

Lehrstuhl für Entwurfsautomatisierung der
Technischen Universität München

Zur Dimensionierung analoger integrierter Schaltungen unter Berücksichtigung struktureller Nebenbedingungen

Robert Schwencker

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. techn. Peter Russer

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Kurt J. Antreich
2. Univ.-Prof. Dr.-Ing. habil. Robert Weigl, Friedrich-Alexander-Universität
Erlangen-Nürnberg

Die Dissertation wurde am 13.12.2001 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 26.06.2002 angenommen.

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als freier Mitarbeiter der Firma Infineon Technologies AG, München und als wissenschaftlicher Mitarbeiter am Lehrstuhl für Entwurfsautomatisierung der Technischen Universität München.

Herrn Prof. Dr.-Ing. Kurt Antreich danke ich sehr herzlich für die Möglichkeit zur Durchführung dieser Arbeit, für seine Unterstützung sowie für die hilfreichen Ratschläge und Diskussionen.

Herrn Prof. Robert Weigl danke ich für die Übernahme des Koreferats und das damit bekundete Interesse an meiner Arbeit.

Desweiteren möchte ich mich bei Herrn Dr.-Ing. Ulf Schlichtmann, Dr.-Ing. Josef Eckmüller, Dr.-Ing. Christian Sporrer und den Mitarbeitern der Abteilung DAT LIB IO von Infineon Technologies bedanken, die durch ihre Unterstützung, Diskussionen und die gute Arbeitsatmosphäre sehr zum Gelingen dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt auch Herrn Dr.-Ing. Helmut Gräb und den Mitarbeitern der Analogentwurfsgruppe des Lehrstuhls, Herrn Volker Glöckel, Herrn Michael Pronath, Herrn Frank Schenkel, Herrn Guido Stehr und Herrn Dr.-Ing. Stephan Zizala die durch ihre kooperative Zusammenarbeit und Diskussionsbereitschaft einen nicht unbedeutenden Beitrag zum Entstehen dieser Arbeit geleistet haben.

Einen nicht unerheblichen Anteil an dieser Arbeit haben auch meine Freundin Fr. Andrea Stenke und meine Familie, denen ich an dieser Stelle nochmals für die Unterstützung und Motivation zur Fertigstellung dieser Arbeit danken will.

München, im Dezember 2001

Robert Schwencker

Inhaltsverzeichnis

1	Einleitung	9
1.1	Überblick	9
1.2	Aufgabenstellung	10
1.3	Stand der Technik	12
1.4	Ziele der Arbeit	17
2	Problemdefinition	21
2.1	Schaltungsparameter und Schaltungseigenschaften	21
2.1.1	Schaltungsparameter	21
2.1.2	Schaltungseigenschaften, Spezifikationen und Ausbeute	24
2.1.3	Strukturelle Nebenbedingungen	26
2.2	Problemstellung des Nominalentwurfs	27
2.2.1	Der Parameterabstand	28
2.2.2	Zielfunktion für den Nominalentwurf	31
2.3	Problemstellung der Entwurfszentrierung	31
2.3.1	Der Worst-Case-Abstand	32
2.3.2	Zielfunktion der Entwurfszentrierung	34
3	Dimensionierung	35
3.1	Überblick	36
3.2	Die charakteristische Grenzkurve	37
3.2.1	Anpassung der Zielfunktion	37
3.2.2	Definition der charakteristischen Grenzkurve	38
3.2.3	Eigenschaften der charakteristischen Grenzkurve	38

3.2.4	Berechnung von Punkten auf der Grenzkurve	40
3.2.5	Berechnung der Steigung der Grenzkurve	41
3.2.6	Abschnittsweise lineare Approximation der Grenzkurve	42
3.2.7	Approximation der Grenzkurve durch Bernstein-Bézier Kurven	47
3.2.8	Bestimmung einzelner Punkte auf der Grenzkurve	50
3.3	Schrittweitenbestimmung	53
3.4	Die verallgemeinerte Grenzkurve	55
3.4.1	Definition der verallgemeinerten Grenzkurve	57
3.4.2	Eigenschaften der verallgemeinerten Grenzkurve	57
3.4.3	Berechnung der verallgemeinerten Grenzkurve	60
3.4.4	Steigung der verallgemeinerten Grenzkurve	61
3.5	Abbruchkriterien	63
4	Ergebnisse	65
4.1	Operationsverstärker	65
4.2	Ausgangstreiber	70
5	Zusammenfassung	75
A	Optimierprobleme	77
A.1	Extremalkriterien für differenzierbare Optimierprobleme mit Nebenbedingungen	77
A.2	Konvexe Optimierprobleme	78
A.2.1	Konvexe Mengen und Gebiete	79
A.2.2	Konvexe Funktionen	79
A.2.3	Konvexe Optimierprobleme	80
B	Bernstein-Bézier Kurven	81
B.1	Allgemeine Definition einer Bernstein-Bézier Kurve	81
B.2	Eigenschaften der BB-Kurve mit drei Kontrollpunkten	81
	Literaturverzeichnis	85
	Abbildungsverzeichnis	93

Tabellenverzeichnis	95
Verzeichnis der Formelzeichen	97
Kurzfassung	99

Kapitel 1

Einleitung

1.1 Überblick

Moderne integrierte Schaltungen ermöglichen durch ihren niedrigen Leistungsverbrauch und ihre große Funktionsvielfalt mobile Telekommunikation [Wei00], sind aber auch aus Haushaltsgeräten [BK00] und Kraftfahrzeugen [Nod00] nicht mehr wegzudenken. Durch die rasante Entwicklung der Mikroelektronik ist mittlerweile die Integration komplexer Systeme auf einem Chip (*System on a Chip, SoC*) möglich [Pin00]. Dabei werden zumeist analoge und digitale Schaltungsteile auf einem Chip integriert (*Mixed-Signal Schaltungen*). Den großen Fortschritten der Halbleitertechnologie stehen auf Seite der Entwurfsautomatisierung keine adäquaten Fortschritte gegenüber (*design productivity gap*) [Sem99]. Um aber mit den steigenden Komplexitäten und den immer kürzeren Entwicklungszyklen Schritt halten zu können, ist ein weitgehend automatisierter Schaltungsentwurf nötig.

Für den Entwurf digitaler Schaltungen und Teilsysteme einer Mixed-Signal-Schaltung existiert ein etabliertes Entwurfsmodell [GK83], das verschiedene Abstraktionsebenen und Repräsentationen (*views*) bietet. Der Übergang zwischen den einzelnen Abstraktionsebenen und Repräsentationen und die Überprüfung dieses Überganges werden durch eine Vielzahl kommerzieller Werkzeuge unterstützt. Die Synthese digitaler Schaltungen basiert auf einer Bibliothek standardisierter Schaltungstypen (*Standardzellen*), die durch wenige Daten (z. B. Verzögerungszeit, Fläche, Leistungsverbrauch) charakterisiert sind. Die Dimensionierung der einzelnen Zellen auf Transistorebene und das Layout erfolgen im Rahmen der Bibliotheksentwicklung nur einmal für jede Technologie. Durch die ausgereiften Werkzeuge und den hohen Grad an Wiederverwendung (*reuse*) ist ein schneller und weitestgehend automatisierter Entwurf digitaler Schaltungen möglich, was auch den Forderungen einer immer kürzeren Time-To-Market entgegenkommt. Dies hat zur Folge, daß der Anteil digitaler Teilsysteme innerhalb einer Mixed-Signal Schaltung meist deutlich über 80% liegt und die Technologien für digitale Schaltungen optimiert sind.

Der Entwurf analoger Schaltungen dagegen erfolgt in der Regel weitgehend manuell durch erfahrene Schaltungsentwickler [LS94] mit Hilfe eines Schaltungssimulators (z. B. SPICE [Nag75], SABER [VVS87] oder TITAN [FWZ⁺92]). Insbesondere die Dimensionierung der Schaltungen auf Transistorebene ist in diesem Zusammenhang eine zeitaufwendige und fehleranfällige Tätigkeit. Desweiteren ist aufgrund der vielfältigen Anforderungen und Einsatzbereiche analoger Schaltungen das

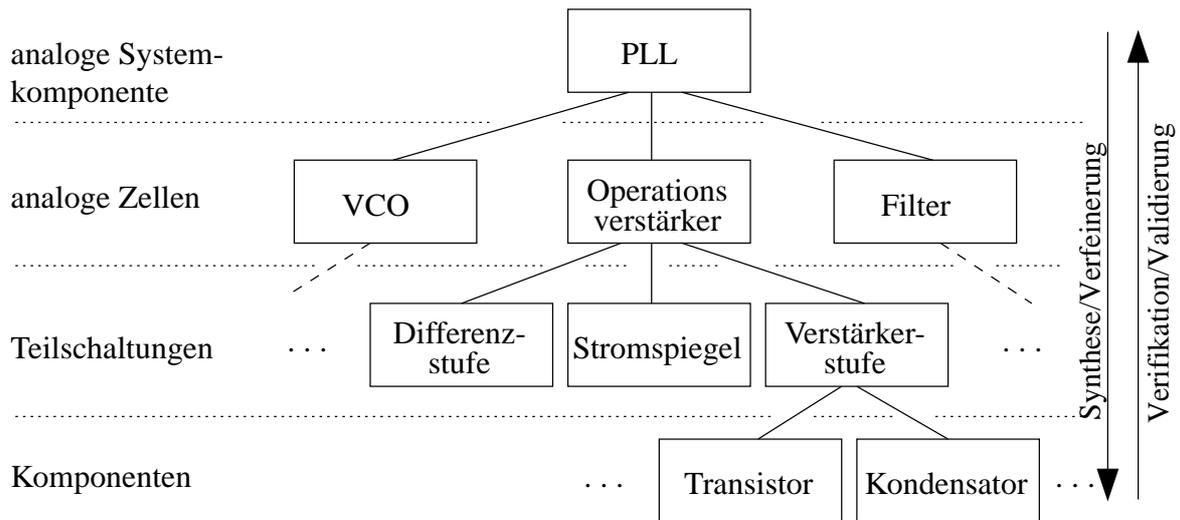


Abbildung 1.1: Abstraktionsebenen beim Entwurf analoger Systeme.

Wiederverwendungspotential gering. Eine Bibliothek wiederverwendbarer analoger Komponenten erscheint nur dann sinnvoll, wenn zugleich Werkzeuge zur automatischen Anpassung der einzelnen Zellen an verschiedene Anforderungen zur Verfügung stehen.

Aufgrund des geringen Wiederverwendungspotentials und des geringen Automatisierungsgrads benötigt der Entwurf analoger Schaltungen trotz des kleinen Flächenanteils meist einen dominanten Anteil der Entwurfskosten einer Mixed-Signal-Schaltung [Bor97, GS91, GRRC00]. Allerdings stellen die analogen Teile oft Schlüsselkomponenten eines Systems dar, die nicht oder nur unter unververtretbaren Kosten (z. B. Fläche oder Leistungsverbrauch) durch digitale Schaltungen ersetzt werden können.

Vor diesem Hintergrund soll im folgenden kurz die Aufgabenstellung der Dimensionierung analoger Schaltungen und die Einordnung in den Entwurfsablauf beschrieben werden.

1.2 Aufgabenstellung

Beim Entwurf eines integrierten Mixed-Signal-Systems (z. B. Mobilfunkchipsatz, Modemchip), wird in einem ersten Schritt das gesamte System in digitale und analoge Teilsysteme zerlegt. Die Unterteilung in digitale und analoge Teilsysteme erfolgt dabei so, daß sich die Systemanforderungen mit minimalem Aufwand (z. B. Fläche, Leistungsverbrauch) innerhalb der gegebenen Entwurfszeit realisieren lassen.

Die vorliegende Arbeit beschäftigt sich mit Aspekten des Entwurfs analoger (Teil-)Systeme. Nach [GS91, Eck98, GRRC00] und wie in Abbildung 1.1 dargestellt, lassen sich beim Entwurf analoger Systeme folgende Abstraktionsebenen unterscheiden:

- *Systemebene:* Eine analoge Systemkomponente stellt einen Schaltungsblock mit einer bestimmten Funktionalität dar. Die Anforderungen an das analoge System ergeben sich aus der Einbettung in das gesamte Mixed-Signal System. Analoge Systeme (z. B. PLL) können in der Größenordnung von einigen Hundert bis mehreren Tausend Transistoren liegen.

- *Zellebene*: Analoge Zellen sind die Bausteine, aus denen ein analoges System aufgebaut ist. Eine analoge Zelle nimmt innerhalb des analogen Systems immer nur eine Funktion wahr (z. B. Verstärkung); für jede Funktion gibt es aber eine Vielzahl von Realisierungen (*Topologien*). Die Zellen werden durch ihre Schaltungseigenschaften beschrieben und bestehen typischerweise aus einer überschaubaren Anzahl von Transistoren (normalerweise weniger als 50 Transistoren).
- *Teilschaltungen (Building blocks)*: Analoge Teilschaltungen von der Größe weniger Transistoren stellen die Grundbausteine dar, aus denen eine analoge Zelle besteht. Die in Abschnitt 2.1.3 vorgestellten strukturellen Nebenbedingungen [Eck98, ZEG98, SEGA99a, GZEA01] garantieren dabei die prinzipielle Funktionsfähigkeit jeder Teilschaltung während des gesamten Entwurfsablaufs.
- *Komponenten (Elektrische Ebene/Transistorebene)*: Auf der Komponentenebene wird die Schaltung durch ihre Bauelemente und Konnektivitäten beschrieben. Für die einzelnen Bauelemente existieren im Simulator Modelle, so daß die Schaltung auf der elektrischen Ebene in einem Schaltungssimulator simuliert werden kann.

In dieser Arbeit werden Teilaspekte des Übergangs von der Zellebene auf die Transistorebene betrachtet. Der Übergang von der Zellebene auf die Transistorebene unterteilt sich in drei Aufgabenstellungen:

1. **Topologieentwurf**: Aufgabe des Topologieentwurfs ist es, eine Schaltungsstruktur (*Topologie*) auszuwählen oder zu entwerfen, die für die gestellten Anforderungen möglichst gut geeignet ist.
2. **Dimensionierung**: Schaltungsdimensionierung bedeutet für eine gegebene Topologie die Parameter der Schaltung (z. B. Längen und Weiten der Transistoren, Kapazitäts- und Widerstandswerte) so einzustellen, daß die Schaltung die an sie gestellten Anforderungen (z. B. bezüglich Slew-Rate oder Verstärkung eines Operationsverstärkers) erfüllt. Zusätzlich sollte die Schaltung so robust wie möglich gegenüber Änderungen der Betriebsumgebung oder Streuungen des Fertigungsprozesses sein.
3. **Layout**: Für die innerhalb der Dimensionierung bestimmten Werte der Schaltungsparameter ist das Layout der Schaltung zu erzeugen. Bei der Layouterstellung sind noch zusätzliche Regeln zu beachten, um z. B. eine Unempfindlichkeit der Schaltung bezüglich Mismatch sicherzustellen.

Die vorliegende Arbeit konzentriert sich auf die Aufgabe der Dimensionierung. Die Dimensionierung einer gegebenen Topologie erfolgt in der industriellen Praxis derzeit weitestgehend manuell und ist daher eine zeitaufwendige und fehleranfällige Tätigkeit. Deshalb soll in dieser Arbeit ein Verfahren zur automatischen Dimensionierung analoger Zellen vorgestellt werden. Die Dimensionierung kann in zwei Schritte unterteilt werden:

- *Nominalentwurf (Nominal design)*: Die Schaltung wird so dimensioniert, daß sie im Nominalpunkt die an sie gestellten Anforderungen erfüllt. Der Nominalpunkt beschreibt typische Werte der Betriebsparameter (z. B. Temperatur, Versorgungsspannung) und einen typischen Fertigungsprozeß (z. B. typische Werte für Oxiddicke und Dotierungen). Ergebnis des Nominalentwurfs sollte eine Dimensionierung sein, die einen möglichst guten Ausgangspunkt für eine nachfolgende simulationsintensive Entwurfszentrierung bietet.

- *Entwurfszentrierung (Design centering)*: Ziel der Entwurfszentrierung ist es, die Schaltung so zu dimensionieren, daß sie robust gegen Einflüsse der Betriebsparameter und Prozeßstreuungen ist. Dies bedeutet zugleich eine hohe Ausbeute unter Berücksichtigung von Parameterschwankungen (*parametric yield*). Katastrophische Fehler (z.B. Spot Defekte) können in der Entwurfszentrierung nicht berücksichtigt werden, da Wahrscheinlichkeiten für diese erst nach Fertigstellung des Layouts ermittelt werden können.

Für moderne Technologien mit ihren immer kleineren Strukturgrößen ($0.25\mu\text{m}$ und kleiner) steigt die Bedeutung der Entwurfszentrierung stark an. Insbesondere Mismatch-Effekte beeinflussen in diesen Technologien dominant das Schaltungsverhalten. Um auch unter Berücksichtigung dieser Effekte hohe Ausbeuten zu erzielen, ist eine automatische Entwurfszentrierung unbedingt notwendig.

Ein weiteres wichtiges Anwendungsgebiet für eine automatische Dimensionierung ist der Entwurf von Bibliothekszellen für den digitalen Schaltungsentwurf (z.B. Flip-Flops, Treiberschaltungen). Für die einzelnen Zellen einer digitalen Bibliothek wird immer die gleiche Struktur (*Topologie*) verwendet. Für jede neue Technologie und jede Änderung der Transistormodellparameter ist eine Neudimensionierung der Zellen notwendig. Die skalierte Dimensionierung in der alten Technologie bietet meist einen guten Ausgangspunkt für eine neue Dimensionierung. Damit sind gute Voraussetzungen für eine Automatisierung dieses Entwurfsschrittes gegeben. Ist eine automatische Layouterstellung vorhanden, können mit Hilfe einer automatischen Dimensionierung auch dynamisch Bibliotheken [Keu99] erzeugt werden.

Insgesamt ergeben sich folgende Erwartungen an ein Werkzeug zur automatischen Dimensionierung und Entwurfszentrierung:

- Verkürzung der Entwurfszeiten durch Automatisierung der zeitaufwendigen Dimensionierung. Das bedeutet eine deutliche Erhöhung der Produktivität der Schaltungsentwickler und eine Verkürzung der Time-To-Market.
- Erstellung robuster Designs, die unempfindlich gegenüber parasitären Effekten sind. Dies soll Schaltungen ermöglichen, die auf Anhieb funktionieren (*first success in silicon*), was ebenfalls den Forderungen nach einer kürzeren Time-To-Market entgegenkommt.
- Höherer Wiederverwendungsgrad von analogen Schaltungsblöcken durch vereinfachte Anpassung analoger Zellen und Blöcke an neue Anforderungen. Durch die Wiederverwendung von bereits produzierten (*silicon proven*) Blöcken wird eine größere Entwurfssicherheit erreicht, was dem Ziel des „first success in silicon“ förderlich ist.

Nicht zuletzt aufgrund des hohen Potentials zur Produktivitätssteigerung existieren zahlreiche Verfahren zur Automatisierung der Dimensionierung analoger Zellen.

1.3 Stand der Technik

Die veröffentlichten Verfahren zur Dimensionierung analoger Schaltungen kann man einerseits in wissensbasierte Verfahren und andererseits in simulationsbasierte Verfahren unterteilen:

- Wissensbasierte Ansätze verwenden Schaltungswissen in Form von Gleichungen, Modellen, Entwurfsplänen, etc., um zeitaufwendige Simulationen zu vermeiden oder deren Anzahl deutlich zu reduzieren.

- Simulationsbasierte Verfahren verwenden einen numerischen Schaltungssimulator zur Bestimmung der Schaltungseigenschaften für einen gegebenen Parametersatz (z. B. Längen und Weiten von Transistoren).

Im folgenden sollen ausgewählte Ansätze aus beiden Bereichen kurz diskutiert werden. Es werden zunächst verschiedene Verfahren zum Nominalentwurf vorgestellt. Anschließend wird auf verschiedene Verfahren zur Entwurfszentrierung eingegangen.

Wissensbasierter Nominalentwurf

Die ersten wissensbasierten Werkzeuge IDAC [De87] und OPASYN [KSG90] unterstützten den gesamten Schritt der Synthese einer analogen Zelle: Topologieauswahl, Dimensionierung und Layout. Für die Dimensionierung müssen für jede Schaltungstopologie (z. B. verschiedene Realisierungen eines Operationsverstärkers) die Schaltungseigenschaften als Gleichungen in Abhängigkeit der Entwurfsparameter und ein Entwurfsplan gegeben sein. Bei Freiheitsgraden (z. B. Anteile der einzelnen Stufen eines Operationsverstärkers an der Gesamtverstärkung) sind teilweise zusätzliche Eingriffe des Schaltungsentwicklers notwendig. Die Integration einer neuen Topologie in diese Werkzeuge bedeutet die Herleitung der Gleichungen für alle interessierenden Eigenschaften und das Erstellen eines Entwurfsplanes für diese Topologie.

OASYS [HRC89] und STAIC [HEL92] verwenden eine hierarchische Beschreibung einer Topologie. Ein Operationsverstärker wird beispielsweise als Zusammenschaltung von Stromspiegeln, Differenzpaaren, etc. beschrieben. Für einen Stromspiegel können wiederum verschiedene Realisierungen eingebunden werden (z. B. Kaskode-Stromspiegel oder einfacher Stromspiegel). Für jede Topologie muß ein Entwurfsplan mit Gleichungen (OASYS) oder nur die Gleichungen (STAIC) gegeben sein. Eine gute Übersicht und ein Vergleich wissensbasierter Dimensionierungswerkzeuge bis 1992 findet sich in [HEL92]

Aufgrund des hohen Aufwands bei der Herleitung der Gleichungen für die Schaltungseigenschaften (z. B. [HRC89] berichtet von 18 Monaten für verschiedene Realisierungen eines ein- und zweistufigen Operationsverstärkers) wurden Möglichkeiten zur Automatisierung dieses Schritts gesucht. ARIADNE [SS93] verwendet den symbolischen Simulator ISAAC [GWS89] zur Generierung vereinfachter Gleichungen für die Kleinsignaleigenschaften. Großsignaleigenschaften müssen nach wie vor manuell hergeleitet werden.

Weitere Ansätze zur Dimensionierung basierend auf symbolischer Analyse sind in [HS95] (ANALOG INSYDES) und [LDGS97] (EF2EF) beschrieben. In [dMHL98] (GPAD) wird gezeigt, daß für Operationsverstärker bei Verwendung geeigneter Transistormodelle viele Kleinsignaleigenschaften als posynomiale Funktionen der Entwurfsparameter beschrieben werden können, womit die Dimensionierung ein konvexes Optimierproblem ist. Konvexe Optimierprobleme wiederum sind dadurch gekennzeichnet, daß jedes Optimum global ist, d. h. es wird unabhängig vom Startpunkt ein globales Optimum gefunden.

Trotz der Fortschritte auf dem Gebiet der symbolischen Analyse, haben die wissensbasierten Verfahren noch immer den Nachteil, daß für jede neue Topologie die Beschreibungen der Großsignaleigenschaften manuell hergeleitet und in die Werkzeuge integriert werden müssen. Erste Ansätze zur Generierung von Verhaltensmodellen, die auch das Großsignalverhalten korrekt beschreiben sind in [Bor98, Bor97] zu finden. Um die Komplexität der Gleichungen klein zu halten, können nur stark

vereinfachte Transistormodelle verwendet werden. Diese geben aber speziell in den modernen Technologien das Schaltungsverhalten häufig nicht mehr mit ausreichender Genauigkeit wieder.

Simulationsbasierter Nominalentwurf

Im Gegensatz zu den wissensbasierten Dimensionierungsverfahren verwenden die simulationsbasierten Werkzeuge einen Schaltungssimulator zur Funktionsauswertung im Rahmen einer Optimierung (Simulator in a loop). Bei den verwendeten Optimieralgorithmen kann man zwischen deterministischen Optimierverfahren (z. B. Gradienten-Verfahren) und statistischen Optimierverfahren (z. B. Simulated Annealing) unterscheiden.

Deterministische Verfahren finden im Allgemeinen nur ein lokales Minimum. Nur für eingeschränkte Problemklassen (z. B. konvexe Optimierprobleme) kann die Konvergenz zu einem globalen Optimum bewiesen werden. Durch ihre Fähigkeit, auch schlechtere Werte der Zielfunktion mit einer endlichen Wahrscheinlichkeit zu akzeptieren, können statistische Optimierverfahren globale Optima finden. „Simulated Annealing“ [KCGV83] beispielsweise versucht, den Kristallisationsprozeß der Natur nachzubilden, der ein globales Energieminimum findet. Ein globales Minimum wird in der Natur aber nur bei einer unendlich langsamen Absenkung der Temperatur gefunden, was einer unendlichen Anzahl von Simulationen in der Schaltungsdimensionierung entspricht.

Das Programm FRIDGE [MFDCRV94], dessen Nachfolger XFRIDGE [PMMDC⁺00], ASTRX OBLX [ORC96, OMRC98] und dessen Nachfolger MAELSTROM [KPRC99] verwenden „Simulated Annealing“ zur Optimierung. ASTRX/OBLX arbeitet zusätzlich noch mit einem speziell angepaßten Simulator. Dieser verwendet *Asymptotic Waveform Evaluation* (AWE) [PR94] für die AC-Analyse. AWE ermöglicht eine deutlich schnellere AC-Analyse bei einer nur geringfügig verringerten Genauigkeit. Zur weiteren Beschleunigung löst der Simulator in der DC-Analyse die Kirchhoff-Gleichungen nicht exakt, sondern integriert sie vielmehr als Optimierziele in die Zielfunktion. Der angepaßte Simulator in ASTRX/OBLX unterstützt keine transiente Analyse, so daß in diesem Fall wieder explizite Gleichungen bereitgestellt werden müssen.

Die durch den speziell angepaßten Schaltungssimulator auferlegten Beschränkungen in ASTRX/OBLX werden in dessen Nachfolger MAELSTROM aufgehoben. Um trotz der dadurch deutlich gestiegenen Simulationszeiten die praktische Anwendbarkeit zu erhalten, wurde ein parallelisierter „Simulated Annealing“ Algorithmus (*parallel recombinative simulated annealing* (PRSA) [MG95]) integriert. Zur Dimensionierung eines Operationsverstärkers (32 Transistoren) mit ASTRX/OBLX sind Zahlen von 100 000 Simulationen veröffentlicht, mit MAELSTROM werden 17 000 Simulationen nur für den Nominalentwurf benötigt.

Aufgrund der starken Konvergenzeigenschaften basieren viele der deterministischen Verfahren zur Schaltungsdimensionierung auf angepaßten Trust-Region-Verfahren nach Levenberg-Marquardt [Lev44, Mar63, Mor77].

CANDI [Pör89] ist ein Werkzeug zur interaktiven Dimensionierung analoger Schaltungen mittels deterministischer Verfahren. CANDI unterstützt den Anwender bei einer sinnvollen Wahl der Schrittweite mit Hilfe der Grenzkurve [Hus82, AH84, Lei86, Pör89, ALP88], die als eine Visualisierung des Levenberg-Marquardt Verfahrens interpretiert werden kann.

In [Eck98] werden darauf aufbauend weitere Verfahren zum interaktiven Nominalentwurf und Entwurfszentrierung behandelt. Insbesondere werden hier die strukturellen Nebenbedingungen vorge-

stellt als ein Mittel, den Suchraum für die Optimierung auf einen schaltungstechnisch sinnvollen Unterraum zu reduzieren. Diese Einschränkung des Suchraumes auf ein Gebiet in dem sich Schaltungen hoher Robustheit befinden, garantiert eine hohe Qualität der berechneten Ergebnisse. Insbesondere für schlecht konditionierte Problemstellungen, die typisch für den analogen Schaltungsentwurf sind, bringt die Einschränkung des Suchraumes auch eine deutliche Verbesserung der Konvergenzgeschwindigkeit des Dimensionierungsalgorithmus.

Weitere Verfahren zur automatischen Dimensionierung analoger Zellen sind in APLSTAP [HZ81] OPSIM [Dog81] und DELIGHT.SPICE [NRSVT88] implementiert. Alle Werkzeuge unterstützen Trust-Region-Verfahren nach Levenberg-Marquardt. DELIGHT.SPICE [NRSVT88] stellt zudem noch weitere unterschiedliche Optimieralgorithmen zur Verfügung.

JIFFYTUNE [CCH⁺98] verwendet das Optimierpaket LANCELOT und den Simulator SPECS zur Berechnung der Funktionswerte und Gradienten. Die Gradienten werden mittels eines adjungierten Netzwerks berechnet, wodurch sich der Aufwand für die Gradientenberechnung deutlich reduziert. Allerdings ist die Gradientenberechnung auf vorgegebene Funktionen beschränkt; derzeit sind in JIFFYTUNE Verzögerungszeit, Rise- und Falltime, Fläche und Leistungsverbrauch implementiert. Das Haupteinsatzgebiet von JIFFYTUNE liegt somit in der Optimierung von digitalen Schaltungen auf Transistorebene.

Entwurfszentrierung

Die Verfahren zur Entwurfszentrierung lassen sich in statistische und deterministische Verfahren einteilen. Die statistischen Verfahren verwenden als Optimierziel zumeist die Ausbeute, die mit Hilfe einer Monte-Carlo-Analyse [HH75, Fis99] ermittelt werden kann. Deterministische Verfahren dagegen beruhen oft auf einer geometrischen Approximation des Akzeptanzgebietes, innerhalb dessen die Schaltung als funktionsfähig gilt.

Eine Monte-Carlo-Analyse liefert für beliebige Schaltungen eine Abschätzung für die Ausbeute. Der Vertrauensbereich σ_{MC} der Ausbeuteabschätzung ist unabhängig von der Anzahl der Parameter [Fis99]: $\sigma_{MC} \sim 1/\sqrt{N_{MC}}$, wobei N_{MC} der Stichprobenumfang ist.

Aus einer Monte-Carlo-Analyse kann ebenfalls der Ausbeutegradient für die statistisch streuenden Parameter mittels *Center of Gravity*-Algorithmen [ZS95, KK98, SS80] bestimmt werden. Eine verbesserte Abschätzung des Gradienten und der zweiten Ableitung (Hesse-Matrix) der Ausbeute bezüglich der statistisch streuenden Parameter und ein darauf aufbauendes Verfahren zur Entwurfszentrierung findet sich in [Kob82, AK82].

Für integrierte Schaltungen sind die Parameter in die statistisch streuenden Parameter der aktiven Elemente (z. B. Transistormodellparameter) und die nicht streuenden Entwurfparameter unterteilt. Die Transistormodellparameter sind nach Abschluß der Technologieentwicklung (*process freeze*) fest und können vom Schaltungsentwickler nicht eingestellt werden, während die Entwurfparameter innerhalb von durch die Technologie vorgegebenen Grenzen frei einstellbar sind. Für diesen Fall der disjunkten Parameterräume liefert die direkte Lösung des Integrals der Wahrscheinlichkeitsdichtefunktion über das Akzeptanzgebiet mittels einer Monte-Carlo-Analyse zunächst keine Informationen über den Gradienten der Ausbeute bezüglich der nicht streuenden Entwurfparameter. Für eine Entwurfszentrierung wird aber dieser Ausbeutegradient benötigt.

Eine Möglichkeit trotz dieser disjunkten Parameterräume einen Ausbeutegradienten für die Entwurfparameter aus der Monte-Carlo Analyse zu berechnen, besteht darin für die Entwurfparameter künstliche Streuungen einzuführen. Werden diese künstlichen Streuungen eingeführt, dann kann aus der Monte-Carlo Analyse aber keine Ausbeute mehr berechnet werden.

In [FD93, FD91, FD90, HCY88] sind Verfahren beschrieben, die durch Verwendung des Gaußschen Integralsatzes die Ausbeute über ein Oberflächenintegral bestimmen. Dieses Oberflächenintegral wird mittels Monte-Carlo Analyse abgeschätzt, wobei zusätzlich zur Abschätzung der Ausbeute auch ein Ausbeutegradient bezüglich der nicht streuenden Entwurfparameter bestimmt werden kann. Für jedes Sample innerhalb der Monte-Carlo Analyse wird eine Liniensuche in Richtung eines Parameters benötigt um, die Grenze des Akzeptanzgebietes in dieser Richtung zu bestimmen. Aufgrund des daher im Vergleich zur reinen Monte-Carlo Analyse nochmals deutlich gestiegenen Rechenaufwandes wird in [FD93, FD91, FD90, HCY88] diese Methode nur mit zumeist linearen Modellen der Schaltung verwendet und nicht mit einem Schaltungssimulator. Insbesondere für lineare Schaltungsmodelle begrenzt der Modellierungsfehler die Genauigkeit der Monte-Carlo Analyse dominant. In [HCY88] wird zudem die Anzahl der statistischen Parameter auf 4 reduziert (Weiten- und Längenreduktion, C_{ox} und V_{FB}), die in den untersuchten digitalen Schaltungen einen dominanten Einfluß auf die Schaltungseigenschaften haben.

In [RMRV98] werden innerhalb der Monte-Carlo Analyse AC-Simulationen mit ähnlichen Parameterwerten zu Simulationsgruppen zusammengefaßt, wodurch die benötigte Rechenzeit für eine Monte-Carlo Analyse auf Kosten der Genauigkeit deutlich reduziert werden kann.

Eine Vielzahl der deterministischen Verfahren zur Entwurfszentrierung sind geometrische Verfahren, die das Akzeptanzgebiet mittels unterschiedlicher geometrischer Körper approximieren. Unter anderem sind Verfahren basierend auf Simplex-Approximation [DH77], Ellipsoiden [AMH91, AMHH99], Polytopen [SVK93] und allgemeinen konvexen Körpern [SPV99] in der Literatur zu finden. Das Zentrum des approximierten Akzeptanzgebietes ist die Dimensionierung mit der höchsten approximierten Ausbeute. Problematisch bei der Approximation des Akzeptanzgebietes ist wiederum, daß beim Entwurf integrierter Schaltungen Entwurfparameter und statistische Parameter disjunkt sind.

[MCR94, OMRC98] ist eine Erweiterung von ASTRX/OBLX [ORC94, ORC96] auf den statistischen Schaltungsentwurf. Für die statistischen Parameter müssen Grenzen gegeben sein (z. B. 3σ), innerhalb derer die Schaltung funktionieren soll, die also implizit das Akzeptanzgebiet definieren. Die Dimensionierung wird darauf basierend als infinites Programmierproblem formuliert und wie in ASTRX/OBLX mittels „Simulated Annealing“ gelöst. Die daraus resultierende extrem hohe Anzahl von Simulationen ist nur aufgrund des speziell angepaßten, schnellen Simulators akzeptabel. Wie im vorherigen Abschnitt beschrieben beherrscht dieser Simulator keine transiente Analyse, so daß Eigenschaften, die eine transiente Analyse benötigen, nicht untersucht werden können.

Neuere Verfahren basieren auf Robustheitsmaßen, welche die Robustheit einer Spezifikation gegenüber den statistischen Parametern und den Betriebsparametern beschreiben. *Performance Scores* [DK95], *Linearized Performance Penalties* (LPP) [KD95], Worst-Case-Abstände [AGW94, Grä93, Wie94, SGA99, AEG⁺00a], *Capability Potential Index* (C_p) und *Capability Performance Index* (C_{pk}) [SA94, DG98] sind als verwendete Robustheitsmaße veröffentlicht.

Die Verfahren basierend auf Robustheitsmaßen können prinzipiell für Nominalentwurf und Entwurfszentrierung verwendet werden. Der Hauptaufwand bei diesen Verfahren liegt in der Berechnung der Robustheitsmaße. Da in diesem Fall für Nominalentwurf und Entwurfszentrierung dieselben Robust-

heitsmaße verwendet werden, können die unterschiedlichen Charakteristiken und Genauigkeitsanforderungen der beiden Entwurfsschritte nicht berücksichtigt werden, was in einem relativ hohen Simulationsaufwand resultiert verglichen mit einer in Nominalentwurf und Entwurfszentrierung getrennten Dimensionierung.

[DG98] reduziert den Simulationsaufwand basierend auf symbolischen Schaltungsgleichungen und einfachen Transistormodellen. Mittels DONALD [SSG90] wird aus den Gleichungen ein Designplan generiert, der zur Dimensionierung der Schaltung dient. Durch die Verwendung symbolischer Gleichungen kann der Simulationsaufwand deutlich reduziert werden. Allerdings unterliegt das Verfahren dadurch aber auch den oben diskutierten Einschränkungen der symbolischen Analyse; die einfachen Transistormodelle limitieren die Genauigkeit der Abschätzung und mit den derzeitigen Methoden können nur lineare bzw. linearisierte Schaltungen innerhalb der symbolischen Analyse untersucht werden.

Die meisten veröffentlichten Verfahren betrachten nur einen der beiden Entwurfsschritte obwohl eine Dimensionierung aus Nominalentwurf und Entwurfszentrierung besteht. Die auf Robustheitsmaßen basierenden Verfahren [KD95, DK95, AGW94, Grä93, Wie94, SGA99, SA94, DG98] vereinheitlichen beide Entwurfsschritte, allerdings auf Kosten eines erhöhten Simulationsaufwandes. Es werden also Verfahren zur Dimensionierung benötigt, die einerseits Nominalentwurf und Entwurfszentrierung unterstützen, andererseits aber auch die Charakteristiken beider Entwurfsschritte berücksichtigen.

Desweiteren liefert eine rein mathematische Lösung des Optimierproblems der Dimensionierung oft eine schaltungstechnisch nicht sinnvolle Lösung. In [MCR94, OMRC98] ist beispielsweise gezeigt, daß ASTRX/OBLX zwar einen Operationsverstärker für den Nominalfall deutlich besser dimensioniert als ein Designer. Bereits bei der Berücksichtigung von kleinen Streuungen der Betriebsparameter aber fällt die Verstärkung extrem ab, da Grundprinzipien des analogen Schaltungsentwurfs verletzt sind. Für die Akzeptanz eines automatischen Dimensionierungsverfahrens ist es also notwendig, diese Grundprinzipien des analogen Schaltungsentwurfs einzuhalten.

1.4 Ziele der Arbeit

Ziel der vorliegenden Arbeit ist es, Verfahren zu entwickeln, die eine effiziente *automatische* Dimensionierung analoger integrierter Schaltung ermöglichen. Insbesondere soll dabei Wert auf eine *allgemeine Einsetzbarkeit* der vorgestellten Verfahren gelegt werden; d.h. sie sollen mit beliebigen Simulatoren (z.B. Schaltungssimulatoren, Leitungssimulatoren) zusammenarbeiten, sie sollen keine Beschränkungen bezüglich der Schaltungsklassen (z.B. nur Operationsverstärker) oder der Art der Eigenschaften aufweisen (z.B. nur Kleinsignaleigenschaften).

Nominalentwurf und Entwurfszentrierung sollen in dieser Arbeit auf eine *gemeinsame mathematische Problemstellung* zurückgeführt werden, wobei die Charakteristiken jedes Entwurfsschrittes in der Formulierung der Zielgrößen berücksichtigt werden. Zielgrößen für den Nominalentwurf sind die *Parameterabstände* [Eck98, SSGA00, SEGA99b], die mit geringem Simulationsaufwand bestimmbar sind und Betriebsparameter sowie Empfindlichkeiten bezüglich der Entwurfparameter berücksichtigen.

Innerhalb der Entwurfszentrierung werden die *Worst-Case-Abstände* [AGW94, Grä93, Wie94] als Zielgrößen verwendet. Basierend auf den Worst-Case-Abständen kann die parametrische Ausbeute

der Schaltung unter Berücksichtigung der Betriebsparameter abgeschätzt werden. Die Zielgrößen in beiden Entwurfsschritten lassen sich als Robustheitsmaße interpretieren, welche die Robustheit der Schaltung gegenüber den in diesem Entwurfsschritt relevanten Parametern beschreiben. Die Zielgrößen werden sowohl im Nominalentwurf als auch in der Entwurfszentrierung in eine Zielfunktion transformiert, die bei Annahme linear von den Parametern abhängender Zielgrößen konvex ist.

Um eine Akzeptanz der Verfahren und eine hohe Effizienz zu erreichen, müssen die besonderen Anforderungen des analogen Schaltungsentwurfs in der Dimensionierung berücksichtigt werden. Gut entworfene analoge Schaltungen zeichnen sich in der Regel durch eine hohe Unempfindlichkeit gegenüber Parameterschwankungen und Parameter mit nahezu identischer Wirkung aus. Dieses für den Betrieb und manuellen Schaltungsentwurf wünschenswerte Schaltungsverhalten bedeutet für eine automatische Schaltungsdimensionierung eine schlechte Kondition des zugrundeliegenden Optimierproblems. Diese schlechte Kondition ist typisch für den Entwurf analoger Schaltungen und muß deshalb in den Algorithmen berücksichtigt werden.

Gleichzeitig ist zu beachten, daß die Simulation analoger Schaltungen zeitaufwendig ist (zwischen mehreren Sekunden bis über eine Stunde für eine Simulation). Jede Auswertung der Zielfunktion innerhalb der Dimensionierung bedeutet zumeist mehrere Simulationen (DC-, AC- und transiente Simulation). Der dominante Teil der Rechenzeit innerhalb der Dimensionierung wird also für Schaltungssimulationen benötigt und nicht für Matrixoperationen innerhalb des Optimieralgorithmus. Dies steht im Gegensatz zu klassischen Optimierproblemen, die Funktionsauswertungen als schnell betrachten und den dominanten Teil der Rechenzeit innerhalb des Optimieralgorithmus verbringen. Dies muß bei der Wahl des Optimieralgorithmus beachtet werden.

Das in dieser Arbeit vorgestellte Optimierverfahren bestimmt in jeder Iteration des Optimieralgorithmus basierend auf einer Linearisierung der Zielgrößen mit Hilfe der *verallgemeinerten Grenzkurve* [SSGA00, SSGA99] eine Schrittweite, die ein günstiges Verhältnis zwischen Auslenkung der Parameter und Reduktion der Zielfunktion hat. Die verallgemeinerte Grenzkurve ist eine Erweiterung der charakteristischen Grenzkurve [ALP88, Lei86, Hus82, Pör89] auf beliebige konvexe Funktionen. Die verallgemeinerte Grenzkurve beschreibt die Menge von paretooptimalen Punkten für die keine gleichzeitige Reduktion von Zielfunktion und Auslenkung möglich ist. In dieser Arbeit wird insbesondere die Anwendung der verallgemeinerten Grenzkurve auf die vorgestellte Zielfunktion diskutiert, sie ist aber nicht darauf limitiert. Wie in den Beispielen in Kapitel 4 gezeigt, ist damit die Bestimmung einer sinnvollen Schrittweite für die verwendete Zielfunktion mit einem geringen Simulationsaufwand möglich.

In der gesamten Dimensionierung werden *strukturelle Nebenbedingungen* [Eck98, ZEG98, GZEA01] konsequent berücksichtigt und der dadurch erreichte große Fortschritt gegenüber dem Stand der Technik wird an Schaltungsbeispielen gezeigt. Strukturelle Nebenbedingungen formalisieren Schaltungswissen und garantieren so die Einhaltung von Grundprinzipien des analogen Schaltungsentwurfs während der gesamten Dimensionierung. Für das der Schaltungsdimensionierung zugrundeliegende Optimierproblem bedeuten die strukturellen Nebenbedingungen einerseits eine deutliche Reduktion der Anzahl der Parameter und andererseits eine Einschränkung des Suchraumes auf ein schaltungstechnisch relevantes Gebiet. Anhand praktischer Schaltungsbeispiele wird in Abschnitt 4 gezeigt, daß erst durch diese Einschränkung des Suchraumes auf ein schaltungstechnisch sinnvolles Gebiet eine automatische Dimensionierung analoger Schaltungen ermöglicht wird.

Im Verlauf der Arbeit entstanden acht Vorveröffentlichungen [SSGA00, AEG⁺00a, SEGA99a, SEGA99b, SSGA99, Sch99, SGLS98, SSLG99, AEG⁺00b]. Die vorgestellten Verfahren sind in dem

Programm WiCkeD [AEG⁺00a, AEG⁺00b] implementiert und werden im industriellen Einsatz zur Schaltungsdimensionierung verwendet. Besonderer Wert wurde darauf gelegt, daß ein Einsatz mit möglichst wenig Vorbereitungsaufwand möglich ist. Industriell verwendete Simulationsumgebungen wie beispielsweise CASAC [Eck98], SIMPILOT oder MATLAB sind ohne wesentlichen Aufwand integrierbar, was eine Grundvoraussetzung für die industrielle Anwendbarkeit ist.

Die weitere Arbeit gliedert sich wie folgt: Kapitel 2 definiert die Grundgrößen für die Schaltungsdimensionierung und führt die Zielfunktion des Nominalentwurf und der Entwurfszentrierung ein. In Kapitel 3 wird das Optimierverfahren für die Minimierung dieser Zielfunktionen diskutiert. Anhand von Beispielschaltungen in Kapitel 4 wird die Anwendung des vorgestellten Verfahrens zur Schaltungsdimensionierung demonstriert. Kapitel 5 stellt die wichtigsten Punkte der Arbeit nochmal im Überblick dar.

Kapitel 2

Problemdefinition

Wie in der Einleitung beschrieben, wird in der vorliegenden Arbeit die Dimensionierung einer Schaltung in Nominalentwurf und Entwurfszentrierung unterteilt. Diese Unterscheidung berücksichtigt insbesondere die charakteristischen Eigenschaften der einzelnen Entwurfsschritte. Um dennoch einen einheitlichen Entwurfsablauf zu ermöglichen, werden als ein wichtiger Beitrag dieser Arbeit beide Entwurfsschritte auf eine gemeinsame mathematische Beschreibung zurückgeführt. Der Unterschied liegt in diesem Fall nur noch in den als Zielgrößen verwendeten Robustheitsmaßen, die sich in Berechnungsaufwand und Interpretation unterscheiden.

Im Folgenden sollen zunächst die Gemeinsamkeiten beider Schritte betrachtet und anschließend die Zielgrößen und das resultierende Optimierproblem für jeden Entwurfsschritt getrennt vorgestellt werden.

2.1 Schaltungsparameter und Schaltungseigenschaften

2.1.1 Schaltungsparameter

Bei einer gegebenen Topologie kann eine Schaltung durch ihre *Schaltungsparameter* $\mathbf{p} \in \mathcal{R}^{n_p}$ beschrieben werden. Diese können wiederum in *Entwurfparameter* $\mathbf{d} \in \mathcal{R}^{n_d}$, *statistische Parameter* $\mathbf{s} \in \mathcal{R}^{n_s}$ und *Betriebsparameter* $\boldsymbol{\theta} \in \mathcal{R}^{n_\theta}$ unterschieden werden:

$$\mathbf{p} = \begin{bmatrix} \mathbf{d} \\ \mathbf{s} \\ \boldsymbol{\theta} \end{bmatrix} \quad (2.1)$$

Entwurfparameter

Die Entwurfparameter \mathbf{d} (auch *Designparameter*) können vom Schaltungsentwickler eingestellt werden und unterliegen keinerlei Streuungen. Die Streuungen werden durch die statistischen Parameter \mathbf{s} modelliert. Entwurfparameter können beispielsweise Transistorlängen/Weiten oder Kapazitätswerte

sein. Für die Entwurfparameter sind Untergrenzen \mathbf{d}_L und Obergrenzen \mathbf{d}_U gegeben. Diese Grenzen definieren einen n_d -dimensionalen Hyperquader, den zulässigen Bereich $\mathcal{E}_d(\mathbf{d}_L, \mathbf{d}_U)$, innerhalb dessen die Parameterwerte liegen müssen:

$$\mathcal{E}_d(\mathbf{d}_L, \mathbf{d}_U) = \{\mathbf{d} \mid \mathbf{d}_L \leq \mathbf{d} \leq \mathbf{d}_U\} \quad (2.2)$$

Die Untergrenzen \mathbf{d}_L sind meist durch die Fertigungstechnologie vorgegeben, beispielsweise definieren die minimalen Strukturgrößen der verwendeten Technologie die Untergrenzen für die Längen und Weiten der Transistoren. Weitere Grenzen ergeben sich durch strukturelle Nebenbedingungen (Abschnitt 2.1.3) oder müssen vom Schaltungsentwickler festgelegt werden. Das Intervall zwischen der Unter- und Obergrenze eines Entwurfparameters definiert den möglichen Änderungsbereich dieses Parameters und kann somit wie unten diskutiert auch zur Normierung der Entwurfparameter verwendet werden.

Statistische Parameter

Die statistischen Parameter \mathbf{s} beschreiben die unvermeidlichen Streuungen des Fertigungsprozesses (z.B. Schwankungen der Oxiddicke t_{ox} , der Dotierung oder der Schwellspannungen V_{th}). Die statistischen Parameter sind in der Regel normalverteilt (z.B. Schwellspannungen V_{th} , Weitenreduktion dw) oder logarithmisch normalverteilt (z.B. Oxiddicke t_{ox} , Beweglichkeit μ_0). Wie in [Esh92] gezeigt, können viele Verteilungen auf eine Gaußsche Normalverteilung zurückgeführt werden. Dies gilt insbesondere auch für eine logarithmische Normalverteilung. Deshalb soll im weiteren für die statistischen Parameter eine Normalverteilung $\mathbf{s} \sim N(\mathbf{s}_0, \mathbf{C})$ mit dem Erwartungswert \mathbf{s}_0 und der Kovarianzmatrix \mathbf{C} angenommen werden.

Die Kovarianzmatrix \mathbf{C} setzt sich aus der Streumatrix $\mathbf{\Sigma}$ und der Korrelationsmatrix \mathbf{R} zusammen.

$$\mathbf{C} = \mathbf{\Sigma} \cdot \mathbf{R} \cdot \mathbf{\Sigma} \quad (2.3)$$

Die Streumatrix $\mathbf{\Sigma}$ enthält die Streuungen σ_i der statistischen Parameter, die im allgemeinen von den Entwurfparametern abhängen können. Dies gilt insbesondere bei Berücksichtigung lokaler (*Intra-die*) Schwankungen, die nach dem Flächengesetz [PDW89] (z.B. $\sigma_{V_{th}} \sim 1/\sqrt{WL}$) modelliert werden.

$$\mathbf{\Sigma}(\mathbf{d}) = \begin{bmatrix} \sigma_1(\mathbf{d}) & 0 & \cdots & 0 \\ 0 & \sigma_2(\mathbf{d}) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{n_s}(\mathbf{d}) \end{bmatrix}, \quad \sigma_i(\mathbf{d}) > 0 \quad (2.4)$$

Ein Element $\rho_{i,j} = \rho_{j,i}$ der symmetrischen und positiv definiten Korrelationsmatrix \mathbf{R} beschreibt die Korrelation zwischen den Parameter s_i und s_j . Für die Korrelationen wird angenommen, daß sie unabhängig von den Entwurfparametern \mathbf{d} sind.

$$\mathbf{R} = \begin{bmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,n_s} \\ \rho_{2,1} & 1 & & \vdots \\ \vdots & & \ddots & \\ \rho_{n_s,1} & \cdots & \rho_{n_s,n_s-1} & 1 \end{bmatrix}, \quad |\rho_{i,j}| \leq 1 \quad (2.5)$$

Die Kovarianzmatrix \mathbf{C} nach Gl. (2.3) ist symmetrisch und positiv definit. Mit den eingeführten Größen ergibt sich für die statistischen Parameter folgende Wahrscheinlichkeitsdichtefunktion der Normalverteilung [JW92]:

$$\text{pdf}(\mathbf{s}, \mathbf{s}_0, \mathbf{C}) = \frac{1}{\sqrt{2\pi}^{n_s} \sqrt{\det \mathbf{C}}} \cdot \exp\left(-\frac{1}{2}(\mathbf{s} - \mathbf{s}_0)^T \cdot \mathbf{C}^{-1} \cdot (\mathbf{s} - \mathbf{s}_0)\right) \quad (2.6)$$

Betriebsparameter

Die Betriebsparameter $\boldsymbol{\theta}$ beschreiben alle Parameter, die sich während des Betriebes der Schaltung ändern können. Bei integrierten Schaltungen können dies beispielsweise die Temperatur oder Versorgungsspannungen sein. Für die Betriebsparameter sind im Datenbuch einer Schaltung Untergrenzen $\boldsymbol{\theta}_L$ und Obergrenzen $\boldsymbol{\theta}_U$ gegeben. Diese Grenzen definieren den Toleranzbereich $\mathcal{T}_\theta(\boldsymbol{\theta}_L, \boldsymbol{\theta}_U)$, innerhalb dessen die Schaltung funktionieren muß.

$$\mathcal{T}_\theta(\boldsymbol{\theta}_L, \boldsymbol{\theta}_U) = \{\boldsymbol{\theta} | \boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U\} \quad (2.7)$$

Desweiteren sind für die Betriebsparameterwerte auch Nominalwerte $\boldsymbol{\theta}_0$ definiert, die typische Betriebsbedingungen beschreiben. Die Nominalwerte müssen innerhalb des Toleranzbereiches \mathcal{T}_θ liegen. Oft werden sie jeweils in der Mitte zwischen Ober- und Untergrenze gewählt.

Normierung der Parameter

Die einzelnen Parameter p_i können unterschiedliche Einheiten haben und von völlig unterschiedlichen Größenordnungen sein (z. B. Transistorweiten im Bereich von μm und Widerstandswerte von $\text{k}\Omega$). Für einen schnellen Erfolg einer Optimierung werden allerdings einheitenlose Parameter benötigt, deren Änderungen sich typischerweise in der gleichen Größenordnung bewegen [NW99, Spe98, GMW81]. Dazu müssen die Parameter auf typische Änderungsbereiche Δp_i normiert werden:

$$\mathbf{P}_{\text{normiert}} = \begin{bmatrix} 1/\Delta p_1 & 0 & \cdots & 0 \\ 0 & 1/\Delta p_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1/\Delta p_{n_p} \end{bmatrix} \cdot \mathbf{P}_{\text{unnormiert}} \quad (2.8)$$

In der Praxis hat sich die in [Eck98, Pör89] vorgeschlagene Normierung auf die Parametergrenzen bzw. Streuungen bewährt und soll deshalb in dieser Arbeit verwendet werden.

$$\Delta d_i = d_{U,i} - d_{L,i} \quad (2.9)$$

$$\Delta s_i = \sigma_i \quad (2.10)$$

$$\Delta \theta_i = \theta_{U,i} - \theta_{L,i} \quad (2.11)$$

Für den Rest der Arbeit werden die Parameter und die von den Parametern abhängigen Matrizen und Vektoren als normiert angenommen.

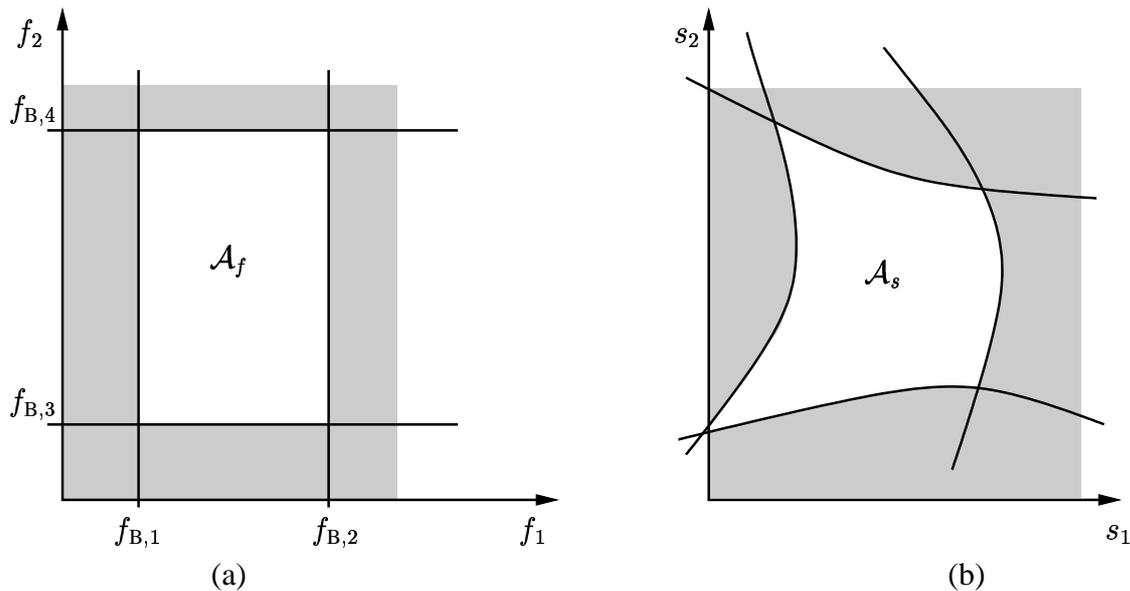


Abbildung 2.1: Akzeptanzgebiet der (a) Eigenschaften und (b) statistischen Parameter ohne Berücksichtigung von Betriebsparametern.

2.1.2 Schaltungseigenschaften, Spezifikationen und Ausbeute

Schaltungseigenschaften

Bei gegebener Schaltungstopologie können die Schaltungsparameter \mathbf{p} mit Hilfe eines Simulators in den Raum der *Schaltungseigenschaften* \mathbf{f} (z. B. Verstärkung, Verzögerungszeit) abgebildet werden:

$$\mathbf{p} \mapsto \mathbf{f}(\mathbf{p}), \quad \mathbf{f} \in \mathcal{R}^{n_f} \quad (2.12)$$

Die Abbildung der Parameter auf die Schaltungseigenschaften ist im allgemeinen eindeutig. Eine eindeutige Rückabbildung dagegen vom Raum der Schaltungseigenschaften in den Parameterraum existiert normalerweise nicht.

Spezifikationen

Die *Spezifikationen* \mathbf{f}_B beschreiben Anforderungen an die Schaltung. Für die Schaltungseigenschaften \mathbf{f} können Untergrenzen \mathbf{f}_L und/oder Obergrenzen \mathbf{f}_U gegeben sein:

$$f_i \stackrel{!}{\geq} f_{L,i} \quad \text{für eine Untergrenze} \quad (2.13)$$

$$f_i \stackrel{!}{\leq} f_{U,i} \quad \text{für eine Obergrenze} \quad (2.14)$$

Ohne Einschränkung der Allgemeinheit kann jede Untergrenze durch Multiplikation mit -1 in eine Obergrenze transformiert werden. Im weiteren werden deshalb alle Spezifikationen \mathbf{f}_B als Obergrenzen angenommen.

Das Akzeptanzgebiet im Eigenschaftsraum $\mathcal{A}_{f,i}$ bezüglich einer Spezifikation $f_{B,i}$ ist das Gebiet, in dem die Spezifikationen $f_{B,i}$ erfüllt ist:

$$\mathcal{A}_{f,i} = \{f_i(\mathbf{p}) | f_i(\mathbf{p}) \leq f_{B,i}\} \quad (2.15)$$

Das Akzeptanzgebiet \mathcal{A}_f aller Schaltungseigenschaften bestimmt sich aus der Schnittmenge der einzelnen Akzeptanzgebiete:

$$\mathcal{A}_f = \bigcap_{i=1, \dots, n_B} \mathcal{A}_{f,i} \quad (2.16)$$

Wie in Abbildung 2.1(a) angedeutet, ist das Akzeptanzgebiet \mathcal{A}_f der Eigenschaften ein n_f -dimensionaler Hyperquader. Ein Ziel des Nominalentwurfs ist es, die Schaltung so zu dimensionieren, daß für typische Betriebsparameter $\boldsymbol{\theta}_0$ und ohne Berücksichtigung von Prozeßstreuungen ($\mathbf{s} = \mathbf{s}_0$) die Schaltungseigenschaften im Akzeptanzgebiet liegen, also alle Spezifikationen erfüllt sind.

Ausbeute

Das Akzeptanzgebiet $\mathcal{A}_{s,i}(\mathbf{d})$ der statistischen Parameter bezüglich einer Spezifikation ist die Menge aller statistischen Parameter, für welche die Schaltung für alle Betriebsparameter aus dem Toleranzbereich \mathcal{T}_θ die Spezifikation $f_{B,i}$ erfüllt:

$$\mathcal{A}_{s,i}(\mathbf{d}) = \left\{ \mathbf{s} \mid \max_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{f_i(\mathbf{d}, \mathbf{s}, \boldsymbol{\theta})\} \leq f_{B,i} \right\} \quad (2.17)$$

Das Akzeptanzgebiet $\mathcal{A}_s(\mathbf{d})$ der statistischen Parameter berechnet sich wiederum als Schnittmenge aller einzelnen Akzeptanzgebiete $\mathcal{A}_{s,i}$.

$$\mathcal{A}_s(\mathbf{d}) = \bigcap_{i=1, \dots, n_B} \mathcal{A}_{s,i}(\mathbf{d}) \quad (2.18)$$

Abbildung 2.1(b) veranschaulicht das Akzeptanzgebiet \mathcal{A}_s der statistischen Parameter. Die *Ausbeute* Y , der Anteil der funktionsfähigen Schaltungen an den gesamten produzierten Schaltungen kann bei Vernachlässigung katastrophischer Fehler (z. B. Kurzschlüsse zwischen zwei Leitbahnen) durch ein Integral berechnet werden.

$$Y(\mathbf{d}) = \int_{\mathcal{A}_s(\mathbf{d})} \text{pdf}(\mathbf{s}, \mathbf{s}_0, \mathbf{C}) \, d\mathbf{s} \quad (2.19)$$

Die Definition der Ausbeute entsprechend Gl. (2.19) wird auch als *parametrische Ausbeute* (*parametric yield*) bezeichnet, da sie außer Parameterschwankungen keine weiteren Fehlerquellen (z. B. Spot-Defekte, Elektromigration, ESD) berücksichtigt. Da für technisch relevante Schaltungen im allgemeinen keine eindeutige Abbildung vom Eigenschaftsraum in den Parameterraum existiert, kann weder das Akzeptanzgebiet $\mathcal{A}_s(\mathbf{d})$, noch die Ausbeute $Y(\mathbf{d})$ analytisch bestimmt werden. Vielmehr wird das Integral (2.19) üblicherweise mit einer Monte-Carlo Analyse abgeschätzt. Ziel einer Dimensionierung ist es, die Ausbeute unter gegebenen Randbedingungen (z. B. gegebene Obergrenze für die Fläche) zu maximieren.

2.1.3 Strukturelle Nebenbedingungen

Schaltungseigenschaften beschreiben das Schaltungsverhalten auf Zell- bzw. Blockebene. Aus einer Systemsimulation auf Zell- resp. Blockebene können deshalb nur Spezifikationen für die Schaltungseigenschaften vorgegeben werden. Schaltungstechnisch sinnvoll dimensionierte analoge Schaltungen zeichnen sich aber nicht nur dadurch aus, daß sie diese Spezifikationen erfüllen. Vielmehr müssen eine Vielzahl weiterer Bedingungen eingehalten werden, die eine hohe Robustheit der Schaltung sicherstellen. Bei einer manuellen Dimensionierung werden diese Regeln zumeist intuitiv durch den Entwickler eingehalten. Um eine erfolgreiche automatische Dimensionierung analoger Schaltungen zu ermöglichen, müssen diese Regeln formalisiert und in den Entwurf integriert werden.

In [EGG98, Eck98, ZEG98, GZEA01] wird gezeigt, daß für den Erfolg einer automatischen rechnergestützten Dimensionierung folgende Überlegungen wichtig sind:

- Schaltungsspezifikationen werden in der Regel nur für die Eigenschaften spezifiziert, die für den geplanten Einsatzzweck der Schaltung kritisch erscheinen. Die Beachtung allein dieser Spezifikationen stellt allerdings noch keinen robusten und physikalisch sinnvollen Schaltungsentwurf sicher und führt teilweise dazu, daß normalerweise nicht spezifizierte Eigenschaften unakzeptable Werte annehmen und somit die Schaltung unbrauchbar für die geplante Anwendung wird.
- Schaltungsspezifikationen werden auf einer höheren Abstraktionsebene definiert und können somit nicht Schaltungszustände auf einer tieferen Abstraktionsebene kontrollieren. Eine robust dimensionierte analoge Schaltung zeichnet sich aber zumeist dadurch aus, daß alle Transistoren die ihnen zugeordneten Funktionen erfüllen (z. B. spannungsgesteuerte Stromquelle), sich also in einem definierten Zustand befinden.
- Ein Grundprinzip des analogen Schaltungsentwurfs ist das Arbeiten mit Transistorpaaren. Um die Empfindlichkeit der Transistorpaare gegenüber lokalen Streuungen (*Mismatch*) klein zu halten, müssen spezielle Regeln eingehalten werden.

Wie in [Eck98, ZEG98, GZEA01] gezeigt, können basierend auf diesen Überlegungen Regeln eingeführt werden, die eine robuste und schaltungstechnisch sinnvolle Dimensionierung sicherstellen. Die Regeln beschreiben Anforderungen an einzelne Transistoren, Transistorpaare oder Teilschaltungen (z. B. Kaskode-Stromspiegel). Für jede Grundstruktur (z. B. Stromspiegel, Level-Shifter, etc.) wird ein Satz von Dimensionierungsnebenbedingungen definiert, die eine sinnvolle und robuste Dimensionierung sicherstellen. Wird eine solche Grundstruktur innerhalb einer Schaltung erkannt, kann ihr der vorher definierte Satz von Nebenbedingungen (*strukturelle Nebenbedingungen*) zugeordnet werden.

Strukturelle Nebenbedingungen können anhand verschiedener Kriterien klassifiziert werden. Im Rahmen der automatischen Dimensionierung ist allerdings vor allen Dingen die Einteilung in Gleichheits- und Ungleichheitsnebenbedingungen von Interesse:

- Jede Gleichheitsnebenbedingung definiert zwei oder mehr Parameter die unmittelbar voneinander abhängen, also durch einen einzelnen unabhängigen Parameter ersetzt werden können. Jede Gleichheitsnebenbedingung reduziert somit die Anzahl der unabhängigen Parameter und verringert so die Komplexität des Problems. Die Gleichheitsnebenbedingungen werden durch eine

Struktur	Ungleichheitsbedingungen	Gleichheitsbedingungen	Summe
Spannungsgesteuerte Stromquelle (VCCS)	3	—	3
Spannungsgesteuerter Widerstand	3	—	3
Stromspiegel	12	2	14
Level-Shifter	13	2	15
Differenzpaar	12	2	14
Komplementärpaar	—	1	1
Kaskodeschaltung	—	1	1

Tabelle 2.1: Überblick über strukturelle Nebenbedingungen für typische analoge CMOS-Grundstrukturen.

Reduktion der Entwurfparameter berücksichtigt und tauchen deshalb in der weiteren Problemformulierung nicht mehr auf. In Kapitel 4 wird an Beispielen gezeigt, wie durch die Gleichheitsnebenbedingungen die Anzahl der unabhängigen Entwurfparameter signifikant reduziert wird.

- Ungleichheitsnebenbedingungen definieren Grenzen im Parameterraum. Dies können zum einen unmittelbar Parametergrenzen sein, die den zulässigen Bereich \mathcal{E}_d der Entwurfparameter definieren. Andere Ungleichheitsnebenbedingungen beschreiben innere Schaltungszustände (z. B. Sättigungsforderungen für Transistoren), welche die prinzipielle Funktionsfähigkeit der Schaltung garantieren. Zur Auswertung dieser Nebenbedingungen ist eine DC-Analyse ausreichend, die im Rahmen der Eigenschaftsauswertung normalerweise auf jeden Fall benötigt wird. Es entsteht also kein zusätzlicher Simulationsaufwand.

Tabelle 2.1 gibt einen Überblick über die strukturellen Nebenbedingungen am Beispiel von CMOS-Grundstrukturen. Ohne Einschränkung der Allgemeinheit können die Ungleichheitsnebenbedingungen $\mathbf{u}(\mathbf{d}, \mathbf{s}, \boldsymbol{\theta})$ so formuliert werden, daß erfüllte Nebenbedingungen positiv sind:

$$\mathbf{u}(\mathbf{d}, \mathbf{s}, \boldsymbol{\theta}) \geq \mathbf{0}, \quad \mathbf{u} \in \mathcal{R}^{n_u} \quad (2.20)$$

Eine gut dimensionierte analoge Schaltung zeichnet sich dadurch aus, daß alle Spezifikationen \mathbf{f}_B und gleichzeitig alle Nebenbedingungen \mathbf{u} erfüllt sind.

2.2 Problemstellung des Nominalentwurfs

Ziel des Nominalentwurfs ist es, die Entwurfparameter so einzustellen, daß die Schaltung für vorgegebene Werte der Betriebsparameter und der statistischen Parameter alle Spezifikationen und Nebenbedingungen erfüllt.

In dieser Arbeit soll die Aufgabenstellung des Nominalentwurfs dahingehend erweitert werden, daß die Eigenschaften unter Berücksichtigung des Einflusses der Betriebsparameter zusätzlich noch so viel Abstand wie möglich zur Spezifikation haben sollen und so ein guter Ausgangspunkt für die

simulationsintensive Entwurfszentrierung bestimmt wird [SSGA00]. Schaltungsbeispiele in Kapitel 4 zeigen, daß auf diese Weise die Gesamtzahl der im Entwurfsablauf benötigten Simulationen reduziert werden kann, obwohl sich die Anzahl der Schaltungssimulationen im Nominalentwurf erhöht.

Zur Bewertung der einzelnen Eigenschaften und zur Bestimmung einer Kostenfunktion im Rahmen des Nominalentwurfs wird eine Zielgröße benötigt, die

- den Einfluß der Betriebsparameter berücksichtigt,
- Eigenschaften mit unterschiedlichen Einheiten und Größenordnungen miteinander vergleichbar macht,
- ein Übertreffen der Spezifikationen unterstützt und
- mit einem möglichst geringen Aufwand berechnet werden kann.

In [Eck98, SEGA99b] wird der Parameterabstand α ohne Berücksichtigung der Betriebsparameter als Zielgröße für den Nominalentwurf vorgeschlagen. Im folgenden soll als ein Beitrag dieser Arbeit die Definition des Parameterabstandes um den Einfluß der Betriebsparameter erweitert werden [SSGA00].

2.2.1 Der Parameterabstand

Der Parameterabstand $\alpha_i(\mathbf{d}_0)$ beschreibt eine Abschätzung für die minimale Parameterauslenkung $\|\Delta \mathbf{d}\|$, die nötig ist um ausgehend von einem Parametersatz \mathbf{d}_0 für eine Schaltungseigenschaft f_i die Spezifikationsgrenze $f_{B,i}$ zu erreichen. Er ist damit ein Maß für den Abstand der Eigenschaft f_i von ihrer Spezifikation $f_{B,i}$. Der Parameterabstand basiert dabei auf einer Empfindlichkeitsanalyse, die im Rahmen des Optimierverfahrens (Kapitel 3) auf jeden Fall gemacht werden muß. Es entsteht deshalb kein zusätzlicher Simulationsaufwand.

In dem hier diskutierten Verfahren sind die statistischen Parameter im gesamten Nominalentwurf auf ihre Erwartungswerte \mathbf{s}_0 gesetzt. Sie können aber auch auf jeden anderen festen Wert gesetzt werden. Werden „eher digitale“ Schaltungen untersucht, so ist es sinnvoll anstatt der Erwartungswerte \mathbf{s}_0 der statistischen Parameter eventuell vorgegebene Eckparametersätze (z.B. *slow* oder *fast* Parametersätze) zu benutzen. Da die statistischen Parameter innerhalb des Nominalentwurfs sich nicht ändern, werden sie in diesem Abschnitt weggelassen.

Zur Berechnung des Parameterabstandes $\alpha_i(\mathbf{d}_0)$ wird ein lineares Modell $\bar{f}_i(\mathbf{d}, \boldsymbol{\theta})$ der Eigenschaft f_i im Punkt $(\mathbf{d}_0, \boldsymbol{\theta}_0)$ bestimmt:

$$\begin{aligned} \bar{f}_i(\mathbf{d}, \boldsymbol{\theta}) &= f_i(\mathbf{d}_0, \boldsymbol{\theta}_0) + \mathbf{g}_{\mathbf{d},i}^T \cdot (\mathbf{d} - \mathbf{d}_0) + \mathbf{g}_{\boldsymbol{\theta},i}^T \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \\ \text{mit } \mathbf{g}_{\mathbf{d},i} &= \nabla_{\mathbf{d}} f_i |_{\mathbf{d}_0, \boldsymbol{\theta}_0}, \quad \mathbf{g}_{\boldsymbol{\theta},i} = \nabla_{\boldsymbol{\theta}} f_i |_{\mathbf{d}_0, \boldsymbol{\theta}_0} \end{aligned} \quad (2.21)$$

In diesem linearen Modell $\bar{f}_i(\mathbf{d}, \boldsymbol{\theta})$ interessiert vor allem die Menge der Grenzparametersätze $\mathcal{D}_{B,i}$. Grenzparametersätze sind die Entwurfsparametersätze \mathbf{d} , für welche die linearisierte Eigenschaft im abgeschätzten Worst-Case-Betriebsparametersatz gleich der Spezifikation $f_{B,i}$ ist:

$$\mathcal{D}_{B,i} = \left\{ \mathbf{d} \mid \max_{\boldsymbol{\theta} \in \mathcal{T}_{\boldsymbol{\theta}}} \{ \bar{f}_i(\mathbf{d}, \boldsymbol{\theta}) \} = f_{B,i} \right\} \quad (2.22)$$

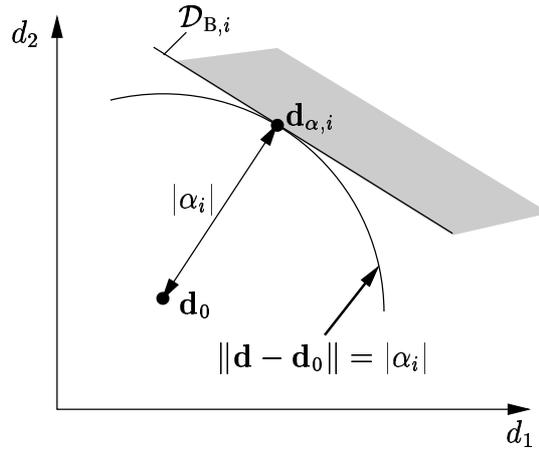


Abbildung 2.2: Veranschaulichung des Parameterabstandes.

Die Menge der Grenzparametersätze ist eine $(n_d - 1)$ -dimensionale Hyperebene, in der veranschaulichenden Abbildung 2.2 mit zwei Entwurfsparametern also eine Gerade. In dieser Menge der Grenzparametersätze $\mathcal{D}_{B,i}$ gibt es wiederum einen ausgezeichneten Parametersatz $\mathbf{d}_{\alpha,i}$, der Parametersatz mit dem geringsten Abstand $\|\mathbf{d} - \mathbf{d}_0\|$ vom Linearisierungspunkt \mathbf{d}_0 . Der Betrag des Abstandes vom Linearisierungspunkt ist der Betrag des Parameterabstandes α_i :

$$\mathbf{d}_{\alpha,i} = \operatorname{argmin}_{\mathbf{d} \in \mathcal{D}_{B,i}} \{(\mathbf{d} - \mathbf{d}_0)^T \cdot (\mathbf{d} - \mathbf{d}_0)\} \quad (2.23)$$

$$|\alpha_i| = \|\mathbf{d}_{\alpha,i} - \mathbf{d}_0\| \quad (2.24)$$

Das Vorzeichen des Parameterabstandes α_i wird so bestimmt, daß der Parameterabstand positiv ist, wenn die Spezifikation im linearen Modell $\bar{f}_i(\mathbf{d}, \boldsymbol{\theta})$ im abgeschätzten Worst-Case-Betriebsparametersatz und $\mathbf{d} = \mathbf{d}_0$ erfüllt ist:

$$\alpha_i(\mathbf{d}_0) = \begin{cases} +|\alpha_i(\mathbf{d}_0)| & \text{für } \max_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{\bar{f}_i(\mathbf{d}_0, \boldsymbol{\theta})\} \leq f_{B,i} \\ -|\alpha_i(\mathbf{d}_0)| & \text{für } \max_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{\bar{f}_i(\mathbf{d}_0, \boldsymbol{\theta})\} > f_{B,i} \end{cases} \quad (2.25)$$

Um sinnvolle Werte für den Parameterabstand α_i zu erhalten, ist zu beachten, daß die Entwurfsparameter normiert sind wie in Kapitel 2.1.1 beschrieben.

Berechnung des Parameterabstandes

Der Parameterabstand basiert auf der Linearisierung einer Schaltungseigenschaft. Die Aufgabe der Bestimmung des Parameterabstandes kann deshalb unterteilt werden in zwei getrennte Problemstellungen; Abschätzen der Worst-Case-Betriebsparameter und Berechnung der Auslenkung der Entwurfsparameter. Der Einfluß der Betriebsparameter kann abgeschätzt werden durch:

$$\bar{\boldsymbol{\theta}}_{\text{WC}_i} = \operatorname{argmax}_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{\bar{f}_i(\mathbf{d}_0, \boldsymbol{\theta})\} \quad (2.26)$$

$$\bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}) = \bar{f}_i(\mathbf{d}, \bar{\boldsymbol{\theta}}_{\text{WC}_i}) \quad (2.27)$$

Gl. (2.26) ist ein separierbares lineares Optimierproblem auf dem Simplex \mathcal{T}_θ . Separierbar bedeutet hier, daß eine Nebenbedingung jeweils parallel zu einer Parameterrichtung und orthogonal zu allen anderen Parameter ist. Daher kann jeder Betriebsparameter unabhängig von den anderen Betriebsparametern auf seine Ober- oder Untergrenze gesetzt werden und somit Gl. (2.26) mit linearer Komplexität $\mathcal{O}(n_\theta)$ gelöst werden. Der abgeschätzte Worst-Case-Betriebsparametersatz $\bar{\theta}_{\text{WC},i}$ liegt in einer Ecke des Toleranzbereiches \mathcal{T}_θ [Dan63] und ist unabhängig von den Entwurfsparametern. Er muß in jedem Schritt des Optimieralgorithmus nur einmal berechnet werden. Die Differenz $\bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}_0) - \bar{f}_i(\mathbf{d}_0, \theta_0)$ ist die Verschiebung des linearisierten Eigenschaftswertes durch die Betriebsparameter.

Der Punkt $\mathbf{d}_{\alpha,i}$ kann mittels eines Lagrange-Ansatzes berechnet werden. Das Optimierproblem zu Gl. (2.23) lautet

$$\mathbf{d}_{\alpha,i} = \underset{\mathbf{d}}{\operatorname{argmin}} \{ \|\mathbf{d} - \mathbf{d}_0\|^2 \mid \bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}) = f_{\text{B},i} \} \quad (2.28)$$

mit der zugehörigen Lagrange-Funktion $L(\mathbf{d}, \lambda)$:

$$L(\mathbf{d}, \lambda) = \|\mathbf{d} - \mathbf{d}_0\|^2 + \lambda \cdot (\bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}) - f_{\text{B},i}) \quad (2.29)$$

Die hinreichenden Extremalbedingungen für ein Minimum lauten [Spe98]:

$$\nabla_{\mathbf{d}} L(\mathbf{d}, \lambda) = \mathbf{0} \quad (\text{Lagrange Bedingung}) \quad (2.30)$$

$$\bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}) - f_{\text{B},i} = 0 \quad (\text{Nebenbedingung}) \quad (2.31)$$

$$\nabla_{\mathbf{d}\mathbf{d}}^2 L(\mathbf{d}, \lambda) > 0 \quad (\text{Minimumbedingung}) \quad (2.32)$$

Die positive Definitheit der Hesse Matrix Gl. (2.32) ist mit $\nabla_{\mathbf{d}\mathbf{d}}^2 L(\mathbf{d}, \lambda) = \mathbf{E}$ für alle $\mathbf{d} \in \mathcal{R}^{n_d}$ garantiert. Setzt man die linearisierte Eigenschaft Gl. (2.21) in die Lagrange Bedingung Gl. (2.30) ein, läßt sich $\mathbf{d}_{\alpha,i}$ bestimmen:

$$\mathbf{d}_{\alpha,i} - \mathbf{d}_0 = -\frac{\lambda}{2} \cdot \mathbf{g}_{\mathbf{d},i} \quad (2.33)$$

Aus der Nebenbedingung Gl. (2.31) unter Berücksichtigung der Linearisierung Gl. (2.21) bestimmt sich λ zu:

$$\lambda = 2 \cdot \frac{\bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}_0) - f_{\text{B},i}}{\mathbf{g}_{\mathbf{d},i}^T \mathbf{g}_{\mathbf{d},i}} \quad (2.34)$$

Verwendet man dieses Ergebnis in Gl. (2.33), so ergibt sich daraus:

$$\mathbf{d}_{\alpha,i} - \mathbf{d}_0 = \frac{|f_{\text{B},i} - \bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}_0)|}{\|\mathbf{g}_{\mathbf{d},i}\|^2} \cdot \mathbf{g}_{\mathbf{d},i} \quad (2.35)$$

Mit dem in Gl. (2.25) definierten Vorzeichen kann der Parameterabstand α_i als explizite Gleichung geschrieben werden:

$$\alpha_i(\mathbf{d}_0) = \frac{f_{\text{B},i} - \bar{f}_{i,\theta_{\text{WC}}}(\mathbf{d}_0)}{\|\mathbf{g}_{\mathbf{d},i}\|} \quad (2.36)$$

Eigenschaften des Parameterabstandes

Der Parameterabstand nach Gl. (2.36) kann interpretiert werden als Abstand der Eigenschaft von ihrer Spezifikation normiert auf die Beeinflußbarkeit, gemessen als Länge des Gradienten.

Mit dem Parameterabstand steht eine Zielgröße für den Nominalentwurf zur Verfügung, die

- unterschiedliche Eigenschaften vergleichbar macht,
- sowohl den Abstand einer Eigenschaft zu ihrer Spezifikation berücksichtigt als auch die Beeinflussbarkeit dieser Eigenschaft,
- den Einfluß der Betriebsparameter auf das Schaltungsverhalten abschätzt,
- ein Übertreffen der Spezifikationen unterstützt und dabei auch auf der Spezifikationsgrenze differenzierbar bleibt und
- keinen zusätzlichen Simulationsaufwand verursacht über eine Empfindlichkeitsanalyse hinaus, die im Rahmen des Optimieralgorithmus auf jeden Fall gemacht werden muß.

Die Berechnung des Parameterabstandes aus den simulierten Empfindlichkeiten ist mit linearer Komplexität $\mathcal{O}(n_d + n_\theta)$ möglich. Die Rechenzeit zur Berechnung des Parameterabstandes kann im Vergleich zur Simulationszeit in der Regel vernachlässigt werden.

2.2.2 Zielfunktion für den Nominalentwurf

Für die Optimierung müssen die Parameterabstände als Zielgrößen des Nominalentwurfs in eine skalare Zielfunktion transformiert werden. Dabei ergibt sich wie später auch in der Entwurfszentrierung die Aufgabe, daß verletzten Spezifikationen – negativen Parameterabständen – ein hohes Gewicht zukommen muß und erfüllte Spezifikationen einen niedrigen Anteil an der Zielfunktion haben sollten.

Mit der in [LD89, Grä93, Wie94, KD95] vorgeschlagene Gewichtung der Robustheitsmaße mit der Exponentialfunktion bestimmt sich die Zielfunktion $\varphi_\alpha(\mathbf{d}_0)$ für den Nominalentwurf zu:

$$\varphi_\alpha(\mathbf{d}_0) = \sum_{i=1}^{n_B} \exp(-a \cdot \alpha_i(\mathbf{d}_0)), \quad a > 0 \quad (2.37)$$

Der positive Faktor a in Gl. (2.37) ist ein Skalierungsfaktor [LD89, KD95] für die Parameterabstände, die Zielgrößen des Nominalentwurfs. Ziel des Nominalentwurfs ist eine Minimierung der Zielfunktion $\varphi_\alpha(\mathbf{d}_0)$ unter Einhaltung aller Nebenbedingungen:

$$\min_{\mathbf{d}_0 \in \mathcal{E}_d} \{ \varphi_\alpha(\mathbf{d}_0) \mid \mathbf{u}(\mathbf{d}_0) \geq \mathbf{0} \} \quad (2.38)$$

Die Lösung von Gl. (2.38) ist eine Dimensionierung, die alle Spezifikationen unter Berücksichtigung des Einflusses der Betriebsparameter so gut wie möglich erfüllt. Insbesondere bricht die Optimierung nach Erreichen aller Spezifikationen nicht ab, sondern versucht sie noch soweit möglich zu übertreffen. Dieses Übertreffen schafft einen Sicherheitsabstand zur Spezifikation, wodurch ein günstiger Ausgangspunkt für eine nachfolgende simulationsaufwendige Entwurfszentrierung bestimmt wird.

2.3 Problemstellung der Entwurfszentrierung

Ziel der Entwurfszentrierung ist es, die parametrische Ausbeute Y Gl. (2.19) zu maximieren. Dies entspricht einer Dimensionierung „tief im Inneren“ des Akzeptanzgebietes \mathcal{A}_s der statistischen Parameter. Mit Hilfe des Worst-Case-Abstands [AGW94, Grä93, Wie94, SSGA99, AEG⁺00a] können das Akzeptanzgebiet $\mathcal{A}_{s,i}$ und die Ausbeute Y_i bezüglich einer Spezifikation approximiert werden und der Gradient der Ausbeute Y_i bestimmt werden.

2.3.1 Der Worst-Case-Abstand

Der Worst-Case-Abstand ist analog zum Parameterabstand definiert, nur daß hier der Abstand zur Spezifikation im Raum der statistischen Parameter bestimmt wird und kein lineares Schaltungsmodell verwendet wird sondern vielmehr die nichtlineare Schaltung selbst.

Zur Formulierung des Worst-Case-Abstands $\beta_{\text{WC},i}$ wird im Raum der statistischen Parameter zunächst die Menge der Grenzparametersätze $\mathcal{S}_{\text{B},i}$ definiert [SGA99, AEG⁺00a]:

$$\mathcal{S}_{\text{B},i}(\mathbf{d}_0) = \left\{ \mathbf{s} \mid \max_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{f_i(\mathbf{d}_0, \mathbf{s}, \boldsymbol{\theta})\} = f_{\text{B},i} \right\} \quad (2.39)$$

Aus der Menge der Grenzparametersätze wird der Punkt bestimmt, der den geringsten Abstand vom Erwartungswert \mathbf{s}_0 hat, gemessen in der mit der inversen Kovarianzmatrix \mathbf{C}^{-1} gewichteten l_2 -Norm. Dieses ist der Worst-Case-Punkt $\mathbf{s}_{\text{WC},i}$ und dessen Abstand wird als Betrag des Worst-Case-Abstandes $|\beta_{\text{WC},i}|$ definiert:

$$\mathbf{s}_{\text{WC},i}(\mathbf{d}_0) = \underset{\mathbf{s} \in \mathcal{S}_{\text{B},i}(\mathbf{d}_0)}{\text{argmin}} \left\{ (\mathbf{s} - \mathbf{s}_0)^{\text{T}} \cdot \mathbf{C}^{-1} \cdot (\mathbf{s} - \mathbf{s}_0) \right\} \quad (2.40)$$

$$|\beta_{\text{WC},i}(\mathbf{d}_0)| = \sqrt{(\mathbf{s}_{\text{WC},i}(\mathbf{d}_0) - \mathbf{s}_0)^{\text{T}} \cdot \mathbf{C}^{-1} (\mathbf{s}_{\text{WC},i}(\mathbf{d}_0) - \mathbf{s}_0)} \quad (2.41)$$

Punkte gleicher Wahrscheinlichkeitsdichtefunktion liegen nach Gl. (2.6) auf Ellipsoiden $(\mathbf{s} - \mathbf{s}_0)^{\text{T}} \mathbf{C}^{-1} (\mathbf{s} - \mathbf{s}_0) = \text{const}$ um den Nominalpunkt \mathbf{s}_0 . Der Worst-Case-Punkt $\mathbf{s}_{\text{WC},i}$ ist daher der Punkt aus der Menge der Grenzparametersätze, der die höchste Auftrittswahrscheinlichkeit hat. Das Vorzeichen des Worst-Case-Abstandes $\beta_{\text{WC},i}$ wird analog zum Parameterabstand so definiert, daß der Worst-Case-Abstand positiv ist, wenn für nominale statistische Parameter ($\mathbf{s} = \mathbf{s}_0$) die Spezifikation für alle Betriebsparameter $\boldsymbol{\theta}$ aus dem Toleranzbereich \mathcal{T}_θ erfüllt ist [Wie94, Grä93]:

$$\beta_{\text{WC},i}(\mathbf{d}_0) = \begin{cases} + |\beta_{\text{WC},i}(\mathbf{d}_0)| & \text{für } \max_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{f_i(\mathbf{d}_0, \mathbf{s}_0, \boldsymbol{\theta})\} \leq f_{\text{B},i} \\ - |\beta_{\text{WC},i}(\mathbf{d}_0)| & \text{für } \max_{\boldsymbol{\theta} \in \mathcal{T}_\theta} \{f_i(\mathbf{d}_0, \mathbf{s}_0, \boldsymbol{\theta})\} > f_{\text{B},i} \end{cases} \quad (2.42)$$

Abbildung 2.3(a) veranschaulicht die Berechnung des Worst-Case-Abstandes. Linearisiert man das Akzeptanzgebiet der statistischen Parameter im Worst-Case-Punkt (Abbildung 2.3(b)), so kann der Worst-Case-Abstand als Streuungsvielfaches interpretiert werden [Grä93, Wie94]. Die abgeschätzte Ausbeute \bar{Y}_i für das linearisierte Akzeptanzgebiet $\bar{\mathcal{A}}_{\mathbf{s},i}$ berechnet sich zu:

$$\bar{Y}_i(\mathbf{d}_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_{\text{WC},i}(\mathbf{d}_0)} \exp\left(-\frac{\xi^2}{2}\right) d\xi \quad (2.43)$$

$\beta_{\text{WC},i}$	-3	-2	-1	0	1	2	3
\bar{Y}_i	0.1%	2.3%	15.9%	50%	84.1%	97.7%	99.9%

Tabelle 2.2: Approximierte Ausbeute \bar{Y}_i in Abhängigkeit vom Worst-Case-Abstand $\beta_{\text{WC},i}$

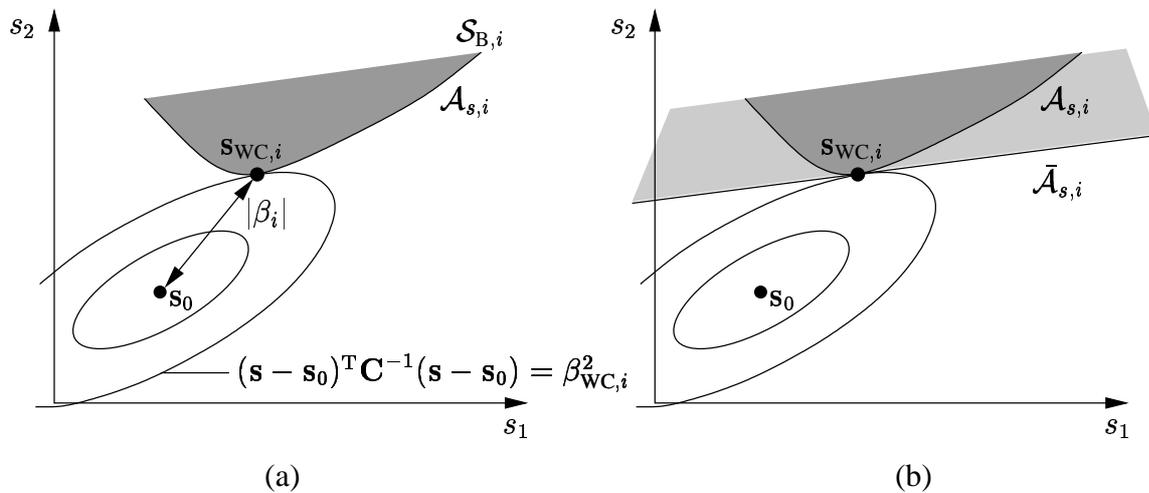


Abbildung 2.3: Berechnung des (a) Worst-Case-Abstand $\beta_{WC,i}$ und (b) linearisierten Akzeptanzgebietes $\bar{\mathcal{A}}_{s,i}$ der statistischen Parameter.

Berechnung des Worst-Case-Abstandes

Der Worst-Case-Abstand ist vergleichbar mit dem Parameterabstand, nur daß er im Vergleich zum Parameterabstand die Empfindlichkeit bezüglich der statistischen Parameter beschreibt und eine direkte Interpretation als Streuungsvielfaches hat. Die exakte Berechnung des Worst-Case-Punktes $\mathbf{s}_{WC,i}$ und damit auch des Worst-Case-Abstandes ist für den Fall einer nichtlinearer Abhängigkeit der Schaltungseigenschaften von den Parametern ein skalares quadratisches Optimierproblem mit einer nichtlinearen Gleichheitsnebenbedingung.

Verfahren zur Berechnung des Worst-Case-Punktes $\mathbf{s}_{WC,i}$ finden sich in [AEG⁺00a, Wie94, AGW94, SSGA99]. In [Wie94, AGW94] wird ein SQP-Verfahren [Sch81] verwendet, [AEG⁺00a] diskutiert ein SQP-Verfahren mit konjugierter Gradienten Eigenschaft, das insbesondere im Fall lokaler Parameterschwankungen vorteilhaft ist. In beiden Fällen wird die benötigte Rechenzeit von der Simulationszeit dominiert. Die Anzahl der benötigten Simulationen ist linear in der Anzahl der statistischen Parameter.

Eigenschaften des Worst-Case-Abstands

Mit dem Worst-Case-Abstand steht eine Zielgröße für die Entwurfszentrierung zur Verfügung mit der

- das Akzeptanzgebiet der statistischen Parameter für eine Eigenschaft approximiert werden kann,
- die Ausbeute ohne weiteren Simulationsaufwand abgeschätzt werden kann (Tabelle 2.2) und
- schaltungs- und eigenschaftsspezifische Worst-Case-Parametersätze bestimmt werden können. Basierend auf diesen Worst-Case-Parametersätzen wiederum können kritische statistische Parameter erkannt werden [AEG⁺00a].

Die Qualität der Approximation des Akzeptanzgebietes und der Ausbeuteabschätzung hängt dabei entscheidend von der Linearität der zugrundeliegenden Eigenschaft ab. Da die Linearisierung aller-

dings in dem Grenzparametersatz mit der höchsten Auftrittswahrscheinlichkeit erfolgt, liefert sie in der Regel eine gute Abschätzung.

2.3.2 Zielfunktion der Entwurfszentrierung

Analog zum Nominalentwurf müssen auch bei der Entwurfszentrierung die Zielgrößen in eine skalare Zielfunktion transformiert werden. Ziel ist es, alle Worst-Case-Abstände so groß wie möglich zu machen, dabei aber insbesondere negative oder kleine Worst-Case-Abstände zu verbessern. Diese schränken die Ausbeute am meisten ein. Die Zielfunktion muß also auch in der Entwurfszentrierung den kleinen bzw. negativen Zielgrößen ein hohes Gewicht geben. Wie in [LD89, KD95, Wie94, Grä93] vorgeschlagen, sollen in der Entwurfszentrierung die Worst-Case-Abstände mittels einer Summe von Exponentialfunktionen in eine skalare Zielfunktion $\varphi_\beta(\mathbf{d}_0)$ transformiert werden:

$$\varphi_\beta(\mathbf{d}_0) = \sum_{i=1}^{n_f} \exp(-a \cdot \beta_{\text{WC},i}(\mathbf{d}_0)), \quad a > 0 \quad (2.44)$$

Der positive Faktor a in Gl. (2.44) ist wie in Gl. (2.37) ein Gewichtungsfaktor für die Worst-Case-Abstände. Ein großer Gewichtungsfaktor a verleiht kleinen und negativen Worst-Case-Abständen ein großes Gewicht und vernachlässigt große Worst-Case-Abstände. Gleichzeitig verschlechtert sich aber die Kondition des Optimierproblems. Für $a \rightarrow \infty$ trägt nur noch der kleinste Worst-Case-Abstand zur Zielfunktion bei. Ziel ist auch hier eine Minimierung der Zielfunktion $\varphi_\beta(\mathbf{d}_0)$ unter Einhaltung aller Nebenbedingungen:

$$\min_{\mathbf{d}_0 \in \mathcal{E}_d} \{\varphi_\beta(\mathbf{d}_0) \mid \mathbf{u}(\mathbf{d}_0) \geq \mathbf{0}\} \quad (2.45)$$

Die Minimierprobleme, die den Nominalentwurf Gl. (2.38) und den Toleranzentwurf Gl. (2.45) beschreiben, sind nichtlineare Optimierprobleme mit Nebenbedingungen. Im nächsten Kapitel wird eine Methodik beschrieben, diese Minimierprobleme mit einem vertretbarem Aufwand zu lösen.

Kapitel 3

Dimensionierung

In Kapitel 2 wurden die Problemstellung des Nominalentwurfs (Abschnitt 2.2) und der Entwurfszentrierung (Abschnitt 2.3) auf ein nichtlineares Minimierproblem mit Nebenbedingungen zurückgeführt. Die Zielfunktion für den Nominalentwurf $\varphi_\alpha(\mathbf{d})$ und die Entwurfszentrierung $\varphi_\beta(\mathbf{d})$ haben den gleichen Aufbau. Der wesentliche Unterschied liegt im Berechnungsaufwand für die Zielgrößen. Innerhalb der Entwurfszentrierung müssen für die Berechnung der Worst-Case Abstände mehrere Optimierprobleme [Wie94, AEG⁺00a] gelöst werden. Dies führt, wie an Schaltungsbeispielen in Kapitel 4 gezeigt, zu einem deutlich höheren Rechenaufwand verglichen mit der Berechnung der Parameterabstände, weshalb eine Trennung der Dimensionierung in Nominalentwurf und Entwurfszentrierung sinnvoll ist.

Die Charakteristiken der beiden Entwurfsschritte sind in der Formulierung der Zielgrößen bereits berücksichtigt. Dieses Kapitel konzentriert sich deshalb auf die Lösung des Optimierproblems Gl. (2.38) bzw. Gl. (2.45) mit Methoden, welche die Eigenheiten des analogen Schaltungsentwurfs berücksichtigen. Die Zielfunktion $\varphi(\mathbf{d}_0)$, die bei Nominalentwurf und Entwurfszentrierung von der Form her gleich ist, wird dabei angenommen zu:

$$\varphi(\mathbf{d}_0) = \sum_{i=1}^{n_\gamma} \exp(-a \cdot \gamma_i(\mathbf{d}_0)), \quad a > 0 \quad (3.1)$$

Das Optimierproblem ergibt sich unter Verwendung von Gl. (3.1) zu:

$$\min_{\mathbf{d}_0 \in E_d} \{\varphi(\mathbf{d}_0) \mid \mathbf{u}(\mathbf{d}_0) \geq \mathbf{0}\} \quad (3.2)$$

In der Nominalpunktdimensionierung werden die Zielgrößen γ_i durch die Parameterabstände α_i ersetzt, in der Entwurfszentrierung durch die Worst-Case Abstände $\beta_{\text{WC},i}$. Dadurch ergibt sich trotz ihrer mathematisch identischen Form jeweils eine unterschiedliche Interpretation der Zielfunktion in beiden Entwurfsschritten. Das vorgestellte Verfahren ist nicht auf Parameterabstände bzw. Worst-Case Abstände beschränkt. Wie später gezeigt, ist es vielmehr auf alle Probleme anwendbar, für welche die linearisierten Zielgrößen in eine skalare konvexe Zielfunktion transformiert werden.

Gl. (3.2) ist ein nichtlineares Minimierproblem mit Nebenbedingungen. Aufgabe der Dimensionierung ist es, eine Lösung dieses Minimierproblems mit einer akzeptablen Simulationszahl zu bestimmen. Ziel ist dabei nicht eine extrem genaue Bestimmung des optimalen Parametersatzes, sondern

vielmehr eine möglichst schnelle Berechnung eines akzeptablen Parametersatzes in der Nähe des Optimums. Die geforderte Genauigkeit ergibt sich beispielsweise aus der Bedingung, daß Transistorweiten und Transistorlängen nur auf einem vordefinierten Gitter liegen dürfen.

Wie bereits in der Einleitung beschrieben ist bei der Lösung des Minimierproblems zu beachten, daß die Auswertung der Zielfunktion rechenzeitintensiv und somit teuer ist. Optimierverfahren, die innerhalb einer schlecht parallelisierbaren Liniensuche eine Vielzahl von Auswertungen der Zielfunktion benötigen, sind deshalb nicht sinnvoll anwendbar.

3.1 Überblick

Nichtlineare Optimierprobleme können außer in Sonderfällen nicht analytisch gelöst werden. Für die Lösung dieser Probleme sind, wie in der Einleitung beschrieben stochastische Methoden (z.B. Simulated Annealing[KCGV83]) und deterministische Methoden verfügbar. Deterministische Methoden haben dabei den Vorteil, zumeist schneller zu konvergieren, d.h. sie benötigen deutlich weniger Funktionsauswertungen, die bei der Dimensionierung zeitaufwendige Simulationen bedeuten. Dafür kann im allgemeinen aber nur die Konvergenz zu einem lokalen Optimum bewiesen werden.

Aufgrund der schnelleren Konvergenzeigenschaften soll in dieser Arbeit ein deterministisches Verfahren Verwendung finden. Von den deterministischen Verfahren sind speziell die *Trust-Region Algorithmen* [Sor82] für schlecht konditionierte Probleme geeignet [DEAM97, DV97]. Zudem benötigen die Trust-Region Algorithmen keine Liniensuche sondern kombinieren vielmehr die Bestimmung einer Suchrichtung und Schrittweite. Der verwendete Algorithmus gliedert sich dabei in folgende Schritte [NW99, Spe98, GMW81]:

1. Setze den Iterationsindex $\mu = 0$ und die Entwurfparameter $\mathbf{d}_0^{(\mu)}$ auf die gegebenen Startwerte.
2. Linearisiere im Linearisierungspunkt $\mathbf{d}_0^{(\mu)}$ die Zielgrößen γ

$$\begin{aligned}\gamma(\mathbf{d}) &= \underbrace{\gamma(\mathbf{d}_0^{(\mu)})}_{\gamma_0^{(\mu)}} + \underbrace{\nabla_{\mathbf{d}}\gamma|_{\mathbf{d}=\mathbf{d}_0^{(\mu)}}}_{\mathbf{S}^{(\mu)}} \cdot \underbrace{(\mathbf{d} - \mathbf{d}_0^{(\mu)})}_{\mathbf{x}} + \dots \\ \bar{\gamma}(\mathbf{x}) &= \gamma_0^{(\mu)} + \mathbf{S}^{(\mu)} \cdot \mathbf{x}\end{aligned}\quad (3.3)$$

und die Nebenbedingungen \mathbf{u}

$$\begin{aligned}\mathbf{u}(\mathbf{d}) &= \underbrace{\mathbf{u}(\mathbf{d}_0^{(\mu)})}_{\mathbf{u}_0^{(\mu)}} + \underbrace{\nabla_{\mathbf{d}}\mathbf{u}|_{\mathbf{d}=\mathbf{d}_0^{(\mu)}}}_{\mathbf{U}^{(\mu)}} \cdot \underbrace{(\mathbf{d} - \mathbf{d}_0^{(\mu)})}_{\mathbf{x}} + \dots \\ \bar{\mathbf{u}}(\mathbf{x}) &= \mathbf{u}_0^{(\mu)} + \mathbf{U}^{(\mu)} \cdot \mathbf{x}\end{aligned}\quad (3.4)$$

3. Bestimme basierend auf der Linearisierung der Zielgrößen eine Parameterkorrektur $\mathbf{x}^{(\mu)}$.
4. Überprüfe die Abbruchkriterien für die Optimierung. Sind die Abbruchkriterien nicht erfüllt, dann bestimme den neuen Linearisierungspunkt $\mathbf{d}_0^{(\mu+1)}$

$$\mathbf{d}_0^{(\mu+1)} = \mathbf{d}_0^{(\mu)} + \mathbf{x}^{(\mu)}, \quad (3.5)$$

erhöhe den Iterationsindex

$$\mu = \mu + 1, \quad (3.6)$$

und starte die nächste Iteration bei Schritt 2.

Der entscheidende Punkt in diesem Algorithmus ist die Bestimmung einer Parameterkorrektur in Schritt 3. In diesem Punkt muß eine Parameterkorrektur berechnet werden, für die sich einerseits eine deutliche Reduktion der Zielfunktion im linearisierten Schaltungsmodell ergibt und andererseits die Linearisierung noch gültig ist, also die Fehlerreduktion auch in der nichtlinearen Schaltung signifikant ist. Insbesondere sollte der Schaltungssimulator im neuen Parametersatz konvergieren. Extreme, schaltungstechnisch nicht sinnvolle Auslenkungen resultieren oft in einer Schaltung für die der Simulator nur sehr schlecht konvergiert. In diesem Fall ist mit sehr langen Simulationszeiten selbst für kleine Schaltungen zu rechnen.

Im Falle einer quadratischen Zielfunktion (*least-square optimization*) hat sich für die Berechnung einer Schrittweite die charakteristische Grenzkurve [ALP88, Hus82, Lei86, Pör89] bewährt. Aufbauend auf den Ideen der charakteristischen Grenzkurve soll in dieser Arbeit ein vergleichbares Verfahren für konvexe Optimierprobleme – eine konvexe Zielfunktion auf einem konvexen Gebiet vorgestellt werden. Dazu wird zunächst die charakteristische Grenzkurve diskutiert. Anschließend wird ausgehend von der charakteristischen Grenzkurve die verallgemeinerte Grenzkurve als eine Methode zur Schrittweitenbestimmung für die gewählte Zielfunktion Gl. (3.1) vorgestellt.

3.2 Die charakteristische Grenzkurve

Die charakteristische Grenzkurve mit Nebenbedingungen dient zur Bestimmung der Schrittweite für ein quadratisches Optimierproblem:

$$\begin{aligned} \min_{\mathbf{x}} \{ \|\boldsymbol{\epsilon}_F + \mathbf{G}\mathbf{x}\|^2 \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \}, \\ \text{mit } \boldsymbol{\epsilon}_F = \boldsymbol{\epsilon}(\mathbf{d}_0), \quad \mathbf{G} = \nabla_{\mathbf{d}}\boldsymbol{\epsilon}(\mathbf{d}_0) \end{aligned} \quad (3.7)$$

Der Vektor $\boldsymbol{\epsilon}$ faßt die „Fehler“ der Schaltung zusammen. Die Formulierung des Problems nach Gl. (3.7) beschreibt die Optimierung des Fehlers auf Wunschwerte, die alle als Null angenommen werden; ein Übertreffen dieser Wunschwerte wird nicht unterstützt.

3.2.1 Anpassung der Zielfunktion

Ziel ist eine Minimierung der Zielfunktion nach Gl. (3.1). Beachtet man, daß jeder einzelne Summand in dieser Zielfunktion immer positiv ist, so kann die Minimierung der Gesamtfunktion als Minimierung des Quadrates aller ihrer Terme beschrieben werden. Paßt man die Zielfunktion derart an, so ergibt sich eine Komponente ϵ_i des Fehlervektors $\boldsymbol{\epsilon}$ zu [SEGA99b]:

$$\epsilon_i(\mathbf{d}_0) = \exp(-a \cdot \gamma_i(\mathbf{d}_0)) \quad (3.8)$$

Die Jacobimatrix \mathbf{G} berechnet sich darauf aufbauend zu:

$$\mathbf{G} = -a \cdot \text{diag}(\boldsymbol{\epsilon}(\mathbf{d}_0)) \cdot \mathbf{S} \quad (3.9)$$

Der $\text{diag}()$ Operator transformiert hierbei einen Vektor in eine Diagonalmatrix.

3.2.2 Definition der charakteristischen Grenzkurve

Wie oben diskutiert ist es nicht das Ziel, in jedem Iterationsschritt das Optimierproblem exakt zu lösen, sondern vielmehr eine Lösung mit deutlicher Fehlerreduktion zu finden, für welche die Linearisierung noch gültig ist. Es muß also ein Kompromiß zwischen Auslenkung und Fehlerreduktion gefunden werden. Dazu wird die Auslenkung gemessen als $\|\mathbf{x}\|^2$ mit einem positiven Faktor λ gewichtet [Hus82, Lei86, Pör89]:

$$\min_{\mathbf{x}} \{ \|\boldsymbol{\epsilon}_F + \mathbf{G}\mathbf{x}\|^2 + \lambda \cdot \|\mathbf{x}\|^2 \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \}, \quad \lambda \geq 0 \quad (3.10)$$

Für $\lambda \rightarrow \infty$ ergibt sich die Lösung mit minimalem Abstand vom Linearisierungspunkt, die alle linearisierten Nebenbedingungen erfüllt; $\lambda = 0$ dagegen löst das ursprüngliche Optimierproblem. Die Lösung von Gl. (3.10) in Abhängigkeit von λ wird mit $\mathbf{x}_c(\lambda)$ und der linearisierte Fehlervektor an dieser Stelle mit $\bar{\boldsymbol{\epsilon}}_c(\lambda)$ bezeichnet:

$$\mathbf{x}_c(\lambda) = \operatorname{argmin}_{\mathbf{x}} \{ \|\boldsymbol{\epsilon}_F + \mathbf{G}\mathbf{x}\|^2 + \lambda \cdot \|\mathbf{x}\|^2 \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \} \quad (3.11)$$

$$\bar{\boldsymbol{\epsilon}}_c(\lambda) = \boldsymbol{\epsilon}_F + \mathbf{G} \cdot \mathbf{x}_c(\lambda) \quad (3.12)$$

Abbildung 3.1(a) zeigt beispielhaft die parametrische Kurve ($\|\mathbf{x}_c(\lambda)\|$, $\|\bar{\boldsymbol{\epsilon}}_c(\lambda)\|$) und veranschaulicht die Extremwerte für $\lambda \rightarrow \infty$ und $\lambda = 0$. Um verschiedene Probleme vergleichen zu können, werden die Auslenkung $\|\mathbf{x}_c(\lambda)\|$ und der linearisierte Schaltungsfehler $\|\bar{\boldsymbol{\epsilon}}_c(\lambda)\|$ so transformiert, daß die normierte Auslenkung $a(\lambda)$ und der normierte Schaltungsfehler $r(\lambda)$ zwischen 0 und 1 liegen:

$$a(\lambda) = \frac{\|\mathbf{x}_c(\lambda)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|}{\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|} \quad (3.13)$$

$$r(\lambda) = \frac{\check{r}(\lambda) - \check{r}(0)}{1 - \check{r}(0)} \quad (3.14)$$

$$\text{mit } \check{r}(\lambda) = \sqrt{\frac{\|\bar{\boldsymbol{\epsilon}}_c(\lambda)\|^2 - \|\bar{\boldsymbol{\epsilon}}_0\|^2}{\|\bar{\boldsymbol{\epsilon}}_c(\lambda \rightarrow \infty)\|^2 - \|\bar{\boldsymbol{\epsilon}}_0\|^2}}, \quad \text{und} \quad (3.15)$$

$$\|\bar{\boldsymbol{\epsilon}}_0\|^2 = \min_{\mathbf{x}} \{ \|\boldsymbol{\epsilon}_F + \mathbf{G} \cdot \mathbf{x}\|^2 \} \quad (3.16)$$

Der Fehler $\|\bar{\boldsymbol{\epsilon}}_0\|$ in Gl. (3.16) ist das Minimum des linearisierten Problems ohne Nebenbedingungen. Für ein unterbestimmtes oder reguläres System ($n_d \geq n_B$) ist dieser Wert immer Null. Für ein überbestimmtes System läßt sich der Fehlervektor $\boldsymbol{\epsilon}$ mittels einer QR-Zerlegung [GvL96] der Empfindlichkeitsmatrix \mathbf{G} so transformieren, daß ein Teil des transformierten Fehlervektors vollständig zu Null reduziert werden kann, während der Rest nicht beeinflussbar ist. Abbildung 3.1 zeigt ein typisches Beispiel für die parametrische Kurve ($a(\lambda)$, $r(\lambda)$), die charakteristische Grenzkurve.

3.2.3 Eigenschaften der charakteristischen Grenzkurve

Im Rahmen eines angepaßten Levenberg-Marquardt Verfahrens [Lev44, Mar63] sind folgende Eigenschaften bedeutend für die Verwendbarkeit der charakteristischen Grenzkurve zur Schrittweitenbestimmung [ALP88, Pör89]:

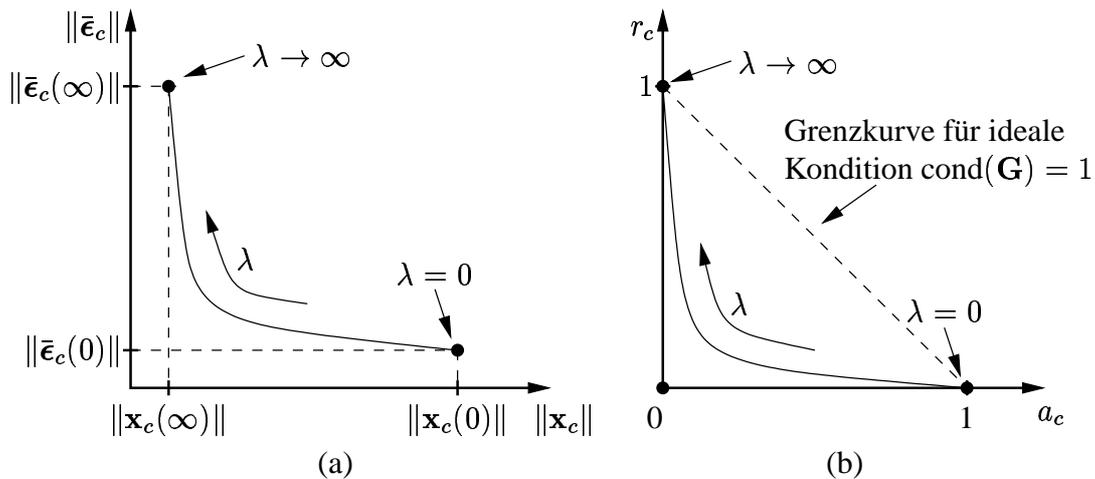


Abbildung 3.1: Darstellung der Auslenkung und des Schaltungsfehlers für die charakteristische Grenzkurve in (a) unnormierter ($\|\mathbf{x}_c\|$, $\|\bar{\mathbf{e}}_c\|$) und (b) normierter (a , r) Darstellung in Abhängigkeit von λ .

- Die normierte Auslenkung und der normierte Schaltungsfehler liegen immer im Intervall $[0 \dots 1]$. Somit sind die Kurven verschiedener Probleme untereinander vergleichbar.
- Die Kurve beschreibt die Lösungen mit minimalem Fehler im linearisierten System für eine gegebene maximale Auslenkung [Mor77]. Im linearisierten System gibt es keine Lösungen, die unterhalb der Kurve liegen. Die Kurve stellt damit die Menge der paretooptimalen Punkte in linearisierten System dar. An diesen Punkten ist es nicht möglich gleichzeitig den Fehler und die Auslenkung zu verringern.
- Die Kurve ist sowohl im normierten als auch im unnormierten System konvex. Je weiter man sich vom Linearisierungspunkt wegbewegt, je größer $\|\mathbf{x}_c(\lambda)\|$, desto flacher wird die Kurve.
- Für ein Problem ohne Nebenbedingungen und mit ideal konditionierter Empfindlichkeitsmatrix \mathbf{G} (Spektrale Konditionszahl [GvL96]: $\text{cond}(\mathbf{G}) = 1$) ist die normierte Kurve $(a(\lambda), r(\lambda))$ eine Gerade. Je schlechter die Kondition der Empfindlichkeitsmatrix, desto stärker ist die Kurve gekrümmt. Wie in der Einleitung angedeutet, ist eine schlechte Kondition der Empfindlichkeitsmatrix aber typisch für den analogen Schaltungsentwurf. Die charakteristische Grenzkurve ist deshalb zumeist stark gekrümmt.

Die Konvexität zusammen mit der typischerweise starken Krümmung der Grenzkurve (Abbildung 3.1) bedeutet, daß am Beginn der Kurve für $\lambda \rightarrow \infty$ mit einer kleinen Auslenkung eine große Fehlerreduktion erreicht werden kann. Bis zum Knickpunkt wird der Fehler r normalerweise stark reduziert. Nach dem Knickpunkt verläuft die Kurve typischerweise relativ flach; eine weitere Auslenkung resultiert nur mehr in einer geringfügigen Fehlerreduktion. Eine erste sinnvolle Schrittweite ist demnach die Auslenkung bis zum Knickpunkt.

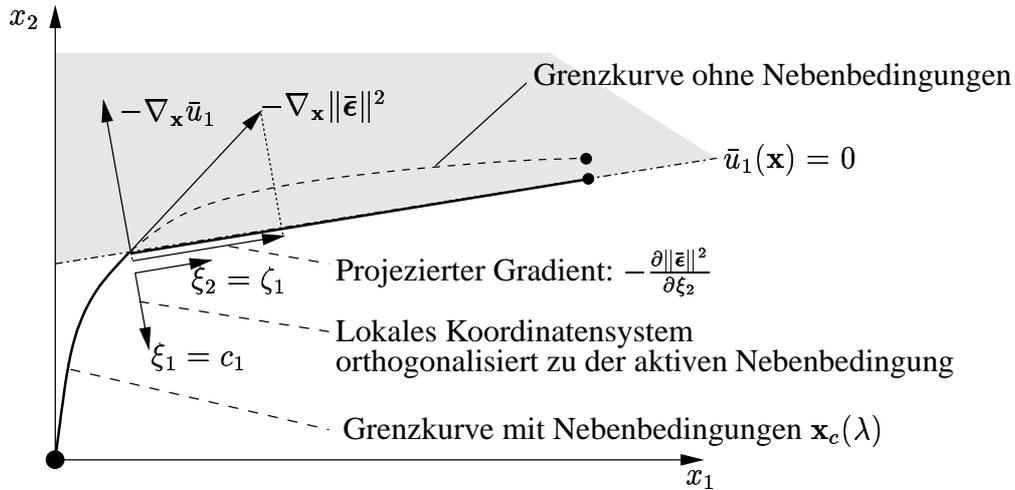


Abbildung 3.2: Grenzcurve mit einer Nebenbedingungen und Veranschaulichung der Transformation orthogonal zu der aktiven Nebenbedingung.

3.2.4 Berechnung von Punkten auf der Grenzcurve

Die Lösung des Minimierproblems Gl. (3.11) kann für ein Dimensionierungsproblem ohne Nebenbedingungen explizit formuliert werden [ALP88, Lei86]:

$$\mathbf{x}_c(\lambda) = -(\mathbf{G}^T \mathbf{G} + \lambda \cdot \mathbf{E})^+ \mathbf{G}^T \boldsymbol{\epsilon}_F, \quad \text{für } n_u = 0 \quad (3.17)$$

Hierbei bezeichnet $(\mathbf{G}^T \mathbf{G} + \lambda \cdot \mathbf{E})^+$ die *Moore-Penrose Inverse (Pseudoinverse)* [GvL96]. Mit Nebenbedingungen kann das Minimierproblems Gl. (3.11) nicht mehr explizit gelöst werden. In [Pör89] wird das Optimierproblem mit dem Beale-Algorithmus [Neu75] gelöst. Der Beale-Algorithmus ist eine Erweiterung des Simplex-Verfahrens [Dan63] auf eine quadratische Zielfunktion. Allerdings kann der Algorithmus wie auch der Simplex Algorithmus im schlimmsten Fall exponentielle Komplexität haben.

Für die Lösung quadratischer Optimierprobleme mit linearen Nebenbedingungen existieren aber auch Algorithmen mit quadratischer Komplexität $\mathcal{O}(n_B \cdot n_d)$, beispielsweise Algorithmen der primalen quadratischen Programmierung (PQP) [Sto71, GM78, Spe98, NW99]. Ist bereits eine Lösung für einen Wert von λ bekannt, ermöglichen diese Algorithmen auch einen schnellen „Restart“ für die Berechnung einer benachbarten Lösung. Ändert sich die Menge der aktiven Nebenbedingungen für die benachbarte Lösung nicht, dann wird die neue Lösung in einem Schritt berechnet. Eine Nebenbedingung u_i ist dann aktiv, wenn sie gerade noch erfüllt ist ($u_i = 0$) und der Gradient der Nebenbedingung und der Zielfunktion in die gleiche Richtung zeigen, ihr Skalarprodukt positiv ist. Dies ist in Abbildung 3.2 veranschaulicht.

Der PQP-Algorithmus liefert zusätzlich zur Lösung des Minimierproblems auch die Lagrange-Multiplikatoren und die Orthogonalisierungsmatrix \mathbf{Q} der aktiven Nebenbedingungen. Mit Hilfe dieser Orthogonalisierungsmatrix kann das Optimierproblem mit aktiven Ungleichheitsnebenbedingungen in ein Optimierproblem ohne aktive Nebenbedingungen transformiert werden. Sind n_{aktiv} Nebenbedingungen aktiv und ist $\mathbf{N} \in \mathcal{R}^{n_{\text{aktiv}} \times n_d}$ die Empfindlichkeitsmatrix (Jacobimatrix) der aktiven

Nebenbedingungen so gilt für die Transformation:

$$\mathbf{Q} \cdot \mathbf{N}^T = \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{R} \in \mathcal{R}^{n_{\text{aktiv}} \times n_{\text{aktiv}}} \text{ obere Dreiecksmatrix} \quad (3.18)$$

$$\boldsymbol{\xi} = \mathbf{Q} \mathbf{x} = \begin{bmatrix} \mathbf{c} \\ \boldsymbol{\zeta} \end{bmatrix}, \quad \mathbf{c} \in \mathcal{R}^{n_{\text{aktiv}}} \quad (3.19)$$

Die Systemmatrix \mathbf{G} und die Empfindlichkeitsmatrix \mathbf{U} der Nebenbedingungen werden ebenfalls entsprechend transformiert. Um innerhalb des linearisierten zulässigen Bereichs zu bleiben, darf der Wert der aktiven Nebenbedingungen nicht mehr verändert werden. Dies ist gleichbedeutend damit, daß im transformierten $\boldsymbol{\xi}$ -Raum die Werte der ersten n_{aktiv} Parameter, in Gl. (3.19) mit \mathbf{c} bezeichnet, nicht variiert werden dürfen. Die restlichen Parameter, in Gl. (3.19) in dem Vektor $\boldsymbol{\zeta}$ zusammengefaßt, dagegen dürfen variiert werden. In Abbildung 3.2 ist die Orthogonalisierung und das lokale $\boldsymbol{\xi}$ -Koordinatensystem veranschaulicht.

Jede aktive Nebenbedingung verringert die Anzahl der freien Entwurfsparameter um eins. Nach [Lei86] können bei der Reduktion der Entwurfsparameter zwei Fälle unterschieden werden:

Überbestimmtes System: In einem überbestimmten System kann bei der Reduktion der Entwurfsparameter eine Konditionsverbesserung der reduzierten Systemmatrix bewiesen werden. Das bedeutet, daß eine aktive Nebenbedingung in einem überbestimmten System immer zu einer Konditionsverbesserung führt. Ein System ist auch dann überbestimmt, wenn die Anzahl der freien Entwurfsparameter $n_d - n_{\text{aktiv}}$ kleiner oder gleich der Anzahl der Zielgrößen ist.

Unterbestimmtes System: In einem unterbestimmten System kann keine Aussage über eine Verbesserung der Kondition der Systemmatrix gemacht werden. Es kann jedoch häufig eine Verbesserung der Kondition erwartet werden.

Bei einer zunehmenden Zahl aktiver Nebenbedingungen kann also eine Verbesserung der Kondition der zu den aktiven Nebenbedingungen orthogonalisierten Systemmatrix erwartet bzw. bewiesen werden.

3.2.5 Berechnung der Steigung der Grenzkurve

Neben den Werten für den transformierten Fehler $r(\lambda)$ und die Auslenkung $a(\lambda)$ ist es ohne wesentlichen Zusatzaufwand auch möglich, die Steigung $m(a, r, \lambda)$ an einem berechneten Punkt $(a(\lambda), r(\lambda))$ zu bestimmen. Im Gegensatz zu [Pör89] soll die Steigung der Grenzkurve hier so bestimmt werden, daß die Ergebnisse später auf die verallgemeinerte Grenzkurve übertragen werden können.

Wie in [Lei86, Pör89] gezeigt, kann der Parameter λ als Lagrange-Parameter für die Nebenbedingung $\|\mathbf{x}\|^2 \leq \|\mathbf{x}(\lambda)\|^2$ interpretiert werden. In [Spe98, NW99] wird gezeigt, daß der zu einer Nebenbedingung gehörige Lagrange Parameter die Empfindlichkeit der Zielfunktion bezüglich dieser Nebenbedingung beschreibt (auch als *Schattenpreise* bezeichnet). Mit der Zielfunktion $\|\bar{\mathbf{e}}_c\|^2$ und der Nebenbedingung $\|\mathbf{x}\|^2 \leq \|\mathbf{x}_c(\lambda)\|^2$ kann die Steigung der unnormierten Grenzkurve bestimmt werden zu:

$$\frac{\partial \|\bar{\mathbf{e}}_c(\lambda)\|^2}{\partial \|\mathbf{x}_c(\lambda)\|^2} = -\lambda \quad (3.20)$$

Die Steigung $m(\lambda)$ der normierten Grenzkurve bestimmt sich durch Differenzierung der Normierungsfunktionen:

$$m(\lambda) = \frac{r(\lambda)}{a(\lambda)} = \frac{\partial r(\lambda)}{\partial \|\bar{\mathbf{e}}_c(\lambda)\|^2} \cdot \frac{\partial \|\bar{\mathbf{e}}_c(\lambda)\|^2}{\partial \|\mathbf{x}_c(\lambda)\|^2} \cdot \frac{\partial \|\mathbf{x}_c(\lambda)\|^2}{\partial a(\lambda)}, \quad \text{mit} \quad (3.21)$$

$$\frac{\partial r(\lambda)}{\partial \|\bar{\mathbf{e}}_c(\lambda)\|^2} = \frac{1}{\check{r}(\lambda) \cdot (1 - \check{r}(0))} \cdot \frac{1}{2 (\|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2 - \|\bar{\mathbf{e}}_0\|^2)} \quad (3.22)$$

$$\frac{\partial \|\mathbf{x}_c(\lambda)\|^2}{\partial a(\lambda)} = 2 \cdot \|\mathbf{x}_c(\lambda)\| \cdot (\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|) \quad (3.23)$$

Aus Gl. (3.20) bis Gl. (3.23) ergibt sich die Steigung zu:

$$m(\lambda) = -\lambda \cdot \frac{\|\mathbf{x}_c(\lambda)\| (\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|)}{(\|\bar{\mathbf{e}}_0\|^2 - \|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2) (1 - \check{r}(0)) \sqrt{\frac{\|\bar{\mathbf{e}}_0\|^2 - \|\bar{\mathbf{e}}_c(\lambda)\|^2}{\|\bar{\mathbf{e}}_0\|^2 - \|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2}}} \quad (3.24)$$

Setzt man die Definitionen für die transformierte Auslenkung a und den Fehler r nach Gl. (3.13) bzw. (3.14) ein, so ergibt sich die in [Pör89] hergeleitete Formel für die Steigung:

$$m(\lambda) = -\lambda \cdot \frac{a(\lambda) \cdot (\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|) + \|\mathbf{x}_c(\lambda \rightarrow \infty)\|}{r(\lambda) \cdot (1 - \check{r}(0)) + \check{r}(0)} \cdot \frac{\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|}{(1 - \check{r}(0)) \cdot (\|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2 - \|\bar{\mathbf{e}}_0\|^2)} \quad (3.25)$$

Da die Steigung in einem bereits berechneten Punkt ohne wesentlichen Zusatzaufwand berechnet werden kann, sollte sie bei der Approximation der Grenzkurve verwendet werden. Dazu wird im nächsten Abschnitt eine Approximation der Grenzkurve basierend auf den Tangenten in den bereits berechneten Punkten vorgestellt werden.

3.2.6 Abschnittsweise lineare Approximation der Grenzkurve

Die Grenzkurve ist nicht als analytische Funktion gegeben. Vielmehr stellt die Berechnung jedes Punktes auf der Grenzkurve die Lösung eines Optimierproblems dar, für das allerdings keine Simulationen benötigt werden. Die Grenzkurve kann also nicht als eine geschlossene Funktion bestimmt werden, sondern vielmehr sind nur einzelne Punkte auf der Grenzkurve und die Steigung der Kurve an diesen Punkten berechenbar. Zur Darstellung der Grenzkurve und zur Berechnung von Punkten mit vorgegebenem Fehler r , vorgegebener Auslenkung a oder Steigung m ist es nötig, die Grenzkurve auf der Basis einzelner berechneter Punkte zu approximieren.

In diesem Abschnitt wird eine in [Pör89] vorgeschlagene abschnittsweise lineare Approximation der Grenzkurve diskutiert. Die Grundlage für dieses Verfahren bildet die Approximation der Grenzkurve zwischen zwei berechneten Punkten (a_1, r_1) und (a_2, r_2) , wobei im folgenden $a_1 < a_2$ angenommen wird.

Abschnittsweise lineare Approximation der Grenzkurve zwischen zwei Punkten

In [Pör89] wird vorgeschlagen, die Kurve zwischen zwei berechneten Punkten durch ihre Tangenten zu approximieren (Abbildung 3.3). Stellt man den approximierten Schaltungsfehler \bar{r}_L in Abhängigkeit der Auslenkung a dar, so ergibt sich:

$$\bar{r}_L(a) = \begin{cases} r_1 + m_1(a - a_1) & \text{für } a \leq a_t \\ r_2 + m_2(a_2 - a) & \text{für } a > a_t \end{cases} \quad (3.26)$$

Die Größen m_1 und m_2 in Gl. (3.26) sind die Steigungen der Grenzkurve nach Gl. (3.24) in den beiden Endpunkten. Der Punkt (a_t, r_t) ist der Tangenschnittpunkt. Für den Tangentenschnittpunkt muß gelten:

$$r_1 + m_1(a_t - a_1) \stackrel{!}{=} r_2 + m_2(a_2 - a_t) \quad (3.27)$$

Als Lösung von Gl. (3.27) ergibt sich der Tangentenschnittpunkt (a_t, r_t) zu:

$$a_t = \frac{r_2 - r_1 + a_1 m_1 - a_2 m_2}{m_1 - m_2} \quad (3.28)$$

$$r_t = r_1 + m_1(a_t - a_1) \quad (3.29)$$

Die stückweise lineare Approximation der Grenzkurve zwischen zwei Punkten ist im gesamten Bereich stetig, hat aber für $m_1 \neq m_2$ im Tangentenschnittpunkt eine Unstetigkeit in der ersten Ableitung.

Bestimmung des maximalen Fehlers für die Approximation mittels Tangenten

Die Grenzkurve ist konvex und verläuft daher innerhalb des Dreiecks, das die Tangenten und die Sekante bilden (Abbildung 3.3). Die Höhe h des Dreiecks ist damit ein Maß für den maximalen Approximationsfehler.

Die Höhe h des Dreiecks schneidet ausgehend vom Tangenschnittpunkt (a_t, r_t) , Gl. (3.28) bzw. (3.29) die Sekante an dem Punkt (a_s, r_s) . Für diesen Punkt muß gelten:

$$r_t + (a_s - a_t) m_h \stackrel{!}{=} r_1 + (a_s - a_1) m_s \quad (3.30)$$

Die Größen m_s und m_h bezeichnen hierbei die Steigung der Sekante bzw. der Höhe, die senkrecht auf der Sekante steht:

$$m_s = \frac{r_1 - r_2}{a_1 - a_2} \quad (3.31)$$

$$m_h = -\frac{1}{m_s} = \frac{a_2 - a_1}{r_1 - r_2} \quad (3.32)$$

Für die Steigung der Sekante m_s gilt aufgrund der Konvexität der Grenzkurve immer:

$$m_1 \leq m_s \leq m_2 \leq 0 \quad (3.33)$$

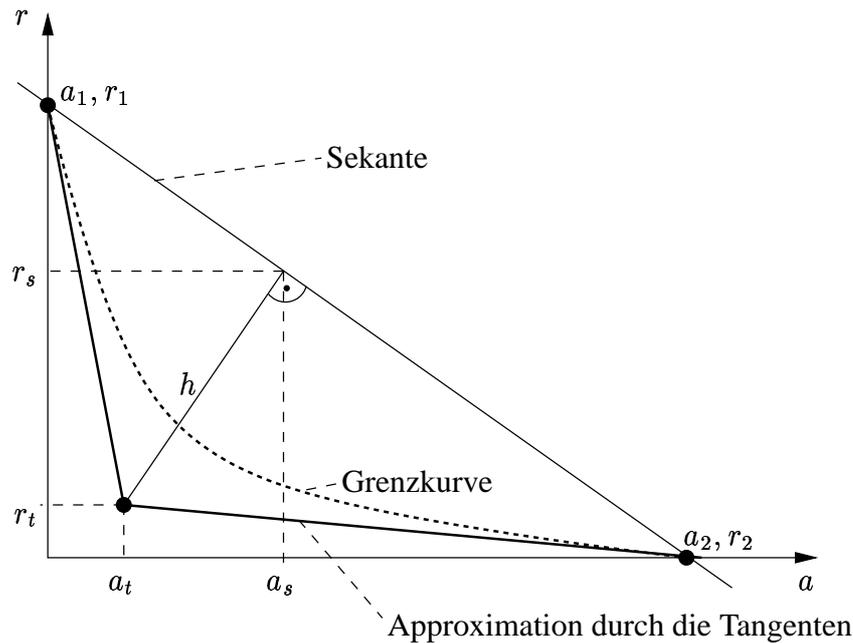


Abbildung 3.3: Bestimmung des Approximationsfehlers der Grenzkurve als Höhe h des Dreiecks zwischen Tangenten und Sekante.

Aus Gl. (3.30), (3.31) und (3.32) können die Werte für a_s bzw. r_s und mit einigen Umformungen die Länge h der Höhe, die den maximalen Approximationsfehler beschreibt, bestimmt werden:

$$\begin{aligned}
 a_s - a_t &= \frac{(m_1 - m_s)(m_2 - m_s)(r_2 - r_1)}{(m_1 - m_2)(1 + m_s^2)} \\
 r_t - r_s &= m_h(a_s - a_t) \\
 h &= \sqrt{(a_s - a_t)^2(1 + m_h^2)} = \frac{m_1 - m_s}{\sqrt{1 + m_s^2}} \cdot \frac{r_1 - r_2 + m_2(a_2 - a_1)}{m_1 - m_2} \quad (3.34)
 \end{aligned}$$

Bei Gl. (3.34) lassen sich drei besondere Fälle betrachten:

- Die Steigungen m_1 und m_2 sind fast gleich ($m_1 \rightarrow m_2$). Unter Berücksichtigung der Definition der Sekantensteigung Gl. (3.31) läßt sich Gl. (3.34) umformen:

$$h = \frac{m_1 - m_s}{m_1 - m_2} \cdot \frac{(m_s - m_2)(a_2 - a_1)}{\sqrt{1 + m_s^2}} \quad (3.35)$$

Da nach Gl. (3.33) der Wert der Sekantensteigung m_s immer zwischen den Tangentensteigungen liegt, gilt:

$$m_1 \rightarrow m_2 \Rightarrow m_s \rightarrow m_2$$

Setzt man dies in Gl. (3.35) ein, so erkennt man, daß die Länge der Höhe linear mit der Differenz zwischen den Steigungen zu Null geht. Bei einer Verringerung des Unterschiedes zwischen den Steigungen verkleinert sich also auch der Approximationsfehler.

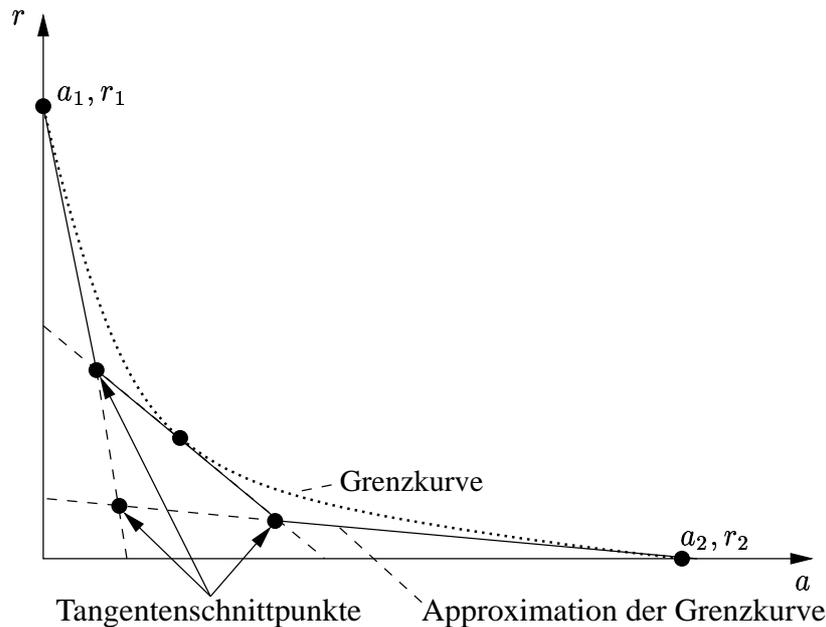


Abbildung 3.4: Stückweise lineare Approximation der Grenzkurve.

- Die transformierten Fehler r_1 und r_2 sind gleich. Formt man Gl. (3.34) für die Höhe geeignet um, so erhält man:

$$h = \frac{m_1 - m_s}{m_1 - m_2} \cdot \frac{(1 - m_2/m_s)(r_1 - r_2)}{m_1 - m_2} \quad (3.37)$$

Aus Gl. (3.37) erkennt man, daß im Grenzfall $r_1 \rightarrow r_2$ die Steigung linear mit der Differenz $r_1 - r_2$ zu Null geht.

- Die transformierten Parameterkorrekturen a_1 und a_2 sind gleich. Setzt man den Grenzwert $a_1 \rightarrow a_2$ in Gl. (3.35) ein, so geht auch für diesen Fall der Approximationsfehler linear mit der Differenz $a_1 - a_2$ zu Null.

Abschnittsweise lineare Approximation der gesamten Grenzkurve

Ausgehend von den Punkten für $\lambda \rightarrow \infty$ ($a = 0, r = 1$) und $\lambda = 0$ ($a = 1, r = 0$) kann die Grenzkurve stückweise linear durch ihre Tangenten angenähert werden. Als eine erste Approximation werden die Tangenten in den beiden Extrempunkten $\lambda \rightarrow \infty$ und $\lambda = 0$ (Abbildung 3.4) verwendet. Der maximale Approximationsfehler in diesem Schritt kann mittels Gl. (3.34) bestimmt werden.

Überschreitet der Approximationsfehler eine vorgegebene Schranke, so wird die Approximation durch die Berechnung eines weiteren Punktes auf der Grenzkurve iterativ verbessert. Wie in oben und in [Pör89] gezeigt, sollte dieser Punkt in der Nähe des Tangentschnittpunktes (a_t, r_t) liegen, um den Approximationsfehler nach der Berechnung des Punktes so klein wie möglich zu halten. Die Aufgabenstellung lautet also einen Wert für λ zu bestimmen, so daß gilt:

$$a(\lambda) \stackrel{!}{=} a_t \quad (3.38)$$

In [Pör89] wird dazu vorgeschlagen, die Steigung m_t im Tangentschnittpunkt als Sekantensteigung m_s der zwei benachbarten, bereits berechneten Punkte anzunehmen. Der Tangentschnittpunkt und die dort angenommene Steigung $m_t = m_s$ werden in Gl. (3.25) eingesetzt und nach λ aufgelöst:

$$\lambda = -m_t \cdot \frac{r_t \cdot (1 - \check{r}(0)) + \check{r}(0)}{a_t \cdot (\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|) + \|\mathbf{x}_c(\lambda \rightarrow \infty)\|} \cdot \frac{(1 - \check{r}(0)) \cdot (\|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2 - \|\bar{\mathbf{e}}_0\|^2)}{\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|} \quad (3.39)$$

Im neu berechneten Punkt, der in der Regel nicht genau im Tangentschnittpunkt liegt wird die Grenzkurve wiederum durch die Tangente angenähert. Wie in Abbildung 3.4 zu erkennen, teilt der neu berechnete Punkt die Grenzkurve in zwei Bereiche für die jeweils ein Approximationsfehler bestimmt werden kann. Überschreitet der Approximationsfehler in einem dieser Bereiche eine vorgegebene Schranke, so kann analog zu vorher die Approximation in diesem Bereich durch einen weiteren berechneten Punkt auf der Grenzkurve iterativ verbessert werden bis ein vorgegebener maximaler Fehler für alle Abschnitte der approximierten Grenzkurve unterschritten wird. Die abschnittsweise lineare Approximation der Grenzkurve durch Tangenten ist insbesondere für die Extremfälle der charakteristischen Grenzkurve gut:

System mit sehr guter Kondition: Für ein System mit idealer Konditionszahl $\text{cond}(\mathbf{G}) = 1$ ist die Grenzkurve eine Gerade. In diesem Fall gilt, daß die Steigung konstant ist ($m(\lambda \rightarrow \infty) = m(\lambda = 0) = 1$). Wie oben gezeigt verschwindet in diesem Fall der Approximationsfehler h nach Gl. (3.34).

System mit sehr schlechter Kondition: Für ein System mit sehr schlechter Konditionszahl $\text{cond}(\mathbf{G}) \rightarrow \infty$ hat die Grenzkurve einen ausgeprägten Knick. Dabei gilt näherungsweise, daß vor dem Knick die Kurve sehr stark abfällt ($m(\lambda \rightarrow \infty) \rightarrow -\infty$) und nach dem Knick nahezu waagrecht verläuft ($m(\lambda = 0) \rightarrow -0$).

Für den Teil der Grenzkurve, der links des Knickpunktes (a_t, r_t) liegt, ist die Differenz in der Parameterkorrektur $a_t - a_1$ klein; für den Teil rechts des Knickpunktes ist die Differenz des normierten Schaltungsfehlers $r_t - r_2$ klein. Nach der Berechnung eines Punktes auf der Grenzkurve in der Nähe des Knickpunktes ist deshalb auch in diesem Fall der Approximationsfehler h nach Gl. (3.34) gering.

Die Approximation der Grenzkurve mittels Tangenten ist überall stetig. Die erste Ableitung hat aber an den Tangentschnittpunkten der einzelnen Bereiche eine Unstetigkeit – die approximierte Kurve also einen Knick (Abbildung 3.4). Wie in [Pör89] gezeigt, kann die Grenzkurve mit Nebenbedingungen Knicke haben, wenn an einem Punkt die Anzahl der aktiven Nebenbedingungen gleich der Anzahl der Entwurfparameter ist – die Kurve verläuft in diesem Fall nach der Orthogonalisierung zu den aktiven Nebenbedingungen in einem Null-dimensionalen Unterraum. Der Fall, daß die Anzahl der Entwurfparameter gleich der Anzahl der aktiven Nebenbedingungen ist, tritt in der Praxis aber nur für große Auslenkungen auf, während in dem interessierenden Bereich kleiner und mittlerer Auslenkungen die Grenzkurve in der Regel glatt verläuft.

Im folgenden Abschnitt soll eine verbesserte abschnittsweise Approximation der Grenzkurve mittels *Bernstein-Bézier Kurven* vorgestellt werden, die stetig in der ersten Ableitung ist, mit geringem Aufwand berechnet und zudem auch zur Abschätzung höherer Ableitungen verwendet werden kann.

3.2.7 Approximation der Grenzkurve durch Bernstein-Bézier Kurven

Wie in Abbildung 3.3 gezeigt, verläuft die Grenzkurve zwischen zwei Punkten aufgrund ihrer Konvexität innerhalb des Dreiecks, das durch die Tangenten und die Sekante aufgespannt wird. Zudem ist an den zwei Endpunkten die Steigung durch die Tangentensteigung vorgegeben. Eine gute Approximation der Grenzkurve sollte deshalb

- in den beiden Randpunkten in Wert und Steigung übereinstimmen,
- innerhalb des Dreiecks verlaufen, das Tangenten und Sekante bilden,
- im Approximationsbereich konvex sein und
- mit möglichst geringem Aufwand berechenbar sein. Insbesondere muß eine Verbesserung der Approximation durch neu berechnete Punkte ohne großen Aufwand möglich sein, um eine iterative Vorgehensweise bei der Bestimmung der Kurve zu unterstützen.

Eine in der Mathematik häufig verwendete Approximation von Funktionen und Meßwerten durch Splines oder Polynome [Sto72] ist für die Grenzkurve nicht sinnvoll, da eine iterative Verbesserung der Approximation nur schwer möglich ist und die Konvexität nicht garantiert werden kann. Vielmehr neigen diese Approximationen leicht zu Oszillationen.

Im Maschinenbau und in der Computergrafik werden für die Approximation von Kurven und Flächen häufig Bernstein-Bézier Kurven (BB-Kurven) [MV90, BSMM95] verwendet, da sie eine in ihren Eigenschaften gut kontrollierbare Approximation von Kurven ermöglichen. Eine Verbesserung der Approximation durch zusätzliche Punkte erfordert dabei nur lokale Änderungen der Koeffizienten und keine komplette Neuberechnung. Bei geeigneter Wahl der Koeffizienten sind die entstehenden Kurven glatt und konvex. Eine detailliertere Definition und Diskussion von BB-Kurven findet sich in Anhang B.

Approximation der Grenzkurve zwischen zwei Punkten durch eine BB-Kurve

Die BB-Kurve geringsten Grades, die zur Approximation der Grenzkurve zwischen zwei Punkten (a_1, r_1) und (a_2, r_2) verwendet werden kann und alle obigen Forderungen an eine sinnvolle Approximation erfüllt, ist eine BB-Kurve mit drei Kontrollpunkten, den beiden Endpunkten (a_1, r_1) , (a_2, r_2) und dem Tangentenschnittpunkt (a_t, r_t) (Gl. (3.28) bzw. (3.29)). Die BB-Kurve für diesen Fall ist eine Konvexkombination (siehe Definition A.2) der drei Kontrollpunkten beschrieben durch eine parametrische Kurve in Abhängigkeit des Parameters $t \in [0, \dots, 1]$ [MV90, BSMM95]:

$$\begin{bmatrix} a_B(t) \\ r_B(t) \end{bmatrix} = (1-t)^2 \begin{bmatrix} a_1 \\ r_1 \end{bmatrix} + 2t(1-t) \begin{bmatrix} a_t \\ r_t \end{bmatrix} + t^2 \begin{bmatrix} a_2 \\ r_2 \end{bmatrix}, \quad 0 \leq t \leq 1 \quad (3.40)$$

Mit diesen Kontrollpunkten hat die BB-Kurve als Approximation der Grenzkurve folgende Eigenschaften (Beweise siehe Anhang B):

- Am Startpunkt $t = 0$ gilt:

$$a_B(t=0) = a_1 \quad \text{und} \quad r_B(t=0) = r_1$$

$$\left. \frac{\partial r_B(t)}{\partial a_B(t)} \right|_{t=0} = m_B(t=0) = m_1$$

- Am Endpunkt $t = 1$ gilt:

$$a_B(t = 1) = a_2 \quad \text{und} \quad r_B(t = 1) = r_2$$

$$\left. \frac{\partial r_B(t)}{\partial a_B(t)} \right|_{t=1} = m_B(t = 1) = m_2$$

- Für $t \in [0, \dots, 1]$ ist die BB-Kurve konvex und verläuft innerhalb des Dreiecks, das durch die Tangenten und die Sekante aufgespannt wird.

Mit diesen Eigenschaften stellt die BB-Kurve eine gute Approximation der Grenzkurve zwischen zwei berechneten Punkten sicher. Die Kurve ist glatt und erhält alle charakteristischen Eigenschaften der Grenzkurve.

Für die Approximation können auch BB-Kurven höheren Grades verwendet werden, bei denen der Tangenschnittpunkt als mehrfacher Kontrollpunkt verwendet wird. Alle in diesem Abschnitt getroffenen Aussagen bleiben auch für BB-Kurven höheren Grades gültig, jedoch ist in diesem Fall die BB-Kurve nicht mehr explizit nach einer Größe a oder r auflösbar und es ergibt sich eine deutlich unübersichtlichere Darstellung. Deshalb soll an dieser Stelle darauf verzichtet werden. In der Praxis hat sich eine BB-Kurve mit vier Kontrollpunkten bewährt, bei welcher der Tangenschnittpunkt als doppelter Kontrollpunkt verwendet wird. Da in diesem Fall die BB-Kurve nicht mehr explizit nach einem Parameter a oder r aufgelöst werden kann, ist eine eindimensionale Liniensuche auf einer monotonen Funktion nötig. Der Aufwand zur Bestimmung einer Nullstelle ist aufgrund der Eigenschaften der BB-Kurve aber sehr gering und fällt bei der Berechnung nicht ins Gewicht.

Bestimmung des Approximationsfehlers für die Approximation mittels BB-Kurve

Um einen Vergleich zwischen der Interpolation der Grenzkurve durch ihre Tangenten und durch eine BB-Kurve zu haben, soll in diesem Abschnitt der Approximationsfehler analog zum vorhergehenden Abschnitt bestimmt werden. Die BB-Kurve trennt das Dreieck, das durch Tangenten und Sekante gebildet wird in zwei Teile (Abbildung 3.5). Damit schneidet die BB-Kurve auch die Höhe h des Dreiecks in dem Punkt (a_b, r_b) . Das Maximum der beiden Anteile h_1 zwischen der Kurve und der Sekante und h_2 zwischen dem Tangenschnittpunkt (a_t, r_t) und der Kurve ist ein Maß für den möglichen maximal möglichen Approximationsfehler. Die einzelnen Anteile berechnen sich dabei unter Berücksichtigung der Steigung der Höhe Gl. (3.32) zu:

$$h_1 = \sqrt{(a_b - a_t)^2 (1 + m_h^2)} \quad (3.41)$$

$$h_2 = \sqrt{(a_s - a_t)^2 (1 + m_h^2)} \quad (3.42)$$

Der maximale Approximationsfehler bestimmt sich als $\max\{h_1, h_2\}$. Berücksichtigt man, daß $h_1 + h_2 = h$ gilt, wobei h die Höhe des Approximationsdreiecks Gl. (3.37) ist, so kann der Fehler abgeschätzt werden zu:

$$h \geq \max\{h_1, h_2\} \geq h/2 \quad (3.43)$$

Der maximale Approximationsfehler für die Interpolation der Grenzkurve mit BB-Kurven liegt in der gleichen Größenordnung verglichen mit der stückweise linearen Approximation. Die Höhe h des

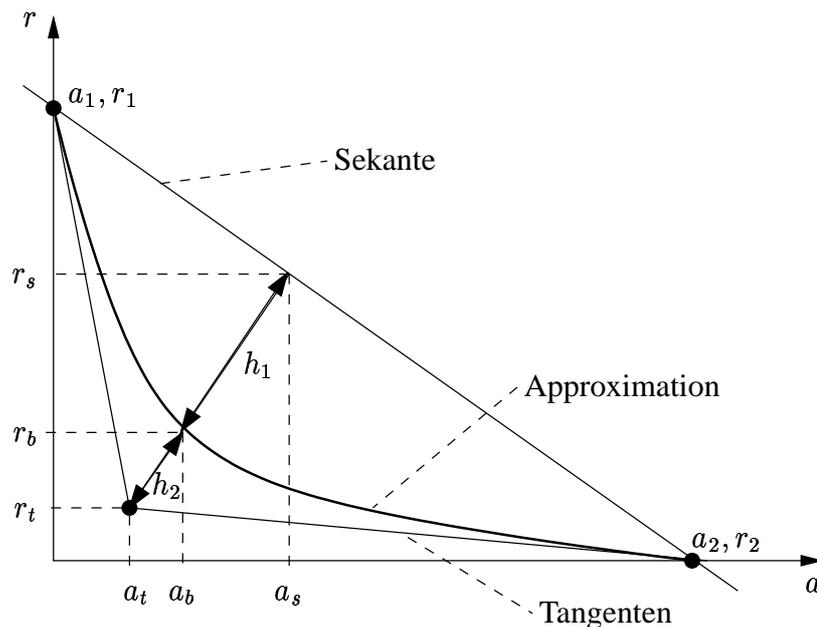


Abbildung 3.5: Bestimmung des Approximationsfehlers für die Interpolation der Grenzkurve durch eine BB-Kurve.

Dreiecks ist auch für diese Approximation ein gutes Maß für den maximal möglichen Approximationsfehler. In der Praxis ergibt sich aber zumeist eine deutlich bessere Approximation, was sich in einer geringeren Iterationszahl bei der Berechnung der Grenzkurve niederschlägt.

Abschnittsweise Approximation der Grenzkurve durch Bernstein-Bézier Kurven

Die oben vorgestellte Approximation der Grenzkurve zwischen zwei bereits berechneten Punkten mittels einer BB-Kurve ist vom Verfahren her vergleichbar mit der stückweise linearen Approximation durch die Tangenten. Die Approximation der gesamten Grenzkurve durch abschnittsweise definierte BB-Kurven erfolgt daher ebenfalls analog zu dem in Abschnitt 3.2.6 vorgestellten Verfahren zur stückweise linearen Approximation der gesamten Grenzkurve.

Die Grenzkurve wird zwischen den beiden Endpunkten $a_1 = a(\lambda \rightarrow \infty)$, $r_1 = r(\lambda \rightarrow \infty)$ und $a_2 = a(\lambda = 0)$, $r_2 = r(\lambda = 0)$ durch eine BB-Kurve nach Gl. (3.40) approximiert. Ist der maximale Approximationsfehler in dem interessierenden Bereich überschritten, kann die Approximation iterativ durch die Berechnung eines neuen Punktes im Bereich zwischen a_1 und a_2 verbessert werden. Um den maximalen Approximationsfehler nach der Berechnung möglichst klein zu halten, sollte der neu berechnete Punkt in der Nähe des Schnittpunktes (a_b, r_b) der Höhe des Dreiecks mit der BB-Kurve liegen (Abbildung 3.5). Es ist also ein t_* zu bestimmen, so daß $a_b \stackrel{!}{=} a_B(t)$ und $r_b \stackrel{!}{=} r_B(t)$ gilt. Dies läßt sich formulieren als:

$$r_B(t_*) \stackrel{!}{=} r_t + m_h(a_B(t_*) - a_t) \quad (3.44)$$

Diese Gleichung läßt sich direkt auflösen:

$$t_* = \frac{(a_1 - a_t)\Delta a + r_1\Delta r(r_1 - r_t) +}{a_1^2 - a_2^2 - 2a_t\Delta a + \Delta r(r_1 + r_2 - 2r_t)} + \frac{\sqrt{((a_t - a_1)\Delta a - \Delta r(r_1 - r_t))((a_2 - a_t)\Delta a + \Delta r(r_2 - r_t))}}{a_1^2 - a_2^2 - 2a_t\Delta a + \Delta r(r_1 + r_2 - 2r_t)}, \quad (3.45)$$

$$\Delta a = a_1 - a_2, \quad \text{und}$$

$$\Delta r = r_1 - r_2$$

Für $t = t_*$ gilt bzw. sei definiert:

$$\text{Auslenkung: } a_b = a_B(t_*), \quad (3.46)$$

$$\text{Schaltungsfehler: } r_b = r_B(t_*), \quad \text{und} \quad (3.47)$$

$$\text{Steigung: } m_b = m_B(t_*) \quad (3.48)$$

Diese Werte können wiederum in die Gleichung für die Steigung (3.25) eingesetzt und nach λ aufgelöst werden:

$$\lambda = -m_b \cdot \frac{r_b \cdot (1 - \check{r}(0)) + \check{r}(0)}{a_b \cdot (\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|) + \|\mathbf{x}_c(\lambda \rightarrow \infty)\|} \cdot \frac{(1 - \check{r}(0)) \cdot (\|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2 - \|\bar{\mathbf{e}}_0\|^2)}{\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|}. \quad (3.49)$$

Mit diesem Wert für λ kann das Optimierproblem (3.7) gelöst und so ein neuer Punkt zur iterativen Verbesserung der Grenzkurve bestimmt werden. Die Integration eines neu berechneten Punktes in die Approximation der gesamten Grenzkurve durch abschnittsweise definierte BB-Kurven erfordert nur, daß die BB-Kurve die bisher die Approximation in der Umgebung des Punktes (a_b, r_b) beschrieben hat in zwei Kurven mit neu zu berechnenden Koeffizienten aufgespalten wird. Alle anderen abschnittsweise definierten Kurven bleiben von der Änderung unberührt. Damit ist eine Verbesserung der Approximation mit einem Aufwand möglich, der unabhängig von der Anzahl der zu interpolierenden Punkte ist. Dies ist beispielsweise ein großer Vorteil gegenüber einer Approximation durch Splines oder Polynome, die bei einer Verbesserung der Approximation durch zusätzliche Punkte einen nicht unerheblichen Rechenaufwand erfordern.

3.2.8 Bestimmung einzelner Punkte auf der Grenzkurve

In diesem Kapitel wird ein iteratives Verfahren zur Bestimmung von Punkten auf der Grenzkurve mit vorgegebener Auslenkung a_* vorgestellt. Die Methode kann mit kleineren Anpassungen, die in den nachfolgenden Abschnitten diskutiert werden auch zur Bestimmung von Punkten mit vorgegebenem Fehler r_* oder Steigung m_* verwendet werden.

Bestimmung von Punkten auf der Grenzkurve mit vorgegebener Auslenkung

Bei dieser Aufgabenstellung ist ein λ_* zu bestimmen, so daß gilt:

$$\lambda_* : a(\lambda_*) \stackrel{!}{=} a_* \quad (3.50)$$

Diese Forderung läßt sich aber im allgemeinen nicht genau erfüllen, da die Grenzkurve nicht als analytische Funktion gegeben ist. Deshalb wird eine Toleranz Δ_{\max} (z.B. $\Delta_{\max} = 0.01$) zugelassen:

$$\lambda_* : |a(\lambda_*) - a_*| \leq \Delta_{\max} \quad (3.51)$$

Im folgenden wird angenommen, daß bereits zwei Punkte (a_1, r_1) und (a_2, r_2) auf der Grenzkurve berechnet wurden, für die gilt:

$$a_1 \leq a_* \leq a_2, \quad \text{und} \quad r_1 \geq r_2$$

Die Punkte können dabei einerseits aus der Approximation der Grenzkurve stammen oder die Extremwerte für $\lambda \rightarrow \infty$ und $\lambda = 0$ sein. Diese beiden Punkte werden als Startpunkte $(a_1^{(0)}, r_1^{(0)}, \lambda_1^{(0)})$ und $(a_2^{(0)}, r_2^{(0)}, \lambda_2^{(0)})$ im nullten Iterationsschritt ($\mu = 0$) des iterativen Verfahrens gewählt. In jedem Iterationsschritt (μ) wird die Grenzkurve durch eine BB-Kurve approximiert:

$$\begin{bmatrix} a_B^{(\mu)}(t) \\ r_B^{(\mu)}(t) \end{bmatrix} = (1-t)^2 \begin{bmatrix} a_1^{(\mu)} \\ r_1^{(\mu)} \end{bmatrix} + 2t(1-t) \begin{bmatrix} a_t^{(\mu)} \\ r_t^{(\mu)} \end{bmatrix} + t^2 \begin{bmatrix} a_2^{(\mu)} \\ r_2^{(\mu)} \end{bmatrix} \quad (3.53)$$

Der Tangentschnittpunkt $(a_t^{(\mu)}, r_t^{(\mu)})$ in jedem Iterationsschritt bestimmt sich nach Gl. (3.28) bzw. Gl. (3.29). Nach der Approximation durch eine BB-Kurve liegt die Grenzkurve als analytische parametrische Funktion vor. Aus dieser analytischen Funktion können der Wert des Schaltungsfehlers r und der Steigung m der Grenzkurve an der Stelle $a_B^{(\mu)}(t) = a_*$ bestimmt werden. Dazu wird zunächst der Parameter $t_*^{(\mu)}$ bestimmt, so daß $a_B^{(\mu)}(t_*^{(\mu)}) \stackrel{!}{=} a_*$ gilt:

$$t_*^{(\mu)} = \frac{a_1^{(\mu)} - a_t^{(\mu)} + \sqrt{a_*(a_1^{(\mu)} - 2a_t^{(\mu)} + a_2^{(\mu)}) + (a_t^{(\mu)})^2 - a_1^{(\mu)}a_2^{(\mu)}}}{a_1^{(\mu)} - 2a_t^{(\mu)} + a_2^{(\mu)}} \quad (3.54)$$

Mit Hilfe von Gl. (3.53) bzw. Gl. (B.5) kann der Wert $r_*^{(\mu)}$ und die Steigung $m_*^{(\mu)}$ in diesem Iterationsschritt bestimmt werden:

$$r_*^{(\mu)} = r_B^{(\mu)}(t_*^{(\mu)}) \quad (3.55)$$

$$m_*^{(\mu)} = m_B^{(\mu)}(t_*^{(\mu)}) \quad (3.56)$$

Analog zu Gl. (3.49) kann mit den Werten $r_*^{(\mu)}$, $m_*^{(\mu)}$ und a_* ein Schätzwert $\lambda_*^{(\mu)}$ aus der Steigungsgleichung (3.25) bestimmt werden:

$$\lambda_*^{(\mu)} = -m_*^{(\mu)} \cdot \frac{r_*^{(\mu)} \cdot (1 - \check{r}(0)) + \check{r}(0)}{a_* \cdot (\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|) + \|\mathbf{x}_c(\lambda \rightarrow \infty)\|} \cdot \frac{(1 - \check{r}(0)) \cdot (\|\bar{\mathbf{e}}_c(\lambda \rightarrow \infty)\|^2 - \|\bar{\mathbf{e}}_0\|^2)}{\|\mathbf{x}_c(0)\| - \|\mathbf{x}_c(\lambda \rightarrow \infty)\|} \quad (3.57)$$

Für $\lambda = \lambda_*^{(\mu)}$ wiederum kann die „wahre“ Auslenkung $a(\lambda)$ und der Fehler $r(\lambda)$ durch Lösung des Optimierproblems Gl. (3.10) bestimmt werden. An dieser Stelle können vier Fälle unterschieden werden:

- $|a(\lambda_*^{(\mu)}) - a_*| \leq \Delta_{\max}$:
Der berechnete Punkt ist im Rahmen der geforderten Genauigkeit der gesuchte Punkt.
- $a_1^{(\mu)} < a(\lambda_*^{(\mu)}) < a_*$:
Der berechnete Punkt hat eine zu kleine Auslenkung. Die Approximation kann iterativ verbessert werden, indem der linke Kontrollpunkt $a_1^{(\mu+1)} = a(\lambda_*^{(\mu)})$, $r_1^{(\mu+1)} = r(\lambda_*^{(\mu)})$ und $\lambda_1^{(\mu+1)} = \lambda_*^{(\mu)}$ gesetzt wird.
- $a_* < a(\lambda_*^{(\mu)}) < a_2^{(\mu)}$:
Der berechnete Punkt hat eine zu große Auslenkung. Die Approximation kann iterativ verbessert werden, indem der rechte Kontrollpunkt $a_2^{(\mu+1)} = a(\lambda_*^{(\mu)})$, $r_2^{(\mu+1)} = r(\lambda_*^{(\mu)})$ und $\lambda_2^{(\mu+1)} = \lambda_*^{(\mu)}$ gesetzt wird.
- $a(\lambda_*^{(\mu)}) = a_1^{(\mu)}$ oder $a(\lambda_*^{(\mu)}) = a_2^{(\mu)}$:
Dieser Sonderfall kann nur auftreten, wenn in $a_1^{(\mu)}$ oder $a_2^{(\mu)}$ die Grenzkurve eine Unstetigkeit in der ersten Ableitung hat, also wenn die Anzahl der aktiven Nebenbedingungen gleich der Anzahl der Entwurfparameter ist. Tritt dieser Sonderfall auf, so wird ein neuer Schätzwert für λ_* folgendermaßen gewählt [Pör89]:

$$\lambda_*^{(\mu)} = \exp \left(\frac{\ln(\lambda_1^{(\mu)}) + \ln(\lambda_2^{(\mu)})}{2} \right) \quad (3.58)$$

Für diesen Schätzwert wird die Auslenkung $a(\lambda_*^{(\mu)})$ und der Fehler $r(\lambda_*^{(\mu)})$ durch Lösung des Optimierproblems (3.10) bestimmt und die Randpunkte $(a_1^{(\mu+1)}, r_1^{(\mu+1)})$ und $(a_2^{(\mu+1)}, r_2^{(\mu+1)})$ des nächsten Iterationsschrittes entsprechend obiger Regeln gewählt.

Die Beachtung dieses Sonderfalles ist nicht zwingend notwendig für eine Konvergenz des Verfahrens. Für $a_1^{(\mu)}, a_2^{(\mu)} \neq a_*$ gilt in jedem Fall:

$$\begin{aligned} \text{Austausch des linken Punktes:} \quad & \lambda_1^{(\mu)} > \lambda_*^{(\mu)} \Rightarrow m_1^{(\mu+1)} < m_1^{(\mu)}, \quad \text{oder} \\ \text{Austausch des rechten Punktes:} \quad & \lambda_2^{(\mu)} < \lambda_*^{(\mu)} \Rightarrow m_2^{(\mu+1)} > m_2^{(\mu)} \end{aligned}$$

Damit ist auch ohne Berücksichtigung dieses Sonderfalles globale Konvergenz des Verfahrens garantiert. In diesem Fall werden allerdings iterativ die Werte der Steigung durchlaufen, die „im Knickpunkt“ liegen, und die Auslenkung ändert sich nicht. Das bedeutet eine sehr langsame, in der Praxis nicht sinnvolle Konvergenzgeschwindigkeit. Durch die heuristische Ermittlung eines neuen Schätzwertes für $\lambda_*^{(\mu)}$ kann eine deutliche Beschleunigung erzielt werden.

Mit den neuen Werten $(a_1^{(\mu+1)}, r_1^{(\mu+1)}, \lambda_1^{(\mu+1)})$ und $(a_2^{(\mu+1)}, r_2^{(\mu+1)}, \lambda_2^{(\mu+1)})$ wird eine neue Iteration $\mu + 1$ des Verfahrens gestartet, bis die Abbruchbedingung (3.51) erfüllt ist.

Bestimmung von Punkten mit vorgegebenem Fehler oder Steigung

Das vorgestellte Verfahren kann ohne größere Änderungen auch zur Lösung anderer Problemstellungen verwendet werden. Zur Bestimmung von Punkten mit vorgegebenem Fehler r_* ist dazu in jedem Iterationsschritt $t_*^{(\mu)}$ analog zu Gl. (3.54) so zu berechnen, daß

$$\begin{aligned} t_*^{(\mu)} : \quad r_B^{(\mu)}(t_*^{(\mu)}) &\stackrel{!}{=} r_* \\ \Rightarrow t_*^{(\mu)} &= \frac{r_1^{(\mu)} - r_t^{(\mu)} + \sqrt{r_* (r_1^{(\mu)} - 2r_t^{(\mu)} + r_2^{(\mu)}) + (r_t^{(\mu)})^2 - r_1^{(\mu)} r_2^{(\mu)}}}{r_1^{(\mu)} - 2r_t^{(\mu)} + r_2^{(\mu)}} \end{aligned} \quad (3.59)$$

Der Rest des Verfahrens kann analog zu oben verwendet werden. Gleiches gilt für die Bestimmung von Punkten mit vorgegebener Auslenkung m_* . In diesem Fall ergibt sich der Wert des Parameters $t_*^{(\mu)}$ der BB-Kurve in jeder Iteration μ aus Gl. (B.5):

$$\begin{aligned} t_*^{(\mu)} : \quad m_B^{(\mu)}(t_*^{(\mu)}) &\stackrel{!}{=} r_* \\ \Rightarrow t_*^{(\mu)} &= \frac{(r_t^{(\mu)} - r_1^{(\mu)}) - m_* (a_t^{(\mu)} - a_1^{(\mu)})}{m_* (a_1^{(\mu)} - 2a_t^{(\mu)} + a_2^{(\mu)}) - (r_1^{(\mu)} - 2r_t^{(\mu)} + r_2^{(\mu)})} \end{aligned} \quad (3.60)$$

Bei der Bestimmung von Punkten mit gegebener Steigung m_* der Kurve ist zu beachten, daß die Beschleunigung des Verfahrens mit Gl. (3.58) nicht möglich ist. Ist die Anzahl der aktiven Nebenbedingungen gleich der Anzahl der Parameter, dann verläuft die Kurve in einem Null-dimensionalen Unterraum. Die Auslenkung a oder der Fehler r können sich in diesem Unterraum nicht ändern, weshalb die Beschleunigung mit Gl. (3.58) die Konvergenz des Verfahrens nicht beeinflußt. Die Steigung kann bzw. muß sich während des Verlaufs in dem Null-dimensionalen Unterraum ändern, was einen Knick im Verlauf der Grenzkurve verursacht. Dies hat zur Folge, daß für unterschiedliche vorgegebene Steigung derselbe Punkt auf der Grenzkurve berechnet wird.

3.3 Schrittweitenbestimmung

Im Gegensatz zu anderen Optimierproblemen besteht bei der Dimensionierung von Schaltungen der Hauptaufwand nicht in den Matrixzerlegungen, die im Laufe der Optimierung nötig sind. Vielmehr dominieren die Schaltungssimulationen zur Bestimmung der Werte der Schaltungseigenschaften und Empfindlichkeiten die Laufzeit des Algorithmus. Bei der Schrittweitenbestimmung ist es deshalb nötig, die Anzahl der Schaltungssimulationen so klein wie möglich zu halten. Eine zeitaufwendige, schlecht parallelisierbare Liniensuche muß deshalb vermieden werden.

Wie oben diskutiert, liegt eine Lösung des linearen Systems mit einem guten Verhältnis von Fehlerreduktion zur Größe der Parameterkorrektur zumeist im Knickpunkt der Grenzkurve. Für diesen Punkt kann eine gute Reduktion der Zielfunktion bei einem, aufgrund der geringen Auslenkung kleinen Linearisierungsfehler erwartet werden. Eine erste Abschätzung einer sinnvollen Schrittweite ist deshalb der Knickpunkt der Grenzkurve.

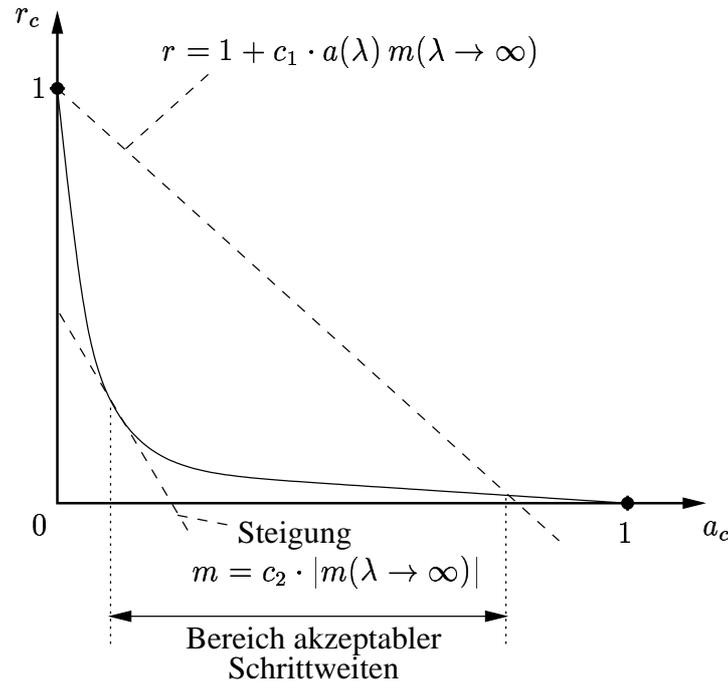


Abbildung 3.6: Veranschaulichung der Wolfe Bedingung (Bedingung des hinreichenden Abstiegs) an der Grenzkurve

Bestimmung des Knickpunktes der Grenzkurve

Für die charakteristische Grenzkurve ohne Nebenbedingungen ist der Knickpunkt der Grenzkurve der Punkt mit dem geringsten Krümmungsradius ρ . Mit Hilfe der BB-Kurve kann der Krümmungsradius in Abhängigkeit vom Parameter t abgeschätzt werden zu [MV90]:

$$\rho(t) = \frac{\sqrt{(\dot{a}_B(t)^2 + \dot{r}_B(t)^2)^3}}{|\dot{a}_B(t)\ddot{r}_B(t) - \ddot{a}_B(t)\dot{r}_B(t)|}, \quad \text{mit } \dot{a}_B = \frac{\partial a_B}{\partial t}, \quad \dot{r}_B = \frac{\partial r_B}{\partial t} \quad (3.61)$$

Wie in Abschnitt 3.2.4 gezeigt, bedeutet jede Änderung in der Menge der aktiven Nebenbedingungen eine Änderung der Systemmatrix der freien Variablen. An den Punkten an denen sich die Menge der aktiven Nebenbedingungen ändert, ist die Grenzkurve stetig, kann aber Knicke haben. Dies bedeutet eine Unstetigkeit im Krümmungsradius, die durch eine δ -Funktion beschrieben werden kann. Bei der Approximation der Grenzkurve durch abschnittsweise definierte BB-Kurven ergibt Gl. (3.61) mehrere lokale Minima. Diese müssen nicht unbedingt besonders günstige Lösungen darstellen. Um eine sinnvolle Lösung für den Knickpunkt zu bekommen, müssen deshalb weitere Forderungen gestellt werden.

Für die Konvergenz eines Optimieralgorithmus werden im allgemeinen Forderungen an die Schrittweite gestellt. Im verwendeten Algorithmus garantiert die *Wolfe-Bedingung (Bedingung des hinreichenden Abstiegs)* [NW99, Spe98] die Konvergenz. Angepaßt auf die Grenzkurve lautet diese Bedingung:

$$r(\lambda) \leq 1 + c_1 \cdot a(\lambda) m(\lambda \rightarrow \infty), \quad \text{und} \quad (3.62)$$

$$|m(\lambda)| \leq c_2 \cdot |m(\lambda \rightarrow \infty)| \quad (3.63)$$

Die positiven Konstanten c_1 und c_2 müssen dabei unabhängig von a , r und λ sein. Will man die Wolfe Bedingung für das lineare Modell der Schaltung einhalten, auf dem die Grenzkurve beruht, ergibt sich eine Einschränkung der Schrittweite wie in Abbildung 3.6 veranschaulicht. Innerhalb dieses Bereichs wird der Knickpunkt gesucht als Minimum des Krümmungsradius nach Gl. (3.61). Gibt es mehrere Minima, so wird das Minimum mit der geringsten Auslenkung a genommen.

Verifikation der Gültigkeit der Linearisierung

Die Schrittweite, die mit Hilfe der Grenzkurve bestimmt wird, stellt eine sinnvolle Auslenkung dar in dem Sinne, daß Fehlerreduktion und Größe der Parameterkorrektur in einem guten Verhältnis zueinander stehen. Damit wird insbesondere die Forderung erfüllt, daß die Lösung möglichst in der Nähe der Startlösung liegen sollte, um schaltungstechnisch nicht sinnvolle Lösungen zu vermeiden.

Eine möglichst kleine Auslenkung bedeutet in der Regel aber auch einen kleinen Linearisierungsfehler. Um eine Konvergenz des Algorithmus zu garantieren, ist es trotzdem auf jeden Fall nötig, in jedem Schritt des Algorithmus die Gültigkeit der Linearisierung zu kontrollieren. Die Wolfe Bedingung nach Gl. (3.62) und Gl. (3.63) liefert die hierzu notwendigen Kriterien.

3.4 Die verallgemeinerte Grenzkurve

Die charakteristische Grenzkurve ist beschränkt auf die Lösung eines Least-Square-Optimierproblems basierend auf einer Linearisierung. Wird die Schaltungsdimensionierung so formuliert wie in Kapitel 2 beschrieben, so ist die eigentliche Aufgabe aber kein Least-square Optimierproblem. Die Transformation der Aufgabenstellung in ein Least-square Optimierproblem entsprechend Gl. (3.8) erzeugt zwei Probleme, die aus der Anpassung der Aufgabenstellung an die charakteristische Grenzkurve resultieren:

- Es wird nicht nur die Abhängigkeit der Zielgrößen γ von den Parametern \mathbf{p} linearisiert, sondern auch die Abhängigkeit der Zielfunktion φ von den Zielgrößen γ . Häufig ist die Abhängigkeit der Zielgrößen von den Parametern nur schwach nichtlinear, die Abhängigkeit der Zielfunktion von Zielgrößen dagegen stark nichtlinear aufgrund der verwendeten Exponentialfunktion. Diese Transformationsfunktion ist bekannt und muß deshalb nicht linearisiert werden.
- Die Formulierung als Least-square Optimieraufgabe fügt eine künstliche Schrittweitenbeschränkung ein. Zur Veranschaulichung wird ein quadratisches System ($n_d = n_f$) mit vollem Rang ohne Nebenbedingungen angenommen. Für dieses System läßt sich die Lösung für $\lambda = 0$ unter Berücksichtigung des Aufbaus der Systemmatrix nach Gl. (3.9) direkt formulieren:

$$\epsilon_F - a \cdot \text{diag}(\epsilon_F) \cdot \mathbf{S} \cdot \mathbf{x}(\lambda = 0) \stackrel{!}{=} \mathbf{0} \quad \Rightarrow \quad \mathbf{x}(0) = 1/a \cdot \mathbf{S}^{-1} \mathbf{1} \quad (3.64)$$

Hierbei bezeichnet $\mathbf{1}$ einen Vektor der Größe n_f , in dem alle Elemente eins sind. Setzt man diese Lösung für \mathbf{x} Gl. (3.64) in die Linearisierung für die Zielgrößen Gl. (3.3) ein, so ergibt sich:

$$\bar{\gamma} = \gamma_0 + \mathbf{S} \cdot \mathbf{x}(0) = \gamma_0 + \mathbf{S}(1/a \cdot \mathbf{S}^{-1} \mathbf{1}) = \gamma_0 + 1/a \cdot \mathbf{1} \quad (3.65)$$

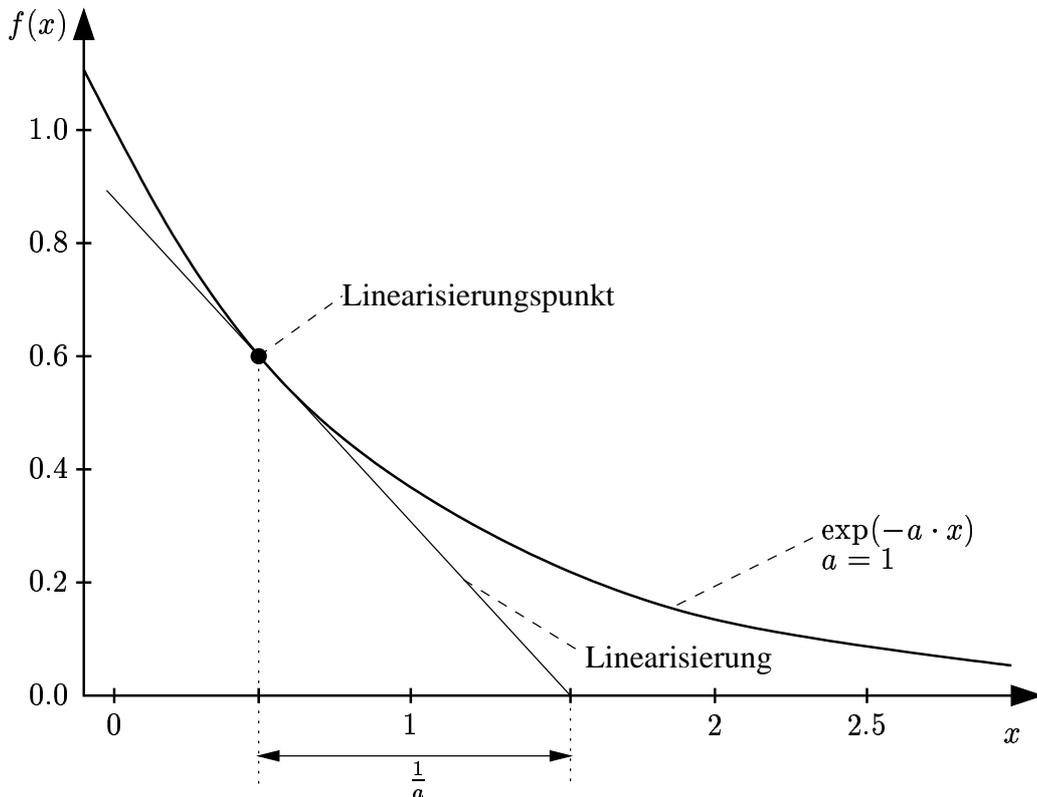


Abbildung 3.7: Künstliche Schrittweitenbeschränkung bei Linearisierung der Exponentialfunktion.

Im linearen Modell ändern sich für den Fall mit der größten Auslenkung und Fehlerreduktion ($\lambda = 0$) die Zielgrößen nur um $1/a$. Wählt man den Skalierungsfaktor a deutlich größer als eins, so wird selbst im linearen Schaltungsmodell nahezu keine Fehlerreduktion erreicht. Diese künstliche Schrittweitenbeschränkung kommt aus der Tatsache, daß für die zur Transformation verwendete Exponentialfunktion an jeder Stelle der Wert der Ableitung und der Funktion gleich sind. Abbildung 3.7 veranschaulicht diese Eigenschaft der Exponentialfunktion für den eindimensionalen Fall.

Oben beschriebene Probleme sind vermeidbar, indem die Grenzkurve an die Problemstellung angepaßt wird. Die Hauptidee ist es, nicht die gesamte Zielfunktion zu linearisieren. Vielmehr wird nur die Abhängigkeit der Zielgrößen γ von den Schaltungsparametern \mathbf{d} linearisiert. Die zur Transformation verwendete stark nichtlineare Exponentialfunktion wird zur Bestimmung einer Schrittweite nicht linearisiert.

Das vorgestellte Verfahren ist dabei nicht auf die Zielfunktion nach Gl. (3.1) beschränkt. Vielmehr ist es auf jede konvexe Zielfunktion, die auf einem konvexen Gebiet minimiert werden soll anwendbar. Im folgenden wird die verallgemeinerte Grenzkurve als ein Verfahren zur Schrittweitenbestimmung für konvexe Optimierprobleme vorgestellt. Anschließend soll gezeigt werden, daß das Optimierproblem Gl. (3.2) nach einer Linearisierung der Zielgrößen γ und der Nebenbedingungen \mathbf{u} ein konvexes Optimierproblem ist.

3.4.1 Definition der verallgemeinerten Grenzkurve

Linearisiert man für die Lösung des Optimierproblems (3.2) nicht die gesamte Zielfunktion nach Gl. (3.1), sondern nur die Zielgrößen γ und Nebenbedingungen \mathbf{u} entsprechend Gl. (3.3) bzw. (3.4), so läßt sich das Optimierproblem schreiben als:

$$\min_{\mathbf{x}} \{ \bar{\varphi}(\mathbf{x}) \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \} \quad (3.66)$$

$$\bar{\varphi}(\mathbf{x}) = \sum_{i=1}^{n_\gamma} \exp(-a \cdot \bar{\gamma}_i(\mathbf{x})) \quad (3.67)$$

Für dieses Optimierproblem kann wiederum ein Ansatz vergleichbar zur charakteristischen Grenzkurve gemacht werden, um einen guten Kompromiß zwischen Auslenkung und Fehlerreduktion zu finden:

$$\min_{\mathbf{x}} \{ \bar{\varphi}(\mathbf{x}) + \lambda \cdot \|\mathbf{x}\|^2 \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \}, \quad \lambda > 0 \quad (3.68)$$

Die Lösung von Gl. (3.68) und der Fehler an dieser Stelle in Abhängigkeit von λ werden analog zur charakteristischen Grenzkurve mit $\mathbf{x}_G(\lambda)$ bzw. $\bar{\varphi}_G(\lambda)$ bezeichnet:

$$\mathbf{x}_G(\lambda) = \underset{\mathbf{x}}{\operatorname{argmin}} \{ \bar{\varphi}(\mathbf{x}) + \lambda \cdot \|\mathbf{x}\|^2 \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \} \quad (3.69)$$

$$\bar{\varphi}_G(\lambda) = \bar{\varphi}(\mathbf{x}_G(\lambda)) \quad (3.70)$$

Der Index G in diesen Gleichungen kennzeichnet die Lösung der verallgemeinerten Grenzkurve (Generalized Boundary Curve – GBC). Wie unten gezeigt, erhält die verallgemeinerte Grenzkurve unter sinnvollen Voraussetzungen alle wichtigen Eigenschaften der charakteristischen Grenzkurve. Die verallgemeinerte Grenzkurve sieht deshalb vergleichbar zur charakteristischen Grenzkurve aus (z. B. Abbildung 3.1(a)). Die charakteristische Grenzkurve kann auch als ein Sonderfall verallgemeinerten Grenzkurve mit geeignet definierter Zielfunktion betrachtet werden.

3.4.2 Eigenschaften der verallgemeinerten Grenzkurve

Unter der Bedingung, daß die Zielfunktion $\bar{\varphi}(\mathbf{x})$ für die Grenzkurve konvex ist, behält die verallgemeinerte Grenzkurve alle wichtigen Eigenschaften der charakteristischen Grenzkurve.

Satz 3.1 Sei $\mathcal{C} = \{ \mathbf{x} \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \}$ das nichtleere Gebiet, auf dem alle linearisierten Nebenbedingungen erfüllt sind. Dann ist die Lösung $\mathbf{x}_G = \mathbf{x}_G(\lambda)$ der Punkt $\mathbf{x} \in \mathcal{C}$ mit minimalem Funktionswert $\bar{\varphi}(\mathbf{x})$ für alle \mathbf{x} mit $\|\mathbf{x}\| \leq \|\mathbf{x}_G\|$.

Beweisidee: Das Gebiet \mathcal{C} ist konvex, da die Grenzen des Gebietes rein linear sind. Aufgrund der Konvexität der Zielfunktion $\bar{\varphi}(x)$ und des zulässigen Gebietes \mathcal{C} ist das Optimierproblem der verallgemeinerten Grenzkurve Gl. (3.69) ein konvexes Optimierproblem. Damit sind die *Karush-Kuhn-Tucker* (KKT) Optimalitätsbedingungen (siehe Anhang A.1) hinreichend für ein globales Optimum.

Im folgenden wird gezeigt, daß der Parameter $\lambda > 0$ auch als Lagrangeparameter für die Nebenbedingung $\|\mathbf{x}\|^2 \leq \|\mathbf{x}_G(\lambda)\|^2$ interpretiert werden kann. Zusammen mit der globalen Optimalität und der

Tatsache, daß der Betrag eines Vektors immer positiv ist, ist dann sichergestellt, daß der Punkt $\mathbf{x}_G(\lambda)$ das Minimum für alle $\mathbf{x} \in \mathcal{C}$, $\|\mathbf{x}\| \leq \|\mathbf{x}_G(\lambda)\|$ ist.

Die Lagrangefunktion (siehe auch Anhang A.1) $L(\mathbf{x}, \lambda, \boldsymbol{\mu})$ für das Optimierproblem

$$\min_{\mathbf{x}} \{ \bar{\varphi}(\mathbf{x}) \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0} \wedge \|\mathbf{x}_G\|^2 - \|\mathbf{x}\|^2 \geq 0 \} \quad (3.71)$$

lautet

$$L(\mathbf{x}, \lambda, \boldsymbol{\mu}) = \bar{\varphi} - \lambda (\|\mathbf{x}_G\|^2 - \|\mathbf{x}\|^2) - \boldsymbol{\mu}^T (\mathbf{U}\mathbf{x} + \mathbf{u}_0). \quad (3.72)$$

Die Lagrange-Faktoren für die Nebenbedingungen wurden hierbei aufgespalten in den Faktor λ für die Bedingung $\|\mathbf{x}\|^2 \leq \|\mathbf{x}_G\|^2$ und die Faktoren $\boldsymbol{\mu}$ für die restlichen Nebenbedingungen $\bar{\mathbf{u}}(\mathbf{x})$. Das Optimierproblem (3.71) ist unter den gegebenen Voraussetzungen ebenfalls konvex, da der zulässige Bereich für die Lösung sich als Schnittmenge des Gebietes \mathcal{C} mit dem abgeschlossenen Inneren der konvexen Hyperkugel $\|\mathbf{x}\|^2 \leq \|\mathbf{x}_G\|^2$ definiert und die Schnittmenge zweier konvexer Menge ebenfalls konvex ist. Die KKT-Optimalitätsbedingungen sind also auch für das Optimierproblem (3.71) hinreichend für ein globales Optimum.

Die KKT-Bedingung (A.4) für Gl. (3.71) lautet:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda, \boldsymbol{\mu}) \stackrel{!}{=} \mathbf{0} \Rightarrow \nabla_{\mathbf{x}} \bar{\varphi}(\mathbf{x}) + 2\lambda \|\mathbf{x}\| - \mathbf{U}^T \boldsymbol{\mu} = \mathbf{0} \quad (3.73)$$

Zusammen mit der primalen Zulässigkeit $\|\mathbf{x}\|^2 \leq \|\mathbf{x}_G\|^2 \wedge \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0}$ ist diese Bedingung bei einem konvexen Optimierproblem hinreichend für ein globales Minimum. Stellt man die Lagrangefunktion für die verallgemeinerte Grenzkurve (3.66) auf, so lautet diese:

$$L_{\text{GBC}}(\mathbf{x}, \lambda, \boldsymbol{\mu}) = \bar{\varphi} + \lambda \|\mathbf{x}\|^2 - \boldsymbol{\mu}^T (\mathbf{U}\mathbf{x} + \mathbf{u}_0). \quad (3.74)$$

Die KKT Bedingung für $L_{\text{GBC}}(\mathbf{x}, \lambda, \boldsymbol{\mu})$ ergibt sich zu:

$$\nabla_{\mathbf{x}} L_{\text{GBC}}(\mathbf{x}, \lambda, \boldsymbol{\mu}) \stackrel{!}{=} \mathbf{0} \Rightarrow \nabla_{\mathbf{x}} \bar{\varphi}(\mathbf{x}) + 2\lambda \|\mathbf{x}\| - \mathbf{U}^T \boldsymbol{\mu} = \mathbf{0} \quad (3.75)$$

Die KKT-Bedingung für die ursprüngliche und die modifizierte Problemstellung (3.71) sind identisch und haben damit auch die gleiche Lösung. An dieser Stelle ist zu beachten, daß für alle Lagrange-Faktoren, aber insbesondere für λ die Komplementaritätsbedingung (A.7) gilt. Ein positiver Lagrange-Faktor λ bedeutet also, daß die Nebenbedingung $\|\mathbf{x}\|^2 \leq \|\mathbf{x}_G\|^2$ aktiv ist. \square

Satz 3.2 Wenn die Funktion $\bar{\varphi}(\mathbf{x})$ konvex ist auf dem durch die erfüllten linearisierten Nebenbedingungen definierten Gebiet $\mathcal{C} = \{\mathbf{x} \mid \mathbf{U}\mathbf{x} + \mathbf{u}_0 \geq \mathbf{0}\}$, dann ist die parametrische Funktion $[\|\mathbf{x}_G(\lambda)\|, \bar{\varphi}_G(\lambda)]$ mit $\lambda \geq 0$ konvex.

Beweisidee: Die Funktion $\bar{\varphi}(\mathbf{x}_G)$ ist dann konvex, wenn gilt:

$$\frac{\partial^2 \bar{\varphi}(\mathbf{x}_G(\lambda))}{\partial \|\mathbf{x}_G(\lambda)\|^2} \geq 0 \quad (3.76)$$

Die Funktion $\bar{\varphi}(\mathbf{x})$ ist im zulässigen Bereich \mathcal{C} nach Voraussetzung konvex. Nach Satz A.2 gilt daher für zwei beliebige Lösungen $\mathbf{x}_1 = \mathbf{x}_G(\lambda_1)$ und $\mathbf{x}_2 = \mathbf{x}_G(\lambda_2)$:

$$\bar{\varphi}(\mathbf{x}_2) \geq \bar{\varphi}(\mathbf{x}_1) + (\mathbf{x}_2 - \mathbf{x}_1)^T \nabla_{\mathbf{x}} \bar{\varphi}(\mathbf{x}_1) \quad (3.77)$$

Mit $\delta = \|\mathbf{x}_2\| - \|\mathbf{x}_1\|$ und $\delta \rightarrow 0$ kann $\bar{\varphi}(\mathbf{x}_2)$ auch mit einer Taylorreihe abgeschätzt werden:

$$\bar{\varphi}_G(\mathbf{x}_2) = \bar{\varphi}_G(\mathbf{x}_1) + \delta \frac{\partial \bar{\varphi}_G(\mathbf{x}_1)}{\partial \|\mathbf{x}_G\|} + \frac{\delta^2}{2} \frac{\partial^2 \bar{\varphi}_G(\mathbf{x}_1)}{\partial \|\mathbf{x}_G\|^2} + \mathcal{O}(\delta^3) \quad (3.78)$$

Der zweite Term dieses Ausdrucks kann wie folgt berechnet werden:

$$\delta \frac{\partial \bar{\varphi}_G(\mathbf{x}_1)}{\partial \|\mathbf{x}_G\|} = \delta \cdot \lim_{\|\mathbf{x}_2\| \rightarrow \|\mathbf{x}_1\|} \frac{(\mathbf{x}_2 - \mathbf{x}_1)^T \nabla_{\mathbf{x}} \varphi(\mathbf{x}_1)}{\|\mathbf{x}_2\| - \|\mathbf{x}_1\|} \quad (3.79)$$

Einsetzen von Gl. (3.77) und Gl. (3.79) in die Taylorreihe (3.78) liefert unter Berücksichtigung von $\delta = \|\mathbf{x}_2\| - \|\mathbf{x}_1\|$ mit $\delta \rightarrow 0$ eine Bedingung für die zweite Ableitung unter Vernachlässigung der Terme $\mathcal{O}(\delta^3)$:

$$\frac{\delta^2}{2} \frac{\partial^2 \bar{\varphi}_G(\mathbf{x}_1)}{\partial \|\mathbf{x}_G\|^2} \geq 0 \quad (3.80)$$

□

Um die verallgemeinerte Grenzkurve auf die Problemstellung der Dimensionierung anwenden zu können, ist noch zu beweisen, daß die Zielfunktion $\bar{\varphi}(\mathbf{x})$ für diesen Anwendungsfall konvex ist.

Konvexität des Optimierproblems

In dem hier vorgestellten Anwendungsfall definiert sich die Zielfunktion $\bar{\varphi}(\mathbf{x})$ nach Gl. (3.67). Da die Summe von konvexen Funktionen wieder eine konvexe Funktion ist (siehe Anhang A.2), reicht es zu zeigen, daß jeweils ein Summand $\bar{\varphi}_i(\mathbf{x})$ konvex ist. Ein Summand setzt sich dabei zusammen aus:

$$\bar{\varphi}_i(\mathbf{x}) = \exp(-a \cdot \bar{\gamma}_i(\mathbf{x})) \quad (3.81)$$

$$\bar{\gamma}_i(\mathbf{x}) = \nabla_{\mathbf{d}}^T \gamma_i \quad (3.82)$$

Die linearisierte Zielgröße γ_i ist eine lineare Funktion und damit konvex. Die Exponentialfunktion, die zur Transformation der Zielgrößen verwendet wird ist ebenfalls konvex. Da die Verknüpfung zweier konvexer Funktionen ebenfalls konvex ist, ist jeder Summand $\bar{\varphi}_i(\mathbf{x})$ der Zielfunktion $\bar{\varphi}(\mathbf{x})$ konvex und somit die Zielfunktion selbst ebenfalls konvex.

Transformation des Schaltungsfehlers und der Parameterauslenkung für die verallgemeinerte Grenzkurve

Für die charakteristische Grenzkurve werden der Fehler $\|\epsilon_c\|$ und die Auslenkung $\|\mathbf{x}_c\|$ so transformiert, daß der transformierte Fehler r und die transformierte Auslenkung a im Intervall $[0, \dots, 1]$ liegen. Somit können unterschiedliche Optimierprobleme untereinander verglichen werden. Für die verallgemeinerte Grenzkurve kann eine ähnliche Transformation vorgenommen werden. Bei der Definition der Transformation ist zu beachten, daß die Transformation die Konvexität des Optimierproblems nicht verletzt.

Für die Transformation der Auslenkung a_G kann wie in der charakteristischen Grenzkurve eine lineare Skalierung analog zu Gl. (3.13) vorgenommen werden:

$$a_G(\lambda) = \frac{\|\mathbf{x}_G(\lambda)\| - \|\mathbf{x}_G(\lambda \rightarrow \infty)\|}{\|\mathbf{x}_G(0)\| - \|\mathbf{x}_G(\lambda \rightarrow \infty)\|} \quad (3.83)$$

Eine rein lineare Skalierung der Auslenkung hat keinen Einfluß auf die Konvexität des Optimierproblems. Die parametrische Kurve $(a_G(\lambda), \bar{\varphi}_G(\lambda))$ ist ebenfalls konvex.

Um die Konvexität des Optimierproblems nicht zu verletzen, ist bei der Transformation des Fehlers zu beachten, daß bei einer Skalierung der Auslenkung nach Gl. (3.83) für die Fehlertransformation $r_G(\bar{\varphi}(\mathbf{x}))$ die Bedingung

$$\nabla_{\mathbf{xx}} r_G(\mathbf{x}) \geq 0, \quad (\text{positiv semidefinit}) \quad (3.84)$$

erfüllt sein muß. Für den in der vorliegenden Arbeit betrachteten Fall hat sich eine lineare Skalierung für $\bar{\varphi}$ bewährt:

$$r_G(\lambda) = \frac{\bar{\varphi}(\lambda) - \bar{\varphi}(0)}{\bar{\varphi}(\lambda \rightarrow \infty) - \bar{\varphi}(0)} \quad (3.85)$$

Es kann natürlich jede andere Transformation verwendet werden, solange diese die Konvexitätsbedingung (3.84) erfüllt.

3.4.3 Berechnung der verallgemeinerten Grenzkurve

Die Aufgabe der Berechnung der verallgemeinerten Grenzkurve unterteilt sich in verschiedene Aspekte:

- **Berechnung von Punkten für einen gegebenen Wert von λ :** Diese Aufgabe erfordert die Lösung des Optimierproblems (3.66) für das gegebene λ . Das Optimierproblem bedeutet die Minimierung einer nichtlinearen Zielfunktion unter Berücksichtigung linearer Nebenbedingungen. Für diesen Anwendungsfall existieren sogenannte Projektionsverfahren [NW99, Spe98], die insbesondere für den Fall linearer Nebenbedingungen sehr effektiv sind. Diese Verfahren sind vergleichbar mit dem für die quadratische Programmierung verwendeten PQP-Verfahren, nur daß sie die Hesse-Matrix iterativ durch ein Update-Verfahren (z. B. BFGS-Update) approximieren. Insbesondere sind bei diesen Verfahren nach der erstmaligen Bestimmung einer zulässigen Lösung, die alle Nebenbedingungen erfüllt, alle weiteren Lösungen im zulässigen Bereich.
- **Approximation der gesamten Kurve:** Die verallgemeinerte Grenzkurve erbt alle wichtigen Eigenschaften von der charakteristischen Grenzkurve. Dazu zählt speziell die Konvexität, die bei der Approximation der charakteristischen Grenzkurve durch Tangenten (Abschnitt 3.2.6) bzw. BB-Kurven (Abschnitt 3.2.7) als Voraussetzung genommen wurde. Damit sind diese Verfahren ohne Änderungen auf die verallgemeinerte Grenzkurve übertragbar. Auch für die verallgemeinerte Grenzkurve müssen bei der Approximation der gesamten Kurve durch BB-Kurven zumeist deutlich weniger Punkte berechnet werden, um einen vorgegebenen maximalen Approximationsfehler unterschreiten.

- **Berechnung von Punkten mit vorgegebenem Fehler, Auslenkung oder Steigung:** Das in Abschnitt 3.2.8 vorgestellte Verfahren zur Berechnung von Punkten auf der Grenzkurve hat als Voraussetzung die Konvexität der Kurve. Damit kann auch dieses Verfahren für die verallgemeinerte Grenzkurve ohne Änderungen übernommen werden.
- **Bestimmung einer sinnvollen Schrittweite:** Für die Bestimmung einer sinnvollen Schrittweite, die ein günstiges Verhältnis zwischen Fehlerreduktion und Größe der Parameterauslenkung hat, können ebenfalls die in Abschnitt 3.3 diskutierten Verfahren verwendet werden.

Alle Verfahren, die für die charakteristische Grenzkurve in Abschnitt 3.2 und 3.3 vorgestellt wurden, können ohne Einschränkungen direkt für die verallgemeinerte Grenzkurve verwendet werden. Einzig die Steigung m an bereits berechneten Punkten der verallgemeinerten Grenzkurve wird zusätzliche benötigt.

3.4.4 Steigung der verallgemeinerten Grenzkurve

In diesem Abschnitt wird die Steigung $m_G = \partial r_G / \partial a_G$ der verallgemeinerten Grenzkurve bestimmt. Die Steigung wird für die drei Fälle $\lambda \rightarrow \infty$, $\lambda \in]\infty, 0[$ und $\lambda = 0$ getrennt betrachtet. Zunächst soll der Normalfall für einen Punkt innerhalb der Kurve betrachtet werden.

Steigung der verallgemeinerten Grenzkurve für $\infty > \lambda > 0$

Die Steigung für diesen Fall kann analog zur charakteristischen Grenzkurve in Abschnitt 3.2.5 bestimmt werden:

$$m_G(\lambda) = \frac{\partial r_G(\lambda)}{\partial a_G(\lambda)} = \frac{\partial r_G(\lambda)}{\partial \bar{\varphi}_G(\lambda)} \cdot \frac{\partial \bar{\varphi}_G(\lambda)}{\partial \|\mathbf{x}_G(\lambda)\|^2} \cdot \frac{\partial \|\mathbf{x}_G(\lambda)\|^2}{\partial a_G(\lambda)} \quad (3.86)$$

Wird die in Gl. (3.83) bzw. Gl. (3.85) vorgeschlagene Transformation für die Auslenkung $\|\mathbf{x}_G(\lambda)\|$ und den Schaltungsfehler $\bar{\varphi}_G(\lambda)$ verwendet, dann können die einzelnen Terme von Gl. (3.86) wie folgt bestimmt werden:

$$\frac{\partial r_G(\lambda)}{\partial \bar{\varphi}_G(\lambda)} = \frac{1}{\bar{\varphi}(\lambda \rightarrow \infty) - \bar{\varphi}(0)} \quad (3.87)$$

$$\frac{\partial \bar{\varphi}_G(\lambda)}{\partial \|\mathbf{x}_G(\lambda)\|^2} = -\lambda \quad (3.88)$$

$$\frac{\partial \|\mathbf{x}_G(\lambda)\|^2}{\partial a_G(\lambda)} = 2 \cdot \|\mathbf{x}_G(\lambda)\| \cdot (\|\mathbf{x}_G(0)\| - \|\mathbf{x}_G(\lambda \rightarrow \infty)\|) \quad (3.89)$$

Und als Steigung ergibt sich dann:

$$m_G(\lambda) = -2\lambda \cdot \|\mathbf{x}_G(\lambda)\| \cdot \frac{(\|\mathbf{x}_G(0)\| - \|\mathbf{x}_G(\lambda \rightarrow \infty)\|)}{\bar{\varphi}(\lambda \rightarrow \infty) - \bar{\varphi}(0)} \quad (3.90)$$

In Gl. (3.88) wurde dabei wieder berücksichtigt, daß nach Gl. (A.11) der Lagrange-Parameter λ die Empfindlichkeit der Zielfunktion gegenüber den Nebenbedingungen $\|\mathbf{x}\|^2 = \|\mathbf{x}_G(\lambda)\|^2$ beschreibt. Wird für den Schaltungsfehler eine andere Transformation als Gl. (3.85) verwendet, so ist zur Bestimmung der Steigung nur der Ausdruck der Ableitung $\partial r_G(\lambda) / \partial \bar{\varphi}_G(\lambda)$ zu ändern.

Steigung der verallgemeinerten Grenzkurve für $\lambda \rightarrow \infty$

Setzt man $\lambda = \infty$ in obige Gleichung für die Steigung (3.90) ein, so lassen sich zwei Fälle unterscheiden:

- Im Startpunkt sind Nebenbedingungen aktiv. In diesem Fall ist $\|\mathbf{x}_G(\lambda)\| \neq 0$ für $\lambda \rightarrow \infty$. Aus Gl. (3.90) ergibt sich der Wert der Steigung zu $m_G(\lambda \rightarrow \infty) = -\infty$.
- Im Startpunkt sind keine Nebenbedingungen aktiv und damit ist $\|\mathbf{x}_G(\lambda)\| = 0$ für $\lambda \rightarrow \infty$. Die Steigung ergibt sich in diesem Fall zu $\infty \cdot 0$. Dieser Fall soll im folgenden behandelt werden.

Für den Fall $\lambda \rightarrow \infty$ und $\mathbf{x}_G(\lambda \rightarrow \infty) = \mathbf{0}$ muß zur Bestimmung der Steigung der Term $\partial\bar{\varphi}_G/\partial a_G$ an der Stelle $\lambda \rightarrow \infty$ abgeschätzt werden. Unter Umgehung des Quadrates $\|\mathbf{x}\|^2$ kann dieser Term aus dem Verhalten der Grenzkurve für $\mathbf{x} = \mathbf{0}$ bestimmt werden.

Im Anfangspunkt ist nach Voraussetzung die Auslenkung gleich Null. Da die Auslenkung quadratisch in die modifizierte Zielfunktion eingeht, ist auch ihr Anteil an der Ableitung der modifizierten Zielfunktion verschwindend. Der Gradient der Zielfunktion $\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x})$ bestimmt daher im Ausgangspunkt die Richtung der Auslenkung. Die Änderung $\Delta\mathbf{x}$ der Entwurfsparameter für eine Auslenkung der Größe δ im Ausgangspunkt $\mathbf{x} = \mathbf{0}$ ergibt sich deshalb zu:

$$\Delta\mathbf{x} = \delta \cdot \left. \frac{-\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x})}{\|\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x})\|} \right|_{\mathbf{x}=\mathbf{0}} \quad (3.91)$$

Die Änderung der Zielfunktion für kleine Auslenkungen δ kann mit einer Linearisierung abgeschätzt werden:

$$\Delta\bar{\varphi}(\Delta\mathbf{x}) = -\delta \cdot \nabla_{\mathbf{x}}^T\bar{\varphi}(\mathbf{x}=\mathbf{0}) \frac{-\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x})}{\|\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x})\|} + \mathcal{O}(\delta^2) \quad (3.92)$$

Mit $\delta \rightarrow 0$ bestimmt sich die Ableitung $\partial\bar{\varphi}_G/\partial\|\mathbf{x}_G\|$ zu:

$$\left. \frac{\partial\bar{\varphi}_G}{\partial\|\mathbf{x}_G\|} \right|_{\lambda \rightarrow \infty} = -\|\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x}=\mathbf{0})\| \quad (3.93)$$

Unter Berücksichtigung der weiteren Ableitungen $\partial r_G/\partial\bar{\varphi}_G$ und $\partial\|\mathbf{x}_G\|/\partial a_G$, die aus Gl. (3.83) bzw. Gl. (3.85) berechnet werden können, ergibt sich die Steigung der verallgemeinerten Grenzkurve für diesen Fall:

$$m_G(\lambda \rightarrow \infty) = \begin{cases} -\infty & \text{für } \mathbf{x}(\lambda \rightarrow \infty) \neq \mathbf{0} \\ \frac{\mathbf{x}_G(0) - \mathbf{x}_G(\lambda \rightarrow \infty)}{\bar{\varphi}_G(\lambda \rightarrow \infty) - \bar{\varphi}_G(0)} \cdot \|\nabla_{\mathbf{x}}\bar{\varphi}(\mathbf{x}=\mathbf{0})\| & \text{für } \mathbf{x}(\lambda \rightarrow \infty) = \mathbf{0} \end{cases} \quad (3.94)$$

Steigung der verallgemeinerten Grenzkurve für $\lambda = 0$

Auch für die Steigung $m_G(\lambda = 0)$ lassen sich wieder zwei Fälle unterscheiden:

- Der zulässige Bereich ist nicht beschränkt und die Lösung für $\lambda = 0$ hat die Auslenkung $\|\mathbf{x}_G(\lambda = 0)\| = \infty$. In diesem Fall kann keine Aussage über die Steigung gemacht werden. Dieser Fall ist allerdings für die Praxis nicht relevant. In der Regel sind für die Parameter Untergrenzen (z. B. minimale Strukturgrößen) und Obergrenzen gegeben. Der zulässige Bereich ist damit in der Praxis immer abgeschlossen.
- Der zulässige Bereich ist abgeschlossen. In diesem Fall kann $\lambda = 0$ in Gl. (3.90) eingesetzt werden und die Steigung ergibt sich zu $m_G(\lambda = 0) = -0$.

Für die in der Praxis relevanten Fälle kann die Steigung im Endpunkt $\lambda = 0$ aufgrund der Abgeschlossenheit des zulässigen Bereiches immer als $m_G(\lambda = 0) = -0$ angenommen werden.

3.5 Abbruchkriterien

Vernünftige Abbruchkriterien sind mitentscheidend für eine Einsetzbarkeit und Akzeptanz einer automatischen Dimensionierung. Bei den Abbruchkriterien sind dabei folgende Aspekte in Betracht zu ziehen:

1. Die Auswertung der Zielfunktion beruht letztendlich auf den Werten der Schaltungseigenschaften, die durch eine Schaltungssimulation gewonnen werden. Die Genauigkeit des Schaltungssimulators liegt dabei weit unterhalb der Maschinengenauigkeit. Eine Verbesserung der Genauigkeit einer Simulation ist zumeist nur unter einer erheblichen Verlängerung der Simulationszeiten möglich. Für eine transiente Analyse bedeutet eine höhere Simulationsgenauigkeit z. B. kleinere Schrittweiten, so daß für die gleiche Simulation mehr Zeitschritte benötigt werden. Diese relativ geringe Genauigkeit bei der Auswertung der Zielfunktion ist bei den Abbruchkriterien zu berücksichtigen.
2. Die Berechnung der benötigten Gradienten erfolgt teilweise mittels finiter Differenzen. Um die Anzahl der hierzu notwendigen Schaltungssimulationen klein zu halten, werden per Voreinstellung nur finite Vorwärtsdifferenzen verwendet und kein zentraler Differenzenquotient. Die Genauigkeit der Gradienten ist damit in der Regel nochmals um eine Größenordnung schlechter als die Berechnungsgenauigkeit der Zielfunktion.
3. Die Entwurfparameter \mathbf{d} sind in der Regel nicht mit beliebiger Genauigkeit einstellbar, sondern müssen auf einem definierten Gitter (*Grid*) liegen. Änderungen der Parameter, die um mehrere Größenordnungen unterhalb der Gitters liegen, sind nicht mehr sinnvoll. Ist für einen Parameter kein Gitter vorgegeben, so gilt in der Realität trotzdem, daß dieser Parameter nicht mit unendlich hoher Genauigkeit produziert werden kann.

Aus diesen Aspekten lassen sich geeignete Abbruchkriterien herleiten. Punkt 1 und 2 obiger Liste begrenzen die Genauigkeit der Zielfunktion und der Gradientenbestimmung. Eine nurmehr geringe Verbesserung der Zielfunktion zwischen zwei Iterationen μ und $\mu + 1$ läßt weitere Iterationen des Dimensionierungsalgorithmus nicht sinnvoll erscheinen.

Eine sehr kleine Schrittweite $\|\mathbf{d}^{(\mu)} - \mathbf{d}^{(\mu+1)}\|$, die deutlich unterhalb der Gittergröße der Entwurfparameter liegt, läßt keine signifikanten Änderungen der Entwurfparameter erwarten. Die zwei verwendeten Abbruchkriterien lauten damit:

Minimale Verbesserung der Zielfunktion: $\varphi(\mathbf{d}^{(\mu)}) - \varphi(\mathbf{d}^{(\mu+1)}) \geq \Delta\varphi_{\min}$ (3.95)

Minimale Größe der Parameterkorrektur: $\|\mathbf{d}^{(\mu)} - \mathbf{d}^{(\mu+1)}\| \geq \Delta d_{\min}$ (3.96)

Ist eines der obigen Kriterien nicht erfüllt, so wird der Dimensionierungsalgorithmus abgebrochen und die in der letzten Iteration bestimmten Werte der Entwurfparameter als Schaltdimensionierung vorgeschlagen.

Kapitel 4

Ergebnisse

In diesem Abschnitt werden die in der vorliegenden Arbeit vorgestellten Verfahren auf zwei Schaltungsbeispiele angewendet. Dabei sollen folgende Aspekte verdeutlicht werden:

- Vergleich der charakteristischen Grenzkurve und verallgemeinerten Grenzkurve für die in dieser Arbeit diskutierte Problemstellung.
- Vergleich der Ergebnisse bei einer frühen Berücksichtigung des Einflusses der Betriebsparameter.
- Einfluß der konsequenten Berücksichtigung der strukturellen Nebenbedingungen während der gesamten Dimensionierung.
- Anwendbarkeit der vorgestellten Verfahren auf Schaltungen aus unterschiedlichen Anwendungsbereichen.

Die in dieser Arbeit vorgestellten Verfahren sind in dem Tool WICKED [AEG⁺00a] implementiert, das bei Infineon Technologies in unterschiedlichen Entwicklungsbereichen verwendet wird. Aus zwei verschiedenen Anwendungsbereichen wurden exemplarisch zwei Schaltungen zur Veranschaulichung der vorgestellten Verfahren ausgewählt. Der Operationsverstärker repräsentiert eine typischen analogen Schaltungsblock während der Ausgangstreiber ein Beispiel für eine digitale Schaltung ist, die im Rahmen des Bibliotheksentwurfs auf Transistorebene dimensioniert wird. Andere untersuchte Schaltungen (u.a. Equalizer, Filter, Band-Gap) zeigen vergleichbare Ergebnisse.

4.1 Operationsverstärker

Der Miller Operationsverstärker (Abbildung 4.1) ist eine typische analoge Standardschaltung. Die Schaltung besteht aus 7 MOS-Transistoren, was 14 potentiellen Entwurfparametern entspricht. Die Betriebsbedingungen werden durch 4 Betriebsparameter (Versorgungsspannung V_{dd} , Temperatur T , Bias-Strom I_{bias} und Lastkapazität C_{load}) modelliert. Die Fertigungsstreuungen beschreiben die 7 statistischen Parameter.

Die Schaltung besteht aus drei Stromspiegeln (M3-M4, M8-M5, M8-M7), einem Differenzpaar (M1-M2) und dem Transistor M6, der als spannungsgesteuerte Stromquelle arbeitet. Nach

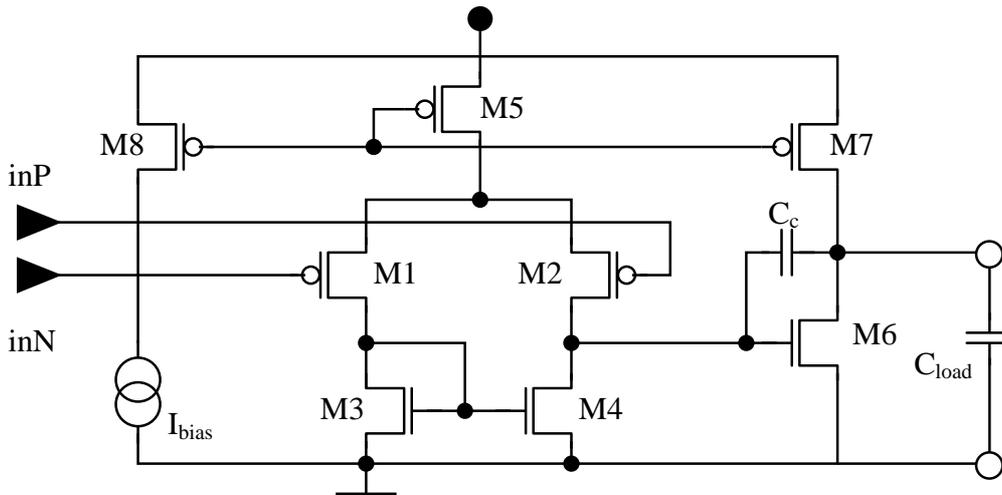


Abbildung 4.1: Beispielschaltung: Miller Operationsverstärker

NMOS	Länge	Weite		PMOS	Länge	Weite
M3	l3	w3		M1	l1	w1
M4	l3	w3		M2	l1	w1
M6	l3	k·w3		M5	l5	w5
				M7	l5	0.5·k·w5
				M8	l5	w8

Tabelle 4.1: Längen und Weiten der Transistoren für den Miller Operationsverstärker.

Abschnitt 2.1.3 ergeben sich somit die in Tabelle 4.2 dargestellten strukturellen Nebenbedingungen. Für die Weiten und Längen der MOS-Transistoren ergibt sich damit die in Tabelle 4.1 angegebene Parametrierung der Schaltung.

Mit den vorgestellten Verfahren wurden für die Dimensionierung des Operationsverstärkers bei Berechnung der Empfindlichkeiten über Differenzenquotienten 1110 Simulationen benötigt. Eine Simulation entspricht hierbei der Extraktion aller Eigenschaften (DC-, AC- und transiente Simulation). Die gesamte für die Simulation benötigte CPU Zeit auf einem Intel Pentium III, 500MHz beträgt ca. 180min. Erfolgen die Simulationen so weit wie möglich parallel auf einem Netzwerk von 10 Intel Pentium III 500MHz Rechnern so werden ca. 60min „Wall-Clock Time“ benötigt. Tabelle 4.3 faßt die Ergebnisse für die Dimensionierung getrennt nach Nominalentwurf und Entwurfszentrierung zusammen.

Die Ergebnisse zeigen einerseits deutlich, daß eine automatische Dimensionierung des Operations-

	Gleichheitsbedingungen	Ungleichheitsbedingungen
3 Stromspiegel	6	36
1 Differenzpaar	2	14
1 Spannungsgesteuerte Stromquelle	—	3

Tabelle 4.2: Strukturelle Nebenbedingungen für den Miller Operationsverstärker.

Eigenschaft	Spezifikation	Ausgangspunkt	Nominalentwurf	Entwurfszentrierung
Verstärkung A_0 [dB]	>80	63	84 ($\beta = 4.5$)	83 ($\beta = 4.3$)
Transitfrequenz f_t [MHz]	>4.0	4.1	4.6 ($\beta = 1.8$)	3.3 ($\beta = 4.6$)
Phasenrandreserve φ_M [°]	>60	52	68 ($\beta > 10$)	73 ($\beta = 5.3$)
Slewrates $\uparrow SR_p$ [V/ μ s]	>4.0	3.4	5.0 ($\beta = 1.2$)	5.3 ($\beta = 3.6$)
Slewrates $\downarrow SR_n$ [V/ μ s]	>4.0	7.2	5.3 ($\beta = 1.6$)	5.8 ($\beta = 3.6$)
Leistung Pwr [mW]	<1.1	0.6	0.84 ($\beta = 4.6$)	0.88 ($\beta = 3.6$)
Iterationen	—	—	10	3
Simulationen	—	—	170	940

Tabelle 4.3: Ergebnisse der Dimensionierung des Operationsverstärkers aufgespalten in Nominalentwurf und Entwurfszentrierung.

verstärkers mit dem vorgestellten Verfahren möglich ist. Andererseits zeigt die Anzahl der Simulationen und die Simulationsdauer im Nominalentwurf und Entwurfszentrierung, daß eine Aufspaltung des Entwurfs in diese zwei Schritte aufgrund des hohen Simulationsaufwandes in der Entwurfszentrierung sinnvoll ist. Obwohl die Iterationszahl in der Entwurfszentrierung deutlich geringer ist als für den Nominalentwurf, wird in der Entwurfszentrierung die fünffache Menge an Simulationen benötigt. Dies ist vor allen Dingen darauf zurückzuführen, daß die Berechnung der Zielgrößen für die Entwurfszentrierung selbst wieder ein Optimierproblem ist. Für eine Beschleunigung des Verfahrens wäre es denkbar, zu Beginn der Entwurfszentrierung mit Worst-Case Abständen basierend auf einem linearisierten Schaltungsmodell vergleichbar den Linearized Performance Penalties (LPPs) [KD95] zu arbeiten. In diesem Fall kann eine Konvergenz des Verfahrens aber nicht mehr bewiesen werden.

Einfluß der frühen Berücksichtigung der Betriebsparameter auf die Dimensionierung

Eigenschaft	Mit Betriebsparameter			Ohne Betriebsparameter		
	Nominal	θ_{WC}	WCD (β_{WC})	Nominal	θ_{WC}	WCD (β_{WC})
A_0	84	83	4.5	87	86	5.1
f_t	4.6	4.1	1.8	5.3	4.8	8.3
φ_M	68	62	10.7	65	58	-5.0
SR_p	5.0	3.9	1.2	5.6	4.0	1.3
SR_n	5.3	4.3	1.6	6.0	5.7	1.8
Pwr	0.84	1.0	4.6	0.95	1.1	4.4

Tabelle 4.4: Vergleich von Nominalentwurf mit und ohne Berücksichtigung der Betriebsparameter anhand der Nominalwerte der Eigenschaften, der Eigenschaftswerte im Worst-Case Betriebsparameterbereich θ_{WC} und der Worst-Case Abstände β_{WC} .

Die frühe Berücksichtigung der Betriebsparameter bereits im Nominalentwurf ermöglicht eine Nominaldimensionierung der Schaltung, die einen guten Ausgangspunkt für die Entwurfszentrierung bildet. Tabelle 4.4 vergleicht einen Nominalentwurf mit und ohne Berücksichtigung der Betriebsparameter bei der Berechnung der Parameterabstände. Es zeigt sich, daß einige Eigenschaften sehr stark von den Betriebsparametern beeinflusst werden, während die Betriebsparameter auf andere Eigenschaften

wenig oder keinen Einfluß haben. Dies zeigt sich auch deutlich an den Worst-Case Abständen, die bei Berücksichtigung der Betriebsparameter bereits im Nominalentwurf deutlich ausgeglichener sind.

Einfluß der strukturellen Nebenbedingungen auf die Dimensionierung

Die konsequente Berücksichtigung der strukturellen Nebenbedingungen während der gesamten Dimensionierung stellt sicher, daß nur schaltungstechnisch sinnvolle Lösungen berechnet werden. Wie unten gezeigt, ermöglicht erst die Beschränkung der Parameter auf ein schaltungstechnisch sinnvolles Gebiet eine sinnvolle automatische Dimensionierung analoger Schaltungen.

Wie in [ZEG98, ZEG99, SPZ⁺01] gezeigt, sind die Schaltungseigenschaften im Bereich erfüllter Nebenbedingungen (*Feasibility region*) zumeist nur eine schwach nichtlineare Funktion der Entwurfparameter. Dies wird nochmals in Abbildung 4.2 verdeutlicht. Diese Abbildung zeigt die Verstärkung und die Sättigungsbedingung für einen Transistor in Abhängigkeit eines Entwurfparameters zwischen dessen vom Schaltungsentwickler vorgegebenen Grenzen. Über einen weiten Bereich zeigt die Schaltung einen nahezu linearen Verlauf der Verstärkung. An dem Punkt an dem die Nebenbedingung verletzt wird, knickt die Verstärkung dagegen ab. Eine Beschränkung der Entwurfparameter auf einen schaltungstechnisch sinnvollen Bereich durch die konsequente Berücksichtigung der strukturellen Nebenbedingungen ermöglicht also erst eine effiziente automatische Dimensionierung robuster analoger Schaltungen.

Dies soll ebenfalls am Miller Operationsverstärker verdeutlicht werden. Dazu wird der Nominalentwurf für den Operationsverstärker mit den in Tabelle 4.3 angegebenen Spezifikationen mit und ohne die Berücksichtigung struktureller Nebenbedingungen verglichen.

Eigenschaft	Spezifikation	ohne Nebenbedingungen	mit Nebenbedingungen
Verstärkung A_0 [dB]	>80	85 ($\beta = -4.0$)	84 ($\beta = 4.5$)
Transitfrequenz f_t [MHz]	>4.0	4.8 ($\beta = 0.5$)	4.6 ($\beta = 1.8$)
Phasenrandreserve φ_M [°]	>60	65 ($\beta = -6.0$)	68 ($\beta > 10$)
Slewrates $\uparrow SR_p$ [V/ μ s]	>4.0	5.3 ($\beta = -1.7$)	5.0 ($\beta = 1.2$)
Slewrates $\downarrow SR_n$ [V/ μ s]	>4.0	5.8 ($\beta = -0.3$)	5.3 ($\beta = 1.6$)
Leistung Pwr [mW]	<1.1	0.87 ($\beta = 8.3$)	0.84 ($\beta = 4.6$)

Tabelle 4.5: Vergleich der Dimensionierung des Miller Operationsverstärkers mit und ohne die Berücksichtigung struktureller Nebenbedingungen.

Der Vergleich ist in Tabelle 4.5 dargestellt. Es zeigt sich, daß ohne die Berücksichtigung struktureller Nebenbedingungen kein sinnvoller Parametersatz berechnet werden kann. Die Nominalwerte der Eigenschaften sind teilweise besser als bei der Dimensionierung mit der Berücksichtigung struktureller Nebenbedingungen. Allerdings zeigt sich eine stark erhöhte Empfindlichkeit gegenüber Schwankungen der Betriebsparameter (siehe Abbildung 4.3). Dies resultiert darin, daß die Spezifikation im Worst-Case Betriebsparameterbereich bereits verletzt ist – der Worst-Case Abstand also negativ ist.

Abbildung 4.3 vergleicht den Verlauf der Verstärkung A_0 des Operationsverstärkers über der Betriebsspannung nach dem Nominalentwurf mit und ohne strukturelle Nebenbedingungen. Obwohl bei der Dimensionierung ohne Berücksichtigung struktureller Nebenbedingungen der Nominalwert der

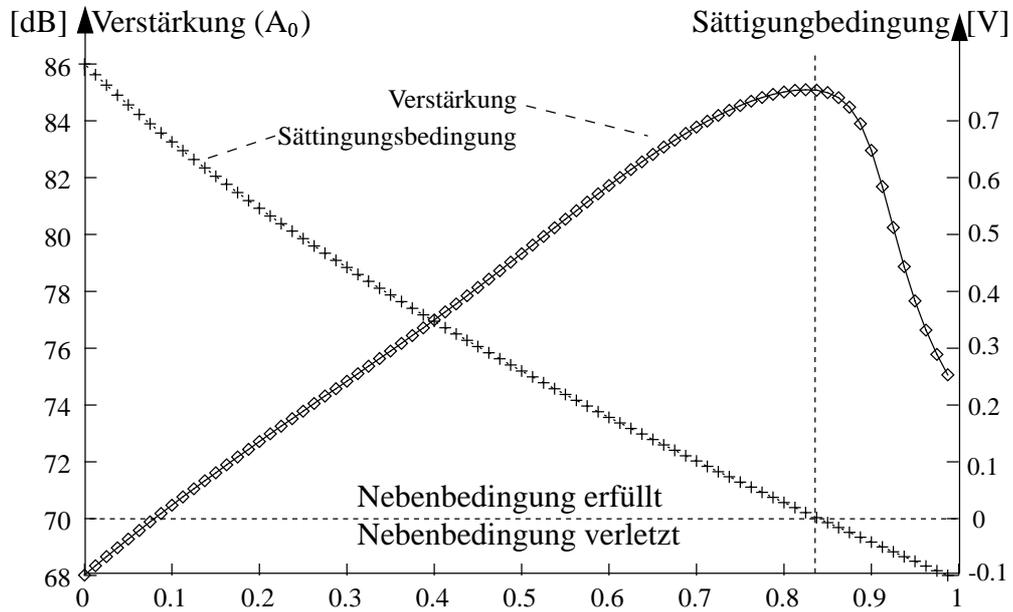


Abbildung 4.2: Verstärkung und eine strukturelle Nebenbedingung in Abhängigkeit der Entwurfparameter.

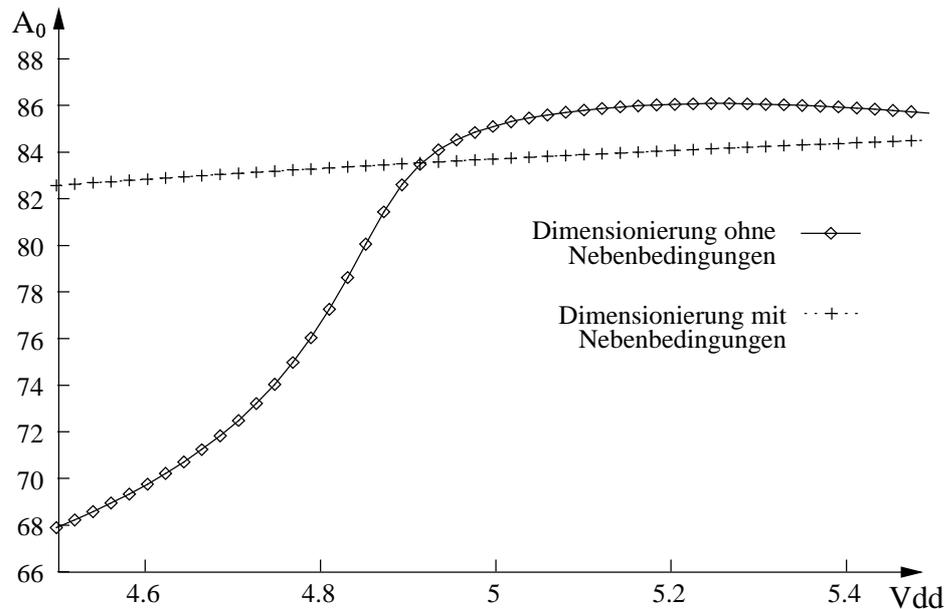


Abbildung 4.3: Verlauf der Verstärkung über der Versorgungsspannung nach dem Nominalentwurf mit und ohne Berücksichtigung der strukturellen Nebenbedingungen.

Eigenschaft besser ist, ist die Schaltung aufgrund der starken Abhängigkeit deutlich schlechter dimensioniert als ohne Berücksichtigung struktureller Nebenbedingungen. Eine konsequente Berücksichtigung struktureller Nebenbedingungen ist also eine Voraussetzung für eine sinnvolle Anwendung einer automatischen Dimensionierung analoger Schaltungen.

4.2 Ausgangstreiber

Der Ausgangstreiber nach Abbildung 4.4 ist eine Standardtopologie, die für unterschiedliche Technologien mit unterschiedlichen minimalen Strukturgrößen Verwendung findet. Durch den Einsatz einer automatischen Dimensionierung erhofft man sich für diesen Fall auch die Möglichkeit eines automatischen Technologietransfers. Ausgehend von der Dimensionierung in einer vorhergehenden Technologie soll die Schaltung so dimensioniert werden, daß sie die an sie gestellten Forderungen in der neuen Technologie erfüllt. Die Startdimensionierung liefert in der Regel bereits einen sinnvollen Ausgangspunkt. Insbesondere konvergiert die Simulation für diese Startdimensionierung, was eine Voraussetzung für den Einsatz der vorgestellten Verfahren ist.

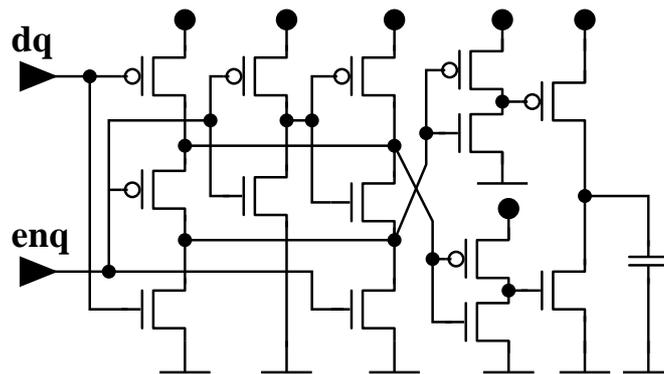


Abbildung 4.4: Beispielschaltung Ausgangstreiber

Eigenschaft	Spezifikation	Ausgangspunkt	Nominalentwurf	Entwurfzentrierung
Delay rise $t_{d,rise}$ [ns]	<4.0	4.1	3.2 ($\beta = 3.9$)	2.8 ($\beta = 4.2$)
Delay fall $t_{d,fall}$ [ns]	<4.0	3.4	2.8 ($\beta > 10$)	3.1 ($\beta = 4.6$)
Anstiegszeit t_r [ns]	<4.0	6.7	3.0 ($\beta = 2.0$)	2.7 ($\beta = 6.4$)
Abfallzeit t_f [ns]	<4.0	5.1	3.0 ($\beta = 1.7$)	2.6 ($\beta = 6.7$)
Noise (vdd) [mV/nH]	<20	12.2	13.4 ($\beta > 6.8$)	14.1 ($\beta = 4.3$)
Noise (vss) [mV/nH]	<20	12.2	13.1 ($\beta = 5.6$)	12.6 ($\beta = 5.7$)
Iterationen	—	—	4	2
Simulationen	—	—	52	342

Tabelle 4.6: Dimensionierungsergebnisse für den Ausgangstreiber unterteilt nach Nominalentwurf und Entwurfzentrierung.

Tabelle 4.6 zeigt die Dimensionierungsergebnisse für den Ausgangstreiber ausgehend von einer vorhergehenden Technologie. Die Längen der Transistoren sind fest auf die minimal vorgegebenen

Längen fixiert, da für den Ausgangstreiber als digitale Schaltungsklasse Mismatch oder erhöhte Streuungen der Einsatzspannung aufgrund von Kurzkanaleffekten keine Rolle spielen. Die Transistorweiten sind als Entwurfparameter definiert, wobei für die Eingangslogik des Treibers jeweils alle p-Kanal und n-Kanal Transistoren die gleich Weite haben. Insgesamt ergeben sich dadurch acht Entwurfparameter. Die Betriebsbedingungen sind durch die Lastkapazität und die Versorgungsspannung modelliert.

Im Endpunkt der Entwurfszentrierung sind die Worst-Case Abstände für das Schaltrauschen auf der Versorgungsspannung (Noise (vdd)) und die Verzögerungszeit für die steigende Flanke (Delay rise $t_{d, \text{rise}}$) nahezu gleich groß. Diese Eigenschaften können nicht miteinander verbessert werden, sondern sind stark negativ miteinander korreliert. Dies zeigt sich auch wenn man die Gradienten der Worst-Case Abstände betrachtet. Der Korrelationsfaktor im Lösungspunkt für die Entwurfszentrierung beträgt 0.983. Die komplette Korrelationsmatrix für alle Worst-Case Abstände im Lösungspunkt ist in Abbildung 4.5 dargestellt.

Linearisierungsfehler für charakteristische und verallgemeinerte Grenzkurve

Der Ausgangstreiber zeigt, daß das vorgestellte Verfahren für den Anwendungszweck des Technologietransfers geeignet ist. Am Beispiel des Ausgangstreibers soll gezeigt werden, wie sich der Linearisierungsfehler der linearisierten Zielfunktion nach Gl. (3.1) und der Zielfunktion nach Gl. (3.67) basierend auf den linearisierten Zielgrößen verhält. Dazu wird zwischen der Lösung des Nominalentwurfs und der Entwurfszentrierung die wirkliche Zielfunktion Gl. (2.44), die linearisierte Zielfunktion Gl. (3.1) und die Zielfunktion Gl. (3.67) basierend auf den linearisierten Worst-Case Abständen in Abbildung 4.6 dargestellt. Die linearisierte Zielfunktion nach Gl. (3.1) – in der Abbildung durch ein '□' gekennzeichnet – bildet die Grundlage der charakteristischen Grenzkurve. Die Zielfunktion basierend auf den linearisierten Worst-Case Abständen nach Gl. (3.67) – in der Abbildung durch ein '+' gekennzeichnet – die Basis für die verallgemeinerte Grenzkurve. Der Linearisierungsfehler dieser Funktionen läßt sich also auf die charakteristische bzw. verallgemeinerte Grenzkurve übertragen.

Die nichtlineare Zielfunktion und die Zielfunktion basierend auf dem linearisierten Worst-Case Abstand sind praktisch nicht voneinander zu unterscheiden. Der maximale Linearisierungsfehler liegt immer unter 2%. Der Linearisierungsfehler für die vollständig linearisierte Zielfunktion dagegen ist bereits für kleine Auslenkungen deutlich zu groß. Mit steigender Auslenkung steigt er exponentiell an. Tabelle 4.7 faßt nochmals den Vergleich der Linearisierung von dem Worst-Case Abstand und den unterschiedlichen Linearisierungen der Zielfunktion zusammen. An diesem Beispiel zeigt sich deutlich, daß der Linearisierungsfehler durch die Berücksichtigung der starken nichtlinearität der Exponentialfunktion deutlich reduziert werden kann.



Abbildung 4.5: Korrelationsmatrix im Lösungspunkt der Entwurfszentrierung.

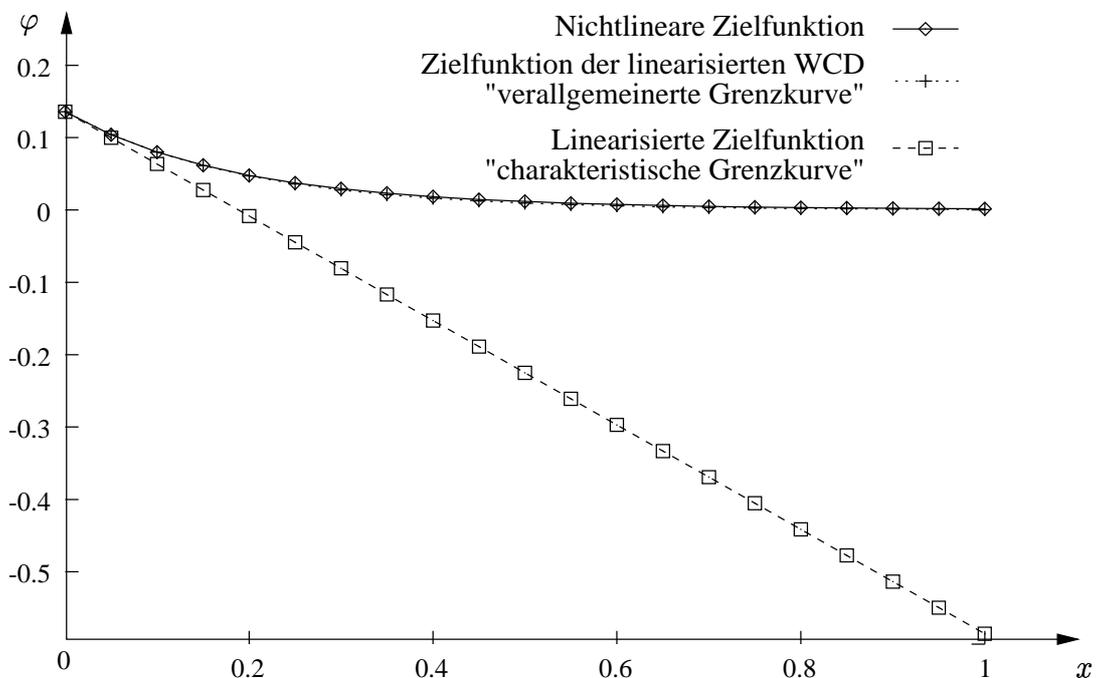


Abbildung 4.6: Vergleich des Linearisierungsfehlers zwischen linearisierter Zielfunktion und Zielfunktion für den linearisierten Worst-Case Abstand der Anstiegszeit t_r basierend auf dem linearisierten Worst-Case Abstand zwischen dem Endpunkt des Nominalentwurfs ($x = 0$) und dem Endpunkt der Entwurfszentrierung ($x = 1$).

Auslenkung	0	0.25	0.5	0.75	1.0
Worst-Case Abstand nichtlinear	2.00	3.29	4.45	5.50	6.44
Worst-Case Abstand linearisiert	2.00	3.32	4.65	6.0	7.31
Linearisierungsfehler	—	2.3%	8.1%	14.2%	19.6%
Nichtlineare Zielfunktion	0.136	0.0374	0.0117	0.0041	0.0016
Linearisierte Zielfunktion und Linearisierungsfehler	0.136	0.0360	0.0095	0.0025	0.0007
	—	1.4%	1.8%	1.1%	0.7%
Zielfunktion der linearisierten WCD und Linearisierungsfehler	0.136	-0.0085	-0.225	-0.405	-0.586
	—	147%	265%	398%	531%

Tabelle 4.7: Linearisierungsfehler des Worst-Case Abstandes der Anstiegszeit t_r zwischen Endpunkt des Nominalentwurfs ($x = 0$) und der Entwurfszentrierung ($x = 1$).

Kapitel 5

Zusammenfassung

Analoge Subsysteme und Schaltungsblöcke stellen in integrierten Schaltungen und Systemen (*System on a Chip*) unverzichtbare Funktionen zur Verfügung. Trotz ihres im Vergleich zum digitalen Schaltungsteil kleinen Flächenanteils an der Gesamtfläche einer integrierten Schaltung verursachen analoge Blöcke einen erheblichen Teil des Entwicklungsaufwandes. Der Entwurf einer analogen Schaltung kann in Topologieentwurf, Dimensionierung und Layout unterteilt werden, wobei die Dimensionierung deutlich mehr als ein Drittel des gesamten Aufwandes verursacht. Die Dimensionierung analoger Schaltungen erfolgt zumeist simulationsbasiert und manuell durch erfahrene Schaltungsentwickler und ist daher eine zeitaufwendige und fehleranfällige Tätigkeit.

Die Dimensionierung einer analogen Schaltung wiederum kann in Nominalentwurf und Entwurfszentrierung unterteilt werden. Ziel des Nominalentwurfs ist es, daß die Schaltung ohne Berücksichtigung des Einflusses von unvermeidlichen Streuungen des Fertigungsprozesses und der Betriebsbedingungen die an sie gestellten Anforderungen so gut wie möglich erfüllt. In der nachfolgenden Entwurfszentrierung wird die Schaltung so dimensioniert, daß sie möglichst robust gegenüber Prozessstreuungen und Variationen der Betriebsumgebung ist. Dies bedeutet auch eine möglichst hohe Schaltungsausbeute.

In dieser Arbeit werden Verfahren zum Nominalentwurf und zur Entwurfszentrierung analoger Schaltungsblöcke (typischerweise zwischen 10 und 30 freie Parameter) auf Transistorebene vorgestellt. Besonderer Augenmerk wurde auf die allgemeine Anwendbarkeit der Verfahren gelegt. Die präsentierten Verfahren unterliegen keinerlei Einschränkungen bezüglich der zu untersuchenden Eigenschaften (z. B. nur Kleinsignaleigenschaften) oder Schaltungsklassen (z. B. nur Operationsverstärker). Um diese Forderungen zu erfüllen, finden in dieser Arbeit simulationsbasierte Verfahren Verwendung. Die Dimensionierung wird als mathematisches Optimierproblem formuliert. Die Auswertung der Zielfunktion und die Bestimmung der benötigten Gradienten erfordert in der Regel mehrere Schaltungssimulationen (*Simulator in a loop*).

Als ein wesentlicher Beitrag dieser Arbeit werden die Aufgabenstellung des Nominalentwurfs und der Entwurfszentrierung auf eine gemeinsame mathematische Problemformulierung zurückgeführt. In beiden Aufgabenstellungen werden eigenschaftsorientierte Robustheitsmaße mittels der Exponentialfunktion in eine skalare Zielfunktion transformiert. Dabei kommt jeweils den Eigenschaften ein hohes Gewicht zu, die ihre Spezifikation verletzen bzw. die die Ausbeute am meisten einschränken. Die Zielgrößen für den Nominalentwurf und die Entwurfszentrierung berücksichtigen dabei jeweils die Charakteristiken der einzelnen Entwurfsschritte. Gegenüber dem Stand der Technik ergibt sich

dadurch eine signifikante Reduktion der benötigten Rechenzeit bei einer gleichzeitigen Vereinheitlichung der mathematischen Beschreibung.

In der Entwurfszentrierung werden die Worst-Case Abstände [AGW94, Grä93, Wie94, AEG⁺00a] als Robustheitsmaße verwendet. Der Worst-Case Abstand ist ein Maß für die parametrische Ausbeute einer Schaltung bezüglich einer Spezifikation. Der Nominalentwurf basiert auf den Parameterabständen als Zielgrößen. Der Parameterabstand beschreibt den Abstand einer Eigenschaft von ihrer Spezifikationsgrenze unter Berücksichtigung der Empfindlichkeit der Eigenschaft und dem Einfluß der Betriebsumgebung. Durch die im Rahmen dieser Arbeit eingeführte frühe Berücksichtigung der Betriebsumgebung wird im Rahmen des Nominalentwurfs ein sehr guter Ausgangspunkt für die nachfolgende Entwurfszentrierung bestimmt und so im Vergleich zu anderen Verfahren die Gesamtzahl der Simulationen im gesamten Entwurfsablauf deutlich reduziert werden.

Das in der Regel schlecht konditionierte Optimierproblem der Dimensionierung wird mit einem Trust-Region-Verfahren gelöst. Als signifikante Verbesserung bestehender Verfahren wird in dieser Arbeit die starke Nichtlinearität, die durch die Exponentialfunktion verursacht wird, besonders berücksichtigt. In jedem Schritt des Dimensionierungsalgorithmus wird nicht die Zielfunktion selbst, sondern nur die Robustheitsmaße linearisiert. Aufbauend auf dieser Linearisierung wird die Menge der paretooptimalen Punkte bestimmt, für die keine gleichzeitige Reduktion von Parameterauslenkung und Zielfunktion möglich ist.

Diese Kurve, die verallgemeinerte Grenzkurve (Generalized boundary curve – GBC), wird benutzt, um einen Parametersatz mit einer möglichst großen Fehlerreduktion bei einer möglichst geringen Fehlerreduktion zu bestimmen. Im Vergleich zu bestehenden Verfahren wird in dieser Arbeit der Berechnungsaufwand für die Grenzkurve durch eine günstige Approximation mit Bernstein-Bézier Kurven deutlich reduziert. Die verallgemeinerte Grenzkurve ist dabei nicht auf die vorgestellte Problemstellung beschränkt sondern vielmehr auf alle konvexen Probleme anwendbar.

Ein weiterer Fortschritt der vorliegenden Arbeit gegenüber dem Stand der Technik ist die erstmalige konsequente Berücksichtigung struktureller Nebenbedingungen [Eck98, ZEG98, SEGA99a, GZEA01] während der gesamten Dimensionierung. Strukturelle Nebenbedingungen schränken den Parameterraum für die Dimensionierung auf ein schaltungstechnisch sinnvolles Gebiet ein, in dem sich robuste Schaltungen hoher Qualität befinden. Strukturelle Nebenbedingungen beschreiben Grundregeln des analogen Schaltungsentwurfs als einfache Bedingungen an Transistoren und Transistorpaare. In den Ergebnissen (Kapitel 4) wird gezeigt, daß erst diese Einschränkung des Parameter-raumes eine sinnvolle automatische Dimensionierung analoger Schaltungen erlaubt.

Die vorgestellten Verfahren sind in dem industriell eingesetzten Entwurfswerkzeug WiCkeD [AEG⁺00a] implementiert. WiCkeD bietet dabei eine leistungsfähige graphische Benutzeroberfläche basierend auf TCL/TK [Ous94, Wel97] zur interaktiven und automatischen Dimensionierung analoger Schaltungen. Daneben können mittels einer TCL/TK-basierten Skriptsprache eigene Oberflächen und Algorithmen in das Tool integriert werden. Für WiCkeD sind Anbindungen an die Simulatoren SABER, SPECTRE, TITAN und SPICE realisiert. An zwei Schaltungsbeispielen, einem Operationsverstärker als klassische analoge Schaltung und einem Ausgangstreiber als digitale Schaltung, wurde die Anwendbarkeit der vorgestellten Verfahren zum Nominalentwurf und zur Entwurfszentrierung gezeigt.

Anhang A

Optimierprobleme

In diesem Abschnitt sind Optimalitätskriterien für differenzierbare Optimierprobleme mit Nebenbedingungen zusammengefaßt. Ein Optimierproblem besteht aus

$$\begin{aligned} \text{Zielfunktion:} & \quad f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{R}^{n_x} \\ \text{Gleichheitsnebenbedingungen:} & \quad \mathbf{g}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{g} \in \mathcal{R}^{n_x \times n_g} \\ \text{Ungleichheitsnebenbedingungen:} & \quad \mathbf{u}(\mathbf{x}) \geq \mathbf{0}, \quad \mathbf{u} \in \mathcal{R}^{n_x \times n_u} \end{aligned} \tag{A.1}$$

Für alle Funktionen wird angenommen, daß sie wenigstens zweimal stetig differenzierbar sind. Zunächst werden notwendige Bedingungen für allgemeine differenzierbare Optimierprobleme diskutiert. Anschliessend wird auf den Spezialfall konvexer Optimierprobleme eingegangen. Beweise und weitere Erklärungen zu den gegebenen Definitionen finden sich z. B. in [Spe98, NW99, GMW81].

A.1 Extremalkriterien für differenzierbare Optimierprobleme mit Nebenbedingungen

Für den allgemeinen Fall der Minimierung einer skalaren Zielfunktion $f(\mathbf{x})$ unter Berücksichtigung von Nebenbedingungen $\mathbf{g}(\mathbf{x})$ und $\mathbf{u}(\mathbf{x})$ können außer in Spezialfällen [HT90] keine Kriterien für eine globale Minimalstelle angegeben werden. Kriterien für lokale Extrema können basierend auf der *Lagrange-Funktion* angegeben werden.

Die Lagrange-Funktion für das Optimierproblem

$$\min_{\mathbf{x}} \{f(\mathbf{x}) \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \wedge \mathbf{u}(\mathbf{x}) \geq \mathbf{0}\} \tag{A.2}$$

mit *Lagrange-Multiplikatoren* $\boldsymbol{\lambda}$ und $\boldsymbol{\mu}$ lautet [Spe98, NW99, GMW81]:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{u}(\mathbf{x}), \quad \boldsymbol{\mu} \geq \mathbf{0} \tag{A.3}$$

Notwendige aber im allgemeinen nicht hinreichende Kriterien für ein Minimum von $f(\mathbf{x})$ auf dem zulässigen Bereich $\mathcal{C} = \{\mathbf{x} \in \mathcal{R}^{n_x} \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \wedge \mathbf{u}(\mathbf{x}) \geq \mathbf{0}\}$ sind:

$$\begin{aligned} \text{Karush-Kuhn-Tucker-Bedingung:} \quad & \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \mathbf{0} \\ & \text{(KKT-Bedingung)} \end{aligned} \quad (\text{A.4})$$

$$\text{Primale Zulässigkeit:} \quad \mathbf{x}^* \in \mathcal{C} \quad (\text{A.5})$$

$$\text{Duale Zulässigkeit:} \quad \boldsymbol{\lambda}^* \geq \mathbf{0} \quad (\text{A.6})$$

$$\text{Komplementarität:} \quad (\boldsymbol{\lambda}^*)^T \mathbf{u}(\mathbf{x}^*) = \mathbf{0} \quad (\text{A.7})$$

$$\text{Positive Definitheit:} \quad \forall_{\mathbf{z} \in \mathcal{Z}(\mathbf{x}^*)} \mathbf{z}^T \nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \mathbf{z} \geq \alpha \mathbf{z}^T \mathbf{z}, \quad \alpha \geq 0 \quad (\text{A.8})$$

Die Menge $\mathcal{Z}(\mathbf{x}^*)$ ist der linearisierende Kegel zu $\mathbf{g}(\mathbf{x}^*), \mathbf{u}(\mathbf{x}^*) \geq \mathbf{0}, \quad \mathbf{x}^* \in \mathcal{C}$:

$$\mathcal{Z}(\mathbf{x}^*) = \{\mathbf{z} \mid \mathbf{z}^T \nabla_{\mathbf{x}} g_i = 0 \wedge \mathbf{z}^T \nabla_{\mathbf{x}} u_j = 0, \quad i = 1, \dots, n_g, \quad j \in \mathcal{A}(\mathbf{x}^*)\} \quad (\text{A.9})$$

$$\mathcal{A}(\mathbf{x}^*) = \{j \in \{1, \dots, n_u\} \mid u_j(\mathbf{x}^*) = 0\} \quad (\text{A.10})$$

Die Menge $\mathcal{A}(\mathbf{x}^*)$ ist die Menge der aktiven Nebenbedingungen in \mathbf{x}^* . Hinreichende Bedingungen für ein Extremum erfordern zusätzliche Qualifikationsbedingungen an die Nebenbedingungen. Diese sind ausführlich in [Spe98, NW99, GMW81] dargestellt.

Eine Sensitivitätsanalyse des Optimierproblems [Spe98] ergibt zusätzlich, daß die Lagrange-Faktoren $\boldsymbol{\lambda}^*$ und $\boldsymbol{\mu}^*$ im Lösungspunkt die Empfindlichkeit der Zielfunktion gegenüber den Nebenbedingungen beschreiben:

$$\nabla_{\mathbf{g}, \mathbf{u}} f(\mathbf{x}^*) = \begin{bmatrix} \boldsymbol{\lambda}^* \\ \boldsymbol{\mu}^* \end{bmatrix} \quad (\text{A.11})$$

Dies bedeutet, daß sich in linearer Näherung beispielsweise eine Verbesserung der Zielfunktion um λ_i erreichen läßt, wenn die Gleichheitsnebenbedingung g_i nicht bei $g_i(\mathbf{x}) = 0$ sondern bei $g_i(\mathbf{x}) = -1$ erfüllt wäre. Aus der Komplementarität Gl. (A.7) ergibt sich wiederum, daß die Empfindlichkeit auf eine Ungleichheitsnebenbedingungen gleich Null ist, wenn diese Nebenbedingung nicht aktiv ist.

A.2 Konvexe Optimierprobleme

Konvexe Gebiete und Funktionen spielen in der Optimierung eine wichtige Rolle. Diese herausragende Stellung ergibt sich aus der Tatsache, daß für ein konvexes Optimierproblem jedes lokale Minimum auch ein globales Minimum ist. Für ein konvexes Optimierproblem kann mit deterministischen Verfahren also nachweislich das globale Minimum gefunden werden [SW70, Spe98, NW99]. In diesem Abschnitt sind deshalb einige wichtige Definitionen und Sätze über konvexe Mengen, Funktionen und Optimierprobleme zusammengefaßt. Eine umfassende Darstellung der Eigenschaften konvexer Funktionen und Gebiete findet sich z. B. in [Roc70].

A.2.1 Konvexe Mengen und Gebiete

Definition A.1 $\mathcal{C} \subset \mathcal{R}^n$ heißt konvex, wenn gilt:

$$\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C} \text{ und } \lambda \in [0, 1] \implies \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in \mathcal{C}.$$

Definition A.1 fordert für eine konvexe Menge \mathcal{C} , daß die Verbindungslinie zwischen zwei beliebigen Punkten aus der Menge vollständig in der Menge \mathcal{C} verläuft. Jede konvexe Menge ist damit automatisch auch ein Gebiet.

Definition A.2 Es sei $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathcal{R}^n$ und $\sum_{i=1}^m w_i = 1$. Dann heißt

$$\mathbf{y} = \sum_{i=1}^m w_i \mathbf{x}_i \tag{A.12}$$

eine Konvexkombination der \mathbf{x}_i . Die Konvexkombination heißt echt, falls $0 < w_i < 1$ für mindestens ein i gilt.

Definition A.3 Sei $\mathcal{C} \subset \mathcal{R}^n$ konvex. Wenn $\mathbf{x} \in \mathcal{C}$ nicht als echte Konvexkombination von Punkten aus \mathcal{C} darstellbar ist, heißt \mathbf{x} Extrempunkt von \mathcal{C} .

Wenn die Menge $\mathcal{C} \subset \mathcal{R}^n$ konvex und kompakt ist, dann gibt es genau $n + 1$ Extrempunkte von \mathcal{C} , so daß $\mathbf{x} \in \mathcal{C}$ eine Konvexkombination dieser Punkte ist [Roc70, Spe98]. Für die BB-Kurve mit drei Kontrollpunkten, die zur Approximation der Grenzkurve verwendet wird (Abschnitt 3.2.7), sind beispielsweise genau diese drei Kontrollpunkte Extrempunkte der konvexen Menge, die durch das Approximationsdreieck begrenzt wird.

A.2.2 Konvexe Funktionen

Eine konvexe Funktion wiederum definiert sich formal über die Eigenschaften konvexer Gebiete:

Definition A.4 Eine Funktion $f : \mathcal{C} \subset \mathcal{R}^n \mapsto \mathcal{R}$ ist konvex, wenn der Epigraph $\text{epi}(f)$

$$\text{epi}(f) = \{(\mathbf{x} \in \mathcal{C}, \alpha) \mid \alpha \geq f(\mathbf{x})\} \supset \mathcal{R}^{n+1} \tag{A.13}$$

konvex ist.

Definition A.4 erfordert, daß der zulässige Parameterraum \mathcal{C} ebenfalls konvex ist. Diese Definition ist allerdings im praktischen Gebrauch eher unhandlich. Deswegen werden hinreichende Kriterien für konvexe Funktionen gesucht.

Satz A.1 Sei $\mathcal{C} \subset \mathcal{R}^n$ nicht leer und konvex und die Funktion $f(\mathbf{x})$ in \mathcal{C} zweimal stetig partiell differenzierbar nach allen Variablen. Die Funktion f ist konvex auf \mathcal{C} genau dann, wenn $\nabla_{\mathbf{xx}}^2 f$ positiv semidefinit auf \mathcal{C} ist.

Im ein-dimensionalen reduziert sich dieser Satz auf die einfach nachprüfbare Bedingung $\frac{d^2 f}{dx^2} \geq 0$. Eine anderes hinreichendes und notwendiges Kriterium für die Konvexität einer differenzierbaren Funktion enthält folgender Satz:

Satz A.2 Sei $\mathcal{C} \subset \mathcal{R}^n$ nicht leer und konvex, f differenzierbar auf $\mathcal{D}_0 \supset \mathcal{C}$. Die Funktion f ist konvex genau dann, wenn für alle $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ gilt:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T \nabla f(\mathbf{x}) \tag{A.14}$$

A.2.3 Konvexe Optimierprobleme

Im Fall einer Minimierung bedeutet ein konvexes Optimierproblem die Minimierung einer konvexen Funktion auf einem konvexen Gebiet. Ein konvexes Optimierproblem hat, wie oben bereits angedeutet, die herausragende Eigenschaft, daß Kriterien zur Bestimmung eines globalen Minimums angegeben werden können. Dazu soll zunächst ein konvexes Optimierproblem formal definiert werden.

Definition A.5 Ein Optimierproblem heißt konvex, wenn der zulässige Bereich \mathcal{C}

$$\mathcal{C} = \{\mathbf{x} \in \mathcal{R}^{n_x} \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \wedge \mathbf{u}(\mathbf{x}) \geq \mathbf{0}\} \quad (\text{A.15})$$

konvex ist und die Zielfunktion $f(\mathbf{x})$ konvex auf \mathcal{C} ist.

Für den in der vorliegenden Arbeit betrachteten Fall der Minimierung einer konvexen Funktion unter Berücksichtigung von Ungleichheitsnebenbedingungen $\mathbf{u}(\mathbf{x})$ und ohne Gleichheitsnebenbedingungen $\mathbf{g}(\mathbf{x})$ sind unter einer milden Zusatzbedingung die in Abschnitt A.1 angegebenen notwendigen Kriterien bereits hinreichend für ein globales Minimum:

Satz A.3 Sei \mathcal{C} konvex, nicht leer und abgeschlossen und die Gleichheitsnebenbedingungen $g_i(\mathbf{x})$ affin linear; die Zielfunktion $f(\mathbf{x})$ sei konvex und differenzierbar auf \mathcal{C} . Die Slater-Bedingung sei erfüllt:

$$\exists_{\mathbf{x}^0 \in \mathcal{C}} \quad \mathbf{u}(\mathbf{x}^0) > \mathbf{0} \wedge \mathbf{g}(\mathbf{x}^0) = \mathbf{0} \quad (\text{A.16})$$

$\mathbf{x}_* \in \mathcal{C}$ ist genau dann globale Lösung des Minimierproblems (A.1) wenn die Multiplikatorregel (KKT-Bedingung) (A.4) erfüllt ist.

Sind keine Gleichheitsnebenbedingungen $\mathbf{g}(\mathbf{x})$ vorhanden, so gilt obiger Satz ebenfalls. Die zusätzlich geforderte Restriktionsqualifikation (Slater-Bedingung) erfordert in diesem Fall, daß das durch die Nebenbedingungen beschriebene konvexe Gebiet nicht nur nicht leer ist, sondern auch ein Inneres hat. In der Praxis ist dieser Fall für die behandelte Problemstellung immer erfüllt.

Anhang B

Bernstein-Bézier Kurven

Bernstein-Bézier Kurven (BB-Kurven) [MV90, BSMM95] sind eine unter anderem in der Grafik sehr beliebte Methode, Gebiete und Kurven zu approximieren. Diese Beliebtheit liegt zum einen an der schnellen und leichten Berechenbarkeit und andererseits an der guten Kontrollierbarkeit der Approximation.

B.1 Allgemeine Definition einer Bernstein-Bézier Kurve

Eine BB-Kurve ist eine parametrische Kurve in Abhängigkeit des Parameters $t \in [0, \dots, 1]$ und der Kontrollpunkte \mathbf{p}_i :

$$\mathbf{p}(t) = \sum_{i=0}^n B_{i,n}(t) \mathbf{p}_i, \quad 0 \leq t \leq 1, \quad B_{i,n} = \binom{n}{i} t^i (1-t)^{n-i} \quad (\text{B.1})$$

An jedem Punkt $t \in [0, \dots, 1]$ gilt dabei für die Summe der Bernstein-Polynome $B_{i,n}$:

$$\sum_{i=0}^n B_{i,n}(t) = 1 \quad (\text{B.2})$$

Mit dieser Eigenschaft der Bernstein-Polynome stellt die BB-Kurve eine Konvexkombination der Kontrollpunkte \mathbf{p}_i nach Definition A.2 dar.

B.2 Eigenschaften der BB-Kurve mit drei Kontrollpunkten

In diesem Abschnitt werden speziell die Eigenschaften der in dieser Arbeit zur Approximation verwendeten BB-Kurve mit drei Kontrollpunkten diskutiert. Zunächst sollen die Eigenschaften ganz allgemeiner BB-Kurven mit drei Kontrollpunkten betrachtet werden, um dann auf den bei der Approximation verwendeten Spezialfall einzugehen. Im folgenden wird angenommen, daß die drei Kontrollpunkte der BB-Kurve nicht auf einer Gerade liegen.

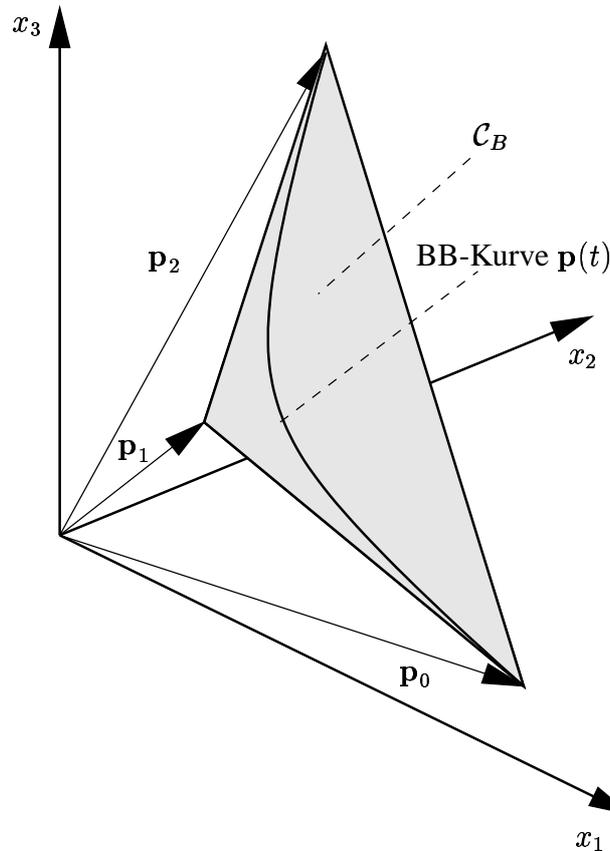


Abbildung B.1: BB-Kurve mit drei Kontrollpunkten im drei-dimensionalen Raum

Der Fall, daß die drei Kontrollpunkte der BB-Kurve auf einer Geraden liegen, entspricht in dem hier untersuchten Anwendungsfall der Approximation einer Grenzkurve mit idealer Konditionszahl und einer sich nicht ändernden Menge aktiver Nebenbedingungen. In diesem Fall ist sowohl die Grenzkurve als auch die BB-Kurve eine einfache Gerade.

Für eine BB-Kurve mit drei Kontrollpunkten p_0 , p_1 und p_2 stellen die Kontrollpunkte Extrempunkte des konvexen Gebietes C_B dar, das durch die Verbindungslinien der Kontrollpunkte gebildet wird (Abbildung B.1). Die drei Kontrollpunkte spannen eine zweidimensionale Ebene in dem im allgemeinen n -dimensionalen Raum auf. Legt man ohne Beschränkung der Allgemeinheit durch eine Verschiebung der Koordinaten einen der drei Kontrollpunkte in den Ursprung, so bilden die Vektoren von diesem Punkt zu den beiden anderen Kontrollpunkten eine (im allgemeinen nicht orthogonale) Basis für den zwei-dimensionalen Unterraum in dem sich die Kurve bewegt. Es reicht also aus, für die BB-Kurve mit drei Kontrollpunkten einen zwei-dimensionalen Raum zu betrachten.

Für den oben ausgeschlossenen entarteten Fall, daß alle drei Kontrollpunkte auf einer Geraden liegen, sind die Kontrollpunkte nicht linear unabhängig und nach der Koordinatenverschiebung eines Kontrollpunktes in den Ursprung spannen die übrigen zwei Kontrollpunkte keinen zwei-dimensionalen sondern nur einen ein-dimensionalen Unterraum auf. Dies ist, wie oben bereits beschrieben, gleichbedeutend mit einer Gerade.

Definition der Kurve

Die drei Kontrollpunkte \mathbf{p}_i der BB-Kurve seien folgendermaßen gewählt:

$$\mathbf{p}_0 = \begin{bmatrix} a_1 \\ r_1 \end{bmatrix}, \quad \mathbf{p}_1 = \begin{bmatrix} a_t \\ r_t \end{bmatrix} \quad \text{und} \quad \mathbf{p}_2 = \begin{bmatrix} a_2 \\ r_2 \end{bmatrix} \quad (\text{B.3})$$

Die BB-Kurve $[a_B(t), r_B(t)]^T$ mit drei Kontrollpunkten, die im weiteren verwendet werden soll, lautet damit:

$$\begin{bmatrix} a_B(t) \\ r_B(t) \end{bmatrix} = (1-t)^2 \begin{bmatrix} a_1 \\ r_1 \end{bmatrix} + 2t(1-t) \begin{bmatrix} a_t \\ r_t \end{bmatrix} + t^2 \begin{bmatrix} a_2 \\ r_2 \end{bmatrix}, \quad 0 \leq t \leq 1 \quad (\text{B.4})$$

Steigung und Krümmung der BB-Kurve

Die Steigung und insbesondere auch die Krümmung der BB-Kurve ist für die Approximation der Grenzkurve von Interesse. Zur Vereinfachung der Schreibweise soll im weiteren die Ableitung nach dem Parameter t durch $\dot{}$ (z.B. $\dot{a}_B = \partial a_B / \partial t$) abgekürzt werden. Für die Steigung $m_B(t)$ und Krümmung $\kappa_B(t)$ der verwendeten BB-Kurve gilt damit:

$$m_B(t) = \frac{\partial r_B(t)}{\partial a_B(t)} = \frac{\dot{r}_B}{\dot{a}_B} = \frac{-r_1(1-t) + r_t(1-2t) + r_2 t}{-a_1(1-t) + a_t(1-2t) + a_2 t} \quad (\text{B.5})$$

$$\begin{aligned} \kappa_B(t) &= \frac{\dot{a}_B \ddot{r}_B - \ddot{a}_B \dot{r}_B}{\sqrt{(\dot{a}_B^2 + \dot{r}_B^2)^3}} = \\ &= \frac{2(a_1(r_t - r_2) + a_t(r_2 - r_1) + a_2(r_1 - a_t))}{\sqrt{(a_1(t-1) + a_t(1-2t) + a_2 t)^2 + (r_1(t-1) + r_t(1-2t) + r_2 t)^2}} \end{aligned} \quad (\text{B.6})$$

Anwendung auf die Approximation der Grenzkurve

Bei der Approximation der Grenzkurve zwischen zwei berechneten Punkten $[a_1, r_1]^T$ und $[a_2, r_2]^T$ bestimmt sich der mittlere Kontrollpunkt $[a_t, r_t]^T$ nach Gl. (3.28) bzw. (3.29) zu:

$$a_t = \frac{r_2 - r_1 + a_1 m_1 - a_2 m_2}{m_1 - m_2} \quad (\text{B.7})$$

$$r_t = r_1 + m_1(a_t - a_1) \quad (\text{B.8})$$

Werden die Endpunkte $[a_1, r_1]^T$ und $[a_2, r_2]^T$ so gewählt, daß $a_1 < a_2$, dann gilt für die Steigungen m_1 und m_2 aufgrund der Konvexität der Grenzkurve immer $m_1 \leq m_2 \leq 0$. Setzt man die Gleichungen für den Tangenschnittpunkt die BB-Kurve ein, so ergibt sich für diesen Fall die Steigung $m_B(t)$ zu:

$$m_B(t) = \frac{m_1 m_2 (a_1 - a_2) (1 - 2t) + (r_1 - r_2) (m_1 (t - 1) + m_2 t)}{(r_1 - r_2) (2t - 1) + (a_1 - a_2) (m_1 t + m_2 (1 - t))} \quad (\text{B.9})$$

Eigenschaften der approximierenden Bernstein-Bézier Kurve

Die BB-Kurve, die die Grenzkurve zwischen zwei berechneten Punkten approximiert, erhält alle wichtigen Eigenschaften der Grenzkurve:

- Für $t = 0$ gilt $a_B(t) = a_1$, $r_B(t) = r_1$ und $m_B(t) = m_1$. Der Beweis erfolgt durch Einsetzen von $t = 0$ in Gl. (B.4) bzw. Gl. (B.9).
- Für $t = 1$ gilt $a_B(t) = a_2$, $r_B(t) = r_2$ und $m_B(t) = m_2$. Der Beweis erfolgt durch Einsetzen von $t = 1$ in Gl. (B.4) bzw. Gl. (B.9).
- Die BB-Kurve verläuft vollständig innerhalb des Dreiecks, daß die Tangenten in den beiden Endpunkten und deren Verbindungslinie bilden. Wie oben diskutiert, ist eine BB-Kurve eine Konvexkombination der einzelnen Kontrollpunkte \mathbf{p}_i . Im zweidimensionalen Fall gibt es genau drei Extrempunkte (siehe Definition A.3) die eine konvexe Menge aufspannen. In diesem Fall sind die drei Kontrollpunkte die drei Extrempunkte. Die konvexe Menge ist das durch die Extrempunkte aufgespannte Dreieck. Mit einer Konvexkombination der Extrempunkte kann das Dreieck nicht verlassen werden.
- Die BB-Kurve ist für $t \in [0, \dots, 1]$ konvex. Dazu ist zu zeigen, daß $\kappa_B(t) \geq 0$ für $t \in [0, \dots, 1]$ gilt. Setzt man die Definition für den Tangenschnittpunkt in die Krümmung Gl. (B.6) ein, so ergibt sich der Zähler des Bruches zu:

$$\dot{a}_B \ddot{r}_B - \ddot{a}_B \dot{r}_B = 4 \cdot \frac{(m_1(a_2 - a_1) - (r_2 - r_1)) \cdot (m_2(a_2 - a_1) - (r_2 - r_1))}{m_1 - m_2} \quad (\text{B.10})$$

Die einzelnen Anteil in dieser Gleichung haben dabei folgende Bedeutung:

- $(m_1 - m_2)$ ist die Änderung der Steigung zwischen den beiden Punkten. Hierbei gilt:

$$m_1 < m_2 \leq 0 \Rightarrow m_1 - m_2 < 0 \quad (\text{B.11})$$

Der Fall $m_1 = m_2$ wurde oben explizit ausgeschlossen und wird deshalb hier nicht behandelt. Für diesen Fall ist die BB-Kurve eine Gerade und ist somit konvex.

- $(r_2 - r_1)$ ist die Änderung des Schaltungsfehlers zwischen den beiden Punkten.
- $(m_1(a_2 - a_1))$ und $(m_2(a_2 - a_1))$ ist die Abschätzung der Änderung des Schaltungsfehlers basierend auf der Linearisierung in den Punkten $[a_1, r_1]^T$ bzw. $[a_2, r_2]^T$. Aufgrund der Konvexität der Grenzkurve gilt für diese Abschätzung immer:

$$m_1(a_2 - a_1) < r_2 - r_1 < m_2(a_2 - a_1) \leq 0 \quad (\text{B.12})$$

$$\Rightarrow (m_1(a_2 - a_1) - (r_2 - r_1)) \cdot (m_2(a_2 - a_1) - (r_2 - r_1)) < 0 \quad (\text{B.13})$$

Aus obigen Betrachtungen Gl. (B.11) und Gl. (B.13) folgt zusammen mit der Definition der Krümmung Gl. (B.6), daß die Krümmung der BB-Kurve immer positiv ist. Somit ist die BB-Kurve konvex.

Literaturverzeichnis

- [AEG⁺00a] K. Antreich, J. Eckmueller, H. Graeb, M. Pronath, F. Schenkel, R. Schwencker, und S. Zizala. WiCkeD: Analog Circuit Synthesis Incorporating Mismatch. In *IEEE Custom Integrated Circuits Conference (CICC)*, 2000.
- [AEG⁺00b] K. Antreich, J. Eckmüller, H. Gräb, M. Pronath, F. Schenkel, R. Schwencker, und S. Zizala. WiCkeD: Analogsynthese unter Berücksichtigung von Mismatch. In *ITG Workshop Mikroelektronik für die Informationstechnik*, November 2000.
- [AGW94] K. Antreich, H. Gräb, und C. Wieser. Circuit analysis and optimization driven by worst-case distances. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 13(1):57–71, Januar 1994.
- [AH84] K. Antreich und S. Huss. An Interactive Optimization Technique for the Nominal Design of Integrated Circuits. *IEEE Transactions on Circuits and Systems CAS*, 31:203–212, 1984.
- [AK82] K. Antreich und R. Koblitz. Design Centering by Yield Prediction. *IEEE Transactions on Circuits and Systems CAS*, 29:88–95, 1982.
- [ALP88] K. Antreich, P. Leibner, und F. Pörnbacher. Nominal design of integrated circuits on circuit level by an interactive improvement method. *IEEE Transactions on Circuits and Systems CAS*, 35:1501–1511, 1988.
- [AMH91] H. Abdel-Malek und A. Hassan. The ellipsoidal technique for design centering and region approximation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 10:1006–1013, 1991.
- [AMHH99] Hany L. Abdel-Malek, Abdel-Karim S. O. Hassan, und Mohamed H. Heaba. A Boundary Gradient Search Techniques and Its Application in Design Centering. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 18(11):1654–1661, November 1999.
- [BK00] M. Bennis und J. Kurpi. Trends:Das neue Zuhause. *c't magazin für computer technik*, (4):170–175, 2000.
- [BSMM95] I. N. Bronstein, K. A. Semendjajew, Musiol, und Mühlig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Frankfurt/Main, 1995.
- [Bor97] C. Borchers. *Automatische Generierung von Verhaltensmodellen für nichtlineare Anlogschaltungen*. Doktorarbeit, Universität Hannover, 1997.

- [Bor98] C. Borchers. Symbolic Behavioral Model Generation of Nonlinear Analog Circuits. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(10):1362, 1998.
- [CCH⁺98] Andrew R. Conn, Paula K. Coulman, Ruud A. Haring, Gregory L. Morill, Chandu Visweswariah, und Chai Wah Wu. JiffyTune: Circuit Optimization using Time-Domain Sensitivities. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 17(12):1292–1309, Dezember 1998.
- [DEAM97] J. E. Dennis, M. El-Alem, und M. C. Maciel. A global convergence theory for general trust-region-based algorithms for equality constrained optimization. *SIAM Journal on Optimization*, 7(1):177–207, Februar 1997.
- [DG98] Geert Debyser und Georges Gielen. Efficient Analog Circuit Synthesis with simultaneous Yield and Robustness Optimization. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 1998.
- [DH77] S. Director und G. Hachtel. The simplicial approximation approach to design centering. *IEEE Transactions on Circuits and Systems CAS*, 24:363–372, 1977.
- [DK95] A. Dharchoudhury und S. M. Kang. Worst-Case Analysis and Optimization of VLSI Circuit Performances. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(4):481–492, April 1995.
- [DV97] J. E. Dennis und L. N. Vicente. On the convergence theory of trust-region-based algorithms for equality-constrained optimization. *SIAM Journal on Optimization*, 7(4):927–950, November 1997.
- [Dan63] C.B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1963.
- [De87] M. Degrauwe und et al. IDAC: An interactive design tool for analog CMOS circuits. *IEEE Journal of Solid-State Circuits SC*, 22:1106–1116, 1987.
- [Dog81] K. Doganis. OPSIM: Circuit optimization capabilities within the circuit simulator. Technical report, Optimum Solutions Inc., 1981.
- [EGG98] Josef Eckmüller, Martin Gröpl, und Helmut Gräb. Hierarchical Characterization of Analog Integrated CMOS Circuits. In *Design, Automation and Test in Europe (DATE)*, Seite 636–643, Paris, France, Februar 1998.
- [Eck98] J. Eckmüller. *Zur rechnergestützten Dimensionierung analoger integrierter Schaltungen unter besonderer Berücksichtigung von Struktureigenschaften*. Doktorarbeit, Technical University of Munich, 1998.
- [Esh92] K. Eshbaugh. Generation of correlated parameters for statistical circuit simulation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 11:1198–1206, 1992.
- [FD90] P. Feldmann und S. Director. Accurate and efficient evaluation of circuit yield and yield gradients. *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Seite 120–123, 1990.

- [FD91] P. Feldmann und S. Director. Improved methods for IC yield and quality optimization using surface integrals. *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Seite 158–161, 1991.
- [FD93] P. Feldmann und S. Director. Integrated Circuit Quality Optimization using Surface Integrals. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 12:1868–1879, 1993.
- [FWZ⁺92] U. Feldmann, U. Wever, Q. Zheng, R. Schultz, und H. Wriedt. Algorithms for modern circuit simulation. *Archiv für Elektronik und Übertragungstechnik (AEÜ)*, 46:274–285, 1992.
- [Fis99] George S. Fishman. *Monte Carlo Concepts, Algorithms, and Applications*. Springer New-York, 1999.
- [GK83] D.D. Gajski und R.H. Kuhn. Guest Editor’s introduction: New VLSI tools. *IEEE Computer*, 16:11–14, 1983.
- [GM78] Ph.E. Gill und W. Murray. Numerically stable methods for quadratic programming. *Mathematical Programming*, 14:349–372, 1978.
- [GMW81] Philip E. Gill, Walter Murray, und Margaret H. Wright. *Practical Optimization*. Academic Press. Inc., London, 1981.
- [GRRC00] G. Gielen, R. Rutenbar, J. Roychowhury, und F. Clement. CAD for Mixed Analogue-Digital and RF ICs: The Current State of the Art. In *ESSCIRC*, Stockholm, September 2000.
- [GS91] G. Gielen und W. Sansen. *Symbolic Analysis for Automated Design of Analog Integrated Circuits*. Kluwer Academic Publishers, Dordrecht, 1991.
- [GWS89] G. Gielen, H. C. Walscharts, und W. C. Sansen. ISAAC: A symbolic simulation for analog integrated circuits. *IEEE Journal of Solid-State Circuits SC*, 24:1587–1597, Dezember 1989.
- [GZEA01] H. Graeb, S. Zizala, J. Eckmueller, und K. Antreich. The Sizing Rules Method for Analog Integrated Circuit Design. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2001.
- [Grä93] H. Gräb. *Schaltkreisoptimierung mit Worst-Case-Abständen als Zielgrößen*. Doktorarbeit, Technical University of Munich, 1993.
- [GvL96] G. H. Golub und C. F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3 Edition, 1996.
- [HCY88] D. Hocevar, P. Cox, und P. Yang. Parametric yield optimization for MOS circuit blocks. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 7:645–658, 1988.
- [HEL92] J. Harvey, M. Elmasry, und B. Leung. STAIC: An interactive framework for synthesizing CMOS and BiCMOS analog circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 11:1402–1417, 1992.

- [HH75] J. M. Hammersley und D. C. Handscomb. *Monte Carlo Methods*. Fletcher, Norwich, 1975.
- [HRC89] R. Harjani, R. Rutenbar, und L. Carley. OASYS: A framework for analog circuit synthesis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 8:1247–1266, 1989.
- [HS95] E. Hennig und R. Sommer. Application of computer algebra methods to analog circuit sizing. In *European Conference on Circuit Theory and Design (ECCTD)*, 1995.
- [HT90] R. Horst und H. Tuy. *Global Optimization*. Springer, 1990.
- [HZ81] G. D. Hachtel und P. Zug. APLSTAP – Circuit Design and Optimization System – User’s Guide. Technical report, IBM Yorktown Research Facility, Yorktown, New York, 1981.
- [Hus82] S. Huss. *Zur interaktiven Optimierung integrierter Schaltungen*. Doktorarbeit, Technical University of Munich, 1982.
- [JW92] R. A. Johnson und D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, Englewood Cliffs, New Jersey, 3 Edition, 1992.
- [KCGV83] S. Kirkpatrick, Jr. C.D. Gelatt, und M.P. Vecchi. Optimization by Simulated Annealing. *Science*, 220:671–680, 1983.
- [KD95] K. Krishna und S. W. Director. The Linearized Performance Penalty (LPP) Method for Optimization of Parametric Yield and Its Reliability. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(12):1557–1568, Dezember 1995.
- [KK98] M. Keramat und R. Kielbasa. Generalized Centers of Gravity Algorithm for Yield Optimization of Integrated Circuits. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, Seite VI–334–337, 1998.
- [KPRC99] M. Krasnicki, R. Phelps, R. A. Rutenbar, und L. R. Carley. MAELSTROM: Efficient Simulation-Based Synthesis for Custom Analog Cells. In *ACM/IEEE Design Automation Conference (DAC)*, 1999.
- [KSG90] H. Koh, C. Sequin, und P. Gray. OPASYN: A compiler for CMOS operational amplifiers. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 9:113–125, 1990.
- [Keu99] Kurt Keutzer. Panel: Cell Libraries - Build vs. Buy; Static vs. Dynamic. In *ACM/IEEE Design Automation Conference (DAC)*, 1999.
- [Kob82] R. Koblitz. *Ein Verfahren zur Entwurfszentrierung elektrischer Schaltungen*. Doktorarbeit, Technical University of Munich, 1982.
- [LD89] K. Low und S. Director. A new methodology for the design centering of IC fabrication processes. *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Seite 194–197, 1989.
- [LDGS97] F. Leyn, W. Daems, G. Gielen, und W. Sansen. A Behavioral Signal Path Modeling Methodology for Qualitative Insight in and Efficient Sizing of CMOS Opamps. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 1997.

- [LS94] K. R. Laker und W. M. C. Sansen. *Design of Analog Integrated Circuits and Systems*. McGraw-Hill, 1994.
- [Lei86] P. Leibner. *Ein interaktives Optimierungsverfahren zur rechnergestützten Dimensionierung integrierter Schaltungen*. Doktorarbeit, Technical University of Munich, 1986.
- [Lev44] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, Seite 164–168, 1944.
- [MCR94] T. Mukherjee, L. Carley, und R. Rutenbar. Synthesis of Manufacturable Analog Circuits. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Seite 586–593, November 1994.
- [MFDCRV94] F. Medeiro, F. V. Fernandez, Dominguez-Castro, und A. Rodriguez-Vasquez. A statistical optimization based approach for automated sizing of analog cells. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 1994.
- [MG95] S.W. Mahfoud und D.D. Goldberg. Parallel Recombinative Simulated Annealing: A Genetic Algorithm. *Parallel Computing*, 21, 1995.
- [MV90] K. Meyberg und P. Vachenauer. *Höhere Mathematik I*. Springer-Lehrbuch, 1990.
- [Mar63] D.W. Marquardt. An algorithm for least squares estimation of non-linear parameters. *SIAM Journal*, Seite 431–441, 1963.
- [Mor77] J.J. More. The Levenberg-Marquardt algorithm: Implementation and theory. *Numerical Analysis*, 1977.
- [NRSVT88] W. Nye, D. Riley, A. Sangiovanni-Vincentelli, und A. Tits. DELIGHT.SPICE: An optimization-based system for the design of integrated circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 7:501–519, 1988.
- [NW99] Jorge Nocedal und Stephen J. Wright. *Numerical Optimization*. Springer, 1999.
- [Nag75] L. Nagel. *SPICE2: A computer program to simulate semiconductor circuits*. Ph. D. dissertation, Univ. of California, Berkeley, 1975.
- [Neu75] K. Neumann. *Operations Research Verfahren*. Carl Hanser Verlag, 1975.
- [Nod00] N. Noda. 21st Century Cars and ICs. In *International Solid-States Circuits Conference (ISSCC)*, volume 43, 2000.
- [OMRC98] E. Ochotta, T. Mukherjee, R.A. Rutenbar, und L.R. Carley. *Practical Synthesis of High-Performance Analog Circuits*. Kluwer Academic Publishers, 1998.
- [ORC94] E. Ochotta, R. Rutenbar, und L. Carley. ASTRX/OBLX: Tools for rapid synthesis of high-performance analog circuits. *ACM/IEEE Design Automation Conference (DAC)*, Seite 24–30, 1994.
- [ORC96] Emil S. Ochotta, Rob A. Rutenbar, und L. Richard Carley. Synthesis of High-Performance Analog Circuits in ASTRX/OBLX. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 15(3):273–294, März 1996.
- [Ous94] John K. Ousterhout. *Tcl and the Tk Toolkit*. Addison-Wesley, 1994.

- [PDW89] M. Pelgrom, A. Duinmaijer, und A. Welbers. Matching properties of MOS transistors. *IEEE Journal of Solid-State Circuits SC*, 24:1433–1440, 1989.
- [PMMDC⁺00] F. M. Perez-Montes, F. Medeiro, R. Dominguez-Castro, F. V. Fernandez, und A. Rodriguez-Vazque. XFRIDGE: A spice-based, portable, user-friendly cell-level sizing tool. In *Design, Automation and Test in Europe (DATE)*, 2000.
- [PR94] L. Pillage und R. Rohrer. The Essence of AWE. *IEEE Circuits and Devices Magazine*, 10(5):12–19, September 1994.
- [Pin00] M.R. Pinto. Atoms To Applets: Building Systems ICs in the 21st Century. In *International Solid-States Circuits Conference (ISSCC)*, volume 43, 2000.
- [Pör89] F. Pörnbacher. CANDI: Ein Programmsystem zur Unterstützung des Entwurfs integrierter Schaltungen. *ntzArchiv*, 11:137–151, 1989.
- [Pör89] F. Pörnbacher. *Zur interaktiven Dimensionierung analoger integrierter Schaltungen unter Berücksichtigung von Nebenbedingungen*. Doktorarbeit, Technical University of Munich, 1989.
- [RMRV98] R. Rodriguez-Macias und A. Rodriguez-Vazquez. A Technique for Fast AC Statistical Analysis of Analog Circuits. In *IEEE Int. Conf. on Electronics, Circuits and Systems*, Seite 81–84, 1998.
- [Roc70] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, 1970.
- [SA94] M. Styblinski und S. Aftab. IC variability minimization using a new Cp and Cpk based variability/performance measure. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, 1994.
- [SEGA99a] R. Schwencker, J. Eckmüller, H. Gräß, und K. Antreich. Automating the Sizing of Analog CMOS-Circuits by Consideration of Structural Constraints. In *Design, Automation and Test in Europe (DATE)*, Munich, März 1999.
- [SEGA99b] R. Schwencker, J. Eckmüller, H. Gräß, und K. Antreich. Automatische Nominalepunktdimensionierung analoger CMOS-Schaltungen mit Parameterabständen als Zielgrößen. In *GME/ITG-Diskussionssitzung Entwicklung von Analogschaltungen mit CAE-Methoden*, Munich, Februar 1999.
- [SGA99] Frank Schenkel, Helmut Gräß, und Kurt Antreich. Ausbeuteanalyse unter Berücksichtigung lokaler und globaler Parameterschwankungen. In *GME/ITG-Diskussionssitzung Entwicklung von Analogschaltungen mit CAE-Methoden*, 1999.
- [SGLS98] Robert Schwencker, Helmut Gräß, Thomas Latzel, und Christian Sporrer. Interface Circuits and Symbolic Analysis. In *Symbolic Methods and Applications in Circuit Design (SMACD)*, Seite 75–79, Oktober 1998.
- [SPV99] Abbas Seifi, K. Ponnambalam, und Jiri Vlach. A Unified Approach to Statistical Design Centering of Integrated Circuits with Correlated Parameters. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 46(1):190–196, Januar 1999.

- [SPZ⁺01] F. Schenkel, M. Pronath, S. Zizala, R. Schwencker, H. Graeb, und K. Antreich. Mismatch Analysis and Direct Yield Optimization by Spec-Wise Linearization and Feasibility-Guided Search. In *ACM/IEEE Design Automation Conference (DAC)*, Juni 2001.
- [SS80] R. Soin und R. Spence. Statistical exploration approach to design centering. *Proceedings of the IEEE*, 127:260–269, 1980.
- [SS93] K. Swings und W. Sansen. ARIADNE: A Constraint-Based Approach to Computer-Aided Synthesis and Modeling of Analog Integrated Circuits. *Analog Integrated Circuits and Signal Processing*, 3:197–215, 1993.
- [SSG90] K. Swings, W. Sansen, und G. Gielen. Donald: An intelligent analog IC design system based on manipulation of design equations. In *IEEE Custom Integrated Circuits Conference (CICC)*, Seite 8.6.1–4, 1990.
- [SSGA99] Robert Schwencker, Frank Schenkel, Helmut Gräß, und Kurt Antreich. Automatische Entwurfszentrierung analoger integrierter Komponenten basierend auf der Verallgemeinerten Grenzkurve von Mehrfach-Robustheits-Maßen. In *E.I.S.-Workshop*, 1999.
- [SSGA00] R. Schwencker, F. Schenkel, H. Gräß, und K. Antreich. The Generalized Boundary Curve – A Common Method for Automatic Nominal Design and Design Centering of Analog Circuits. In *Design, Automation and Test in Europe (DATE)*, Seite xx–xx, 2000.
- [SSLG99] C. Sporrer, R. Schwencker, T. Latzel, und H. Gräß. Generierung und Anwendungen von deklarativen Modellen für Treiberschaltungen. In *GME/ITG-Diskussionssitzung Entwicklung von Analogschaltungen mit CAE-Methoden*, 1999.
- [SVK93] S. Sapatnekar, P. Vaidya, und S. Kang. Convexity-based algorithms for design centering. *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Seite 206–209, 1993.
- [SW70] J. Stoer und Ch. Witzigall. *Convexity and Optimization in finite Dimensions I*. Springer, 1970.
- [Sch81] K. Schittkowski. The nonlinear programming method of Wilson, Han, and Powell with an augmented Lagrangian type line search function. *Numer. Math.*, 38:83–114, 1981.
- [Sch99] Robert Schwencker. Automatic Design Centering of Analog Integrated Circuits Based on the Generalized Boundary Curve of Multiple Robustness Objectives. Technical report, Institute of Electronic Design Automation TUM-LEA-99-1, Technical University Munich, 1999.
- [Sem99] Semiconductor Industry Association. *International Technology Roadmap for Semiconductors*, 1999.
- [Sor82] D. Sorenson. Trust region methods for unconstrained minimization. *Nonlinear Optimization*, 1982.
- [Spe98] Peter Spellucci. *Numerische Verfahren der nichtlinearen Optimierung*. Birkhäuser, 1998.

- [Sto71] J. Stoer. On the numerical solution of constrained least squares problems. *SINUM*, 8:382–411, 1971.
- [Sto72] J. Stoer. *Einführung in die Numerische Mathematik I*. Springer, Berlin, 1972.
- [VVS87] M. Vlach, J. Vlach, und K. Singhal. *SABER: A design tool for analog systems*. Analog Inc., Beaverton, U.S.A., 1987.
- [Wei00] G. Weinberger. The New Millennium: Wireless Technologies for a Truly Mobile Society. In *International Solid-States Circuits Conference (ISSCC)*, volume 43, 2000.
- [Wel97] Brent B. Welch. *Practical Programming in Tcl and Tk*. Prentice Hall PTR, 1997.
- [Wie94] Claudia Wieser. *Schaltkreisanalyse mit Worst-Case Abständen*. Doktorarbeit, Technical University of Munich, 1994.
- [ZEG98] Stephan Zizala, Josef Eckmüller, und Helmut Gräb. Fast Calculation of Analog Circuits' Feasibility Regions by Low Level Functional Measures. In *IEEE Int. Conf. on Electronics, Circuits and Systems*, Seite 85–88, September 1998.
- [ZEG99] Stephan Zizala, Josef Eckmüller, und Helmut Gräb. Effiziente Modellierung integrierter analoger CMOS-Schaltungen durch Berücksichtigung von Struktureigenschaften. In *GME/ITG-Diskussionssitzung Entwicklung von Analogschaltungen mit CAE-Methoden*, 1999.
- [ZS95] J. C. Zhang und M. A. Styblinski. *Yield and Variability Optimization of Integrated Circuits*. Kluwer Academic Publishers, 1995.
- [dMHBL98] Maria del Mar Hershenson, Stephen P. Boyd, und Thomas H. Lee. GPCAD: A Tool for CMOS Op-Amp Synthesis. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 1998.

Abbildungsverzeichnis

1.1	Abstraktionsebenen beim Entwurf analoger Systeme.	10
2.1	Akzeptanzgebiet der Eigenschaften und statistischen Parameter.	24
2.2	Veranschaulichung des Parameterabstandes.	29
2.3	Berechnung des Worst-Case-Abstand $\beta_{WC,i}$ und des linearisierten Akzeptanzgebietes $\bar{\mathcal{A}}_{s,i}$	33
3.1	Darstellung der Auslenkung und des Schaltungsfehlers für die charakteristische Grenzkurve in unnormierter und normierter Darstellung in Abhängigkeit von λ	39
3.2	Grenzkurve mit einer Nebenbedingungen und Veranschaulichung der Transformation orthogonal zu der aktiven Nebenbedingung.	40
3.3	Bestimmung des Approximationsfehlers der charakteristischen Grenzkurve.	44
3.4	Stückweise lineare Approximation der Grenzkurve.	45
3.5	Bestimmung des Approximationsfehlers für die Interpolation der Grenzkurve durch eine BB-Kurve.	49
3.6	Veranschaulichung der Wolfe Bedingung (Bedingung des hinreichenden Abstiegs) an der Grenzkurve	54
3.7	Künstliche Schrittweitenbeschränkung bei Linearisierung der Exponentialfunktion.	56
4.1	Beispielschaltung: Miller Operationsverstärker	66
4.2	Verstärkung und eine strukturelle Nebenbedingung in Abhängigkeit der Entwurfparameter.	69
4.3	Verlauf der Verstärkung über der Versorgungsspannung nach dem Nominalentwurf mit und ohne Berücksichtigung der strukturellen Nebenbedingungen.	69
4.4	Beispielschaltung Ausgangstreiber	70
4.5	Korrelationsmatrix im Lösungspunkt der Entwurfszentrierung.	72
4.6	Vergleich des Linearisierungsfehlers	72
B.1	BB-Kurve mit drei Kontrollpunkte im drei-dimensionalen Raum	82

Tabellenverzeichnis

2.1	Überblick über strukturelle Nebenbedingungen für typische analoge CMOS-Grundstrukturen.	27
2.2	Approximierte Ausbeute \bar{Y}_i in Abhängigkeit vom Worst-Case-Abstand $\beta_{WC,i}$	32
4.1	Längen und Weiten der Transistoren für den Miller Operationsverstärker.	66
4.2	Strukturelle Nebenbedingungen für den Miller Operationsverstärker.	66
4.3	Ergebnisse der Dimensionierung des Operationsverstärkers aufgespalten in Nominalentwurf und Entwurfszentrierung.	67
4.4	Vergleich von Nominalentwurf mit und ohne Betriebsparameter für den Miller Operationsverstärker	67
4.5	Vergleich der Dimensionierung des Miller Operationsverstärkers mit und ohne die Berücksichtigung struktureller Nebenbedingungen.	68
4.6	Dimensionierungsergebnisse Ausgangstreiber	70
4.7	Vergleich des Linearisierungsfehlers des Worst-Case Abstandes und der Zielfunktion	73

Verzeichnis der Formelzeichen

Mengen und Gebiete

- \mathcal{A}_f Akzeptanzgebiet der Eigenschaften
- \mathcal{A}_s Akzeptanzgebiet der statistischen Parameter
- \mathcal{E}_d Zulässige Bereich der Entwurfparameter
- $\mathcal{S}_{B,i}$ Menge der Grenzparametersätze im Raum der statistischen Parameter
- \mathcal{T}_θ Toleranzbereich der Betriebsparameter
- $\mathcal{D}_{B,i}$ Menge der Grenzparametersätze im Raum der Entwurfparameter

Matrizen und Vektoren

- \mathbf{C} Kovarianzmatrix der statistischen Parameter
- \mathbf{R} Korrelationsmatrix der statistischen Parameter
- \mathbf{U} Empfindlichkeitsmatrix der Nebenbedingungen
- $\mathbf{\Sigma}$ Streuungsmatrix
- \mathbf{d} Vektor der Entwurfparameter
- \mathbf{d}_L Vektor der Untergrenzen der Entwurfparameter
- \mathbf{d}_U Vektor der Obergrenzen der Entwurfparameter
- $\mathbf{g}_{d,i}$ Gradient $\nabla_{\mathbf{d}} f_i$
- $\mathbf{g}_{\theta,i}$ Gradient $\nabla_{\boldsymbol{\theta}} f_i$
- $\mathbf{g}_{s,i}$ Gradient $\nabla_{\mathbf{s}} f_i$
- \mathbf{f} Vektor der Schaltungseigenschaften
- \mathbf{f}_B Vektor der spezifizierten Obergrenzen
- \mathbf{p} Vektor der Schaltungsparameter
- \mathbf{s} Vektor der statistischen Parameter
- \mathbf{s}_0 Vektor der Erwartungswerte der statistischen Parameter
- \mathbf{u} Vektor der Nebenbedingungen
- $\boldsymbol{\alpha}$ Vektor der Parameterabstände
- $\boldsymbol{\beta}_{WC}$ Vektor der Worst-Case Abstände
- $\boldsymbol{\epsilon}$ Vektor der Fehler der Schaltungseigenschaften
- $\boldsymbol{\sigma}$ Vektor der Streuungen der statistischen Parameter
- $\boldsymbol{\theta}$ Vektor der Betriebsparameter
- $\boldsymbol{\theta}_L$ Vektor der Untergrenzen der Betriebsparameter
- $\boldsymbol{\theta}_U$ Vektor der Obergrenzen der Betriebsparameter

Skalare und Funktionen

Y	Parametrische Ausbeute
\bar{Y}_i	Approximierte parametrische eigenschaftsorientierte Ausbeute
α	Skalierungsfaktor für die Zielgrößen in der Zielfunktion
$\varphi(\mathbf{d}_0)$	Zielfunktion für die Optimierung
$\varphi_\alpha(\mathbf{d}_0)$	Zielfunktion für den Nominalentwurf
$\varphi_\beta(\mathbf{d}_0)$	Zielfunktion für die Entwurfszentrierung
σ_i	Streuung des statistischen Parameters s_i
$\rho_{i,j}$	Korrelationsfaktor $\in [-1, 1]$ zwischen den statistischen Parameters s_i und s_j
$L(\cdot)$	Lagrange-Funktion
λ	Lagrange Parameter
μ	Iterationsschritt
a	Normierte Auslenkung der Grenzkurve
m	Steigung $\partial r / \partial a$ der normierten Grenzkurve
$n_{\mathbf{d}}$	Anzahl der Entwurfsparameter
$n_{\mathbf{f}}$	Anzahl der Eigenschaften/Spezifikationsgrenzen
$n_{\mathbf{p}}$	Anzahl der Parameter
$n_{\mathbf{s}}$	Anzahl der statistischen Parameter
$n_{\boldsymbol{\theta}}$	Anzahl der Betriebsparameter
a	Normierter Fehler der Grenzkurve
t	Parameter $\in [0, 1]$ der parametrischen Berstein-Bézier Kurve

Kurzfassung

Diese Arbeit stellt Verfahren zur automatischen simulationsbasierten Dimensionierung analoger integrierter Schaltungen vor. Die Dimensionierung wird in Nominalentwurf und Entwurfszentrierung unterschieden wodurch die Gesamtzahl der Schaltungssimulationen möglichst klein gehalten werden kann. Nominalentwurf und Entwurfszentrierung werden als ein vergleichbares Optimierproblem formuliert, für das ein effizienter Lösungsalgorithmus präsentiert wird. Dabei wurden strukturelle Nebenbedingungen konsequent berücksichtigt. Anhand von praktischen Schaltungsbeispielen wurde nachgewiesen, daß die entwickelten Verfahren vollautomatisch zu schaltungstechnisch sinnvollen und robusten Dimensionierungen führen.