Kapitel 1

Einleitung und Motivation

Im Bereich der Strukturmechanik wird die folgende Definition des Begriffes Leichtbau verwendet [1]:

> Von Leichtbau spricht man, wenn eine Konstruktion so gebaut ist, daß sie mit einem Minimum an Masse die gestellten Anforderungen wie Festigkeit oder Steifigkeit erfüllt.

Ähnliche Definitionen des Begriffs finden sich bei Wiedemann [2], Hertel [3] und anderen. Wird also Leichtbau betrieben, so ist zumindest implizit ein Optimierungsproblem zu lösen. In der Praxis geschieht dies meistens durch die Auswahl geeigneter Bauformen, den sogenannten Formleichtbau [1], und/oder Materialien, dem Materialleichtbau [1]. Das Optimierungsproblem im mathematischen Sinne wird dabei oft nicht betrachtet, die Lösung erfolgt durch „ingenieursmäßiges“ Vorgehen. Dies bedeutet, dass eine viel versprechende Bauweisen- oder Materialauswahl getroffen wird, die auf Erfahrungen, Literaturquellen oder innovativen Ideen beruht.

Um ein Problem unter Leichtbauaspekten zu lösen muss zunächst ein geeignetes Berechnungsverfahren für die relevanten Struktureigenschaften bestimmt werden. Es soll die vorliegende Struktur möglichst genau berechnen. Die unvermeidbaren Ungenauigkeiten in folge von Vereinfachungen und Vernachlässigungen sind so einzustellen, dass sich ein konservatives Ergebnis ergibt. Mit einem solchen Verfahren lassen sich Voraussagen machen, die zum einen möglichst nahe an realen Ergebnissen, beispielsweise aus Versuchen liegen, zum anderen jedoch unterhalb der realen Versagenslasten. Die stetige Weiterentwicklung von Berechnungsverfahren führt dazu, dass die realen Versagenslasten immer genauer vorausgesagt werden können. Unter der Voraussetzung, dass die Verfahren konservative Vorhersagen liefern, entsteht dadurch ein Masseneinsparungspotenzial für Strukturen, die nach früheren Verfahren ausgelegt worden sind. Um diese Einsparmöglichkeit für abgeleitete Strukturen zu nutzen, ist der Einsatz von mathematischen Optimierungsalgorithmen sinnvoll, da sie vorhandene Reserven möglichst effektiv in Massenreduktion umsetzen können.

Das Vorgehen bei der Berechnung in der Strukturmechanik geht in der Regel von einfachen Strukturelementen aus, die in verschiedenen Ausführungen zusammengesetzt eine Gesamtstruktur ergeben. Als Beispiel dazu sei die Methode der finiten Elemente genannt [4], bei der sehr komplexe Strukturen wie Automobilkarosserien, Motoren oder Flugzeugrümpfe aus einfachen (finiten) Elementen zusammengesetzt werden, die durch einige kleine Matrizen

(lineare Steifigkeit, Masse und nichtlineare Steifigkeit) beschrieben werden. Ein weiteres Beispiel ist die Behandlung eines Flugzeugrumpfes als Balken, wie sie Kapitel in 4 verwendet wird. Der Balken ist ein einfaches Element der Strukturmechanik, im Fall eines Rumpfes setzt sich sein Querschnitt jedoch aus mehr als 100 Stringern und 100 Hautfeldern zusammen, so dass sich ein komplexes Gesamtsystem ergibt. Nach der Definition des Leichtbaus muss dafür eine Auslegung gefunden werden, die ein Minimum der Masse unter den gegebenen Randbedingungen darstellt. Für reale Strukturen ist die Zahl der Variablen meistens so groß, dass ein Mensch nicht in der Lage sein wird sie optimal zu bestimmen.

Zusätzlich zur großen Zahl der möglichen Variablen sind die strukturmechanischen Optimierungsprobleme oft stark nichtlinear. Dies fängt bei der Berücksichtigung des plastischen Materialverhaltens an, ohne das keine massenoptimale Auslegung gefunden wird [1]. Weitere Nichtlinearitäten werden durch große Verformungen, Kontakt oder Reibung verursacht. Die Berechnungsverfahren an sich können ebenfalls Nichtlinearitäten in den Optimierungsprozess einbringen. Ein Beispiel hierfür ist das Hautbeulen von Flugzeugrümpfen, das zwar nicht zum Versagen der Struktur führt, wohl aber zu Spannungsumlagerungen, die zum einen nichtlinear von der Spannung abhängen, zum anderen plötzlich bei Überschreiten der Beulspannung einsetzen. An dieser Stelle ist das Berechnungsverfahren nicht differenzierbar und nicht linear. Dies entspricht nicht dem realen Strukturverhalten, bei dem das Beulen kontinuierlich stärker wird, sondern wird durch die Modellierung des Strukturverhaltens verursacht. An sich ist damit die Modellierung nicht ausreichend genau. Bessere Verfahren sind aber teilweise nicht bestimmbar oder führen zu einem zu großen Rechenaufwand. Sie sind auch für die Berechnung der Versagenslast nicht nötig, da die vorhandene Modellierung ausreicht. Nur für eine Optimierung ist die Differenzierbarkeit unter Umständen erforderlich. Eine letzte Nichtlinearität wird durch die Optimierungsvariablen selbst verursacht. Einige Variablen können nur ganzzahlige Werte annehmen (z. B. die Zahl der Stringer im Rumpf) oder sie dürfen nur einige bestimmte Werte zugewiesen bekommen. Diese Werte lassen sich unter Umständen nicht ordnen, so dass sich eine nichtlineare Abhängigkeit des Berechnungsergebnis von diesen Variablen ergibt. Ein Beispiel hierfür ist die Optimierung des Materials. Mit der Variablen kann nur eines von wenigen zur Verfügung stehenden Materialien ausgewählt werden. Diese Wahl beeinflusst eine Reihe von Größen, die für das Berechnungsergebnis wichtig sind, z. B. E-Modul, Fließgrenze, Schubmodul oder Dichte. Ein Material mit höherer Fließgrenze kann eine niedrigere Dichte oder einen niedrigeren E-Modul haben als ein anderes Material. Damit ist eine eindeutige Ordnung der Variablenwerte im Sinne von besserem und schlechterem Material nicht möglich.

Aus diesen Überlegungen ergeben sich Anforderungen an eine rechnergestützte Optimierung. Sie muss in der Lage sein eine große Zahl von Variablen für unstetige, nicht differenzierbare Zielfunktionen zu bestimmen. Dabei sind Restriktionen einzuhalten, die ebenfalls unstetig und nicht differenzierbar sein können. Die Variablen können Werte aus dem Bereich der reelen Zahlen sein, sie können aber auch auf die natürlichen Zahlen oder diskrete Wertemengen begrenzt sein.

Es ist direkt offensichtlich, dass diese Forderungen nicht allgemein erfüllt werden können. Es gibt kein Optimierungsverfahren, das in der Lage ist, für ein so allgemein formuliertes Optimumproblem das globale Optimum [5] zu identifizieren. Die Zielfunktion und die Restriktionen sind geprägt von einer großen Zahl lokaler Optima, die teilweise sehr ähnliche Funktionswerte aufweisen und somit kaum zu unterscheiden sind.

Das Ziel dieser Arbeit ist es, ein Optimierungsverfahren zu entwickeln, das mit geringen Anpassungen auf eine Vielzahl von Optimierungsaufgaben der oben beschriebenen Art an-

wendbar ist. Dabei kann die Identifikation des globalen Optimums nicht zwingend Aufgabe dieses Verfahrens sein, da dies für eine Reihe von Problemen nur durch Ausprobieren aller Möglichkeiten sicher identifiziert werden kann. Ein bekanntes Beispiel dafür ist das Travelling-Salesman-Problem [6]. Sind einige der Variablen reele Zahlen, so ergibt sich eine unendliche Zahl an Möglichkeiten und das Problem ist nicht mehr geschlosssen lösbar. Vielmehr soll das Verfahren zuverlässig gute Ergebnisse in Form von lokalen Optima oder, wenn möglich, des globalen Optimums liefern. Das Optimierungsprogramm soll die bisherige Arbeitsweise in der Strukturmechanik nicht ersetzen, sondern ergänzen. Die Erfahrung und Intuition des Ingenieurs sollen einfließen und zu einem guten Ergebnis beitragen.

Um diese Zielsetzung zu erreichen wird der Optimierungsprozess als Ganzes betrachtet. Im Einzelnen umfasst er die folgenden Schritte:

- Definition der zu optimierenden Zielfunktion und der Restriktionsfunktionen

  Durch die Wahl der Berechnungsverfahren wird bereits festgelegt, welche Variablen optimierbar sind. Des Weiteren können die gefundenen optimalen Konfigurationen nur innerhalb des Gültigkeitsbereichs der verwendeten Verfahren liegen. Für eine erfolgreiche Optimierung sollten deshalb die Berechnungsverfahren sorgfältig ausgewählt und gegebenenfalls erweitert werden. Im Beispiel der Optimierung von Mikrometeoriten und Space Debris Schutzsystemen (Kapitel 3) wird gezeigt, wie mit einer Erweiterung des Verfahrens zusätzliches Optimierungspotenzial zugänglich gemacht werden kann.

- Auswahl der zu optimierenden Variablen und ihrer Grenzen

  Offensichtlich hat die Festlegung der Variablengrenzen entscheidenden Einfluss auf das erzielte Ergebnis. Durch ungeeignete Grenzen können gute Konfigurationen ausgeschlossen werden. Durch zu große Intervalle wird das zu durchsuchende Gebiet und damit der Rechenaufwand vergrößert. Bei der Festlegung der Grenzen ist man daher stark auf die Erfahrung und die Kenntnisse des Anwenders angewiesen. Dies gilt auch bei der Auswahl der Variablen. Durch das Weglassen von Variablen kann der Rechenaufwand verringert werden und/oder das Ergebnis negativ beeinflusst werden. Neben der Erfahrung des Nutzers bietet sich hier eine systematische Sensitivitätsanalyse vor der eigentlichen Optimierung an, wie sie in Kapitel 2.1.3 vorgestellt wird. Damit werden dem Nutzer Zahlenwerte zur Verfügung gestellt, die eine gezielte Auswahl der Variablen ermöglichen. Darüber hinaus sind aus der Sensitivitätsanalyse Tendenzen ableitbar, die qualitative Aussagen für die optimale Konfiguration erlauben.

- Auswahl und Anwendung eines Optimierungsalgorithmus

  Mit dem Optimierungsergebnis werden bereits die Eigenschaften der Zielfunktion festgelegt, die optimiert werden können. Für die oben beschriebenen, nicht differenzierbaren Zielfunktionen sind zum Beispiel gradientenbasierte Verfahren nicht anwendbar. Grundsätzlich kann für jedes zu optimierende Problem ein spezieller Algorithmus entwickelt werden, der besonders effektiv und schnell arbeitet. In dieser Arbeit soll dieser Weg jedoch nicht beschritten werden, da dann schon bei kleinen Änderungen am Berechnungsverfahren gravierende Änderungen am Optimierer notwendig werden können. Es soll im Gegenteil ein Verfahren entwickelt werden, das bei geringen Modifikationen des Berechnungsverfahrens keine Änderungen erfordert. Zusammen mit den oben skizzierten hohen Anforderungen bezüglich der Variablenzahl und Charakteristiken der Funktionen, stellt diese Zielsetzung große Ansprüche an das zu entwickelnde

Verfahren. Daher werden in Kapitel 2 eine Reihe möglicher Kandidaten für solche Algorithmen diskutiert. Die wenigen, die die Zielsetzung erfüllen können, werden zu einem kombinierten Algorithmus zusammengestellt, der ihre Stärke verbindet und die Schwächen minimiert. Dieser kombinierte Algorithmus ist in der Lage die Zielsetzung zu erfüllen, was an sich schon ein großer Vorteil ist. Dies wird durch eine vergleichsweise hohe Rechenzeit erkauft. Hierbei muss jedoch beachtet werden, dass die Rechenleistung stetig steigt [7, 8] und damit in wenigen Jahren Probleme lösbar werden, die heute mangels Rechenkapazität nicht bearbeitet werden können. Damit wird auch dieser Nachteil des kombinierten Algorithmus relativiert.

- Überprüfung und Vergleich des erzielten Ergebnisses

  Die Überprüfung der Ergebnisse ist eine Selbstverständlichkeit in der Ingenieurspraxis. Für jede optimierte Konfiguration muss geprüft werden, ob die Resultate plausibel sind, ob die gemachten Annahmen und Vereinfachungen zulässig sind und die Konfiguration die Voraussetzungen der Verfahren erfüllt. In der Praxis und damit in dieser Arbeit müssen die optimierten Konfigurationen bezüglich ihrer Optimalität bewertet werden. Da in der Regel das globale Optimum nicht bekannt und nicht bestimmbar ist, müssen Referenzkonfigurationen verwendet werden, die durch ingenieursmäßiges Vorgehen oder „scharfes Hingucken“ bestimmt werden.

Es wird zunächst ein Optimierungsverfahren definiert, das die Zielsetzungen erfüllt. Anschließend wird es auf drei Beispiele angewendet. Dies sind zum einen die Massenoptimierung von Mikrometeoriten und Space Debris Schutzsystemen, die Querschnittsoptimierung von orthotrop versteiften Rumpfschalen und die Massenoptimierung von Sandwichzylinderschalen. In allen Beispielen müssen Nebenbedingungen eingehalten werden, die im weitesten Sinne als Sicherheit gegen Versagen bezeichnet werden können. Für jedes Beispiel wird zunächst das Berechnungsverfahren dargestellt, das während der Optimierung verwendet wird. Dabei werden zum Teil notwendige Verfahren erst erarbeitet, ohne die eine Optimierung nicht möglich wäre. Anschließend werden ein oder mehrere Testfälle optimiert, wobei jeweils die Variablen und ihre Grenzen unter Zuhilfenahme einer Sensitivitätsanalyse festgelegt werden. Die erzielten Optimierungsergebnisse werden mit geeigneten Referenzkonfigurationen, wie sie sich in der Praxis ergeben, verglichen und bewertet. Am Ende der Arbeit erfolgt eine Zusammenfassung, aus der sich Ausblicke für zukünftige Arbeiten ergeben.

Kapitel 2

Optimumsprobleme

2.1 Extremwertprobleme

Unter Optimierung im Allgemeinen versteht man Vorgänge, bei denen unter gegebenen Voraussetzungen eine bestmögliche Lösung im Hinblick auf ein zu erreichendes Ziel gesucht wird.

Optimierung in der Strukturmechanik befasst sich mit mathematischen Methoden für die optimale Auslegung von Bauteilen, Tragwerken und anderen mechanischen Systemen. Unter mathematischen Methoden werden dabei die Ermittlung und Beschreibung der besten Auswahl aller unter vorgegebenen Bedingungen und Anforderungen in Frage kommenden Alternativen verstanden. Dieser Entscheidungsprozess kommt mithilfe mathematisch-numerischer Algorithmen zustande. Im speziellen bedeutet das: „Es wird ein Satz von freien Parametern (die Optimierungs- oder Entwurfsvariablen) eines Systems (mechanische Struktur) so bestimmt, dass ein oder mehrere Gütekriterien (Zielfunktionen) bestmöglich erfüllt sind und gleichzeitig zu beachtende physikalisch-technische Anforderungen (Restriktionen) eingehalten werden“ [9].

In der Strukturmechanik wird häufig auf numerische Hilfsmittel wie die FEM-Analyse zurückgegriffen. Diese führen bei Optimierungsaufgaben auf einen nur numerisch darstellbaren Zusammenhang von Optimierungs- und Antwortvariablen. Man spricht dann von einer impliziten Optimierungsaufgabe. Im Gegensatz dazu ergibt sich bei kleineren Aufgaben oft ein formelmäßiger Zusammenhang, dann ist eine explizite Aufgabe gegeben.

2.1.1 Mathematische Grundlagen

Die Anwendung von Verfahren zur Parameteroptimierung erfordert eine mathematische Modellbildung. Damit wird das Problem in eine mathematische Beschreibungsweise überführt, auf die spezielle Optimierungsalgorithmen angewandt werden können. Nach Kölsch [5] lässt sich das mathematische Modell der Parameteroptimierung allgemein wie folgt formulieren:

$$\text{Minimiere} \quad f(\vec{X}) \tag{2.1}$$

$$\text{mit} \quad g_\alpha(\vec{X}) = 0 \qquad \alpha = 1, 2, \ldots, m' \tag{2.2}$$

$$\text{und} \quad g_\alpha(\vec{X}) \geq 0 \qquad \alpha = m'+1, \ldots, m \tag{2.3}$$

$$\text{und} \quad \vec{X}_{\beta,u} \leq \vec{X} \leq \vec{X}_{\beta,o} \qquad \beta = 1, 2, \ldots, n' \tag{2.4}$$

$$\text{und} \quad \vec{X} \in D_{i,\beta} \qquad \beta = n'+1, \ldots, n \tag{2.5}$$

$$D_{i,\beta} = \left\{ \vec{X}_{\beta,1}, \vec{X}_{\beta,2}, \ldots \right\}$$

Die Zielfunktion Gleichung (2.1) beschreibt das zu optimierende Merkmal in mathematischer Form. Sie ist dabei abhängig von den Optimierungsparametern, die im Vektor $\vec{X}$ zusammengefasst werden. Zudem müssen noch Restriktionsfunktionen erfüllt werden, die allgemein in den Gleichungen (2.2) und den Ungleichungen (2.3) angegeben sind. Sie werden zusammen von 1 bis m durchnummeriert, mit m' als letzte Gleichung und $m'+1$ als erste Ungleichung. Mögliche Grenzen, zwischen denen sich die Optimierungsparameter bewegen dürfen, sowie diskrete Werte der Optimierungsparameter werden mit (2.4) und (2.5) erfasst. Die Nummerierung erfolgt analog zu den Gleichungen und Ungleichungen, mit dem Buchstaben n. Es wird hier nur die Minimierung einer Zielfunktion betrachtet, da sich jedes Maximierungsproblem durch Multiplikation der Zielfunktion mit -1 in ein Minimumproblem überführen lässt.

Mit Hilfe dieser allgemeinen Formulierung lassen sich Optimierungsprobleme entsprechend der Eigenschaften ihrer Gleichungen klassifizieren:

Eigenschaften der Zielfunktion	Eigenschaften der Restriktionsfunktionen
Funktion einer Variablen	Keine Beschränkungen (m=0)
Lineare Funktion	Einfache Parameterbeschränkungen
Quadratische Funktion	Lineare Beschränkungen
Stetige nichtlineare Funktion	Nichtlineare Beschränkungen
Nichtlineare Funktion	

Tabelle 2.1: Klassifizierung von Optimierungsproblemen anhand ihrer Eigenschaften [10]

Die Zielfunktion kann von einer oder mehreren Variablen abhängen. Ist $\vec{X} \in \mathbb{R}^1$, so können die aus der Analysis bekannten Methoden angewandt werden. Dazu wird die Funktion zweimal abgeleitet, werden die Nullstellen der ersten Ableitung bestimmt, und mit der Hilfe der zweiten Ableitung entschieden, ob ein Minimum, Maximum oder ein Sattelpunkt vorliegt. Die Funktionswerte an den Minima und die Werte an den eventuell vorhanden Grenzen werden dann verglichen und so das globale Minimum bestimmt.

Bei Zielfunktionen mehrerer Veränderlicher ist dieses Verfahren anwendbar, wenn die Aufgabe nicht restringiert ist. Die erste Ableitung muss Null sein und die zweite Ableitung muss positiv definit sein. Bei beschränkten Problemen gilt aber gerade die letzte Bedingung nicht

immer. Dann müssen andere Algorithmen angewandt werden, die in ihrer Anwendbarkeit von den Eigenschaften der Zielfunktion und der Restriktionen abhängen. Solche Algorithmen werden in Kapitel 2.2 erläutert und vorgestellt.

Die Restriktionen stellen eine weitere Komplikation des Problems dar, denn durch sie wird der Lösungsraum definiert. Dies ist die Menge aller $\vec{X}$, die alle Restriktionen erfüllt. Nur Elemente von $\vec{X}$, die innerhalb des Lösungsraumes liegen, sind als Lösung des Optimierungsproblems zulässig.

Für den Fall, dass keine Restriktionen gegeben sind, stehen eine Vielzahl von effizienten und erprobten Verfahren zur Verfügung. Daher kann es sinnvoll sein Restriktionen zu eliminieren. Dies wird durch sogenannte Straffunktionen erreicht. Dabei wird die Zielfunktion so modifiziert, dass bei Verletzung einer Restriktion ein Strafterm addiert wird. Diese Lösung ist dann in ihrer Qualität herabgesetzt und wird als nicht optimal bewertet.

Straffunktionen können auf unterschiedliche Weise realisiert werden, von der Addition konstanter Strafterme über den linearen bis hin zum exponentiellen Anstieg in Abhängigkeit von der Größe der Verletzung.

Problematisch beim Einsatz von Straffunktionen ist, dass Kenntnisse der Originalzielfunktion nötig sind. Sind die gewählten Strafterme zu klein, so kann es vorkommen, dass der Lösungsalgorithmus bei einem unzulässigen Punkt konvergiert. Zudem gibt es Probleme, wenn der Lösungsraum in diskrete Bereiche unterteilt ist. Bei schrittweisen Lösungsverfahren führt der Weg zum gesuchten, globalen Minimum je nach Startvektor nur über unzulässige Punkte, die durch zu hohe Strafterme den Algorithmus in lokale Minima zurückdrängen.

Optimierungsprobleme werden als linear bezeichnet, wenn sowohl die Zielfunktion als auch die Restriktionsfunktionen linear von den Optimierungsparametern abhängen. Trifft dies in einer Gleichung nicht zu, so liegt ein nichtlineares Problem vor. Die meisten Probleme aus der Strukturmechanik sind nichtlinear. Zum Beispiel ist oft das Gewicht zu beachten, welches vom Volumen abhängt, auch Schwingungseinflüsse sind nichtlinearer Natur. Nichtlineare Probleme sind im Allgemeinen multimodal, was bedeutet, dass sie neben einem globalen auch noch lokale Minima im Lösungsraum besitzen.

Eine wichtige geometrische Eigenschaft für Optimierungsprobleme ist die Konvexität von Funktionen und Bereichen. Durch sie wird bestimmt, ob ein Problem nur eine Lösung hat, oder ob mehrere Lösungen existieren können. Von konvexen Optimierungsproblemen spricht man immer, wenn sowohl die Zielfunktion konvex ist als auch die Restriktionsfunktionen einen konvexen Bereich beschreiben und alle Funktionen stetig sind.

Eine Funktion $f\left(\vec{X}\right)$ heißt konvex, wenn für zwei Punkte $\vec{X}_1$ und $\vec{X}_2$ mit $0 < a < 1$ gilt [9, 11]:

$$f\left(a\vec{X}_1 + (1-a)\,\vec{X}_2\right) \;\leq\; a \cdot f\left(\vec{X}_1\right) + (1-a)\,f\left(\vec{X}_2\right) \tag{2.6}$$

Eine geometrische Interpretation der Gleichung (2.6) kann am Beispiel der Funktionsverläufe in Abbildung 2.1 vorgenommen werden.

Aus dieser Darstellung erkennt man, dass eine Funktion dann konvex ist, wenn alle Punkte auf einer Geraden, deren Anfangs- und Endwerte von den Funktionswerten bei $\vec{X}_1$ und $\vec{X}_2$ gebildet werden, größere Werte annehmen als die Funktion in diesem Intervall.

Ist die Funktion von einer Variablen abhängig und zweimal differenzierbar, so gilt für Konvexität, dass $\frac{\partial^2 f}{\partial X^2} \geq 0$ ist. Bei Funktionen mehrerer Veränderlicher gilt entsprechend, dass

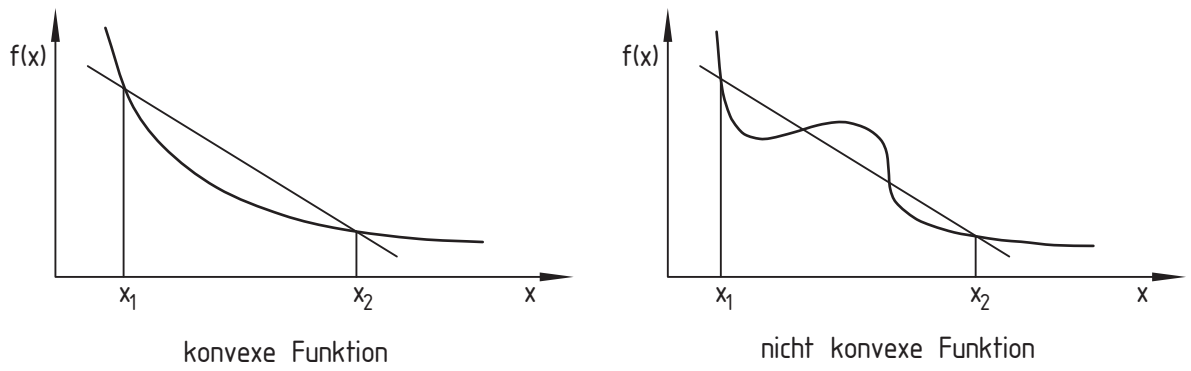

Abbildung 2.1: Konvexität von Funktionen

die Hesse-Matrix $H = \frac{\partial^2 f}{\partial x_i \partial x_j}$ positiv definit ist. Eine Matrix ist positiv definit, wenn sie symmetrisch ist und alle ihre Eigenwerte größer Null sind [11].

Ein Bereich ist dann konvex, wenn für beliebige zwei Punkte $\vec{X}_1$ und $\vec{X}_2$ des Bereiches gilt, dass auch jeder Punkt der Verbindungslinie zwischen den beiden Punkten innerhalb des Bereichs liegt [5]. Dies ist in Abbildung 2.2 dargestellt.

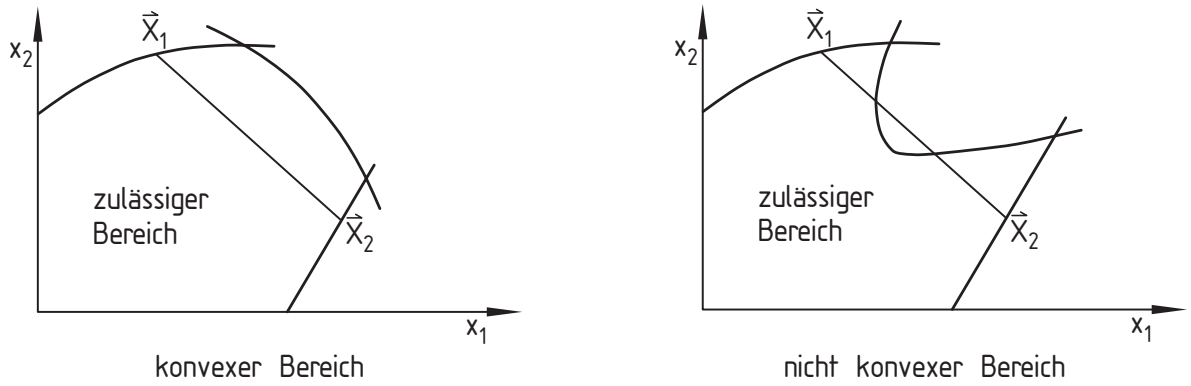

Abbildung 2.2: Konvexität von Lösungsräumen

Aus diesen Betrachtungen wird deutlich, dass bei impliziten Optimierungsaufgaben eine Verifikation der Konvexität meistens unmöglich ist, da die Positivdefinitheit der Hesse-Matrix für alle Funktionen bestenfalls nur für wenige Punkte nachweisbar ist. Daher ist von Nichtkonvexität und damit von der möglichen Existenz lokaler Minima auszugehen. Viele Sätze der mathematischen Optimierungsrechnung und der theoretischen Konvergenzbetrachtung von Algorithmen gelten aber nur für konvexe Probleme [9]. Ihre Umsetzung und Anwendbarkeit für nicht konvexe Aufgaben ist daher sorgfältig zu prüfen.

Aus den bisherigen Betrachtungen ergibt sich die Definition der Lösung der Optimierungsaufgabe: Ein zulässiger Vektor $\vec{X}_{opt}$ heißt Lösung von Gleichungssystem (2.1)-(2.5), wenn in seiner Umgebung U kein weiterer zulässiger Vektor $\vec{X}$ mit einem niedrigeren Zielfunktionswert existiert (vgl. [9]), also:

$$f\left(\vec{X}_{opt}\right) \leq f\left(\vec{X}\right) \ \forall \vec{X} \in U \tag{2.7}$$

Kann die Umgebung U beliebig groß sein, also den ganzen Lösungsraum umfassen, so ist $\vec{X}_{opt}$ die globale, ansonsten nur eine lokale Lösung.

In der Praxis begegnet man häufig auch sogenannten flachen Optima, in deren Nachbarschaft sich die Werte der Zielfunktion nur wenig ändert. Bei Aufgaben mit mehreren Veränderlichen ist dies durch den gegenläufigen Einfluss der Variablen möglich, die Vergrößerung einer Variablen führt durch Restriktionen zur Verkleinerung einer anderen, so dass sich die Zielfunktion nur wenig ändert.

Mathematische Kriterien, die eine sichere und eindeutige Identifikation des globalen Optimums erlauben, sind für allgemeine Optimierungsaufgaben nicht bekannt [5]. Für konvexe Probleme repräsentieren aber die Kuhn-Tucker-Bedingungen notwendige und hinreichende Optimalitätsbedingungen.

Zur Formulierung der Kuhn-Tucker-Bedingungen wird die Optimierungsaufgabe (2.1)-(2.5) als Lagrangefunktion geschrieben. Sie stellt eine Möglichkeit dar, Restriktionen und Zielfunktion in einer Funktion zu berücksichtigen. Die Funktion

$$L\left(\vec{X},\vec{\lambda}\right) = f\left(\vec{X}\right) - \sum_{\alpha=1}^{m} \lambda_\alpha \cdot g_\alpha\left(\vec{X}\right) \tag{2.8}$$

mit

$$\lambda_\alpha \;\geq\; 0 \tag{2.9}$$

heißt Lagrangefunktion der nichtlinearen Optimierungsaufgabe (2.1)-(2.5). λ_α sind die zu den Restriktionen gehörenden Lagrangeparameter. Die wesentliche Eigenschaft der Lagrangefunktion ist, dass ein Sattelpunkt von L gleichzeitig Lösung des zugehörigen Optimumproblems ist. Verdeutlicht werden soll dies an einer Optimierungsaufgabe, die nur eine Ungleichheitsrestriktion enthält. Die zugehörige Lagrangefunktion lautet:

$$L\left(\vec{X},\vec{\lambda}\right) = f\left(\vec{X}\right) - \lambda_1 \cdot g_1\left(\vec{X}\right) \tag{2.10}$$

Ein Sattelpunkt $\vec{X}_{opt}$, $\vec{\lambda}_{opt}$ ist nach [12] wie folgt definiert:

$$L\left(\vec{X}_{opt},\vec{\lambda}\right) \leq L\left(\vec{X}_{opt},\vec{\lambda}_{opt}\right) \leq L\left(\vec{X},\vec{\lambda}_{opt}\right) \tag{2.11}$$

Da $\lambda_1 \geq 0$, ist der linke Teil der Ungleichung (2.11) nur erfüllbar, wenn $g_1\left(\vec{X}\right) \geq 0$ ist. Damit sorgt diese Ungleichung für die Zulässigkeit der gefundenen Lösung. Die rechte Ungleichung sorgt für ein Minimum von f an der Stelle $\vec{X}_{opt}$, da $g_1\left(\vec{X}_{opt}\right) = 0$ [9]. Wäre g_1 nicht Null, so wäre diese Restriktion im Optimum nicht aktiv und könnte für die Bestimmung des Optimums ignoriert werden.

Bei nichtkonvexen Aufgaben ist diese Sattelpunkteigenschaft hinreichend und nicht notwendig. Das heißt, dass in diesem Fall nicht zu jeder Lösung der Optimierungsaufgabe ein entsprechender Sattelpunkt der zugehörigen Lagrangefunktion existieren muss. Für konvexe Probleme ist die Eigenschaft notwendig und hinreichend, für die einzige Lösung $\vec{X}_{opt}$ existiert genau ein Sattelpunkt mit $\vec{X}_{opt}$, λ_{opt}.

Die Kuhn-Tucker-Bedingungen sind notwendig für nichtkonvexe Probleme und notwendig und hinreichend für konvexe Aufgaben. Sei $\vec{X}_{opt}$, eine Lösung der Optimierungsaufgabe (2.1)-(2.5), dann gibt es ein $\vec{\lambda}$, so dass gilt [9, 12]: