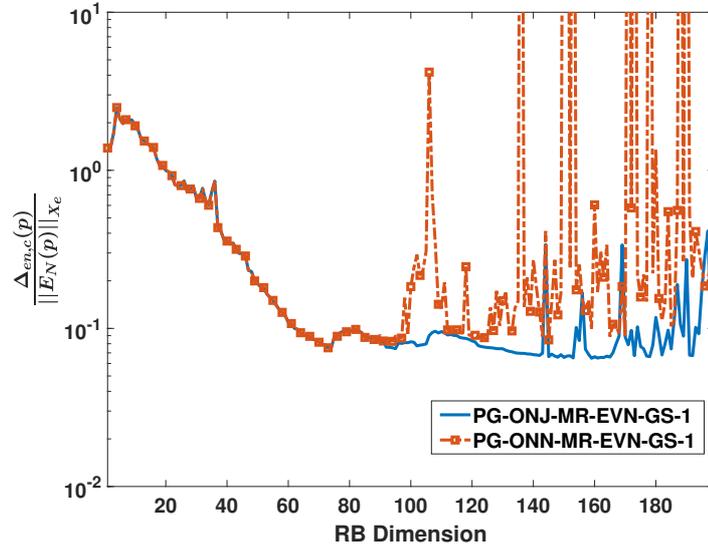


Abbildung 4.8: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase ohne Kontrolle der Kondition des reduzierten Systems.

Bei der Berechnung der Konditionen wurde die interne MATLAB-Funktion `cond()` verwendet und die Orthonormalisierung wurde mit einem einfachen Gram-Schmidt-Verfahren durchgeführt. In der Abbildung 4.8 wird der Effekt einer schlecht konditionierten RB-Basismatrix aufgezeigt. Dabei wurde keine Kontrolle der Kondition der reduzierten Systemmatrizen durchgeführt. Um den Effekt deutlicher aufzuzeigen, wurde hierfür der Fehlerschätzer zur Reduktion des getrennten Systems zum ersten Antennenport betrachtet. Bei zusätzlicher Verwendung der Kondition des reduzierten Systems als Abbruchkriterium terminieren die beiden Varianten aus Abbildung 4.8 bei den folgenden RB-Dimensionen:

PG-ONJ-MR-EVN-GS-1: 91,

PG-ONN-MR-EVN-GS-1: 58.

Während der Fehlerschätzer für das nichtorthonormalisierte RB-System ab der Dimension von 100 eine Instabilität aufweist, tritt dieser Effekt bei einer Orthonormalisierung erst ab der RB-Dimension von 140 auf.

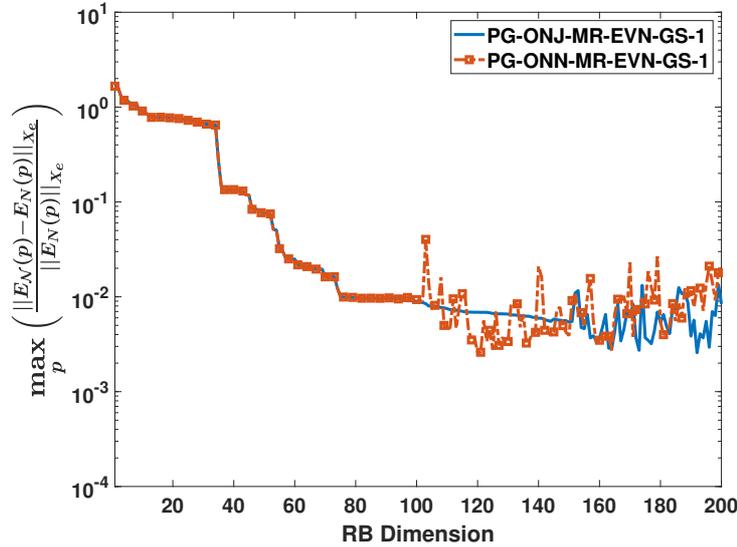


Abbildung 4.9: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

Dass dieser Effekt nicht nur Auswirkungen auf den Fehlerschätzer hat, sondern auch auf die echte Approximationsgüte des reduzierten Modells ist in Abbildung 4.9 zu sehen. Dabei wurde zur Ermittlung des maximalen relativen Fehlers eine Parametermenge  $\mathbb{P}_{100}^{test}$  mit

$$\mathbb{P}_{100}^{test} \cap \mathbb{P}_{300}^{train} = \emptyset$$

verwendet. Die Abbildung 4.9 zeigt dabei auf, dass mit der Orthonormalisierung der RB-Matrix eine verbesserte Approximationsgüte durch eine höhere RB-Dimension erreicht werden kann. Zudem erhalten wir bei der Nichtorthonormalisierung ab einer RB-Dimension von 100 ein instabiles reduziertes Modell, was mit der Instabilität des Fehlerschätzers in Abbildung 4.8 übereinstimmt. Dies zeigt die Notwendigkeit der Orthonormalisierung der RB-Basismatrix auf.

Als nächstes wollen wir die Auswirkungen der verschiedenen Varianten des Fehlerschätzers, wiedergegeben in Abbildung 4.10, analysieren. Die Varianten mit Eigenvektorenerweiterung weisen in Abbildung 4.10 ein instabiles Verhalten auf. Es wird darauf hingewiesen, dass bei diesen Auswertungen die Eigenvektorenmatrix nicht orthonormalisiert wurde. Auf den möglichen Effekt der Orthonormalisierung der Eigenvektorenmatrix gehen wir zu einem späteren Zeitpunkt ein. Zunächst wollen wir den Grund für das instabile Verhalten finden und betrachten deshalb zusätzlich neben den Fehlerschätzern auch noch den maximalen relativen Fehler, wobei wie zuvor eine Testparametermenge  $\mathbb{P}_{100}^{test}$  verwendet wird.

#### 4. Reduzierte-Basis-Methode

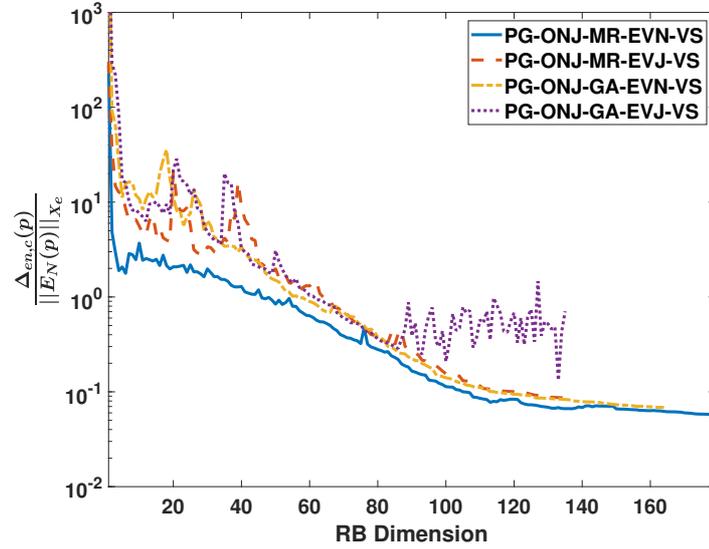


Abbildung 4.10: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase

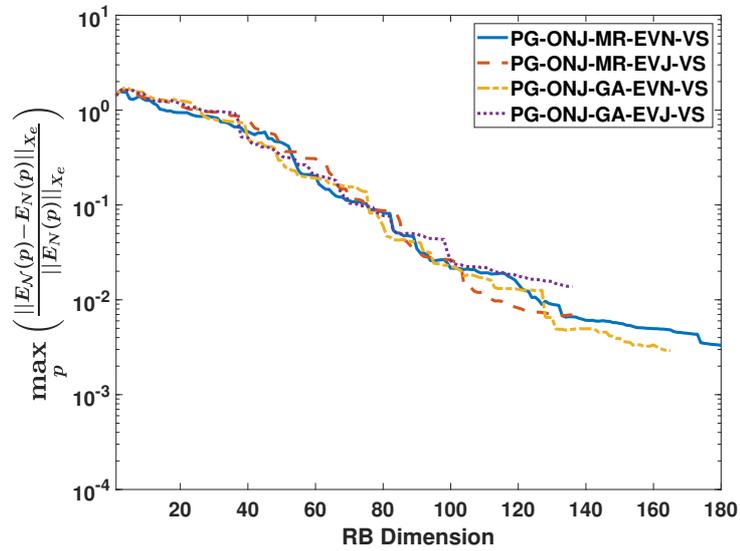


Abbildung 4.11: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

Nach Betrachtung der Abbildung 4.11 kann festgehalten werden, dass keines der reduzierten Systeme bei der RB-Dimension von 1 bis 100 eine nennenswert bessere Approximation im Vergleich zu den jeweiligen anderen Varianten liefert. Erst ab der RB-Dimension von 100 lässt sich ein Unterschied erkennen. Dies liegt an der Wahl der Parameter für die Erstellung der RB-Basismatrix, die mit dem Greedy-Algorithmus 1 aus dem maximalen Wert des Fehlerschätzers abgeleitet werden. Auch wenn die RB-Basismatrix  $\Xi_c$  orthonormalisiert wird, benötigen wir als Testraum bei der Petrov-Galerkin-Projektion den Supremumsoperator  $\mathcal{T}_p^{\mathcal{N}_{d,c}}$  angewendet auf die Basisvektoren. Mit der Zerlegung (4.67) ergeben sich deshalb innerhalb des Gleichungssystems (4.68) die Teilmatrizen  $\mathcal{T}_k^{\mathcal{N}_{d,c}}[\Xi_c]$ ,  $1 \leq k \leq Q_a$ . Diese Teilmatrizen dürfen aufgrund der affinen Darstellung (4.67) des Supremumsoperators weder normiert noch orthogonalisiert werden, da sonst die für die Petrov-Galerkin-Projektion notwendige Testmatrix  $\mathcal{T}_k^{\mathcal{N}_{d,c}}[\Xi_c]$  nicht mit der Summe der affinen Darstellung übereinstimmt. Dies führt bei nicht optimaler Wahl der Parameter für die Erstellung der RB-Basismatrix zu einer schlecht konditionierten Testmatrix  $\mathcal{T}_k^{\mathcal{N}_{d,c}}[\Xi_c]$  und damit zu schlecht konditionierten reduzierten Systemmatrizen, weshalb der Algorithmus aufgrund des Kondition-Abbruchkriteriums früher terminiert. Um das unterschiedliche Verhalten in Abbildung 4.10 zu ergründen, betrachten wir nun den Fehlerschätzer im Detail.

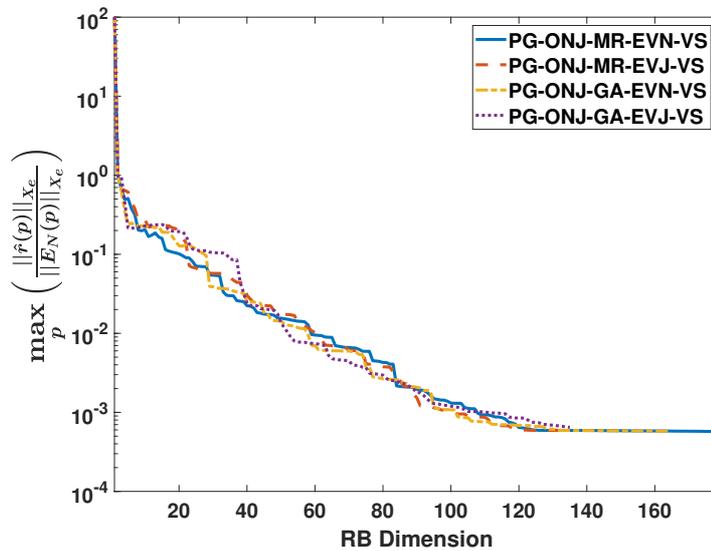


Abbildung 4.12: Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge  $\mathbb{P}_{300}^{train}$  während der Offline-Phase.

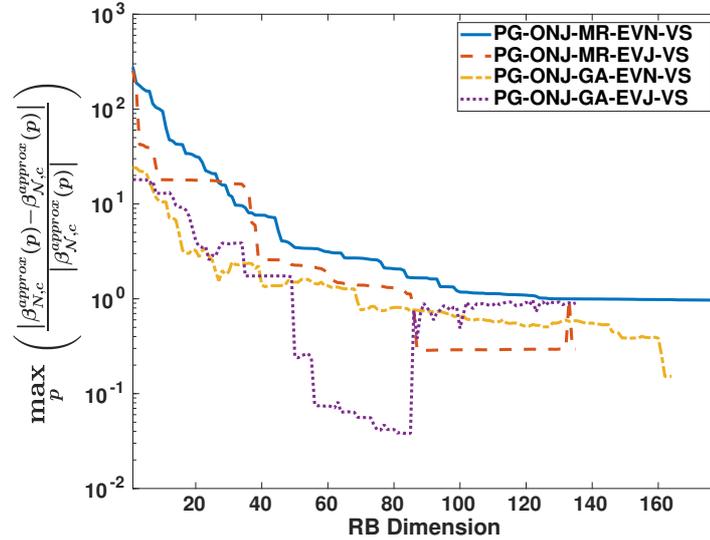


Abbildung 4.13: Maximaler relativer Fehler bezüglich der inf-sup-Konstanten  $\beta_{N,c}^{approx}(p)$  der Parametermenge  $\mathbb{P}_{300}^{train}$  während der Offline-Phase.

Nach Betrachtung der Abbildung 4.12 kann festgehalten werden, dass das unterschiedliche Verhalten des Fehlerschätzers nicht von der Berechnung (4.80) des Riesz-Repräsentanten abhängt. Wir widmen uns deshalb den verschiedenen Varianten zur Approximation der inf-sup-Konstanten. Hierzu wird der maximale, minimale und durchschnittliche relative Fehler der Approximation in den nachfolgenden Abbildungen 4.13-4.15 angegeben. Betrachten wir die Konvergenz der verschiedenen Varianten, dann sehen wir, dass die Galerkin-Approximation der inf-sup-Konstanten bei den RB-Dimensionen von 1 bis circa 80 eine bessere Konvergenz im Vergleich zur MinRes-Approximation aufweist. Der Sprung im Fehler bei der Galerkin-Approximation mit Verwendung der Eigenvektoren folgt aufgrund der zunehmenden Verschlechterung der Kondition der Eigenwertprobleme, da die Eigenvektorenmatrix nicht orthonormalisiert wurde. Die Verschlechterung der Approximationsgüte, aufgrund der fehlenden Orthonormalisierung, kann auch bei der MinRes-Approximation in Abbildung 4.14 beobachtet werden. Dabei wirkt sich der Effekt aufgrund der stabileren Approximation nicht so stark aus.

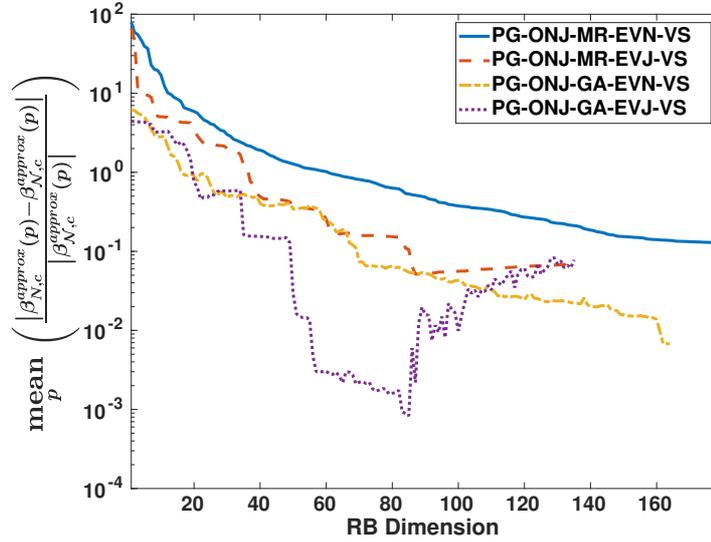


Abbildung 4.14: Durchschnittlicher relativer Fehler bezüglich der inf-sup-Konstanten  $\beta_{\mathcal{N},c}^{\text{approx}}(p)$  der Parametermenge  $\mathbb{P}_{300}^{\text{train}}$  während der Offline-Phase.

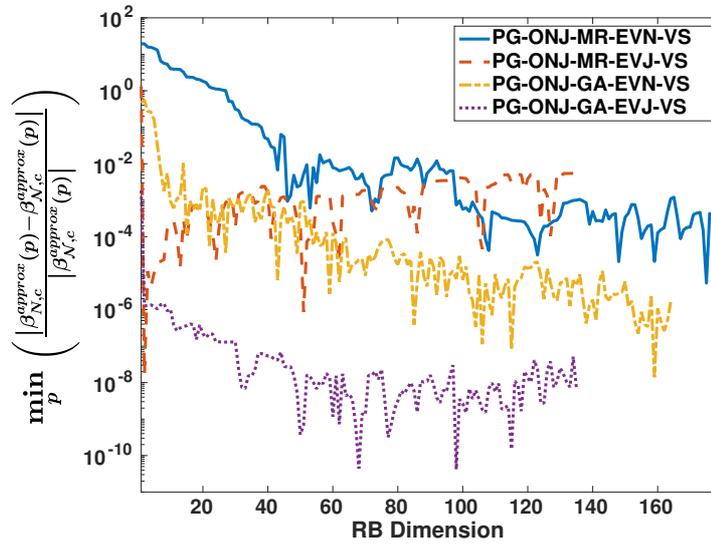


Abbildung 4.15: Minimaler relativer Fehler bezüglich der inf-sup-Konstanten  $\beta_{\mathcal{N},c}^{\text{approx}}(p)$  der Parametermenge  $\mathbb{P}_{300}^{\text{train}}$  während der Offline-Phase.

#### 4. Reduzierte-Basis-Methode

Die Frage, die sich nun stellt, ist, weshalb trotz einer besseren Approximation der inf-sup-Konstanten eine schlechtere Wahl der Parameter durchgeführt wird, die wiederum schneller zu einer schlecht konditionierten RB-Systemmatrix führt. Für diese Frage kann keine eindeutige Erklärung abgegeben werden. Betrachten wir aber die Verläufe der beiden Abbildungen 4.13 und 4.14, dann sehen wir bei bestimmten RB-Dimensionen, dass bei der zusätzlichen Verwendung der Eigenvektoren große Sprünge in der Approximationsgüte auftreten, während dies bei den anderen beiden Varianten nicht der Fall ist. Diese großen Sprünge in der Approximationsgüte haben einen großen Einfluss auf den Fehlerschätzer und damit auf die Wahl der Parameter, denn ein zunächst als gut approximiert gewerteter Parameter kann im Vergleich zu anderen Parametern bei Erweiterung der RB-Dimension auf einmal als deutlich schlechter approximierter Parameter gewertet werden und das nur, weil die inf-sup-Approximation für den einen Parameter schlagartig verbessert wird. Dies hat zur Folge, dass ein zur Ermittlung der nächsten Basisfunktion besser geeigneter Parameter, dessen Approximation der inf-sup-Konstante nicht so gut ist, nicht gewählt wird. Bei der MinRes-Approximation ohne Eigenvektoren hingegen haben wir einen größeren Fehler, aber dessen Auswirkung auf den Fehlerschätzer mit zunehmender RB-Dimension ist gleichmäßiger und stabiler. Um weitere Eigenschaften der unterschiedlichen Approximation herauszuarbeiten, geben wir nachfolgend den relativen Abstand zu zwei ausgewählten Parametern an.

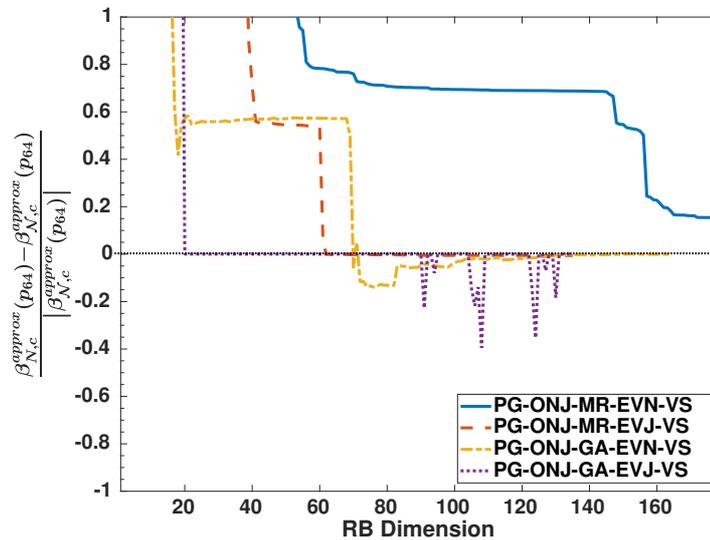


Abbildung 4.16: Relativer Abstand bezüglich der inf-sup-Konstanten  $\beta_{N,c}^{approx}(p_{64})$  mit  $p_{64} \in \mathbb{P}_{300}^{train}$  während der Offline-Phase.

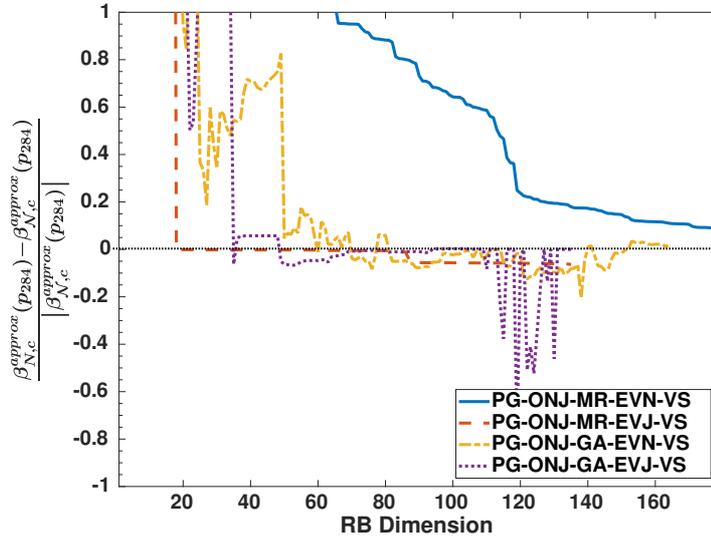


Abbildung 4.17: Relativer Abstand bezüglich der inf-sup-Konstanten  $\beta_{N,c}^{approx}(p_{284})$  mit  $p_{284} \in \mathbb{P}_{300}^{train}$  während der Offline-Phase.

In der Abbildung 4.16 liegt die gute Approximation der inf-sup-Konstanten der Kurve PG-ONJ-GA-EVJ-VS ab der RB-Dimension 20 daran, dass der Parameter  $p_{64}$  zu diesem Zeitpunkt in die RB-Basismatrix aufgenommen wurde, weshalb der entsprechende Eigenvektor für diesen Parameter bestimmt und für die nachfolgenden Approximationen verwendet wurde. Dies führt wiederum zu der guten Approximation bis zur RB-Dimension von 90, ab der die schlechte Kondition der Eigenvektorenmatrix und RB-Basismatrix wieder für einen größeren Fehler sorgt. Der gleiche Approximationseffekt tritt auch in Abbildung 4.17 bei der Kurve PG-ONJ-MR-EVJ-VS auf. Auch hier wurde der Parameter  $p_{284}$  bei der Berechnung des achtzehnten RB Basis- und Eigenvektors verwendet. Zudem lässt sich die Stabilität des MinRes-Schätzers gegenüber dem Galerkin-Schätzer gut erkennen.

Als nächstes betrachten wir zum einen die Orthonormalisierung der Eigenvektorenmatrix  $v[\Xi_c]$ , gekennzeichnet mit „ORTHEV“, und zum anderen die Orthonormalisierung der gemeinsamen Matrix  $(\Xi_c, v[\Xi_c])$ , bestehend aus der Eigenvektorenmatrix  $v[\Xi_c]$  und der RB-Basismatrix  $\Xi_c$ , gekennzeichnet mit „ORTHEVRB“.

#### 4. Reduzierte-Basis-Methode

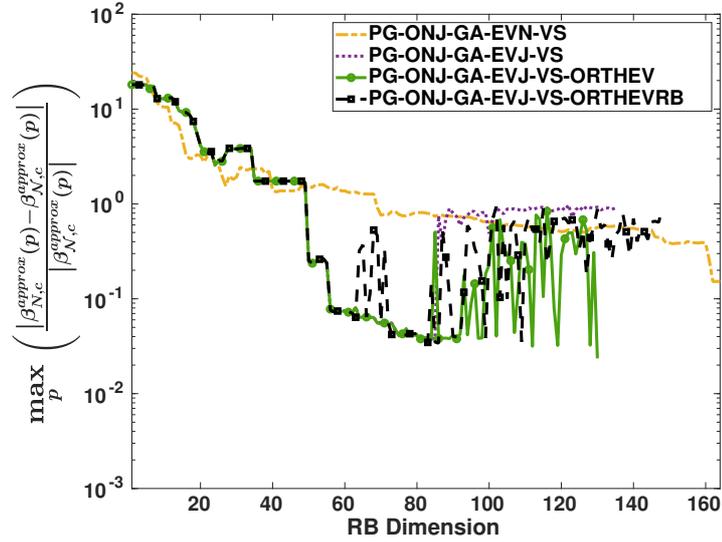


Abbildung 4.18: Maximaler relativer Fehler bezüglich der inf-sup-Konstanten  $\beta_{N,c}^{approx}(p)$  der Parametermenge  $\mathbb{P}_{300}^{train}$  während der Offline-Phase.

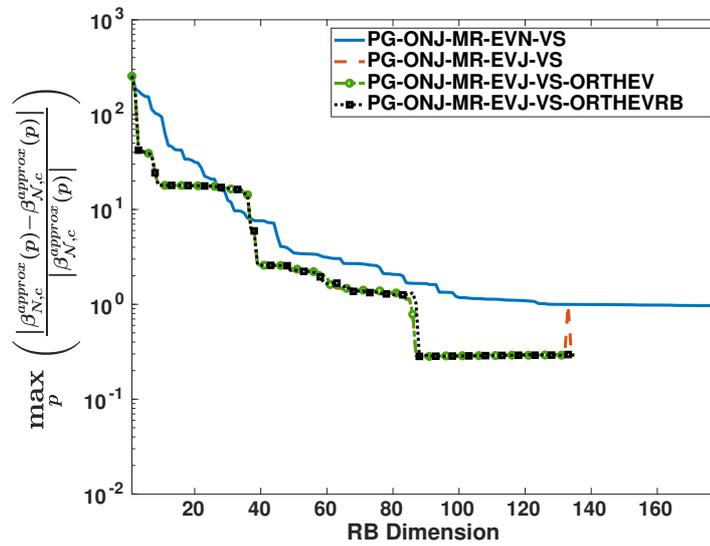


Abbildung 4.19: Maximaler relativer Fehler bezüglich der inf-sup-Konstanten  $\beta_{N,c}^{approx}(p)$  der Parametermenge  $\mathbb{P}_{300}^{train}$  während der Offline-Phase.

Im Fall des MinRes-Schätzers sehen wir in der Abbildung 4.19, dass die Varianten mit der Eigenvektorenerweiterung alle das gleiche Verhalten aufweisen und die Reduktion zum selben Zeitpunkt aufgrund der Konditionsbedingung abbricht. Auch beim Galerkin-Schätzer weisen die verschiedenen Varianten bis zur RB-Dimension 60 dasselbe Verhalten auf, siehe Abbildung 4.18. Danach weist die nichtorthonormalisierte Variante ein minimal stabileres Verhalten auf, während die Varianten mit Orthonormalisierung bereits instabil werden. Im Gegensatz dazu ist der relative Fehler mit Orthonormalisierung kleiner als ohne Orthonormalisierung. Es kann festgehalten werden, dass in dem hier betrachteten Problem keine Verbesserung durch die Orthonormalisierung der Eigenvektoren erreicht werden kann. Will man die Eigenvektorenerweiterung verwenden, dann muss hierfür ebenfalls eine Konditionsbedingung oder ein dafür entwickelter Fehlerschätzer verwendet werden, um die Anzahl zu verwendender Eigenvektoren zu begrenzen.

Als nächstes wollen wir die unterschiedlichen Projektions-Varianten miteinander vergleichen. Dabei weisen wir nochmals darauf hin, dass der Fehlerschätzer in Abbildung 4.7 für den Fall der Petrov-Galerkin-Projektion eine schärfere Schranke an den Fehler darstellt. Während zuvor die Approximationsart der inf-sup-Konstante der Auslöser für das unterschiedliche Verhalten war, ist dies bei Abbildung 4.7 nicht der Fall. Wie durch die Abbildungen 4.20 und 4.21 ersichtlich, unterscheidet sich der Fehlerschätzer der verschiedenen Projektionsvarianten diesmal in der relativen Norm des Riesz-Repräsentanten  $\hat{r}(p)$  mit der reduzierten Lösung  $u_N(p)$ .

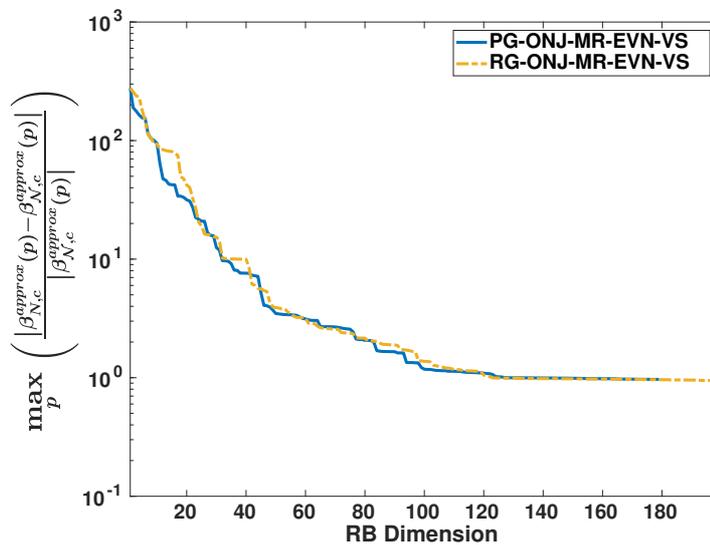


Abbildung 4.20: Maximaler relativer Fehler bezüglich der inf-sup-Konstanten  $\beta_{N,c}^{approx}(p)$  der Parametermenge  $\mathbb{P}_{300}^{train}$  während der Offline-Phase.

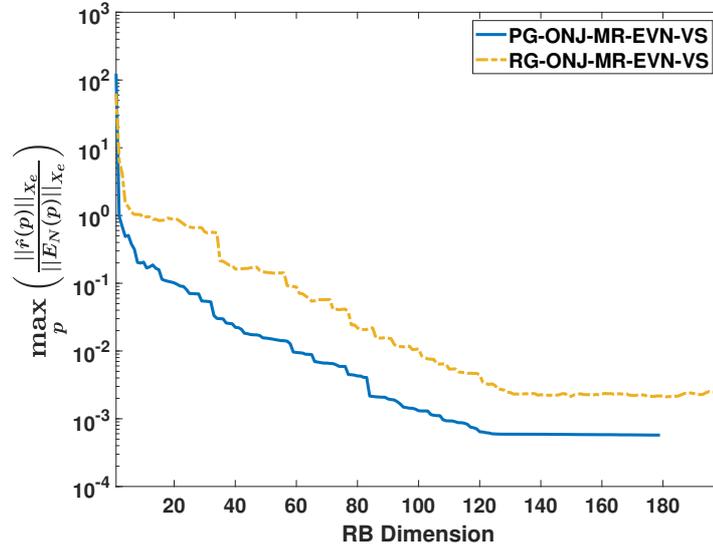


Abbildung 4.21: Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge  $\mathbb{P}_{300}^{train}$  während der Offline-Phase.

Dieser Unterschied in Abbildung 4.21 darf nicht in direkten Zusammenhang mit dem Residuum  $r_c$  des Problems, gegeben durch (4.71), gesetzt werden, denn die Norm des Riesz-Repräsentanten, die dargestellt werden kann als

$$\|\hat{r}_c(p)\|_{V_e} = \sup_{\varphi \in \text{Kern}(\mathcal{D}_f)} \frac{|r_c(\varphi; p)|}{\|\varphi\|_{V_e}},$$

ist das Supremum über das relative Residuum für alle  $\varphi \in \text{Kern}(\mathcal{D}_f)$ . Dass der Fehlerschätzer kein Indiz für den relativen Approximationsfehler darstellt, wird in Abbildung 4.22 verdeutlicht.

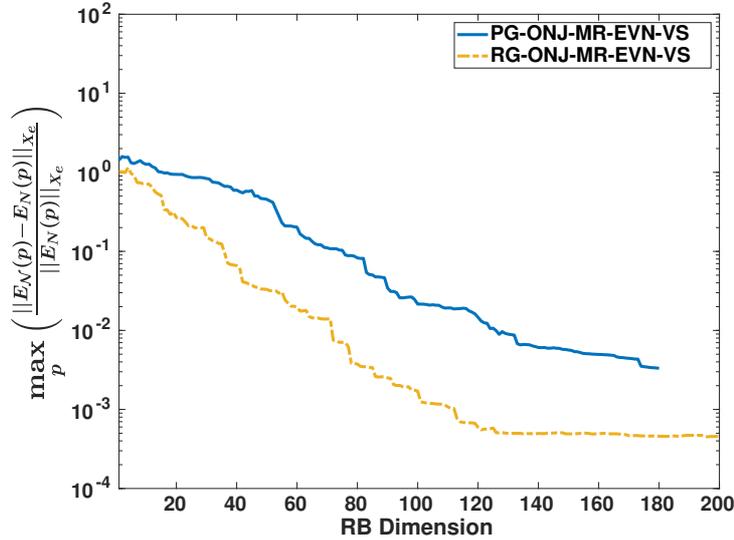


Abbildung 4.22: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

Zunächst einmal kann festgehalten werden, dass in beiden Fällen die Fehlerschätzer in Abbildung 4.7 zu der Parametermenge  $\mathbb{P}_{300}^{train}$  oberhalb der relativen Fehler der Abbildung 4.22 zu der Testparametermenge  $\mathbb{P}_{100}^{test}$  liegen. Auch wenn der Fehlerschätzer in Abbildung 4.7 einen möglicherweise größeren Fehler in der Ritz-Galerkin-Projektion prognostiziert, ergibt sich bei dem hier betrachteten Problem mit der Ritz-Galerkin-Projektion eine bessere Approximation der Lösungen bezüglich  $\mathbb{P}_{100}^{test}$ . Auch hier kann der Grund für die verbesserte Approximationsgüte nicht eindeutig geklärt werden, aber möglicherweise spielt der Erhalt der Hermitizität hierbei eine wichtige Rolle. Mit der Verwendung der RB-Basismatrix auch als Testmatrix werden die Eigenwerte der Systemmatrix und die damit verbundenen Informationen des FEM-Systems bei der Ritz-Galerkin-Projektion besser approximiert. Jedoch lässt sich im Gegensatz dazu nur für die Petrov-Galerkin-Projektion die Existenz von Lösungen zeigen und es kann die Stabilität der Projektion garantiert werden. Aus Stabilitätsgründen verwenden wir im Folgenden die Petrov-Galerkin-Projektionen in Kombination mit dem MinRes-Schätzer für die Betrachtung der getrennten Systeme, auch wenn die Ritz-Galerkin-Projektion eine bessere Approximation ermöglicht. Wir widmen uns nun der Betrachtung der getrennten Systeme. Die Werte der Tabelle 1 sind aus den Daten der Abbildung 4.23 abgeleitet.

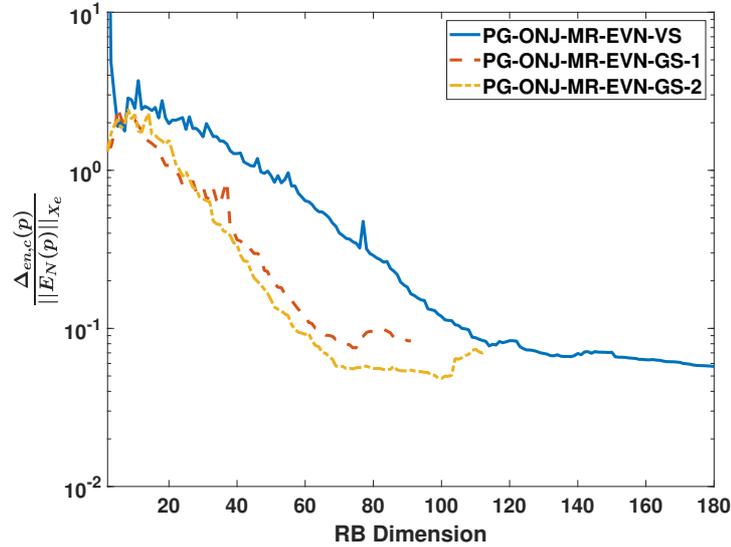
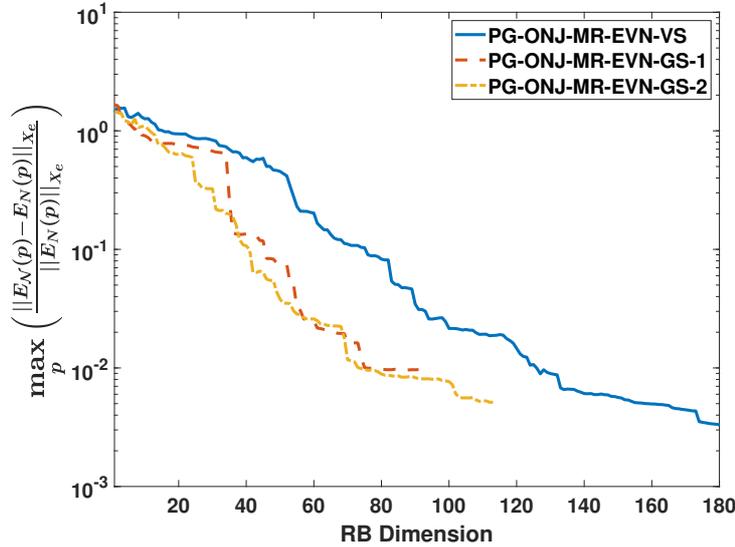


Abbildung 4.23: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

RB-Variante	$\frac{\Delta_{en,c}(p)}{\ E_N(p)\ _{X_e}} \leq 1.0$	$\frac{\Delta_{en,c}(p)}{\ E_N(p)\ _{X_e}} \leq 0.1$
PG-ONJ-MR-EVN-VS	46	106
PG-ONJ-MR-EVN-GS-1	20	63
PG-ONJ-MR-EVN-GS-2	23	56

Tabelle 1: RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der Fehlerschätzer-Schranken zu Abbildung 4.23.

Mit der Abbildung 4.23 bestätigt sich die verbesserte Konvergenz der einzelnen Teilsysteme, wobei dies aufgrund der kleineren Parametermenge zu erwarten war. Vergleichen wir jedoch die RB-Dimensionen zu den in Tabelle 1 angegebenen Schranken, dann benötigen wir in Summe für die getrennten Systeme tendenziell mehr RB-Basisfunktionen wie bei der Verwendung des vollen Systems. Wir wollen diese Aussage auch noch mit der Betrachtung der relativen Fehler bestätigen.

Abbildung 4.24: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

RB-Variante	$\frac{\ E_{\mathcal{N}}(p) - E_N(p)\ _{X_e}}{\ E_N(p)\ _{X_e}} \leq 0.1$	$\frac{\ E_{\mathcal{N}}(p) - E_N(p)\ _{X_e}}{\ E_N(p)\ _{X_e}} \leq 0.01$
PG-ONJ-MR-EVN-VS	76	126
PG-ONJ-MR-EVN-GS-1	46	76
PG-ONJ-MR-EVN-GS-2	41	74

Tabelle 2: RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der maximalen relativen Fehler-Schranken zu Abbildung 4.24.

Bei Betrachtung von Tabelle 2 und Abbildung 4.24 bestätigt sich die eben geäußerte Vermutung. Der Grund hierfür ist wahrscheinlich, dass eine Basisfunktion im vollen System Informationen beider Teilsysteme beinhaltet, während für dieselbe Information in den getrennten Systemen eine Basisfunktion für jedes Teilsystem benötigt wird. Um ein tieferes Verständnis für den relativen Fehler des reduzierten Modells zu entwickeln, betrachten wir für die RB-Variante PG-ONJ-MR-EVN-VS den vektorwertigen ortsabhängigen relativen Fehler

$$\frac{|E_{\mathcal{N}}(x) - E_N(x)|}{|E_N(x)|},$$

wobei  $x \in \Omega$  oder  $x \in \Omega_E$ . Die nachfolgenden ortsabhängigen Abbildungen wurden dabei

mit COMSOL generiert. Dabei ist vor allem der Approximationsfehler im Epoxidharz  $\Omega_E$  für die im Kapitel 5 beschriebene Optimalsteuerung wichtig. Für die nachfolgenden Abbildungen 4.25 und 4.26 wurde der Parametervektor

$$p_{\max} = \arg \max_{p \in \mathbb{P}_{100}^{test}} \left( \frac{\|E_{\mathcal{N}}(p) - E_N(p)\|_{X_e}}{\|E_N(p)\|_{X_e}} \right)$$

gewählt, wobei  $\mathbb{P}_{100}^{test}$  die in Abbildung 4.24 verwendete Testparametermenge ist. Zunächst vergleichen wir die elektromagnetischen Felder der FEM-Lösung  $E_{\mathcal{N}}(x, p_{\max})$  und der RB-Lösung  $E_N(x, p_{\max})$ . Wie in Abbildung 4.25 zu erkennen, sehen die Felddbilder der FEM- und RB-Lösung genau gleich aus. Um den Unterschied besser darzustellen, betrachten wir den entstandenen absoluten und relativen vektorielle Fehler. Sowohl zwischen den absoluten Fehlern und den Felddbildern in Abbildung 4.25 als auch zwischen den absoluten und relativen Fehlern ist kein direkter Zusammenhang erkennbar. Vergleichen wir hingegen die relativen Fehler in Abbildung 4.26 mit den Feldintensitäten in Abbildung 4.25, dann sehen wir, dass die größten relativen Fehler an den Stellen mit geringer Feldintensität entstehen. Wie in Abbildung 4.26 zu sehen ist, ist der absolute Fehler im Epoxidharz kleiner als in der Umgebung und den Antennen. Dies wiederum lässt vermuten, dass die in Kapitel 5 hergeleiteten Zielfunktionale mit Verwendung der reduzierten Modelle eine gute Approximation der Zielfunktionale bei Verwendung der FEM darstellen. Damit kann die Verwendung der ausgangsorientierten RBM sinnvoll sein um eine frühere Terminierung der Offline-Phase zu einem vorgegebenen Ausgangsfehler zu erzielen. Dennoch hat die Betrachtung feldbasierender Fehlerschranken, wie sie in dieser Arbeit verwendet werden, ihre Berechtigung. Beispielsweise werden in der Hochfrequenztechnik Modelle, wie sie hier verwendet werden, als  $N$ -Port-Netzwerke definiert, deren Eingangs-Ausgangsverhalten durch die vorher erwähnte  $S$ -Parametermatrix dargestellt wird. Um auch die auf den Antennenports definierten  $S$ -Parameter im reduzierten Modell gut zu approximieren muss der Fehler im kompletten System klein sein. Ist das reduzierte Modell demnach eine gute Approximation des kompletten Systems, dann kann das reduzierte Modell verwendet werden um Zusammenhänge zwischen  $S$ -Parametern und dem Parameterraum der RB oder den optimalen Steuerungsvektoren herauszuarbeiten.

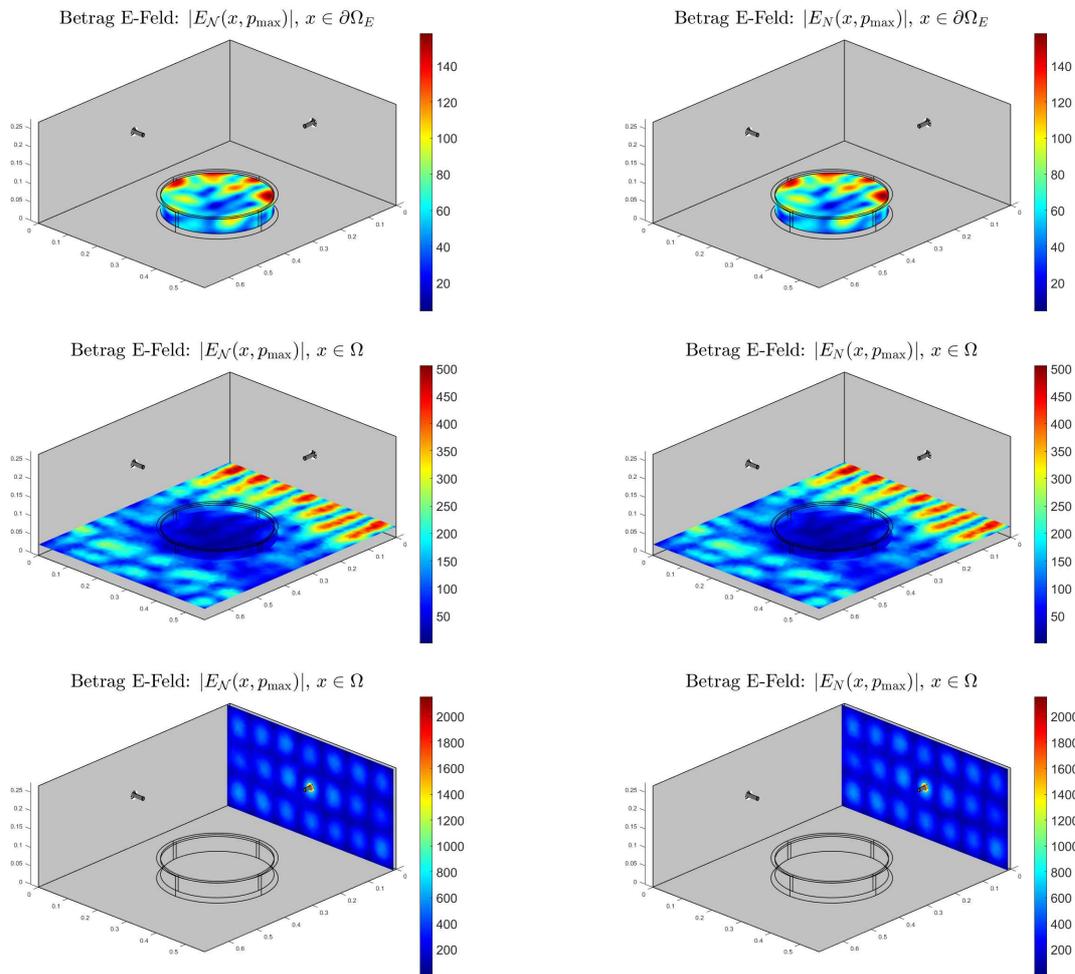


Abbildung 4.25: Komplexer Betrag der ortsabhängigen elektromagnetischen Feldvektoren zu  $p_{\max} \in \mathbb{P}_{100}^{test}$ .

#### 4. Reduzierte-Basis-Methode

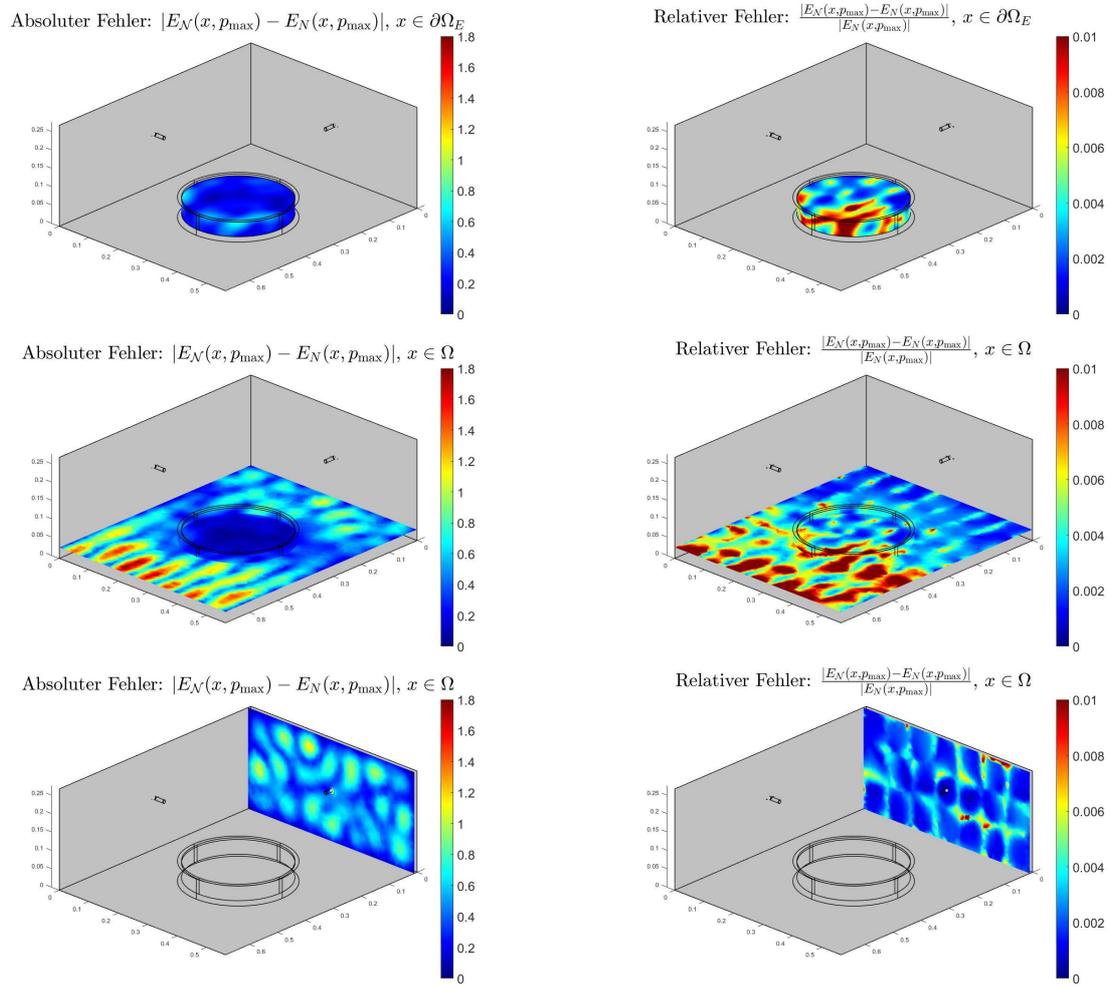


Abbildung 4.26: Absoluter und relativer Fehler der ortsabhängigen elektromagnetischen Feldvektoren zu  $p_{\max} \in \mathbb{P}_{100}^{test}$ .

## 4.3.2. Modell II: Vier Antennen, ein Objekt

## Modellbeschreibung

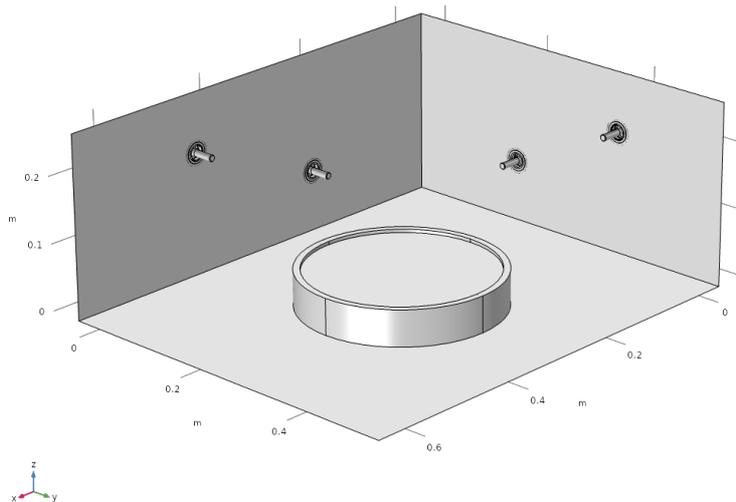


Abbildung 4.27: Geometrie des zweiten Modells.

Das zweite Modell beinhaltet vier Monopol-Antennen und ein zu erwärmendes Objekt in einer Schale, dargestellt in den Abbildungen 4.27 und 4.28. Dabei sind die vier Monopol-Antennen wie folgt an den entsprechenden Randflächen der Kavität angebracht:

1. Antenne:  $\frac{1}{3}$  der Länge ( $x$ -Achse),  $\frac{1}{3}$  der Höhe ( $z$ -Achse)
2. Antenne:  $\frac{2}{3}$  der Länge ( $x$ -Achse),  $\frac{2}{3}$  der Höhe ( $z$ -Achse)
3. Antenne:  $\frac{1}{3}$  der Breite ( $y$ -Achse),  $\frac{1}{3}$  der Höhe ( $z$ -Achse)
4. Antenne:  $\frac{2}{3}$  der Breite ( $y$ -Achse),  $\frac{2}{3}$  der Höhe ( $z$ -Achse)

#### 4. Reduzierte-Basis-Methode

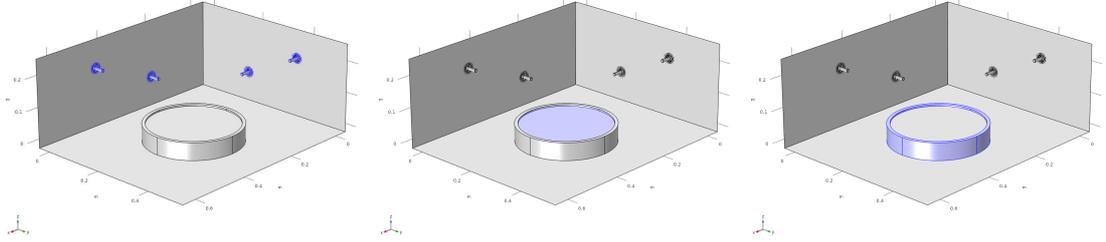


Abbildung 4.28: Geometrie des zweiten Modells: vier Antennen (links), ein zu erwärmendes Objekt (mitte), eine Schale (rechts).

Das zu erwärmende Objekt und die Schale haben dieselben Maße (4.93) wie im ersten Modell. Daher haben wir für das zu erwärmende Objekt ebenfalls das gerundete Volumen  $V = 2463.0\text{cm}^3$ . Wir wählen jegliche Gebiete und Materialparameter wie in Modell eins, wobei der einzige Unterschied die Anzahl der Portoberflächen ist. Wir bezeichnen die vier Portoberflächen der vier Antennen mit  $\Gamma_{P,1}$ ,  $\Gamma_{P,2}$ ,  $\Gamma_{P,3}$  und  $\Gamma_{P,4}$ . Daher definieren wir für die RBM den Parametervektor

$$p = (p_1, \dots, p_{11})^T = (\omega, \text{Re}(\varepsilon_{r,E}), \text{Im}(\varepsilon_{r,E}), P_{in,1}, P_{in,2}, P_{in,3}, P_{in,4}, \theta_1, \theta_2, \theta_3, \theta_4)^T$$

und die zugehörige Parametermenge

$$[4.8\pi, 5\pi] \times [3, 3.7] \times [2, 3.5] \times [0, 1]^4 \times [0, 2\pi]^4.$$

Wir erhalten damit die folgende affine Darstellung der Sesquilinearform  $a$  der Variationsproblemformulierung  $[V_D]$  für das Modell II:

$$\begin{aligned} \Theta_1^a(p) &= 1, & a_1(E_N, \varphi) &= \int_{\Omega_L} \frac{1}{\mu_{r,L}} (\nabla \times E_N) \cdot (\nabla \times \varphi) \, d\Omega \\ & & &+ \int_{\Omega_G} \frac{1}{\mu_{r,G}} (\nabla \times E_N) \cdot (\nabla \times \varphi) \, d\Omega \\ & & &+ \int_{\Omega_E} \frac{1}{\mu_{r,E}} (\nabla \times E_N) \cdot (\nabla \times \varphi) \, d\Omega, \\ \Theta_2^a(p) &= p_1^2, & a_2(E_N, \varphi) &= - \int_{\Omega_L} (10^9)^2 \mu_0 \varepsilon_0 \varepsilon_{r,L} E_N \cdot \varphi \, d\Omega \\ & & &- \int_{\Omega_G} (10^9)^2 \mu_0 \varepsilon_0 \varepsilon_{r,G} E_N \cdot \varphi \, d\Omega, \end{aligned}$$

$$\begin{aligned}
 \Theta_3^a(p) &= p_1^2(p_2 + ip_3), & a_3(E_{\mathcal{N}}, \varphi) &= - \int_{\Omega_E} (10^9)^2 \mu_0 \varepsilon_0 E_{\mathcal{N}} \cdot \varphi \, d\Omega, \\
 \Theta_4^a(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,S}(p_1)}{\mu_0 \mu_{r,S}}}, & a_4(E_{\mathcal{N}}, \varphi) &= - \int_{\Gamma_{IMP}} 10^9 \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_5^a(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}}, & a_5(E_{\mathcal{N}}, \varphi) &= - \sum_{k=1}^4 \int_{\Gamma_{P,k}} 10^9 \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi] \, d\Gamma.
 \end{aligned}$$

Mit derselben Herleitung der Normierungsfaktoren  $P_{m,k}$ ,  $1 \leq k \leq 4$ , wie im ersten Modell, erhalten wir die folgende affine Darstellung der rechten Seite  $f$  der Variationsproblemformulierung  $[V_{D,\mathcal{N}}]$  für das zweite Modell:

$$\begin{aligned}
 \Theta_1^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_4}}{\sqrt{P_{m,1}}} e^{-ip_8}, & f_1(\varphi) &= -2 \int_{\Gamma_{P,1}} 10^9 \mu_0 \pi_t[E_{0,1}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_2^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_5}}{\sqrt{P_{m,2}}} e^{-ip_9}, & f_2(\varphi) &= -2 \int_{\Gamma_{P,2}} 10^9 \mu_0 \pi_t[E_{0,2}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_3^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_6}}{\sqrt{P_{m,3}}} e^{-ip_{10}}, & f_3(\varphi) &= -2 \int_{\Gamma_{P,3}} 10^9 \mu_0 \pi_t[E_{0,3}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_4^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_7}}{\sqrt{P_{m,4}}} e^{-ip_{11}}, & f_4(\varphi) &= -2 \int_{\Gamma_{P,4}} 10^9 \mu_0 \pi_t[E_{0,4}] \cdot \pi_t[\varphi] \, d\Gamma.
 \end{aligned}$$

Aufgrund der Materialwahl Luft an der Antenne sind  $P_{m,1}$ ,  $P_{m,2}$ ,  $P_{m,3}$  und  $P_{m,4}$  von der Frequenz unabhängig. Mit der Lagrange-Formulierung in Abschnitt 3.4, sowie der diskretisierten und evaluierten FEM-Probleme in COMSOL erhalten wir mit der eben hergeleiteten affinen Darstellung das Lagrange-System

$$\left( \begin{array}{cc} \sum_{k=1}^5 \Theta_k^a(p) A_k^{\mathcal{N}} & D^H \\ D & 0 \end{array} \right) \left( \begin{array}{c} E^{\mathcal{N}} \\ \lambda^{\mathcal{N}_{PEC}} \end{array} \right) = \left( \begin{array}{c} \sum_{k=1}^4 \Theta_k^f(p) f_k^{\mathcal{N}} \\ 0 \end{array} \right). \quad (4.98)$$

Dabei ist die FEM-Dimension  $\mathcal{N} = 385562$  und  $\mathcal{N}_c = \mathcal{N} - \mathcal{N}_{PEC} = 379312$ . Wie im ersten Modell erhalten wir mit (4.96) das zu reduzierende Gleichungssystem

$$\left( \sum_{k=1}^5 \Theta_k^a(p) B_D^H A_k^{\mathcal{N}} B_D \right) E_c^{\mathcal{N}_c} = \sum_{k=1}^4 \Theta_k^f(p) B_D^H f_k^{\mathcal{N}}. \quad (4.99)$$

Analog zu der Herleitung bei Modell I erhalten wir hier die vier getrennten Systeme

$$\left( \sum_{k=1}^5 \Theta_k^a(\tilde{p}) B_D^H A_k^N B_D \right) E_c^N = \tilde{\Theta}_l^f(\tilde{p}) B_D^H f_l^N, \quad 1 \leq l \leq 4,$$

wobei wiederum

$$\tilde{p} = (p_1, p_2, p_3)^T = (\omega, \operatorname{Re}(\varepsilon_{r,E}), \operatorname{Im}(\varepsilon_{r,E}))^T$$

und

$$\begin{aligned} \Theta_1^f(p) &= \tilde{\Theta}_1^f(\tilde{p}) \sqrt{p_4} e^{-ip_8}, & \Theta_2^f(p) &= \tilde{\Theta}_2^f(\tilde{p}) \sqrt{p_5} e^{-ip_9}, \\ \Theta_3^f(p) &= \tilde{\Theta}_3^f(\tilde{p}) \sqrt{p_6} e^{-ip_{10}}, & \Theta_4^f(p) &= \tilde{\Theta}_4^f(\tilde{p}) \sqrt{p_7} e^{-ip_{11}}, \end{aligned}$$

gilt.

### Ergebnisse der RBM

Für das zweite Modell werden wir im Folgenden bestimmte Aussagen aus den Ergebnissen des ersten Modells bestätigen und betrachten die Auswirkung der Verwendung von vier Antennen auf die RB-Dimension. Aufgrund der größeren Anzahl von Parametern im zweiten Modell wurde für die Offline-Phase die Parametermenge  $\mathbb{P}_{400}^{train} \subset \mathbb{P}$  mit 400 zufälligen Parametern verwendet und die maximale Dimension des reduzierten Systems auf dreihundert festgesetzt. Zudem wurde nach der Reduktion der tatsächliche relative Fehler auf der Testparametermenge  $\mathbb{P}_{100}^{test}$  mit

$$\mathbb{P}_{100}^{test} \cap \mathbb{P}_{400}^{train} = \emptyset$$

betrachtet.

In Abbildung 4.30 sehen wir wieder das unterschiedliche Verhalten des Galerkin-Schätzers im Vergleich zum MinRes-Schätzer wie in Kapitel 4.3.1. Wie wir bereits in Abbildung 4.12 gesehen haben, liegt das unterschiedliche Verhalten der Fehlerschätzer an der Approximation der inf-sup-Konstante und nicht an der Norm des Riesz-Repräsentanten des Residuums, wie in Abbildung 4.29 zu sehen ist.

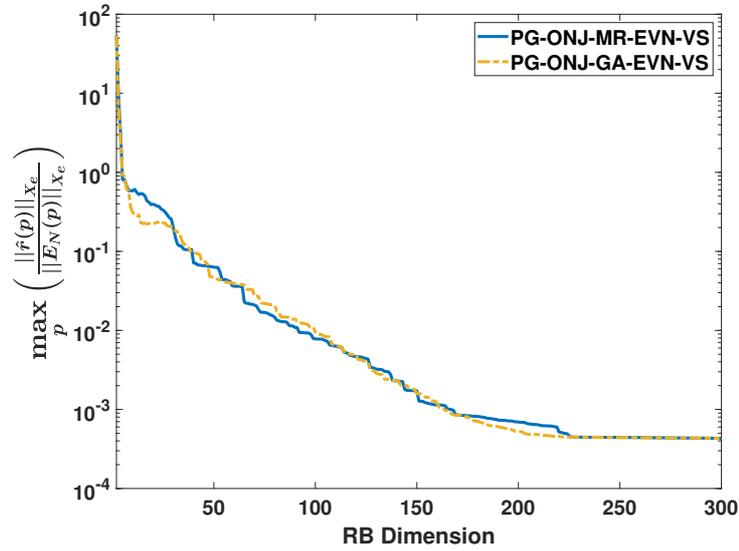


Abbildung 4.29: Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge  $\mathbb{P}_{400}^{train}$  während der Offline-Phase.

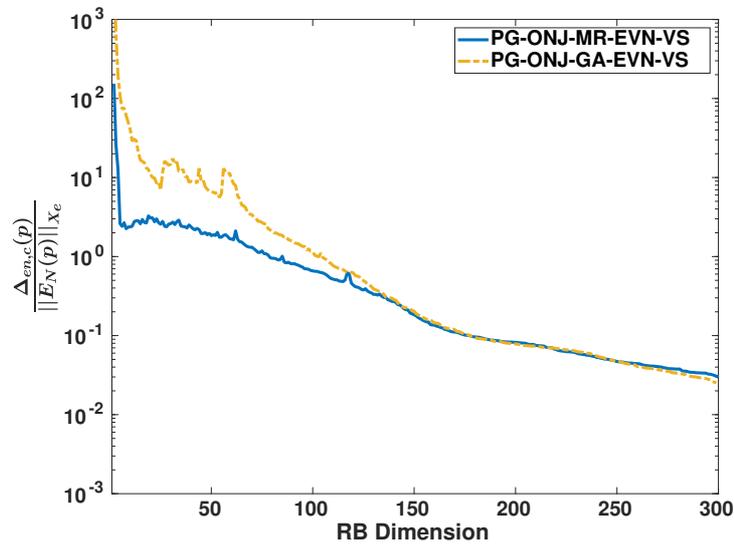


Abbildung 4.30: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

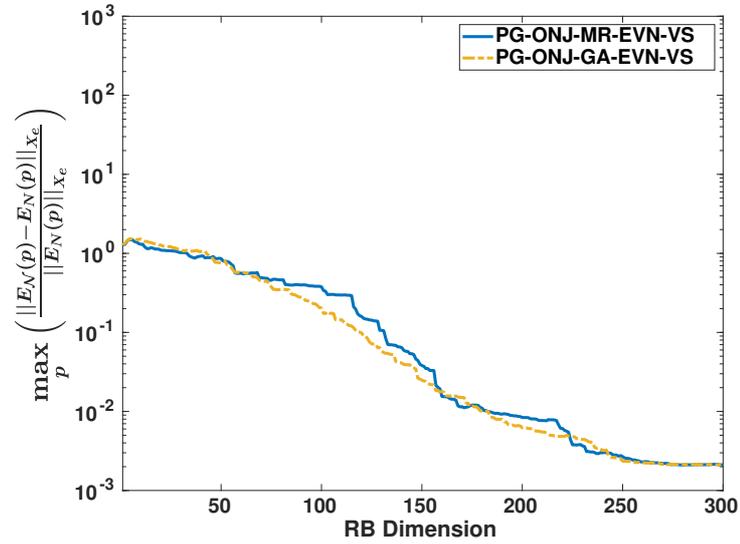


Abbildung 4.31: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

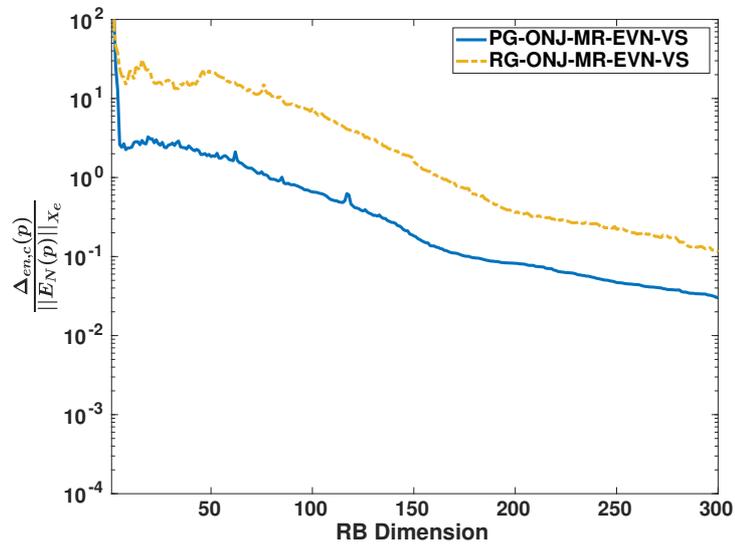
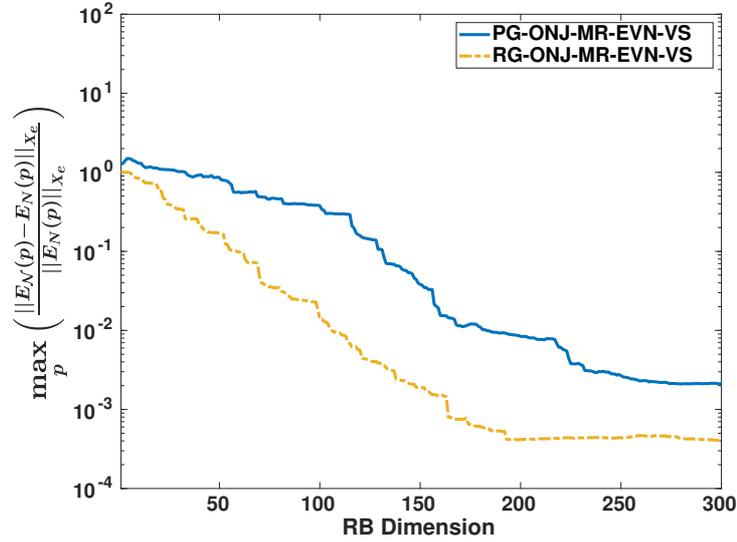
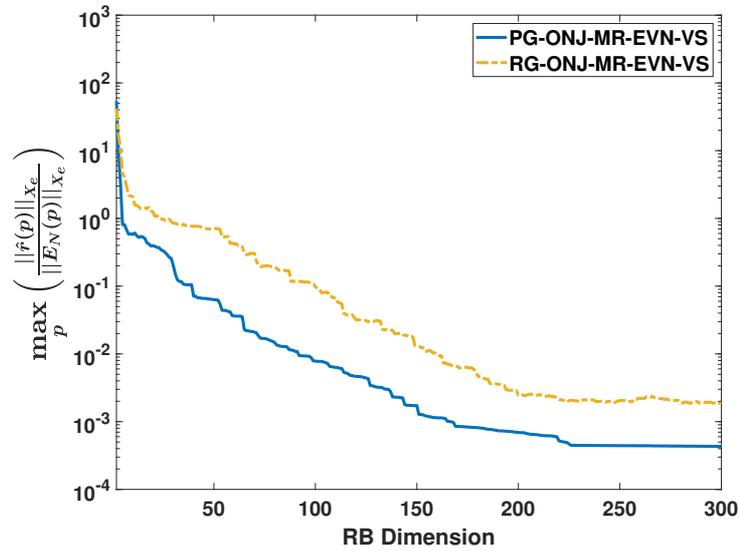


Abbildung 4.32: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

Abbildung 4.33: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .Abbildung 4.34: Maximum bezüglich der relativen Norm des Riesz-Repräsentanten und der reduzierten Lösung zu der Parametermenge  $\mathbb{P}_{400}^{train}$  während der Offline-Phase.

Wie in Kapitel 4.3.1 erhalten wir die gleichen Effekte bei der Betrachtung von Fehlerschätzer, siehe Abbildung 4.30, und relativen Fehler, siehe Abbildung 4.31, bei dem Vergleich von Ritz-Galerkin- und Petrov-Galerkin-Projektion.

Wie in Abbildung 4.34 zu sehen, ergibt sich der Unterschied im Fehlerschätzer aus der Berechnung des Riesz-Repräsentanten. Aber auch für das Modell zwei haben wir bei der Betrachtung des echten relativen Fehlers in Abbildung 4.33 den Effekt, dass die Ritz-Galerkin-Projektion die FEM-Lösung trotz schlechterem Fehlerschätzerwert besser approximiert. Nach Betrachtung der Abbildungen 4.29 bis 4.34 lassen sich die in Kapitel 4.3.1 getätigten Aussagen bestätigen. Wir widmen uns nun der Betrachtung der getrennten Systeme und dabei zunächst einmal dem Fehlerschätzerverhalten.

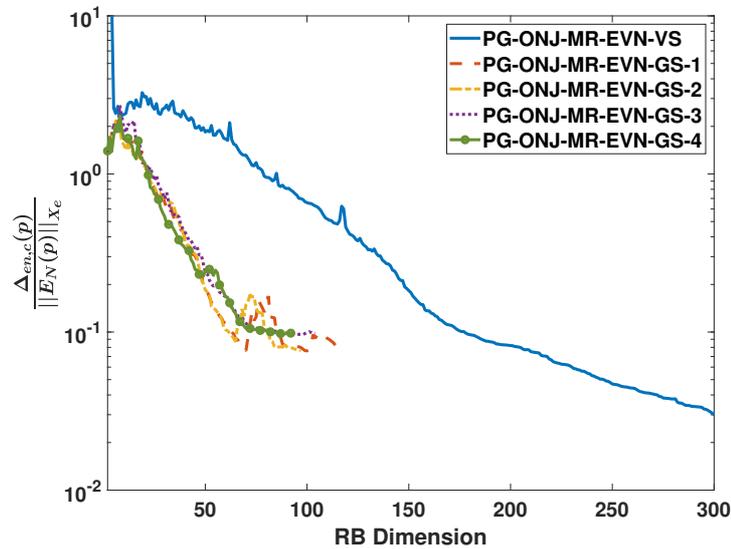


Abbildung 4.35: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

Zunächst einmal vergleichen wir die RB-Dimensionen in Tabelle 3 mit denen in Tabelle 1. Betrachten wir die Werte zur ersten angegebenen Schranke von 1.0, dann unterschreiten wir bei den getrennten Systemen die Schranke bei den gleichen RB-Dimensionen. Bei der zweiten Schranke hingegen benötigen wir für das System drei und vier eine RB-Dimension von mehr als achtzig, während die ersten beiden Systeme eine RB-Dimension von wenig mehr als sechzig benötigen. Betrachtet man hierzu die Positionierungen der Antennen und deren Nummerierung, siehe Abbildung 4.36, dann sehen wir, dass die Antennen auf der linken Seitenwand die geringere RB-Dimension benötigen. Zudem sind die RB-Dimensionen der ersten beiden Antennen des zweiten Modells und der ersten Antenne des ersten Modells, welche alle auf der linken Seitenwand positioniert sind,

RB-Variante	$\frac{\Delta_{en,c}(p)}{\ E_N(p)\ _{X_e}} \leq 1.0$	$\frac{\Delta_{en,c}(p)}{\ E_N(p)\ _{X_e}} \leq 0.1$
PG-ONJ-MR-EVN-VS	78	177
PG-ONJ-MR-EVN-GS-1	23	63
PG-ONJ-MR-EVN-GS-2	21	61
PG-ONJ-MR-EVN-GS-3	23	86
PG-ONJ-MR-EVN-GS-4	21	82

Tabelle 3: RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der Fehlerschätzer-Schranken zu Abbildung 4.35.

vergleichbar. Die beiden anderen Antennen auf der Hinterwand benötigen im Vergleich zu der Antenne im ersten Modell eine höhere RB-Dimension. Der Unterschied in der benötigten RB-Dimension hängt dabei von der Positionierung der Antennen und deren Auswirkungen zueinander ab.

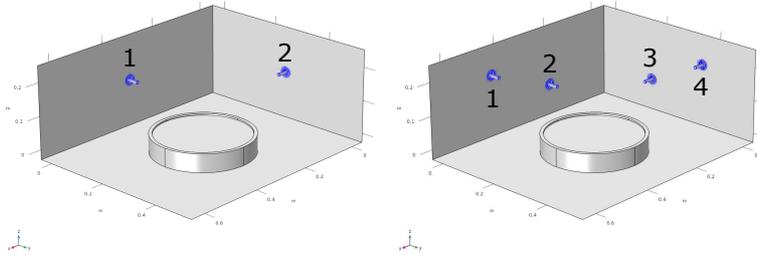


Abbildung 4.36: Position und Nummerierung der Antennen des ersten (links) und zweiten (rechts) Modells.

Wir wollen die eben getätigten Aussagen auch noch mit der Betrachtung der relativen Fehler bestätigen.

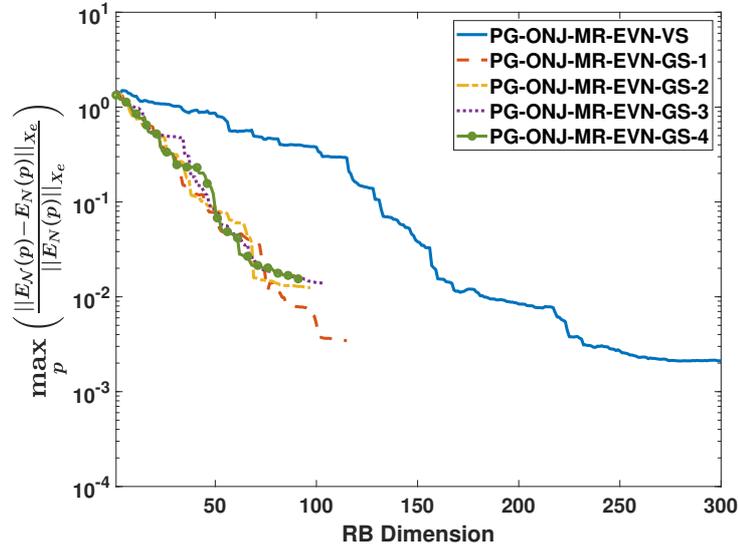


Abbildung 4.37: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

RB-Variante	$\frac{\ E_{\mathcal{N}}(p) - E_N(p)\ _{X_e}}{\ E_N(p)\ _{X_e}} \leq 0.1$	$\frac{\ E_{\mathcal{N}}(p) - E_N(p)\ _{X_e}}{\ E_N(p)\ _{X_e}} \leq 0.01$
PG-ONJ-MR-EVN-VS	132	183
PG-ONJ-MR-EVN-GS-1	45	83
PG-ONJ-MR-EVN-GS-2	45	nicht unterschritten
PG-ONJ-MR-EVN-GS-3	47	nicht unterschritten
PG-ONJ-MR-EVN-GS-4	50	nicht unterschritten

Tabelle 4: RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der echten maximalen relativen Fehler-Schranken zu Abbildung 4.37.

Auch in der Tabelle 4 unterschreiten wir den Fehler bei den getrennten Systemen zu annähernd gleichen RB-Dimensionen. Als weiteren Punkt können wir zudem wieder sehen, dass die getrennten Systeme in Summe mehr RB-Lösungen benötigen als das volle System. Damit kann auch die entsprechende Aussage aus Kapitel 4.3.1 bestätigt werden.

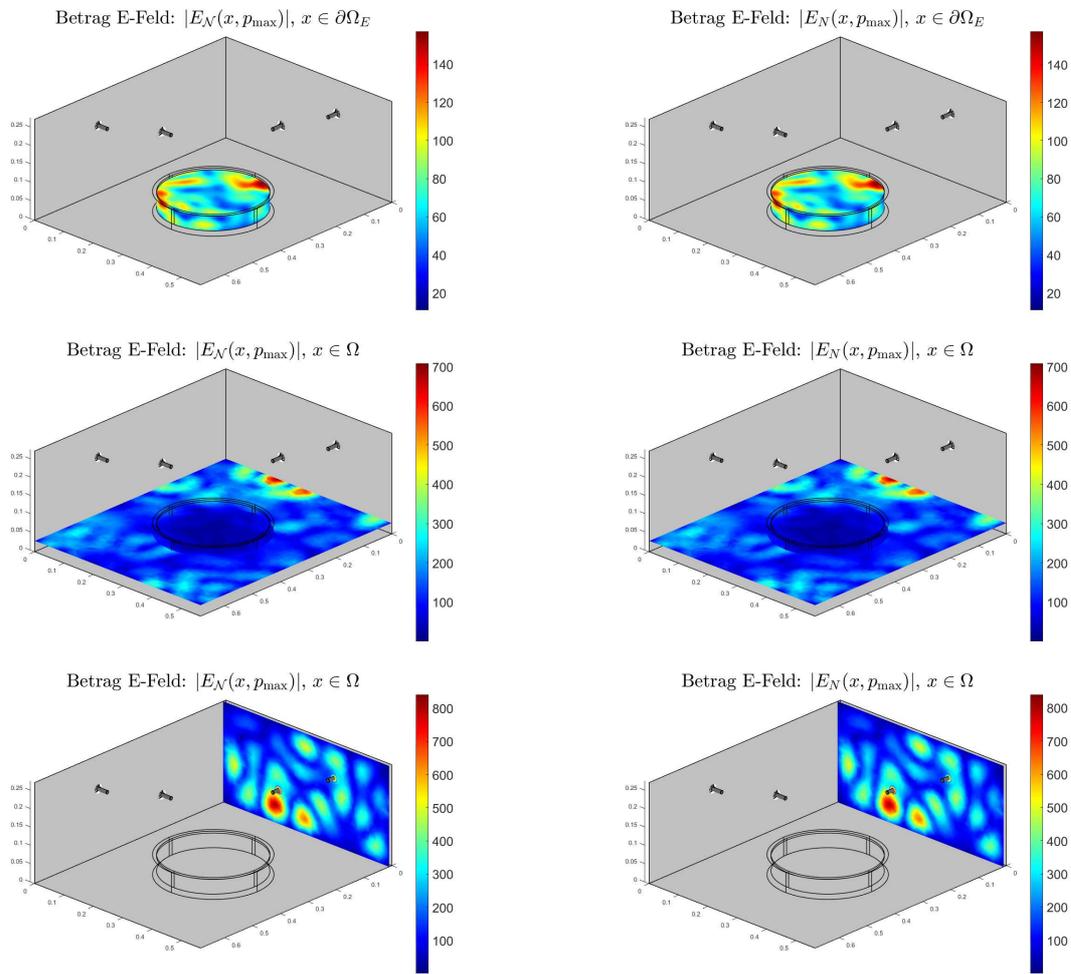


Abbildung 4.38: Komplexer Betrag der ortsabhängigen elektromagnetischen Feldvektoren zu  $p_{\max} \in \mathbb{P}_{100}^{test}$ .

#### 4. Reduzierte-Basis-Methode

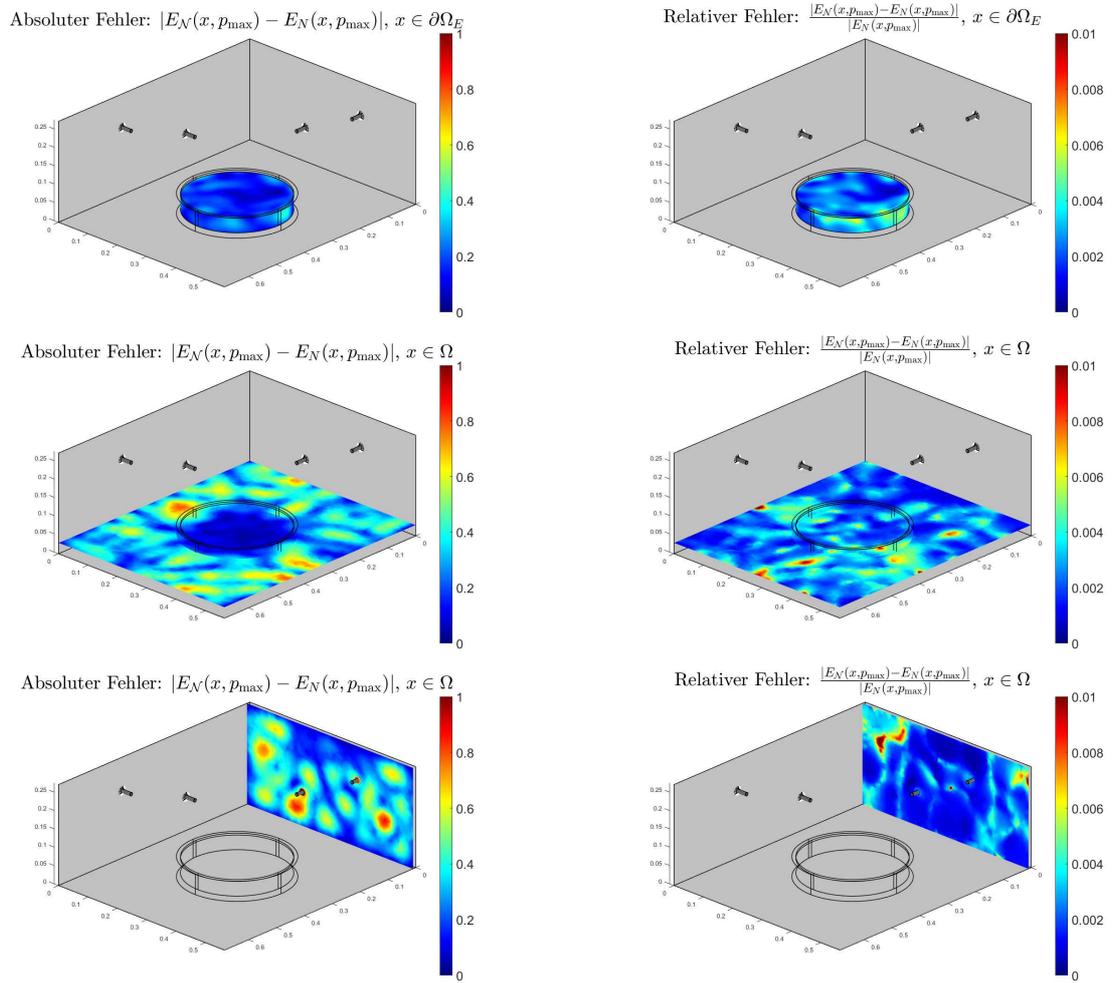


Abbildung 4.39: Absoluter und relativer Fehler der ortsabhängigen elektromagnetischen Feldvektoren zu  $p_{\max} \in \mathbb{P}_{100}^{\text{test}}$ .

Zum Abschluss widmen wir uns wieder den vektorwertigen ortsabhängigen relativen Fehlern. Hierzu betrachten wir wieder das argumentative Maximum  $p_{\max}$  des relativen Fehlers bezüglich der Testparametermenge  $\mathbb{P}_{100}^{\text{test}}$ . Zunächst vergleichen wir wieder die tatsächlichen elektromagnetischen Felder der FEM-Lösung  $E_N(x, p_{\max})$  und der RB-Lösung  $E_N(x, p_{\max})$  bei der RB-Dimension 300 und der RB-Variante PG-ONJ-MR-EVN-VS. Wie in Abbildung 4.25 zu erkennen ist, ist wieder kein Unterschied in den Feldbildern der FEM- und RB-Lösung erkennbar. Um den Unterschied besser darzustellen, geben wir den entstandenen absoluten und relativen vektoriellen Fehler an. Wie wir bereits in Kapitel 4.3.1 herausgearbeitet haben, entsteht auch hier der größte ortsabhängige relative Fehler an den Stellen mit geringer Feldintensität.

## 4.3.3. Modell III: Vier Antennen, zwei Objekte

## Modellbeschreibung

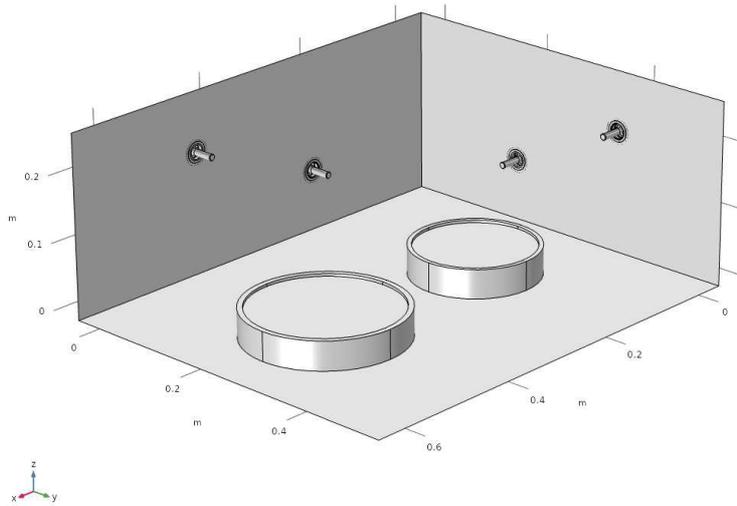


Abbildung 4.40: Geometrie des dritten Modells.

Das dritte Modell beinhaltet vier Monopol-Antennen und zwei zu erwärmende Objekte, dargestellt in den Abbildungen 4.40 und 4.41. Die Position der vier Antennen entspricht dabei den Positionen aus Modell II.

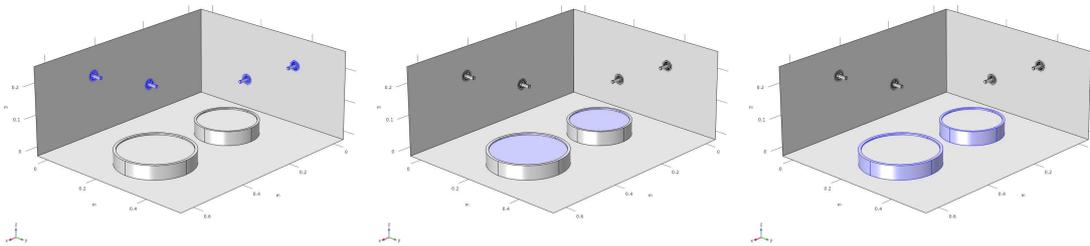


Abbildung 4.41: Geometrie des dritten Modells: vier Antennen (links), zwei zu erwärmende Objekte (mitte), zwei Schalen (rechts).

#### 4. Reduzierte-Basis-Methode

---

Die zu erwärmenden Objekte und deren Schalen in Modell III sind skalierte Versionen der entsprechenden Objekte aus Modell I und II. Das größere der beiden zu erwärmenden Objekte und dessen Schale haben die nach Abbildung 4.6 definierten Maße:

$$h_a = 4.4\text{cm}, \quad h_i = 3.2\text{cm}, \quad b = 0.8\text{cm}, \quad r = 11.2\text{cm}. \quad (4.100)$$

Für das kleinere Objekt und dessen Schale haben wir die Maße:

$$h_a = 4.4\text{cm}, \quad h_i = 3.2\text{cm}, \quad b = 0.8\text{cm}, \quad r = 8.96\text{cm}. \quad (4.101)$$

Damit ergeben sich für die zu erwärmenden Objekte die Volumina  $V_{E,1} = 1261.1\text{cm}^3$  und  $V_{E,2} = 807.1\text{cm}^3$ . Für die beiden Epoxidharzobjekte definieren wir zwei verschiedene Materialparameter

$$\text{Re}(\varepsilon_{r,E,1}(\omega)) \in [3, 3.7], \quad \text{Im}(\varepsilon_{r,E,1}(\omega)) \in [2, 3.5], \quad \mu_{r,E,1} = 1$$

für das größere Objekt  $\Omega_{E,1}$ , und

$$\text{Re}(\varepsilon_{r,E,2}(\omega)) \in [3, 3.7], \quad \text{Im}(\varepsilon_{r,E,2}(\omega)) \in [2, 3.5], \quad \mu_{r,E,2} = 1$$

für das kleinere Objekt  $\Omega_{E,2}$ . Ansonsten wählen wir die Gebiete und Materialparameter wie in Modell II. Daher definieren wir für die RBM den Parametervektor

$$p = (p_1, \dots, p_{13})^T = (\omega, \text{Re}(\varepsilon_{r,E,1}), \text{Im}(\varepsilon_{r,E,1}), \text{Re}(\varepsilon_{r,E,2}), \text{Im}(\varepsilon_{r,E,2}), \\ P_{in,1}, P_{in,2}, P_{in,3}, P_{in,4}, \theta_1, \theta_2, \theta_3, \theta_4)^T$$

und die zugehörige Parametermenge

$$\mathbb{P} = [4.8\pi, 5\pi] \times [3, 3.7] \times [2, 3.5] \times [3, 3.7] \times [2, 3.5] \times [0, 1]^4 \times [0, 2\pi]^4.$$

Es ergibt sich damit die folgende affine Darstellung der Sesquilinearform  $a$  der Variationsproblemformulierung  $[V_{D,\mathcal{N}}]$  für das Modell III:

$$\Theta_1^a(p) = 1, \quad a_1(E_{\mathcal{N}}, \varphi) = \int_{\Omega_L} \frac{1}{\mu_{r,L}} (\nabla \times E_{\mathcal{N}}) \cdot (\nabla \times \varphi) \, d\Omega \\ + \int_{\Omega_G} \frac{1}{\mu_{r,G}} (\nabla \times E_{\mathcal{N}}) \cdot (\nabla \times \varphi) \, d\Omega \\ + \sum_{k=1}^2 \int_{\Omega_{E,k}} \frac{1}{\mu_{r,E,k}} (\nabla \times E_{\mathcal{N}}) \cdot (\nabla \times \varphi) \, d\Omega,$$

$$\begin{aligned}
 \Theta_2^a(p) &= p_1^2, & a_2(E_{\mathcal{N}}, \varphi) &= - \int_{\Omega_L} (10^9)^2 \mu_0 \varepsilon_0 \varepsilon_{r,L} E_{\mathcal{N}} \cdot \varphi \, d\Omega \\
 & & & - \int_{\Omega_G} (10^9)^2 \mu_0 \varepsilon_0 \varepsilon_{r,G} E_{\mathcal{N}} \cdot \varphi \, d\Omega, \\
 \Theta_3^a(p) &= p_1^2 (p_2 + ip_3), & a_3(E_{\mathcal{N}}, \varphi) &= - \int_{\Omega_{E,1}} (10^9)^2 \mu_0 \varepsilon_0 E_{\mathcal{N}} \cdot \varphi \, d\Omega, \\
 \Theta_4^a(p) &= p_1^2 (p_4 + ip_5), & a_4(E_{\mathcal{N}}, \varphi) &= - \int_{\Omega_{E,2}} (10^9)^2 \mu_0 \varepsilon_0 E_{\mathcal{N}} \cdot \varphi \, d\Omega, \\
 \Theta_5^a(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,S}(p_1)}{\mu_0 \mu_{r,S}}}, & a_5(E_{\mathcal{N}}, \varphi) &= - \int_{\Gamma_{IMP}} 10^9 \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_6^a(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}}, & a_6(E_{\mathcal{N}}, \varphi) &= - \sum_{k=1}^4 \int_{\Gamma_{P,k}} 10^9 \mu_0 \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\varphi] \, d\Gamma.
 \end{aligned}$$

Mit derselben Herleitung der Normierungsfaktoren  $P_{m,k}$ ,  $1 \leq k \leq 4$ , wie im ersten Modell, erhalten wir die folgende affine Darstellung der rechten Seite  $f$  der Variationsproblemformulierung  $[\text{V}_{D,\mathcal{N}2}]$  für das Modell III:

$$\begin{aligned}
 \Theta_1^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_6}}{\sqrt{P_{m,1}}} e^{-ip_{10}}, & f_1(\varphi) &= -2 \int_{\Gamma_{P,1}} 10^9 \mu_0 \pi_t[E_{0,1}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_2^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_7}}{\sqrt{P_{m,2}}} e^{-ip_{11}}, & f_2(\varphi) &= -2 \int_{\Gamma_{P,2}} 10^9 \mu_0 \pi_t[E_{0,2}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_3^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_8}}{\sqrt{P_{m,3}}} e^{-ip_{12}}, & f_3(\varphi) &= -2 \int_{\Gamma_{P,3}} 10^9 \mu_0 \pi_t[E_{0,3}] \cdot \pi_t[\varphi] \, d\Gamma, \\
 \Theta_4^f(p) &= ip_1 \sqrt{\frac{\varepsilon_0 \varepsilon_{r,L}}{\mu_0 \mu_{r,L}}} \frac{\sqrt{p_9}}{\sqrt{P_{m,4}}} e^{-ip_{13}}, & f_4(\varphi) &= -2 \int_{\Gamma_{P,4}} 10^9 \mu_0 \pi_t[E_{0,4}] \cdot \pi_t[\varphi] \, d\Gamma.
 \end{aligned}$$

Aufgrund der Materialwahl Luft an der Antenne sind  $P_{m,1}$ ,  $P_{m,2}$ ,  $P_{m,3}$  und  $P_{m,4}$  von der Frequenz unabhängig. Mit der Lagrange-Formulierung in Abschnitt 3.4 der diskretisierten und evaluierten FEM-Probleme in COMSOL können wir mit der eben hergeleiteten

affinen Darstellung das Lagrange-System

$$\begin{pmatrix} \sum_{k=1}^6 \Theta_k^a(p) A_k^{\mathcal{N}} & D^H \\ D & 0 \end{pmatrix} \begin{pmatrix} E^{\mathcal{N}} \\ \lambda^{\mathcal{N}_{PEC}} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^4 \Theta_k^f(p) f_k^{\mathcal{N}} \\ 0 \end{pmatrix} \quad (4.102)$$

herleiten. Dabei ist die FEM-Dimension  $\mathcal{N} = 384975$  und  $\mathcal{N}_c = \mathcal{N} - \mathcal{N}_{PEC} = 378725$ . Wie in Modell I leiten wir mit (4.96) das zu reduzierende Gleichungssystem

$$\left( \sum_{k=1}^6 \Theta_k^a(p) B_D^H A_k^{\mathcal{N}} B_D \right) E_c^{\mathcal{N}_c} = \sum_{k=1}^4 \Theta_k^f(p) B_D^H f_k^{\mathcal{N}} \quad (4.103)$$

her und mit den entsprechend definierten Parameterfunktionen  $\tilde{\Theta}_l$  und der Parameterdefinition von  $\tilde{p}$  die vier getrennten Systeme

$$\left( \sum_{k=1}^6 \Theta_k^a(\tilde{p}) B_D^H A_k^{\mathcal{N}} B_D \right) E_c^{\mathcal{N}_c} = \tilde{\Theta}_l^f(\tilde{p}) B_D^H f_l^{\mathcal{N}}, \quad 1 \leq l \leq 4.$$

### Ergebnisse der RBM

Im Folgenden geben wir die Ergebnisse zu dem dritten Modell an, die die Aussagen aus den ersten beiden Modellen bestätigen. Da dabei aber keine neuen Erkenntnisse abgeleitet werden, geben wir lediglich deren Graphen ohne weitere Kommentare an. Dabei wurde in der Offline-Phase wiederum eine Parametermenge  $\mathbb{P}_{550}^{train} \subset \mathbb{P}$  mit 550 zufälligen Parametern verwendet und die maximale Dimension des reduzierten Systems auf 400 festgesetzt. Zudem wurde nach der Reduktion der tatsächliche relative Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$  mit

$$\mathbb{P}_{100}^{test} \cap \mathbb{P}_{550}^{train} = \emptyset$$

betrachtet.

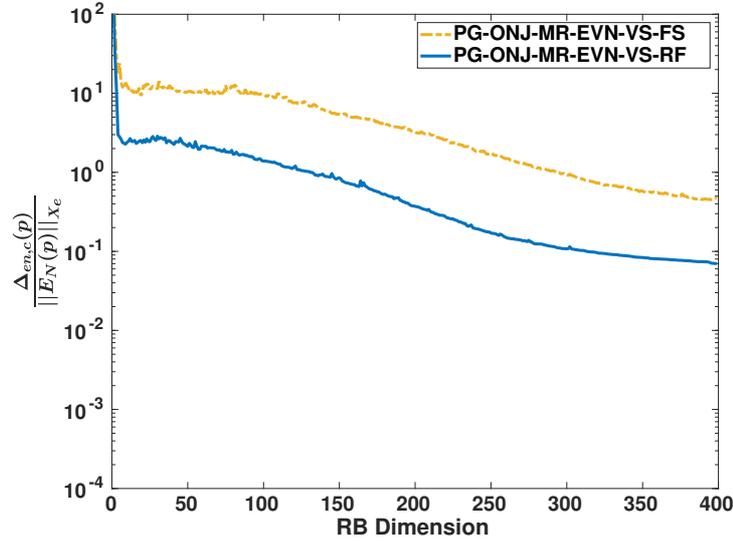


Abbildung 4.42: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

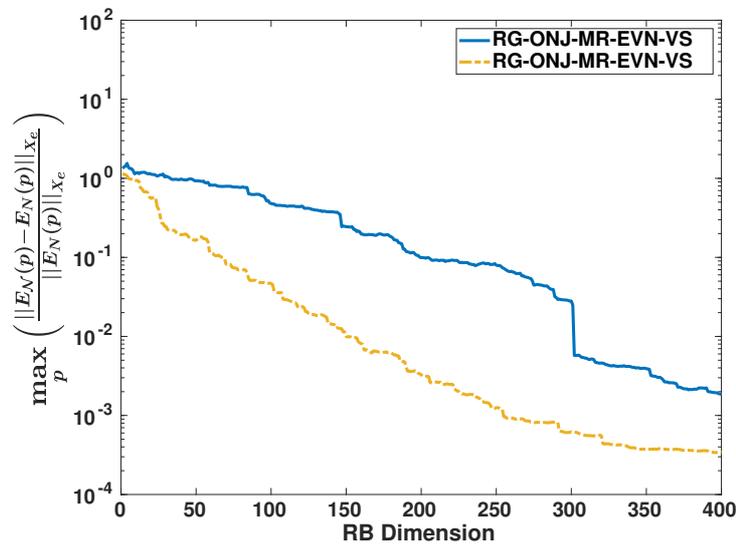


Abbildung 4.43: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{\text{test}}$ .

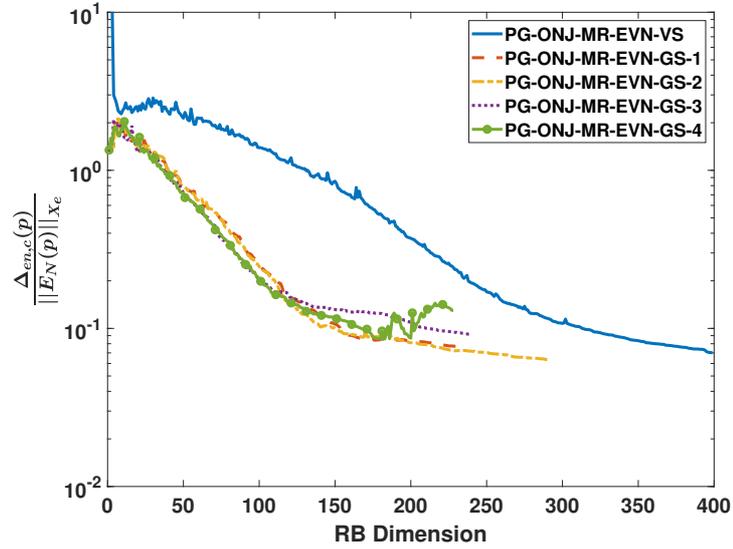
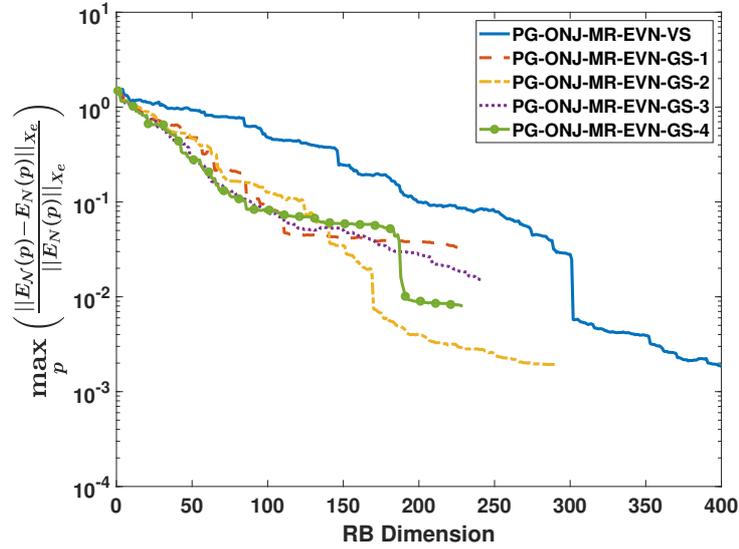


Abbildung 4.44: Maximale Werte des relativen Fehlerschätzers während der Offline-Phase.

RB-Variante	$\frac{\Delta_{en,c}(p)}{\ E_N(p)\ _{X_e}} \leq 1.0$	$\frac{\Delta_{en,c}(p)}{\ E_N(p)\ _{X_e}} \leq 0.1$
PG-ONJ-MR-EVN-VS	131	313
PG-ONJ-MR-EVN-GS-1	42	154
PG-ONJ-MR-EVN-GS-2	41	151
PG-ONJ-MR-EVN-GS-3	38	211
PG-ONJ-MR-EVN-GS-4	37	168

Tabelle 5: RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der Fehlerschätzer-Schranken zu Abbildung 4.44

Abbildung 4.45: Maximaler relativer Fehler zu der Testparametermenge  $\mathbb{P}_{100}^{test}$ .

RB-Variante	$\frac{\ E_{\mathcal{N}}(p) - E_N(p)\ _{X_e}}{\ E_N(p)\ _{X_e}} \leq 0.1$	$\frac{\ E_{\mathcal{N}}(p) - E_N(p)\ _{X_e}}{\ E_N(p)\ _{X_e}} \leq 0.01$
PG-ONJ-MR-EVN-VS	200	302
PG-ONJ-MR-EVN-GS-1	95	nicht unterschritten
PG-ONJ-MR-EVN-GS-2	124	170
PG-ONJ-MR-EVN-GS-3	91	nicht unterschritten
PG-ONJ-MR-EVN-GS-4	86	192

Tabelle 6: RB-Dimension der verschiedenen Varianten bei erstmaliger Unterschreitung der echten maximalen relativen Fehler-Schranken zu Abbildung 4.45.

##### 4.3.4. Zusammenfassung

Die in dieser Arbeit dargelegten Ergebnisse zeigen einige interessante Aspekte bei der Verwendung der RBM im Bereich der Hochfrequenztechnik auf. Der erste wichtige Aspekt bezieht sich dabei auf die Notwendigkeit der Orthonormalisierung der RB-Basismatrix  $\Xi_c$ , welche zu einer verbesserten Stabilität des Fehlerschätzers und des Reduktionsalgorithmus führt. Zudem konnte herausgearbeitet werden, dass das Konditionskriterium der reduzierten Systemmatrizen, die in direktem Zusammenhang mit Kondition der RB-Basismatrix steht, in der Offline-Phase zusammen mit dem Fehlerschätzer verwendet werden kann, um ein Abbruchkriterium für die maximale Anzahl an Basisvektoren zu kreieren. Ein weiterer wichtiger Aspekt bezieht sich auf die Approximation der inf-sup-Konstante zur Berechnung des Fehlerschätzers. Hierbei konnte festgestellt werden, dass unabhängig von der Wahl der Approximation die reduzierten Modelle alle eine annähernd gleiche Approximationsgüte aufweisen. Interessant dabei ist jedoch, dass trotz der besseren Approximation der inf-sup-Konstanten durch die Verwendung der Eigenvektorenerweiterung nicht nur keine besseren RB-Modelle generiert wurden, sondern dies auch noch zu einer früheren Terminierung der Offline-Phase aufgrund des Konditionskriteriums führte. Der nächste interessante Aspekt bezieht sich auf die Betrachtung der Ritz-Galerkin- und Petrov-Galerkin-Projektion. Während der Fehlerschätzer für die Ritz-Galerkin-Projektion in der Offline-Phase einen größeren Fehler vorraussagt, konnte eine bessere Approximation des RB-Modells im Vergleich zur Petrov-Galerkin-Projektion festgestellt werden. Dies zeigt gut auf, dass bessere Fehlerschätzer-Werte nicht als Indiz für eine bessere Approximationsgüte des reduzierten Systems verwendet werden kann. Der letzte Aspekt betrifft die Betrachtung der getrennten Systeme, denn in Summe werden mehr RB-Basisvektoren für die getrennten Systeme benötigt als bei den vollen Systemen.

---

## 5. Optimalsteuerung elektromagnetischer Felder

Elektromagnetische Wellen innerhalb einer geschlossenen Kavität führen zu stehenden Wellen. Für die Erwärmung von Epoxidharzen bedeutet dies, dass sich Zonen mit mehr Erwärmung und Zonen mit geringerer Erwärmung ergeben und diese verändern sich je nach Anregung des Systems. Um die elektromagnetischen Felder für den Erwärmungsprozess optimal zu steuern, leiten wir in Abschnitt 5.1 kontinuierliche Optimalsteuerungsprobleme her und diskutieren dazu die Existenz von deren Lösungen. In Abschnitt 5.2 erörtern wir die diskreten Optimalsteuerungsprobleme und betrachten den resultierenden Fehler zwischen der kontinuierlichen und der diskreten Lösung. Bei einer komplexen Geometrie wird ein hochdimensionaler FE-Raum benötigt und die Berechnung einer Lösung ist sehr rechenintensiv, was bei den dazugehörigen Optimalsteuerungsproblemen zu sehr hohen Rechenzeiten führt. Um diese Rechenzeit zu verringern beschreiben wir in Abschnitt 5.3 die reduzierten Optimalsteuerungsprobleme allgemein und diskutieren den Fehler zwischen der reduzierten und kontinuierlichen Lösung. Weiterhin betrachten wir auch kurz die getrennten Systeme, welche zusätzlich die Möglichkeit der Parallelisierung ermöglichen.

### 5.1. Kontinuierliche Optimalsteuerung

Bei der Betrachtung kontinuierlicher Optimalsteuerungsprobleme werden Zielfunktionale bezüglich der Steuerparametervektoren und der Lösung der kontinuierlichen Variationsformulierung  $[V_K]$  betrachtet, wobei die Variationsformulierung  $[V_K]$  als Nebenbedingung an das Optimierungsproblem gestellt wird. Innerhalb dieser Arbeit beschäftigen wir uns mit der Betrachtung von kohärenten Feldern, wobei kohärent bedeutet, dass an den einzelnen Antennen jeweils dieselbe Frequenz angeregt wird. Neben der Frequenz sind die weiteren Steuergrößen gegeben durch die eingespeiste Leistung und die Phase der einzelnen Antennen.

**Definition 5.1.** Der Steuerparametervektor  $p_S$  besteht aus der Frequenz  $\omega$ , den eingespeisten Leistungen  $P_{in,l}$  und den Phasen  $\theta_l$  der einzelnen Antennen, wobei  $1 \leq l \leq K_A$  und  $K_A$  die Anzahl der Antennen. Die Steuerparameter sind demnach

$$\omega \in \mathbb{R}, \quad [P_{in,k}]_{k=1}^{K_A} \in \mathbb{R}^{K_A}, \quad [\theta_k]_{k=1}^{K_A} \in \mathbb{R}^{K_A}.$$

Somit sind der Steuerparametervektor und der Steuerparameterraum gegeben durch

$$p_S = \left( \omega, \left( [P_{in,k}]_{k=1}^{K_A} \right)^T, \left( [\theta_k]_{k=1}^{K_A} \right)^T \right)^T \in \mathbb{P}_S = [\omega_{\min}, \omega_{\max}] \times [0, P_{\max}]^{K_A} \times [0, 2\pi]^{K_A}.$$

Nach Definition 5.1 ist die Parametermenge  $\mathbb{P}_S$  kompakt.

### 5.1.1. Stetige Abhängigkeit der Lösung bezüglich der Steuerparameter

Um die Stetigkeit der Lösung  $E \in X_e$  der variationellen Formulierung  $[V_K]$ , gegeben in Kapitel 3, bezüglich des Steuerparametervektors  $p_S$  zu beweisen, benötigen wir zunächst die uniforme Beschränktheit der Lösung bezüglich der Steuerparameter.

**Lemma 5.2.** *Es existiert eine von den Steuerparametern unabhängige uniforme Konstante  $C > 0$  mit*

$$\|E\|_{X_e} \leq C$$

für alle Lösungen  $E \in X_e$  von  $[V_K]$  zu dem Parametervektor  $p_S \in \mathbb{P}_S$ .

*Beweis.* Nach Satz 3.10 erfüllt jede Lösung  $E \in X_e$  von  $[V_K]$  zu dem Parametervektor  $p_S \in \mathbb{P}_S$  die Ungleichung

$$\|E\|_{X_e} \leq \tilde{C} \|\mathcal{F}\|_{X'_e},$$

wobei  $\mathcal{F}$  das Dualraumelement darstellt, welches gegeben ist als die rechte Seite von  $[V_K]$ . Setzen wir das zum Parametervektor  $p_S$  zugehörige Funktional ein und verwenden die Voraussetzung 2.10, so ergibt sich

$$\begin{aligned} \|E\|_{X_e} &\leq 2\tilde{C}\mu_0 \sum_{l=1}^{K_A} \sup_{\varphi \in X_e} \left( \frac{\left| \int_{\Gamma_{P,l}} i\omega Y(\omega) \frac{\sqrt{P_{m,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma \right|}{\|\varphi\|_{X_e}} \right) \\ &\leq C_1 \sum_{l=1}^{K_A} \sup_{\varphi \in X_e} \left( \frac{\int_{\Gamma_{P,l}} \frac{1}{\sqrt{|P_{m,l}(\omega)|}} |\pi_t[E_{0,l}] \cdot \pi_t[\varphi]| \, d\Gamma}{\|\varphi\|_{X_e}} \right) \end{aligned}$$

mit

$$C_1 = 2\tilde{C}\mu_0\omega_{\max} \sqrt{Y_{\max,\text{Re}}^2 + Y_{\max,\text{Im}}^2} \sqrt{P_{\max}}.$$

Um nun weitere Abschätzungen zu treffen, verwenden wir die in Satz 2.11 hergeleitete Darstellung des Leistungsnormierungsfaktors

$$P_{m,l}(\omega) = -\frac{1}{2} \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \int_{\Gamma_{P,l}} \text{Re} \left( \sqrt{\varepsilon_r(\omega)} \right) (\pi_t[E_{0,l}] \times \gamma_t[\pi_t[E_{0,l}]] \cdot n) \, dS. \quad (5.1)$$

Um den Normierungsfaktor  $P_{m,l}(\omega)$  abzuschätzen, betrachten wir die komplexe Wurzel der relativen Permittivität, für die bereits  $\text{Re}(\sqrt{\varepsilon_r(\omega)}) > 0$  in Abschnitt 2.5 gezeigt

wurde. Daraus erhalten wir mit Voraussetzung 2.10 die folgenden beiden Bedingungen

$$\begin{aligned} \left(\operatorname{Re}\left(\sqrt{\varepsilon_r(\omega)}\right)\right)^2 - \left(\operatorname{Im}\left(\sqrt{\varepsilon_r(\omega)}\right)\right)^2 &= \operatorname{Re}\left(\varepsilon_r(\omega)\right) > 0, \\ 2\operatorname{Re}\left(\sqrt{\varepsilon_r(\omega)}\right)\operatorname{Im}\left(\sqrt{\varepsilon_r(\omega)}\right) &= \operatorname{Im}\left(\varepsilon_r(\omega)\right) \geq 0 \end{aligned}$$

indem wir den Real- und Imaginärteil des Quadrates der komplexen Zahl  $\sqrt{\varepsilon_r(\omega)}$  mit dem Real- und Imaginärteil der komplexen Zahl  $\varepsilon_r(\omega)$  vergleichen. Verwenden wir die beiden Bedingungen, dann können wir die Darstellung

$$\operatorname{Re}\left(\sqrt{\varepsilon_r(\omega)}\right) = \sqrt{\frac{1}{2}\left(\operatorname{Re}\left(\varepsilon_r(\omega)\right) + \sqrt{\operatorname{Re}\left(\varepsilon_r(\omega)\right)^2 + \operatorname{Im}\left(\varepsilon_r(\omega)\right)^2}\right)}$$

herleiten. Die Korrektheit dieser Darstellung lässt sich leicht nachvollziehen, indem in die rechte Seite die beiden Bedingungsgleichungen eingesetzt und umgeformt werden. Wir erhalten somit

$$\operatorname{Re}\left(\sqrt{\varepsilon_r(\omega)}\right) \geq \sqrt{\varepsilon_{\min, \operatorname{Re}}}.$$

Mit der Vektoridentität (4.94) folgern wir aus der obigen Darstellung

$$P_{m,l}(\omega) \geq \frac{1}{2} \sqrt{\frac{\varepsilon_0 \varepsilon_{\min, \operatorname{Re}}}{\mu_0 \mu_{\max}}} \|\gamma_t[E_{0,l}]\|_{[L_t^2(\Gamma_{P,l})]}^2, \quad (5.2)$$

wobei dies wiederum bedeutet, dass  $P_{m,l}(\omega) > 0$ . Mit (5.2) ergibt sich somit

$$\frac{1}{\sqrt{|P_{m,l}(\omega)|}} \leq \frac{1}{\sqrt{\frac{1}{2} \sqrt{\frac{\varepsilon_0 \varepsilon_{\min, \operatorname{Re}}}{\mu_0 \mu_{\max}}} \|\gamma_t[E_{0,l}]\|_{[L_t^2(\Gamma_{P,l})]}^3}}.$$

Verwenden wir nun die Hölder-Ungleichung und

$$\|\pi_t[\varphi]\|_{[L_t^2(\Gamma_{P,l})]}^3 \leq \|\varphi\|_{X_e},$$

dann folgt

$$\|E\|_{X_e} \leq \frac{C_1}{\sqrt{\frac{1}{2} \sqrt{\frac{\varepsilon_0 \varepsilon_{\min, \operatorname{Re}}}{\mu_0 \mu_{\max}}}}} \sum_{l=1}^{K_A} \frac{\|\pi_t[E_{0,l}]\|_{[L_t^2(\Gamma_{P,l})]}^3}{\|\gamma_t[E_{0,l}]\|_{[L_t^2(\Gamma_{P,l})]}^3}.$$

□

Mit Lemma 5.2 können wir nun folgende Aussage über die Konvergenz von Lösungen  $E \in X_e$  bezüglich der Steuerparametervektoren  $p_S$  beweisen.

**Satz 5.3.** Sei  $p_{S,k} = (p_S)_k$  konvergente Folge in  $\mathbb{P}_S$  mit Grenzwert  $\hat{p}_S \in \mathbb{P}_S$ . Zu der Folge an Lösungen  $E_k$  von  $[V_K]$  zu den Parametervektoren  $p_{S,k}$  existiert eine konvergente Teilfolge mit Grenzwert  $\hat{E}$  und  $\hat{E}$  ist Lösung von  $[V_K]$  zu dem Parametervektor  $\hat{p}_S$ .

*Beweis.* Nach Satz 3.10 existieren für alle  $p_{S,k}$ ,  $k \in \mathbb{N}$ , und  $\hat{p}_S$  die eindeutigen Lösungen  $E_k, \hat{E} \in X_e$  zu dem Variationsproblem  $[V_K]$ . Wegen Lemma 5.2 ist die Folge  $(E_k)_k$  uniform durch eine von den Steuerparametern unabhängige Konstante  $C$  beschränkt. Da  $X_e$  ein Hilbertraum ist, existiert eine schwach konvergente Teilfolge  $(E_{k_l})_l$  mit Grenzwert  $E \in X_e$ . Zur Vereinfachung der Notation bezeichnen wir die Teilfolge  $(E_{k_l})_l$  wiederum mit  $(E_k)_k$ . Wir zeigen nun, dass aus der schwachen Konvergenz der Lösungen

$$|a(E_k, \varphi, p_{S,k}) - a(E, \varphi, \hat{p}_S)| \xrightarrow{k \rightarrow \infty} 0 \quad (5.3)$$

und

$$|f(\varphi, p_{S,k}) - f(\varphi, \hat{p}_S)| \xrightarrow{k \rightarrow \infty} 0 \quad (5.4)$$

folgt. Mit diesen beiden Konvergenzaussagen erhalten wir dann

$$a(E, \varphi, \hat{p}_S) = \lim_{k \rightarrow \infty} a(E_k, \varphi, p_{S,k}) = \lim_{k \rightarrow \infty} f(\varphi, p_{S,k}) = f(\varphi, \hat{p}_S), \quad (5.5)$$

was beweist, dass der Grenzwert  $E = \hat{E}$  und die Lösung von  $[V_K]$  zum Parametervektor  $\hat{p}_S$  ist. Zuerst beweisen wir die Aussage (5.3) und erhalten hierfür aufgrund der Linearität der Sesquilinearform und der Dreiecksungleichung

$$\begin{aligned} & |a(E_k, \varphi, p_{S,k}) - a(E, \varphi, \hat{p}_S)| \quad (5.6) \\ & \leq \frac{1}{\mu_{\min}} \left| \int_{\Omega} (\nabla \times (E_k - E)) \cdot (\nabla \times \varphi) \, d\Omega \right| \\ & \quad + \varepsilon_0 \mu_0 \left| \int_{\Omega} \omega_k^2 \varepsilon_r(\omega_k) E_k \cdot \varphi - \hat{\omega}^2 \varepsilon_r(\hat{\omega}) E \cdot \varphi \, d\Omega \right| \\ & \quad + \mu_0 \left| \int_{\Gamma_{IMP}} i\omega_k Y(\omega_k) \pi_t[E_k] \cdot \pi_t[\varphi] - i\hat{\omega} Y(\hat{\omega}) \pi_t[E] \cdot \pi_t[\varphi] \, d\Gamma \right| \\ & \quad + \mu_0 \sum_{l=1}^{K_A} \left| \int_{\Gamma_{P,l}} i\omega_k Y(\omega_k) \pi_t[E_k] \cdot \pi_t[\varphi] - i\hat{\omega} Y(\hat{\omega}) \pi_t[E] \cdot \pi_t[\varphi] \, d\Gamma \right| \\ & =: \frac{1}{\mu_{\min}} I_k + \varepsilon_0 \mu_0 II_k + \mu_0 III_k + \mu_0 IV_k. \end{aligned}$$

Im folgenden beweisen wir die Konvergenz der einzelnen Terme gegen Null. Als erstes betrachten wir den Term  $I_k$  und definieren hierfür den Operator

$$\begin{aligned} \mathcal{L}_\varphi : X_e &\rightarrow \mathbb{C}, \\ E &\mapsto \mathcal{L}_\varphi[E] = \int_{\Omega} (\nabla \times E) \cdot (\nabla \times \varphi) \, d\Omega. \end{aligned} \quad (5.7)$$

Aufgrund der Linearität des Integrals und des curl-Operators ist der Operator  $\mathcal{L}_\varphi$  linear. Die Stetigkeit des Operators folgt aus der Hölder-Ungleichung und der Definition der Norm (3.2) von  $X_e$ . Damit folgt  $\mathcal{L}_\varphi \in X_e'$  und somit

$$I_k = \left| \int_{\Omega} (\nabla \times (E_k - E)) \cdot (\nabla \times \varphi) \, d\Omega \right| = |\mathcal{L}_\varphi[E_k - E]| \xrightarrow{k \rightarrow \infty} 0,$$

aufgrund der schwachen Konvergenz der Folge  $E_k$ . Für den zweiten Term gilt mit der Dreiecksungleichung und der Hölder-Ungleichung

$$\begin{aligned} II_k &= \left| \int_{\Omega} \omega_k^2 \varepsilon_r(\omega_k) E_k \cdot \varphi - \hat{\omega}^2 \varepsilon_r(\hat{\omega}) E \cdot \varphi \, d\Omega \right| \\ &\leq \left| \int_{\Omega} (\omega_k^2 \varepsilon_r(\omega_k) - \hat{\omega}^2 \varepsilon_r(\hat{\omega})) E_k \cdot \varphi \, d\Omega \right| + \left| \int_{\Omega} \hat{\omega}^2 \varepsilon_r(\hat{\omega}) (E_k - E) \cdot \varphi \, d\Omega \right| \\ &\leq \sup_{x \in \Omega} |\omega_k^2 \varepsilon_r(\omega_k) - \hat{\omega}^2 \varepsilon_r(\hat{\omega})| \|E_k\|_{[L^2(\Omega)]^3} \|\varphi\|_{[L^2(\Omega)]^3} + \left| \hat{\mathcal{L}}_\varphi[E_k - E] \right| \end{aligned} \quad (5.8)$$

mit dem Operator

$$\begin{aligned} \hat{\mathcal{L}}_\varphi : X_e &\rightarrow \mathbb{C}, \\ E &\mapsto \hat{\mathcal{L}}_\varphi[E] = \int_{\Omega} \hat{\omega}^2 \varepsilon_r(\hat{\omega}) E \cdot \varphi \, d\Omega. \end{aligned}$$

Nach Voraussetzung 2.10 ist die relative frequenzabhängige Permittivität  $\varepsilon_r(\omega)$  stetig bezüglich  $\omega$ , weswegen

$$|\omega_k^2 \varepsilon_r(\omega_k) - \hat{\omega}^2 \varepsilon_r(\hat{\omega})| \xrightarrow{k \rightarrow \infty} 0$$

in jedem Punkt  $x \in \Omega$  gilt und somit konvergiert auch das Supremum gegen Null. Da die beiden Normen aufgrund von  $\varphi \in X_e$  und der uniformen Beschränktheit der Folge  $(E_k)_k$  beschränkt sind, konvergiert der erste Term in (5.8) gegen Null. Der Beweis für die Konvergenz des zweiten Terms gegen Null folgt analog zum Beweis des Termes  $I_k$ ,

womit sich

$$II_k \xrightarrow{k \rightarrow \infty} 0$$

ergibt. Analog zu dem Beweis der Konvergenz des Termes  $II_k$  erhalten wir über die Stetigkeit von  $Y(\omega)$  nach Satz 2.11 und der Stetigkeit der Tangentialoperatoren nach Satz 2.8 das auch die Terme  $III_k$  und  $IV_k$  gegen Null konvergieren. Damit folgt in (5.6) die zu zeigende Konvergenz (5.3).

Bevor wir (5.4) beweisen, betrachten wir die Abbildungen  $h_l : \mathbb{P}_S \rightarrow \mathbb{C}$ ,  $1 \leq l \leq K_A$ , definiert durch

$$h_l(p_S) = \omega \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l}.$$

Aufgrund der Stetigkeit der Wurzelfunktion und da wegen (5.2) gerade  $P_{m,l} > 0$  gilt, ist auch die Abbildung  $h_l(p_S)$  stetig bezüglich des Steuerparametervektors  $p_S$ . Wir erhalten somit für die rechte Seite

$$\begin{aligned} & |f(\varphi, p_{S,k}) - f(\varphi, \hat{p}_S)| \tag{5.9} \\ &= \left| 2i \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} (Y(\omega_k) h_l(p_{S,k}) - Y(\hat{\omega}) h_l(\hat{p}_S)) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] d\Gamma \right| \\ &\leq 2 \sum_{l=1}^{K_A} \hat{\omega} \frac{\sqrt{\hat{P}_{in,l}}}{\sqrt{|P_{m,l}(\hat{\omega})|}} \sup_{x \in \Gamma_{P,l}} |Y(\omega_k) - Y(\hat{\omega})| \|\pi_t[E_{0,l}]\|_{L_t^2(\Gamma_{P,l})} \|\varphi\|_{X_e} \\ &\quad + 2 \sum_{l=1}^{K_A} \sqrt{Y_{\max, \text{Re}}^2 + Y_{\max, \text{Im}}^2} |h_l(p_{S,k}) - h_l(\hat{p}_S)| \|\pi_t[E_{0,l}]\|_{L_t^2(\Gamma_{P,l})} \|\varphi\|_{X_e}. \end{aligned}$$

Mit der Stetigkeit von  $h_l(p_S)$  und  $Y(\omega)$ , nach Satz 2.11, sowie der Definition des Vektorfeldes  $E_{0,l}$  auf der beschränkten Oberfläche  $\Gamma_{P,l}$  und  $\varphi \in X_e$  folgt mit den gleichen Argumenten wie zuvor die Behauptung (5.4). Somit haben wir nun (5.5) und daher die schwache Konvergenz der Feldlösungen bewiesen.

Aufgrund der schwachen Konvergenz der Lösungen erhalten wir den schwachen Grenzwert  $E_k - \hat{E} \xrightarrow{k \rightarrow \infty} 0$ , für den wir nun noch die starke Konvergenz beweisen müssen. Um nun die starke Konvergenz und damit die stetige Abhängigkeit zu beweisen, betrachten wir die Gleichung

$$a(E_k, \varphi, p_{S,k}) - a(\hat{E}, \varphi, \hat{p}_S) = f(\varphi, p_{S,k}) - f(\varphi, \hat{p}_S), \tag{5.10}$$

zu der schwach konvergenten Folge an Lösungen  $E_k \xrightarrow{k \rightarrow \infty} \hat{E}$  und der stark konvergenten Folge der Steuerparametervektoren  $p_{S,k} \xrightarrow{k \rightarrow \infty} \hat{p}_S$ . Wegen der Linearität der Sesquilinearform, der Linearität der Spurooperatoren und mit dem Einfügen von entsprechenden

Zwischentermen erhalten wir für die linke Seite

$$\begin{aligned}
 & a(E_k, \varphi, p_{S,k}) - a(\hat{E}, \varphi, \hat{p}_S) \\
 &= \frac{1}{\mu_r} \int_{\Omega} (\nabla \times (E_k - \hat{E})) \cdot (\nabla \times \varphi) \, d\Omega - \int_{\Omega} \varepsilon_0 \mu_0 \hat{\omega}^2 \varepsilon_r(\hat{\omega}) (E_k - \hat{E}) \cdot \varphi \, d\Omega \\
 &\quad - \int_{\Gamma_{IMP}} i \mu_0 \hat{\omega} Y(\hat{\omega}) \pi_t[E_k - \hat{E}] \cdot \pi_t[\varphi] \, d\Gamma - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i \mu_0 \hat{\omega} Y(\hat{\omega}) \pi_t[E_k - \hat{E}] \cdot \pi_t[\varphi] \, d\Gamma \\
 &\quad - \int_{\Omega} \varepsilon_0 \mu_0 (\omega_k^2 \varepsilon_r(\omega_k) - \hat{\omega}^2 \varepsilon_r(\hat{\omega})) E_k \cdot \varphi \, d\Omega \\
 &\quad - \int_{\Gamma_{IMP}} i \mu_0 (\omega_k Y(\omega_k) - \hat{\omega} Y(\hat{\omega})) \pi_t[E_k] \cdot \pi_t[\varphi] \, d\Gamma \\
 &\quad - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i \mu_0 (\omega_k Y(\omega_k) - \hat{\omega} Y(\hat{\omega})) \pi_t[E_k] \cdot \pi_t[\varphi] \, d\Gamma \\
 &= a(E_k - \hat{E}, \varphi, \hat{p}_S) + R(\omega_k, \hat{\omega}, E_k).
 \end{aligned}$$

Betrachten wir nun den Grenzwert und verwenden den Satz von Lebesgue, dann erhalten wir aufgrund der Stetigkeit der Parameterfunktionen nach Satz 2.11 gerade

$$a(E_k - \hat{E}, \varphi, \hat{p}_S) + R(\omega_k, \hat{\omega}, E_k) \xrightarrow{k \rightarrow \infty} a(\lim_{k \rightarrow \infty} (E_k - \hat{E}), \varphi, \hat{p}_S), \quad (5.11)$$

da der Restterm  $R(\omega_k, \hat{\omega}, E_k)$  wegen der starken Konvergenz in den Parameterfunktionen entsprechend gegen Null konvergiert. Des Weiteren erhalten wir

$$\lim_{k \rightarrow \infty} (f(\varphi, p_{S,k}) - f(\varphi, \hat{p}_S)) = 0 \quad (5.12)$$

aufgrund der Stetigkeit der Parameterfunktionen der rechten Seite. Nach Grenzübergang in (5.10), indem wir (5.11) und (5.12) verwenden, ergibt sich die Variationsgleichung

$$a(\lim_{k \rightarrow \infty} (E_k - \hat{E}), \varphi, \hat{p}_S) = 0, \quad \varphi \in X_e.$$

Nach Satz 3.10 zur Existenz von Lösungen der Variationsprobleme  $[V_K]$  gilt, dass die eindeutige Lösung der obigen Variationsgleichung  $E = 0$  ist, also

$$\lim_{k \rightarrow \infty} (E_k - \hat{E}) = 0$$

in  $X_e$  gilt, und dies entspricht gerade der starken Konvergenz der Lösungen.  $\square$

**Korollar 5.4.** *Der Lösungsraum*

$$\mathbb{L}_S = \{E \in X_e : E \text{ Lösung von } [V_K] \text{ zu } p_S \in \mathbb{P}_S\}$$

des Variationsproblems  $[V_K]$  zur Steuerparametermenge  $\mathbb{P}_S$  ist kompakt.

*Beweis.* Betrachten wir eine Folge an Lösungen  $(E_k)_{k \in \mathbb{N}} \subset \mathbb{L}_S$ , dann sind diese nach Satz 3.10 eindeutig über die Variationsformulierung  $[V_K]$  und eine Folge  $(p_{S,k})_{k \in \mathbb{N}} \subset \mathbb{P}_S$  bestimmt. Da  $\mathbb{P}_S \subset \mathbb{R}^{1+2KA}$  eine kompakte Teilmenge ist, existiert eine konvergente Teilfolge  $(p_{S,k_n})_{n \in \mathbb{N}}$  mit Grenzwert  $\hat{p}_S$ . Nach Übergang auf eine weitere Teilfolge erhalten wir mit Satz 5.3 die Existenz einer konvergenten Teilfolge der Lösungsfolge  $(E_k)_{k \in \mathbb{N}}$  und damit die Kompaktheit.  $\square$

Wir haben somit den wichtigsten Satz für die Optimalsteuerung bewiesen und führen im nächsten Abschnitt zwei verschiedene Zielfunktionale ein und definieren die dazugehörigen Optimalsteuerungsprobleme.

### 5.1.2. Optimalsteuerungsprobleme

Die gleichmäßige Erwärmung eines Objektes wird durch ein konstantes elektromagnetisches Feld in einem Gebiet  $V \subset \Omega$  erzeugt. Dieses elektrische Feld bezeichnen wir im Folgenden mit  $E_V$ . Wir wollen nun Felder erzeugen, die dieses Feld bestmöglichst approximieren. Hierzu definieren wir das erste Optimalsteuerungsproblem:

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$  und  $E \in X_e$ , wobei  $E$  Lösung von  $[V_K]$  zum Parametervektor  $p_S$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[O_1]_K \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ E \in X_e}} Z_1(p_S, E) = \frac{1}{2} \frac{1}{|V|} \|E - E_V\|_{[L^2(V)]^3}^2$$

unter  $a(E, \varphi, p_S) = f(\varphi, p_S), \quad \varphi \in X_e.$

Eine direkte Folgerung aus der Definition des Zielfunktional  $Z_1$  ist deren Stetigkeit in Bezug auf ihre beiden Komponenten.

**Lemma 5.5.** *Das Zielfunktional  $Z_1 : \mathbb{P}_S \times X_e \rightarrow \mathbb{R}_{\geq 0}$  ist stetig in jeder Komponente.*

*Beweis.* Die Stetigkeit bezüglich des Steuerparametervektors  $p_S$  ist klar, da diese nicht im Zielfunktional auftritt. Die Stetigkeit in der anderen Komponente  $E$  folgt direkt aus der entsprechenden Dreiecksungleichung für die Norm, wenn die Stetigkeit der Wurzelfunktion ausgenutzt wird.  $\square$

Neben der eben angegebenen Feldsteuerung wollen wir die elektromagnetischen Felder für die Erwärmung von Objekten nutzen. Hierfür haben wir in Abschnitt 2.3 bereits eine Leistungserhaltungsgleichung (2.59) hergeleitet, in der das Integral wie in (2.60) die elektromagnetischen Verluste darstellt. Nach Voraussetzung 2.10 betrachten wir innerhalb dieser Arbeit Materialien ohne magnetische Verluste, also  $\text{Im}(\mu_r) = 0$ , weswegen die durch die elektromagnetischen Felder erzeugten Verluste gegeben sind als

$$P_{Q,e}(V) = \frac{1}{2} \int_V \omega \varepsilon_0 \text{Im}(\varepsilon_r(\omega)) |E|^2 d\Omega.$$

Wir verwenden diese Darstellung und beweisen, dass diese entsprechende Normeigenschaften besitzt. Damit eine entsprechende Definition sinnvoll ist, muss für das betrachtete Gebiet  $V$  der Imaginärteil der Permittivität  $\text{Im}(\varepsilon_r(\omega))$  ungleich Null sein.

**Definition 5.6.** Wir bezeichnen mit  $\|\cdot\|_{P(V,p_S)}$  die *Verlustleistung im Gebiet  $V \subset \Omega$* , die gegeben ist durch

$$\|E\|_{P(V,p_S)}^2 = \frac{1}{2} \int_V \omega \varepsilon_0 \text{Im}(\varepsilon_r(\omega)) |E|^2 d\Omega.$$

**Lemma 5.7.** *Sei  $V \subset \Omega$ . Gilt  $\text{Im}(\varepsilon_r(\omega)) > 0$  in jedem Punkt innerhalb des Gebiets  $V$ , dann ist die Verlustleistung  $\|\cdot\|_{P(V,p_S)}$  eine Norm auf  $[L^2(V)]^3$ .*

*Beweis.* Als erstes müssen wir zeigen, dass aus  $\|E\|_{P(V,p_S)} = 0$  bereits  $E = 0$  in  $[L^2(V)]^3$  folgt. Nach der Voraussetzung des Lemmas gilt

$$\text{Im}(\varepsilon_r(\omega)) > 0 \tag{5.13}$$

im kompletten Gebiet  $V$ . Damit ergibt sich

$$0 = \|E\|_{P(V,p_S)} = \frac{\sqrt{2}}{2} \sqrt{\omega \varepsilon_0} \|\sqrt{\text{Im}(\varepsilon_r(\omega))} E\|_{[L^2(V)]^3}.$$

Da (5.13) gilt, folgt sogleich  $E = 0$ , aufgrund der Normeigenschaft in  $[L^2(V)]^3$ . Die anderen beiden Normbedingungen folgen ebenfalls aus der Normeigenschaft von  $\|\cdot\|_{[L^2(V)]^3}$  und der obigen Darstellung.  $\square$

## 5. Optimalsteuerung elektromagnetischer Felder

---

Für das zweite Optimalsteuerungsproblem verwenden wir Definition 5.6 mit der Voraussetzung  $\text{Im}(\varepsilon_r(\omega)) > 0$  in  $V$  und erhalten:

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$  und  $E \in X_e$ , wobei  $E$  Lösung von  $[\text{V}_K]$  zum Parametervektor  $p_S$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[\text{O}_2]_K \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ E \in X_e}} Z_2(p_S, E) = \frac{1}{2} \frac{1}{|V|} \|E - E_V\|_{P(V, p_S)}^2$$

unter  $a(E, \varphi, p_S) = f(\varphi, p_S), \quad \varphi \in X_e.$

Vergleichen wir  $[\text{O}_2]_K$  mit  $[\text{O}_1]_K$ , dann unterscheiden sich die beiden Zielfunktionale um den parameterabhängigen Faktor  $\omega \text{Im}(\varepsilon_r(\omega))$ . Dies stellt gerade eine Skalierung bezüglich der elektromagnetischen Verluste dar. Aus der Definition des Zielfunktional  $Z_2$  erhalten wir wiederum die Stetigkeit in Bezug auf die beiden Komponenten.

**Lemma 5.8.** *Das Zielfunktional  $Z_2 : \mathbb{P}_S \times X_e \rightarrow \mathbb{R}_{\geq 0}$  ist stetig in jeder Komponente.*

*Beweis.* Die Stetigkeit bezüglich des Steuerparametervektors  $p_S$  folgt aus der Stetigkeit der Abbildung  $\varepsilon_r(\omega)$  nach Voraussetzung 2.10. Die Stetigkeit in der anderen Komponente folgt wiederum direkt aus der entsprechenden Dreiecksungleichung für die Norm, wenn die Stetigkeit der Wurzelfunktion ausgenutzt wird.  $\square$

Als nächstes geben wir noch die Definition optimaler Steuerungen und der dazugehörigen Zustände an.

**Definition 5.9.** Sei  $k = 1, 2$ ,  $(\hat{p}_S, \hat{E}) \in \mathbb{O} \subset \mathbb{P}_S \times X_e$  ein Element der zulässigen Menge  $\mathbb{O}$  zu  $[\text{O}_k]_K$ . Ist das Tupel  $(\hat{p}_S, \hat{E})$  lokal optimal zu  $E_V$ , also

$$Z_k(p_S, E) \geq Z_k(\hat{p}_S, \hat{E}), \quad (p_S, E) \in U_\epsilon(\hat{p}_S, \hat{E}) \subset \mathbb{O},$$

mit

$$U_\epsilon(\hat{p}_S, \hat{E}) = \left\{ (p_S, E) \in \mathbb{O} : d\left((\hat{p}_S, \hat{E}), (p_S, E)\right) = \|\hat{p}_S - p_S\|_{\mathbb{R}^{1+2K_A}} + \|\hat{E} - E\|_{X_e} \leq \epsilon \right\},$$

dann heißt  $\hat{p}_S$  *lokal optimale Steuerung* und der dazugehörige Zustand  $\hat{E}$  *lokal optimaler Zustand*. Ist das Tupel  $(\hat{p}_S, \hat{E})$  global optimal zu  $E_V$ , also

$$Z_k(p_S, E) \geq Z_k(\hat{p}_S, \hat{E}), \quad (p_S, E) \in \mathbb{O},$$

dann heißt  $\hat{p}_S$  *global optimale Steuerung* und der dazugehörige Zustand  $\hat{E}$  *global optimaler Zustand*.

Für die zulässige Menge  $\mathbb{O}$  definieren wir noch

$$\mathbb{O} = \mathbb{O}_S \times \mathbb{O}_e, \quad \mathbb{O}_S \subset \mathbb{P}_S, \quad \mathbb{O}_e \subset X_e, \quad (5.14)$$

mit der Bedingung  $P_{in,l} > 0$ ,  $1 \leq l \leq K_A$ , an die Leistungen zu dem Parameter  $p_S \in \mathbb{O}_S$ . Mit Lemma 5.5 und Lemma 5.8 können wir nun die Existenz globaler optimaler Steuerungen zu den Optimalsteuerungsproblemen  $[\mathbb{O}_1]_K$  und  $[\mathbb{O}_2]_K$  beweisen.

**Satz 5.10.** *Es existiert mindestens ein Lösungstupel  $(\hat{p}_S, \hat{E}) \in \mathbb{O}$  zu dem Optimalsteuerungsproblem  $[\mathbb{O}_k]_K$ ,  $k = 1, 2$ .*

*Beweis.* Nach Satz 5.3, bzw. Korollar 5.4, erhalten wir für jede Folge an Lösungstupel  $(p_{S,n}, E_n)_{n \in \mathbb{N}} \subset \mathbb{O}$  die Existenz einer konvergenten Folge an Lösungstupel in der zulässigen Menge  $\mathbb{O}$ . Damit ist  $\mathbb{O}$  kompakt. Die Zielfunktionale  $Z_k$ ,  $k = 1, 2$ , sind nach Lemma 5.5 und Lemma 5.8 stetig. Da stetige Funktionen auf kompakten Mengen nach dem Satz von Weierstraß immer ein Minimum annehmen, ist dies gleichbedeutend mit der Existenz einer globalen optimalen Steuerung nach Definition 5.9.  $\square$

In der Optimalsteuerung wird die Lösung der Variationsformulierung, welche als Nebenbedingung fungiert, auch Zustand genannt. Dieser Zustand ist daher indirekt abhängig von der Steuerung, weshalb ein entsprechender Steuerungs-Zustands-Operator definiert wird.

**Definition 5.11.** Der Operator

$$\begin{aligned} \mathcal{E} : \mathbb{P}_S &\rightarrow X_e, \\ p_S &\mapsto \mathcal{E}[p_S], \end{aligned}$$

gegeben als Lösung des Variationsproblems  $[V_K]$  zum Parametervektor  $p_S \in \mathbb{P}_S$ , heißt *Steuerungs-Zustands-Operator*. Dieser bildet die Parametermenge  $\mathbb{P}_S$  eindeutig auf die Lösungsmenge  $\mathbb{L}_s \subset X_e$  ab. Sei  $\mathcal{T} : X_e \rightarrow [L^2(V)]^3$  der lineare Einschränkungoperator auf  $V$ . Der Operator

$$\begin{aligned} \mathcal{S} : \mathbb{P}_S &\rightarrow [L^2(V)]^3, \\ p_S &\mapsto \mathcal{S}[p_S] = \mathcal{T}[\mathcal{E}[p_S]], \end{aligned}$$

bildet den Zustand  $\mathcal{E}[p_S]$  auf das für das Zielfunktional entsprechende Teilgebiet  $V$  ab.

Die Eindeutigkeit des Operators  $\mathcal{E}$  in Definition 5.11 ergibt sich wegen der Lösungseigenschaft in Satz 3.10. Aufgrund der nichtlinearen Abhängigkeit des Variationsproblems  $[V_K]$  von den Steuerparametervektoren  $p_S$  sind der Steuerungs-Zustands-Operator  $\mathcal{E}$  und der Operator  $\mathcal{S}$  nichtlinear, jedoch ist  $\mathcal{E}$  wegen Satz 5.3 stetig. Die Stetigkeit von  $\mathcal{S}$  folgt dabei aus der Stetigkeit der einzelnen Operatoren  $\mathcal{E}$  und  $\mathcal{T}$ . Zur Vereinfachung

der Notation werden wir weiterführend den Einschränkungoperator  $\mathcal{T}$  und damit auch  $\mathcal{S}$  nicht verwenden, da deren Verwendung aus dem Kontext heraus klar ist. Mit Definition 5.11 können wir die beiden Optimierungsprobleme  $[O_k]_{\mathbb{K}}$ ,  $k = 1, 2$ , entsprechend umformulieren zu

$$\min_{p_S \in \mathbb{P}_S} \mathcal{Z}_1[p_S] = \frac{1}{2} \frac{1}{|V|} \|\mathcal{E}[p_S] - E_V\|_{[L^2(V)]^3}^2 \quad (5.15)$$

und

$$\min_{p_S \in \mathbb{P}_S} \mathcal{Z}_2[p_S] = \frac{1}{2} \frac{1}{|V|} \|\mathcal{E}[p_S] - E_V\|_{P(V, p_S)}^2. \quad (5.16)$$

Die beiden Zielfunktionale  $\mathcal{Z}_1$  und  $\mathcal{Z}_2$  werden auf den Steuerungsparameter  $p_S$  reduzierte Zielfunktionale genannt. Die Stetigkeit bezüglich der Steuerparameter folgt dabei direkt aus der Stetigkeit des Operators  $\mathcal{E}$ , Lemma 5.5 und Lemma 5.8.

### 5.1.3. Erste Variation und Gâteaux-Differenzierbarkeit

Nachdem wir im vorherigen Abschnitt die beiden Optimalsteuerungsprobleme beschrieben und die Existenz von Lösungen bewiesen haben, betrachten wir innerhalb dieses Abschnitts entsprechende Ableitungen bezüglich der Steuerung. Damit können dann auch gradientenbasierende Optimierungsalgorithmen verwendet werden. Hierzu führen wir als erstes folgende Definitionen ein.

**Definition 5.12.** Seien  $X, Y$  Banachräume,  $\mathcal{X} \subset X$  offen und  $F : \mathcal{X} \rightarrow Y$ . Existiert zu gegebenen Elementen  $u \in \mathcal{X}$ ,  $v \in X$  der Grenzwert

$$\delta F[u, v] = \lim_{h \rightarrow 0} \frac{F[u + hv] - F[u]}{h} \in Y, \quad (5.17)$$

so heißt dieser *Richtungsableitung* von  $F$  an der Stelle  $u$  in Richtung  $v$ . Existiert dieser Grenzwert für alle  $v \in X$ , dann heißt die Abbildung  $v \mapsto \delta F[u, v]$  *erste Variation* von  $F$  an der Stelle  $u$  in Richtung  $v$ .

**Definition 5.13.** Seien  $X, Y$  Banachräume. Existieren die erste Variation  $\delta F[u, v]$  an der Stelle  $u$  und ein linearer stetiger Operator  $A : X \rightarrow Y$ , sodass

$$\delta F[u, v] = Av, \quad v \in X, \quad (5.18)$$

dann heißt  $F$  an der Stelle  $u$  *Gâteaux-differenzierbar* und  $A$  *Gâteaux-Ableitung* von  $F$  an der Stelle  $u$ . Man schreibt dann auch  $A = \nabla F[u]$ .

*Bemerkung 5.14.* Sind  $X, Y$  endlichdimensional, dann entspricht  $\delta F[u, v]$  der klassischen Richtungsableitung.

Um die erste Variation und die Gâteaux-Ableitung der Optimalsteuerungsprobleme  $[O_k]_{\mathbb{K}}$ ,  $k = 1, 2$ , in Bezug auf den Steuerparametervektor  $p_S$  herzuleiten, benötigen wir zusätzlich zu Voraussetzung 2.10 die folgenden Voraussetzungen.

**Voraussetzung 5.15.** Für die Abbildung  $\omega \mapsto \varepsilon_r(\omega)$  existiert die erste Variation  $\delta\varepsilon_r[\omega, \tilde{\omega}]$  für alle  $\omega \in [\omega_{\min}, \omega_{\max}]$ ,  $\tilde{\omega} \in \mathbb{R}$  und in jedem Punkt aus  $\bar{\Omega}$ .

**Voraussetzung 5.16.** Für die Abbildung  $\omega \mapsto \varepsilon_r(\omega)$  existiert die Gâteaux-Ableitung  $\nabla\varepsilon_r[\omega]$  für alle  $\omega \in [\omega_{\min}, \omega_{\max}]$  und in jedem Punkt aus  $\bar{\Omega}$ .

Hierbei ist Voraussetzung 5.16 eine stärkere Bedingung als Voraussetzung 5.15 und kann als Spezialfall betrachtet werden. Wir erhalten damit folgende Aussagen für die anderen frequenzabhängigen Parameterfunktionen.

**Lemma 5.17.** *Gilt Voraussetzung 5.15, dann existiert die erste Variation zu den frequenzabhängigen Abbildungen  $\omega \mapsto Y(\omega)$  und  $\omega \mapsto P_{m,l}(\omega)$ ,  $1 \leq l \leq K_A$ , an der Stelle  $\omega \in [\omega_{\min}, \omega_{\max}]$  in Richtung  $\tilde{\omega} \in \mathbb{R}$ . Gilt zudem Voraussetzung 5.16, dann sind die Abbildungen sogar Gâteaux-differenzierbar, also*

$$\delta Y[\omega, \tilde{\omega}] = \nabla Y[\omega] \tilde{\omega}, \quad \delta P_{m,l}[\omega, \tilde{\omega}] = \nabla P_{m,l}[\omega] \tilde{\omega}, \quad 1 \leq l \leq K_A.$$

*Beweis.* Wir betrachten die erste Variation der Abbildung  $Y$  in Zusammenhang mit ihrer Definition und erhalten

$$\delta Y[\omega, \tilde{\omega}] = \lim_{h \rightarrow \infty} \frac{Y(\omega + h\tilde{\omega}) - Y(\omega)}{h} = \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \lim_{h \rightarrow \infty} \frac{1}{h} \left( \sqrt{\varepsilon_r(\omega + h\tilde{\omega})} - \sqrt{\varepsilon_r(\omega)} \right).$$

Da die Wurzelfunktion reell differenzierbar ist, folgt

$$\delta Y[\omega, \tilde{\omega}] = \frac{1}{2} \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \frac{1}{\sqrt{\varepsilon_r(\omega)}} \delta\varepsilon_r[\omega, \tilde{\omega}].$$

Für den frequenzabhängigen Normierungsfaktor  $P_{m,l}(\omega)$  betrachten wir die Darstellung (5.1) und wenden wiederum die Vektoridentität (4.94) an, um die Darstellung

$$P_{m,l}(\omega) = \frac{1}{2} \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \int_{\Gamma_{P,l}} \operatorname{Re} \left( \sqrt{\varepsilon_r(\omega)} \right) |\gamma_t[E_{0,l}]|^2 d\Gamma \quad (5.19)$$

zu erhalten. Aus der Summationseigenschaft und der Linearitätseigenschaft der Realteil-Funktion  $\operatorname{Re}$  bei reellen Faktoren erhalten wir mit (5.19) die erste Variation

$$\delta P_{m,l}[\omega, \tilde{\omega}] = \frac{1}{2} \sqrt{\frac{\varepsilon_0}{\mu_0 \mu_r}} \int_{\Gamma_{P,l}} \operatorname{Re} \left( \frac{1}{2} \frac{1}{\sqrt{\varepsilon_r(\omega)}} \delta\varepsilon_r[\omega, \tilde{\omega}] \right) |\gamma_t[E_{0,l}]|^2 d\Gamma.$$

In Kombination mit Voraussetzung 5.15 erhalten wir somit die erste Aussage. Gilt zudem noch Voraussetzung 5.16, dann erhalten wir direkt aus den Darstellungen die Existenz der Gâteaux-Ableitungen von  $Y$  und  $P_{m,l}$ .  $\square$

## 5. Optimalsteuerung elektromagnetischer Felder

---

*Bemerkung 5.18.* Für den Leistungsfaktor  $P_{m,l}(\omega)$  in (5.19) gilt  $P_{m,l}(\omega) > 0$  wegen (5.2), weswegen der Betrag nicht notwendig ist und somit bei der Differenzierung kein Problem darstellt.

Bevor wir nun Ableitungen des Zielfunktionals bezüglich der Steuerparametervektoren  $p_S$  betrachten, leiten wir die erste Variation des Steuerungs-Zustands-Operators  $\mathcal{E}$  her.

**Satz 5.19.** *Gilt Voraussetzung 5.15, dann ist die erste Variation  $\delta\mathcal{E}[p_S, \tilde{p}_S] \in X_e$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  die Lösung des folgenden Variationsproblems*

$$a(\delta\mathcal{E}[p_S, \tilde{p}_S], \varphi, p_S) = f_2(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S), \quad \varphi \in X_e, \quad (5.20)$$

mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\begin{aligned} & f_2(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S) \quad (5.21) \\ &= \int_{\Omega} \varepsilon_0 \mu_0 (\omega^2 \delta \varepsilon_r[\omega, \tilde{\omega}] + 2\omega \tilde{\omega} \varepsilon_r(\omega)) \mathcal{E}[p_S] \cdot \varphi \, d\Omega \\ &+ \int_{\Gamma_{IMP}} i\mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ &+ \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ &- 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 \sqrt{P_{in,l}} e^{-i\theta_l} \left( \frac{\omega \delta Y[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)}} + \frac{\tilde{\omega} Y(\omega)}{\sqrt{P_{m,l}(\omega)}} \right. \\ &\quad \left. - \frac{\omega}{2} Y(\omega) \frac{\delta P_{m,l}[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)^3}} \right) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma \\ &- \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 \omega Y(\omega) \frac{\tilde{P}_{in,l}}{\sqrt{P_{in,l}} \sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma \\ &- 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \mu_0 \omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \tilde{\theta}_l \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma. \end{aligned}$$

*Beweis.* Wir betrachten zunächst die beiden Zustände  $\mathcal{E}[p_S]$  und  $\mathcal{E}[p_S + h\tilde{p}_S]$ , wobei  $h$  klein genug und  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  so gewählt wird, dass  $p_S + h\tilde{p}_S \in \mathbb{O}_S$  gilt. Damit folgt aufgrund der Definition 5.11 die Gleichung

$$a(\mathcal{E}[p_S + h\tilde{p}_S], \varphi, p_S + h\tilde{p}_S) - a(\mathcal{E}[p_S], \varphi, p_S) = f(\varphi, p_S + h\tilde{p}_S) - f(\varphi, p_S), \quad (5.22)$$

wobei  $a$  und  $f$  durch  $[V_K]$  gegeben sind. Als erstes betrachten wir die linke Seite dieser Gleichung, fügen einige Zwischenterme ein und nutzen die entsprechenden Linearitäten der Operatoren aus, um

$$\begin{aligned}
 & a(\mathcal{E}[p_S + h\tilde{p}_S], \varphi, p_S + h\tilde{p}_S) - a(\mathcal{E}[p_S], \varphi, p_S) \\
 &= \int_{\Omega} \frac{1}{\mu_r} (\nabla \times (\mathcal{E}[p_S + h\tilde{p}_S] - \mathcal{E}[p_S])) \cdot (\nabla \times \varphi) \, d\Omega \\
 &\quad - \int_{\Omega} \varepsilon_0 \mu_0 ((\omega + h\tilde{\omega})^2 \varepsilon_r(\omega + h\tilde{\omega}) - \omega^2 \varepsilon_r(\omega)) \mathcal{E}[p_S + h\tilde{p}_S] \cdot \varphi \, d\Omega \\
 &\quad - \int_{\Omega} \varepsilon_0 \mu_0 \omega^2 \varepsilon_r(\omega) (\mathcal{E}[p_S + h\tilde{p}_S] - \mathcal{E}[p_S]) \cdot \varphi \, d\Omega \\
 &\quad - \int_{\Gamma_{IMP}} i\mu_0 ((\omega + h\tilde{\omega})Y(\omega + h\tilde{\omega}) - \omega Y(\omega)) \pi_t[\mathcal{E}[p_S + h\tilde{p}_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &\quad - \int_{\Gamma_{IMP}} i\mu_0 \omega Y(\omega) \pi_t[\mathcal{E}[p_S + h\tilde{p}_S] - \mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &\quad - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 ((\omega + h\tilde{\omega})Y(\omega + h\tilde{\omega}) - \omega Y(\omega)) \pi_t[\mathcal{E}[p_S + h\tilde{p}_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &\quad - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 \omega Y(\omega) \pi_t[\mathcal{E}[p_S + h\tilde{p}_S] - \mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma
 \end{aligned} \tag{5.23}$$

zu erhalten. Betrachten wir nun die Parameterfunktionen in den Integralen, dann gilt zum einen

$$(\omega + h\tilde{\omega})^2 \varepsilon_r(\omega + h\tilde{\omega}) - \omega^2 \varepsilon_r(\omega) = (\omega + h\tilde{\omega})^2 (\varepsilon_r(\omega + h\tilde{\omega}) - \varepsilon_r(\omega)) + (2h\omega\tilde{\omega} + h^2\tilde{\omega}^2) \varepsilon_r(\omega)$$

und zum anderen

$$(\omega + h\tilde{\omega})Y(\omega + h\tilde{\omega}) - \omega Y(\omega) = (\omega + h\tilde{\omega})(Y(\omega + h\tilde{\omega}) - Y(\omega)) + h\tilde{\omega}Y(\omega).$$

Multiplizieren wir diese mit  $\frac{1}{h}$  und bilden den Grenzwert für  $h \rightarrow 0$ , dann haben wir zum einen

$$\lim_{h \rightarrow 0} \frac{1}{h} ((\omega + h\tilde{\omega})^2 \varepsilon_r(\omega + h\tilde{\omega}) - \omega^2 \varepsilon_r(\omega)) = \omega^2 \delta \varepsilon_r[\omega, \tilde{\omega}] + 2\omega\tilde{\omega} \varepsilon_r(\omega)$$

und zum anderen

$$\lim_{h \rightarrow 0} \frac{1}{h} ((\omega + h\tilde{\omega})Y(\omega + h\tilde{\omega}) - \omega Y(\omega)) = \omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega),$$

wobei diese Grenzwerte wegen Voraussetzung 5.15 nach Lemma 5.17 existieren. Wie eben multiplizieren wir (5.23) mit  $\frac{1}{h}$  und betrachten wiederum den Grenzwert für  $h$  gegen Null, wobei wir zunächst die Existenz der ersten Variation  $\delta \mathcal{E}[p_S, \tilde{p}_S]$  annehmen und später noch zeigen werden, dass diese Annahme korrekt ist. Wir erhalten damit mit den obigen Konvergenzen der Parameterfunktionen, Satz 5.3 und dem Satz von Lebesgue den Grenzwert

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{1}{h} (a(\mathcal{E}[p_S + h\tilde{p}_S], \varphi, p_S + h\tilde{p}_S) - a(\mathcal{E}[p_S], \varphi, p_S)) & (5.24) \\ & = a(\delta \mathcal{E}[p_S, \tilde{p}_S], \varphi, p_S) - \int_{\Omega} \varepsilon_0 \mu_0 (\omega^2 \delta \varepsilon_r[\omega, \tilde{\omega}] + 2\omega \tilde{\omega} \varepsilon_r(\omega)) \mathcal{E}[p_S] \cdot \varphi \, d\Omega \\ & \quad - \int_{\Gamma_{IMP}} i \mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ & \quad - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i \mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma. \end{aligned}$$

Verfahren wir für die rechte Seite in (5.22) analog zur linken Seite in (5.22), dann ergibt sich

$$\begin{aligned} & f(\varphi, p_S + h\tilde{p}_S) - f(\varphi, p_S) \\ & = -2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i(\omega + h\tilde{\omega}) Y(\omega + h\tilde{\omega}) \mu_0 \frac{\sqrt{P_{in,l} + h\tilde{P}_{in,l}}}{\sqrt{|P_{m,l}(\omega + h\tilde{\omega})|}} e^{-i(\theta_l + h\tilde{\theta}_l)} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma \\ & \quad + 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma. \end{aligned}$$

Werden nun wiederum verschiedene Zwischenterme eingefügt, die Gleichung mit  $\frac{1}{h}$  mul-

tippliziert und dann der Grenzwert für  $h$  gegen Null betrachtet, erhalten wir

$$\begin{aligned}
 & \lim_{h \rightarrow 0} \frac{1}{h} (f(\varphi, p_S + h\tilde{p}_S) - f(\varphi, p_S)) & (5.25) \\
 &= -2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 \sqrt{P_{in,l}} e^{-i\theta_l} \left( \frac{\omega \delta Y[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)}} + \frac{\tilde{\omega} Y(\omega)}{\sqrt{P_{m,l}(\omega)}} \right. \\
 & \quad \left. - \frac{\omega}{2} Y(\omega) \frac{\delta P_{m,l}[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)^3}} \right) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] d\Gamma \\
 & - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 \omega Y(\omega) \frac{\tilde{P}_{in,l}}{\sqrt{P_{in,l}} \sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] d\Gamma \\
 & - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \mu_0 \omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \tilde{\theta}_l \pi_t[E_{0,l}] \cdot \pi_t[\varphi] d\Gamma,
 \end{aligned}$$

wobei wir die Differenzierbarkeit der Abbildungen  $\theta_l \mapsto e^{-i\theta_l}$  und  $P_{in,l} \mapsto \sqrt{P_{in,l}}$ , wobei  $P_{in,l} > 0$ ,  $1 \leq l \leq K_A$ , wegen der Bedingung an die zulässigen Menge  $\mathbb{O}_S$  in (5.14), benutzt und Bemerkung 5.18 verwendet haben. Kombinieren wir nun (5.24) und (5.25), so erhalten wir die Variationsgleichung (5.20).

Für die Korrektheit der Annahme der Existenz der ersten Variation  $\delta\mathcal{E}[p_S, \tilde{p}_S]$  definieren wir den Operator

$$\begin{aligned}
 \mathcal{L}_{f_2} : X_e &\rightarrow \mathbb{C}, \\
 \varphi &\mapsto \mathcal{L}_{f_2}[\varphi] = f_2(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S),
 \end{aligned}$$

wobei  $f_2$  gegeben ist durch (5.21). Es ist einfach zu sehen, dass der Operator komplex linear und in der Dualraumnorm beschränkt ist, weswegen der Operator Element des Dualraums ist. Daher erhalten wir mit derselben Herangehensweise wie im Beweis von Satz 3.10 die Existenz und Eindeutigkeit der ersten Variation durch Satz 3.1.  $\square$

**Korollar 5.20.** *Gilt Voraussetzung 5.16, dann ist die Abbildung  $p_S \mapsto \mathcal{E}[p_S]$  für die Steuerparameter  $p_S \in \mathbb{O}_S$  Gâteaux-differenzierbar, also*

$$\delta\mathcal{E}[p_S, \tilde{p}_S] = \nabla\mathcal{E}[p_S]\tilde{p}_S, \quad (5.26)$$

wobei

$$\nabla\mathcal{E}[p_S] = \left( \nabla_\omega \mathcal{E}[p_S], [\nabla_{P_{in,l}} \mathcal{E}[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l} \mathcal{E}[p_S]]_{l=1}^{K_A} \right)$$

mit  $\nabla_\omega \mathcal{E}[p_S], \nabla_{P_{in,l}} \mathcal{E}[p_S], \nabla_{\theta_l} \mathcal{E}[p_S] \in X_e$ ,  $1 \leq l \leq K_A$ , wobei  $[\nabla_{P_{in,l}} \mathcal{E}[p_S]]_{l=1}^{K_A}$  bzw.  $[\nabla_{\theta_l} \mathcal{E}[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}} \mathcal{E}[p_S]$  bzw.  $\nabla_{\theta_l} \mathcal{E}[p_S]$  sind, und die ein-

## 5. Optimalsteuerung elektromagnetischer Felder

---

zeln Gradienten folgende Variationsprobleme lösen:

1. Der Gradient  $\nabla_\omega \mathcal{E}[p_S]$  löst das Variationsproblem

$$a(\nabla_\omega \mathcal{E}[p_S], \varphi, p_S) = f_{G_\omega}(\varphi, \mathcal{E}[p_S], p_S), \quad \varphi \in X_e, \quad (5.27)$$

mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\begin{aligned} f_{G_\omega}(\varphi, \mathcal{E}[p_S], p_S) &= \int_{\Omega} \varepsilon_0 \mu_0 (\omega^2 \nabla \varepsilon_r[\omega] + 2\omega \varepsilon_r(\omega)) \mathcal{E}[p_S] \cdot \varphi \, d\Omega \\ &+ \int_{\Gamma_{IMP}} i\mu_0 (\omega \nabla Y[\omega] + Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ &+ \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 (\omega \nabla Y[\omega] + Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ &- 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\mu_0 \sqrt{P_{in,l}} e^{-i\theta_l} \left( \frac{\omega \nabla Y[\omega]}{\sqrt{P_{m,l}(\omega)}} + \frac{Y(\omega)}{\sqrt{P_{m,l}(\omega)}} \right. \\ &\quad \left. - \frac{\omega}{2} Y(\omega) \frac{\nabla P_{m,l}[\omega]}{\sqrt{P_{m,l}(\omega)^3}} \right) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma. \end{aligned}$$

2. Für den Gradient  $\nabla_{P_{in,l}} \mathcal{E}[p_S]$ ,  $1 \leq l \leq K_A$ , gilt

$$\nabla_{P_{in,l}} \mathcal{E}[p_S] = \frac{1}{2P_{in,l}} \mathcal{E}[\hat{p}_S],$$

wobei  $\mathcal{E}[\hat{p}_S]$  Lösung von  $[V_K]$  zu dem Parametervektor  $\hat{p}_S$  gegeben als

$$\hat{\omega} = \omega, \quad \hat{\theta}_l = \theta_l, \quad \hat{P}_{in,l} = P_{in,l},$$

sowie  $\hat{\theta}_k = \hat{P}_{in,k} = 0$  für  $1 \leq k \leq K_A$  mit  $k \neq l$ .

3. Für den Gradient  $\nabla_{\theta_l} \mathcal{E}[p_S]$ ,  $1 \leq l \leq K_A$ , gilt

$$\nabla_{\theta_l} \mathcal{E}[p_S] = -i\mathcal{E}[\hat{p}_S],$$

wobei  $\mathcal{E}[\hat{p}_S]$  Lösung von  $[V_K]$  zu dem Parametervektor  $\hat{p}_S$  gegeben als

$$\hat{\omega} = \omega, \quad \hat{\theta}_l = \theta_l, \quad \hat{P}_{in,l} = P_{in,l},$$

sowie  $\hat{\theta}_k = \hat{P}_{in,k} = 0$  für  $1 \leq k \leq K_A$  mit  $k \neq l$ .

*Beweis.* Für die Aussage zum Gradienten  $\nabla_\omega \mathcal{E}[p_S]$ , betrachten wir zunächst die erste Variation  $\delta_\omega \mathcal{E}[p_S, \tilde{p}_S]$  nur in Richtung der Frequenz, also die erste Variation  $\delta \mathcal{E}[p_S, \tilde{p}_S]$  aus Satz 5.19 mit

$$\tilde{p}_S = (\tilde{\omega}, ([0]_{k=1}^{K_A})^T, ([0]_{k=1}^{K_A})^T)^T.$$

Wir erhalten damit, dass die erste Variation  $\delta_\omega \mathcal{E}[p_S, \tilde{p}_S]$  in Richtung  $\tilde{\omega}$  das folgende Variationsproblem löst:

$$a(\delta_\omega \mathcal{E}[p_S, \tilde{p}_S], \varphi, p_S) = \tilde{f}(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S), \quad \varphi \in X_e,$$

mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\begin{aligned} & \tilde{f}(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S) \\ &= \int_{\Omega} \varepsilon_0 \mu_0 (\omega^2 \delta \varepsilon_r[\omega, \tilde{\omega}] + 2\omega \tilde{\omega} \varepsilon_r(\omega)) \mathcal{E}[p_S] \cdot \varphi \, d\Omega \\ &+ \int_{\Gamma_{IMP}} i \mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ &+ \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i \mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\ &- 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i \mu_0 \sqrt{P_{in,l}} e^{-i\theta_l} \left( \frac{\omega \delta Y[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)}} + \frac{\tilde{\omega} Y(\omega)}{\sqrt{P_{m,l}(\omega)}} \right. \\ &\quad \left. - \frac{\omega}{2} Y(\omega) \frac{\delta P_{m,l}[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)^3}} \right) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma. \end{aligned}$$

Wegen Voraussetzung 5.16 und Lemma 5.17 können wir die ersten Variationen der Parameterterme entsprechend durch deren Gâteaux-Ableitungen ersetzen und erhalten

$$\tilde{f}(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S) = f_{G_\omega}(\varphi, \mathcal{E}[p_S], p_S) \tilde{\omega}.$$

Betrachten wir demnach eine Lösung  $\nabla_\omega \mathcal{E}[p_S]$  zu (5.27), dann erhalten wir die Gâteaux-Differenzierbarkeit

$$\delta_\omega \mathcal{E}[p_S, \tilde{p}_S] = \nabla_\omega \mathcal{E}[p_S] \tilde{\omega}.$$

Für die zweite Aussage zum Gradienten  $\nabla_{P_{in,l}} \mathcal{E}[p_S]$ , betrachten wir wieder die erste Variation  $\delta_{P_{in,l}} \mathcal{E}[p_S, \tilde{p}_S]$  nur in Richtung der  $l$ -ten Leistung  $P_{in,l}$ , also die erste Variation  $\delta \mathcal{E}[p_S, \tilde{p}_S]$  aus Satz 5.19 mit

$$\tilde{p}_S = (0, 0, \dots, \tilde{P}_{in,l}, \dots, 0, ([0]_{k=1}^{K_A})^T)^T.$$

## 5. Optimalsteuerung elektromagnetischer Felder

Wir erhalten damit wiederum, dass die erste Variation  $\delta_{P_{in,l}}\mathcal{E}[p_S, \tilde{p}_S]$  in Richtung  $\tilde{P}_{in,l}$  das folgende Variationsproblem löst:

$$a(\delta_{P_{in,l}}\mathcal{E}[p_S, \tilde{p}_S], \varphi, p_S) = \hat{f}(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S), \quad \varphi \in X_e,$$

mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\hat{f}(\varphi, \mathcal{E}[p_S], p_S, \tilde{p}_S) = -\tilde{P}_{in,l} \int_{\Gamma_{P,l}} i\mu_0\omega Y(\omega) \frac{1}{\sqrt{P_{in,l}}\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\varphi] d\Gamma.$$

Vergleichen wir nun das hier angegebene Variationsproblem zu  $\delta_{P_{in,l}}\mathcal{E}[p_S, \tilde{p}_S]$  mit dem Variationsproblem  $[V_K]$ , dann unterscheidet sich das Integral in  $\hat{f}$  von dem  $l$ -ten Summanden der rechten Seite in  $[V_K]$  gerade um den Faktor  $\frac{\tilde{P}_{in,l}}{2P_{in,l}}$ . Betrachten wir also eine Lösung  $\mathcal{E}[\hat{p}_S]$  von  $[V_K]$  zu dem Parametervektor  $\hat{p}_S$  gegeben als

$$\hat{\omega} = \omega, \quad \hat{\theta}_l = \theta_l, \quad \hat{P}_{in,l} = P_{in,l},$$

sowie  $\hat{\theta}_k = \hat{P}_{in,k} = 0$  für  $1 \leq k \leq K_A$  mit  $k \neq l$ , dann gilt

$$\delta_{P_{in,l}}\mathcal{E}[p_S, \tilde{p}_S] = \frac{1}{2P_{in,l}}\mathcal{E}[\hat{p}_S]\tilde{P}_{in,l}$$

und somit auch

$$\nabla_{P_{in,l}}\mathcal{E}[p_S] = \frac{1}{2P_{in,l}}\mathcal{E}[\hat{p}_S].$$

Die Aussage zu  $\nabla_{\theta_l}\mathcal{E}[p_S]$  folgt dabei komplett analog zu eben.  $\square$

Mit Satz 5.19 und Korollar 5.20 können wir nun die erste Variation der beiden Zielfunktionale  $\mathcal{Z}_1$  in (5.15) und  $\mathcal{Z}_2$  in (5.16) bezüglich des Steuerparametervektors  $p_S$  bestimmen.

**Satz 5.21.** *Gilt Voraussetzung 5.15, dann existiert die erste Variation  $\delta\mathcal{Z}_1[p_S, \tilde{p}_S]$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  und ist gegeben durch*

$$\delta\mathcal{Z}_1[p_S, \tilde{p}_S] = \frac{1}{|V|} \int_V \operatorname{Re}(\delta\mathcal{E}[p_S, \tilde{p}_S] \cdot (\mathcal{E}[p_S] - E_V)) d\Omega, \quad (5.28)$$

wobei  $\delta\mathcal{E}[p_S, \tilde{p}_S]$  durch Satz 5.19 gegeben ist. Gilt zudem Voraussetzung 5.16, dann ist

$$\delta\mathcal{Z}_1[p_S, \tilde{p}_S] = \nabla\mathcal{Z}_1[p_S]\tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_1[p_S] = \frac{1}{|V|} \int_V \operatorname{Re}(\nabla \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V)) \, d\Omega, \quad (5.29)$$

wobei die Gâteaux-Ableitung  $\nabla \mathcal{E}[p_S]$  in Korollar 5.20 gegeben ist.

*Beweis.* Wir betrachten zunächst die Differenz aus den beiden Zielfunktionalen

$$\mathcal{Z}_1[p_S + h\tilde{p}_S] - \mathcal{Z}_1[p_S] = \frac{1}{2} \frac{1}{|V|} \left( \int_V |\mathcal{E}[p_S + h\tilde{p}_S] - E_V|^2 \, d\Omega - \int_V |\mathcal{E}[p_S] - E_V|^2 \, d\Omega \right).$$

Betrachten wir anstatt  $|\cdot|^2$  das komplexe Skalarprodukt und fügen noch einen Zwischenterm ein, dann erhalten wir

$$\begin{aligned} \mathcal{Z}_1[p_S + h\tilde{p}_S] - \mathcal{Z}_1[p_S] = \frac{1}{2} \frac{1}{|V|} & \left( \int_V (\mathcal{E}[p_S + h\tilde{p}_S] - \mathcal{E}[p_S]) \cdot (\mathcal{E}[p_S + h\tilde{p}_S] - E_V) \, d\Omega \right. \\ & \left. + \int_V (\mathcal{E}[p_S] - E_V) \cdot (\mathcal{E}[p_S + h\tilde{p}_S] - \mathcal{E}[p_S]) \, d\Omega \right). \end{aligned}$$

Multiplizieren wir die Gleichung wieder mit dem Faktor  $\frac{1}{h}$  und betrachten den Grenzwert für  $h$  gegen Null, dann folgt mit der starken Konvergenz

$$\mathcal{E}[p_S + h\tilde{p}_S] \xrightarrow{h \rightarrow 0} \mathcal{E}[p_S]$$

nach Satz 5.3 und der Existenz der ersten Variation  $\delta \mathcal{E}[p_S, \tilde{p}_S]$  nach Satz 5.19 die Gleichung

$$\delta \mathcal{Z}_1[p_S, \tilde{p}_S] = \frac{1}{2} \frac{1}{|V|} \int_V \delta \mathcal{E}[p_S, \tilde{p}_S] \cdot (\mathcal{E}[p_S] - E_V) + (\mathcal{E}[p_S] - E_V) \cdot \delta \mathcal{E}[p_S, \tilde{p}_S] \, d\Omega.$$

Wird nun noch eine einfache Rechnung für das komplexe Skalarprodukt durchgeführt, dann ergibt sich gerade die Darstellung (5.28). Die zweite Behauptung folgt dann direkt aus (5.28) bei Verwendung der Darstellung (5.26) aus Korollar 5.20.  $\square$

Mit denselben Beweisschritten folgt komplett analog dazu auch die Behauptung für das zweite Zielfunktional.

**Satz 5.22.** *Gilt Voraussetzung 5.15, dann existiert die erste Variation  $\delta\mathcal{Z}_2[p_S, \tilde{p}_S]$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  und ist gegeben durch*

$$\begin{aligned} \delta\mathcal{Z}_2[p_S, \tilde{p}_S] &= \frac{1}{4} \frac{1}{|V|} \int_V \varepsilon_0 (\omega \operatorname{Im}(\delta\varepsilon_r[\omega, \tilde{\omega}] + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}[p_S] - E_V|^2 \, d\Omega \\ &\quad + \frac{1}{2} \frac{1}{|V|} \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}(\delta\mathcal{E}[p_S, \tilde{p}_S] \cdot (\mathcal{E}[p_S] - E_V)) \, d\Omega, \end{aligned} \quad (5.30)$$

wobei  $\delta\mathcal{E}[p_S, \tilde{p}_S]$  in Satz 5.19 gegeben ist. Gilt zudem Voraussetzung 5.16, dann ist

$$\delta\mathcal{Z}_2[p_S, \tilde{p}_S] = \nabla\mathcal{Z}_2[p_S] \tilde{p}_S$$

mit

$$\begin{aligned} \nabla\mathcal{Z}_2[p_S] &= \frac{1}{4} \frac{1}{|V|} \int_V \varepsilon_0 (\omega \operatorname{Im}(\nabla\varepsilon_r[\omega]) + \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}[p_S] - E_V|^2 \, d\Omega \\ &\quad + \frac{1}{2} \frac{1}{|V|} \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}(\nabla\mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V)) \, d\Omega, \end{aligned} \quad (5.31)$$

wobei  $\nabla\mathcal{E}[p_S]$  in Korollar 5.20 gegeben ist.

Nachdem wir nun die Ableitungen hergeleitet haben, gehen wir die in der Optimierung oft verwendete Theorie des adjungierten Problems an, siehe [27], und leiten diese im nächsten Abschnitt her. Ein Vorteil der Lösung des adjungierten Problems liegt darin, dass dadurch eine weitere Formel zur Bestimmung der ersten Variation und des Gradienten der Zielfunktionale  $\mathcal{Z}_1$  und  $\mathcal{Z}_2$  hergeleitet werden kann, die zuweilen einen geringeren Rechenaufwand besitzt.

#### 5.1.4. Adjungierte Probleme

Um das adjungierte Problem der beiden Optimalsteuerungsprobleme  $[O_k]_K$ ,  $k = 1, 2$ , herzuleiten, verwenden wir das in [27, Kapitel 2.10] vorgestellte formale Lagrangeprinzip. Wir zeigen die Vorgehensweise an  $[O_1]_K$  und geben danach das adjungierte Problem von  $[O_2]_K$  ohne Herleitung an, da dies analog erfolgt. Um das formale Lagrangeprinzip anzuwenden, betrachten wir  $[O_1]_K$  nicht mit der PDE im variationellen Sinne als Nebenbedingung, sondern in der ursprünglichen Form:

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$  und  $E \in H(\text{curl}; \Omega)$ , sodass diese Lösung des folgenden Optimierungsproblems sind:

$$\begin{aligned} \min_{\substack{p_S \in \mathbb{P}_S, \\ E \in H(\text{curl}; \Omega)}} Z_1(p_S, E) &= \frac{1}{2} \frac{1}{|V|} \int_V |E - E_V|^2 d\Omega \\ &\text{unter (2.81), (2.82), (2.83), (2.84).} \end{aligned} \quad (5.32)$$

Die Existenz einer Lösung  $E \in H(\text{curl}; \Omega)$  zu diesem Optimierungsproblem ist nicht klar, da  $\nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times E \right)$  nicht existieren muss. Wir können jedoch dieses Optimierungsproblem verwenden um eine wohldefinierte Lagrange-Funktion herzuleiten. Aufgrund der Vektoridentität der komplexen Zahlen  $\mathbb{C}$  mit dem reellen Vektorraum  $\mathbb{R}^2$ , beschreiben wir die komplexe Nebenbedingung als zwei reelle Nebenbedingungen. Die zum eben definierten Optimierungsproblem lautende formale Lagrange-Funktion ist demnach gegeben durch

$$\begin{aligned} &\mathcal{L}(p_S, E, [\lambda_k^{\text{Re}}]_{k=1}^{K_A+3}, [\lambda_k^{\text{Im}}]_{k=1}^{K_A+3}) \\ &= \frac{1}{2} \frac{1}{|V|} \int_V |E - E_V|^2 d\Omega + \sum_{\mathcal{G} \in \{\text{Re}, \text{Im}\}} \left[ - \int_{\Gamma_{PEC}} (\mathcal{G}(\gamma_t[E]), \lambda_2^{\mathcal{G}})_{\mathbb{R}^3} d\Gamma \right. \\ &\quad - \int_{\Omega} \left( \mathcal{G} \left( \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times E \right) - \omega^2 \varepsilon_0 \varepsilon_r(\omega) E \right), \lambda_1^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Omega \\ &\quad - \int_{\Gamma_{IMP}} \left( \mathcal{G} \left( \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] - i\omega Y(\omega) \pi_t[E] \right), \lambda_3^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Gamma \\ &\quad + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \left( \mathcal{G} \left( i\omega \left( Y(\omega) \pi_t[E] - \frac{2Y(\omega) \sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \right) \right), \lambda_{3+l}^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Gamma \\ &\quad \left. - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \left( \mathcal{G} \left( \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] \right), \lambda_{3+l}^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Gamma \right], \end{aligned}$$

## 5. Optimalsteuerung elektromagnetischer Felder

---

wobei  $\lambda_k^{\text{Re}}, \lambda_k^{\text{Im}} \in H(\text{curl}; \Omega)$  reelle Vektorfelder im  $\mathbb{R}^3$  sind. Verwenden wir die Summen- und Integraleigenschaften der Real- und Imaginärteil-Operatoren erhalten wir

$$\begin{aligned}
& \mathcal{L}(p_S, E, [\lambda_k^{\text{Re}}]_{k=1}^{K_A+3}, [\lambda_k^{\text{Im}}]_{k=1}^{K_A+3}) \\
&= \frac{1}{2} \frac{1}{|V|} \int_V |E - E_V|^2 d\Omega + \sum_{\mathcal{G} \in \{\text{Re}, \text{Im}\}} \left[ \mathcal{G} \left( - \int_{\Gamma_{PEC}} (\gamma_t[E], \lambda_2^{\mathcal{G}})_{\mathbb{R}^3} d\Gamma \right. \right. \\
&\quad - \int_{\Omega} \left( \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times E \right) - \omega^2 \varepsilon_0 \varepsilon_r(\omega) E, \lambda_1^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Omega \\
&\quad - \int_{\Gamma_{IMP}} \left( \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] - i\omega Y(\omega) \pi_t[E], \lambda_3^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Gamma \\
&\quad \left. + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \left( i\omega \left( Y(\omega) \pi_t[E] - \frac{2Y(\omega) \sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \right), \lambda_{3+l}^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Gamma \right. \\
&\quad \left. - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \left( \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right], \lambda_{3+l}^{\mathcal{G}} \right)_{\mathbb{R}^3} d\Gamma \right).
\end{aligned}$$

Des Weiteren folgern wir

$$\begin{aligned}
& \mathcal{L}(p_S, E, [\lambda_k]_{k=1}^{K_A+3}) \\
&= \frac{1}{2} \frac{1}{|V|} \int_V |E - E_V|^2 d\Omega + \text{Re} \left( - \int_{\Gamma_{PEC}} \gamma_t[E] \cdot \lambda_2 d\Gamma \right. \\
&\quad - \int_{\Omega} \left( \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times E \right) - \omega^2 \varepsilon_0 \varepsilon_r(\omega) E \right) \cdot \lambda_1 d\Omega \\
&\quad - \int_{\Gamma_{IMP}} \left( \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] - i\omega Y(\omega) \pi_t[E] \right) \cdot \lambda_3 d\Gamma \\
&\quad \left. + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega \left( Y(\omega) \pi_t[E] - \frac{2Y(\omega) \sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \right) \cdot \lambda_{3+l} d\Gamma \right. \\
&\quad \left. - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] \cdot \lambda_{3+l} d\Gamma \right),
\end{aligned}$$

indem wir die Summanden der Lagrange-Funktion entsprechend der Eigenschaft

$$\begin{aligned} \operatorname{Re}((a + ib) \cdot c) + \operatorname{Im}((a + ib) \cdot d) &= \operatorname{Re}((a + ib) \cdot c) + \operatorname{Re}(-i(a + ib) \cdot d) \\ &= \operatorname{Re}((a + ib) \cdot (c + id)) \end{aligned} \quad (5.33)$$

komplexer Zahlen und des komplexen Skalarproduktes zusammenführen und deswegen die Lagrange-Multiplikatoren als komplexe Multiplikatoren

$$\lambda_k = \lambda_k^{\operatorname{Re}} + i\lambda_k^{\operatorname{Im}}, \quad 1 \leq k \leq K_A + 3,$$

definieren. Mit zweifachem Anwenden der Greenschen-Integrationsformel (2.55) erhalten wir

$$\begin{aligned} &\mathcal{L}(p_S, E, [\lambda_k]_{k=1}^{K_A+3}) \\ &= \frac{1}{2} \frac{1}{|V|} \int_V |E - E_V|^2 d\Omega \\ &\quad + \operatorname{Re} \left( - \int_{\Omega} E \cdot \left( \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right) - \omega^2 \varepsilon_0 \overline{\varepsilon_r(\omega)} \lambda_1 \right) d\Omega \right. \\ &\quad - \int_{\Gamma_{PEC}} -\gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] \cdot \pi_t[\lambda_1] + \gamma_t [E] \cdot \left( \lambda_2 - \pi_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] \right) d\Gamma \\ &\quad - \int_{\Gamma_{IMP}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] \cdot (\lambda_3 - \pi_t[\lambda_1]) d\Gamma \\ &\quad + \int_{\Gamma_{IMP}} \pi_t[E] \cdot \left( \omega i \overline{Y(\omega)} \lambda_3 - \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] \right) d\Gamma \\ &\quad - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times E \right] \cdot (\lambda_{3+l} - \pi_t[\lambda_1]) d\Gamma \\ &\quad + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \pi_t[E] \cdot \left( \omega i \overline{Y(\omega)} \lambda_{3+l} - \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] \right) d\Gamma \\ &\quad - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \left( i\omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \lambda_{3+l} d\Gamma \right), \end{aligned}$$

## 5. Optimalsteuerung elektromagnetischer Felder

---

wobei wir hier die Vektoridentität

$$\gamma_t[E] \cdot \pi_t[\tilde{E}] = -\pi_t[E] \cdot \gamma_t[\tilde{E}]$$

sowohl für die vierte Zeile

$$\int_{\Gamma_{IMP}} \gamma_t[E] \cdot \pi_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] d\Gamma = - \int_{\Gamma_{IMP}} \pi_t[E] \cdot \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] d\Gamma$$

als auch für die sechste Zeile

$$\int_{\Gamma_{P,l}} \gamma_t[E] \cdot \pi_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] d\Gamma = - \int_{\Gamma_{P,l}} \pi_t[E] \cdot \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] d\Gamma$$

mit  $1 \leq l \leq K_A$  angewendet haben. Für ein optimales elektrisches Feld  $\hat{E}$  des formalen Optimalsteuerungsproblems

$$\min \mathcal{L}(p_S, E, [\lambda_k]_{k=1}^{K_A+3})$$

verschwindet die erste Variation, also

$$0 = \lim_{h \rightarrow 0} \frac{1}{h} \left( \mathcal{L}(p_S, \hat{E} + h\tilde{E}, [\lambda_k]_{k=1}^{K_A+3}) - \mathcal{L}(p_S, \hat{E}, [\lambda_k]_{k=1}^{K_A+3}) \right),$$

für beliebiges  $\tilde{E} \in H(\text{curl}; \Omega)$ . Aus der Summeneigenschaft der Real- und Imaginärteil-Operatoren, der Linearität der verschiedenen Operatoren und dem Satz von Lebesgue erhalten wir daher

$$\begin{aligned} 0 &= \frac{1}{2} \frac{1}{|V|} \int_V \tilde{E} \cdot (\hat{E} - E_V) + (\hat{E} - E_V) \cdot \tilde{E} d\Omega & (5.34) \\ &+ \text{Re} \left( - \int_{\Omega} \tilde{E} \cdot \left( \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right) - \omega^2 \varepsilon_0 \overline{\varepsilon_r(\omega)} \lambda_1 \right) d\Omega \right. \\ &+ \int_{\Gamma_{PEC}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \tilde{E} \right] \cdot \pi_t[\lambda_1] d\Gamma - \int_{\Gamma_{PEC}} \gamma_t [\tilde{E}] \cdot \left( \lambda_2 - \pi_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] \right) d\Gamma \\ &\left. - \int_{\Gamma_{IMP}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \tilde{E} \right] \cdot (\lambda_3 - \pi_t[\lambda_1]) d\Gamma \right) \end{aligned}$$

$$\begin{aligned}
 & + \int_{\Gamma_{IMP}} \pi_t[\tilde{E}] \cdot \left( \overline{\omega i Y(\omega)} \lambda_3 - \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] \right) d\Gamma \\
 & - \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \tilde{E} \right] \cdot (\lambda_{3+l} - \pi_t[\lambda_1]) d\Gamma \\
 & + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} \pi_t[\tilde{E}] \cdot \left( \overline{\omega i Y(\omega)} \lambda_{3+l} - \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right] \right) d\Gamma.
 \end{aligned}$$

Nach [32] liegt  $[C_0^\infty(\Omega)]^3$  dicht in  $H_0(\text{curl}; \Omega)$ , weswegen wir zunächst  $\tilde{E} \in [C_0^\infty(\Omega)]^3$  wählen. Dadurch fallen in (5.34) alle Randterme weg und wir erhalten

$$\begin{aligned}
 & \frac{1}{2} \frac{1}{|V|} \int_V \tilde{E} \cdot (\hat{E} - E_V) + (\hat{E} - E_V) \cdot \tilde{E} d\Omega \\
 & = \text{Re} \left( \int_{\Omega} \tilde{E} \cdot \left( \nabla \times \left( \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right) - \omega^2 \varepsilon_0 \overline{\varepsilon_r(\omega)} \lambda_1 \right) d\Omega \right).
 \end{aligned}$$

Als nächstes wählen wir  $\tilde{E} \in H(\text{curl}; \Omega)$  beliebig aber fest, mit den zusätzlichen Eigenschaften  $\gamma_t[\tilde{E}] = \pi_t[\tilde{E}] = 0$  auf  $\Gamma_{PEC} \cup \Gamma_{IMP} \cup \left( \bigcup_{l=1}^{K_A} \Gamma_{P,l} \right)$ , sowie  $\gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \tilde{E} \right] = 0$  auf  $\Gamma_{IMP} \cup \left( \bigcup_{l=1}^{K_A} \Gamma_{P,l} \right)$ , und erhalten damit

$$0 = \int_{\Gamma_{PEC}} \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \tilde{E} \right] \cdot \pi_t[\lambda_1] d\Gamma.$$

Dies ist aufgrund der beliebigen Wahl von  $\tilde{E}$  nach dem Fundamentalsatz der Variationsrechnung äquivalent zu

$$\pi_t[\lambda_1] = 0,$$

bzw.

$$\gamma_t[\lambda_1] = 0$$

auf  $\Gamma_{PEC}$ . Wählen wir nun  $\tilde{E} \in H(\text{curl}; \Omega)$  mit  $\pi_t[\tilde{E}] = 0$  auf  $\Gamma_{IMP} \cup \left( \bigcup_{l=1}^{K_A} \Gamma_{P,l} \right)$ , sowie

$$\gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \tilde{E} \right] = 0$$

auf  $\Gamma_{PEC} \cup \Gamma_{IMP} \cup \left( \bigcup_{l=1}^{K_A} \Gamma_{P,l} \right)$ , dann ergibt sich gerade

$$\lambda_2 = \pi_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right].$$

Analog verfahren wir für die anderen Terme und erhalten die Gleichungen

$$\begin{aligned} \lambda_3 &= \pi_t[\lambda_1], & \overline{\omega i Y(\omega)} \lambda_3 &= \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right], \\ \lambda_{3+l} &= \pi_t[\lambda_1], & \overline{\omega i Y(\omega)} \lambda_{3+l} &= \gamma_t \left[ \frac{1}{\mu_0 \mu_r} \nabla \times \lambda_1 \right], \quad 1 \leq l \leq K_A. \end{aligned}$$

Für den ersten Term in (5.34) erhalten wir zudem

$$\frac{1}{2} \frac{1}{|V|} \int_V \tilde{E} \cdot (\hat{E} - E_V) + (\hat{E} - E_V) \cdot \tilde{E} \, d\Omega = \frac{1}{|V|} \int_V \operatorname{Re} \left( \tilde{E} \cdot (\hat{E} - E_V) \right) \, d\Omega, \quad (5.35)$$

aufgrund des komplexen Skalarproduktes. Wenden wir nun die Greensche Integrationsformel (2.55) in (5.34) an und verwenden die hergeleiteten Erkenntnisse, so lässt sich das adjungierte Problem zum Steuerungsproblem  $[O_1]_K$  ableiten und die korrekte Darstellung der Lagrange-Funktion angeben:

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $E \in X_e$  des Optimierungsproblems  $[O_1]_K$  die adjungierte Lösung  $\lambda \in X_e$ , sodass

$$[D_1]_K \quad a(\varphi, \lambda, p_S) = f_{D_1}(\varphi, E, E_V, p_S), \quad \varphi \in X_e,$$

gilt mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$f_{D_1}(\varphi, E, E_V, p_S) = \frac{1}{|V|} \int_V \varphi \cdot (E - E_V) \, d\Omega. \quad (5.36)$$

Zudem können wir die Lagrange-Funktion zu dem Optimierungsproblem  $[O_1]_K$  mit Anwendung von (5.33) entsprechend umschreiben.

**Definition 5.23.** Die Lagrange-Funktion zu dem Optimierungsproblem  $[O_1]_K$  ist gegeben durch

$$\begin{aligned}
 \mathcal{L}_1(p_S, E, \lambda) & \quad (5.37) \\
 &= \frac{1}{2} \frac{1}{|V|} \int_V |E - E_V|^2 \, d\Omega + \operatorname{Re} \left( - \int_{\Omega} \frac{1}{\mu_0 \mu_r} (\nabla \times E) \cdot (\nabla \times \lambda) \, d\Omega \right. \\
 & \quad + \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) E \cdot \lambda \, d\Omega + \int_{\Gamma_{IMP}} i\omega Y(\omega) \pi_t[E] \cdot \pi_t[\lambda] \, d\Gamma \\
 & \quad + \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \pi_t[E] \cdot \pi_t[\lambda] \, d\Gamma \\
 & \quad \left. - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda] \, d\Gamma \right),
 \end{aligned}$$

wobei  $p_S \in \mathbb{P}_S$ ,  $E, E_V, \lambda \in X_e$ .

Um die rechte Seite (5.36) aus (5.34) herzuleiten, verwenden wir die rechte Seite von (5.35) ohne den Realteil-Operator. Der Operator wird dabei nicht fallen gelassen, sondern im Satz 5.25 verwendet, um die adjungierte Lösung mit der ersten Variation des Zielfunktional in Verbindung zu bringen. Die Existenz einer Lösung zu  $[D_1]_K$  erhalten wir dabei wieder völlig analog zu Abschnitt 3.1. Die Lagrange-Funktion in (5.37) ist damit wohldefiniert. Verwenden wir die formale Lagrangetechnik für das Optimierungsproblem  $[O_2]_K$ , erhalten wir das folgende adjungierte Problem und die Lagrange-Funktion zum zweiten Zielfunktional.

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $E \in X_e$  des Optimierungsproblems  $[O_2]_K$  die adjungierte Lösung  $\lambda \in X_e$ , sodass

$$[D_2]_K \quad a(\varphi, \lambda, p_S) = f_{D_2}(\varphi, E, E_V, p_S), \quad \varphi \in X_e,$$

gilt mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$f_{D_2}(\varphi, E, E_V, p_S) = \frac{1}{|V|} \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) \varphi \cdot (E - E_V) \, d\Omega. \quad (5.38)$$

**Definition 5.24.** Die Lagrange-Funktion zu dem Optimierungsproblem  $[O_2]_K$  ist gegeben durch

$$\begin{aligned}
 \mathcal{L}_2(p_S, E, \lambda) & \qquad \qquad \qquad (5.39) \\
 &= \frac{1}{2} \frac{1}{|V|} \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E - E_V|^2 d\Omega \\
 &+ \operatorname{Re} \left( - \int_{\Omega} \frac{1}{\mu_0 \mu_r} (\nabla \times E) \cdot (\nabla \times \lambda) d\Omega \right. \\
 &+ \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) E \cdot \lambda d\Omega + \int_{\Gamma_{IMP}} i\omega Y(\omega) \pi_t[E] \cdot \pi_t[\lambda] d\Gamma \\
 &+ \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \pi_t[E] \cdot \pi_t[\lambda] d\Gamma \\
 &\left. - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda] d\Gamma \right),
 \end{aligned}$$

wobei  $p_S \in \mathbb{P}_S$ ,  $E, E_V, \lambda \in X_e$ .

Mit Verwendung der adjungierten Lösungen können wir die erste Variation der Zielfunktionale der Optimierungsprobleme  $[O_1]_K$  und  $[O_2]_K$  über die rechte Seite (5.21) des Variationsproblems (5.20) darstellen.

**Satz 5.25.** *Gilt Voraussetzung 5.15 und  $p_S \in \mathbb{O}_S$ , dann folgt:*

1. Die erste Variation  $\delta \mathcal{Z}_1[p_S, \tilde{p}_S]$  des Kostenfunktionals  $\mathcal{Z}_1[p_S]$  in (5.15) ist gegeben durch

$$\delta \mathcal{Z}_1[p_S, \tilde{p}_S] = \operatorname{Re} (f_2(\lambda_1, \mathcal{E}[p_S], p_S, \tilde{p}_S)),$$

wobei  $\lambda_1 \in X_e$  Lösung zu  $[D_1]_K$ .

2. Die erste Variation  $\delta \mathcal{Z}_2[p_S, \tilde{p}_S]$  des Kostenfunktionals  $\mathcal{Z}_2[p_S]$  in (5.16) ist gegeben durch

$$\begin{aligned}
 \delta \mathcal{Z}_2[p_S, \tilde{p}_S] &= \frac{1}{4} \frac{1}{|V|} \int_V \varepsilon_0 (\omega \operatorname{Im}(\delta \varepsilon_r[\omega, \tilde{\omega}]) + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\omega))(\omega)) |\mathcal{E}[p_S] - E_V|^2 d\Omega \\
 &+ \operatorname{Re} (f_2(\lambda_2, \mathcal{E}[p_S], p_S, \tilde{p}_S)),
 \end{aligned}$$

wobei  $\lambda_2 \in X_e$  Lösung zu  $[D_2]_K$ .

*Beweis.* Wir beweisen hier nur die erste Aussage, da die zweite Aussage analog folgt. Ziehen wir den Realteil-Operator in (5.28) aus dem Integral und wenden  $[D_1]_K$  an, dann gilt

$$\begin{aligned}\delta \mathcal{Z}_1[p_S, \tilde{p}_S] &= \operatorname{Re} \left( \frac{1}{|V|} \int_V \delta \mathcal{E}[p_S, \tilde{p}_S] \cdot (\mathcal{E}[p_S] - E_V) \, d\Omega \right) \\ &= \operatorname{Re} (f_{D_1}(\delta \mathcal{E}[p_S, \tilde{p}_S], \mathcal{E}[p_S], E_V, p_S)) \\ &= \operatorname{Re} (a(\delta \mathcal{E}[p_S, \tilde{p}_S], \lambda_1, p_S))\end{aligned}$$

Verwenden wir nun die Lösungseigenschaft der Funktion  $\delta \mathcal{E}[p_S, \tilde{p}_S]$  zu  $\mathcal{E}[p_S]$  und der Variationsgleichung (5.20), so erhalten wir

$$\delta \mathcal{Z}_1[p_S, \tilde{p}_S] = \operatorname{Re} (f_2(\lambda_1, \mathcal{E}[p_S], p_S, \tilde{p}_S)).$$

□

Ebenso dazu erhalten wir die folgende Aussage für die Gâteaux-Ableitung mit Korollar 5.20.

**Satz 5.26.** *Gilt Voraussetzung 5.16 und  $p_S \in \mathbb{O}_S$ , dann folgt für  $k = 1, 2$  die Gleichung*

$$\delta \mathcal{Z}_k[p_S, \tilde{p}_S] = \nabla \mathcal{Z}_k[p_S] \tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_k[p_S] = \left( \nabla_\omega \mathcal{Z}_k[p_S], [\nabla_{P_{in,l}} \mathcal{Z}_k[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l} \mathcal{Z}_k[p_S]]_{l=1}^{K_A} \right),$$

wobei  $[\nabla_{P_{in,l}} \mathcal{Z}[p_S]]_{l=1}^{K_A}$  und  $[\nabla_{\theta_l} \mathcal{Z}[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}} \mathcal{Z}[p_S]$  bzw.  $\nabla_{\theta_l} \mathcal{Z}[p_S]$  sind. Sei  $\lambda_k$  die entsprechende adjungierte Lösung von  $[D_k]_K$ ,  $k = 1, 2$ , dann gelten die folgenden Gleichungen:

1. Für den Gradienten  $\nabla_\omega \mathcal{Z}_1[p_S]$  gilt

$$\nabla_\omega \mathcal{Z}_1[p_S] = \operatorname{Re} (f_{G_\omega}(\lambda_1, \mathcal{E}[p_S], p_S))$$

und für den Gradienten  $\nabla_\omega \mathcal{Z}_2[p_S]$  gilt

$$\begin{aligned}\nabla_\omega \mathcal{Z}_2[p_S] &= \frac{1}{4} \frac{1}{|V|} \int_V \varepsilon_0(\omega \operatorname{Im}(\nabla \varepsilon_r[\omega]) + \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}[p_S] - E_V|^2 \, d\Omega \\ &\quad + \operatorname{Re} (f_{G_\omega}(\lambda_2, \mathcal{E}[p_S], p_S))\end{aligned}$$

mit  $f_{G_\omega}$  wie in Korollar 5.20.

## 5. Optimalsteuerung elektromagnetischer Felder

---

2. Für den Gradienten  $\nabla_{P_{in,l}} \mathcal{Z}_k[p_S]$ ,  $1 \leq l \leq K_A$  und  $k = 1, 2$ , gilt

$$\nabla_{P_{in,l}} \mathcal{Z}_k[p_S] = -\operatorname{Re} \left( \int_{\Gamma_{P,l}} i\mu_0\omega Y(\omega) \frac{1}{\sqrt{P_{in,l}}\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_k] d\Gamma \right).$$

3. Für den Gradienten  $\nabla_{\theta_l} \mathcal{Z}_k[p_S]$ ,  $1 \leq l \leq K_A$  und  $k = 1, 2$ , gilt

$$\nabla_{\theta_l} \mathcal{Z}_k[p_S] = -2 \operatorname{Re} \left( \int_{\Gamma_{P,l}} \mu_0\omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_k] d\Gamma \right).$$

*Beweis.* Der Beweis dieses Satzes folgt ziemlich analog zu dem des Satzes 5.25. Wir beweisen zunächst die Aussagen nur für  $k = 1$ , denn für  $k = 2$  kann der Beweis komplett analog geführt werden. Wir betrachten demnach (5.29) mit  $\tilde{p}_S$  und ziehen den Realteil-Operator aus dem Integral, also

$$\nabla \mathcal{Z}_1[p_S] \tilde{p}_S = \operatorname{Re} \left( \frac{1}{|V|} \int_V \nabla \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V) d\Omega \right) \tilde{p}_S.$$

Verwenden wir nun die Definition von  $\nabla \mathcal{E}[p_S]$  in Korollar 5.20 und die Summeneigenschaft des Realteil-Operators, so erhalten wir

$$\begin{aligned} \nabla \mathcal{Z}_1[p_S] \tilde{p}_S &= \operatorname{Re} \left( \frac{1}{|V|} \int_V \nabla_\omega \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V) d\Omega \right) \tilde{\omega} \\ &\quad + \sum_{l=1}^{K_A} \operatorname{Re} \left( \frac{1}{|V|} \int_V \nabla_{P_{in,l}} \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V) d\Omega \right) \tilde{P}_{in,l} \\ &\quad + \sum_{l=1}^{K_A} \operatorname{Re} \left( \frac{1}{|V|} \int_V \nabla_{\theta_l} \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V) d\Omega \right) \tilde{\theta}_l. \end{aligned} \tag{5.40}$$

Wir definieren den ersten Faktor des ersten Terms in (5.40) als

$$\nabla_\omega \mathcal{Z}_1[p_S] = \operatorname{Re} \left( \frac{1}{|V|} \int_V \nabla_\omega \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V) d\Omega \right).$$

Verwenden wir nun wie in Satz 5.25 die Definition  $[D_1]_K$  für die Lagrange-Lösung  $\lambda_1$ ,

dann ergibt sich gerade

$$\begin{aligned}\nabla_{\omega} \mathcal{Z}_1[p_S] &= \operatorname{Re} (f_{D_1}(\nabla_{\omega} \mathcal{E}[p_S], \mathcal{E}[p_S], E_V, p_S)) \\ &= \operatorname{Re} (a(\nabla_{\omega} \mathcal{E}[p_S], \lambda_1, p_S)).\end{aligned}$$

Mit der Lösungseigenschaft (5.27) des Gradienten  $\nabla_{\omega} \mathcal{E}[p_S]$  erhalten wir

$$\nabla_{\omega} \mathcal{Z}_1[p_S] = \operatorname{Re} (f_{G_{\omega}}(\lambda_1, \mathcal{E}[p_S], p_S))$$

und damit den ersten Teil unserer Aussage. Analog definieren wir nun

$$\nabla_{P_{in,l}} \mathcal{Z}_1[p_S] = \operatorname{Re} \left( \frac{1}{|V|} \int_V \nabla_{P_{in,l}} \mathcal{E}[p_S] \cdot (\mathcal{E}[p_S] - E_V) \, d\Omega \right), \quad 1 \leq l \leq K_A,$$

und erhalten wiederum

$$\begin{aligned}\nabla_{P_{in,l}} \mathcal{Z}_1[p_S] &= \operatorname{Re} (f_{D_1}(\nabla_{P_{in,l}} \mathcal{E}[p_S], \mathcal{E}[p_S], E_V, p_S)) \\ &= \operatorname{Re} (a(\nabla_{P_{in,l}} \mathcal{E}[p_S], \lambda_1, p_S)),\end{aligned}$$

für  $1 \leq l \leq K_A$ , mit der Definition  $[D_1]_K$  des Lagrange-Problems. Verwenden wir nun die zweite Aussage von Korollar 5.20 zu  $\nabla_{P_{in,l}} \mathcal{E}[p_S]$ , dann erfolgt

$$\nabla_{P_{in,l}} \mathcal{Z}_1[p_S] = \frac{1}{2P_{in,l}} \operatorname{Re} (a(\mathcal{E}[\hat{p}_S], \lambda_1, p_S)),$$

wobei der Parametervektor  $\hat{p}_S$  gegeben ist als

$$\hat{\omega} = \omega, \quad \hat{\theta}_l = \theta_l, \quad \hat{P}_{in,l} = P_{in,l},$$

sowie  $\hat{\theta}_k = \hat{P}_{in,k} = 0$  für  $1 \leq k \leq K_A$  mit  $k \neq l$ . Wird nun noch die Lösungseigenschaft von  $\mathcal{E}[\hat{p}_S]$  zu  $[V_K]$  verwendet, ergibt sich abschließend

$$\nabla_{P_{in,l}} \mathcal{Z}_1[p_S] = \frac{1}{2P_{in,l}} \operatorname{Re} \left( -2 \int_{\Gamma_{P,l}} i\mu_0 \omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_1] \, d\Gamma \right)$$

und damit die zu zeigende Aussage für  $\nabla_{P_{in,l}} \mathcal{Z}_1[p_S]$ . Analog kann nun auch die Aussage zu  $\nabla_{\theta_l} \mathcal{Z}_1[p_S]$  gezeigt werden. Der Beweis für  $k = 2$  folgt ziemlich analog, indem lediglich die Gleichung (5.31) des Satzes 5.22 und die Definition des Lagrange-Problems  $[D_2]_K$  benutzt wird.  $\square$

Wie wir bereits mit Definition 5.11 wissen, gilt für die Probleme  $[O_k]_K$ ,  $k = 1, 2$ , die indirekte Abhängigkeit der elektrischen Felder  $\mathcal{E}[p_S]$  von den Steuerparametern. Dies bedeutet, dass das Optimum durch den Steuerparametervektor  $p_S$  beschrieben wird, weshalb wir mit der Betrachtung der Lagrange-Funktionen (5.37) und (5.39) in Bezug auf den Steuerparametervektor  $p_S$  abschließend die folgenden Optimalitätsbedingungen herleiten können.

**Satz 5.27.** *Gelte Voraussetzung 5.15. Ist das Tupel  $(\hat{p}_S, \hat{E}) \in \mathbb{O} \subset \mathbb{P}_S \times X_e$  lokales Minimum der Funktion  $\mathcal{L}_k(p_S, E, \lambda_k)$ ,  $k = 1, 2$ , definiert durch (5.37) und (5.39) zu gegebenem  $E_V \in X_e$ , wobei  $\hat{E}$  nach Definition 5.11 definiert ist als der Zustand  $\mathcal{E}[\hat{p}_S]$ , sowie*

$$\tilde{p}_S = p_S - \hat{p}_S \in \mathbb{P}_S,$$

für beliebiges  $p_S \in \mathbb{P}_S$ , dann existiert eine eindeutige adjungierte Lösung  $\lambda_k \in X_e$  von  $[D_k]_K$ ,  $k = 1, 2$ , zu  $(\hat{p}_S, \hat{E})$ , sodass die folgenden Variationsungleichungen erfüllt sind

$k = 1$ :

$$\operatorname{Re} \left( f_2(\lambda_1, \hat{E}, \hat{p}_S, \tilde{p}_S) \right) \geq 0. \quad (5.41)$$

$k = 2$ :

$$\frac{1}{4} \frac{1}{|V|} \int_V \varepsilon_0 (\hat{\omega} \operatorname{Im}(\delta \varepsilon_r[\hat{\omega}, \tilde{\omega}]) + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\hat{\omega}))) \left| \hat{E} - E_V \right|^2 d\Omega + \operatorname{Re} \left( f_2(\lambda_2, \hat{E}, \hat{p}_S, \tilde{p}_S) \right) \geq 0. \quad (5.42)$$

Dabei ist  $f_2$  gegeben durch (5.21). Gilt zudem Voraussetzung 5.16, dann lassen sich die Bedingungen (5.41) und (5.42) auch folgendermaßen ausdrücken

$$\nabla \mathcal{Z}_k(\hat{p}_S)(p_S - \hat{p}_S) \geq 0 \quad k = 1, 2. \quad (5.43)$$

*Beweis.* Wir beweisen wiederum nur die Aussage für  $k = 1$ , denn die Aussage für  $k = 2$  folgt analog. Wie zuvor bereits erwähnt, erhalten wir die Eindeutigkeit der adjungierten Lösung  $\lambda_1$  des adjungierten Problems  $[D_1]_K$  zu dem Tupel  $(\hat{p}_S, \hat{E})$  analog zu Abschnitt 3.1. Aus der Minimaleigenschaft des Tupels  $(\hat{p}_S, \hat{E})$  und der Definition des Zustandes  $\hat{E} = \mathcal{E}[\hat{p}_S]$  erhalten wir zudem

$$\mathcal{L}_1(\hat{p}_S, \mathcal{E}[\hat{p}_S], \lambda_1) \leq \mathcal{L}_1(\hat{p}_S + h\tilde{p}_S, \mathcal{E}[\hat{p}_S + h\tilde{p}_S], \lambda_1)$$

für  $h \in [0, 1]$  klein genug. Bringen wir den linken Term auf die andere Seite, multiplizieren mit dem Faktor  $\frac{1}{h}$  und betrachten den Grenzwert  $h \rightarrow 0$ , so ergibt sich

$$0 \leq \lim_{h \rightarrow 0} \frac{1}{h} (\mathcal{L}_1(\hat{p}_S + h\tilde{p}_S, \mathcal{E}[\hat{p}_S + h\tilde{p}_S], \lambda_1) - \mathcal{L}_1(\hat{p}_S, \mathcal{E}[\hat{p}_S], \lambda_1)).$$

Wir erhalten somit die Bedingung an die erste Variation der Lagrange-Funktion

$$0 \leq \delta \mathcal{L}_1[\hat{p}_S, \tilde{p}_S]$$

an der Stelle  $\hat{p}_S \in \mathbb{O}_S$  in Richtung des Parametervektors  $\tilde{p}_S$ . Leiten wir analog zu Satz 5.19 die Darstellung der ersten Variation der Lagrange-Funktion her und verwenden die Lösungseigenschaften von  $\mathcal{E}[\hat{p}_S]$  und  $\lambda_1$ , dann ergibt sich mit denselben Argumenten wie im Beweis von Satz 5.25 gerade

$$\delta \mathcal{L}_1[\hat{p}_S, \tilde{p}_S] = \operatorname{Re}(f_2(\lambda_1, \mathcal{E}[\hat{p}_S], \hat{p}_S, \tilde{p}_S)) \quad (5.44)$$

und damit die erste Aussage. Der zweite Teil des Satzes ergibt sich dann aus Satz 5.25 und Satz 5.26.  $\square$

### 5.1.5. Getrennte Systeme

Bei der Theorie der getrennten Systeme spielt die in der Physik bekannte kohärente Überlagerung von Wellen eine Rolle. Dabei bedeutet kohärente Überlagerung, dass das gesamte Wellenfeld, welches durch verschiedene Einspeisepunkte bei gleicher Frequenz erzeugt wird, als Summe der Wellenfelder der einzelnen Ports dargestellt werden kann. Ist das betrachtete System geschlossen, das heißt die Felderzeugung ergibt sich nur in den Ports, dann kann das elektromagnetische Feld als Summe über die elektromagnetischen Felder der einzelnen Ports dargestellt werden. Wird hingegen ein nicht geschlossenes System betrachtet, muss noch der verbleibende Anteil in der Summe mitbetrachtet werden. In unserem Fall des Variationsproblems  $[V_K]$  erhalten wir das folgende Resultat über getrennte Systeme.

**Satz 5.28.** *Das zur Lösung  $E \in X_e$  zu  $[V_K]$  äquivalente getrennte System ist gegeben durch die Darstellung*

$$E = \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} E_l, \quad (5.45)$$

wobei  $E_l$ ,  $1 \leq l \leq K_A$ , Lösung des folgenden Variationsproblems ist:

- Bestimme  $E_l \in X_e$ , sodass

$$[V_{K,Div_l}] \quad a(E_l, \varphi; \omega) = f_l(\varphi; \omega), \quad \varphi \in X_e,$$

gilt mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$f_l(\varphi; \omega) = -2 \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \frac{1}{\sqrt{|P_{m,l}(\omega)|}} \pi_t [E_{0,l}] \cdot \pi_t [\varphi] \, d\Gamma. \quad (5.46)$$

*Beweis.* Betrachten wir das Variationsproblem  $[V_K]$  und verwenden dann die Darstellung (5.45), dann erhalten wir mit der Linearität der Sesquilinearform  $a$  die Variationsgleichung

$$\sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} a(E_l, \varphi; \omega) = -2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \mu_0 \frac{\sqrt{P_{in,l}}}{\sqrt{|P_{m,l}(\omega)|}} e^{-i\theta_l} \pi_t [E_{0,l}] \cdot \pi_t [\varphi] \, d\Gamma,$$

mit  $\varphi \in X_e$ . Ziehen wir den Faktor  $\sqrt{P_{in,l}} e^{-i\theta_l}$  auf der rechten Seite aus dem Integral und definieren die Lösungen  $E_l$  über die Variationsgleichungen  $[V_{K,Div_l}]$ , dann erhalten wir damit gerade unsere Aussage.  $\square$

Die Besonderheit der getrennten Systeme  $[V_{K, \text{Div}_l}]$  ist, dass die einzelnen Variationsprobleme nur noch von der Frequenz und nicht mehr von den Leistungen und Phasen abhängig sind. Verwenden wir nun Satz 5.28 in den Optimalsteuerungsproblemen von Abschnitt 5.1.2, dann erhalten wir die zu  $[O_k]_K$ ,  $k = 1, 2$ , äquivalenten Optimalsteuerungsprobleme:

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$ , wobei  $E_l \in X_e$ ,  $1 \leq l \leq K_A$  und  $E_l$  Lösung von  $[V_{K, \text{Div}_l}]$  zum Parameter  $\omega$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$[O_1]_{K, \text{Div}}$

$$\min_{\substack{p_S \in \mathbb{P}_S, \\ [E_l]_{l=1}^{K_A} \in X_e}} Z_1(p_S, [E_l]_{l=1}^{K_A}) = \frac{1}{2} \frac{1}{|V|} \int_V \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} E_l - E_V \right|^2 d\Omega$$

unter  $a(E_l, \varphi, \omega) = f_l(\varphi, \omega)$ ,  $\varphi \in X_e$ ,  $1 \leq l \leq K_A$ .

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$ , wobei  $E_l \in X_e$ ,  $1 \leq l \leq K_A$  und  $E_l$  Lösung von  $[V_{K, \text{Div}_l}]$  zum Parameter  $\omega$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$[O_2]_{K, \text{Div}}$

$$\min_{\substack{p_S \in \mathbb{P}_S, \\ [E_l]_{l=1}^{K_A} \in X_e}} Z_2(p_S, [E_l]_{l=1}^{K_A}) = \frac{1}{4} \frac{\omega \varepsilon_0}{|V|} \int_V \text{Im}(\varepsilon_r(\omega)) \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} E_l - E_V \right|^2 d\Omega$$

unter  $a(E_l, \varphi, \omega) = f_l(\varphi, \omega)$ ,  $\varphi \in X_e$ ,  $1 \leq l \leq K_A$ .

Entsprechend der Theorie in Abschnitt 5.1.3 benötigen wir für die Herleitung der ersten Variation und der Gâteaux-Ableitung den folgenden Satz, welcher als Analogon zu Satz 5.19 und Korollar 5.20 verstanden werden kann.

**Satz 5.29.** Sei  $1 \leq l \leq K_A$ . Gilt Voraussetzung 5.15, dann ist die erste Variation  $\delta \mathcal{E}_l(\omega, \tilde{\omega}) \in X_e$ ,  $\omega \in [\omega_{\min}, \omega_{\max}]$  und  $\tilde{\omega} \in \mathbb{R}$ , Lösung des folgenden Variationsproblems

$$a(\delta \mathcal{E}_l[\omega, \tilde{\omega}], \varphi, \omega) = f_{2,l}(\varphi, \mathcal{E}_l[\omega], \omega, \tilde{\omega}), \quad \varphi \in X_e,$$

## 5. Optimalsteuerung elektromagnetischer Felder

---

mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\begin{aligned}
 f_{2,l}(\varphi, \mathcal{E}_l[\omega], \omega, \tilde{\omega}) &= \int_{\Omega} \varepsilon_0 \mu_0 (\omega^2 \delta \varepsilon_r[\omega, \tilde{\omega}] + 2\omega \tilde{\omega} \varepsilon_r(\omega)) \mathcal{E}_l[p_S] \cdot \varphi \, d\Omega \\
 &+ \int_{\Gamma_{IMP}} i\mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}_l[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &+ \sum_{n=1}^{K_A} \int_{\Gamma_{P,n}} i\mu_0 (\omega \delta Y[\omega, \tilde{\omega}] + \tilde{\omega} Y(\omega)) \pi_t[\mathcal{E}_l[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &- 2 \int_{\Gamma_{P,l}} i\mu_0 \left( \frac{\omega \delta Y[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)}} + \frac{\tilde{\omega} Y(\omega)}{\sqrt{P_{m,l}(\omega)}} \right. \\
 &\quad \left. - \frac{\omega}{2} Y(\omega) \frac{\delta P_{m,l}[\omega, \tilde{\omega}]}{\sqrt{P_{m,l}(\omega)^3}} \right) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma.
 \end{aligned}$$

Gilt zudem Voraussetzung 5.16, dann ist die Abbildung  $\omega \mapsto \mathcal{E}_l[\omega]$  Gâteaux-differenzierbar, also

$$\delta \mathcal{E}_l[\omega, \tilde{\omega}] = \nabla_{\omega} \mathcal{E}_l[\omega] \tilde{\omega},$$

wobei  $\nabla_{\omega} \mathcal{E}_l[\omega]$  das folgende Variationsproblem löst

$$a(\nabla_{\omega} \mathcal{E}_l[\omega], \varphi, \omega) = f_{G,l}(\varphi, \mathcal{E}_l[\omega], \omega), \quad \varphi \in X_e,$$

mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\begin{aligned}
 f_{G,l}(\varphi, \mathcal{E}_l[\omega], \omega) &= \int_{\Omega} \varepsilon_0 \mu_0 (\omega^2 \nabla \varepsilon_r[\omega] + 2\omega \varepsilon_r(\omega)) \mathcal{E}_l[p_S] \cdot \varphi \, d\Omega \\
 &+ \int_{\Gamma_{IMP}} i\mu_0 (\omega \nabla Y[\omega] + Y(\omega)) \pi_t[\mathcal{E}_l[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &+ \sum_{n=1}^{K_A} \int_{\Gamma_{P,n}} i\mu_0 (\omega \nabla Y[\omega] + Y(\omega)) \pi_t[\mathcal{E}_l[p_S]] \cdot \pi_t[\varphi] \, d\Gamma \\
 &- 2 \int_{\Gamma_{P,l}} i\mu_0 \left( \frac{\omega \nabla Y[\omega]}{\sqrt{P_{m,l}(\omega)}} + \frac{Y(\omega)}{\sqrt{P_{m,l}(\omega)}} \right. \\
 &\quad \left. - \frac{\omega}{2} Y(\omega) \frac{\nabla P_{m,l}[\omega]}{\sqrt{P_{m,l}(\omega)^3}} \right) \pi_t[E_{0,l}] \cdot \pi_t[\varphi] \, d\Gamma.
 \end{aligned}$$

Weiterhin erhalten wir mit denselben Beweisschritten wie in Satz 5.21 und Satz 5.22 auch für  $[O_k]_{K, \text{Div}}$ ,  $k = 1, 2$ , entsprechende Aussagen.

**Satz 5.30.** *Gilt Voraussetzung 5.15, dann existiert die erste Variation  $\delta \mathcal{Z}_1[p_S, \tilde{p}_S]$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  und ist gegeben durch*

$$\delta \mathcal{Z}_1[p_S, \tilde{p}_S] = \frac{1}{|V|} \int_V \text{Re} \left( \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \delta \mathcal{E}_l[\omega, \tilde{\omega}] + \frac{1}{2\sqrt{P_{in,l}}} \tilde{P}_{in,l} e^{-i\theta_l} \mathcal{E}_l[\omega] - i\sqrt{P_{in,l}} e^{-i\theta_l} \tilde{\theta}_l \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega,$$

wobei  $\delta \mathcal{E}_l[\omega, \tilde{\omega}]$  durch Satz 5.29 gegeben ist. Gilt zudem Voraussetzung 5.16, dann ist

$$\delta \mathcal{Z}_1[p_S, \tilde{p}_S] = \nabla \mathcal{Z}_1[p_S] \tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_1[p_S] = \left( \nabla_\omega \mathcal{Z}_1[p_S], [\nabla_{P_{in,l}} \mathcal{Z}_1[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l} \mathcal{Z}_1[p_S]]_{l=1}^{K_A} \right),$$

wobei  $[\nabla_{P_{in,l}} \mathcal{Z}_1[p_S]]_{l=1}^{K_A}$  und  $[\nabla_{\theta_l} \mathcal{Z}_1[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}} \mathcal{Z}_1[p_S]$  bzw.  $\nabla_{\theta_l} \mathcal{Z}_1[p_S]$  sind, und es gelten die folgenden Gleichungen:

1. Für den Gradienten  $\nabla_\omega \mathcal{Z}_1[p_S]$  und  $\nabla_\omega \mathcal{E}_l[\omega]$  gegeben durch Satz 5.29 gilt

$$\nabla_\omega \mathcal{Z}_1[p_S] = \frac{1}{|V|} \int_V \text{Re} \left( \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \nabla_\omega \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega.$$

2. Für den Gradienten  $\nabla_{P_{in,l}} \mathcal{Z}_1[p_S]$ ,  $1 \leq l \leq K_A$ , gilt

$$\nabla_{P_{in,l}} \mathcal{Z}_1[p_S] = \frac{1}{|V|} \int_V \text{Re} \left( \left( \frac{1}{2\sqrt{P_{in,l}}} e^{-i\theta_l} \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega.$$

3. Für den Gradienten  $\nabla_{\theta_l} \mathcal{Z}_1[p_S]$ ,  $1 \leq l \leq K_A$ , gilt

$$\nabla_{\theta_l} \mathcal{Z}_1[p_S] = -\frac{1}{|V|} \int_V \text{Re} \left( \left( i\sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega.$$

**Satz 5.31.** *Gilt Voraussetzung 5.15, dann existiert die erste Variation  $\delta \mathcal{Z}_2[p_S, \tilde{p}_S]$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  und ist gegeben durch*

$$\begin{aligned} & \delta \mathcal{Z}_2[p_S, \tilde{p}_S] \\ &= \frac{1}{4|V|} \int_V (\omega \operatorname{Im}(\delta \varepsilon_r[\omega, \tilde{\omega}] + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\omega))) \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right|^2 d\Omega \\ &+ \frac{1}{2|V|} \int_V \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re} \left( \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \delta \mathcal{E}_l[\omega, \tilde{\omega}] + \frac{1}{2\sqrt{P_{in,l}}} \tilde{P}_{in,l} e^{-i\theta_l} \mathcal{E}_l[\omega] \right. \right. \\ &\quad \left. \left. - i\sqrt{P_{in,l}} e^{-i\theta_l} \tilde{\theta}_l \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega, \end{aligned}$$

wobei  $\delta \mathcal{E}_l[\omega, \tilde{\omega}]$  durch Satz 5.29 gegeben ist. Gilt zudem Voraussetzung 5.16, dann ist

$$\delta \mathcal{Z}_2[p_S, \tilde{p}_S] = \nabla \mathcal{Z}_2[p_S] \tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_2[p_S] = \left( \nabla_\omega \mathcal{Z}_2[p_S], [\nabla_{P_{in,l}} \mathcal{Z}_2[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l} \mathcal{Z}_2[p_S]]_{l=1}^{K_A} \right),$$

wobei  $[\nabla_{P_{in,l}} \mathcal{Z}_2[p_S]]_{l=1}^{K_A}$  und  $[\nabla_{\theta_l} \mathcal{Z}_2[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}} \mathcal{Z}_2[p_S]$  bzw.  $\nabla_{\theta_l} \mathcal{Z}_2[p_S]$  sind, und es gelten die folgenden Gleichungen:

1. Für den Gradienten  $\nabla_\omega \mathcal{Z}_2[p_S]$  und  $\nabla_\omega \mathcal{E}_l[\omega]$  gegeben durch Satz 5.29 gilt

$$\begin{aligned} \nabla_\omega \mathcal{Z}_2[p_S] &= \frac{1}{4|V|} \int_V (\omega \operatorname{Im}(\nabla_\omega \varepsilon_r(\omega)) + \operatorname{Im}(\varepsilon_r(\omega))) \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right|^2 d\Omega \\ &+ \frac{1}{2|V|} \int_V \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re} \left( \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \nabla_\omega \mathcal{E}_l[\omega] \right) \right. \\ &\quad \left. \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega. \end{aligned}$$

2. Für den Gradienten  $\nabla_{P_{in,l}} \mathcal{Z}_2[p_S]$ ,  $1 \leq l \leq K_A$ , gilt

$$\nabla_{P_{in,l}} \mathcal{Z}_2[p_S] = \frac{1}{2|V|} \int_V \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re} \left( \left( \frac{1}{2\sqrt{P_{in,l}}} e^{-i\theta_l} \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega.$$

3. Für den Gradienten  $\nabla_{\theta_l} \mathcal{Z}_1[p_S]$ ,  $1 \leq l \leq K_A$ , gilt

$$\nabla_{\theta_l} \mathcal{Z}_1[p_S] = -\frac{1}{2|V|} \int_V \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re} \left( \left( i\sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega.$$

Verwenden wir, wie in Abschnitt 5.1.4, die formale Lagrangetechnik zur Herleitung der adjungierten Probleme, dann ergeben sich folgende Probleme und die dazugehörigen Lagrange-Funktionen.

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $E_m \in X_e$ ,  $1 \leq m \leq K_A$ , des Optimierungsproblems  $[\mathbf{O}_1]_{K, \operatorname{Div}}$  die adjungierten Lösungen  $\lambda_l \in X_e$ ,  $1 \leq l \leq K_A$ , sodass

$$[\mathbf{D}_1]_{K, \operatorname{Div}_l} \quad a(\varphi, \lambda_l, \omega) = f_{D_1,l}(\varphi, [E_m]_{m=1}^{K_A}, p_S), \quad \varphi \in X_e,$$

gilt, mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[\mathbf{V}_K]$  und

$$f_{D_1,l}(\varphi, [E_m]_{m=1}^{K_A}, p_S) = \frac{1}{|V|} \int_V (\sqrt{P_{in,l}} e^{-i\theta_l} \varphi) \cdot \left( \sum_{m=1}^{K_A} \sqrt{P_{in,m}} e^{-i\theta_m} E_m - E_V \right) d\Omega.$$

**Definition 5.32.** Die Lagrange-Funktion zu dem Optimierungsproblem  $[O_1]_{K, \text{Div}_l}$  ist gegeben durch

$$\begin{aligned} \mathcal{L}_{1, \text{Div}}(p_S, [E_l]_{l=1}^{K_A}, [\lambda_l]_{l=1}^{K_A}) \\ = \frac{1}{2} \frac{1}{|V|} \int_V \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} E_l - E_V \right|^2 d\Omega - \sum_{l=1}^{K_A} \text{Re}(a(E_l, \lambda_l, \omega)) \\ - 2 \sum_{l=1}^{K_A} \text{Re} \left( \int_{\Gamma_{P,l}} i\omega Y(\omega) \frac{1}{\sqrt{P_{m,l}(\omega)}} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_l] d\Gamma \right), \end{aligned}$$

wobei  $p_S \in \mathbb{P}_S$  und  $[E_l]_{l=1}^{K_A}, [\lambda_l]_{l=1}^{K_A}, E_V \in X_e$ .

Zudem erhalten wir auch noch die Variationsprobleme der adjungierten Lösungen des Optimierungsproblems  $[O_2]_{K, \text{Div}}$ :

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $E_m \in X_e$ ,  $1 \leq m \leq K_A$ , des Optimierungsproblems  $[O_2]_{K, \text{Div}}$  die adjungierten Lösungen  $\lambda_l \in X_e$ ,  $1 \leq l \leq K_A$ , sodass

$$[D_2]_{K, \text{Div}_l} \quad a(\varphi, \lambda_l, \omega) = f_{D_2,l}(\varphi, [E_m]_{m=1}^{K_A}, p_S), \quad \varphi \in X_e,$$

gilt, mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_K]$  und

$$\begin{aligned} f_{D_2,l}(\varphi, [E_m]_{m=1}^{K_A}, p_S) \\ = \frac{1}{2|V|} \int_V \omega \varepsilon_0 \text{Im}(\varepsilon_r(\omega)) (\sqrt{P_{in,l}} e^{-i\theta_l} \varphi) \cdot \left( \sum_{m=1}^{K_A} \sqrt{P_{in,m}} e^{-i\theta_m} E_m - E_V \right) d\Omega. \end{aligned}$$

**Definition 5.33.** Die Lagrange-Funktion zu dem Optimierungsproblem  $[O_2]_{K, \text{Div}_l}$  ist gegeben durch

$$\begin{aligned} \mathcal{L}_{2, \text{Div}}(p_S, [E_l]_{l=1}^{K_A}, [\lambda_l]_{l=1}^{K_A}) \\ = \frac{1}{4} \frac{1}{|V|} \int_V \omega \varepsilon_0 \text{Im}(\varepsilon_r(\omega)) \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} E_l - E_V \right|^2 d\Omega - \sum_{l=1}^{K_A} \text{Re}(a(E_l, \lambda_l, \omega)) \\ - 2 \sum_{l=1}^{K_A} \text{Re} \left( \int_{\Gamma_{P,l}} i\omega Y(\omega) \frac{1}{\sqrt{P_{m,l}(\omega)}} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_l] d\Gamma \right), \end{aligned}$$

wobei  $p_S \in \mathbb{P}_S$  und  $[E_l]_{l=1}^{K_A}, [\lambda_l]_{l=1}^{K_A}, E_V \in X_e$ .

Mit den Variationsproblemen  $[D_k]_{K, \text{Div}_l}$ ,  $k = 1, 2$  und  $1 \leq l \leq K_A$ , erhalten wir wie in Abschnitt 5.1.4 abschließend eine andere Darstellung der ersten Variation und der Gâteaux-Ableitungen der Zielfunktionale.

**Satz 5.34.** *Gilt Voraussetzung 5.15,  $p_S \in \mathbb{O}_S$  und  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$ , dann folgt:*

1. Die erste Variation  $\delta \mathcal{Z}_1[p_S, \tilde{p}_S]$  ist gegeben durch

$$\delta \mathcal{Z}_1[p_S, \tilde{p}_S] = \sum_{l=1}^{K_A} \text{Re} (f_{2,l}(\lambda_l, \mathcal{E}_l[\omega], \omega, \tilde{\omega})) + \frac{1}{|V|} \int_V \text{Re} \left( \left( \sum_{l=1}^{K_A} \frac{\tilde{P}_{in,l} e^{-i\theta_l}}{2\sqrt{P_{in,l}}} \mathcal{E}_l[\omega] - i\sqrt{P_{in,l}} e^{-i\theta_l} \tilde{\theta}_l \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega,$$

wobei  $[\lambda_l]_{l=1}^{K_A} \in X_e$  die adjungierten Lösungen zu  $[D_1]_{K, \text{Div}_l}$ ,  $1 \leq l \leq K_A$ , sind.

2. Die erste Variation  $\delta \mathcal{Z}_2[p_S, \tilde{p}_S]$  ist gegeben durch

$$\begin{aligned} & \delta \mathcal{Z}_2[p_S, \tilde{p}_S] \\ &= \frac{1}{4|V|} \int_V (\omega \text{Im}(\delta \varepsilon_r[\omega, \tilde{\omega}] + \tilde{\omega} \text{Im}(\varepsilon_r(\omega))) \left| \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right|^2 d\Omega \\ &+ \sum_{l=1}^{K_A} \text{Re} (f_{2,l}(\lambda_l, \mathcal{E}_l[\omega], \omega, \tilde{\omega})) + \frac{1}{2|V|} \int_V \omega \varepsilon_0 \text{Im}(\varepsilon_r(\omega)) \text{Re} \left( \left( \sum_{l=1}^{K_A} \frac{\tilde{P}_{in,l} e^{-i\theta_l}}{2\sqrt{P_{in,l}}} \mathcal{E}_l[\omega] - i\sqrt{P_{in,l}} e^{-i\theta_l} \tilde{\theta}_l \mathcal{E}_l[\omega] \right) \cdot \left( \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right) \right) d\Omega, \end{aligned}$$

wobei  $[\lambda_l]_{l=1}^{K_A} \in X_e$  die adjungierten Lösungen zu  $[D_2]_{K, \text{Div}_l}$ ,  $1 \leq l \leq K_A$ , sind.

Gilt zudem Voraussetzung 5.16, dann gilt für  $k = 1, 2$  die Gleichung

$$\delta \mathcal{Z}_k[p_S, \tilde{p}_S] = \nabla \mathcal{Z}_k[p_S] \tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_k[p_S] = \left( \nabla_{\omega} \mathcal{Z}_k[p_S], [\nabla_{P_{in,l}} \mathcal{Z}_k[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l} \mathcal{Z}_k[p_S]]_{l=1}^{K_A} \right),$$

wobei  $[\nabla_{P_{in,l}} \mathcal{Z}_k[p_S]]_{l=1}^{K_A}$  und  $[\nabla_{\theta_l} \mathcal{Z}_k[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}} \mathcal{Z}_k[p_S]$  bzw.  $\nabla_{\theta_l} \mathcal{Z}_k[p_S]$  sind. Dabei ist  $\nabla_{P_{in,l}} \mathcal{Z}_k[p_S], \nabla_{\theta_l} \mathcal{Z}_k[p_S]$ ,  $k = 1, 2$  und  $1 \leq l \leq K_A$ , wie in Satz 5.30 und Satz 5.31, sowie

## 5. Optimalsteuerung elektromagnetischer Felder

---

3. Die Gâteaux-Ableitung  $\nabla_{\omega} \mathcal{Z}_1[p_S]$  ist gegeben durch

$$\nabla_{\omega} \mathcal{Z}_1[p_S] = \sum_{l=1}^{K_A} \operatorname{Re} (f_{G,l}(\lambda_l, \mathcal{E}_l[\omega], \omega)),$$

wobei  $[\lambda_l]_{l=1}^{K_A} \in X_e$  die adjungierten Lösungen zu  $[D_1]_{K, \operatorname{Div}_l}$ ,  $1 \leq l \leq K_A$ , sind.

4. Die Gâteaux-Ableitung  $\nabla_{\omega} \mathcal{Z}_2[p_S]$  ist gegeben durch

$$\begin{aligned} \nabla_{\omega} \mathcal{Z}_2[p_S] = & \frac{1}{4|V|} \int_V (\omega \operatorname{Im}(\nabla_{\omega} \varepsilon_r(\omega)) + \operatorname{Im}(\varepsilon_r(\omega))) \left| \sum_{l=1}^{K_A} \sqrt{P_{m,l}} e^{-i\theta_l} \mathcal{E}_l[\omega] - E_V \right|^2 d\Omega \\ & + \sum_{l=1}^{K_A} \operatorname{Re} (f_{G,l}(\lambda_l, \mathcal{E}_l[\omega], \omega)), \end{aligned}$$

wobei  $[\lambda_l]_{l=1}^{K_A} \in X_e$  die adjungierten Lösungen zu  $[D_2]_{K, \operatorname{Div}_l}$ ,  $1 \leq l \leq K_A$ , sind.

## 5.2. Diskrete Optimalsteuerungsprobleme

Um  $[O_k]_K$ ,  $k = 1, 2$ , numerisch zu lösen, diskretisieren wir das Problem und verwenden als Nebenbedingung das Variationsproblem  $[V_{D,\mathcal{N}}]$ :

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$  und  $E_{\mathcal{N}} \in X_{\mathcal{N}}$ , wobei  $E_{\mathcal{N}}$  Lösung von  $[V_{D,\mathcal{N}}]$  zum Parametervektor  $p_S$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[O_1]_{D,\mathcal{N}} \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ E_{\mathcal{N}} \in X_{\mathcal{N}}}} Z_1(p_S, E_{\mathcal{N}}) = \frac{1}{2|V|} \|E_{\mathcal{N}} - E_V\|_{[L^2(V)]^3}^2$$

unter  $a(E_{\mathcal{N}}, \varphi, p_S) = f(\varphi, p_S), \quad \varphi \in X_{\mathcal{N}}.$

- Finde zu  $E_V \in X_e$  jenes  $p_S \in \mathbb{P}_S$  und  $E_{\mathcal{N}} \in X_{\mathcal{N}}$ , wobei  $E_{\mathcal{N}}$  Lösung von  $[V_{D,\mathcal{N}}]$  zum Parametervektor  $p_S$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[O_2]_{D,\mathcal{N}} \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ E_{\mathcal{N}} \in X_{\mathcal{N}}}} Z_2(p_S, E_{\mathcal{N}}) = \frac{1}{2|V|} \|E_{\mathcal{N}} - E_V\|_{P(V,p_S)}^2$$

unter  $a(E_{\mathcal{N}}, \varphi, p_S) = f(\varphi, p_S), \quad \varphi \in X_{\mathcal{N}}.$

Neben der optimalen Steuerung zu dem Vektorfeld  $E_V \in X_e$  können wir dieselben Probleme auch mit diskretisiertem Feld  $E_{V,\mathcal{N}} \in X_{\mathcal{N}} \subset X_e$  betrachten, was vor allem für Abschnitt 5.4 von Interesse ist. Analog zu Definition 5.11 führen wir den Steuerungs-Zustands-Operator

$$\begin{aligned} \mathcal{E}_{\mathcal{N}} : \mathbb{P}_S &\rightarrow X_{\mathcal{N}}, \\ p_S &\mapsto \mathcal{E}_{\mathcal{N}}[p_S], \end{aligned}$$

für die Lösung der diskreten Variationsformulierung  $[V_{D,\mathcal{N}}]$  ein. Mithilfe von Satz 3.20 über die Existenz und Eindeutigkeit einer Lösung  $\mathcal{E}_{\mathcal{N}}[p_S]$  erhalten wir dann folgende Bemerkung.

*Bemerkung 5.35.* Alle Aussagen aus den Abschnitten 5.1.1 bis 5.1.5 sind ebenfalls gültig für die Steuerungsprobleme  $[O_k]_{D,\mathcal{N}}$ ,  $k = 1, 2$ , wenn jede Funktion in  $X_e$  durch ihre Diskretisierung in  $X_{\mathcal{N}}$  ersetzt wird.

### 5.2.1. Fehlertheorie

Wir betrachten zunächst den Fehler zwischen kontinuierlicher und diskreter Lösung des Optimierungsproblems  $[O_1]_K$  zu  $E_V \in X_e$  und erhalten hierfür die folgende Aussage.

**Satz 5.36.** 1. Sei  $\hat{p}_S \in \mathbb{P}_S$  globale Lösung von  $[O_1]_{D, \mathcal{N}}$  zum konformen Gitter  $\mathcal{T}_h(\Omega)$  und dazugehörigem Raum  $X_{\mathcal{N}} \subset X_e$ , sowie  $\tilde{p}_S \in \mathbb{P}_S$  globale Lösung von  $[O_1]_K$ , dann ist der Fehler in der Zielfunktion bezüglich  $\hat{p}_S$  und  $\tilde{p}_S$  für  $[O_1]_K$  gegeben durch

$$|\mathcal{Z}_1[\hat{p}_S]^{\frac{1}{2}} - \mathcal{Z}_1[\tilde{p}_S]^{\frac{1}{2}}| \leq \frac{C_2}{\sqrt{2|V|}} \left( \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\hat{p}_S] - U_{\mathcal{N}}\|_{X_e} \right. \\ \left. + \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\tilde{p}_S] - U_{\mathcal{N}}\|_{X_e} \right), \quad (5.47)$$

mit  $\mathcal{Z}_1$  aus (5.15) und  $C_2$  aus Satz 3.20.

2. Sei  $\mathcal{T}_{h_n}(\Omega)$  eine reguläre Folge konformer Gitter, wobei  $h_n \xrightarrow{n \rightarrow \infty} 0$  gilt, und  $\mathcal{N}_n$  die dazugehörige Anzahl an Freiheitsgraden, für die  $\mathcal{N}_n \xrightarrow{n \rightarrow \infty} \infty$  gilt. Gibt es einen Parameter  $\delta \in (\frac{1}{2}, 1]$ , sodass für alle Lösungen

$$\mathcal{E}[p_S] \in H^\delta(\text{curl}; \Omega), \quad \pi_t[\mathcal{E}[p_S]] \in H^\delta(\text{curl}; \Gamma)$$

mit  $p_S \in \mathbb{P}_S$  gilt, dann existiert eine konvergente Teilfolge  $(p_{S, n_m})_{n_m}$  zu den globalen Lösungen  $(p_{S, n})_n \in \mathbb{P}_S$  von  $[O_1]_{D, \mathcal{N}_n}$  und diese konvergiert gegen die globale Lösung  $\tilde{p}_S$  von  $[O_1]_K$ .

*Beweis.* Da  $\tilde{p}_S$  globale Lösung von  $[O_1]_K$  ist, erhalten wir mit der normalen und umgekehrten Dreiecksungleichung

$$\begin{aligned} 0 &\leq \mathcal{Z}_1[\hat{p}_S]^{\frac{1}{2}} - \mathcal{Z}_1[\tilde{p}_S]^{\frac{1}{2}} \\ &= \frac{1}{\sqrt{2|V|}} \left( \|\mathcal{E}[\hat{p}_S] - E_V\|_{[L^2(V)]^3} - \|\mathcal{E}[\tilde{p}_S] - E_V\|_{[L^2(V)]^3} \right) \\ &\leq \frac{1}{\sqrt{2|V|}} \left( \|\mathcal{E}[\hat{p}_S] - \mathcal{E}_{\mathcal{N}}[\hat{p}_S]\|_{[L^2(V)]^3} + \|\mathcal{E}_{\mathcal{N}}[\hat{p}_S] - E_V\|_{[L^2(V)]^3} - \|\mathcal{E}[\tilde{p}_S] - E_V\|_{[L^2(V)]^3} \right) \\ &\leq \frac{1}{\sqrt{2|V|}} \left( \|\mathcal{E}[\hat{p}_S] - \mathcal{E}_{\mathcal{N}}[\hat{p}_S]\|_{[L^2(V)]^3} + \|\mathcal{E}_{\mathcal{N}}[\tilde{p}_S] - E_V\|_{[L^2(V)]^3} - \|\mathcal{E}[\tilde{p}_S] - E_V\|_{[L^2(V)]^3} \right) \\ &\leq \frac{1}{\sqrt{2|V|}} \left( \|\mathcal{E}[\hat{p}_S] - \mathcal{E}_{\mathcal{N}}[\hat{p}_S]\|_{[L^2(V)]^3} + \|\mathcal{E}[\tilde{p}_S] - \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]\|_{[L^2(V)]^3} \right) \\ &\leq \frac{1}{\sqrt{2|V|}} \left( \|\mathcal{E}[\hat{p}_S] - \mathcal{E}_{\mathcal{N}}[\hat{p}_S]\|_{X_e} + \|\mathcal{E}[\tilde{p}_S] - \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]\|_{X_e} \right), \end{aligned}$$

wobei wir die Minimierungseigenschaft der globalen Lösung  $\hat{p}_S$  des diskreten Optimierungsproblems  $[O_1]_{D,\mathcal{N}}$  in der Ungleichung von Zeile drei auf Zeile vier ausgenutzt haben. Verwenden wir nun die Approximationseigenschaft (3.67) aus Satz 3.20, so erhalten wir die erste Aussage.

Für den Beweis der zweiten Aussage betrachten wir die Folge an globalen Lösungen  $(p_{S,n})_n \in \mathbb{P}_S$ , für die aufgrund der Kompaktheit von  $\mathbb{P}_S$  eine konvergente Teilfolge  $(p_{S,n_m})_{n_m}$  mit Grenzwert  $\tilde{p}_S \in \mathbb{P}_S$  existiert. Um die Notation zu vereinfachen bezeichnen wir die konvergente Teilfolge wieder mit  $(p_{S,n})_n$ . Nach Satz 5.3 für den diskretisierten Fall gilt für die Folge an Lösungen  $(\mathcal{E}_{\mathcal{N}_n}[p_{S,n}])_n \in X_{\mathcal{N}_n}$  zu einem festen konformen Gitter  $\mathcal{T}_{h_n}(\Omega)$  gerade

$$\|\mathcal{E}_{\mathcal{N}_n}[p_{S,n}] - \mathcal{E}_{\mathcal{N}_n}[\tilde{p}_S]\|_{X_e} \xrightarrow{n \rightarrow \infty} 0.$$

Des Weiteren haben wir für einen festen Parametervektor  $p_S \in \mathbb{P}_S$  nach Voraussetzung und Satz 3.20

$$\|\mathcal{E}_{\mathcal{N}_n}[p_S] - \mathcal{E}[p_S]\|_{X_e} \leq C_3 h_n^\delta (\|\mathcal{E}[p_S]\|_{H^\delta(\text{curl};\Omega)} + \|\pi_t[\mathcal{E}[p_S]]\|_{H^\delta(\text{curl}_\Gamma;\Gamma)}) \xrightarrow{n \rightarrow \infty} 0.$$

Kombinieren wir die beiden Konvergenzaussagen, dann erhalten wir mit einem Diagonalfolgenargument

$$\|\mathcal{E}_{\mathcal{N}_n}[p_{S,n}] - \mathcal{E}[\tilde{p}_S]\|_{X_e} \xrightarrow{n \rightarrow \infty} 0.$$

Daraus folgt direkt

$$|\mathcal{Z}_1[p_{S,n}]^{\frac{1}{2}} - \mathcal{Z}_1[\tilde{p}_S]^{\frac{1}{2}}| \leq \frac{1}{\sqrt{2|V|}} \|\mathcal{E}_{\mathcal{N}_n}[p_{S,n}] - \mathcal{E}[\tilde{p}_S]\|_{X_e} \xrightarrow{n \rightarrow \infty} 0$$

und mit der Stetigkeit der Wurzelfunktion

$$\mathcal{Z}_1[p_{S,n}] \xrightarrow{n \rightarrow \infty} \mathcal{Z}_1[\tilde{p}_S].$$

Als Letztes zeigen wir noch die globale Lösungseigenschaft des Steuerparametervektors  $\tilde{p}_S$ . Diese folgt durch die eben hergeleitete Konvergenz und der Lösungseigenschaft der Folge  $p_{S,n}$ , denn

$$\mathcal{Z}_1[\tilde{p}_S] = \lim_{n \rightarrow \infty} \mathcal{Z}_1(p_{S,n}, \mathcal{E}_{\mathcal{N}_n}[p_{S,n}]) \leq \lim_{n \rightarrow \infty} \mathcal{Z}_1(p_S, \mathcal{E}_{\mathcal{N}_n}[p_S]) = \mathcal{Z}_1[p_S]$$

gilt für alle  $p_S \in \mathbb{P}_S$ , was wiederum die globale Lösungseigenschaft ist.  $\square$

Mit demselben Beweis und unter Beachtung der Stetigkeit der Abbildung  $\text{Im}(\varepsilon_r(\omega))$ , sowie Lemma 5.7 erhalten wir analog zu Satz 5.36 die Aussage über den Fehler für das Optimierungsproblem  $[O_2]_K$ .

**Satz 5.37.** 1. Sei  $\hat{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}_2]_{\mathbf{D},\mathcal{N}}$  zum konformen Gitter  $\mathcal{T}_h(\Omega)$  und dazugehörigem Raum  $X_{\mathcal{N}} \subset X_e$ , sowie  $\tilde{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}_2]_{\mathbf{K}}$ , dann ist der Fehler in der Zielfunktion bezüglich  $\hat{p}_S$  und  $\tilde{p}_S$  für  $[\mathbf{O}_2]_{\mathbf{K}}$  gegeben durch

$$\begin{aligned} |\mathcal{Z}_2[\hat{p}_S]^{\frac{1}{2}} - \mathcal{Z}_2[\tilde{p}_S]^{\frac{1}{2}}| &\leq \frac{C_2}{\sqrt{2|V|}} \left( \hat{\omega}^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} \left( \max_{x \in V} (\operatorname{Im}(\varepsilon_r[\hat{\omega}])) \right) \right)^{\frac{1}{2}} \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\hat{p}_S] - U_{\mathcal{N}}\|_{X_e} \\ &\quad + \tilde{\omega}^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} \left( \max_{x \in V} (\operatorname{Im}(\varepsilon_r[\tilde{\omega}])) \right)^{\frac{1}{2}} \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\tilde{p}_S] - U_{\mathcal{N}}\|_{X_e}, \end{aligned} \quad (5.48)$$

mit  $\mathcal{Z}_2$  aus (5.16) und  $C_2$  aus Satz 3.20.

2. Sei  $\mathcal{T}_{h_n}(\Omega)$  eine reguläre Folge konformer Gitter, wobei  $h_n \xrightarrow{n \rightarrow \infty} 0$  gilt, und  $\mathcal{N}_n$  die dazugehörige Anzahl an Freiheitsgraden, für die  $\mathcal{N}_n \xrightarrow{n \rightarrow \infty} \infty$  gilt. Gibt es einen Parameter  $\delta \in (\frac{1}{2}, 1]$ , sodass für alle Lösungen

$$\mathcal{E}[p_S] \in H^\delta(\operatorname{curl}; \Omega), \quad \pi_t[\mathcal{E}[p_S]] \in H^\delta(\operatorname{curl}_\Gamma; \Gamma)$$

mit  $p_S \in \mathbb{P}_S$  gilt, dann existiert eine konvergente Teilfolge  $(p_{S,n_m})_{n_m}$  zu den globalen Lösungen  $(p_{S,n})_n \in \mathbb{P}_S$  von  $[\mathbf{O}_2]_{\mathbf{D},\mathcal{N}_n}$  und diese konvergiert gegen die globale Lösung  $\tilde{p}_S$  von  $[\mathbf{O}_2]_{\mathbf{K}}$ .

Als nächstes betrachten wir den Fehler in der Approximation des Zielfeldes  $E_V$  durch das diskretisierte Feld  $E_{V,\mathcal{N}}$ .

**Satz 5.38.** Für  $k = 1, 2$  gilt:

1. Sei  $\hat{p}_S^{E_V,\mathcal{N}} \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}_k]_{\mathbf{D},\mathcal{N}}$  zu  $E_{V,\mathcal{N}}$  zum konformen Gitter  $\mathcal{T}_h(\Omega)$  und dazugehörigem Raum  $X_{\mathcal{N}} \subset X_e$ , sowie  $\tilde{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}_k]_{\mathbf{K}}$  zu  $E_V$ , dann ist der Fehler in der Zielfunktion bezüglich  $\hat{p}_S^{E_V,\mathcal{N}}$  und  $\tilde{p}_S$  für  $[\mathbf{O}_k]_{\mathbf{K}}$  gegeben durch

$$\begin{aligned} |\mathcal{Z}_1[\hat{p}_S^{E_V,\mathcal{N}}]^{\frac{1}{2}} - \mathcal{Z}_1[\tilde{p}_S]^{\frac{1}{2}}| &\leq \frac{1}{\sqrt{2|V|}} \left( C_2 \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\hat{p}_S^{E_V,\mathcal{N}}] - U_{\mathcal{N}}\|_{X_e} + C_2 \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\tilde{p}_S] - U_{\mathcal{N}}\|_{X_e} \right. \\ &\quad \left. + 2\|E_V - E_{V,\mathcal{N}}\|_{[L^2(V)]^3} \right) \end{aligned} \quad (5.49)$$

und

$$\begin{aligned}
 & |\mathcal{Z}_2[\hat{p}_S^{E_{V,\mathcal{N}}}]^{\frac{1}{2}} - \mathcal{Z}_2[\tilde{p}_S, E_{V,\mathcal{N}}]^{\frac{1}{2}}| \\
 & \leq \frac{1}{\sqrt{2|V|}} \left( C_2 \hat{\omega}^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} \left( \max_{x \in V} (\operatorname{Im}(\varepsilon_r[\hat{\omega}])) \right)^{\frac{1}{2}} \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\hat{p}_S^{E_{V,\mathcal{N}}}] - U_{\mathcal{N}}\|_{X_e} \right. \\
 & \quad + C_2 \tilde{\omega}^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} \left( \max_{x \in V} (\operatorname{Im}(\varepsilon_r[\tilde{\omega}])) \right)^{\frac{1}{2}} \inf_{U_{\mathcal{N}} \in X_{\mathcal{N}}} \|\mathcal{E}[\tilde{p}_S] - U_{\mathcal{N}}\|_{X_e} \\
 & \quad + \hat{\omega}^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} \left( \max_{x \in V} (\operatorname{Im}(\varepsilon_r[\hat{\omega}])) \right)^{\frac{1}{2}} \|E_V - E_{V,\mathcal{N}}\|_{[L^2(V)]^3} \\
 & \quad \left. + \tilde{\omega}^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} \left( \max_{x \in V} (\operatorname{Im}(\varepsilon_r[\tilde{\omega}])) \right)^{\frac{1}{2}} \|E_V - E_{V,\mathcal{N}}\|_{[L^2(V)]^3} \right),
 \end{aligned} \tag{5.50}$$

mit  $C_2$  aus Satz 3.20.

2. Sei  $\mathcal{T}_{h_n}(\Omega)$  eine reguläre Folge konformer Gitter, wobei  $h_n \xrightarrow{n \rightarrow \infty} 0$  gilt, und  $\mathcal{N}_n$  die dazugehörige Anzahl an Freiheitsgraden, für die  $\mathcal{N}_n \xrightarrow{n \rightarrow \infty} \infty$  gilt. Gibt es einen Parameter  $\delta \in (\frac{1}{2}, 1]$ , sodass  $E_V \in H^\delta(\operatorname{curl}; \Omega)$ , sowie  $E_{V,\mathcal{N}} = \Pi_{\mathcal{N}}[E_V]$  mit Interpolationsoperator (3.60) und für alle Lösungen

$$\mathcal{E}[p_S] \in H^\delta(\operatorname{curl}; \Omega), \quad \pi_t[\mathcal{E}[p_S]] \in H^\delta(\operatorname{curl}; \Gamma)$$

mit  $p_S \in \mathbb{P}_S$  gilt, dann existiert eine konvergente Teilfolge  $(p_{S,n_m})_{n_m}$  zu den globalen Lösungen  $(p_{S,n})_n \in \mathbb{P}_S$  von  $[O_k]_{\mathbb{D}, \mathcal{N}_n}$  und diese konvergiert gegen die globale Lösung  $\tilde{p}_S$  von  $[O_k]_{\mathbb{K}}$ .

*Beweis.* Die erste Aussage folgt mit denselben Argumenten wie im Beweis von Satz 5.36, indem neben  $\mathcal{E}_{\mathcal{N}}[\hat{p}_S^{E_{V,\mathcal{N}}}]$  auch  $E_{V,\mathcal{N}}$  als Zwischenterm eingefügt wird. Wir erhalten demnach nach der Anwendung der Dreiecksungleichung anstatt

$$\|\mathcal{E}_{\mathcal{N}}[\hat{p}_S^{E_{V,\mathcal{N}}}] - E_V\|_{[L^2(V)]^3},$$

den Anteil

$$\|\mathcal{E}_{\mathcal{N}}[\hat{p}_S^{E_{V,\mathcal{N}}}] - E_{V,\mathcal{N}}\|_{[L^2(V)]^3} + \|E_{V,\mathcal{N}} - E_V\|_{[L^2(V)]^3}.$$

Für den ersten Term lässt sich dann die Lösungseigenschaft für eine weitere Abschätzung ausnutzen, wobei dann  $\hat{p}_S^{E_{V,\mathcal{N}}}$  durch  $\tilde{p}_S$  ersetzt wird. Daraufhin wird wieder  $E_{V,\mathcal{N}}$  als Zwischenterm eingefügt und die Dreiecksungleichung angewandt. Danach folgt die Aussage wieder analog zu Satz 5.36. Auf dieselbe Art folgt auch die Behauptung für  $k = 2$ . Der zweite Teil des Satzes ergibt sich wiederum analog zu Satz 5.36, nur dass zusätzlich Satz 3.20 für die starke Konvergenz von  $E_{V,\mathcal{N}}$  gegen  $E_V$  verwendet wird.  $\square$

Wir sehen damit, dass der Fehler, den wir in unseren Zielfunktionalen generieren, direkt abhängig von der Approximationseigenschaft der FEM-Diskretisierung ist, weshalb wir bei der Diskretisierung entsprechend feine Gitter  $\mathcal{T}_h(\Omega)$  verwenden. Wir erhalten analog ähnliche Aussagen wie eben auch für die diskretisierten getrennten Systeme  $[O_k]_{K, \text{Div}_l}$ ,  $k = 1, 2$ . Hierzu bezeichnen wir mit  $[V_{D, \text{Div}_l, \mathcal{N}_l}]$  die diskretisierte Formulierung von  $[V_{K, \text{Div}_l}]$  zu den jeweiligen Gittern  $\mathcal{T}_{h_l}(\Omega)$ , wobei  $\mathcal{N}_l$  die Anzahl an Freiheitsgraden bezeichnet. In den einzelnen Sätzen ergibt sich dann eine Summe der Approximationseigenschaften

$$\inf_{U_{\mathcal{N}_l} \in X_{\mathcal{N}_l}} \|\mathcal{E}_l[\hat{\omega}] - U_{\mathcal{N}_l}\|_{X_e}$$

der einzelnen getrennten Probleme  $[V_{D, \text{Div}_l, \mathcal{N}_l}]$ ,  $1 \leq l \leq K_A$ .

### 5.2.2. $L^2$ -Feldoptimierung

Die in Abschnitt 5.1.2 hergeleiteten Optimierungsprobleme  $[O_k]_{K}$ ,  $k = 1, 2$ , suchen das optimale elektromagnetische Feld  $E$  zu einem gegebenen Vektorfeld  $E_V$ . Betrachten wir jedoch die aus dem Poynting-Satz in Abschnitt 2.3 hergeleitete Definition 5.6 nicht im Integral sondern in jedem Punkt im Gebiet  $V \subset \Omega$ , also den Integranden, dann können zwei komplett unterschiedliche Vektorfelder  $E$  dieselbe Verlustleistung in jedem Punkt in  $V$  erzeugen. Beispielsweise liefern die Vektorfelder  $E_1 = (1, 0, 0)^T$  und  $E_2 = (0, 1, 0)^T$  denselben Betrag, damit dieselbe Verlustleistung und somit auch dasselbe über das Gebiet betrachtete Erwärmungsmuster. Dabei sollte darauf hingewiesen werden, dass dieses Beispiel lediglich für den Fall linearer isotroper Materialien richtig ist, denn bei anisotropen Materialien ist die relative Permittivität matrixwertig, siehe Definition 2.4. Jedoch lassen sich auch für anisotrope Materialien entsprechende Vektorfelder mit gleicher Verlustleistung herleiten. Dies bedeutet, dass für die Erwärmung das exakte elektromagnetische Feld in jeder Komponente eine untergeordnete Rolle spielt und wir deshalb ein Zielfunktional herleiten, welches die Verlustleistungsfunktion beinhaltet.

**Definition 5.39.** Der Operator

$$P_Q : \mathbb{R}_+ \times X_e \rightarrow L^1(V, \mathbb{R}), \tag{5.51}$$

$$(\omega, E) \mapsto P_Q[\omega, E] = \frac{1}{2} \omega \varepsilon_0 \text{Im}(\varepsilon_r(\omega)) |E|^2,$$

heißt *punktuelle Verlustleistung* oder auch *punktuelles Leistungsmuster*. Die physikalische Einheit von  $P_Q[\omega, E]$  ist  $\left[\frac{W}{m^3}\right]$ .

Wie in (5.51) werden wir auch im weiteren Verlauf auf die reellwertigkeit der Funktionen  $P_Q[\omega, E]$  hinweisen, indem wir explizit  $L^1(V, \mathbb{R})$  schreiben. Mit Voraussetzung 2.10 und der Definition des Raumes  $X_e$  ist die Wohldefiniertheit des Operators  $P_Q$  klar. Der limitierende Faktor für die Regularität  $P_Q[\omega, E] \in L^1(V, \mathbb{R})$  liegt an  $E \in X_e$  und die

Ermittlung höherer Regularitätseigenschaften für  $|E|$  ist derzeitiges Forschungsgebiet. In der Hyperthermie, siehe [52, 53], wird für die Steuerung die gesamte Verlustleistung  $P(V)$  im Gebiet  $V$  in Zusammenhang mit den getrennten Systemen betrachtet, denn das ergibt

$$P(V) = \int_V P_Q[\omega, E] d\Omega = \sum_{l=1}^{K_A} \sum_{k=1}^{K_A} \sqrt{P_{in,l}} \sqrt{P_{in,k}} e^{-i\theta_l} e^{i\theta_k} \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) E_l \cdot E_k d\Omega.$$

Die elektromagnetischen Felder  $E_l$ ,  $1 \leq l \leq K_A$ , sind damit nur von der Frequenz abhängig und werden dadurch ermittelt, dass jeder Port einzeln mit einem Watt und der Phase Null angeregt wird. Da die Verlustleistung direkt in Wärme umgewandelt wird, kann bei Beobachtung der Wärmeentwicklung, beispielsweise durch eine Wärmebildkamera, das Integral ohne Wissen der Felder  $E_l$  und der Absorptionsrate  $\operatorname{Im}(\varepsilon_r(\omega))$  ermittelt werden. Diesen Umstand macht sich die Hyperthermie zu nutze, um das Integral durch eine gebietsbezogene Matrix  $Q(V)$  mit Einträgen

$$(Q(V))_{k,l} = \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) E_l \cdot E_k d\Omega$$

zu beschreiben, die sogenannte  $Q$ -Matrix. Mithilfe der  $Q$ -Matrix lässt sich damit

$$P(V) = a^H Q(V) a \quad (5.52)$$

mit  $a = [\sqrt{P_{in,l}} e^{-i\theta_l}]_{l=1}^{K_A} \in \mathbb{C}^{K_A}$  definieren. Damit lassen sich bei fester Frequenz die Phase und Leistung über Matrixgleichungen optimal steuern. Hierzu gibt es bereits seit Anfang der Neunziger Jahre einiges an Theorie, siehe dazu [52, 53, 54]. Wir werden im weiteren Verlauf keine  $Q$ -Matrizen verwenden und wollen hingegen das Leistungsmuster (5.51) in der entsprechenden  $L^1(V)$ -Norm betrachten, also

$$\|P_Q[\omega, E] - Q_V\|_{L^1(V, \mathbb{R})} = \int_V |P_Q[\omega, E] - Q_V| d\Omega = \int_V \left| \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E|^2 - Q_V \right| d\Omega$$

mit gegebenem Leistungsmuster  $Q_V \in L^2(V, \mathbb{R})$ . Um diesen Ausdruck nach der Diskretisierung berechnen zu können, benötigen wir eine Quadraturformel, welche wiederum die genaue Definition der FE-Ansatzfunktionen beinhaltet. Diese wollen wir hingegen vermeiden, weshalb wir obigen Ausdruck im  $L^2(V)$ -Sinne

$$\|P_Q[\omega, E] - Q_V\|_{L^2(V, \mathbb{R})}^2 = \int_V \left| \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E|^2 - Q_V \right|^2 d\Omega$$

betrachten wollen. Wie bereits in der Diskussion zur Definition 5.39 genannt, ist diese Betrachtung nicht wohldefiniert. Um diese Betrachtung dennoch zu verwenden und den Vorteil auszunutzen, dass wir in  $L^2(V)$  ein Skalarprodukt zur Definition der Norm haben, betrachten wir den Ausdruck nur für die diskretisierten elektrischen Felder  $E_{\mathcal{N}}$ .

**Satz 5.40.** *Die Einschränkung des Operators  $P_Q[\omega, \cdot]$  von  $X_e$  auf  $X_{\mathcal{N}} \subset X_e$  zum Gitter  $\mathcal{T}_h(\Omega)$  bildet nach  $L^2(V, \mathbb{R})$  ab, das heißt der Operator ist gegeben durch*

$$\begin{aligned} P_Q : \mathbb{R}_+ \times X_{\mathcal{N}} &\rightarrow L^2(V, \mathbb{R}), \\ (\omega, E_{\mathcal{N}}) &\mapsto P_Q[\omega, E_{\mathcal{N}}] = \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E_{\mathcal{N}}|^2. \end{aligned} \quad (5.53)$$

*Beweis.* Für  $E_{\mathcal{N}} \in X_{\mathcal{N}}$  zu dem konformen Gitter  $\mathcal{T}_h(\Omega)$  gilt

$$E_{\mathcal{N}}|_K \in [C^\infty(K)]^3,$$

für jeden Tetraeder  $K \in \mathcal{T}_h(\Omega)$  des Gitters, da nach Einschränkung auf den Tetraeder  $E_{\mathcal{N}}|_K$  durch eine Summe von lokal definierten Polynomen beschrieben wird. Damit ist  $E_{\mathcal{N}}$  ein stückweise stetiges Vektorfeld und daher gilt auch  $E_{\mathcal{N}} \in [L^4(\Omega)]^3$ . Wir erhalten damit

$$\begin{aligned} \|P_Q[\omega, E_{\mathcal{N}}]\|_{L^2(V, \mathbb{R})} &\leq \frac{1}{2} \omega_{\max} \varepsilon_0 \varepsilon_{\max, \operatorname{Im}} \| |E_{\mathcal{N}}|^2 \|_{L^2(\Omega, \mathbb{R})} \\ &= \frac{1}{2} \omega_{\max} \varepsilon_0 \varepsilon_{\max, \operatorname{Im}} \|E_{\mathcal{N}}\|_{[L^4(\Omega)]^3}^2 \\ &< \infty \end{aligned}$$

□

Dieser Satz und die Betrachtung des Operators (5.53) sind gute Approximationen an die echten Leistungsmuster, insofern die Approximationseigenschaft des Raumes  $X_{\mathcal{N}}$  zu  $X_e$  nach Satz 3.20 gut genug ist, also die Diskretisierungsfehler  $\|E_{\mathcal{N}} - E\|_{X_e}$  zu dem Gitter  $\mathcal{T}_h(\Omega)$  in der Norm (3.67) sehr klein werden.

*Bemerkung 5.41.* Seien  $K_1, K_2 \in \mathcal{T}_h(\Omega)$  mit  $\overline{K_1} \cap \overline{K_2} \neq \emptyset$ , dann wissen wir aus dem Beweis von Satz 5.40, dass  $E_{\mathcal{N}}|_{\overline{K_1}}, E_{\mathcal{N}}|_{\overline{K_2}} \in C_S^\infty(\overline{K_1 \cup K_2})$ , also stückweise stetige  $C^\infty$  Funktionen sind. Diese Regularität kann noch weiter verbessert werden, indem die Randbedingung (2.62) auf der gemeinsamen Randfläche  $\overline{K_1} \cap \overline{K_2}$  betrachtet wird. Nehmen wir an, dass die Oberflächenladungsdichte  $\rho_S$  Null ist, dann ist

$$E_{\mathcal{N}}|_{\overline{K_1 \cup K_2}} \in [C(\overline{K_1 \cup K_2})]^3 \quad (5.54)$$

genau dann, wenn  $\varepsilon_{r,1}(\omega) = \varepsilon_{r,2}(\omega)$  auf  $\overline{K_1} \cap \overline{K_2}$  gilt.

Das heißt, mit der Stetigkeit der relativen Permittivitätsabbildung  $\varepsilon_r(\omega) : V \rightarrow \mathbb{C}$  erhalten wir  $E_{\mathcal{N}} \in [C(V, \mathbb{C})]^3$  und somit die Erweiterung von Satz 5.40 durch die Raumzugehörigkeit  $P_Q[\omega, E_{\mathcal{N}}] \in C(V, \mathbb{R})$ . Wir definieren das dritte Optimierungsproblem:

- Finde zu  $Q_V \in L^2(V, \mathbb{R})$  jenes  $p_S \in \mathbb{P}_S$  und  $E_{\mathcal{N}} \in X_{\mathcal{N}}$ , wobei  $E_{\mathcal{N}}$  Lösung von  $[V_{D, \mathcal{N}}]$  zum Parametervektor  $p_S$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[\text{O}_3]_{\text{K}} \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ E_{\mathcal{N}} \in X_{\mathcal{N}}}} Z_3(p_S, E_{\mathcal{N}}) = \frac{1}{2|V|} \|P_Q[\omega, E_{\mathcal{N}}] - Q_V\|_{L^2(V, \mathbb{R})}^2$$

unter  $a(E_{\mathcal{N}}, \varphi, p_S) = f(\varphi, p_S), \quad \varphi \in X_{\mathcal{N}}$ .

Eine direkte Folgerung aus der Definition des Zielfunktional  $Z_3$  ist deren Stetigkeit in Bezug auf deren zwei Komponenten.

**Lemma 5.42.** *Das Zielfunktional  $Z_3 : \mathbb{P}_S \times L^2(V, \mathbb{R}) \rightarrow \mathbb{R}_+$  ist stetig in jeder Komponente.*

*Beweis.* Wegen Bemerkung 5.35, Satz 5.3 und Satz 5.40 erhalten wir die Stetigkeit der Abbildung (5.53) in beiden Komponenten und mit der Normstetigkeit die obige Aussage.  $\square$

Aufgrund der  $L^2(V)$ -Norm ist  $Z_3$  schwach unterhalbstetig in  $E_{\mathcal{N}}$  und mit entsprechender Definition 5.9 der optimalen Steuerungen für  $[\text{O}_3]_{\text{K}}$  erhalten wir analog zu Satz 5.10 die folgende Aussage.

**Satz 5.43.** *Es existiert mindestens ein Lösungs-Tupel  $(\hat{p}_S, \hat{E}_{\mathcal{N}}) \in \mathbb{O}$  in der zulässigen Menge  $\mathbb{O} \subset \mathbb{P}_S \times X_{\mathcal{N}}$  zu dem Optimalsteuerungsproblem  $[\text{O}_3]_{\text{K}}$ .*

Um  $[\text{O}_3]_{\text{K}}$  numerisch zu lösen, führen wir folgendes Variationsproblem ein:

- Bestimme  $P_Q[\omega, E_{\mathcal{N}}] \in L^2(V, \mathbb{R})$ , sodass

$$[\text{V}_{\text{K}, Q}] \quad a_Q(P_Q[\omega, E_{\mathcal{N}}], \phi) = f_Q(\phi, E_{\mathcal{N}}, p_S), \quad \phi \in L^2(V, \mathbb{R}),$$

gilt mit

$$a_Q(P_Q[\omega, E_{\mathcal{N}}], \phi) = \int_V P_Q[\omega, E_{\mathcal{N}}] \phi \, d\Omega, \quad (5.55)$$

$$f_Q(\phi, E_{\mathcal{N}}, p_S) = \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E_{\mathcal{N}}|^2 \phi \, d\Omega. \quad (5.56)$$

Dieses Variationsproblem werden wir in der numerischen Betrachtung verwenden, doch zuvor benötigen wir die Wohldefiniertheit, also die Existenz und Eindeutigkeit.

**Satz 5.44.** *Es existiert eine eindeutige Lösung  $P_Q[\omega, E_{\mathcal{N}}] \in L^2(V, \mathbb{R})$  des Variationsproblems  $[V_{K,Q}]$ .*

*Beweis.* Nach Definition (5.55) der Bilinearform  $a_Q$ , ist  $a_Q$  koerziv und stetig. Mit Satz A.17 existiert damit eine eindeutige Lösung  $P_Q[\omega, E_{\mathcal{N}}] \in L^2(V, \mathbb{R})$  genau dann, wenn die Abbildung  $\phi \mapsto f_Q[\phi] = f_Q(\phi, E_{\mathcal{N}}, p_S)$  ein Element des Dualraums darstellt. Wir müssen deshalb zeigen, dass  $f_Q$  linear und in der Dualraumnorm beschränkt ist. Die Linearität folgt direkt aus (5.56), während für die Beschränktheit

$$\begin{aligned} \|f_Q\|_{L^2(V, \mathbb{R})'} &= \sup_{\phi \in L^2(V, \mathbb{R})} \frac{|\int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E_{\mathcal{N}}|^2 \phi \, d\Omega|}{\|\phi\|_{L^2(V, \mathbb{R})}} \\ &\leq \frac{1}{2} \omega_{\max} \varepsilon_0 \varepsilon_{\max, \operatorname{Im}} \|E_{\mathcal{N}}\|_{[L^4(V)]^3}^2 \end{aligned}$$

durch Anwendung der Hölder-Ungleichung folgt. Der Beweis von Satz 5.40 liefert uns  $E_{\mathcal{N}} \in [L^4(\Omega)]^3$ , wodurch wir die Aussage erhalten.  $\square$

**Korollar 5.45.** *Die eindeutige Lösung  $P_Q[\omega, E_{\mathcal{N}}] \in L^2(V, \mathbb{R})$  von  $[V_{K,Q}]$  zu  $\omega$  und  $E_{\mathcal{N}}$  ist gegeben durch die Abbildung  $P_Q[\omega, E_{\mathcal{N}}]$  in (5.53).*

**Satz 5.46.** *Sei  $p_{S,k} = (p_S)_k$  Folge in  $\mathbb{P}_S$  und  $\hat{p}_S \in \mathbb{P}_S$  mit*

$$\|p_{S,k} - \hat{p}_S\|_{\mathbb{R}^{1+2K_A}} \xrightarrow{k \rightarrow \infty} 0,$$

*sowie  $E_{\mathcal{N},k} = (E_{\mathcal{N}})_k$  Folge in  $X_{\mathcal{N}}$  und  $\hat{E}_{\mathcal{N}}$  mit*

$$\|E_{\mathcal{N},k} - \hat{E}_{\mathcal{N}}\|_{X_e} \xrightarrow{k \rightarrow \infty} 0.$$

*Zu der Folge an Lösungen  $P_{Q,k}[p_{S,k}, E_{\mathcal{N},k}]$  von  $[V_{K,Q}]$  zu den Parametervektoren  $p_{S,k}$  und  $E_{\mathcal{N},k}$  existiert eine konvergente Teilfolge mit Grenzwert  $\hat{P}_Q$ , mit anderen Worten*

$$\|P_{Q,k_l}[p_{S,k_l}, E_{\mathcal{N},k_l}] - \hat{P}_Q\|_{L^2(V)} \xrightarrow{l \rightarrow \infty} 0, \quad (5.57)$$

*und  $\hat{P}_Q$  ist die Lösung  $P_Q[\hat{p}_S, \hat{E}_{\mathcal{N}}]$  von  $[V_{K,Q}]$  zu  $\hat{p}_S$  und  $\hat{E}_{\mathcal{N}}$ .*

*Beweis.* Da  $\mathbb{P}_S$  kompakt ist, existiert eine konvergente Teilfolge  $p_{S,k_m}$  mit Grenzwert  $\hat{p}_S$ . Des Weiteren ist  $X_{\mathcal{N}} \subset X_e$  als Unterraum reflexiv, weshalb eine schwach konvergente Teilfolge  $E_{\mathcal{N},k_m}$  mit Grenzwert  $\hat{E}_{\mathcal{N}}$  existiert. Da  $X_{\mathcal{N}}$  zudem endlich dimensional ist, konvergiert  $E_{\mathcal{N},k_m}$  sogar stark gegen  $\hat{E}_{\mathcal{N}}$ . Für eine Teilfolge  $k_m$ ,  $m \in \mathbb{N}$ , erhalten wir mit

der Stetigkeit des Operators (5.53) in beiden Komponenten die Konvergenz

$$\|P_{Q,k_m}[\omega_{k_m}, E_{\mathcal{N},k_m}] - \hat{P}_Q\|_{L^2(V)} \xrightarrow{m \rightarrow \infty} 0.$$

Zusammen mit Korollar 5.45 ergibt dies gerade die zu beweisende Aussage.  $\square$

Mit Korollar 5.45 können wir  $[O_3]_K$  auch folgendermaßen umschreiben:

- Finde zu  $Q_V \in L^2(V, \mathbb{R})$  jenes  $p_S \in \mathbb{P}_S$ ,  $P_Q[\omega, E_{\mathcal{N}}] \in L^2(V, \mathbb{R})$  und  $E_{\mathcal{N}} \in X_{\mathcal{N}}$ , wobei  $P_Q[\omega, E_{\mathcal{N}}]$  und  $E_{\mathcal{N}}$  die Lösungen von  $[V_{K,Q}]$  bzw.  $[V_{D,\mathcal{N}}]$  zum Parametervektor  $p_S$  sind, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[O_3]_K \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ P_Q[\omega, E_{\mathcal{N}}] \in L^2(V, \mathbb{R})}} Z_3(p_S, P_Q[\omega, E_{\mathcal{N}}]) = \frac{1}{2} \frac{1}{|V|} \|P_Q[\omega, E_{\mathcal{N}}] - Q_V\|_{L^2(V, \mathbb{R})}^2$$

unter

$$a_Q(P_Q[\omega, E_{\mathcal{N}}], \phi) = f_Q(\phi, E_{\mathcal{N}}, p_S), \quad \phi \in L^2(V, \mathbb{R}),$$

$$a(E_{\mathcal{N}}, \varphi, p_S) = f(\varphi, p_S), \quad \varphi \in X_{\mathcal{N}}.$$

Analog zu Definition 5.11 definieren wir zu dem Steuerungs-Zustands-Operator  $\mathcal{E}_{\mathcal{N}}$  für  $[V_K]$  auch den Zustands-Operator

$$P_Q : [\omega_{\min}, \omega_{\max}] \times X_{\mathcal{N}} \rightarrow L^2(V, \mathbb{R}),$$

$$(\omega, E_{\mathcal{N}}) \mapsto P_Q[\omega, E_{\mathcal{N}}],$$

der gegeben ist als Lösung des Variationsproblems  $[V_{K,Q}]$ . Nach Korollar 5.45 entspricht dieser gerade der Definition (5.53). Wir kombinieren die beiden Operatoren und erhalten den Steuerungs-Leistungs-Operator

$$\mathcal{P}_Q : \mathbb{P}_S \rightarrow L^2(V, \mathbb{R}),$$

$$p_S \mapsto \mathcal{P}_Q[p_S] = P_Q[\omega, \mathcal{E}[p_S]].$$

Damit haben wir die notwendigen Definitionen eingeführt, um die entsprechenden Sätze aus den vorherigen Abschnitten herzuleiten. Da die Beweise zu diesen analog geführt werden können, werden im folgenden keine Beweise angegeben. Als erstes geben wir den Satz zur Existenz der ersten Variation des Steuerungs-Leistungs-Operators und der damit einhergehenden Existenz der ersten Variation des undefinierten Optimierungsproblems

$$\min_{p_S \in \mathbb{P}_S} \mathcal{Z}_3[p_S] = Z_3(p_S, \mathcal{P}_Q[p_S]) \tag{5.58}$$

für das Optimierungsproblem  $[O_3]_K$  an.

**Satz 5.47.** *Gilt Voraussetzung 5.15, dann ist die erste Variation  $\delta\mathcal{P}_Q[p_S, \tilde{p}_S] \in L^2(V, \mathbb{R})$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  Lösung des folgenden Variationsproblems*

$$a_Q(\delta\mathcal{P}_Q[p_S, \tilde{p}_S], \phi) = f_{Q,2}(\phi, \delta\mathcal{E}_N[p_S, \tilde{p}_S], \mathcal{E}_N[p_S], p_S, \tilde{p}_S), \quad \phi \in L^2(V, \mathbb{R}), \quad (5.59)$$

mit Sesquilinearform  $a_Q$  gegeben durch (5.55) des Variationsproblems  $[V_{K,Q}]$  und

$$\begin{aligned} f_{Q,2}(\phi, \delta\mathcal{E}_N[p_S, \tilde{p}_S], \mathcal{E}_N[p_S], p_S, \tilde{p}_S) & \quad (5.60) \\ &= \frac{1}{2} \int_V \varepsilon_0(\omega \operatorname{Im}(\delta\varepsilon_r[\omega, \tilde{\omega}]) + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}_N[p_S]|^2 \phi \, d\Omega \\ &+ \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}(\delta\mathcal{E}_N[p_S, \tilde{p}_S] \cdot \mathcal{E}_N[p_S]) \phi \, d\Omega, \end{aligned}$$

wobei die erste Variation  $\delta\mathcal{E}_N[p_S, \tilde{p}_S]$  an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  die diskretisierte Version der Gleichungen in Satz 5.19 löst.

**Korollar 5.48.** *Gilt Voraussetzung 5.16, dann ist die Abbildung  $p_S \mapsto \mathcal{P}_Q[p_S]$  Gâteaux-differenzierbar, also*

$$\delta\mathcal{P}_Q[p_S, \tilde{p}_S] = \nabla\mathcal{P}_Q[p_S]\tilde{p}_S, \quad (5.61)$$

mit

$$\nabla\mathcal{P}_Q[p_S] = \left( \nabla_\omega\mathcal{P}_Q[p_S], [\nabla_{P_{in,l}}\mathcal{P}_Q[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l}\mathcal{P}_Q[p_S]]_{l=1}^{K_A} \right),$$

wobei  $[\nabla_{P_{in,l}}\mathcal{P}_Q[p_S]]_{l=1}^{K_A}$  und  $[\nabla_{\theta_l}\mathcal{P}_Q[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}}\mathcal{P}_Q[p_S]$  bzw.  $\nabla_{\theta_l}\mathcal{P}_Q[p_S]$  sind, und die einzelnen Gradienten folgende Variationsprobleme lösen:

1. Der Gradient  $\nabla_\omega\mathcal{P}_Q[p_S]$  löst das Variationsproblem

$$\begin{aligned} a_Q(\nabla_\omega\mathcal{P}_Q[p_S], \phi) &= \frac{1}{2} \int_V \varepsilon_0(\omega \operatorname{Im}(\nabla\varepsilon_r[\omega]) + \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}_N[p_S]|^2 \phi \, d\Omega \\ &+ \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}(\nabla_\omega\mathcal{E}_N[p_S] \cdot \mathcal{E}_N[p_S]) \phi \, d\Omega, \quad \phi \in L^2(V, \mathbb{R}). \end{aligned}$$

2. Der Gradient  $\nabla_{P_{in,l}}\mathcal{P}_Q[p_S]$ ,  $1 \leq l \leq K_A$ , löst das Variationsproblem

$$a_Q(\nabla_{P_{in,l}}\mathcal{P}_Q[p_S], \phi) = \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}(\nabla_{P_{in,l}}\mathcal{E}_N[p_S] \cdot \mathcal{E}_N[p_S]) \phi \, d\Omega, \quad \phi \in L^2(V, \mathbb{R}).$$

3. Der Gradient  $\nabla_{\theta_l} \mathcal{P}_Q[p_S]$ ,  $1 \leq l \leq K_A$ , löst das Variationsproblem

$$a_Q(\nabla_{\theta_l} \mathcal{P}_Q[p_S], \phi) = \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}(\nabla_{\theta_l} \mathcal{E}_N[p_S] \cdot \mathcal{E}_N[p_S]) \phi \, d\Omega, \quad \phi \in L^2(V, \mathbb{R}).$$

Dabei löst die Gâteaux-Ableitung  $\nabla \mathcal{E}_N[p_S]$  an der Stelle  $p_S$  die diskretisierte Version der Gleichungen in Korollar 5.20.

Mit Satz 5.47 und dem daraus ableitbarem Korollar 5.48 erhalten wir dann die Existenz der ersten Variation und der Gâteaux-Ableitung des Zielfunktional von  $[\mathcal{O}_3]_K$ .

**Satz 5.49.** *Gilt Voraussetzung 5.15, dann existiert die erste Variation  $\delta \mathcal{Z}_3[p_S, \tilde{p}_S]$  an der Stelle  $p_S \in \mathcal{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^{1+2K_A}$  und ist definiert durch*

$$\delta \mathcal{Z}_3[p_S, \tilde{p}_S] = \frac{1}{|V|} \int_V \delta \mathcal{P}_Q[p_S, \tilde{p}_S] (\mathcal{P}_Q[p_S] - Q_V) \, d\Omega, \quad (5.62)$$

wobei  $\delta \mathcal{P}_Q[p_S, \tilde{p}_S]$  durch Satz 5.47 gegeben ist. Gilt zudem Voraussetzung 5.16, dann ist

$$\delta \mathcal{Z}_3[p_S, \tilde{p}_S, Q_V] = \nabla \mathcal{Z}_3[p_S, Q_V] \tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_3[p_S] = \frac{1}{|V|} \int_V \nabla \mathcal{P}_Q[p_S] (\mathcal{P}_Q[p_S] - Q_V) \, d\Omega, \quad (5.63)$$

wobei  $\nabla \mathcal{P}_Q[p_S]$  durch Korollar 5.48 gegeben ist.

Wie in Abschnitt 5.1.4 lässt sich mit der formalen Lagrange-Technik die Lagrange-Funktion und das adjungierte Problem zu  $[\mathcal{O}_3]_K$  herleiten:

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $P_Q \in L^2(V, \mathbb{R})$  des Optimierungsproblems  $[\mathcal{O}_3]_K$  die adjungierte Lösung  $\lambda_1 \in L^2(V, \mathbb{R})$ , sodass

$$[\mathcal{D}_3]_K \quad a_Q(\phi, \lambda_1) = f_{Q, D_3}(\phi, P_Q, Q_V, p_S), \quad \phi \in L^2(V, \mathbb{R}),$$

gilt mit Bilinearform  $a_Q$  gegeben durch (5.55) des Variationsproblems  $[\mathcal{V}_{K, Q}]$  und

$$f_{Q, D_3}(\phi, P_Q, Q_V, p_S) = \frac{1}{|V|} \int_V \phi (P_Q - Q_V) \, d\Omega. \quad (5.64)$$

Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$ ,  $E_N \in X_N$  und der adjungierten

Lösung  $\lambda_1 \in L^2(V, \mathbb{R})$  zu  $[D_3]_K$  die adjungierte Lösung  $\lambda_2 \in X_{\mathcal{N}}$ , sodass

$$a(\varphi, \lambda_2, p_S) = f_{D_3}(\varphi, E_{\mathcal{N}}, \lambda_1, p_S), \quad \varphi \in X_{\mathcal{N}}, \quad (5.65)$$

gilt mit Sesquilinearform  $a$  gegeben durch (3.3) des Variationsproblems  $[V_{D, \mathcal{N}}]$  und

$$f_{D_3}(\varphi, E_{\mathcal{N}}, \lambda_1, p_S) = \int_V \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) (\varphi \cdot E_{\mathcal{N}}) \lambda_1 \, d\Omega. \quad (5.66)$$

**Definition 5.50.** Die Lagrange-Funktion zu dem Optimierungsproblem  $[O_3]_K$  ist gegeben durch

$$\begin{aligned} & \mathcal{L}_3(p_S, E_{\mathcal{N}}, P_Q[\omega, E_{\mathcal{N}}], \lambda_1, \lambda_2) \quad (5.67) \\ &= \frac{1}{2} \frac{1}{|V|} \int_V |P_Q[\omega, E_{\mathcal{N}}] - Q_V|^2 \, d\Omega - \int_V P_Q[\omega, E_{\mathcal{N}}] \lambda_1 \, d\Omega \\ &+ \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E_{\mathcal{N}}|^2 \lambda_1 \, d\Omega + \operatorname{Re} \left( - \int_{\Omega} \frac{1}{\mu_0 \mu_r} (\nabla \times E_{\mathcal{N}}) \cdot (\nabla \times \lambda_2) \, d\Omega \right. \\ &+ \int_{\Omega} \omega^2 \varepsilon_0 \varepsilon_r(\omega) E_{\mathcal{N}} \cdot \lambda_2 \, d\Omega + \int_{\Gamma_{IMP}} i\omega Y(\omega) \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\lambda_2] \, d\Gamma \\ &+ \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \pi_t[E_{\mathcal{N}}] \cdot \pi_t[\lambda_2] \, d\Gamma \\ &\left. - 2 \sum_{l=1}^{K_A} \int_{\Gamma_{P,l}} i\omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_2] \, d\Gamma \right), \end{aligned}$$

wobei  $p_S \in \mathbb{P}_S$ ,  $P_Q, \lambda_1, Q_V \in L^2(V, \mathbb{R})$  und  $E_{\mathcal{N}}, \lambda_2 \in X_{\mathcal{N}}$ .

Die Existenz einer Lösung zu  $[D_3]_K$  erhalten wir dabei wieder völlig analog zu Satz 5.44 und Kapitel 3. Mit der Aussage in Korollar 5.45 ergibt sich für  $[D_3]_K$  die Darstellung der adjungierten Lösung

$$\lambda_1 = \frac{1}{|V|} (P_Q[\omega, E_{\mathcal{N}}] - Q_V).$$

Kombinieren wir Satz 5.19 und Satz 5.49 mit  $[D_3]_K$ , ergibt sich die veränderte Darstellung der ersten Variation (5.62) und eine Bedingung für die lokale Minimierungseigenschaft.

**Satz 5.51.** Seien  $\lambda_1, \lambda_2$  Lösungen zu  $[D_3]_K$ . Gilt Voraussetzung 5.15, dann folgt

$$\begin{aligned} \delta \mathcal{Z}_3[p_S, \tilde{p}_S] &= f_{Q,2}(\lambda_1, \delta \mathcal{E}[p_S, \tilde{p}_S], \mathcal{E}[p_S], p_S, \tilde{p}_S) \\ &= \frac{1}{2} \int_V \varepsilon_0(\omega \operatorname{Im}(\delta \varepsilon_r[\omega, \tilde{\omega}]) + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}_N[p_S]|^2 \lambda_1 d\Omega \\ &\quad + \operatorname{Re}(f_2(\lambda_2, \mathcal{E}_N[p_S], p_S, \tilde{p}_S)) \end{aligned}$$

mit  $f_2$  gegeben durch (5.21). Gilt Voraussetzung 5.16, dann folgt

$$\delta \mathcal{Z}_3[p_S, \tilde{p}_S] = \nabla \mathcal{Z}_3[p_S] \tilde{p}_S,$$

mit

$$\nabla \mathcal{Z}_3[p_S] = \left( \nabla_\omega \mathcal{Z}_3[p_S], [\nabla_{P_{in,l}} \mathcal{Z}_3[p_S]]_{l=1}^{K_A}, [\nabla_{\theta_l} \mathcal{Z}_3[p_S]]_{l=1}^{K_A} \right),$$

wobei  $[\nabla_{P_{in,l}} \mathcal{Z}_3[p_S]]_{l=1}^{K_A}$  und  $[\nabla_{\theta_l} \mathcal{Z}_3[p_S]]_{l=1}^{K_A}$  Matrizen mit Spalteneinträgen  $\nabla_{P_{in,l}} \mathcal{Z}_3[p_S]$  bzw.  $\nabla_{\theta_l} \mathcal{Z}_3[p_S]$  sind. Seien  $\lambda_1, \lambda_2$  die adjungierten Lösungen zu  $[D_3]_K$ , dann gelten die folgenden Gleichungen:

1. Der Gradient  $\nabla_\omega \mathcal{Z}_3[p_S]$  ist gegeben durch

$$\begin{aligned} \nabla_\omega \mathcal{Z}_3[p_S] &= \frac{1}{2} \int_V \varepsilon_0(\omega \operatorname{Im}(\nabla \varepsilon_r[\omega]) + \operatorname{Im}(\varepsilon_r(\omega))) |\mathcal{E}_N[p_S]|^2 \lambda_1 d\Omega \\ &\quad + \operatorname{Re}(f_{G_\omega}(\lambda_2, \mathcal{E}_N[p_S], p_S)), \end{aligned}$$

mit  $f_{G_\omega}$  wie in Korollar 5.20.

2. Der Gradient  $\nabla_{P_{in,l}} \mathcal{Z}_3[p_S, Q_V]$ ,  $1 \leq l \leq K_A$ , ist gegeben durch

$$\begin{aligned} \nabla_{P_{in,l}} \mathcal{Z}_3[p_S, Q_V] &= \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re} \left( \left( \frac{1}{2P_{in,l}} \mathcal{E}_N[\hat{p}_S] \right) \cdot \mathcal{E}_N[p_S] \right) \lambda_1 d\Omega \\ &= - \operatorname{Re} \left( \int_{\Gamma_{P,l}} i \mu_0 \omega Y(\omega) \frac{e^{-i\theta_l}}{\sqrt{P_{in,l}} \sqrt{P_{m,l}(\omega)}} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_2] d\Gamma \right), \end{aligned}$$

wobei  $\mathcal{E}[\hat{p}_S]$  Lösung von  $[V_K]$  zu dem Parametervektor  $\hat{p}_S$  definiert als

$$\hat{\omega} = \omega, \quad \hat{\theta}_l = \theta_l, \quad \hat{P}_{in,l} = P_{in,l},$$

sowie  $\hat{\theta}_k = \hat{P}_{in,k} = 0$  für  $1 \leq k \leq K_A$  mit  $k \neq l$ .

3. Der Gradient  $\nabla_{\theta_l} \mathcal{Z}_3[p_S]$ ,  $1 \leq l \leq K_A$ , ist gegeben durch

$$\begin{aligned} \nabla_{\theta_l} \mathcal{Z}_3[p_S] &= \int_V \varepsilon_0 \omega \operatorname{Im}(\varepsilon_r(\omega)) \operatorname{Re}((-i \mathcal{E}_N[\hat{p}_S]) \cdot \mathcal{E}_N[p_S]) \lambda_1 \, d\Omega \\ &= -2 \operatorname{Re} \left( \int_{\Gamma_{P,l}} \mu_0 \omega Y(\omega) \frac{\sqrt{P_{in,l}}}{\sqrt{P_{m,l}(\omega)}} e^{-i\theta_l} \pi_t[E_{0,l}] \cdot \pi_t[\lambda_2] \, d\Gamma \right), \end{aligned}$$

wobei  $\mathcal{E}[\hat{p}_S]$  Lösung von  $[V_K]$  zu dem Parametervektor  $\hat{p}_S$  definiert als

$$\hat{\omega} = \omega, \quad \hat{\theta}_l = \theta_l, \quad \hat{P}_{in,l} = P_{in,l},$$

sowie  $\hat{\theta}_k = \hat{P}_{in,k} = 0$  für  $1 \leq k \leq K_A$  mit  $k \neq l$ .

*Bemerkung 5.52.* In Satz 5.51 existieren zur Bestimmung der Gradienten  $\nabla_{P_{in,l}} \mathcal{Z}_3[p_S]$  und  $\nabla_{\theta_l} \mathcal{Z}_3[p_S]$ ,  $1 \leq l \leq K_A$ , jeweils zwei unterschiedliche Darstellungen. Die erste Darstellung bezieht sich auf die Anwendung der Lösungseigenschaft der adjungierten Lösung  $\lambda_1$  des ersten Teils des adjungierten Problems  $[D_3]_K$  und der Anwendung der Gradientendarstellungen  $\nabla_{P_{in,l}} \mathcal{E}[p_S]$  und  $\nabla_{\theta_l} \mathcal{E}[p_S]$ ,  $1 \leq l \leq K_A$ , in Korollar 5.20. Dabei besitzt die erste Darstellung einen Vorteil bei der Betrachtung der getrennten Systeme, bei denen die Teillösungen  $\mathcal{E}[\hat{p}_S]$  bereits in den einzelnen Subsystemen  $[V_{K,Div_l}]$  bestimmt werden. Die zweite Darstellung ergibt sich aus der erweiterten Anwendung der Lösungseigenschaft der adjungierten Lösung  $\lambda_2$  des zweiten Teils des adjungierten Problems  $[D_3]_K$  und der Lösungseigenschaften  $\mathcal{E}[\hat{p}_S]$  zu  $[V_K]$ , wie in Satz 5.26, und liefert eine Darstellung für die Betrachtung der vollen Systeme.

**Satz 5.53.** Gelte Voraussetzung 5.15 und sei das Tripel

$$(\hat{p}_S, \hat{E}_N, \hat{P}_Q[\hat{\omega}, \hat{E}_N]) \in \mathbb{P}_S \times X_N \times L^2(V, \mathbb{R})$$

ein lokales Minimum der Lagrange-Funktion  $\mathcal{L}_3(p_S, E_N, P_Q[\omega, E_N], \lambda_1, \lambda_2)$  zu gegebenem  $Q_V \in L^2(V)$  definiert durch (5.67), sowie

$$\tilde{p}_S = p_S - \hat{p}_S \in \mathbb{P}_S,$$

für beliebiges  $p_S \in \mathbb{P}_S$ . Dann existieren eindeutige adjungierte Lösungen  $\lambda_1 \in L^2(V, \mathbb{R})$  und  $\lambda_2 \in X_N$  von  $[D_3]_K$ , sodass die folgenden Variationsungleichungen erfüllt sind

$$\frac{1}{2} \int_V \varepsilon_0 (\hat{\omega} \operatorname{Im}(\delta \varepsilon_r[\hat{\omega}, \tilde{\omega}] + \tilde{\omega} \operatorname{Im}(\varepsilon_r(\hat{\omega}))) \left| \hat{E}_N \right|^2 \lambda_1 \, d\Omega + \operatorname{Re} \left( f_2(\lambda_2, \hat{E}_N, \hat{p}_S, \tilde{p}_S) \right) \geq 0, \quad (5.68)$$

wobei  $f_2$  gegeben durch (5.21) aus Satz 5.19 ist. Gilt zudem Voraussetzung 5.16 und

definieren wir den optimalen Zustand  $\hat{E}_N$  nach Definition 5.11 als  $\mathcal{E}_N[\hat{p}_S]$ , dann lässt sich die Bedingung (5.68) auch folgendermaßen ausdrücken

$$\nabla \mathcal{Z}_3[\hat{p}_S](p_S - \hat{p}_S) \geq 0. \quad (5.69)$$

Wir widmen uns nun der Fehlerrechnung, indem wir  $[O_3]_K$  diskretisieren. Die Diskretisierung der Felder  $E_N \in X_N$  haben wir bereits vorausgesetzt und deshalb bedarf es der Diskretisierung von  $P_Q[\omega, E_N] \in L^2(V, \mathbb{R})$ . Betrachten wir die Definition (5.55) der Bilinearform  $a_Q$ , dann ist klar, dass  $a_Q$  stetig und koerziv ist, mit Stetigkeitskonstante  $C = 1$  und Koerzivitätskonstante  $\alpha = 1$ . Nach dem Lemma von Céa erhalten wir hierfür die Bestapproximationseigenschaft

$$\|P_Q[\omega, E_N] - P_{Q, \mathcal{N}_L}[\omega, E_N]\|_{L^2(V, \mathbb{R})} = \inf_{u_{\mathcal{N}_L} \in L_{\mathcal{N}_L}} \|P_Q[\omega, E_N] - u_{\mathcal{N}_L}\|_{L^2(V, \mathbb{R})}, \quad (5.70)$$

wenn  $P_{Q, \mathcal{N}_L}[\omega, E_N]$  Lösung des folgenden diskretisierten Variationsproblems ist:

- Bestimme  $P_{Q, \mathcal{N}_L}[\omega, E_N] \in L_{\mathcal{N}_L} \subset L^2(V, \mathbb{R})$ , sodass

$$[V_{D, Q, \mathcal{N}_L}] \quad a_Q(P_{Q, \mathcal{N}_L}[\omega, E_N], \phi_{\mathcal{N}_L}) = f_Q(\phi_{\mathcal{N}_L}, E_N, p_S), \quad \phi_{\mathcal{N}_L} \in L_{\mathcal{N}_L},$$

gilt mit

$$a_Q(P_{Q, \mathcal{N}_L}[\omega, E_N], \phi_{\mathcal{N}_L}) = \int_V P_{Q, \mathcal{N}_L}[\omega, E_N] \phi_{\mathcal{N}_L} \, d\Omega, \quad (5.71)$$

$$f_Q(\phi_{\mathcal{N}_L}, E_N, p_S) = \int_V \frac{1}{2} \omega \varepsilon_0 \operatorname{Im}(\varepsilon_r(\omega)) |E_N|^2 \phi_{\mathcal{N}_L} \, d\Omega. \quad (5.72)$$

Für die Aussage (5.70) ist die Darstellung des diskretisierten Unterraumes  $L_{\mathcal{N}_L}$  beliebig. Wir diskretisieren hier unsere Funktion wieder mit der FE-Methode, indem wir Lagrange-Elemente als Ansatzfunktionen verwenden. Hierzu sei nun  $L_{\mathcal{N}_L} \subset L^2(V, \mathbb{R})$  der durch das Gitter  $\mathcal{T}_{h_L}(V)$  erzeugte und mit Lagrange-Elementen versehene Unterraum. Dabei sei  $\mathcal{N}_L$  die Anzahl an Freiheitsgraden zur Beschreibung von  $L_{\mathcal{N}_L}$ . Wegen Korollar 5.45, der stückweise stetigen relativen Permittivität  $\varepsilon_r(\omega) \in C_S(\Omega)$ , Voraussetzung 2.10 und  $E_N \in [C_S^\infty(\Omega)]^3$ , wobei  $[C_S^\infty(\Omega)]^3$  der Raum der stückweise unendlich oft differenzierbaren Vektorfelder, folgt für die Lösung des Problems  $[V_{K, Q}]$  gerade  $P_Q \in C_S(V, \mathbb{R})$ . Für  $L_{\mathcal{N}_L}$  kann damit ein Interpolationsoperator

$$\Pi_{\mathcal{N}_L} : L^2(V, \mathbb{R}) \subset H^1(V, \mathbb{R}) \rightarrow L_{\mathcal{N}_L}$$

definiert werden, für den wir folgendes erhalten.

**Lemma 5.54.** Sei  $\mathcal{T}_{h_{L,n}}(V)$  reguläre Folge konformer Gitter, wobei  $h_{L,n} \xrightarrow{n \rightarrow \infty} 0$  gilt,  $L_{\mathcal{N}_L}$  der zu  $\mathcal{T}_{h_{L,n}}(V)$  diskretisierte Lagrange-Raum und  $\mathcal{N}_{L,n}$  die dazugehörige Anzahl an Freiheitsgraden mit  $\mathcal{N}_{L,n} \xrightarrow{n \rightarrow \infty} \infty$ . Dann gilt

$$\lim_{h_{L,n} \xrightarrow{n \rightarrow \infty} 0} \|P_Q[\omega, E_{\mathcal{N}}] - \Pi_{\mathcal{N}_L}[P_Q[\omega, E_{\mathcal{N}}]]\|_{L^2(V, \mathbb{R})} = 0.$$

Diese Approximationseigenschaft der Interpolationsfunktion  $\Pi_{\mathcal{N}_L}$  kann in [55] nachgelesen werden. Damit erhalten wir das diskrete Optimierungsproblem:

- Finde zu  $Q_V \in L^2(V, \mathbb{R})$  jenes  $p_S \in \mathbb{P}_S$ ,  $P_{Q, \mathcal{N}_L}[\omega, E_{\mathcal{N}}] \in L_{\mathcal{N}_L}$  und  $E_{\mathcal{N}} \in X_{\mathcal{N}}$ , wobei  $P_{Q, \mathcal{N}_L}[\omega, E_{\mathcal{N}}]$  und  $E_{\mathcal{N}}$  Lösung von  $[\mathbf{V}_{D, Q, \mathcal{N}_L}]$  bzw.  $[\mathbf{V}_{D, \mathcal{N}}]$  zum Parametervektor  $p_S$  sind, sodass diese Lösung des folgenden Optimierungsproblems sind:

$$\begin{aligned} & [\mathbf{O}_3]_{D, \mathcal{N}_L} \\ & \min_{\substack{p_S \in \mathbb{P}_S, \\ P_{Q, \mathcal{N}_L}[\omega, E_{\mathcal{N}}] \in L_{\mathcal{N}_L}}} Z_3(p_S, P_{Q, \mathcal{N}_L}, Q_V) = \frac{1}{2|V|} \|P_{Q, \mathcal{N}_L}[\omega, E_{\mathcal{N}}] - Q_V\|_{L^2(V, \mathbb{R})}^2 \\ & \text{unter} \quad \begin{aligned} a_Q(P_{Q, \mathcal{N}_L}, \phi_{\mathcal{N}_L}) &= f_Q(\phi_{\mathcal{N}_L}, E_{\mathcal{N}}, p_S), & \phi_{\mathcal{N}_L} &\in L_{\mathcal{N}_L}, \\ a(E_{\mathcal{N}}, \varphi, p_S) &= f(\varphi, p_S), & \varphi &\in X_{\mathcal{N}}. \end{aligned} \end{aligned}$$

Kombinieren wir nun Lemma 5.54 mit (5.70) dann erhalten wir folgenden Fehler.

**Satz 5.55.** 1. Sei  $\hat{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}_3]_{D, \mathcal{N}_L}$  zu den Gittern  $\mathcal{T}_h(\Omega)$ ,  $\mathcal{T}_{h_L}(V)$  und dazugehörigen Räumen  $X_{\mathcal{N}}$ ,  $L_{\mathcal{N}_L} \subset L^2(V)$ , sowie  $\tilde{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}_3]_K$ , dann ist der Fehler in der Zielfunktion bezüglich  $\hat{p}_S$  und  $\tilde{p}_S$  für  $[\mathbf{O}_3]_K$  gegeben durch

$$\begin{aligned} |\mathcal{Z}_3[\hat{p}_S]^{\frac{1}{2}} - \mathcal{Z}_3[\tilde{p}_S]^{\frac{1}{2}}| &\leq \frac{1}{\sqrt{2|V|}} \left( \inf_{u_{\mathcal{N}_L} \in L_{\mathcal{N}_L}} \|P_Q[\hat{p}_S] - u_{\mathcal{N}_L}\|_{L^2(V, \mathbb{R})} \right. \\ &\quad \left. + \inf_{u_{\mathcal{N}_L} \in L_{\mathcal{N}_L}} \|P_Q[\tilde{p}_S] - u_{\mathcal{N}_L}\|_{L^2(V, \mathbb{R})} \right). \end{aligned} \quad (5.73)$$

2. Sei  $\mathcal{T}_{h_{L,n}}(\Omega)$  reguläre Folge konformer Gitter, wobei  $h_{L,n} \xrightarrow{n \rightarrow \infty} 0$  gilt, und  $\mathcal{N}_{L,n}$  die dazugehörige Anzahl an Freiheitsgraden ist, für die  $\mathcal{N}_{L,n} \xrightarrow{n \rightarrow \infty} \infty$  gilt. Dann existiert eine konvergente Teilfolge  $(p_{S, n_m})_{n_m}$  zu den globalen Steuerungslösungen  $(p_{S, n})_n \in \mathbb{P}_S$  von  $[\mathbf{O}_3]_{D, \mathcal{N}_{L,n}}$ , welche gegen die globale Lösung  $\tilde{p}_S$  von  $[\mathbf{O}_3]_K$  konvergiert.

Der Beweis zu diesem Satz folgt analog zum Beweis von Satz 5.36. Des Weiteren lässt sich eine äquivalente Aussage wie in (5.38) auch für die Probleme  $[\mathbf{O}_3]_K$  herleiten, wobei

wiederum der zusätzliche Summand

$$2\|Q_V - Q_{V,\mathcal{N}_L}\|_{L^2(V,\mathbb{R})}$$

zu (5.73) hinzugefügt wird. In Abschnitt 5.1.5 haben wir die getrennten Systeme  $[V_{K,\text{Div}_l}]$ ,  $1 \leq l \leq K_A$ , betrachtet. Indem wir nun diese als Nebenbedingungen verwenden, erhalten wir:

- Finde zu  $Q_V \in L^2(V,\mathbb{R})$  jenes  $p_S \in \mathbb{P}_S$ ,  $P_Q[\omega, E_{\mathcal{N}}] \in L^2(V,\mathbb{R})$  und  $E_{\mathcal{N}_l} \in X_{\mathcal{N}_l}$ ,  $1 \leq l \leq K_A$ , wobei  $P_Q[\omega, E_{\mathcal{N}}]$  Lösung von  $[V_{K,Q}]$  zu

$$E_{\mathcal{N}} = \sum_{l=1}^{K_A} \sqrt{P_{in,l}} e^{-i\theta_l} E_{\mathcal{N}_l},$$

und  $E_{\mathcal{N}_l}$  Lösung zu  $[V_{D,\text{Div}_l,\mathcal{N}_l}]$  ist, sodass diese Lösung des folgenden Optimierungsproblems sind

$[O_3]_{K,\text{Div}}$

$$\min_{p_S \in \mathbb{P}_S, P_Q \in L^2(V)} Z_3(p_S, P_Q) = \frac{1}{2|V|} \|P_Q[\omega, E_{\mathcal{N}}] - Q_V\|_{L^2(V,\mathbb{R})}^2$$

$$\text{unter} \quad \begin{aligned} a_Q(P_Q, \phi) &= f_Q(\phi, E_{\mathcal{N}}, p_S), & \phi &\in L^2(V,\mathbb{R}), \\ a(E_{\mathcal{N}_l}, \varphi_{\mathcal{N}_l}, p_S) &= f(\varphi_{\mathcal{N}_l}, p_S), & \varphi_{\mathcal{N}_l} &\in X_{\mathcal{N}_l}, \quad 1 \leq l \leq K_A. \end{aligned}$$

Wie bereits in Abschnitt 5.1.5 für die beiden Probleme  $[O_k]_{K,\text{Div}}$ ,  $k = 1, 2$ , gezeigt, können ähnliche Aussagen wie in Satz 5.47 und Korollar 5.48 auch für  $[O_3]_{K,\text{Div}}$  hergeleitet werden. Ebenso lassen sich mehrere adjungierte Probleme definieren, welche in Bezug zu den einzelnen Feldern  $E_{\mathcal{N}_l}$  zu setzen sind. Diese Sätze möchten wir hier nicht formulieren, da die Herangehensweise dieselbe bleibt. Wir haben damit die Betrachtung der drei verschiedenen Optimierungsprobleme  $[O_k]_K$ ,  $k = 1, 2, 3$ , und ihrer Diskretisierung abgeschlossen und widmen uns nun den durch die RBM reduzierten Optimalsteuerungsprobleme.

### 5.3. RB Optimalsteuerungsprobleme

Die Berechnung der elektrischen Felder  $E_{\mathcal{N}}$  für die in Abschnitt 5.2 hergeleiteten diskreten Optimierungsprobleme  $[O_1]_{D,\mathcal{N}}$ ,  $[O_2]_{D,\mathcal{N}}$  und  $[O_3]_{D,\mathcal{N}_L}$  kann je nach Dimension des Problems sehr rechenintensiv sein und viel Zeit benötigen. Nachdem in der Optimierung eine Vielzahl an Lösungen  $\mathcal{E}_{\mathcal{N}}[p_S]$  bestimmt werden müssen, ergeben sich für den Optimierungsprozess sehr lange Rechenzeiten. Dies ist in einem System, welches in Echtzeit fungieren soll, nicht praktikabel, weshalb zunächst die Variationsgleichung  $[V_{D,\mathcal{N}}]$  mit der in Kapitel 4 eingeführten RBM reduziert und daraufhin eine Optimierung auf Basis des reduzierten Systems durchgeführt wird. Dabei betrachten wir innerhalb dieses Kapitels den reduzierten Formalismus zu einem allgemein formulierten Optimalsteuerungsproblem:

- Finde jene Steuerung  $p_S \in \mathbb{P}_S \subset \mathbb{R}^m$  und Zustand  $u_{\mathcal{N}} \in X_{\mathcal{N}}$ , sodass diese Lösung des folgenden Optimierungsproblems sind:

$$[O]_{D,\mathcal{N}} \quad \min_{\substack{p_S \in \mathbb{P}_S, \\ u_{\mathcal{N}} \in X_{\mathcal{N}}}} Z(p_S, p, u_{\mathcal{N}})$$

unter  $a(u_{\mathcal{N}}, \varphi, p_S, p) = f(\varphi, p_S, p), \quad \varphi \in X_{\mathcal{N}}.$

Um die RBM zur Reduktion der Nebenbedingung sinnvoll anwenden zu können, benötigen wir entsprechende Voraussetzungen an die Sesquilinear- und Linearformen bezüglich der Steuerparametervektoren.

**Voraussetzung 5.56.** Sei  $(p_S, p) \in \mathbb{P}_S \times \mathbb{P}$  die in der RBM betrachtete Parametermenge, wobei  $p_S$  die Steuerungsparameter darstellt. Die Sesquilinearform  $a(u_{\mathcal{N}}, \varphi, p_S, p)$ , die Linearform  $f(\varphi, p_S, p)$  und die Zielfunktion  $Z(p_S, p, u_{\mathcal{N}})$  besitzen die affinen Darstellungen

$$a(u_{\mathcal{N}}, \varphi, p_S, p) = \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) a_k(u_{\mathcal{N}}, \varphi), \quad f(\varphi, p_S, p) = \sum_{k=1}^{Q_f} \Theta_k^f(p_S, p) f_k(\varphi),$$

$$Z(p_S, p, u_{\mathcal{N}}) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) Z_k(u_{\mathcal{N}}),$$

wobei  $\Theta_k^Z : \mathbb{P}_S \times \mathbb{P} \rightarrow \mathbb{R}$  und  $Z_k : X_{\mathcal{N}} \rightarrow \mathbb{R}$ ,  $1 \leq k \leq Q_Z$  gilt.

Wir haben bereits in Abschnitt 4.1 über die Ritz-Galerkin- und Petrov-Galerkin-Projektion gesprochen. Nun geben wir die zu den beiden Projektionen gehörenden Optimalsteuerungsprobleme unter Voraussetzung 5.56 an:

- Finde jene Steuerung  $p_S \in \mathbb{P}_S$  und Zustand  $u_N^{RG} \in X_N^{RG}$ , sodass diese Lösung des folgenden Optimierungsproblems sind:

$$\begin{aligned}
 & [\text{O}]_{\text{D},N}^{RG} \\
 & \min_{\substack{p_S \in \mathbb{P}_S, \\ u_N^{RG} \in X_N^{RG}}} Z(p_S, u_N^{RG}) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) Z_k(u_N^{RG}) \\
 & \text{unter } \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) a_k(u_N^{RG}, \varphi) = \sum_{k=1}^{Q_f} \Theta_k^f(p_S, p) f_k(\varphi), \quad \varphi \in X_N^{RG}.
 \end{aligned}$$

- Finde jene Steuerung  $p_S \in \mathbb{P}_S$  und Zustand  $u_N^{PG} \in X_N^{PG}$ , sodass diese Lösung des folgenden Optimierungsproblems sind:

$$\begin{aligned}
 & [\text{O}]_{\text{D},N}^{PG} \\
 & \min_{\substack{p_S \in \mathbb{P}_S, \\ u_N^{PG} \in X_N^{PG}}} Z(p_S, u_N^{PG}) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) Z_k(u_N^{PG}) \\
 & \text{unter } \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p_S, p) \overline{\Theta_l^a(p_S, p)} a_k(u_N^{PG}, \mathcal{T}_N^l[\varphi]) \\
 & \quad = \sum_{k=1}^{Q_f} \sum_{l=1}^{Q_a} \Theta_k^f(p_S, p) \overline{\Theta_l^a(p_S, p)} f_k(\mathcal{T}_N^l[\varphi]), \quad \varphi \in X_N^{PG},
 \end{aligned}$$

wobei  $\mathcal{T}_N^l$ ,  $1 \leq l \leq Q_a$ , die durch (4.43) definierten Supremumsoperatoren sind.

*Bemerkung 5.57.* Gilt  $\varphi \in Y_N$  mit  $Y_N \neq X_N$  in  $[\text{O}]_{\text{D},N}$ , dann kann lediglich das reduzierte Problem  $[\text{O}]_{\text{D},N}^{PG}$  betrachtet werden.

Wie bereits zuvor in Abschnitt 5.1.1 benötigen wir die Konvergenz der Steuerungsprobleme  $[\text{O}]_{\text{D},N}^{RG}$ ,  $[\text{O}]_{\text{D},N}^{PG}$  bezüglich des Steuerparametervektors  $p_S$ . Bereits in der Definition der beiden reduzierten Probleme ist zu erkennen, dass diese Konvergenz lediglich von den Konvergenzeigenschaften der Parameterfunktionen  $\Theta_k^a$ ,  $\Theta_k^f$  und  $\Theta_k^Z$  abhängt.

**Voraussetzung 5.58.** Die Parameterfunktionen  $\Theta_k^a$ ,  $1 \leq k \leq Q_a$ , und  $\Theta_k^f$ ,  $1 \leq k \leq Q_f$ , sind bezüglich des Steuerparametervektors  $p_S$  stetig.

**Satz 5.59.** Sei  $p_{S,k} = (p_S)_k$  Folge in  $\mathbb{P}_S$  und  $\hat{p}_S \in \mathbb{P}_S$  mit

$$\|p_{S,k} - \hat{p}_S\|_{\mathbb{R}^{1+2K_A}} \xrightarrow{k \rightarrow \infty} 0.$$

## 5. Optimalsteuerung elektromagnetischer Felder

Gilt Voraussetzung 5.58, dann konvergieren die Lösungen  $u_{N,k}^{RG}$  und  $u_{N,k}^{PG}$  der Variationsprobleme in  $[O]_{D,N}^{RG}$  und  $[O]_{D,N}^{PG}$  zu  $p_{S,k}$  gegen die Lösungen  $\hat{u}_N^{RG}$  und  $\hat{u}_N^{PG}$  der Variationsprobleme in  $[O]_{D,N}^{RG}$  und  $[O]_{D,N}^{PG}$  zu  $\hat{p}_S$ . Gilt zudem die Stetigkeit der Parameterfunktionen  $\Theta_k^Z$ ,  $1 \leq k \leq Q_Z$ , bezüglich des Steuerparametervektors  $p_S$  und die stetige Abhängigkeit der affinen Komponenten  $Z_k$ ,  $1 \leq k \leq Q_Z$ , von der Lösungskomponente  $u_N$ , dann konvergieren auch die Zielfunktionale  $Z(p_S, u_{N,k}^{RG})$  bzw.  $Z(p_S, u_{N,k}^{PG})$  gegen  $Z(\hat{p}_S, \hat{u}_N^{RG})$  bzw.  $Z(\hat{p}_S, \hat{u}_N^{PG})$ .

Die Existenz einer Lösung zu den reduzierten Problemen  $[O]_{D,N}^{RG}$  und  $[O]_{D,N}^{PG}$  hängt direkt mit der Existenz einer Lösung  $u_N^{RG}$  bzw.  $u_N^{PG}$  des entsprechenden Variationsproblems zusammen. Hierzu müssen wir nun zwei unterschiedliche Fälle betrachten, welche wir bereits in Abschnitt 4.1 diskutiert haben. Setzen wir in die Variationsprobleme die entsprechenden Basisdarstellungen der reduzierten Räume  $X_N^{RG}$  und  $X_N^{PG}$  ein, so ergeben sich die zu lösenden linearen Gleichungssysteme

$$\left( \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) A_k \right) u_{RG}^N = \sum_{k=1}^{Q_f} \Theta_k^f(p_S, p) f_k$$

bzw.

$$\left( \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p_S, p) \overline{\Theta_l^a(p_S, p)} A_{k,l} \right) u_{PG}^N = \sum_{k=1}^{Q_f} \sum_{l=1}^{Q_a} \Theta_k^f(p_S, p) \overline{\Theta_l^a(p_S, p)} f_{k,l},$$

wobei  $A_k, A_{k,l} \in \mathbb{C}^{N \times N}$  und  $f_k, f_{k,l} \in \mathbb{C}^N$  durch die Basisdarstellung in den reduzierten Räumen generiert wird und  $u_{RG}^N, u_{PG}^N \in \mathbb{C}^N$  die Koeffizientenvektoren der entsprechenden Lösungen sind. Damit hängt die Lösbarkeit von der Matrix des linearen Gleichungssystems ab. Wie bereits in Abschnitt 4.1.1 diskutiert, kann für stetige koerzive Sesquilinearformen  $a$  die Existenz der reduzierten Lösung  $u_N^{RG}$  aus der Existenz der Lösung  $u_N$  des diskreten Systems gefolgert werden, denn für die Ritz-Galerkin-Projektion einer stetigen koerziven Sesquilinearform gilt für die Koerzivitätskonstante

$$\alpha_N^{RG}(p_S, p) = \inf_{u_N^{RG} \in X_N^{RG} \setminus \{0\}} \frac{|a(u_N^{RG}, \varphi, p_S, p)|}{\|u_N^{RG}\|^2} \geq \inf_{u_N \in X_N \setminus \{0\}} \frac{|a(u_N, \varphi, p_S, p)|}{\|u_N\|^2} = \alpha_N(p_S, p). \quad (5.74)$$

Aus dieser Ungleichung heraus folgt  $\alpha_N^{RG}(p_S, p) > 0$  und mit Satz A.17 die Existenz der Lösung  $u_N^{RG}$  und die Stabilität des reduzierten Systems. Für inf-sup-stabile Sesquilinearformen  $a$  erhalten wir die Existenz einer Lösung und die Stabilität der Petrov-Galerkin-Projektion  $u_N^{PG}$  aus Lemma 4.1 und (A.4). Auch wenn keine Existenz- und Stabilitätsaussagen für die Ritz-Galerkin-Projektion einer inf-sup-stabilen Sesquilinearform  $a$  gezeigt werden können, gibt es aus numerischer Sicht einige Vorteile, denn

die Verwendung der Ritz-Galerkin-Projektion erhält bei einer symmetrischen Sesquilinearform die Symmetrie und es wird deutlich weniger Speicherplatz benötigt, da weniger Teilmatrizen gespeichert werden müssen. Deshalb definieren wir folgende Voraussetzung, damit die Ritz-Galerkin-Projektionen für inf-sup-stabile Sesquilinearformen wohldefiniert ist.

**Voraussetzung 5.60.** Für eine stetige und inf-sup-stabile Sesquilinearform  $a$  existiert eine von der Parametermenge  $\mathbb{P}_S \times \mathbb{P}$  unabhängige untere Schranke  $\beta_{LB} > 0$  an die inf-sup-Konstante des reduzierten Systems

$$\beta_N^{RG}(p_S, p) = \inf_{u_N^{RG} \in X_N^{RG} \setminus \{0\}} \sup_{v_N^{RG} \in X_N^{RG} \setminus \{0\}} \frac{|a(u_N^{RG}, v_N^{RG}; p_S, p)|}{\|u_N^{RG}\|_{X_e} \|v_N^{RG}\|_{X_e}} \geq \beta_{LB}, \quad (p_S, p) \in \mathbb{P}_S \times \mathbb{P}.$$

Mit Voraussetzung 5.60 und (A.4) erhält man dann wiederum die Existenz einer Lösung der Ritz-Galerkin-Projektion, aber keine Stabilität, denn es ist nicht klar in welchem Verhältnis die inf-sup-Konstante  $\beta_N^{RG}(p_S, p)$  zu der diskreten inf-sup-Konstante  $\beta_N(p_S, p)$  steht. Wie in Abschnitt 5.1.2 definieren wir wieder die Steuerungs-Zustands-Operatoren  $\mathcal{E}_N$ ,  $\mathcal{E}_N^{RG}$  und  $\mathcal{E}_N^{PG}$  der einzelnen Variationsformulierungen. Je nach Eigenschaft der Sesquilinearform ergibt sich somit das folgende Fehlerresultat.

- Satz 5.61.**
1. Ist die Sesquilinearform koerziv oder gilt Voraussetzung 5.60 für eine inf-sup-stabile Sesquilinearform, dann existieren Lösungen zu  $[\mathbf{O}]_{D,N}^{RG}$ .
  2. Ist die Sesquilinearform inf-sup-stabil, dann existieren Lösungen zu  $[\mathbf{O}]_{D,N}^{PG}$ .
  3. Sei  $\hat{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}]_{D,N}^G$  mit  $X_N^G \subset X_N$ , wobei  $G$  für  $RG$  oder  $PG$  steht. Weiterhin sei  $\tilde{p}_S \in \mathbb{P}_S$  globale Lösung von  $[\mathbf{O}]_{D,N}$  zum Gitter  $\mathcal{T}_h(\Omega)$  und dazugehörigem Raum  $X_N \subset X_e$ . Dann ist der Fehler in der Zielfunktion bezüglich  $\hat{p}_S$  und  $\tilde{p}_S$  für  $[\mathbf{O}]_{D,N}$  gegeben durch

$$\begin{aligned} & |Z(\hat{p}_S, \mathcal{E}_N^G[\hat{p}_S]) - Z(\tilde{p}_S, \mathcal{E}_N[\tilde{p}_S])| \\ & \leq \sum_{k=1}^{Q_Z} |\Theta_k^Z(\hat{p}_S, p) - \Theta_k^Z(\tilde{p}_S, p)| Z_k(\mathcal{E}_N[\tilde{p}_S]) \\ & \quad + \sum_{k=1}^{Q_Z} \Theta_k^Z(\hat{p}_S, p) |Z_k(\mathcal{E}_N[\hat{p}_S]) - Z_k(\mathcal{E}_N^G[\hat{p}_S])| \\ & \quad + \sum_{k=1}^{Q_Z} \Theta_k^Z(\tilde{p}_S, p) |Z_k(\mathcal{E}_N[\tilde{p}_S]) - Z_k(\mathcal{E}_N^G[\tilde{p}_S])|, \end{aligned} \tag{5.75}$$

wobei  $Z_k$  die affinen Komponenten von  $Z$  wie in Voraussetzung 5.56 sind.

Der Beweis folgt mit denselben Schritten wie im Beweis von Satz 5.36, indem wir Zwischenterme einfügen und die Lösungseigenschaft ausnutzen. Wenn (5.75) betrachtet wird, dann ist der Fehler direkt von dem Verhalten der Parameterfunktionen  $\Theta_k^Z(p_S, p)$ ,  $1 \leq k \leq Q_Z$ , und der Approximationseigenschaft des reduzierten Problems abhängig. Für den Fall einer koerziven Sesquilinearform erhalten wir mit [14, Proposition 4.2] eine explizite Fehlerdarstellung

$$\|\mathcal{E}_{\mathcal{N}}[p_S] - \mathcal{E}_N^{RG}[p_S]\|_{X_e} \leq \frac{\|\hat{r}(p_S, p)\|_{X_e}}{\alpha_{LB}}$$

zwischen der hochdimensionalen Lösung  $\mathcal{E}_{\mathcal{N}}[p_S]$  und der reduzierten Lösung  $\mathcal{E}_N^{RG}[p_S]$ , wobei  $\hat{r}(p_S, p)$  der Riesz-Repräsentant des Residuums, wie in Abschnitt 4.1.4, und  $\alpha_{LB}$  eine untere Schranke an die Koerzivitätskonstante ist. Mit Proposition 4.3 erhalten wir die Fehlerdarstellung (4.29) für Petrov-Galerkin-Projektionen  $u_N^{PG}$  einer inf-sup-stabilen Sesquilinearform  $a$ . Mit (5.74) und Lemma 4.1 können die unteren Schranken  $\alpha_{LB}$  und  $\beta_{LB}$  unabhängig von der Wahl des Unterraumes  $X_N^{RG}$  bzw.  $X_N^{PG}$  bestimmt werden, weshalb der Fehler nur von dem Residuum abhängig ist. Wählen wir den Greedy-Algorithmus zur Bestimmung der Basis des reduzierten Raumes, dann hängt die Approximationseigenschaft nach Abschnitt 4.1.3 von der Kolmogorov-N-Breite (4.23) ab. Bewiesen wird diese Aussage in [50]. Das bedeutet, dass der Fehler (5.75) für einen reduzierten Raum mit guten Approximationseigenschaften sehr klein wird. Betrachten wir eine Ritz-Galerkin-Projektion einer inf-sup-stabilen Sesquilinearform in Zusammenhang mit Voraussetzung 5.60, dann lässt sich keine Aussage in Bezug auf die Approximationseigenschaft und den Fehler zeigen. In diesem Fall können wir nicht sagen ob eine Dimensionsvergrößerung des RB-Raumes eine Verbesserung darstellt oder nicht. Die Herleitung der ersten Variation und der Gateaux-Ableitung funktioniert analog zu Abschnitt 5.1.3, weswegen wir für deren Existenz eine Regularitäts-Voraussetzung an die Parameterfunktionen benötigen.

**Voraussetzung 5.62.** Es existieren die ersten Variationen  $\delta\Theta_k^a[p_S, \tilde{p}_S]$ ,  $1 \leq k \leq Q_a$ , und  $\delta\Theta_k^f[p_S, \tilde{p}_S]$ ,  $1 \leq k \leq Q_f$ , der Parameterfunktionen an der Stelle  $p_S \in \mathbb{O}_S \subset \mathbb{P}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^m$ .

**Voraussetzung 5.63.** Es existieren die Gateaux-Ableitungen  $\nabla\Theta_k^a[p_S]$ ,  $1 \leq k \leq Q_a$ , und  $\nabla\Theta_k^f[p_S]$ ,  $1 \leq k \leq Q_f$ , der Parameterfunktionen an der Stelle  $p_S$ .

**Satz 5.64.** Gilt Voraussetzung 5.62, dann existiert die erste Variation der Ritz-Galerkin-Projektion  $\delta\mathcal{E}_N^{RG}[p_S, \tilde{p}_S] \in X_N^{RG}$  und die erste Variation der Petrov-Galerkin-Projektion  $\delta\mathcal{E}_N^{PG}[p_S, \tilde{p}_S] \in X_N^{PG}$  an der Stelle  $p_S \in \mathbb{O}_S \subset \mathbb{P}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^m$ . Diese erfüllen die folgenden Gleichungen:

RG: Falls die Sesquilinearform  $a$  koerziv ist oder Voraussetzung 5.60 erfüllt, dann gilt:

$$\begin{aligned} \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) a_k(\delta \mathcal{E}_N^{RG}[p_S, \tilde{p}_S], \varphi) &= - \sum_{k=1}^{Q_a} \delta \Theta_k^a[p_S, \tilde{p}_S] a_k(\mathcal{E}_N^{RG}[p_S], \varphi) \\ &+ \sum_{k=1}^{Q_f} \delta \Theta_k^f[p_S, \tilde{p}_S] f_k(\varphi), \quad \varphi \in X_N^{RG}. \end{aligned} \quad (5.76)$$

PG: Falls die Sesquilinearform  $a$  inf-sup-stabil ist, dann gilt:

$$\begin{aligned} \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p_S, p) \overline{\Theta_l^a(p_S, p)} a_k(\delta \mathcal{E}_N^{PG}[p_S, \tilde{p}_S], \mathcal{T}_N^l[\varphi]) & \quad (5.77) \\ &= - \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p_S, p) \overline{\delta \Theta_l^a[p_S, \tilde{p}_S]} a_k(\mathcal{E}_N^{PG}[p_S], \mathcal{T}_N^l[\varphi]) \\ &- \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \delta \Theta_k^a[p_S, \tilde{p}_S] \overline{\Theta_l^a(p_S, p)} a_k(\mathcal{E}_N^{PG}[p_S], \mathcal{T}_N^l[\varphi]) \\ &+ \sum_{k=1}^{Q_f} \sum_{l=1}^{Q_a} \Theta_k^f(p_S, p) \overline{\delta \Theta_l^a[p_S, \tilde{p}_S]} f_k(\mathcal{T}_N^l[\varphi]) \\ &+ \sum_{k=1}^{Q_f} \sum_{l=1}^{Q_a} \delta \Theta_k^f[p_S, \tilde{p}_S] \overline{\Theta_l^a(p_S, p)} f_k(\mathcal{T}_N^l[\varphi]), \quad \varphi \in X_N^{PG}. \end{aligned}$$

Gilt zudem Voraussetzung 5.63, dann sind auch  $\delta \mathcal{E}_N^{RG}[p_S, \tilde{p}_S]$  und  $\delta \mathcal{E}_N^{PG}[p_S, \tilde{p}_S]$  Gâteaux-differenzierbar. Es gilt

$$\delta \mathcal{E}_N^G[p_S, \tilde{p}_S] = \nabla \mathcal{E}_N^G[p_S] \tilde{p}_S,$$

mit  $\nabla \mathcal{E}_N^G[p_S]$ , wobei  $G$  entweder für RG oder für PG steht, gegeben durch die Variationsprobleme (5.76) und (5.77), indem die ersten Variationen der Parameterfunktionen durch deren Gâteaux-Ableitungen ersetzt werden.

Der Beweis von Satz 5.64 folgt mit denselben Schritten wie in Satz 5.19 und dem daraus folgenden Korollar 5.20. Weiterführend erhalten wir mit weiteren Bedingungen an die affinen Komponenten  $Z_k$ ,  $1 \leq k \leq Q_Z$ , auch die Differenzierbarkeitseigenschaften für die Zielfunktion  $Z$ .

**Satz 5.65.** *Existieren die ersten Variationen  $\delta \Theta_k^Z[p_S, \tilde{p}_S]$ ,  $1 \leq k \leq Q_Z$ , und sind die affinen Komponenten  $Z_k$ ,  $1 \leq k \leq Q_Z$ , stetig abhängig von der Lösungskomponente  $u_N^G$ ,*

## 5. Optimalsteuerung elektromagnetischer Felder

---

dann existiert die erste Variation des Zielfunktional  $\mathcal{Z}^G[p_S] = Z(p_S, \mathcal{E}_N^G[p_S])$  und ist gegeben durch

$$\delta \mathcal{Z}^G[p_S, \tilde{p}_S] = \lim_{h \rightarrow 0} \left( \frac{Z(p_S + h\tilde{p}_S, \mathcal{E}_N^G[p_S + h\tilde{p}_S]) - Z(p_S, \mathcal{E}_N^G[p_S])}{h} \right),$$

wobei  $G$  entweder für  $RG$  oder für  $PG$  steht, an der Stelle  $p_S \in \mathbb{O}_S$  in Richtung  $\tilde{p}_S \in \mathbb{R}^m$  genau dann, wenn zudem die ersten Variationen

$$\delta \mathcal{Z}_k^G[p_S, \tilde{p}_S] = \lim_{h \rightarrow 0} \left( \frac{Z_k(\mathcal{E}_N^G[p_S + h\tilde{p}_S]) - Z_k(\mathcal{E}_N^G[p_S])}{h} \right), \quad 1 \leq k \leq Q_Z,$$

der affinen Komponenten existieren. Existieren zudem die Gâteaux-Ableitungen  $\nabla \Theta_k^Z[p_S]$ ,  $1 \leq k \leq Q_Z$ , dann existiert die Gâteaux-Ableitung  $\nabla \mathcal{Z}^G[p_S]$  mit

$$\delta \mathcal{Z}^G[p_S, \tilde{p}_S] = \nabla \mathcal{Z}^G[p_S] \tilde{p}_S$$

genau dann, wenn die affinen Komponenten  $Z_k$  Gâteaux-differenzierbar bezüglich des Steuerparametervektors  $p_S$  sind.

*Beweis.* Wir betrachten die Differenz des Kostenfunktional zu den beiden Parametern

$$\begin{aligned} \mathcal{Z}^G[p_S + h\tilde{p}_S] - \mathcal{Z}^G[p_S] &= \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S + h\tilde{p}_S, p) Z_k(\mathcal{E}_N^G[p_S + h\tilde{p}_S]) - \Theta_k^Z(p_S, p) Z_k(\mathcal{E}_N^G[p_S]) \\ &= \sum_{k=1}^{Q_Z} (\Theta_k^Z(p_S + h\tilde{p}_S, p) - \Theta_k^Z(p_S, p)) Z_k(\mathcal{E}_N^G[p_S + h\tilde{p}_S]) \\ &\quad + \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) (Z_k(\mathcal{E}_N^G[p_S + h\tilde{p}_S]) - Z_k(\mathcal{E}_N^G[p_S])), \end{aligned}$$

wobei  $h$  klein genug, sodass  $p_S + h\tilde{p}_S \in \mathbb{O}_S$ . Multiplizieren wir die Gleichung mit  $h^{-1}$  und bilden den Grenzwert  $h \rightarrow 0$ , dann erhalten wir mit den Voraussetzungen des Satzes und Satz 5.59

$$\begin{aligned} \lim_{h \rightarrow 0} \left( \frac{\mathcal{Z}^G[p_S + h\tilde{p}_S] - \mathcal{Z}^G[p_S]}{h} \right) &= \sum_{k=1}^{Q_Z} \delta \Theta_k^Z[p_S, \tilde{p}_S] Z_k(\mathcal{E}_N^G[p_S]) \\ &\quad + \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) \delta \mathcal{Z}_k^G[p_S, \tilde{p}_S] \end{aligned}$$

und damit die erste Aussage. Die zweite Aussage folgt dann direkt aus der obigen Gleichung durch das Ersetzen der ersten Variationen mit den Gâteaux-Ableitungen.  $\square$

Um das adjungierte Problem herzuleiten, betrachten wir eine besondere Klasse an Zielfunktionalen, welche dadurch gekennzeichnet ist, dass die ersten Variationen der affinen Komponenten folgende Form besitzen

$$\delta Z_k^G[p_S, \tilde{p}_S] = \tilde{Z}_k[p_S, \tilde{p}_S, \mathcal{E}_N^G[p_S], \delta \mathcal{E}_N^G[p_S, \tilde{p}_S]]. \quad (5.78)$$

Sie können also durch weitere Abbildungen  $\tilde{Z}_k$ ,  $1 \leq k \leq Q_Z$ , dargestellt werden, die wiederum von der ersten Variation  $\delta \mathcal{E}_N^G[p_S, \tilde{p}_S]$  abhängig sind. Die explizite Abhängigkeit von  $\delta \mathcal{E}_N^G[p_S, \tilde{p}_S]$  hilft uns im späteren Verlauf die ersten Variationen und die Gâteaux-Ableitungen aus dem Satz 5.65 mithilfe der noch zu definierenden adjungierten Probleme zu beschreiben.

*Bemerkung 5.66.* Die Optimalsteuerungsprobleme der Abschnitte 5.1 und 5.2 besitzen die eben genannte Abbildungseigenschaft.

Indem wir als Lagrange-Multiplikatoren  $\lambda_N^{RG} \in X_N^{RG}$  verwenden, wenn wir Ritz-Galerkin Optimierungsprobleme  $[O]_{D,N}^{RG}$  betrachten, und  $\lambda_N^{PG} \in \mathcal{T}_N^{p_S,p}[X_N^{PG}]$  bei Petrov-Galerkin Optimierungsproblemen  $[O]_{D,N}^{PG}$ , lassen sich die adjungierten Probleme und die Lagrange-Funktionen leicht herleiten:

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $u_N^{RG} \in X_N^{RG}$  des Optimierungsproblems  $[O]_{D,N}^{RG}$  die adjungierte Lösung  $\lambda_N^{RG} \in X_N^{RG}$ , sodass

$$[D]_{D,N}^{RG} \quad \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) a_k(\varphi, \lambda_N^{RG}) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) \tilde{Z}_k[p_S, \tilde{p}_S, \mathcal{E}_N^{RG}[p_S], \varphi], \quad \varphi \in X_N^{RG}.$$

- Bestimme zu den Optimierungsgrößen  $p_S \in \mathbb{P}_S$  und  $u_N^{PG} \in X_N^{PG}$  des Optimierungsproblems  $[O]_{D,N}^{PG}$  die adjungierte Lösung  $\lambda_N^{PG} \in \mathcal{T}_N^{p_S,p}[X_N^{PG}]$ , sodass

$$[D]_{D,N}^{PG} \quad \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) a_k(\varphi, \lambda_N^{PG}) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) \tilde{Z}_k[p_S, \tilde{p}_S, \mathcal{E}_N^{PG}[p_S], \varphi], \quad \varphi \in X_N^{PG}.$$

**Definition 5.67.** Die Lagrange-Funktion zu dem Optimierungsproblem  $[O]_{D,N}^G$ , wobei  $G$  entweder für  $RG$  oder  $PG$  steht, ist gegeben durch

$$\begin{aligned} \mathcal{L}^G(p_S, u_N^G, \lambda_N^G) = & \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) Z_k(u_N^G) + \operatorname{Re} \left( - \sum_{k=1}^{Q_a} \Theta_k^a(p_S, p) a_k(u_N^G, \lambda_N^G) \right. \\ & \left. + \sum_{k=1}^{Q_f} \Theta_k^f(p_S, p) f_k(\lambda_N^G) \right), \end{aligned} \quad (5.79)$$

## 5. Optimalsteuerung elektromagnetischer Felder

---

wobei  $p_S \in \mathbb{P}_S$ ,  $u_N^G \in X_N^G$  und  $\lambda_N^{RG} \in X_N^{RG}$  bzw.  $\lambda_N^{PG} \in \mathcal{T}_N^{p_S, p}[X_N^{PG}]$ .

Im Folgenden betrachten wir das Variationsproblem  $[D]_{D, N}^{PG}$  näher. Die adjungierte Lösung  $\lambda_N^{PG} \in \mathcal{T}_N^{p_S, p}[X_N^{PG}]$  wird nach Definition des Supremumsoperators  $\mathcal{T}_N^{p_S, p}$  eindeutig durch eine Funktion  $v_N^{PG} \in X_N^{PG}$  dargestellt, d.h.

$$\lambda_N^{PG} = \mathcal{T}_N^{p_S, p}[v_N^{PG}].$$

Mit der affinen Darstellung des Supremumsoperators  $\mathcal{T}_N^{p_S, p}$  durch die affinen Komponenten  $\mathcal{T}_N^k$ ,  $1 \leq k \leq Q_a$ , analog zu (4.42), lässt sich die adjungierte Lösung auch über

$$\lambda_N^{PG} = \sum_{l=1}^{Q_a} \Theta_l^a(p_S, p) \mathcal{T}_N^k[v_N^{PG}]$$

darstellen, wobei  $v_N^{PG} \in X_N^{PG}$  dann Lösung des Variationsproblems

$$\sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p_S, p) \overline{\Theta_l^a(p_S, p)} a_k(\varphi, \mathcal{T}_N^k[v_N^{PG}]) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) \tilde{Z}_k[p_S, \tilde{p}_S, \mathcal{E}_N^{PG}[p_S], \varphi] \quad (5.80)$$

mit  $\varphi \in X_N^{PG}$  ist. Wie bereits in Abschnitt 5.1.4 gesehen, können wir mit den adjungierten Lösungen die Berechnung der ersten Variation und der Gâteaux-Ableitung der Zielfunktion in Bezug auf die Steuerparametervektoren anpassen.

**Satz 5.68.** *Sei  $\lambda_N^G \in X_N^G$  die adjungierte Lösung zu  $[D]_{D, N}^G$ , wobei  $G$  entweder für  $RG$  oder  $PG$  steht. Gelten Voraussetzung 5.62, die Voraussetzungen von Satz 5.65 und existieren die ersten Variationen der affinen Komponenten  $\delta Z_k^G[p_S, \tilde{p}_S]$ ,  $1 \leq k \leq Q_Z$ , wobei diese die Darstellungen (5.78) besitzen, dann ist die erste Variation  $\delta Z^G[p_S, \tilde{p}_S]$  gegeben durch*

$$\delta Z^G[p_S, \tilde{p}_S] = \sum_{k=1}^{Q_Z} \delta \Theta_k^Z[p_S, \tilde{p}_S] Z_k(\mathcal{E}_N^G[p_S]) + \operatorname{Re} \left( \tilde{f}^G(p_S, \tilde{p}_S, \mathcal{E}_N^G[p_S], \lambda_N^G) \right),$$

wobei  $\tilde{f}^G$  die rechte Seite von (5.76) im Falle einer Ritz-Galerkin-Projektion und (5.77) im Falle der Petrov-Galerkin-Projektion ist.

Betrachten wir die Diskretisierungen von  $[D]_{D, N}^{RG}$  und (5.80), dann ergeben sich dieselben Matrizen wie in der Diskretisierung der Variationsprobleme in  $[O]_{D, N}^{RG}$  und  $[O]_{D, N}^{PG}$ , nur dass wir die liegenden Koeffizientenvektoren  $(v_{RG}^N)^H$  und  $(v_{PG}^N)^H$  suchen. Beispielsweise

erhalten wir zu  $v_{RG}^N$  das lineare Gleichungssystem

$$(v_{RG}^N)^H \left( \sum_{k=1}^{Q_a} \sum_{l=1}^{Q_a} \Theta_k^a(p_S, p) \overline{\Theta_l^a(p_S, p)} A_{k,l} \right) = \sum_{k=1}^{Q_Z} \Theta_k^Z(p_S, p) \tilde{M}_k,$$

mit  $(\tilde{M}_k)^H \in \mathbb{C}^N$  gegeben durch die Einträge  $\tilde{Z}_k[p_S, \tilde{p}_S, \mathcal{E}_N^G[p_S], \mathcal{E}_{N,l}]$  bezüglich der Basisfunktionen  $(\mathcal{E}_{N,l})_{l=1}^N$  des Raumes  $X_N^{PG}$ .

Zum Abschluss betrachten wir noch den Fall der getrennten Systeme, das heißt die Parameterfunktionen der Variationsformulierung in  $[O]_{D,\mathcal{N}}$  erfüllen die Voraussetzung 4.25 aus dem Abschnitt 4.2.5. Damit lässt sich das allgemeine Optimalsteuerungsproblem  $[O]_{D,\mathcal{N}}$  entsprechend der Theorie für getrennte Systeme auch folgendermaßen schreiben:

- Finde jene Steuerung  $p_S = [p_{S,1}^T, p_{S,2}^T]^T \in \mathbb{P}_S$  und Zustände  $[u_{\mathcal{N}}^{f,l}]_{l=1}^{Q_f} \in X_{\mathcal{N}}$ , sodass diese für  $p = [p_1^T, p_2^T]^T \in \mathbb{P}$  Lösung des folgenden Optimierungsproblems sind:

$[O]_{D,\mathcal{N},\text{Div}}$

$$\min_{\substack{p_S \in \mathbb{P}_S, \\ [u_{\mathcal{N}}^{f,l}]_{l=1}^{Q_f} \in X_{\mathcal{N}}}} Z \left( p_S, p, \sum_{l=1}^{Q_f} \Theta_l^{f,2}(p_{S,2}, p_2) u_{\mathcal{N}}^{f,l} \right)$$

$$\text{unter} \quad \sum_{k=1}^{Q_a} \Theta_k^a(p_{S,1}, p_1) a_k(u_{\mathcal{N}}^{f,l}, \varphi) = \Theta_l^{f,1}(p_{S,1}, p_1) f_l(\varphi),$$

$$\varphi \in X_{\mathcal{N}}, \quad 1 \leq l \leq Q_f.$$

Wie zuvor in Abschnitt 5.1.5 gezeigt, lässt sich die komplette Theorie bezüglich der adjungierten Probleme und der Differentiationstheorie auf dieselbe Weise auch auf die getrennten Systeme anwenden, weswegen wir dies nicht im Detail besprechen werden. Wollen wir für  $[O]_{D,\mathcal{N},\text{Div}}$  ein reduziertes Optimierungsproblem betrachten, dann müssen  $Q_f$  viele RB-Systeme der verschiedenen Nebenbedingungen bestimmt werden. Der Vorteil hierbei ist die Möglichkeit der Parallelisierung, indem die RB-Systeme simultan gelöst werden. Da zudem der betrachtete Parameterraum eines jeden RB-Systems kleiner ist, benötigen wir weniger Basisvektoren in den einzelnen RB-Systemen im Vergleich zur Betrachtung des RB-Systems zum nicht getrennten System. Dies resultiert möglicherweise in einer schnelleren Berechnung im Online-Phase und eines höheren Speicheraufwandes, welcher durch die Anzahl der RB-Systeme entsteht. Eine Betrachtung dieser Thematik widmen wir uns in Abschnitt 5.5 über die numerischen Ergebnisse zu den in Abschnitt 4.3.1 definierten Testproblemen.

#### 5.4. Steuerungsprobleme für den Erwärmungsprozess

Für ein optimales Aushärtungsergebnis ist in erster Linie eine gleichmäßige Erwärmung des Gebietes notwendig, welche wir durch Verwendung elektromagnetischer Wellen erreichen wollen. Wie wir jedoch bereits in den numerischen Ergebnissen zur RBM in Abbildung 4.25 gesehen haben, ist das durch eine Steuerung  $p_S \in \mathbb{P}_S$  generierte elektrische Feld  $E$  nicht unbedingt gleichmäßig. Dies ist bei größeren Objekten aus physikalischer Sicht auch nicht zu erwarten. Das bedeutet, dass mehrere optimale Steuerungen verwendet werden und diese zudem zueinander konträre Leistungsmuster liefern müssen. Neben der gleichmäßigen Erwärmung wollen wir zudem viel Energie in das Objekt einbringen um den Aushärtungseffekt zu beschleunigen und den Wirkungsgrad zu erhöhen. Wir beschreiben deshalb den Prozess des Findens  $K_{opt} \in \mathbb{N}$  komplementärer optimaler Steuerungen  $[p_{S,m}]_{m=1}^{K_{opt}}$  als iterativen Prozess und definieren zu einem einzelnen Steuerparameter  $p_S$  die folgenden Benchmark-Werte.

**Definition 5.69.** Sei  $V \subset \Omega$ . Sei zudem  $V_H \subset V$  gegeben als das Gebiet, welches aus dem heißesten 1% des Gebietes  $V$  besteht. Die Benchmark-Kriterien sind gegeben als:

- MP (Mean Power), durchschnittlich eingebrachte Leistung (physikalische Einheit von MP ist  $[\frac{W}{m^3}]$ ):

$$\text{MP} = \left( \frac{1}{|V|} \int_V |\mathcal{P}_{Q,\mathcal{N}_L}[p_S]|^2 d\Omega \right)^{\frac{1}{2}}$$

- RSDHR (Relative Standard Deviation in Heated Region), Gleichmäßigkeit:

$$\text{RSDHR} = \frac{\left( \frac{1}{|V|} \int_V |\mathcal{P}_{Q,\mathcal{N}_L}[p_S] - \text{MP}|^2 d\Omega \right)^{\frac{1}{2}}}{\text{MP}}$$

- HTQ (Hotspot to Target Quotient), relativer Wert zwischen Hotspot und zu erwärmenden Gebiet:

$$\text{HTQ} = \frac{\left( \frac{1}{|V_H|} \int_{V_H} |\mathcal{P}_{Q,\mathcal{N}_L}[p_S]|^2 d\Omega \right)^{\frac{1}{2}}}{\text{MP}}$$

Das Ziel bei diesen Benchmark-Werten ist es, dass der Wert MP so groß wie möglich ist, während die Werte RSDHR und HTQ so klein wie möglich sein sollen. Dabei korrelieren die Werte RSDHR und HTQ sehr stark miteinander, was wir im späteren Verlauf in

Kapitel 5.5.3 noch sehen werden. Wir beschränken uns deshalb zunächst auf die Werte RSDHR und MP. Die Betrachtung des HTQ-Wertes ist vor allem bei kleinen zu erwärmenden Gebieten sehr interessant und wird deshalb im Bereich der Hyperthermie, wobei hier HTQ für Hotspot Tumor Quotient steht, anstatt des RSDHR-Wertes verwendet, siehe [53, 56]. Bevor wir aus den Kostenfunktionalen der Kapitel 5.1 und 5.2 die zur Erwärmung benötigten iterativen Kostenfunktionale definieren, gehen wir im Folgenden auf ein Problem bei der Transformation in den Frequenzbereich ein. Diese Probleme treten dabei nur bei der Betrachtung von Kostenfunktionalen wie in  $[O_1]_K$  und  $[O_2]_K$  auf, welche nur vom elektrischen Feld  $E$  aber nicht von der Leistungsverteilung  $P_Q$  abhängig sind, auf. Für die Erklärung der Frequenzbereichsproblematik betrachten wir die stehende eindimensionale Welle im Frequenzbereich

$$\tilde{y}(\theta, x) = \sin(x) e^{-i\theta},$$

wobei hier  $\theta = \omega t$  für die Phase zur Frequenz  $\omega$  und dem Zeitpunkt  $t$  wie in Kapitel 2 steht. Diese stehende Welle ist grafisch in Abbildung 5.1 im Zeitbereich zu verschiedenen Phasenzeitpunkten dargestellt. Verwenden wir für die Frequenztransformation die Pointer-Arithmetik mit dem Faktor  $e^{-i\theta}$  anstatt von  $e^{i\theta}$ , dann ist

$$\tilde{y}(t, x) = \operatorname{Re}(\sin(x) e^{-i\theta}) = \operatorname{Re}(\sin(x) e^{-i\omega t})$$

die Wellenfunktion im Zeitbereich. Um das Problem zu verdeutlichen, betrachten wir nun das folgende Minimierungsproblem im Frequenzbereich

$$\min_{\substack{[\theta_k]_{k=1}^{K_A} \in [0, 2\pi], \\ [y_k]_{k=1}^{K_A} \in C([0, 2\pi])}} \int_0^{2\pi} \left| y \left( [y_k]_{k=1}^{K_A}, [\theta_k]_{k=1}^{K_A}, x \right) - \tilde{y}(0, x) \right|^2 d\Omega, \quad (5.81)$$

mit der Integration über die Raum-Komponente  $x$  und der kohärenten Überlagerung

$$y \left( [y_k]_{k=1}^{K_A}, [\theta_k]_{k=1}^{K_A}, x \right) = \sum_{k=1}^{K_A} y_k(x) e^{-i\theta_k}$$

der Wellenfunktionen  $[y_k]_{k=1}^{K_A}$  zu den Phasenzeitpunkten  $\theta_k$ . Das Problem bei dieser Optimierung ist der Phasenversatz möglicher Lösungen  $y$ , denn betrachten wir die um  $\pi$  in der Phase verschobene Funktion

$$\tilde{y}(\pi, x) = \sin(x) e^{-i\pi} = \sum_{k=1}^{K_A} \frac{1}{K_A} \sin(x) e^{-i\pi} = y([y_k]_{k=1}^{K_A}, [\pi]_{k=1}^{K_A}, x)$$

## 5. Optimalsteuerung elektromagnetischer Felder

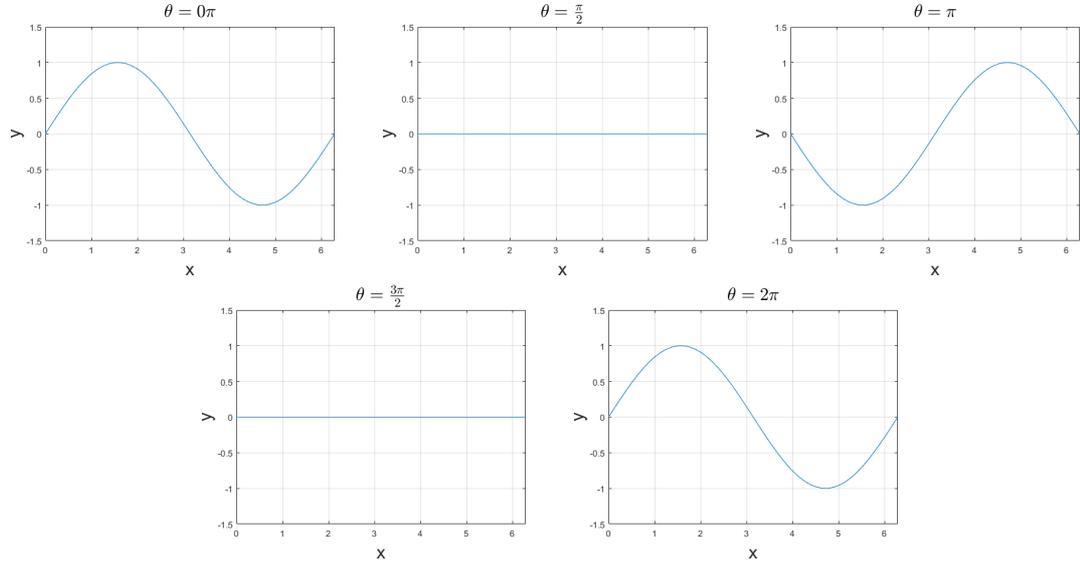


Abbildung 5.1: Stehende Welle  $\tilde{y}(t, x)$  zu verschiedenen Phasenzeitpunkten.

als mögliche Lösung, dann ergibt sich

$$\int_0^{2\pi} |\tilde{y}(\pi, x) - \tilde{y}(0, x)|^2 d\Omega = 4\pi.$$

Damit ist  $\tilde{y}(\pi, x)$  keine optimale Lösung des Problems (5.81), denn die optimale Lösung

$$\hat{y} \left( \left[ \frac{1}{K_A} \sin \right]_{k=1}^{K_A}, [0]_{k=1}^{K_A}, x \right) = \sum_{k=1}^{K_A} \frac{1}{K_A} \sin(x) e^0 = \tilde{y}(0, x)$$

liefert

$$\int_0^{2\pi} \left| \hat{y} \left( \left[ \frac{1}{K_A} \sin \right]_{k=1}^{K_A}, [0]_{k=1}^{K_A}, x \right) - \tilde{y}(0, x) \right|^2 d\Omega = 0.$$

Während wir bei der Optimierung von (5.81) den genauen Phasenzeitpunkt finden wollen, ist es für den Erwärmungsprozess eines Objektes unerheblich, ob die zu approximierende Funktion  $\tilde{y}$  zum Zeitpunkt  $\theta = 0$  oder  $\theta = \pi$  betrachtet wird, da beide Funktionen dasselbe Leistungsmuster liefern. Dies bedeutet, dass für den Erwärmungsprozess  $\tilde{y}(\pi, x)$

und  $\tilde{y}(0, x)$  als äquivalent gelten. Hierzu betrachten wir die Phasenverschiebung

$$y \left( [y_k]_{k=1}^{K_A}, [\theta_k + \tilde{\theta}]_{k=1}^{K_A}, x \right) = y \left( [y_k]_{k=1}^{K_A}, [\theta_k]_{k=1}^{K_A}, x \right) e^{-i\tilde{\theta}} \quad (5.82)$$

einer Überlagerung stehender Wellen um die Phase  $\tilde{\theta}$ . Diese einfache Multiplikation können wir uns zu Nutze machen, um erstens eine Art Äquivalenzbetrachtung durchzuführen und zweitens die Rechenzeit des Optimierungsprozesses von (5.81) zu reduzieren. Im Beispiel (5.81) ermöglichen wir dies, indem wir

$$\min_{\substack{[\theta_k]_{k=1}^{K_A} \in [0, 2\pi], \\ [y_k]_{k=1}^{K_A} \in C([0, 2\pi])}} \min_{\tilde{\theta} \in [0, 2\pi]} \int_0^{2\pi} \left| y \left( [y_k]_{k=1}^{K_A}, [\theta_k]_{k=1}^{K_A}, x \right) e^{-i\tilde{\theta}} - \tilde{y}(0, x) \right|^2 d\Omega \quad (5.83)$$

betrachten. Die Verringerung der Rechenzeit ergibt sich dabei aus der Multiplikation einer in der Phase  $\tilde{\theta}$  verschobenen Überlagerung  $y$  in (5.82), denn die Rechenzeit der Multiplikation und Evaluierung des Zielfunktional in (5.83) ist meist deutlich geringer als die Rechenzeit des zur Optimierung verwendeten Algorithmus. Auch wenn der exakte Phasenwert damit kein echter Steuerparameter für den Erwärmungsprozess ist, sind vor allem die Phasendifferenzen zwischen den zu überlagernden Wellen entscheidend. Beispielsweise löschen sich die beiden stehenden Wellen  $\tilde{y}(\pi, x)$  und  $\tilde{y}(0, x)$  bei der Überlagerung aus, während die Überlagerung von  $\tilde{y}(0, x)$  und  $\tilde{y}(0, x)$  zu einer Verdoppelung der Amplitude führt. Als Darstellung für die Verschiebung der Phasenwerte des Steuerparametervektors  $p_S$  führen wir deshalb die Phasenversatzfunktion  $\rho(p_S, \theta)$  ein, welche alle Phasenparameter in  $p_S$  um die Phase  $\theta$  verschiebt. Zudem verwenden wir anstatt der Minimierung über das komplette Intervall  $[0, 2\pi]$  im Folgenden eine äquidistante Zerlegung des Intervalls  $[0, 2\pi]$  und bezeichnen die Menge der Stützstellen mit  $I_\theta$ . Wie bereits zum Anfang des Kapitels erklärt, benötigen wir für die gleichmäßige Erwärmung mehrere zueinander komplementäre Steuerungen. Eine erste Möglichkeit diesbezüglich wäre die Verwendung von mehreren lokalen Minima. Das Problem hierbei ist, dass dadurch nicht gewährleistet werden kann, dass die Felder zueinander ergänzend sind. Das heißt unterschiedliche Steuerungen könnten ein vergleichbares elektrisches Feld im zu erwärmenden Gebiet erzeugen. Das Suchen von optimalen Steuerungen  $[p_{S,l}^{opt}]_{l=1}^{K_{opt}}$  beschreiben wir deshalb als iterative Optimierungsprobleme und definieren hierfür den

Strafterm zum  $l$ -ten Iterationsschritt als

$$\begin{aligned} \mathcal{R}_{\max} \left( p_{S,l}, \left[ \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}] \right]_{m=1}^{l-1} \right) &= \max_{m \in \{1, \dots, l-1\}} \frac{\mathcal{Z}_k[p_{S,m}^{opt}; 0]}{\min_{\theta \in I_\theta} Z_k(\rho(p_{S,l}, \theta), \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}])}, \quad \text{für } k = 1, 2, \\ \mathcal{R}_{\max} \left( p_{S,l}, \left[ \mathcal{P}_{\mathcal{N}_L}[p_{S,m}^{opt}] \right]_{m=1}^{l-1} \right) &= \max_{m \in \{1, \dots, l-1\}} \frac{\mathcal{Z}_k[p_{S,m}^{opt}; 0]}{Z_k(p_{S,l}, \mathcal{P}_{Q, \mathcal{N}_L}[p_{S,m}^{opt}])}, \quad \text{für } k = 3, \end{aligned} \quad (5.84)$$

wobei  $\mathcal{Z}_k[p_{S,m}^{opt}; 0]$  das Kostenfunktional  $Z_k(p_{S,m}^{opt}, \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}])$  zu  $E_{V, \mathcal{N}} = 0$  darstellt. Anstatt der Verwendung des Maximums könnten wir als Strafterm auch die Summe der Werte

$$\begin{aligned} \mathcal{R}_{\Sigma} \left( p_{S,l}, \left[ \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}] \right]_{m=1}^{l-1} \right) &= \sum_{m=1}^{l-1} \frac{\mathcal{Z}_k[p_{S,m}^{opt}; 0]}{\min_{\theta \in I_\theta} Z_k(\rho(p_{S,l}, \theta), \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}])}, \quad \text{für } k = 1, 2, \\ \mathcal{R}_{\Sigma} \left( p_{S,l}, \left[ \mathcal{P}_{\mathcal{N}_L}[p_{S,m}^{opt}] \right]_{m=1}^{l-1} \right) &= \sum_{m=1}^{l-1} \frac{\mathcal{Z}_k[p_{S,m}^{opt}; 0]}{Z_k(p_{S,l}, \mathcal{P}_{Q, \mathcal{N}_L}[p_{S,m}^{opt}])}, \quad \text{für } k = 3, \end{aligned} \quad (5.85)$$

betrachten, jedoch ergibt sich dann bei der Betrachtung der Ableitungen ein erhöhter Rechenaufwand. Neben den beiden Straftermen (5.84) und (5.85) gibt es noch viele weitere Möglichkeiten, die wir aber nicht genauer untersuchen werden. Die Idee der Strafterme (5.84) und (5.85) ist es den Abstand des aktuellen Zustands  $\mathcal{E}_{\mathcal{N}}[p_{S,l}]$  zu allen vorherigen Zuständen  $[\mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}]]_{m=1}^{l-1}$  zu maximieren. Zudem werden die Abstände mit dem Zielfunktionswert der vorherigen Zustände zu  $E_{V, \mathcal{N}} = 0$  skaliert. Dabei kann die Evaluierung der Kostenfunktionale  $Z_k$ ,  $k = 1, 2, 3$ , zu dem elektromagnetischen Feld  $E_{V, \mathcal{N}} = 0$  oder dem Leistungsmuster  $Q_{V, \mathcal{N}_L} = 0$  als Äquivalent zur eingebrachten Energie in das zu erwärmende Gebiet betrachtet werden. Diesen Umstand werden wir auch später bei der endgültigen Definition der Zielfunktionale für den Erwärmungsprozess verwenden. Im Folgenden schreiben wir für die Strafterme kurz  $\mathcal{R}_{\max}^{k,l}$  mit  $k = 1, 2, 3$ . Um die iterativen Optimierungsprobleme und deren Kostenfunktionale zu definieren betrachten wir zwei unterschiedliche Fälle:

1.  $E_{V, \mathcal{N}} \neq 0, Q_{V, \mathcal{N}_L} \neq 0,$
2.  $E_{V, \mathcal{N}} = 0, Q_{V, \mathcal{N}_L} = 0.$

Dabei beschreibt der erste Fall die Approximation eines explizit vorgegebenen Feldes, während wir beim zweiten Fall Zielfunktionale definieren wollen, die ohne vorgegebenes Feld  $E_{V, \mathcal{N}}$  die Benchmark-Werte RSDHR und MP optimieren. Wir definieren nun die iterativen Optimalsteuerungsprobleme für die oben genannten Fälle:

**Definition 5.70.** Sei  $E_{V,\mathcal{N}} \neq 0$ ,  $Q_{V,\mathcal{N}_L} \neq 0$  gegeben. Die  $l$ -ten Optimalsteuerungsprobleme des iterativen Prozesses für die gleichmäßige Erwärmung sind:

$[\mathbf{I}_1]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \min_{\theta \in I_\theta} [\mathcal{Z}_k[\rho(p_S, \theta); E_{V,\mathcal{N}}]] + \mathcal{R}_{\max}^{k,l} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \mathcal{Z}_k[p_S; Q_{V,\mathcal{N}_L}] + \mathcal{R}_{\max}^{k,l} \right]$$

$[\mathbf{I}_1^{rel}]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\min_{\theta \in I_\theta} [\mathcal{Z}_k[\rho(p_S, \theta); E_{V,\mathcal{N}}]] + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\mathcal{Z}_k[p_S; Q_{V,\mathcal{N}_L}] + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

$[\mathbf{I}_2]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \min_{\theta \in I_\theta} \left[ \mathcal{Z}_k \left[ \rho(p_S, \theta); E_{V,\mathcal{N}} - \sum_{m=1}^{l-1} \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}] \right] \right] + \mathcal{R}_{\max}^{k,l} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \mathcal{Z}_k \left[ p_S; Q_{V,\mathcal{N}_L} - \sum_{m=1}^{l-1} \mathcal{P}_{\mathcal{N}_L}[p_{S,m}^{opt}] \right] + \mathcal{R}_{\max}^{k,l} \right]$$

$[\mathbf{I}_2^{rel}]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\min_{\theta \in I_\theta} \left[ \mathcal{Z}_k \left[ \rho(p_S, \theta); E_{V,\mathcal{N}} - \sum_{m=1}^{l-1} \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}] \right] \right] + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\mathcal{Z}_k \left[ p_S; Q_{V,\mathcal{N}_L} - \sum_{m=1}^{l-1} \mathcal{P}_{\mathcal{N}_L}[p_{S,m}^{opt}] \right] + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

Bei den eben definierten Kostenfunktionalen in Definition 5.70 betrachten wir im Gegensatz zu (5.15), (5.16) und (5.58) die expliziten Abhängigkeit hinsichtlich der Zielfunktionen  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$ . Wir betrachten nun die eben definierten Steuerungsprobleme

hinsichtlich der Benchmark-Werte RSDHR und MP. Die Gleichmäßigkeit erhalten wir zum einen durch die Wahl von  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  als gleichmäßige Feldverteilungen und durch den Strafterm  $\mathcal{R}_{\max}^{k,l}$ , der gleiche Felder stark bestraft. Der Unterschied zwischen den Steuerungsproblemen  $[I_1]_k^l$  und  $[I_2]_k^l$  ist, dass wir den Zielwert im ersten Summanden aktualisieren. Durch die Aktualisierung versuchen wir das zu den vorher angeregten Feldern komplementäre Feldbild zu approximieren, was wiederum die Gleichmäßigkeit erhöhen soll. Um zudem die mittlere Leistung MP zu erhöhen, teilen wir in den relativen Fällen durch den aktuellen Wert  $\mathcal{Z}_k[p_S; 0]$  des verwendeten Kostenfunktional, der als Äquivalent zur eingebrachten Leistung gesehen werden kann. Der Grund für die Verwendung dieser Kostenfunktionale anstatt der direkten Verwendung der Benchmark-Werte RSDHR und MP ist deren sehr aufwändige Berechnung, was wir in Abschnitt 5.5 noch genauer betrachten werden. Als nächstes widmen wir uns der Herleitung einer möglichen Wahl gleichmäßiger Steuerungsfelder  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$ . Die Herleitung funktioniert dabei in zwei Schritten. Für die Herleitung des elektrischen Feldes  $E_{V,\mathcal{N}}$  betrachten wir im ersten Schritt die diskrete Integralgleichung:

1.Schritt: Seien  $\tilde{e}_x, \tilde{e}_y, \tilde{e}_z \in \mathbb{C}$  beliebig aber fest. Bestimme

$$\tilde{E}_{V,\mathcal{N}} \in \{U_{\mathcal{N}} \in Y_{\mathcal{N}} : \gamma_t[U_{\mathcal{N}}] = [\tilde{e}_x, \tilde{e}_y, \tilde{e}_z]^T \text{ auf } \Gamma_T\} \subset Y_{\mathcal{N}},$$

sodass

$$\int_{\Omega} \left( \tilde{E}_{V,\mathcal{N}} - [\tilde{e}_x, \tilde{e}_y, \tilde{e}_z]^T \right) \cdot \varphi \, d\Omega = 0, \quad \varphi \in Y_{\mathcal{N}},$$

gilt.

Durch Lösen dieser Integralgleichung erhalten wir ein gleichmäßiges Feld in der FEM-Darstellung. Da wir die maximal mögliche Energie jeder Antenne auf 1 Watt beschränkt haben, müssen wir das Feld  $\tilde{E}_{V,\mathcal{N}}$  noch entsprechend der Anzahl an Antennen  $K_A$  und bezüglich des zu erwärmenden Gebiets  $V \subset \Omega$  normalisieren. Hierzu lösen wir in einem zweiten Schritt die Integralgleichung:

2.Schritt: Sei  $V \subset \Omega$  das zu erwärmende Gebiet. Seien  $e_x, e_y, e_z \in \mathbb{C}$  gegeben durch

$$e_m = \tilde{e}_m \sqrt{\frac{K_A}{\int_V |\frac{1}{2} \omega_{\max} \varepsilon_0 \varepsilon_{\max, \text{Im}} |\tilde{E}_{V,\mathcal{N}}|^2 \, d\Omega}},$$

mit  $m \in \{x, y, z\}$ . Bestimme

$$E_{V,\mathcal{N}} \in \{U_{\mathcal{N}} \in Y_{\mathcal{N}} : \gamma_t[U_{\mathcal{N}}] = [e_x, e_y, e_z]^T \text{ auf } \Gamma_T\} \subset Y_{\mathcal{N}},$$

sodass

$$\int_{\Omega} \left( E_{V,\mathcal{N}} - [e_x, e_y, e_z]^T \right) \cdot \varphi \, d\Omega = 0, \quad \varphi \in Y_{\mathcal{N}},$$

gilt.

Damit erhalten wir das Feld  $E_{V,\mathcal{N}}$ , welches zum einen gleichmäßig bezüglich jeder Raumkomponente ist und deren absorbierte Leistung im betrachteten Gebiet  $V$  genau  $K_A$  Watt entspricht. Analog erhalten wir so auch die gleichmäßige und normierte Leistungsverteilung  $Q_{V,\mathcal{N}_L}$  durch die beiden Schritte:

1.Schritt: Seien  $\tilde{q} \in \mathbb{R}$  fest aber beliebig. Bestimme

$$\tilde{Q}_{V,\mathcal{N}_L} \in L_{\mathcal{N}_L} \subset L^2(V, \mathbb{R})$$

sodass

$$\int_V (\tilde{Q}_{V,\mathcal{N}_L} - \tilde{q}) \phi_{\mathcal{N}_L} \, d\Omega = 0, \quad \phi_{\mathcal{N}_L} \in L_{\mathcal{N}_L},$$

gilt.

2.Schritt: Sei  $q \in \mathbb{R}$  gegeben durch

$$q = \tilde{q} \frac{K_A}{\int_V |\tilde{Q}_{V,\mathcal{N}_L}| \, d\Omega}.$$

Bestimme

$$Q_{V,\mathcal{N}_L} \in L_{\mathcal{N}_L} \subset L^2(V, \mathbb{R})$$

sodass

$$\int_V (Q_{V,\mathcal{N}_L} - q) \phi_{\mathcal{N}_L} \, d\Omega = 0, \quad \phi_{\mathcal{N}_L} \in L_{\mathcal{N}_L},$$

gilt.

Wie wir noch in Abschnitt 5.5 sehen werden, ist die Wahl der Funktionen  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  ein kritischer Faktor bei der Betrachtung der in Definition 5.70 beschriebenen Steuerungsprobleme. Wir beschreiben deshalb im Folgenden weitere von der Wahl der Funktionen  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  unabhängige Steuerungsprobleme.

**Definition 5.71.** Sei  $E_{V,\mathcal{N}} = 0$ ,  $Q_{V,\mathcal{N}_L} = 0$  gegeben. Die  $l$ -ten Optimalsteuerungsprobleme des iterativen Prozesses für die gleichmäßige Erwärmung sind:

$[\mathbf{I}_3]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{1}{\min_{\theta \in I_\theta} [\mathcal{Z}_k[\rho(p_S, \theta); 0]]} + \mathcal{R}_{\max}^{k,l} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{1}{\mathcal{Z}_k[p_S; 0]} + \mathcal{R}_{\max}^{k,l} \right]$$

$[\mathbb{I}_3^{rel}]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\frac{1}{\min_{\theta \in I_\theta} [\mathcal{Z}_k[\rho(p_S, \theta); 0]]} + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\frac{1}{\mathcal{Z}_k[p_S; 0]} + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

$[\mathbb{I}_4]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{1}{\min_{\theta \in I_\theta} \left[ \mathcal{Z}_k \left[ \rho(p_S, \theta); \sum_{m=1}^{l-1} \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}] \right] \right]} + \mathcal{R}_{\max}^{k,l} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{1}{\mathcal{Z}_k \left[ p_S; \sum_{m=1}^{l-1} \mathcal{P}_{\mathcal{N}_L}[p_{S,m}^{opt}] \right]} + \mathcal{R}_{\max}^{k,l} \right]$$

$[\mathbb{I}_4^{rel}]_k^l$   $k = 1, 2$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\frac{1}{\min_{\theta \in I_\theta} [\mathcal{Z}_k[\rho(p_S, \theta); \sum_{m=1}^{l-1} \mathcal{E}_{\mathcal{N}}[p_{S,m}^{opt}]]]} + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

$k = 3$ :

$$\min_{p_S \in \mathbb{P}_S} \left[ \frac{\frac{1}{\mathcal{Z}_k[p_S; \sum_{m=1}^{l-1} \mathcal{P}_{\mathcal{N}_L}[p_{S,m}^{opt}]]} + \mathcal{R}_{\max}^{k,l}}{\mathcal{Z}_k[p_S; 0]} \right]$$

Während wir in den angegebenen Steuerungsproblemen in Definition 5.70 den Abstand zu den gleichmäßig definierten Funktionen  $E_{V, \mathcal{N}}$ ,  $Q_{V, \mathcal{N}_L}$  minimieren wollen, versuchen wir bei den angegebenen Steuerungsproblemen in Definition 5.71 zunächst viel Energie einzubringen und daraufhin über den Strafterm die Gleichmäßigkeit herzustellen. Die Existenz von Lösungen der in diesem Abschnitt definierten iterativen Steuerungsprobleme erhalten wir aufgrund der Stetigkeit der Zielfunktionale  $\mathcal{Z}_k$ ,  $k = 1, 2, 3$ , in Lemma 5.5, 5.8 und 5.42. Durch die Existenzaussagen in Abschnitt 5.1 und 5.2 erhalten wir ebenso die Existenz der ersten Variation und der Gâteaux-Ableitung der Steuerungsfunktionen aus Definition 5.70 und 5.71 unter den Voraussetzungen 5.15 und 5.16. Die Darstellungen der ersten Variation und der Gâteaux-Ableitung ergeben sich dabei wiederum

über die Zielfunktionale  $Z_k$ ,  $k = 1, 2, 3$ , und den Differentiationsregeln. Um nun die iterativen Steuerungsprobleme miteinander zu vergleichen und die iterative Verbesserung bei Überlagerung von Leistungsmustern zu beschreiben, betrachten wir in den numerischen Ergebnissen anstatt der Benchmark-Werte in Definition 5.69 die nachfolgenden kumulativ-iterativen Summenwerte.

**Definition 5.72.** Sei  $1 \leq l \leq K_{opt}$  und  $V \subset \Omega$ . Sei zudem  $V_H \subset V$  gegeben als das Gebiet, welches aus dem heißesten 1% des Gebietes  $V$  besteht. Die Benchmark-Kriterien sind gegeben als:

- MP Summe (Mean Power), kumulativ-iterative durchschnittlich eingebrachte Leistung (physikalische Einheit von MP Summe ist  $[\frac{W}{m^3}]$ ):

$$\text{MP Summe} = \left( \frac{1}{|V|} \int_V \left| \sum_{m=1}^l \mathcal{P}_{Q, \mathcal{N}_L} [p_{S,m}] \right|^2 d\Omega \right)^{\frac{1}{2}} \quad (5.86)$$

- RSDHR (Relative Standard Deviation in Heated Region), kumulativ-iterative Gleichmäßigkeit:

$$\text{RSDHR Summe} = \frac{\left( \frac{1}{|V|} \int_V \left| \sum_{m=1}^l \mathcal{P}_{Q, \mathcal{N}_L} [p_{S,m}] - \text{MP Summe} \right|^2 d\Omega \right)^{\frac{1}{2}}}{\text{MP Summe}} \quad (5.87)$$

- HTQ Summe (Hotspot to Target Quotient), kumulativ-iterativer relativer Wert zwischen Hotspot und zu erwärmenden Gebiet:

$$\text{HTQ Summe} = \frac{\left( \frac{1}{|V_H|} \int_{V_H} \left| \sum_{m=1}^l \mathcal{P}_{Q, \mathcal{N}_L} [p_{S,m}] \right|^2 d\Omega \right)^{\frac{1}{2}}}{\text{MP Summe}} \quad (5.88)$$

Der Wert  $l$  in Definition 5.72 beschreibt dabei entweder die Betrachtung von  $l$  beliebigen Steuerparametern oder die nach  $l$  Iterationen erhaltenen optimalen Steuerparameter  $[p_{S,m}]_{m=1}^l$  für die Überlagerung der Leistungsmuster und die Auswertung der Benchmark-Kriterien. Wir haben damit verschiedene Optimalsteuerungsprobleme für den optimalen Erwärmungsprozess und die für den Benchmark notwendigen Kriterien definiert. Die unterschiedlichen Optimalsteuerungsprobleme werden wir im nächsten Kapitel anhand der reduzierten Modelle aus Abschnitt 4.3.1 auf ihre Funktionsweise bezüglich verschiedener existierender Optimierungsalgorithmen testen.

### 5.5. Numerische Ergebnisse

Um die Steuerungsprobleme und deren Lösungen bezüglich der Benchmark-Werte in Definition 5.72 bewerten zu können, benötigen wir zum einen Vergleichsalgorithmen und zum anderen die gleichen Ziele. Als Ziel setzen wir uns deshalb die Bestimmung von  $K_{opt} = 11$  Steuerungen  $[p_{S,l}^{opt}]_{l=1}^{K_{opt}}$ , die wir daraufhin bezüglich der Benchmark-Werte RSDHR und MP vergleichen werden. Zudem betrachten wir zwei unterschiedliche Fälle:

- a)  $P_{in,m} \in [0, 1], 1 \leq m \leq K_A$ .
- b)  $P_{in,m} = 1, 1 \leq m \leq K_A$ .

In dem ersten Fall lassen wir die Variation der Leistung an den einzelnen Antennen zu, während wir diese im zweiten Fall auf 1 Watt beschränken. Wie wir noch im Laufe dieses Abschnitts sehen werden, können wir eine bessere Gleichmäßigkeit erreichen, indem wir die Leistungen variieren. Beispielsweise ist das konstante elektrische Feld  $E_{\mathcal{N}} = 0$  bezüglich dem Gleichmäßigkeitswert RSDHR optimal. Da aber in vielen Anwendungsfällen die volle Leistungskapazität der Antennen ausgenutzt werden soll, betrachten wir auch den zweiten Fall mit der Einschränkung auf einen Watt pro Antenne. Als Vergleichsalgorithmen für die beiden Fälle werden wir die folgenden Algorithmen verwenden:

- a)
  - Random: Tendenz bei zufälliger Auswahl an Steuerparametervektoren (Mittelwerte über 10 Durchläufe)
  - SVD: Singulärwertzerlegung der S-Parametermatrix an festgelegten Frequenzpunkten  $f_k = 2.4 + 0.01k, 1 \leq k \leq 11$ , und Verwendung des Singulärvektors zum kleinsten Eigenwert
- b)
  - Random: Tendenz bei zufälliger Auswahl an Steuerparametervektoren (Mittelwerte über 10 Durchläufe)
  - RSDHR-Opt: Optimierung bezüglich des Benchmark-Wertes RSDHR Summe

Bis auf den Vergleichsalgorithmus SVD ist die Funktionsweise der verwendeten Algorithmen aus dem Kontext heraus klar, weshalb wir die Idee hinter dem SVD-Algorithmus noch genauer betrachten werden. Wie bereits in Abschnitt 4.3 erwähnt, stellen die in dieser Arbeit betrachteten Modelle sogenannte  $N$ -Port Netzwerke dar, deren Eingangs-Ausgangsverhalten über die S-Parametermatrizen dargestellt werden können

$$s = Se,$$

wobei  $e \in \mathbb{C}^N$  den Eingang,  $s \in \mathbb{C}^N$  den Ausgang und  $S \in \mathbb{C}^{N \times N}$  die S-Parametermatrix darstellt. Dabei werden die einzelnen S-Parameter definiert über den relativen Wert aus

reflektierter, bzw. transmittierter, Leistung und der eingespeisten Leistung. Betrachtet man demnach zu festen Frequenzen die Minimierung des Quotienten

$$\min_{\substack{s, e \in \mathbb{C}^N \\ s = Se}} \frac{s^H s}{e^H e} = \min_{e \in \mathbb{C}^N} \frac{e^H S^H S e}{e^H e}$$

der reflektierten und eingespeisten Leistungen, wobei die reflektierten Leistungen mit der S-Parametermatrix über die Eingangsleistungen definiert werden können, dann soll dieser Quotient minimiert werden. Das entstehende Optimierungsproblem zur Bestimmung der Phasen und Leistungen an den Antennen, gegeben durch den komplexen Eingangsvektor  $e \in \mathbb{C}^N$ , wird gelöst durch Wahl des Singulärvektors zum kleinsten Singulärwert der S-Parametermatrix. Dies entspricht also der Minimierung der Reflexion und damit auch der Maximierung der Energie innerhalb der Kavität relativ zur eingespeisten Leistung. Für die weitere Theorie der S-Parameter eines  $N$ -Port Netzwerks verweisen wir wieder auf [37]. Die Betrachtung des RSDHR-Optimierung-Algorithmus für den Fall a) wurde aufgrund des zu großen Suchraums  $\mathbb{P}_S$  im Fall a) und der sehr langen Laufzeit, wie wir noch im späteren Verlauf sehen werden, nicht durchgeführt. Ebenso wurde der SVD-Algorithmus nicht für den Fall b) angewandt, da der SVD-Algorithmus für die Optimierung bezüglich der Reflexion die Leistungen variiert. Für die Vergleichbarkeit zwischen den verschiedensten Auswertungen wurden alle Optimalsteuerungsauswertungen mit der folgenden CPU und folgendem Arbeitsspeicher durchgeführt:

CPU: Intel<sup>®</sup> Xeon<sup>®</sup> CPU E5-1630 v4 @ 3.70 GHz

RAM: 128 GB

Zudem wurde die Parallel Computing Toolbox von MATLAB verwendet um die Berechnungen zu beschleunigen.

### 5.5.1. Modell I: Zwei Antennen, ein Objekt

Im Folgenden betrachten wir verschiedene Aspekte der Optimalsteuerung bezüglich des ersten Modells in Abschnitt 4.3.1. Dabei ist das zu erwärmende Objekt  $V \subset \Omega$ , welches in Abbildung 4.5 dargestellt wird, gegeben als das Epoxidharz Rapid. Als reduzierte Modelle für die Steuerungsprobleme aus Abschnitt 5.4 wurde die Reduktionsvariante PG-ONJ-MR-EVN sowohl für das volle System (VS) als auch für die getrennten Systeme (GS) betrachtet. Die einzelnen RB-Dimensionen sind:

RB-Variante	RB-Dimension
PG-ONJ-MR-EVN-VS	180
PG-ONJ-MR-EVN-GS-1	91
PG-ONJ-MR-EVN-GS-2	113

Tabelle 7: RB-Dimensionen des vollen und der getrennten Systeme für die Verwendung in der Optimalsteuerung.

Bevor die verschiedenen Aspekte der Optimalsteuerung hinsichtlich der Benchmark-Werte in Definition 5.72 beurteilt werden, wollen wir zunächst die Rechenzeit der einzelnen Bausteine für die Berechnung der Kostenfunktionale betrachten.

RB-Variante	$\mathcal{E}_N$	$\mathcal{P}_{N_L}$	$\mathcal{Z}_1$	$\mathcal{Z}_2$	$\mathcal{Z}_3$
VS	$\sim 0.869$	$\sim 23.761$	$\sim 0.017$	$\sim 0.032$	$\sim 0.005$
GS	$\sim 20.107$	$\sim 24.013$	$\sim 0.024$	$\sim 0.032$	$\sim 0.004$

Tabelle 8: Durchschnittliche Rechenzeit der einzelnen Bausteine für die Berechnung der Kostenfunktionale  $\mathcal{Z}_k$ ,  $k = 1, 2, 3$ , in Sekunden.

RB-Variante	$\nabla \mathcal{E}_N$	$\nabla \mathcal{P}_{N_L}$	$\nabla \mathcal{Z}_1$	$\nabla \mathcal{Z}_2$	$\nabla \mathcal{Z}_3$
VS	$\sim 42.209$	$\sim 200.224$	$\sim 0.050$	$\sim 0.095$	$\sim 0.010$
GS	$\sim 19.825$	$\sim 201.595$	$\sim 0.053$	$\sim 0.086$	$\sim 0.009$

Tabelle 9: Durchschnittliche Rechenzeit der einzelnen Bausteine für die Berechnung der Ableitung der Kostenfunktionale  $\nabla \mathcal{Z}_k$ ,  $k = 1, 2, 3$ , bei Verwendung der Ableitungen  $\nabla \mathcal{E}_N$  und  $\nabla \mathcal{P}_{N_L}$  in Sekunden.

RB-Variante	$\lambda_{1,N}$	$\lambda_{2,N}$	$\lambda_{3,N}^1, \lambda_{3,N_L}^2$	$\nabla \mathcal{Z}_1$	$\nabla \mathcal{Z}_2$	$\nabla \mathcal{Z}_3$
VS	$\sim 0.945$	$\sim 1.828$	$\sim 1.017$	$\sim 0.032$	$\sim 0.038$	$\sim 0.115$
GS	$\sim 1.261$	$\sim 2.380$	$\sim 2.602$	$\sim 0.079$	$\sim 0.113$	$\sim 0.484$

Tabelle 10: Durchschnittliche Rechenzeit der einzelnen Bausteine für die Berechnung der Ableitung der Kostenfunktionale  $\nabla \mathcal{Z}_k$ ,  $k = 1, 2, 3$ , bei Verwendung der adjungierten Lösungen  $\lambda$  in Sekunden.

Zunächst einmal sei darauf hingewiesen, dass die Berechnungen der Zielfunktionale  $\mathcal{Z}_k$  abhängig sind von den Zuständen  $\mathcal{E}_N$  und  $\mathcal{P}_{N_L}$ . Das heißt zu einer Steuerung  $p_S$  muss beispielsweise zunächst  $\mathcal{E}_N[p_S]$  bestimmt werden, um daraufhin

$$\mathcal{Z}_1[p_S] = Z_1(p_S, \mathcal{E}_N[p_S])$$

zu berechnen. Damit ergibt sich dann die kumulative Rechenzeit von

$$0.869 + 0.017 = 0.886$$

aus der Tabelle 8. Analog verhält es sich bei den Tabellen 9 und 10. Bezüglich der Tabelle 10 sei darauf hingewiesen, dass zur Berechnung der Ableitung des Kostenfunktionals  $\nabla \mathcal{Z}_k$  die jeweilige adjungierte Lösung  $\lambda_{k,N}$ , bzw. bei  $k = 3$  die adjungierten Lösungen  $\lambda_{3,N}^1, \lambda_{3,N_L}^2$ , benötigt werden. Vergleichen wir die Rechenzeiten für das volle System und für die getrennten Systeme in Tabelle 8, dann benötigen wir zur Berechnung des Zustandes  $\mathcal{E}_N$  bei den getrennten Systemen deutlich länger. Dies liegt an der Auslastung des Prozessors und der damit verbundenen Einschränkung für eine gute Parallelisierung. Die Rechenzeit zur Bestimmung des Zustandes  $\mathcal{P}_{N_L}$  und der Zielfunktionale weisen dabei keine nennenswerten Unterschiede auf. Betrachten wir die Berechnung der Ableitungen in Tabelle 9, dann ergibt sich eine schnellere Berechnung der Gâteaux-Ableitung  $\nabla \mathcal{E}_N$  bei Verwendung der getrennten Systeme. Der Grund hierfür ist, dass wir durch die Verwendung der getrennten Systeme bereits Teillösungen bestimmt haben, mit denen die Berechnung der Ableitungen in den Leistungen und Phasen schnell durchgeführt werden kann, siehe Korollar 5.20. Dabei entspricht der Unterschied in der Berechnung  $\nabla \mathcal{E}_N$  zwischen VS und GS ungefähr dem Zeitunterschied aus Tabelle 8. Berechnen wir demnach in der Optimierung sowohl den Zielfunktionswert als auch die Ableitung, dann ist die Rechenzeit des vollen Systems vergleichbar mit der der getrennten Systeme. Außerdem können wir aus Tabelle 8 ablesen, dass die Berechnung der Leistungsverteilung  $\mathcal{P}_{N_L}$  zu dem elektrischen Feld  $\mathcal{E}_N$  sehr zeitintensiv ist. Dies ist auch die Begründung in Abschnitt 5.4, weshalb wir die Benchmark-Werte RSDHR und MPHR nicht als Zielfunktionen verwenden wollen. Zudem sehen wir den in Abschnitt 5.4 erwähnten Vorteil

in der Rechenzeit bei Phasenverschiebungen des Zustandes

$$\mathcal{E}_N[\rho(p_S, \theta)] = \mathcal{E}_N[p_S] e^{-i\theta}$$

bei der Betrachtung der Steuerungsprobleme in den Definitionen 5.70 und 5.71, denn die Berechnung der Kostenfunktionale  $\mathcal{Z}_k$ ,  $k = 1, 2$ , beträgt nur wenige Millisekunden. Vergleichen wir nun noch die Art der Ableitungsberechnungen aus den Tabellen 9 und 10, dann können wir festhalten, dass die Berechnung mit der adjungierten Lösung deutlich schneller als die Berechnungen mithilfe der Gâteaux-Ableitungen der Zustände ist. Wenn jedoch eine große Anzahl an Zielfunktionsauswertungen zu verschiedenen Zielwerten  $E_{V,\mathcal{N}}$  betrachtet wird, kann die Verwendung der Gâteaux-Ableitung der Zustände sinnvoll sein. Während für jeden neuen Zielwert  $E_{V,\mathcal{N}}$  eine eigenständige adjungierte Lösung berechnet werden muss, wird bei der Verwendung der Gâteaux-Ableitung der Zustände die Auswertung nur einmal durchgeführt und die Evaluierung der von den Zielwerten  $E_{V,\mathcal{N}}$  abhängigen Zielfunktion  $\nabla \mathcal{Z}_k$  benötigt kaum Rechenzeit. Aufgrund der kürzeren Rechenzeiten, wurden in den folgenden Abschnitten die adjungierten Lösungen zur Berechnung der Gâteaux-Ableitungen verwendet.

### Vergleich lokaler und globaler Optimierungsalgorithmen in MATLAB

Bevor wir uns den Aspekten der einzelnen Steuerungsprobleme in Abschnitt 5.4 und der Vergleichsalgorithmen widmen, wollen wir zunächst die von MATLAB bereitgestellten Optimierungsalgorithmen miteinander vergleichen. Da die Kostenfunktionale zu den Steuerungsproblemen in Abschnitt 5.4 Abbildungen von  $\mathbb{R}^{1+2K_A}$  nach  $\mathbb{R}$  darstellen, verwenden wir die in der Optimization Toolbox und Global Optimization Toolbox implementierten Algorithmen. Als lokale Optimierungsalgorithmen für die Funktion `fmincon` können wir innerhalb von MATLAB die folgenden Algorithmen wählen:

- Interior-Point (IPO),
- Trust-Region-Reflective (TRR),
- SQP (SQP),
- SQP-Legacy (SPL),
- Active-Set (ASE).

Als globale Optimierungsalgorithmen können wir innerhalb von MATLAB die folgenden Algorithmen wählen:

- Pattern-Search (PSE),
- Genetic (GEN),

- Particle-Swarm (PAS),
- Surrogate-Optimization (SUO),
- Simulated-Annealing (SAN).

Die Akronyme der Algorithmen in den Klammern werden bei bestimmten nachfolgenden Abbildungen verwendet. Die ausführliche Beschreibung der lokalen Algorithmen ist in [57, 58] gegeben und die Beschreibung der globalen Algorithmen ist in [59, 60, 61, 62] zu finden. Eine genauere Beschreibung der Algorithmen kann zudem auch in der MATLAB Dokumentation<sup>9</sup> nachgelesen werden. Um die Algorithmen miteinander zu vergleichen, wurden die iterativen Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und jeweils für die beiden Fälle  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , und  $P_{in,m} = 1$ ,  $m = 1, 2$ , betrachtet. Bis auf die Verwendung der Gradienten in den lokalen Optimierungsvarianten wurden die verwendeten Optimierungsalgorithmen mit den `default`-Einstellung ausgeführt.

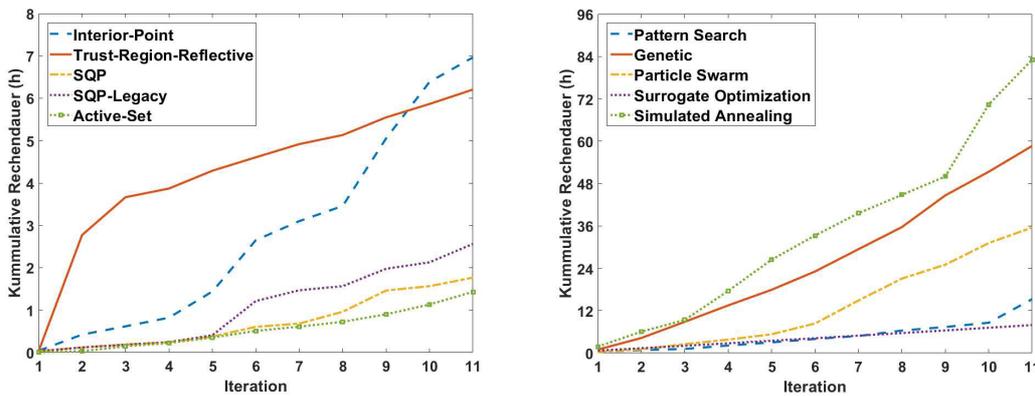


Abbildung 5.2: Kumulativ-iterative Rechenzeit der lokalen und globalen Optimierungsalgorithmen unter Verwendung der Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ .

Es wird darauf hingewiesen, dass für die beiden Abbildungen in Abbildung 5.2 aus Darstellungsgründen jeweils eine unterschiedliche Zeit-Skalierung gewählt wurde. Vergleichen wir zunächst die Rechenzeit zwischen lokalen und globalen Algorithmen, dann erkennen wir sofort, dass die globalen Algorithmen eine deutlich größere Rechenzeit aufweisen. Lediglich der Surrogate-Optimization-Algorithmus benötigt für elf Iterationen die annähernd gleiche Rechenzeit wie der Interior-Point- und der Trust-Region-Reflective-Algorithmus. Die geringere Rechenzeit der lokalen Algorithmen basiert dabei auf der Information der Ableitungen, wodurch deutlich weniger Zielfunktionsauswertungen  $\mathcal{Z}_1$

<sup>9</sup><https://de.mathworks.com/help/matlab/>

und Berechnungen des elektromagnetischen Feldes  $\mathcal{E}_N$  durchgeführt werden müssen. Auch innerhalb der einzelnen Abbildungen weisen die unterschiedlichen lokalen und globalen Algorithmen deutlichere Zeitunterschiede auf. Bei den globalen Optimierungsalgorithmen sind die Rechenzeit des Genetic- und des Simulated-Annealing-Algorithmus bis zur dritten Iteration recht ähnlich, wobei der Simulated-Annealing-Algorithmus tendenziell längere Rechenzeiten besitzt. Auch die beiden globalen Algorithmen Surrogate-Optimization und Pattern-Search weisen eine annähernd gleiche Rechenzeit auf. Betrachten wir nun die Rechenzeit der lokalen Algorithmen, dann sind sich der Interior-Point- und der Trust-Region-Reflective-Algorithmus hinsichtlich der Rechenzeiten ähnlich, denn in bestimmten Iterationen ist eine hohe Rechenzeit zu verzeichnen und in anderen wiederum ergeben sich auch kürzere Rechenzeiten. Die anderen drei lokalen Algorithmen besitzen ebenfalls wieder annähernd gleiche Rechenzeiten, wobei diese deutlich geringer sind als die Rechenzeiten des Interior-Point- und des Trust-Region-Reflective-Algorithmus. Die nachfolgenden Abbildungen geben die Ergebnisse der lokalen und globalen Algorithmen bezüglich der Benchmark-Werte RSDHR Summe und MP Summe nach Definition 5.72 an, wobei für die Vergleichbarkeit zwischen den unterschiedlichen Modellen, welche noch später diskutiert werden, und deren Anzahl an Antennen der MP-Wert pro Antenne betrachtet wird. Hierbei wird noch einmal darauf hingewiesen, dass der MP-Wert die physikalische Einheit  $\left[\frac{W}{m^3}\right]$  besitzt. Bei Betrachtung der kumulativ-iterativen Ergebnisse der lokalen Optimierungsalgorithmen in Abbildung 5.3, lässt sich der iterative Prozess sowohl bei dem RSDHR Summe Wert als auch bei dem MP Summe Wert sehr gut erkennen. Wir sehen dabei, dass die Gleichmäßigkeit bei Überlagerung der Leistungsmuster mit zunehmender Iteration tendenziell besser wird. Auffällig hierbei sind die Ergebnisse des Active-Set-Algorithmus der ab der vierten Iteration kaum noch Energie einbringt. Dies führt auch zu den geringeren Rechenzeiten im Vergleich zum SQP- oder zum SQP-Legacy-Algorithmus. Bei höherem Energieeintrag ist die Auswirkung bei Veränderungen der Leistungsmuster  $\mathcal{P}_{Q, \mathcal{N}_L}[p_S]$  durch eine Veränderung des Steuerparametervektors  $p_S$  deutlich größer, wodurch wiederum das Optimierungsproblem schwieriger zu lösen ist.

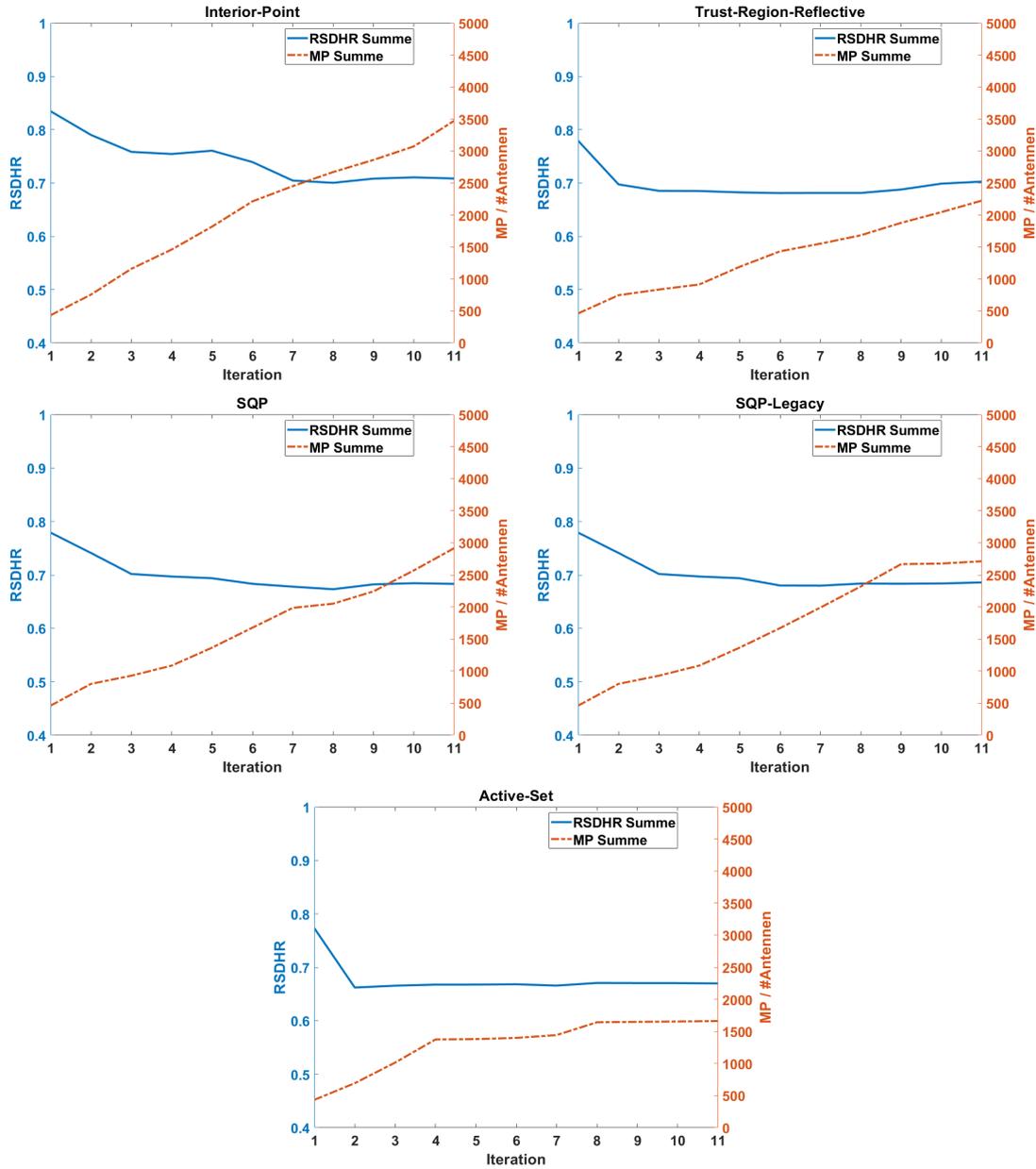


Abbildung 5.3: Kumulativ-iterative RSDHR- und MP-Werte der lokalen Optimierungsalgorithmen unter Verwendung der Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ .

## 5. Optimalsteuerung elektromagnetischer Felder

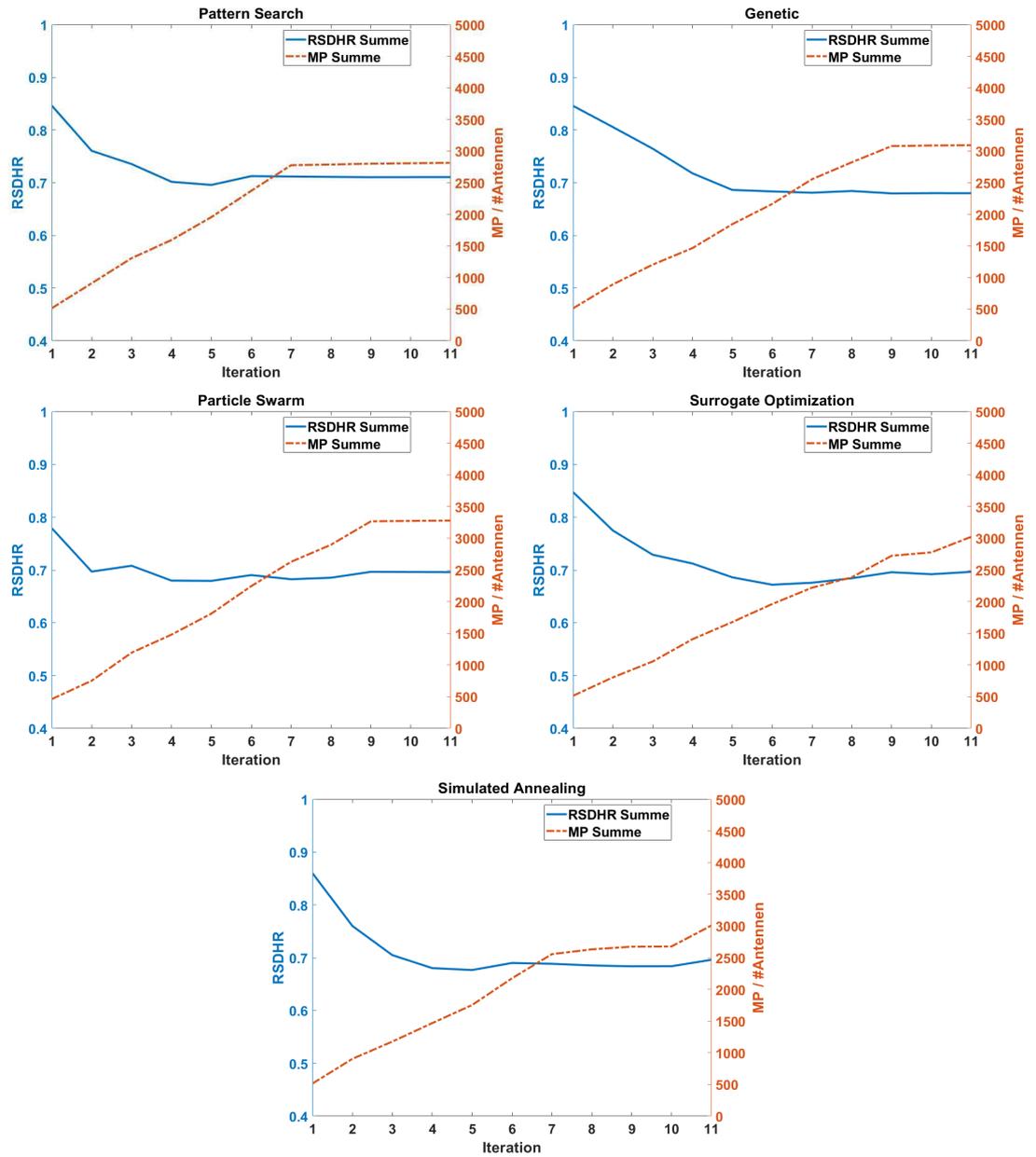


Abbildung 5.4: Kumulativ-iterative RSDHR- und MP-Werte der globalen Optimierungsalgorithmen unter Verwendung der Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ .

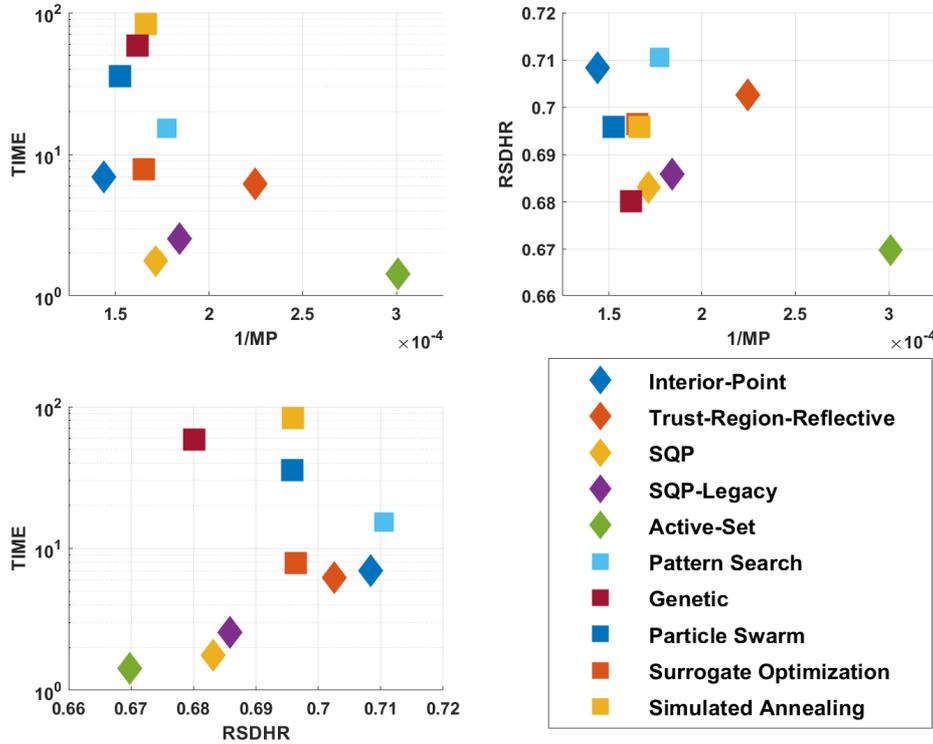


Abbildung 5.5: Kumulativ-iterative Vergleichswerte der lokalen und globalen Optimierungsalgorithmen der elften Iteration für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ .

Auch bei der Betrachtung der globalen Ergebnisse in Abbildung 5.4 sehen wir den Effekt der iterativen Verbesserung bei Überlagerung mehrerer Leistungsmuster. Dabei können wir zudem gut erkennen, dass nach einer anfänglichen Verbesserung des RSDHR Summe-Wertes sich eine Art Grenzwert einstellt. Diesen Effekt und seine Gründe werden wir innerhalb dieses Abschnittes 5.5 immer wieder aufgreifen und genauer analysieren. Bei der vergleichenden Betrachtung der beiden Abbildungen 5.3 und 5.4 lässt sich zudem kein Vorteil in der Gleichmäßigkeit oder dem Energieeintrag bei Verwendung globaler Algorithmen erkennen. Neben den kumulativ-iterativen Werten zu jeder Iteration geben wir noch die kumulativ-iterativen RSDHR Summe und MP Summe Werte zu der elften Iteration in einer vergleichenden Darstellung in Abbildung 5.5 an. In den Subgrafiken wurde anstatt dem direkten MP Summe Wert sein Inverses verwendet. Dies hat den Vorteil, dass bei Betrachtung der einzelnen Subgrafiken jener Algorithmus am besten abgeschnitten hat dessen Werte am weitesten links bezüglich der x-Achse und am wei-

testen unten bezüglich der  $y$ -Achse liegen. Vergleichen wir die Benchmark-Werte der lokalen Optimierungsalgorithmen und deren Rechenzeit im Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , dann liefert der SQP-Algorithmus die besten Ergebnisse. Die kurze Rechenzeit und der relativ gute RSDHR-Wert beim Active-Set-Algorithmus entsteht dadurch, dass ab der vierten Iteration kaum Energie eingespeist wird. Da dieser also kaum Energie einbringt, wird dieser für den Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq K_A$ , nicht weiter betrachtet.

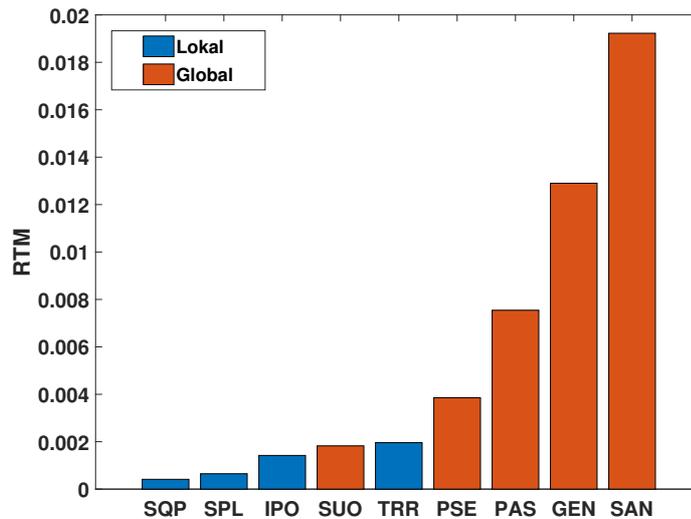


Abbildung 5.6: RTM-Werte der lokalen und globalen Optimierungsalgorithmen für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ .

Für den Vergleich der verschiedenen lokalen und globalen Algorithmen betrachten wir in Abbildung 5.6 für die kumulativ-iterativen Werte der elften Iteration noch den kombinierten Vergleichswert

$$RTM = \frac{RSDHR \text{ Summe} \cdot TIME}{MP \text{ Summe}},$$

wobei TIME die benötigte Rechenzeit der Optimierungsalgorithmen zur Bestimmung von  $K_{opt} = 11$  optimalen Steuerungen in Stunden darstellt. Werden die beiden Algorithmen SQP und SQP-Legacy mit den anderen beiden lokalen Optimierungsalgorithmen verglichen, dann kann festgehalten werden, dass die SQP-Algorithmen bei kürzerer Rechenzeit bessere Gleichmäßigkeit liefern und im Fall des Trust-Region-Reflective-Algorithmus auch mehr Leistung einbringen. Vergleichen wir zudem die Ergebnisse in Abbildung 5.4 für die globalen Optimierungsalgorithmen, dann erhalten wir zum Teil

Unterschiede in den Verläufen. Dies bedeutet wiederum, dass die globalen Algorithmen in der ersten Iteration aufgrund ihrer Rechenvorschriften unterschiedliche globale Optima errechnen. Die Suche nach einem globalen Optimum erweist sich demnach als sehr schwierig, was zum einen auf den großen Suchraum  $\mathbb{P}_S$  zurückzuführen ist und zum anderen mit der Komplexität der Lösungsmannigfaltigkeit zusammenhängt. Vergleichen wir die kumulativ-iterativen Werte in Abbildung 5.5, dann ergeben sich relativ vergleichbare Leistungswerte MP Summe und Gleichmäßigkeitswerte RSDHR Summe, weswegen lediglich die Rechenzeit als Auswahlkriterium verwendet werden kann. Die schnellste Rechenzeit aller Optimierungsalgorithmen liefert hierbei der Surrogate-Optimization-Algorithmus. Aufgrund der guten Benchmark-Werte und der vergleichsweise kurzen Rechenzeit verwenden wir ab sofort den SQP-Algorithmus als Algorithmus für den Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq K_A$ . Wir widmen uns nun dem Benchmark der lokalen und global Optimierungsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ .

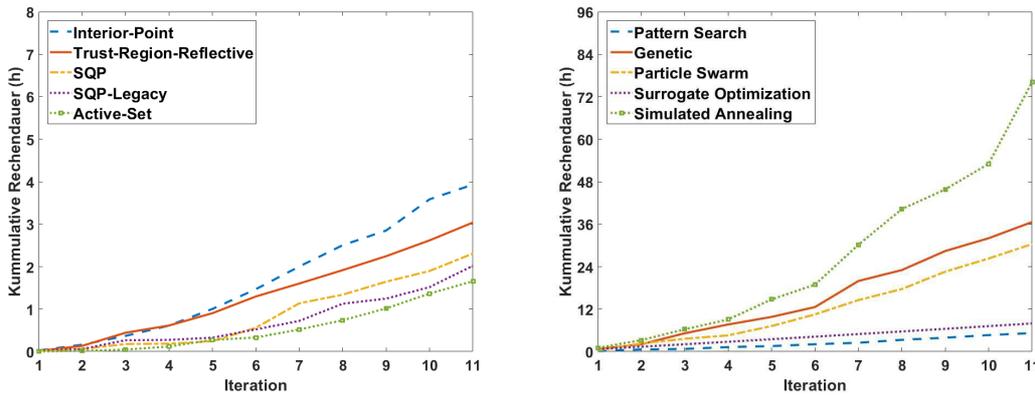


Abbildung 5.7: Kumulativ-iterative Rechenzeit der lokalen und globalen Optimierungsalgorithmen unter Verwendung der Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ .

## 5. Optimalsteuerung elektromagnetischer Felder

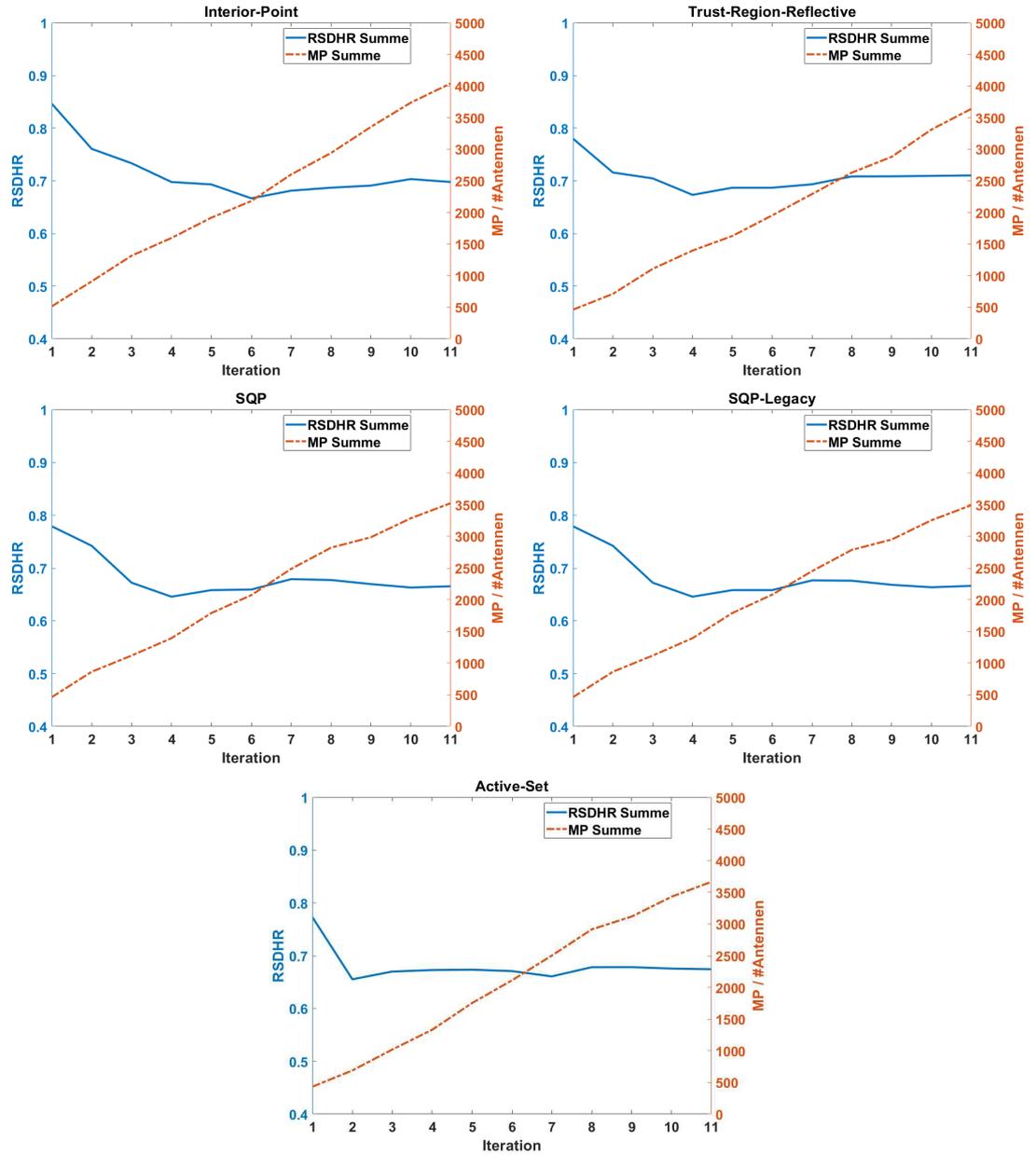


Abbildung 5.8: Kumulativ-iterative RSDHR- und MP-Werte der lokalen Optimierungsalgorithmen unter Verwendung der Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ .

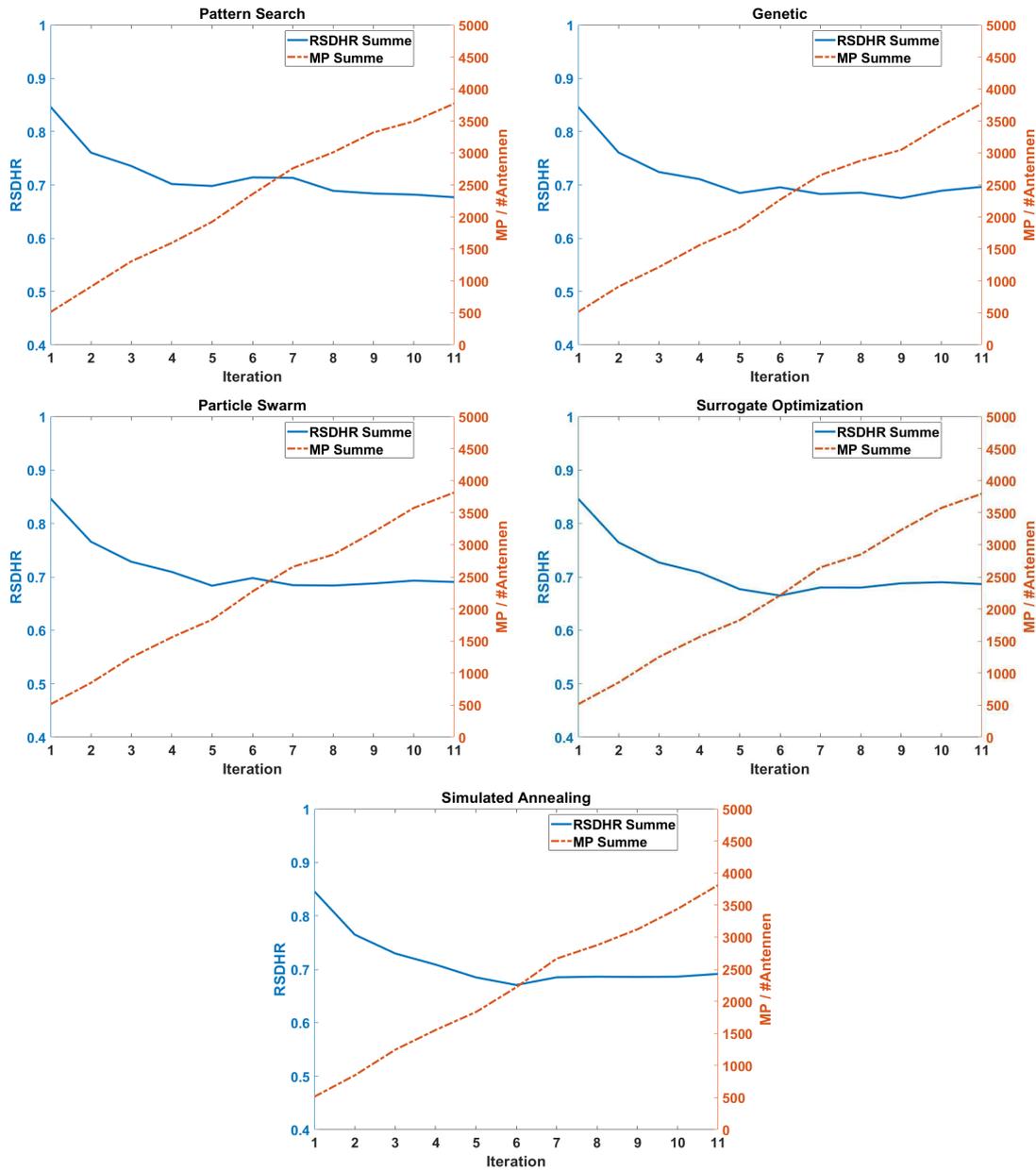


Abbildung 5.9: Kumulativ-iterative RSDHR- und MP-Werte der globalen Optimierungsalgorithmen unter Verwendung der Probleme  $[I_3]_1^l$  für  $K_{opt} = 11$  und für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ .

## 5. Optimalsteuerung elektromagnetischer Felder

Vergleichen wir zuerst die Rechenzeit in Abbildung 5.7 mit denen in Abbildung 5.2, dann ergibt sich zum Teil eine sehr deutliche Verkürzung der Rechenzeit bei den einzelnen Algorithmen. Wie bereits zur Beschreibung von Abbildung 5.2 erklärt wurde, gilt auch für den Fall  $P_{in,m} = 1, m = 1, 2$ , dass bei den globalen Algorithmen der Simulated-Annealing- und bei den lokalen Algorithmen der Interior-Point-Algorithmus die jeweiligen längsten Rechenzeiten vorweisen. Auch die weitere Rangfolge bleibt annähernd gleich. Betrachten wir nun die Ergebnisse der kumulativ-iterativen Benchmark-Werte RSDHR Summe und MP Summe in den Abbildungen 5.8 und 5.9, so ergibt sich wie zuvor für jeden betrachteten Algorithmus eine Verbesserung in der Gleichmäßigkeit. Im Gegensatz zum ersten Fall  $P_{in,m} \in [0, 1], m = 1, 2$ , erhalten wir hier jedoch deutlich größere Leistungseinträge MP Summe, was jedoch hinsichtlich der Wahl  $P_{in,m} = 1, m = 1, 2$ , verständlich ist.

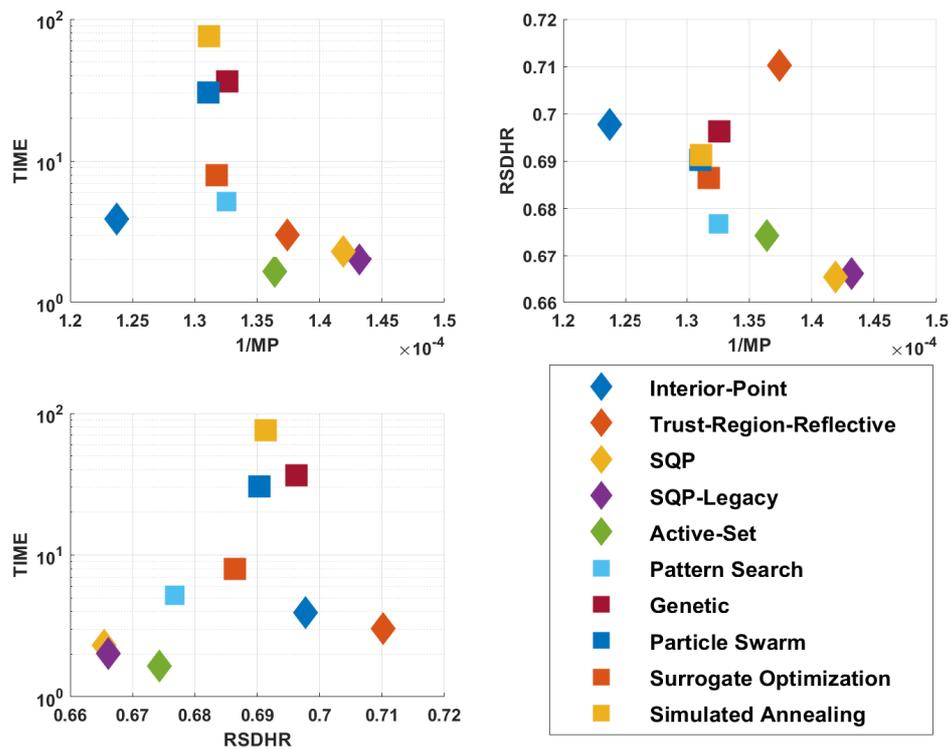


Abbildung 5.10: Kumulativ-iterative Vergleichswerte der lokalen und globalen Optimierungsalgorithmen der elften Iteration für den Fall  $P_{in,m} = 1, m = 1, 2$ .

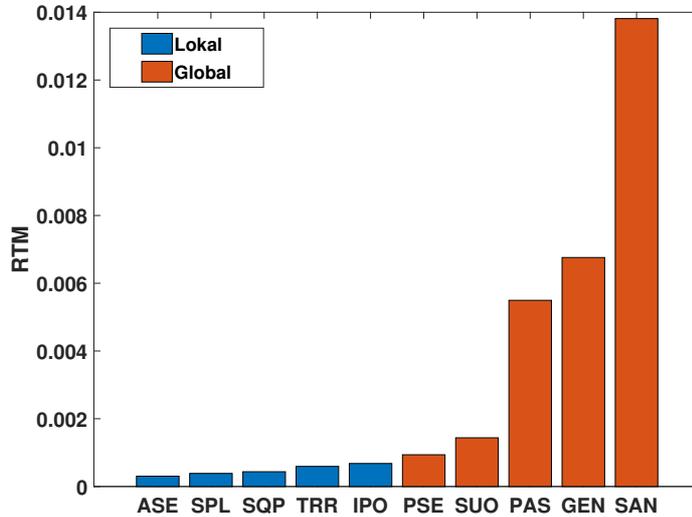


Abbildung 5.11: RTM-Werte der lokalen und globalen Optimierungsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ .

In Abbildung 5.10 geben wir wieder die zuvor diskutierten kumulativ-iterativen Auswertungen zur elften Iteration auch für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , an. Im Gegensatz zu Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , kann für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , auch der Active-Set-Algorithmus für den Vergleich herangezogen werden, da er im Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , ebenso viel Energie einbringt wie die anderen Algorithmen. Vergleichen wir die globalen Algorithmen für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , dann liefern diese wie bereits im Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , annähernd vergleichbare Benchmark-Werte. Des Weiteren besitzen die globalen Algorithmen eine deutlich höhere Rechenzeit als die lokalen Algorithmen. Für die Wahl des Algorithmus für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , betrachten wir wiederum die RTM-Werte in Abbildung 5.11. Nach Betrachtung der RTM-Werte und der Ergebnisse in Abbildung 5.10 liefert der Active-Set die besten Ergebnisse in kürzester Zeit, aber auch die SQP-Algorithmen liefern bei geringfügig längerer Rechenzeit gute Ergebnisse, weswegen diese ebenso gewählt werden können. Um die beiden Fälle  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , und  $P_{in,m} = 1$ ,  $m = 1, 2$ , vergleichen zu können, verwenden wir auch im Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , den SQP-Algorithmus. Eine genaue Analyse der Verläufe wie in den Abbildungen 5.8 und 5.9 führen wir nun in den nachfolgenden Abschnitten durch.

Vergleich der Steuerungsprobleme bezüglich Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$

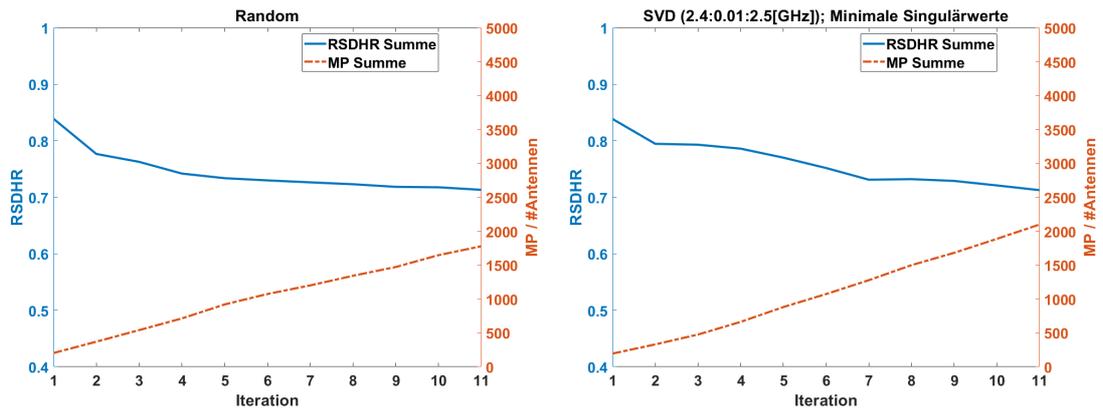


Abbildung 5.12: Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . (Modell I)

Sowohl die RSDHR-Kurve des Random-Algorithmus, als auch die des SVD-Algorithmus, in Abbildung 5.12 fällt stetig ab und führt somit zu einer stetigen Verbesserung der Gleichmäßigkeit. Des Weiteren lässt sich bei der Betrachtung festhalten, dass der SVD-Algorithmus trotz eines leicht größeren Leistungseintrags denselben Gleichmäßigkeitswert liefert und damit nur geringfügig besser als der Random-Algorithmus ist. Neben den Benchmark-Werten des SVD-Algorithmus geben wir in Abbildung 5.13 noch eine grafische Darstellung der Werte der S-Parametermatrix des ersten Modells bezüglich der in der Hochfrequenz weit verbreiteten dB-Darstellung (4.92) an. Aus der Darstellung der  $S_{11}$ - und  $S_{22}$ -Werte in Abbildung 5.13 können wir sehen, dass eine sehr gute Einspeisung der Energie bei ungefähr 2.465 GHz vorliegt und zudem im höheren Frequenzbereich tendentiell mehr Energie eingespeist werden kann. Des Weiteren können wir aus den Darstellungen der Transmissionswerte  $S_{12}$  und  $S_{21}$  ableiten, dass im Bereich von 2.42 bis 2.44 GHz eine erhöhte Verkopplung der beiden Antennen auftritt. Weitere Interpretationen der in Abbildung 5.13 dargestellten S-Parameter können in der Fachliteratur der Hochfrequenztechnik, wie [37], nachgelesen werden.

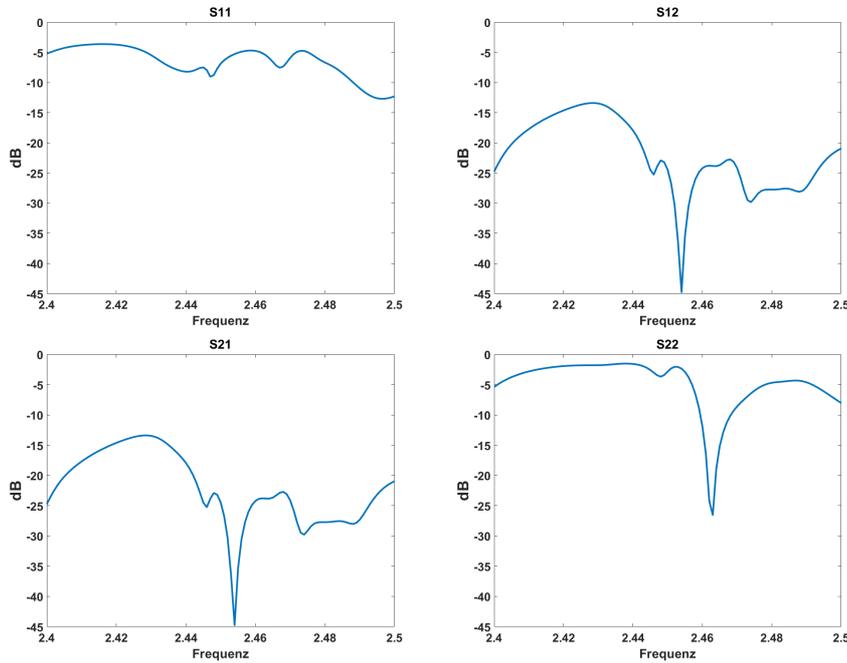


Abbildung 5.13: Darstellung der S-Parametermatrix in dB. (Modell I)

Wir widmen uns nachfolgend den Ergebnissen der optimalen Steuerungsfunktionen aus Abschnitt 5.4 für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . Neben den kumulativ-iterativen Ergebnissen für die Benchmark-Werte, wie beispielsweise in Abbildung 5.14, geben wir in den Ergebnistabellen für die verschiedenen Steuerungsfunktionen, wie Tabelle 11, auch die Anzahl der optimalen Steuerungen und die dafür benötigte Zeit an. Für die Subgrafiken der Abbildungen wie beispielsweise Abbildung 5.14 gilt, wie bereits im vorherigen Abschnitt erklärt, je näher der Marker am Ursprung, bzw. der linken unteren Ecke, ist, desto besser ist das Ergebnis. In Abbildung 5.14 wurde auf die Darstellung der Steuerungsprobleme  $[I_1]_k$  und  $[I_2]_k$ ,  $k = 1, 2, 3$ , verzichtet, denn nach Betrachtung der Ergebnisse in den Tabellen 11 und 12 werden nur wenige optimale Steuerungen gefunden und die Steuerungsprobleme eignen sich demnach nicht für die weitere Betrachtung im Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . Die Terminierung des iterativen Optimierungsprozesses erfolgt dabei durch das Abbruchkriterium bezüglich der optimalen Leistungsparameter  $P_{in,l}^{opt}$ ,  $1 \leq l \leq K_A$ , falls diese nahe an Null sind. Sind die Leistungsparameter  $P_{in,l}^{opt}$  klein, dann ist die Feldlösung nahe an der Null-Lösung und es wird kaum Energie eingebracht. Dies führt wiederum zu den annähernd vergleichbaren optimalen Steuerparameter im nächsten Iterationsschritt. Lassen wir demnach die Optimierung laufen, dann stellen wir damit unnötige Berechnungen an, bei denen keine Verbesserung zu erwarten ist.

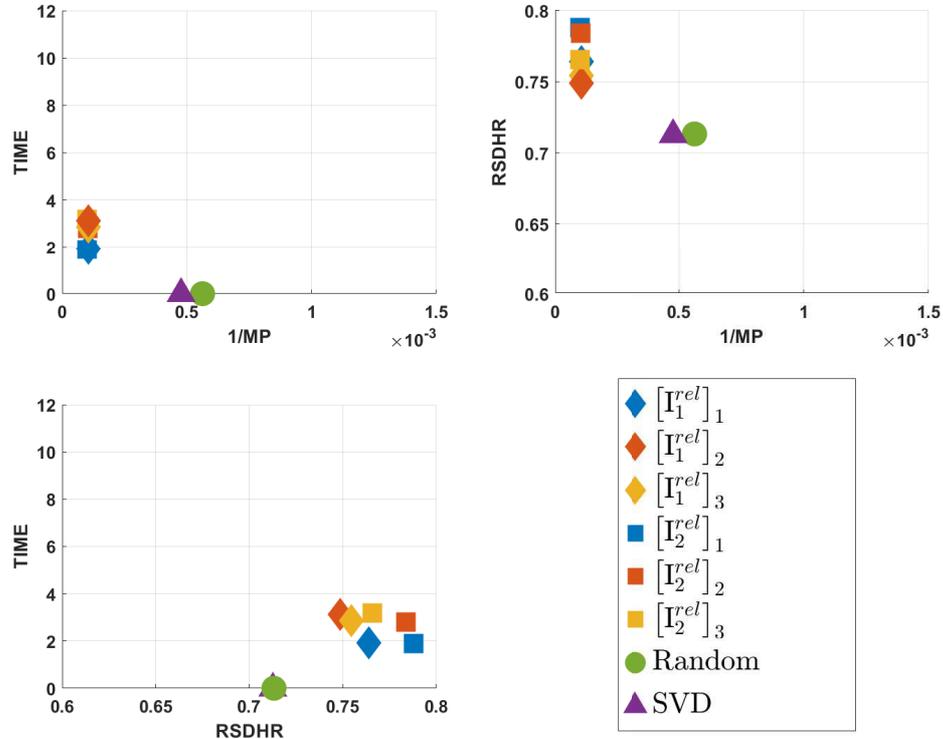


Abbildung 5.14: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_1^{rel}]_k, [I_2^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_1]_1$	1	$\sim 1.02$ min	0.8479	$1.7325 \cdot 10^{-13}$
$[I_1]_2$	1	$\sim 1.43$ min	0.7100	$1.4163 \cdot 10^{-13}$
$[I_1]_3$	11	$\sim 16.11$ h	0.7543	$6.4179 \cdot 10^3$
$[I_1^{rel}]_1$	11	$\sim 1.92$ h	0.7641	$9.5393 \cdot 10^3$
$[I_1^{rel}]_2$	11	$\sim 3.12$ h	0.7487	$9.4822 \cdot 10^3$
$[I_1^{rel}]_3$	11	$\sim 2.85$ h	0.7545	$9.4268 \cdot 10^3$

Tabelle 11: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_1]$  für den Fall  $P_{in,m} \in [0, 1], m = 1, 2$ . (Modell I)

SP	# $p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_2]_1$	1	$\sim 0.98$ min	0.8479	$1.7325 \cdot 10^{-13}$
$[I_2]_2$	1	$\sim 1.37$ min	0.7100	$1.4163 \cdot 10^{-13}$
$[I_2]_3$	4	$\sim 1.53$ h	0.7524	$6.1900 \cdot 10^2$
$[I_2^{rel}]_1$	11	$\sim 1.88$ h	0.7880	$9.9278 \cdot 10^3$
$[I_2^{rel}]_2$	11	$\sim 2.81$ h	0.7839	$9.8550 \cdot 10^3$
$[I_2^{rel}]_3$	11	$\sim 3.17$ h	0.7656	$9.9277 \cdot 10^3$

Tabelle 12: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_2]$  für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . (Modell I)

Betrachten wir im Gegenzug die relativen optimalen Steuerungsprobleme  $[I_1^{rel}]_k$  und  $[I_2^{rel}]_k$ ,  $k = 1, 2, 3$ , in Abbildung 5.14, dann bringen diese mehr Leistung als die Vergleichsalgorithmen ein, aber generieren schlechte Gleichmäßigkeitswerte. Dieser Effekt und das Finden von elf optimalen Steuerungen kann dabei direkt mit der relativen Betrachtung korreliert werden. Durch den zusätzlichen Divisor  $\mathcal{Z}_k(p_S)$  mit  $E_V = 0$  wird nach Steuerungen mit hohem Energieeintrag gesucht.

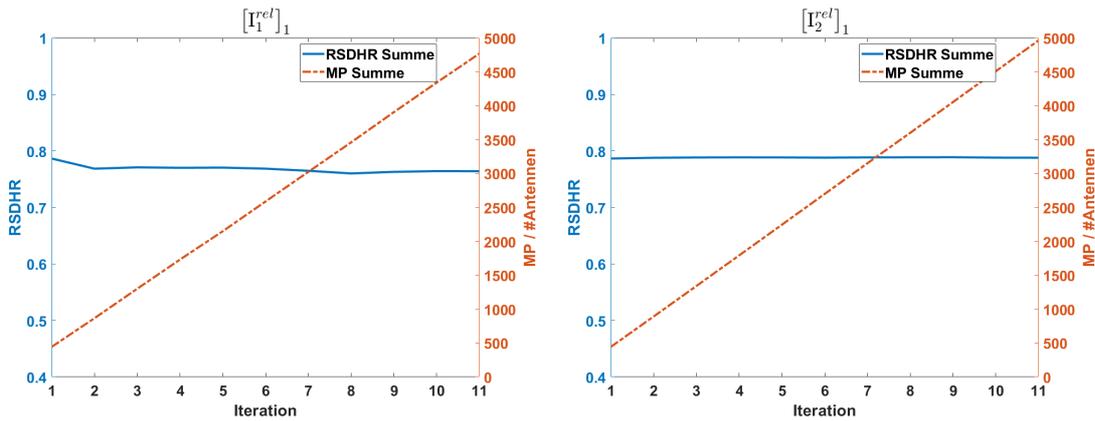


Abbildung 5.15: Kumulativ-iterative RSDHR- und MP-Werte der Steuerungsprobleme  $[I_1^{rel}]_1$  und  $[I_2^{rel}]_1$  für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . (Modell I)

5. Optimalsteuerung elektromagnetischer Felder

SP	$f$	$P_{in,1}$	$P_{in,2}$	$\theta_1$	$\theta_2$
$[I_1^{rel}]_1^1$	2.5000	1	1	6.1384	3.5691
$[I_1^{rel}]_1^2$	2.5000	1	1	0.2920	3.0382
$[I_1^{rel}]_1^3$	2.5000	1	1	5.8496	4.0119
$[I_1^{rel}]_1^4$	2.5000	1	1	0.0138	3.2436
$[I_1^{rel}]_1^5$	2.4670	1	1	5.0940	1.6604
$[I_1^{rel}]_1^6$	2.4671	1	1	4.7160	1.9132
$[I_1^{rel}]_1^7$	2.4675	1	1	4.9400	1.7855
$[I_1^{rel}]_1^8$	2.4670	1	1	4.3452	2.0853
$[I_1^{rel}]_1^9$	2.4651	1	1	4.6890	1.5292
$[I_1^{rel}]_1^{10}$	2.4662	1	1	5.2889	1.2254
$[I_1^{rel}]_1^{11}$	2.4650	1	1	5.4962	0.9258

Tabelle 13: Optimale Steuerparameter des iterativen Steuerungsproblems (SP)  $[I_1^{rel}]_1^l$ ,  $1 \leq l \leq 11$ , für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . (Modell I)

SP	$f$	$P_{in,1}$	$P_{in,2}$	$\theta_1$	$\theta_2$
$[I_2^{rel}]_1^1$	2.5000	1	1	6.1384	3.5691
$[I_2^{rel}]_1^2$	2.5000	1	1	0.0445	3.4824
$[I_2^{rel}]_1^3$	2.5000	1	1	6.1011	3.7503
$[I_2^{rel}]_1^4$	2.5000	1	1	6.1247	3.8582
$[I_2^{rel}]_1^5$	2.5000	1	1	6.0416	3.5463
$[I_2^{rel}]_1^6$	2.5000	1	1	6.1135	3.4614
$[I_2^{rel}]_1^7$	2.5000	1	1	5.9458	3.5226
$[I_2^{rel}]_1^8$	2.5000	1	1	6.0739	3.3712
$[I_2^{rel}]_1^9$	2.5000	1	1	6.0845	3.5881
$[I_2^{rel}]_1^{10}$	2.5000	1	1	5.9879	3.5647
$[I_2^{rel}]_1^{11}$	2.5000	1	1	5.9630	3.9984

Tabelle 14: Optimale Steuerparameter des iterativen Steuerungsproblems (SP)  $[I_2^{rel}]_1^l$ ,  $1 \leq l \leq 11$ , für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . (Modell I)

Beziehen wir nun noch die Verläufe der Steuerungsprobleme  $[I_1^{rel}]_1$  und  $[I_2^{rel}]_1$  aus Abbildung 5.15 und die dazugehörigen optimalen Steuerparameter in den Tabellen 13 und 14 bei der Erklärung mit ein, dann werden trotz des Fehlerterms  $\mathcal{R}_{\max}^{1,l}$  in jeder Iteration annähernd vergleichbare Steuerparameter gewählt. Dabei entsprechen die verwendeten optimalen Frequenzen gerade denen aus Abbildung 5.13, bei denen die Antennen viel Energie in die Kavität einspeisen. Die Ursache für die schlechten RSDHR-Werte ist

wie schon bei den Steuerungsproblemen  $[I_1]_k$  und  $[I_2]_k$ ,  $k = 1, 2, 3$ , der dominierende erste Summand im Zähler. Dies liegt wiederum an dem großen Abstand zwischen den Feldlösungen  $\mathcal{E}_{\mathcal{N}}[p_S]$  und  $\mathcal{P}_{Q, \mathcal{N}_L}[p_S]$ ,  $p_S \in \mathbb{P}_S$ , und des Vergleichfeldes  $E_{V, \mathcal{N}}$ , bzw. der Leistungsverteilung  $Q_{V, \mathcal{N}_L}$ , nach Abschnitt 5.4. Aufgrund der eben dargestellten Gründe und den obigen Ergebnissen eignen sich die Steuerungsprobleme  $[I_1]_k$ ,  $[I_2]_k$ ,  $[I_1^{rel}]_k$  und  $[I_2^{rel}]_k$ ,  $k = 1, 2, 3$ , nicht für den Fall  $P_{in, m} \in [0, 1]$ ,  $m = 1, 2$ , und werden bei den noch folgenden numerischen Ergebnissen der anderen Modelle nicht mehr betrachtet. Wir widmen uns nun den Steuerungsproblemen  $[I_3]_k$ ,  $[I_4]_k$ ,  $[I_3^{rel}]_k$  und  $[I_4^{rel}]_k$ ,  $k = 1, 2, 3$ , die ohne vorgegebene Vergleichsfelder  $E_{V, \mathcal{N}}$  und  $Q_{V, \mathcal{N}_L}$  auskommen.

Wir widmen uns zunächst der Betrachtung der Rechenzeiten der verschiedenen Steuerungsprobleme für verschiedene Indizes  $k$ , welcher sich auf die entsprechende Kostenfunktion  $\mathcal{Z}_k$  bezieht. Sowohl in Abbildung 5.16 als auch in Abbildung 5.17 weisen die Steuerungsprobleme für  $k = 1$  kürzere Rechenzeiten auf als für  $k = 2$ . Da zudem die Benchmark-Werte annähernd gleich sind, bringt demnach die Verwendung der skalierten Funktionen  $\mathcal{Z}_2$  keine nennenswerten Vorteile mit sich. Anders verhält es sich hierbei bei dem Vergleich mit  $k = 3$ , denn während die Rechenzeit für  $[I_3]_3$ ,  $[I_3^{rel}]_3$  deutlich höher ausfällt, erhalten wir relativ kurze Rechenzeiten bei  $[I_4]_3$ ,  $[I_4^{rel}]_3$ . Dabei ergibt sich die erhöhte Rechenzeit bei  $[I_3]_3$ ,  $[I_3^{rel}]_3$  aufgrund der Berechnung von  $\mathcal{P}_{Q, \mathcal{N}_L}$ , siehe Tabelle 8. Der gute Zeitwert von  $[I_4]_3$  wiederum liegt an der geringen Anzahl an optimalen Steuerungen. Vergleichen wir die Steuerungsprobleme  $[I_3]_k$ ,  $[I_4]_k$  mit deren relativen Varianten  $[I_3^{rel}]_k$ ,  $[I_4^{rel}]_k$ ,  $k = 1, 2, 3$ , dann speisen die relativen Varianten immer mehr Energie in das zu erwärmende Gebiet ein als ihr Pendant, was wiederum den Gedanken der relativen Betrachtung der Steuerungsprobleme unterstützt. Möchte man demnach mehr Energie einspeisen, dann sollten die relativen Steuerungsprobleme verwendet werden. Eine weitere Gemeinsamkeit der Abbildungen 5.16 und 5.17 sind die beiden Ausreißer  $[I_3]_3$  und  $[I_4]_3$ , die auf den Leistungsverteilungen  $\mathcal{P}_{Q, \mathcal{N}_L}$  basieren. Dabei erhalten wir bei  $k = 3$  eine Verbesserung der Benchmark-Werte, indem wir das relative Pendant zu diesen verwenden.

Zum Abschluss betrachten wir noch die kumulativ-iterativen Darstellungen der Werte RSDHR Summe und MP Summe von  $[I_3]_1$  und den Vergleichsalgorithmen. Wie wir in Abbildung 5.18 sehen können, besitzt das Steuerungsproblem  $[I_3]_1$  neben der besseren Gleichmäßigkeit auch einen höheren Leistungseintrag. Dass dies kein Zufall ist, sondern an der Definition des Steuerungsproblems liegt, ergibt sich bei der Betrachtung der Abbildungen 5.3 und 5.4, die alle vergleichbare Benchmark-Werte generieren.

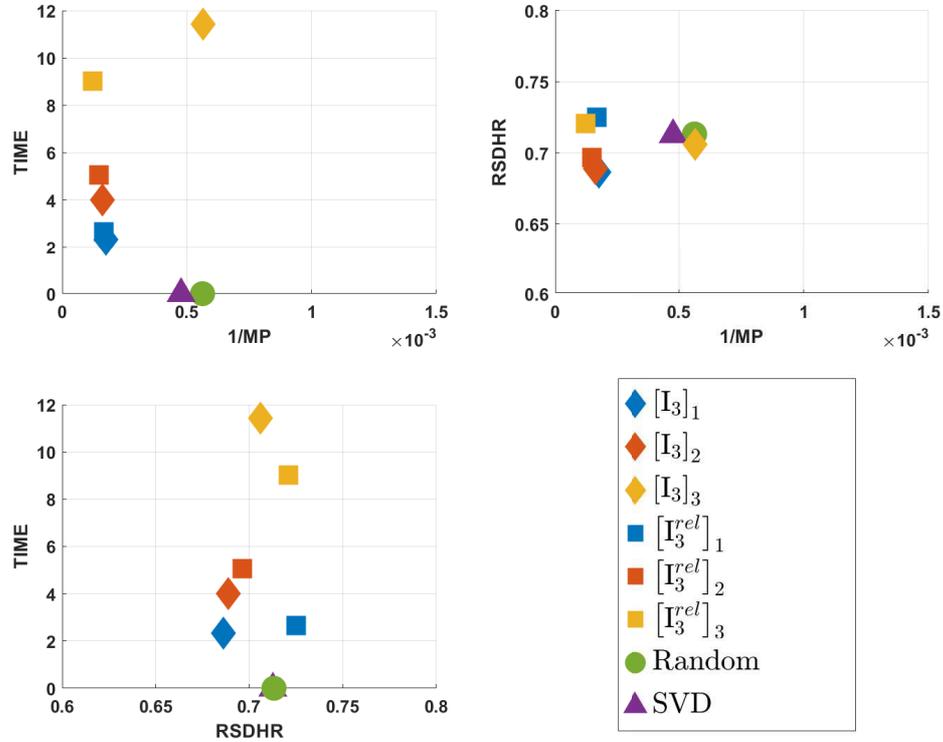


Abbildung 5.16: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_3]_1$	11	$\sim 2.32$ h	0.6862	$5.6897 \cdot 10^3$
$[I_3]_2$	11	$\sim 4.01$ h	0.6888	$6.1572 \cdot 10^3$
$[I_3]_3$	11	$\sim 11.42$ h	0.7058	$1.7735 \cdot 10^3$
$[I_3^{rel}]_1$	11	$\sim 2.65$ h	0.7251	$6.0207 \cdot 10^3$
$[I_3^{rel}]_2$	11	$\sim 5.07$ h	0.6964	$6.7771 \cdot 10^3$
$[I_3^{rel}]_3$	11	$\sim 9.02$ h	0.7208	$8.2372 \cdot 10^3$

Tabelle 15: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_3]$  für den Fall  $P_{in,m} \in [0, 1], m = 1, 2$ . (Modell I)

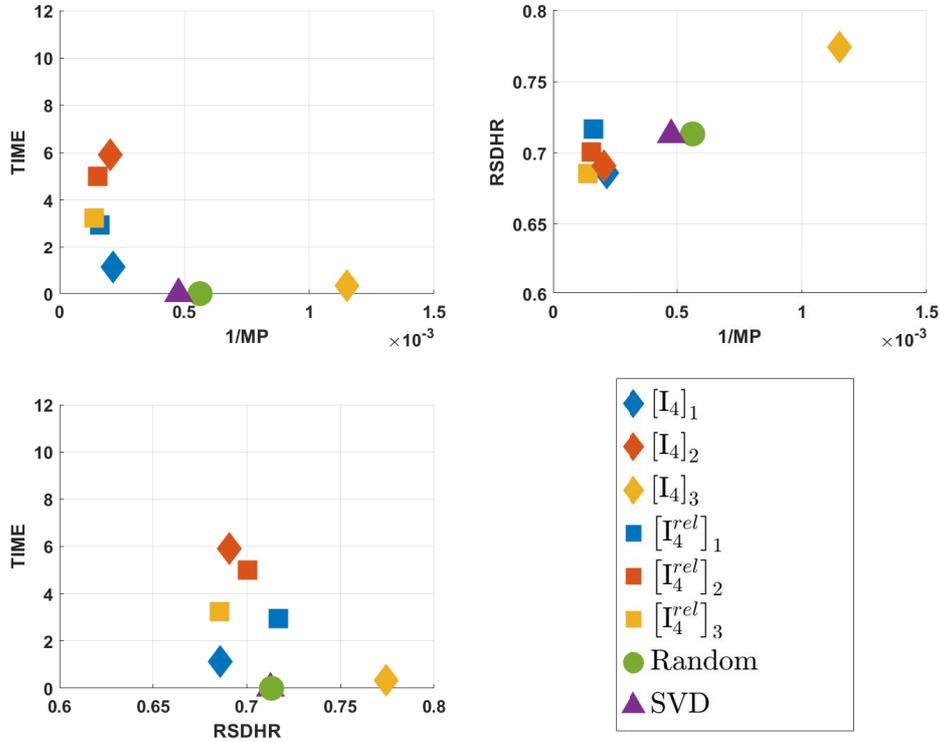


Abbildung 5.17: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_4]_1$	9	$\sim 1.12$ h	0.6859	$4.6786 \cdot 10^3$
$[I_4]_2$	11	$\sim 5.91$ h	0.6906	$4.9242 \cdot 10^3$
$[I_4]_3$	2	$\sim 0.34$ h	0.7743	$8.6782 \cdot 10^2$
$[I_4^{rel}]_1$	11	$\sim 2.94$ h	0.7168	$6.2164 \cdot 10^3$
$[I_4^{rel}]_2$	11	$\sim 4.99$ h	0.7003	$6.5099 \cdot 10^3$
$[I_4^{rel}]_3$	11	$\sim 3.26$ h	0.6854	$7.2562 \cdot 10^3$

Tabelle 16: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} \in [0, 1], m = 1, 2$ . (Modell I)

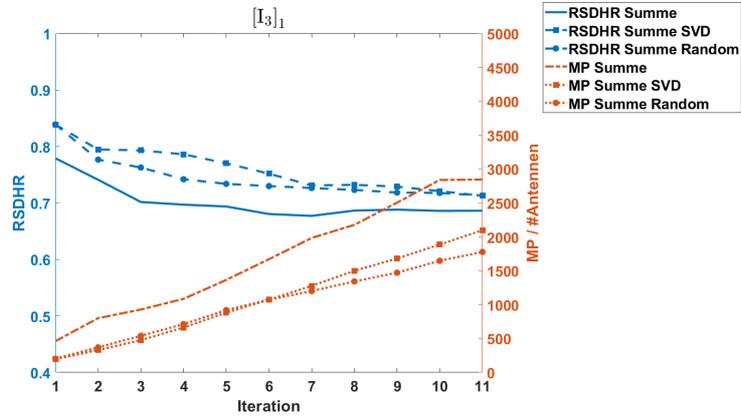


Abbildung 5.18: Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems  $[I_3]_1$  und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ . (Modell I)

**Vergleich der Steuerungsprobleme bezüglich Fall  $P_{in,m} = 1$ ,  $m = 1, 2$**

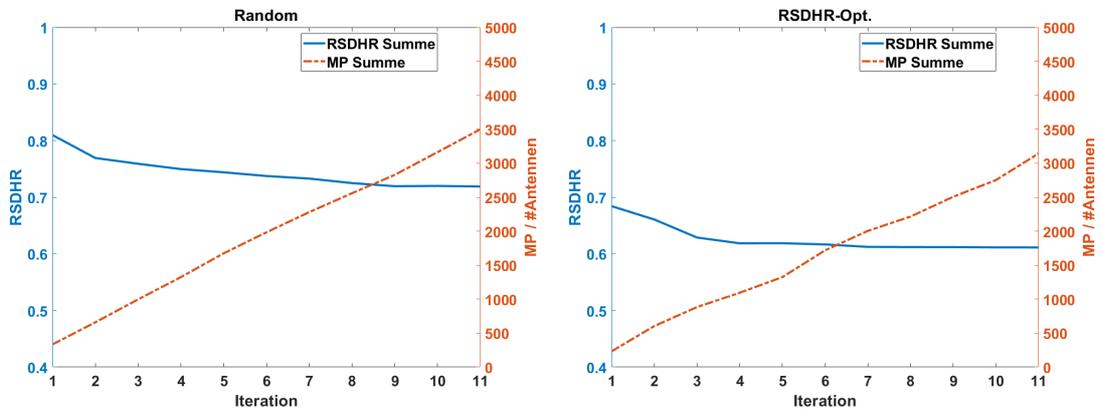


Abbildung 5.19: Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ . (Modell I)

Vergleichen wir die Grafiken der Abbildung 5.19, dann erhalten wir beim RSDHR-Optimierung-Algorithmus eine deutlich bessere Gleichmäßigkeit bei nur geringem Unter-

schied in der gemittelten Leistung. Der RSDHR-Optimierung-Algorithmus ist demnach ein sehr guter Vergleichsalgorithmus, bei dem jedoch die Rechenzeit im realen Anwendungsfall sehr hoch und damit nicht praktikabel ist. Vergrößert sich zudem der Parameterraum  $\mathbb{P}_S$  durch die Verwendung weiterer Antennen, dann ergibt sich ein exorbitanter Anstieg in der Rechenzeit. Bevor wir die Benchmark-Werte RSDHR Summe und MP Summe zu den Optimalsteuerungsproblemen aus Abschnitt 5.4 mit denen der Vergleichsalgorithmen aus Abbildung 5.19 vergleichen, betrachten wir zunächst die Hyperflächen der Benchmark-Werte RSDHR und MP nach Definition 5.69 zu den Steuerparametervektoren  $p_S \in \mathbb{P}_S$  unter Berücksichtigung des Falles  $P_{in,m} = 1, m = 1, 2$ .

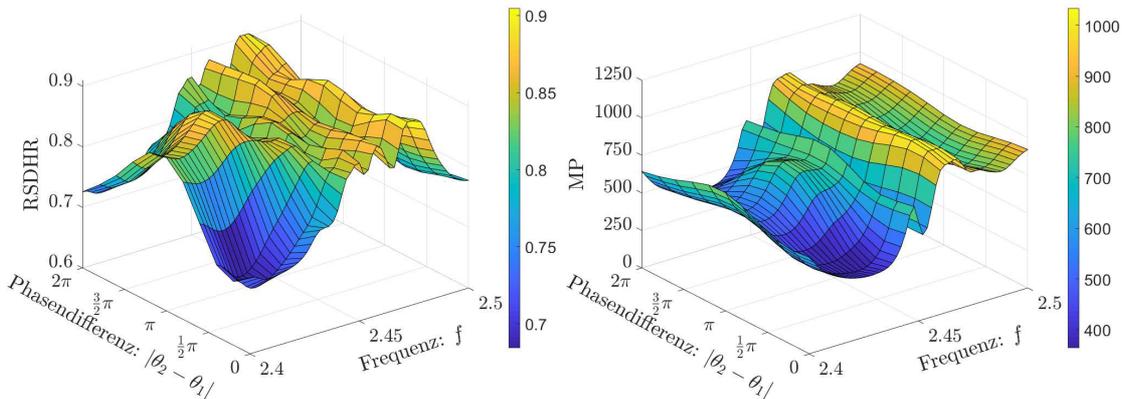


Abbildung 5.20: RSDHR- und MP-Werte des Steuerparametervektors  $p_S \in \mathbb{P}_S$  für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

Vergleichen wir die beiden Hyperflächen der Abbildung 5.20, dann können wir festhalten, dass zum Teil eine annähernd indirekte Korrelation zwischen RSDHR- und MP-Wert existiert. Das heißt, dass bei geringerem Leistungseintrag tendentiell auch die Gleichmäßigkeit besser ist, also der RSDHR-Wert klein ist. Dass dies aber nicht immer der Fall sein muss, sehen wir beispielsweise bei  $f = 2.5$ , denn zum einen wird viel Leistung eingebracht und zum anderen haben wir einen vergleichsmäßig kleinen und damit guten RSDHR-Wert.

## 5. Optimalsteuerung elektromagnetischer Felder

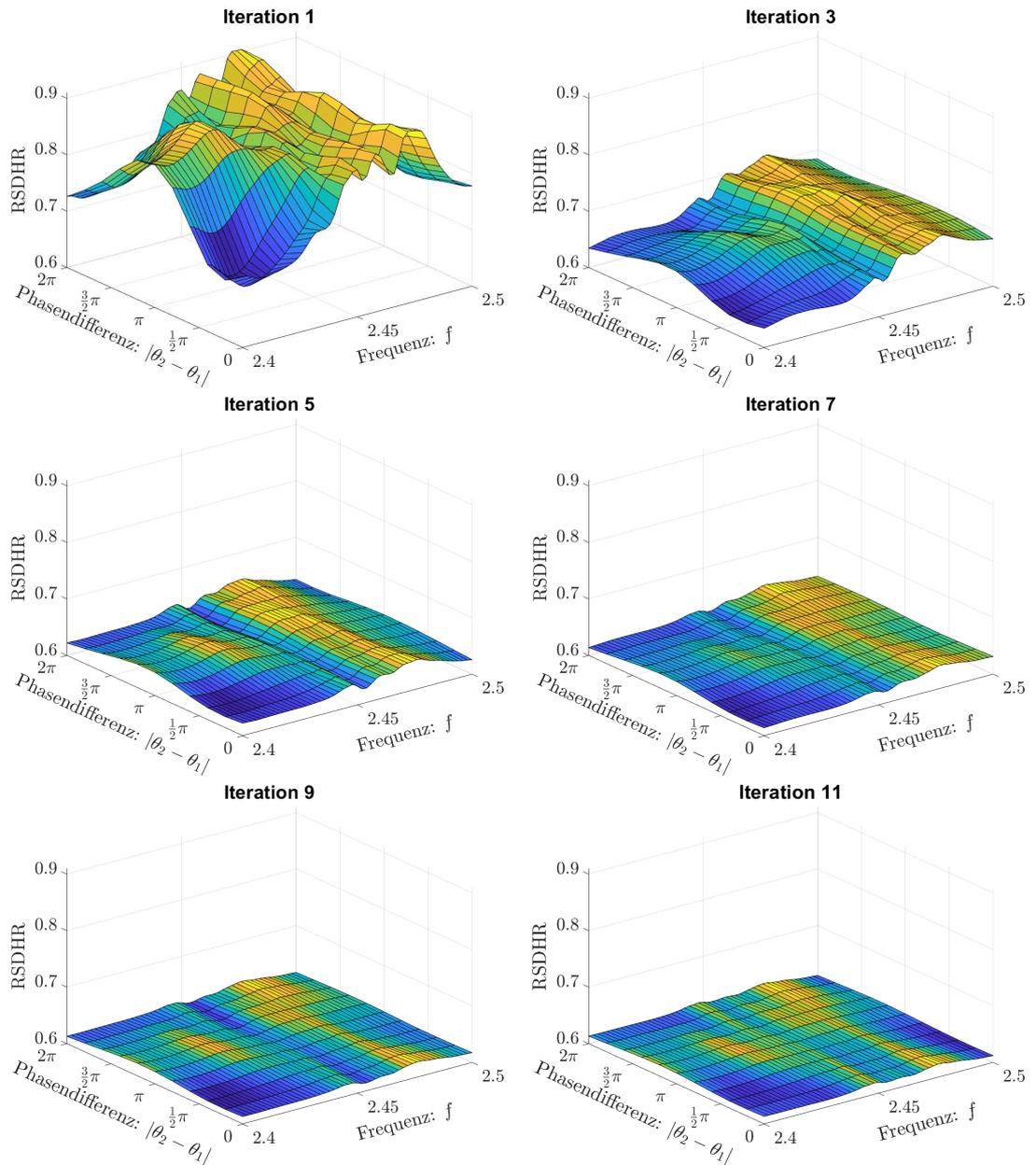


Abbildung 5.21: Hyperfläche des steuerungsabhängigen kumulativ-iterativen RSDHR-Wertes für den Vergleichsalgorithmus RSDHR-Optimierung. (Modell I)

In den Abbildungen 5.19 und 5.21 können wir sehr gut erkennen, dass die Verwendung unterschiedlicher Steuerungen den Gleichmäßigkeitswert RSDHR iterativ verbessert. Außerdem lässt sich aufgrund der immer gleichmäßiger werdenden Hyperfläche in Abbildung 5.21 eine untere Schranke an den RSDHR-Wert feststellen. Diese Schranke ergibt sich aufgrund der beschränkenden Variation von zu überlagernder Felder zweier Antennen und den in der Kavität ausbreitungsfähigen elektromagnetischen Feldern. Des Weiteren hat die Phasendifferenz  $|\theta_2 - \theta_1|$  je nach Frequenz  $f$  großen Einfluss sowohl auf die Gleichmäßigkeit als auch auf den Energieeintrag. Beispielsweise erhalten wir bei einer Frequenz von  $f = 2.42$ , dass in der ersten Iteration bei einer Phasendifferenz von  $\pi$  ein hoher Energieeintrag und schlechte Gleichmäßigkeit erhalten wird. Wählen wir jedoch die Phasendifferenz  $\frac{\pi}{4}$ , dann erhalten wir eine sehr gute Gleichmäßigkeit bei geringerem Energieeintrag. Zudem können wir in der ersten Iteration gut erkennen, dass zum Teil kleine Frequenzänderungen große Auswirkungen auf die Gleichmäßigkeit und den Energieeintrag haben.

Wir widmen uns nun den numerischen Ergebnissen der Steuerungsprobleme  $[I_1]_k$ ,  $[I_2]_k$ ,  $[I_1^{rel}]_k$  und  $[I_2^{rel}]_k$ ,  $k = 1, 2, 3$ , für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , in den Abbildungen 5.22 und 5.23, sowie den Tabellen 17 und 18, und wollen sehen, ob auch nach der Einschränkung des Parameterraumes die Werte weiterhin schlecht bleiben. Zu Beginn fällt auf, dass anders als bei  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , alle Steuerungsprobleme elf optimale Steuerungen generiert haben, was lediglich an der Wahl  $P_{in,m} = 1$ ,  $m = 1, 2$ , liegt. Wie schon im Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , erhalten wir auch für  $P_{in,m} = 1$ ,  $m = 1, 2$ , keine guten RSDHR-Werte. Auch hier liegt dies wiederum am dominierenden ersten Summanden im Zähler, der dafür sorgt, dass der Strafterm nicht ins Gewicht fällt. Damit sind die iterativen Steuerungsprobleme  $[I_1]_k$ ,  $[I_2]_k$ ,  $[I_1^{rel}]_k$  und  $[I_2^{rel}]_k$ ,  $k = 1, 2, 3$ , für die hergeleiteten Vergleichsfelder  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  aus Abschnitt 5.4 auch für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , nicht geeignet. Um dennoch mehr über die Effekte bei den Steuerungsproblemen  $[I_1]_k$ ,  $[I_2]_k$ ,  $[I_1^{rel}]_k$  und  $[I_2^{rel}]_k$ ,  $k = 1, 2, 3$ , zu erfahren, untersuchen wir diese im nächsten Abschnitt für Felder  $E_{V,\mathcal{N}} = \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]$ , bzw.  $Q_{V,\mathcal{N}_L} = \mathcal{P}_{Q,\mathcal{N}_L}[\tilde{p}_S]$ , wobei  $\tilde{p}_S \in \mathbb{P}_S$  eine beliebige Steuerung darstellt.

Wir betrachten als nächstes die Abbildungen 5.24 und 5.25, sowie die Tabellen 19 und 20, zu den Optimalsteuerungsproblemen  $[I_3]_k$ ,  $[I_4]_k$ ,  $[I_3^{rel}]_k$  und  $[I_4^{rel}]_k$ ,  $k = 1, 2, 3$ , die ohne vorgegebene Vergleichsfelder  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  auskommen. Wie bereits im Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , zu sehen war, erkennen wir in den Tabellen 17, 18, 19 und 20 sofort, dass die relativen Steuerungsprobleme wieder mehr Leistung in das Gebiet einspeisen als die jeweiligen Pendanten. Des Weiteren besitzt in jeder Gruppe von Steuerungsproblemen  $[I_3]_k$ ,  $[I_4]_k$ ,  $[I_3^{rel}]_k$  und  $[I_4^{rel}]_k$  das jeweilige Problem für  $k = 3$  die höchste Rechenzeit, ohne dabei bessere Ergebnisse zu liefern.

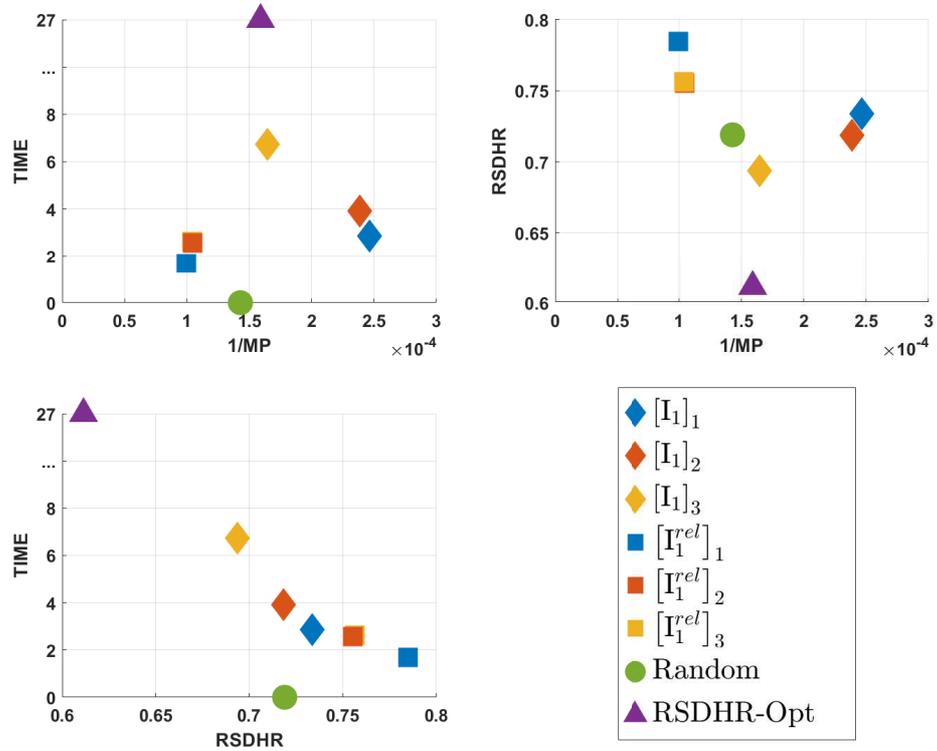


Abbildung 5.22: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_1]_k, [I_1^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_1]_1$	11	$\sim 2.86$ h	0.7338	$4.0597 \cdot 10^3$
$[I_1]_2$	11	$\sim 3.92$ h	0.7185	$4.1884 \cdot 10^3$
$[I_1]_3$	11	$\sim 6.74$ h	0.6935	$6.0749 \cdot 10^3$
$[I_1^{rel}]_1$	11	$\sim 1.70$ h	0.7847	$1.0071 \cdot 10^4$
$[I_1^{rel}]_2$	11	$\sim 2.57$ h	0.7553	$9.5379 \cdot 10^3$
$[I_1^{rel}]_3$	11	$\sim 2.62$ h	0.7563	$9.5957 \cdot 10^3$

Tabelle 17: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_1]$  für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

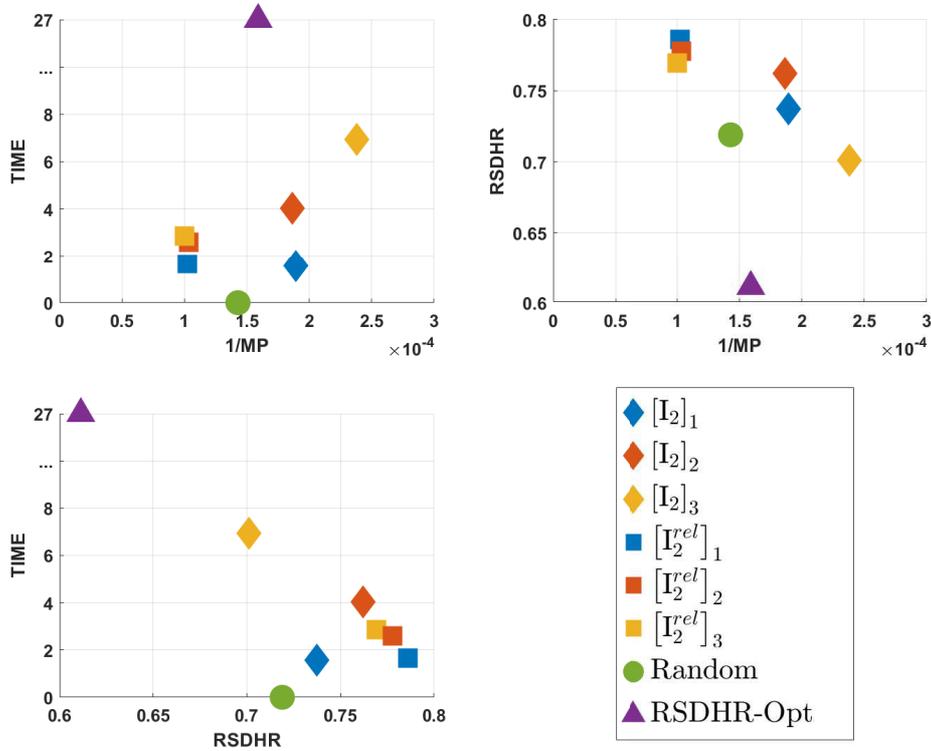


Abbildung 5.23: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_2]_k, [I_2^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_2]_1$	11	$\sim 1.57$ h	0.7373	$5.2829 \cdot 10^3$
$[I_2]_2$	11	$\sim 4.02$ h	0.7621	$5.3630 \cdot 10^3$
$[I_2]_3$	11	$\sim 6.94$ h	0.7010	$4.1961 \cdot 10^3$
$[I_2^{rel}]_1$	11	$\sim 1.66$ h	0.7861	$9.7718 \cdot 10^3$
$[I_2^{rel}]_2$	11	$\sim 2.59$ h	0.7777	$9.6909 \cdot 10^3$
$[I_2^{rel}]_3$	11	$\sim 2.86$ h	0.7692	$9.9840 \cdot 10^3$

Tabelle 18: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_2]$  für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

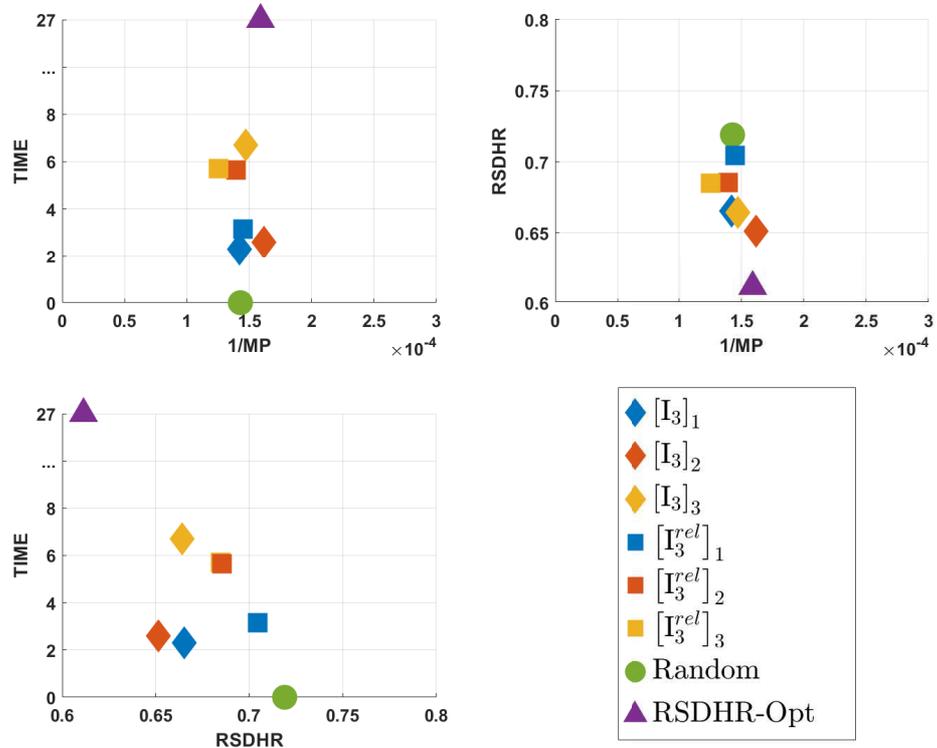


Abbildung 5.24: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_3]_1$	11	$\sim 2.30$ h	0.6654	$7.0471 \cdot 10^3$
$[I_3]_2$	11	$\sim 2.59$ h	0.6512	$6.1921 \cdot 10^3$
$[I_3]_3$	11	$\sim 6.72$ h	0.6643	$6.8022 \cdot 10^3$
$[I_3^{rel}]_1$	11	$\sim 3.15$ h	0.7043	$6.8855 \cdot 10^3$
$[I_3^{rel}]_2$	11	$\sim 5.66$ h	0.6852	$7.1903 \cdot 10^3$
$[I_3^{rel}]_3$	11	$\sim 5.71$ h	0.6846	$7.9736 \cdot 10^3$

Tabelle 19: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_3]$  für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

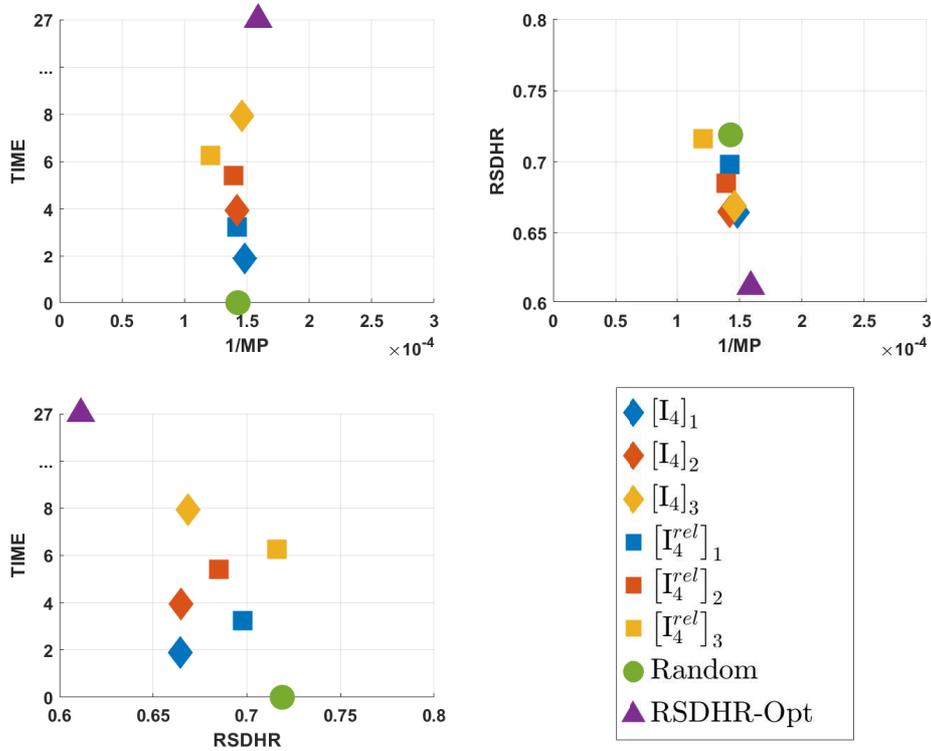


Abbildung 5.25: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_4]_1$	11	$\sim 1.88$ h	0.6645	$6.7500 \cdot 10^3$
$[I_4]_2$	11	$\sim 3.96$ h	0.6648	$7.0274 \cdot 10^3$
$[I_4]_3$	11	$\sim 7.94$ h	0.6686	$6.8470 \cdot 10^3$
$[I_4^{rel}]_1$	11	$\sim 3.24$ h	0.6978	$7.0404 \cdot 10^3$
$[I_4^{rel}]_2$	11	$\sim 5.42$ h	0.6849	$7.1826 \cdot 10^3$
$[I_4^{rel}]_3$	11	$\sim 6.27$ h	0.7161	$8.2745 \cdot 10^3$

Tabelle 20: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} = 1, m = 1, 2$ . (Modell I)

Wir betrachten wieder die kumulativ-iterativen Darstellungen der Werte RSDHR Summe und MP Summe von  $[I_3]_1$  und den Vergleichsalgorithmen und wollen daraus Unterschiede zwischen den beiden Fällen  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , und  $P_{in,m} = 1$ ,  $m = 1, 2$ , herausarbeiten.

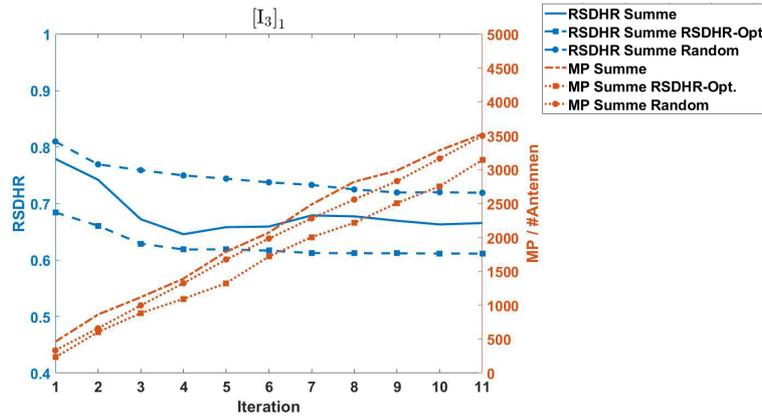


Abbildung 5.26: Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems  $[I_3]_1$  und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ . (Modell I)

Vergleichen wir das Steuerungsproblem  $[I_3]_1$  mit dem Random-Algorithmus, dann erhalten wir eine bessere Gleichmäßigkeit bei gleichem Energieeintrag. Im Gegensatz dazu erhalten wir mit dem RSDHR-Optimierung-Algorithmus die zu erwartenden besseren RSDHR-Werte, aber speisen zeitgleich auch weniger Energie ein. Damit ist das Steuerungsproblem  $[I_3]_1$  eine Kombination aus Energieeintrag und Gleichmäßigkeit, so wie wir es bei der Definition in Abschnitt 5.4 auch beschrieben haben. Weitere sehr interessante Ergebnisse erhalten wir bei dem Vergleich der Abbildungen 5.18 und 5.26. Betrachten wir dazu die Benchmark-Werte in den Tabellen 15 und 19, dann erhalten wir für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , einen höheren Energieeintrag und zudem eine bessere Gleichmäßigkeit. Dies korreliert wieder mit der Einschränkung  $P_{in,m} = 1$ ,  $m = 1, 2$ , die den Suchraum  $\mathbb{P}_S$  deutlich verkleinert. Durch die Fixierung auf ein Watt pro Antenne legen wir schon fest, dass mehr Energie eingespeist werden soll. Dies wiederum verbessert die Wirkung des Strafterms, denn bei der Wahl  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , lässt sich der Abstand zweier Felder im Strafterm vergrößern indem die Leistungen  $[P_{in,m}]_{m=1}^2$  skaliert werden. Dies ist bei  $P_{in,m} = 1$ ,  $m = 1, 2$ , nicht möglich, weshalb der Strafterm damit mehr ins Gewicht fällt.

**Vergleich der Steuerungsprobleme bei Wahl von  $E_{V,\mathcal{N}} = \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]$ ,  $\tilde{p}_S \in \mathbb{P}_S$** 

Betrachten wir Abbildung 5.15, welches sinnbildlich für die Probleme der Steuerungsprobleme  $[I_1]_k$ ,  $[I_1^{rel}]_k$ ,  $[I_2]_k$ ,  $[I_2^{rel}]_k$ ,  $k = 1, 2, 3$ , steht, dann lässt sich sowohl für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , als auch für den Fall  $P_{in,m} = 1$ ,  $m = 1, 2$ , feststellen, dass sich trotz des Fehlerterms  $\mathcal{R}_{\max}^{k,l}$  keine iterative Verbesserung des RSDHR-Wertes ergibt. Wird zudem die Wahl der optimalen Steuerparameter betrachtet, kann festgestellt werden, dass entweder keine Energie eingespeist wird, siehe Tabellen 11 und 12, oder in den einzelnen Iterationen immer annähernd vergleichbare Steuerparameter verwendet werden, siehe Tabellen 13 und 14. Dies liegt an der Wahl des Referenzfeldes  $E_{V,\mathcal{N}}$ , bzw. der Leistungsverteilung  $Q_{V,\mathcal{N}_L}$ , und dem damit verbundenen dominierenden ersten Summanden im Zähler. Innerhalb dieses Abschnitts betrachten wir deshalb die erwähnten Steuerungsprobleme bezüglich der Zielwerte  $E_{V,\mathcal{N}} = \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]$ , bzw.  $Q_{V,\mathcal{N}_L} = \mathcal{P}_{Q,\mathcal{N}_L}[\tilde{p}_S]$ , wobei  $\tilde{p}_S \in \mathbb{P}_S$  eine beliebige Steuerung darstellt. Hierbei beschränken wir uns nur auf die Suche nach einem optimalen Steuerparametervektor  $K_{opt} = 1$  für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$ , und verwenden wiederum den SQP-Algorithmus. Es wird darauf hingewiesen, dass bei der Betrachtung von lediglich einem optimalen Steuerparametervektor die Steuerungsprobleme  $[I_1]$  und  $[I_2]$  identisch sind. Wir geben in der nachfolgenden Tabelle sowohl die erhaltenen Steuerparameter der einzelnen Steuerungsprobleme als auch den Steuerparameter des Referenzfeldes und den Startparameter an.

SP	$f$	$P_{in,1}$	$P_{in,2}$	$\theta_1$	$\theta_2$
$E_{V,\mathcal{N}}, Q_{V,\mathcal{N}_L}$	2.4326	0.8048	0.4105	4.4465	3.4205
$p_S^{start}$	2.4968	0.0117	0.2121	2.2816	6.1078
$[I_1]_1, [I_2]_1$	2.4000	$2.22 \cdot 10^{-16}$	$2.22 \cdot 10^{-16}$	0.7540	0.7540
$[I_1]_2, [I_2]_2$	2.4305	0.9133	$2.22 \cdot 10^{-16}$	0.2914	2.4764
$[I_1]_3, [I_2]_3$	2.4302	1	$2.22 \cdot 10^{-16}$	3.3572	0
$[I_1^{rel}]_1, [I_2^{rel}]_1$	2.5000	1	1	4.2165	1.0100
$[I_1^{rel}]_2, [I_2^{rel}]_2$	2.5000	1	1	4.4678	1.2614
$[I_1^{rel}]_3, [I_2^{rel}]_3$	2.4953	0.5613	0.8525	2.7432	5.5348

Tabelle 21: Optimale Steuerparameter der iterativen Steuerungsprobleme (SP) bei Verwendung von Referenzfeld  $E_{V,\mathcal{N}} = \mathcal{E}_{\mathcal{N}}[\tilde{p}_S]$ , bzw.  $Q_{V,\mathcal{N}_L} = \mathcal{P}_{Q,\mathcal{N}_L}[\tilde{p}_S]$ , mit  $\tilde{p}_S \in \mathbb{P}_S$ , für den Fall  $P_{in,m} \in [0, 1]$ ,  $m = 1, 2$  (Modell I)

Betrachten wir zunächst die Ergebnisse der relativen Steuerungsprobleme in Tabelle 21. In den beiden ersten Fällen wird die Leistung beider Antennen auf eins gesetzt, was

mit der Definition aus Abschnitt 5.4 und der Überlegung der Erhöhung der mittleren Leistung MP einhergeht. Der dritte Fall führt dabei, aufgrund der Verwendung der Leistungsverteilungen, zu einem anderen Ergebnis und variiert hierbei auch noch die Leistungen. Die Gründe für dieses Verhalten können nicht einwandfrei geklärt werden. Bei Betrachtung der Ergebnisse für die Steuerungsprobleme  $[I_1]_k$ ,  $[I_2]_k$ ,  $k = 1, 2, 3$ , erhalten wir im ersten Fall als beste Lösung die Null-Lösung, während wir in den anderen beiden Fällen schon recht nahe an die richtige Frequenz kommen. Dabei unterscheiden sich die verschiedenen Steuerungsprobleme lediglich um die Frequenzinformation im Zielfunktional, welche auch ausschlaggebend für die bessere Frequenzapproximation ist. Die Ergebnisse der Tabelle 21 unterstützen damit die in Abschnitt 5.4 diskutierten Ideen hinter den Kostenfunktionalen der verschiedenen Steuerungsprobleme. Bei einer besseren Wahl der Vergleichsfelder  $E_{V,N}$ , bzw.  $Q_{V,N_L}$ , im Vergleich zu der Herleitung in Abschnitt 5.4 können die Steuerungsprobleme  $[I_1]_k$  und  $[I_2]_k$ ,  $k = 1, 2, 3$ , möglicherweise doch verwendet werden um mehrere optimale Steuerungen zu generieren, welche wiederum gute Benchmark-Werte aufweisen.

### 5.5.2. Modell II: Vier Antennen, ein Objekt

Vergleich der Steuerungsprobleme bezüglich Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$

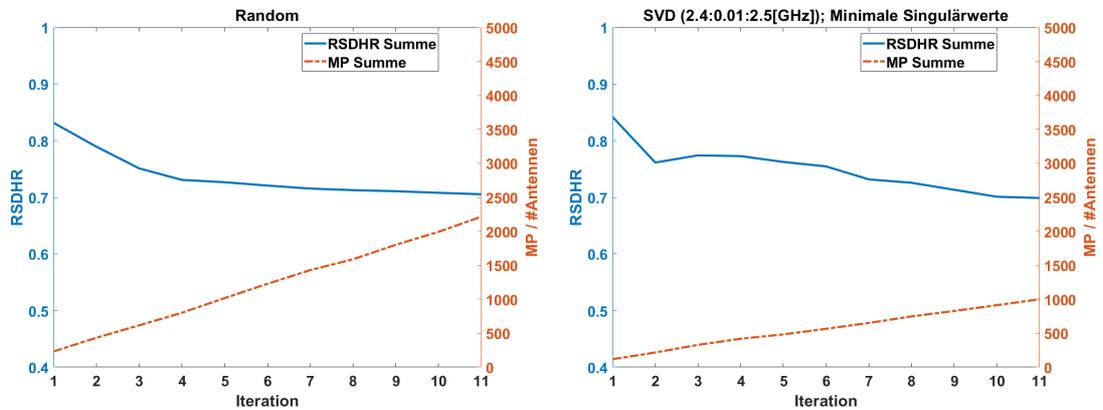


Abbildung 5.27: Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$ . (Modell II)

Bei Betrachtung der Abbildung 5.27 fällt sofort auf, dass der SVD-Algorithmus diesmal deutlich weniger Leistung pro Antenne einspeist. Dies liegt an der Art des SVD-Algorithmus, der die Summe der Leistungen der zu der S-Parametermatrix zugehörigen

Singulärvektoren auf ein Watt normiert. Da sich aber die Gleichmäßigkeit bei Erhöhung der Leistungen mit gleichbleibenden Leistungsunterschieden zwischen den Antennen nicht verändert, kann durch Skalierung der Singulärvektoren eine Variation des SVD-Algorithmus bestimmt werden, der wiederum mehr Leistung einspeist. Des Weiteren sehen wir im Vergleich mit Abbildung 5.12, dass wir bei Verwendung von vier Antennen sowohl mit dem Random- als auch mit dem SVD-Algorithmus keine nennenswerte Verbesserung in der Gleichmäßigkeit erreichen. Wie beim ersten Modell geben wir nachfolgend wieder die zur SVD zugehörige S-Parametermatrix in der dB-Darstellung an.

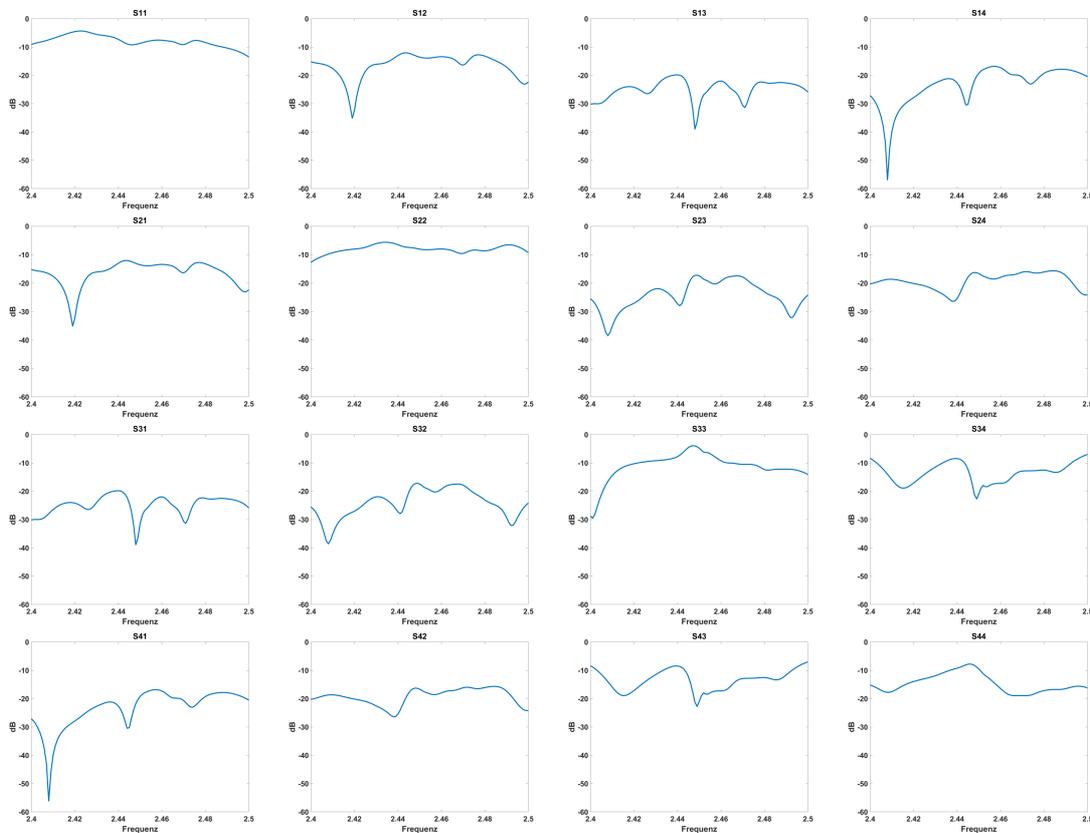


Abbildung 5.28: Darstellung der S-Parametermatrix in dB. (Modell II)

Im Folgenden betrachten wir vergleichend die Ergebnisse zu den Steuerungsproblemen  $[I_3]_k$ ,  $[I_3^{rel}]_k$ ,  $[I_3]_k$  und  $[I_3^{rel}]_k$  mit  $k = 1, 2, 3$  für den Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$ .

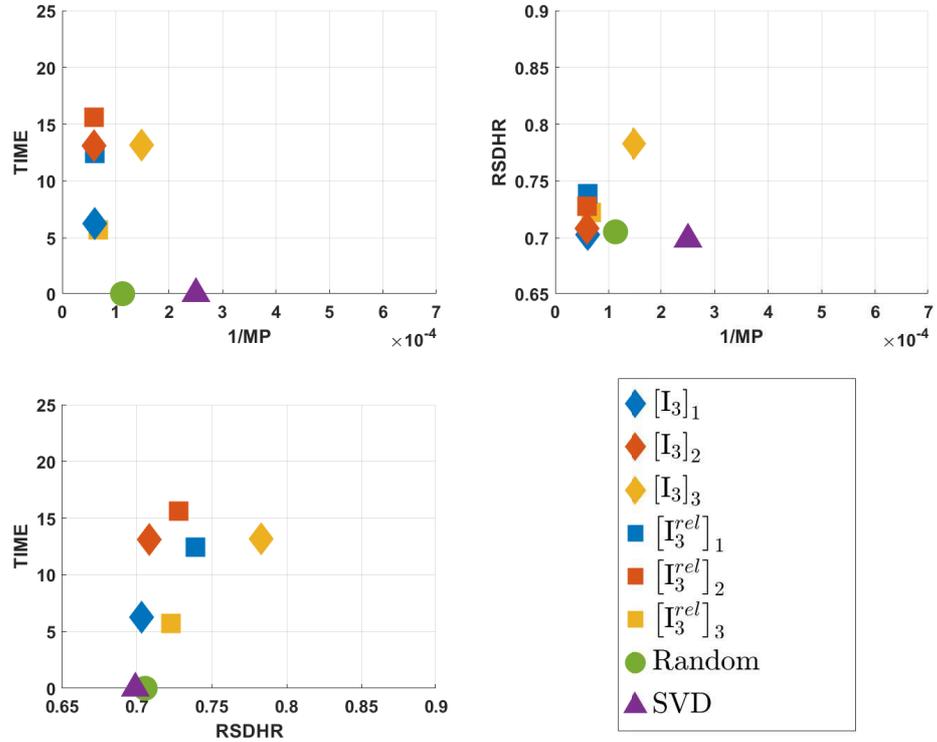


Abbildung 5.29: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell II)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_3]_1$	11	$\sim 6.25$ h	0.7031	$1.6404 \cdot 10^4$
$[I_3]_2$	11	$\sim 13.11$ h	0.7082	$1.6899 \cdot 10^4$
$[I_3]_3$	11	$\sim 13.18$ h	0.7831	$6.7390 \cdot 10^3$
$[I_3^{rel}]_1$	11	$\sim 12.47$ h	0.7389	$1.6430 \cdot 10^4$
$[I_3^{rel}]_2$	11	$\sim 15.62$ h	0.7278	$1.6801 \cdot 10^4$
$[I_3^{rel}]_3$	11	$\sim 5.70$ h	0.7226	$1.4968 \cdot 10^4$

Tabelle 22: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell II)

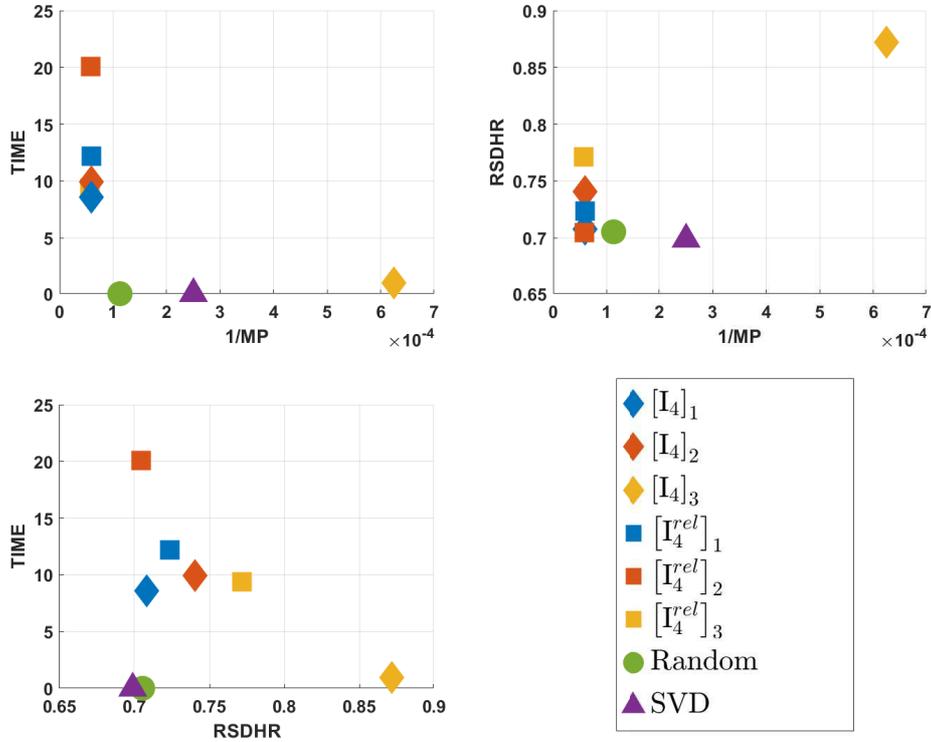


Abbildung 5.30: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell II)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_4]_1$	11	$\sim 8.62$ h	0.7080	$1.6785 \cdot 10^4$
$[I_4]_2$	11	$\sim 9.94$ h	0.7406	$1.6839 \cdot 10^4$
$[I_4]_3$	3	$\sim 0.94$ h	0.8719	$1.5996 \cdot 10^3$
$[I_4^{rel}]_1$	11	$\sim 12.18$ h	0.7236	$1.6861 \cdot 10^4$
$[I_4^{rel}]_2$	11	$\sim 20.05$ h	0.7046	$1.7260 \cdot 10^4$
$[I_4^{rel}]_3$	11	$\sim 9.40$ h	0.7716	$1.7757 \cdot 10^4$

Tabelle 23: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell II)

Bei der Betrachtung der Abbildungen 5.29 und 5.30 fällt auf, dass die Steuerungsprobleme  $[I_3]_3$  und  $[I_4]_3$  im Vergleich zu den anderen Steuerungsproblemen schlechtere Ergebnisse liefern. Dies haben wir auch schon im ersten Modell bei den Abbildungen 5.16 und 5.17 ableiten können, weshalb wir  $[I_3]_3$  und  $[I_4]_3$  aus der folgenden Diskussion ausnehmen. Vergleichen wir demnach die anderen Steuerungsprobleme mit den Vergleichsalgorithmen, dann liefern alle Steuerungsprobleme einen größeren und annähernd vergleichbaren Leistungseintrag. Wir speisen damit wiederum mehr Energie ein als die Vergleichsalgorithmen, was, wie bereits bei der Diskussion im ersten Modell erwähnt, an der Definition der Kostenfunktionale festzumachen ist. Während wir einen fast doppelt so hohen Energieeintrag wie der Random-Algorithmus generieren, bekommen wir anders als im ersten Modell keine bessere Gleichmäßigkeit. Diesen Vergleich stellen wir wieder mit dem Steuerungsproblem  $[I_3]_1$  in Abbildung 5.31 bezüglich der kumulativ-iterativen Darstellung der Benchmark-Werte RSDHR Summe und MP Summe dar. Vergleichen wir die beiden Abbildungen 5.18 und 5.31 des ersten und zweiten Modells, dann haben wir für  $[I_3]_1$  im zweiten Modell deutlich mehr Energie pro Antenne eingebracht als im ersten Modell. Im Gegensatz dazu speist der SVD diesmal kaum Energie ein und der Random-Algorithmus kommt wiederum auf annähernd denselben MP-Wert. Bei der Gleichmäßigkeit hingegen liefern alle drei Algorithmen fast den gleichen Wert. Da in der Definition der Steuerungsprobleme der erste Term den Energieeintrag maximiert und dieser Term aufgrund des deutlich größeren Suchraums dominiert, wirkt der Strafterm  $\mathcal{R}_{\max}$ , der für eine bessere Gleichmäßigkeit sorgt, nicht mehr so stark wie im ersten Modell. Eine Verbesserung im RSDHR-Wert können wir aber auch im zweiten Modell erkennen. Zusammengefasst erhalten wir im Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$ , des zweiten Modells mit  $[I_3]_1$  die besten Ergebnisse.

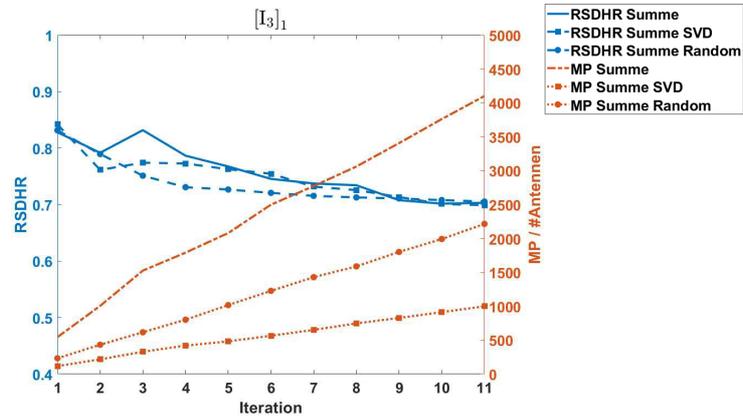


Abbildung 5.31: Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems  $[I_3]_1$  und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$ . (Modell II)

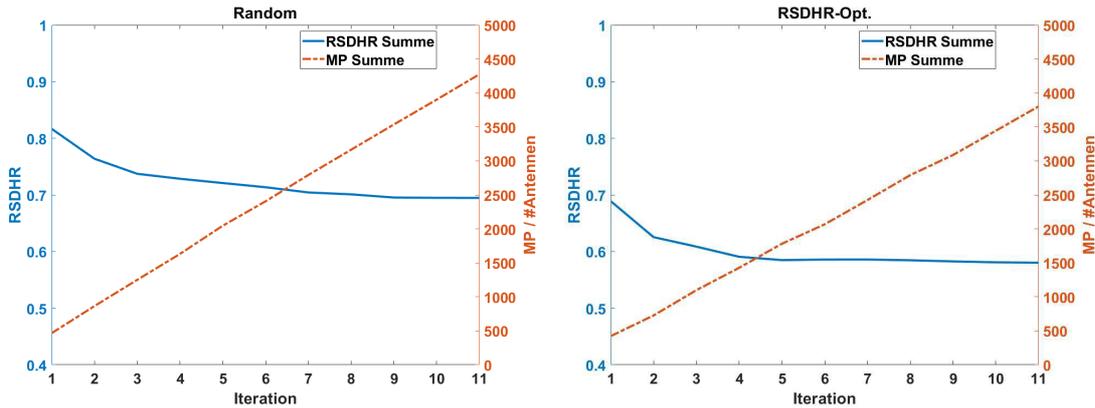
**Vergleich der Steuerungsprobleme bezüglich Fall  $P_{in,m} = 1, 1 \leq m \leq 4$** 


Abbildung 5.32: Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II)

Vergleichen wir die Abbildungen 5.19 und 5.32 miteinander, dann sehen wir zum einen die geringe Verbesserung der Gleichmäßigkeit durch Verwendung von vier Antennen und den deutlich höheren Energieeintrag pro Antenne. Während die verbesserten RSDHR-Werte mit der Anzahl an Antennen und der damit variablen Überlagerung von elektromagnetischen Feldern korreliert werden kann, erhalten wir die deutlich besseren Energieeinträge wohl durch die veränderte Positionierung der Antennen in der Kavität. Diese Abhängigkeit bezüglich der Einkoppelposition ist in der Hochfrequenztechnik bekannt und wir können damit durch Betrachtung der Energie pro Antenne die Qualität der Einkoppelposition auch für andere Modelle verifizieren. Auch wenn wir bei vier Antennen eine verbesserte Gleichmäßigkeit erreichen können, sind wir durch die Geometrie des zu erwärmenden Gebiets und der Kavität in den RSDHR-Werten aufgrund der ausbreitungsfähigen elektromagnetischen Felder beschränkt.

Wir widmen uns nun den Ergebnissen zu den Steuerungsproblemen  $[I_3]_k$ ,  $[I_4]_k$ ,  $[I_3^{rel}]_k$  und  $[I_4^{rel}]_k$ ,  $k = 1, 2, 3$ , und vergleichen diese mit denen aus Kapitel 5.5.1. Bei Verwendung von vier Antennen und der Betrachtung des Falles  $P_{in,m} = 1, 1 \leq m \leq 4$ , erhalten wir in Abbildung 5.33 und Abbildung 5.34 durchwegs eine schlechtere Gleichmäßigkeit als bei Verwendung des Random-Algorithmus, aber andererseits speisen wir dafür mehr Energie ein.

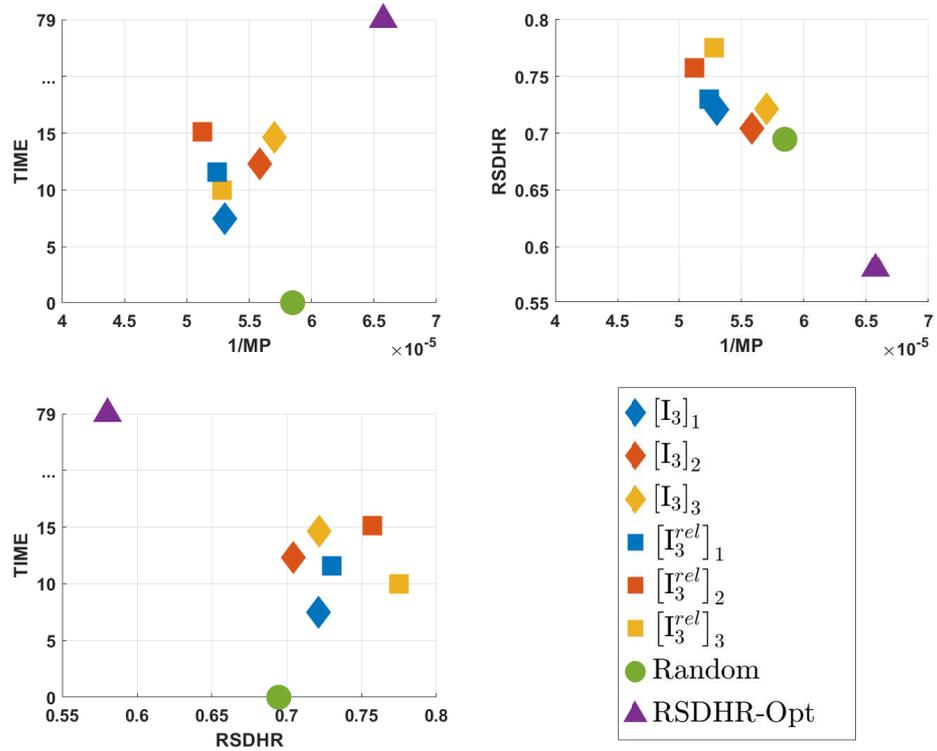


Abbildung 5.33: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_3]_1$	11	$\sim 7.48$ h	0.7211	$1.8860 \cdot 10^4$
$[I_3]_2$	11	$\sim 12.29$ h	0.7044	$1.7915 \cdot 10^4$
$[I_3]_3$	11	$\sim 14.63$ h	0.7217	$1.7545 \cdot 10^4$
$[I_3^{rel}]_1$	11	$\sim 11.57$ h	0.7302	$1.9070 \cdot 10^4$
$[I_3^{rel}]_2$	11	$\sim 15.11$ h	0.7575	$1.9518 \cdot 10^4$
$[I_3^{rel}]_3$	11	$\sim 10.02$ h	0.7750	$1.8943 \cdot 10^4$

Tabelle 24: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_3]$  für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II)

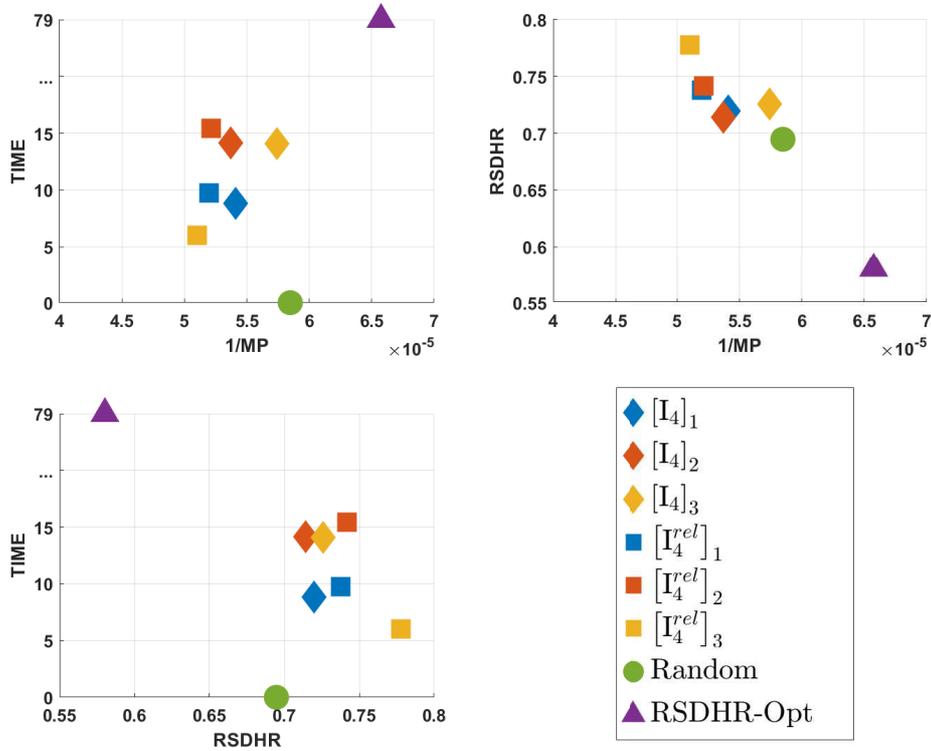


Abbildung 5.34: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II)

SP	$\# p_S^{opt}$	TIME	RSDHR Summe	MP Summe
$[I_4]_1$	11	$\sim 8.86$ h	0.7199	$1.8478 \cdot 10^4$
$[I_4]_2$	11	$\sim 14.14$ h	0.7143	$1.8615 \cdot 10^4$
$[I_4]_3$	11	$\sim 14.07$ h	0.7260	$1.7413 \cdot 10^4$
$[I_4^{rel}]_1$	11	$\sim 9.75$ h	0.7377	$1.9234 \cdot 10^4$
$[I_4^{rel}]_2$	11	$\sim 15.44$ h	0.7419	$1.9176 \cdot 10^4$
$[I_4^{rel}]_3$	11	$\sim 6.01$ h	0.7780	$1.9611 \cdot 10^4$

Tabelle 25: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II)

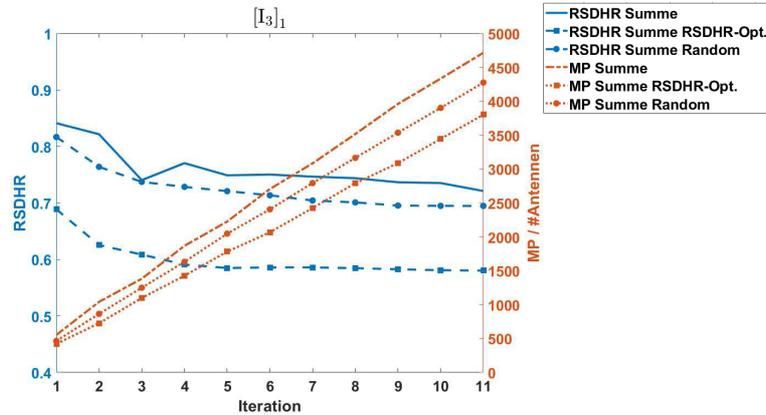


Abbildung 5.35: Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems  $[I_3]_1$  und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell II)

Eine mögliche Ursache für die schlechten RSDHR-Werte ist der höhere Leistungseintrag pro Antenne. Wie zuvor bereits beschrieben, ist der Leistungseintrag pro Antenne bei dem zweiten Modell mit vier Antennen höher als beim dem des ersten Modells. Da die Steuerungsprobleme so definiert sind, dass viel Energie eingebracht werden soll, dominiert hier der Effekt der Leistungsmaximierung gegenüber dem der Gleichmäßigkeit. Der im Vergleich zum ersten Modell größere Suchraum, wegen der Verwendung von vier Antennen, führt dabei zu einer Mehrzahl an Steuerungen mit hohem Leistungseintrag, die meist eine schlechtere Gleichmäßigkeit aufweisen. Es kann jedoch auch an dem Strafterm liegen, denn dieser bestraft zueinander ähnliche elektromagnetische Felder. Wenn sich beispielsweise zwei Felder ähneln, wobei eines eine gute Gleichmäßigkeit aufweist und das andere nicht, und zunächst das Feld mit schlechterem RSDHR-Wert gewählt wird, dann würde das Feld mit einer guten Gleichmäßigkeit nicht mehr gewählt werden, da der Strafterm hierfür einen größeren Wert ausgibt. Die Ergebnisse zu den Abbildungen 5.33, 5.34 und 5.35 zeigen bereits auf, dass noch weitere Forschung bei der Wahl der Steuerungsfunktionen für eine bessere Gleichmäßigkeit notwendig ist. Bestimmte Effekte können wir jedoch sowohl für das erste Modell als auch für das zweite Modell herausarbeiten. Zum einen speisen die relativen Varianten mehr Leistung in das zu erwärmende Objekt ein und zum anderen benötigen wir für  $k = 1, 2$  tendentiell mehr Rechenzeit bei der Berechnung der relativen Varianten als bei deren Pendanten. Dies konnte in beiden Fällen  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ , und  $P_{in,m} = 1, 1 \leq m \leq 4$ , festgestellt werden. Eine Ausnahme hierzu stellt wiederum die Betrachtung von  $k = 3$  dar, bei der keine Tendenzen erkennbar sind und deren Verhalten bisher nicht endgültig geklärt werden kann.

## 5.5.3. Modell III: Vier Antennen, zwei Objekte

Vergleich der Steuerungsprobleme bezüglich Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$

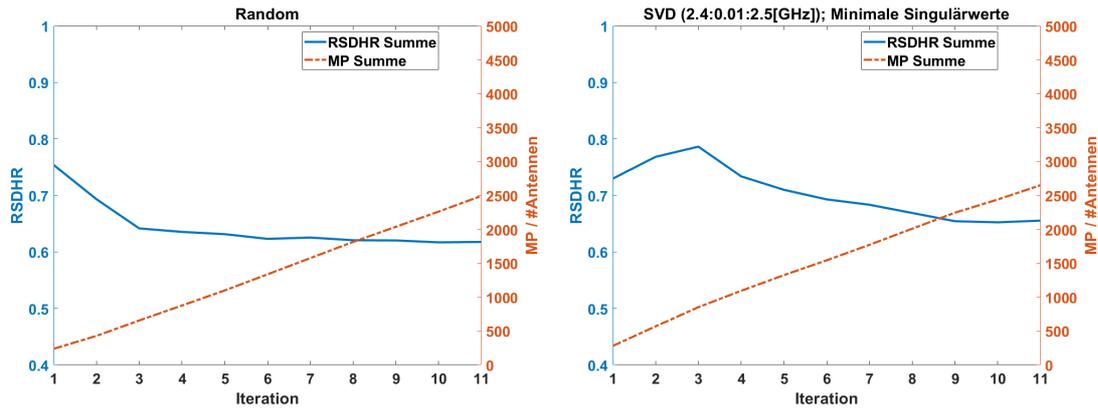


Abbildung 5.36: Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III)

Im Gegensatz zum zweiten Modell erhalten wir bei dem dritten Modell wieder einen annähernd vergleichbaren Leistungseintrag bei Random- und SVD-Algorithmus. Da in beiden Modellen jeweils vier Antennen derselben Art und Position verwendet wurden, kann die Verbesserung des Leistungseintrages pro Antenne eindeutig auf die Geometrien der zu erwärmenden Epoxidharze zurückgeführt werden. Das bedeutet also, dass der SVD-Algorithmus und somit die S-Parameter, sowie die dazugehörigen SVD-Vektoren, Informationen über die Geometrie innerhalb der Kavität besitzen. Damit kann möglicherweise ein Zusammenhang aus den optimalen Steuerungen und den S-Parametern abgeleitet werden, die wiederum als Heuristik in Echtzeit für jegliche Art von Kavität verwendet werden kann. Wir geben nachfolgend wieder die zur SVD zugehörige S-Parametermatrix in der dB-Darstellung an.

## 5. Optimalsteuerung elektromagnetischer Felder

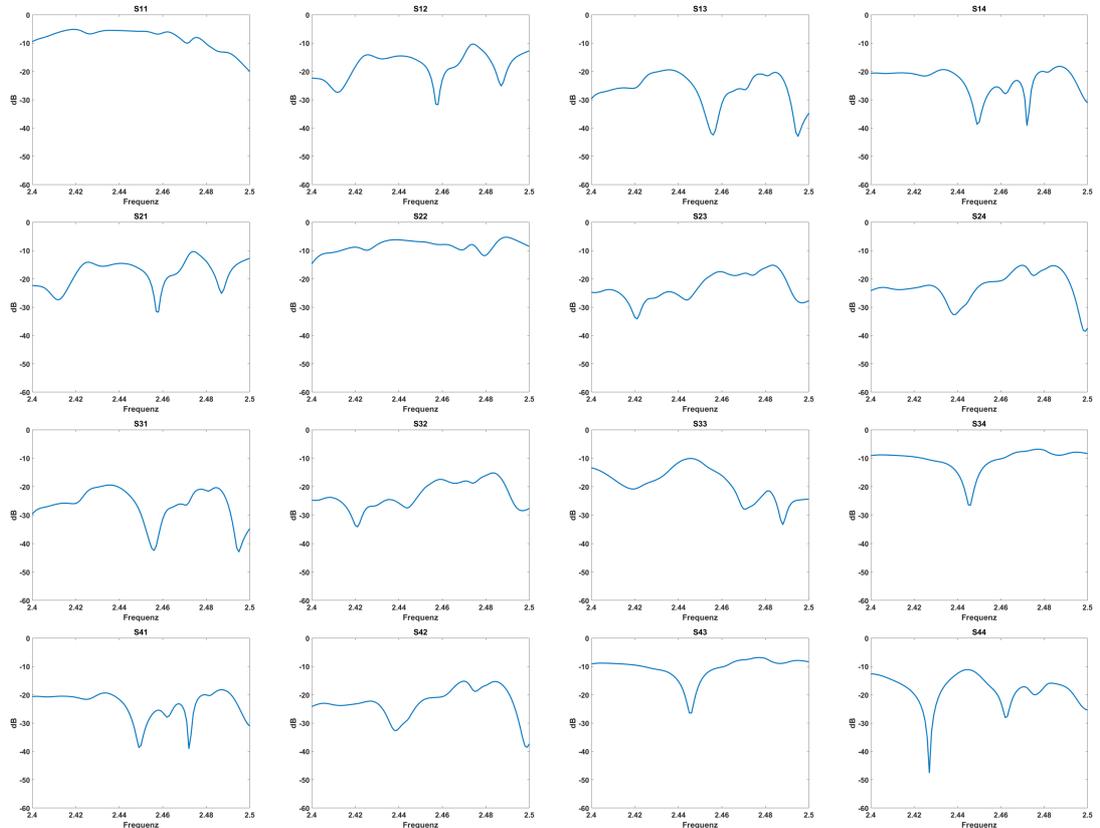


Abbildung 5.37: Darstellung der S-Parametermatrix in dB. (Modell III)

Im Folgenden betrachten wir vergleichend die Ergebnisse zu den Steuerungsproblemen  $[I_3]_k$ ,  $[I_3^{rel}]_k$ ,  $[I_3]_k$  und  $[I_3^{rel}]_k$ , mit  $k = 1, 2, 3$ , für den Fall  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$ . Dabei werden wir die Rechenzeit der Steuerungsprobleme nicht mehr mit angeben, da keine neuen Informationen hieraus abgeleitet werden können. Als erstes bleibt auch für das dritte Modell festzuhalten, dass die Steuerungsprobleme mit  $k = 3$  wiederum die schlechtesten Ergebnisse liefern und sich komplett anders verhalten als der Rest. Für die Werte  $k = 1, 2$  erhalten wir zudem wieder eine bessere Gleichmäßigkeit bei höherem Energieeintrag, was wiederum an den Geometrien der Epoxidharze liegt. Demnach wirkt der Strafterm aufgrund der veränderten Geometrie innerhalb der Kavität wiederum deutlich besser. Möglicherweise führt eine adaptive und von der Geometrie abhängige Skalierung des Strafterms zu deutlich besseren Steuerungsproblemen, die dann auch im zweiten Modell sehr gute Ergebnisse liefern.

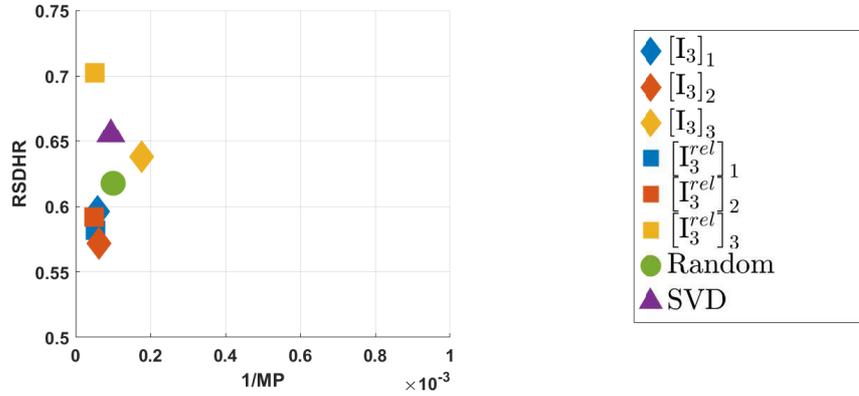


Abbildung 5.38: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III)

SP	# $p_S^{opt}$	RSDHR Summe	MP Summe
$[I_3]_1$	11	0.5962	$1.7122 \cdot 10^4$
$[I_3]_2$	11	0.5720	$1.6191 \cdot 10^4$
$[I_3]_3$	11	0.6378	$5.6488 \cdot 10^3$
$[I_3^{rel}]_1$	11	0.5817	$1.8710 \cdot 10^4$
$[I_3^{rel}]_2$	11	0.5921	$1.9909 \cdot 10^4$
$[I_3^{rel}]_3$	11	0.7028	$1.9721 \cdot 10^4$

Tabelle 26: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III)

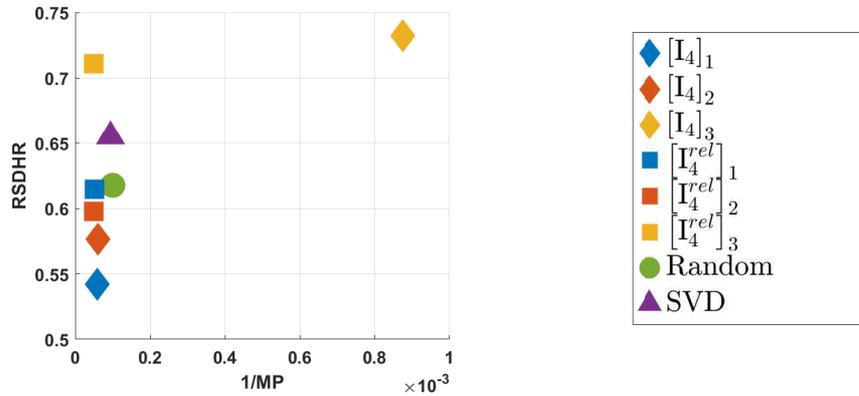


Abbildung 5.39: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III)

SP	$\# p_S^{opt}$	RSDHR Summe	MP Summe
$[I_4]_1$	11	0.5421	$1.7005 \cdot 10^4$
$[I_4]_2$	11	0.5766	$1.6338 \cdot 10^4$
$[I_4]_3$	2	0.7320	$1.1412 \cdot 10^3$
$[I_4^{rel}]_1$	11	0.6149	$1.9432 \cdot 10^4$
$[I_4^{rel}]_2$	11	0.5979	$2.0239 \cdot 10^4$
$[I_4^{rel}]_3$	11	0.7110	$1.9876 \cdot 10^4$

Tabelle 27: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} \in [0, 1], 1 \leq m \leq 4$ . (Modell III)

Betrachten wir wieder die vergleichende Darstellung in Abbildung 5.40 der verschiedenen Algorithmen, dann erhalten wir wie schon im zweiten Modell einen deutlich höheren Leistungseintrag pro Antenne im Vergleich zum Random- und SVD-Algorithmus und zudem diesmal auch noch eine bessere Gleichmäßigkeit. Dabei ist der RSDHR-Wert bei allen drei Algorithmen deutlich geringer als noch beim zweiten Modell, was wiederum auf die Geometrie innerhalb der Kavität zurückgeführt werden kann. Die Geometrie ist also maßgeblich für die maximal mögliche Gleichmäßigkeit und auch für den maximal möglichen Leistungseintrag verantwortlich.

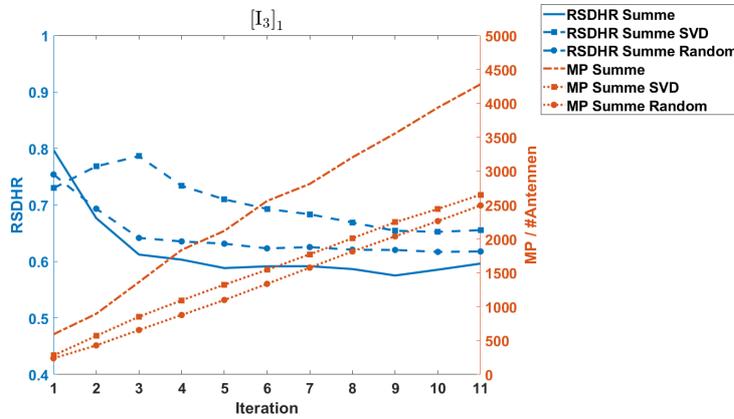


Abbildung 5.40: Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems  $[I_3]_1$  und der Vergleichsalgorithmen für den Fall a)  $P_{in,m} \in [0, 1]$ ,  $1 \leq m \leq 4$ . (Modell III)

**Vergleich der Steuerungsprobleme bezüglich Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$**

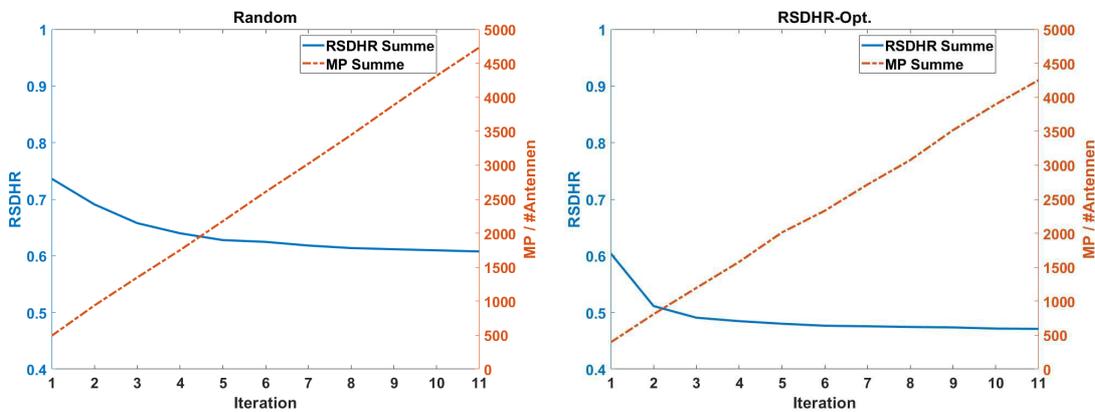


Abbildung 5.41: Kumulativ-iterative RSDHR- und MP-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ . (Modell III)

Wie bereits am Ende des letzten Abschnitts genannt, erhalten wir auch für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ , wiederum deutlich bessere RSDHR-Werte aufgrund der Geometrie

## 5. Optimalsteuerung elektromagnetischer Felder

der zu erwärmenden Objekte. Zudem sehen wir in Abbildung 5.41, dass der Random-Algorithmus bei schlechterer Gleichmäßigkeit mehr Energie pro Antenne einspeist.

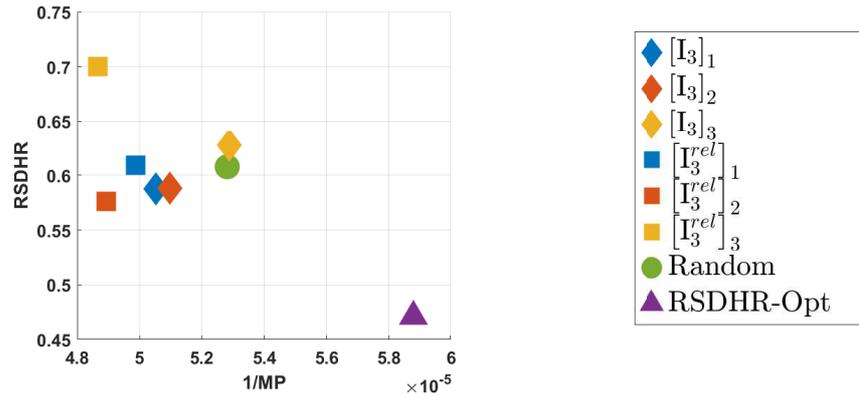


Abbildung 5.42: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_3]_k, [I_3^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III)

SP	$\# p_S^{opt}$	RSDHR Summe	MP Summe
$[I_3]_1$	11	0.5877	$1.9790 \cdot 10^4$
$[I_3]_2$	11	0.5884	$1.9617 \cdot 10^4$
$[I_3]_3$	11	0.6274	$1.8912 \cdot 10^4$
$[I_3^{rel}]_1$	11	0.6090	$2.0047 \cdot 10^4$
$[I_3^{rel}]_2$	11	0.5763	$2.0434 \cdot 10^4$
$[I_3^{rel}]_3$	11	0.6999	$2.0550 \cdot 10^4$

Tabelle 28: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_3]$  für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III)

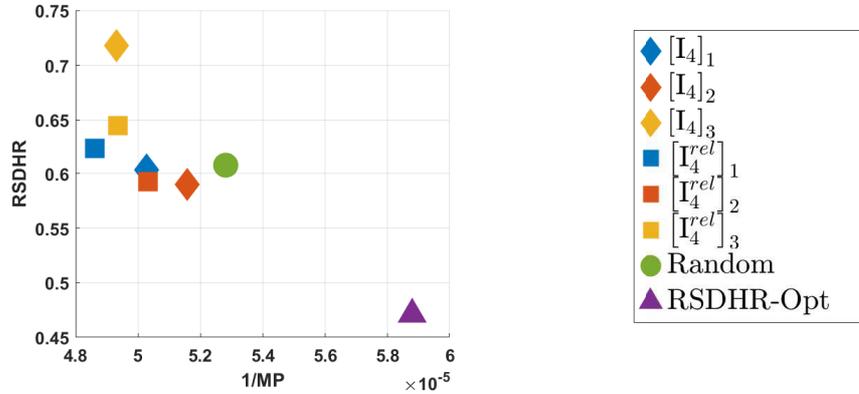


Abbildung 5.43: Kumulativ-iterative Vergleichswerte bezüglich der iterativen Steuerungsprobleme  $[I_4]_k, [I_4^{rel}]_k, k = 1, 2, 3$ , und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III)

SP	# $p_S^{opt}$	RSDHR Summe	MP Summe
$[I_4]_1$	11	0.6090	$1.9887 \cdot 10^4$
$[I_4]_2$	11	0.5899	$1.9384 \cdot 10^4$
$[I_4]_3$	11	0.7177	$2.0276 \cdot 10^4$
$[I_4^{rel}]_1$	11	0.6233	$2.0566 \cdot 10^4$
$[I_4^{rel}]_2$	11	0.5926	$1.9870 \cdot 10^4$
$[I_4^{rel}]_3$	11	0.6445	$2.0258 \cdot 10^4$

Tabelle 29: Kumulativ-iterative Benchmark-Werte bezüglich der iterativen Steuerungsprobleme (SP)  $[I_4]$  für den Fall  $P_{in,m} = 1, 1 \leq m \leq 4$ . (Modell III)

Für jedes Steuerungsproblem  $[I_3]_k, [I_3^{rel}]_k, [I_4]_k$  und  $[I_4^{rel}]_k$  mit  $k = 1, 2$  erhalten wir annähernd denselben RSDHR-Wert wie der Random-Algorithmus, speisen dafür aber mehr Leistung ein.

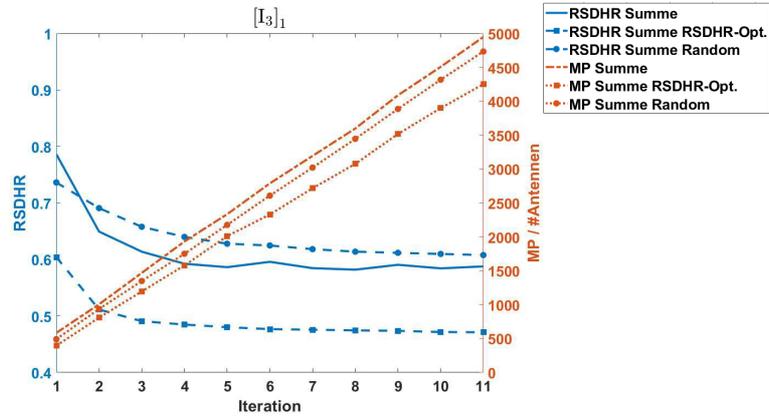


Abbildung 5.44: Kumulativ-iterative RSDHR- und MP-Werte des Steuerungsproblems  $[I_3]_1$  und der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ . (Modell III)

In Abbildung 5.44 können wir wiederum gut den Effekt des Strafterms erkennen, der zur iterativen Verbesserung des RSDHR-Wertes führt. Des Weiteren können wir festhalten, dass wir auch hier wiederum mehr Energie als die Vergleichsalgorithmen einspeisen. Damit liefert  $[I_3]_1$  auch im dritten Modell wieder sehr gute Ergebnisse.

### Vergleich der Steuerungsprobleme bezüglich gezielter Erwärmung

In diesem letzten Abschnitt der numerischen Ergebnisse wollen wir noch einen weiteren letzten Punkt ansprechen, der von regem Interesse für die aktuelle Forschung und die Industrie ist und die Stärke der Optimalsteuerung zeigen wird. In unserem dritten Modell haben wir zwei zu erwärmende Epoxidharze unterschiedlicher Größe. Falls beide Objekte zum selben Zeitpunkt ausgehärtet sein sollen, muss mehr Energie in das größere Gebiet eingebracht werden, um die Aushärtungszeit im Vergleich zum kleineren Gebiet zu verkürzen. Diesen Vorgang bezeichnen wir folgend als gezielte Erwärmung. Hierbei kann für die Verteilung der eingespeisten Energie auch gemeint sein, dass keine Energie in einen der Bereiche eingebracht werden soll. Um Algorithmen für eine gezielte Erwärmung bezüglich ihres Verhaltens zu vergleichen, benötigen wir einen weiteren Benchmark-Wert.

**Definition 5.73.** Sei  $V \subset \Omega$  und  $V_H, V_K \subset V$ . Das Gebiet  $V_H$  symbolisiert jenes Gebiet welchem mehr Energie zugeführt werden soll, während  $V_K$  für das kältere mit weniger Energie bedachte Gebiet steht. Das Benchmark-Kriterium für eine gezielte Erwärmung zwischen zwei Gebieten ist gegeben als:

- MPDR (Mean Power Density Ratio), Relation der durchschnittlich eingebrachten Leistungen:

$$\text{MPDR} = \frac{\left( \frac{1}{|V_K|} \int_{V_K} |\mathcal{P}_{Q, \mathcal{N}_L}[p_S]|^2 d\Omega \right)^{\frac{1}{2}}}{\left( \frac{1}{|V_H|} \int_{V_H} |\mathcal{P}_{Q, \mathcal{N}_L}[p_S]|^2 d\Omega \right)^{\frac{1}{2}}} \quad (5.89)$$

**Definition 5.74.** Sei  $1 \leq l \leq K_{opt}$ ,  $V \subset \Omega$  und  $V_H, V_K \subset V$ . Das Gebiet  $V_H$  symbolisiert jenes Gebiet welchem mehr Energie zugeführt werden soll, während  $V_K$  für das kältere mit weniger Energie bedachte Gebiet steht. Das Benchmark-Kriterium für eine gezielte Erwärmung zwischen zwei Gebieten ist gegeben als:

- MPDR Summe (Mean Power Density Ratio), kumulativ-iterative Relation der durchschnittlich eingebrachten Leistungen:

$$\text{MPDR Summe} = \frac{\left( \frac{1}{|V_K|} \int_{V_K} \left| \sum_{m=1}^l \mathcal{P}_{Q, \mathcal{N}_L}[p_{S,m}] \right|^2 d\Omega \right)^{\frac{1}{2}}}{\left( \frac{1}{|V_H|} \int_{V_H} \left| \sum_{m=1}^l \mathcal{P}_{Q, \mathcal{N}_L}[p_{S,m}] \right|^2 d\Omega \right)^{\frac{1}{2}}} \quad (5.90)$$

Um die bisher betrachteten Algorithmen hinsichtlich ihrer MPDR-Werte zu vergleichen, definieren wir  $V_H = V_{E,1}$ ,  $V_K = V_{E,2}$  bezüglich der in Abschnitt 4.3.3 definierten Gebiete  $V_{E,1}$ ,  $V_{E,2}$  und wollen lediglich den oben genannten Fall betrachten, in dem viel Energie in  $V_H$  und wenig Energie in  $V_K$  eingebracht werden soll. Um hierzu einen Vergleichsalgorithmus für den minimal erreichbaren MPDR-Wert zu erhalten, betrachten wir für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ , den zusätzlichen Optimierungsalgorithmus:

- MPDR-Opt: Optimierung bezüglich des Benchmark-Wertes MPDR Summe

Wir geben nachfolgend zum einen die kumulativ-iterative Entwicklung der Werte MPDR Summe und RSDHR Summe der beiden zu erwärmenden Gebiete und zum anderen noch die schlussendlich erhaltenen Werte HTQ Summe, RSDHR Summe und MP Summe bezogen auf die einzelnen Gebiete  $V_K$  und  $V_H$  an.

## 5. Optimalsteuerung elektromagnetischer Felder

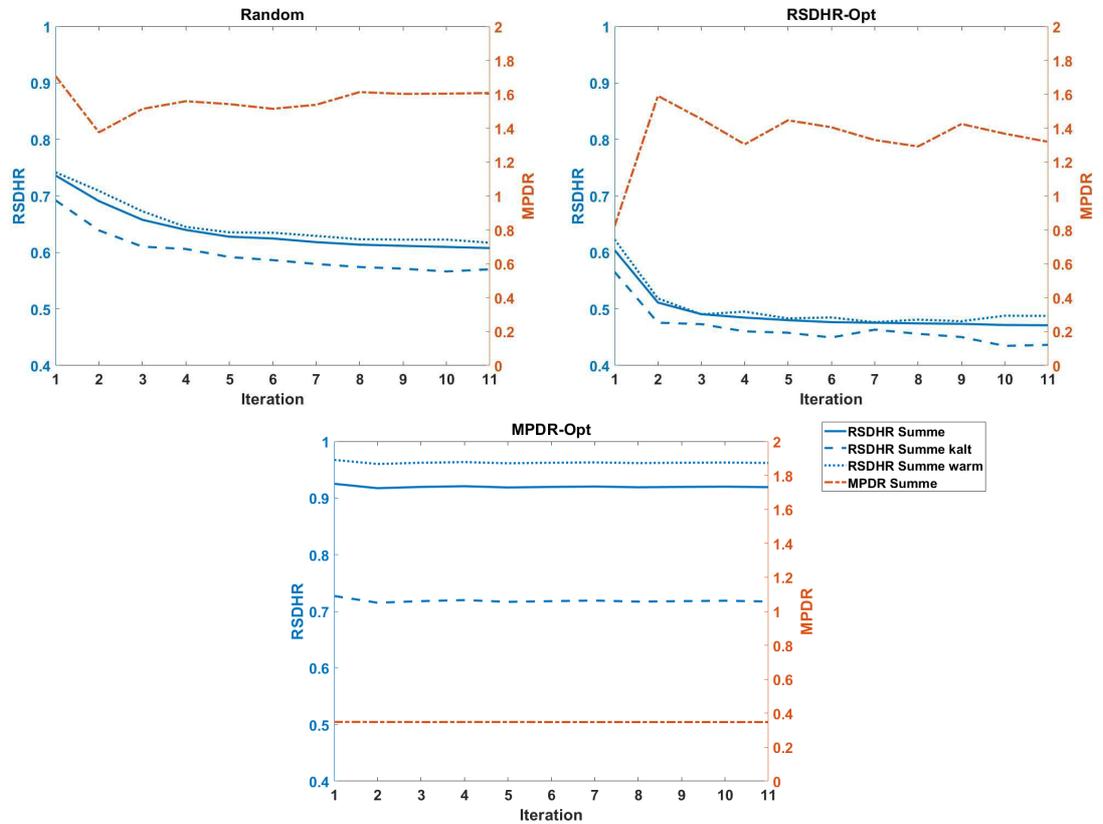


Abbildung 5.45: Kumulativ-iterative RSDHR- und MPDR-Werte der Vergleichsalgorithmen für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ .

Während der RSDHR-Optimierung-Algorithmus eine Verbesserung der Gleichmäßigkeit durch Anregung mehrerer verschiedener elektromagnetischer Felder generiert, ergeben sich beim MPDR-Optimierung-Algorithmus in Abbildung 5.45 kaum Veränderungen in den Benchmark-Werten. Dies liegt an der Wahl der Steuerungen, die selbst bei der kumulativ-iterativen Betrachtung der Benchmark-Werte nur zwei verschiedene Felder anregt. Dies bedeutet wiederum, dass keine Verbesserung der Werte MPDR Summe durch Verwendung vieler verschiedener Steuerungen erhalten werden kann. In Abbildung 5.46 lässt sich der deutliche Unterschied des Energieeintrags bei dem MPDR-Optimierung-Algorithmus feststellen. Sehr gut zu erkennen ist dabei auch die Korrelation zwischen RSDHR-, HTQ- und MP-Werten. Wie bereits am Anfang des Kapitels erwähnt, gibt es eine starke Korrelation zwischen RSDHR- und HTQ-Werten, denn eine gute Gleichmäßigkeit in der Erwärmung erzeugt auch weniger starke Hotspots. Des Weiteren führt die Verbesserung des MPDR-Wertes zu einer schlechteren Gleichmäßigkeit in der Erwärmung.

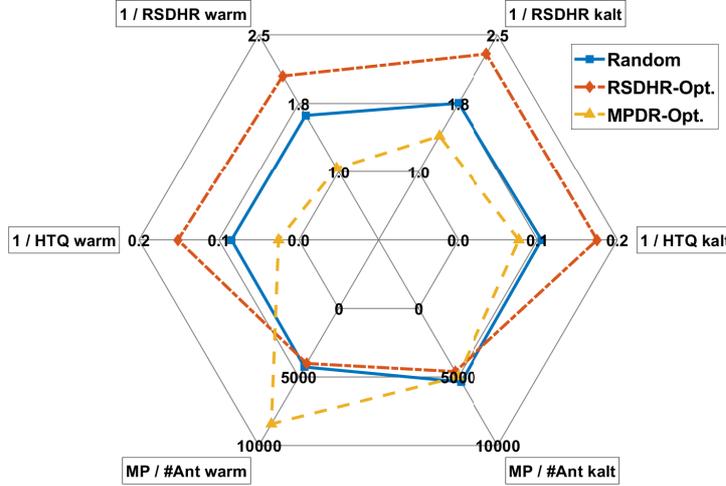


Abbildung 5.46: Kumulativ-iterative RSDHR-, HTQ und MP-Werte der Vergleichsalgorithmen im warmen und kalten Gebiet für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ .

Wir wollen diese nun mit einem weiteren Steuerungsproblem vergleichen und geben hierzu folgende Definition.

**Definition 5.75.** Sei  $E_{V,\mathcal{N}} = 0$  gegeben und  $V \subset \Omega$  und  $V_H, V_K \subset V$ . Die  $l$ -ten Optimalsteuerungsprobleme des iterativen Prozesses für die gezielte Erwärmung sind:

$$[\mathbb{I}_{3,mpdr}]_1^l \quad \min_{p_S \in \mathbb{P}_S} \left[ \frac{\min_{\theta \in I_\theta} [\mathcal{Z}_1[\rho(p_S, \theta)]]|_{V_K}}{\min_{\theta \in I_\theta} [\mathcal{Z}_1[\rho(p_S, \theta)]]|_{V_H}} + \mathcal{R}_{\max}^{1,l} \right]$$

Wir geben ein weiteres Mal die kumulativ-iterative Entwicklung der Werte MPDR Summe und RSDHR Summe und der auf die einzelnen Volumina bezogenen kumulativ-iterativen Benchmark-Werte zu bestimmten Iterationen an und diskutieren diese im Anschluss.

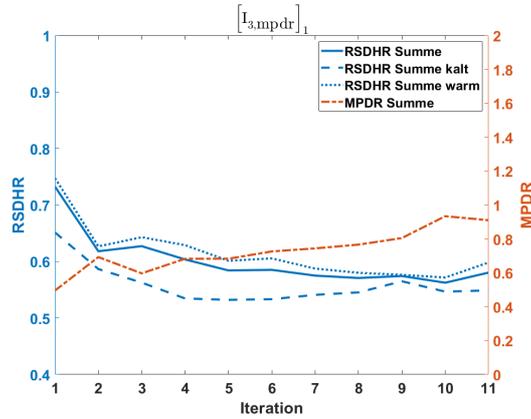


Abbildung 5.47: Kumulativ-iterative RSDHR- und MPDR-Werte des Steuerungsproblems  $[I_{3,mpdr}]_1$  für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ .

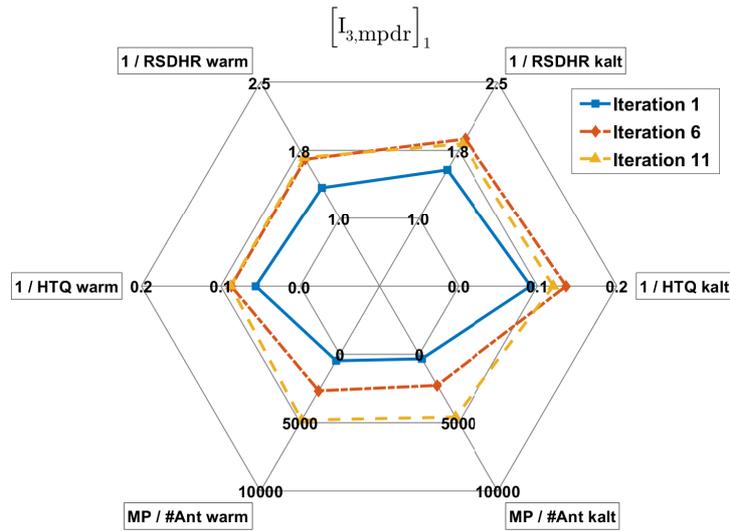


Abbildung 5.48: Kumulativ-iterative RSDHR-, HTQ und MP-Werte des Steuerungsproblems  $[I_{3,mpdr}]_1$  im warmen und kalten Gebiet für den Fall  $P_{in,m} = 1$ ,  $1 \leq m \leq 4$ .

Vergleichen wir die Abbildungen 5.45 und 5.47, dann sehen wir sofort, dass das Steuerungsproblem  $[I_{3,mpdr}]_1^l$  sowohl eine sehr gute Gleichmäßigkeit als auch einen relativ guten MPDR Summe Wert erreicht. Damit ist  $[I_{3,mpdr}]_1^l$  gerade eine Mischung aus RSDHR-Optimierung- und MPDR-Optimierung-Algorithmus und optimiert auf beide

Benchmark-Werte. Zudem sehen wir wieder sehr gut, dass die Verbesserung der Gleichmäßigkeit eine Verschlechterung bei der gezielten Erwärmung generiert. Betrachten wir dazu auch noch Abbildung 5.48, dann erkennen wir gut wie die Benchmark-Werte iterativ verbessert werden. Im Vergleich mit Abbildung 5.46 zeigt dies sehr gut auf, dass  $[I_{3,mpdr}]_1^l$  den anderen Algorithmen überlegen ist.

#### 5.5.4. Zusammenfassung

Wir fassen die Ergebnisse aus Abschnitt 5.5 zusammen. Die ersten wichtigen Erkenntnisse aus Abschnitt 5.5.1 zeigen den Vorteil in der Rechenzeit bei Verwendung der adjungierten Probleme zur Berechnung der Gâteaux-Ableitung. Zudem konnte herausgearbeitet werden, dass wegen der Berechnung der Gâteaux-Ableitungen die lokalen Optimierungsalgorithmen bei gleichwertigen Ergebnissen schneller als globale Algorithmen sind, wobei hierbei der SQP-Algorithmus die beste Performance zeigt. Bei dem Vergleich der iterativen Optimalsteuerungsprobleme  $[I_m]_k$ ,  $1 \leq m \leq 4$  und  $1 \leq k \leq 3$ , und  $[I_m^{rel}]_k$ ,  $1 \leq m \leq 4$  und  $1 \leq k \leq 3$ , konnten wir feststellen, dass die in Abschnitt 5.4 hergeleiteten Vergleichsfelder  $E_{V,\mathcal{N}}$ ,  $Q_{V,\mathcal{N}_L}$  für die Steuerungsprobleme  $[I_1]_k$ ,  $[I_1^{rel}]_k$ ,  $[I_2]_k$ ,  $[I_2^{rel}]_k$ , mit  $1 \leq k \leq 3$ , keine guten Ergebnisse liefern. Bei der Wahl  $E_{V,\mathcal{N}} = 0$ ,  $Q_{V,\mathcal{N}_L} = 0$  und der Betrachtung der Steuerungsprobleme  $[I_3]_k$ ,  $[I_3^{rel}]_k$ ,  $[I_4]_k$ ,  $[I_4^{rel}]_k$ , mit  $1 \leq k \leq 3$ , konnten wir jedoch einige interessante Ergebnisse herausarbeiten. Zum einen konnten wir sehen, dass durch Anregung von verschiedenen optimalen Steuerungen die Gleichmäßigkeit iterativ verbessert werden kann und damit eine Steuerung der elektromagnetischen Felder auch in der Anwendung Sinn macht, und zum anderen, dass die Geometrie und die Anzahl an Antennen eine untere Grenze an die Gleichmäßigkeit definieren. Diese Grenze hängt von den unterschiedlich anregbaren Moden innerhalb der Kavität und der Position des zu erwärmenden Objektes ab. Zudem wurde gezeigt, dass die hergeleiteten iterativen Steuerungsprobleme tendentiell eine bessere Gleichmäßigkeit bei mehr Energieeintrag liefert als die Vergleichsalgorithmen Random und SVD. Die Einschränkung auf  $P_{in,m} = 1$ ,  $1 \leq m \leq K_A$ , lieferte zum Teil bessere Ergebnisse als bei variabler Amplitude, was an dem kleineren Suchraum und der damit besseren Vergleichbarkeit der elektromagnetischen Felder liegt. Zum Abschluss des numerischen Kapitels haben wir noch einen Ausblick bezüglich der gezielten Erwärmung eines Objektes gemacht, falls mehrere Objekte innerhalb der Kavität vorhanden sind. Dabei konnten wir den Vorteil bei der Betrachtung von Optimalsteuerungsproblemen im Gegensatz zu den Vergleichsalgorithmen sehen, denn das Optimalsteuerungsproblem optimiert sowohl auf die Relation der durchschnittlich eingebrachten Leistung als auch auf die Gleichmäßigkeit.

## 6. Zusammenfassung und Ausblick

### 6.1. Zusammenfassung

In dieser Arbeit wurde die optimale Steuerung elektromagnetischer geschlossener Systeme mit koaxialer Anregung betrachtet. Hierfür wurden zunächst die benötigten partiellen Differentialgleichungen hergeleitet und die Existenz und Eindeutigkeit der dazugehörigen variationellen Formulierung bewiesen. Für die Evaluierung der optimalen Steuerungsprobleme wurde das diskretisierte System zudem zunächst mit der RBM reduziert, wobei die Diskretisierung mithilfe der Simulationssoftware COMSOL Multiphysics durchgeführt wurde.

Für die Reduktion der Lagrange-Systeme wurde die Theorie der RBM entsprechend erweitert. Dabei konnte numerisch gezeigt werden, dass die RBM bei Verwendung eines Greedy-Algorithmus eine exponentielle Konvergenz aufweist. Aufgrund der Betrachtung großdimensionaler Systeme und der damit einhergehenden Approximation der inf-sup-Konstante wurden verschiedene Approximationsvarianten getestet. Dabei konnte festgestellt werden, dass die Verwendung von Eigenwertinformationen zur Berechnung der approximativen inf-sup-Konstante, trotz einer besseren Konvergenz in der Approximationsgüte, keinen Vorteil im Reduktionsprozess der RBM für die in der Arbeit betrachteten Probleme liefert. Des Weiteren wurden unterschiedliche Projektionsvarianten im Reduktionsprozess betrachtet. Beim Vergleich des Fehlerschätzers und des echten relativen Fehlers bezüglich der Ritz-Galerkin- und der Petrov-Galerkin-Projektion konnte festgestellt werden, dass trotz eines schlechteren Fehlerschätzer-Wertes der echte relative Fehler einer Ritz-Galerkin-Projektion kleiner ist als der Fehler bei einer Petrov-Galerkin-Projektion.

Für die optimale Steuerung der elektromagnetischen Prozesse wurden als erstes verschiedene Kostenfunktionen hergeleitet und diskutiert. Neben dem Beweis der Existenz optimaler Steuerungen für die einzelnen Kostenfunktionale wurde für die Funktionale und das elektromagnetische Feld die starke Konvergenz bezüglich der Steuerparameter gezeigt. Aus der starken Konvergenz konnte daraufhin die Existenz der ersten Variation und der Gâteaux-Ableitung bewiesen werden. Zudem wurden die adjungierten Probleme hergeleitet, deren Lösung wiederum für die Berechnung der Gâteaux-Ableitungen der Kostenfunktionale verwendet werden kann. Erweitert wurde die Theorie noch durch die Fehlertheorie bezüglich der diskretisierten und reduzierten Steuerungsprobleme. Numerisch konnte dabei gezeigt werden, dass ein deutlicher Zeitvorteil in der Evaluierung der Gâteaux-Ableitung der Kostenfunktionale bei Verwendung adjungierter Lösungen erreicht werden kann. Die in der Theorie hergeleiteten Kostenfunktionale wurden daraufhin verwendet um Steuerungsprobleme für einen Erwärmungsprozess zu definieren. Diese Steuerungsprobleme wurden dann numerisch gegenüber entsprechenden Vergleichsalgorithmen und eigens dafür hergeleitete Benchmark-Werten getestet und miteinander verglichen. Hierbei wurde festgestellt, dass die Verwendung von Steuerungsproblemen

auf Basis vordefinierter elektromagnetische Felder  $E_V$  und Leistungsverteilungen  $Q_V$  stark von der Wahl dieser abhängig sind und keine guten Ergebnisse liefern. Die anderen Steuerungsprobleme, die hingegen den Energieeintrag maximieren, liefern meist bessere Ergebnisse gegenüber den in der Praxis verwendeten Algorithmen. Besonders bei der Betrachtung der gezielten Erwärmung in Abschnitt 5.5.3 ist der enorme Vorteil bei der Verwendung der optimalen Steuerungsprobleme erkennbar.

## 6.2. Ausblick

Neben den in dieser Arbeit verwendeten Modellen und den betrachteten numerischen Aspekten der Modellreduktion und Optimalsteuerung ergeben sich noch viele weitere Probleme und interessante Aufgaben basierend auf dieser Arbeit. Nachfolgend sind einige von diesen angegeben:

- Stabilität der Ritz-Galerkin-Projektion elektromagnetischer Systeme:

Wie in Abschnitt 4.3 erläutert, ist es für die in dieser Arbeit betrachteten Modelle aus numerischer Sicht möglich eine Ritz-Galerkin-Projektion durchzuführen. Für diese kann aber bisher keine Stabilität garantiert werden. Jedoch weist die Ritz-Galerkin-Projektion nach Abbildung 4.22 eine bessere Konvergenz im Reduktionsprozess und in der Approximationsgüte auf. Insofern die Stabilität der Ritz-Galerkin-Projektion gezeigt werden kann, wird demnach die Verwendung dieser präferiert.

- RB-Reduktion der Lagrange-Systeme mit  $u_{\mathcal{N},r}(p) \neq 0$ :

Mit der in Abschnitt 4.2 hergeleiteten Theorie, können partielle Differentialgleichungen mit parameterabhängigen Nebenbedingungen betrachtet werden. In den Modellen dieser Arbeit gilt stets  $u_{\mathcal{N},r}(p) = 0$ , weswegen die Betrachtung der RB-Reduktion der Lagrange-Systeme mit

$$u_{\mathcal{N},r}(p) = \sum_{k=1}^{Q_g} \Theta_k^g(p) u_{\mathcal{N},r,k} \neq 0$$

und die damit einhergehende Fehlertheorie in Abschnitt 4.2.2 hinsichtlich ihrer Funktionalität erst noch weiter betrachtet werden muss.

- Numerische Parameter- und Fehleranalyse der RBM:

Für den Greedy-Algorithmus 1 ist die Wahl der Parametermenge  $\mathbb{P}_{N_{RB}}$  hinsichtlich der Approximationsgüte des resultierenden Systems und der Rechendauer des Re-

duktionsprozesses von enormer Bedeutung. Um den Reduktionsprozess bei gleichbleibender Approximationsgüte für neue Modelle zu beschleunigen, muss die Testparametermenge  $\mathbb{P}_{NRB}$  klein und die in der Menge enthaltenen Parameter gut gewählt sein. Hierfür kann eine Parameteranalyse durchgeführt werden, um den Diskretisierungsgrad eines jeden einzelnen Parameters aus der Analyse abzuleiten. Eine weitere zu betrachtende Möglichkeit ist die adaptive Verfeinerung und Erweiterung der Parametermenge  $\mathbb{P}_{NRB}$ .

- Herleitung optimaler Vergleichsfunktionen  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$ :

Im Abschnitt 5.5 wurde aufgezeigt, dass die Steuerungsprobleme  $[I_1]_k^l$ ,  $[I_1^{rel}]_k^l$ ,  $[I_2]_k^l$  und  $[I_2^{el}]_k^l$ ,  $k = 1, 2, 3$ , stark abhängig von der Wahl der Vergleichsfunktionen  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  sind. Möglicherweise bieten diese bei der gezielten Erwärmung Vorteile, weshalb die Herleitung sinnvoller Vergleichsfunktionen  $E_{V,\mathcal{N}}$  und  $Q_{V,\mathcal{N}_L}$  von Interesse sein kann.

- Detaillierte Betrachtung der Optimierungsalgorithmen und derer Einstellungsoptionen zur Rechenzeitoptimierung:

Die Rechenzeiten der Steuerungsprobleme in Abschnitt 5.5.1 hinsichtlich der Optimierungsalgorithmen sind zum Teil sehr lang. Eine Analyse der Einstellmöglichkeiten bietet demnach eine Möglichkeit die Rechendauer bei gleichbleibenden Ergebnissen zu verbessern.

- Herleitung von besseren Steuerungsproblemen:

Werden die Rechenzeiten und die erhaltenen Benchmark-Werte der einzelnen Steuerungsprobleme aus Abschnitt 5.4 für die in der Arbeit betrachteten Modelle miteinander verglichen, dann ergeben sich zum Teil lange Rechenlaufzeiten. Diese entstehen aufgrund der komplexen Darstellungen der Kostenfunktionale der Steuerungsprobleme aus Abschnitt 5.4 und der damit einhergehenden aufwändigen Berechnung der Ableitungen. Hierbei kann jedoch nicht ausgeschlossen werden, dass bessere Steuerungsfunktionen existieren, die kürzere Rechenlaufzeiten bei vergleichbaren Ergebnissen in den Benchmark-Werten aufweisen.

- Betrachtung der geometrischen Abhängigkeit elektromagnetischer Felder, der reduzierten Modelle und der optimalen Steuerungen:

In der Realität können selbst kleine geometrische Veränderungen zum Teil großen Einfluss auf die elektromagnetischen Felder innerhalb einer Kavität haben. Um den Effekt der geometrischen Einflüsse auf die optimalen Steuerungen und den

dazugehörigen elektromagnetischen Felder zu untersuchen, bedarf es dabei ebenso der Betrachtung von geometrischen Parametern in den Modellen selbst und der Auswirkung auf die RBM.

- Analyse des Zusammenhangs optimaler Steuerungen und der S-Parameter:

Hinsichtlich der im vorherigen Punkt genannten starken Abhängigkeit gegenüber geometrischen Veränderungen ist vor allem die Analyse zwischen den optimalen Steuerungen und den S-Parametern von großem Interesse. Der Grund hierfür ist, dass die Messung der S-Parameter relativ einfach ist und in der Praxis lediglich ein paar Millisekunden benötigt. Wäre es also möglich einen Zusammenhang zwischen den optimalen Steuerungen, die die besten Benchmark-Werte erreichen, und den Werten der S-Parameter, oder daraus abgeleiteten Eigenschaften, herzuleiten, dann ließe sich ein äußerst effektiver und schneller Auswahlalgorithmus entwickeln, der in Echtzeit die optimalen Steuerungen bestimmt.

## A. Anhang

### A.1. Funktionenräume und Operatoren

Die nachfolgenden Definitionen und Beweise finden sich in Büchern der Analysis, wie beispielsweise [39]. Im Folgenden sei  $\Omega \subset \mathbb{R}^d$  und die Funktionenräume bilden nach  $\mathbb{C}$  ab, weswegen dies nicht explizit angegeben wird.

**Definition A.1.** Wir definieren die normierten Räume:

- $(C^0(\Omega), \|\cdot\|_{C^0})$  der Raum der stetigen Funktionen mit  $\|f\|_{C^0} = \max_{x \in \Omega} |f(x)|$ .
- $(C^k(\Omega), \|\cdot\|_{C^k})$  der Raum der stetig differenzierbaren Funktionen bis zum Grad  $k$  mit  $\|f\|_{C^k} = \max_{|s| \leq k} \|\partial^s f(x)\|_{C^0}$ .
- $(C^{k,\alpha}(\Omega), \|\cdot\|_{C^{k,\alpha}})$  der Raum der  $C^k(\Omega)$ -Funktionen, deren Ableitungen der Ordnung  $k$  Hölderstetig sind. In diesem Raum liegen alle Funktionen  $f \in C^k(\Omega)$  mit

$$\|f\|_{C^{k,\alpha}} = \sum_{|s| \leq k} \|\partial^s f\|_{C^0} + \sum_{|s|=k} [\partial^s f]_{C^{0,\alpha}} \leq C,$$

mit der Hölder-Halbnorm

$$[\partial^s f]_{C^{0,\alpha}} = \sup_{x,y \in \Omega, x \neq y} \left\{ \frac{|\partial^s f(x) - \partial^s f(y)|}{|x - y|^\alpha} \right\}.$$

- $(L^p(\Omega), \|\cdot\|_{L^p})$  der Raum der Lebesgue-integrierbaren Funktionen mit  $p \in [1, \infty)$  und  $\|f\|_{L^p(\Omega)}^p = \int_{\Omega} |f(x)|^p dx$ . Es gilt die Äquivalenzrelation, dass zwei Funktionen  $u, v \in L^p(\Omega)$  gleich sind, falls  $u = v$  für alle  $K \subset \Omega$  mit  $\mathcal{L}^n(\Omega \setminus K) = 0$  bezüglich des Lebesguemaßes. Für lokal integrierbare Funktionen  $f \in L^p_{loc}(\Omega)$  soll

$$\int_K |f(x)|^p dx \leq C_K$$

für alle kompakten Teilmengen  $K \subset \Omega$  gelten. Auf dem Raum  $L^2(\Omega)$  kann zudem das Skalarprodukt

$$(u, v)_{L^2(\Omega)} = \int_{\Omega} u \cdot v \, d\Omega$$

definiert werden.

- $(L^\infty(\Omega), \|\cdot\|_{L^\infty})$  der Raum der essentiell beschränkten Funktionen mit

$$\|f\|_{L^\infty(\Omega)} = \operatorname{ess\,sup}_{\Omega} |f(x)| = \inf_{N \subset \Omega, \mathcal{L}^n(N)=0} \sup_{\Omega \setminus N} |f(x)|.$$

Wie bei den  $L^p(\Omega)$ -Funktionen ist auch auf  $L^\infty(\Omega)$  eine analoge Äquivalenzrelation definiert.

Jeder dieser Räume ist mit seiner Norm ein Banachraum. Der Raum  $L^2(\Omega)$  mit seinem Skalarprodukt ist sogar ein Hilbertraum. In Hilberträumen kann wegen der Existenz eines Skalarproduktes zudem das orthogonale Komplement definiert werden.

**Definition A.2.** Sei  $(X, (\cdot, \cdot)_X)$  Hilbertraum und  $M \subset X$ . Das *orthogonale Komplement*  $M^\perp$  von  $M$  wird definiert durch

$$M^\perp = \{u \in X : (u, v)_X = 0 \text{ für alle } v \in M\}.$$

Bei der Betrachtung der variationellen Formulierung partieller Differentialgleichungen spielt der folgende Raum eine entscheidende Rolle.

**Definition A.3.** Die Menge  $\mathcal{D}(\Omega) = C_0^\infty(\Omega)$  der unendlich oft stetig differenzierbaren Funktionen mit kompaktem Träger in  $\Omega$  wird als die *Menge der Testfunktionen* bezeichnet.

In der Literatur wird für  $C_0^\infty(\Omega)$  auch oft  $C_c^\infty(\Omega)$  verwendet.

**Definition A.4.** Seien  $u, v \in L_{loc}^1(\Omega)$ ,  $\alpha$  Multiindex mit  $|\alpha| \leq k$ . Die Funktion  $v$  heißt  $\alpha$ -te schwache Ableitung von  $u$ , wenn gilt

$$\int_{\Omega} u \cdot \partial^\alpha \varphi \, d\Omega = (-1)^{|\alpha|} \int_{\Omega} v \cdot \varphi \, d\Omega, \quad \varphi \in C_0^\infty(\Omega).$$

**Definition A.5** (Sobolev-Raum  $W^{k,p}(\Omega)$ ). Eine Funktion  $u \in L^p(\Omega)$  liegt im Raum  $W^{k,p}(\Omega)$ , falls alle schwachen Ableitungen  $\partial^s u$  für  $|s| \leq k$  existieren und  $\partial^s u \in L^p(\Omega)$  für alle Multiindizes mit  $|s| \leq k$ . Die Norm ist gegeben durch

$$\|u\|_{W^{k,p}(\Omega)} = \begin{cases} \left( \sum_{|s| \leq k} \|\partial^s u\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \sum_{|s| \leq k} \|\partial^s u\|_{L^\infty(\Omega)}, & p = \infty. \end{cases}$$

Für  $p = 2$  schreibt man auch  $W^{k,2}(\Omega) = H^k(\Omega)$ . Für diesen Raum existiert das Skalarprodukt

$$(u, v)_{H^k(\Omega)} = \sum_{|s| \leq k} (\partial^s u, \partial^s v)_{L^2(\Omega)}.$$

Der Sobolev-Raum  $W^{k,p}(\Omega)$  ist mit seiner Norm ein Banachraum und  $H^k(\Omega)$  mit dem Skalarprodukt ein Hilbertraum. Für Abbildungen zwischen den Funktionenräumen definieren wir Operatoren.

**Definition A.6.** Seien  $X, Y$  normierte Räume. Sei  $V \subset X$ . Abbildungen  $T : V \rightarrow Y$  heißen *Operatoren*. Ein Operator heißt *linearer Operator*, wenn

$$T[u + \lambda v] = T[u] + \lambda T[v]$$

für alle  $\lambda \in \mathbb{K}$  und alle  $u, v \in V$  gilt. Ein Operator heißt *stetig* in  $v \in V$ , wenn für alle konvergenten Folgen  $(v_n), v_n \in V$ , mit  $\lim_{n \rightarrow \infty} v_n = v$  gerade

$$\lim_{n \rightarrow \infty} T[v_n] = T[v]$$

gilt. Die Menge aller linearen Operatoren wird mit  $L(X, Y)$  bezeichnet. Ist  $Y = \mathbb{K}$  der zugrundeliegende Zahlenkörper  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ , so spricht man vom *Dualraum*

$$X' = L(X, \mathbb{K}).$$

Elemente im Dualraum werden *Funktionale* genannt.

Mit der Operatornorm

$$\|T\| = \inf \{C > 0 \mid \|T[u]\|_Y \leq C\|u\|_X\},$$

oder der äquivalenten Formulierung

$$\|T\| = \sup_{\substack{u \in X \\ u \neq 0}} \frac{\|T[u]\|_Y}{\|u\|_X},$$

werden  $L(X, Y)$  und der Dualraum  $X'$  zu normierten Räumen. Es ist weiterhin bekannt, dass für lineare Operatoren die Stetigkeit äquivalent zur Beschränktheit

$$\exists C > 0 \text{ mit } \|T[v]\|_Y \leq C\|v\|_X, \quad v \in X,$$

ist. Des Weiteren wird ein Operator  $T \in L(X, Y)$  *Isomorphismus* genannt, wenn  $T$  invertierbar und  $T^{-1}$  stetig ist. Gilt zusätzlich  $\|T[v]\|_Y = \|v\|_X$  wird der Isomorphismus *isometrisch* genannt. Für die Existenz von optimalen Lösungen geben wir die Definition der Unterhalbstetigkeit an

**Definition A.7.** Sei  $S \subset X$  Teilmenge eines normierten Raumes  $X$  und  $f : S \rightarrow \mathbb{R}$  ein Funktional. Das Funktional  $f$  heißt *schwach unterhalbstetig* wenn für jede schwach konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $S$  mit Grenzwert  $\hat{x}$  gilt, dass

$$\liminf_{n \rightarrow \infty} f(x_n) \geq f(\hat{x}).$$

Die Funktionenräume aus Definition A.1 und Definition A.5 lassen sich ebenso auch für Gebiete  $\Gamma \subset \mathbb{R}^d$  niedriger Dimension, beispielsweise Flächen im dreidimensionalen

Raum, definieren. Während für  $v \in C^0(\Omega)$  Auswertungen auf Oberflächen ohne Probleme möglich ist, ist dies für  $v \in L^p(\Omega)$  nicht klar, denn für das Lebesgue-Maß von niederdimensionalen Objekten gilt

$$\mathcal{L}^n(\Gamma) = 0.$$

Für Sobolev-Räume bestimmter Ordnung können wir hingegen Spuroperatoren von Gebieten auf dessen Rand definieren, insofern das Gebiet bestimmte Eigenschaften besitzt.

**Satz A.8.** *Sei  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 2$ , mit Lipschitz-Rand  $\Gamma$ . Dann besitzen die Funktionen  $u \in W^{1,p}(\Omega)$ ,  $p \in [1, \infty)$ , Randwerte in folgendem Sinne. Es gibt genau einen linearen stetigen Operator  $\gamma : W^{1,p}(\Omega) \rightarrow L^p(\Gamma)$ , der für Funktionen  $u \in C(\bar{\Omega}) \cap W^{1,p}(\Omega)$  die klassischen Randwerte  $\gamma[u](x) = u(x)|_{x \in \Gamma}$  liefert.*

Für weitere Eigenschaften des Spuroperators  $\gamma$ , führen wir die Sobolev-Slobodeckij-Räume ein.

**Definition A.9** (Sobolev-Slobodeckij-Raum  $H^s(\Omega)$ ). Sei  $s \geq 0$  mit  $s = k + t$ ,  $k \in \mathbb{N}_0$ ,  $0 < t < 1$ . Der Raum

$$H^s(\Omega) = \left\{ v \in H^k(\Omega) : [\partial^\alpha v]_{H^t(\Omega)} < \infty, |\alpha| = k \right\}$$

besitzt die Halbnorm

$$[w]_{H^t(\Omega)} = \int_{\Omega} \int_{\Omega} \frac{(w(x) - w(y))^2}{\|x - y\|^{d+2t}} d\Omega d\Omega.$$

Mit Definition A.9 erhalten wir folgenden Spursatz.

**Satz A.10.** *Sei  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 2$ , mit Lipschitz-Rand  $\Gamma$ . Dann besitzen die Funktionen  $u \in H^1(\Omega)$  Randwerte in folgendem Sinne. Es gibt genau einen linearen stetigen Operator  $\gamma : H^1(\Omega) \rightarrow H^{\frac{1}{2}}(\Gamma)$ , der für Funktionen  $u \in C(\bar{\Omega}) \cup H^1(\Omega)$  die klassischen Randwerte  $\gamma[u](x) = u(x)|_{x \in \Gamma}$  liefert. Zudem gilt, dass der Spuroperator  $\gamma$  surjektiv ist und damit existiert eine rechtsinverse  $\gamma^{-1}$ .*

Auch für  $s > \frac{1}{2}$  kann die Existenz eines linearen stetigen Spuroperators

$$\gamma : H^s(\Omega) \rightarrow H^{s-\frac{1}{2}}(\Gamma)$$

nachgewiesen werden.

## A.2. Einbettungssätze

Zunächst geben wir die Definition der Einbettung und einige allgemeine Resultate zu Sobolev-Einbettungen an, siehe [39].

**Definition A.11.** Seien  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$  Banachräume. Existiert eine lineare injektive Abbildung  $j : X \rightarrow Y$ , so heißt  $X$  in  $Y$  *eingebettet*. Ist die Abbildung  $j$  zudem stetig, so heißt  $X$  in  $Y$  *stetig eingebettet* und wir schreiben  $X \hookrightarrow Y$ . Ist die Abbildung zudem kompakt, d.h. sie ist stetig und beschränkte Teilmengen von  $X$  werden auf relativ kompakte Teilmengen von  $Y$  abgebildet, so heißt  $X$  in  $Y$  *kompakt eingebettet*. Wir schreiben dafür  $X \hookrightarrow_C Y$ .

Zwei wichtige Einbettungsergebnisse beziehen sich auf die Sobolev-Räume  $W^{k,p}(\Omega)$  und deren Abhängigkeiten zueinander für unterschiedliche Werte  $k$ ,  $p$  und können im Detail in [39] nachgelesen werden.

**Satz A.12.** Sei  $\Omega \subset \mathbb{R}^d$  beschränktes Gebiet mit Lipschitz-Rand  $\Gamma$ . Seien  $k, l \in \mathbb{N}$  und  $p \in \mathbb{R}$ ,  $1 \leq p \leq \infty$ . Des Weiteren bezeichne  $\Omega^m$  den Schnitt einer  $m$ -dimensionalen Hyperfläche mit  $\Omega$ . Dann gibt es die folgenden Einbettungen:

1. Sei  $lp < d$  und  $d - lp < m \leq d$ , dann gilt

$$W^{k+l,p}(\Omega) \hookrightarrow W^{k,q}(\Omega^m), \quad p \leq q \leq \frac{mp}{d - lp}.$$

2. Sei  $lp = d$  und  $1 \leq m \leq d$ , dann gilt für  $p \leq q < \infty$

$$W^{k+l,p}(\Omega) \hookrightarrow W^{k,q}(\Omega^m).$$

3. Sei  $lp > d \geq (l - 1)p$ , dann gilt

$$W^{k+l,p}(\Omega) \hookrightarrow C^k(\overline{\Omega}).$$

**Satz A.13.** Sei  $\Omega \subset \mathbb{R}^d$  ein beschränktes Gebiet mit Lipschitz-Rand  $\Gamma$  und  $\Omega_0 \subseteq \Omega$ . Seien  $k, l \in \mathbb{N}$  mit  $k \geq 0$ ,  $l \geq 1$  und  $p \in \mathbb{R}$ ,  $1 \leq p < \infty$ . Des Weiteren bezeichne  $\Omega_0^m$  den Schnitt einer  $m$ -dimensionalen Hyperfläche mit  $\Omega_0$ . Dann sind die folgenden Einbettungen kompakt:

1. Sei  $lp \leq d$  und  $0 < d - lp < m \leq d$ , dann gilt

$$W^{k+l,p}(\Omega) \hookrightarrow_C W^{k,q}(\Omega_0^m), \quad 1 \leq q \leq \frac{mp}{d - lp}.$$

2. Sei  $lp = d$  und  $1 \leq m \leq d$ , dann gilt

$$W^{k+l,p}(\Omega) \hookrightarrow_C W^{k,q}(\Omega_0^m), \quad 1 \leq q < \infty.$$

3. Sei  $lp > d$ , dann gilt

$$W^{k+l,p}(\Omega) \hookrightarrow_C C^k(\overline{\Omega_0}).$$

Auch für die in Abschnitt 2.2 hergeleiteten Räume  $H_0(\text{curl}; \Omega)$  existieren wichtige Einbettungsergebnisse deren Beweise in [32, 38] nachzulesen sind.

**Lemma A.14.** *Sei  $\Omega \subset \mathbb{R}^3$  ein polyedrisches Gebiet mit Lipschitz-Rand  $\Gamma$ , dann ist*

$$\{u \in H_0(\text{curl}; \Omega) : \nabla \cdot u = 0\} \hookrightarrow [L^2(\Omega)]^3$$

*stetig und kompakt.*

**Lemma A.15.** *Sei  $\Omega \subset \mathbb{R}^3$  ein polyedrisches Gebiet mit Lipschitz-Rand  $\Gamma$ , dann existiert ein  $s > \frac{1}{2}$ , sodass*

$$\{u \in H_0(\text{curl}; \Omega) : \nabla \cdot u = 0\} \hookrightarrow [H^s(\Omega)]^3$$

*stetig ist.*

### A.3. Existenztheorie

Wir geben hier wichtige Resultate für die Existenz von Lösungen an, welche im Detail in [32, 39] nachgelesen werden können.

**Satz A.16** (Rieszscher Darstellungssatz). *Sei  $(X, (\cdot, \cdot)_X)$  ein Hilbertraum über  $\mathbb{C}$ . Dann gibt es zu jedem Funktional  $u' \in X'$  genau ein  $u \in X$  mit*

$$u'(v) = (v, u)_X, \quad v \in X.$$

*Zusätzlich gilt  $\|u\|_X = \|u'\|_{X'}$ . Mit anderen Worten bedeutet dies, dass die Abbildung  $\phi : X \rightarrow X', u \mapsto (\cdot, u)_X$  bijektiv, isometrisch und isomorph ist.*

Sei  $u \in X$  und wir betrachten die Variationsformulierung

$$a(u, v) = f(v), \quad v \in X, \tag{A.1}$$

mit Sesquilinearform  $a$  und Linearform  $f \in X'$ . Die Sesquilinearform  $a$  ist *stetig*, wenn eine Konstante  $\gamma > 0$  existiert, sodass

$$|a(u, v)| \leq \gamma \|u\|_X \|v\|_X, \quad u, v \in X,$$

gilt. Existiert eine Konstante  $\alpha > 0$ , die Koerzivitätskonstante, sodass

$$\alpha \|u\|_X^2 \leq |a(u, u)|, \quad u \in X,$$

gilt, dann ist die Sesquilinearform  $a$  *koerziv*. Ist  $X$  ein Hilbertraum, dann lässt sich folgendes Satz beweisen.

**Satz A.17** (Lax-Milgram). Sei  $(X, (\cdot, \cdot)_X)$  ein Hilbertraum,  $a : X \times X \rightarrow \mathbb{C}$  eine stetige koerzive Sesquilinearform und  $f \in X'$ . Dann existiert genau ein  $u \in X$ , sodass

$$a(u, v) = f(v), \quad v \in X, \quad (\text{A.2})$$

gilt. Des Weiteren folgt

$$\|u\|_X \leq \frac{\gamma}{\alpha} \|f\|_{X'},$$

mit Stetigkeitskonstante  $\gamma$  und Koerzivitatskonstante  $\alpha$ .

Mit Satz A.17 erhalten wir die Existenz und Eindeutigkeit von (A.1) fur stetige koerzive Sesquilinearformen. Eine Verallgemeinerung des Satzes von Lax-Milgram ist durch die Babuška-Brezzi-Theorie gegeben. Hierfur wird eine Konstante

$$\beta = \inf_{u \in X} \sup_{v \in X} \frac{|a(u, v)|}{\|u\|_X \|v\|_X} > 0 \quad (\text{A.3})$$

definiert, welche als inf-sup-Konstante bezeichnet wird. Fur die numerische Betrachtung partieller Differentialgleichungen werden Ansatzraume  $X$  und Testraume  $Y$  verwendet, fur die nicht notwendigerweise  $X = Y$  gilt. Deswegen definieren wir die inf-sup-Konstante auch fur Sesquilinearformen  $a : X \times Y \mapsto \mathbb{C}$

$$\beta = \inf_{u \in X} \sup_{v \in Y} \frac{|a(u, v)|}{\|u\|_X \|v\|_Y} > 0. \quad (\text{A.4})$$

Falls so eine Konstante existiert, dann heit die Sesquilinearform  $a$  *inf-sup-stabil*. Wir geben nun die verallgemeinerte Form von Satz A.17 an.

**Satz A.18** (Verallgemeinerter Lax-Milgram-Satz). Sei  $a : X \times Y \rightarrow \mathbb{C}$  eine stetige und inf-sup-stabile Sesquilinearform. Dann existiert fur jedes  $f \in Y'$  eine eindeutige Losung  $u \in X$ , sodass gilt

$$a(u, v) = f(v), \quad v \in Y,$$

mit

$$\|u\|_X \leq \frac{\gamma}{\beta} \|f\|_{Y'}.$$

Wir geben noch einen Spezialfall fur zwei Sesquilinearformen an. Hierfur seien  $(X, (\cdot, \cdot)_X)$  und  $(Y, (\cdot, \cdot)_Y)$  zwei Hilbertraume und

$$a : X \times X \rightarrow \mathbb{C}, \quad b : X \times Y \rightarrow \mathbb{C}$$

beschrankte Sesquilinearformen, also

$$\begin{aligned} |a(u, v)| &\leq C \|u\|_X \|v\|_X, & u, v \in X, \\ |b(u, v)| &\leq C \|u\|_X \|v\|_Y, & u \in X, v \in Y. \end{aligned}$$

**Definition A.19.** Sei

$$Z = \{u \in X : b(u, v) = 0, v \in Y\}.$$

Die Sesquilinearform  $a$  heißt  $Z$ -koerziv, falls eine Konstante  $\alpha > 0$  existiert, sodass

$$|a(u, u)| \geq \alpha \|u\|_X^2, \quad u \in Z. \quad (\text{A.5})$$

**Definition A.20.** Die Sesquilinearform  $b$  erfüllt die *Babuška-Brezzi-Bedingung*, falls eine Konstante  $\beta > 0$  existiert, sodass

$$\beta \|z\|_Y \leq \sup_{u \in X} \frac{|b(u, z)|}{\|u\|}, \quad z \in Y, \quad (\text{A.6})$$

wobei  $\beta$  unabhängig von  $z$  ist.

Mit diesen beiden Definitionen ist der nächste Satz eine Folgerung aus Satz A.18, siehe [32].

**Satz A.21.** Seien  $(X, (\cdot, \cdot)_X)$  und  $(Y, (\cdot, \cdot)_Y)$  Hilberträume und  $a : X \times X \rightarrow \mathbb{C}$  und  $b : X \times Y \rightarrow \mathbb{C}$  beschränkte Sesquilinearformen. Des Weiteren ist  $a$   $Z$ -koerziv und  $b$  erfüllt die Babuška-Brezzi-Bedingung. Sei  $f \in X'$  und  $g \in Y'$ . Dann besitzt das Variationsproblem

$$\begin{aligned} a(u, v) + b(v, z) &= f(v), & v \in X, \\ b(u, w) &= g(w), & w \in Y, \end{aligned}$$

eine eindeutige Lösung  $(u, z) \in X \times Y$  und es gilt

$$\|u\|_X + \|z\|_Y \leq C(\|f\|_{X'} + \|g\|_{Y'}). \quad (\text{A.7})$$

## Literatur

- [1] P. Lechner. *Untersuchungen zum mikrowellenunterstützten Aushärten von vernetzten Harzsystemen auf Polyurethan- und Kohlenwasserstoffbasis*. PhD thesis, Technische Universität Darmstadt, 2012.
- [2] P. Benner, V. Mehrmann, and D. C. Sorensen. *Dimension Reduction of Large-Scale Systems*, volume 45. Springer-Verlag, Berlin Heidelberg, 2005.
- [3] W. H. Schilders and H. A. van der Vorst and J. Rommes. *Model Order Reduction: Theory, Research Aspects and Applications*, volume 13. Springer-Verlag, Berlin Heidelberg, 2008.
- [4] A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. *Comput. Methods Appl. Mech. Eng.*, 34:955–985, 1980.
- [5] A. K. Noor. Recent advances in reduction methods for nonlinear problems. *Comput. Struct.*, 13:31–44, 1981.
- [6] A. K. Noor. On making large scale nonlinear problems small. *Comput. Methods Appl. Mech. Eng.*, 34:955–985, 1982.
- [7] A. K. Noor, C. D. Balch, and M. A. Shibus. Reduction methods for nonlinear steady-state thermal analysis. *Int. J. Numer. Methods Eng.*, 20:1323–1348, 1983.
- [8] J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM J. Sci. Stat. Comput.*, 10:777–786, 1989.
- [9] M.-Y. L. Lee. Estimation of the error in the reduced basis method solution of differential algebraic equation systems. *SIAM J. Numer. Anal.*, 28:512–528, 1991.
- [10] C. Prudhomme, D. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parameterized partial differential equations: reduced-basis output bound methods. *J. Fluids Eng.*, 124:70–80, 2002.
- [11] C. Prudhomme, D. Rovas, K. Veroy, and A. T. Patera. A posteriori error bounds for reduced-basis approximation of parameterized noncoercive and nonlinear elliptic partial differential equations. In *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, 2003.
- [12] Y. Maday, A. T. Patera, and D. V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. *Stud. in Math. and its Appl.*, 31, 2002.
- [13] G. Rozza. Reduced-basis methods for elliptic equations in sub-domains with a posteriori error bounds and adaptivity. *Appl. Numer. Math.*, 55:403–424, 2005.

- 
- [14] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified Reduced Basis Methods for Parameterized Partial Differential Equations*. Springer-Verlag, Heidelberg, 2016.
- [15] Y. Chen, J. S. Hesthaven, Y. Maday, and J. Rodríguez. Certified reduced basis methods and output bounds for the harmonic Maxwell's equations. *SIAM J. Sci. Comput.*, 32(2):970–996, 2010.
- [16] J. Pomplun. *Reduced Basis method for electromagnetic scattering problems*. PhD thesis, Free University Berlin, 2010.
- [17] M. W. Hess and P. Benner. Fast evaluation of time-harmonic Maxwell's equations using the reduced basis method. *IEEE Trans. Microw. Theory Tech.*, 2012.
- [18] M. W. Hess and P. Benner. A reduced basis method for microwave semiconductor devices with geometrical variations. *Int. J. Comput. Math. Electrical Electronic Eng.*, 33(4):1071–1081, 2014.
- [19] M. W. Hess. *Reduced Basis Approximations for Electromagnetic Applications*. PhD thesis, Otto-von-Guericke-Universität Magdeburg, 2016.
- [20] G. Rozza. *Shape Design by Optimal Flow Control and Reduced Basis Techniques: Applications to Bypass Configurations in Haemodynamics*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2016.
- [21] T. Tonn, K. Urban, and S. Volkwein. Optimal control of parameter-dependent convection-diffusion problems around rigid bodies. *SIAM J. Sci. Comput.*, 32:1237–1260, 2010.
- [22] T. Tonn, K. Urban, and S. Volkwein. Comparison of the reduced-basis and POD: a posteriori error estimators for an elliptic linear-quadratic optimal control problem. *Math. Comput. Mod. Dyn. Systems*, 17:355–369, 2011.
- [23] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni. Reduced basis method for parametrized elliptic optimal control problems. *SIAM J. Sci. Comput.*, 35:2316–2340, 2013.
- [24] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko. *The Mathematical Theory of Optimal Processes*. Wiley, New York, 1962.
- [25] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag, Berlin, 1971.
- [26] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer-Verlag, Berlin, 2009.

- [27] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg+Teubner, Wiesbaden, 2009.
- [28] G. N. Gatica and S. Meddahi. Finite element analysis of a time harmonic Maxwell problem with an impedance boundary condition. *IMA J. Numer. Anal.*, 32, 2011.
- [29] J. C. Maxwell. A Dynamical Theory of the electromagnetic Field. *Phil. Trans. R. Soc. Lond.*, 155:459–512, 1865.
- [30] K. Jänich. *Vektoranalysis*. Springer-Verlag, 2005.
- [31] J. Billingham and A. C. King. *Wave Motion*. Cambridge University Press, Cambridge, 2006.
- [32] P. Monk. *Finite Element Methods for Maxwell's Equation*. Clarendon Press, Oxford, 2003.
- [33] A. Buffa and P. Ciarlet. On traces for functional spaces related to Maxwell's equations Part I: An integration by parts formula in Lipschitz polyhedra. *Math. Methods Appl. Sci.*, 24:9–30, 2001.
- [34] A. Buffa, M. Costabel, and D. Sheen. On traces for  $H(\text{curl}, \Omega)$  in Lipschitz domains. *Math. Anal. Appl.*, 276:845–867, 2002.
- [35] J.D. Jackson. *Klassische Elektrodynamik*. De Gruyter, San Francisco, 4 edition, 2006.
- [36] T. B. A. Senior. Approximate Boundary Conditions, Part I. Technical report, University of Michigan, 1990.
- [37] D. M. Pozar. *Microwave Engineering*. John Wiley & Sons, Inc., Hoboken, 2011.
- [38] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault. Vector Potentials in Three-Dimensional Nonsmooth Domains. *Math. Methods Appl. Sci.*, 21:823–864, 1998.
- [39] D. Werner. *Funktionalanalysis*. Springer-Verlag, Berlin Heidelberg, 2007.
- [40] J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numer. Math.*, 94:195–202, 2003.
- [41] R. Hiptmaier. Coupling of finite elements and boundary elements in electromagnetic scattering. *SIAM J. Math. Anal.*, 41:919–944, 2003.
- [42] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159. Springer-Verlag, New York, 2004.

- 
- [43] A. Schneebeli. An  $H(\text{curl};\Omega)$ -conforming FEM: Nédélec's elements of first type. Technical report, University of Trier, 2003.
- [44] P. Jacobsson. Nédélec elements for computational electromagnetics. In *Inproceeding*, 2007.
- [45] A. Alonso and A. Valli. An optimal domain decomposition preconditioner for low-frequency time-harmonic Maxwell equations. *Math. Comput.*, 68:607–631, 1999.
- [46] J.-M. Jin. *The Finite Element Method in Electromagnetics*. Wiley, Hoboken, 2014.
- [47] A. Buffa and R. Hiptmaier. Galerkin boundary element methods for electromagnetic scattering. *Lecture Notes Comput. Sci. Eng.*, 31:83–124, 2003.
- [48] A. T. Patera and G. Rozza. *Reduced Basis Approximation and A Posteriori error Estimation for parameterized Partial Differential Equations*. MIT Pappalardo Graduate Monographs in Mechanical Engineering, 2007.
- [49] Y. Chen, J. S. Hesthaven, and Y. Maday. *A seamless reduced Basis Element Method for 2D Maxwell's Problem: An Introduction*, volume 76. Springer-Verlag, 2011.
- [50] P. Binev, A. Cohen, W. Dahmen, R. Devore, P. Guergana, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *SIAM J. Math. Anal.*, 43:1457–1472, 2010.
- [51] R. Schaback and H. Wendland. *Numerische Mathematik*. Springer, 2006.
- [52] O. M. Bucci, C. D'Elia, G. Mazzarella, and G. Panariello. Antenna pattern synthesis: a new general approach. *Proceedings of the IEEE*, 82:358–371, 1994.
- [53] T. Köhler, P. Maass, P. Wurst, and M. Seebass. A fast algorithm to find optimal controls of multi-antenna applicators in regional hyperthermia. *Phys. Med. Bio.*, 46:2503, 2001.
- [54] B. Fuchs, A. Skrivervik, and J. R. Mosig. Shaped beam synthesis of arrays via sequential convex optimizations. *IEEE Antennas Wireless Propagation Letters*, 12:1049–1052, 2013.
- [55] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York, 1994.
- [56] R. A. M. Canters. *Optimization and Control in Deep Hyperthermia*. PhD thesis, Erasmus Universität Rotterdam, 2013.
- [57] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer-Verlag, New York, 2006.

- [58] C. Geiger and C. Kanzow. *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer-Verlag, Berlin Heidelberg, 1999.
- [59] R. Hooke and T.A. Jeeves. Direct Search Solution of numerical and statistical Problems. *J. of the Ass. for Comput. Mach.*, 8(2):212–229, 1961.
- [60] W. Banzhaf, P. Nordin, R.E. Keller, and F.D. Francone. *Genetic Programming An Introduction*. Morgan Kaufmann Publishers Inc., San Francisco, 1998.
- [61] J. Kennedy and R. Eberhart. Particle Swarm Optimization. *Proceedings of ICNN'95*, 1995.
- [62] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *SCIENCE*, 220(4598):671–680, 1983.
- [63] B. Schweizer. *Partielle Differentialgleichungen*. Springer-Verlag, Berlin Heidelberg, 2013.
- [64] C. Eck, H. Garcke, and P. Knabner. *Mathematische Modellierung*. Springer-Verlag, Berlin Heidelberg, 2011.
- [65] J. Jahn. *Introduction to the Theory of Nonlinear Optimization*. Springer-Verlag, Berlin Heidelberg, 2007.
- [66] J. L. Volakis, A. Chatterjee, and L. C. Kempel. *Finite Element Method for Electromagnetics*. IEEE Press, New York, 1998.
- [67] M. Ohlberger and S. Rave. Reduced basis method: success, limitations and future challenges. In *Proceeding of ALGORITHMY 2016*, pages 1–12, 2016.
- [68] T. A. Porsching and M.-Y. L. Lee. The reduced basis method for initial value problems. *SIAM J. Numer. Anal.*, 24:1277–1287, 1987.
- [69] Y. Chen, J. S. Hesthaven, Y. Maday, and J. Rodríguez. Improved successive constraint method based a posteriori error estimate for reduced basis approximation. *ESAIM: M2AN*, 43:1099–1116, 2009.
- [70] M. E. Rognes, R. C. Kirby, and A. Logg. Efficient assembly of  $H(\text{div})$  and  $H(\text{curl})$  conforming finite elements. *SIAM J. Sci. Comput.*, 31(6):4130–4151, 2009.
- [71] Y. Maday and E. M. Rønquist. A reduced-basis element method. *Numer. Anal.*, 335:195–200, 2002.
- [72] R. Hiptmaier. Finite elements in computational electromagnetism. *Acta Numerica*, 11:237–339, 2002.

- [73] K. Stavrakakis. *Model Order Reduction Methods for Parameterized Systems in Electromagnetic Field Simulations*. PhD thesis, Technische Universität Darmstadt, 2012.