

MATHEMATIK

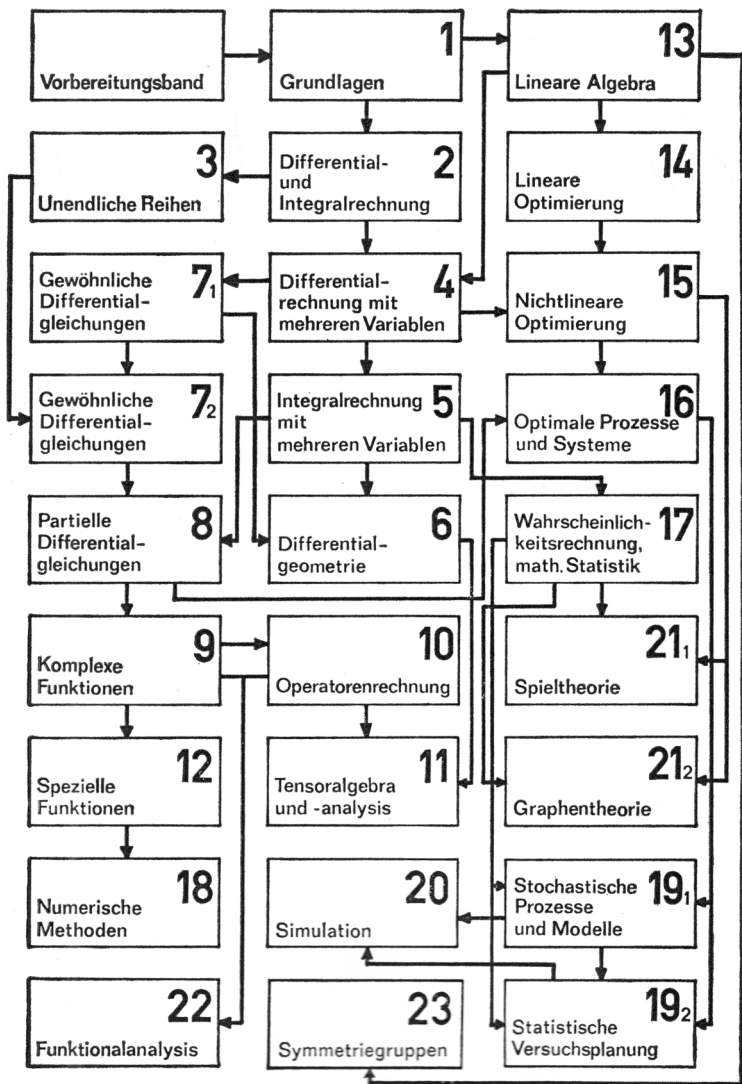
FÜR INGENIEURE
NATURWISSENSCHAFTLER
ÖKONOMEN
LANDWIRTE

16

BIESS · ERFURTH · ZEIDLER

Optimale Prozesse
und Systeme

Abhängigkeitsgraph



MATHEMATIK FÜR INGENIEURE, NATURWISSENSCHAFTLER,
ÖKONOMEN UND LANDWIRTE · BAND 16

Herausgeber: Prof. Dr. O. Beyer, Magdeburg · Prof. Dr. H. Erfurth, Merseburg
Prof. Dr. O. Greuel †, · Prof. Dr. H. Kadner, Dresden
Prof. Dr. K. Manteuffel, Magdeburg · Doz. Dr. G. Zeidler, Berlin

PROF. DR. G. BIESS
PROF. DR. H. ERFURTH
DOZ. DR. G. ZEIDLER

Optimale Prozesse und Systeme

2. AUFLAGE



BSB B. G. TEUBNER VERLAGSGESELLSCHAFT
1980

Verantwortlicher Herausgeber:

Dr. sc. O. Beyer, ordentlicher Professor an der Technischen Hochschule „Otto von Guericke“,
Magdeburg

Autor der Abschnitte 1., 3.3.–3.5.:

Dr. rer. nat. habil. H. Erfurth, ordentlicher Professor an der Technischen Hochschule „Carl Schor-
lemmer“ Leuna-Merseburg

Autor der Abschnitte 2.–3.2. und 4.1.–4.2.:

Dr. rer. nat. G. Zeidler, Dozent an der Hochschule für Ökonomie „Bruno Leuschner“, Berlin

Autor der Abschnitte 4.3.–4.4.:

Dr. sc. G. Bieß, ordentlicher Professor an der Technischen Hochschule „Carl Schorlemmer“ Leuna-
Merseburg

Als Lehrbuch für die Ausbildung an Universitäten und Hochschulen der DDR anerkannt.

Berlin, Dezember 1979

Minister für Hoch- und Fachschulwesen

Anerkanntes Lehrbuch seit der 1. Auflage 1974

© BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1974

2. Auflage

VLN 294-375/23/80 · LSV 1084

Lektor: Dorothea Ziegler

Printed in the German Democratic Republic

Gesamtherstellung:

ITERDRUCK Graphischer GroßbetriebN Leipzig – III/18/97

Bestell-Nr. 665 719 8

DDR 8,- M

Inhalt

1.	Einleitung	5
2.	Optimale Prozesse und damit zusammenhängende Begriffe	6
3.	Stetige Prozesse	9
3.1.	Grundaufgabe der Optimierung	10
3.2.	Pontrjaginsches Maximumprinzip	14
3.3.	Andere Aufgabentypen	20
3.3.1.	Probleme mit beweglichem Endpunkt	20
3.3.2.	Probleme mit anderem Optimierungsziel	24
3.3.3.	Nichtautonome Systeme	26
3.3.4.	Probleme mit fester Endzeit	29
3.3.5.	Einige weitere Probleme	35
3.4.	Numerische Methoden	36
3.4.1.	Vorbemerkungen	36
3.4.2.	Einige Abstiegsverfahren	37
3.4.3.	Behandlung als Zwei-Punkte-Randwertaufgabe	38
3.4.4.	Anwendung der Methode von Newton-Raphson	38
3.5.	Systeme mit verteilten Parametern	39
3.5.1.	Zur Problemstellung	39
3.5.2.	Ein hyperbolisches System	40
3.5.3.	Numerische Methoden	45
4.	Diskrete Prozesse	49
4.1.	Aufgabenstellung	49
4.2.	Diskretes Maximumprinzip	53
4.3.	Dynamische Optimierung	57
4.3.1.	Zusammenstellung der Grundlagen	58
4.3.2.	Einige andere Aufgabenstellungen	64
4.3.3.	Probleme, bei denen die Zielfunktion keine Summenform hat	68
4.3.4.	Probleme mit allgemeinerer Struktur	75
4.3.5.	Suboptimale Steuerungen	89
4.3.6.	Stetige Probleme	91
4.4.	Kombinierte Anwendung der dynamischen Optimierung und des Maximumprinzips	96
4.4.1.	Vergleich der Methoden	96
4.4.2.	Möglichkeiten der kombinierten Anwendung beider Methoden	98
	Lösungen der Aufgaben	104
	Literatur	106
	Namen- und Sachregister	107

1. Einleitung

Optimierungsaufgaben werden in der Mathematik seit langem betrachtet. So gehören notwendige Optimalitätsbedingungen für Extremalaufgaben ohne Nebenbedingungen bzw. mit Nebenbedingungen in Gleichungsform und für Extremalprobleme, wie sie in der Variationsrechnung betrachtet werden, schon zur klassischen Mathematik. In den letzten Jahrzehnten erfuhr jedoch die Theorie der Optimierung und der Lösungsmethoden einen gewaltigen Aufschwung. Ursachen dafür waren einerseits wichtige Aufgabenstellungen in Technik, Ökonomie und Naturwissenschaften — es gibt heute kaum ein Gebiet, in dem nicht Optimierungsprobleme vorkommen —, andererseits die Entwicklung der elektronischen Rechentechnik; erst dadurch wurde die Möglichkeit geschaffen, eine Vielfalt von Problemen bis zum numerischen Resultat zu bearbeiten.

Diese Etappe begann 1939, als der sowjetische Mathematiker L. W. Kantorowitsch Methoden zur Lösung linearer Optimierungsprobleme veröffentlichte; es handelte sich dabei um einen völlig neuen Aufgabentyp. In der folgenden Zeit wurden nicht nur Theorie und Lösungsmethoden der linearen Optimierung durch eine sehr große Zahl von Arbeiten ausgebaut und weitergeführt, sondern es entwickelte sich auch die nichtlineare Optimierung, wobei eine zentrale Stellung die sogenannte konvexe Optimierung einnimmt.

Vor etwa zwei Jahrzehnten begann auch die Entwicklung der Theorie der optimalen Steuerung von Prozessen (optimale Prozesse). Seit der Formulierung von notwendigen Optimalitätsbedingungen in Form des Maximumprinzips von L. S. Pontrjagin ist die Anzahl von Veröffentlichungen zu dieser Problematik sprunghaft angestiegen. Anfangs betrachtete man dabei Prozesse, die durch gewöhnliche Differentialgleichungen beschrieben werden. Später untersuchte man auch Probleme, bei denen das mathematische Modell aus partiellen Differentialgleichungen oder Integralgleichungen bestand. Parallel dazu wurden Optimierungsprobleme im Zusammenhang mit Modellen, die durch Differenzengleichungen beschrieben werden, untersucht.

Die ersten Arbeiten beschäftigten sich zunächst mit theoretischen Fragestellungen, d. h. mit notwendigen und hinreichenden Optimalitätsbedingungen und Existenzaussagen. Da die gewonnenen Optimalitätsbedingungen nicht immer einen praktikablen Algorithmus zur Berechnung der optimalen Steuerung liefern, wurde die Entwicklung numerischer Lösungsmethoden erforderlich.

Es sei hier darauf hingewiesen, daß eine wichtige Aufgabe der gegenwärtigen mathematischen Forschung darin besteht, von den konkreten Aufgabenstellungen der verschiedenen Optimierungsgebiete zu abstrahieren und unter Verwendung funktionalanalytischer Methoden eine einheitliche Theorie aufzubauen. Eine solche Optimierungstheorie in allgemeinen Räumen gestattet es, eine Vielzahl von Problemen unter einem einheitlichen Gesichtspunkt zu betrachten (vgl. Bd. 22).

Das Ziel des vorliegenden Buches besteht darin, dem Nichtmathematiker eine Einführung in die Grundlagen der Theorie optimaler Prozesse zu geben. Voraussetzung zum Verständnis sind Kenntnisse aus der linearen Algebra, der Differentialrechnung für Funktionen mit mehreren Variablen und über gewöhnliche und partielle Differentialgleichungen. Auf eine Beweisführung wird verzichtet. Wesentliche Ergebnisse werden durch Beispiele erläutert.

Die Autoren bedanken sich für zahlreiche Hinweise bei Herrn Professor Dr. Sieber und dem Herausgeber, Herrn Professor Dr. Beyer.

2. Optimale Prozesse und damit zusammenhängende Begriffe

Ziel dieses Abschnittes ist es, die Grundlagen der Theorie optimaler Prozesse zu entwickeln. Hierzu müssen zunächst die beiden Begriffe „Prozeß“ und „Optimalität“ näher erläutert werden. Dabei wird es sich erweisen, daß weitere Begriffe erforderlich sind. Von diesen seien vorab solche genannt wie „Steuerung“, „Steuerbereich“ und „Phasenraum“.

Unter einem *Prozeß* versteht man im allgemeinen eine Folge sich in der Zeit ändernder Zustände eines Systems. Der Begriff des *Systems* wird in seiner intuitiv verständlichen Form verwendet (vgl. [11]).

Es sei erwähnt, daß ein System dem Charakter seiner wesentlichen Bestandteile nach von physikalischer, technischer, ökonomischer, biologischer, chemischer u.a. Natur sein kann.

Die Prozesse werden in entsprechender Weise unterschieden. Als Beispiele von Prozessen seien hier der Flug einer Rakete, der Lauf einer Turbinenwelle, die Produktion von bestimmten Erzeugnissen sowie die Vorgänge in chemischen Reaktoren genannt; aber auch die Vorgänge in einem Telefonkabel sowie die Veränderungen auf einem Parkplatz, gemessen in freien Parkplätzen, sind Prozesse. Die Vielfalt dieser wenigen Beispiele weist schon darauf hin, daß der Begriff des Prozesses im obigen Sinne sehr allgemein ist. Für unsere Zwecke ist es erforderlich, ihn zu spezifizieren und zu präzisieren.

Prozesse können unter verschiedenen Gesichtspunkten klassifiziert werden. Wir nennen drei Klassifizierungen.

So kann man in Abhängigkeit von den zeitlichen Veränderungen zwischen diskontinuierlichen und kontinuierlichen bzw. — wie man auch sagt — zwischen *diskreten* und *stetigen* Prozessen unterscheiden. Ein Prozeß der ersten Art ist dadurch gekennzeichnet, daß die Zeit nur diskrete (im allgemeinen äquidistante) Werte annimmt, und der Prozeß somit eine Folge endlich vieler oder höchstens abzählbar vieler unterschiedlicher Zustände bzw. getrennter Stufen darstellt. Als Beispiel eines solchen Prozesses kann die Benutzung des Parkplatzes angesehen werden. Demgegenüber durchläuft die Zeit bei einem stetigen Prozeß alle Werte eines Intervalls, so daß er im allgemeinen aus überabzählbar unendlich vielen Zuständen bestehen kann.

Der Flug der Rakete, der Lauf der Turbinenwelle und die Produktion eines Erzeugnisses (z.B. unter den Bedingungen des Dreischichtsystems) sind Beispiele stetiger Prozesse. Die Abgrenzung stetiger Prozesse von diskreten schließt jedoch nicht aus, daß stetige Prozesse auch als diskrete aufgefaßt und dargestellt werden. Das macht der Physiker, wenn er die Schwingungen einer Saite modelliert und dabei ein gewisses Stückchen dieser Saite Δs während einer endlichen Zeitdauer Δt betrachtet; das macht auch der Ökonom, wenn er die Ergebnisse des Produktionsprozesses kumulativ für einen Tag, für eine Dekade oder für einen Monat erfaßt.

Eine zweite Möglichkeit der Klassifizierung ergibt sich, wenn man den Zustand des Prozesses in einem fixierten Augenblick in Abhängigkeit von einem Anfangszustand und den zwischenzeitlichen Einwirkungen auf den Prozeß untersucht. Ist diese Abhängigkeit eindeutig bestimmt, so spricht man von einem *deterministischen* Prozeß. Als Beispiele hierfür können wir den Flug der Rakete sowie den Lauf der Turbinenwelle auffassen. Ist dagegen der Endzustand eines Prozesses durch seinen Anfangszustand und die zwischenzeitlichen Einwirkungen nicht eindeutig bestimmt, son-

dem unterliegt gewissen Wahrscheinlichkeitsverteilungen, so bezeichnet man ihn als *stochastischen Prozeß*.

Wenn hier von Einwirkungen auf den Prozeß gesprochen wurde, so ist damit bereits das dritte Klassifizierungsmerkmal genannt worden. Wir werden einen Prozeß nämlich *steuerbar* nennen, wenn man auf seinen Verlauf einwirken kann und die Ergebnisse dieses Einwirkens kennt. In diesem Sinne ist durchaus nicht jeder Prozeß steuerbar: So kann z. B. der Wetterverlauf gegenwärtig noch nicht zu den steuerbaren Prozessen gerechnet werden.

Einen Hauptgegenstand dieses Buches bilden *steuerbare deterministische Prozesse* und ihre mathematische Untersuchung. Dabei werden im 3. Abschnitt stetige und im 4. Abschnitt diskrete Prozesse betrachtet. Stochastische Prozesse werden im Band 19/1 behandelt.

Wir gehen zur mathematischen Beschreibung unseres Untersuchungsobjektes über. Unter einem *steuerbaren deterministischen Prozeß* – im weiteren kurz Prozeß oder steuerbarer Prozeß genannt – wollen wir einen solchen Prozeß verstehen, dessen Zustand in jedem Augenblick durch eine gewisse Anzahl von Parametern – den sogenannten *Phasenkoordinaten* – eindeutig bestimmt ist. Wir nehmen an, es seien n solcher Phasenkoordinaten erforderlich, und bezeichnen diese mit $x_i(t)$, $i = 1, 2, \dots, n$. Dabei bezeichnet in der Regel t die Zeit. Sie stellt in Abhängigkeit davon, ob der Prozeß stetig oder diskret ist, eine stetige (vgl. 3. Abschn.) oder diskrete Größe (vgl. 4. Abschn.) dar. Die Abhängigkeit der Phasenkoordinaten von der Zeit ist berechtigt und erforderlich, weil sich die Zustände in der Zeit ändern. Weiterhin soll der Prozeß über gewisse Parameter – die sogenannten *Steuerparameter* oder auch *Steuerungen* – verfügen, durch die auf seinen Verlauf Einfluß genommen werden kann. Wir nehmen an, es seien r solcher Steuerungen vorhanden und bezeichnen sie mit $u_j(t)$, $j = 1, 2, \dots, r$. Schließlich wird vorausgesetzt, daß der Prozeß vollständig bestimmt ist, wenn sein Anfangszustand $x_i(t_0)$, $i = 1, 2, \dots, n$, sowie seine Steuerungen $u_j(t)$, $j = 1, 2, \dots, r$, bekannt sind.

Gewöhnlich werden die Phasenkoordinaten $x_i(t)$ als Komponenten eines Vektors $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))$ aufgefaßt. Für jedes feste t stellt der Vektor $\mathbf{x}(t)$ einen Punkt im n -dimensionalen Raum R^n dar. Er wird in diesem Zusammenhang *Phasenraum* genannt. Es ist üblich, jede Kurve in dem Phasenraum als *Trajektorie* zu bezeichnen. Somit beschreibt speziell der Phasenvektor $\mathbf{x}(t)$ eines Prozesses bei stetig veränderlicher Zeit t eine Trajektorie im Raum R^n . Handelt es sich bei dem Prozeß um die Bewegung eines Objektes im dreidimensionalen euklidischen Raum – wie z. B. die der Rakete –, so unterscheiden sich im allgemeinen der Phasenraum und die Trajektorie dieses Prozesses grundsätzlich von dem Bewegungsraum R^3 und der tatsächlichen Bewegungslinie des Objektes. Analog wie bei den Phasenkoordinaten eines Prozesses wird aus seinen Steuerungen der Steuervektor $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_r(t))$ gebildet, der für jedes fixierte t einen Punkt im r -dimensionalen *Steuerraum* R^r darstellt. Die Menge aller der Punkte, die vom Steuervektor eines Prozesses angenommen werden bzw. für ihn zugelassen sind, bilden den *Steuerbereich* $U \subseteq R^r$ des Prozesses. In Abhängigkeit von der Art des Prozesses kann der Steuerbereich eine abgeschlossene oder offene, eine beschränkte oder unbeschränkte Menge sein (vgl. Bd. I). Betrachtet man z. B. bei der Rakete die Beschleunigung als Steuerung und geht von der durchaus realen Annahme aus, daß die Beschleunigung dem Betrage nach eine gewisse maximale Größe zwar erreichen, sie jedoch auf Grund physikalischer und technischer Gegebenheiten nicht überschreiten darf, so erhält man als Steuerbereich eine

beschränkte, abgeschlossene Menge. Es wird sich noch zeigen, daß sich die mathematischen Aufgaben und Probleme für abgeschlossene oder offene Steuerbereiche wesentlich unterscheiden.

Unter einem *Prozeß* werden wir also im weiteren ein Paar verstehen, das aus dem Phasenvektor $\mathbf{x}(t)$ und dem Steuervektor $\mathbf{u}(t)$ gebildet wird, wobei letzterer zusammen mit einem Anfangszustand $\mathbf{x}(t_0)$ den Phasenvektor $\mathbf{x}(t)$ und damit den Prozeß eindeutig festlegt. Wie das erfolgt, wird in Abschnitt 3. und 4. dargelegt (siehe z. B. Prozeßgleichungen (3.6)).

Es ist nun noch der Begriff der *Optimalität* näher zu erläutern. Um hierbei auch grafische Darstellungen benutzen zu können, beschränken wir uns auf stetige Prozesse mit nur zwei Phasenkoordinaten, so daß der Phasenraum die x_1, x_2 -Ebene ist.

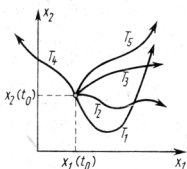


Bild 2.1

Für diese Prozesse ergeben sich bei fixiertem Anfangszustand $\mathbf{x}(t_0) = (x_1(t_0), x_2(t_0))$ in Abhängigkeit von der Wahl des Steuervektors $\mathbf{u}_k(t)$ im allgemeinen unterschiedliche Phasenvektoren $\mathbf{x}_k(t)$ und damit unterschiedliche Trajektorien T_k (s. Bild 2.1). Die Vielzahl der möglichen Trajektorien gestattet es nun, die Frage nach einer „besten“ Trajektorie zu stellen. Das ist eine ähnliche Situation, wie sie uns von der linearen Optimierung her (vgl. Bd. 14) schon bekannt ist. Auch dort gestattete erst die Vielzahl der im Rahmen gegebener Auflagen möglichen Produktionsvarianten oder die Vielzahl der im Rahmen geforderter Nährstoffgehalte möglichen Diätpläne die Auswahl von „besten“ (optimalen) Varianten. Hier wie dort sind für eine solche Auswahl natürlich Kriterien erforderlich. In der linearen Optimierung hingen die Kriterien von der jeweils betrachteten Aufgabe ab und konnten z. B. die maximale Auslastung der Produktionskapazitäten oder den minimalen Preis beinhalten. Analog verhält es sich bei den Prozessen. So können solche Kriterien z. B. für die Rakete der minimale Treibstoffverbrauch oder der kürzeste Weg oder die geringste Zeit beim Flug von einem Punkt zum anderen sein. Für die Turbinenwelle kann das Kriterium darin bestehen, ihren Lauf in kürzester Zeit in vorgegebener Weise zu verändern. Derartige Kriterien werden *Optimalitätskriterien* genannt. Sie gestatten es, aus der Vielzahl der möglichen Trajektorien eines Prozesses diejenige auszuwählen, für die eine charakteristische Prozeßgröße (z. B. Treibstoff, Weg, Zeit o. a.) ihren kleinsten oder größten Wert annimmt. Die so ausgewählte Trajektorie sowie die Steuerung, die sie erzeugt, und der zugehörige Prozeß werden optimal genannt.

3. Stetige Prozesse

Von den steuerbaren, deterministischen Prozessen werden jetzt diejenigen untersucht, die im Abschnitt 2. als stetige bezeichnet worden sind. Dabei sei zunächst bemerkt, daß die Stetigkeit eines Prozesses im obigen Sinne sich nur auf die Zeit bezieht, so daß die für die Zeit stehende Variable t beliebige Werte eines Intervalls annehmen kann, dieses Intervall also stetig durchläuft. Die Stetigkeit des Prozesses bedeutet jedoch keineswegs, daß auch die Prozeßparameter wie Phasenkoordinaten und Steuerungen selbst stetig sind. Im Gegenteil, es zeigt sich sogar, daß für eine Reihe von praktischen Problemen die optimalen Prozesse gerade dadurch gekennzeichnet sind, daß für sie gewisse Parameter, und zwar vorrangig die Steuerungen, Sprungstellen besitzen. Dieser Hinweis ist von prinzipieller Bedeutung für das Verständnis der folgenden Darlegungen.

Im Kapitel 3. werden die Grundaufgabe der Theorie optimaler Prozesse einschließlich ihrer Modifizierungen (siehe Abschn. 3.1., 3.3. und 3.5.) formuliert sowie ein notwendiges Kriterium für die Optimalität entwickelt und Lösungsverfahren dargestellt.

Zunächst muß jedoch etwas über den Zusammenhang zwischen den Steuerungen und den Phasenkoordinaten eines Prozesses gesagt werden. Wir gehen bekanntlich davon aus, daß die Phasenkoordinaten durch die Steuerungen eindeutig bestimmt werden. Also müssen zwischen ihnen gewisse Beziehungen bestehen, deren mathematische Form Gleichungen sind. Beachtet man nun, daß die von uns betrachteten stetigen Prozesse im Prinzip Bewegungen sind, so darf man erwarten, daß die Gleichungen durch die Phasenkoordinaten und Steuerungen verknüpft sind, die Prozeßbewegung beschreiben. Aus der Physik ist aber bekannt, daß Bewegungsgleichungen gewöhnlich gewisse Differentialgleichungen sind (vgl. Band 7.1, Beispiel 1.8). Daher kann erwartet werden, daß die Phasenkoordinaten eines Prozesses mit seinen Steuerungen durch Differentialgleichungen verknüpft sind.

Beispiel 3.1: Wir betrachten — beginnend vom Zeitpunkt t_0 — die geradlinige Bewegung eines Körpers mit der Masse m . Dieser Körper sei mit einem Antrieb versehen, durch den in jedem Augenblick t eine Kraft entwickelt werden kann, die mit $u(t)$ bezeichnet wird. Auf die Gerade, entlang der sich der Körper bewegt, wird ein Nullpunkt und ein Maßstab aufgetragen, so daß der Ort, in dem sich der Körper im Augenblick t befindet, mit $x(t)$ bezeichnet werden kann. Dann sind mit $\dot{x}(t)$ die Geschwindigkeit und mit $\ddot{x}(t)$ die Beschleunigung des Körpers im Augenblick t gegeben. Auf den Körper möge

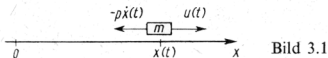


Bild 3.1

neben der vom Antrieb entwickelten Kraft noch die Reibung wirken. Sie ist der Größe nach bekanntlich proportional zur Geschwindigkeit $\dot{x}(t)$ und der Bewegungsrichtung immer entgegengesetzt (siehe Bild 3.1). Bezeichnet man den Proportionalitätsfaktor mit p , so gilt nach dem Gesetz von Newton folgende Bewegungsgleichung für den Körper:

$$m\ddot{x}(t) = u(t) - p\dot{x}(t) \quad (3.1)$$

oder

$$\ddot{x}(t) = a\dot{x}(t) + bu(t) \quad \text{mit} \quad a = -\frac{p}{m}, \quad b = \frac{1}{m}. \quad (3.2)$$

Mit (3.1) bzw. (3.2) haben wir bereits Differentialgleichungen erhalten. Für unsere Zwecke ist es vorteilhaft, die Gleichung (3.2) noch umzuformen. Hierzu werden die Bezeichnungen

$$x_1(t) = x(t), \quad x_2(t) = \dot{x}(t), \quad t \geq t_0 \quad (3.3)$$

eingeführt. Dabei sind diese Bezeichnungen nicht willkürlich, sondern entsprechen der Tatsache, daß die Bewegung des Körpers in jedem Augenblick vollständig beschrieben wird durch den Ort, in dem er sich befindet (entspricht $x_1(t)$), sowie durch seine Geschwindigkeit (entspricht $x_2(t)$). Somit liegt es nahe, $x_1(t)$ und $x_2(t)$ als Phasenkoordinaten des betrachteten Prozesses aufzufassen. Weiterhin kann der Prozeß offensichtlich durch die vom Antrieb entwickelte Kraft $u(t)$ gesteuert werden. Verwendet man nun anstelle von $x(t)$ die durch (3.3) eingeführten Größen, so kann (3.2) ersetzt werden durch

$$\left. \begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= ax_2(t) + bu(t) \end{aligned} \right\} t \geq t_0. \quad (3.4)$$

Damit ist unsere obige Hypothese bestätigt worden: Für den betrachteten Prozeß sind die Phasenkoordinaten und die Steuerung durch Differentialgleichungen miteinander verknüpft.

Das Ergebnis (3.4) ist allgemeingültig für eine ganze Klasse stetiger Prozesse. Ihr charakteristisches Merkmal besteht darin, daß die Steuerungen $u_j(t)$, $j = 1, \dots, r$, mit den Phasenkoordinaten $x_i(t)$, $i = 1, \dots, n$, durch ein System gewöhnlicher Differentialgleichungen erster Ordnung der Art

$$\frac{dx_i(t)}{dt} = f_i(x_1(t), \dots, x_n(t), u_1(t), \dots, u_r(t)), \quad i = 1, \dots, n, \quad (3.5)$$

oder kürzer

$$\dot{x}_i(t) = f_i(\mathbf{x}(t), \mathbf{u}(t)), \quad i = 1, \dots, n,$$

bzw.

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad (3.6)$$

verknüpft sind. Es ist bekannt (vgl. Band 7.1), daß (3.6) bei gegebenen Funktionen $u_j(t)$, $j = 1, 2, \dots, r$, bezüglich $\mathbf{x}(t)$ lösbar ist und zusammen mit den Anfangswerten x_i^0 , $i = 1, 2, \dots, n$, die Phasenkoordinaten $x_i(t)$, $i = 1, 2, \dots, n$, eindeutig bestimmt, wenn die Funktionen f_i die entsprechenden Eigenschaften besitzen. Daher werden wir in (3.6) die *Prozeß-* oder auch *Bewegungsgleichung* des Prozesses nennen.

Im weiteren betrachten wir solche Prozesse, deren Bewegungsgleichung die Form (3.6) hat, wobei vorausgesetzt wird, daß die Funktionen f_i , $i = 1, 2, \dots, n$, auf $R^n \times U$ definiert und dort einschließlich ihrer partiellen Ableitungen erster Ordnung nach x_k , $k = 1, 2, \dots, n$, stetig sind. Die Funktionen f_i ergeben sich aus der Spezifik des jeweiligen Prozesses. Man erhält sie im Ergebnis der mathematischen Modellierung des realen Prozesses. Für das Beispiel 3.1 lauten diese Funktionen (vgl. (3.4))

$$f_1(\mathbf{x}(t), \mathbf{u}(t)) = x_2(t), \quad f_2(\mathbf{x}(t), \mathbf{u}(t)) = ax_2(t) + bu(t).$$

3.1. Grundaufgabe der Optimierung

Das mathematische Modell der betrachteten Prozesse umfaßt bisher das Differentialgleichungssystem (3.6). Es wurde schon bemerkt, daß dieses System für verschiedene Steuervektoren $\mathbf{u}(t)$ im allgemeinen unterschiedliche Phasenvektoren $\mathbf{x}(t)$ als Lösungen besitzt und damit unterschiedliche Trajektorien des Prozesses liefert

(vgl. Bild 2.1). Die Vielzahl der Trajektorien erlaubt die Frage nach einer optimalen Steuerung und der zugehörigen optimalen Trajektorie. Dabei muß jetzt näher erläutert werden, was als Optimalitätskriterium gewählt wird. In der linearen Optimierung wird das Optimalitätskriterium durch eine lineare Zielfunktion dargestellt. Für stetige Prozesse nimmt das Optimalitätskriterium eine allgemeinere Form an und wird durch ein Funktional J ausgedrückt, dessen Definitionsbereich die Steuer- und Phasenvektoren des Prozesses sind. Der Wertebereich eines Funktional ist bekanntlich eine Teilmenge der reellen Zahlen, so daß man nach dem kleinsten oder größten Wert des Funktional fragen kann. In der Theorie der optimalen Prozesse ist es üblich, die Aufgaben als Minimumprobleme zu formulieren (vgl. [6], [11], [13], [14], [19]). Die Ursache hierfür ist in den ursprünglichen physikalischen und technischen Problemen zu sehen, die zur Theorie optimaler Prozesse führten. Das Wesen dieser Probleme bestand darin, einen Prozeß durch Energiezufuhr so zu steuern, daß er einen im voraus festgelegten Verlauf nimmt oder diesem Verlauf möglichst nahe kommt und dabei der Energieverbrauch minimal ist.

Im weiteren werden wir uns auf die Betrachtung solcher Prozesse beschränken, für die das Funktional des Optimalitätskriteriums die Form eines bestimmten Integrals

$$J(\mathbf{u}) = \int_{t_0}^{t_1} f_0(\mathbf{x}(t), \mathbf{u}(t)) dt \quad (3.7)$$

(Aufgabe von Lagrange) annimmt, wobei f_0 wiederum eine bekannte Funktion ist, die sich aus der Spezifik des jeweiligen Prozesses ergibt, und von der vorausgesetzt wird, daß sie die gleichen Stetigkeits- und Differenzierbarkeitseigenschaften wie die Funktionen f_i der Bewegungsgleichung (3.6) besitzt.

Beispiel 3.2: Wir betrachten erneut den in Beispiel 3.1 eingeführten Prozeß der geradlinigen Bewegung eines Körpers. Hier ist es sinnvoll, die Frage zu stellen, wie die Bewegung aus einem Zustand $\mathbf{x}_0 = (x_1^0, x_2^0)$ in einen anderen Zustand $\mathbf{x}_1 = (x_1^1, x_2^1)$ zu vollziehen ist, daß dabei der Treibstoffverbrauch möglichst gering wird. In grober Näherung kann man annehmen, daß der Treibstoffverbrauch im Intervall $(t, t + \Delta t)$ proportional zu der Größe $u^2(t)\Delta t$ ist. Dann ergibt sich als Maß für den Treibstoffverbrauch

$$J(u) = \int_{t_0}^{t_1} cu^2(t) dt,$$

wobei c ein gewisser Proportionalitätsfaktor ist und t_i die Zeitpunkte sind, in denen der Körper die Zustände \mathbf{x}_i , $i = 0, 1$, durchläuft. Somit hängt in diesem Beispiel die Funktion f_0 aus (3.7) gar nicht vom Phasenvektor $\mathbf{x}(t)$ ab und hat die konkrete Form

$$f_0(\mathbf{x}(t), \mathbf{u}(t)) = cu^2(t).$$

Bevor nun die Grundaufgabe formuliert wird, muß noch der für die Theorie optimaler Prozesse fundamentale Begriff der zulässigen Steuerung erklärt werden.

Definition 3.1: Es sei ein Prozeß mit fixiertem Steuerbereich $U \subseteq R^r$ gegeben. Dann wird $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_r(t))$ ein **zulässiger Steuervektor** (bezüglich U) genannt, wenn ein Intervall $[t_0, t_1]$ derart existiert, daß gilt:

1. $\mathbf{u}(t) \in U$ für alle $t \in [t_0, t_1]$,

2. alle Komponenten $u_j(t)$, $j = 1, 2, \dots, r$, und damit $\mathbf{u}(t)$ selbst, sind stückweise stetige Funktionen in $[t_0, t_1]$, wobei sie in den Sprungstellen $\tau \in (t_0, t_1)$ linksseitig stetig sein müssen

$$\lim_{t \rightarrow \tau - 0} u_j(t) = u_j(\tau), \quad j = 1, 2, \dots, r. \quad (3.8)$$

In manchen Publikationen wird der Begriff des zulässigen Steuervektors noch allgemeiner gefaßt (vgl. [20]). Aber schon mit der von uns angegebenen Definition des zulässigen Steuervektors sind einige Probleme verbunden, die einer Erläuterung bedürfen.

Erinnern wir uns daran, daß die Steuerungen $u_j(t)$ als solche Parameter eines Prozesses eingeführt worden sind, durch die man auf ihn Einfluß nehmen kann, so besagt die 2. Forderung der obigen Definition, daß die Steuerungen Sprungstellen haben können. Hier ergeben sich natürlich sofort zweierlei Einwände. Einerseits kann man fragen, ob denn nicht stetige Steuerungen die Realität besser widerspiegeln, und andererseits ergibt sich die Frage, wie die Sprünge der Steuerparameter praktisch realisiert werden können.

Der erste Einwand kann durch Hinweis auf entsprechende Ergebnisse entkräftet werden. Es zeigt sich, daß für eine umfangreiche Menge von Prozessen die optimalen Lösungen gerade Steuerungen mit Sprungstellen sind. Hierzu gehören die sogenannten schnelligkeitsoptimalen Prozesse (siehe unten). Ein einfacher Vertreter hiervon ergibt sich als Spezialfall des Prozesses aus Beispiel 3.1. Setzt man dort voraus, daß $p = 0$ ist und der Steuerparameter der Bedingung $|u(t)| \leq 1$ genügt, so hat die Aufgabe, den Körper aus einem Zustand $\mathbf{x}_0 = (x_1^0, x_2^0)$ in einen anderen Zustand $\mathbf{x}_1 = (x_1^1, x_2^1)$ in der kürzesten Zeit zu überführen, eine optimale Lösung. Dabei erweist sich die optimale Steuerung als eine stückweise konstante Funktion, die nur die Werte ± 1 annimmt und eine Sprungstelle besitzt.

Der zweite Einwand ist schon schwerwiegender und hat Einfluß auf die Modellierung. Steuerungen mit Sprungstellen bedeuten für die Realität nämlich, daß die entsprechenden Steuerparameter trägheitslos sind. Die Parameter eines realen Prozesses besitzen aber im allgemeinen eine gewisse Trägheit. Deshalb kommt es darauf an, die Steuerparameter so zu wählen, daß sie zwei Bedingungen erfüllen: Der Prozeß muß durch sie gesteuert werden können, und ihre Trägheit muß so gering sein, daß sie im Modell vernachlässigt werden kann. Die Erfahrungen lehren, daß das immer möglich ist (vgl. [6]). Betrachtet man z. B. den Winkel $\alpha(t)$, den das Ruder eines Schiffes mit der Längsachse des Schiffes bildet, so kann das Schiff zwar mit dieser Größe gesteuert werden, als Steuerparameter ist $\alpha(t)$ jedoch ungeeignet. Die Trägheit der Änderung des Winkels $\alpha(t)$ ist nämlich relativ groß, weil das Ruder den Wasserwiderstand überwinden muß. Dagegen ist die Trägheit der Änderung der Winkelgeschwindigkeit $\dot{\alpha}(t)$ schon geringer und die von $\ddot{\alpha}(t)$ kann praktisch vernachlässigt werden. Es würde sich also empfehlen, nicht $\alpha(t)$, sondern $\ddot{\alpha}(t)$ als Steuerung zu wählen.

Zu dem Intervall $[t_0, t_1]$ muß erläuternd bemerkt werden, daß bei gegebenem Steuerbereich U zu zwei verschiedenen zulässigen Steuervektoren $\mathbf{u}(t)$ und $\hat{\mathbf{u}}(t)$ im allgemeinen auch unterschiedliche Zeitintervalle $[t_0, t_1]$ bzw. $[t_0, \hat{t}_1]$ gehören werden. So sind z. B. im Fall der geradlinigen Bewegung des Körpers (vgl. Beispiele 3.1 und 3.2) für $U = \{u: -1 \leq u \leq 1\}$ die Funktionen $u(t) = \cos t$, $t \in [0, t_1]$, mit beliebig fixiertem t_1 , und $u(t) = \frac{1}{2}t - 3$, $t \in [4, 8]$, zulässige Steuerungen, deren zugehörige Zeitintervalle $[0, t_1]$ bzw. $[4, 8]$ unterschiedlich sind. Somit hat das Integral in (3.7) im allgemeinen für verschiedene Steuervektoren $\mathbf{u}(t)$ auch unterschiedliche Integrationsgrenzen.

Schließlich führt die 2. Forderung aus der Definition 3.1 noch zu einer mathematischen Problematik. Wird nämlich in die rechte Seite der Bewegungsgleichung (3.6) eines Prozesses ein zulässiger Steuervektor $\mathbf{u}(t)$ mit einer Sprungstelle $\tau \in (t_0, t_1)$ eingesetzt, so sind i. a. auch die Funktionen $\varphi_i(t) = f_i(\mathbf{x}(t), \mathbf{u}(t))$ nur noch stückweise stetig. Damit sind aber die Bedingungen des Existenz- und Eindeutigkeitssatzes (vgl.

Bd. 7/1) nur noch in den Teilintervallen $[t_0, \tau]$ und $[\tau, t_1]$, jedoch nicht mehr in $[t_0, t_1]$ erfüllt, und für (3.6) kann in $[t_0, t_1]$ die Existenz einer stetig differenzierbaren Lösung $\mathbf{x}(t)$ nicht mehr garantiert werden. Diese Schwierigkeit wird behoben, indem man (3.6) zunächst für das Intervall $[t_0, \tau]$ löst und den dabei erhaltenen Vektor $\mathbf{x}(\tau)$ als Anfangszustand für die Lösung von (3.6) in $[\tau, t_1]$ benutzt. Analog wird bei mehreren Sprungstellen verfahren. Damit ist gesichert, daß (3.6) bei gegebenen Anfangswerten $x_i^0, i = 1, 2, \dots, n$, für jeden zulässigen Steuervektor $\mathbf{u}(t)$ eindeutig lösbar ist.

Definition 3.2: Es sei $\mathbf{x}(t)$ die Lösung von (3.6), die man bei gegebenen Anfangswerten $x_i^0, i = 1, 2, \dots, n$, und fixiertem zulässigem Steuervektor $\mathbf{u}(t)$ mit $t \in [t_0, t_1]$ erhalten hat. Dann werden wir $\mathbf{x}(t)$ die dem zulässigen Steuervektor $\mathbf{u}(t)$ entsprechende **Trajektorie** des Prozesses nennen und sagen, daß $\mathbf{u}(t)$ den Phasenpunkt $\mathbf{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)$ in den Phasenpunkt $\mathbf{x}_1 = (x_1^1, x_2^1, \dots, x_n^1)$ mit $x_i^1 = x_i(t_1)$ überführt.

Existieren nun für zwei fixierte Phasenpunkte \mathbf{x}_0 und \mathbf{x}_1 mehrere Steuerungen, die \mathbf{x}_0 in \mathbf{x}_1 überführen, dann ergibt sich sofort die Frage nach einer „besten“ oder „optimalen“ Steuerung. Mathematisch wird diese Frage in Form der folgenden **Grundaufgabe** zur Bestimmung optimaler Prozesse formuliert.

Definition 3.3: Gegeben sei ein Prozeß mit der Bewegungsgleichung (3.6), dem Steuerbereich $U \subseteq R^r$ und mit einer Funktion $f_0(\mathbf{x}, \mathbf{u})$ die – genau wie die Funktionen f_i in (3.6) – für alle $\mathbf{x} = (x_1, x_2, \dots, x_n)$ des Phasenraumes R^n sowie für alle $\mathbf{u} \in U$ definiert und einschließlich ihrer partiellen Ableitungen erster Ordnung nach $x_i, i = 1, 2, \dots, n$, stetig ist. Dann ist für zwei vorgegebene Phasenpunkte $\mathbf{x}_0, \mathbf{x}_1 \in R^n$ unter allen zulässigen Steuervektoren $\mathbf{u}(t)$ mit $t \in [t_0, t_1]$, die \mathbf{x}_0 in \mathbf{x}_1 überführen, derjenige zu bestimmen, für den das Funktional

$$J(\mathbf{u}) = \int_{t_0}^{t_1} f_0(\mathbf{x}(t), \mathbf{u}(t)) dt \quad (3.7)$$

den kleinsten Wert annimmt. Hierbei ist $\mathbf{x}(t)$ die Trajektorie, die dem Steuervektor $\mathbf{u}(t)$ entspricht und den Bedingungen $\mathbf{x}(t_0) = \mathbf{x}_0$ sowie $\mathbf{x}(t_1) = \mathbf{x}_1$ genügt.

Definition 3.4: Jede Lösung $\mathbf{u}(t)$ der Grundaufgabe wird **optimaler Steuervektor** von \mathbf{x}_0 nach \mathbf{x}_1 genannt; die ihm entsprechende Trajektorie $\mathbf{x}(t)$ und der aus beiden gebildete Prozeß heißen ebenfalls **optimal**.

Somit besteht die Grundaufgabe optimaler Prozesse in der Bestimmung optimaler Steuervektoren und der ihnen entsprechenden optimalen Trajektorien. Es sei noch besonders bemerkt, daß in der Grundaufgabe t_0 und t_1 nicht gegeben sind und sich erst mit der Bestimmung der optimalen Steuerung ergeben.

In Abhängigkeit von der Art des Steuerbereiches U sowie den Eigenschaften der Funktionen $f_i, i = 0, 1, \dots, n$, besitzt die Grundaufgabe unterschiedlichen mathematischen Schwierigkeitsgrad. Ist z. B. U eine offene Teilmenge¹⁾ von R^r oder $U = R^r$, so kann gezeigt werden, daß die Grundaufgabe in diesem Falle äquivalent zu Aufgaben der Variationsrechnung ist (vgl. [20]). Damit steht für diese Teilklasse von Grundaufgaben optimaler Prozesse das umfangreiche Arsenal bekannter Aussagen

¹⁾ Beispiele hierfür wären (vgl. Bd. 1): $(-\infty, 0)$ Teilmenge des R^1 ; $(-1, +1)$ Teilmenge des R^1 ; $\{(x_1, x_2) \in R^2 \mid x_1^2 + x_2^2 < 4\}$ Teilmenge des R^2 u. ä.

und numerischer Verfahren der Variationsrechnung zur Verfügung (vgl. [22], [16]). Ist der Steuerbereich U dagegen eine abgeschlossene Teilmenge von R^r , so versagt die Variationsrechnung, und die Grundaufgabe ist mit prinzipiell neuen Schwierigkeiten verbunden. Aber auch innerhalb dieser Klasse von Grundaufgaben gibt es wiederum solche, die sich durch relative Einfachheit auszeichnen. Dazu gehören die sogenannten *schnelligkeitsoptimalen* Prozesse, die man als Spezialfall der Grundaufgabe bei

$$f_0(\mathbf{x}, \mathbf{u}) = 1$$

erhält. Für sie lautet das Funktional (3.7) einfach $J(\mathbf{u}) = t_1 - t_0$, so daß das Wesen dieser Prozesse darin besteht, den Steuervektor $\mathbf{u}(t)$ zu bestimmen, der einen Phasenpunkt \mathbf{x}_0 in kürzester Zeit in einen anderen gegebenen Punkt \mathbf{x}_1 überführt. Sind darüber hinaus die rechten Seiten der Bewegungsgleichung linear in allen Argumenten d.h., gilt auch noch

$$f_i(\mathbf{x}, \mathbf{u}) = \sum_{j=1}^n a_{ij} x_j + \sum_{k=1}^r b_{ik} u_k, \quad i = 1, 2, \dots, n,$$

so ergibt sich bei gewissen zusätzlichen Bedingungen an U die Teilklasse der sogenannten linearen schnelligkeitsoptimalen Prozesse, für die die umfassendsten Aussagen bekannt sind (vgl. [6]).

Optimale Steuervektoren und die ihnen entsprechenden optimalen Trajektorien besitzen einige allgemeine Eigenschaften, von denen wir hier die beiden folgenden nennen:

1. *Die Eigenschaft eines zulässigen Steuervektors, optimal zu sein, ist invariant gegenüber Parallelverschiebungen entlang der t -Achse.* Mit anderen Worten: Ist $\mathbf{u}(t)$ für $t \in [t_0, t_1]$ optimaler Steuervektor, dann ist es auch der Steuervektor $\hat{\mathbf{u}}(t) = \mathbf{u}(t + h)$ für $t \in [t_0, \hat{t}_1]$ mit $\hat{t}_1 = t_1 + h$. Das beruht darauf, daß keine der Funktionen f_i , $i = 0, 1, \dots, n$, explizit von t abhängt, und gestattet es, den Anfangspunkt t_0 des Zeitintervalls für alle zulässigen Steuerungen einheitlich zu wählen.

2. *Jeder Abschnitt einer optimalen Trajektorie ist selbst wieder optimale Trajektorie.* Ist also $\mathbf{u}^*(t)$ mit $t \in [t_0, t_1]$ optimaler Steuervektor, der \mathbf{x}_0 in \mathbf{x}_1 überführt, und $\mathbf{x}^*(t)$ die ihm entsprechende optimale Trajektorie, so gilt für zwei beliebig fixierte Punkte $\mathbf{x}^*(\tau_0)$ und $\mathbf{x}^*(\tau_1)$, $t_0 \leq \tau_0 \leq \tau_1 \leq t_1$, der optimalen Trajektorie folgendes: $\mathbf{u}^*(t)$ ist für $t \in [\tau_0, \tau_1]$ optimaler Steuervektor, der $\mathbf{x}^*(\tau_0)$ in $\mathbf{x}^*(\tau_1)$ überführt, und $\mathbf{x}^*(t)$, $t \in [\tau_0, \tau_1]$, ist die ihm entsprechende optimale Trajektorie. Diese Eigenschaft ist ein gewisses Analogon zum Bellmanschen Optimalprinzip (vgl. Abschn. 4.3.).

3.2. Pontrjaginsches Maximumprinzip

Es wurde schon erwähnt, daß die Methoden der Variationsrechnung für Grundaufgaben mit abgeschlossenen Mengen als Steuerbereich nicht angewendet werden können. Daher muß für sie ein neuer Lösungsweg entwickelt werden. Die Basis hierfür ist das *Pontrjaginsche Maximumprinzip*¹⁾ (vgl. [20] und [5]). Seine Rolle und Be-

¹⁾ Hier ist die Bezeichnung „Prinzip“ historisch (vgl. [5]) und im Sinn eines Ausgangspunktes zu sehen. Keinesfalls darf angenommen werden, daß mit ihm nur ein Grundsatz des Handels formuliert wird, der nicht bewiesen ist. Im Gegenteil, das Maximumprinzip ist streng bewiesen (vgl. [20]).

deutung für die Theorie optimaler Prozesse kann verglichen werden mit den notwendigen Bedingungen für das Extremum von differenzierbaren Funktionen. Letztere liefern bekanntlich zwar die Extrema nicht direkt, gestatten jedoch, einerseits zu überprüfen, ob die Funktion in einem Punkt ein Extremum besitzen kann und andererseits die extremwertverdächtigen Punkte zu ermitteln. Analog hierzu besitzt auch das Pontrjaginsche Maximumprinzip eine zweifache Bedeutung. Erstens kann mit seiner Hilfe geprüft werden, ob eine fixierte zulässige Steuerung sowie die ihr entsprechende Trajektorie optimal sein kann. Zweitens stellt es einen Ausgangspunkt für die Ermittlung der optimalitätsverdächtigen Steuerungen dar.

Zur Formulierung des Maximumprinzips benötigen wir noch einige Begriffe und Symbole, die zunächst eingeführt werden.

Durch die Differentialgleichung

$$\dot{x}_0(t) = f_0(\mathbf{x}(t), \mathbf{u}(t)) \quad (3.9)$$

wird eine neue Phasenkoordinate $x_0(t)$ eingeführt. Neben der Bewegungsgleichung des Prozesses in der Form (3.6) werden wir das um die Gleichung (3.9) erweiterte System

$$\dot{x}_i(t) = f_i(\mathbf{x}(t), \mathbf{u}(t)), \quad i = 0, 1, \dots, n, \quad (3.10)$$

betrachten. Dabei gilt für jede fixierte zulässige Steuerung $\mathbf{u}(t)$, $t \in [t_0, t_1]$, und einer ihr entsprechenden Trajektorie $\mathbf{x}(t)$ wegen (3.9) und (3.7) immer

$$x_0(t_0) = 0, \quad x_0(t_1) = J(\mathbf{u}). \quad (3.11)$$

Weiterhin benötigen wir Hilfsvariablen $p_i(t)$, $i = 0, 1, \dots, n$, die durch das Differentialgleichungssystem

$$\dot{p}_i(t) = - \sum_{k=0}^n \frac{\partial f_k(\mathbf{x}, \mathbf{u})}{\partial x_i} p_k(t), \quad i = 0, 1, \dots, n, \quad (3.12)$$

definiert werden. Das System (3.12) wird als das zum System (3.6) *adjungierte System* bezeichnet. Man überzeugt sich leicht davon, daß (3.12) ein lineares Differentialgleichungssystem in $p_i(t)$ ist, dessen Koeffizienten für jeden fixierten zulässigen Steuervektor $\mathbf{u}(t)$, $t \in [t_0, t_1]$, und für die ihm entsprechende Trajektorie $\mathbf{x}(t)$ stückweise stetige Funktionen der Zeit t sind; daher besitzt (3.12) für beliebige Anfangswerte $p_i(t_0) = p_i^0$, $i = 0, 1, \dots, n$, eine eindeutige Lösung

$$\mathbf{P}(t) = (p_0(t), p_1(t), \dots, p_n(t)), \quad (3.13)$$

die für alle $t \in [t_0, t_1]$ definiert und stetig ist. Natürlich hängt $\mathbf{P}(t)$ von dem Steuervektor $\mathbf{u}(t)$ und der Trajektorie $\mathbf{x}(t)$ ab. Jeden auf diese Weise ermittelten Vektor $\mathbf{P}(t)$ werden wir Lösung von (3.12) nennen, die dem zulässigen Steuervektor $\mathbf{u}(t)$ und einer seiner Trajektorien $\mathbf{x}(t)$ entspricht.

Schließlich führen wir durch die Definitionsgleichung

$$\mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}) = \sum_{i=0}^n p_i f_i(\mathbf{x}, \mathbf{u}) \quad (3.14)$$

nach die sogenannte *Hamiltonfunktion* ein. Mit ihr lassen sich die Systeme (3.10) und (3.12) zu dem folgenden *Hamiltonsystem* vereinigen:

$$\dot{x}_i(t) = \frac{\partial \mathcal{H}}{\partial p_i}, \quad i = 0, 1, \dots, n, \quad (3.15)$$

$$\dot{p}_i(t) = -\frac{\partial \mathcal{H}}{\partial x_i}, \quad i = 0, 1, \dots, n. \quad (3.16)$$

Darüber hinaus erweist sich die Hamiltonfunktion für die Theorie optimaler Prozesse und insbesondere für die Formulierung der notwendigen Bedingungen in folgendem Sinne von fundamentaler Bedeutung. Setzt man für \mathbf{P} und \mathbf{x} in (3.14) fixierte Vektoren $\tilde{\mathbf{P}}$ und $\tilde{\mathbf{x}}$ ein, so ergibt sich eine Funktion, die nur noch von \mathbf{u} abhängt. Für sie kann man die Frage nach dem Supremum über $\mathbf{u} \in U$ stellen. Dieses Supremum wird mit $\mathcal{M}(\tilde{\mathbf{P}}, \tilde{\mathbf{x}})$ bezeichnet:

$$\mathcal{M}(\tilde{\mathbf{P}}, \tilde{\mathbf{x}}) = \sup_{\mathbf{u} \in U} \mathcal{H}(\tilde{\mathbf{P}}, \tilde{\mathbf{x}}, \mathbf{u}). \quad (3.17)$$

Wird das Supremum für ein gewisses $\mathbf{u} \in U$ angenommen, so gibt $\mathcal{M}(\tilde{\mathbf{P}}, \tilde{\mathbf{x}})$ einfach das Maximum von \mathcal{H} bei fixierten $\tilde{\mathbf{P}}$ und $\tilde{\mathbf{x}}$ an. Das ist übrigens auf Grund der Konstruktion von \mathcal{H} und der Stetigkeitseigenschaften von f_i immer dann der Fall, wenn U eine abgeschlossene Menge ist. Deshalb wird das folgende Kriterium über die notwendigen Bedingungen auch Pontrjaginsches Maximumprinzip genannt (vgl. [20]).

Satz 3.1 (notwendige Optimalitätsbedingung): *Dafür, daß ein zulässiger Steuervektor $\mathbf{u}(t)$, $t \in [t_0, t_1]$, der \mathbf{x}_0 in \mathbf{x}_1 überführt, und die ihm entsprechende Trajektorie $\mathbf{x}(t)$ optimal (im Sinne der Definition 3.4) sind, ist notwendig, daß eine ihnen entsprechende stetige Lösung $\mathbf{P}(t) = (p_0(t), p_1(t), \dots, p_n(t))$ von (3.12) mit folgenden Eigenschaften existiert:*

1. Die Lösung $\mathbf{P}(t)$ besitzt wenigstens eine nicht identisch verschwindende Komponente, d. h. $\mathbf{P}(t) \neq \mathbf{0}$ für $t \in [t_0, t_1]$;

2. Für beliebig fixiertes $t \in [t_0, t_1]$ nimmt $\mathcal{H}(\mathbf{P}(t), \mathbf{x}(t), \mathbf{u})$ als Funktion von $\mathbf{u} \in U$ in $\mathbf{u} = \mathbf{u}(t)$ ihr Maximum an

$$\mathcal{H}(\mathbf{P}(t), \mathbf{x}(t), \mathbf{u}(t)) = \mathcal{M}(\mathbf{P}(t), \mathbf{x}(t)); \quad (3.18)$$

3. Im Endzeitpunkt t_1 gelten die Beziehungen

$$p_0(t_1) \leq 0 \quad \text{und} \quad \mathcal{M}(\mathbf{P}(t_1), \mathbf{x}(t_1)) = 0. \quad (3.19)$$

Weiterhin gilt: erfüllen $\mathbf{P}(t)$, $\mathbf{x}_0(t)$, $\mathbf{x}(t)$ und $\mathbf{u}(t)$ das Hamiltonsystem (3.15), (3.16) sowie die obige zweite Bedingung, dann sind $p_0(t)$ und $\mathcal{M}(\mathbf{P}(t), \mathbf{x}(t))$ als Funktionen der Zeit konstant, so daß es genügt, die Bedingungen (3.19) für ein beliebiges $t \in [t_0, t_1]$ nachzuprüfen.

Dieses Prinzip hat eine umfangreiche Verbreitung gefunden. Von den vielen Veröffentlichungen, die seiner Anwendung und theoretischen Weiterführung gewidmet sind, seien hier nur einige genannt. In [12] und vor allem in [2] sind zahlreiche praktische Beispiele aus der Technik und Physik behandelt und bis zur numerischen Lösung

entwickelt worden. Eine Anwendung auf chemische Prozesse zeigt [13]. Ansätze zu möglichen Anwendungen des Maximumprinzips in der Ökonomie findet man in [1], [3], [21] und [23]. In [17] ist eine umfassende Zusammenfassung numerischer Verfahren gegeben.

Um dem Leser eine Vorstellung davon zu vermitteln, wie man das Maximumprinzip anwenden kann, betrachten wir das folgende Beispiel.

Beispiel 3.3: Es wird die Grundaufgabe untersucht, bei der $U = \mathbb{R}^2$, $\mathbf{x}_0 = (2, 1)$, $\mathbf{x}_1 = (6, 4)$ ist, die Bewegungsgleichung die Gestalt

$$\dot{x}_i = u_i, \quad i = 1, 2, \quad (3.20)$$

hat und das Funktional

$$J(u) = \int_{t_0}^{t_1} \sqrt{u_1^2 + u_2^2} \, dt$$

gegeben ist. Mit anderen Worten, es ist der zulässige Steuervektor $\mathbf{u}(t) = (u_1(t), u_2(t))$, $t \in [t_0, t_1]$ zu bestimmen, für den das Funktional $J(u)$ sein Minimum annimmt und außerdem die zugehörige Lösung $\mathbf{x}(t) = (x_1(t), x_2(t))$ von (3.20) den Phasenpunkt \mathbf{x}_0 in \mathbf{x}_1 überführt, d. h., es muß $\mathbf{x}(t_i) = \mathbf{x}_i$, $i = 0, 1$, gelten. Dabei sind t_0 und t_1 ebenfalls nicht gegebene, sondern gesuchte Größen.

Zur Lösung dieser Aufgabe wenden wir das Maximumprinzip an und bilden zunächst die Hamiltonfunktion

$$\mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}) = p_0 \sqrt{u_1^2 + u_2^2} + p_1 u_1 + p_2 u_2.$$

Somit ergibt sich für die Definition der Hilfsfunktionen $p_i(t)$ das System (vgl. (3.16)):

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} = 0, \quad i = 0, 1, 2.$$

Hieraus folgt sofort, daß die Hilfsfunktionen konstant sind:

$$p_i(t) = C_i,$$

wobei die C_i , $i = 0, 1, 2$, zunächst noch frei wählbare Integrationskonstanten sind. Die Hamiltonfunktion nimmt also die Form

$$\mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}) = C_0 \sqrt{u_1^2 + u_2^2} + C_1 u_1 + C_2 u_2$$

an. Nun suchen wir (vgl. mit der zweiten Bedingung in Satz 3.1) das Maximum dieser Funktion bezüglich \mathbf{u} .

Aus den für das Maximum von \mathcal{H} notwendigen Bedingungen

$$\frac{\partial \mathcal{H}}{\partial u_i} = \frac{C_0 u_i}{\sqrt{u_1^2 + u_2^2}} + C_i = 0, \quad i = 1, 2,$$

folgt sofort, daß $C_0 \neq 0$ sein muß, denn sonst würde auch $C_1 = C_2 = 0$ folgen, und das widerspricht dem Maximumprinzip (siehe erste Bedingung). Da außerdem nach dem Maximumprinzip $p_0(t_i) \leq 0$ gelten soll, erfüllen wir sicher alle Forderungen, wenn wir $C_0 = -1$ setzen. Somit ergibt sich zunächst

$$\mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}) = -\sqrt{u_1^2 + u_2^2} + C_1 u_1 + C_2 u_2 \quad (3.21)$$

mit den notwendigen Maximumbedingungen

$$u_i^* = C_i \sqrt{u_1^{*2} + u_2^{*2}}, \quad i = 1, 2, \quad (3.22)$$

aus denen sich die Relation

$$\frac{u_2^*}{u_1^*} = \frac{C_2}{C_1} \quad (3.23)$$

ergibt. Nehmen wir nun einmal an, wir hätten u_i^* bereits explizit ermittelt und setzen diese Werte in (3.21) ein. Dann ergibt sich unter Beachtung von (3.19) und (3.22) die Relation

$$0 = \mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}^*) = \sqrt{u_1^{*2} + u_2^{*2}} (-1 + C_1^2 + C_2^2) \quad \text{oder} \quad C_1^2 + C_2^2 = 1. \quad (3.24)$$

Eine weitere Bedingung für die C_i ergibt sich, wenn wir die u_i^* in die Bewegungsgleichung (3.20) einsetzen und dabei (3.23) sowie die beiden Phasenpunkte \mathbf{x}_i , $i = 0, 1$, benutzen. Wir erhalten nämlich

$$x_1(t) = 2 + \int_{t_0}^t u_1^* d\tau \quad \text{sowie} \quad x_2(t) = 1 + \frac{C_2}{C_1} \int_{t_0}^t u_1^* d\tau, \quad (3.25)$$

woraus speziell für $t = t_1$ nach entsprechender Division folgt

$$\frac{C_2}{C_1} = \frac{x_2(t_1) - 1}{x_1(t_1) - 2} = \frac{4 - 1}{6 - 2} = \frac{3}{4}. \quad (3.26)$$

Man beachte, daß bei diesen Rechnungen u_i^* sowie t_0, t_1 noch nicht bekannt waren. Aus (3.24) und (3.26) folgt nun als eine Möglichkeit

$$C_1 = \frac{4}{5}, \quad C_2 = \frac{3}{5}. \quad (3.27)$$

Somit nimmt die Hamilton-Funktion die Form

$$\mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}) = -\sqrt{u_1^2 + u_2^2} + \frac{4}{5}u_1 + \frac{3}{5}u_2$$

an. Man kann nun durch einige elementare Abschätzungen zeigen, daß diese Funktion immer kleiner oder gleich null bleibt. Daher liefern die Lösungen $\mathbf{u}^* = (u_1^*, u_2^*)$ von (3.22) nicht nur die für relative Extrema der Funktion \mathcal{H} verdächtigen Punkte, sondern (beachte hierzu (3.24)) auch deren absolutes Maximum. Hiernach kann bereits festgestellt werden: wenn eine zulässige Steuerung $\mathbf{u}(t) = (u_1(t), u_2(t))$ nicht der Bedingung (vgl. (3.23))

$$\frac{u_2(t)}{u_1(t)} = \frac{3}{4}$$

genügt, dann kann sie nicht optimal sein. Zur endgültigen Bestimmung des optimalen Steuervektors $\mathbf{u}^*(t)$ einschließlich seines Definitionsbereiches $[t_0, t_1]$ bemerken wir zunächst, daß im gegebenen Falle $\mathbf{u}^*(t)$ nicht eindeutig bestimmt ist. Wählt man z. B. für $u_1^*(t)$ eine beliebige positive stetige Funktion und setzt dann $u_2^*(t) = \frac{3}{4}u_1^*(t)$, so erfüllen diese beiden Steuerparameter die Bedingungen (3.22) und (3.23), wobei die C_i gemäß (3.27) gewählt sind, und erteilen daher der Funktion \mathcal{H} ihren maximalen Wert 0. t_0 kann man wegen der oben erwähnten ersten Eigenschaft der Steuervektoren beliebig wählen (z. B. $t_0 = 0$) und t_1 muß dann der Relation

$$4 = \int_0^{t_1} u_1^*(t) dt$$

genügen. Der einfachste Fall ergibt sich, wenn man $u_1^*(t) = u_1^* = \text{const}$ wählt. Dann ist $t_1 = \frac{4}{u_1^*}$.

Es sei schließlich noch erwähnt, daß die Parameterdarstellung (3.25) der optimalen Trajektorie mit $x = x_1(t)$ und $y = x_2(t)$ sowie unter Verwendung der Gleichung (3.26) umgeformt werden kann auf $y = \frac{3}{4}x - \frac{1}{2}$. Dieses Ergebnis entspricht voll dem geometrischen Inhalt unserer Aufgabe. Wir haben nämlich unter allen Kurven mit beliebigem Anstieg, die die beiden Punkte \mathbf{x}_0 und \mathbf{x} verbinden, diejenige mit der geringsten Länge gesucht.

Es sei im Rahmen dieses Abschnittes noch auf zwei Abwandlungen der Grundaufgabe optimaler Prozesse verwiesen. Die eine ergibt sich, wenn man statt zwei fester Punkte \mathbf{x}_0 und \mathbf{x}_1 nur einen, etwa \mathbf{x}_0 , vorgibt und gleichzeitig jedoch den Beginn t_0 und das Ende t_1 des Prozesses als gegeben annimmt. Man spricht dann von einem *Prozeß mit fixierter Zeit und freiem rechtem Ende* (siehe 3.3.4.). Für derartige Prozesse lautet das Maximumprinzip genau wie Satz 3.1, nur mit dem Unterschied, daß der erste Teil der Bedingung (3.19) durch

$$p_0(t) \equiv -1, \quad t \in [t_0, t_1],$$

ersetzt werden kann und die Bedingungen

$$p_i(t_1) = 0, \quad i = 1, 2, \dots, n,$$

hinzugefügt werden müssen.

Die andere Abwandlung der Grundaufgabe wurde bereits erwähnt. Sie ist ein Spezialfall, der sich für

$$f_0(\mathbf{x}(t), \mathbf{u}(t)) \equiv 1$$

ergibt, und wird Aufgabe der *Zeitoptimalität* genannt. Das Wesen dieser Aufgabe wurde oben (s. S. 14) schon erläutert. Für sie gilt natürlich auch ein Maximumprinzip, das aus dem allgemeinen Satz 3.1 abgeleitet werden kann. Wegen $f_0(\mathbf{x}, \mathbf{u}) \equiv 1$ erhält man nämlich

$$\mathcal{H} = p_0 + \sum_{i=1}^n p_i f_i(\mathbf{x}, \mathbf{u}).$$

Mit den neuen Bezeichnungen

$$\mathbf{p} = (p_1, p_2, \dots, p_n) \quad \text{und} \quad H(\mathbf{p}, \mathbf{x}, \mathbf{u}) = \sum_{i=1}^n p_i f_i(\mathbf{x}, \mathbf{u})$$

ergeben sich dann

$$\mathcal{H} = p_0 + H \quad \text{bzw.} \quad H = \mathcal{H} - p_0 \quad (3.28)$$

und außerdem das folgende *reduzierte Hamiltonsystem*:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \quad i = 1, 2, \dots, n, \quad (3.29)$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i}, \quad i = 1, 2, \dots, n. \quad (3.30)$$

Die dabei vorgenommene Reduzierung des Vektors \mathbf{P} um die Komponente p_0 und die damit verbundene Reduzierung der Gleichungssysteme (3.15) und (3.16) um die Gleichung für $i = 0$ erweist sich als vorteilhaft, weil man das notwendige Kriterium der Zeitoptimalität ohne p_0 formulieren kann. Dazu wird wie für \mathcal{H} auch für H bei fixierten $\tilde{\mathbf{p}}$ und $\tilde{\mathbf{x}}$ eine Größe $M(\tilde{\mathbf{p}}, \tilde{\mathbf{x}})$ eingeführt:

$$M(\tilde{\mathbf{p}}, \tilde{\mathbf{x}}) = \sup_{\mathbf{u} \in U} H(\tilde{\mathbf{p}}, \tilde{\mathbf{x}}, \mathbf{u}).$$

Wegen (3.28) gilt dabei

$$M(\tilde{\mathbf{p}}, \tilde{\mathbf{x}}) = \mathcal{M}(\mathbf{P}, \tilde{\mathbf{x}}) - p_0 \quad \text{mit} \quad \mathbf{P} = (p_0, \mathbf{p}),$$

so daß sich die beiden Relationen (3.19) durch

$$M(\mathbf{p}(t_1), \mathbf{x}(t_1)) = -p_0(t_1) \geq 0$$

zusammenfassen lassen. Somit ergibt sich

Satz 3.2 (*Notwendiges Kriterium für Zeitoptimalität*): *Dafür, daß ein zulässiger Steuervektor $\mathbf{u}(t)$, $t \in [t_0, t_1]$, der \mathbf{x}_0 in \mathbf{x}_1 überführt, und die ihm entsprechende Trajektorie $\mathbf{x}(t)$ einen schnelligkeitsoptimalen Prozeß bilden, ist notwendig, daß eine ihnen entsprechende stetige Lösung $\mathbf{p}(t) = (p_1(t), p_2(t), \dots, p_n(t))$ von (3.30) mit folgenden Eigenschaften existiert:*

1. Die Lösung $\mathbf{p}(t)$ besitzt wenigstens eine nicht identisch verschwindende Komponente, d.h. $\mathbf{p}(t) \neq \mathbf{0}$ für $t \in [t_0, t_1]$;
2. für beliebig fixiertes $t \in [t_0, t_1]$ nimmt $H(\mathbf{p}(t), \mathbf{x}(t), \mathbf{u})$ als Funktion von $\mathbf{u} \in U$ in $\mathbf{u} = \mathbf{u}(t)$ ihr Maximum an

$$H(\mathbf{p}(t), \mathbf{x}(t), \mathbf{u}(t)) = M(\mathbf{p}(t), \mathbf{x}(t));$$

3. im Endzeitpunkt t_1 besteht die Relation

$$M(\mathbf{p}(t_1), \mathbf{x}(t_1)) \geq 0.$$

Weiterhin gilt: Erfüllen $\mathbf{p}(t)$, $\mathbf{x}(t)$ und $\mathbf{u}(t)$ das Hamiltonsystem (3.29), (3.30) sowie die obige zweite Bedingung, dann ist $M(\mathbf{p}(t), \mathbf{x}(t))$ als Funktion von t konstant, so daß es genügt, die dritte Bedingung für irgendein $t \in [t_0, t_1]$ zu überprüfen.

3.3. Andere Aufgabentypen

3.3.1. Probleme mit beweglichem Endpunkt

Bei der Formulierung der Grundaufgabe (Def. 3.3) in 3.1. wurde davon ausgegangen, daß ein gegebener Punkt \mathbf{x}_0 in einen gegebenen Punkt \mathbf{x}_1 des Phasenraumes zu überführen ist. Gegeben seien nun im Phasenraum die Bedingungen

$$F_k(\mathbf{x}) = 0, \quad k = 1, \dots, m \leq n. \quad (3.31)$$

Es werde vorausgesetzt, daß alle $F_k(\mathbf{x})$ zweimal differenzierbar nach allen x_i sind und die Jacobische Matrix

$$\left(\frac{\partial F_k}{\partial x_i} \right)$$

den Maximalrang m besitzt. Man sagt dann auch, durch $F_k(\mathbf{x}) = 0$ ist eine glatte Mannigfaltigkeit S^1 im R^n bestimmt. Wegen der Forderung, daß die Jacobische

¹⁾ Die Mannigfaltigkeit S besteht aus der Punktmenge des R^n , die durch den Schnitt der Hyperflächen (3.31) bestimmt wird.

Matrix Maximalrang besitzen soll, sind die Vektoren $\text{grad } F_k$ linear unabhängig, sie bilden eine Basis des Normalenraumes zu S , d.h., jeder Vektor

$$\sum_{k=1}^m \lambda_k \text{grad } F_k$$

mit beliebigen Koeffizienten λ_k (nicht alle gleichzeitig gleich null) ist orthogonal zu S .

Wir betrachten nun in Verallgemeinerung der Grundaufgabe die folgende Aufgabe:

Unter allen zulässigen Steuervektoren $\mathbf{u}(t)$ mit $t \in [t_0, t_1]$ ist derjenige (optimale) zu bestimmen, der $\mathbf{x}(t_0) = \mathbf{x}_0$ in $\mathbf{x}(t_1)$ überführt, wobei

$$F_k[\mathbf{x}(t_1)] = 0, \quad k = 1, \dots, m,$$

gilt (Überführung von \mathbf{x}_0 in einen beliebigen Punkt von S) und für den das Funktional

$$J = \int_{t_0}^{t_1} f_0(\mathbf{x}, \mathbf{u}) \, dt \quad (3.32)$$

den kleinsten Wert annimmt. Dabei genüge $\mathbf{x}(t)$ dem System (3.6)

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}).$$

Ist T der Tangentialraum an die Mannigfaltigkeit S im Punkte $\mathbf{x}(t_1)$, dann sagt man auch:

Der Vektor $\mathbf{p}(t)$ erfüllt die Transversalitätsbedingung im rechten Endpunkt der Trajektorie $\mathbf{x}(t)$, wenn $\mathbf{p}(t_1)$ zu T orthogonal ist, d.h. für jeden Vektor $\mathbf{q}^T = (q_1, \dots, q_n)$, der in T liegt, gilt

$$\sum_{i=1}^n p_i(t_1) q_i = (\mathbf{p}(t_1), \mathbf{q}) = 0^1. \quad (3.33)$$

Äquivalent dazu ist

$$\mathbf{p}(t_1) = - \sum_{k=1}^m \lambda_k \text{grad } F_k|_{t=t_1} \quad (\text{nicht alle } \lambda_k \text{ gleich null}). \quad (3.34)$$

Angenommen, wir hätten unsere Aufgabe gelöst, dann ergibt sich ein bestimmter Endpunkt $\mathbf{x}(t_1)$. Daraus folgt, daß die zur Grundaufgabe formulierten Bedingungen des Maximumprinzips auch hier erfüllt sein müssen. Das Maximumprinzip wird bei Problemen mit beweglichem Endpunkt lediglich ergänzt durch die Transversalitätsbedingung (3.34).

Satz 3.3: *Es sei $\mathbf{u}(t)$, $t \in [t_0, t_1]$, eine zulässige Steuerung, die $\mathbf{x}(t_0) = \mathbf{x}_0$ in $\mathbf{x}(t_1) \in S$ überführt, und $\mathbf{x}(t)$ die zugehörige Trajektorie. Notwendig dafür, daß $\mathbf{x}(t)$ und $\mathbf{u}(t)$ Lösungen des oben formulierten Optimierungsproblems sind, ist die Existenz eines nicht identisch verschwindenden stetigen Vektors $\mathbf{p}(t)$, so daß das Maximumprinzip (Satz 3.1) und im rechten Endpunkt die Transversalitätsbedingung erfüllt werden.*

¹⁾ Wenn zwei Vektoren orthogonal sind, dann muß bekanntlich ihr Skalarprodukt gleich null sein.

Bemerkung: Entsprechend folgt eine Transversalitätsbedingung für den Anfangspunkt im Fall, daß $\mathbf{x}(t_0)$ nicht fest vorgeschrieben ist, sondern in einer gegebenen Mannigfaltigkeit liegt (*Aufgabe:* Schreiben Sie die entsprechenden Bedingungen auf!).

Betrachten wir noch den Fall, daß lediglich einige Endwerte, etwa

$$x_i(t_1) = x_i^1, \quad i = 1, \dots, m < n,$$

$$x_i(t_1) \text{ beliebig, } i = m + 1, \dots, n,$$

gegeben sind.

Wir setzen dann

$$F_k(\mathbf{x}) = x_k - x_k^1 = 0, \quad k = 1, \dots, m, \quad (3.35)$$

und erhalten

$$\frac{\partial F_k}{\partial x_i} = \begin{cases} 0 & \text{für } i \neq k, \\ 1 & \text{für } i = k. \end{cases}$$

Damit ergibt sich die Transversalitätsbedingung

$$p_i(t_1) = \begin{cases} -\lambda_i \text{ (beliebig)} & \text{für } i \leq m, \\ 0 & \text{für } i > m, \end{cases} \quad (3.36)$$

d.h., die Endwerte $p_i(t_1)$ sind nicht gegeben für diejenigen Indizes, für welche die $x_i(t_1)$ gegeben sind. Für das Differentialgleichungssystem für die $2n$ Funktionen $x_i(t)$ und $p_i(t)$, also

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}),$$

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{x}},$$

sind somit insgesamt n Endwerte gegeben, nämlich

$$x_i(t_1) \quad \text{für } i \leq m,$$

$$p_i(t_1) \quad \text{für } i > m.$$

Beispiel 3.4: Gegeben sei das System

$$\dot{x}_1 = x_2, \quad x_1(0) = \frac{1}{4},$$

$$\dot{x}_2 = u, \quad x_2(0) = 0,$$

$$|u| \leq 1.$$

Gesucht ist eine solche optimale Steuerung, die den Punkt $\mathbf{x}^0 = \mathbf{x}(0)$ in kürzester Zeit t_1 in einen Punkt der Geraden $x_1 = x_2$ überführt. Für den Endpunkt gilt offensichtlich die Bedingung

$$F_1(\mathbf{x}(t_1)) \equiv x_1(t_1) - x_2(t_1) = 0.$$

Damit erhalten wir die Transversalitätsbedingung

$$p_1(t_1) = -\lambda,$$

$$p_2(t_1) = \lambda.$$

Aus $H = x_2 p_1 + u p_2$ erhalten wir mit Satz 3.2 als notwendige Bedingung für die optimale Steuerung $u = \operatorname{sgn} p_2$ (falls $p_2 \neq 0$). Weiter ist $\dot{p}_1 = 0$, $\dot{p}_2 = -p_1$ und damit unter Berücksichtigung der Transversalitätsbedingung

$$p_1(t) = -\lambda, \quad p_2(t) = \lambda t + \lambda(1 - t_1) = \lambda(t - t_1 + 1).$$

$p_2(t)$ ist also eine lineare Funktion von t ; es erfolgt ein Vorzeichenwechsel und damit höchstens eine Umschaltung bezüglich der optimalen Steuerung. Die Umschaltzeit ist dann $t_s = t_1 - 1$. Angenommen, wir würden vollständig mit $u = +1$ steuern. Dann folgt

$$x_1 = \frac{t^2}{2} + \frac{1}{7},$$

$$x_2 = t$$

und als Trajektorie die Parabel

$$x_2^2 = 2(x_1 - \frac{1}{7}).$$

Für den Endpunkt und die Endzeit ergibt sich wegen $x_1(t_1) = x_2(t_1)$

$$t_1 = \frac{t_1^2}{2} + \frac{1}{7}$$

d. h.

$$t_1 = 1 \pm \sqrt{\frac{5}{7}}.$$

Die zwei Werte entsprechen dabei den zwei Schnittpunkten der Parabel mit der Geraden $x_1 = x_2$ (vgl. Bild 3.2).

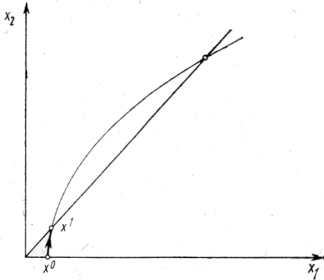


Bild 3.2
Optimale Trajektorie für Beispiel 3.4

Nur der kleinere Wert ist für uns von Interesse (warum?). Bezüglich der Hamiltonfunktion folgt

$$H|_{t=t_1} = -\lambda x_2(t_1) + \lambda u(t_1) = -\lambda t_1 + \lambda = 1$$

für $\lambda = \frac{1}{1 - t_1} = \sqrt{\frac{7}{5}}$ und damit $\operatorname{sgn} p_2 = +1$ für $t \in [0, t_1]$.

Wird nur mit $u = -1$ gesteuert, folgt (führen Sie die Rechnung aus!)

$$t_1 = 1 \pm \sqrt{\frac{9}{7}}.$$

Wegen $1 + \sqrt{\frac{9}{7}} > 1 - \sqrt{\frac{5}{7}}$ und $1 - \sqrt{\frac{9}{7}} < 0$ liegen also ein schlechterer und ein unbrauchbarer (warum?) Wert vor.

Zeigen Sie, daß man auch für die Steuerungen mit einer Umschaltung größere Werte für t_1 erhält!

3.3.2. Probleme mit anderem Optimierungsziel

Bei vielen Aufgaben ist das Optimierungsziel nicht in der Form (3.32) (*Aufgabe von Lagrange*) sondern als Funktion des Endpunktes

$$F[\mathbf{x}(t_1)]$$

gegeben (*Aufgabe von Mayer*). Dabei soll $F[\mathbf{x}(t)]$ nach allen Argumenten zweimal differenzierbar sein. In diesem Fall gilt

Satz 3.4: Ist das Optimierungsziel in der Form $F[\mathbf{x}(t_1)]$ gegeben, dann ist notwendig für Optimalität einer Steuerung $\mathbf{u}(t)$ die Existenz eines stetigen, nichtidentisch verschwindenden adjungierten Vektors $\mathbf{p}(t)$ als Lösung des Systems

$$\dot{p}_i = - \sum_{j=1}^n p_j \frac{\partial f_j}{\partial x_i}, \quad i = 1, \dots, n, \quad (3.37)$$

mit den Transversalitätsbedingungen

$$p_i(t_1) = - \left. \frac{\partial F}{\partial x_i} \right|_{t=t_1} - \sum_{k=1}^m \lambda_k \left. \frac{\partial F_k}{\partial x_i} \right|_{t=t_1}, \quad (3.38)$$

so daß gilt

$$\max_{\mathbf{u}} H(\mathbf{x}(t), \mathbf{p}(t), \mathbf{u}) = H(\mathbf{x}(t), \mathbf{p}(t), \mathbf{u}(t)), \quad (3.39)$$

$$\max H|_{t=t_1} = 0 \quad (3.40)$$

mit

$$H = \sum_{i=1}^n p_i f_i. \quad (3.41)$$

Für die Anwendungen bedeutsam ist u. a. der Fall, daß

$$F[\mathbf{x}(t_1)] = \sum_{i=1}^n c_i x_i(t_1) \quad (3.42)$$

zu minimieren ist. Hier ist

$$\frac{\partial F}{\partial x_i} = c_i, \quad (3.43)$$

und damit folgt aus (3.38) die Endbedingung

$$p_i(t_1) = -c_i - \sum_{k=1}^m \lambda_k \left. \frac{\partial F_k}{\partial x_i} \right|_{t=t_1}. \quad (3.44)$$

Aufgabe 3.1: Zeigen Sie, daß man auch die Aufgabe von Lagrange auf die Aufgabe von Mayer zurückführen kann!

Aufgabe 3.2: Hat das zu minimierende Funktional die Form

$$J = F[x(t_1)] + \int_{t_0}^{t_1} f_0(x, u) dt,$$

spricht man auch von der Aufgabe von Bolza. Führen Sie diese Aufgabe auf die Aufgabe von a) Lagrange, b) Mayer zurück!

Beispiel 3.5: Gegeben sei die chemische Folgereaktion 1. Ordnung $A \rightarrow B \rightarrow C$ mit den Folgeprodukten A, B, C , den zugehörigen Konzentrationen x_1, x_2, x_3 ($x_1 + x_2 + x_3 = 1$) und den Geschwindigkeitskoeffizienten k_1, k_2 . Das kinetische System (Prozeßgleichungen) ist dann

$$\frac{dx_1}{dt} = -k_1 x_1,$$

$$\frac{dx_2}{dt} = k_1 x_1 - k_2 x_2$$

mit den Anfangsbedingungen

$$x_1(0) = 1, \quad x_2(0) = 0.$$

Die Geschwindigkeitskoeffizienten werden in der Form

$$k_i = k_{i\infty} \exp\left(-\frac{E_i}{RT}\right), \quad i = 1, 2,$$

angenommen, wobei $T(t)$ der Temperaturverlauf, E_i die Aktivierungsenergien und R die Gaskonstante sind. Gesucht ist ein solcher Temperaturverlauf $T(t)$ mit $T_1 \leq T(t) \leq T_2$ (T_1 : Zündtemperatur der Reaktion, T_2 obere Grenztemperatur), so daß maximale Ausbeute hinsichtlich B , d.h. $\max x_2$, erreicht wird.

In diesem Fall liegt das Optimierungsziel in der Form (3.42) mit $c_1 = 0, c_2 = 1$ vor. Es folgt die Hamiltonfunktion

$$H = -k_1 x_1 p_1 + (k_1 x_1 - k_2 x_2) p_2 = k_1 x_1 (p_2 - p_1) - k_2 x_2 p_2$$

und das adjungierte System

$$\frac{dp_1}{dt} = k_1 (p_1 - p_2), \quad p_1(t_1) = -c_1 = 0,$$

$$\frac{dp_2}{dt} = k_2 p_2, \quad p_2(t_1) = -c_2 = -1.$$

Untersuchen wir zunächst H auf stationäre Stellen bezüglich T . Es ist

$$\frac{\partial H}{\partial T} = k_1' x_1 (p_2 - p_1) - k_2' x_2 p_2,$$

wobei

$$k_i' = \frac{dk_i}{dT} = k_{i\infty} \frac{E_i}{RT^2} \exp\left(-\frac{E_i}{RT}\right) = \frac{E_i}{RT^2} k_i$$

gilt. Aus $\frac{\partial H}{\partial T} = 0$ erhalten wir

$$E_1 k_1 x_1 (p_2 - p_1) - E_2 k_2 x_2 p_2 = 0,$$

d. h.

$$\frac{k_1}{k_2} = \frac{E_2 x_2 p_2}{E_1 x_1 (p_2 - p_1)}$$

oder

$$\exp \frac{E_2 - E_1}{RT_s} = \frac{k_{2\infty} E_2 x_2 p_2}{k_{1\infty} E_1 x_1 (p_2 - p_1)}.$$

Unter der Annahme, daß die rechte Seite dieser Beziehung positiv ist, ergibt sich schließlich

$$T_s = \frac{E_2 - E_1}{R \ln \frac{k_{2\infty} E_2 x_2 p_2}{k_{1\infty} E_1 x_1 (p_2 - p_1)}}.$$

Setzen wir

$$q = \frac{x_2 p_2}{x_1 (p_2 - p_1)},$$

so folgt

$$T_s = \frac{E_2 - E_1}{R \ln \frac{k_{2\infty} E_2}{k_{1\infty} E_1} q}.$$

Zu überprüfen wäre noch, daß $\left. \frac{\partial^2 H}{\partial T^2} \right|_{T=T_s} > 0$ gilt (Aufgabe!).

Zur Bestimmung von T_s ist damit die Kenntnis von q erforderlich. Kinetisches und adjungiertes System sind jedoch nicht geschlossen integrierbar, so daß zur weiteren Lösung ein numerisches Verfahren eingesetzt werden muß. Falls T_s reell und $T_1 \leq T_s \leq T_2$ gilt, ist somit $T_{\text{opt}} = T_s$. Anderenfalls ist mit T_1 bzw. T_2 zu steuern, je nachdem, für welchen Wert sich das Minimum von H einstellt.

3.3.3. Nichtautonome Systeme

Bisher hatten wir Systeme (3.6) betrachtet, bei denen die rechten Seiten nicht explizit von t abhängen, man spricht dann auch von einem *autonomen* System. Systeme der Form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \quad (3.45)$$

nennt man auch *nichtautonom*. Dabei kann das Zielfunktional in der Form

$$J = \int_{t_0}^{t_1} f_0(\mathbf{x}, \mathbf{u}, t) dt \quad (3.46)$$

gegeben sein.

Führen wir die zusätzliche Phasenkoordinate

$$x_{n+1} = t$$

ein, dann ergibt sich als weitere Differentialgleichung

$$\frac{dx_{n+1}}{dt} = 1, \quad x_{n+1}(t_0) = t_0. \quad (3.47)$$

Im Raum der Variablen x_1, x_2, \dots, x_{n+1} liegt damit das autonome System

$$\begin{aligned}\frac{dx_i}{dt} &= f_i(\mathbf{x}, \mathbf{u}, x_{n+1}), \quad i = 1, \dots, n, \\ \frac{dx_{n+1}}{dt} &= 1,\end{aligned}\tag{3.48}$$

$$J = \int_{t_0}^{t_1} f_0(\mathbf{x}, \mathbf{u}, x_{n+1}) dt \tag{3.49}$$

vor, auf das unmittelbar das Maximumprinzip angewandt werden kann. Die Hamiltonfunktion für das neue System lautet

$$\mathcal{H}^* = \sum_{i=0}^n p_i f_i + p_{n+1} \tag{3.50}$$

und mit

$$\begin{aligned}\mathcal{H} &= \sum_{i=0}^n p_i f_i \\ \mathcal{H}^* &= \mathcal{H} + p_{n+1}.\end{aligned}\tag{3.51}$$

Entsprechend ergibt sich das adjungierte System

$$\begin{aligned}\frac{dp_i}{dt} &= - \sum_{j=0}^n \frac{\partial f_j}{\partial x_i} p_j, \quad i = 0, 1, \dots, n, \\ \frac{dp_{n+1}}{dt} &= - \sum_{j=0}^n \frac{\partial f_j}{\partial t} p_j.\end{aligned}\tag{3.52}$$

Die Bedingung des Maximumprinzips

$$\max \mathcal{H}^* = 0$$

kann deshalb ersetzt werden durch

$$\max \mathcal{H} = - p_{n+1}.\tag{3.53}$$

Sind bezüglich des Endpunktes die Bedingungen

$$F_k[\mathbf{x}(t_1), t_1] = 0, \quad k = 1, \dots, m \leq n,\tag{3.54}$$

zu erfüllen, so kann man die Transversalitätsbedingungen (3.34) unmittelbar anwenden, wobei zu beachten ist, daß gilt

$$\mathbf{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_n \\ p_{n+1} \end{bmatrix}, \quad \text{grad } F_k = \begin{bmatrix} \frac{\partial F_k}{\partial x_1} \\ \vdots \\ \frac{\partial F_k}{\partial x_n} \\ \frac{\partial F_k}{\partial t} \end{bmatrix}.$$

Ist z.B. der Endpunkt $\mathbf{x}(t_1) = \mathbf{x}^1$ fest vorgegeben, so ist lediglich $x_{n+1}(t_1)$ im neuen Phasenraum beliebig. Entsprechend (3.36) folgt somit

$$p_{n+1}(t_1) = 0.$$

Beispiel 3.6: Gegeben sei

$$\dot{x}_1 = 2u + t, \quad |u| \leq 1.$$

Gesucht ist eine zulässige Steuerung, für die die zugehörige Trajektorie den Punkt

$$x_1(0) = 1 \text{ in den Punkt } x_1(t_1) = 0$$

überführt, wobei

$$J = \int_0^{t_1} (x_1 + t) dt$$

zum Minimum wird.

Mit $x_2 = t$ erhalten wir die Differentialgleichung

$$\dot{x}_2 = 1, \quad x_2(0) = 0,$$

die Hamiltonfunktion

$$\mathcal{H}^* = p_0(x_1 + x_2) + p_1(2u + x_2) + p_2$$

und das adjungierte System sowie die Transversalitätsbedingungen

$$\dot{p}_0 = 0, \quad p_0(t_1) = -1,$$

$$\dot{p}_1 = -p_0 = 1,$$

$$\dot{p}_2 = -p_0 - p_1 = 1 - p_1, \quad p_2(t_1) = 0.$$

Das Maximum von \mathcal{H}^* ergibt sich für $u = \operatorname{sgn} p_1$ (falls $p_1 \neq 0$). Integration der entsprechenden Differentialgleichung liefert

$$p_1 = t + c_1$$

und somit

$$u = \operatorname{sgn}(t + c_1).$$

Die Funktion $p_1 = t + c_1$ besitzt genau eine Nullstelle, d.h., es erfolgt ein Vorzeichenwechsel. Somit kann es bezüglich der optimalen Steuerung höchstens eine Umschaltung geben. Dann muß jedoch $c_1 < 0$ sein. Folgende drei Fälle sind damit zunächst möglich:

1. es wird nur mit $u = 1$ gesteuert,
2. es wird in $[0, t_2]$ mit $u = -1$ und in (t_2, t_1) mit $u = +1$ gesteuert, wobei $t_2 = -c_1$ gilt,
3. es wird nur mit $u = -1$ gesteuert.

Für den 1. Fall ergibt sich

$$x_1 = \frac{t^2}{2} + 2t + 1.$$

Wir überzeugen uns leicht, daß der Endpunkt $x_1(t_1) = 0$ für $t_1 > 0$ nicht erreicht werden kann. Beginnen wir also mit $u = -1$. Dann folgt

$$x_1 = \frac{t^2}{2} - 2t + 1.$$

Nehmen wir an, es ist keine Umschaltung erforderlich. Aus $x_1(t_1) = 0$ ergibt sich

$$t_1 = 2 \pm \sqrt{2}.$$

Weiter ist

$$\mathcal{H}^*|_{t_1} = -t_1 + (t_1 - 2)(t_1 + c_1) = -(2 \pm \sqrt{2}) - (\pm \sqrt{2})(2 \pm \sqrt{2} + c_1) = 0$$

und damit

$$c_1 = -1$$

also

$$p_1 = t - 1.$$

Für $t \leq t_1 = 2 - \sqrt{2}$ ist $p_1 < 0$, also tatsächlich

$$u = \operatorname{sgn} p_1 = -1.$$

Der zugehörige Wert von J ist dann

$$J = \int_0^{t_1} (x_1 + t) dt = \frac{1}{3} (7 - 4\sqrt{2}).$$

3.3.4. Probleme mit fester Endzeit

Im Unterschied zu den bisher betrachteten Problemstellungen soll nun neben t_0 auch t_1 fest vorgegeben sein. Gegeben seien also das System (3.45) mit dem Optimierungsziel (3.46), t_0 und t_1 sowie entsprechende Anfangs- und Endbedingungen. Setzt man wieder $x_{n+1} = t$, dann folgt die zusätzliche Differentialgleichung

$$\frac{dx_{n+1}}{dt} = 1 \quad \text{mit} \quad x_{n+1}(t_0) = t_0, \quad x_{n+1}(t_1) = t_1, \quad (3.55)$$

d.h., der Endwert von x_{n+1} ist jetzt fest gegeben. Damit können wir in (3.46) t_1 als variabel ansehen. Auswirkungen ergeben sich bezüglich der Transversalitätsbedingungen.

Ist beispielsweise $\mathbf{x}(t_0)$ in $\mathbf{x}(t_1)$ zu überführen, so sind im $(n+1)$ -dimensionalen neuen Phasenraum alle Endpunkte gegeben, d.h., es ist keiner der Werte $p_i(t_1)$, $i = 1, \dots, n+1$, gegeben. Aus

$$\mathcal{H}^*|_{t_1} = \mathcal{H}|_{t_1} + p_{n+1}(t_1) = 0 \quad (3.56)$$

folgt somit

$$\mathcal{H}|_{t_1} = -p_{n+1}(t_1).$$

Die Bedingung $\mathcal{H}|_{t_1} = 0$ des Maximumprinzips entfällt, alle anderen Aussagen bleiben gültig.

Ist $\mathbf{x}(t_0)$ in einen Punkt der Mannigfaltigkeit

$$F_k[\mathbf{x}(t_1), t_1] = 0, \quad k = 1, \dots, m,$$

zu überführen, ergeben sich die Transversalitätsbedingungen

$$\begin{aligned} p_i(t_1) &= - \sum_{k=1}^m \lambda_k \frac{\partial F_k}{\partial x_i} \Big|_{t=t_1}, \quad i = 1, \dots, n, \\ p_{n+1}(t_1) &= - \sum_{k=1}^m \lambda_k \frac{\partial F_k}{\partial t} \Big|_{t=t_1} - \lambda_{m+1}. \end{aligned} \quad (3.57)$$

Beispiel 3.7: Gegeben sei

$$\dot{x}_1 = x_2, \quad x_1(0) = x_2(0) = 1,$$

$$\dot{x}_2 = u, \quad x_1(1) = 0.$$

Gesucht ist eine Steuerung mit $|u| \leq 6$, so daß

$$J = \int_0^1 u^2 dt$$

minimal wird. Man nennt Aufgaben mit solchem Funktional auch Probleme mit minimalem Energieverbrauch.

Wir erhalten mit $p_0 = -1$ entsprechend Bedingung (3.16) und Satz 3.1

$$\begin{aligned} \mathcal{H} &= p_0 u^2 + p_1 x_2 + p_2 u = -u^2 + p_1 x_2 + p_2 u \\ &= - \left(u - \frac{p_2}{2} \right)^2 + p_1 x_2 + \frac{p_2^2}{4} \end{aligned}$$

und

$$\dot{p}_1 = 0, \quad \dot{p}_2 = -p_1, \quad p_2(1) = 0$$

ohne Endbedingung für p_1 . Maximierung von \mathcal{H} liefert (quadratische Funktion!)

$$u = \begin{cases} \frac{p_2}{2} & \text{falls } |p_2| \leq 12, \\ 6 \operatorname{sgn} p_2 & \text{falls } |p_2| > 12. \end{cases}$$

Integration des adjungierten Systems ergibt unter Beachtung von $p_2(1) = 0$

$$p_1 = c_1, \quad p_2 = -c_1 t + c_2 = -c_1(t-1).$$

Unter Berücksichtigung des Verhaltens von $p_2(t)$ erhalten wir damit folgende Möglichkeiten:

1. Falls $c_1 > 12$, dann ist

$$u = \begin{cases} 6 & \text{für } 0 \leq t \leq t_2, \\ -\frac{1}{2} c_1(t-1) & \text{für } t_2 \leq t \leq 1, \end{cases}$$

wobei

$$c_1 = -\frac{12}{t_2 - 1} \text{ gilt.}$$

2. Falls $-12 \leq c_1 \leq 12$, dann ist

$$u = -\frac{1}{2} c_1(t-1).$$

3. Falls $c_1 < -12$, dann ist

$$u = \begin{cases} -6 & \text{für } 0 \leq t \leq t_2, \\ -\frac{1}{2}c_1(t-1) & \text{für } t_2 \leq t \leq 1, \end{cases}$$

wobei

$$c_1 = \frac{12}{t_2 - 1}$$

gilt.

Betrachten wir zunächst den 2. Fall. Durch Integration des Zustandssystems erhalten wir

$$x_2 = -\frac{1}{2}c_1\left(\frac{t^2}{2} - t\right) + 1,$$

$$x_1 = -\frac{1}{2}c_1\left(\frac{t^3}{6} - \frac{t^2}{2}\right) + t + 1.$$

Aus $x_1(1) = 0$ ergibt sich $c_1 = -12$, d.h., die erforderliche Bedingung ist erfüllt. Das Zielfunktional hat für $u = 6(t-1)$ den Wert

$$J = \int_0^1 u^2 dt = 12.$$

Den 1. und 3. Fall behandeln wir gemeinsam, indem wir $a = 6$ bzw. -6 setzen. Dann folgt

$$x_2 = \begin{cases} at + 1 & \text{für } t \leq t_2, \\ -\frac{1}{2}c_1\left(\frac{t^2}{2} - t - \frac{t_2^2}{2} + t_2\right) + at_2 + 1 & \text{für } t \geq t_2, \end{cases}$$

$$x_1 = \begin{cases} \frac{a}{2}t^2 + t + 1 & \text{für } t \leq t_2, \\ -\frac{1}{2}c_1\left[\frac{t^3}{6} - \frac{t^2}{2} - t\left(\frac{t_2^2}{2} - t_2\right) - \frac{t_2^3}{6} + \frac{t_2^2}{2} + t_2\left(\frac{t_2^2}{2} - t_2\right)\right] \\ \quad + (at_2 + 1)(t - t_2) + \frac{a}{2}t_2^2 + t_2 + 1 & \text{für } t \geq t_2. \end{cases}$$

Wegen $x_1(1) = 0$ und $c_1 = -\frac{2a}{t_2 - 1}$ erhalten wir

$$\frac{a}{t_2 - 1} \left(-\frac{1}{3} + t_2 - t_2^2 + \frac{t_2^3}{3} \right) - a \frac{t_2^2}{2} + at_2 + 2 = 0,$$

und daraus durch Umformen

$$a \left(\frac{t_2^2}{3} - 2\frac{t_2}{3} + \frac{1}{3} \right) - a \frac{t_2^2}{2} + at_2 + 2 = 0,$$

$$t_2^2 - 2t_2 - 2 - \frac{12}{a} = 0,$$

$$t_2 = 1 \pm \sqrt{3 + \frac{12}{a}}.$$

Für $a = 6$ folgt $t_2 = 1 \pm \sqrt{5}$; beide Werte liegen nicht im Intervall $[0, 1]$. Für $a = -6$ folgt $t_2 = 2$ bzw. 0, der 1. Wert liegt nicht im entsprechenden Intervall; $t_2 = 0$ ist identisch mit dem 2. Fall.

Aufgabe 3.3: Lösen Sie das in Beispiel 3.7 gestellte Problem mit dem Optimierungsziel

$$J = \int_0^1 u \, dt!$$

Aufgabe 3.4: Gegeben sei $\dot{x}_1 = 2u + t$, $x_1(0) = 1$, $|u| \leq 1$. Gesucht ist eine solche zulässige Steuerung, für die

a) $\int_0^1 x_1 \, dt,$

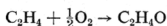
c) $\int_0^1 (x_1 + u^2) \, dt,$

b) $\int_0^1 (x_1 + u) \, dt,$

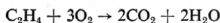
d) $x_1(1)$

minimal wird.

Beispiel 3.8: In einem Reaktor (als ideales Strömungsrohr angesehen) soll Äthylenoxid durch Überleitung eines Äthylen-Luft-Gemisches über einen Ag-Katalysator hergestellt werden. Dabei betrachten wir nur die stationäre Betriebsweise und vernachlässigen Volumenänderungen des Reaktionsgemisches. Neben der erwünschten Reaktion



findet noch eine Totaloxydation



statt. Die Steuerung des Prozesses soll über die Abführung der Reaktionswärme durch einen Kühlmantel, d.h., durch das Temperaturprofil des Reaktionsrohres erfolgen. Setzen wir eine einfache Langmuir-Hinshelwood-Kinetik voraus, ergibt sich das Prozeßmodell¹⁾

$$\frac{dx_1}{dt} = (k_1 + k_2) \frac{1 - x_1}{1 + ax_1},$$

$$\frac{dx_2}{dt} = k_1 \frac{1 - x_1}{1 + ax_1}$$

mit $x_1(0) = x_2(0) = 0$, $0 \leq t \leq t_1$ (t_1 fest gegeben). Dabei bedeuten

$x_1(t)$ Gesamtumsetzungsgrad des C_2H_4 nach beiden Reaktionen,

$x_2(t)$ Umsetzungsgrad des C_2H_4 zu $\text{C}_2\text{H}_4\text{O}$,

$T(t)$ Temperaturverteilung mit $T_1 \leq T \leq T_2$ (T_1 Zündtemperatur, T_2 obere Grenztemperatur),

$k_i(T)$ Geschwindigkeitskoeffizienten der einzelnen Reaktionen ($i = 1, 2$) mit

$$k_i(T) = k_{i\infty} \exp\left(-\frac{E_i}{RT}\right), \quad E_2 > E_1,$$

R universelle Gaskonstante,

E_i Aktivierungsenergien der Reaktionen,

t Verweilzeit,

$a > 0$ Prozeßkonstante.

¹⁾ Vgl. Wolin, J. M.; G. M. Ostrowski, M. G. Slinko, Kinetika i Katalis (Кинетика и катализ), IV/5, 1963, S. 760 bis 767.

Gesucht ist eine solche zulässige Steuerung $T(t)$, für die sich maximale Ausbeute bezüglich C_2H_4O ergibt, also ist $x_2(t_1)$ zu maximieren (Spezialfall der Aufgabe von Mayer).

Wir erhalten zunächst

$$H = p_1(k_1 + k_2) \frac{1 - x_1}{1 + ax_1} + p_2 k_1 \frac{1 - x_1}{1 + ax_1} = \frac{1 - x_1}{1 + ax_1} [(k_1 + k_2)p_1 + k_1 p_2],$$

$$\frac{dp_1}{dt} = -\frac{\partial H}{\partial x_1} = \frac{1 + a}{(1 + ax_1)^2} [(k_1 + k_2)p_1 + k_1 p_2], \quad p_1(t_1) = 0,$$

$$\frac{dp_2}{dt} = 0, \quad p_2(t_1) = -1, \quad \text{also} \quad p_2(t) = -1.$$

Wegen der Maximierung von $x_2(t_1)$ ist H zu minimieren. Falls ein Minimum im Innern von $[T_1, T_0]$ liegt, muß dort $\frac{\partial H}{\partial T} = 0$ gelten. Es ist

$$\frac{\partial H}{\partial T} = \frac{1 - x_1}{1 + ax_1} [k_1'(p_1 - 1) + k_2' p_1]$$

mit

$$k_i' = \frac{dk_i}{dT} = k_{i\infty} \frac{E_i}{RT^2} \exp\left(-\frac{E_i}{RT}\right) = \frac{E_i}{RT^2} k_i.$$

Aus $\frac{\partial H}{\partial T} = 0$ folgt (die Lösung werde mit T_s bezeichnet)

$$k_1'(p_1 - 1) + k_2' p_1 = 0,$$

also

$$E_1 k_1(p_1 - 1) + E_2 k_2 p_1 = 0$$

oder

$$\exp \frac{E_2 - E_1}{RT_s} = \frac{k_{2\infty} E_2 p_1}{k_{1\infty} E_1 (1 - p_1)} = q$$

und, falls $q > 0$,

$$T_s = \frac{E_2 - E_1}{R \cdot \ln q}.$$

Zeigen Sie, daß $\frac{\partial^2 H}{\partial T^2} \Big|_{T=T_s} > 0$ gilt, also tatsächlich ein Minimum von H vorliegt!

Wir haben nur T_s in Abhängigkeit von der noch nicht bekannten Funktion $p_1(t)$ dargestellt. Dabei ist $p_1(t)$ nicht durch geschlossene Integration zu gewinnen, hier müßte ein numerisches Verfahren eingesetzt werden. Machen wir zunächst einige qualitative Aussagen. Bezüglich $T_s = f(p_1)$ gilt:

p_1	q	T_s
$p_1 > 1$	$q < 0$	nicht reell
$p_1 \rightarrow 1 - 0$	$q \rightarrow \infty$	$T_s \rightarrow +0$
$p_1 \rightarrow +0$	$q \rightarrow +0$	$T_s \rightarrow -0$
$p_1 < 0$	$q < 0$	nicht reell
$p_1 \rightarrow \frac{k_{1\infty} E_1}{k_{1\infty} E_1 + k_{2\infty} E_2} \pm 0$	$q \rightarrow 1 \pm 0$	$T_s \rightarrow \pm \infty$

Damit können wir $T_s = f(p_1)$ schematisch darstellen (vgl. Bild 3.3, dabei wurde $\frac{k_{1\infty}E_1}{k_{1\infty}E_1 + k_{2\infty}E_2} = p_{1\infty}$ gesetzt). Für $T_s \in [T_1, T_2]$ entspricht T_s der optimalen Temperaturführung. Das zugehörige Intervall für p_1 ist $[p_{12}, p_{11}]$ mit

$$p_{1i} = \frac{E_1 k_1(T_i)}{E_1 k_1(T_i) + E_2 k_2(T_i)}, \quad i = 1, 2.$$

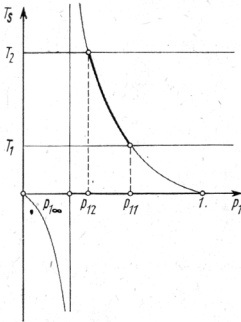


Bild 3.3

Schematische Darstellung von $T_s = f(p_1)$ in Beispiel 3.8

Für $p_1 < p_{12}$ und $p_1 > p_{11}$ ist entweder mit T_1 oder mit T_2 zu steuern. Zur Entscheidung muß überprüft werden, für welchen Wert H zum Minimum wird (Aufgabe!). Schließlich erhalten wir

$$T_{\text{opt}} = \begin{cases} T_2 & \text{für } p_1 < p_{12}, \\ T_s & \text{für } p_{12} \leq p_1 \leq p_{11}, \\ T_1 & \text{für } p_{11} < p_1, \end{cases}$$

d.h., T_{opt} ist in Abhängigkeit von p_1 monoton fallend. Untersuchen wir $p_1(t)$ weiter. Es ist

$$\text{sgn } \frac{dp_1}{dt} = \text{sgn } [(k_1 + k_2)p_1 - k_1].$$

Wegen $p_1(t_1) = 0$ ergibt sich $\left. \frac{dp_1}{dt} \right|_{t_1} < 0$, d.h., bei t_1 ist die Funktion monoton fallend. Es läßt sich sogar weiter zeigen, daß solange $p_1 \leq p_{11} = \frac{E_1 k_1}{E_1 k_1 + E_2 k_2}$ ist, gilt

$$(k_1 + k_2)p_1 - k_1 \leq (k_1 + k_2) \frac{E_1 k_1}{E_1 k_1 + E_2 k_2} - k_1 = \frac{(E_1 - E_2)k_1 k_2}{E_1 k_1 + E_2 k_2} < 0.$$

Damit ist $p_1(t)$ in einem Intervall $[t_2, t_1]$ monoton fallend. Angenommen bei t_2 ändert sich das Verhalten, d.h., für $t < t_2$ ist $p_1(t)$ monoton zunehmend. Dann muß $p_1(t) \leq p_1(t_2)$ gelten, es ergibt sich also ein Widerspruch. Folgende qualitative Aussage ist damit möglich:

$p_1(t)$ ist monoton fallend, es gilt $0 \leq p_1(t) \leq \frac{k_1}{k_1 + k_2} < 1$. Da auch $T_{\text{opt}}(p_1)$ monoton fallend ist, folgt sofort:

$$T_{\text{opt}}(t) = T_{\text{opt}}[p_1(t)]$$

ist monoton zunehmend. Je nach dem Verhalten von p_1 sind also drei Reaktorfahrweisen möglich:

1. Es ist stets $p_1 \leq p_{12}$. Dann gilt

$$T_{\text{opt}} = T_2.$$

2. Es ist $p_1 \leq p_{11}$, aber es gibt Werte mit $p_{12} < p_1$.

Dann gilt

$$T_{\text{opt}} = \begin{cases} T_s & \text{für } 0 \leq t \leq t_a, \\ T_2 & \text{für } t_a < t \leq t_1, \end{cases}$$

wobei t_a aus $p_1(t_a) = p_{12}$ zu bestimmen ist.

3. Es gibt Werte mit $p_{11} < p_1$. Dann gilt

$$T_{\text{opt}} = \begin{cases} T_1 & \text{für } 0 \leq t \leq t_b, \\ T_s & \text{für } t_b < t \leq t_a, \\ T_2 & \text{für } t_a < t \leq t_1, \end{cases}$$

wobei t_b aus $p_1(t_b) = p_{11}$ zu bestimmen ist.

Es soll aber noch einmal darauf hingewiesen werden, daß die Bestimmung von p_1 und damit T_{opt} den Einsatz von numerischen Verfahren erfordert, da die entsprechenden Differentialgleichungssysteme nicht geschlossen integriert werden können.

3.3.5. Einige weitere Probleme

Neben den hier behandelten Aufgabenstellungen sind im Zusammenhang mit optimalen Prozessen noch eine Reihe anderer Probleme von Bedeutung. Wir können hier keine umfassende Darstellung geben. Erwähnt werden sollen lediglich drei weitere Aufgaben. Bezüglich ihrer mathematischen Behandlung sei auf entsprechende Literatur verwiesen, z.B. [6], [20].

- a) *Probleme mit unbeschränktem Zeitintervall.* Diese Probleme sind dadurch gekennzeichnet, daß als Zeitintervall $[t_0, \infty]$ gegeben ist und eine solche zulässige Steuerung gesucht wird, für die das uneigentliche Integral

$$\int_{t_0}^{\infty} f_0(x, u) dt \quad (3.58)$$

konvergiert und minimal wird.

- b) *Optimale Prozesse mit Verzögerung.* Bei manchen Problemen tritt ein Verzögerungseffekt auf, z.B. kann für die Weitergabe von entsprechenden Signalen eine zu berücksichtigende feste Zeit T erforderlich sein. Nehmen wir an, daß diese Verzögerung lediglich den Zustand beeinflußt, so läßt sich das Problem durch

$$\frac{dx}{dt} = f[x(t), x(t-T), u(t)] \quad (T > 0, \text{ fest}) \quad (3.59)$$

beschreiben. Das Optimierungsziel kann dann z.B. in der Form

$$\int_{t_0}^{t_1} f_0[x(t), x(t-T), u(t)] dt \quad (3.60)$$

vorliegen.

c) *Optimale Prozesse mit Parametern.* Hängt ein Prozeß nicht nur von einem zeitabhängigen Steuervektor $u(t)$ ab, sondern zusätzlich von einem in einem gewissen Bereich frei wählbaren zeitunabhängigen (d. h. über das gesamte Zeitintervall konstanten) Parametervektor w , so spricht man auch von einem Prozeß mit Parametern. Für das System

$$\frac{dx}{dt} = f[x(t), u(t), w]$$

mit entsprechenden Anfangs- und Endbedingungen ist dann beispielsweise ein solcher Steuervektor $u(t)$ und Parametervektor w gesucht, daß

$$\int_{t_0}^{t_1} f_0[x(t), u(t), w] dt \quad (3.61)$$

minimal wird.

3.4. Numerische Methoden

3.4.1. Vorbemerkungen

Die numerischen Verfahren zur Bestimmung optimaler Steuerungen können wir in zwei Klassen einteilen

1. Verfahren, die auf Maximumprinzipien aufbauen,
2. Verfahren, bei denen Maximumprinzipien nicht benutzt werden.

Es sei $J(u)$ zu minimieren, der zulässige Steuerbereich sei U . In vielen Verfahren wird dann eine sog. *Minimalfolge* $\{u^k\}$ konstruiert. Dabei heißt die Folge $\{u^k\}$ *Minimalfolge*, wenn für die zugehörige Folge $\{J(u^k)\}$ gilt

$$\lim_{k \rightarrow \infty} J(u^k) = \inf_{u \in U} J(u).$$

Die Minimalfolge selbst braucht dabei nicht zu konvergieren. Auch falls keine optimale Steuerung existiert, aber $J(u)$ stetig und nach unten beschränkt ist, existiert eine Minimalfolge.

Die Existenz einer optimalen Steuerung ist gesichert, wenn die Konvergenz einer Minimalfolge gegen eine zulässige Steuerung u und die Konvergenz der zugehörigen $J(u^k)$ gegen $\min J(u)$ vorliegt. Konvergenzaussagen sind nur unter relativ starken Voraussetzungen möglich. Wir gehen deshalb bei der Beschreibung einiger Verfahren auf dieses Problem nicht näher ein.

Ist $\{u^k\}$ eine Minimalfolge, dann kann zu jedem $\varepsilon > 0$ ein $k_0(\varepsilon)$ angegeben werden, so daß

$$0 \leq J(u^k) - \inf J(u) < \varepsilon \quad \text{für alle} \quad k \geq k_0$$

gilt. Eine solche Steuerung (*fast-optimaler Steuerung*) ist für praktische Probleme im allgemeinen ausreichend. Manche Verfahren liefern eine Minimalfolge, für die außerdem noch

$$J(u^{k+1}) \leq J(u^k) \quad \forall k$$

gilt.

Wir wollen im folgenden lediglich einige Verfahren kurz skizzieren; mehr darüber findet man in Band 18. Beim Einsatz sollte man eine Beratung mit einem erfahrenen Mathematiker unbedingt durchführen.

3.4.2. Einige Abstiegsverfahren

Wir wollen an folgende Problemstellung anknüpfen: Gegeben sei

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \quad (3.62)$$

$$\mathbf{x}(t_0) = \mathbf{x}^0, \quad t \in [t_0, t_1].$$

Gesucht ist eine solche zulässige Steuerung (bei zunächst unbeschränktem Steuerbereich), daß

$$J = \sum_{i=1}^n c_i x_i(t_1) \quad (c_i = \text{const}) \quad (3.63)$$

minimal wird. Bei anderen Aufgabenstellungen sind die in Abschnitt 3.3. dargelegten Probleme zu berücksichtigen.

Wir bilden zunächst die Hamiltonfunktion

$$H(\mathbf{x}, \mathbf{p}, \mathbf{u}, t) = \sum_{i=1}^n p_i f_i \quad (3.64)$$

und das adjungierte System

$$\dot{p}_i = - \frac{\partial H}{\partial x_i}, \quad (3.65)$$

$$p_i(t_1) = -c_i, \quad i = 1, \dots, n.$$

Nun wenden wir die Idee des *Gradientenverfahrens* zur Maximierung von H an (vgl. Band 15). Es sei \mathbf{u}^1 eine Ausgangsteuerung. Wir setzen in (3.62) ein, integrieren und erhalten die zugehörige Trajektorie $\mathbf{x}^1(t)$. Einsetzen von $\mathbf{u}^1(t)$ und $\mathbf{x}^1(t)$ in (3.65) liefert nach Integration $\mathbf{p}^1(t)$. Schließlich berechnen wir den Gradienten von H bezüglich \mathbf{u} für $\mathbf{u}^1, \mathbf{x}^1, \mathbf{p}^1$, d.h.

$$\text{grad}_{\mathbf{u}} H|_{\mathbf{x}=\mathbf{x}^1, \mathbf{u}=\mathbf{u}^1, \mathbf{p}=\mathbf{p}^1}. \quad (3.66)$$

Mit geeigneten $k(t) > 0$ (Schrittweite des Gradientenverfahrens) wird als nächste Steuerung

$$\mathbf{u}^2 = \mathbf{u}^1 + k(t) \text{grad}_{\mathbf{u}} H \quad (3.67)$$

bestimmt und das Verfahren solange wiederholt, bis eine der im Gradientenverfahren üblichen Abbruchbedingungen erfüllt ist. Es empfiehlt sich dabei, die entsprechenden Werte von (3.63) mit zu vergleichen. Sind Beschränkungen für die Steuerungen gegeben, kann man z.B. das *projizierte Gradientenverfahren* sinngemäß anwenden.

Ein Verfahren, bei dem die Bildung des Gradienten durch die Maximierung von H ersetzt wird, ergibt sich folgendermaßen:

Nachdem zu \mathbf{u}^1 wie im Gradientenverfahren \mathbf{x}^1 und \mathbf{p}^1 bestimmt wurden, berechnet man zunächst $\hat{\mathbf{u}}$ aus

$$\max_{\mathbf{u} \in U} H(\mathbf{x}^1, \mathbf{p}^1, \mathbf{u}, t) = H(\mathbf{x}^1, \mathbf{p}^1, \hat{\mathbf{u}}, t) \quad (3.68)$$

und bildet die neue Steuerung

$$\mathbf{u}^2 = \mathbf{u}^1 + k(t)(\hat{\mathbf{u}} - \mathbf{u}^1), \quad k(t) > 0. \quad (3.69)$$

Für $k(t) \equiv 1$ folgt z.B. $\mathbf{u}^2 = \hat{\mathbf{u}}^1$. Auch hier ist $k(t)$ so zu wählen, daß

$$H(\mathbf{x}^2, \mathbf{p}^2, \mathbf{u}^2, t) \geq H(\mathbf{x}^1, \mathbf{p}^1, \mathbf{u}^1, t) \quad (3.70)$$

gilt. Zur Maximierung von H ist ein geeignetes Optimierungsverfahren einzusetzen, wodurch natürlich zusätzlich Schwierigkeiten entstehen.

3.4.3. Behandlung als Zwei-Punkte-Randwertaufgabe

Wir gehen von der gleichen Aufgabe wie in 3.4.2. aus. Durch Maximierung der Hamiltonfunktion sei die optimale Steuerung in Abhängigkeit von \mathbf{x} , \mathbf{p} und t bestimmt.

Es sei also

$$\mathbf{u}(\mathbf{x}, \mathbf{p}, t) \quad (3.71)$$

die optimale Steuerung. Durch Einsetzen in (3.62) und (3.65) erhalten wir

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, \mathbf{p}, t), t) = \mathbf{v}(\mathbf{x}, \mathbf{p}, t), \quad (3.72)$$

$$\dot{\mathbf{p}} = \mathbf{g}(\mathbf{x}, \mathbf{p}, \mathbf{u}, t) = \mathbf{g}(\mathbf{x}, \mathbf{p}, \mathbf{u}(\mathbf{x}, \mathbf{p}, t), t) = \mathbf{w}(\mathbf{x}, \mathbf{p}, t),$$

mit den Randbedingungen $x_i(t_0) = x_i^0$, $p_i(t_1) = -c_i$, $i = 1, \dots, n$. Setzen wir

$$\mathbf{z} = \begin{pmatrix} \mathbf{x} \\ \mathbf{p} \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix},$$

so erhalten wir das $2n$ -dimensionale System

$$\dot{\mathbf{z}} = \mathbf{F}(\mathbf{z}, t) \quad (3.73)$$

mit den Randbedingungen

$$\begin{aligned} z_i(t_0) &= x_i^0, \\ z_{n+i}(t_1) &= -c_i, \end{aligned} \quad i = 1, \dots, n,$$

(sog. *Zwei-Punkte-Randwertaufgabe* vgl. Band 18; 5.3.1.). Stellen wir die Situation bezüglich der Randbedingungen noch einmal zusammen:

$t = t_0$	$\underbrace{z_1, z_2, \dots, z_n}_{\text{gegeben}}$	$\underbrace{z_{n+1}, \dots, z_{2n}}_{\text{nicht gegeben}}$
$t = t_1$	$\underbrace{z_1, z_2, \dots, z_n}_{\text{nicht gegeben}}$	$\underbrace{z_{n+1}, \dots, z_{2n}}_{\text{gegeben}}$

Die Schwierigkeiten liegen offensichtlich darin, daß sowohl für eine Vorwärts- als auch für eine Rückwärtsintegration von (3.73) nicht alle erforderlichen Randwerte bekannt sind. Schätzt man beispielsweise die fehlenden Werte $z_{n+i}(t_0)$, so erhält man nach Integration die Werte $z_{n+i}(t_1)$, die im allgemeinen von den gegebenen Werten abweichen. Das Problem könnte gelöst werden, wenn es gelänge, eine geeignete Strategie zur Abänderung der Schätzwerte anzugeben. Eine Möglichkeit besteht im Einsatz der sog. *Regula falsi* (vgl. Band 2; 7.7.2.).

3.4.4. Anwendung der Methode von Newton-Raphson

Die *Methode von Newton-Raphson* (vgl. Band 18; 2.1.3.3.) wurde entwickelt zur Lösung eines (nichtlinearen) Gleichungssystems

$$\mathbf{g}(\mathbf{x}) = \mathbf{0}.$$

Das Wesen besteht darin, daß man zu einer Näherungslösung \mathbf{x}^n eine verbesserte Lösung aus

$$\mathbf{g}(\mathbf{x}^{n+1}) - \mathbf{g}(\mathbf{x}^n) \approx \frac{\partial \mathbf{g}}{\partial \mathbf{x}} (\mathbf{x}^{n+1} - \mathbf{x}^n) = -\mathbf{g}(\mathbf{x}^n)$$

oder kürzer

$$\Delta g \approx \frac{\partial g}{\partial x} \Delta x = -g(x^n) \quad (3.74)$$

bestimmt. Dabei ist $\frac{\partial g}{\partial x}$ die Matrix der partiellen Ableitungen $\frac{\partial g_j}{\partial x_i}$ an der Stelle x^n .

Die Lösung wird also auf eine Folge zu lösender linearer Gleichungssysteme zurückgeführt. Liegen keine Beschränkungen für die Steuerungen vor, dann ergeben sich nach Anwendung des Maximumprinzips auf die in 3.4.2. formulierte Aufgabe die notwendigen Bedingungen

$$\left. \begin{aligned} \dot{x}_i &= \frac{\partial H}{\partial p_i} = H_{p_i} = f_i, & x_i(t_0) &= x_i^0, \\ \dot{p}_i &= -\frac{\partial H}{\partial x_i} = -H_{x_i}, & p_i(t_1) &= -c_i, \end{aligned} \right\} i = 1, \dots, n, \quad (3.75)$$

$$\frac{\partial H}{\partial u_i} = H_{u_i} = 0, \quad i = 1, \dots, m,$$

also in vektorieller Schreibweise das System von $2n + m$ Gleichungen

$$\begin{aligned} \dot{x} - H_p &= 0, \\ \dot{p} + H_x &= 0, \\ H_u &= 0. \end{aligned} \quad (3.76)$$

Auf dieses System läßt sich nun die Methode von Newton-Raphson anwenden.

3.5. Systeme mit verteilten Parametern

3.5.1. Zur Problemstellung

Es seien im folgenden

x ein m -dimensionaler Vektor mit den Koordinaten x_1, \dots, x_m ,

t die Zeit,

$Q(x, t)$ eine n -dimensionale Vektorfunktion mit den Koordinaten $Q_1(x, t), \dots, Q_n(x, t)$ (Zustandsvektor oder Phasenvektor),

$u(x, t)$ eine r -dimensionale Vektorfunktion mit den Koordinaten $u_1(x, t), \dots, u_r(x, t)$ (Steuervektor),

U ein abgeschlossener Bereich eines r -dimensionalen linearen Vektorraumes (Steuerbereich),

D der Definitionsbereich von $Q(x, t)$ (bezüglich x und t).

Wir wollen in Übertragung des in 3.1 eingeführten Begriffs der zulässigen Steuerung sagen: eine Steuerung $u(x, t)$ heißt zulässig, wenn sie auf D definiert (und beschränkt) ist, eine Zerlegung von D in Teilbereiche D_j mit $\bigcup D_j = D$ existiert, so daß $u(x, t)$ stetig auf jedem D_j ist (u „bereichsweise“ stetig auf D) und wenn sie nur Werte aus U annimmt.

Ist ein System partieller Differentialgleichungen für $Q(x, t)$ mit entsprechenden Rand- und Anfangsbedingungen gegeben, das von $u(x, t)$ abhängt, so sagt man auch, es liegt ein System mit verteilten Parametern vor. Im Unterschied zu den in 3.1. bis 3.4. behandelten Problemen wird hier also der Zustand durch eine Vektorfunktion mit mehr als einer unabhängigen Variablen beschrieben.

Beispiel 3.9: Die eindimensionale induktive Erwärmung eines Körpers wird beschrieben durch

$$Q_t = aQ_{xx} + u(t)e^{-Sx},$$

$$Q(x, 0) = h(x) \quad (\text{Anfangsbedingung}),$$

dabei sind $Q(x, t)$ die Temperaturverteilung im Werkstück, x der Abstand von der Oberfläche des Werkstücks mit $0 \leq x \leq l$, t die Zeit mit $t \geq 0$, S eine Kenngröße für das Abklingen der Heizwirkung, $u(t)$ die der Induktorspannung proportionale Steuerung. Auch hier lassen sich wichtige Optimierungsforderungen stellen, z.B. kann eine solche Steuerung gesucht sein, so daß in vorgegebener Zeit t_1 eine gewünschte Temperaturverteilung $R(x)$ möglichst gut angenähert wird, d.h., daß

$$\int_0^1 [Q(x, t_1) - R(x)]^2 dx$$

minimal wird.

Es ist offensichtlich, daß man bei solchen Problemen mit noch größeren Schwierigkeiten zu rechnen hat als bei Systemen, die durch gewöhnliche Differentialgleichungen beschrieben werden. Wir wollen deshalb im folgenden lediglich ein spezielles System betrachten, ein entsprechendes Maximumprinzip angeben und kurz zwei numerische Verfahren erläutern.

3.5.2. Ein hyperbolisches System

Gegeben sei das System partieller Differentialgleichungen

$$\frac{\partial^2 Q_i}{\partial x \partial t} = f_i(x, t, Q, \frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial t}, u(x, t)), \quad i = 1, \dots, n, \quad (3.77)$$

mit $0 \leq x \leq a$, $0 \leq t \leq b$, dem Steuerbereich U , den Rand- und Anfangsbedingungen

$$Q_i(0, t) = g_i(t), \quad Q_i(x, 0) = h_i(x) \quad (3.78)$$

und den Anschlußbedingungen

$$Q_i(0, 0) = h_i(0) = g_i(0). \quad (3.79)$$

Vektoriell schreiben wir kürzer

$$\begin{aligned} \frac{\partial^2 Q}{\partial x \partial t} &= Q_{xt} = f(x, t, Q, Q_x, Q_t, u), \\ Q(0, t) &= g(t), \quad Q(x, 0) = h(x), \\ g(0) &= h(0). \end{aligned} \quad (3.80)$$

Wir setzen voraus:

1. Alle f_i seien auf D stetig-differenzierbar bezüglich x und t und bezüglich der restlichen Variablen zweimal stetig differenzierbar.
2. Alle $h_i(x)$ und $g_i(t)$ seien stetig und stückweise stetig differenzierbar.

Gesucht ist eine solche zulässige Steuerung $\mathbf{u}(x, t)$, daß die zugehörige Vektorfunktion $\mathbf{Q}(x, t)$ das Funktional

$$J = \sum_{i=1}^n c_i Q_i(a, b) \quad (3.81)$$

(c_i gegebene Konstanten) minimiert.

Zunächst wollen wir zeigen, daß einige Aufgabenstellungen, bei denen das Zielfunktional in anderer Form vorliegt, sich auf die o. g. Aufgabenstellung zurückführen lassen.

Ist

$$\int_0^a \int_0^b f_0(x, t, \mathbf{Q}, \mathbf{Q}_x, \mathbf{Q}_t, \mathbf{u}) dt dx \quad (3.82)$$

zu minimieren, führen wir durch

$$Q_0(x, t) = \int_0^x \int_0^t f_0 d\tau d\xi \quad (3.83)$$

eine zusätzliche „Zustandsfunktion“ ein. Es ergibt sich dann die zusätzliche Differentialgleichung

$$\frac{\partial^2 Q_0}{\partial x \partial t} = f_0(x, t, \mathbf{Q}, \mathbf{Q}_x, \mathbf{Q}_t, \mathbf{u}) \quad (3.84)$$

mit den Rand- bzw. Anfangsbedingungen

$$Q_0(0, t) = Q_0(x, 0) = 0 \quad (3.85)$$

und dem Optimierungsziel in der Form (3.81), d. h.

$$Q_0(a, b). \quad (3.86)$$

Ist

$$\int_0^a f_0[x, \mathbf{Q}(x, b), \mathbf{Q}_x(x, b)] dx \quad (3.87)$$

zu minimieren, setzen wir

$$Q_0(x, t) = \int_0^x f_0 d\xi \quad (3.88)$$

und erhalten

$$\frac{\partial^2 Q_0}{\partial x \partial t} = \sum_{i=1}^n \left[\frac{\partial f_0}{\partial Q_i} \cdot \frac{\partial Q_i}{\partial t} + \frac{\partial f_0}{\partial Q_{ix}} f_i \right] \quad (3.89)$$

mit

$$Q_0(0, t) = 0,$$

$$Q_0(x, 0) = \int_0^x f_0[\xi, \mathbf{h}(\xi), \mathbf{h}'(\xi)] d\xi \quad (3.90)$$

und

$$Q_0(a, b) \quad (3.91)$$

als Optimierungsziel.

Aufgabe 3.5: Es sei $F[\mathbf{Q}(a, b)]$ zu minimieren, wobei $F(\mathbf{Q})$ stetige partielle Ableitungen 2. Ordnung nach allen Q_i besitzen möge.

Führen Sie durch $Q_0(x, t) = F[\mathbf{Q}(x, t)]$ eine zusätzliche Zustandsfunktion ein. Geben Sie die entsprechende Differentialgleichung, die Anfangs- und Randbedingungen und das Optimierungsziel an.

Hinweis: Vergleichen Sie mit der entsprechenden Problemstellung bei Systemen mit konzentrierten Parametern (Abschnitt 3.3.2.).

Zur Formulierung des Maximumprinzips bezüglich der eingangs angegebenen Aufgabenstellung führen wir die Hamiltonfunktion

$$H(x, t, \mathbf{N}, \mathbf{Q}, \mathbf{Q}_x, \mathbf{Q}_t, \mathbf{u}) = \sum_{i=1}^n N_i f_i(x, t, \mathbf{Q}, \mathbf{Q}_x, \mathbf{Q}_t, \mathbf{u}) \quad (3.92)$$

ein. Dabei genüge die adjungierte Vektorfunktion $\mathbf{N}(x, t)$ dem System

$$\frac{\partial^2 N_i}{\partial x \partial t} = \frac{\partial H}{\partial Q_i} - \frac{d}{dx} \left(\frac{\partial H}{\partial Q_{ix}} \right) - \frac{d}{dt} \left(\frac{\partial H}{\partial Q_{it}} \right), \quad i = 1, \dots, n, \quad (3.93)$$

mit den Randbedingungen

$$\frac{\partial N_i}{\partial t} = - \frac{\partial H}{\partial Q_{ix}} \quad \text{für} \quad x = a, \quad (3.94)$$

den Endbedingungen

$$\frac{\partial N_i}{\partial x} = - \frac{\partial H}{\partial Q_{it}} \quad \text{für} \quad t = b \quad (3.95)$$

und den Anschlußbedingungen

$$N_i(a, b) = -c_i. \quad (3.96)$$

Dann gilt das Maximumprinzip in folgender Form

Satz 3.4: Gegeben sei (3.80) mit dem Optimierungsziel (3.81). Wenn $\mathbf{u}^0(x, t)$ optimale Steuerung ist und $\mathbf{Q}^0(x, t)$, $\mathbf{N}^0(x, t)$ die dieser Steuerung entsprechenden Lösungen von (3.80) bzw. (3.93) – (3.96) sind, so gilt

$$H(x, t, \mathbf{N}^0, \mathbf{Q}^0, \mathbf{Q}_x^0, \mathbf{Q}_t^0, \mathbf{u}^0) = \max_{\mathbf{u} \in U} H(x, t, \mathbf{N}^0, \mathbf{Q}^0, \mathbf{Q}_x^0, \mathbf{Q}_t^0, \mathbf{u}) \quad (3.97)$$

für fast alle $(x, t) \in D$.

Bemerkung: Für lineare Systeme der Form

$$\mathbf{Q}_{xt} = \mathbf{C}(x, t) \mathbf{Q}_x + \mathbf{D}(x, t) \mathbf{Q}_t + \mathbf{E}(x, t) \mathbf{Q} + \mathbf{f}(\mathbf{u}) \quad (3.98)$$

(\mathbf{C} , \mathbf{D} , \mathbf{E} sind dabei (n, n) - bzw. (n, r) -Matrizen) ist die Bedingung von Satz 3.4 nicht nur notwendig, sondern auch hinreichend.

Beispiel 3.10: Gegeben sei $Q_{xt} = Q_t + u$ mit $Q(x, 0) = Q(0, t) = 0$. Gesucht ist eine solche zulässige Steuerung u mit $|u| \leq 1$, daß

$$\int_0^1 \int_0^1 (Q + u^2) \, d\tau \, dx$$

minimal wird.

Mit

$$Q_0 = \int_0^x \int_0^t (Q + u^2) \, d\tau \, d\xi$$

erhalten wir zunächst

$$Q_{0xt} = Q + u^2, \quad Q_0(0, t) = Q_0(x, 0) = 0.$$

Die Hamiltonfunktion lautet

$$H = (Q + u^2)N_0 + (Q_t + u)N,$$

das adjungierte System

$$N_{0xt} = 0,$$

$$N_{xt} = N_0 - N_t$$

mit

$$N_{0t} = 0 \quad \text{für} \quad x = 1, \quad N_{0x} = 0 \quad \text{für} \quad t = 1,$$

$$N_t = 0 \quad \text{für} \quad x = 1, \quad N_x = -N \quad \text{für} \quad t = 1,$$

$$N(1, 1) = 0, \quad N_0(1, 1) = -1.$$

Es folgt sofort

$$N_0(x, t) = -1$$

und folglich

$$H = -Q + NQ_t - u^2 + uN = -Q + NQ_t - \left(u - \frac{N}{2}\right)^2 + \frac{N^2}{4}.$$

Damit wird H maximal für

$$u^0 = \begin{cases} \frac{N}{2} & \text{falls } |N| \leq 2, \\ \operatorname{sgn} N & \text{falls } |N| > 2. \end{cases}$$

Wir bestimmen deshalb zunächst N durch Lösung von

$$N_{xt} + N_t + 1 = 0,$$

$$N_t(1, t) = 0, \quad N_x(x, 1) = -N(x, 1), \quad N(1, 1) = 0.$$

Suchen Sie die Lösung, indem Sie die Hilfsfunktion $N_x + N = P$ einführen!

Wir wollen einen Separationsansatz verwenden. Mit

$$N(x, t) = \varphi(x) \cdot \psi(t)$$

erhalten wir

$$\begin{aligned}\varphi'(x) \cdot \psi'(t) + \varphi(x) \cdot \psi'(t) + 1 &= 0, \\ \varphi(1) \cdot \psi'(t) &= 0, \quad \varphi'(x) \cdot \psi(1) = -\varphi(x) \cdot \psi(1), \\ \varphi(1) \cdot \psi(1) &= 0.\end{aligned}$$

Durch Separation der Differentialgleichung nach den zwei unabhängigen Variablen folgt

$$\varphi' + \varphi = \lambda, \quad \psi' = -\frac{1}{\lambda}$$

mit den allgemeinen Lösungen

$$\varphi = \lambda - c_1 e^{-x}, \quad \psi = -\frac{1}{\lambda} t + c_2$$

und unter Berücksichtigung der zusätzlichen Bedingungen

$$\lambda = c_1 e, \quad c_2 = \frac{1}{\lambda}$$

schließlich

$$N(x, t) = (1 - t)(1 - e^{1-x}).$$

Für $0 \leq t \leq 1$ und $0 \leq x \leq 1$ gilt $0 \leq 1 - t \leq 1$ und $1 - e \leq 1 - e^{1-x} \leq 0$, also $-2 < 1 - e \leq N \leq 0$. Damit ist

$$u^0 = \frac{N}{2} = \frac{1}{2}(1 - t)(1 - e^{1-x}).$$

Aufgabe 3.6: Zeigen Sie, daß gilt

$$Q^0 = \frac{1}{8} t(2 - t)(e^{1-x} - e^{1+x} + 2e^x - 2),$$

$$Q_0^0(1, 1) = -\frac{1}{24}(e^2 - 4e + 5).$$

Entsprechende Maximumprinzipien lassen sich auch für andere Aufgabenstellungen formulieren. Von Interesse, insbesondere auch für praktische Problemstellungen, sind beispielsweise Optimierungsprobleme bei Systemen von partiellen Differentialgleichungen der Form

$$\frac{\partial Q_i}{\partial t} = f_i(x, t, \mathbf{Q}, \mathbf{Q}_x, \mathbf{u}), \quad i = 1, \dots, n, \quad (3.99)$$

oder auch

$$\frac{\partial^2 Q_i}{\partial t^2} = f_i(x, t, \mathbf{Q}, \mathbf{Q}_x, \mathbf{u}), \quad i = 1, \dots, n, \quad (3.100)$$

mit entsprechenden Zusatzbedingungen. Viele Prozesse werden auch über Anfangs- bzw. Randbedingungen gesteuert, wobei sogar drei unabhängige Steuereinflüsse möglich sind, und zwar über die Prozeßgleichungen, Anfangsbedingungen und Randbedingungen. Eine eingehende Behandlung ist an dieser Stelle nicht möglich.

3.5.3. Numerische Methoden

Bezüglich der numerischen Bestimmung optimaler Steuerungen bei Systemen mit verteilten Parametern gelten die in 3.4.1. gemachten Bemerkungen, wobei naturgemäß die Schwierigkeiten noch zunehmen. So kann man beispielsweise versuchen, Gradientenverfahren, bedingte oder projizierte Gradientenverfahren auszunutzen. Eine andere Möglichkeit besteht darin, daß man das gegebene Problem durch Diskretisierung bezüglich gewisser Variabler oder aller Variabler auf ein Problem mit konzentrierten Parametern oder auf ein vollständig diskretes System zurückführt. Wir wollen dieses Vorgehen lediglich an zwei Beispielen erläutern.

Beispiel 3.11: Gegeben sei die Problemstellung von Beispiel 3.9. Wir führen eine Diskretisierung bezüglich t durch, dazu bezeichnen wir

$$t_k - t_{k-1} = l, \quad k = 1, \dots, K,$$

$$Q(x, t_k) = Q_k(x), \quad u(x, t_k) = u_k(x)$$

und ersetzen $Q_t(x, t_k)$ durch

$$\frac{1}{l} [Q_k(x) - Q_{k-1}(x)].$$

Die gegebene partielle Differentialgleichung für $Q(x, t)$ geht dann über in das System gewöhnlicher Differentialgleichungen für die Funktionen $Q_k(x)$

$$Q_{kx} = Q_{k-1,x} + Q_k - Q_{k-1} + lu_k, \quad k = 1, \dots, K,$$

$$Q_0 = 0.$$

Dieses Differentialgleichungssystem können wir leicht in ein System überführen, wie es bei Systemen mit konzentrierten Parametern betrachtet wurde. Dazu schreiben wir einmal k Differentialgleichungen auf:

$$\begin{aligned} Q_{1x} &= Q_1 + lu_1, \\ Q_{2x} &= Q_{1x} + Q_2 - Q_1 + lu_2, \\ Q_{3x} &= Q_{2x} + Q_3 - Q_2 + lu_3, \\ &\dots\dots\dots \\ Q_{kx} &= Q_{k-1,x} + Q_k - Q_{k-1} + lu_k. \end{aligned}$$

Addition liefert sofort

$$\sum_{i=1}^k Q_{ix} = \sum_{i=1}^{k-1} Q_{ix} + Q_k + l \sum_{i=1}^k u_i$$

also

$$Q_{kx} = Q_k + l \sum_{i=1}^k u_i, \quad k = 1, \dots, K,$$

d.h. bezüglich der Funktionen $Q_k(x)$ ein System gewöhnlicher Differentialgleichungen. Durch Anwendung der Trapezformel (vgl. Band 2; Satz 10.13, oder Band 18; 4.3.4.) wird

$$\int_0^1 (Q + u^2) dt$$

ersetzt durch

$$\frac{l}{2} [(Q_0 + u_0^2) + (Q_K + u_K^2)] + l \sum_{i=1}^{K-1} (Q_i + u_i^2).$$

Damit erhalten wir als neues Optimierungsziel

$$\frac{1}{2} l \int_0^1 [Q_0 + u_0^2 + 2(Q_1 + u_1^2) + \dots + 2(Q_{K-1} + u_{K-1}^2) + Q_K + u_K^2] dx.$$

Für das somit entstandene System mit konzentrierten Parametern lautet die Hamiltonfunktion

$$\begin{aligned} H &= \frac{1}{2} l N_0 [Q_0 + u_0^2 + 2(Q_1 + u_1^2) + \dots + 2(Q_{K-1} + u_{K-1}^2) + Q_K + u_K^2] \\ &\quad + N_1(Q_1 + l u_1) \\ &\quad + N_2(Q_2 + l(u_1 + u_2)) \\ &\quad + \dots \\ &\quad + N_K(Q_K + l(u_1 + u_2 + \dots + u_K)) \\ &= \frac{1}{2} l N_0 Q_0 + \sum_{i=1}^{K-1} (l N_0 + N_i) Q_i + (\frac{1}{2} l N_0 + N_K) Q_K \\ &\quad + \frac{1}{2} l N_0 u_0^2 + l N_0 u_1^2 + l u_1 \sum_{i=1}^K N_i \\ &\quad + l N_0 u_2^2 + l u_2 \sum_{i=2}^K N_i \\ &\quad + \dots \\ &\quad + l N_0 u_{K-1}^2 + l u_{K-1} \sum_{i=K-1}^K N_i + \frac{1}{2} l N_0 u_K^2 + l u_K N_K \end{aligned}$$

und das adjungierte System

$$\begin{aligned} N_{0x} &= 0, \quad N_{0x}(1) = -1, \\ N_{1x} &= l - N_1, \\ N_{2x} &= l - N_2, \\ &\dots \\ N_{K-1,x} &= l - N_{K-1}, \\ N_{Kx} &= \frac{1}{2} l - N_K, \quad N_i(1) = 0, \quad i = 1, \dots, K. \end{aligned}$$

Als Lösung des adjungierten Systems ergibt sich

$$\begin{aligned} N_0 &= -1, \quad N_i = l(1 - e^{1-x}), \quad i = 1, \dots, K-1, \\ N_K &= \frac{1}{2} l(1 - e^{1-x}). \end{aligned}$$

Setzen wir

$$S_i = \sum_{j=i}^K N_j = l \left(\frac{1}{2} + K - i \right) (1 - e^{1-x}),$$

dann erhalten wir durch Maximierung von H für die optimale Steuerung

$$\begin{aligned} u_0 &= 0, \\ u_i &= \begin{cases} \frac{1}{2} S_i, & \text{falls } |S_i| \leq 2, \\ \operatorname{sgn} S_i, & \text{falls } |S_i| > 2, \end{cases} \quad i = 1, \dots, K-1, \\ u_K &= \begin{cases} S_K, & \text{falls } |S_K| \leq 1, \\ \operatorname{sgn} S_K, & \text{falls } |S_K| > 1. \end{cases} \end{aligned}$$

Mit $l \cdot K = 1$ und $i \geq 1$ folgt wegen $1 - e < 1 - e^{1-x} \leq 0$

$$-1 < l(i - \frac{1}{2}) - 1 = l(K - i + \frac{1}{2}) < \frac{1}{2} l(K - i + \frac{1}{2}) (1 - e) \leq \frac{1}{2} S_i \leq 0$$

und

$$-1 < -l \leq S_K \leq 0,$$

also

$$u_i = \begin{cases} \frac{1}{2} S_i & \text{für } i = 1, \dots, K-1, \\ S_i & \text{für } i = K. \end{cases}$$

Berücksichtigen wir noch $l \cdot i = t_i$, so ergibt sich

$$\begin{aligned} \lim_{l \rightarrow 0} u_i &= \lim_{l \rightarrow 0} \frac{1}{2} l(K - i + \frac{1}{2}) (1 - e^{1-x}) \\ &= \lim_{l \rightarrow 0} \frac{1}{2} (1 - t_i + \frac{1}{2} l) (1 - e^{1-x}) = \frac{1}{2} (1 - t_i) (1 - e^{1-x}) \end{aligned}$$

für $i < K$ und

$$\lim_{l \rightarrow 0} u_K = \lim_{l \rightarrow 0} \frac{1}{2} l(1 - e^{1-x}) = 0.$$

Die Lösung unseres Näherungsproblems konvergiert also gegen die in Beispiel 3.9 entwickelte exakte Lösung.

Beispiel 3.12: Wir betrachten die gleiche Problemstellung und diskretisieren durch

$$t_k - t_{k-1} = l, \quad k = 1, \dots, K, \quad K \cdot l = 1,$$

$$x_m - x_{m-1} = h, \quad m = 1, \dots, M, \quad M \cdot h = 1,$$

hinsichtlich beider Variablen. Zur Abkürzung setzen wir noch

$$Q(x_m, t_k) = Q(m \cdot h, k \cdot l) = Q_{mk},$$

$$u(x_m, t_k) = u(m \cdot h, k \cdot l) = u_{mk}.$$

Die Ableitungen werden durch die entsprechenden Differenzenquotienten ersetzt, also

$$Q_t(x_m, t_k) = \frac{1}{l} (Q_{mk} - Q_{m,k-1}),$$

$$Q_{xt}(x_m, t_k) = \frac{1}{n \cdot l} (Q_{mk} - Q_{m-1,k} - Q_{m,k-1} + Q_{m-1,k-1}),$$

wobei noch

$$Q_{0k} = Q_{m0} = 0$$

gilt. Damit geht die Differentialgleichung über in das Gleichungssystem für die $M \cdot K$ Funktionswerte Q_{mk}

$$\frac{1}{n \cdot l} (Q_{mk} - Q_{m-1,k} - Q_{m,k-1} + Q_{m-1,k-1}) = \frac{1}{l} (Q_{mk} - Q_{m,k-1}) + u_{mk},$$

d.h.

$$Q_{mk} - Q_{m,k-1} = \frac{1}{1-h} (Q_{m-1,k} - Q_{m-1,k-1}) + \frac{h \cdot l}{1-h} u_{mk}.$$

Die Lösung dieses Systems von $M \cdot K$ Gleichungen für $M \cdot K$ Unbekannte ist leicht möglich. Es ergibt sich (Entwickeln Sie die Lösung!)

$$Q_{mk} = \frac{h \cdot l}{1-h} \sum_{\alpha=1}^m \left[\frac{1}{(1-h)^{m-\alpha}} \sum_{\beta=1}^k u_{\alpha\beta} \right].$$

Nun diskretisieren wir das Optimierungsziel, indem wir das Doppelintegral

$$J = \int_0^1 \int_0^1 (Q + u^2) dt dx$$

durch

$$J_{MK} = h \cdot l + \sum_{m=1}^M \sum_{k=1}^K (Q_{mk} + u_{mk}^2)$$

ersetzen. Es folgt weiter (wobei $g = \frac{1}{1-h}$ gesetzt wurde)

$$\frac{\partial J_{MK}}{\partial u_{\alpha\beta}} = hl \left(2u_{\alpha\beta} + \sum_{m=1}^M \sum_{k=1}^K \frac{\partial Q_{mk}}{\partial u_{\alpha\beta}} \right)$$

mit

$$\frac{\partial Q_{mk}}{\partial u_{\alpha\beta}} = \begin{cases} hlg^{m-\alpha+1} & \text{für } \alpha \leq m, \beta \leq k, \\ 0 & \text{sonst.} \end{cases}$$

Aus $\frac{\partial J_{MK}}{\partial u_{\alpha\beta}} = 0$ erhalten wir somit

$$u_{\alpha\beta} = \frac{1}{2} (K - \beta + 1) (1 - g^{M-\alpha+1}).$$

Aufgabe 3.7: Überprüfen Sie, daß die hinreichende Bedingung für ein Minimum tatsächlich erfüllt ist, $u_{\alpha\beta} \leq 1$ gilt und die Näherungslösung für $l \rightarrow 0$, $h \rightarrow 0$ gegen die exakte Lösung konvergiert.

4. Diskrete Prozesse

Von den steuerbaren, deterministischen Prozessen werden jetzt diejenigen untersucht, die im Abschnitt 2. als diskrete bezeichnet worden sind. Dabei wird die Grundaufgabe der Theorie optimaler diskreter Prozesse einschließlich gewisser Modifizierungen (siehe Abschn. 4.1.) formuliert sowie zwei Kriterien für die Optimalität und damit zusammenhängende Lösungsmethoden entwickelt (siehe Abschn. 4.2. und 4.3.). Von den beiden Methoden wird insbesondere die dynamische Optimierung ausführlicher dargelegt und ihre Anwendung an praktischen Beispielen demonstriert.

Bei der Behandlung stetiger Prozesse bezeichnete die Variable t i. a. die Zeit. Davon müssen wir uns im Zusammenhang mit diskreten Prozessen lösen. Hier numeriert t einfach die Folge endlich vieler oder abzählbar unendlich vieler unterschiedlicher Zustände bzw. getrennter Stufen. Dabei kann der Unterschied zeitlicher Natur sein, muß es aber nicht. Es sei z. B. eine Mehrstufenrakete mit begrenztem Startgewicht zu konstruieren, die einen Flugkörper in den Kosmos tragen soll. Jede Stufe enthält einen gewissen Treibstoffvorrat; ist er verbraucht, wird die Stufe abgeworfen und die nächste gezündet. Gewicht des Flugkörpers und Anzahl der Raketenstufen seien bekannt. Dann ist es eine natürliche Aufgabe, den Treibstoff so über die einzelnen Stufen der Rakete zu verteilen, daß die Geschwindigkeit des Flugkörpers nach Verbrauch des gesamten Treibstoffes maximal ist. Diese Aufgabe läßt sich als Optimierungsproblem eines diskreten Prozesses formulieren (vgl. [7]), wobei seine „unterschiedlichen Zustände“ einfach aus den einzelnen Stufen der Rakete gebildet werden. Numeriert man nun diese unterschiedlichen Zustände mit t , $t = 1, 2, \dots, N$ (N – Anzahl der Raketenstufen), so hat t offensichtlich keinerlei Beziehungen zur Zeit.

4.1. Aufgabenstellung

Die Grundaufgabe der Theorie optimaler diskreter Prozesse und die mit ihr in Zusammenhang stehenden Begriffe haben eine Reihe Gemeinsamkeiten mit denen stetiger Prozesse, wie sie am Anfang von Kapitel 3. und im Abschnitt 3.1. dargelegt worden sind. Auf Grund dieser Analogien kann das Wesentliche hier in geraffter Form ohne ausführliche Erläuterungen dargelegt werden. Lediglich die Unterschiede, die sich zwischen den Darlegungen diskreter und stetiger Prozesse ergeben, werden stärker herausgearbeitet.

Zur Entwicklung der Aufgabenstellung wird zunächst ein Beispiel betrachtet, auf das wir uns im weiteren häufig beziehen.

Beispiel 4.1: In einer Lösung befindet sich ein gewisser Stoff, der durch Spülung mit Wasser herausgefiltert werden soll (vgl. [9] oder [7]). Es wird vorausgesetzt, daß sich dabei das Wasser nicht mit der Lösung vermischt. Die Spülung wird in N gleichartigen Aggregaten durchgeführt, die die Lösung nacheinander durchfließt (siehe Bild 4.1 a)). Es sei bekannt, daß die durch ein Aggregat ausgespülte Stoffmenge gleich $m(x, u)$ ist, wobei x die Stoffmenge angibt, die zu Beginn des Spülprozesses mit

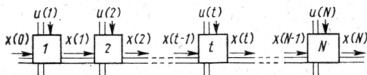


Bild 4.1 a) System aus N in Reihe geschalteten Aggregaten

der Lösung in das Aggregat geflossen ist, und u die bei der Spülung im Aggregat verwendete Wassermenge bezeichnet. Dann ist die Stoffmenge $x(t)$, die das t -te Aggregat nach der Spülung zusammen mit der Lösung verläßt, gleich der Stoffmenge $x(t-1)$, die mit der Lösung in ihn eingeflossen ist, vermindert um die Menge $m(x(t-1), u(t))$, die durch ihn ausgefiltert wurde:

$$x(t) = x(t-1) - m(x(t-1), u(t)), \quad t = 1, 2, \dots, N, \quad (4.1)$$

wobei $x(0) = x_0$ die Stoffmenge bezeichnet, die in der Lösung enthalten ist, welche in das erste Aggregat gegeben wird. Weiterhin möge das Ergebnis des gesamten Spülungsprozesses gemessen werden können durch die Größe

$$I = a_1 [x(0) - x(N)] - a_2 \sum_{t=1}^N u(t), \quad (4.2)$$

wobei a_1 und a_2 bekannte Parameter sind; hierin gibt die Differenz $x(0) - x(N)$ die insgesamt ausgespülte Stoffmenge an, und die Summe $\sum_{t=1}^N u(t)$ stellt die verwendete Wassermenge dar. Offensichtlich hängt dieses Ergebnis von der Größe der in den einzelnen Aggregaten verwendeten Wassermengen $u(t)$ ab und ist durch sie vermittels (4.1) eindeutig bestimmt. Somit kann die Aufgabe gestellt werden diejenigen $u(t)$, $t = 1, 2, \dots, N$, zu bestimmen, für die das Spülungsergebnis I maximiert wird. Beachtet man nun noch, daß die verwendeten Wassermengen einer Bedingung der Art

$$0 \leq u(t) \leq b(x(t-1)), \quad t = 1, 2, \dots, N, \quad (4.3)$$

unterworfen werden müssen, wobei $b(x)$ eine bekannte Funktion ist, die sich aus der Konstruktion der Filteraggregate ergibt, so nimmt die obige Aufgabe folgende mathematische Form an: es ist eine solche Folge von Werten $u(1), u(2), \dots, u(N)$ und die ihnen gemäß (4.1) entsprechende Folge $x(1), x(2), \dots, x(N)$ zu bestimmen, die der Beschränkung (4.3) genügen und der Größe (4.2) ihren maximalen Wert erteilen.

Erinnert man sich nun an die allgemeinen Ausführungen über steuerbare deterministische Prozesse des Kapitels 2, so kann festgestellt werden, daß die mathematische Aufgabe, zu der das Beispiel 4.1 führte, bereits alle wesentlichen Elemente der Theorie optimaler diskreter Prozesse enthält. Tatsächlich, es liegen endlich viele — nämlich N — unterschiedliche „Zustände“ vor, $u(t)$ kann als Steuerung, $x(t)$ als Phasenkoordinate und I als Optimalitätskriterium aufgefaßt werden. Dann ist durch (4.3) der Steuerbereich und durch (4.1) die Prozeßgleichung gegeben. Somit liefert das Beispiel 4.1 eine Vorstellung von der Grundaufgabe optimaler diskreter Prozesse.

Im allgemeinen treten als Steuerungen bzw. Phasenkoordinaten bei diskreten Prozessen — genau wie bei stetigen — nicht nur jeweils eine, sondern mehrere Größen $u_1(t), u_2(t), \dots, u_r(t)$ bzw. $x_1(t), x_2(t), \dots, x_n(t)$ auf, die in gewohnter Weise als Komponenten von Vektoren

$$\mathbf{u}(t) = (u_1(t), \dots, u_r(t)) \quad \text{bzw.} \quad \mathbf{x}(t) = (x_1(t), \dots, x_n(t))$$

eines r - bzw. n -dimensionalen Raumes aufgefaßt werden. Jetzt nimmt t jedoch nur diskrete Werte $t = 0, 1, \dots, N$ an, wobei im weiteren vorausgesetzt wird, daß N eine fixierte natürliche Zahl ist. Bei stetigen Prozessen waren Steuer- und Phasenvektor durch das Differentialgleichungssystem (3.6) miteinander verknüpft. Bei diskreten Prozessen wird die Rolle von (3.6) durch das Differenzengleichungssystem

$$\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t-1), \mathbf{u}(t)), \quad t = 1, 2, \dots, N, \quad (4.4)$$

übernommen, wobei $\mathbf{f}(\mathbf{x}, \mathbf{u}) = (f_1(\mathbf{x}, \mathbf{u}), \dots, f_n(\mathbf{x}, \mathbf{u}))$ ein Vektor ist, dessen Komponenten bekannte Funktionen sind. Man erhält diese Funktionen für jeden konkret gege-

benen diskreten Prozeß durch die mathematische Modellierung seiner charakteristischen Eigenschaften. So ergab sich im Beispiel 4.1 für das Differenzengleichungssystem (4.4) nur die eine Gleichung (4.1), und es könnte

$$f(x, u) = x - m(x, u)$$

gesetzt werden.

Ist nun der Anfangszustand $\mathbf{x}(0)$ eines diskreten Prozesses bekannt – und das wird im folgenden immer vorausgesetzt –, so sind nach Wahl einer Steuerung $\mathbf{u}(t)$, $t = 1, 2, \dots, N$, alle weiteren Zustände $\mathbf{x}(t)$, $t = 1, 2, \dots, N$, des diskreten Prozesses durch das System (4.4) eindeutig bestimmt. Daher wird (4.4) auch hier (vergleiche mit (3.6) in Abschn. 3.1.) Prozeß- oder Bewegungsgleichung genannt. Der Terminus „Bewegung“ ist nun allerdings im weiteren Sinne aufzufassen, weil ja t nicht mehr in jedem Falle die Zeit repräsentiert.

Es sind noch die Begriffe „Steuerbereich“ und „zulässige Steuerungen“ präziser zu fassen. Im allgemeinen wird bei diskreten Prozessen davon ausgegangen, daß im r -dimensionalen Raum der Steuerungen R^r für jedes \mathbf{x} aus dem Phasenraum und jedes $t = 1, 2, \dots, N$ eine nichtleere Menge $U_t(\mathbf{x})$ gegeben ist. Diese Menge wird Steuerbereich genannt, der dem Wert t und dem Phasenzustand \mathbf{x} entspricht. In Anlehnung an Beispiel 4.1 und an weitere praktische Probleme (vgl. [7], [9], [19]) beschränken wir uns hier auf Steuerbereiche der Art $U_t(\mathbf{x}(t-1))$. Dementsprechend wird folgende Begriffsbildung vorgenommen:

Definition 4.1: Gegeben sei ein diskreter Prozeß mit dem Anfangszustand $\mathbf{x}(0)$ und den Steuerbereichen $U_t(\mathbf{x}) \subseteq R^r$ ($\mathbf{x} \in R^n$, $t = 1, 2, \dots, N$). Dann wird $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_r(t))$ eine **zulässige Steuerung** (bezüglich $\mathbf{x}(0)$ und $U_t(\mathbf{x})$) genannt, wenn

$$\mathbf{u}(t) \in U_t(\mathbf{x}(t-1)), \quad t = 1, 2, \dots, N, \quad (4.5)$$

gilt; dabei ist $\mathbf{x}(t)$ die Trajektorie, die in $\mathbf{x}(0)$ beginnt und durch (4.4) bestimmt ist.

Im Beispiel 4.1 ist der Steuerbereich durch (4.3) gegeben, hängt nicht explizit von t ab und hat die Form

$$U(\mathbf{x}(t-1)) = \{u \in R^1 \mid 0 \leq u \leq b(\mathbf{x}(t-1))\}.$$

Jede Lösung von (4.4) und (4.5), bestehend aus $\mathbf{u}(t)$ und $\mathbf{x}(t)$, $t = 1, 2, \dots, N$, stellt einen diskreten Prozeß dar. Man erhält eine solche Lösung ausgehend von einem gegebenen Anfangszustand $\mathbf{x}(0)$, indem man zunächst gemäß (4.5) eine zulässige Steuerung $\mathbf{u}(1) \in U_1(\mathbf{x}(0))$ wählt, diese dann in (4.4) einsetzt und danach $\mathbf{x}(1) = \mathbf{f}(\mathbf{x}(0), \mathbf{u}(1))$ bestimmt; nun wird $\mathbf{x}(1)$ wie $\mathbf{x}(0)$ zur Bestimmung von $\mathbf{u}(2)$ verwendet, und dieses Vorgehen wird bis zur Ermittlung von $\mathbf{u}(N)$ und $\mathbf{x}(N)$ fortgesetzt.

Für das weitere wird noch folgender Begriff benötigt:

Definition 4.2: Es sei $\mathbf{x}(t)$, $t = 1, 2, \dots, N$, die Lösung von (4.4), die man bei gegebenem Anfangszustand $\mathbf{x}(0)$ und für eine zulässige Steuerung $\mathbf{u}(t)$ erhalten hat. Dann werden wir $\mathbf{x}(t)$ die der zulässigen Steuerung $\mathbf{u}(t)$ entsprechende Trajektorie nennen, die in $\mathbf{x}(0)$ beginnt.

Es sei bemerkt, daß im Unterschied zu den stetigen Prozessen die Trajektorie hier nur eine diskrete Punktfolge im n -dimensionalen Phasenraum bildet.

Nach der mathematischen Beschreibung diskreter Prozesse durch (4.4) und (4.5) kann nun zur Frage der optimalen Steuerung derartiger Prozesse übergegangen werden. Bezüglich des dazu erforderlichen Optimalitätskriteriums wird angenommen, daß eine Größe der Form

$$J = \sum_{t=1}^N f_{0t}(\mathbf{x}(t-1), \mathbf{u}(t)) \quad (4.6)$$

maximiert bzw. minimiert werden soll. Dabei sind $f_{0t}(\mathbf{x}, \mathbf{u})$, $t = 1, 2, \dots, N$, bekannte Funktionen von $n + r$ Argumenten, die sich aus der Spezifik des jeweiligen Prozesses ergeben und gewissermaßen das Ergebnis des Prozesses auf jeder seiner Stufe messen. So kann im Beispiel 4.1 die zu maximierende Größe (4.2) unter Berücksichtigung von (4.1) umgeformt werden auf

$$J = \sum_{t=1}^N [a_1 m(x(t-1), u(t)) - a_2 u(t)];$$

hier hängen also die in (4.6) auftretenden Funktionen gar nicht von t ab, und es könnte gesetzt werden

$$f_0(\mathbf{x}(t-1), \mathbf{u}(t)) = a_1 m(\mathbf{x}(t-1), \mathbf{u}(t)) - a_2 u(t).$$

Existieren nun für einen diskreten Prozeß mit gegebenem Anfangszustand $\mathbf{x}(0)$ mehrere zulässige Steuerungen, so entsteht die Frage nach der „besten“ oder „optimalen“ Steuerung. Mathematisch führt diese Frage zur folgenden Grundaufgabe der Bestimmung optimaler diskreter Prozesse:

Definition 4.3: Gegeben sei ein diskreter Prozeß mit dem Anfangszustand $\mathbf{x}(0)$, der Bewegungsgleichung (4.4), den Steuerbereichen $U_t(\mathbf{x}) \subseteq R^r$ und dem Optimalitätskriterium (4.6). Dann ist unter allen zulässigen Steuerungen $\mathbf{u}(t)$, $t = 1, 2, \dots, N$, diejenige zu bestimmen, die der Größe (4.6) ein Maximum (bzw. ein Minimum) erteilt. Dabei ist in (4.6) die der Steuerung $\mathbf{u}(t)$ entsprechende Trajektorie $\mathbf{x}(t)$ einzusetzen, die in $\mathbf{x}(0)$ beginnt. Wir werden sagen, daß eine zulässige Steuerung und die ihr entsprechende Trajektorie, die die Grundaufgabe lösen, einen **optimalen Prozeß** bilden.

Die Grundaufgabe der Theorie optimaler stetiger Prozesse (vgl. Def. 3.3) war charakterisiert durch freie Zeit und feste Enden. Im Gegensatz dazu kann die Grundaufgabe optimaler diskreter Prozesse als Aufgabe mit fester Zeit (N ist fixiert und gegeben), einem festen Ende ($\mathbf{x}(0)$) und einem freien Ende ($\mathbf{x}(N)$) betrachtet werden. Der wesentliche Unterschied ist vor allem darin zu sehen, daß wir es einmal mit freier Zeit und zum anderen mit fester Zeit zu tun haben. Die Ursache dafür liegt darin, daß für diskrete Prozesse meistens die Anzahl der Zustände gegeben ist und nicht der Optimierung unterliegt.

Genau wie in Kapitel 3. existieren natürlich auch Modifizierungen für die in der Definition 4.3 formulierte Grundaufgabe. Einige seien hier kurz genannt. Weitere findet man in [7] und [9].

So treten z. B. Aufgaben mit beweglichen Enden auf. In diesem Falle sind zwei Mengen X_0 und X_N im Phasenraum R^n gegeben, und es sind ein solcher Anfangszustand $\mathbf{x}(0) \in X_0$ sowie eine solche zulässige Steuerung $\mathbf{u}(t)$ (bezüglich $\mathbf{x}(0)$) gesucht, die zusammen mit der ihr entsprechenden Trajektorie

$\mathbf{x}(t)$, die in $\mathbf{x}(0)$ beginnt, der Größe (4.6) ihren maximalen Wert erteilt, wobei außerdem $\mathbf{x}(N) \in X_N$ gelten muß. Entarten hierbei X_0 und X_N in zwei fixierte Punkte $\mathbf{x}(0)$ und $\mathbf{x}(N)$, so erhält man als Spezialfall die Aufgabe mit festen Enden.

Auch Aufgaben der folgenden Art ergeben sich: Für jedes $t = 0, 1, \dots, N$ sei eine Menge X_t im Phasenraum gegeben. Dann sind ein solcher Anfangszustand $\mathbf{x}(0) \in X_0$ sowie eine solche (bezüglich $\mathbf{x}(0)$) zulässige Steuerung $\mathbf{u}(t)$ gesucht, die zusammen mit der ihr entsprechenden Trajektorie $\mathbf{x}(t)$, die in $\mathbf{x}(0)$ beginnt, der Größe (4.6) ihren maximalen Wert erteilt, wobei außerdem $\mathbf{x}(t) \in X_t$ gelten muß. Das sind Aufgaben mit beschränkten Phasenkoordinaten. Entarten hierbei die Mengen X_t zum ganzen Phasenraum: $X_t = R^n$, $t = 0, 1, \dots, N$, so ergibt sich als Spezialfall die Aufgabe mit freien Enden.

Schließlich seien noch Prozesse erwähnt, bei denen sich die in (4.4) auftretenden Funktionen \mathbf{f} von Stufe zu Stufe ändern und damit explizit von t abhängen. Die Bewegungsgleichung (4.4) nimmt dann die Form

$$\mathbf{x}(t) = \mathbf{f}_t(\mathbf{x}(t-1), \mathbf{u}(t)), \quad t = 1, 2, \dots, N,$$

an. Das ist ein Analogon zu den nichtautonomen stetigen Prozessen (vgl. Abschn. 3.3.3.). Hätten wir im Beispiel 4.1 nicht gleichartige, sondern von Stufe zu Stufe unterschiedliche Aggregate vorausgesetzt, dann wäre auch die ausgefilterte Stoffmenge von den Stufen abhängig gewesen, und anstelle von (4.1) hätte sich

$$x(t) = x(t-1) - m_t(x(t-1), u(t)), \quad t = 1, 2, \dots, N,$$

ergeben. Damit wäre ein konkreter Prozeß mit einer Bewegungsgleichung der obigen Art vorgestellt.

Im Zusammenhang mit der zu entwickelnden dynamischen Optimierung als einer Lösungsmethode der Grundaufgabe sei auf einige Änderungen in den Bezeichnungen von (4.4)–(4.6) hingewiesen. Dem Vorgehen der dynamischen Optimierung entspricht es besser, wenn die Stufen des Prozesses entgegen dem Prozeßablauf numeriert werden (vgl. Bild 4.1 b, Seite 58). Außerdem werden die in der Prozeßgleichung (4.4) auftretenden Funktionen mit T bezeichnet; diese Bezeichnung geht davon aus, daß in jeder Stufe die Eingangsgrößen zusammen mit der Steuerung einer Transformation unterworfen werden. Schließlich werden die Funktionen im Optimalitätskriterium (4.6) mit G bezeichnet, was damit zusammenhängt, daß es bei einer Reihe praktischer diskreter Prozesse darauf ankommt, in jeder Stufe einen Gewinn zu erzielen. Mit diesen Änderungen nehmen die Relationen (4.4) bis (4.6) dann die folgende Form an:

$$\mathbf{x}_n = \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n), \quad n = 1, 2, \dots, N, \quad (4.7)$$

$$\mathbf{u}_n \in U_n(\mathbf{x}_{n+1}), \quad n = 1, 2, \dots, N, \quad (4.8)$$

$$G = \sum_{n=1}^N G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) = \max! \quad (4.9)$$

Die Formulierung der Grundaufgabe selbst bleibt unverändert; es wird lediglich anstelle von (4.4) – (4.6) auf die entsprechenden neuen Relationen (4.7) – (4.9) Bezug genommen.

4.2. Diskretes Maximumprinzip

Es wird jetzt eine erste Methode zur Lösung der in 4.1. formulierten Grundaufgabe entwickelt. Sie hat sehr enge Beziehungen zu dem für stetige Prozesse bereits bekannten Pontrjaginschen Maximumprinzip (vgl. Abschn. 3.2.) und kann aus diesem sogar durch einen formalen Übergang von stetigen zu diskreten Prozessen hergeleitet werden

(siehe [7]). Das Wesen der Methode besteht genau wie beim Pontrjaginschen Maximumprinzip darin, die Grundaufgabe durch eine andere Aufgabe zu ersetzen, die i. a. leichter gelöst werden kann und besser für eine numerische Behandlung geeignet ist. Dabei erweist sich die „Ersatzaufgabe“ ebenfalls als eine Optimierungsaufgabe, und mit ihrer Lösung erhält man gleichzeitig die Lösung der Grundaufgabe.

Im allgemeinen besteht die Methode in der Anwendung eines notwendigen Optimalitätskriteriums für zulässige Steuerungen. Dabei unterscheidet sich dieses notwendige Kriterium für beliebige diskrete Prozesse jedoch in wesentlichen Teilen von dem Pontrjaginschen Maximumprinzip für stetige Prozesse. Das ist im Wesen der Grundaufgabe diskreter Prozesse begründet (vgl. Anlage 3 in [9]). Allerdings kann man Klassen diskreter Prozesse angeben, für die ein Analogon zum Pontrjaginschen Maximumprinzip gültig ist. Um einerseits diese Analogie zu den Ergebnissen von Abschnitt 3.2. herzustellen und um andererseits Vereinfachungen in der Darstellung zu erreichen, werden wir im weiteren voraussetzen, daß die Funktionen f in (4.4) sowie f_0 in (4.6) linear in den Phasenkoordinaten x sind und f_0 nicht explizit von t abhängt. Bezüglich der Steuerbereiche wird vorausgesetzt, daß sie weder von t noch von x abhängen, also für alle Stufen des Prozesses und für alle Phasenpunkte x gleich sind: $U_i(x) = U$, $t = 1, 2, \dots, N$, $x \in R^n$ beliebig.

Wir betrachten also jetzt die Grundaufgabe (4.4)–(4.6) mit den soeben getroffenen Vereinfachungen. Zur Formulierung ihrer „Ersatzaufgabe“ werden genau wie beim Pontrjaginschen Maximumprinzip einige neue Begriffe benötigt.

Zunächst werden durch das Differenzengleichungssystem

$$p_i(t-1) = \sum_{k=0}^n \frac{\partial f_k(x(t-1), u(t))}{\partial x_i} p_k(t), \quad (4.10)$$

$$t = 1, 2, \dots, N, \quad i = 1, 2, \dots, n,$$

gewisse Hilfsfunktionen eingeführt. Dabei wird angenommen, daß

$$p_0 = 1, \quad (4.11)$$

$$p_i(N) = 0, \quad i = 1, 2, \dots, n, \quad (4.12)$$

gilt. Bezüglich dieser Bedingungen vergleiche man die Ausführungen über die Aufgabe mit fester Zeit und freiem rechten Ende am Schluß von Abschnitt 3.2. Mit den Bedingungen (4.11) und (4.12) besitzt (4.10) für jede zulässige Steuerung $u(t)$, $t = 1, 2, \dots, N$, und die ihr entsprechende Trajektorie $x(t)$, $t = 1, 2, \dots, N$, die in $x(0)$ beginnt, eine eindeutig bestimmte Lösung

$$P(t) = (p_0, p_1(t), \dots, p_n(t)), \quad t = 0, 1, \dots, N.$$

Diese Lösung erhält man, indem man in (4.10) zunächst $t = N$ setzt und die Bedingungen (4.11) und (4.12) sowie die gewählte zulässige Steuerung und die ihr entsprechende Trajektorie verwendet. Dadurch findet man $p_i(N-1)$, $i = 1, 2, \dots, n$. Danach wird in (4.10) $t = N-1$ gesetzt, und unter Verwendung der bereits gefundenen $p_i(N-1)$ können nun die Werte $p_i(N-2)$ bestimmt werden. Auf diese Weise wird die gesamte Lösung $P(t)$, $t = 0, 1, 2, \dots, N$, sukzessive – beginnend mit $t = N-1$ und fortschreitend bis $t = 0$ – ermittelt. Jede auf diese Weise ermittelte Folge von Vektoren $P(t)$, $t = 0, 1, \dots, N$, werden wir Lösung von (4.10) – (4.12) nennen; die der zulässigen Steuerung $u(t)$ und ihrer Trajektorie $x(t)$ entspricht.

Auch für diskrete Prozesse wird wieder durch die Definitionsgleichung

$$\mathcal{H}(\mathbf{P}, \mathbf{x}, \mathbf{u}) = \sum_{i=0}^n p_i f_i(\mathbf{x}, \mathbf{u}) \quad (4.13)$$

eine Hamiltonfunktion eingeführt. Mit ihr kann dann das System (4.10) auch in der Form

$$p_i(t-1) = \frac{\partial \mathcal{H}(\mathbf{P}(t), \mathbf{x}(t-1), \mathbf{u}(t))}{\partial x_i}, \quad i = 1, 2, \dots, N,$$

geschrieben werden. Wesentlich ist jedoch, daß mit ihrer Hilfe unmittelbar die „Ersatzaufgabe“ der Grundaufgabe formuliert werden kann. Es erweist sich nämlich, daß folgende Aussage gültig ist.

Satz 4.1 (diskretes Maximumprinzip): *Dafür, daß eine zulässige Steuerung $\mathbf{u}(t)$, $t = 1, 2, \dots, N$, und die ihr entsprechende Trajektorie $\mathbf{x}(t)$, die in $\mathbf{x}(0)$ beginnt, einer Grundaufgabe mit zu maximierender Größe (4.6) optimal (im Sinne der Definition 4.3) sind, ist notwendig und hinreichend, daß für sie zusammen mit der ihnen entsprechenden Lösung $\mathbf{P}(t)$, $t = 1, 2, \dots, N$, von (4.10) – (4.12) gilt:*

$$\mathcal{H}(\mathbf{P}(t), \mathbf{x}(t-1), \mathbf{u}(t)) = \max_{\mathbf{u} \in U} \mathcal{H}(\mathbf{P}(t), \mathbf{x}(t-1), \mathbf{u}), \quad t = 1, 2, \dots, N. \quad (4.14)$$

Wenn die Größe (4.6) in der Grundaufgabe nicht maximiert, sondern minimiert werden soll, dann ist (4.11) zu ersetzen durch

$$p_0 = -1. \quad (4.11')$$

Mit (4.14) ist die Ersatzaufgabe formuliert, von der oben bereits gesprochen wurde. Sie erweist sich im Gegensatz zur entsprechenden Aufgabe des Pontrjaginschen Maximumprinzips sogar als äquivalent zur Grundaufgabe, weil das diskrete Maximumprinzip als notwendiges und hinreichendes Kriterium formuliert werden konnte. In diesem Zusammenhang muß unbedingt noch einmal betont werden, daß das diskrete Maximumprinzip in Form des Satzes 4.1 keine Allgemeingültigkeit besitzt, sondern nur auf die Klasse der diskreten Prozesse angewendet werden kann, die die obigen vereinfachenden Voraussetzungen erfüllen. Die Äquivalenz zwischen Grundaufgabe und Ersatzaufgabe (4.14) wurde also erkaufte durch die Einschränkung der Klasse der zugelassenen diskreten Prozesse. Wie wesentlich diese Einschränkung für die Gültigkeit des diskreten Maximumprinzips in der Formulierung von Satz 4.1 ist, möge das folgende Beispiel (vgl. Beispiel 10.3 in [7]) demonstrieren. Gleichzeitig wird im zweiten Teil dieses Beispiels angedeutet, wie man mit Hilfe des diskreten Maximumprinzips optimale Prozesse ermitteln kann.

Beispiel 4.2: Betrachtet wird die Grundaufgabe, deren Phasen- und Steuerungsraum eindimensional sind, deren Bewegungsgleichung folglich nur aus einer Gleichung besteht, die von der Form

$$\dot{x}(t) = u(t), \quad t = 1, 2, \dots, N, \quad (4.15)$$

sei; weiter möge bezüglich des Steuerbereiches gelten: $U = [-1, 1]$, und das Optimalitätskriterium (4.6) möge die Form

$$I = \sum_{t=1}^N f_0(x(t-1), u(t)) = \max!$$

mit

$$f_0(x(t-1), u(t)) = u^2(t) - 2x^2(t-1) \quad (4.16)$$

haben. Schließlich laute die Anfangsbedingung $x(0) = 0$.

Für diese Grundaufgabe sind die obigen vereinfachenden Voraussetzungen nicht erfüllt, denn die Funktion f_0 ist offensichtlich nicht linear in x . Diese Verletzung der Voraussetzungen gestattet es uns, eine optimale Steuerung der betrachteten Grundaufgabe anzugeben, für die die Hamiltonfunktion nicht ihr Maximum auf U annimmt. Damit ist dann gezeigt, daß das diskrete Maximumprinzip für diese Aufgabe keine notwendige Bedingung darstellt. Eine Steuerung mit den gewünschten Eigenschaften ist

$$u^*(t) = 0, \quad t = 1, 2, \dots, N-1, \quad u^*(N) = 1. \quad (4.17)$$

Sie ist auf jeden Fall zulässig; ihre entsprechende Trajektorie lautet wegen (4.15) und der Anfangsbedingung

$$x^*(t) = 0, \quad t = 0, 1, \dots, N-1, \quad x^*(N) = 1. \quad (4.18)$$

Mit diesen Werten ergibt sich $I = 1$. Allgemein folgt aber für jede zulässige Steuerung $u(t) \in U$ und der ihr entsprechenden Trajektorie $x(t)$ unter Berücksichtigung von (4.15) und (4.16) sowie der Anfangsbedingung $x(0) = 0$ die Abschätzung

$$\begin{aligned} I &= u^2(1) + \sum_{t=2}^N [u^2(t) - 2u^2(t-1)] \\ &= u^2(1) + \sum_{t=2}^N u^2(t) - 2 \sum_{t=2}^N u^2(t-1) = u^2(N) - \sum_{t=1}^{N-1} u^2(t) \\ &\leq u^2(N) \leq 1. \end{aligned}$$

Damit ist zunächst einmal gezeigt, daß die Steuerung (4.17) mit der ihr entsprechenden Trajektorie $x^*(t)$ einen optimalen Prozeß der betrachteten Grundaufgabe bilden.

Was ergibt nun das diskrete Maximumprinzip? Für die Hilfsfunktionen gilt allgemein $p_0 = 1$, $p_1(N) = 0$ und

$$p_1(t-1) = \frac{\partial f_0(x(t-1), u(t))}{\partial x} = -4x(t-1).$$

Daher ergibt sich für die gewählte Steuerung (4.17) und die ihr entsprechende Trajektorie $x^*(t)$

$$p_1^*(t) = 0, \quad t = 1, 2, \dots, N-1. \quad (4.19)$$

Die Hamiltonfunktion lautet wegen (4.15) und (4.16)

$$\mathcal{H}(P(t), x(t-1), u(t)) = [u^2(t) - 2x^2(t-1)] + p_1(t)u(t).$$

Daher folgt für die der optimalen Steuerung (4.17) entsprechenden Trajektorie (4.18) und der zugehörigen Lösung (4.19)

$$\max_{u \in U} \mathcal{H}(P^*(t), x^*(t-1), u) = \max_{u \in U} u^2 = 1, \quad t = 1, 2, \dots, N-1,$$

wogegen andererseits

$$\mathcal{H}(P^*(t), x^*(t-1), u^*(t)) = 0, \quad t = 1, 2, \dots, N-1,$$

gilt. Damit ist aber das Maximumprinzip für diese Grundaufgabe auf den Stufen $t = 1, 2, \dots, N-1$ nicht erfüllt.

Ändern wir die zu Beginn des Beispiels formulierte Grundaufgabe nun einmal so ab, daß die vereinfachenden Voraussetzungen erfüllt sind. Dazu genügt es, (4.16) durch

$$f_0(x(t-1), u(t)) = u^2(t) - 2x(t-1) \quad (4.20)$$

zu ersetzen. Dann kann das diskrete Maximumprinzip angewendet werden und liefert den optimalen Prozeß wie folgt: Gemäß (4.10) finden wir

$$p_1(t-1) = -2, \quad t = 1, 2, \dots, N.$$

Zusammen mit $p_1(N) = 0$ liefert das die Hamiltonfunktion

$$\mathcal{H}(P(t), x(t-1), u) = u^2 - 2x(t-1) - 2u, \quad t = 1, 2, \dots, N-1,$$

$$\mathcal{H}(P(N), x(N-1), u) = u^2 - 2x(N-1).$$

Damit folgt aber unmittelbar

$$\max_{u \in U} \mathcal{H}(P(t), x(t-1), u) = \begin{cases} 3 - 2x(t-1), & t = 1, 2, \dots, N-1, \\ 1 - 2x(N-1), \end{cases}$$

wobei

$$u^*(t) = -1, \quad t = 1, 2, \dots, N,$$

gewählt werden kann. Aus (4.15) findet man damit die dieser Steuerung entsprechende Trajektorie $x^*(t) = -1$, $t = 1, 2, \dots, N$, mit dem Anfangszustand $x(0) = 0$. Beide zusammen stellen den optimalen diskreten Prozeß der geänderten Aufgabe dar.

4.3. Dynamische Optimierung

Zur Lösung von Optimierungsaufgaben, insbesondere für diskrete (mehrstufige) dynamische Modelle, hat sich neben dem Maximumprinzip die Methode der *dynamischen Optimierung* bewährt. Der Grundgedanke dieser auf dem *Optimalprinzip* von Bellman beruhenden Methode besteht darin, die optimalen Werte nicht für alle Variablen eines Modells zugleich, sondern nacheinander für die einzelnen Stufen des Modells aufzusuchen. Diese Vorgehensweise hat eine beträchtliche Senkung des Rechenaufwandes zur Folge.

Die Methode der dynamischen Optimierung ist auf einen sehr breiten Kreis von Modelltypen und Problemstellungen anwendbar. Es können sowohl diskrete als auch stetige Modelle bearbeitet werden, und die Schaltungsstruktur der Stufen im Gesamtmodell kann kompliziert sein. Nebenbedingungen für die Steuer- und Zustandsvariablen, auch von komplizierter Art, stellen im allgemeinen kein Hindernis für die Anwendbarkeit der Methode dar, ja das Vorhandensein solcher Restriktionen wirkt sich oft günstig auf den rechnerischen Aufwand aus. Wegen des zyklischen Charakters des Rechenablaufs ist die dynamische Optimierung gut für eine Programmierung geeignet.

Diese Aufzählung von Vorteilen darf aber nicht dazu führen, in der dynamischen Optimierung eine ideale Methode zu sehen, die sich auf jedes noch so hochdimensionale und komplexe Problem ohne Schwierigkeiten anwenden läßt. Grenzen der praktischen Anwendbarkeit werden vor allem durch den mit der Zahl der Variablen stark ansteigenden rechentechnischen Aufwand, d.h., Bedarf an Speicherplatz und Rechenzeit, gesetzt.

Im Abschnitt 4.3.1. werden die Grundlagen der dynamischen Optimierung zusammengestellt. Anschließend werden verschiedene Ergänzungen und Erweiterungen der Methode besprochen, die für ihre Anwendung auf praktische Problemstellungen wichtig erscheinen. Dabei wird Wert darauf gelegt, die Handhabung der Algorithmen an einfachen Beispielen zu demonstrieren, die der Leser mit vertretbarem Zeitaufwand nachrechnen kann.

4.3.1. Zusammenstellung der Grundlagen

Grundaufgabe der optimalen Steuerung¹ für ein N -stufiges reihenförmiges System:

Wir betrachten ein System, das aus N in Reihe geschalteten Teilsystemen (Stufen) besteht (Abb. 4.1b). Dabei bedeutet \mathbf{x}_n den Vektor des Austrittszustandes der Stufe

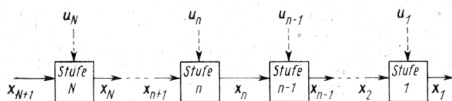


Bild 4.1b) System aus N in Reihe geschalteten Stufen

Nr. n , und \mathbf{u}_n den Vektor der auf Stufe n einwirkenden Steuergrößen. In bezug auf dieses System wird folgende Aufgabe gestellt, die wir als *Grundaufgabe der optimalen Steuerung* bezeichnen wollen:

Gegeben seien

- für alle Stufen Modellgleichungen (Transformationsgleichungen) der Form¹⁾

$$\mathbf{x}_n = \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n), \quad n = 1, 2, \dots, N; \quad (4.21)$$

- zulässige Bereiche für die Steuer- und Zustandsvektoren

$$\mathbf{x}_n \in X_n, \quad \mathbf{u}_n \in U_n, \quad n = 1, 2, \dots, N; \quad (4.22)$$

- der Anfangszustand

$$\mathbf{x}_{N+1} = \alpha. \quad (4.23)$$

Gesucht werden

- aus der Menge der zulässigen Steuerungen diejenigen optimalen $(\mathbf{u}_1^*, \dots, \mathbf{u}_N^*)$, für die gilt

$$G = \sum_{n=1}^N G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) = \max! \quad (4.24)$$

- die zugehörigen optimalen Zustände $(\mathbf{x}_1^*, \dots, \mathbf{x}_N^*)$.

¹⁾ Auf die gegenüber den früheren Abschnitten geänderte Bezeichnungsweise wurde schon in Zusammenhang mit den Formeln (4.7–4.9) hingewiesen. Insbesondere wird im folgenden n (früher Dimension des Zustandsraumes) zur Bezeichnung der Stufennummer verwendet, dafür gibt s die Dimension von Zustandsvektoren an.

Diese Aufgabenstellung ist noch sehr allgemein. So ist offengelassen, von welcher Form die Modellgleichungen (4.21) sind. Diese müssen nicht unbedingt in analytischer Form vorliegen; es kann sich z.B. auch um eine tabellarisch gegebene eindeutige Zuordnung von Werten (\mathbf{x}_{n+1} , \mathbf{u}_n) zu \mathbf{x}_n handeln, analoges gilt für die Summanden der Zielfunktion (4.24). Die Art der Beschränkungen (4.22) ist ebenfalls offengelassen, die Zustands- und Steuerbereiche X_n bzw. U_n können stetige oder auch diskrete Punktmengen sein usw.

Der Algorithmus der dynamischen Optimierung:

Zur Lösung der formulierten Grundaufgabe soll nun die Methode der dynamischen Optimierung angewendet werden. Grundlage dieser Methode ist, wie schon erwähnt, das **Optimalprinzip von Bellman**, das in bezug auf die Grundaufgabe folgendermaßen formuliert werden kann:

Ist $(\mathbf{u}_1^, \dots, \mathbf{u}_{N-1}^*, \mathbf{u}_N^*)$ eine optimale Steuerung des N -stufigen Systems mit dem Anfangszustand $\mathbf{x}_{N+1} = \alpha$, und ist \mathbf{x}_N^* der zugehörige Ausgangszustand von Stufe N , so ist gleichzeitig $(\mathbf{u}_1, \dots, \mathbf{u}_{N-1}^*)$ eine optimale Steuerung des um die Stufe N verkürzten Systems mit dem Anfangszustand \mathbf{x}_N^* .*

Es sei darauf hingewiesen, daß dieses Prinzip beweisbar ist, also einen mathematischen Satz darstellt. Anschaulich besagt das Prinzip, wenn man es auf die verkürzten Systeme wieder anwendet: Die Lösung der Grundaufgabe für die N -stufige Kette enthält die optimalen Steuerungen aller kürzeren Ketten (die entstehen, wenn man von links her eine oder mehrere Stufen wegstreicht), vorausgesetzt, daß man als deren Anfangszustand den jeweiligen optimalen Zustandsvektor der Gesamtlösung verwendet.

Der in Ausnutzung dieser Eigenschaft entwickelte Algorithmus der dynamischen Optimierung umfaßt zwei Arbeitsabschnitte, von denen jeder aus N gleichartigen Schritten besteht.

1. Arbeitsabschnitt: Die Einbettungstechnik¹⁾

Beginnend mit Stufe 1 wird „entgegen der Stromrichtung“ folgende Rekursionsformel der Reihe nach für $n = 1, 2, \dots, N$ angewendet:

$$f_n(\mathbf{x}_{n+1}) = \max_{\mathbf{u}_n} \{G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n))\},$$

$$\mathbf{x}_{n+1} \in X_{n+1}, \quad \mathbf{u}_n \in U_n. \quad (4.25)$$

Dabei ist zu beachten

$$f_0(\mathbf{x}_1) \equiv 0, \quad X_{N+1} = \{\alpha\}. \quad (4.26)$$

In jedem Schritt der Einbettungstechnik gewinnt man zwei Funktionen:

$f_n(\mathbf{x}_{n+1})$ gibt den Maximalwert des zur Stufenfolge $n \dots 1$ gehörenden Teils der Zielfunktion (4.24) in Abhängigkeit vom Eingangszustand \mathbf{x}_{n+1} an,

$\mathbf{u}_n(\mathbf{x}_{n+1})$ gibt die zugehörige Steuerung in der Stufe n an, ebenfalls in Abhängigkeit vom Eingangszustand \mathbf{x}_{n+1} in die Teilkette.

¹⁾ Diese Bezeichnung geht auf Bellman zurück (imbedding-technique).

Die Optimalfunktion $f_n(\mathbf{x}_{n+1})$ wird nur für den folgenden Schritt benötigt, $\mathbf{u}_n(\mathbf{x}_{n+1})$ muß für den zweiten Abschnitt des Algorithmus aufgehoben (z.B. als Tabelle abgespeichert) werden. Im letzten Schritt ergibt sich der Zahlenwert $f_N(\mathbf{x}_{N+1}) = f_N(\boldsymbol{\alpha})$, dies ist der gesuchte Optimalwert der Zielfunktion G . Die zugehörige Steuerung $\mathbf{u}_N(\boldsymbol{\alpha}) = \mathbf{u}_N^*$ ist die tatsächliche optimale Steuerung der gesuchten Gesamtlösung in der Stufe N .

2. Arbeitsabschnitt: Zusammenstellen der optimalen Lösung¹⁾

Durch abwechselnde Benutzung der Modellgleichungen (4.21) und der mit der Einbettungstechnik „auf Vorrat“ berechneten und (z.B. als Tabellen) abgespeicherten Funktionen $\mathbf{u}_n(\mathbf{x}_{n+1})$ wird, diesmal „in Stromrichtung“, durch die Stufenkette hindurchgerechnet. Das erfordert wiederum N Schritte:

$$\begin{aligned} \mathbf{u}_n^* &= \mathbf{u}_n(\mathbf{x}_{n+1}^*), & n &= N, N-1, \dots, 1, \\ \mathbf{x}_n^* &= \mathbf{T}_n(\mathbf{x}_{n+1}^*, \mathbf{u}_n^*), \end{aligned} \quad (4.27)$$

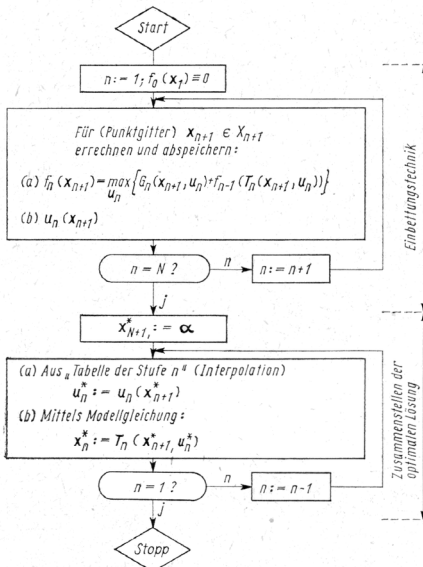


Bild 4.2 Ablaufplan des Algorithmus der dynamischen Optimierung

¹⁾ In der englischsprachigen Literatur wird für diesen Abschnitt die sehr treffende Bezeichnung „table-entry-technique“ verwendet.

wobei als Anfangswert zu benutzen ist:

$$\mathbf{x}_{N+1}^* = \alpha. \quad (4.28)$$

In Abb. 4.2 ist der Algorithmus nochmals in kurzer Form als Ablaufplan dargestellt. Man erkennt deutlich den zyklischen Charakter, der eine Programmierung erleichtert.

Die Arbeitsweise des Algorithmus soll nun noch anhand eines einfachen Beispiels demonstriert werden.

Beispiel 4.3: Der Druck eines Gases soll durch N hintereinander geschaltete Kompressoren von p_a auf p_1 erhöht werden. Die Anfangstemperatur sei ϑ . Nach jeder adiabatischen Kompression werde das Gas isobar wieder auf diese Temperatur abgekühlt. Die Druckstufen sollen so gewählt werden, daß die zur Kompression insgesamt benötigte Energie minimal ist. Um diese Problemstellung in die Form der Grundaufgabe zu bringen, führen wir $x_n = p_n$ als Zustandsvariable und $u_n = p_n/p_{n+1}$ als Steuervariable ein (die Vektoren \mathbf{x} und \mathbf{u} sind hier also eindimensional). Damit erhalten wir

– Modellgleichungen

$$x_n = T_n(x_{n+1}, u_n) = x_{n+1} \cdot u_n; \quad n = 1, 2, \dots, N,$$

– zulässige Bereiche

$$1 \leq u_n \leq p_1/x_{n+1}; \quad n = 1, 2, \dots, N,$$

– den Anfangszustand

$$x_{N+1} = p_a,$$

– das Optimierungsziel

$$G = \sum_{n=1}^N G_n(x_{n+1}, u_n) = \sum_{n=1}^N \frac{mR\vartheta}{a} (u_n^a - 1) = \min!$$

G_n stellt den Energieverbrauch der Stufe n dar. m bezeichnet die Gasmenge, R die Gaskonstante, das Verhältnis $a = \frac{c_p - c_v}{c_p}$ der spezifischen Wärmen werde als konstant angenommen. Wir setzen

$$\text{noch } K = \frac{mR\vartheta}{a}.$$

Für die Durchführung der Einbettungstechnik muß in unserem Fall in der Formel (4.25) natürlich „max“ durch „min“ ersetzt werden. Im ersten Schritt ist das Problem

$$f_1(x_2) = \min_{u_1} \{K(u_1^a - 1)\},$$

$$1 \leq u_1 \leq p_1/x_2, \quad x_1 = x_2 \cdot u_1 = p_1,$$

zu lösen. Weil p_1 vorgegeben ist, besteht der Steuerbereich für jeden x_2 -Wert nur aus einem Punkt: $u_1 = p_1/x_2$. Wir erhalten demnach

$$u_1(x_2) = p_1/x_2, \quad f_1(x_2) = K \left[\left(\frac{p_1}{x_2} \right)^a - 1 \right].$$

Im zweiten Schritt (Hinzunahme der Stufe 2) besteht die Aufgabe

$$f_2(x_3) = \min_{u_2} \left\{ K(u_2^a - 1) + K \left[\left(\frac{p_1}{x_2} \right)^a - 1 \right] \right\},$$

$$1 \leq u_2 \leq p_1/x_3.$$

Wir setzen zunächst $x_2 = x_3 \cdot u_2$ ein und suchen dann das Minimum durch Nullsetzen der Ableitung des Klammerausdrucks nach u_2

$$Kau_2^{a-1} - Ka \left(\frac{p_1}{x_3} \right) \cdot u_2^{-(a+1)} = 0.$$

So erhalten wir

$$u_2(x_3) = \left(\frac{p_1}{x_3} \right)^{\frac{1}{2}}, \quad f_2(x_3) = 2K \left[\left(\frac{p_1}{x_3} \right)^{\frac{a}{2}} - 1 \right]$$

Entsprechend ergibt sich für $n = 3, \dots, N-1$

$$u_n(x_{n+1}) = \left(\frac{p_1}{x_{n+1}} \right)^{\frac{1}{n}}, \quad f_n(x_{n+1}) = nk \left[\left(\frac{p_1}{x_{n+1}} \right)^{\frac{a}{n}} - 1 \right]$$

und im letzten Schritt

$$u_N(x_{N+1}) = u_N^* = \left(\frac{p_1}{p_\alpha} \right)^{\frac{1}{N}}, \quad f_N(x_{N+1}) = G^* = NK \left[\left(\frac{p_1}{p_\alpha} \right)^{\frac{a}{N}} - 1 \right].$$

Damit ist die Einbettungstechnik abgeschlossen.

Im zweiten Arbeitsabschnitt wird nun durch abwechselnde Verwendung der Modellgleichungen und der Funktionen $u_n(x_{n+1})$ die optimale Lösung zusammengestellt. So ergibt sich

– im ersten Schritt:

$$u_N^* = \left(\frac{p_1}{p_\alpha} \right)^{\frac{1}{N}}; \quad x_N^* = p_\alpha \cdot u_N^*, \quad \text{d.h.,} \quad p_N = \sqrt[N]{p_\alpha^{N-1} \cdot p_1};$$

– im zweiten Schritt:

$$u_{N-1}^* = u_{N-1}(x_N^*) = \left(\frac{p_1}{p_\alpha} \right)^{\frac{1}{N-1}};$$

$$x_{N-1}^* = x_N^* \cdot u_{N-1}^*, \quad \text{d.h.,} \quad p_{N-1} = \sqrt[N]{p_\alpha^{N-2} \cdot p_1^2};$$

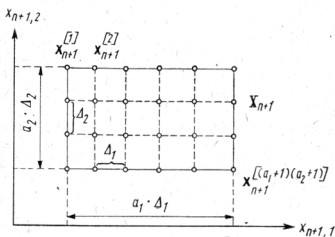
– allgemein:

$$u_n^* = u_n(x_{n+1}^*) = \left(\frac{p_1}{p_\alpha} \right)^{\frac{1}{N}} \quad (\text{unabh. von } n)$$

$$p_n = \sqrt[N]{p_\alpha^{n-1} \cdot p_1^{N-n+1}}; \quad n = 1, 2, \dots, N.$$

Bemerkungen zur numerischen Durchführung und zum rechentechnischen Aufwand:

Bei der Anwendung des Algorithmus auf praktische Probleme wird es oft nicht möglich sein, die Funktionen $f_n(\mathbf{x}_{n+1})$, $\mathbf{u}_n(\mathbf{x}_{n+1})$ in geschlossener Form zu berechnen. Man muß vielmehr meist so vorgehen, daß man den Zustandsbereich X_{n+1} durch ein *Punktgitter* überdeckt, vgl. Abb. 4.3, und dann für jeden Gitterpunkt gemäß Formel (4.25) die Maximierung vornimmt. (Hierbei ist zu betonen, daß durch den Algorithmus der dynamischen Optimierung nicht festgelegt ist, welches Suchverfahren zur Bestimmung des Maximums benutzt wird.) Die so erhaltene Menge diskreter f_n -Werte wird zusammen mit den zugehörigen \mathbf{u}_n -Werten als *Tabelle der Stufe n* abge-



* Tabelle der Stufe n				
Gitterpunkte	$x_{n+1}^{[1]}$	$x_{n+1}^{[2]}$...	$x_{n+1}^{[(a_1+1)(a_2+1)]}$
Werte der Funktion $f_n(x_{n+1})$	$f_n(x_{n+1})^{[1]}$	$f_n(x_{n+1})^{[2]}$...	
zugehörige Vektoren $u_n(x_{n+1})$	$u_n(x_{n+1})^{[1]}$	$u_n(x_{n+1})^{[2]}$...	

Bild 4.3 Überdeckung des zulässigen Bereiches X_{n+1} eines (zweidimensionalen) Zustandsvektors x_{n+1} durch ein Punktgitter und Inhalt der „Tabelle der Stufe n“

speichert. Bei dem in (4.25) erforderlichen Zurückgreifen auf f_{n-1} und beim Hineingehen in die Tabelle $u_n(x_{n+1})$ während des zweiten Arbeitsabschnittes muß dann öfter interpoliert werden, wobei i. a. lineare Interpolationsformeln verwendet werden.

Man kann sich vorstellen, daß mit zunehmender Dimension der Vektoren x_n und u_n sowohl der Umfang der abzuspeichernden Tabellen als auch die Rechenzeit sehr stark ansteigen. Da sich die Grenzen für die praktische Anwendbarkeit der dynamischen Optimierung hauptsächlich vom rechen-technischen Aufwand her ergeben, ist es nützlich, die Größenordnung dieses Aufwandes abzuschätzen. Zu diesem Zweck nehmen wir an, daß jeder Zustandsvektor x_{n+1} des N -stufigen Systems s Koordinaten, jeder Steuervektor u_n r Koordinaten besitzt und daß die durch (4.22) festgelegten zulässigen Bereiche X_{n+1} s -dimensionale Quader sind, Abb. 4.3. Zerlegt man die i -te Kante des Quaders in a_i Teile, $i = 1 \dots s$, so entsteht ein Gitter aus

$$A = (a_1 + 1) \cdot (a_2 + 1) \cdot \dots \cdot (a_s + 1)$$

Punkten. Im n -ten Schritt der Einbettungstechnik werden der Reihe nach für jeden Gitterpunkt errechnet

- ein Zahlenwert für die Optimalfunktion $f_n(x_{n+1})$,
- r Koordinaten des zugehörigen Vektors $u_n(x_{n+1})$.

Die „Tabelle der Stufe n“ benötigt deshalb A Speicherplätze für die f_n -Werte und rA Plätze für die Vektoren u_n . Letztere werden während der Einbettungstechnik nicht weiter verwendet; man benötigt sie erst wieder im zweiten Abschnitt des Algorithmus. Man kann die u_n -Tabellen deshalb zunächst

auf externen Speichermedien der Rechanlage (z.B. Magnetbändern) deponieren. Im Gegensatz dazu werden die entstandenen f_n -Werte gemäß Rekursionsformel (4.25) bereits im nächsten Schritt wieder gebraucht (danach allerdings nicht mehr). Zusammenfassend kann man sagen, daß folgender Bedarf an Speicherplatz besteht:

$N_{ss} \approx 2 \cdot A$ Plätze im schnellen Speicher (Kernspeicher) der Anlage, nämlich A für die entstehenden f_n -Werte und A für diejenigen des vorangegangenen Schrittes,

$N_{sl} \approx N \cdot r \cdot A$ Plätze im langsamen (externen) Speicher der Anlage für die u_n -Tabellen der insgesamt N Stufen.

Eine grobe Abschätzung für die Rechenzeit läßt sich folgendermaßen gewinnen: Der bestimmende Anteil kommt zweifellos durch die punktweise Berechnung der Werte $f_n(x_{n+1})$, $u_n(x_{n+1})$ zustande. Wird für jeden Rechengang die Zeit $\Delta\tau$ benötigt, so ergibt sich

$$T \approx \Delta\tau \cdot N \cdot A.$$

Aus den so entwickelten groben Abschätzungsbeziehungen lassen sich folgende *Faustregeln* ablesen (wenn man noch die Teilungen a_i der Quaderkanten als ungefähr gleich voraussetzt).

- Der Bedarf an schnellen Speichern wächst exponentiell mit der Dimension s der Zustandsvektoren. Von der Zahl der Stufen hängt er nicht ab. (Das exponentielle Anwachsen wird ersichtlich, wenn man in der obigen Formel für A $a_1 = a_2 = \dots = a_s = a$ setzt.)
- Der Bedarf an langsamen Speichern wächst ebenfalls exponentiell mit s , zur Stufenzahl N und zur Dimension der Steuervektoren ist er proportional.
- Der Bedarf an Rechenzeit nimmt exponentiell mit s zu und ist proportional zur Stufenzahl.

Wie man bemerkt, ist als „kritische Größe“ hinsichtlich der praktischen Durchführbarkeit der dynamischen Optimierung die Dimension der Zustandsvektoren anzusehen, während Stufenzahl N und Zahl der Steuergrößen r den Bedarf an schnellen Speichern, der immer noch den „Engpaß“ bei den verfügbaren Rechanlagen bildet, wenig beeinflussen.

Die Abschätzungen machen die Erfahrungstatsache verständlich, daß Anwendungen der dynamischen Optimierung über Probleme mit mehr als 3 bis 4 Zustandsvariablen je Stufe zur Zeit normalerweise nicht hinausgehen.

4.3.2. Einige andere Aufgabenstellungen

Die in der Praxis im Zusammenhang mit mehrstufigen Systemen auftretenden Optimierungsprobleme können in verschiedenster Weise von der Form der Grundaufgabe abweichen. Aufgaben, bei denen die Zielfunktion nicht die Summenform (4.24) hat, werden im Abschnitt 4.3.3. besprochen. Die Ausdehnung der Methode auf den Fall, daß die Struktur der Stufenkopplung nicht eine reine Reihenschaltung darstellt, ist Gegenstand des Abschnittes 4.3.4. Zunächst wollen wir auf einige nicht so einschneidende Abwandlungen der Grundaufgabe eingehen und die entsprechenden Modifikationen des Algorithmus kennenlernen.

Minimierungsaufgaben:

Wenn die Abweichung zur Grundaufgabe nur darin besteht, daß anstelle von (4.24) die Forderung gestellt ist, das *Minimum von G* zu bestimmen, so ist dies gleichbedeutend damit, $G' = -G$ zu maximieren. Auf dasselbe läuft es hinaus, wenn man statt (4.25) als Rekursionsformel

$$f_n(\mathbf{x}_{n+1}) = \min_{\mathbf{u}_n} \{G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n))\} \quad (4.29)$$

anwendet.

Bestimmung des Optimums als Funktion des Anfangszustandes:

In der Grundaufgabe war angenommen, daß der Anfangszustand zahlenmäßig vorgegeben ist: $\mathbf{x}_{N+1} = \alpha$. Bei praktischen Problemen wird man oft daran interessiert sein, die optimale Lösung für verschiedene denkbare \mathbf{x}_{N+1} zu kennen, um auch bei eintretenden Abweichungen von α die optimale Entscheidung treffen zu können. In diesem Fall ist dann anstelle von (4.23) für \mathbf{x}_{N+1} ein bestimmter Wertebereich X_{N+1} vorgegeben, und es besteht die Aufgabe, den Optimalwert von G sowie die optimalen Steuerungen für alle Punkte dieses Bereiches zu ermitteln:

$$G^* = f_N(\mathbf{x}_{N+1}); \quad \mathbf{u}_n^* = \mathbf{u}_n^*(\mathbf{x}_{N+1}); \quad \mathbf{x}_{N+1} \in X_{N+1}.$$

Am Algorithmus ändert sich dabei nichts. Man hat lediglich im letzten Schritt der Einbettungstechnik die Maximierung über \mathbf{u}_N nicht nur für einen Wert $\mathbf{x}_{N+1} = \alpha$, sondern für ein ganzes Punktgitter von \mathbf{x}_{N+1} -Werten durchzuführen.

Optimierung über den Anfangszustand:

In anderen Fällen ist der Anfangszustand \mathbf{x}_{N+1} nicht als für den Prozeß gegeben zu betrachten, sondern er soll, ebenso wie die \mathbf{u}_n , optimal bestimmt werden. \mathbf{x}_{N+1} spielt dann die Rolle von zusätzlichen Steuervariablen, und man muß an den letzten Schritt der Einbettungstechnik noch eine Maximierung

$$G^* = \max_{\mathbf{x}_{N+1}} \{f_N(\mathbf{x}_{N+1})\}; \quad \mathbf{x}_{N+1} \in X_{N+1}, \quad (4.30)$$

anschließen.

Bestimmung des Optimums als Funktion des Endzustandes:

In Abweichung zur Grundaufgabe liege der Fall vor, daß nicht der Anfangszustand \mathbf{x}_{N+1} , sondern der Endzustand \mathbf{x}_1 des Prozesses festgelegt ist. Bei der Suche nach dem Maximum von G sind dann alle die Steuerungen zugelassen, die den Prozeß von irgendeinem Anfangszustand \mathbf{x}_{N+1} in den festgelegten Endzustand \mathbf{x}_1 führen. Jetzt ist also

$$G^* = f_N(\mathbf{x}_1)$$

gesucht. Auch hier könnte verlangt sein, dies nicht nur für einen \mathbf{x}_1 -Wert, sondern für einen ganzen Bereich X_1 möglicher Werte zu tun.

Der Unterschied zwischen diesem Problem und der früher besprochenen Bestimmung des Optimums als Funktion des Anfangszustandes liegt nur in der „Richtung“ des Prozesses. Wenn man sich nämlich in Abb. 4.1b) die zwischen den Stufen verlaufenden Pfeile umgekehrt denkt, so wird aus dem Endzustand \mathbf{x}_1 der (jetzt gegebene) Anfangszustand, und aus dem Anfangszustand \mathbf{x}_{N+1} wird der Endzustand der Stufenkette. Die Richtung des Prozesses kommt aber im mathematischen Modell (4.21–4.24) nur durch die dort gewählte Form der Modellgleichungen (4.21) zum Ausdruck,

$$\mathbf{x}_n = \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n),$$

die so geschrieben sind, daß jeweils der Ausgangszustand der Stufe als Funktion des Eingangszustandes und der Steuerung vorliegt. Wenn es also gelingt, die Beziehungen (4.21) nach \mathbf{x}_{n+1} aufzulösen, so geht unser jetziges Problem genau in eine Aufgabe des schon besprochenen Typs über und kann auf die übliche Weise behandelt werden. Man bezeichnet diese Verfahrensweise als *Zustandsinversion*. Formal ergibt sich dabei folgender Rechengang:

– Auflösen der Modellgleichungen (4.21) nach den Eingangsgrößen:

$$\mathbf{x}_n = \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n) \rightarrow \mathbf{x}_{n+1} = \boldsymbol{\theta}_n(\mathbf{x}_n, \mathbf{u}_n); \quad n = 1, 2, \dots, N, \quad (4.31)$$

– Umschreiben der Zielfunktion:

$$G_n(\boldsymbol{\theta}_n(\mathbf{x}_n, \mathbf{u}_n), \mathbf{u}_n) = \Gamma_n(\mathbf{x}_n, \mathbf{u}_n); \quad n = 1, 2, \dots, N, \quad (4.32)$$

– Anwendung der sinngemäß umgeschriebenen Rekursionsgleichungen. Anstelle von (4.25, 4.26) haben diese jetzt folgende Form:

$$\begin{aligned} \varphi_n(\mathbf{x}_n) &= \max_{\mathbf{u}_n} \{ \Gamma_n(\mathbf{x}_n, \mathbf{u}_n) + \varphi_{n+1}(\boldsymbol{\theta}_n(\mathbf{x}_n, \mathbf{u}_n)) \}, \\ \mathbf{x}_n &\in X_n, \quad \mathbf{u}_n \in U_n, \quad n = N, N-1, \dots, 1, \quad \varphi_{N+1}(\mathbf{x}_{N+1}) \equiv 0. \end{aligned} \quad (4.33)$$

Diese Formeln sind jetzt aber der Reihe nach für $N, N-1, \dots, 1$ anzuwenden; d.h., wenn wir Abb. 4.1b) zugrunde legen, verläuft die Einbettungstechnik jetzt „in Stromrichtung“. Man spricht deshalb auch von einer Optimierung durch *Vorwärtsrekursion*.

– Entsprechend diesem Unterschied läuft die zweite Phase des Algorithmus, die Zusammenstellung der optimalen Lösung, diesmal entgegen der Stromrichtung.

Ein solcher Übergang zur Vorwärtsrechnung erscheint im Prinzip einfach. Doch darf man natürlich die Schwierigkeiten nicht übersehen, die sich beim Umkehren der Modellgleichungen zeigen können. Die notwendige *eindeutige* Umkehrung der Transformation (4.21) wird nicht immer möglich sein. Eine Rolle spielt auch die Dimension der Zustandsvektoren, die ja für \mathbf{x}_{n+1} und \mathbf{x}_n keineswegs immer gleich sein muß. Nehmen wir an, es sei

$$\dim \mathbf{x}_{n+1} = r, \quad \dim \mathbf{x}_n = s,$$

und denken wir daran, daß (4.21) eine Vektorgleichung ist, die ausführlicher geschrieben folgendes beinhaltet:

$$\begin{aligned} x_{n,1} &= T_{n,1}(x_{n+1,1}, \dots, x_{n+1,r}; \mathbf{u}_n), \\ x_{n,2} &= T_{n,2}(x_{n+1,1}, \dots, x_{n+1,r}; \mathbf{u}_n), \\ &\vdots \\ x_{n,s} &= T_{n,s}(x_{n+1,1}, \dots, x_{n+1,r}; \mathbf{u}_n). \end{aligned} \quad (4.34)$$

Wenn wir eine Zustandsinversion durchführen wollen, so ist also ein System von s Gleichungen nach den $x_{n+1,i}$ aufzulösen.

Im Falle $r > s$ (mehr Variable am Eingang als am Ausgang der Stufe), ergibt sich etwa, falls eine Auflösung überhaupt möglich ist,

$$x_{n+1,i} = \theta_{n,i}(x_{n,1}, \dots, x_{n,s}, \mathbf{u}_n, x_{n+1,s+1}, \dots, x_{n+1,r}), \quad i = 1, 2, \dots, s, \quad (4.35)$$

d. h., neben \mathbf{u}_n treten jetzt rechts $(r - s)$ Größen $x_{n+1,s+1} \dots x_{n+1,r}$ auf, die auch als Steuervariable zu behandeln sind.

Im Falle $r < s$ (weniger Variable am Eingang als am Ausgang) kann man etwa die ersten r Gleichungen von (4.34) benutzen, um nach den $x_{n+1,i}$ aufzulösen:

$$x_{n+1,i} = \theta_{n,i}(x_{n,1}, \dots, x_{n,r}; \mathbf{u}_n), \quad i = 1, \dots, r. \quad (4.36)$$

Setzt man dies in die verbleibenden $(s - r)$ Gleichungen von (4.34) ein, so entstehen Beziehungen

$$g_j(x_{n,1}, \dots, x_{n,s}; \mathbf{u}_n) = 0, \quad j = 1, \dots, (s - r), \quad (4.37)$$

die als zusätzliche Nebenbedingungen für $\mathbf{x}_n, \mathbf{u}_n$ zu behandeln sind.

Im Falle $r = s$ ändert sich bei einer Inversion die Zahl der Steuervariablen und Nebenbedingungen nicht.

Die Möglichkeit, die Berechnungsrichtung der dynamischen Optimierung umzukehren, hängt also davon ab, ob die Zustandsinversion durchführbar und vom Rechenaufwand her zweckmäßig ist.

Falls eine Zustandsinversion nicht möglich ist, kann $f_N(\mathbf{x}_1)$ auch durch die normale Rückwärtsrechnung ermittelt werden. Man muß dann allerdings bei den Schritten der Einbettungstechnik \mathbf{x}_1 als zusätzlichen Parameter mitführen: Vor dem ersten Schritt wird zunächst mittels der Modellgleichung

$$\mathbf{x}_1 = \mathbf{T}_1(\mathbf{x}_2, \mathbf{u}_1)$$

der Endzustand \mathbf{x}_1 in den zur Stufe 1 gehörigen Summanden der Zielfunktion eingeführt. Nehmen wir an, \mathbf{u}_1 habe p Koordinaten

$$\mathbf{u}_1^T = (u_{1,1} \dots u_{1,p}).$$

Wenn wir die für $n = 1$ aufgeschriebenen Gleichungen (4.34) nach den $u_{1,i}$ auflösen, so ergibt sich

– im Fall $s = p$

$$\mathbf{u}_1 = \omega(\mathbf{x}_1, \mathbf{x}_2); \quad G_1(\mathbf{x}_2, \omega(\mathbf{x}_1, \mathbf{x}_2)) = \bar{G}_1(\mathbf{x}_2, \mathbf{x}_1).$$

– im Fall $s > p$ (durch Auflösen der ersten p Gleichungen)

$$\mathbf{u}_1 = \omega(x_{1,1}, \dots, x_{1,p}, \mathbf{x}_2); \quad G_1(\mathbf{x}_2, \omega) = \bar{G}_1(\mathbf{x}_2, x_{1,1}, \dots, x_{1,p}).$$

Die restlichen $(s - p)$ Gleichungen von (4.34) liefern zusätzliche Nebenbedingungen für $\mathbf{x}_1, \mathbf{x}_2$.

– im Fall $s < p$

$$u_{1,i} = \omega_i(\mathbf{x}_1, \mathbf{x}_2), \quad i = 1, \dots, s,$$

$$G_1(\mathbf{x}_2, \omega_1, \dots, \omega_s, u_{1,s+1}, \dots, u_{1p}) = \bar{G}_1(\mathbf{x}_2, \mathbf{x}_1, u_{1,s+1}, \dots, u_{1p}).$$

Der erste Schritt der Einbettungstechnik, Formel (4.25), liefert dann

$$f_1(\mathbf{x}_2, \mathbf{x}_1) = \max_{\mathbf{u}_1} \{\bar{G}(\mathbf{x}_2, \mathbf{x}_1, \mathbf{u}_1)\}, \quad (4.38)$$

wobei im Fall $s \geq p$ die Maximierung entfällt, während im Fall $s < p$ nur noch über einen Teil der Steuervariablen zu maximieren ist. Die übrigen Schritte verlaufen dann wie gewohnt, wobei Funktionen

$$f_n(\mathbf{x}_{n+1}, \mathbf{x}_1), \quad \mathbf{u}_n(\mathbf{x}_{n+1}, \mathbf{x}_1), \quad n = 1, 2, \dots, N, \quad (4.39)$$

entstehen. Zuletzt erhält man mittels

$$f_N(\mathbf{x}_1) = \max_{\mathbf{x}_{N+1}} \{f_N(\mathbf{x}_{N+1}, \mathbf{x}_1)\} \quad (4.40)$$

die gesuchte Optimalfunktion. Das Unangenehme bei dieser Vorgehensweise besteht aber darin, daß die Tabellen der Einbettungstechnik jetzt außer von \mathbf{x}_{n+1} auch noch von \mathbf{x}_1 abhängen; sie sind also umfangreicher zu berechnen und abzuspeichern. Dies kann ein ernstes Hindernis für die Durchführbarkeit der Methode werden.

4.3.3. Probleme, bei denen die Zielfunktion keine Summenform hat

Bei der im Abschnitt 4.3.1. formulierten Grundaufgabe war für die Zielfunktion G die Summenform (4.24) vorausgesetzt. Wenn wir G als „Gewinn“ oder „Ertrag“ des zu optimierenden Systems auffassen, so bedeutet dies, daß sich die in den einzelnen Stufen erzielten Gewinne additiv zum Gesamtgewinn zusammensetzen. Dieser Fall wird in den praktischen Anwendungen tatsächlich oft vorliegen, wobei die Abgrenzung der einzelnen Stufen durch die gegebene Struktur des Systems nahegelegt wird. Das muß aber nicht immer so sein. Es werden Fälle auftreten, bei denen es zunächst viele Möglichkeiten gibt, das System in Stufen aufzuteilen, d.h. eine Dekomposition für die rechnerische Behandlung vorzunehmen. Es ist dann eine Frage der Zweckmäßigkeit, welche Variante man benutzt, und zu einzelnen Varianten kann eine Zielfunktion gehören, die von der Summenform abweicht. Einige solche Fälle wollen wir im folgenden besprechen.

Optimierung des Endzustandes:

Im Prinzip sehr einfach zu behandeln ist der Fall, daß das Optimierungskriterium im Gegensatz zur Grundaufgabe die Form hat

$$G = \Phi(\mathbf{x}_1) = \max! \quad (4.41)$$

Wir sprechen dann von einer Optimierung des Endzustandes.

Dieser Problemtyp läßt sich ohne Schwierigkeit auf die Grundaufgabe zurückführen. Setzt man die Modellgleichung der Stufe 1

$$\mathbf{x}_1 = \mathbf{T}_1(\mathbf{x}_2, \mathbf{u}_1)$$

in (4.41) ein, so hat man mit

$$\Phi(\mathbf{T}_1(\mathbf{x}_2, \mathbf{u}_1)) = G_1(\mathbf{x}_2, \mathbf{u}_1),$$

$$G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) \equiv 0, \quad n = 2, 3, \dots, N, \quad (4.42)$$

das Problem bereits auf die Form der Grundaufgabe umgeschrieben.

Verallgemeinerte Dekomposition:

Wir betrachten jetzt folgende Problematik: Für die mathematische Behandlung eines Systems ist eine Zerlegung in N Stufen entsprechend Abb. 4.1b) ins Auge gefaßt worden. Es soll eine Optimierung bezüglich der Variablen \mathbf{u} durchgeführt werden, die zu maximierende Zielfunktion G setzt sich in gegebener Weise aus den Gewinnfunktionen $G_n(\mathbf{x}_{n+1}, \mathbf{u}_n)$ der einzelnen Stufen zusammen.

$$G = F[G_1(\mathbf{x}_2, \mathbf{u}_1), G_2(\mathbf{x}_3, \mathbf{u}_2), \dots, G_N(\mathbf{x}_{N+1}, \mathbf{u}_N)] = \max_{\mathbf{u}_1 \dots \mathbf{u}_N} \quad (4.43)$$

Bei der Optimierung sind Nebenbedingungen (4.21–4.23) zu beachten. Es wird nun die Frage gestellt, unter welchen Voraussetzungen man die Optimierungsaufgabe (4.21 bis 4.23, 4.43) zerlegen kann in N den einzelnen Stufen zugeordnete Teilprobleme, wobei in jedem Teilproblem jeweils nur über die Steuervariablen der betreffenden Stufe zu maximieren ist. Außerdem wird natürlich die Form der Teilprobleme gesucht.

Für den Spezialfall, daß G eine Summenform (4.24) hat, wissen wir bereits, daß eine solche Zerlegung möglich ist, wobei die Formeln (4.25, 4.26) die Teilaufgaben beschreiben. Für den allgemeinen Fall (4.43) wurde von Mitten [18] eine hinreichende Bedingung gefunden. Um sie formulieren zu können, erinnern wir uns zunächst an den Begriff der Monotonie einer Funktion:

Die reellwertige Funktion $G = F(g)$ sei auf einer Teilmenge M der reellen Zahlen definiert. Sie heißt dort *monoton nicht abnehmend*, wenn gilt:

$$\text{Aus } g < g' \text{ folgt stets } F(g) \leq F(g'); g, g' \in M. \quad (4.44)$$

Wir führen noch folgenden Begriff ein:

Definition 4.4: Die reellwertige Funktion $G = F(g_1, g_2, \dots, g_N)$ heiße *separabel* bezüglich g_N , wenn zwei reellwertige Funktionen Φ_N, F_{N-1} existieren, so daß gilt:

$$F(g_1, g_2, \dots, g_N) = \Phi_N[g_N, F_{N-1}(g_1, \dots, g_{N-1})]. \quad (4.45)$$

Zur Erläuterung ein Beispiel: Für $F(g_1, g_2, g_3, g_4) = g_4 + g_3 \cdot (g_2 + g_1)$ gilt mit

$$\Phi_4 = g_4 + F_3; \quad F_3 = g_3 \cdot (g_2 + g_1)$$

die Eigenschaft (4.45), d.h., F ist bezüglich g_4 separabel. Dagegen ist $\hat{F} = g_4 \cdot g_3 + g_2 \cdot g_1$ offenbar nicht separabel bezüglich g_4 .

Unter Verwendung der Begriffe Monotonie und Separabilität kann nun folgender Satz hinsichtlich des Optimierungsproblems (4.21–4.23, 4.43) bewiesen werden:

Satz 4.2: Ist $G = F(G_1, G_2, \dots, G_N)$ separabel bezüglich G_N ,

$$F(G_1, \dots, G_N) = \Phi_N[G_N, F_{N-1}(G_1, \dots, G_{N-1})], \quad (4.46)$$

und ist weiter $\Phi_N[G_N, F_{N-1}]$ für jeden Wert von G_N eine monoton nicht abnehmende Funktion des Arguments F_{N-1} , so gilt

$$\begin{aligned} \max_{u_1, \dots, u_N} F(G_1, \dots, G_N) = \\ \max_{u_N} \Phi_N[G_N(x_{N+1}, u_N), \max_{u_1, \dots, u_{N-1}} F_{N-1}(G_1(x_2, u_1), \dots, G_{N-1}(x_N, u_{N-1}))]. \end{aligned} \quad (4.47)$$

Der Satz besagt, daß unter den genannten Voraussetzungen die Optimierung in zwei Schritten ausgeführt werden kann. Im ersten Schritt werden die optimalen Steuerungen der Stufen $N-1, \dots, 1$ ermittelt, wobei sich als Ergebnis eine Funktion $F_{N-1}^*(x_N)$ ergibt. Danach wird im zweiten Schritt durch

$$\max_{u_N} \Phi_N[G_N(x_{N+1}, u_N), F_{N-1}^*(T_N(x_{N+1}, u_N))] \quad (4.48)$$

die optimale Steuerung der Stufe N separat berechnet. Man hat damit das Gesamtproblem in zwei äquivalente Teilprobleme zerlegt.

Der obige Satz zeigt, wie man die Optimierung eines N -stufigen Systems aufteilen kann in die Optimierung der Stufe N und die Optimierung einer zur Stufenfolge $N-1, \dots, 1$ gehörigen Funktion F_{N-1} . Auf F_{N-1} kann der Satz nochmals angewendet werden, wenn für diese Funktion wieder Separabilität (bezüglich G_{N-1}) und die entsprechende Monotonieeigenschaft vorliegt. In günstigen Fällen erreicht man so tatsächlich die Zerlegung des Gesamtproblems in N den einzelnen Stufen zugeordnete Teilaufgaben. Die Formulierung der zugehörigen Rekursionsgleichung ist jetzt leicht zu finden:

Vorgelegt sei also ein Optimierungsproblem (4.21–4.23, 4.43), wobei die Zielfunktion folgende Eigenschaften aufweist:

(a) $F(G_1, G_2, \dots, G_N)$ ist fortlaufend separabel, d.h., es existieren Funktionen F_N, F_{N-1}, \dots, F_1 , so daß gilt

$$F_n(G_1, G_2, \dots, G_n) = \Phi_n[G_n, F_{n-1}(G_1, \dots, G_{n-1})], \quad n = N, N-1, \dots, 2, \quad (4.49)$$

$$F_N \equiv F.$$

(b) $\Phi_n[G_n, F_{n-1}]$ ist für jeden zulässigen Wert von G_n eine monoton nicht abnehmende Funktion des Arguments F_{n-1} , $n = N, N-1, \dots, 2$.

Entsprechend (4.47) kann dann folgende Rekursionsgleichung zur Lösung der N -stufigen Optimierungsaufgabe angewendet werden:

$$f_n(\mathbf{x}_{n+1}) = \begin{cases} \max_{\mathbf{u}_1} F_1(G_1(\mathbf{x}_2, \mathbf{u}_1)); & n = 1, \\ \max_{\mathbf{u}_n} \Phi_n[G_n(\mathbf{x}_{n+1}, \mathbf{u}_n), f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n))]; & n = 2, 3, \dots, N; \end{cases} \quad (4.50)$$

$$\mathbf{x}_{n+1} \in X_{n+1}; \quad \mathbf{u}_n \in U_n;$$

$$X_{N+1} = \{\alpha\}.$$

Im übrigen unterscheidet sich der Rechengang nicht von dem im Abschnitt 4.3.1. geschilderten Ablauf. Man hat also zunächst die N Schritte der Einbettungstechnik mittels (4.50) durchzuführen, anschließend die N Schritte der Zusammenstellung der optimalen Lösung, wobei ungeändert die Formeln (4.27, 4.28) zu verwenden sind.

Wir wollen die Rekursionsformeln noch an einigen Beispielen erläutern:

Beispiel 4.4: $F(G_1, \dots, G_N) = \sum_{n=1}^N G_n(\mathbf{x}_{n+1}, \mathbf{u}_n)$. F ist fortlaufend separabel, denn man kann schreiben

$$\Phi_n[G_n, F_{n-1}(G_1, \dots, G_{n-1})] = G_n + F_{n-1}, \quad (4.51)$$

$$F_{n-1} = \sum_{i=1}^{n-1} G_i, \quad n = N, N-1, \dots, 2.$$

Die Monotonieeigenschaft ist erfüllt

$$F_{n-1} \leq F'_{n-1} \rightarrow G_n + F_{n-1} \leq G_n + F'_{n-1}. \quad (4.52)$$

Aus (4.50) entsteht die wohlbekannte Rekursionsgleichung (4.25).

Beispiel 4.5: $F(G_1, \dots, G_N) = \prod_{n=1}^N G_n(\mathbf{x}_{n+1}, \mathbf{u}_n)$. (4.53)

Auch diese Form ist fortlaufend separabel; man kann nämlich setzen:

$$\Phi_n[G_n, F_{n-1}(G_1, \dots, G_{n-1})] = G_n \cdot F_{n-1},$$

$$F_{n-1} = G_{n-1} \cdot G_{n-2} \cdot \dots \cdot G_1; \quad n = N, N-1, \dots, 2.$$

Die Monotonieeigenschaft ist sicher dann erfüllt, wenn gilt:

$$G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) \geq 0 \quad \text{für alle zulässigen } \mathbf{x}_{n+1}, \mathbf{u}_n, \quad n = 1, 2, \dots, N.$$

Als Rekursionsgleichung ergibt sich aus (4.50) in diesem Fall:

$$f_n(\mathbf{x}_{n+1}) = \begin{cases} \max_{\mathbf{u}_1} G_1(\mathbf{x}_2, \mathbf{u}_1); & n = 1, \\ \max_{\mathbf{u}_n} G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) \cdot f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n)); & n = 2, 3, \dots, N. \end{cases} \quad (4.54)$$

Beispiel 4.6: $F(G_1, \dots, G_4) = G_4(\mathbf{x}_5, \mathbf{u}_4) + G_3(\mathbf{x}_4, \mathbf{u}_3) \cdot [G_2(\mathbf{x}_3, \mathbf{u}_2) + G_1(\mathbf{x}_2, \mathbf{u}_1)]$.

Hier kann man setzen:

$$F = F_4 = G_4 + F_3; \quad F_3 = G_3 \cdot F_2; \quad F_2 = G_2 + F_1; \quad F_1 = G_1.$$

Die Monotonieeigenschaft ist erfüllt, wenn $G_3(x_4, u_3) \geq 0$ für alle erlaubten x_4, u_3 . Man erhält die folgenden Rekursionsgleichungen:

$$f_1(x_2) = \max_{u_1} G_1(x_2, u_1),$$

$$f_2(x_3) = \max_{u_2} [G_2(x_3, u_2) + f_1(T_2(x_3, u_2))],$$

$$f_3(x_4) = \max_{u_3} [G_3(x_4, u_3) + f_2(T_3(x_4, u_3))],$$

$$f_4(x_5) = \max_{u_4} [G_4(x_5, u_4) + f_3(T_4(x_5, u_4))].$$

Beispiel 4.7: Auf vier Personengruppen, deren Stärke in Tabelle 4.1 angegeben ist, sollen 26 (gleichartige) Dinge „möglichst gerecht“ verteilt werden. Als ideal wäre diejenige Verteilung anzusehen, bei der auf jede Person dieselbe Stückzahl, nämlich 0,05 käme. Nun sind die betreffenden Dinge aber nicht teilbar, d. h., es kommt nur eine ganzzahlige Aufteilung in Betracht. Die ideale Verteilung ist also nicht zu realisieren. Als „möglichst gerecht“ wird diejenige ganzzahlige Verteilung angesehen, bei der die maximale Abweichung der tatsächlichen pro-Kopf-Zahl vom idealen Wert 0,05 möglichst klein ausfällt.

Tabelle 4.1 Verteilungsproblem

Gruppe	Anzahl Personen	Ideale Verteilung		Optimale Verteilungen			
		pro Kopf	Stückzahl	pro Kopf	Stückzahl	pro Kopf	Stückzahl
$n=I$	$a_n = 131$	0,05	6,55	0,046	6	0,053	7
II	89	0,05	4,45	0,045	4	0,045	4
III	249	0,05	12,45	0,052	13	0,048	12
IV	51	0,05	2,55	0,059	3	0,059	3
Σ	520		26		26		26

Wir können diese Aufgabe als 4-stufiges Entscheidungsproblem darstellen, vgl. Abb. 4.4. Es bedeuten

u_n die der Gruppe n zugeteilte Anzahl von Dingen,

x_n die den Gruppen 4 bis n insgesamt zugeteilte Stückzahl.

Als Modellgleichungen haben wir also:

$$x_n = T_n(x_{n+1}, u_n) = x_{n+1} + u_n; \quad n = 1, 2, 3, 4. \quad (4.55)$$

Dabei sind folgende Beschränkungen zu beachten:

$$x_1 = 26; \quad x_5 = 0; \quad u_n \geq 0 \text{ und ganzzahlig,} \quad n = 1, 2, 3, 4. \quad (4.56)$$

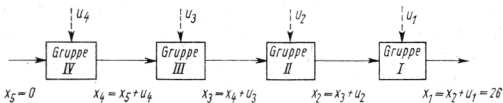


Bild 4.4 Verteilungsproblem

Die Bedingung $\sum_1^4 u_n = 26$ ist mit (4.55, 4.56) automatisch erfüllt.

Zu jeder zulässigen Steuerung lassen sich Abweichungen A_n der Pro-Kopf-Zahl vom Idealwert berechnen:

$$A_n = \left| \frac{u_n}{a_n} - 0,05 \right|, \quad n = 1, 2, 3, 4,$$

und als Maßzahl für die „Güte der Steuerung“ wird wie gesagt die größte Abweichung verwendet:

$$M = \max \{A_1, A_2, A_3, A_4\}.$$

Gesucht wird als optimale Steuerung diejenige, bei der M einen möglichst kleinen Wert hat:

$$\min_u M = \min_u [\max \{A_1, \dots, A_4\}].$$

Dieses Optimierungskriterium ist offenbar gleichbedeutend mit

$$\max_u [\min \{-A_1, \dots, -A_4\}],$$

und wir erhalten, wenn wir zu unserer bisher verwendeten Bezeichnungsweise übergehen,

$$\begin{aligned} G &= F[G_1(u_1), \dots, G_4(u_4)] = \max_u! \\ F[G_1, \dots, G_4] &= \min \{G_1, \dots, G_4\}; \quad G_n = - \left| \frac{u_n}{a_n} - 0,05 \right|. \end{aligned} \quad (4.57)$$

Mit (4.55–4.57) liegt ein Problem vor, welches der Aufgabenstellung (4.21–4.23, 4.43) entspricht. Unsere jetzige Zielfunktion ist fortlaufend separabel, denn offenbar gilt:

$$\begin{aligned} \Phi_n[G_n, F_{n-1}(G_1 \cdots G_{n-1})] &= \min(G_n, F_{n-1}), \\ F_{n-1} &= \min(G_{n-1}, \dots, G_1); \quad n = 5, 4, 3, 2. \end{aligned}$$

Die Monotonieeigenschaft ist ebenfalls erfüllt, denn aus

$$F'_{n-1} \leq F''_{n-1} \quad \text{folgt} \quad \min(G_n, F'_{n-1}) \leq \min(G_n, F''_{n-1}).$$

Wir erhalten deshalb folgende Rekursionsbeziehung

$$f_n(x_{n+1}) = \begin{cases} \max_{u_1} G_1; & n = 1, \\ \max_{u_n} [\min(G_n, f_{n-1}(T_n(x_{n+1}, u_n)))] & n = 2, 3, 4, \end{cases} \quad (4.58)$$

die unter Beachtung der Nebenbedingungen (4.56) in der bekannten Weise auszuwerten ist.

Nun ist, wenn man die ideale Verteilung der Tabelle 4.1 ins Auge faßt, die Lösung fast zu erraten (Auf- bzw. Abrunden auf die nächstgelegene ganze Zahl). Das Beispiel wird eigentlich nur deshalb angeführt, um auf den breiten Anwendungsbereich der dynamischen Optimierung hinzuweisen. Man sieht, daß die Zielfunktion G z.B. auch durch eine Vorschrift „Suche die kleinste von N Zahlen“ gegeben sein kann. Auch Nebenbedingungen der Form (4.56), d.h. Forderungen der Ganzzahligkeit usw. stören die Anwendbarkeit der Methode nicht.

Wir wollen den Rechengang bei der Lösung noch andeuten: Zunächst kann man offenbar den Rechenaufwand dadurch beträchtlich verringern, daß man in Hinblick auf die ungefähre Kenntnis der Lösung die Bereiche der zulässigen Entscheidungen in vernünftiger Weise einengt, etwa:

$$\begin{aligned} u_1 &\in \{5, 6, 7, 8\}; & u_2 &\in \{4, 5, \}; & u_3 &\in \{11, 12, 13, 14\}; \\ u_4 &\in \{1, 2, 3, 4\}. \end{aligned}$$

Die Zielfunktion habe Summenform

$$G = \sum_{n=1}^{N+M+L} G_n, \quad (4.60)$$

wobei G_n jeweils vom Eingangszustand und vom Steuervektor der Stufe n abhängig sei. Das Maximum der Zielfunktion

$$\max_u G = f_{N+M+L}(\mathbf{x}_{N+M+L+1}) \quad (4.61)$$

soll in Abhängigkeit vom Anfangszustand $\mathbf{x}_{N+M+L+1}$ unter Beachtung gegebener zulässiger Bereiche X_n, U_n für die Zustands- und Steuergrößen bestimmt werden.

Die Einbettungstechnik ist hier in folgenden vier Abschnitten durchzuführen:

I: Rückwärtsrechnung im oberen Zweig (Stufen 1 bis N):

$$\begin{aligned} f_1(\mathbf{x}_2) &= \max_{\mathbf{u}_1} G_1(\mathbf{x}_2, \mathbf{u}_1), \\ f_n(\mathbf{x}_{n+1}) &= \max_{\mathbf{u}_n} \{G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n))\}, \quad n = 2, 3, \dots, N, \end{aligned} \quad (4.62)$$

(wobei im letzten Schritt \mathbf{x}_{N+1} als $\hat{\mathbf{x}}_{N+M+1}$ zu lesen ist). Als Ergebnis erhält man $f_N(\hat{\mathbf{x}}_{N+M+1})$ und natürlich die zu den Stufen gehörigen optimalen Steuerungen $\mathbf{u}_n(\mathbf{x}_{n+1})$.

II: Rückwärtsrechnung im unteren Zweig (Stufen $N+1$ bis $N+M$):

$$\begin{aligned} f_{N+1}(\mathbf{x}_{N+2}) &= \max_{\mathbf{u}_{N+1}} G_{N+1}(\mathbf{x}_{N+2}, \mathbf{u}_{N+1}), \\ f_n(\mathbf{x}_{n+1}) &= \max_{\mathbf{u}_n} \{G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n))\}, \\ n &= N+1, N+2, \dots, N+M. \end{aligned} \quad (4.63)$$

Dies liefert zuletzt $f_{N+M}(\mathbf{x}_{N+M+1})$.

III: Hinzunahme der Verzweigungsstufe $N+M+1$:

$$\begin{aligned} f_{N+M+1}(\mathbf{x}_{N+M+2}) &= \max_{\mathbf{u}_{N+M+1}} \{G_{N+M+1}(\mathbf{x}_{N+M+2}, \mathbf{u}_{N+M+1}) \\ &\quad + f_N(\hat{\mathbf{T}}_{N+M+1}(\mathbf{x}_{N+M+2}, \mathbf{u}_{N+M+1}) \\ &\quad + f_{N+M}(\mathbf{T}_{N+M+1}(\mathbf{x}_{N+M+2}, \mathbf{u}_{N+M+1}))\}. \end{aligned} \quad (4.64)$$

IV: Rückwärtsrechnung durch die Restkette (Stufen $N+M+2$ bis $N+M+L$)

$$\begin{aligned} f_n(\mathbf{x}_{n+1}) &= \max_{\mathbf{u}_n} \{G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_n(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n))\}, \\ n &= N+M+2, \dots, N+M+L. \end{aligned} \quad (4.65)$$

Dies liefert schließlich die gesuchte Abhängigkeit des Optimalwertes $f_{N+M+L}(\mathbf{x}_{N+M+L+1})$ der Zielfunktion. Mit Hilfe der abgespeicherten Beziehungen für die Optimalwerte

$\mathbf{u}_n(\mathbf{x}_{n+1})$ der Steuergrößen kann man dann in üblicher Weise durch Vorwärtsrechnung die Optimallösung zusammenstellen.

Man sieht, daß sich der Fall einer verzweigten Kette etwa mit dem gleichen Aufwand bearbeiten läßt wie eine unverzweigte Kette gleicher Stufenzahl.

Die Vereinigung zweier Ketten:

Das System habe jetzt die in Abb. 4.6 gezeigte Struktur. Die Modellgleichungen der Stufen lauten

$$\begin{aligned}\mathbf{x}_n &= \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n); & n &= 1, \dots, N+M+L & (n \neq N, N+M), \\ \mathbf{x}_N &= \mathbf{T}_N(\mathbf{x}_{N+1}, \mathbf{x}_{N+M+1}, \mathbf{u}_N), \\ \mathbf{x}_{N+M} &= \mathbf{T}_{N+M}(\hat{\mathbf{x}}_{N+M+1}, \mathbf{u}_{N+M}).\end{aligned}\quad (4.66)$$

Die Zielfunktion hat wieder die Gestalt (4.60).

Zu beachten ist, daß

$$G_N = G_N(\mathbf{x}_{N+1}, \mathbf{x}_{N+M+1}, \mathbf{u}_N) \quad (4.67)$$

von zwei Eingangsvektoren abhängig ist. Wir nehmen an, daß die Anfangszustände $\hat{\mathbf{x}}_{N+M+1}$ und $\mathbf{x}_{N+M+L+1}$ als fest betrachtet werden, d.h., daß

$$\max_{\mathbf{u}} G = f_{N+M+L+1}(\hat{\mathbf{x}}_{N+M+1}, \mathbf{x}_{N+M+L+1}) \quad (4.68)$$

gesucht wird.

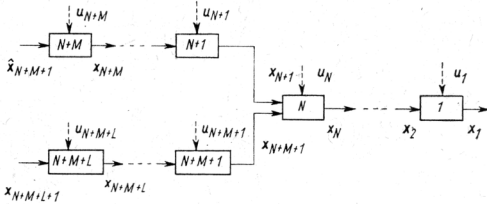


Bild 4.6 Vereinigung zweier Ketten

Für die Anwendung der dynamischen Optimierung auf dieses System kommen zwei Varianten in Frage, je nachdem, ob eine Zustandsinversion möglich bzw. rechnerisch mit vertretbarem Aufwand durchführbar ist oder nicht.

Variante A: Es sei möglich, in einem der zusammentreffenden Zweige, etwa im oberen (Stufen $N+1$ bis $N+M$) die Zustandsinversion durchzuführen. Dann stehen die Modellgleichungen in der Form

$$\begin{aligned}\mathbf{x}_{n+1} &= \theta_n(\mathbf{x}_n, \mathbf{u}_n); & n &= N+1, \dots, N+M-1, \\ \hat{\mathbf{x}}_{N+M+1} &= \theta_{N+M}(\mathbf{x}_{N+M}, \mathbf{u}_{N+M})\end{aligned}\quad (4.69)$$

zur Verfügung, und entsprechend sind die Summanden der Zielfunktion umgeschrieben (vgl. hierzu 4.3.2.):

$$\begin{aligned} G_{N+M} &= \Gamma_{N+M}(\hat{\mathbf{x}}_{N+M+1}, \mathbf{x}_{N+M}, \mathbf{u}_{N+M}); & G_n &= \Gamma_n(\mathbf{x}_n, \mathbf{u}_n), \\ n &= N + M - 1, \dots, N + 1. \end{aligned} \quad (4.70)$$

Wir haben damit den schon betrachteten Fall einer sich verzweigenden Stufenfolge hergestellt, wobei die Aufgabenstellung von der damaligen nur dadurch abweicht, daß außer dem Anfangszustand $\mathbf{x}_{N+M+L+1}$ auch noch der „Endzustand“ $\hat{\mathbf{x}}_{N+M+1}$ als fest zu betrachten ist.

Die Einbettungstechnik umfaßt folgende Arbeitsschritte:

I: *Rückwärtsrechnung durch die Stufenfolge* 1, 2, ..., $N - 1$ unter Verwendung der Formeln (4.62). Dies liefert $f_{N-1}(\mathbf{x}_N)$.

II: *Vorwärtsrekursion im oberen Zweig (Stufen* $N + M$ *bis* $N + 1$ *)*. Dabei sind (vgl. 4.3.2.) folgende Beziehungen zu verwenden:

$$\begin{aligned} \varphi_{N+M}(\mathbf{x}_{N+M}, \hat{\mathbf{x}}_{N+M+1}) &= \max_{\mathbf{u}_{N+M}} \Gamma_{N+M}(\hat{\mathbf{x}}_{N+M+1}, \mathbf{x}_{N+M}, \mathbf{u}_{N+M}), \\ \varphi_n(\mathbf{x}_n, \hat{\mathbf{x}}_{N+M+1}) &= \max_{\mathbf{u}_n} \{ \Gamma_n(\mathbf{x}_n, \mathbf{u}_n) + \varphi_{n+1}(\theta_n(\mathbf{x}_n, \mathbf{u}_n), \hat{\mathbf{x}}_{N+M+1}) \}, \\ n &= N + M - 1, \dots, N + 1. \end{aligned} \quad (4.71)$$

Zu bemerken ist, daß im ersten Schritt die Maximierung entfällt, wenn durch die Modellgleichung der Stufe $N + M$

$$\hat{\mathbf{x}}_{N+M+1} = \theta_{N+M}(\mathbf{x}_{N+M}, \mathbf{u}_{N+M}) \rightarrow \mathbf{u}_{N+M} = \omega(\hat{\mathbf{x}}_{N+M+1}, \mathbf{x}_{N+M})$$

bei festem $\hat{\mathbf{x}}_{N+M+1}$ für jeden Wert \mathbf{x}_{N+M} die Steuerung eindeutig festgelegt ist. Als Ergebnis dieses Arbeitsabschnittes erhält man

$$\varphi_{N+1}(\mathbf{x}_{N+1}, \hat{\mathbf{x}}_{N+M+1}).$$

III: *Hinzunahme der Vereinigungsstufe* N :

$$\begin{aligned} f_N(\mathbf{x}_{N+M+1}, \hat{\mathbf{x}}_{N+M+1}) &= \max_{\mathbf{u}_N, \mathbf{x}_{N+1}} \{ G_n(\mathbf{x}_{N+1}, \mathbf{x}_{N+M+1}, \mathbf{u}_N) \\ &+ f_{N-1}(\mathbf{T}_N(\mathbf{x}_{N+1}, \mathbf{x}_{N+M+1}, \mathbf{u}_N)) + \varphi_{N+1}(\mathbf{x}_{N+1}, \hat{\mathbf{x}}_{N+M+1}) \}. \end{aligned} \quad (4.72)$$

IV: *Rückwärtsrechnung durch die Stufen* $N + M + 1, \dots, N + M + L$ *des unteren Zweiges*:

$$\begin{aligned} f_n(\mathbf{x}_{n+1}, \hat{\mathbf{x}}_{N+M+1}) &= \max_{\mathbf{u}_n} \{ G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n), \hat{\mathbf{x}}_{N+M+1}) \}, \\ n &= N + M + 1, \dots, N + M + L. \end{aligned} \quad (4.73)$$

Dies ergibt schließlich die gesuchte Abhängigkeit

$$f_{N+M+L}(\mathbf{x}_{N+M+L+1}, \hat{\mathbf{x}}_{N+M+1}).$$

Variante B: Wir betrachten jetzt den Fall, daß eine Inversion der Modellgleichungen nicht ausgeführt werden kann. Die Einbettungstechnik verläuft dann folgendermaßen.

I: Wie bei Variante A. Man erhält $f_{N-1}(\mathbf{x}_N)$.

II: Rückwärtsrechnung im oberen Zweig (Stufen $N+1$ bis $N+M$) mit dem Ziel, $f_{N+M}(\hat{\mathbf{x}}_{N+M+1}, \mathbf{x}_{N+1})$ zu erhalten. Dabei ist als erster Schritt die Funktion

$$f_{N+1}(\mathbf{x}_{N+2}, \mathbf{x}_{N+1}) = \max_{\mathbf{u}_{N+1}} G_{N+1}(\mathbf{x}_{N+2}, \mathbf{u}_{N+1}) \quad (4.74)$$

unter Beachtung der Nebenbedingung,

$$\mathbf{x}_{N+1} = \mathbf{T}_{N+1}(\mathbf{x}_{N+2}, \mathbf{u}_{N+1}) \quad (4.75)$$

zu bestimmen. Letzteres bedeutet, daß man zunächst mittels (4.75) die Steuervariablen \mathbf{u}_{N+1} soweit wie möglich durch die Zustandsgrößen \mathbf{x}_{N+1} , \mathbf{x}_{N+2} ausdrückt und dies in G_{N+1} einsetzt (vgl. 4.3.2.). Danach wird gemäß (4.74) über die nicht eliminierbaren Steuergrößen maximiert. Wenn sich \mathbf{u}_{N+1} vollständig durch \mathbf{x}_{N+1} , \mathbf{x}_{N+2} ausdrücken läßt, entfällt die Maximierung.

Die weiteren Schritte sind in der üblichen Art auszuführen:

$$f_n(\mathbf{x}_{n+1}, \mathbf{x}_{N+1}) = \max_{\mathbf{u}_n} \{G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n), \mathbf{x}_{N+1})\},$$

$$n = N+2, \dots, N+M. \quad (4.76)$$

Sie liefern zuletzt $f_{N+M}(\hat{\mathbf{x}}_{N+M+1}, \mathbf{x}_{N+1})$. Diese Funktion ist mit φ_{N+1} aus Variante A identisch.

III: *Hinzunahme der Vereinigungsstufe N.* Hierbei gilt Formel (4.72), wenn man dort φ_{N+1} durch f_{N+M} ersetzt.

IV: Wie bei Variante A. Man erhält $f_{N+M+L}(\mathbf{x}_{N+M+L+1}, \hat{\mathbf{x}}_{N+M+1})$.

Vergleicht man beide Arten des Vorgehens, so stellt man folgendes fest: Wenn in der Aufgabenstellung der Anfangszustand $\hat{\mathbf{x}}_{N+M+1}$ zahlenmäßig fest vorgegeben ist, so ist Variante A mit geringerem rechentechnischen Aufwand verbunden, da dann die in Abschnitt II entstehenden Funktionen φ_n nur in Abhängigkeit von *einem* Zustand \mathbf{x}_n abzuspeichern sind. Bei Variante B kommen dagegen *zwei* Zustände vor: $f_n(\mathbf{x}_{n+1}, \mathbf{x}_{N+1})$. Falls die optimale Lösung für einen ganzen Bereich von $\hat{\mathbf{x}}_{N+M+1}$ -Werten gesucht wird, dürfte der Aufwand beider Varianten etwa gleich sein.

Beispiel 4.8: Zur Erläuterung des Rechenablaufes betrachten wir das in Abb. 4.7 dargestellte 6-stufige System. Alle Zustands- und Steuervektoren sind eindimensional, die Modellgleichungen der Stufen und die Summanden der Zielfunktion sind in Tabelle 4.3 zusammengestellt. Beschränkungen für die Zustands- und Steuervariablen seien nicht vorhanden. Die Eingangszustände \hat{x}_5 und x_7 sind als gegeben zu betrachten. Es soll das Minimum der Zielfunktion

$$G = \sum_{n=1}^6 G_n$$

bestimmt werden, und zwar in Abhängigkeit von \hat{x}_5 , x_7 .

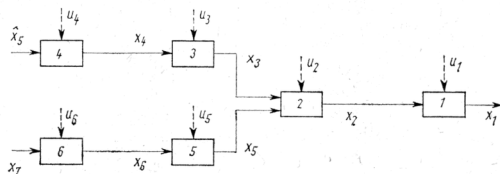


Bild 4.7 Beispiel für eine Vereinigung zweier Ketten

Tabelle 4.3 Beispiel für eine Vereinigung zweier Ketten

Stufennummer n	Modellgleichung	Summand G_n der Zielfunktion
1	$x_1 = x_2 - \frac{1}{2} u_1$	$(u_1 + x_2)^2$
2	$x_2 = x_3 + \frac{1}{2} x_5 + u_2$	$x_3^2 + x_5^2 + u_2^2$
3	$x_3 = u_3 - x_4$	$(x_4 - 2)^2 + (u_3 - 1)^2$
4	$x_4 = \hat{x}_5 + u_4$	$(\hat{x}_5 - u_4)^2$
5	$x_5 = x_6 + u_5$	$x_6^2 + \frac{1}{2} u_5^2$
6	$x_6 = x_7 - u_6$	$(x_7 - 1)^2 + u_6^2$

Stufennummer n	optimale Entscheidung	optimaler Zustand
1	$u_1 = \frac{2}{5} \hat{x}_5 + \frac{1}{14} x_7$	$x_1 = -\frac{3}{5} \hat{x}_5 + \frac{3}{28} x_7$
2	$u_2 = 0$	$x_2 = -\frac{2}{5} \hat{x}_5 + \frac{1}{14} x_7$
3	$u_3 = \frac{2}{5} \hat{x}_5 + 1$	$x_3 = -\frac{2}{5} \hat{x}_5$
4	$u_4 = -\frac{1}{5} \hat{x}_5 + 1$	$x_4 = \frac{4}{5} \hat{x}_5 + 1$
5	$u_5 = -\frac{2}{7} x_7$	$x_5 = \frac{1}{7} x_7$
6	$u_6 = \frac{4}{7} x_7$	$x_6 = \frac{3}{7} x_7$

Zum Vergleich führen wir die Berechnung mit beiden Varianten durch. Natürlich ist in den Formeln jetzt „max“ durch „min“ zu ersetzen.

Berechnung nach Variante A

I: Anwendung von (4.62) auf die Stufe $n = 1$ ergibt:

$$f_1(x_2) = \min_{u_1} (u_1 + x_2)^2.$$

Da für u_1 keine Beschränkungen zu beachten sind, können wir das Minimum durch Nullsetzen der ersten Ableitung suchen. Man erhält:

$$\frac{\partial G_1}{\partial u_1} = 2(u_1 + x_2) = 0 \rightarrow u_1(x_2) = -x_2$$

und weiter durch Einsetzen in G_1

$$f_1(x_2) \equiv 0.$$

II: In den Stufen 3, 4 wird die Zustandsinversion ausgeführt. Man erhält für Stufe 4 aus der Tabelle 4.3:

$$x_4 = \hat{x}_5 + u_4 \rightarrow u_4 = x_4 - \hat{x}_5,$$

$$G_4 = (\hat{x}_5 - u_4)^2 \rightarrow I_4 = (2\hat{x}_5 - x_4)^2.$$

Da I_4 nicht von der Entscheidungsgröße u_4 abhängt (durch die Nebenbedingung $x_4 = \hat{x}_5 + u_4$ ist bei festem \hat{x}_5 für jeden Wert x_4 die Entscheidung eindeutig festgelegt), entfällt die Minimierung, d. h., es gilt:

$$\varphi_4(x_4, \hat{x}_5) \equiv I_4 = (2\hat{x}_5 - x_4)^2,$$

$$u_4(x_4, \hat{x}_5) = x_4 - \hat{x}_5.$$

Im nächsten Schritt der Vorwärtsrekursion erhält man gemäß (4.71) für die Stufe 3:

$$x_3 = u_3 - x_4 \rightarrow x_4 = u_3 - x_3,$$

$$G_3 = (x_4 - 2)^2 + (u_3 - 1)^2 \rightarrow I_3 = (u_3 - x_3 - 2)^2 + (u_3 - 1)^2,$$

$$\varphi_3(x_3, \hat{x}_5) = \min_{u_3} \{I_3 + (2\hat{x}_5 - u_3 + x_3)^2\}.$$

Mittels $\frac{\partial}{\partial u_3} (I_3 + \varphi_4) = 0$ ergibt sich:

$$\varphi_3(x_3, \hat{x}_5) = \frac{2}{3}x_3^2 + \frac{8}{3}\hat{x}_5^2 + \frac{4}{3}x_3 \cdot \hat{x}_5 - 4\hat{x}_5 + 2,$$

$$u_3(x_3, \hat{x}_5) = \frac{2}{3}x_3 + \frac{2}{3}\hat{x}_5 + 1.$$

III: Hinzunahme der Vereinigungsstufe 2. Wir müssen die Formel (4.72) anwenden:

$$f_2(x_5, \hat{x}_5) = \min_{x_3, u_2} \{(x_3^2 + x_5^2 + u_2^2) + \varphi_3(x_3, \hat{x}_5) + f_1(x_2)\},$$

wobei noch $x_2 = x_3 + \frac{1}{2} \cdot x_5 + u_2$ einzusetzen ist. Nullsetzen der partiellen Ableitungen nach x_3 und u_2 ergibt

$$x_3(x_5, \hat{x}_5) = -\frac{2}{5} \cdot \hat{x}_5, \quad u_2(x_5, \hat{x}_5) \equiv 0,$$

$$f_2(x_5, \hat{x}_5) = \frac{12}{5}\hat{x}_5^2 - 4\hat{x}_5 + x_5^2 + 2.$$

IV: Es schließt sich eine normale Rückwärtsrechnung durch die Stufen 5, 6 des unteren Zweiges an. Mit (4.73) erhält man

$$f_5(x_6, \hat{x}_5) = \frac{4}{3}x_6^2 + \frac{12}{5}\hat{x}_5^2 - 4\hat{x}_5 + 2,$$

$$u_5(x_6, \hat{x}_5) = -\frac{2}{3}x_6$$

und im letzten Schritt

$$f_6(x_7, \hat{x}_5) = \frac{12}{5} \hat{x}_5^2 - 4\hat{x}_5 + \frac{11}{7} x_7^2 - 2x_7 + 3,$$

$$u_6(x_7, \hat{x}_5) = \frac{4}{7} x_7.$$

Die Einbettungstechnik ist damit zu Ende geführt. $f_6(x_7, \hat{x}_5)$ stellt den gesuchten Optimalwert der Zielfunktion in Abhängigkeit von den Eingangsgrößen x_7 und \hat{x}_5 dar. Die Zusammenstellung der optimalen Lösung kann jetzt durch abwechselnde Benutzung der Optimalfunktionen für die u_n und der Modellgleichungen in üblicher Weise erfolgen. Die Ergebnisse sind in Tabelle 4.3 aufgeführt.

Berechnung nach Variante B

I: Verläuft wie bei Variante A.

II: Diesmal wird keine Inversion der Modellgleichungen vorgenommen, sondern eine normale Rückwärtsrechnung durch die Stufen 3, 4 des oberen Zweiges ausgeführt. Gemäß (4.74, 4.75) besteht im ersten Schritt das Ziel, $f_3(x_4, x_3)$ zu gewinnen. Zu diesem Zweck muß man zunächst mittels der Modellgleichung der Stufe 3

$$x_3 = u_3 - x_4 \rightarrow u_3 = x_3 + x_4$$

und Einsetzen in den Summanden der Zielfunktion in diesen die Abhängigkeit von x_3 hineinbringen:

$$G_3 = (x_4 - 2)^2 + (x_3 + x_4 - 1)^2.$$

Wie man sieht, entfällt dann eine Minimierung bzgl. u_3 und man hat

$$f_3(x_4, x_3) = 2x_4^2 - 6x_4 + x_3^2 + 2x_3x_4 - 2x_3 + 5.$$

$$u_3(x_4, x_3) = x_3 + x_4.$$

Der zweite Schritt geht normal vonstatten:

$$f_4(\hat{x}_5, x_3) = \min_{u_4} \{(\hat{x}_5 - u_4)^2 + f_3(x_3, T_4(\hat{x}_5, u_4))\}.$$

Durch Nullsetzen der Ableitung nach u_4 ergibt sich:

$$u_4(x_3, \hat{x}_5) = 1 - \frac{1}{3} \hat{x}_5 - \frac{1}{3} x_3,$$

$$f_4(x_3, \hat{x}_5) = \frac{2}{3} x_3^2 + \frac{8}{3} \hat{x}_5^2 + \frac{4}{3} x_3 \cdot \hat{x}_5 - 4\hat{x}_5 + 2.$$

Letzteres ist mit $\varphi_3(x_3, \hat{x}_5)$ aus Variante A identisch.

III: Verläuft wie bei Variante A.

IV: Verläuft wie bei Variante A.

Die vorwärts gerichtete Schleife:

Die vorwärts gerichtete Schleife, Abb. 4.8, enthält in der Stufe $N + M + L + 1$ eine Verzweigung, in der Stufe N eine Vereinigung. Da wir eine Kombination der schon behandelten Fälle vor uns haben, wird sich auch der Rechengang aus Abschnitten der damals ausführlich beschriebenen Art zusammensetzen. Dabei ist eine ganze Reihe von Varianten möglich, je nachdem, ob man den oberen oder den unteren Zweig der Schleife zunächst für sich optimiert, ob man dieses Teilergebnis in den Rest der

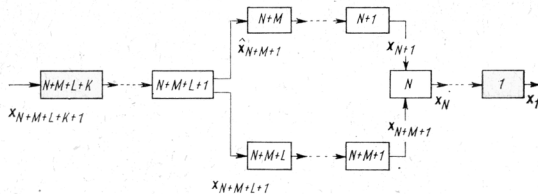


Bild 4.8 Vorwärts gerichtete Schleife

Zielfunktion an der Vereinigungsstelle N oder an der Verzweigungsstelle $N + M + L + 1$ einbezieht oder ob man für einen Teil des Systems eine Zustandsinversion vornimmt.

Wir erläutern im folgenden nur eine dieser Möglichkeiten, wobei wir uns auf die Beschreibung des prinzipiellen Ablaufs der Rechnung beschränken. Die Formulierung der detaillierten Rekursionsbeziehungen dürfte auf Grund der früheren Erörterungen keine Schwierigkeiten machen. Wie üblich wird angenommen, daß die Modellgleichungen jeweils den Ausgangszustand der Stufe als Funktion des Eingangszustandes und der Steuerung darstellen, daß als Zielfunktion eine Summe über alle Stufen vorliegt, und daß ihr Optimalwert in Abhängigkeit vom als fest zu betrachtenden Anfangszustand $\mathbf{x}_{N+M+L+K+1}$ des Gesamtsystems gesucht ist. Die Einbettungstechnik verläuft dann in folgenden Abschnitten:

I: *Rückwärtsrechnung* durch die Stufen 1 bis $N - 1$. Als Ergebnis erhält man die Optimalfunktion $f_{N-1}(\mathbf{x}_N)$.

II: *Separate Optimierung des Teils der Zielfunktion*, der zu einem der Zweige der Schleife gehört (wir wählen beispielsweise den oberen Zweig mit den Stufen $N + 1$ bis $N + M$). Die entstehende Optimalfunktion bezeichnen wir mit f_{N+M} , sie ist in Abhängigkeit von $\hat{\mathbf{x}}_{N+M+1}$ und \mathbf{x}_{N+1} zu ermitteln. Dabei hat man so vorzugehen, wie bei der Vereinigung zweier Ketten, Variante B, beschrieben. Als Ergebnis erhält man $f_{N+M}(\mathbf{x}_{N+1}, \hat{\mathbf{x}}_{N+M+1})$.

III: Einbeziehung des oberen Zweiges auf der Stufe N :

$$f_N(\mathbf{x}_{N+M+1}, \hat{\mathbf{x}}_{N+M+1}) = \max_{\mathbf{u}_{N+1}, \mathbf{u}_N} \{G_N(\mathbf{x}_{N+1}, \mathbf{x}_{N+M+1}, \mathbf{u}_N) + f_{N-1}(\mathbf{T}_N(\mathbf{x}_{N+1}, \mathbf{x}_{N+M+1}, \mathbf{u}_N)) + f_{N+M}(\mathbf{x}_{N+1}, \hat{\mathbf{x}}_{N+M+1})\}. \quad (4.77)$$

Danach Rückwärtsrechnung durch die Stufen $N + M + 1, \dots, N + M + L$ des unteren Zweiges mit dem Ergebnis

$$f_{N+M+L}(\mathbf{x}_{N+M+L+1}, \hat{\mathbf{x}}_{N+M+1}). \quad (4.78)$$

Schließlich Einbeziehung der Verzweigungsstufe $N + M + L + 1$:

$$f_{N+M+L+1}(\mathbf{x}_{N+M+L+2}) = \max_{\mathbf{u}_{N+M+L+1}} \{G_{N+M+L+1}(\mathbf{x}_{N+M+L+2}, \mathbf{u}_{N+M+L+1}) + f_{N+M+L}(\mathbf{T}_{N+M+L+1}(\mathbf{x}_{N+M+L+2}, \mathbf{u}_{N+M+L+1})) + \hat{\mathbf{T}}_{N+M+L+1}(\mathbf{x}_{N+M+L+2}, \mathbf{u}_{N+M+L+1})\}. \quad (4.79)$$

IV. *Rückwärtsrechnung* durch die restlichen Stufen $N + M + L + 2$ bis $N + M + L + K$ mit dem gesuchten Ergebnis

$$f_{N+M+L+K}(\mathbf{x}_{N+M+L+K+1}). \quad (4.80)$$

Beispiel 4.9: Wir erweitern das als Beispiel 4.8 berechnete System (vgl. Abb. 4.7 und Tab. 4.3) durch Anfügen einer Stufe zu einer Vorwärtsschleife, wie in Abb. 4.9 dargestellt. Das Minimum des siebenstufigen Systems soll in Abhängigkeit vom Anfangszustand x_8 bestimmt werden.

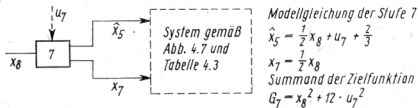


Bild 4.9 Beispiel für eine vorwärts gerichtete Schleife

Der Rechenablauf ist im wesentlichen mit dem identisch, den wir im Beispiel 4.8 (Variante B) hatten. Der einzige Unterschied besteht darin, daß wir mit dem damaligen Endergebnis noch einen Schritt weiterrechnen müssen, da jetzt noch die Verzweigungsstufe 7 einzubeziehen ist (letzte Maßnahme im Schritt III):

$$f_7(x_8) = \min_{u_7} \{G_7(x_8, u_7) + f_6(x_7, \hat{x}_5)\}.$$

Tabelle 4.4 Lösung des Beispiels einer vorwärts gerichteten Schleife

Stufe	optimale Entscheidung	optimaler Zustand
1	$u_1 = \frac{11}{84}x_8 + \frac{5}{18}$	$x_1 = -\frac{11}{56}x_8 - \frac{5}{12}$
2	$u_2 = 0$	$x_2 = -\frac{11}{84}x_8 - \frac{5}{18}$
3	$u_3 = \frac{1}{6}x_8 + \frac{23}{18}$	$x_3 = -\frac{1}{6}x_8 - \frac{5}{18}$
4	$u_4 = -\frac{1}{12}x_8 + \frac{31}{36}$	$x_4 = \frac{1}{3}x_8 + \frac{14}{9}$
5	$u_5 = -\frac{3}{7}x_8$	$x_5 = \frac{1}{14}x_8$
6	$u_6 = \frac{3}{7}x_8$	$x_6 = \frac{3}{14}x_8$
7	$u_7 = -\frac{1}{12}x_8 + \frac{1}{36}$	$\begin{cases} \hat{x}_5 = \frac{5}{12}x_8 + \frac{25}{36} \\ x_7 = \frac{1}{2}x_8 \end{cases}$

Wenn wir x_7, \hat{x}_8 durch die in Abb. 4.9 notierten Modellgleichungen ausdrücken, ergibt sich durch Nullsetzen der Ableitung nach u_7

$$u_7(x_8) = -\frac{1}{12}x_8 + \frac{1}{36}$$

und damit weiter

$$f_7(x_8) = \frac{53}{28}x_8^2 - \frac{4}{3}x_8 + \frac{25}{18}.$$

Die optimalen Steuerungen und Zustände, die sich bei der anschließenden Zusammenstellung der Gesamtlösung ergeben, sind in Tabelle 4.4 aufgeführt.

Die Rückkopplungsschleife:

Die in Abb. 4.10 dargestellte Rückkopplungsschleife geht aus der vorwärts gerichteten Schleife (Abb. 4.8) hervor, indem man dort die Stromrichtung des oberen Zweiges umkehrt. Auch jetzt gibt es wieder verschiedene Möglichkeiten für die Durchführung der dynamischen Optimierung. Wir erläutern hier nur den Fall, daß der Beitrag des oberen, rückwärts gerichteten Zweiges bei der Stufe N einbezogen wird. Bezüglich der Aufgabenstellung wird wieder vorausgesetzt, daß für jede Stufe die Modellgleichung den Ausgangsstrom als Funktion des Eingangsstromes und des Steuervektors darstellt und daß die Zielfunktionsbeiträge dieselbe Abhängigkeit zeigen. Gesucht wird das Maximum der Summe dieser Beiträge über alle Stufen, und zwar in Abhängigkeit vom Eingangsvektor $\mathbf{x}_{N+M+L+K+1}$ des Gesamtsystems.

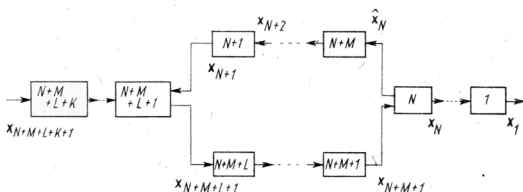


Bild 4.10 Rückkopplungsschleife

Man kann dann die Arbeitsabschnitte bei der Einbettungstechnik folgendermaßen beschreiben:

I. *Rückwärtsrechnung* durch die Stufen 1 bis $N-1$ mit dem Ergebnis $f_{N-1}(\mathbf{x}_N)$.

II. *Separate Behandlung des oberen Zweiges* durch Rückwärtsrechnung durch die Stufenfolge $N+1, \dots, N+M$ mit dem Ziel, den optimalen Beitrag der zugehörigen Teilsumme der Zielfunktion in Abhängigkeit von \mathbf{x}_{N+1} und $\hat{\mathbf{x}}_N$ zu erhalten. Hierbei ist wieder so vorzugehen, daß zunächst über den zur Stufe $N+1$ gehörigen Summanden zu optimieren ist:

$$f_{N+1}(\mathbf{x}_{N+2}, \mathbf{x}_{N+1}) = \max_{\mathbf{u}_{N+1}} G_{N+1}(\mathbf{x}_{N+2}, \mathbf{u}_{N+1}); \quad (4.81)$$

$$\text{Nebenbedingung: } \mathbf{x}_{N+1} = \mathbf{T}_{N+1}(\mathbf{x}_{N+2}, \mathbf{u}_{N+1}).$$

Wenn durch Auflösen der Nebenbedingung der Entscheidungsvektor als Funktion von \mathbf{x}_{N+2} , \mathbf{x}_{N+1} dargestellt werden kann, entfällt die Maximierung. Anschließend werden der Reihe nach die restlichen Stufen $N+2$ bis $N+M$ einbezogen, man erhält schließlich

$$f_{N+M}(\hat{\mathbf{x}}_N, \mathbf{x}_{N+1}).$$

III. Einbeziehung des Rückwärtszweiges an der Stufe N :

$$\begin{aligned} f_N(\mathbf{x}_{N+M+1}, \mathbf{x}_{N+1}) = \max_{\mathbf{u}_N} \{ & G_N(\mathbf{x}_{N+M+1}, \mathbf{u}_N) + f_{N-1}(\mathbf{T}_N(\mathbf{x}_{N+M+1}, \mathbf{u}_N)) \\ & + f_{N+M}(\hat{\mathbf{T}}_N(\mathbf{x}_{N+M+1}, \mathbf{u}_N), \mathbf{x}_{N+1}) \}. \end{aligned} \quad (4.82)$$

Danach wird rückwärts durch die Stufen des unteren Zweiges hindurch gerechnet, wobei man im Ergebnis erhält:

$$f_{N+M+L}(\mathbf{x}_{N+M+L+1}, \mathbf{x}_{N+1}). \quad (4.83)$$

Schließlich ist die Stufe $N+M+L+1$ einzubeziehen ($V = N+M+L$):

$$\begin{aligned} f_{V+1}(\mathbf{x}_{V+2}) = \max_{\mathbf{u}_{V+1}, \mathbf{x}_{N+1}} \{ & G_{V+1}(\mathbf{x}_{V+2}, \mathbf{x}_{N+1}, \mathbf{u}_{V+1}) \\ & + f_V(\mathbf{T}_{V+1}(\mathbf{x}_{V+2}, \mathbf{x}_{N+1}, \mathbf{u}_{V+1}), \mathbf{x}_{N+1}) \}. \end{aligned} \quad (4.84)$$

IV. Zum Abschluß der Einbettungstechnik ist eine *Rückwärtsrechnung durch die Stufen $N+M+L+2$ bis $N+M+L+K$* durchzuführen, die schließlich das geforderte Ergebnis $f_{N+M+L+K}(\mathbf{x}_{N+M+L+K+1})$ liefert.

Mit den während des Rechenganges I.–IV. ebenfalls erhaltenen Abhängigkeiten der Optimalwerte von \mathbf{u}_n und den Modellgleichungen kann dann in üblicher Weise die gesamte Optimallösung zusammengestellt werden.

Beispiel 4.10: Für das in Abb. 4.11 gezeigte fünfstufige System mit Rückkopplungsschleife sind die Modellgleichungen sowie die Summanden der Zielfunktion in Tabelle 4.5 angegeben. Die entsprechenden Funktionen sind wieder einfach gewählt, um die Rechenergebnisse in geschlossener Form erhalten zu können und nicht durch Wertetabellen darstellen zu müssen. Es wird die Aufgabe gestellt,

$$f_5(x_6) = \min_{u_1, \dots, u_5} \sum_{n=1}^5 G_n \quad (4.85)$$

zu bestimmen. Da außer den Modellgleichungen keine Nebenbedingungen für die Entscheidungsvariablen u_i vorliegen, können die im Rechenverlauf notwendigen Minimierungen durch Nullsetzen

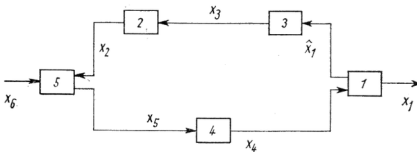


Bild 4.11 Beispiel eines Systems mit Rückkopplungsschleife

Tabelle 4.5 Beispiel eines Systems mit Rückkopplungsschleife

Stufe n	Modellgleichung	Summand G_n der Zielfunktion
1	$\begin{cases} x_1 = x_4 - u_1 \\ \hat{x}_1 = u_1 \end{cases}$	$(x_4 - u_1)^2$
2	$x_2 = x_3 - u_2$	$x_3^2 + u_2^2$
3	$x_3 = u_3 - \hat{x}_1$	$\hat{x}_1^2 + u_3^2$
4	$x_4 = x_5 + u_4$	$x_5^2 + \frac{1}{2} u_4^2$
5	$x_5 = x_6 + x_2 + u_5$	$(x_6 - 1)^2 + u_5^2$

Stufe n	optimale Entscheidung	optimaler Zustand
1	$u_1 = \frac{3}{25} x_6$	$x_1 = \frac{1}{25} x_6; \quad \hat{x}_1 = \frac{3}{25} x_6$
2	$u_2 = \frac{7}{25} x_6$	$x_2 = -\frac{12}{25} x_6$
3	$u_3 = -\frac{2}{25} x_6$	$x_3 = -\frac{5}{25} x_6$
4	$u_4 = -\frac{2}{25} x_6$	$x_4 = \frac{4}{25} x_6$
5	$u_5 = -\frac{7}{25} x_6$	$x_5 = \frac{6}{25} x_6$

der entsprechenden Ableitungen erreicht werden. Wenden wir uns nun dem Rechengang der Einbettungstechnik zu:

I. Entfällt bei diesem Beispiel, da der Verzweigungsstelle keine Stufen vorgelagert sind.

II. Rückwärtsrechnung durch die Stufen 2, 3 des oberen Zweiges.

Zunächst ist auszuführen

$$f_2(x_3, x_2) = \min_{u_2} \{x_3^2 + u_2^2\};$$

Nebenbedingung: $x_2 = x_3 - u_2$.

Da hier durch die Nebenbedingung die Entscheidungsgröße eindeutig durch x_3, x_2 festgelegt wird,

$$u_2 = x_3 - x_2,$$

entfällt die Minimierung, und man erhält:

$$f_2(x_3, x_2) = 2x_3^2 + x_2^2 - 2x_3x_2.$$

Jetzt wird Stufe 3 einbezogen:

$$f_3(\hat{x}_1, x_2) = \min_{u_3} \{(\hat{x}_1^2 + u_3^2) + f_2(x_3, x_2)\}.$$

Trägt man hier zunächst $x_3 = u_3 - \hat{x}_1$ ein und setzt danach die Ableitung nach u_3 gleich null, so ergibt sich

$$u_3(\hat{x}_1, x_2) = \frac{1}{3}x_2 + \frac{2}{3}\hat{x}_1$$

und damit weiter

$$f_3(\hat{x}_1, x_2) = \frac{2}{3}x_2^2 + \frac{5}{3}\hat{x}_1^2 + \frac{2}{3}x_2 \cdot \hat{x}_1.$$

III. Die Beiträge des oberen Zweiges und der Stufe 1 werden gekoppelt:

$$f_1(x_4, x_2) = \min_{u_1} \{(x_4 - u_1)^2 + \frac{2}{3}x_2^2 + \frac{5}{3}u_1^2 + \frac{2}{3}x_2u_1\} = \min_{u_1} \bar{f},$$

wobei $\hat{x}_1 = u_1$ eingesetzt wurde. Aus $\frac{\partial f}{\partial u_1} = 0$ folgt:

$$u_1(x_4, x_2) = \frac{3}{8}x_4 - \frac{1}{8}x_2$$

und weiter

$$f_1(x_4, x_2) = \frac{5}{8}x_4^2 + \frac{1}{4}x_4x_2 + \frac{15}{24}x_2^2.$$

Hinzunahme der Stufe 4 ergibt durch entsprechende Rechnung

$$u_4(x_5, x_2) = -\frac{1}{9}x_2 - \frac{5}{9}x_5,$$

$$f_4(x_5, x_2) = \frac{23}{18}x_5^2 + \frac{1}{9}x_5x_2 + \frac{11}{18}x_2^2.$$

Nun muß noch Stufe 5 einbezogen werden:

$$f_5(x_6) = \min_{u_5, x_2} \{(x_6 - 1)^2 + u_5^2 + f_4(x_5, x_2)\},$$

wobei vor der Minimierung $x_5 = x_6 + x_2 + u_5$ einzusetzen ist. Man erhält

$$u_5 = -\frac{7}{25}x_6; \quad x_2 = -\frac{12}{25}x_6$$

und weiter als Optimum der Gesamtziel Funktion

$$f_5(x_6) = \frac{32}{25}x_6^2 - 2x_6 + 1.$$

Damit ist die Einbettungstechnik abgeschlossen, der Rechenabschnitt IV. entfällt hier, da links von der Vereinigungsstelle 5 keine Stufen mehr vorhanden sind.

Die Zusammenstellung der optimalen Lösung kann nun leicht vorgenommen werden, die Ergebnisse sind in Tab. 4.5 aufgeschrieben.

Abschließend ist folgende Bemerkung angebracht: Alle im Abschnitt über nicht reihenförmige Systeme enthaltenen Beispiele wurden bewußt einfach gewählt, um den Lösungsverlauf übersichtlich und für den Leser mit erträglichem Zeitaufwand nachrechenbar zu gestalten. Die Optimalfunktionen konnten in allen Fällen in geschlossener Form erhalten werden, und die Extremwertsuche machte wegen des Fehlens von Nebenbedingungen in Ungleichungsform keine Schwierigkeiten. Bei praktischen

Problemen werden solche günstigen Bedingungen oft nicht vorhanden sein. Meist wird man die Optimalfunktionen nur an diskreten Stellen, über einem Punktgitter berechnen können, d. h. in Tabellenform abspeichern müssen. Mit steigender Stufenzahl und vor allem bei ansteigender Dimension der Zustandsvektoren \mathbf{x} erhöht sich der Speicherbedarf entsprechend stark und kann die praktische Durchführbarkeit einer Berechnung in Frage stellen. Das Speicherproblem tritt vor allem auch dann auf, wenn die Optimalfunktionen teilweise in Abhängigkeit von mehreren Zustandsvektoren tabellarisch abzuspeichern sind. Hinzu kommt die Schwierigkeit, bei den Rekursionsschritten in „mehrdimensionalen“ Tabellen interpolieren zu müssen, wodurch die Rechenzeit entsprechend ansteigt. Der bei umfangreicheren Problemen stark ansteigende numerische und rechentechnische Aufwand darf also keineswegs unterschätzt werden und ist bei der Planung einer Problembearbeitung sorgfältig abzuschätzen.

4.3.5. Suboptimale Steuerungen

Die Problemstellung:

Als Lösung der in 4.3.1. formulierten Grundaufgabe erhält man die bezüglich des verwendeten Optimierungskriteriums beste Steuerung. Daneben wird man sich in vielen Fällen auch für *suboptimale Lösungen*, d. h. für die „zweitbeste, drittbeste, ...“ Entscheidung interessieren. Die Kenntnis suboptimaler Lösungen kann aus verschiedenen Gründen nützlich sein. Zunächst hat man, falls die optimale Variante aus irgendeinem Grund nicht realisierbar ist, entsprechende Alternativen in der Hand. Zum anderen wird der Einblick in die Eigenschaften des modellierten Prozesses erweitert, wenn man weiß, wie er sich in der Nachbarschaft des optimalen Punktes verhält. Ein drittes Argument ist mehr rechnerischer Art: Wie wir gesehen haben, ist es für die praktische Durchführung der dynamischen Optimierung meist notwendig, die Zustandsbereiche X_n zu diskretisieren. Aus Gründen der Speicherkapazität wird man bei größeren Problemen oft mit relativ groben Punktgittern arbeiten müssen, so daß die Frage auftritt, inwieweit die errechnete Optimallösung numerisch zuverlässig ist. Die Kenntnis suboptimaler Lösungen erleichtert eine solche Einschätzung. Weichen die ersten suboptimalen Entscheidungen stark von der optimalen ab, so wird dies oft ein Hinweis dafür sein, daß die verwendete Näherung zu grob war. Sind andererseits die Abweichungen gering, so erscheint die optimale Lösung verläßlich.

Für das folgende setzen wir voraus, daß die Steuervariablen jeder Stufe nur endlich viele Werte annehmen können.¹⁾ Die Menge

$$V = \{\mathbf{u} \mid \mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N); \quad \mathbf{u}_n \in U_n, n = 1, \dots, N\} \quad (4.86)$$

enthält dann ebenfalls nur endlich viele Elemente.

Definition 4.5: *V sei zerlegt in Untermengen $V = V_1 \cup V_2 \cup \dots; V_i \cap V_k = \emptyset (i \neq k)$, so daß gilt:*

$$G(\mathbf{u}) = \text{konstant} = g_k, \quad \forall \mathbf{u} \in V_k, g_k > g_{k+1}, k = 1, 2, \dots \quad (4.87)$$

Dann heißt jedes $\mathbf{u} \in V_k$ k.-beste Steuerung.

¹⁾ Gegebenenfalls kann man diese Voraussetzung durch Diskretisierung der Entscheidungsbereiche U_n schaffen.

V_1 enthält also speziell die optimalen Steuerungen. Eine Möglichkeit zur Ermittlung suboptimaler Steuerungen wurde bereits von Bellman und Kalaba aufgezeigt, indem diese von einer erweiterten Form des Optimalprinzips ausgingen. Der Algorithmus wurde von Fan und Wang [9] in rechentechnischer Hinsicht weiter ausgebaut.

Der Algorithmus:

Ein Rechenverfahren, mit dem man außer der optimalen auch die 2.- bis k .-besten Steuerungen ermitteln kann, ergibt sich analog zum Verfahren der normalen dynamischen Optimierung, wenn man von folgender Erweiterung des Optimalprinzips ausgeht (wir betrachten ein N -stufiges System gemäß Abb. 4.1b).

Eine k .-beste Steuerung hat folgende Eigenschaft: Unabhängig vom Eintrittszustand in die erste Stufe (Nummer N) und der in dieser gewählten Steuerung stellen die Steuerungen der folgenden Stufen (Nummern $N-1$ bis 1) eine der 1.- bis k .-besten Steuerungen des um die erste Stufe verkürzten Systems bez. des aus der ersten Stufe resultierenden Zustandes (\mathbf{x}_N) dar.

Aus dem erweiterten Optimalprinzip folgt nun analog zum einfachen Fall die entsprechende Rekursionsgleichung. Im Gegensatz zu Formel (4.25) haben wir im n -ten Schritt der Einbettungstechnik bei der Maximierung über \mathbf{u}_n nicht nur die Optimalwerte f_{n-1} des vorangegangenen Schrittes zuzulassen, sondern – wenn wir die k .-beste Entscheidung suchen – auch die 2., 3., ..., k .-besten Werte. Im folgenden soll bedeuten:

$\max^{(j)}y$ den j -größten Wert von y ,

$f_n^{(j)}(\mathbf{x}_{n+1})$ den j -größten Wert des zur Stufenfolge $n, \dots, 1$ gehörenden Teils der Zielfunktion (4.24) in Abhängigkeit vom Eingangszustand \mathbf{x}_{n+1} ,

$u_n^{(j)}(\mathbf{x}_{n+1})$ die zugehörige Steuerung auf der Stufe n , ebenfalls in Abhängigkeit vom Eingangszustand \mathbf{x}_{n+1} .

Dann gilt

$$f_n^{(j)}(\mathbf{x}_{n+1}) = \max_{u_n, i}^{(j)} \{ G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}^{(i)}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n)) \}; \quad (4.88)$$

$$\mathbf{x}_{n+1} \in X_{n+1}; \quad \mathbf{u}_n \in U_n; \quad i \in \{1, 2, \dots, j\}.$$

Diese Rekursionsformel ist mit $j = 1, 2, \dots, k$ der Reihe nach für die Stufen $n = 1, 2, \dots, N$ anzuwenden, wobei zu beachten ist:

$$f_0^{(j)}(\mathbf{x}_1) \equiv 0, \quad j = 1, 2, \dots, k; \quad X_{N+1} = \{\alpha\}. \quad (4.89)$$

Im Gegensatz zum einfachen Algorithmus aus Abschnitt 4.3.1. entstehen jetzt in jedem Schritt der Einbettungstechnik k Tabellen der in Tab. 4.6 angedeuteten Art. Dabei vermerkt $i_n^{(j)}(\mathbf{x}_{n+1})$, daß der Maximalwert $f_n^{(j)}(\mathbf{x}_{n+1})$ in (4.88) unter Verwendung von $f_{n-1}^{(i)}$, d. h. der i -besten Steuerung der Stufenfolge $n-1, \dots, 1$ erhalten wurde. Im N -ten Schritt ergeben sich die Zahlenwerte

$$f_N^{(j)}(\mathbf{x}_{N+1}) = f_N^{(j)}(\alpha), \quad j = 1, \dots, k,$$

Tabelle 4.6 Tabelle Nr. j der Stufe n (j -beste Werte)

\mathbf{x}_{n+1}	...
$f_n^{(j)}(\mathbf{x}_{n+1})$...
$\mathbf{u}_n^{(j)}(\mathbf{x}_{n+1})$...
$i_n^{(j)}(\mathbf{x}_{n+1})$...

Diese Tabelle ist für $j = 1, 2, \dots, k$ zu berechnen
(im Falle $j = 1$ fehlt der Bestandteil $i_1^{(j)}(x_2)$)

und dies sind die j -größten Werte der Zielfunktion G , speziell ist $f_N^{(1)}(\alpha)$ der Optimalwert. Die zugehörigen Vektoren

$$\mathbf{u}_N^{(j)} = \mathbf{u}_N^{(j)}(\alpha), \quad j = 1, \dots, k,$$

bedeuten die tatsächlichen auf Stufe N anzuwendenden Steuerungen für die optimalen und suboptimalen Lösungen.

Die zweite Phase des Algorithmus, die Zusammenstellung der optimalen und suboptimalen Lösungen, erfolgt dann analog wie im einfachen Fall des Abschnitts 4.3.1.

Zusammenstellung der j -besten Steuerung:

$$(a) \text{ Anfangswerte: } \mathbf{x}_{N+1}^{(j)} = \alpha, \quad i_{N+1} = j. \quad (4.90)$$

(b) Für $n = N, N-1, \dots, 1$ ist der Reihe nach folgendes auszuführen:

$$\begin{aligned} \mathbf{u}_n^{(j)} &= \mathbf{u}_n^{(i_{n+1})}(\mathbf{x}_{n+1}^{(j)}) \text{ aus Tab. } i_{n+1} \text{ der Stufe } n, \\ \mathbf{x}_n^{(j)} &= \mathbf{T}_n(\mathbf{x}_{n+1}^{(j)}, \mathbf{u}_n^{(j)}) \text{ mittels Modellgleichung der Stufe } n, \\ i_n &= i_n^{(j)}(\mathbf{x}_{n+1}^{(j)}) \text{ aus Tab. } j \text{ der Stufe } n. \end{aligned} \quad (4.91)$$

Mit $j = 1$ erhält man so die optimale, mit $j = 2, 3, \dots, k$ die suboptimalen Lösungen. Auch der hier beschriebene Algorithmus ist offenbar leicht zu programmieren. Im Vergleich zur „normalen“ dynamischen Optimierung erhöht sich im wesentlichen nur die Anzahl der Größenvergleiche, so daß die Rechenzeit nicht sehr stark ansteigt, wenn man außer der optimalen Lösung auch die ersten suboptimalen Steuerungen bestimmt. Der für die Tabellen der Einbettungstechnik benötigte Speicherplatz steigt aber auf das k -fache, wenn man bis zur k -besten Variante geht.

4.3.6. Stetige Probleme

Die stetige Form der Grundaufgabe und der Rekursionsbeziehung:

Wir wenden uns jetzt Systemen zu, bei denen die Koordinaten des Zustandsvektors \mathbf{x} und des Steuervektors \mathbf{u} in stetiger Weise von einem Parameter t abhängen, wie das in Abb. 4.12 veranschaulicht ist. Die sinnngemäße Übertragung der in Abschnitt 4.3.1.

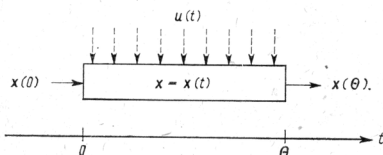


Bild 4.12 Stetiger Entscheidungsprozeß

für den diskreten Fall formulierten Grundaufgabe auf den stetigen Fall führt zu folgender Problemstellung, vgl. Abb. 4.12: Gegeben seien

- als Modellgleichungen des Systems ein Satz von gewöhnlichen Differentialgleichungen der Form¹⁾

$$\frac{dx}{dt} = T(t, x, u), \quad 0 \leq t \leq \theta, \quad (4.92)$$

- Beschränkungen für die Steuer- und Zustandsvariablen

$$x(t) \in X, \quad u(t) \in U, \quad 0 \leq t \leq \theta, \quad (4.93)$$

- der Anfangszustand

$$x(0) = \alpha. \quad (4.94)$$

Gesucht werden

- aus der Menge der zulässigen Steuerungen diejenige optimale $u^*(t)$, $0 \leq t \leq \theta$, für die gilt

$$G = \int_0^\theta g(t, x, u) dt = \max! \quad (4.95)$$

- der zugehörige optimale Zustandsverlauf $x^*(t)$, $0 \leq t \leq \theta$.

In konkreten Fällen wird der Parameter t oft die Bedeutung der Zeit (vom Beginn des Prozesses an gerechnet) oder einer Länge (vom Eingang in die Stufe an gerechnet) haben. t entspricht dem Stufenindex n , den wir bei den diskreten Modellen hatten, wobei wir uns nicht daran stören wollen, daß t jetzt „in Stromrichtung“ zunimmt, während wir n seinerzeit entgegen der Stromrichtung zählten. Aus der früheren Summenform der Zielfunktion ist ein Integral geworden, wobei $g(t, x, u) dt$ den Beitrag beschreibt, der zu einem Stufenabschnitt der „Länge dt “ gehört.

Das Optimalprinzip von Bellman kann jetzt folgendermaßen formuliert werden (der Beweis ergibt sich analog zum diskreten Fall):

Eine optimale Steuerung hat folgende Eigenschaft: Unabhängig vom Anfangszustand $x(0)$ und der im Intervall $0 \leq t \leq \tau$ verwendeten Steuerung $u(t)$ ist die Steuerfunktion

$$u(t), \quad \tau < t \leq \theta,$$

¹⁾ Die nach der Ableitung des Zustandsvektors aufgelöste Form der Modellgleichung wurde hier gewählt, um die Darstellung zu vereinfachen, die Methode ist auch in anderen Fällen durchführbar.

optimal bezüglich des zum Parameterintervall $(\tau, \theta]$ gehörigen Abschnitt des Prozesses, wenn als dessen Anfangszustand der aus dem vorangegangenen Intervall $[0, \tau]$ resultierende Vektor $\mathbf{x}(\tau)$ genommen wird.

Hieraus ergibt sich durch Überlegungen, die denen beim diskreten Modell ebenfalls völlig analog sind, die folgende Funktionalgleichung:

$$f(\tau, \mathbf{x}(\tau)) = \max_{\substack{\mathbf{u}(t) \in U \\ t \in [\tau, \tau + \Delta t]}} \left\{ \int_{\tau}^{\tau + \Delta t} g(t, \mathbf{x}, \mathbf{u}) dt + f(\tau + \Delta t, \mathbf{x}(\tau + \Delta t)) \right\};$$

$$f(\theta, \mathbf{x}(\theta)) = 0; \quad 0 \leq \tau \leq \theta - \Delta t. \quad (4.96)$$

Diese Beziehung entspricht der Rekursionsgleichung, die bei den diskreten mehrstufigen Prozessen anzuwenden ist. Indem man den Klammersausdruck in eine Reihe nach Potenzen von Δt entwickelt, einen Grenzübergang $\Delta t \rightarrow 0$ vornimmt, und die für das Maximum notwendigen Bedingungen ausnutzt, kann man ein System von partiellen Differentialgleichungen ableiten, welches im Prinzip eine Ermittlung der optimalen Lösung gestattet [18], [19]. Zur Gewinnung numerischer Lösungen ist dieser Weg bei den meisten praktischen Aufgaben wenig geeignet. Zweckmäßiger ist es, das stetige Problem durch ein diskretes mehrstufiges Problem anzunähern. Dieser Bearbeitungsweise wollen wir uns jetzt zuwenden.

Approximation durch ein diskretes mehrstufiges System:

Zur Diskretisierung des durch (4.92–4.95) beschriebenen Problems zerlegen wir das t -Intervall $[0, \theta]$ zunächst in N Abschnitte

$$\Delta t = \frac{\theta}{N}; \quad t_{n+1} = \theta - n \cdot \Delta t, \quad n = 0, \dots, N, \quad (4.97)$$

die wir der Einfachheit halber als gleichlang annehmen wollen. Durch diese Maßnahme wird die stetige Stufe in N Teilstufen zerschnitten, vgl. Abb. 4.13. In jeder Teilstufe ersetzen wir den tatsächlichen Verlauf der Steuerung $\mathbf{u}(t)$ und des Zustandes $\mathbf{x}(t)$ durch einen konstanten Verlauf:

$$\mathbf{u}(t) \approx \mathbf{u}(t_n) = \text{konstant} = \mathbf{u}_n;$$

$$\mathbf{x}(t) \approx \mathbf{x}(t_{n+1}) = \text{konstant} = \mathbf{x}_{n+1}; \quad t_{n+1} < t \leq t_n. \quad (4.98)$$

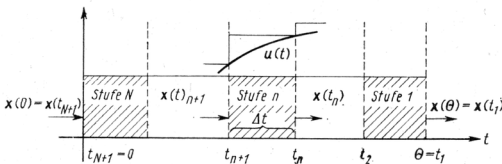


Bild 4.13 Diskretisierung des stetigen Problems

Wir nähern also den i. allg. stetigen Verlauf der Kurven durch Treppenkurven an. In jedem Teilintervall n werden weiter die Differentialgleichungen (4.92) durch Differenzgleichungen approximiert:

$$\frac{dx}{dt} \approx \frac{x(t_n) - x(t_{n+1})}{\Delta t} = \frac{x_n - x_{n+1}}{\Delta t},$$

$$T(t, x, u) \approx T(t_n, x_{n+1}, u_n) = \hat{T}_n(x_{n+1}, u_n). \quad (4.99)$$

Durch Gleichsetzen der Näherungsausdrücke erhalten wir als Modellgleichung der Stufe n :

$$x_n = x_{n+1} + \hat{T}_n(x_{n+1}, u_n) \cdot \Delta t = T_n(x_{n+1}, u_n). \quad (4.100)$$

Schließlich nähern wir das Integral (4.95) folgendermaßen durch eine Summe an:

$$G = \int_0^{\theta} g(t, x, u) dt = \sum_{n=1}^N \int_{t_{n+1}}^{t_n} g(t, x, u) dt$$

$$\approx \sum_{n=1}^N g(t_n, x_{n+1}, u_n) \cdot \Delta t = \sum_{n=1}^N G_n(x_{n+1}, u_n). \quad (4.101)$$

Entsprechend ergeben sich aus (4.93) Beschränkungen für die Teilstufen

$$x_n \in X_n, \quad u_n \in U_n, \quad n = 1, \dots, N. \quad (4.102)$$

In den Formeln (4.100–4.102) und $x(0) = x_{N+1} = \alpha$ haben wir das beabsichtigte diskrete Näherungsmodell vor uns, das nun mit dem im Abschnitt 4.3.1. beschriebenen Algorithmus bearbeitet werden kann.

Diese Näherungstechnik ist natürlich auch für Problemtypen anwendbar, die von der Grundaufgabe abweichen. Durch Lösung des Näherungsproblems erhalten wir Folgen von Approximationspunkten für die optimalen Kurvenverläufe der Steuerung $u(t)$ und des Zustandes $x(t)$, Zwischenwerte kann man bei Bedarf durch Interpolation berechnen. Die Güte der Näherung kann durch die Wahl einer hinreichend kleinen Schrittlänge Δt gesteigert werden.

Beispiel 4.11: Es bestehe die Aufgabe, die optimale Steuerung zu folgendem Problem zu bestimmen (x und u eindimensional):

$$\frac{dx}{dt} = u, \quad 0 \leq t \leq 5,$$

$$0 \leq x(t) \leq 3, \quad -2 \leq u(t) \leq 2,$$

$$x(0) = 2,$$

$$G = \int_0^5 (x^2 + u^2) dt = \min!$$

Wir zerlegen das Intervall $[0, 5]$ durch die Teilpunkte

$$t_{n+1} = 5 - n \cdot \Delta t; \quad n = 0, \dots, N,$$

in N Abschnitte gleicher Länge Δt . Gemäß den Formeln (4.98–4.102) ergibt sich das folgende N -stufige Näherungsproblem:

$$x_n = x_{n+1} + u_n \cdot \Delta t, \quad n = 1, \dots, N,$$

$$0 \leq x_n \leq 3, \quad -2 \leq u_n \leq 2,$$

$$x_{N+1} = 2,$$

$$G = \sum_{n=1}^N (x_{n+1}^2 + u_n^2) \cdot \Delta t = \min!$$

Dieses kann nun mit dem Algorithmus der diskreten dynamischen Optimierung in der üblichen Art behandelt werden. Wir deuten für den Fall $N = 5, \Delta t = 1$, die ersten Schritte der Einbettungstechnik noch an:

1. Schritt:

$$f_1(x_2) = \min_{u_1 \in [-2, 2]} \{x_2^2 + u_1^2\}.$$

Der Klammerausdruck stellt eine Parabel dar, deren Scheitel bei $u_1 = 0$ liegt. Das Minimum befindet sich also, unabhängig vom x_2 -Wert, an der Stelle Null, d. h.,

$$u_1(x_2) = 0,$$

$$f_1(x_2) = x_2^2, \quad 0 \leq x_2 \leq 3.$$

Tabelle 4.7 Lösung eines Problems der optimalen Steuerung

t	Exakte Lösung		Lösung für 5-stufiges Näherungsmodell		Lösung für 10-stufiges Näherungsmodell	
	u_{opt}	x_{opt}	u_{opt}	x_{opt}	u_{opt}	x_{opt}
0	-2,000	2,000		2,000	-1,561	2,000
0,5	-1,213	1,213	-1,235		-0,952	1,219
1	-0,735	0,736		0,765	-0,580	0,744
1,5	-0,446	0,447	-0,471		-0,353	0,454
2	-0,271	0,270		0,294	-0,214	0,277
2,5	-0,163	0,165	-0,176		-0,129	0,170
3	-0,098	0,101		0,118	-0,076	0,106
3,5	-0,057	0,063	-0,059		-0,042	0,068
4	-0,032	0,042		0,059	-0,019	0,047
4,5	-0,014	0,030	0		0	0,037
5	0	0,027		0,059		0,037
G_{opt}	4,0		6,47		5,12	

2. Schritt:

$$\begin{aligned}
 f_2(x_3) &= \min_{u_2 \in [-2, 2]} \{ (x_3^2 + u_2^2) + (x_3 + u_2)^2 \} \\
 &= \min_{u_2} 2 \cdot \left\{ (u_2 + \frac{1}{2} x_3)^2 + \frac{3}{4} x_3^2 \right\}.
 \end{aligned}$$

Diesmal liegt eine Parabel mit dem Scheitel bei $u_1 = -\frac{1}{2}x_3$ vor. Da für alle $x_3 \in [0, 3]$ diese Stelle im Inneren des Steuerbereiches $[-2, 2]$ liegt, erhalten wir die Optimalfunktion:

$$\begin{aligned}
 u_2(x_3) &= -\frac{1}{2}x_3; \\
 f_2(x_3) &= \frac{3}{4}x_3^2; \quad 0 \leq x_3 \leq 3.
 \end{aligned}$$

Entsprechend ist die Einbettungstechnik über die restlichen drei Stufen fortzusetzen, wonach die optimale Lösung durch Vorwärtsrechnung zusammengestellt werden kann. Die Ergebnisse sind in Tab. 4.7 zusammengestellt. Zum Vergleich sind dort auch die Zahlenwerte für ein Modell mit $N = 10$, Schrittlänge $\Delta t = 0,5$ und für die exakte Lösung aufgeführt. Letztere kann durch Anwendung des Maximumprinzips gefunden werden (vgl. Abschn. 4.4.2.). Die Werte der Näherungslösungen weichen bei diesem Beispiel noch stark von den exakten Ergebnissen ab, offenbar ist die verwendete Intervallteilung nicht fein genug. Aber selbst solche groben Näherungen können von praktischem Nutzen sein, beispielsweise dann, wenn man ein Steuerungsproblem mit Hilfe des Maximumprinzips lösen will und dort als Ausgangspunkt für die numerische Behandlung den ungefähren Verlauf der Steuerfunktion $u(t)$ benötigt.

4.4. Kombinierte Anwendung der dynamischen Optimierung und des Maximumprinzips

In den vorangegangenen Abschnitten haben wir in den vom Maximumprinzip einerseits und vom Bellmanschen Optimalprinzip andererseits ausgehenden Methoden wirksame Hilfsmittel zur Lösung von Problemen der optimalen Steuerung kennengelernt. Dabei zeigte es sich, daß nicht etwa jede dieser Methoden eindeutig einer bestimmten Problemklasse zugeordnet ist, sondern daß in vielen Fällen prinzipiell beide Vorgehensweisen anwendbar sind. Man kann sicher keine Aussage derart machen, daß eine dieser Methoden der anderen grundsätzlich überlegen ist, vielmehr hat jede ihre Vor- und Nachteile. Indem wir im folgenden beide Methoden von verschiedenen Gesichtspunkten aus qualitativ vergleichen, werden sich aber Hinweise dafür ergeben, wann in der Regel die eine, wann die andere zweckmäßiger einzusetzen ist. Im Anschluß daran erläutern wir Möglichkeiten der kombinierten Anwendung beider Methoden.

4.4.1. Vergleich der Methoden

Um einige Vor- und Nachteile sichtbar zu machen, gehen wir von der folgenden kurzen Charakterisierung beider Methoden aus.

Dynamische Optimierung: Das Problem wird in N nacheinander zu lösende Teilprobleme für die Systemstufen zerlegt. In jedem Schritt wird der gesamte zulässige Bereich des Eingangszustandes der betreffenden Stufe durch ein Punktgitter überdeckt, für jeden Gitterpunkt wird „auf Vorrat“ das Teiloptimierungsproblem gelöst, und alle Ergebnisse werden abgespeichert. In den entstehenden Tabellen ist das

gesuchte optimale Regime eingebettet, es wird anschließend, ausgehend vom gegebenen Anfangszustand des Gesamtprozesses, herausgesucht.

Maximumprinzip: Durch eine Reihe notwendiger Bedingungen (zu bilden u.a. durch Ableiten der Hamiltonfunktion) wird aus der Menge aller zulässigen Lösungen sofort eine kleine Teilmenge abgegrenzt, in der (falls existent) die optimale Lösung enthalten ist. Die numerische Auswertung der Bedingungen erfolgt normalerweise in Form einer iterativen oder Trial-error-Rechnung, d.h., es sind anfangs einige Variablenwerte (Randwerte) zu schätzen und danach bis zum Abgleich zu verbessern.

Diese Gegenüberstellung weist auf folgende Unterschiede und Aspekte der praktischen Anwendbarkeit beider Methoden hin [9]:

(a) *Rechentechnischer Aufwand:* Durch Verwendung der „Einbettungstechnik“ ergibt sich bei der dynamischen Optimierung mit steigender Dimension des Problems schnell ein außerordentlich hoher Bedarf an Speicherplatz und Rechenzeit. Im Abschnitt 4.3.1. wurde auf diesen Nachteil schon ausführlicher eingegangen, wobei sich zeigte, daß der Aufwand in erster Linie durch die Zahl der Zustandsvariablen je Stufe bestimmt wird. Man kann natürlich den Speicherbedarf dadurch senken, daß man die Punktgitter über die Zustandsbereiche X_n relativ weitmaschig anlegt. Dadurch wird aber die Zuverlässigkeit der in beiden Abschnitten des Algorithmus notwendigen Interpolationsrechnungen fragwürdig. Wie schon bemerkt, kann man einschätzen, daß aus diesen Gründen die praktische Anwendung der dynamischen Optimierung problematisch wird, wenn mehr als 3 bis 4 Zustandsvariable je Stufe vorhanden sind. Aus dieser Sicht dürfte die dynamische Optimierung oft schlecht zur Behandlung von Systemen geeignet sein, deren Schaltungsstruktur Schleifen enthält (vgl. Abschn. 4.3.4.), da dann die Rekursionsrechnung u.U. durch zusätzliche Zustandsvariable belastet ist. Auch die Anwendung auf stetige Probleme wird oft an rechentechnischen Schwierigkeiten scheitern. Hier führt das Bellmansche Prinzip auf partielle Differentialgleichungen. Diese können zwar durch Diskretisierung in Differenzgleichungen umgewandelt werden, aber man wird ein diskretes Näherungsmodell mit vielen Stufen verwenden müssen, um eine ausreichende Genauigkeit zu erzielen. Bei Verwendung des Maximumprinzips sind diese Schwierigkeiten hinsichtlich des rechentechnischen Aufwands, insbesondere was den Speicherbedarf betrifft, nicht in dem Maße vorhanden. Vor allem bei stetigen Modellen wird es deshalb oft vorzuziehen sein.

(b) *Nebenbedingungen für die Zustandsvariablen:* Das Vorhandensein von Beschränkungen der Form $x_n \in X_n$ macht bei der Methode der dynamischen Optimierung keine Schwierigkeiten. Solche Nebenbedingungen werden im Rechengang automatisch dadurch eingehalten, daß man die Punktgitter der Einbettungstechnik entsprechend abgrenzt. Beim Maximumprinzip führen Zustandsbeschränkungen dagegen zu ersten Komplikationen. Es ist zwar möglich, die Bedingungen des Maximumprinzips auf diesen Fall auszudehnen, aber die Beziehungen werden dann sehr kompliziert und sind numerisch schlecht auszuwerten. Da bei praktischen Problemen sehr häufig Zustandsbeschränkungen vorliegen, wird hier ein wesentlicher Vorteil der dynamischen Optimierung sichtbar.

(c) *Form der Modellgleichungen:* Wie wir gesehen haben, erhält man beim Maximumprinzip die Gleichungen für die adjungierten Variablen durch Bilden der partiellen Ableitungen der Hamiltonfunktion. Dies setzt voraus, daß die Modellgleichungen

in einer analytischen, stetig nach den Zustandsvariablen differenzierbaren Form vorliegen. Für die Anwendbarkeit des Algorithmus der dynamischen Optimierung ist dies keineswegs notwendig. Hier ist es nur erforderlich, daß durch die Beziehungen $\mathbf{x}_n = \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n)$ eine eindeutige Zuordnung $(\mathbf{x}_{n+1}, \mathbf{u}_n) \rightarrow \mathbf{x}_n$ gegeben ist, die aber durchaus auch in Tabellenform oder durch verbale Regeln vorliegen kann. Beispiele dieser Art haben wir in den Abschnitten 4.3.3., 4.3.5. behandelt.

(d) *Sicherung des globalen Extremwertes*: Wenn bei einem Problem im zulässigen Bereich mehrere lokale Extremwerte vorhanden sind, so ist bei Anwendung des Maximumprinzips (Auswertung notwendiger Bedingungen!) i. allg. nicht gesichert, daß die Behandlung den globalen Extremwert, d. h., wirklich das Optimum liefert. Die dynamische Optimierung hingegen führt im Prinzip immer zum Optimum, da ja durch die Einbettungstechnik der gesamte zulässige Bereich abgesucht wird. (Praktisch kann allerdings durch ein zu grobes Punktgitter und die damit verbundenen Interpolationsfehler das Optimum „übersehen“ werden).

4.4.2. Möglichkeiten der kombinierten Anwendung beider Methoden

Wenn wir eben die dynamische Optimierung und das Maximumprinzip miteinander verglichen, ihre Vor- und Nachteile herausgestellt haben, so liegt die Frage nahe, ob man nicht durch eine Kombination beider Methoden in gewissen Fällen die Vorteile vereinigen kann. Auf zwei derartige Möglichkeiten gehen wir im folgenden kurz ein.

Die parallele Verbindung der Methoden:

Die Bezeichnung parallele Verbindung von dynamischer Optimierung und Maximumprinzip wurde von Fan [9] eingeführt. Gemeint ist damit die folgende Vorgehensweise. Es wird beabsichtigt, die optimale Lösung eines vorgelegten Problems mit Hilfe des Maximumprinzips zu ermitteln. Um der Gefahr zu entgehen, hierbei eventuell nur ein relatives, vom Optimum verschiedenes Extremum zu erhalten, wird zunächst die ungefähre Lage des Optimums mit Hilfe der dynamischen Optimierung berechnet (Verwendung relativ weitmaschiger Punktgitter). Von dieser Näherungslösung ausgehend wird anschließend die genaue optimale Lösung unter Anwendung des Maximumprinzips bestimmt.

Die sequentielle Verbindung der Methoden:

Der Vorschlag zu einer sog. sequentiellen Verbindung der beiden Methoden geht auf Lee zurück [14]. Er geht von der Tatsache aus, daß durch den Algorithmus der dynamischen Optimierung in keiner Weise festgelegt wird, in welcher Weise das zu optimierende Gesamtsystem in Stufen zerlegt wird und mit welchem Verfahren in den einzelnen Schritten der Einbettungstechnik die Teiloptimierungen in den Stufen durchgeführt werden. So kann man, wenn dies zweckmäßig erscheint, in einzelnen Stufen z. B. auch das Maximumprinzip zur Optimierung verwenden. Dadurch wird folgende Vorgehensweise möglich:

- (1) Das Gesamtsystem wird in zweckmäßiger Weise in Subsysteme zerlegt (wobei ein Subsystem aus einer einzigen Stufe oder auch aus einer Kombination mehrerer Stufen bestehen kann).

- (2) Die Schaltungsstruktur der Subsysteme wird mittels der Methode der dynamischen Optimierung bearbeitet.
- (3) Bei den Teiloptimierungen in den Subsystemen, die im Verlaufe der Einbettungstechnik durchzuführen sind, wird je nach Zweckmäßigkeit die dynamische Optimierung oder das Maximumprinzip (oder eine andere geeignete Optimierungsmethode) angewendet.

Wir erläutern diese Vorgehensweise an zwei Formen von Reihenschaltungen, damit dürfte das Prinzip dann auch für kompliziertere Strukturen klar sein.

Zuerst wird das aus einer stetigen und N nachgeschalteten diskreten Stufen bestehende System der Abb. 4.14a betrachtet. Das Optimierungsproblem sei durch folgende Beziehungen beschrieben:

Modellgleichungen:

$$\mathbf{x}_n = \mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n), \quad n = 1, 2, \dots, N, \quad (4.103)$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{T}(t, \mathbf{x}, \mathbf{u}), \quad 0 \leq t \leq T. \quad (4.104)$$

Beschränkungen:

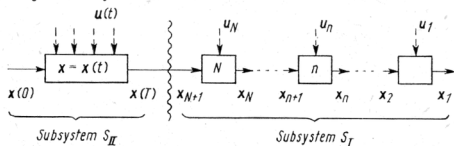
$$\mathbf{x}_n \in X_n; \quad \mathbf{u}_n \in U_n; \quad n = 1, 2, \dots, N, \quad (4.105)$$

$$\mathbf{u}(t) \in U; \quad 0 \leq t \leq T, \quad (4.106)$$

sowie

$$\mathbf{x}_{N+1} \in X_{N+1}; \quad \mathbf{x}(T) = \mathbf{x}_{N+1}; \quad \mathbf{x}(0) = \alpha \quad (\text{fest vorgegeben}). \quad (4.107)$$

a) Stetige Stufe vorgeschaltet



b) Stetige Stufe nachgeschaltet

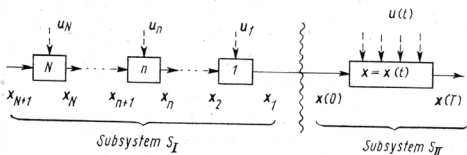


Bild 4.14 Systeme mit N diskreten und einer stetigen Stufe

Optimierungsziel:

$$G = \int_0^T g(t, \mathbf{x}, \mathbf{u}) dt + \sum_{n=1}^N G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) = \max! \quad (4.108)$$

Wir zerlegen das System in der angegebenen Art in die beiden Subsysteme I, II und beschließen, die Teiloptimierung in S_I mit der dynamischen Optimierung, diejenige in S_{II} mit dem Maximumprinzip auszuführen.

Die Anwendung des Bellmanschen Optimalprinzips auf das aus den zwei Bestandteilen S_I , S_{II} zusammengesetzte Gesamtsystem liefert als Rekursionsformel:

$$G^* = \max_{\mathbf{u}(t)} \left\{ \int_0^T g(t, \mathbf{x}, \mathbf{u}) dt + f_N(\mathbf{x}(T)) \right\}, \quad (4.109)$$

wobei $f_N(\mathbf{x}_{N+1})$ die Optimalfunktion von S_{II} in Abhängigkeit vom Eingangszustand in dieses Subsystem darstellt und $\mathbf{x}_{N+1} = \mathbf{x}(T)$ berücksichtigt wird. Zunächst muß also $f_N(\mathbf{x}_{N+1})$ für alle $\mathbf{x}_{N+1} \in X_{N+1}$ in üblicher Weise durch Rückwärtsrechnung mittels der dynamischen Optimierung errechnet werden. Nur in einfachen Fällen wird es gelingen, diese Funktion in geschlossener Form zu erhalten, normalerweise fällt aber $f_N(\mathbf{x}_{N+1})$ als Tabelle an. Dann ist es, um anschließend die Optimierung (4.109) mit dem Maximumprinzip vornehmen zu können, erforderlich, die Tabellenwerte durch einen analytischen Ausdruck zu approximieren (z. B. durch ein Polynom). Ist dies geschehen, so kann man die zu (4.109) gehörenden Bedingungen des Maximumprinzips aufstellen. Der Zustandsvektor $\mathbf{x}(t)$ besitze s Koordinaten: $x^{(1)}(t), \dots, x^{(s)}(t)$. In üblicher Weise führen wir eine zusätzliche Koordinate ein

$$x^{(s+1)}(t) = \int_0^t g(t, \mathbf{x}, \mathbf{u}) dt + f_N(\mathbf{x}(t)). \quad (4.110)$$

Als zugehörige Differentialgleichung erhalten wir

$$\frac{dx^{(s+1)}}{dt} = g(t, \mathbf{x}, \mathbf{u}) + \sum_{i=1}^s \frac{\partial f_N}{\partial x^{(i)}} \cdot \frac{dx^{(i)}}{dt}$$

und weiter mit (4.104)

$$\frac{dx^{(s+1)}}{dt} = g(t, \mathbf{x}, \mathbf{u}) + \text{grad}_x f_N \cdot \mathbf{T}(t, \mathbf{x}, \mathbf{u}), \quad (4.111)$$

$$x^{(s+1)}(0) = \alpha^{(s+1)} = 0. \quad (4.112)$$

Dadurch bekommt die Zielfunktion die neue Gestalt

$$G = x^{(s+1)}(T) = \max! \quad (4.113)$$

Auf das durch die Formeln (4.104, 4.106, 4.107, 4.111–4.113) beschriebene Problem der optimalen Steuerung kann nun das Maximumprinzip in üblicher Form angewendet

werden, worauf hier nicht näher eingegangen zu werden braucht. Als Lösung ergeben sich die Funktionen

$$\mathbf{u}^*(t), \quad \mathbf{x}^*(t), \quad 0 \leq t \leq T.$$

Insbesondere erhält man auch

$$\mathbf{x}^*(T) = \mathbf{x}_{N+1}^*. \quad (4.114)$$

Von (4.114) ausgehend kann man schließlich durch abwechselnde Verwendung der Modellgleichungen (4.103) und der in der Einbettungstechnik erhaltenen Funktionen $\mathbf{u}_n(\mathbf{x}_{n+1})$ die Optimalwerte für das Subsystem S_1 zusammenstellen.

Die Bearbeitung des in Abb. 4.14b dargestellten Systems, bei dem jetzt die stetige Stufe nachgeschaltet ist, verläuft ganz analog. Die Problemstellung wird jetzt durch die Formeln (4.103–4.106, 4.108) und

$$\mathbf{x}_{N+1} = \beta \text{ (fest vorgegeben); } \quad \mathbf{x}_1 = \mathbf{x}(0) \quad (4.107')$$

gegeben. Anstelle von (4.109) liefert das Optimalprinzip die Rekursionsbeziehung

$$G^* = \max_{\mathbf{u}_1, \dots, \mathbf{u}_N} \left\{ \sum_{n=1}^N G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f(\mathbf{T}_1(\mathbf{x}_2, \mathbf{u}_1)) \right\}, \quad (4.115)$$

wobei

$$f(\mathbf{x}(0)) = \max_{\mathbf{u}(t) \in U} \int_0^T g(t, \mathbf{x}, \mathbf{u}) \, dt \quad (4.116)$$

die Optimalfunktion der stetigen Stufe ist und $\mathbf{x}(0) = \mathbf{x}_1 = \mathbf{T}_1(\mathbf{x}_2, \mathbf{u}_1)$ eingesetzt wurde. Zur Berechnung von (4.116) wird das Maximumprinzip benutzt. Mit

$$\begin{aligned} \mathbf{x}^{(s+1)}(t) &= \int_0^t g(t, \mathbf{x}, \mathbf{u}) \, dt, \\ \frac{d\mathbf{x}^{(s+1)}}{dt} &= g(t, \mathbf{x}, \mathbf{u}), \quad \mathbf{x}^{(s+1)}(0) = 0 \end{aligned} \quad (4.117)$$

wird

$$f(\mathbf{x}(0)) = \max_{\mathbf{u}(t) \in U} \{ \mathbf{x}^{(s+1)}(T) \}. \quad (4.118)$$

Hiervon ausgehend werden in bekannter Weise die Bedingungen des Prinzips aufgestellt und gelöst. Um anschließend die Teiloptimierung des Subsystems S_1 vornehmen zu können, muß allerdings (4.118) für alle $\mathbf{x}(0) = \mathbf{x}_1 \in X_1$, d. h. praktisch für ein Punktgitter von $\mathbf{x}(0)$ -Werten, ermittelt werden. Ist dies geschehen, so kann man die Rekursion über die Stufen von S_1 anschließen:

$$\begin{aligned} f_n(\mathbf{x}_{n+1}) &= \max_{\mathbf{u}_n} \{ G_n(\mathbf{x}_{n+1}, \mathbf{u}_n) + f_{n-1}(\mathbf{T}_n(\mathbf{x}_{n+1}, \mathbf{u}_n)) \}, \quad n = 1, 2, \dots, N, \\ f_0 &\equiv f(\mathbf{x}(0)). \end{aligned} \quad (4.119)$$

Beispiel 4.12: Wir demonstrieren die sequentielle Verbindung der Methoden abschließend noch durch Lösung des in Abb. 4.15 dargestellten Beispiels. Zunächst wird das Minimum des zur stetigen Stufe S_{II} gehörenden Summanden der Zielfunktion bestimmt. Nach Einführen der Zusatzvariablen

$$x^{(2)} = \int_0^t (x^2 + u^2) dt; \quad \frac{dx^{(2)}}{dt} = x^2 + u^2; \quad x^{(2)}(0) = 0$$

erhalten wir als Hamiltonfunktion

$$H = pu + p^{(2)} \cdot (x^2 + u^2). \quad (4.120)$$

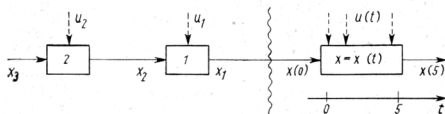


Bild 4.15 Ein Beispiel

Subsystem S_I

Subsystem S_{II}

$$x_1 = x_2 + u_1; \quad x_2 = x_3 - 2u_2, \quad \frac{dx}{dt} = u; \quad 0 \leq t \leq 5,$$

$$1 \leq x_1 \leq 3; \quad 0 \leq x_2 \leq 2; \quad x_3 = 1, \quad -2 \leq u(t) \leq +2,$$

$$-1 \leq u_1 \leq 2; \quad -1 \leq u_2 \leq 1,$$

$$G = (x_3 + u_2)^2 + (x_2^2 + u_1^2) + \int_0^5 (x^2 + u^2) dt = \min!$$

Die adjungierten Differentialgleichungen lauten

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = -2p^{(2)} \cdot x, \quad p(5) = 0,$$

$$\frac{dp^{(2)}}{dt} = -\frac{\partial H}{\partial x^{(2)}} = 0, \quad p^{(2)}(5) = 1. \quad (4.121)$$

Sofort ergibt sich

$$p^{(2)}(t) = \text{konstant} = 1. \quad (4.122)$$

Damit reduziert sich das Problem auf die Lösung der folgenden Beziehungen:

$$\begin{aligned} \dot{x} &= u, & x(0) &= x_1, \\ p &= -2x, & p(5) &= 0, \\ p \cdot u + x^2 + u^2 &= \min! \end{aligned} \quad (4.123)$$

Das Minimum der Hamiltonfunktion kann hier durch Nullsetzen der Ableitung gefunden werden:

$$\frac{\partial H}{\partial u} = p + 2u = 0 \rightarrow u = -\frac{1}{2}p. \quad (4.124)$$

Nach Einsetzen von (4.124) lassen sich die Differentialgleichungen in geschlossener Form lösen:

$$\begin{aligned}x(t) &= A \cdot [e^{t-10} + e^{-t}], \\u(t) &= A \cdot [e^{t-10} - e^{-t}], \\A &= \frac{x_1}{1 + e^{-10}}.\end{aligned}\tag{4.125}$$

Als Optimalfunktion der stetigen Stufe erhalten wir durch Integration

$$\begin{aligned}f(x_1) &= \int_0^5 (x^2 + u^2) dt = C \cdot x_1^2, \\C &= \frac{1 - e^{-10}}{1 + e^{-10}} \approx 1.\end{aligned}\tag{4.126}$$

Nun kann die Rekursion über die diskreten Stufen fortgeführt werden. Einbeziehung der Stufe 1 ergibt

$$\begin{aligned}f_1(x_2) &= \min_{u_1} \{(x_2^2 + u_1^2) + C \cdot (x_2 + u_1)^2\}, \\u_1(x_2) &= -\frac{C}{C+1} \cdot x_2, \\f_1(x_2) &= \frac{2C+1}{C+1} \cdot x_2^2.\end{aligned}\tag{4.127}$$

Schließlich wird die Stufe 2 dazugenommen

$$\begin{aligned}f_2(x_3) &= \min_{u_2} \left\{ (x_3 + u_2)^2 + \frac{2C+1}{C+1} \cdot (x_3 - 2u_2)^2 \right\}, \\u_2(x_3) &= \frac{3C+1}{9C+5} \cdot x_3, \\f_2(x_3) &= \frac{9 \cdot (2C+1)}{9C+5} \cdot x_3^2.\end{aligned}\tag{4.128}$$

Von $x_3 = 1$ ausgehend kann nun das optimale Regime zusammengestellt werden. Dies ergibt schließlich

$$\begin{aligned}G^* &= \frac{9(2C+1)}{9C+5}; \\u_2^* &= \frac{3C+1}{9C+5}; \quad x_2^* = \frac{3C+3}{9C+5}; \\u_1^* &= -\frac{3C}{9C+5}; \quad x_1^* = \frac{3}{9C+5}; \\u^*(t) &= \frac{3}{(9C+5)(1+e^{-10})} \cdot [e^{t-10} - e^{-t}]; \\x^*(t) &= \frac{3}{(9C+5)(1+e^{-10})} \cdot [e^{t-10} + e^{-t}].\end{aligned}\tag{4.129}$$

Lösungen der Aufgaben

3.1: Mit $x_0 = \int_{t_0}^t f_0(\mathbf{x}, \mathbf{u}) \, dt$ folgt

$$\dot{x}_0 = f_0(\mathbf{x}, \mathbf{u}), \quad x_0(t_0) = 0.$$

Im $(n+1)$ -dimensionalen Zustandsraum der Variablen (x_0, x_1, \dots, x_n) ist dann $x_0(t_1)$ zu minimieren.

3.2: Mit $x_0 = F[\mathbf{x}(t)] + \int_{t_0}^t f_0(\mathbf{x}, \mathbf{u}) \, dt$ folgt

$$\dot{x}_0 = \sum_{i=1}^n \frac{\partial F}{\partial x_i} f_i(\mathbf{x}, \mathbf{u}) + f_0(\mathbf{x}, \mathbf{u}) = F_0(\mathbf{x}, \mathbf{u}).$$

Im n -dimensionalen Zustandsraum der Variablen (x_1, \dots, x_n) ist dann

$$\int_{t_0}^{t_1} F_0(\mathbf{x}, \mathbf{u}) \, dt$$

zu minimieren.

3.3: $\mathcal{H} = p_1 x_2 + (p_2 - 1)u, \quad u = 6 \operatorname{sgn}(p_2 - 1), \quad p_2 = -\sqrt[5]{\frac{6}{5}}(t-1),$

$$u = \begin{cases} 6 & \text{für } 0 \leq t \leq 1 - \sqrt[5]{\frac{5}{6}}, \\ -6 & \text{für } 1 - \sqrt[5]{\frac{5}{6}} < t \leq 1. \end{cases}$$

3.4: a) $\mathcal{H} = -x_1 + (2u + t)p_1, \quad u = \operatorname{sgn} p_1 = \operatorname{sgn}(t-1) = -1,$

b) $\mathcal{H} = -x_1 + tp_1 + (2p_1 - 1)u, \quad u = \operatorname{sgn}(2p_1 - 1) = \operatorname{sgn}(2t - 3) = -1,$

c) $\mathcal{H} = -x_1 + tp_1 + p_1^2 - (u - p_1)^2, \quad u = p_1 = t - 1,$

d) $H = (2u + t)p_1, \quad u = \operatorname{sgn} p_1 = -1.$

3.5: $Q_{\text{opt}} = \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial^2 F}{\partial Q_i \partial Q_j} Q_{jt} Q_{it} + \frac{\partial F}{\partial Q_i} f_i \right),$

$$Q_0(0, t) = F[\mathbf{g}(t)], \quad Q_0(x, 0) = F[\mathbf{h}(x)],$$

Opt.-Ziel: $Q_0(a, b).$

3.6: Es ist

$$Q_t^0 = \frac{1}{4}(1-t)(2e^x + e^{1-x} - e^{1+x} - 2)$$

$$Q_{xt}^0 = \frac{1}{4}(1-t)(2e^x - e^{1-x} - e^{1+x})$$

und damit $Q_{xt}^0 = Q_t^0 + u^0$ erfüllt. Weiter ist

$$Q_0^0(1, 1) = \int_0^1 \int_0^1 (Q^0 + u^{02}) \, dt \, dx = \frac{1}{24} (4e - 5 - e^2).$$

3.7: Es ist

$$\frac{\partial^2 J}{\partial u_{\alpha\beta} \partial u_{\gamma\delta}} = \begin{cases} 2nl & \text{für } \alpha = \gamma, \beta = \delta, \\ 0 & \text{sonst.} \end{cases}$$

Die entsprechende Matrix der Ableitungen 2. Ordnung ist somit positiv definit.

Unter Benutzung von

$$lK = 1, \quad l \leq l(K - \beta + 1) \leq 1,$$

$$1 \leq M - \alpha + 1 \leq M, \quad -\frac{1}{2} \leq g < 0,$$

folgt sofort $\frac{31}{8} \leq u_{\alpha\beta} \leq \frac{3}{4}$.

Mit $l\beta = t$, $n\alpha = x$ ergibt sich

$$\lim_{l \rightarrow 0} l(K - \beta + 1) = 1 - t, \quad \lim_{n \rightarrow 0} g^{M - \alpha + 1} = e^{1-t}$$

und damit

$$\lim_{\substack{n \rightarrow 0 \\ l \rightarrow 0}} u_{\alpha\beta} = \frac{1}{2}(1 - t)(1 - e^{1-t}).$$

Literatur

- [1] *Aganbegjan, A. G.* u. a.: Sistema modelej narodnochozjajstvennogo planirovanja. Izd. Mysl', Moskva 1972.
- [2] *Athans, M.; Falb, P.*: Optimal'noe upravlenie. Izd. Maschinostrojenie, Moskva 1968.
- [3] Autorenkollektiv: Operationsforschung, Mathematische Grundlagen, Methoden und Modelle, Bd. 3. Berlin: VEB Deutscher Verlag der Wissenschaften 1972.
- [4] *Bojarinov, A. I.; Kafarov, V. V.*: Metody optimizacii v chimičeskoj tehnologii. Izd. Chimija, Moskva 1969.
- [5] *Boltjanski, W. G.*, u. a.: K teorii optimal'nych prozessov. DAN SSSR, Bd. 110, Nr. 1 (1956), S. 7–10.
- [6] *Boltjanski, W. G.*: Mathematische Methoden der optimalen Steuerung (Übers. a. d. Russ.). Leipzig: Akademische Verlagsgesellschaft Geest & Portig K.-G. 1971.
- [7] *Boltjanski, W. G.*: Optimale Steuerung diskreter Systeme (Übers. a. d. Russ.). Leipzig: Akademische Verlagsgesellschaft Geest & Portig K.-G. 1976.
- [8] *Butkovskij, A. G.*: Teorija optimal'nogo upravlenija sistemami s raspredeľennymi parametrami. Izd. Nauka, Moskva 1965.
- [9] *Fan, L. T.; Wang, C. S.*: Diskretnyj princip maksimuma. Izd. Mir, Moskva 1967.
- [10] *Fan, L. T.*: The continuous maximum principle. John Wiley & Sons, Inc., New York-London-Sydney 1964.
- [11] *Fedorenko, N. P.*, u. a. (Red.): Mathematik und Kybernetik in der Ökonomie; Nachschlagewerk. Berlin: Verlag Die Wirtschaft 1973.
- [12] *Feldbaum, A. A.*: Osnovy teorii optimal'nych avtomatičeskich system. Izd. Nauka, Moskva 1966.
- [13] *Jasinski, H.; Wagner, P.*: Zur Optimierung eines chemischen Rohrreaktors bei Polymerisationsreaktionen. Wiss. Zeitschrift der TH Leuna-Merseburg **13** (1971) H. 2, S. 179–183.
- [14] *Lee, E. B.; Markus, L.*: Osnovy teorii optimal'nogo upravlenija. Izd. Nauka, Moskva 1972.
- [15] *Leitmann, G.*: Vvedenie v teoriju optimal'nogo upravlenija. Izd. Nauka, Moskva 1968.
- [16] *Michlin, S. M.*: Numerische Realisierung von Variationsmethoden (Übers. a. d. Russ.). Berlin: Akademie-Verlag 1969.
- [17] *Moisseev, N. N.*: Čislennye metody v teorii optimal'nych sistem. Izd. Nauka, Moskva 1971.
- [18] *Nemhauser, G. L.*: Einführung in die Praxis der dynamischen Programmierung. München-Wien: R. Oldenbourg Verlag 1969.
- [19] *Piehler, J.*: Einführung in die dynamische Optimierung. Leipzig: B. G. Teubner Verlagsgesellschaft 1966.
- [20] *Pontrjagin, L. S.*, u. a.: Mathematische Theorie optimaler Prozesse (Übers. a. d. Russ.) Berlin: Akademie-Verlag 1964.
- [21] *Smirnov, A. D.*: Modelirovanie i prognozirovanie socialističeskogo vosproizvodstva. Izd. Ekonomika, Moskva 1970.
- [22] *Smirnow, W. I.*: Lehrgang der höheren Mathematik, Teil IV (Übers. a. d. Russ.). 6. Aufl. Berlin: VEB Deutscher Verlag der Wissenschaften 1973.
- [23] *Zeidler, G.*: Das Pontrjaginsche Maximumprinzip und einige seiner Anwendungsmöglichkeiten in der Ökonomie. Wiss. Zeitschr. d. Hochsch. f. Ökon., Berlin, **4** (1969), S. 573–586.

Sachregister

- adjungiertes System 15
- Algorithmus der dynamischen Optimierung 59, 60
- Anwendung, kombinierte 98
- autonomes System 26
- Bellman, Optimalprinzip von 57, 59, 92
- Bewegungsgleichung 10, 51
- Bolza, Aufgabe von 25
- Dekomposition 68
 - , verallgemeinerte 69
- deterministischer Prozeß 6
- diskontinuierlicher Prozeß 6
- diskreter Prozeß 6
- diskretes Maximumprinzip 55
- Diskretisierung des stetigen Problems 93
- dynamische Optimierung 57, 96, 98
 - –, Algorithmus 59, 60
- Einbettungstechnik 59
- Endzeit, feste 29
- Erweiterung des Optimalprinzips 90
- Extremwert, globaler 98
- fast-optimale Steuerung 36
- feste Endzeit 29
- Funktion, Monotonie einer 69
 - , separable 69
- glatte Mannigfaltigkeit 20
- globaler Extremwert 98
- Gradientenverfahren 37
 - , projiziertes 37
- Grundaufgabe der optimalen Steuerung 13, 58
- Hamiltonfunktion 16, 42
- Hamiltonsystem 16
 - , reduziertes 19
- hinreichendes Kriterium 55
- Ketten, Vereinigung zweier 77
- kombinierte Anwendung 98
- kontinuierlicher Prozeß 6
- Kriterium für Zeitoptimalität, notwendiges 20
 - , hinreichendes 55
- Lagrange, Aufgabe von 24
- Lösung, suboptimale 89
- Mannigfaltigkeit, glatte 20
- Maximumprinzip 14, 16, 42, 96, 98
- Maximumprinzip, diskretes 55
- Mayer, Aufgabe von 24
- mehrstufige Modelle 57
- Methode von Newton-Raphson 38
- Minimalfolge 36
- Modelle, mehrstufige 57
- Monotonie einer Funktion 69
- Newton-Raphson-Methode 38
- nichtautonomes System 26
- notwendige Optimalitätsbedingung 16
- Optimale Lösung, Zusammenstellung 60
 - Prozesse 52
 - – mit Parametern 36
 - – – Verzögerung 35
 - Steuerung, Grundaufgabe 58
 - Trajektorie 13
- optimaler Prozeß 8
 - Steuervektor 13
- Optimalität 8
- Optimalitätsbedingung, notwendige 16
- Optimalitätskriterium 8
- Optimalprinzip, Erweiterung 90
 - von Bellman 57, 59, 92
- Optimierung, dynamische 57, 96, 98
- Parameter, verteilte 39, 40
- Phasenkoordinaten 7
- Phasenraum 7
- Pontrjaginsches Maximumprinzip 14, 16
- Problem mit beweglichem Endpunkt 20
 - – minimalem Energieverbrauch 30
 - – unbeschränktem Zeitintervall 35
- , stetiges 91
- projiziertes Gradientenverfahren 37
- Prozeß 6, 8
 - , deterministischer 6
 - , diskontinuierlicher 6
 - , diskreter 6
 - , kontinuierlicher 6
 - mit fixierter Zeit und freiem rechten Ende 19
 - , optimaler 8, 52
 - , schnelligkeitsoptimaler 14
 - , stetiger 6, 9
 - , steuerbarer 7
 - , – deterministischer 7
 - , stochastischer 7
- Prozeßgleichung 10, 51
- Punktgitter 62

Randwertaufgabe, Zwei-Punkte- 38
reduziertes Hamiltonsystem 19
Reihenschaltung 64
Rückkopplungsschleife 85

Schleife, vorwärts gerichtete 82
schnelligkeitsoptimaler Prozeß 14
separable Funktion 69
stetiger Prozeß 6
stetiges Problem 91
— —, Diskretisierung 93
steuerbarer Prozeß 7
Steuerbereich 7, 39, 51
Steuerparameter 7
Steuerraum 7
Steuerung 7
—, fast-optimale 36
—, k.-beste 89
—, suboptimale 89
—, zulässige 39, 51
Steuervektor 39
—, optimaler 13
—, zulässiger 11
stochastischer Prozeß 7
Stufenfolge, verzweigende 75
suboptimale Lösung 89
suboptimale Steuerung 89

Subsystem 98
System 6
—, adjungiertes 15
—, autonomes 26
—, mit verteilten Parametern 39, 40
—, nichtautonomes 26

Tabelle der Stufe n 62
Trajektorie 7
Transversalitätsbedingung 21

unbeschränktes Zeitintervall 35

verallgemeinerte Dekomposition 69
Vereinigung zweier Ketten 77
verzweigende Stufenfolge 75
vorwärts gerichtete Schleife 82
Vorwärtsrekursion 66

Zeitoptimalität 19, 20
zulässige Steuerung 39, 51
zulässiger Steuervektor 11
Zusammenstellen der optimalen Lösung 60
Zustandsinversion 66
Zustandsvektor 39
Zwei-Punkte-Randwertaufgabe 38