

INHALT

	<i>Seite</i>
BIRKLE, M. Angewandte Forschung in Europa - Chancen neuer Technologien	1
VAN DE LOGT, G.; HOFFMANN, K. Numerische Planung von Trajektorien für Roboter	3
WEINMANN, A. Iterative Approximation of Arbitrary Functions in Time and Frequency Domain by Basis Functions Subject to Constraints	14
GE, J.-H.; FRANK, P.M. H _∞ Control via Measurement Feedback for Jump Linear Systems	21
STEINKOGLER, A.; ODENBREIT, M. Genetische Algorithmen erstellen Faustformeln zur Reglereinstellung	29
SCHRUTEK, C.; DEMATTIO, R. Verknüpfung von Echtzeitsystemen mit nichtechtzeitkritischen Verwaltungssystemen	41
LEICHTFRIED, J.; HEISS, M. Fuzzy-Regler als glättender Regel-Interpolator	47
GSCHWENDTNER, G. Integration der automatisierten Demontage im Konstruktions- und Planungsprozess	62
SCHLAGLICHTER	73
DISSERTATIONSKURZFASSUNGEN	77
BUCHBESPRECHUNGEN	83
EIN INSTITUT STELLT SICH VOR Abteilung für Regelungsmathematik, Hybridrechen- und Simulationstechnik am Institut 114 der TU Wien	84
VERANSTALTUNGSKALENDER	88

Angewandte Forschung in Europa - Chancen neuer Technologien

Prof.Dr. M. Birkle

Fraunhofer-Institut Informations- und Datenverarbeitung (IIB), Karlsruhe

In den modernen Industriegesellschaften kommt der angewandten Forschung die entscheidende Transferfunktion für Ergebnisse aus der Grundlagenforschung in die Industrielle Anwendung zu. Unter angewandter Forschung versteht man dabei die Innovationsleistung, die erbracht werden muß, um Ergebnisse aus der Grundlagenforschung für die industrielle Anwendung nutzbar zu machen. Die angewandte Forschung ist damit ein wichtiges Kriterium für die Innovationskraft einer Industriegesellschaft und zumindest kurzfristig für die Wettbewerbsfähigkeit einer Volkswirtschaft als die Grundlagenforschung. In Europa ist die angewandte Forschung dezentral organisiert, mit im wesentlichen indirekter staatlicher Einflußnahme über die Forschungsförderung.

In Europa wird die angewandte Forschung von einer ganzen Reihe von Institutionen geleistet, die einen mehr oder weniger großen Anteil ihrer Haushalte aus Forschungsaufträgen erlösen müssen und sich dadurch von den voll staatlich finanzierten Grundlagenforschungseinrichtungen unterscheiden. Der Zwang, zur Finanzierung der Haushalte Forschungsaufträge einwerben zu müssen, hat sich dabei als wirksames Mittel der Erfolgskontrolle erwiesen.

Als Vertreter der Grundlagenforschung werden gemeinhin die Universitäten, die Großforschungseinrichtungen und Wissenschaftsorganisationen, wie in Deutschland die Max-Planck-Gesellschaft, genannt. Dabei muß aber beachtet werden, daß diese Organisationen in der Regel auch mehr oder weniger ausgeprägt angewandte Forschung betreiben. Ein großer Teil der Drittmittelforschung der Universitäten ist z.B. dieser angewandten Forschung zuzurechnen. Insofern ist die Abgrenzung der angewandten Forschung von der Grundlagenforschung schwierig. Die Übergänge sind fließend. Dies gilt insbesondere auch für den Bereich der militärischen Forschung und der Weltraumforschung. Auf der anderen Seite ist die industrielle Forschung natürlich nicht auf die Entwicklung neuer Produkte beschränkt. Insbesondere die Forschungsabteilungen der Großindustrie bemühen sich natürlich auch darum, Ergebnisse aus der Grundlagenforschung aufzugreifen und in industrielle Anwendungen umzusetzen. Insofern sind Angaben zu den Aufwendungen für angewandte Forschung in Europa mit großen Unsicherheiten und Interpretationsspielräumen versehen.

Die Größen der Institutionen der angewandten Forschung in den einzelnen Mitgliedsländern der Europäischen Gemeinschaft sind dabei recht unterschiedlich. Die vier größten Institutionen sind die Fraunhofer-Gesellschaft in der Bundesrepublik mit etwa 1 Milliarde DM, die TNO in den Niederlanden mit ca. 0,6 Milliarden DM, die VTT in Finnland mit ca. 0,35 Milliarden DM und die SINTEF in Norwegen mit ca. 0,3 Milliarden DM Haushaltsvolumen pro Jahr.

Die verschiedenen Institutionen der angewandten Forschung sind in unterschiedlichen Fachbereichen tätig. Die angewandte Forschung in Europa beschränkt sich jedoch mit ganz wenigen Ausnahmen auf die Ingenieur- und Naturwissenschaften.

Bedingt durch die Nachfrage liegen die Schwerpunkte dabei auf Verfahrenstechnik, Produktionstechnik, Mikroelektronik, Informatik, Materialforschung, Logistik, Meßtechnik, Umwelttechnik und Medizintechnik. Aktivitäten im Grenzbereich zu den Sozialwissenschaften, wie z.B. Arbeitsorganisation, Technologiefolgenabschätzung, Untersuchungen von wirtschaftlichen Trends runden das Angebot ab.

Trotz unterschiedlicher Industriestrukturen und unterschiedlicher administrativer Randbedingungen in den einzelnen Mitgliedsländern der EG ist die strukturelle Gliederung der angewandten Forschung weitgehend gleich. Die Leistungen der angewandten Forschungen werden in der Regel in Instituten oder institutsähnlichen Organisationseinheiten mit im Mittel 30 bis 200 Akademikern erbracht, die fachlich methodisch orientiert sind. Diese Institute agieren am Markt für angewandte Forschung und decken dabei jeweils einen ganz bestimmten Fachbereich ab. Die großen Institutionen vereinigen mehrere dieser Institute unter einem organisatorischen Dach.

Mit Ausnahme der TNO in den Niederlanden, die in der Vergangenheit schon international tätig war, sind die Institutionen der angewandten Forschung im wesentlichen national orientiert, d.h. der überwiegende Anteil des Geschäftsvolumens wird mit nationalen Partnern (Kunden) abgewickelt. Diese nationale Ausrichtung beginnt sich aufzulösen. Bedingt durch die Förderung multinationaler Projekte innerhalb der Europäischen Gemeinschaft durch die EG-Kommission sind grenzüberschreitende Kontakte und Kundenbeziehungen entstanden. Dazu kommt, daß insbesondere die großen Institutionen der angewandten Forschung sich der Globalisierungstendenz ihrer Hauptkunden anschließen und Aktivitäten im nordamerikanischen und asiatisch-pazifischen Markt aufzubauen beginnen.

Diese sich abzeichnende Tendenz der Internationalisierung wird zu Veränderungen auf der Seite der Anbieter angewandter Forschungsleistungen führen. Zum einen kann die Internationalisierung finanziell nicht von den bestehenden Instituten oder institutsähnlichen Organisationseinheiten getragen werden, zum anderen ist die Angebotsbreite einzelner Institute für internationale Projekte in der Regel zu gering. Die Internationalisierung muß daher von größeren Organisationseinheiten, wie sie in einigen Ländern bereits bestehen, getragen werden und sie muß in die nationalstaatliche Forschungs- und Förderpolitik eingebunden werden.

Des weiteren wird man dann auf eine Internationalisierung der Institutionen der angewandten Forschung verzichten müssen, wenn man nicht akzeptieren will, daß mit nationalen Steuergeldern geförderte Transferleistungen auch ausländischen Konkurrenten zugute kommen. Für die rasche Umsetzung von Ergebnissen der Grundlagenforschung in Anwendungen und industrielle Produkte wäre die Internationalisierung der angewandten Forschung jedoch ein großer Gewinn.

Numerische Planung von Trajektorien für Roboter

G. van de Logt, K. Hoffmann

Lehrstuhl für Systemtheorie der Elektrotechnik

Universität des Saarlandes, Saarbrücken, Deutschland

Eingelangt am 8. Mai 1995

Zusammenfassung

Die vorliegende Arbeit beschreibt einen neuen Ansatz zur Planung von Trajektorien für Roboter. Basierend auf einem allgemeinen Verfahren für die Rückwärtsrechnung werden sechs numerische Verfahren für die Planung von Trajektorien entwickelt. Bei der numerischen Rückwärtsrechnung wird der Abstand zwischen zwei Raumpunkten, einem Ziel- und einem aktuellen Suchpunkt, durch ein numerisches Optimierungsverfahren minimiert, so daß Ziel- und Suchpunkt am Ende der Optimierung übereinstimmen. Alle während der Optimierung bestimmten Suchpunkte werden bei der Berechnung der Trajektorien als Stützstellen verwendet. Die numerischen Verfahren werden in Form von Simulationen am Beispiel des Roboters Puma 560 untersucht. Um die Trajektorien nicht nur grafisch sondern auch bezüglich der Bewegungszeit miteinander vergleichen zu können, werden die zugehörigen Bahnen mit einer unsynchronisierten Punkt-zu-Punkt-Bewegung berechnet.

1 Einleitung

Der steigende Einsatz von Robotern in der Automatisierung von Produktionsabläufen in der Industrie hat zu einem verstärkten Einsatz von Robotersimulationssystemen in der Planung und der Projektierung geführt. Um die Bewegungen von Robotern zu simulieren, sind neben der Beschreibung des kinematischen Aufbaus von Robotern auch die Koordinatentransformationen, Vorwärts- und Rückwärtsrechnung, wichtig. Die Vorwärtsrechnung bestimmt die Position und die Orientierung der Roboterhand aus den Achsvariablen des Roboters. Die Rückwärtsrechnung bestimmt die Achsvariablen des Roboters aus der Position und der Orientierung der Roboterhand. Während für die Vorwärtsrechnung ein

eindeutiges, analytisches Verfahren verwendet wird, wird in der vorliegenden Arbeit für die Rückwärtsrechnung eine neues, numerisches Verfahren vorgeschlagen. Dabei wird eine Fehlerfunktion, die den Abstand zwischen zwei Raumpunkten, einem Ziel- und einem aktuellen Suchpunkt, berechnet, von einem numerischen Optimierungsverfahren minimiert, so daß am Ende der Optimierung Ziel- und Suchpunkt übereinstimmen.

In den numerischen Verfahren zur Planung von Trajektorien werden alle während einer numerischen Rückwärtsrechnung erzeugten Suchpunkte als Stützstellen für die Berechnung der Trajektorien benutzt. Es wird insbesondere der Einfluß der maximalen Achsvariablenänderung, einer getrennten Positionierung und Orientierung und einer veränderten Fehlerfunktion auf die Trajektorien untersucht. Für die Berechnung der Bahnen zu den Trajektorien wird eine unsynchronisierte Punkt-zu-Punkt-Bewegung (PTP-Bewegung) verwendet.

2 Numerische Rückwärtsrechnung

Die in dieser Arbeit verwendete numerische Rückwärtsrechnung führt die Rückwärtsrechnung auf eine Vorwärtsrechnung zurück. Für die Vorwärtsrechnung wird ein allgemeines, analytisches Verfahren, das auf den Arbeiten von Denavit und Hartenberg (Denavit, J. und Hartenberg, R.S., 1955) und von Paul (Paul, R.P., 1981) beruht, benutzt. Die numerische Rückwärtsrechnung geht von zwei Raumpunkten, einem Zielpunkt und einem aktuellen Suchpunkt, aus. Das Ziel der numerischen Rückwärtsrechnung ist die Bestimmung der Achsvariablen für einen vorgegebenen Raumpunkt, hier für den Zielpunkt. Das für die numerische Rückwärtsrechnung verwendete numerische Optimierungsverfahren ist kein selbststartendes Verfahren. Deshalb ist die Vorgabe eines Anfangspunktes für die Optimierung, hier in Form eines Startachsvariablensatzes, notwendig. Für die Berechnung des Abstandes (Fehlers) zwischen Ziel- und Suchpunkt wird folgende Fehlerfunktion verwendet:

$$F = \sum_{i=1}^3 (t_{z,i} - t_{s,i})^2 + \sum_{j=1}^3 (w_{z,j} - w_{s,j})^2 \quad (1)$$

Der Vektor \underline{t}_z gibt die Verschiebungen des Zielpunktes, der Vektor \underline{t}_s die Verschiebungen des Suchpunktes in den einzelnen Koordinatenrichtungen bezüglich eines Referenzkoordinatensystems an. Die Vektoren \underline{w}_z und \underline{w}_s geben die Drehungen von Ziel- und Suchpunkt bezüglich des Referenzkoordinatensystems an. Für die Drehungswinkel werden die RPY-Winkel benutzt. Der Fehler zwischen Ziel- und Suchpunkt hat demnach zwei Anteile: einen Positions- und einen Orientierungsfehler.

Bei der numerischen Rückwärtsrechnung wird ausgehend von einem Startachsvariablensatz mit der Vorwärtsrechnung der Startsuchpunkt bestimmt (Iteration 1). Mit Gleichung (1) wird der Fehler zwischen Ziel- und Startsuchpunkt berechnet. Ausgehend von diesem aktuellen Fehler und dem alten Achsvariablensatz erzeugt ein numerisches Optimierungsverfahren einen neuen Achsvariablensatz. Für diesen Achsvariablensatz wird mit der Vorwärtsrechnung wieder ein Suchpunkt bestimmt (Iteration 2). Dieser aktuelle Suchpunkt wird dann wieder durch Gleichung (1) mit dem Zielpunkt verglichen, usw. . Bei allen nachfolgenden Iterationen des numerischen Optimierungsverfahrens wird jeweils ein Suchpunkt mit einem geringeren Abstand zum Zielpunkt erzeugt. Die Optimierung wird abgebrochen, wenn der Fehler zwischen dem Ziel- und dem aktuellen Suchpunkt kleiner als eine vorgegebene Schranke ist, also Ziel- und aktueller Suchpunkt übereinstimmen. Der letzte Achsvariablensatz ist dann die Lösung der Rückwärtsrechnung für den Zielpunkt (van de Logt, G., 1993).

3 Numerische Planung von Trajektorien

Die in den folgenden Abschnitten erläuterten Verfahren zur numerischen Planung von Trajektorien arbeiten alle nach dem gleichen Prinzip. Die während einer numerischen Rückwärtsrechnung erzeugten Suchpunkte werden als Stützstellen bei der Berechnung der Trajektorien verwendet. Eine einfache Möglichkeit zur Berechnung der Trajektorien, ist die direkte Verbindung aufeinanderfolgender Stützstellen.

Bei der Anwendung der in Abschnitt 2 erläuterten numerischen Rückwärtsrechnung wird der Startsuchpunkt als Anfangs- und der Zielpunkt als Endpunkt der Trajektorie benutzt.

3.1 Basisverfahren

Beim Basisverfahren V1 wird eine numerische Rückwärtsrechnung durchgeführt. Alle bei den einzelnen Iterationen der numerischen Rückwärtsrechnung erzeugten Suchpunkte werden als Stützstellen für die Trajektorien benutzt.

3.2 Verfahren mit Begrenzung der maximalen Achsvariablenänderung

Wenn ein Roboter von einem Startpunkt zu einem Zielpunkt fährt, müssen im allgemeinen alle Achsen bewegt werden. Ist die Änderung einer Achsvariablen zu groß, werden von der Robotersteuerung Zwischenpunkte erzeugt, um die Bahn zwischen den Raumpunkten zu

berechnen. Beim Verfahren V2 mit Begrenzung der maximalen Achsvariablenänderung wird deshalb die Achsvariablenänderung begrenzt, um zu vermeiden, daß die Robotersteuerung neue Zwischenpunkte berechnet.

Die maximale Achsvariablenänderung wird durch Vorgabe einer festen Schranke Δq begrenzt. Jede Achsvariable eines neu erzeugten Achsvariablensatzes wird mit der zugehörigen Achsvariable des zuvor erzeugten Achsvariablensatzes verglichen. Ist für eine Achsvariable die Änderung größer als Δq , wird die alte Achsvariable um Δq erhöht und als neue Achsvariable weiterverwendet. Mit allen neuen Achsvariablen wird dann der neue Suchpunkt bestimmt.

3.3 Verfahren mit getrennter Positionierung und Orientierung

Bei den zuvor erläuterten Verfahren werden alle Achsvariablen gleichzeitig geändert. Die Positionierung und die Orientierung der Roboterhand werden gleichzeitig durchgeführt. Beim Verfahren V3 mit getrennter Positionierung und Orientierung wird die Roboterhand erst positioniert und dann orientiert. Um die Roboterhand grob zu positionieren, werden bei einer ersten numerischen Rückwärtsrechnung nur die Achsvariablen der Hauptachsen (q_1, q_2, q_3) verändert. In der nachfolgenden numerischen Rückwärtsrechnung werden nun wieder alle Achsvariablen verändert, um die Roboterhand fein zu positionieren und zu orientieren.

Dieses Verfahren zur numerischen Planung von Trajektorien wird auch in Verbindung mit dem Verfahren mit Begrenzung der maximalen Achsvariablenänderung verwendet (Verfahren V4). Die prinzipielle Arbeitsweise bleibt bis auf die Berücksichtigung der maximalen Achsvariablenänderung Δq gleich.

3.4 Verfahren mit veränderter Fehlerfunktion

Die in Abschnitt 2 beschriebene Fehlerfunktion nach Gleichung (1) ist die Summe aus einem Positions- und einem Orientierungsfehler. Beim Verfahren V5 mit veränderter Fehlerfunktion wird bei einer ersten numerischen Rückwärtsrechnung nur der Positionsfehler berücksichtigt. Bei der nachfolgenden numerischen Rückwärtsrechnung wird dann wieder die Summe aus Positions- und Orientierungsfehler benutzt, um den Fehler zwischen dem Ziel- und dem aktuellen Suchpunkt zu berechnen.

Wie das Verfahren mit getrennter Positionierung und Orientierung wird auch das Verfahren mit veränderter Fehlerfunktion in Verbindung mit dem Verfahren mit Begrenzung der maximalen Achsvariablenänderung verwendet (Verfahren V6). Auch hier ändert sich an der prinzipiellen Arbeitsweise bis auf die Berücksichtigung der maximalen Achsvara-

blenänderung Δq nichts.

4 Roboter Puma 560

Für die Untersuchungen der in Abschnitt 3 erläuterten Verfahren zur numerischen Planung von Trajektorien wird der sechssachsige Roboter Puma 560, der in Abb. 1 dargestellt ist, verwendet. Die DH-Parameter und Arbeitsbereiche des Roboters Puma 560 finden sich

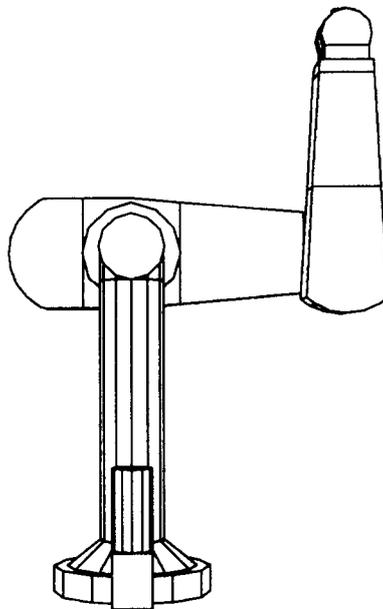


Abbildung 1: Roboter Puma 560

in Tab. 1.

Für die Untersuchungen wird das Davidon-Fletcher-Powell-Verfahren (DFP-Verfahren) (Spellucci, P., 1993) aus der Gruppe der verbesserten Gradientenverfahren, den Quasi-Newton-Verfahren, benutzt. Da dieses Optimierungsverfahren während der Optimierung eindimensionale Optimierungsverfahren verwendet, wird dafür das Verfahren des goldenen Schnittes (Papageorgiou, M., 1991) eingesetzt.

Als Zielpunkt wird ein Raumpunkt mit folgenden Daten verwendet:

- Position: Verschiebung in x -Richtung um $t_1 = 400$ mm, Verschiebung in y -Richtung um $t_2 = 400$ mm, Verschiebung in z -Richtung um $t_3 = 400$ mm,

i	Θ_i	d_i	a_i	α_i	$B_{min,i}$	$B_{max,i}$
1	0	0	0	-90°	-160°	160°
2	0	149 mm	432 mm	0	-225°	45°
3	0	0	0	90°	-45°	225°
4	0	432 mm	0	-90°	-110°	170°
5	0	0	0	90°	-100°	100°
6	0	56 mm	0	0	-266°	266°

Tabelle 1: DH-Parameter und Arbeitsbereiche des Roboters Puma 560

- Orientierung: Drehung um den RPY-Winkel $w_1 = 42.80^\circ$, Drehung um den RPY-Winkel $w_2 = 28.02^\circ$, Drehung um den RPY-Winkel $w_3 = 11.17^\circ$

Als Startpunkt wird die Nullstellung des Roboters Puma 560 ($q_1 = \dots = q_6 = 0^\circ$) benutzt. In den folgenden Abbildungen ist der Startpunkt mit einem x gekennzeichnet. In Abb. 2 findet sich der Verlauf der Trajektorien aller Verfahren zur numerischen Planung von Trajektorien bei Verwendung des DFP-Verfahrens.

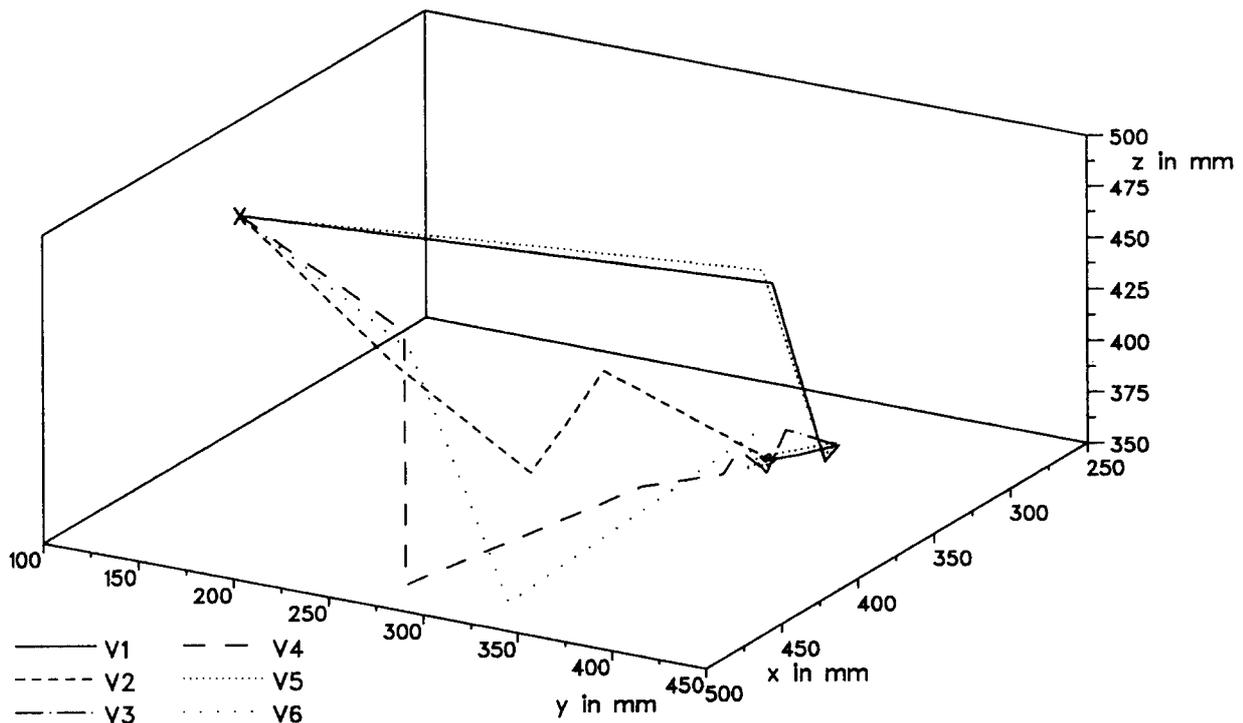


Abbildung 2: Verlauf der Trajektorien

Beim Verfahren V2 wird die maximale Achsvariablenänderung auf $\Delta q = 10^\circ$ gesetzt. Der Verlauf der Trajektorien in Abb. 2 zeigt, daß alle Verfahren in der Lage sind, den

Zielpunkt vom gewählten Startpunkt aus zu erreichen. Die Verfahren V2, V4, V5 und V6 erzeugen unterschiedliche Trajektorien zwischen Start- und Zielpunkt. Die weitgehende Übereinstimmung der Trajektorien der Verfahren V1 und V3 ist im Aufbau der Fehlerfunktion nach Gleichung (1) als Summe aus Positions- und Orientierungsfehler begründet. Wegen des großen Anteils des Positionsfehlers am Gesamtfehler zu Beginn der numerischen Rückwärtsrechnung werden auch bei Verfahren V1 erst die Achsvariablen der Hauptachsen (q_1, q_2, q_3) verändert, weil sie fast allein für die Positionierung der Roboterhand verantwortlich sind. Am Ende der numerischen Rückwärtsrechnung, wenn Positions- und Orientierungsfehler in der gleichen Größenordnung liegen, erzeugen beide Verfahren dann unterschiedliche Stützstellen (Abb. 2).

Um die Trajektorien nicht nur grafisch sondern auch bezüglich der Bewegungszeit t_B miteinander vergleichen zu können, werden ausgehend von den Stützstellen der Trajektorien die zugehörigen Bahnen berechnet. Die Bahn eines Roboters ist der von der Roboterhand zurückgelegte Weg zwischen Start- und Zielpunkt. Für die Berechnung der Bahn wird eine unsynchronisierte PTP-Bewegung verwendet, bei der sich alle Achsen gleichzeitig mit maximaler Geschwindigkeit und mit maximaler Beschleunigung bewegen. Da die einzelnen Verfahren in der Nähe des Zielpunktes viele, eng beieinander liegende Stützstellen erzeugen, ist die Zahl dieser Stützstellen zu reduzieren, um eine zeitoptimale Bahn zu berechnen. Die Achsvariablen der einzelnen Stützstellen werden miteinander verglichen. Eine Stützstelle wird nur dann beibehalten, wenn für mindestens eine Achse die Achsvariablenänderung über einer vorgegebenen Schranke liegt. Liegen alle Achsvariablenänderungen unter der vorgegebenen Schranke, wird die Stützstelle für die Berechnung der Bahn nicht berücksichtigt. Den Verlauf der Bahnen zeigt Abb. 3.

In Tab. 2 finden sich die Bewegungszeiten des Roboters für die mit den einzelnen Verfahren zur numerischen Planung von Trajektorien (VznPT) erzeugten Bahnen zwischen Start- und Zielpunkt. Die Ergebnisse in Tab. 2 zeigen, daß Verfahren V2 mit einer maximalen Achsvariablenänderung $\Delta q = 10^\circ$ die kleinste Bewegungszeit liefert. Im folgenden soll nun der Einfluß der maximalen Achsvariablenänderung auf den Verlauf der Trajektorien und Bahnen untersucht werden. Für die maximale Achsvariablenänderung werden fünf verschiedene Werte gewählt: $\Delta q = 2^\circ$, $\Delta q = 4^\circ$, $\Delta q = 6^\circ$, $\Delta q = 8^\circ$ und $\Delta q = 10^\circ$. In Abb. 4 finden sich die Verläufe der Trajektorien für die unterschiedlichen maximalen Achsvariablenänderungen.

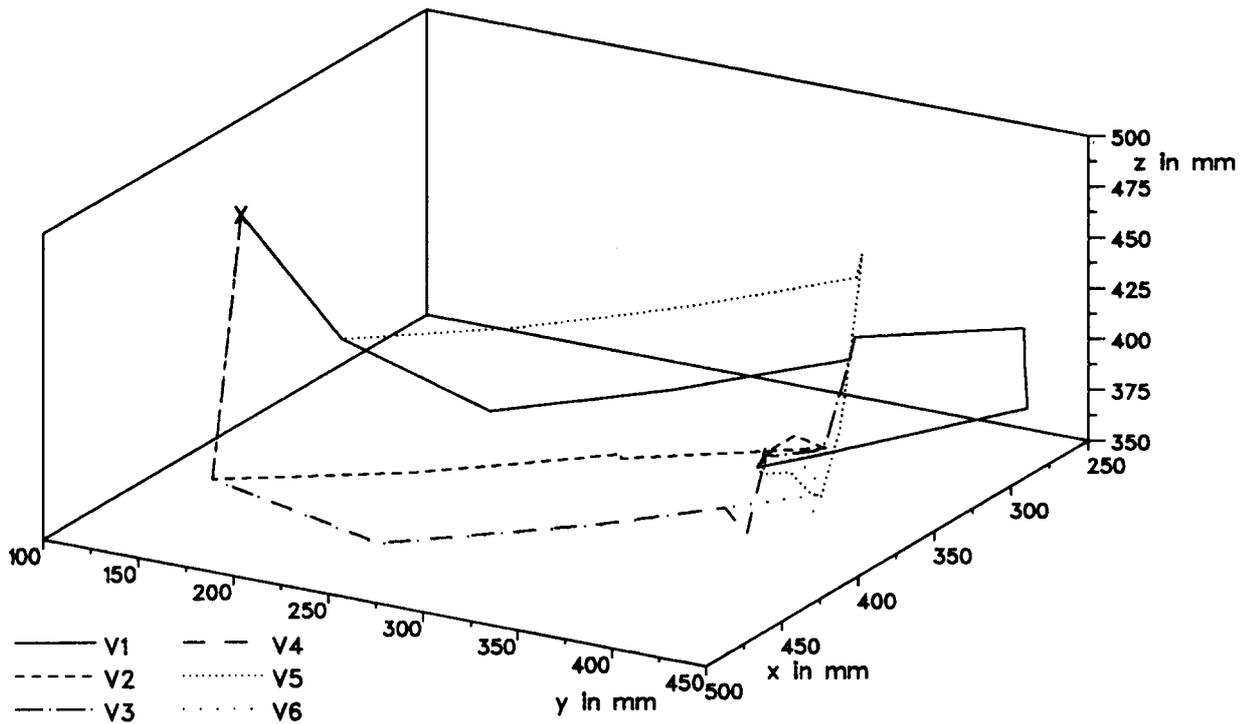


Abbildung 3: Verlauf der Bahnen

VznTP	t_B
V1	12.9 s
V2	11.0 s
V3	14.2 s
V4	13.0 s
V5	14.4 s
V6	11.4 s

Tabelle 2: Bewegungszeiten des Roboters für die Bahnen zwischen Start- und Zielpunkt

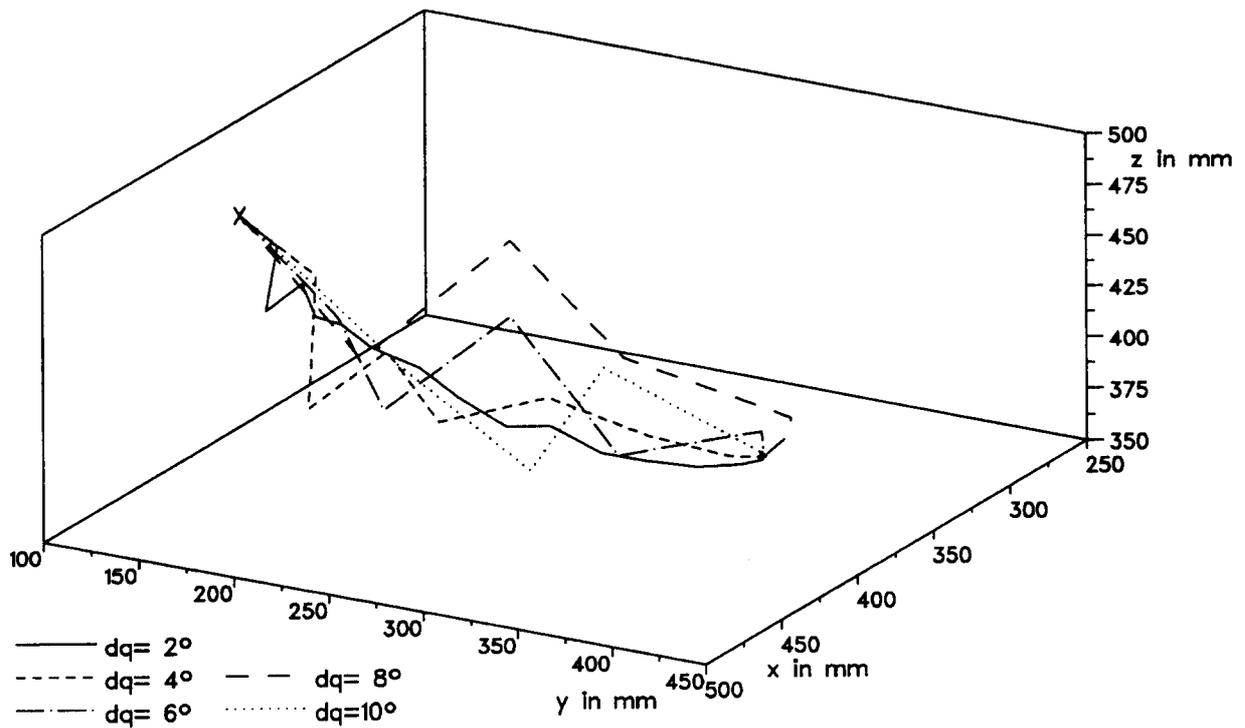


Abbildung 4: Verlauf der Trajektorien bei unterschiedlichen maximalen Achsvariablenänderungen

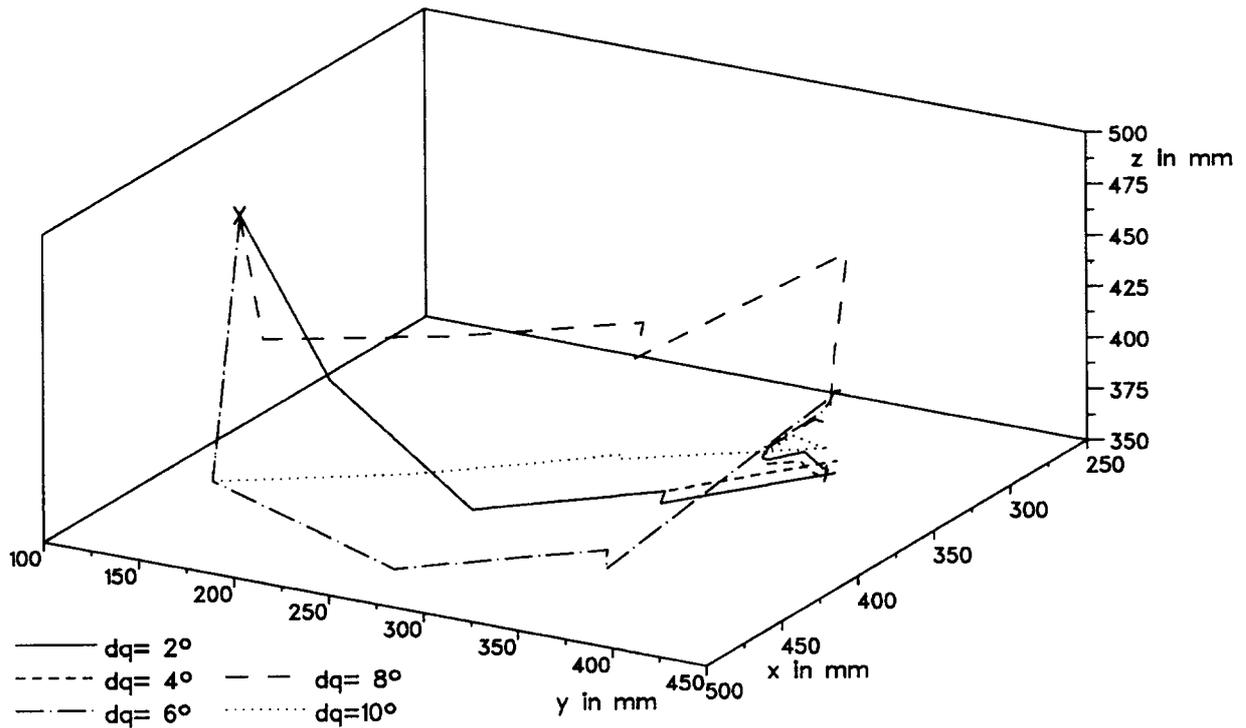


Abbildung 5: Verlauf der Bahnen bei unterschiedlichen maximalen Achsvariablenänderungen

Aus dem Verlauf der Trajektorien ist zu sehen, daß je kleiner die maximale Achsvariablenänderung ist, desto direkter ist die Verbindung zwischen Start- und Zielpunkt. Eine direktere Verbindung zwischen Start- und Zielpunkt muß jedoch nicht unbedingt zu einer Verkürzung der Bewegungszeit führen. Um diesen Sachverhalt genauer zu untersuchen, werden wieder die Bahnen mit einer unsynchronisierten PTP-Bewegung berechnet. Vor der Berechnung der Bahnen wird die Anzahl der Stützstellen reduziert. Den Verlauf der Bahnen zeigt Abb. 5. In Tab. 3 finden sich die Bewegungszeiten des Roboters für die Bahnen zwischen Start- und Zielpunkt bei unterschiedlichen maximalen Achsvariablenänderungen. Die Ergebnisse in Tab. 3 zeigen, daß ein direkterer Verlauf der Trajektorien nicht unbedingt zu einer kürzeren Bewegungszeit der zugehörigen Bahnen führen muß.

Δq	t_B
2°	13.4 s
4°	10.4 s
6°	16.8 s
8°	18.0 s
10°	11.0 s

Tabelle 3: Bewegungszeiten des Roboters für die Bahnen zwischen Start- und Zielpunkt bei unterschiedlichen maximalen Achsvariablenänderungen

5 Zusammenfassung und Ausblick

Die in dieser Arbeit entwickelten Verfahren zur numerischen Planung von Trajektorien basieren alle auf dem Verfahren zur numerischen Rückwärtsrechnung. Sie unterscheiden sich vor allem in der Durchführung der numerischen Rückwärtsrechnung. Mit einer Begrenzung der maximalen Achsvariablenänderung, einer getrennt durchgeführten Positionierung und Orientierung und einer veränderten Fehlerfunktion werden unterschiedliche Vorgehensweisen bei der numerischen Planung der Trajektorien zur Verfügung gestellt. Die Ergebnisse am Beispiel des Roboters Puma 560 zeigen, daß alle Verfahren in der Lage sind, Trajektorien zwischen einem Start- und einem Zielpunkt zu erzeugen.

Um die Trajektorien nicht nur grafisch miteinander vergleichen zu können, werden mit der unsynchronisierten PTP-Bewegung zugehörige Bahnen berechnet, deren Bewegungszeiten dann zur Beurteilung der einzelnen Verfahren zur numerischen Planung von

Trajektorien herangezogen werden.

Die Eignung weiterer numerischer Optimierungsverfahren für die Verfahren zur numerischen Planung von Trajektorien wird in (Hoffmann, K., 1995) ausführlich untersucht.

Literatur

- Denavit, J., Hartenberg, R.S., 1955, A Kinematic Notation for Lower Pair Mechanisms based on Matrices, Journal of Applied Mechanics 22 vol. 2, pp. 215–221*
- Hoffmann, K., 1995, Entwicklung eines numerischen Verfahrens zur Planung von Trajektorien für Roboter, Diplomarbeit, Systemtheorie der Elektrotechnik, Universität des Saarlandes*
- van de Logt, G., 1993, Entwicklung eines numerischen Verfahrens zur inversen kinematischen Transformation für Industrieroboter, Diplomarbeit, Systemtheorie der Elektrotechnik, Universität des Saarlandes*
- Papageorgiou, M., 1991, Optimierung, (R. Oldenburg Verlag München Wien)*
- Paul, R.P., 1981, Robot Manipulators: Mathematics, Programming and Control, (MIT-Press Cambridge)*
- Spellucci, P., 1993, Numerische Verfahren der nichtlinearen Optimierung, (Birkhäuser Verlag Basel Boston Berlin)*

Iterative Approximation of Arbitrary Functions in Time and Frequency Domain by Basis Functions Subject to Constraints

Alexander Weinmann, ÖVE, Senior Member IEEE

Manuscript received June 8, 1995

Abstract

The problem of optimally reducing the least squares residual by a one-shot method is considered from an unusual perspective. This viewpoint is characterized by selecting those model functions which are obtained by maximizing the gain in residual and keeping parameters fixed to those values which are achieved in a previous L_2 norm minimization step. The optimal model fitting is also studied in the case of inequality constraint models.

1 Introduction

Identification algorithms frequently are based on linear regression (ordinary least squares). A parameter vector is selected such that a given set of functions weighted by the parameters approximates a given process output function optimally. For the sake of recalling basic concepts of least-squares methods, *Björk, Å, 1990*, and *Fiacco, A.V., 1983*, are cited. Actuator constraints as used in the present paper are also mentioned in Internal Model Control implementations by *Morari, M., and Zafriou, E., 1989*.

In a recent paper (*Weinmann, A., 1995*) an unusual idea was imparted, executing a supplementary variation of the model functions in order to reduce the residual. A matricial gradient is utilized in order to reduce the residual to a maximum extent. This method enables feedforward control and identification to be carried out when using modifications of functions or functions subject to constraints. The change of the model functions is based on frozen parameters as achieved by the previous regression.

Revisiting the basic concept and result of the matricial gradient method, the standard least squares problem (L_2 optimal approximation)

$$C_2 \triangleq \int_0^T [f(t) - \mathbf{g}^T(t)\mathbf{p}]^2 w(t) dt \rightarrow \min_{\mathbf{p}} \quad (1)$$

is considered where $f(t)$ is the function to be approximated, $\mathbf{p} \in \mathcal{R}^{n_p}$ is a parameter vector, $\mathbf{g}(t)$ is the vector of model functions $g_i(t) \forall i = 1, 2, \dots, n_p$, i.e., $\mathbf{g}(t) = \text{vec} [g_i(t)] \in \mathcal{R}^{n_p}$. The

problem can also be reformulated as an approximation in frequency domain. For simplicity but without loss in generality, we choose the weighting function $w(t) \equiv 1$. By defining

$$\mathbf{N} \triangleq \int_0^T \mathbf{g}(t)\mathbf{g}^T(t) dt \in \mathcal{R}^{n_p \times n_p} \quad \mathbf{y} \triangleq \int_0^T f(t)\mathbf{g}(t) dt \in \mathcal{R}^{n_p}, \quad (2)$$

and by $\partial C_2 / \partial \mathbf{p} = \mathbf{0}$ the optimum parameter is

$$\mathbf{p}^* = \mathbf{N}^{-1}\mathbf{y}. \quad (3)$$

Combining Eq.(1) and (3) yields the residual $\min_{\mathbf{p}} C_2 \triangleq C_2^*$

$$C_2^* = \int_0^T [f(t) - \mathbf{g}^T(t)\mathbf{p}^*]^2 dt = f_o - \mathbf{y}^T \mathbf{N}^{-1}\mathbf{y} \quad \text{where} \quad f_o \triangleq \int_0^T f(t)^2 dt. \quad (4)$$

By discretizing with respect to time, the interval T is divided into n_m equal parts, i.e., $\Delta T = T/n_m$, and the column-like matrix \mathbf{M} is defined by

$$\mathbf{M} = \text{matrix}[m_{ki}] \triangleq (\mathbf{g}_1 : \mathbf{g}_2 : \dots : \mathbf{g}_i : \dots : \mathbf{g}_{n_p}) \in \mathcal{R}^{n_m \times n_p} \quad \text{and} \quad m_{ki} = g_i(t_k) = g_{ik}. \quad (5)$$

Then

$$\mathbf{N} \doteq \mathbf{M}^T \mathbf{M} \Delta T \in \mathcal{R}^{n_p \times n_p} \quad \mathbf{y} \doteq \mathbf{M}^T \mathbf{f} \Delta T \in \mathcal{R}^{n_p} \quad \mathbf{p}^* \doteq \mathbf{N}^{-1}\mathbf{y} \in \mathcal{R}^{n_p}. \quad (6)$$

Using

$$\mathbf{F} \triangleq \mathbf{f} \mathbf{f}^T \in \mathcal{R}^{n_m \times n_m} \quad (7)$$

and applying the trace, from Eq.(4),

$$C_2^* = \text{tr}\{\mathbf{F}[\mathbf{I}_{n_m} - \mathbf{M}(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T]\} \Delta T. \quad (8)$$

Finally, from Eq.(8) the matricial gradient is

$$\frac{\partial C_2^*}{\partial \mathbf{M}} = 2[\mathbf{M}\mathbf{N}^{-1}\mathbf{y} - \mathbf{f}]\mathbf{y}^T \mathbf{N}^{-1} \Delta T = 2[\mathbf{M}(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T - \mathbf{I}_{n_m}]\mathbf{F}\mathbf{M}(\mathbf{M}^T \mathbf{M})^{-1} \Delta T. \quad (9)$$

For optimal curve fitting, the increment $\Delta \mathbf{M}$ can be set as

$$\Delta \mathbf{M} = -\kappa \frac{\partial C_2^*}{\partial \mathbf{M}} \quad (10)$$

where κ is a constant factor. In order to obtain optimal matricial gradient step, by using the definition

$$\mathbf{U} \triangleq \mathbf{M}(\mathbf{M}^T \mathbf{M})^{-2} \mathbf{M}^T, \quad (11)$$

the optimum result in κ is (Weinmann, A., 1995)

$$\kappa^* = \frac{\text{tr}\{[\mathbf{I}_{n_m} - \mathbf{M}(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T]\mathbf{F}\mathbf{U}\mathbf{F}\}}{2 \text{tr}\{[\mathbf{I}_{n_m} - \mathbf{M}(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T]\mathbf{F}\mathbf{U}\mathbf{F}\mathbf{U}\mathbf{F}\} \Delta T}. \quad (12)$$

Moreover, by employing the fact that \mathbf{F} of Eq.(7) is a dyadic product, one has

$$\text{tr}\{\mathbf{X}\mathbf{F}\} = \text{tr}\{\mathbf{X}\mathbf{f}\mathbf{f}^T\} = \text{tr}\{\mathbf{f}^T \mathbf{X}\mathbf{f}\} \quad (13)$$

for any \mathbf{X} , and Eq.(12) can be rewritten to the short version

$$\kappa^* = \frac{1}{2\mathbf{f}^T \mathbf{U}\mathbf{f} \Delta T} = \frac{1}{2\text{tr}\{\mathbf{U}\mathbf{F}\} \Delta T}. \quad (14)$$

2 Minimum Norm of the Model Matrix Increment

Now, the result of Eqs.(9) and (12) is derived based on a very different consideration. Using the model

$$\mathbf{M}\mathbf{p} \doteq \mathbf{f} , \quad (15)$$

the well-known ordinary least squares result \mathbf{p}^* resulting from the given model matrix \mathbf{M} and measuring vector \mathbf{f} is

$$\mathbf{p}^* = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{f} . \quad (16)$$

Let us consider an increment $\Delta\mathbf{M}$, which satisfies

$$(\mathbf{M} + \Delta\mathbf{M})\mathbf{p}^* = \mathbf{f} \quad (17)$$

for keeping \mathbf{p}^* unchanged. Then

$$\Delta\mathbf{M} \mathbf{p}^* = \mathbf{f} - \mathbf{M}\mathbf{p}^* . \quad (18)$$

Demanding Frobenius norm $\|\Delta\mathbf{M}\|_F \rightarrow \min$ and valuating that this aim is imposed on the model condition Eq.(17), one has

$$\|\Delta\mathbf{M}\|_F^2 + \lambda^T(\Delta\mathbf{M} \mathbf{p}^* - \mathbf{f} + \mathbf{M}\mathbf{p}^*) \rightarrow \min_{\Delta\mathbf{M}} . \quad (19)$$

From $\partial/\partial\Delta\mathbf{M}$

$$2\Delta\mathbf{M} + \lambda\mathbf{p}^{*T} = \mathbf{0} \rightsquigarrow \Delta\mathbf{M} = -0.5 \lambda\mathbf{p}^{*T} , \quad (20)$$

and then from the model condition, the Lagrange multiplier λ

$$(\mathbf{M} + \Delta\mathbf{M})\mathbf{p}^* - \mathbf{f} = \mathbf{M}\mathbf{p}^* - 0.5 \lambda\mathbf{p}^{*T}\mathbf{p}^* - \mathbf{f} = \mathbf{0} \quad (21)$$

$$\lambda = -\frac{2}{\mathbf{p}^{*T}\mathbf{p}^*}(\mathbf{f} - \mathbf{M}\mathbf{p}^*) \quad (22)$$

results. Eventually, one gets the optimum model increment

$$\Delta\mathbf{M} = -0.5 \lambda\mathbf{p}^{*T} = \frac{1}{\mathbf{p}^{*T}\mathbf{p}^*}(\mathbf{f} - \mathbf{M}\mathbf{p}^*)\mathbf{p}^{*T} . \quad (23)$$

For keeping \mathbf{p}^* unchanged, this solution is unique.

Using the definition (identity matrix minus hat matrix)

$$\mathbf{R} \triangleq \mathbf{I}_{n_m} - \mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T \quad (24)$$

and Eq.(16), an alternative formula is

$$\Delta\mathbf{M} = \frac{\mathbf{R}\mathbf{F}\mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1}}{\text{tr}\{\mathbf{U}\mathbf{F}\}} . \quad (25)$$

Since the product $\kappa^* \frac{\partial C_2^*}{\partial \mathbf{M}}$ based on Eqs.(9), (10) and (14) is equal to Eq.(25), this result points out a useful identity: The one-step-ahead matricial gradient can be considered not only as a first-order approximation of the residual to be minimized with respect to \mathbf{M} but also as a precise result for $\Delta\mathbf{M}$ causing minimum residual for fixed original \mathbf{p}^* .

3 Additional Conditions Imposed on the Model Matrix Increment

Consider the case of Eq.(17) with the single additional restriction $M_{ki} + \Delta M_{ki} - V_{ki} = 0$ for some given k, i . Then, by applying an additional Lagrange multiplier μ_{ki}

$$\|\Delta \mathbf{M}\|_F^2 + \boldsymbol{\lambda}^T (\Delta \mathbf{M} \mathbf{p}^* - \mathbf{f} + \mathbf{M} \mathbf{p}^*) + \mu_{ki} \times (M_{ki} + \Delta M_{ki} - V_{ki}) \rightarrow \min_{\Delta \mathbf{M}, \boldsymbol{\lambda}, \mu_{ki}} . \quad (26)$$

By differentiating with respect to $\Delta \mathbf{M}$,

$$2\Delta \mathbf{M} + \boldsymbol{\lambda} \mathbf{p}^{*T} + \mu_{ki} \frac{\partial \Delta M_{ki}}{\partial \Delta \mathbf{M}} = 0 \quad \rightsquigarrow \quad \Delta \mathbf{M} = -0.5 \boldsymbol{\lambda} \mathbf{p}^{*T} - 0.5 \mu_{ki} \frac{\partial \Delta M_{ki}}{\partial \Delta \mathbf{M}} . \quad (27)$$

The derivative of ΔM_{ki} with respect to $\Delta \mathbf{M}$ is obtained as a matrix $\Theta_{k,i} \in \mathcal{R}^{n_m \times n_p}$ which contains zeros as its entries except the (i, j) -position unity¹

$$\Theta_{k,i} \triangleq \frac{\partial \Delta M_{ki}}{\partial \Delta \mathbf{M}} = \text{matrix}_{\alpha, \beta} \left[\frac{\partial \Delta M_{ki}}{\partial \Delta M_{\alpha\beta}} \right] , \quad (28)$$

presupposing that there are no interdependencies in the elements of $\Delta \mathbf{M}$. From Eqs.(27), (28) and (17)

$$\mathbf{M} \mathbf{p}^* + \Delta \mathbf{M} \mathbf{p}^* = \mathbf{M} \mathbf{p}^* - 0.5 \boldsymbol{\lambda} \mathbf{p}^{*T} \mathbf{p}^* - 0.5 \mu_{ki} \Theta_{k,i} \mathbf{p}^* = \mathbf{f} \quad (29)$$

$$\boldsymbol{\lambda} = \frac{1}{\mathbf{p}^{*T} \mathbf{p}^*} (2\mathbf{f} - 2\mathbf{M} \mathbf{p}^* + \mu_{ki} \Theta_{k,i} \mathbf{p}^*) . \quad (30)$$

Substituting Eq.(30) into Eq.(27), one gets

$$\Delta \mathbf{M} = \frac{1}{\mathbf{p}^{*T} \mathbf{p}^*} \left[\mathbf{f} \mathbf{p}^{*T} - \mathbf{M} \mathbf{p}^* \mathbf{p}^{*T} + 0.5 \mu_{ki} \Theta_{k,i} (\mathbf{p}^* \mathbf{p}^{*T} - \mathbf{p}^{*T} \mathbf{p}^* \mathbf{I}_{n_p}) \right] . \quad (31)$$

Equating $\Delta M_{ki} = V_{ki} - M_{ki}$, i.e., imposing the conditions on ΔM_{ki} , one finds

$$V_{ki} - M_{ki} = \frac{1}{\mathbf{p}^{*T} \mathbf{p}^*} \left[(\mathbf{f} \mathbf{p}^{*T} - \mathbf{M} \mathbf{p}^* \mathbf{p}^{*T})_{ki} + 0.5 \mu_{ki} (\mathbf{p}^* \mathbf{p}^{*T} - \mathbf{p}^{*T} \mathbf{p}^* \mathbf{I}_{n_p})_{ii} \right] . \quad (32)$$

Using the definitions

$$a_{ki} \triangleq (\mathbf{f} \mathbf{p}^{*T} - \mathbf{M} \mathbf{p}^* \mathbf{p}^{*T})_{ki} \quad (33)$$

$$b_i \triangleq 0.5 (\mathbf{p}^* \mathbf{p}^{*T} - \mathbf{p}^{*T} \mathbf{p}^* \mathbf{I}_{n_p})_{ii} , \quad (34)$$

from Eq.(32)

$$\mu_{ki} = \frac{(V_{ki} - M_{ki}) \mathbf{p}^{*T} \mathbf{p}^* - a_{ki}}{b_i} \quad (35)$$

¹Note that $\Theta_{k,i}$ is a zero matrix with the exception 1 at the (k, i) -position but V_{ki} and ΔM_{ki} are entries of \mathbf{V} and $\Delta \mathbf{M}$ at the (k, i) -position. In general, $(\)_{ki}$ denotes the (k, i) -position of a matrix.

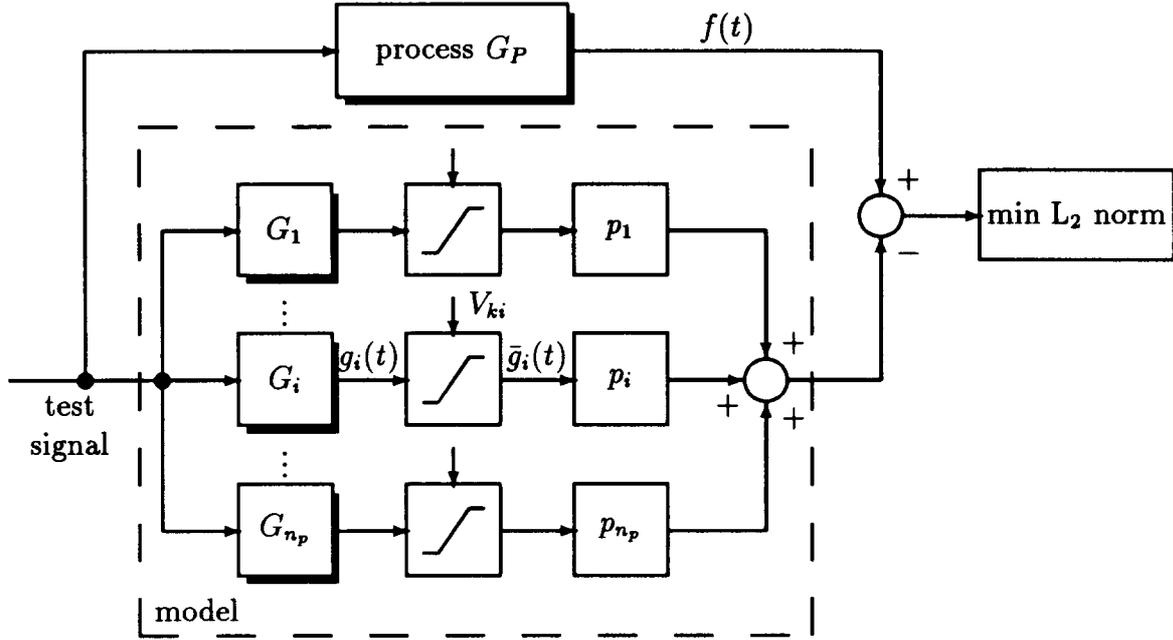


Figure 1: Identification based on constrained functions

results. Combining with Eq.(31) and evaluating yields

$$\Delta \mathbf{M} = \frac{1}{\mathbf{p}^{*T} \mathbf{p}^*} \left[(\mathbf{f} - \mathbf{M} \mathbf{p}^*) \mathbf{p}^{*T} + [(V_{ki} - M_{ki}) \mathbf{p}^{*T} \mathbf{p}^* - a_{ki}] \Theta_{ki} \right]. \quad (36)$$

Since \mathbf{p}^* is fixed the result agrees with the result Eq.(23) except the (k, i) -position.

Now, consider n_r constraints. Eq.(26) is extended to

$$\|\Delta \mathbf{M}\|_F^2 + \lambda^T (\Delta \mathbf{M} \mathbf{p}^* - \mathbf{f} + \mathbf{M} \mathbf{p}^*) + \sum_{k,i} \mu_{ki} \times (M_{ki} + \Delta M_{ki} - V_{ki}) \rightarrow \min_{\Delta \mathbf{M}, \lambda, \mu_{ki}}. \quad (37)$$

and Eq.(31) turns out as

$$\Delta \mathbf{M} = \frac{1}{\mathbf{p}^{*T} \mathbf{p}^*} \left[\mathbf{f} \mathbf{p}^{*T} - \mathbf{M} \mathbf{p}^* \mathbf{p}^{*T} + 0.5 \Theta_{\mu} (\mathbf{p}^* \mathbf{p}^{*T} - \mathbf{p}^{*T} \mathbf{p}^* \mathbf{I}_{n_p}) \right] \quad (38)$$

where Θ_{μ} is a matrix with μ_{ki} on the (k, i) -positions constrained to $V_{ki} - M_{ki}$ and with otherwise zeros. Hence,

$$\frac{1}{\mathbf{p}^{*T} \mathbf{p}^*} \left[(\mathbf{f} \mathbf{p}^{*T} - \mathbf{M} \mathbf{p}^* \mathbf{p}^{*T})_{ki} + 0.5 \sum_{\nu} \mu_{k\nu} (\mathbf{p}^* \mathbf{p}^{*T} - \mathbf{p}^{*T} \mathbf{p}^* \mathbf{I}_{n_p})_{\nu i} \right] = V_{ki} - M_{ki} \quad \forall k, i. \quad (39)$$

This is a system of n_r linear equations in μ_{ki} . The sum covers only those k, i which are constrained to the equality condition, i.e., where the inequality condition is not fulfilled. Since $\mathbf{p}^* \mathbf{p}^{*T} - \mathbf{p}^{*T} \mathbf{p}^* \mathbf{I}_{n_p}$ is singular for $n_p \geq 3$, the limitation must not concern more than two functions i at the same instant k .

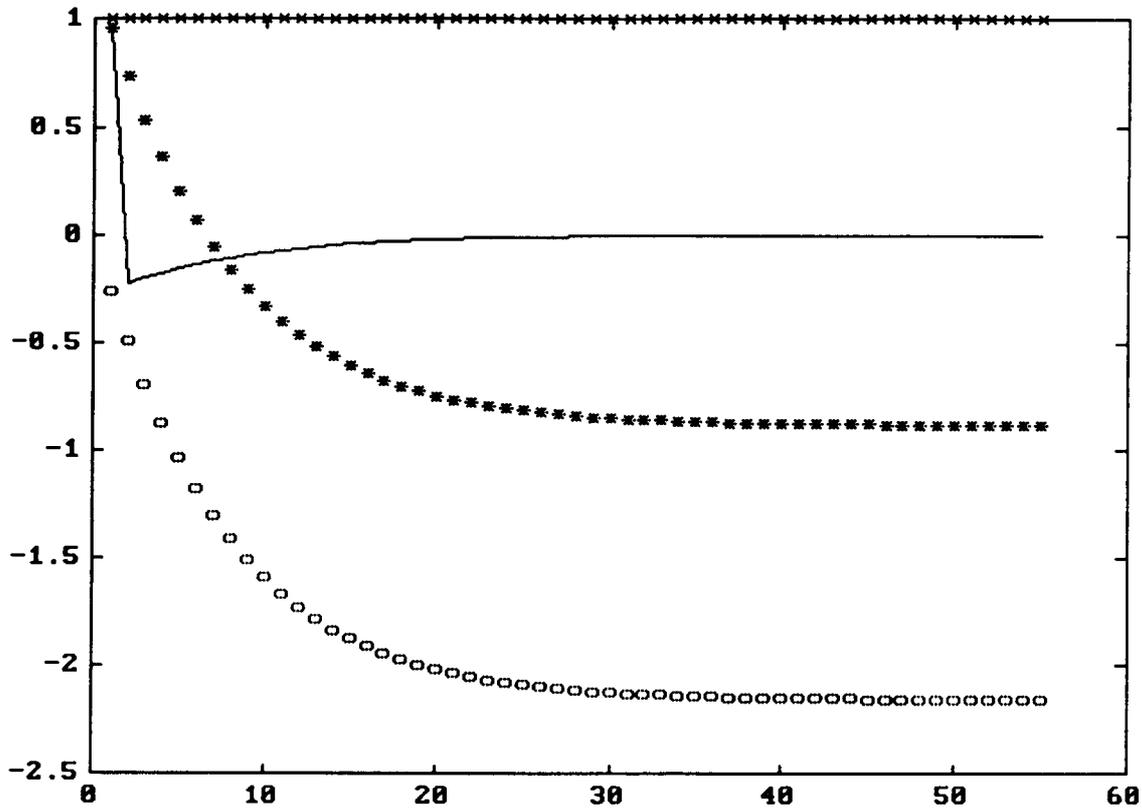


Figure 2: Plot of $(\Delta M)_{31}$ (solid), $\sum_r(\Delta M)_{31}$ (\star), $\sum_r(\Delta M)_{41}$ (o) and $\sum_r(\Delta M)_{21}$ (x) versus step number

4 Application

Consider the identification problem as depicted in Fig. 1.

A test signal input accomplishes the excitation of the real process G_P and of the model in parallel. The output of G_P is $f(t)$ in time domain. The model is a parallel connection of systems G_i . The output signals $g_i(t)$ are subject to constraints $v_i(t_k)$. Referring to Eq.(5), let $v_i(t_k) = V_{ki}$. The outputs $\bar{g}_i(t)$ are weighted by p_i . Eq.(36) is based on \mathbf{p}^* resulting from a previous calculation of Eq.(6). For fixed \mathbf{p}^* , the improvement in g_i is obtained. Owing to changes in g_i in consequence of the constraint, the calculation has to be repeated. Instead of using the matricial gradient, the optimum g_i including the constraint is computed directly.

The computation has to be repeated for \mathbf{p}^* and ΔM alternately, using Eq.(16) and Eq.(36) [or Eq.(39)], respectively. The resulting iteration is continued until a specified stopping condition is satisfied.

In Fig. 1 the components g_i are limited individually. The parallel structure of this model is closely related to the partial fraction decomposition of the transfer function. The reason for partitioning the model by functions under constraints is given by the fact that in

many practical systems the operation is also characterized by constraints in state variables and/or in the output.

Example. Constrained model fitting: Given the original data M_o and f

$$\begin{array}{l}
 n_p \triangleq 2 \\
 n_m \triangleq 10 \\
 \Delta T \triangleq 10 \\
 n_r \triangleq 1
 \end{array}
 \quad
 M_o \triangleq
 \begin{pmatrix}
 10 & 100 \\
 20 & 400 \\
 30 & 900 \\
 40 & 1600 \\
 50 & 2500 \\
 60 & 3600 \\
 70 & 4900 \\
 80 & 6400 \\
 90 & 8100 \\
 100 & 10000
 \end{pmatrix}
 \quad
 f \triangleq
 \begin{pmatrix}
 21 \\
 34 \\
 49 \\
 66 \\
 85 \\
 106 \\
 129 \\
 154 \\
 181 \\
 210
 \end{pmatrix},
 \quad (40)$$

a constrained model fitting is carried out with constraint $M_{21} \leq 21$. The procedure of model adapting is accomplished in 55 steps of iteration. The results are plotted versus step number of iteration in Fig. 2.

5 Conclusion

An optimum result in unconstrained curve fitting could be obtained in merely one single step of approximation. Minimizing the residual with respect to the model function while keeping the parameters of the previous regression fixed leads to the same result as accomplishing one step of the matricial gradient of the least-squares residual. For constrained curve fitting, the iteration is of the same quality of convergence as in the case of using the matricial gradient method but the calculus is simpler and offers an easier way of interpretation. In fact, an algorithm corresponding to a steepest descent method only arises in the constrained case.

References

- Björk, Å., 1990, Least Squares Methods. In: Ciarlet, P.G., and Lions, J.L., (Eds.), Handbook of Numerical Analysis, Vol. 1 (North-Holland, Amsterdam), pp. 465-647*
- Fiacco, A.V., 1989, Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, (Academic Press, New York)*
- Morari, M., and Zafriou, E., 1989, Robust Process Control (Prentice Hall, Englewood Cliffs)*
- Weinmann, A., 1995, Matricial gradient for the least-squares residual and its application to optimal model fitting, Int. J. Systems Sci. 26, pp. 147-160*

H_{∞} Control via Measurement Feedback for Jump Linear Systems

Jian-Hua Ge and P. M. Frank
Department of Measurement and Control,
University of Duisburg, Bismarckstr. 81 BB
47057 Duisburg, Germany
Fax: 00492033792928

June 1, 1995

Abstract

This paper addresses the problem of H_{∞} control via measurement feedback for continuous-time linear systems with Markov jump parameters. A sufficient condition is first developed such that the closed-loop jump linear systems are exponentially stable in the mean square and guarantee a prescribed H_{∞} norm bound of the closed-loop transfer function matrix from the disturbance to the controlled output. Further, a full order observer-based H_{∞} controller is constructed by solving two sets of coupled algebraic Riccati equations. An example is given to demonstrate the applicability of the proposed approach.

1 Introduction

The linear stochastic systems with jump parameters can be used to formulate systems that are subject to abrupt changes in their structure and components, due to failures, repairs or changing subsystem interconnections. The control of such systems has potential applications involving systems requiring fault-tolerant control (Ge and Frank, 1995; Srichander and Walker, 1993). Significant research efforts have been devoted to the analysis and synthesis for jump linear systems; see, for example, Ji and Chizeck (1992), Mariton (1990), Wonham (1971), and the references therein. Recently, motivated by the significant advances in the H_{∞} control theory (Doyle *et al.*, 1989; Francis, 1987; Stoorvogel, 1992), de Souza and Fragoso (1993) studied the problem of H_{∞} state feedback control for jump linear systems. However, the problems of H_{∞} control via measurement feedback for jump linear systems are still open.

The goal of this paper is to construct a full order observer-based H_{∞} controller for the continuous-time linear systems with jump parameters which can be described by an underlying finite-state Markov chain. A sufficient condition is proposed such that the closed-loop jump linear systems are exponentially stable in the mean square and guarantee the prescribed H_{∞} norm bound of the closed-loop transfer function matrix from the disturbance to the controlled output. Based on this derivation, an H_{∞} controller via measurement feedback is constructed by solving two sets of coupled algebraic Riccati equations, respectively.

The organization of this paper is as follows: Section 2 provides a precise problem formulation. Main results make up Section 3. In Section 4, an illustrative example gives into the results obtained. Section 5 presents some concluding remarks.

2 Problem Formulation

Fix a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and consider continuous-time linear systems with jump parameters described by

$$\begin{cases} \dot{x}(t) = A(\theta_t)x(t) + G(\theta_t)w_1(t) + B(\theta_t)u(t), \\ x(0) = 0, \theta_0 = i, \\ z(t) = \begin{pmatrix} H(\theta_t)x(t) \\ u(t) \end{pmatrix}, \\ y(t) = C(\theta_t)x(t) + w_2(t), \end{cases} \quad (1)$$

where $x(t) \in R^n$ is the state, $u(t) \in R^m$ is the control input, $y(t) \in R^l$ is the measurement, $w_1(t) \in R^p$ and $w_2(t) \in R^l$ are the disturbance inputs which belong to $L_2[0, \infty)$, and $z(t) \in R^{q+m}$ is the controlled output. $\{\theta_t, t \in [0, \infty)\}$ is a homogeneous Markov process with right continuous trajectories and taking values on the finite set $\Phi = \{1, 2, \dots, N\}$. We assume here that the values of θ_t are available at each time t , and systems (1) are stabilizable and observable. We also assume that $P_t := (P_t^1, \dots, P_t^N)$, with $P_t^i := p(\theta_t = i)$, $i = 1, \dots, N$, satisfies the Kolmogorov forward equation $dP_t/dt = \Lambda P_t$; $P_0 = \bar{P}$, $0 \leq t < \infty$, where $\Lambda := [(\lambda_{ij})]^T$ is the stationary transition rate matrix of $\{\theta_t\}$ with $\lambda_{ij} \geq 0$, $i \neq j$, and

$$\lambda_{ii} = - \sum_{\substack{j=1 \\ j \neq i}}^N \lambda_{ij}. \quad (2)$$

$A(\theta_t)$, $B(\theta_t)$, $C(\theta_t)$, $G(\theta_t)$, and $H(\theta_t)$ are appropriately dimensioned matrices. These matrices are deterministic for given θ_t value.

In order to put the H_∞ control issue in a stochastic setting, we bring to bear the space $L^2([\Omega, \mathcal{F}, \mathcal{P}], [0, \infty))$ of \mathcal{F} -measurable processes, $z(t)$, for which

$$\|z\|_2 = E \left\{ \int_0^T z^T(t)z(t)dt \right\}^{1/2} < \infty,$$

where $E\{\cdot\}$ denotes the expectation operator.

In the remainder of this paper, we use indistinctly $\|\cdot\|_2$ to denote the norm either in $L^2[0, \infty)$ or in $L^2([\Omega, \mathcal{F}, \mathcal{P}], [0, \infty))$, whenever the context makes it clear which one we are referring to.

The objective is to design an output feedback controller

$$u(t) = K(\theta_t)y(t), \quad (3)$$

such that for all non-zero $w \in L^2[0, \infty)$,

$$\|z\|_2 < \gamma \|w\|_2, \quad (4)$$

where $w = [w_1^T w_2^T]^T$, $z(t)$ is the controlled output defined by (1), and $\gamma > 0$ is a prescribed level of disturbance attenuation to be achieved.

3 Main Results

We construct the dynamic compensator as follows:

$$\begin{cases} \dot{\xi}(t) = (A(\theta_t) + B(\theta_t)K(\theta_t) + G(\theta_t)K_d(\theta_t) - L(\theta_t)C(\theta_t))\xi(t) + L(\theta_t)y(t), \\ u(t) = K(\theta_t)\xi(t), \end{cases} \quad (5)$$

where $K(\theta_t) \in R^{m \times n}$ is the feedback gain, $K_d(\theta_t) \in R^{p \times n}$ is the disturbance estimate gain, and $L(\theta_t) \in R^{n \times l}$ is the observer gain. $K(\theta_t)$, $K_d(\theta_t)$ and $L(\theta_t)$ are the design parameters to be determined.

To get a closed-loop system description, we define

$$x_e^T(t) = \begin{bmatrix} x(t) \\ \xi(t) - x(t) \end{bmatrix}^T, \quad (6)$$

and apply the controller (5) to the systems (1) to obtain

$$\begin{cases} \dot{x}_e(t) = \tilde{F}(\theta_t)x_e(t) + \tilde{G}(\theta_t)w(t), \\ z(t) = \tilde{H}(\theta_t)x_e(t), \end{cases} \quad (7)$$

where

$$\begin{aligned} \tilde{F}(\theta_t) &= \begin{bmatrix} A(\theta_t) + B(\theta_t)K(\theta_t) & B(\theta_t)K(\theta_t) \\ G(\theta_t)K_d(\theta_t) & A(\theta_t) + G(\theta_t)K_d(\theta_t) - L(\theta_t)C(\theta_t) \end{bmatrix}, \\ \tilde{G}(\theta_t) &= \begin{bmatrix} G(\theta_t) & 0 \\ -G(\theta_t) & L(\theta_t) \end{bmatrix}, \quad \tilde{H}(\theta_t) = \begin{bmatrix} H(\theta_t) & 0 \\ K(\theta_t) & K(\theta_t) \end{bmatrix}. \end{aligned} \quad (8)$$

For notational simplicity, we will denote $A_i := A(\theta_t = i)$ and the same for the others matrices in (8), where $\{A_i, B_i, C_i, G_i, H_i, K_i, K_{di}, L_i, \tilde{F}_i, \tilde{G}_i, \tilde{H}_i, i \in \Phi\}$ are constant matrices. The following lemma gives a necessary and sufficient condition of exponential stability in the mean square for the closed-loop systems (7).

Lemma 1 (*Ji and Chizeck, 1990*). *The systems (7) are exponentially stable in the mean square, if and only if for any given set of symmetric positive definite matrices Q_i , $i \in \Phi$, there exist a set of symmetric positive definite matrices P_i , $i \in \Phi$, which satisfy the following matrix equations:*

$$\tilde{F}_i^T P_i + P_i \tilde{F}_i + \sum_{h=1}^N \lambda_{ih} P_h = -Q_i, \quad i \in \Phi. \quad (9)$$

The next theorem is essential in deriving H_∞ -norm bounding controls for systems (1).

Theorem 1 . *Let $T(s) = \tilde{H}_i(sI - \tilde{F}_i)^{-1} \tilde{G}_i$, $i \in \Phi$, if there exist a set of symmetric positive definite matrices P_i , $i \in \Phi$, and a positive scalar γ such that*

$$\tilde{F}_i^T P_i + P_i \tilde{F}_i + \frac{1}{\gamma^2} P_i \tilde{G}_i \tilde{G}_i^T P_i + \sum_{h=1}^N \lambda_{ih} P_h + \tilde{H}_i^T \tilde{H}_i < 0, \quad i \in \Phi, \quad (10)$$

then $T(s)$ is exponentially stable in the mean square, and $T(s)$ satisfies

$$\|T\|_\infty < \gamma. \quad (11)$$

Proof. By condition (10), we obtain

$$\tilde{F}_i^T P_i + P_i \tilde{F}_i + \sum_{h=1}^N \lambda_{ih} P_h < 0, \quad i \in \Phi. \quad (12)$$

It follows from Lemma 1 that $T(s)$ is exponentially stable in the mean square.

Let $\omega \in \mathbb{R}$, add and subtract $j\omega P_i$ to obtain from (10)

$$-(-j\omega I - \tilde{F}_i^T) P_i - P_i (j\omega I - \tilde{F}_i) + \frac{1}{\gamma^2} P_i \tilde{G}_i \tilde{G}_i^T P_i + \sum_{h=1}^N \lambda_{ih} P_h + \tilde{H}_i^T \tilde{H}_i < 0. \quad (13)$$

Define

$$K(j\omega) = \frac{1}{\gamma^2} \tilde{G}_i^T P_i (j\omega I - \tilde{F}_i)^{-1} \tilde{G}_i.$$

Premultiply (13) by $\frac{1}{\gamma^2} \tilde{G}_i^T (-j\omega I - \tilde{F}_i)^{-1}$, and postmultiply by $\frac{1}{\gamma^2} (j\omega I - \tilde{F}_i)^{-1} \tilde{G}_i$ to obtain

$$\begin{aligned} & -K(j\omega) - K^T(-j\omega) + K^T(-j\omega)K(j\omega) \\ & + \sum_{h=1}^N \lambda_{ih} \frac{1}{\gamma^4} \tilde{G}_i^T (-j\omega I - \tilde{F}_i)^{-1} P_h (j\omega I - \tilde{F}_i)^{-1} \tilde{G}_i + \frac{1}{\gamma^2} T^T(-j\omega)T(j\omega) < 0 \end{aligned} \quad (14)$$

or

$$\begin{aligned} I - \frac{1}{\gamma^2} T^T(-j\omega)T(j\omega) & > [I - K(-j\omega)]^T [I - K(j\omega)] \\ & + \frac{1}{\gamma^4} \sum_{h=1}^N \lambda_{ih} \tilde{G}_i^T (-j\omega I - \tilde{F}_i)^{-1} P_h (j\omega I - \tilde{F}_i)^{-1} \tilde{G}_i \\ & > 0 \end{aligned} \quad (15)$$

which implies (11). The proof is completed.

We are now in a position to present the main results of this paper.

Theorem 2 . Let $\gamma > 0$ be a prescribed level of disturbance attenuation, and Q_{1i} , Q_{2i} , $i \in \Phi$ are any given symmetric positive definite matrices. Suppose

$$K_i = -B_i^T X_i, \quad K_{di} = \frac{1}{\gamma^2} G_i^T X_i, \quad (16)$$

where $X_i > 0$, $i \in \Phi$, satisfy the following set of coupled algebraic Riccati equations

$$A_i^T X_i + X_i A_i + X_i \left(\frac{1}{\gamma^2} G_i G_i^T - B_i B_i^T \right) X_i + \sum_{h=1}^N \lambda_{ih} P_h + H_i^T H_i + Q_{1i} = 0, \quad i \in \Phi. \quad (17)$$

Suppose also

$$L_i = (I - \gamma^{-2} Y_i X_i)^{-1} Y_i C_i^T, \quad (18)$$

where $Y_i > 0$, $i \in \Phi$, satisfy the following set of coupled algebraic Riccati equations

$$A_i Y_i + Y_i A_i^T + \frac{1}{\gamma^2} Y_i H_i^T H_i Y_i - Y_i C_i^T C_i Y_i + \sum_{h=1}^N \lambda_{ih} Y_h + G_i G_i^T + Q_{2i} = 0, \quad i \in \Phi, \quad (19)$$

with $\sigma_{\max}\{Y_i X_i\} < \gamma^2$. Then the dynamic compensator (5) stabilizes the systems (1), and the closed-loop transfer function matrix $T(s) = \tilde{H}_i(sI - \tilde{F}_i)^{-1} \tilde{G}_i$ satisfies $\|T\|_\infty < \gamma$.

Proof. Define $W_i > 0$, $i \in \Phi$, by

$$W_i = \left(Y_i^{-1} - \frac{1}{\gamma^2} X_i \right), \quad i \in \Phi, \quad (20)$$

where the positive definiteness of W_i is guaranteed by the condition $\sigma_{\max}\{Y_i X_i\} < \gamma^2$, $i \in \Phi$.

Define $P_i > 0$ and $Q_i > 0$ by

$$P_i = \begin{pmatrix} X_i & 0 \\ 0 & \gamma^2 W_i \end{pmatrix}, \quad Q_i = \begin{pmatrix} Q_{1i} & 0 \\ 0 & Q_{2i} \end{pmatrix}, \quad i \in \Phi, \quad (21)$$

respectively. Then, routine algebraic computations show that

$$\tilde{F}_i^T P_i + P_i \tilde{F}_i + \frac{1}{\gamma^2} P_i \tilde{G}_i \tilde{G}_i^T P_i + \sum_{h=1}^N \lambda_{ih} P_h + \tilde{H}_i^T \tilde{H}_i = \begin{pmatrix} M_{11i} & M_{21i}^T \\ M_{21i} & M_{22i} \end{pmatrix}, \quad i \in \Phi, \quad (22)$$

where

$$\begin{aligned} M_{21i} &= K_i^T B_i^T X_i + \gamma^2 W_i G_i K_{di} - W_i G_i G_i^T X_i + K_i^T K_i, \\ M_{11i} &= (A_i + B_i K_i)^T X_i + X_i (A_i + B_i K_i) \\ &\quad + \frac{1}{\gamma^2} X_i G_i G_i^T X_i + \sum_{h=1}^N \lambda_{ih} X_h + H_i^T H_i + K_i^T K_i, \\ M_{22i} &= \gamma^2 (A_i + G_i K_{di} - L_i C_i)^T W_i + \gamma^2 W_i (A_i + G_i K_{di} - L_i C_i) \\ &\quad + \gamma^2 W_i (G_i G_i^T + L_i L_i^T) W_i + \gamma^2 \sum_{h=1}^N \lambda_{ih} W_h + K_i^T K_i \end{aligned}$$

for $i \in \Phi$. Using (16)-(20), we obtain

$$\begin{aligned} M_{21i} &= 0, \\ M_{11i} &= -Q_{1i} < 0, \\ M_{22i} &= -Q_{2i} < 0, \end{aligned}$$

for $i \in \Phi$. It follows that

$$\tilde{F}_i^T P_i + P_i \tilde{F}_i + \frac{1}{\gamma^2} P_i \tilde{G}_i \tilde{G}_i^T P_i + \sum_{h=1}^N \lambda_{ih} P_h + \tilde{H}_i^T \tilde{H}_i < 0, \quad i \in \Phi,$$

By Theorem 1, the dynamic compensator (5) stabilizes the systems (1), and the closed-loop transfer function matrix $T(s)$ satisfies $\|T\|_{\infty} < \gamma$. The proof is completed.

Remark 1: Since only a sufficient condition is given, standard assumptions of controllability and observability of the systems (1) are not explicitly required, but are necessary for existence of appropriate solutions of the design equations.

Remark 2: The design equations (17) and (19) can be solved by iterative methods proposed in Abou-Kandil *et al.* (1994).

4 Example

Consider systems (1) having three operating modes as follows

Mode 1:

$$\begin{aligned} A_1 &= \begin{pmatrix} 0.3 & 0.3 & 0.5 \\ 1 & -1 & 0.2 \\ 0.1 & 0.5 & -0.3 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0.7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ C_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad G_1 = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}, \\ H_1 &= \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \end{pmatrix}, \end{aligned}$$

Mode 2:

$$\begin{aligned} A_2 &= \begin{pmatrix} 0.5 & 0.5 & 0.3 \\ -0.5 & -2 & -1 \\ 0.4 & 0.2 & -0.1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0.7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.7 \end{pmatrix}, \\ C_2 &= \begin{pmatrix} 0.8 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0.4 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}, \\ H_2 &= \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.4 & 0 \end{pmatrix}, \end{aligned}$$

Mode 3:

$$\begin{aligned} A_3 &= \begin{pmatrix} -2 & 1.0 & -0.5 \\ 1 & -2 & 0.7 \\ 0.2 & 0.9 & -0.2 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0.7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ C_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.4 \end{pmatrix}, \\ H_3 &= \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.5 & 0 \end{pmatrix}, \end{aligned}$$

and the transition density matrix is given by

$$\Lambda = \begin{pmatrix} -2 & 0.5 & 1.5 \\ 1 & -2 & 1 \\ 0.5 & 1 & -1.5 \end{pmatrix}.$$

Let $\gamma = 1.5$, and

$$Q_{1i} = Q_{2i} = \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{pmatrix}, i \in \{1, 2, 3\}.$$

By solving eq.(17), we obtain

$$X_1 = \begin{pmatrix} 1.0170 & 0.3371 & 0.4179 \\ 0.3371 & 0.2188 & 0.1702 \\ 0.4179 & 0.1702 & 0.2493 \end{pmatrix} > 0, X_2 = \begin{pmatrix} 1.0472 & 0.2626 & 0.2937 \\ 0.2626 & 0.1379 & 0.0881 \\ 0.2937 & 0.0881 & 0.1604 \end{pmatrix} > 0,$$

$$X_3 = \begin{pmatrix} 0.3533 & 0.1916 & 0.1478 \\ 0.1916 & 0.1898 & 0.1191 \\ 0.1478 & 0.1191 & 0.1446 \end{pmatrix} > 0.$$

By solving eq.(19), we obtain

$$Y_1 = \begin{pmatrix} 0.8609 & 0.3425 & 0.5582 \\ 0.3425 & 0.3423 & 0.3721 \\ 0.5582 & 0.3721 & 1.4752 \end{pmatrix} > 0, Y_2 = \begin{pmatrix} 1.0016 & -0.1654 & 0.6564 \\ -0.1654 & 0.2339 & -0.2728 \\ 0.6564 & -0.2728 & 1.6592 \end{pmatrix} > 0,$$

$$Y_3 = \begin{pmatrix} 0.2937 & 0.0758 & 0.1461 \\ 0.0758 & 0.2413 & 0.3880 \\ 0.1461 & 0.3880 & 1.7244 \end{pmatrix} > 0,$$

and

$$\sigma_{max}(Y_1 X_1) = 2.2319 < \gamma^2, \sigma_{max}(Y_2 X_2) = 1.7882 < \gamma^2,$$

$$\sigma_{max}(Y_3 X_3) = 0.6395 < \gamma^2.$$

Hence, there exists the required dynamic compensator (5) with feedback gains

$$K_1 = \begin{pmatrix} -0.7119 & -0.2360 & -0.2925 \\ -0.3371 & -0.2188 & -0.1702 \\ -0.4179 & -0.1702 & -0.2493 \end{pmatrix}, K_2 = \begin{pmatrix} -0.7331 & -0.1838 & -0.2056 \\ -0.2626 & -0.1379 & -0.0881 \\ -0.2056 & -0.0617 & -0.1123 \end{pmatrix},$$

$$K_3 = \begin{pmatrix} -0.2473 & -0.1341 & -0.1034 \\ -0.1916 & -0.1898 & -0.1191 \\ -0.1478 & -0.1191 & -0.1446 \end{pmatrix},$$

disturbance estimate gains

$$K_{d1} = \begin{pmatrix} 0.2260 & 0.0749 & 0.0929 \\ 0.0749 & 0.0486 & 0.0378 \\ 0.0929 & 0.0378 & 0.0554 \end{pmatrix}, K_{d2} = \begin{pmatrix} 0.1862 & 0.0467 & 0.0522 \\ 0.0467 & 0.0245 & 0.0157 \\ 0.0653 & 0.0196 & 0.0357 \end{pmatrix},$$

$$K_{d3} = \begin{pmatrix} 0.0785 & 0.0426 & 0.0328 \\ 0.0426 & 0.0422 & 0.0265 \\ 0.0263 & 0.0212 & 0.0257 \end{pmatrix},$$

and observer gains

$$L_1 = \begin{pmatrix} -0.0500 & -0.1387 \\ -0.1387 & 0.0991 \\ -0.5022 & -0.1610 \end{pmatrix}, \quad L_2 = \begin{pmatrix} -0.2276 & 0.0568 \\ 0.0454 & 0.2060 \\ -0.5147 & -0.0796 \end{pmatrix},$$

$$L_3 = \begin{pmatrix} -0.3915 & -0.6067 \\ -0.7583 & -0.5955 \\ -2.4883 & -2.1659 \end{pmatrix}.$$

5 Conclusions

An output feedback controller has been derived such that the resulting closed-loop jump systems are exponentially stable in the mean square and guarantee a prescribed H_∞ norm bound of the closed-loop transfer function matrix. The proposed full order observer-based H_∞ controller can be obtained by solving two sets of coupled algebraic Riccati equations. An illustrative example is given to show the effectiveness of the proposed approach.

Acknowledgement

The first author would like to acknowledge the support of the Alexander von Humboldt Foundation, Germany.

6 References

- Abou-Kandil, H., Freiling G., and Jank, G., 1994, Solution and asymptotic behavior of coupled Riccati equations in jump linear systems, *IEEE Trans. on Aut. Contr.* **39**, 1631-1636.
- de Souza, C. E., Fragoso, M. D., 1993, H_∞ control for linear systems with markovian jumping parameters, *Control-Theory and Advanced Technology* **9**, 457-466.
- Doyle, J. C., Glover, K., Khargonekar, P. P., and Francis, B. A., 1989, State space solutions to the standard H^2 and H_∞ control problems, *IEEE Trans. on Aut. Contr.* **34**, 831-847.
- Francis, B. A., 1987, *A Course in H_∞ Control Theory*, vol. **88**. Springer-Verlag, New York.
- Ge, J. H., Frank, P. M., 1995, Stochastic stability for discrete time linear active fault tolerant control systems, *Proc. of 8th Int. Conf. on System, Modeling and Control, Zakopane, Poland*, vol. **1**, 273-278.
- Ji, Y., Chizeck, H. J., 1990, Controllability, stabilizability, and continuous time markovian jump linear quadratic control, *IEEE Trans. on Aut. Contr.* **35**, 777-788.
- Ji, Y., Chizeck, H. J., 1992, Jump linear quadratic gaussian control in continuous time, *IEEE Trans. on Aut. Contr.* **37**, 1884-1892.
- Mariton, M., 1990, *Jump Linear Systems in Automatic Control*, Marcel Dekker, New York.
- Srichander, R., Walker, B. K., 1993, Stochastic stability analysis for continuous-time fault tolerant systems, *Int. J. Control* **57**, 433-452.
- Stoorvogel, A., 1992, *The H_∞ Control Problem: A State Space Approach*, Prentice-Hall, New York.
- Wonham, W. M., 1971, Random Differential Equations in Control Theory, in *Probabilistic Methods in Applied Mathematics*, Bharucha, A. T. ed., vol. **2**, Academic, New York, pp. 131-213.

Genetische Algorithmen erstellen Faustformeln zur Reglereinstellung

A. Steinkogler, M. Odenbreit

Lehrstuhl für Systemtheorie der Elektrotechnik

Universität des Saarlandes

Saarbrücken, Deutschland

Eingelangt am 8. Mai 1995

Kurzfassung

Das Problem, Faustformeln zur Reglereinstellung zu finden, kann als ein nichtkonvexes nichtlineares Parameteroptimierungsproblem formuliert werden. Zur Lösung dieses Problems werden hier genetische Algorithmen eingesetzt. Es wird dabei eine reellwertige Codierung mit arithmetischen Kreuzungs- und Mutationsoperatoren verwendet.

Am Beispiel der Kombination von PID- T_1 -Regler und schaltendem Stellglied an I- T_2 -Strecken werden Faustformeln zur Reglereinstellung automatisch entworfen. Anhand von Simulationen wird die Qualität der mit diesen Faustformeln eingestellten Regler diskutiert.

1 Einleitung

In vielen Fällen befinden sich Regelungs-Ingenieure bei ihrer Arbeit in einem Dilemma. Auf der einen Seite lassen sich beim Reglerentwurf mit den Erkenntnissen aus der Theorie der Regelungstechnik oft sehr gute Ergebnisse erzielen. Auf der anderen Seite jedoch steht dem Entwicklungsingenieur meist nicht die Zeit zur Verfügung, ein Problem hinreichend zu durchleuchten, um diese Erkenntnisse nutzen zu können. So werden in vielen Fällen Regler nach wie vor durch den Einsatz von Faustformeln zur Reglereinstellung entworfen. Die bekanntesten solcher Faustformeln dürften die von Ziegler und Nichols oder die von Kessler sein (Merz, L., Jaschek, H., 1993). Der Prozeß des Reglerentwurfs vereinfacht sich

auf diese Weise erheblich, und es können durch Versuche an der realen Anlage oft noch Verbesserungen des Regelverhaltens erzielt werden. Leider existieren solche Faustformeln jedoch nur für einfache Strecken- und Reglertypen.

Es wird hier ein Verfahren vorgeschlagen, in dem die Suche von Faustformeln als ein Optimierungsproblem der Parameteroptimierung formuliert wird. Da dieses Problem einerseits nicht analytisch gelöst werden kann, andererseits aber relativ viele lokale Optima aufweist, und sich deshalb einer Lösung durch Abstiegsverfahren widersetzt, wird das Problem mit dem Verfahren der genetischen Algorithmen gelöst. Dieses findet zwar meist nicht das exakte Optimum, jedoch gute Näherungen dafür.

2 Formulierung als Optimierungsproblem

Im folgenden wird erläutert, wie sich ein Problem der Parameteroptimierung formulieren läßt, dessen Lösung die gesuchte Faustformel liefert. Eine solche soll für vorgegebene Strecken- und Reglertypen die einzustellenden Reglerparameter als Funktion der Streckenparameter darstellen.

Die Regelstrecke gehorcht der Gleichung

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}_S(\mathbf{x}, \mathbf{u}, \mathcal{S}, t) \\ \mathbf{v} &= \mathbf{g}_S(\mathbf{x}, \mathbf{u}, \mathcal{S}, t)\end{aligned}\tag{1}$$

Dabei stellen t die Zeit, $\mathbf{x} \in \mathbb{R}^n$ den Zustands-, $\mathbf{u} \in \mathbb{R}^p$ den Steuer-, $\mathbf{v} \in \mathbb{R}^q$ den Ausgangsvektor und $\mathcal{S} \in \mathbb{R}^\ell$ den Vektor der Streckenparameter dar. Handelt es sich beim vorgegebenen Streckentyp z.B. um P-T₁-Strecken, enthält \mathcal{S} die Streckenverstärkung und die Verzögerungszeitkonstante.

Der Regler läßt sich durch die Gleichungen

$$\begin{aligned}\dot{\mathbf{x}}_R &= \mathbf{f}_R(\mathbf{x}_R, \mathbf{x}, \mathbf{v}, \mathbf{w}, \mathcal{P}, t) \\ \mathbf{u} &= \mathbf{g}_R(\mathbf{x}_R, \mathbf{x}, \mathbf{v}, \mathbf{w}, \mathcal{P}, t)\end{aligned}\tag{2}$$

darstellen. Es sind $\mathbf{x}_R \in \mathbb{R}^r$ der Zustandsvektor des Reglers, $\mathbf{w} \in \mathbb{R}^q$ der Führungsgrößenvektor und $\mathcal{P} \in \mathbb{R}^m$ der Vektor der einzustellenden Reglerparameter. Eine Faustformel stellt nun den Vektor der optimalen Reglerparameter \mathcal{P}^* als Funktion des Streckenparametervektors \mathcal{S} dar

$$\mathcal{P}^* = \mathcal{P}^*(\mathcal{S})\tag{3}$$

Eine der Faustformeln von Ziegler und Nichols stellt z.B. die optimalen Parameter eines PID-Reglers $\mathcal{P}_{ZN}^* = [K_{PR}, T_n, T_v]^T$ in Abhängigkeit vom Streckenparametervektor $\mathcal{S}_{ZN} = [K_{PS}, T_u, T_g]^T$ dar.

Um bei Faustformeln zur Reglereinstellung vom Begriff der Optimalität sprechen zu können, wird zuerst ein Zielkriterium $I(\mathcal{P}, \dots)$ festgelegt, welches die Qualität eines einzelnen Regelvorgangs bewertet. Im technischen Bereich handelt es sich dabei meist um Kostenkriterien, die es zu minimieren gilt, z.B. die mit dem Zeitquadrat gewichtete quadratische Regelfläche

$$I_{QE} = \int_{t_0}^{t_{end}} [t^2 (\mathbf{v} - \mathbf{w})^T (\mathbf{v} - \mathbf{w})] dt \quad (4)$$

Für das weitere Vorgehen wird gefordert, daß sich die Qualität einer Reglereinstellung anhand einiger weniger Regelvorgänge \mathcal{R}_ν — im Idealfall eines einzelnen — bewerten läßt.

Eine Faustformel ist dann optimal, wenn für sie bei allen möglichen Streckenparametervektoren \mathcal{S} , allen zu untersuchenden Regelvorgängen \mathcal{R}_ν und unter Berücksichtigung von (1) und (2) gilt

$$\min_{\mathcal{P}(\mathcal{S})} I(\mathcal{P}, \mathcal{S}, \mathcal{R}_\nu) \rightsquigarrow \mathcal{P}^*(\mathcal{S}) \quad (5)$$

Es handelt sich hierbei um ein Problem der dynamischen Optimierung, das sich nur in speziellen Fällen analytisch lösen läßt. Um das Problem numerisch lösen zu können, wird es in ein Parameteroptimierungsproblem umgewandelt. Dazu wird die Funktion \mathcal{P} durch eine Funktion $\tilde{\mathcal{P}}(\mathcal{S}, \mathcal{V})$ mit a priori gewählter Struktur angenähert. Die Funktion $\tilde{\mathcal{P}}$ hängt dabei außer von \mathcal{S} noch vom Entwurfparametervektor $\mathcal{V} \in \mathbb{R}^\gamma$ ab. Eine mögliche Funktion ist z.B. die Linearisierung

$$\tilde{\mathcal{P}}(\mathcal{S}, \mathcal{V}) = \begin{bmatrix} \mathbf{v}_1^T \cdot \mathcal{S} \\ \vdots \\ \mathbf{v}_m^T \cdot \mathcal{S} \end{bmatrix} + \mathbf{v}_0 \quad (6)$$

wobei \mathcal{V} dann die Form

$$\mathcal{V} = [\mathbf{v}_0^T, \dots, \mathbf{v}_m^T]^T \quad \text{mit} \quad \gamma = m \cdot (\ell + 1) \quad (7)$$

annimmt. Für diese Vereinfachung ergibt sich das Parameteroptimierungsproblem

$$\min_{\mathcal{V}} I(\tilde{\mathcal{P}}(\mathcal{V}), \mathcal{S}, \mathcal{R}_\nu) \rightsquigarrow \mathcal{V}^*$$

für alle möglichen Streckenparametervektoren \mathcal{S} und alle Regelvorgänge \mathcal{R}_ν unter Berücksichtigung der oben genannten Nebenbedingungen (1) und (2).

Hier zeigt sich ein Problem: durch die Vereinfachung dürfte es in den meisten Fällen unmöglich werden, einen Vektor \mathcal{V}^* zu bestimmen, der für alle möglichen Streckenparametervektoren \mathcal{S} optimale Reglerkennwerte liefert. Speziell für den Fall der Linearisierung wird sofort deutlich, daß geklärt werden muß, um welchen Arbeitspunkt \mathcal{S} linearisiert wird. Das Festlegen eines Arbeitspunkts ist jedoch äußerst unerwünscht, denn wenn die Faustformeln nur für diesen Punkt und seine nähere Umgebung gelten, lassen sich mit deutlich weniger Aufwand optimale Einstellwerte für speziell diesen Punkt berechnen. Es ist aber wünschenswert, Faustformeln zu erstellen, die für einen, normalerweise konvexen, Arbeitsbereich $\mathbf{A} \subset \mathbb{R}^\ell$ gute Ergebnisse liefern.

Dies kann erreicht werden, indem im Rahmen der Parameteroptimierung eine ausreichende Anzahl sinnvoll gewählter Streckenparametervektoren \mathcal{S} aus dem Arbeitsbereich \mathbf{A} , die Testfälle \mathcal{S}_κ , betrachtet werden. Für jeden einzelnen Testfall \mathcal{S}_κ werden die zur Bewertung benötigten Regelvorgänge \mathcal{R}_ν simuliert; die resultierende Zielfunktion ist die gegebenenfalls gewichtete Summe der einzelnen Zielfunktionen der Testfälle.

Um das unterschiedlich schnelle Einschwingen der verschiedenen Strecken zu berücksichtigen, werden die Endzeiten $t_{end,\kappa}$ der Simulation für jeden Testfall \mathcal{S}_κ ermittelt. Werden diese als konstant angenommen, besteht die Gefahr, daß die Einstellregeln besser an langsamere Strecken angepaßt sind, da diese mehr Zeit zum Einschwingen benötigen. Jede der Zeitkonstanten $t_{end,\kappa}$ sollte dabei von einer die zugehörige Strecke charakterisierenden Zeitkonstante abhängen. Als solche bieten sich zum Beispiel die Ausregelzeit für den zeitoptimalen Regelvorgang, die Ausregelzeit bei Regelung mit $K_{PR,krit}/2$ oder ähnliche an. Der Einfluß der Endzeit auf die Zielfunktionen muß anschließend eliminiert werden. Bei Verwendung des Zielkriteriums aus (4) geschieht dies durch eine Division durch $t_{end,\kappa}^3$.

Es ergibt sich dann als Zielfunktion

$$J(\dots) = \sum_{\mathcal{S}_\kappa} \sum_{\mathcal{R}_\nu} \left\{ \frac{1}{t_{end,\kappa}^3} \int_{t_0}^{t_{end,\kappa}} \left[t^2 \cdot (\mathbf{v} - \mathbf{w})^T (\mathbf{v} - \mathbf{w}) \right] \Big|_{\mathcal{R}_\nu} dt \right\} \quad (8)$$

Die gesuchte Faustformel ergibt sich damit als

$$\tilde{\mathcal{P}}^*(\mathcal{S}) = \mathcal{P}(\mathcal{S}, \mathcal{V}^*) \quad (9)$$

wobei sich die Größe \mathcal{V}^* aus

$$\mathcal{V}^* = \min_{\mathcal{V}} [J(\dots)] \quad (10)$$

unter Berücksichtigung von (1) und (2) berechnet.

Das Problem läßt sich nicht lösen, indem mit der Methode der minimalen Fehlerquadrate die Faustformeln aus vorab bestimmten optimalen Reglerparametern berechnet werden. Auf solche Art entstandene Faustformeln können die unterschiedlichen Empfindlichkeiten der Regelgüte gegenüber den einzelnen Reglerparametern nicht berücksichtigen und führen deshalb meist zu sehr viel schlechteren Reglereinstellungen.

3 Genetische Algorithmen

Da die im vorigen Abschnitt entwickelte Zielfunktion meist stark nichtlinear ist und viele lokale Minima aufweist, widersetzt sie sich in vielen Fällen einer Optimierung durch konventionelle Abstiegsverfahren. Aus diesem Grund wird sie hier mit genetischen Algorithmen (Goldberg, D.E., 1989) optimiert, die im folgenden kurz beschrieben werden.

Das Verfahren der genetischen Algorithmen simuliert in der Natur vorkommende Prinzipien der Evolution und kann zur Optimierung technischer Prozesse eingesetzt werden. Im Gegensatz zu anderen Optimierungsverfahren werden große Anzahlen von Punkten \mathbf{c}_i des Suchraums $S \subset \mathbb{R}^p$, also möglichen Lösungen, quasiparallel untersucht. Dabei wird mit Codierungen der möglichen Lösungen gearbeitet. Jede mögliche Lösung wird dabei als *Individuum*, die Menge der Individuen als *Population* bezeichnet. Die Anzahl der Individuen in einer Population heißt *Populationsgröße* μ . Die einzelnen Komponenten der \mathbf{c}_i werden als *Eigenschaften* des i -ten Individuums bezeichnet. Jedes Individuum einer Population wird durch die *Fitnessfunktion* $\mathcal{F}(\mathbf{c}_i)$ bezüglich seiner Qualität bewertet. Dabei bedeutet eine hohe Fitness eine gute Lösung des Problems, eine geringere Fitness eine schlechtere Lösung. Handelt es sich bei der Zielfunktion J um eine zu minimierende Zielfunktion, kann die Fitness z.B. als $\mathcal{F} = C - J$ durch Subtraktion der Zielfunktion von einer geeigneten Konstante C berechnet werden.

Genetische Operatoren modellieren einzelne im Rahmen der Evolution stattfindende Prozesse. Durch die Anwendung der genetischen Operatoren verändern sich die Individuen der Population und sie durchläuft den Übergang von einer *Generation* zur nächsten. Nach den Gesetzen der Wahrscheinlichkeit verbessert sich dabei die mittlere Fitness der Population. Die wichtigsten genetischen Operatoren sind:

Überleben der Besten: In der Natur überleben die besser an ihre Umwelt angepaßten Individuen länger und vermehren sich häufiger als ihre schlechter angepaßten Kon-

kurrenten. In genetischen Algorithmen wird durch den Operator der *Selektion* sichergestellt, daß Individuen, deren Fitness höher ist, mit einem höheren Erwartungswert in der nächsten Generation vertreten sind als solche mit geringerer Fitness. Manche Implementierungen des Selektionsoperators benötigen einen Parameter p_r , der angibt, welcher Anteil der Population der Selektion unterworfen werden soll. Dieser Parameter wird *Selektionsrate* genannt.

Vermischung von Erbgut: Verschiedene Lebewesen können Erbgut miteinander austauschen und das neu entstandene Erbgut an ihre Nachkommen weitergeben. In genetischen Algorithmen wird durch den Operator der *Kreuzung* das Erbgut mehrerer Individuen vermischt. Jedes Individuum wird mit der *Kreuzungswahrscheinlichkeit* p_c an einer Kreuzung beteiligt.

Zufällige Veränderung von Erbgut: Bei Lebewesen kann durch äußere Einflüsse das Erbgut verändert werden. In genetischen Algorithmen wird diese Tatsache durch den Operator der *Mutation* modelliert. Die Wahrscheinlichkeit, daß ein Individuum mutiert wird, beträgt p_m . Die *Mutationswahrscheinlichkeit* p_m wird hier auf ein Individuum und nicht auf ein Zeichen des Codes, also bei binärer Codierung ein bit, wie bei (Goldberg, D.E., 1989) bezogen. Dadurch wird erreicht, daß die Wahl des Parameters p_m beim Wechsel auf eine Codierung durch Zeichen mit anderem Informationsgehalt beibehalten werden kann.

Auch andere genetische Operatoren sind möglich, es können zum Beispiel geschlechtliche Fortpflanzung, dominante und rezessive Vererbung durch genetische Operatoren nachgebildet werden (Goldberg, D.E., 1989). Die Wirkung dieser Operatoren ist jedoch nicht genau bekannt, aus diesem Grund werden sie hier nicht benutzt.

Die in dieser Arbeit verwendeten genetischen Algorithmen sind mit reellcodierten Individuen implementiert. Jede Eigenschaft des Individuums wird durch eine reelle Zahl codiert, dies geschieht durch die identische Abbildung. Als Selektion wird die Turnier-Selektion verwendet. Bei dieser werden $p_r \cdot \mu / 2$ Paare nacheinander zufällig ausgewählt. Jedes dieser Paare führt ein ‚Turnier‘ durch. Dabei wird das Individuum mit der schlechteren Fitness durch das mit der besseren ersetzt. Sowohl Kreuzungs- als auch Mutationsoperator werden durch arithmetische Operationen realisiert. Die Kreuzung wird dabei auf folgende Weise realisiert: es wird $p_c \cdot \mu / 2$ mal ein Paar zufällig ausgewählt. Für jedes Paar von Eigenschaften $(c_{i1,j}, c_{i2,j})$ des ausgewählten Paares wird eine Zufallszahl $\alpha_j \in [-0.5; 1.5]$

ermittelt. Die neuen Eigenschaften ergeben sich als

$$\begin{aligned}c'_{i1,j} &= (1 - \alpha_j) \cdot c_{i1,j} + \alpha_j \cdot c_{i2,j} \\c'_{i2,j} &= (1 - \alpha_j) \cdot c_{i2,j} + \alpha_j \cdot c_{i1,j}\end{aligned}\tag{11}$$

Die neuen Individuen ersetzen ihre Vorfahren. Die Mutation wird dabei durch Multiplikation jeder Eigenschaft $c_{i,j}$ mit einer Zufallszahl $\beta_j \in [0.8; 1.25]$ durchgeführt. Auch hier ersetzt das mutierte Individuum das, aus dem es entsteht. Wird durch die Kreuzung oder die Mutation ein Individuum erzeugt, das nicht im Suchraum S liegt, werden die Änderungen an diesem Individuum rückgängig gemacht und das Individuum wird unverändert in die nächste Generation übernommen.

Obwohl bei binärer Codierung der Individuen Berechnungen der Konvergenzwahrscheinlichkeiten genetischer Algorithmen vorliegen (Goldberg, D.E., 1989), bei reellwertiger Codierung nicht, wird letztere bevorzugt, da sich in vorangegangenen Untersuchungen bei ähnlichen Problemen für diese besseres Konvergenzverhalten ergab: nach gleicher Anzahl von Generationen waren bei gleicher Populationsgröße meist bessere Lösungen vorhanden und auch die Vielfalt der Individuen der Population war bei reeller Codierung größer.

4 Beispiel eines Entwurfs

Um die Vorgehensweise beim Entwurf von Faustformeln zur Reglereinstellung zu erläutern, wird hier eine Art von Regler betrachtet, die sich mit herkömmlichen Faustformeln nur unbefriedigend einstellen läßt. Das Konzept dieses Reglers wird im folgenden kurz erläutert.

Sollen große Leistungen geregelt werden, verbietet sich der Einsatz linearer Stellglieder oft aus Kostengründen. Schaltende Stellglieder sind in vielen Fällen um Größenordnungen billiger in der Anschaffung und im Betrieb. Andererseits tritt häufig der Fall auf, daß der Einsatz von Zwei- oder Dreipunktreglern zu äußerst unbefriedigendem Regelverhalten führt. Aus diesem Grund kann erwogen werden, einen linearen Regler, z.B. einen PID-T₁-Regler an einem schaltenden Stellglied einzusetzen. Der PID-T₁-Regler kann dann kostengünstig realisiert werden, da er nur auf niedrigem Leistungsniveau arbeiten muß. Ein so aufgebauter Regelkreis kann ähnlich gutes Regelverhalten erzielen wie einer mit linearem Stellglied. Die Einstellung des Reglers darf dann aber nicht mit herkömmlichen Faustformeln erfolgen, da diese nicht für das betrachtete Problem entworfen wurden. Besonders in der Antriebstechnik, wo häufig schaltende Stellglieder eingesetzt werden,

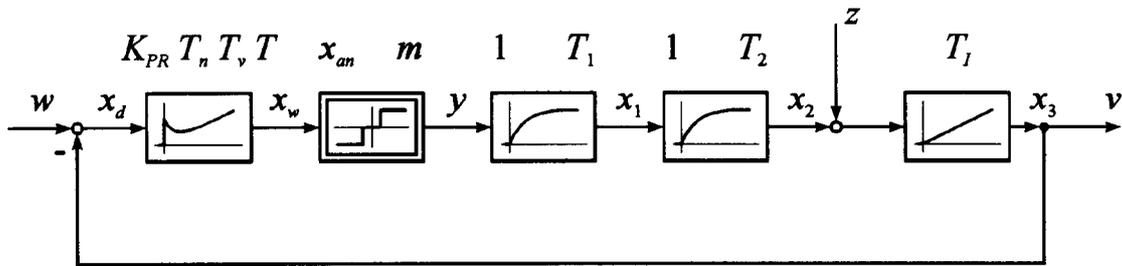


Abbildung 1: Signalflußplan des Regelkreises

verspricht der Einsatz von speziell auf das schaltende Stellglied abgestimmten PID- T_1 -Reglern große Vorteile.

Im hier betrachteten Fall hat der PID- T_1 -Regler die Übertragungsfunktion

$$F_R(p) = K_{PR} \left(\frac{1}{T_n p} + \frac{1 + T_v p}{1 + T p} \right) \quad (12)$$

Die Funktion \mathcal{P} nach Gl. 2 enthält die Parameter K_{PR} , T_n und T_v . Die parasitäre Zeitkonstante T ist hier auf den Wert $T = 0.1 \cdot T_v$ festgelegt, da dieser Wert sich auch mit einfachen elektronischen Regelschaltungen noch realisieren läßt, und nicht anzunehmen ist, daß sich Verbesserungen durch größeres T erzielen lassen.

Das Stellglied hat das Schaltgesetz

$$y = \begin{cases} m & \text{für } x_w > x_{an} \\ 0 & \text{für } -x_{an} \leq x_w \leq x_{an} \\ -m & \text{für } x_w < -x_{an} \end{cases}, \quad (13)$$

dabei werden immer $x_{an} = 1$ und $m = 2$ angenommen. Es macht wenig Sinn, m für die Optimierung freizugeben, da dieser Wert meist durch die Aufgabenstellung festgelegt ist. Der Parameter x_{an} ist redundant und läßt sich durch K_{PR} skalieren, x_{an} kann also ohne Verlust der Allgemeingültigkeit auf jeden beliebigen Wert festgelegt werden.

Als Typ der Musterstrecke wird hier eine I- T_2 -Struktur gewählt, die drei reelle Pole besitzt. Der Integrierer liegt am Ausgang der Strecke, an seinem Eingang greift die Störgröße $z(t)$ an. Der Signalflußplan des gesamten Regelkreises ist in Abb.1 dargestellt.

Die hier zu erstellende Faustformel zur Reglereinstellung soll die optimalen Werte für \mathcal{P} in Abhängigkeit von den Streckenparametern T_1 , T_2 und T_I darstellen. Eine der Streckenzeitkonstanten kann dabei jedoch auf einen beliebigen Wert festgelegt werden, da sie durch

κ	1	2	3	4	5	6	7	8	9	10
T_1/s	1.0	1.0	1.0	1.0	0.7	0.7	0.7	0.4	0.4	0.1
T_2/s	1.0	0.7	0.4	0.1	0.7	0.4	0.1	0.4	0.1	0.1

 Tabelle 1: Auswahl der Zeitkonstanten von I-T₂-Regelstrecken

eine Skalierung aller Zeitkonstanten auf jeden anderen Wert umgerechnet werden kann (Odenbreit, M., 1994), entscheidend sind nur die Verhältnisse der Zeitkonstanten untereinander. Deshalb wird die Integrationszeitkonstante T_I auf den Wert 1.0 s festgelegt. Der Streckenparametervektor \mathcal{S} lautet also $\mathcal{S} = [T_1, T_2]$. Desweiteren wird ohne Verlust der Allgemeingültigkeit davon ausgegangen, daß $T_1 \geq T_2$ gilt.

Als Testfälle \mathcal{S}_κ werden die in Tabelle 1 gezeigten zehn gewählt. Für jeden dieser Testfälle wird sowohl das Führungs- als auch das Störverhalten des Regelkreises bei einem Einheitssprung simuliert und die jeweiligen Gütefunktionen addiert. Es werden also zwei Regelvorgänge \mathcal{R}_ν betrachtet.

Werden an die Strecke gleichzeitig eine Störung $z = 1$ und die Steuergröße $y = -m = -2$ angelegt, ergibt sich eine Zeitkonstante T_0 , nach der der Streckenausgang wieder den Wert $v = 0$ erreicht. Das Streckenverhalten kann für jeden Testfall durch Simulation oder analytisch, der Nulldurchgang des Ausgangssignals aus der graphischen Darstellung oder numerisch bestimmt werden. Für die so ermittelten Zeitkonstanten gilt etwa $T_{0,\kappa} \approx 2 \cdot (T_{1,\kappa} + T_{2,\kappa})$. Die Zeit $t_{end,\kappa}$ wird um den Faktor 5 höher, also zu $t_{end,\kappa} = 10 \cdot (T_{1,\kappa} + T_{2,\kappa})$ gewählt.

Als Gütefunktion ergibt sich so

$$J = \sum_{\mathcal{S}_\kappa} \left\{ \frac{1}{t_{end,\kappa}^3} \int_{t_0}^{t_{end,\kappa}} [t^2 \cdot (v - w)^2] dt \Big|_{\substack{w(t)=1 \\ z(t)=0}} + \frac{1}{t_{end,\kappa}^3} \int_{t_0}^{t_{end,\kappa}} [t^2 \cdot (v - w)^2] dt \Big|_{\substack{w(t)=0 \\ z(t)=1}} \right\} \quad (14)$$

Eine Optimierung mit 500 Individuen und 250 Generationen liefert als Faustformeln für die Einstellung der Reglerparameter in Abhängigkeit von den Streckenparametern

$$\begin{aligned} K_{PR} &= 37.3 - 12.7 \cdot T_1 - 16.9 \cdot T_2 \\ T_n &= 36.8 - 9.99 \cdot T_1 - 1.31 \cdot T_2 \\ T_v &= 0.0103 + 0.244 \cdot T_1 + 0.708 \cdot T_2 \end{aligned} \quad (15)$$

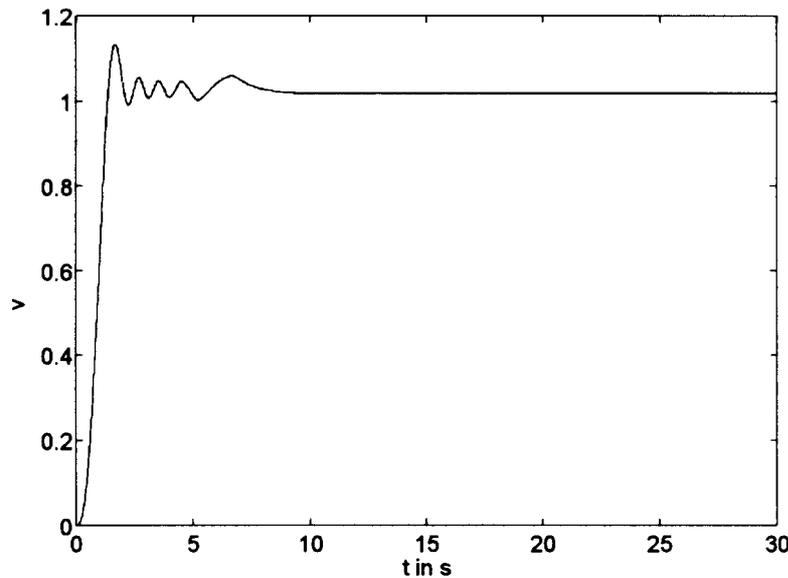


Abbildung 2: Verhalten des Regelkreises nach (Odenbreit, M., 1994) bei einem Führungssprung von $\Delta w = 1.0$

Um diese Faustformeln zuverlässig zu testen, werden sie an einer Strecke angewandt, die im Arbeitsbereich liegt, jedoch keinem der Testfälle \mathcal{S}_k entspricht. Es wird hier die Strecke mit den Parametern $T_1 = 0.8$ s und $T_2 = 0.2$ s ausgewählt. Wird an dieser ein PID- T_1 -Regler mit schaltendem Stellglied eingesetzt und mit Faustformeln nach (15) eingestellt, zeigt sich das in Abb.2 dargestellte Führungsverhalten des Regelkreises bei einem Führungssprung von $\Delta w = 1.0$. Die in Abb.2 erkennbare Regeldifferenz von etwa 2.5 % ist jedoch keine bleibende — der integrierende Anteil des Reglers verringert sie im Mittel auf Null. Durch das schaltende Stellglied tritt zwar in vielen Fällen eine Arbeitsbewegung auf, diese hat jedoch sehr kleine Amplitude und Frequenz.

Das Störverhalten des Regelkreises als Reaktion auf einen Störsprung von $z = 1$ ist in Abb.3 gezeigt. Auch hier zeigt sich ein sehr gutes Verhalten des Regelkreises. Die Arbeitsbewegung ist hier jedoch deutlich stärker ausgeprägt als beim Führungsverhalten, da das Stellglied nicht längere Zeit in der Null-Lage verharren kann. Das Störverhalten ist durchaus vergleichbar mit dem eines Regelkreises mit realem PID-Regler, der nach den Faustformeln von Kessler eingestellt wurde (Abb.4). Der nach (Odenbreit, M., 1994) eingestellte Regelkreis erreicht wesentlich schneller das 5 %-Toleranzband und verharrt dort. Erst bei Betrachtung des 2 %-Toleranzbandes erzielt der nach Kessler eingestellte Regelkreis besseres Verhalten, da die Arbeitsbewegungen in Abb. 3 fast 2 % betragen.

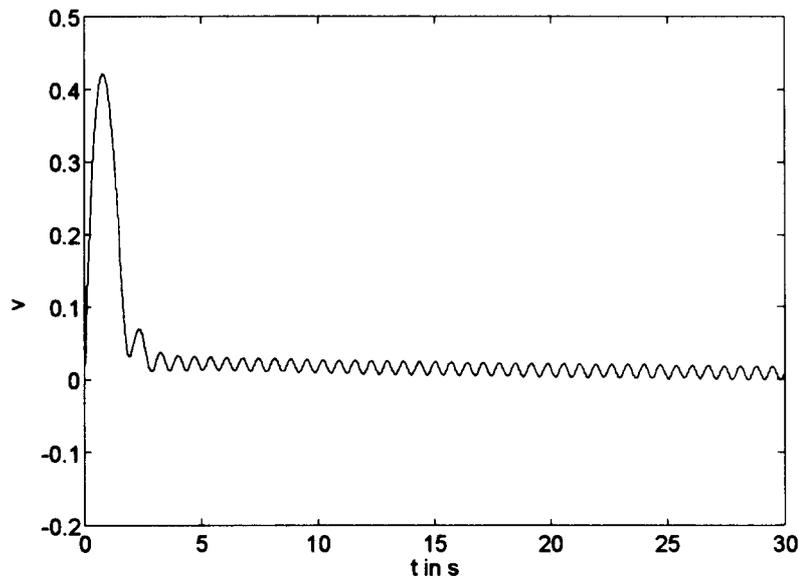


Abbildung 3: Verhalten des Regelkreises nach (Odenbreit, M., 1994) bei einem Störsprung von $\Delta z = 1.0$

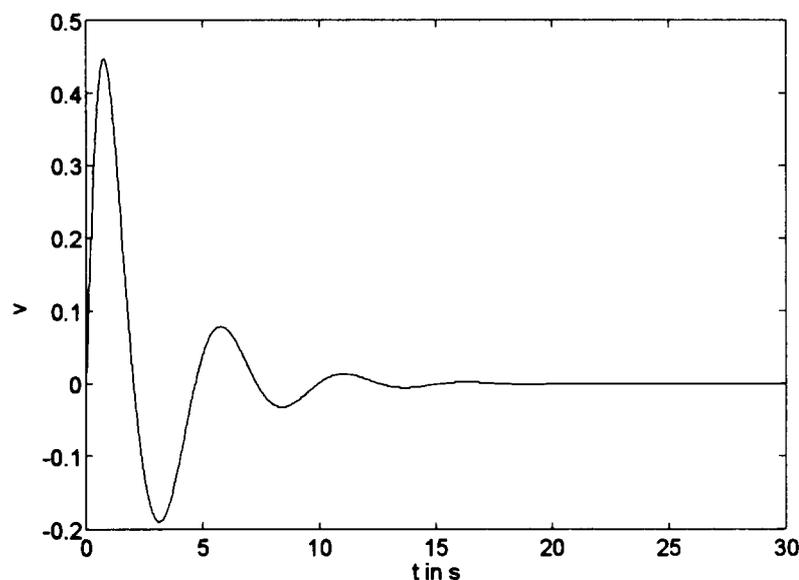


Abbildung 4: Verhalten des Regelkreises nach Kessler bei einem Störsprung von $\Delta z = 1.0$

5 Zusammenfassung und Ausblick

In dieser Arbeit wird ein universelles Verfahren beschrieben, um Faustformeln zur Reglereinstellung automatisch zu generieren. Für die Faustformeln wird eine Struktur a priori festgelegt, die von einer Anzahl von Entwurfsparametern abhängt. Durch Simulationen von Regelkreisen mit verschiedenen in einem Arbeitsbereich liegenden Regelstrecken und verschiedenen Regelvorgängen läßt sich eine Gütefunktion berechnen, die die Qualität

einer Faustformel festlegt. Das entstehende Parameteroptimierungsproblem wird mit genetischen Algorithmen gelöst. Die so gefundenen Lösungen weisen bei Regelstrecken im Arbeitsbereich hohe Qualität auf.

Einer vollständigen Automatisierung des Entwurfs der Faustformeln stehen jedoch noch zwei Probleme entgegen: die Struktur der Formeln muß vor der Optimierung festgelegt werden, und die Testfälle und Regelvorgänge müssen sowohl in ausreichender Anzahl als auch sinnvoll gewählt werden. Durch die Verwendung der genetischen Programmierung (Koza, J.R., 1992) oder einer Kombination dieser mit genetischen Algorithmen läßt sich wohl das erste Problem und durch eine Co-Evolution der Testfälle und Regelvorgänge (Koza, J.R., 1992) wohl das zweite lösen.

Literaturverzeichnis

- Goldberg, D.E., 1989, *Genetic Algorithms in Search, Optimization and Machine Learning*, (Addison-Wesley, Reading, Mass.)
- Koza, J.R., 1992, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, (MIT Press, Cambridge, Mass.)
- Merz, L.; Jaschek, H., 1993, *Grundkurs der Regelungstechnik*, 12. Auflage, (R. Oldenbourg Verlag, München)
- Odenbreit, M., 1994, *Entwurf von Einstellregeln für schaltende Regler mit Hilfe genetischer Algorithmen*, Diplomarbeit, (Lehrstuhl für Systemtheorie der Elektrotechnik, Universität des Saarlandes)

Verknüpfung von Echtzeitsystemen mit nichtechtzeit-tauglichen Verwaltungssystemen

Carl Schrotek und Richard Demattio
ELIN Energieanwendung GmbH, Wien

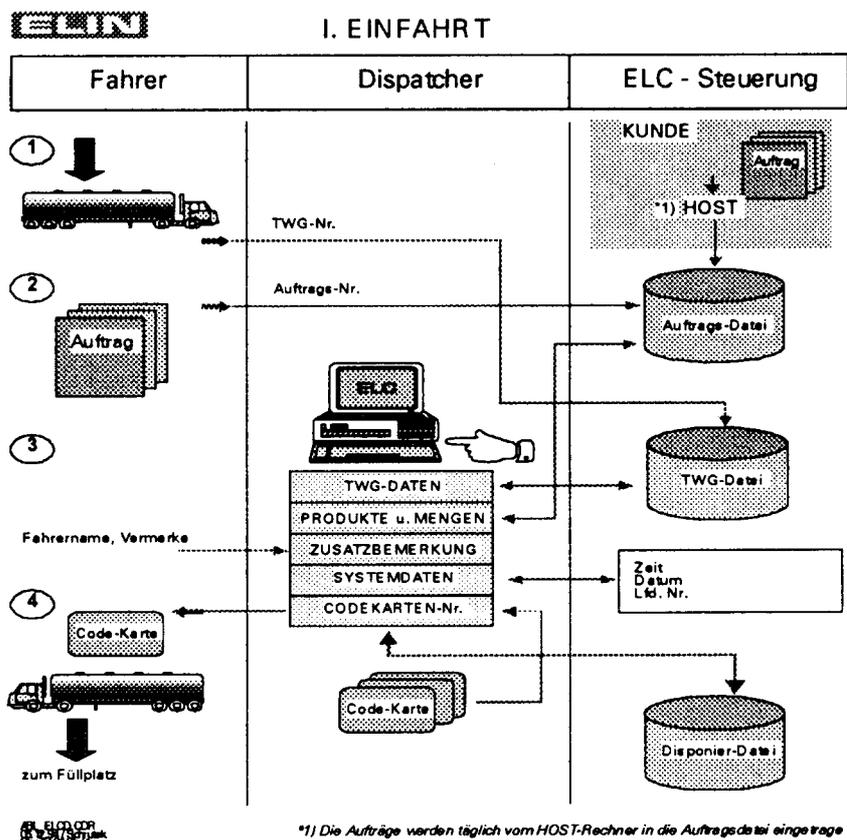
Eingelangt am 12. Juli 1995

Zusammenfassung

Bei der Automatisierung von technischen Anlagen, bei denen neben dem technisch einwandfreien Ablauf der Prozesse auch eine kaufmännische Verwaltung in den automatischen Betrieb miteinzubeziehen ist, sind die unterschiedlichen Datenqualitäten der technischen Echtzeitsysteme und der nichtechtzeit-tauglichen Verwaltungssysteme miteinander zu verknüpfen. Darüberhinaus sind die Redundanzkonzepte für die Betriebstauglichkeit derartiger Systeme von entscheidender Bedeutung.

EINLEITUNG

Bei der Automatisierung von Anlagen, bei denen technische Abläufe von Verwaltungsabläufen und kaufmännischen Daten abhängig sind, ergibt sich die Notwendigkeit der Datenverknüpfung zwischen zwei qualitativ unterschiedlichen Systemen. Die daraus entstehende Problematik soll an Hand der Automatisierung von Abfüllanlagen für Mineralölprodukte dargestellt werden. Für Raffinerien und Tanklagerbetreiber besteht die Notwendigkeit, die Produkte an die Transportmittel ihrer Kunden abzufüllen. Diese Transportmittel sind im allgemeinen Kesselwagen oder Tankwagen. Die Entladung der Produkte beim Endverbraucher ist üblicherweise sehr einfach, die Verladung hingegen ist sowohl bezüglich des Steuerungssystems als auch der Meßwerterfassung und vor allem bezüglich der Datenverarbeitung wesentlich komplexer. Die Automatisierungstechnik, insbesondere die Datenverarbeitung, spielt dabei eine besondere Rolle.



Die Entladung der Produkte beim Endverbraucher ist üblicherweise sehr einfach, die Verladung hingegen ist sowohl bezüglich des Steuerungssystems als auch der Meßwerterfassung und vor allem bezüglich der Datenverarbeitung wesentlich komplexer. Die Automatisierungstechnik, insbesondere die Datenverarbeitung, spielt dabei eine besondere Rolle.

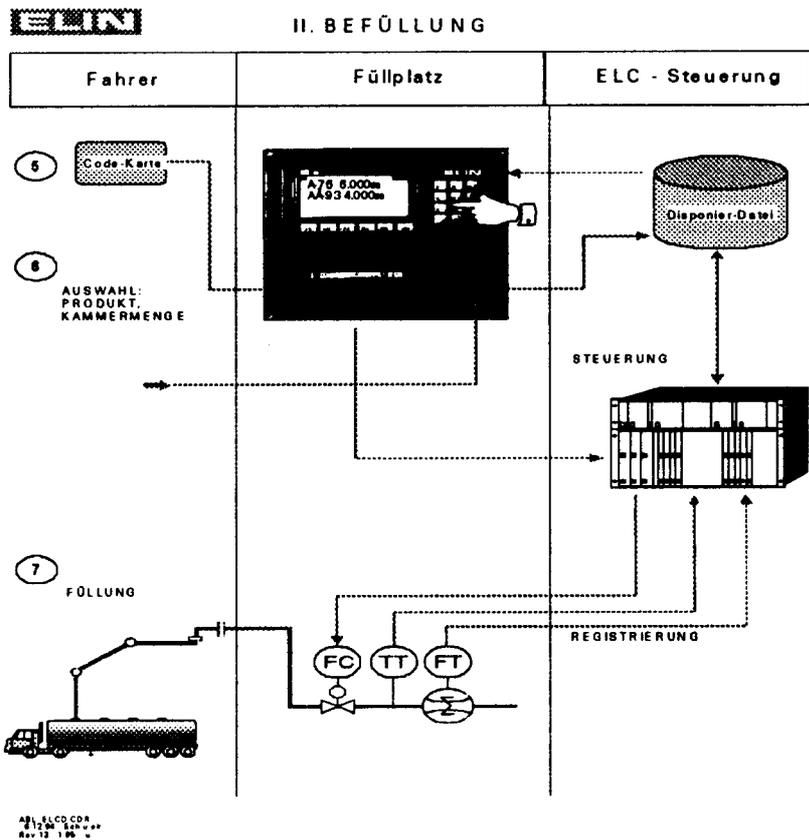
DER ABLAUF EINES ABFÜLLVORGANGS

Die Kesselwagenfüllstation (KWG):

Bei der Kesselwagenfüllstation wird ein gesamter Kesselwagenzug im Verband mittels Verschublokomotive auf die Füllstation zugestellt. Da jeder Kesselwagen einzeln befüllt wird, muß jeder Kesselwagen auf den eigentlichen Verladepunkt (Waage und Befülleinrichtung) rangiert werden (daher auch die Bezeichnung ON SPOT Verladung). Diese Positioniersteuerung wird von einer speicherprogrammierbaren Steuerung (SPS) ausgeführt. Dies erfordert bereits eine geeignete Software, da während der Positionierung die Art und Größe des Kesselwagens zu messen und zu berechnen ist. Aufgrund dieser Berechnung wird die Waage, eine geteilte Brückenwaage, ausgewählt und der Kesselwagen mittig positioniert. Danach wird der Kesselwagen gewogen (Leergewicht), das Füllrohr eingeführt, die Pumpen, Ventile etc. gesteuert und somit der Kesselwagen befüllt. Nach Befüllung wird der Kesselwagen neuerlich gewogen und dadurch das Verladegewicht festgestellt. Während dieses gesamten Vorganges werden die notwendigen technischen Daten mit dem Verladerechner ausgetauscht.

Die Tankwagenfüllstation (TWG):

Bei der Tankwagenfüllstation fährt der Tankwagen-Fahrer in die Verladeanlage ein und meldet sich bei der Einfahrt mittels Codekarte an (Dispatcher). Die Anmeldung wird durch den Verladerechner entgegengenommen und abgewickelt (Kontrolle über Auftrag, Erfassung des Tankwagens etc.). Ab diesem Zeitpunkt wird der Fahrer mittels Codekarte identifiziert. Somit kann der Tankwagen auf einen beliebigen Füllplatz fahren und dort selbständig das gewünschte Produkt abfüllen. Am Füllplatz muß sich der Tankwagen-Fahrer mittels Codekarte nochmals identifizieren, worauf die für die Steuerung vorgesehene SPS vom Verladerechner die zulässigen Produkte und Mengen übermittelt bekommt. Mittels eines an die SPS angeschlossenen Füllbühnenterminals kann der Tankwagen-Fahrer die Daten kontrollieren und eine gewisse Auswahl treffen. Der Füllvorgang (Steuerung der Pumpen, Ventile etc.) wird von der SPS abgewickelt. Am Ende des Füllvorganges werden die Daten an den Verladerechner übermittelt und gespeichert. Verläßt der Fahrer wieder die Füllstation, muß er sich bei der Ausfahrt abmelden, wo er auch die Verladepapiere ausgehändigt bekommt. Wie aus dem Vorgang ersichtlich, hat bei der TWG der Verladerechner neben der Aufgabe, die Aufträge zu speichern und die Verlademengen zu registrieren, auch die Aufgabe, den gesamten Vorgang in der Verladeanlage zu überwachen.



DER VERLADERECHNER

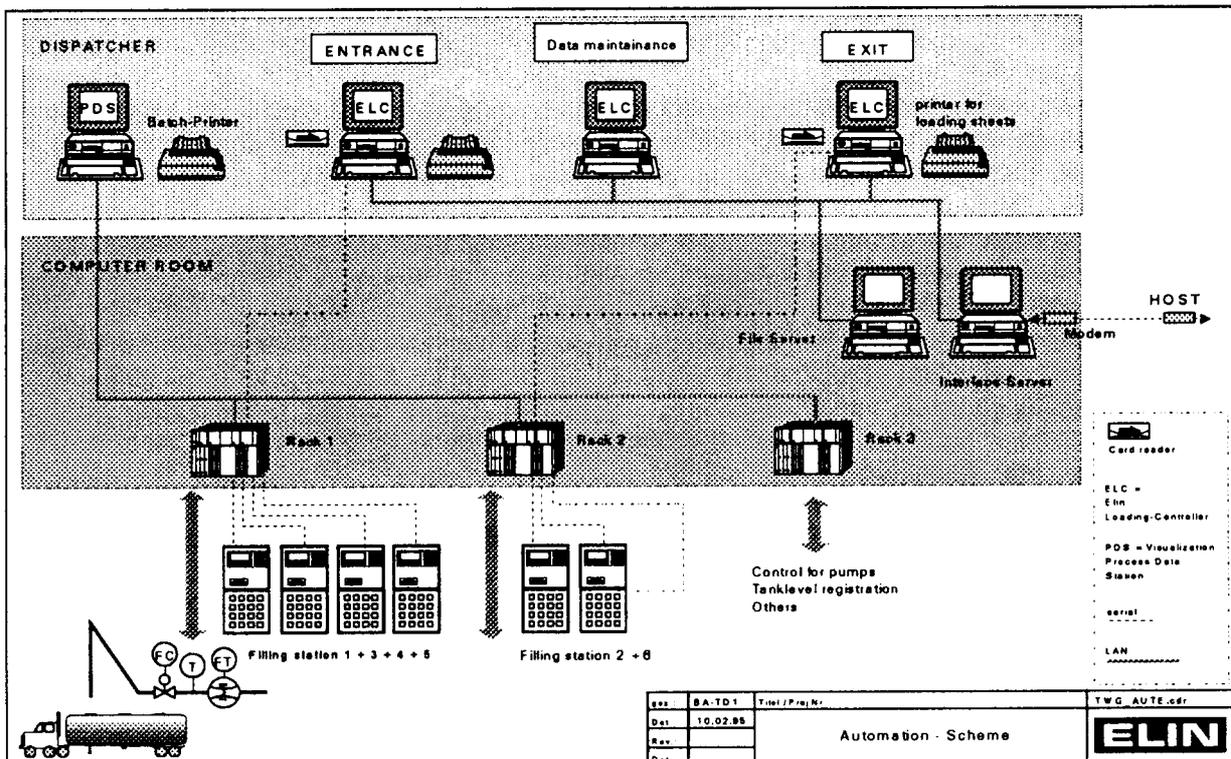
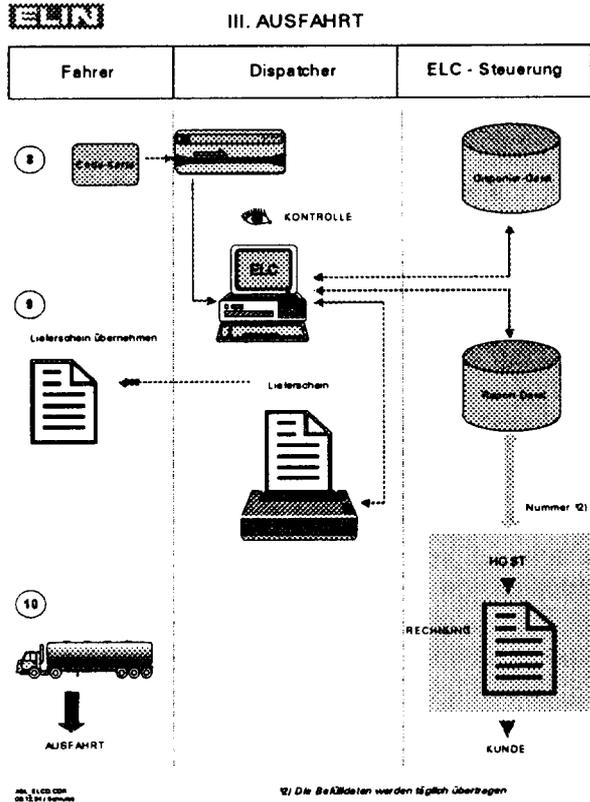
Der Aufbau des Systems ist für die KWG und TWG grundsätzlich gleich, im folgenden wird das Konzept an Hand einer TWG betrachtet.

Die Ausgangssituation:

Die langjährige Beschäftigung mit Verladeanlagen hat bereits einen beträchtlichen Erfahrungsschatz für den Bau derartiger Anlagen anwachsen lassen. So wurde bereits bei einigen Anlagen ein einfacher Verladerechner auf PC-Basis, jedoch noch unter DOS, installiert. Auch Erfahrungen beim Aufbau von Netzwerken liegen vor.

Das Automatisierungskonzept:

Nur eine PC-Lösung kann im Hinblick auf die Preissituation die hohen Automatisierungserwartungen befriedigen. Da „Windows“ den Stand der Technik darstellt, wird die Bedienoberfläche entsprechend ausgeführt. Das mit gutem Erfolg bereits bei früher realisierten Anlagen eingesetzte Datenbanksystem FOXPRO wird von Microsoft in einer Windows-Version unterstützt. Für die Vernetzung der PC's wird NOVELL eingesetzt, das sich auch mit WINDOWS gut verträgt. Novell bietet verschiedene Sicherheitsstufen für den SERVER an. (Je nach Kundenwunsch kann eine Spiegelplatte oder ein Doppelsystem eingesetzt werden, ohne daß dies in den Applikationen berücksichtigt werden muß). Die Betriebssicherheit hat eine überaus große Bedeutung,



denn der Betrieb der gesamten Anlage darf nicht von einem PC abhängig sein. Dabei bildet der Server naturgemäß jene Stelle, an der besondere Vorkehrungen getroffen werden müssen. Daher wird je nach gewünschter Betriebssicherheit entweder ein sofort einsetzbarer Ersatz-Server (Ausfallzeit ca 30 Minuten) oder die zuvor erwähnte Spiegelplatte bzw. Spiegel-PC ausgeführt. Dabei entstehen keine Ausfallszeiten, die Reparatur kann während eines Anlagenstillstandes durchgeführt werden.

Das Redundanzkonzept zielt auf die Aufrechterhaltung der Lieferfähigkeit der Abfüllanlage hinsichtlich aller Produkte, die in dieser Anlage abgefüllt werden können, ab. Das Konzept sieht vor, daß alle Füllbühnen für alle Produkte zur Verfügung stehen müssen, unabhängig davon, welchem Steuerungssegment die einzelne Füllbühne zugeteilt ist. Bei Störung einer Funktion auf einer Füllbühne (Pumpenausfall, Ventilfehler, etc.) kann das Produkt an einer anderen Füllbühne abgefüllt werden. Bei Ausfall eines Loading Controllers wird der Abfüllvorgang auf der dieser Steuerungseinheit zugeteilten Füllbühne über eine andere Steuerungseinheit abgewickelt, ohne daß dadurch die Vernetzung mit den kaufmännischen Transaktionen eine Störung erfährt. Die Betriebssicherheit bezieht sich daher vorwiegend auf den ungestörten Ablauf der Geschäftsfunktionen, die wesentlich von der Betriebssicherheit des Verladerechners beeinflusst wird. Ziel des Redundanzkonzepts ist also die 100 %ige Verfügbarkeit des Abfüllprozesses als Geschäftsfall, der zu beliefernde Kunde erleidet keinen Ausfall durch nicht von ihm selbst verschuldete Vorkommnisse auf der Abfüllanlage. Es ist daher Aufgabe des Verladerechners, daß die ohnehin auf der Anlage in mehrfacher Ausführung vorhandenen Einzelkomponenten in gegenseitigem Austausch einsetzbar sind, ohne die vielen bereits erwähnten gegenseitigen Abhängigkeiten, Vernetzungen und Verriegelungen zu verletzen.

Der modulare Aufbau:

Da keine Anlage der anderen gleicht, muß die Software modular aufgebaut und leicht durch kundenspezifische Sondermodule erweiterbar sein. Einige Softwaremodule sind allerdings vielseitig verwendbar, wie z.B. jenes zur Auswertung von Meldungen.

Die wichtigsten Module sind:

PARAMETRIERUNG

Hier können die anlagen-spezifischen Einstellungen vorgenommen werden.

DATENPFLEGE

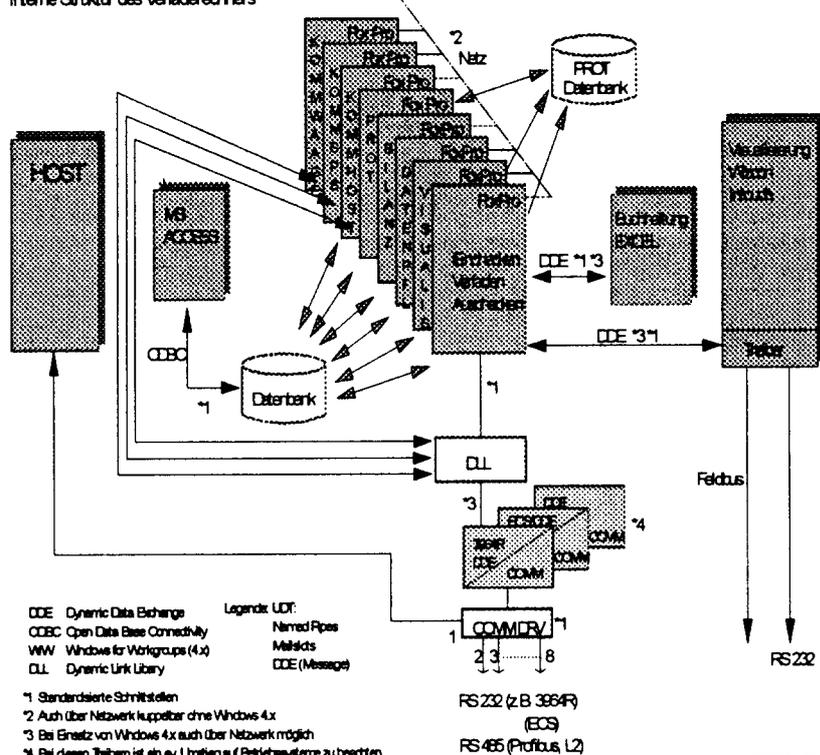
Eingabe und Anzeige von Kundendateien, Tankwagen-dateien etc.

EINCHECKEN

Erfassung des Tankwagen-fahrers mittels Codekarte und Einspeicherung (Disp-nierung) der gewünschten Produkte und Verlade-mengen.

Interne Struktur des Verladerechners

65194



BEFÜLLUNG

Abwicklung des Füllablaufes auf der Füllinsel (Schwerpunkt ist die Kommunikation mit der SPS) sowie Speicherung der Verladedaten.

AUSCHECKEN

Drucken der Verladepapiere und Speicherung der Daten für die Verrechnung.

PROTOKOLLIERUNG

Anzeige der Disponiervorgänge, Störmeldungen etc.

STATUSANZEIGE

Grafische Anzeige des Anlagenzustandes.

AMPELSTEUERUNG

Verkehrslenkung mittels Ampel. Die Auslastung der Anlage wird überwacht. Mittels produktzugeordneter Ampeln wird die Zufahrt zur Anlage geregelt.

PRODUKTINFORMATION

Verwaltung der Daten des Tanklagers (Produktspezifikationen). Gewährleistung der Verbindung vom Tank bis zum Füllplatz.

Die Vernetzung mit dem kaufmännischen System:

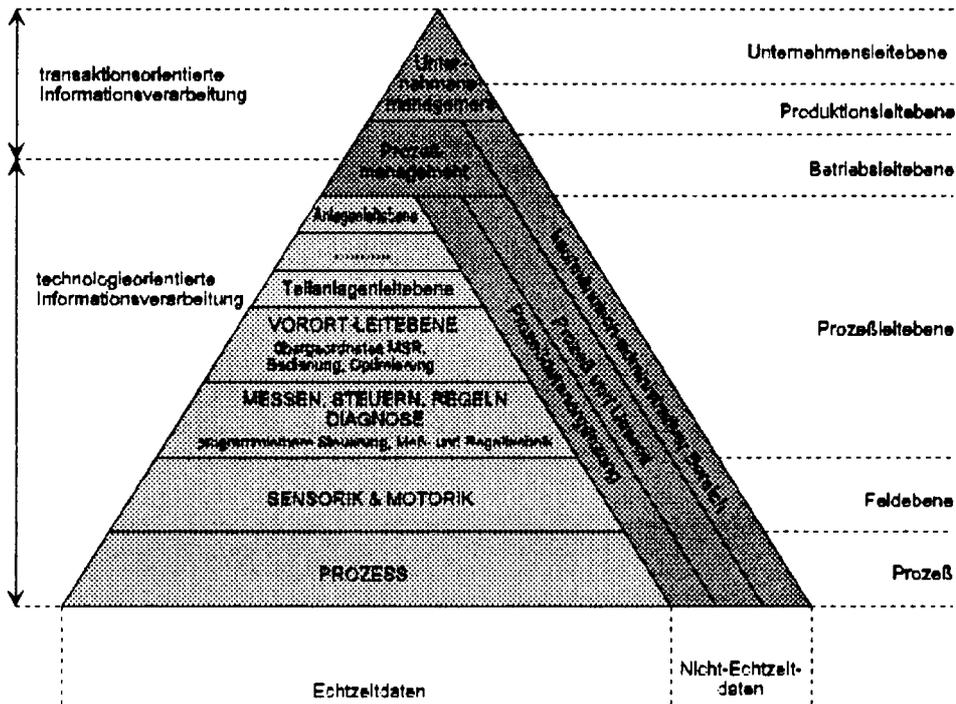
Alle Daten, die für die kaufmännische Verwaltung und Verrechnung der Abfüllvorgänge benötigt werden und vom technischen Systemteil generiert wurden, sind in das kaufmännische System zu übertragen. Weiters sind Daten aus dem kaufmännischen Bereich für die Freigabe der technischen Abläufe im technischen Systemteil notwendig. Die wesentliche Problematik bei diesem Datenaustausch ist der Umstand, daß kaufmännische Programme nicht echtzeitfähig sind, d.h. daß es keine garantierten Zeitintervalle gibt, nach deren Ablauf das Ergebnis der Berechnung vorliegt, sodaß bei der Kombination von echtzeitfähigen technischen Systemen mit nicht-echtzeitfähigen kaufmännischen Systemen der Konflikt der gegenseitigen Blockierung zu lösen ist.

Wie aus dem „Schichtenmodell“ hervorgeht, gibt es im Bereich der Automatisierung quasi zwei „Welten“ : Die transaktionsorientierte Informationsverarbeitung (in unserem Fall entspricht das der kaufmännischen Abarbeitung jedes Abfüllauftrags) und die technologieorientierte Informationsverarbeitung (in unserem Fall entspricht das der Abfüllautomatisierung). Wie aus dem „Schichtenmodell“ weiters ersichtlich ist, gibt es einen Bereich, wo die beiden „Welten“ mit ihren unterschiedlichen Datenqualitäten aufeinanderstoßen. Ob diese Problemsituationen gelöst werden können oder nicht, ist in jedem Einzelfall zu untersuchen. Es gibt Situationen, für die es keine Lösung gibt, d.h. wo ein Zusammenführen dieser beiden „Welten“ nicht konfliktfrei möglich ist. Einige Anforderungen, die es zu erfüllen gilt, sind:

- * Das sichere und richtige Erkennen und Melden, ob die Übernahme bzw. Übergabe von Daten von einer „Welt“ in die andere richtig und vollständig bzw. überhaupt stattgefunden hat, auch wenn in derartigen Übernahme- / Übergabe-Phasen einzelne Systemteile ausfallen bzw. ausgefallen sind. Schlimmer als ein erkanntes Fehlverhalten ist ein „verdecktes“ Fehlverhalten, das als „Wohlverhalten“ interpretiert werden kann.



Schichtenmodell



ELIN ENERGIEANWENDUNG - BEREICH ELECTRONIC



- * Verriegelungsmechanismen, die ein Verändern der Auftragsstammdaten während des vollautomatischen Abfüllvorgangs verhindern, sodaß weder im technischen noch im kaufmännischen Teil undefinierte Zustände eintreten können.
- * Verriegelungsmechanismen, die bei Ausfall von Systemteilen auch wieder ordnungsgemäß aufgelöst werden können.
- * Fehler müssen vermieden werden wie z. B.:
 - Belege für nicht oder nur teilweise ausgeführte Aufträge,
 - Fehlen von Belegen für ordnungsgemäß ausgeführte Aufträge.
- * Backtracingmechanismen, zumindest rudimentär, auch in der SPS.
- * im Rahmen generierbare Reaktionszeiten (echtzeitähnlich), auch im Bereich der kaufmännischen Transaktionen.

Literatur:

atp-Sonderheft NAMUR-Statusbericht '90, R.Oldenbourg-Verlag München, 1990

Fuzzy-Regler als glättender Regel-Interpolator

Josef Leichtfried und Michael Heiss

Manuskript eingelangt: 27. März 1995

Kurzfassung

Fuzzy-Regler können als Kennfeldregler betrachtet werden: Die Regeln geben Stützstellen vor, und das gewählte Fuzzy-Konzept legt die Art der Interpolation (bzw. Approximation) fest. Diese Interpolation soll den Erwartungen des Experten entsprechen, der die Regeln formuliert hat, also möglichst glatt (smooth), überschwingfrei, monotonieerhaltend und gegebenenfalls geradentreu sein. Durch diese Eigenschaften können jene Effekte vermieden werden, die bei einigen klassischen Fuzzy-Konzepten auftreten, jedoch bei der Formulierung der Regeln keinesfalls beabsichtigt waren.

Die hier präsentierte Arbeit kann auf zwei Arten gelesen werden: einerseits als Vorstellung eines neuen Fuzzy-Konzeptes, das die genannten hochgesteckten Forderungen erfüllt, und andererseits als Vorstellung eines neuen, mehrdimensionalen Interpolationsverfahrens, das diese Forderungen erfüllt.

Abstract

Fuzzy controllers can be seen as memoryless input-output maps. The rules represent interpolation nodes and the chosen fuzzy concept defines the way of interpolating or approximating these nodes. The interpolation should satisfy the expectations of the expert who defined the rules, and therefore be smooth, avoid overshooting, conserve monotony and conserve straight lines. These properties avoid effects which were never intended by the expert but which nevertheless are known from some classical fuzzy concepts.

The presented paper can be read in two ways: first, as a new fuzzy concept fulfilling the above requirements, and second, as a presentation of a new interpolation method fulfilling the above requirements.

1 Einleitung

Bei vielen Fuzzy-Regler Anwendungen wird aus wenigen Eingangsgrößen über eine vollständig ausgefüllte Matrix von Fuzzy-Regeln die Stellgröße ermittelt [1, 2, 3, 4, 5]. Ausgehend von diesem typischen Einsatzgebiet läßt sich allgemein aussagen:

Das Ein-Ausgangsverhalten eines Fuzzy-Reglers läßt sich durch ein Kennfeld beschreiben.

Ein Fuzzy-Regler ermittelt zu jeder Eingangsgrößenkombination $\mathbf{e} = (e_1, e_2, \dots, e_n)^T$ deterministisch und reproduzierbar einen Ausgangswert $u(\mathbf{e})$ [6, 1, 7, 8, 9, 10]. Ein Fuzzy-Regler ist daher in seinem Ein-Ausgangsverhalten äquivalent zu einer n -dimensionalen Regelkennfläche $u(\mathbf{e})$.

Jede Fuzzy-Regel läßt sich als Punkt in diesem Kennfeld darstellen.

Die Regelkennfläche wird bei Fuzzy-Reglern in Form von linguistischen WENN-DANN-Regeln beschrieben. Zugehörigkeitsfunktionen legen die Bedeutung der linguistischen Terme fest. Betrachtet man zunächst nur die typischen Werte der in den Regeln verwendeten Terme, also die Lage der Maxima der entsprechenden Zugehörigkeitsfunktionen, so läßt sich jede Regel als Punkt im n -dimensionalen Kennfeld betrachten [11, 12]. So erhält man mit der Regel „WENN e_1 =heiß (typ. 30°) UND e_2 =klein (typ. 4) DANN u =stark kühlen (typ. 500)“ einen Kennfeldpunkt $u(30, 4) = 500$.

Das gewählte Fuzzy-Konzept legt die Art der Interpolation dieser Punkte fest.

Die Form der Zugehörigkeitsfunktionen, der Inferenzmechanismus und die Defuzzifizierungsmethode definieren das sogenannte *Fuzzy-Konzept*. In Abhängigkeit vom gewählten Fuzzy-Konzept sind die Punkte des Kennfeldes definiert, die zwischen den durch die Regeln vorgegebenen Punkten liegen (siehe auch [13, 14]). Weiters ist es vom Fuzzy-Konzept abhängig, ob das Kennfeld durch die vorgegebenen Punkte geht oder nicht (Interpolation oder Approximation).

Diese Tatsachen können noch ergänzt werden durch Wunschvorgaben:

Die durch die Regeln dargestellten Punkte sollen auf dem Kennfeld liegen.

Zu diesem Wunsch gibt es geteilte Meinungen: Wenn zufällig alle Eingangsgrößen im Maximum der entsprechenden Zugehörigkeitsfunktionen liegen und somit genau eine Regel den aktuellen Zustand bestmöglich beschreibt, dann soll ausschließlich die Schlußfolgerung dieser einen Regel ausgeführt werden, wie dies bei Fuzzy Konzepten mit dreieckförmigen Zugehörigkeitsfunktionen meist selbstverständlich ist [1]. In obiger Notation bedeutet dies, daß der Punkt auf der Kennfläche liegen soll.

Es ist auch die Meinung vertretbar, daß obwohl im obigen Fall genau eine Regel auf den aktuellen Zustand zugeschnitten ist, trotzdem auch die Nachbarregeln einen Einfluß auf die Regelschlußfolgerung haben sollen. Die Punkte würden dann approximiert [11, Kap. 8] statt interpoliert werden. Leser mit der letzteren Meinung werden die hier vorgestellte Arbeit nur als neue Interpolationsmethode, und nicht als Beitrag zur Fuzzy-Regelung werten können.

Das Fuzzy-Konzept darf in seinem Interpolationsverhalten keine unerwünschten Effekte wie z.B. Überschwingen zeigen.

Obwohl der Experte linguistische Regeln formuliert hat, sind diese in unserer Betrachtungsweise nur Stützstellen eines Kennfeldes. Für die Qualität des Fuzzy-Reglers ist dabei von entscheidender Bedeutung, welches Fuzzy-Konzept gewählt wurde, wie sorgsam also diese Experteninformation genützt — oder konkret — interpoliert wird. Aufgabe unserer Arbeit ist es, diese Regeln so konservativ (informationserhaltend) wie möglich zu interpolieren, d.h. daß keine Effekte auftreten sollen, die vom Experten nicht beabsichtigt waren. Ein Überschwingen zwischen den Stützstellen würde beispielsweise bedeuten, daß obwohl die aktuelle Eingangsgrößenkombination e fast genau einer der Regeln entspricht, trotzdem eine Stellgröße $u(e)$ ausgegeben wird, die viel höher ist als die Schlußfolgerung dieser Regel, eventuell sogar höher als in irgendeiner der Regeln vorgeschlagen.

Aus obigen Aussagen ergibt sich für die Autoren folgende Schlußfolgerung:

Der Anwender sollte keinen Einfluß auf die Form der Zugehörigkeitsfunktionen haben, sondern nur die Lage der Maxima der Zugehörigkeitsfunktionen und natürlich die Regeln frei wählen dürfen.

Würde der Anwender die Form der Zugehörigkeitsfunktionen selbst verändern, ist die Gefahr zu groß, daß dabei unerwünschte (Interpolations-)Effekte auftreten. Er soll nur prinzipiell vorgeben, ob eine glatte Regelkennfläche notwendig ist oder nicht¹. In dieser Arbeit werden glatte Kennflächen durch Verwendung von B-Splines als Zugehörigkeitsfunktionen erreicht [17]. Anstatt auf die Form dieser fix vorgegebenen Zugehörigkeitsfunk-

¹Knicke im Regelkennfeld können beispielsweise zu unerwünschten Grenzyklen führen [15, 16].

tionen Einfluß zu nehmen, kann durch ein dichteres Legen von Regeln (=Stützstellen) jede glatte Regelkennfläche realisiert werden [18].

Der Experte formuliert in gewohnter Weise die linguistischen Regeln. Er braucht jedoch nicht mehr den gesamten Verlauf der Zugehörigkeitsfunktionen zu bestimmen, sondern nur noch die typischen Werte der linguistischen Terme anzugeben.²

Eine wesentliche Eigenschaft des vorzustellenden Konzeptes ist die Möglichkeit, ein n -dimensionales Kennfeld unter Beibehaltung der gewünschten Eigenschaften durch mehrmalige Berechnung des eindimensionalen Falles zu ermitteln.

Deswegen wird zunächst in den Abschnitten 2 bis 5 ausführlich der eindimensionale Fall behandelt, um dann im Abschnitt 6 nur noch die Erweiterung auf den n -dimensionalen Fall besprechen zu müssen.

Die eindimensionale Kennlinie wird aus Stücken zusammengesetzt. Zwischen zwei benachbarten Stützstellen ist die Kennlinie durch die Koordinaten der Stützstellen und den Anstiegen in den Stützstellen definiert. Da der Experte in Form von Regeln nur die Stützstellen definiert, soll der Fuzzy-Regler die Anstiege als Freiheitsgrade nutzen, um die Regeln in der erwähnten konservativen Weise zu interpolieren. Wie diese Anstiege berechnet werden, ist das zentrale Thema dieser Arbeit.

Ein Hinweis zur Notation: Abgesehen von den Termen $(2\xi - 3)^2$ und ξ^2 in Gl. (1), 4^n in Gl. (28), 3^n in Gl. (29) und der Gl. (30) werden in der gesamten Arbeit auch Hochzahlen zur Indizierung verwendet. So gibt zum Beispiel die Hochzahl k die Nummer der Regel an.

2 Grundkonzept (eindimensionaler Fall)

In [11] wurde ein normierter Fuzzy-Regler mit B-Spline Zugehörigkeitsfunktionen μ^k vorgestellt. Die B-Spline Zugehörigkeitsfunktionen sind stückweise aus quadratischen Polynomen zusammengesetzt und werden durch die glatte Funktion

$$\mu^k(e) = f_{\text{spline}}(\xi) \Big|_{(\xi = \frac{e - e^k}{b})} = \begin{cases} 0 & \text{für } \xi \geq 1.5 \\ \frac{1}{8}(2\xi - 3)^2 & \text{für } 0.5 < \xi < 1.5 \\ \frac{3}{4} - \xi^2 & \text{für } \xi \leq 0.5 \end{cases} \quad (1)$$

beschrieben, wobei e die Eingangsgröße, e^k der typische Wert der Zugehörigkeitsfunktion (jener Wert von e , bei dem das Maximum auftritt) und b der Abstand der gleichmäßig angeordneten Zugehörigkeitsfunktionen ist ($e^{k+1} = e^k + b$). Die gesamte Breite einer Zugehörigkeitsfunktion ist $3b$ (Abb. 1).

Als UND-Operator wird die Multiplikation eingesetzt, als ODER-Operator die Summe [20, 21]. Als Ausgangszugehörigkeitsfunktionen werden Singletons verwendet und zur Defuzzifizierung die Schwerpunktmethod. Die Zugehörigkeitsfunktionen sind so gewählt,

²Wie aus der Literatur ersichtlich ist [8, 1, 4, 3, 5, 19], war der Experte mit der Wahl des genauen Verlaufes der Zugehörigkeitsfunktionen überfordert, weshalb meist kein Gebrauch davon gemacht wurde und einfachheitshalber Dreiecke verwendet wurden, von denen auch nur die Lage der Maxima angegeben werden mußte.

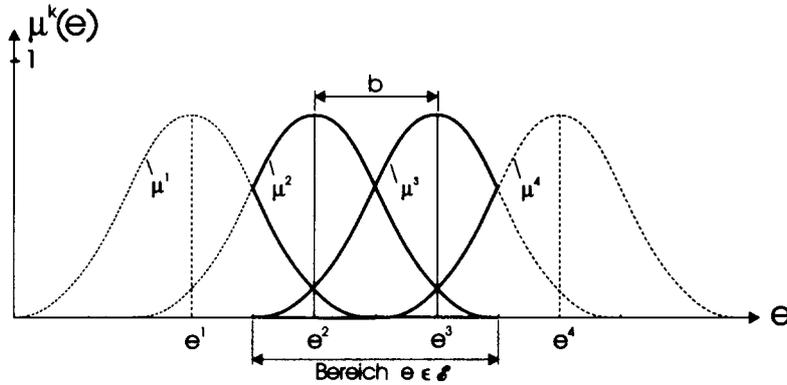


Abb. 1: B-Spline Zugehörigkeitsfunktionen nach (1).

daß deren Summe gleich eins ist (siehe Abb. 1, bzw. *normierter Fuzzy-Regler* in [11]). Im Fall von nur einer Eingangsgröße und einem vollständigen Regelsatz ist die Stellgröße

$$u(e) = \sum_{k=1}^L s^k f_{spline}\left(\frac{|e - e^k|}{b}\right) \quad \forall e \in \mathcal{E} \quad (2)$$

und somit eine Linearkombination der Zugehörigkeitsfunktionen, wobei L die Anzahl der Zugehörigkeitsfunktionen (für eine Eingangsgröße ident mit der Anzahl der Regeln), s^k die Schlußfolgerung der k -ten Regel und $\mathcal{E} = \left\{e \in \mathbf{R} \mid \left\{\frac{e^1 + e^2}{2} \leq e \leq \frac{e^{L-1} + e^L}{2}\right\}\right\}$ der Definitionsbereich ist. Die Einschränkung der Eingangsgröße auf $e \in \mathcal{E}$ ist notwendig, da nur in diesem Bereich die Summe der Zugehörigkeitsfunktionen gleich eins ist.

Der Fall $L = 4$ mit den Stützstellen $S^1(e^1, s^1)$, $S^2(e^2, s^2)$ und $S^3(e^3, s^3)$, $S^4(e^4, s^4)$ kann als Regler-Grundelement angesehen werden. Die Regler-Kennlinie kann aus solchen Grundelementen zusammengesetzt werden. Wie noch gezeigt wird, können Kennlinien auch dann aus diesem Grundelement zusammengesetzt werden, wenn die Stützstellen nicht äquidistant liegen.

Wie in [11] gezeigt, würde durch (2) eine Kennlinie entstehen, die im allgemeinen nicht durch die Stützstellen verläuft. Im hier besprochenen Fall ist es jedoch wünschenswert, daß die vorgegebenen Stützstellen (=Regeln) wirklich interpoliert und nicht nur approximiert werden, die Kennlinie also durch die Stützstellen hindurch geht. Andererseits würde man gerne die einfache Form von Gleichung (2) beibehalten. Mit einem Trick ist beides möglich:

In [11] wurde nämlich weiters gezeigt, daß für den Spezialfall, daß eine Teilmenge der Stützstellen auf einer Geraden (allg. Hyperebene) liegt, jener Teil der Kennlinie (allg. Kennfeld) auch auf dieser Geraden (Hyperebene) liegt, der ausschließlich aus dieser Teilmenge von Regeln (Stützstellen) berechnet wird [11, Satz 6].

Es mag zunächst so klingen, als hätte dieser Satz keine praktische Bedeutung, denn wer will schon Geraden-Kennlinien darstellen. Andererseits kann man behaupten, daß zwei benachbarte Punkte, z.B. S^1 und S^2 , auf derselben Geraden, nämlich ihrer Verbindungsgeraden, liegen. Aus (1), oder besser Abb. 1 ist ersichtlich, daß die Reichweite der Zugehörigkeitsfunktionen vom Zentrum aus nur $1,5 b$ in beide Richtungen ist und daher obiger Satz nur für einen Punkt sicher anwendbar ist: Genau in der Mitte zwischen S^1 und S^2 gilt, daß der zugehörige Kennfeldpunkt ausschließlich aus den Stützstellen S^1 und S^2 berechnet wird, da alle anderen Zugehörigkeitsfunktionen an dieser Stelle Null sind.

Der dort berechnete Kennlinienpunkt muß deshalb auf der Verbindungsgeraden liegen. Definiert man daher für das Grundelement entsprechend Abb. 2 zwei Halbierungspunkte $H^a(eH^a, sH^a, kH^a)$ und $H^b(eH^b, sH^b, kH^b)$ mit den Koordinaten eH und sH sowie den Anstiegen³ kH ,

$$eH^a = \frac{e^1 + e^2}{2}, \quad sH^a = \frac{s^1 + s^2}{2}, \quad kH^a = \frac{s^2 - s^1}{e^2 - e^1} \quad (3)$$

und

$$eH^b = \frac{e^3 + e^4}{2}, \quad sH^b = \frac{s^3 + s^4}{2}, \quad kH^b = \frac{s^4 - s^3}{e^4 - e^3}, \quad (4)$$

dann muß gelten, daß die Halbierungspunkte H^a, H^b auf der Kennlinie

$$u(eH^a) = sH^a \text{ bzw. } u(eH^b) = sH^b \quad (5)$$

liegen. Ebenso kann gezeigt werden, daß in den Halbierungspunkten die Anstiege der Verbindungsgerade mit jenen der Kennlinie

$$\left. \frac{du(e)}{de} \right|_{e=eH^a} = kH^a \text{ bzw. } \left. \frac{du(e)}{de} \right|_{e=eH^b} = kH^b \quad (6)$$

übereinstimmen (siehe Abb. 2).

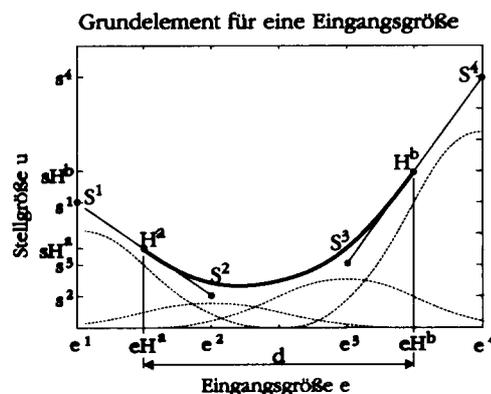


Abb. 2: Fuzzy-Regler Grundelement, das durch die beiden Halbierungspunkte H^a, H^b und den Anstieg in diesen Halbierungspunkten festgelegt ist. Strichliert sind die mit s^k multiplizierten Zugehörigkeitsfunktionen eingetragen, die aufsummiert die Kennlinie ergeben. Auf den Achsen sind die Koordinaten der Stützpunkte S und H aufgetragen.

Der Trick ist nun, die Betrachtungen umzukehren: Dem Anwender gegenüber werden nur noch die Halbierungspunkte H sichtbar gemacht, die ab nun als Stützstellen H bezeichnet werden; eH^k ist dann der typische Wert der k -ten Zugehörigkeitsfunktion und sH^k die Schlußfolgerung der k -ten Regel des Fuzzy-Reglers. Aus den Stützstellen H werden passende Punkte S abgeleitet (offline), und mit den Punkten S wird die Ausgangsgröße

³Man beachte, daß in die Definition der Halbierungspunkte auch der Anstieg aufgenommen wurde.

$u(e)$ berechnet (online), wodurch der Vorteil der einfachen Berechnung nach (2) erhalten bleibt.

Bei vorgegebenen Stützpunkten H und zugehörigen Anstiegen können in Umkehrung der Gleichungen (3) und (4) die Punkte S^1 bis S^4 mit der Abkürzung $d = eH^b - eH^a = 2b$ als

$$e^1 = eH^a - \frac{d}{4}, \quad e^2 = eH^a + \frac{d}{4}, \quad e^3 = eH^b - \frac{d}{4}, \quad e^4 = eH^b + \frac{d}{4} \quad (7)$$

und

$$s^1 = sH^a - \frac{d}{4}kH^a, \quad s^2 = sH^a + \frac{d}{4}kH^a, \quad s^3 = sH^b - \frac{d}{4}kH^b, \quad s^4 = sH^b + \frac{d}{4}kH^b \quad (8)$$

berechnet werden. Mit der Rechteckfunktion

$$\text{rec}(e, H^a, H^b) = \begin{cases} 1 & \text{für } eH^a < e \leq eH^b \\ 0 & \text{sonst} \end{cases}, \quad (9)$$

die nur innerhalb der beiden Stützpunkte H^a und H^b gleich 1 ist, kann nun das Fuzzy-Regler Grundelement

$$f_H(e, H^a, H^b) = \text{rec}(e, H^a, H^b) \sum_{k=1}^4 s^k f_{\text{spline}}\left(\frac{e - e^k}{d/2}\right) \quad (10)$$

definiert werden, das durch die beiden Stützpunkte H^a , H^b festgelegt ist (siehe Abb. 2), wobei e^k und s^k mit (7) und (8) aus H^a und H^b berechnet werden.

Aus einer beliebigen (auch ungleichmäßig verteilten) Folge von $N + 1$ Stützpunkten H^k mit $eH^k < eH^{k+1}$ kann mit

$$u(e) = \sum_{k=1}^N f_H(e, H^k, H^{k+1}) \quad \forall e \in \mathbf{R} | \{eH^1 < e \leq eH^{N+1}\} \quad (11)$$

die Kennlinie $u(e)$ durch Zusammensetzen aus N Teilstücken berechnet werden, wobei die Stützpunkte H auf der Kennlinie liegen und die Anstiege kH wesentlichen Einfluß auf die Kennlinie haben (siehe Abb. 3).

Die Anstiege kH der Stützpunkte H sind Freiheitsgrade und sollen so gewählt werden, daß die Kennlinie des Fuzzy-Reglers die gewünschten konservativen Eigenschaften erhält. Damit der Anwender nicht selbst diese Anstiege wählen muß, wird im folgenden ein Algorithmus vorgestellt, der eine automatische Berechnung der Anstiege kH ermöglicht [22].

Die Breite der Zugehörigkeitsfunktionen dieses Fuzzy-Reglers ist durch die Abstände zu den benachbarten Regeln (Stützpunkten H) bestimmt. Bereiche starker Änderungen im Regelverhalten können daher durch eng beisammenliegende Begriffe beschrieben werden, und in Bereichen, in denen sich wenig ändert, genügen weiter „entfernt“ liegende Begriffe. Soll der Fuzzy-Regler auch außerhalb des Definitionsbereiches sinnvolle Werte liefern, so kann dies durch ein automatisches Generieren von zusätzlichen Regeln am Rand erreicht werden, die so berechnet werden, daß das Regelkennfeld wunschgemäß extrapoliert wird.

Zusammenfassung: Der Experte formuliert in gewohnter Weise sein Prozeßwissen in Form von linguistischen Regeln. Auch schon bestehende Regelsätze können verwendet werden. Ohne daß der Experte davon zu wissen braucht, werden seine Regeln als Stützpunkte H im Regelkennfeld betrachtet. Zur Vereinfachung der numerischen Auswertung werden intern die Stützpunkte H in Punkte S umgerechnet.

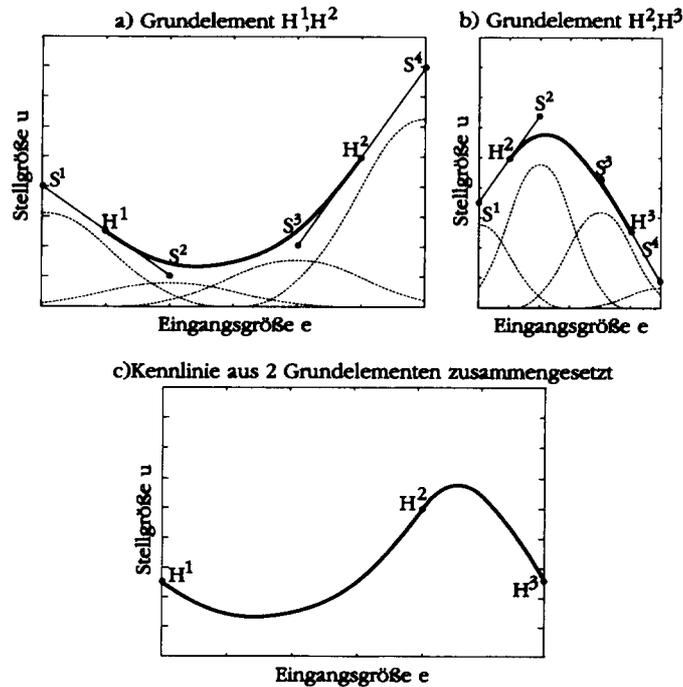


Abb. 3: Zusammensetzen eines Fuzzy-Reglers mit 3 Stützpunkten H aus zwei Grundelementen. In a) und b) sind die beiden Grundelemente, die im Stützpunkt H^2 gleichen Anstieg haben, dargestellt. In c) ist die Kennlinie des Fuzzy-Reglers dargestellt, der aus der Kombination dieser beiden Grundelemente besteht.

3 Geforderte Kennlinieneigenschaften

Um einen Fuzzy-Regler zu erhalten, der die durch Experten formulierten Regeln im erwähnten Sinn konservativ interpoliert, wird der Term „konservativ“ durch die gewünschten Eigenschaften im folgenden definiert. Die gewünschten Eigenschaften sind durch geeignete Wahl der Anstiege kH realisierbar.

3.1 Monotonie (kein Überschwingen):

Der Mensch neigt dazu, die Extremwerte des Regelverhaltens durch verbale Regeln zu charakterisieren. Es erscheint daher sinnvoll, die Regler-Kennlinie im Bereich zwischen zwei benachbarten Stützpunkten (Regeln) auf monotone Funktionen zu beschränken, so daß die Extremwerte nur an Stützpunkten H auftreten können. Dadurch erhält man einen Algorithmus, der ein Überschwingen der Kennlinie vermeidet. Wird der Anstieg der Verbindungsgeraden von H^k und seinem rechten Nachbarn von H^{k+1}

$$kr^k = \frac{sH^{k+1} - sH^k}{eH^{k+1} - eH^k} \quad \forall k \in \{1 \dots N\} \quad (12)$$

genannt, lautet die geforderte Monotonieeigenschaft:

$$\frac{du(e)}{de} \left\{ \begin{array}{ll} \geq 0 & \text{wenn } kr^k > 0 \\ = 0 & \text{wenn } kr^k = 0 \\ \leq 0 & \text{wenn } kr^k < 0 \end{array} \right\} \quad \text{für } \begin{array}{l} eH^k < e \leq eH^{k+1} \\ \forall k \in \{1 \dots N\} \end{array} \quad (13)$$

3.2 Parabeltreue

Die Parabeltreue gewährleistet eine exakte Darstellung linearer und quadratischer Kennlinien(abschnitte) und eine gute Modellierung eines weiten Bereiches praktisch vorkommender Kennlinienverläufe. Da die B-Spline Funktionen stückweise durch quadratische Polynome beschrieben werden, kann auch die aus B-Splines zusammengesetzte Kennlinie stückweise durch quadratische Polynome beschrieben werden. Es kann einfach gezeigt werden, daß ein quadratisches Polynom mit beliebigen Koeffizienten a , b und c bei entsprechender Wahl von zwei Stützpunkten H^k , H^{k+1} durch einen Fuzzy-Regler nach (11) beschrieben werden kann, sodaß

$$u(e) = a + be + ce^2 = f_H(e, H^k, H^{k+1}) \quad \text{für } eH^k < e \leq eH^{k+1} \quad (14)$$

gilt.

3.3 Erweiterte Geradentreue

Durch eine zusätzliche Schranke für den Anstieg,

$$\min(kr^{k-1}, kr^k) \leq kH^k \leq \max(kr^{k-1}, kr^k), \quad (15)$$

wird ein „Schwingen“ der Kennlinie im Randbereich eines geraden Kennlinienabschnittes vermieden (siehe Abb. 4). Liegen drei Stützpunkte H^{k-1} , H^k und H^{k+1} auf einer Geraden, so ist mit (15) $kH^k = kr^{k-1} = kr^k$. Es genügt also, daß drei benachbarte Punkte auf einer Geraden liegen, daß im Mittleren der drei Punkte diese Gerade die Tangente an die Kennlinie ist.

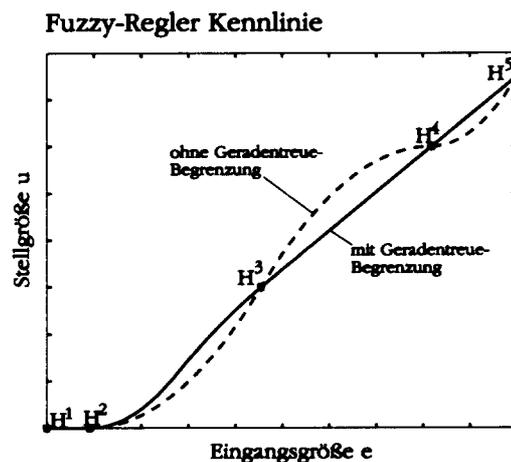


Abb. 4: Vergleich mit und ohne Geradentreue-Begrenzung. Die Stützpunkte H^2 bis H^5 liegen auf einer Geraden. Mit Geradentreue-Begrenzung (15) schmiegt sich die Kennlinie ab H^3 an diese Gerade. Ohne Geradentreue-Begrenzung entsteht ein Schwingen der Kennlinie, da der Parabelinterpolations-Algorithmus für die Berechnung des Anstieges im Stützpunkt H^3 den Stützpunkt H^4 nicht berücksichtigt (siehe Abschnitt 5.2).

4 Klassifizierung der Stützpunkte H

Um die Untersuchung beliebiger Stützpunktfolgen H zu vereinfachen, wird jeder Stützpunkt H^k entsprechend der Definition

$$Typ(H^k) = \begin{cases} \text{Rp (Randpunkt)} & \text{wenn } k = 1 \text{ oder } k = N + 1 \\ \text{Ex (Extremwert)} & \text{wenn } Typ(H^k) \neq \text{Rp und } sign(kr^{k-1}) \neq sign(kr^k) \\ \text{Ss (Steigung steigend)} & \text{wenn } Typ(H^k) \neq (\text{Rp oder Ex}) \text{ und } |kr^{k-1}| \leq |kr^k| \\ \text{Sf (Steigung fallend)} & \text{sonst} \end{cases}, \quad (16)$$

genau einem Typ zugeordnet. Beispielsweise gilt für Abb. 4, daß $Typ(H^1) = Typ(H^5) = \text{Rp}$, $Typ(H^2) = \text{Ex}$, $Typ(H^3) = Typ(H^4) = \text{Ss}$.

5 Berechnung der Anstiege (kH -Algorithmus)

Zur Berechnung der Anstiege werden zuerst die Betragsminima des Kennlinienanstieges gesucht und anschließend, ausgehend von den Anstiegsminima, die restlichen Anstiege berechnet. Die im Abschnitt 3 gestellten Forderungen sind hier integriert und müssen nicht getrennt berücksichtigt werden.

5.1 Festlegung der Betragsminima des Anstieges

Die Betragsminima des Anstieges einer Funktion müssen entweder Extremwerte, Wendepunkte oder Randpunkte der Funktion sein.

5.1.1 Betragsminima in den Extremwerten

Ist der Stützpunkt H^k ein Extremwert ($Typ(H^k) = \text{Ex}$), so muß auf Grund der Monotonieforderung der Kennlinienanstieg kH^k im Punkt H^k gleich Null sein:

$$kH^k = 0 \text{ wenn } Typ(H^k) = \text{Ex}. \quad (17)$$

Durch (17) wird die Geradentreue nicht verletzt, jedoch hat diese Festlegung entscheidende Auswirkungen auf die Parabeltreue (siehe Abschnitt 5.2).

5.1.2 Betragsminima in Wendepunkten

Wenn die Kennlinie einen Wendepunkt aufweist, in dem der Anstiegsbetrag ein Minimum hat, dann wird der Anstieg im Wendepunkt gleich dem Anstieg der Verbindungsgeraden der benachbarten Stützpunkte gewählt. Für diese Wendepunkte (es gibt ja auch solche mit Betragsmaximum des Anstieges) gilt

$$kH^k = kH^{k+1} = kr^k \text{ wenn } Typ(H^k) = \text{Sf und } Typ(H^{k+1}) = \text{Ss}. \quad (18)$$

Auch durch (18) wird die Geradentreue nicht verletzt. Einfluß auf Parabeltreue wie Gleichung (17).

5.1.3 Betragsminima in Randpunkten

Der linke Randpunkt H^1 weist dann ein Betragsminimum des Anstieges auf, wenn H^2 vom Typ Steigung steigend ist. In diesem Fall wird der Anstieg kH^1 mittels einer Parabelinterpolation festgelegt. Es gilt

$$kH^1 = \begin{cases} \frac{kr^1(2d^2 + d^1) - kr^2 d^2}{d^1 + d^2} & \text{für } |kr^1(2d^2 + d^1)| > |kr^2 d^2| \text{ wenn } Typ(H^2) = \text{Ss} \\ 0 & \text{sonst} \end{cases} \quad (19)$$

wobei $d^k = eH^{k+1} - eH^k$ der Abstand zum rechten Nachbarpunkt ist. Die Fallunterscheidung in (19) ist auf Grund der Monotonie-Bedingung notwendig. Entsprechend gilt für den rechten Randpunkt H^{N+1} :

$$kH^{N+1} = \begin{cases} \frac{kr^N(2d^{N-1} + d^N) - kr^{N-1} d^{N-1}}{d^N + d^{N-1}} & \text{für } |kr^N(2d^{N-1} + d^N)| > |kr^{N-1} d^{N-1}| \\ 0 & \text{sonst} \end{cases} \quad (20)$$

wenn $Typ(H^N) = \text{Sf}$.

5.2 Berechnung der restlichen Anstiege ausgehend von den Betragsminima des Anstieges

Ausgehend von diesen Anstiegs-Betragsminima nach (17), (18), (19) und (20), kann ein erster Teil der Anstiege in den restlichen Stützpunkten mit

$$kH^k = \begin{cases} 2kr^{k-1} - kH^{k-1} & \text{für } |2kr^{k-1} - kH^{k-1}| < |kr^k| \\ kr^k & \text{sonst} \end{cases} \quad (21)$$

wenn $Typ(H^k) = (Ss \text{ oder } Rp)$ und kH^{k-1} bereits berechnet

berechnet werden. Die dann noch fehlenden Anstiege können mit

$$kH^k = \begin{cases} 2kr^k - kH^{k+1} & \text{für } |2kr^k - kH^{k+1}| < |kr^{k-1}| \\ kr^{k-1} & \text{sonst} \end{cases} \quad (22)$$

wenn $Typ(H^k) = (Sf \text{ oder } Rp)$ und kH^{k+1} bereits berechnet

berechnet werden. Da zu diesem Zeitpunkt der Anstieg zumindest in einem der Nachbarpunkte bereits festliegt, ist das Parabelstück durch den Punkt H^k , den Nachbarpunkt und dessen Anstieg festgelegt und nicht vom anderen Nachbarpunkt abhängig. Durch wiederholtes Anwenden von (21) bzw. (22) kann jeweils aus dem Stützpunkt mit festgelegtem Anstieg der Anstieg im Nachbarpunkt berechnet werden. Da dieser Algorithmus nur den festgelegten Nachbarpunkt zur Berechnung benützt, ist er in Richtung des nicht festgelegten Punktes sozusagen „fast blind“. Er weiß nur, daß der Anstieg in Richtung dieses Punktes größer wird. Durch die mittels Fallunterscheidung in (21) und (22) integrierte Geradentreue-Begrenzung sind die unerwünschten Folgen dieser „Blindheit“ ausgeschlossen (siehe Abb. 4).

Es läßt sich zeigen, daß auf Grund der Monotonie und der Zusammensetzung der Kennlinienstücke aus jeweils vier B-Spline Funktionen das Betragsmaximum des Kennlinienanstieges nie größer als das Doppelte des Anstieges der Verbindungsgeraden der benachbarten Stützpunkte H^k, H^{k+1} sein kann. Dies ist wieder ein Zeichen für die Konservativität der Interpolationsmethode.

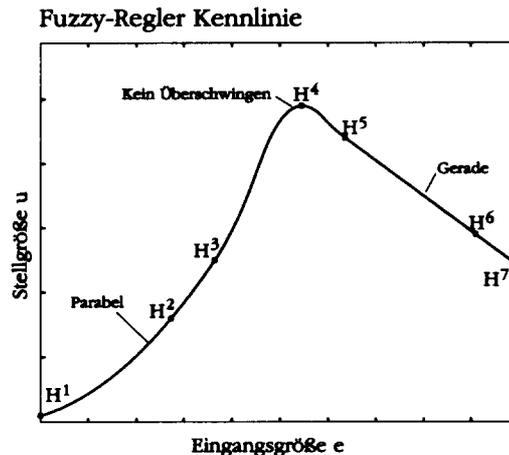


Abb. 5: Glatte, stückweise monotone, überschwingfreie, geraden- und parabeltreue Regelinterpolation mit 100%-igem Erreichen der ungleichmäßig angeordneten Stützstellen H .

Mit (17) bis (22) sind alle Anstiege kH^k eindeutig bestimmt. Mit (7) und (8) können aus den Stützpunkten H^k die Stützpunkte S^k der entsprechenden Grundelemente berechnet werden. Diese Berechnung kann offline, also vor der Inbetriebnahme des Reglers, erfolgen. Sind die Stützstellen S^k der Grundelemente berechnet, so benötigt der Fuzzy-Regler nur (11) bzw. (10) zur Berechnung der Stellgröße, wodurch eine zeiteffiziente Berechnung gewährleistet ist⁴ (Beispiel: siehe Abb. 5).

⁴Eine noch effizientere Berechnung ist möglich, wenn die Stützpunkte H äquidistant gelegt werden. Dann würden nämlich im Gegensatz zum Abb. 3 S^3, S^4 von einem Grundelement und S^1, S^2 vom nächsten Grundelement zusammenfallen und die zugehörigen Zugehörigkeitsfunktionen müßten jeweils nur einmal ausgewertet werden.

6 n -dimensionaler Fall

Für den eindimensionalen Fall gilt, daß der Halbierungspunkt H von zwei benachbarten Punkten S immer auf der Kennlinie liegt. Im Fall von zwei Eingangsgrößen gilt ebenso auf Grund der begrenzten Reichweite der Zugehörigkeitsfunktionen, daß der Mittelpunkt H eines durch Stützstellen S aufgespannten räumlichen Rechtecks laut Satz 6 aus [11] auf der Kennfläche liegt.

Auch hier kann wieder die umgekehrte Betrachtungsweise verwendet werden: Durch die Regeln werden direkt die Mittelpunkte H vorgegeben. Zu jedem Mittelpunkt H werden die vier Eckpunkte S des umgebenden Rechtecks als Hilfspunkte für die interne Darstellung ermittelt. Ein Grundelement (Abb. 6) wird aus 4 Mittelpunkten H , also aus $4 \times 4 = 16$ Punkten S zusammengesetzt.

Grundelement für 2 Eingangsgrößen

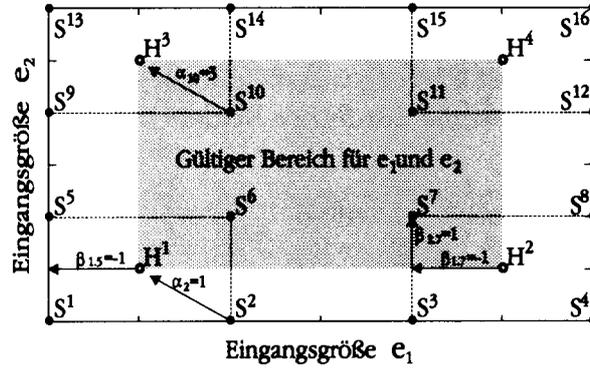


Abb. 6: Grundelement für zwei Eingangsgrößen. Durch die Stützpunkte H^1 bis H^4 und die Tangentialebenen in diesen Stützpunkten sind die Stützpunkte S^1 bis S^{16} , die zur Stellgrößenberechnung verwendet werden, festgelegt.

Im allgemeinen Fall von n Eingangsgrößen benötigt man für ein Grundelement 2^n Mittelpunkte H und jeder dieser Mittelpunkte wird intern durch 2^n Punkte S repräsentiert. Ein Grundelement enthält daher $2^n \times 2^n = 4^n$ Punkte S .

Analog zum eindimensionalen Fall werden auch hier die Mittelpunkte, die ja durch die Regeln des Fuzzy-Reglers vorgegeben werden, als Stützpunkte H^j bezeichnet. Im Fall von n Eingangsgrößen besitzt ein Stützpunkt $H^j(eH^j, sH^j, kH^j)$ die Koordinaten $eH^j = (eH_1^j \dots eH_n^j)^T$ und sH^j . Die Tangentialebene (Hyperebene) der Kennfläche ist durch die Ableitungen $kH^j = (kH_1^j \dots kH_n^j)^T$ in Richtung der Eingangsgrößen $e_1 \dots e_n$ bestimmt. In jedem Stützpunkt H^j gilt für die Kennfläche $u(e)$ analog zu (5) und (6):

$$u(eH^j) = sH^j \quad \forall j \in \{1 \dots 2^n\} \quad (23)$$

und

$$\left. \frac{\partial u(e)}{\partial e_i} \right|_{e=eH^j} = kH_i^j \quad \forall i \in \{1 \dots n\}, j \in \{1 \dots 2^n\}. \quad (24)$$

Ähnlich zu (7) können die Zentren e_i^k der Zugehörigkeitsfunktionen allgemein mit

$$e_i^k = eH_i^{\alpha_k} + \frac{d_i}{4}\beta_{i,k} \quad \forall k \in \{1 \dots 4^n\} \quad (25)$$

berechnet werden, wobei $\alpha_k \in \{1 \dots 2^n\}$ den Index j jenes Stützpunktes H^j angibt, der zum Stützpunkt S^k gehört (z.B. Abb. 6: $\alpha_{10} = 3$ zeigt an, daß S^{10} dem Stützpunkt H^3 zugeordnet ist) und $\beta_{i,k} \in \{-1, 1\}$ angibt, ob der Punkt S^k von H^j aus gesehen in positiver ($\beta_{i,k} = 1$) oder negativer ($\beta_{i,k} = -1$) Richtung der i -ten Eingangsgröße liegt (siehe Abb. 6). Ähnlich zu (8) können die Regelschlußfolgerungen s^k mit

$$s^k = sH^{\alpha_k} + \sum_{i=1}^n \frac{d_i}{4}\beta_{i,k} kH_i^{\alpha_k} \quad \forall k \in \{1 \dots 4^n\} \quad (26)$$

aus den Stützpunkten H^j berechnet werden. Definiert man

$$\mathcal{E}_i = \{e_i \in \mathbf{R} \mid \min_{j=1 \dots 2^n} (eH_i^j) \leq e_i \leq \max_{j=1 \dots 2^n} (eH_i^j)\}, \quad (27)$$

dann läßt sich der Gültigkeitsbereich des durch die Punkte $S^1 \dots S^{4^n}$ beschriebenen Grundelement des Fuzzy-Reglers mit $\mathcal{E}^n = \mathcal{E}_1 \times \mathcal{E}_2 \times \dots \times \mathcal{E}_n$ beschreiben. Innerhalb dieses Grundelementes läßt sich die Kennfläche des Fuzzy-Reglers durch

$$u(\mathbf{e}) = \sum_{k=1}^{4^n} \left(s^k \prod_{i=1}^n f_{\text{spline}}\left(\frac{e_i - e_i^k}{d_i/2}\right) \right) \quad \forall \mathbf{e} \in \mathcal{E}^n \quad (28)$$

berechnen.

6.1 Anstiegsberechnung im n -dimensionalen Fall

Die Berechnung aller Anstiege kH_i^j ($i = 1 \dots n$) für jeden der Stützpunkte H^j kann auf die Anstiegsberechnung im eindimensionalen Fall zurückgeführt werden. Zum besseren Verständnis sei zunächst der zweidimensionale Fall ($n = 2$) am Beispiel vom Abb. 7 erklärt. Es wird dabei der Begriff *Gitterlinie* verwendet. Gitterlinien sind Linien, entlang denen die Stützstellen H angeordnet sind. Jene Gitterlinien, die parallel zur Koordinatenachse e_i sind, werden als e_i -Gitterlinien bezeichnet (siehe Abb. 7b strichlierte Linien).

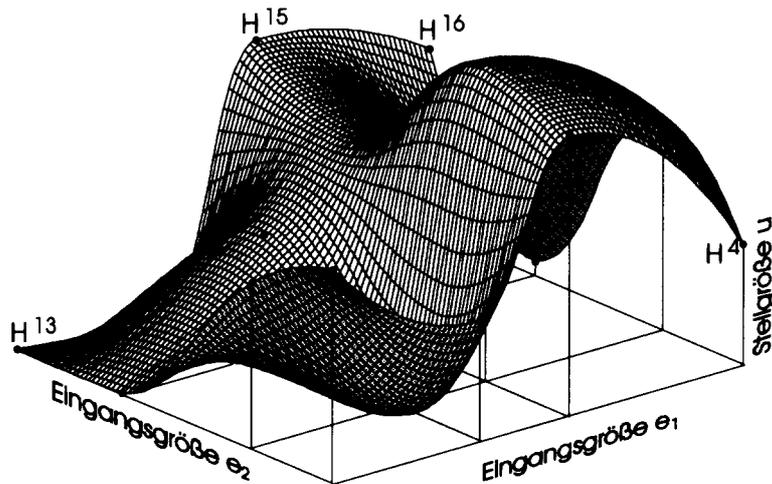
- Zuerst werden entlang der ersten e_1 -Gitterlinie, auf der die Stützpunkte H^1, H^2, H^3, H^4 liegen, die e_1 -Anstiege ($kH_1^1, kH_1^2, kH_1^3, kH_1^4$) mit dem eindimensionalen kH -Algorithmus ($e = e_1$) berechnet. Danach werden in gleicher Weise die e_1 -Anstiege kH_1^j entlang aller anderen e_1 -Gitterlinien berechnet (im Abb. 7: $kH_1^5, kH_1^6, kH_1^7, kH_1^8$, danach $kH_1^9, kH_1^{10}, kH_1^{11}, kH_1^{12}, \dots$).
- Dann werden entlang der e_2 -Gitterlinien wieder mit dem eindimensionalen kH -Algorithmus ($e = e_2$) die e_2 -Anstiege kH_2^j berechnet (im Abb. 7: $kH_2^1, kH_2^2, kH_2^3, kH_2^4$, dann $kH_2^5, kH_2^6, kH_2^{10}, kH_2^{14}, \dots$).

Bei n Eingangsgrößen muß in gleicher Weise fortgefahren werden, bis entlang aller Gitterlinien die Anstiege berechnet wurden und somit alle Anstiege kH_i^j bekannt sind.

6.2 Kennflächenberechnung

Um die Kennfläche wie bei einer Eingangsgröße aus mehreren Grundelementen zusammensetzen zu können, muß für einen Eingangsgrößenvektor \mathbf{e} der Bereich \mathcal{E}^n und somit die beteiligten (benachbarten) Stützpunkte H^j ermittelt werden. Anschließend können mit den berechneten Anstiegen und (25), (26) für jeden Bereich \mathcal{E}^n die Stützstellen S^k berechnet werden. Mit (28) kann dann die Stellgröße des Fuzzy-Reglers in Abhängigkeit von den Eingangsgrößen \mathbf{e} berechnet werden.

a) Fuzzy-Regler Kennfläche



b) Konturbild der Reglerkennfläche

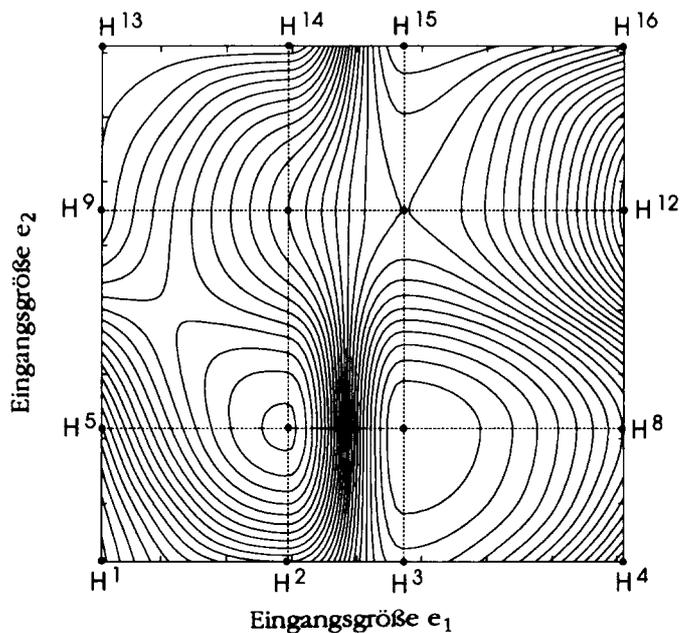


Abb. 7: Kennfläche eines Fuzzy-Reglers mit 2 Eingangsgrößen und 16 Regeln (Stützpunkte H). In b) ist das Konturbild (Linien konstanter Stellgröße) der in a) dargestellten Kennfläche zu sehen. Im Konturbild werden die Glattheit und die konservativen Interpolationseigenschaften des Fuzzy-Konzeptes besonders deutlich.

6.3 Berechnungsaufwand für n Eingangsgrößen

Da für jeden Wert einer Eingangsgröße e_i nur 3 Zugehörigkeitsfunktionen ungleich Null sein können, genügt die Berechnung von $3n$ Zugehörigkeitsfunktionen. Die dafür notwendige Zeit ist $3n t_{MF}$. Wenn die

Schlußfolgerungen s^k schon vorab berechnet wurden, läßt sich die Rechenzeit für die Stellgröße nach (28) mit

$$t_{\text{rechen}} \approx 3n t_{MF} + 3^n t_{\text{addition}} + n 3^n t_{\text{multipl.}} \quad (29)$$

abschätzen. Es fällt auf, daß die Rechenzeit nur von der Anzahl der Eingangsgrößen, nicht aber von der Anzahl der Zugehörigkeitsfunktionen einer Eingangsgröße abhängt. Für den Fall von zwei Eingangsgrößen sind somit nur 6 Zugehörigkeitsfunktionswerte, 9 Additionen und 18 Multiplikationen zu berechnen. Zusätzlich ist noch ein Aufwand notwendig, um den Bereich \mathcal{E}^n zu bestimmen.

Zur Reduktion des Rechenaufwandes kann in Anlehnung an die Vorgangsweise von [4, 3] auch hier der Eingangsgrößenbereich in Unterbereiche aufgeteilt werden, in denen sich (28) und (1) geschlossen auswerten lassen. Im zweidimensionalen Fall müßte dann jedes Grundelement noch weiter auf 16 Unterbereiche aufgeteilt werden, und für jeden Unterbereich gilt dann

$$u = a_1 + a_2 e_1 + a_3 e_2 + a_4 e_1 e_2 + a_5 e_1^2 + a_6 e_2^2 + a_7 e_1^2 e_2 + a_8 e_1 e_2^2 + a_9 e_1^2 e_2^2 \quad (30)$$

als Regelgesetz mit den für jeden Unterbereich konstanten Parametern a_j . Es sind dann nur noch 8 Additionen und 14 Multiplikationen notwendig, um die Stellgröße u zu berechnen, jedoch ein höherer Verwaltungs- und Speicheraufwand.

Für besonders rechenzeitkritische Anwendungen kann bei niedrigen Dimensionen n noch ein Schritt weiter gegangen werden und für jede im Rahmen der Quantisierung mögliche Eingangsgrößenkombination e der zugehörige Kennfeldpunkt $u(e)$ direkt abgespeichert werden [23]. In diesem Fall sind dann gar keine Berechnungen mehr notwendig.

7 Zusammenfassung

Es wurde gezeigt, daß ein Fuzzy-Regler als Interpolator der durch Regeln vorgegebenen Stützstellen betrachtet werden kann. Ein Fuzzy-Konzept wurde vorgestellt, das folgende anwendungsfreundliche Interpolationseigenschaften besitzt:

- Die Stützstellen, die durch die typischen Werte der Zugehörigkeitsfunktionen und die Regelschlußfolgerungen festgelegt sind, liegen auf der Kennlinie (Kennfläche).
- Die Kennlinie (Kennfläche) ist glatt und stückweise (halber Stützpunktabstand) durch quadratische Polynome darstellbar.
- Die Kennlinie (Kennfläche) ist im Bereich zwischen zwei benachbarten Stützstellen monoton, wodurch ein Überschwingen vermieden wird. Die Extremwerte der Kennlinie liegen daher immer an den Stützpunkten und sind somit durch die Regeln genau festlegbar.
- Der Anstieg der Kennlinie (Kennfläche) ist zwischen zwei benachbarten Stützpunkten betragsmäßig nie größer als der doppelte Anstieg der Verbindungsgerade dieser Stützpunkte. Dadurch kann sehr einfach eine obere Grenze für die Regelverstärkung bestimmt werden.
- Der Interpolator ist, soweit es die anderen Forderungen erlauben, geradentreu, ebenentreu und parabeltreu (siehe Abb. 4, 5 und 7).

Für die Berechnung des n -dimensionalen Falles wird mehrmals der eindimensionale Fall herangezogen. Die Eigenschaften gelten daher sinngemäß auch im n -dimensionalen Fall.

Dank

Dem österreichischen „Fonds zur Förderung der wissenschaftlichen Forschung“ (FWF) wird für die finanzielle Unterstützung der Arbeit im Rahmen des FWF-Projektes P8914-TEC und Herrn Dipl.-Ing. Dr. Georg Brasseur für die ideelle Unterstützung gedankt.

Literatur

- [1] Koch, M., Kuhn, T., und Wernstedt, J. Ein neues Entwurfskonzept für Fuzzy-Regelungen. *Automatisierungstechnik at*, 41(5):152–158, Mai 1993.
- [2] Kuhn, T. und Wernstedt, J. Regelungstechnischer Entwurf, Implementierung und Anwendung von Fuzzy-Steuerungen. *9. Österreichischer Automatisierungstag*, S. 20–36, Sept. 1993.
- [3] Lukas, P., Rehfeldt, K., und Schöne, A. Vergleich von Fuzzy-Reglern mit konventionellen digitalen Reglern. In *2. Workshop Fuzzy Control des GMA-U.A. 1.4.2*, Dortmund, November 1992. ISSN 0941-4169.
- [4] Rehfeldt, K. Funktionale Beschreibung von Fuzzy-Regelalgorithmen. In *DFMRS Jahrestagung*. Forschungsbericht Nr.93-1, ISSN 0944-694X, 1993.
- [5] Bertram, T. und Svaricek, F. Zur Fuzzy-Regelung eines aufrechtstehenden Pendels. *Automatisierungstechnik at*, 40(8):308–310, 1992.
- [6] Kiendl, H. und Fritsch, M. Fuzzy Control - Theorie für den Anwender. *Automatisierungstechnik at*, 41(2):A5–A8, Feb. 1993.
- [7] Driankov, D., Hellendoorn, H., und Reinfrank, M. *An Introduction to Fuzzy Control*. Springer-Verlag, New York, 1993.
- [8] v. Altrock, C. *Fuzzy Logic*. Oldenbourg, München, 1993.
- [9] Mayer-Gramann, K. D. und Jüngst, E.-W. Fuzzy Control - schnell und kostengünstig implementiert auf Standard-Hardware. *Automatisierungstechnik at*, 41(5):166–172, Mai 1993.
- [10] Zeng, X.-J. und Singh, M. G. Approximation theory of fuzzy systems — SISO case. *IEEE Transactions on Fuzzy Systems*, 2(2):162–176, 1994.
- [11] Leichtfried, J. und Heiss, M. Ein kennfeldorientiertes Konzept für Fuzzy-Regler. *Automatisierungstechnik at*, 43(1):31–40, 1995.
- [12] Heiss, M. und Leichtfried, J. Selflearning fuzzy controller with smooth transfer characteristic and guaranteed convergence. In *IEEE Conference on Control Applications (CCA 94)*, S. 1251–1256, Glasgow, 1994.
- [13] Drechsel, D., Strauch, W., und Pandit, M. Fuzzy Control im Vergleich zu einem regelbasierten Interpolations-Controller. In *Forschungsbericht Nr. 0293 zum 3. Workshop Fuzzy Control des GMA-U.A. 1.4.2*, S. 127–140, Dortmund, 1993. Universität Dortmund, Fak. f. Elektrotechnik.
- [14] Drechsel, D., Strauch, W., und Pandit, M. RIP Control im Vergleich zu Fuzzy Control. In *VDI Berichte Nr. 1113, GMA Aussprachetag Fuzzy Control*, S. 41–56, Düsseldorf, 1994. VDI-Verlag.
- [15] Föllinger, O. *Nichtlineare Regelungen I, 6. Auflage*. R. Oldenbourg Verlag, München, 1991.
- [16] Weinmann, A. *Regelungen, Analyse und technischer Entwurf, Band 2*. Springer Verlag, Wien, 3. Auflage, 1993.
- [17] Lane, S. H., Handelman, D. A., und Gelfand, J. J. Theory and development of higher-order CMAC neural networks. *IEEE Control Systems Magazine*, 12(2):23–30, April 1992.
- [18] Wang, L. X. und Mendel, Jerry M. Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. *IEEE Transactions on Neural Networks*, 3(5):807–814, 1992.
- [19] Higgins, C. M. und Goodman, R. M. Fuzzy rule-based networks for control. *IEEE Transactions on Fuzzy Systems*, 2(1):82–88, 1994.
- [20] Wang, L.-X. *Adaptive Fuzzy Systems and Control*. Prentice Hall, New Jersey, 1994.
- [21] Wang, L.-X. Stable adaptive fuzzy control of nonlinear systems. *IEEE Transactions on Fuzzy Systems*, 1(2):146–155, May 1993.
- [22] Leichtfried, J. Kennfeldorientierte Konzepte für Fuzzy-Regler. Diplomarbeit, Technische Universität Wien, Februar 1994.
- [23] Shandle, J. Fuzzy-logic compiler means faster, cheaper controllers. *Electronic Design*, S. 28, December 17, 1992.

Integration der automatisierten Demontage im Konstruktions- und Planungsprozeß

G. Gschwendtner

Institut für Handhabungsgeräte und Robotertechnik
Technische Universität Wien

Kurzfassung: Die Demontage ist ein relativ neues Gebiet, das immer mehr an Bedeutung gewinnen wird. Eine Automatisierung ist unumgänglich, um ein effizientes Recyceln durchführen zu können. Nachfolgender Bericht analysiert deshalb die möglichen Demontagearten und -operationen und beschreibt darauf aufbauend entwickelte Gestaltungsrichtlinien zur demontagegerechten Konstruktion. Zur Integration der automatisierten Demontage in den Konstruktions- und Planungsprozeß wird zugleich eine Methode vorgestellt, mit der alle Aspekte der Demontage in allen Phasen mitberücksichtigt werden können.

1 Einleitung

Eines der am meisten diskutierten Themen der heutigen Zeit ist der Umweltschutz. Anfangs von vielen nur als Modeerscheinung belächelt, hat er sich im letzten Jahrzehnt in den Industrieländern etabliert, was sich auch sehr leicht durch die Wahlergebnisse belegen läßt. Recycling ist eines der großen Schlagworte im Umweltschutz. Dieser Bereich, im Gegensatz zu anderen Themen wie z.B. Toxikologie und Ressourcenströme, ist für jedermann mittlerweile fast in allen Lebenslagen sichtbar. Der ständig steigende Konsum, welcher mit einer expandierenden Produktion verbunden ist, führte in den letzten Jahren und Jahrzehnten dazu, daß die Müllberge immer weiter gewachsen sind und damit geeigneter Raum zur Deponierung von Abfällen immer knapper wurde, wodurch die Abfallentsorgungskosten exponentiell ansteigen werden. Durch diese Bedrohung reagiert auch die Öffentlichkeit immer sensibler auf dieses Thema. Dieses verstärkte Umweltbewußtsein wird auch den Unternehmer immer mehr in seinen Aktivitäten beeinflussen. Dies sowohl durch den nicht zu unterschätzenden öffentlichen Druck, der ja in einem großen Ausmaß das Firmenimage bestimmt, als mittlerweile auch durch den Staat, der viele neue Umweltgesetze, wie z.B. die neue Abfallverordnung und die Verpackungsverordnung, beschließt, durch die unter anderem die Entsorgungskosten in die Bilanz des Unternehmens zurücktransferiert werden.

Für die Abfälle kommt es dabei zu einer Verantwortungsverlagerung von den öffentlichen Stellen hin zu den Verbrauchern und Produzenten. Umweltschutz muß deshalb zum integralen Bestandteil der Unternehmenskultur werden. Die Verantwortung des Unternehmers für seine Produkte darf aber nicht mit dem Verkauf an den Konsumenten enden. Für den Produzenten und Kunden wird es in Zukunft wahrscheinlich viel leichter

sein ein neues Produkt herzustellen und zu erwerben als ein ausgedientes auf anständige Art und Weise loszuwerden.

Zu diesem Zweck ist es notwendig, daß Wertschöpfungskreisläufe entstehen, um die bereits vorhandenen Wertschöpfungsketten zu schließen. Dem gesamten Life Cycle des Produktes muß zu allen Zeiten Beachtung geschenkt werden. Dies bedeutet, daß Altprodukte wiederverwertet bzw. noch wichtiger wiederverwendet - d.h. recycelt - werden müssen, um Rohstoffe zu schonen und der Explosion der Abfallmenge Einhalt zu bieten.

Die Demontage stellt beim Recycling den ersten Schritt dar. Ohne Demontage kann kein sinnvolles Recycling durchgeführt werden. Je besser und effizienter die Demontage realisiert werden kann, desto eher kommt man von der bisher durchgeführten Wiederverwertung hin zur Wiederverwendung bzw. zur Optimierung der Instandhaltung. Das Zukunftspotential des Recyclings liegt in der Wiederverwendung oder, um eine neues Schlagwort zu benutzen, im Upcycling. Um letzteres zu erreichen, ist es erforderlich, daß die Bauteile zerstörungsfrei demontiert werden oder eine gezielte zerstörende Demontage zur Optimierung der Verwertung realisiert wird.

Vom Nutzen ist die Demontage ist allerdings nur, wenn sie ökonomisch und ökologisch wirtschaftlich durchgeführt werden kann. Es sollte für jeden klar sein, daß eine Demontage, Aufbereitung und Remontage sinnlos ist, wenn dabei insgesamt mehr Energie und Ressourcen verbraucht werden als bei der Neuproduktion.

Die Wirtschaftlichkeit der Demontage, welche die Grundvoraussetzung des Recyclings ist, wird in erster Linie durch die demontagegerechte Konstruktion und den Entwurf der Demontagezelle bestimmt, d.h. durch die Produkt- und die Produktionsplanung. Da wir uns hier auf einem ganz neuen Gebiet befinden - die Automatisierung der Montage hat sich auch erst in den letzten Jahren durchgesetzt, wobei der Anstieg aber in exponentieller Weise vor sich ging und mittlerweile Automatisierungsgrade von bis zu 95% keine Seltenheit mehr sind - besteht hier noch ein großer Handlungsbedarf, der ein beachtliches Rationalisierungspotential enthält. Das wird noch dadurch verstärkt, da die Konstruktion und Planung selbst nur einen kleinen Teil der Produktionskosten ausmachen, diese aber zum größten Teil bestimmen.

Deshalb wird in Zukunft ein wesentlicher Punkt sein bereits bestehende Systeme der rechnergestützten Konstruktion zu verbessern und die Erkenntnisse der Demontage, die man größtenteils neu gewinnen muß bzw. zum Teil durch eine möglichst wirkungsvolle Synthese bereits vorhandener Verfahren oder übertragbarer Erkenntnisse aus verwandten Gebiet, besonders der Montage, erlangt, in diese einzubringen. Die ersten Ergebnisse dazu werden nachfolgend kurz erläutert.

2 Das Fachgebiet Demontage

Bisher war die Demontage als erster Schritt der Instandsetzung bekannt und wurde dabei nur von Hand durchgeführt. In letzter Zeit entwickelte sich die Demontage aber als ein eigener Wirtschaftszweig im Rahmen vom Recycling als Bestandteil der Austauscherzeugnisfertigung und zur Wiederverwertung der Materialien. Einige Anwender sagen auch "Rückbau" oder "Schadstoffentfrachtung" dazu. Im Automobilbau, einem der Hauptanwender, heißt es auch "Trockenlegung". Damit ist das Entleeren von

Benzin und Ölen, Bremsflüssigkeit und Kühlsysteminhalt, die Entnahme von FCKW aus der Klimaanlage und der pyrotechnischen Treibsatzchemie aus dem Airbag gemeint.

Trotz dieser Vielfalt der Namen, überwiegt bei vielen Leuten die Einfalt der Vorstellungen darüber, daß Demontage nur eine Montage rückwärts sei. Ein Rückschluß der Montage, besonders bezüglich der Automatisierbarkeit, auf die Demontage ist jedoch nur zum Teil zulässig. Eine logische Umkehrung der einzelnen Vorgänge, bei der die Technologien und Einrichtungen zur automatisierten Montage auch für die Demontage eingesetzt werden, gilt nur für die Abläufe bei der Bildung bzw. Auflösung der Baustruktur eines Produktes, d.h. für die Reihenfolge, in der die Montage- bzw. Demontageoperationen durchgeführt werden. In den seltensten Fällen herrscht eine Übereinstimmung bei den Verfahren zum Fügen bzw. Lösen der Verbindungen der einzelnen Bauteile. Man denke hier nur an die Schweißverbindungen, die zwar sehr leicht automatisiert gefügt, aber nicht mit der gleichen technischen Einrichtung wieder gelöst werden können.

Den Unterschied zwischen der Montage und der Demontage sieht man sehr deutlich, wenn wir die uns die einzelnen Operationen ansehen, die diese Vorgänge beinhalten.

Die **Demontageoperationen** bestehen dabei aus:

- Handhabungstätigkeiten (VDI 2860)
- Fügetätigkeiten (DIN 8590) und zusätzlich
- Trenntätigkeiten (DIN 8580)

Grundsätzlich kann man den automatisierten Demontagevorgang in die Aufnahme aus dem Bereitstellungsort (sowohl mittels Zubringeeinrichtungen als auch gleich mit dem Roboter selber), dem Aufbringen in die vorgegebene Position, der Durchführung der Trennoperation und das Ablegen gliedern. Tabellarisch kann dieser nach Tabelle 1 dargestellt werden. Im Gegensatz zur Montage fehlt allerdings die Bereitstellung der Einzelteile.

Die im **Demontagevorgang** benutzten Operationen des Trennen sind in der DIN 8580 (Fertigungsverfahren) genau definiert. Dabei meint man das "Fertigen durch Ändern der Form eines festen Körpers, wobei der Zusammenhalt örtlich aufgehoben, d.h. im ganzen vermindert wird. Dabei ist die Endform in der Ausgangsform enthalten; auch wird das Zerlegen zusammengesetzter Körper dazugerechnet."

Trennen wird in der weiteren Folge noch unterteilt in folgende Tätigkeiten:

- Zerteilen
- Spanen
- Abtragen
- Zerlegen
- Reinigen

Bei der Demontage stehen hauptsächlich das **Zerteilen** (Scherschneiden, Messerschneiden, Reißschneiden, Spalten, Reißen und Brechen), das **Zerlegen** (Auseinandernehmen, Entleeren, Lösen kraftschlüssiger Verbindungen, Zerlegen von durch Urformen gefügten Teilen, Ablöten, Lösen von Klebeverbindungen und das Zerlegen von textilen Verbunden) und das **Reinigen** (Reinigungsstrahlen, Mechanisches Reinigen, Strömungstechnisches

Reinigen, Reinigen mit Lösungsmitteln, Chemisches Reinigen und Thermisches Reinigen) im Vordergrund.

Teiloperation	Art der Teiloperation	Erklärung
Speichern	Handhaben	Meist geordnete Lagerung der zu demontierenden Teile
Bewegen	Fügen	Transport der zu demontierenden Teile in die Aufnahmestation
Erkennen	Handhaben	Wahrnehmung der Demontageteile, unter Beachtung geometrischer, funktioneller und physikalischer Charakteristiken, der Materialkombination, der Verschmutzung und der Art der Verbindung
Ergreifen	Handhaben	Aufnahme der Demontageteile
Orientieren	Handhaben	Der Demontageteil wird in die Lage gebracht, in welcher der Trennvorgang richtungsmäßig durchgeführt werden kann (Bearbeitungsrichtung)
Ausrichten	Handhaben	Drehen des Demontageteils in die Lage, in der ein korrektes Anbringen der Werkzeuge ermöglicht wird
Sichern	Handhaben	Sicherung des Werkstückes während des Trennen
Demontieren/-Zerlegen	Trennen	Ändern der Form des festen Körpers durch Aufhebung des örtlichen Zusammenhangs
Kontrollieren	Handhaben	Überprüfen der Ordnungsmäßigkeit des Trennvorganges
Bewegen	Fügen	Ablegen der demontierten Teile in Zwischenspeicher (laut Norm ein Fügevorgang)
Speichern	Handhaben	Geordnete oder ungeordnete Lagerung der demontierten Teile bzw. Komponenten für weitere Prozesse

Tab. 1 Demontagevorgang

Die **Trennverfahren** lassen sich weiters einteilen in:

- **Umkehrbare Fügeverfahren** z.B. das Lösen von Schraubverbindungen und
- **Nicht umkehrbare Fügeverfahren** bzw. an Fügeverfahren angepaßte Trennverfahren. Dabei kann eine zusätzliche Unterteilung in *zerstörende* (z.B. Schweißen) und *nicht zerstörende* Trennoperationen (z.B. federnd einspreitzen) erfolgen.

Zerstörende Trennoperationen können unterteilt werden in solche, welche die Verbindung selbst, den Basisteil oder den abzunehmenden Teil zerstören.

Die für umkehrbare Fügeverfahren benutzten Werkzeuge sind meistens die, welche auch beim Fügen eingesetzt werden, wodurch ein relativ geringer Aufwand entsteht. So kann zum Beispiel eine automatisierte Schraubvorrichtung und deren Überwachungssensoren sehr schnell zu einer Demontageanlage umprogrammiert werden (deren Effizienz ist in dieser Ausführung jedoch äußerst fragwürdig). Komplizierter dagegen ist die Auswahl der anderen Trennverfahren und deren Geräte. Wie schon bei der Auswahl der Fugemittel bei der Montage sollte hier auf eine breite Wissensbasis zurückgegriffen werden, um für diese Sache ein möglichst optimales Ergebnis erhalten zu können. Um diese Probleme gleich

vom Anfang an in den Griff zu bekommen, war es deshalb notwendig eine Methode zur Produktentwicklung zu erarbeiten, welche diese Anforderungen bestmöglich erfüllt.

3 Concurrent Engineering for Disassembling

Basis für diese Vorgangsweise der Produktentwicklung ist der Lebenszyklus des Produktes, von dem während der Auslegung immer alle Phasen (Produktion, Gebrauch und Entsorgung) zu berücksichtigen sind. Die Ziele des Life Cycle Engineering sind mit der herkömmlichen sequentiellen Produktentwicklung, bei denen die einzelnen aktiven Prozesse in den drei Phasen des Lebenszyklus jeweils voneinander unabhängig behandelt werden, jedoch kaum zu verwirklichen. Zusätzlich hat sie noch den erheblichen Nachteil, daß sie sehr zeitintensiv und fehlerbehaftet ist. Deshalb war es notwendig die Entwicklung eines Produktes radikal umzustellen, was mit Concurrent Engineering for Disassembling realisiert wird.

Damit können durch eine simultane Ausführung der Planung die Einflüsse aller einzelnen Prozesse des Lebenszyklus untereinander auf die Art und Weise optimiert werden, daß man sie frühest möglich erfaßt und mit minimalen Aufwand untereinander abstimmen kann. Vorstellbare Fehler werden ehest erkannt, alle sonst nie oder erst viel zu spät bemerkten Einflüsse im frühest möglichen Stadium berücksichtigt und damit Zeit und Kosten, auch solche die vielleicht erst in Zukunft entstehen, erheblich reduziert. Dadurch kann das Produkt bezüglich allen Anforderungen bestmöglich gestaltet werden.

Der Konstruktions- und Planungsprozeß für die Produktion wurde in vier verschiedene Hauptprozesse (Abbildung 1), die z.T. überlappend alle Phasen des Lebenszykluses berücksichtigen, klassifiziert. Durch die ersten beiden wird die erste Phase des Lebenszyklus berücksichtigt, die zweite Phase durch alle Prozesse und die dritte durch die letzten zwei, wobei das Recycling ohne der Demontage nicht durchgeführt werden kann. Deshalb ist es erforderlich, daß dieser Punkt selbständig behandelt wird.

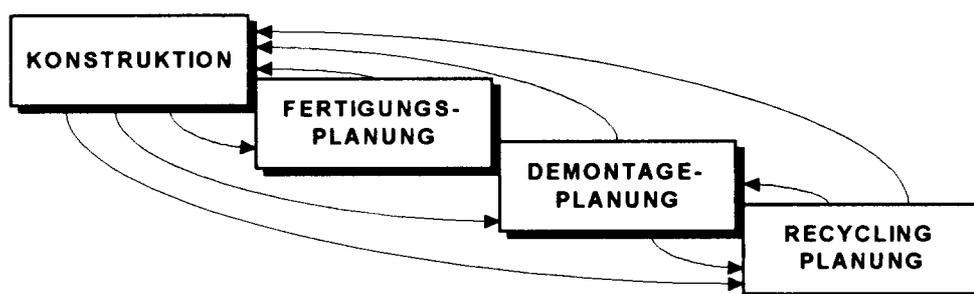


Abb. 1 Hauptprozesse bei der Produktionsplanung

Basierend auf diesen Informationsfluß zwischen den Komponenten, der maschinell meist erst in einer konkreteren Phase der Produktentwicklung einsetzt, ist es auch obligatorisch die Zusammenarbeit zwischen den einzelnen Abteilungen eines Unternehmens diesen Schritt anzupassen. Das geschieht durch die **Bildung eines Projektteams**, das vom ersten Zeitpunkt an in einer methodischen Art und Weise zusammenarbeitet.

Dieses Projektteam soll - beruhend auf den vier Hauptprozessen der Produktionsplanung - in zusätzliche spezialisierte Gruppen gegliedert werden, um alle Anforderungen erfüllen zu können.

Durch die Bildung eines Projektteams ist es möglich, daß die Arbeitsteiligkeit erheblich reduziert wird. Dadurch können Informationsverluste an den Schnittstellen zwischen den einzelnen Arbeitsschritten vermieden und die "geistige Rüstzeit" verringert werden. Auf diese Weise steigert sich die wertschöpfende Tätigkeit der Mitarbeiter beträchtlich. Durch die Aufgabenintegration in dezentrale Einheiten wird von den Mitarbeitern des Projektteams, neben fundierten Spezialkenntnissen, auch eine Erweiterung der bereichsübergreifenden Kenntnisse gefordert. Somit sind sie einem kontinuierlichen Lernprozeß ausgesetzt. In der Folge reduzieren sich die Routinetätigkeiten womit gleichermaßen die Arbeitsmotivation gefördert wird.

Zeitlich wird das Vorgehen mit dieser Methode in fünf Phasen gegliedert, die jeweils zufriedenstellend abzuschließen sind, bevor mit der nächsten begonnen werden darf.

Ab der zweiten Phase ist es möglich und für die weitere Entwicklung sogar dringend erforderlich das Produkt in verschiedene voneinander nahezu unabhängige **Module zu gliedern**. Diese können aufgrund ihrer Funktion, des Materials, der Fertigungs- und/oder Recyclingprozesse oder wie dieser Ansatz empfiehlt **aufgrund der Montage- und Demontagestruktur** festgelegt werden. Sowohl für das Unternehmen als auch für die Umwelt stellt das die beste Lösung dar, da für derartige Module auch die Fertigung und das Recycling sehr leicht organisiert werden kann.

Aufgrund dieser modularen Auslegung, bei der die montage- und demontagegerechte Auslegung besonderes beachtet werden sollte, ist es ab diesen Zeitpunkt möglich für das zu entwickelnde Produkt nahezu simultan ein Fertigungs-, Demontage- und Recyclingkonzept zu entwickeln.

Ein weiterer Vorteil der modularen Produktentwicklung ist die unproblematische Änderung bzw. Modernisierung einzelner Gruppen des Produktes bzw. die Konstruktion und Planung eines ähnlichen Neuen.

Die Konstruktion, als erster Punkt der Methode, bestimmt wesentlich die drei anderen Punkte. Deshalb wird nachfolgend ausgehend, von den allgemeinen Konstruktionsanforderungen, der Konstruktionsprozeß unter dem Aspekt der Demontage erörtert.

4 Demontagegerechte Konstruktion

Dem Konstrukteur obliegt die wesentliche Verantwortung die einzelnen Lebensphasen eines Produktes zu bestimmen. Grundlage für alle Arbeitsschritte sind die **Grundregeln**, daß eindeutig, einfach und sicher zu konstruieren sei und damit Voraussetzung für den Erfolg der späteren Lösung sind. Diese Anweisungen zur Gestaltung leiten sich aus folgenden generellen Zielsetzungen (Pahl, G., 93) ab:

- Erfüllung der technischen Funktion,
- wirtschaftliche Realisierung und
- Sicherheit für Mensch und Umgebung.

Ihre Nichtbeachtung kann zu großen Nachteilen, Fehlern, Schäden oder schlimmstenfalls zu Unglücken führen.

Die **Eindeutigkeit** ermöglicht Wirkung und Verhalten der Konstruktion zuverlässig vorauszusagen. Für die **Demontage** bedeutet dies, daß eine klare Demontagerihenfolge mit eindeutig zu bestimmenden Operationen und Werkzeugen aufgrund der konstruktiven Gestaltung vorgegeben wird.

Durch die **Einfachheit** wird eine wirtschaftliche Lösung erreicht. Dieser Punkt ist für die **Demontage** besonders wichtig. Anzustreben sind an dieser Stelle eine geringe Anzahl von Teilen und Verbindungen, einfacher Gestaltungsformen und Erkennbarkeit wodurch sich die Operationen schneller und besser durchführen lassen.

Zuletzt ist bezüglich der **Demontage**, aufgrund dem Einsatz hoher Kräfte, der Verschmutzung und dem oftmaligen Umgang mit gefährlichen Gütern und Prozessen, besonders die Beachtung der **Sicherheit** zu berücksichtigen. Diese bewirkt generell sowohl die zuverlässige Erfüllung einer technischen Funktion und Produktion (in unserem Fall Demontage) als auch die Gefahrenminderung für den Menschen und die Umgebung.

Neben diesen drei Grundregeln unterliegt die Lösung auch **Bedingungen** und **Anforderungen**, die sich sowohl **allgemein** (diese werden bei einer Aufgabe nicht explizit angegeben, aber stillschweigend vorausgesetzt und sind daher zu beachten) als auch aus der **konkreten** Aufgabenstellung ergeben. Durch Merkmale lassen sich diese übersichtlich und umfassend angeben, wobei sie sich dabei in der Regel gegenseitig beeinflussen.

Abhängig vom Konkretisierungsgrad werden jene, die auch als die **Eigenschaften** des Produktes gelten, während des Konstruktionsprozesses als zu beachtende **Leitlinien** bzw. **Gestaltungsrichtlinien** zusammengefaßt und verwendet. Diese sollten dann sowohl für Einzelteile, Komponenten und das gesamte Produkt gelten. Abbildung 2 zeigt eine Zusammenfassung möglicher Richtlinien zur Erfüllung bestimmter Anforderungen, die mehr oder weniger alle zufriedengestellt werden sollten.

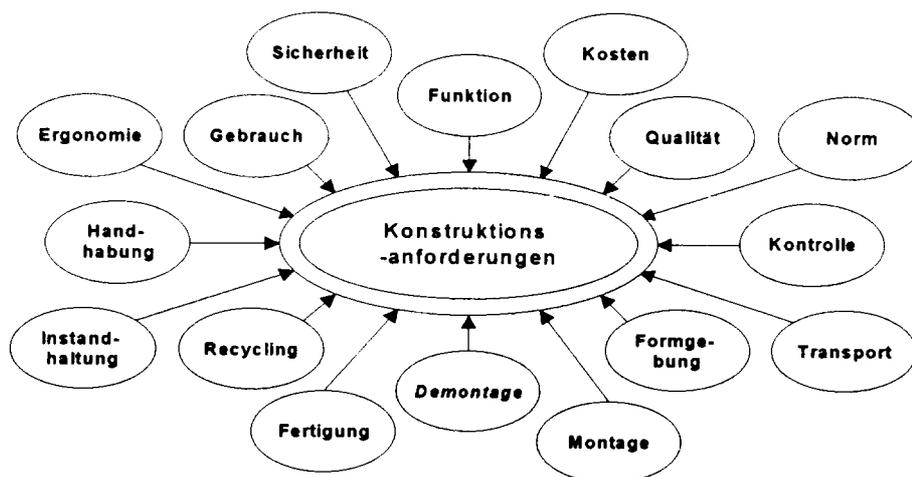


Abb. 2 Gestaltungsrichtlinien

Bei einer genaueren Betrachtung der Gestaltungsrichtlinien fällt auf, daß manche Punkte ein sehr enges Verhältnis untereinander aufweisen und ohne den anderen nicht erfüllt werden können. So ist die Demontage die Basis für das Recyceln und auch der

Instandhaltung bzw. Wartung. Auch die Verwandtschaft zur Montage ist logischerweise klar erkennbar und beide sollen gleichermaßen sicher, mit geringen Kosten und funktionsfähig ablaufen.

Die Demontage hat im Vergleich zur Montage den Vorteil, daß die **Teilebereitstellung fehlt**, was sich wesentlich auf die Anzahl der benötigten Operationen auswirkt, da bei der Montage alle zu montierenden Teile und Komponenten richtig bereitgestellt werden müssen. Deshalb ist der Ausgangspunkt der Überlegungen bezüglich Regeln zur montagegerechten Produktgestaltung immer der Einzelteil. Danach werden in der Regel die Baugruppen untersucht und zuletzt der Produktaufbau analysiert. In der Folge können noch spezielle Erwägungen bezüglich der Automatisierung, falls sie nicht schon in den vorherigen Punkten integriert sind, angestellt werden.

Ausgangspunkt bei der Demontage ist allerdings das **fertige Produkt**. Dieses muß so gestaltet sein, daß die einzelnen **Operationen**, die für die automatisierte Demontage benötigt werden, erfüllt werden können. Das schließt gleichzeitig mit sich ein, daß dabei eine Synthese der Überlegungen bezüglich der Einzelteile, der Baugruppen und des gesamten Produktaufbaues durchgeführt werden muß, und deshalb eine weitere Unterteilung bezüglich dieser Unterpunkte hier nicht sinnvoll erscheint.

Es ist also notwendig sämtliche Einflußfaktoren zu berücksichtigen. Die Wissensbasis zur Unterstützung des demontageorientierten Konstruktionsprozeß kann in **drei Hauptpunkte** gegliedert werden:

- Wissen zur Begünstigung der einzelnen automatisiert durchführbaren **Demontageoperationen**
- Allgemeine Regeln zur **Produktstrukturierung** zum Zwecke der Begünstigung der Automatisierten Demontage
- Wissen zur Erfüllung speziellen **Randbedingungen** basierend auf der Analyse des Fachgebietes Demontage und den daraus gezogenen Schlußfolgerungen

Bei dieser Gliederung ist zu beachten, daß diese Punkte zwar unabhängig voneinander untersucht werden können, aber bei der Produktentwicklung nur sinnvoll sind, wenn sie **alle im gleichen Ausmaß beachtet** werden. Es nützt beispielsweise nichts, daß das Produkt bis zu einem gewissen Grad automatisiert demontiert werden kann, die gewonnenen Komponenten und Teile aber keinen Recyclingprozeß (Randbedingung) zugeführt werden können. Diese Randbedingungen des demontageorientierten Konstruktionsprozesses sind deshalb sehr wichtig.

Grundsätzlich kann für alle diese festgestellt werden, daß diese **Wissensbasis sowohl für die automatisierte Demontage als auch der manuellen Demontage** Gültigkeit besitzt, da erstere ein Spezialfall der letzteren ist. D.h. auch ohne der Inanspruchnahme der Automatisierungsmöglichkeit wird die Demontage des Produktes begünstigt.

Die erste der aufgezählten Gruppe enthält dazu Regeln, die sich nach den in Tabelle 1 determinierten Operationen gliedern. Vereinfacht ausgesprochen, kann dazu gesagt werden, es ist so zu konstruieren, daß man alle Operationen möglichst einfach ausführen kann.

Als besonders zu beachtendes Problem ergibt sich dabei die Verbindungstechnik. Der Erfolg der Demontageoperation hängt im wesentlichen von zwei voneinander unabhängigen Punkten ab.

- Die **Zugänglichkeit** der Demontagewerkzeuge zu allen Verbindungen muß gewährleistet werden.
- Abhängig von den geplanten Recyclingverfahren muß eine geeignete Auswahl der **Verbindungen** erfolgen.

Hierzu wurden in der Folge die möglichen Verbindungstechniken auf deren Lösbarkeit, Automatisierbarkeit und den anzustrebenden Recyclingformen hin untersucht und bewertet. Für die automatisierte Demontage am besten geeignet erwiesen sich dabei Schnappverbindungen und Schrauben. Die günstigste Schraubenform ist eine mit Gewindeloch mit einem Innentorx- bzw. Innensechskantschraubenkopf. Diese Wertung gilt dabei gleichzeitig als Auswahlvorschrift für die durchzuführenden Trennoperationen, die abhängig von der eingesetzten Verbindung und der angestrebten Recyclingform festgelegt wird.

Konstruktive Regeln zur Begünstigung der automatisierten Demontageoperationen bieten sich, da sie die allgemeine Produktstruktur nicht so sehr beeinflussen, vor allem bei einer kurzfristigen Umkonstruktion eines bestehenden Produktes als erste durchzuführende Verbesserungsmaßnahme an. Bei einer Neukonstruktion, bei der letztere selbstverständlich auch allezeit zu beachten sind, gibt es aber noch wirksamere Möglichkeiten Einfluß auf die Demontagerechtigkeit auszuüben, nämlich durch nachhaltige Maßnahmen, die sich global auf die **Produktstrukturierung** auswirken. Diese gliedern sich dabei in Maßnahmen zur Verbesserung der allgemeinen Produktstruktur, an den Baugruppen und an den Einzelteilen bzw. produktstrukturelle Maßnahmen zur Vereinfachung der Demontageoperationen und zur Vereinfachung der Handhabungsoperationen.

Nach den produktstrukturellen und den Maßnahmen zur Vereinfachung der Demontageoperationen besteht bereits die grundsätzliche Möglichkeit zur automatisierten Demontage. Aufgrund von gewissen **Randbedingungen** kann es jedoch sein, daß diese Möglichkeit im Bereich der Theorie bleibt oder aber der dabei eingegangene Aufwand sich nicht lohnt. Diesen für die Praxis wichtigen Bereichen wird deshalb zuletzt in den Gestaltungsrichtlinien zur Erfüllung der Randbedingungen nachgegangen: Jene gliedern sich dabei in: korrosionsgerecht, kriechgerecht, verschleißgerecht, trockenlegungsgerecht, reinigungsgerecht, prüf- und sortiergerecht, aufarbeitungsgerecht und recyclinggerecht. Durch das Befolgen all dieser Gestaltungsrichtlinien sollte es möglich sein das Produkt einfach und automatisiert zerlegen und effizient recyceln zu können.

5 Ausblick und Zusammenfassung

Die Demontage ist ein noch neues Themengebiet, das aufgrund seiner Bedeutung im Rahmen von Recycling überproportional stark an Wichtigkeit gewinnen wird und deshalb für immer mehr Firmen interessant wird. Aufgrund der Komplexität des gesamten Vorganges ist es freilich von großem Vorteil, daß dieses Themengebiet schon in den ersten Phasen des Konstruktions- und Planungsprozeß integriert wird und dann ständig beachtet

wird. Eine Gliederung des Produktes aufgrund der Demontagestruktur ist ebenfalls von großer Bedeutung. Nur so ist es möglich das Produkt zu zerlegen und wirtschaftlich zu recyceln.

Durch das Befolgen demontagegerechter Gestaltungsrichtlinien im Rahmen von Concurrent Engineering for Disassembling kann dies effizient realisiert werden. Als nächstes zu bearbeitendes Themengebiet wird es in Zukunft für die Planung der Demontagezelle erforderlich sein, die notwendigen Eigenschaften für Industrieroboter und der zugehörigen Sensorik, Demontagewerkzeuge und Transfersysteme festzulegen. Erschwerend bei diesem Punkt kommt hinzu, daß bei der Demontage die Produkte verschmutzt, beschädigt oder durch Verschleiß oder Korrosion beeinträchtigt sein können, weshalb an die eingesetzten Geräte viel höhere Anforderungen als bei der Montage gestellt werden.

6 Literatur

- Bullinger H.J., Warschat J., Bopp R.: Methods and Tools to Support Design for Recycling. Proceedings of the 2. International Seminar on Life Cycle Engineering, Erlangen, 1994, pp 106-127
- Gschwendtner G., Kopacek P.: Concurrent Engineering for Disassembling. Accepted paper for the eight Symposium on Information Control Problems in Manufacturing, Beijing 11.-13. Oct. 1995.
- Olesen J., Keldmann T.: Design for Environment - a Framework. Journal of Engineering Design, Vol.5, No.1 (1992), pp. 45-54
- Pahl G. & Beitz W.: Konstruktionslehre. Springer Verlag Berlin, 1993
- Schmaus T., Kahemeyer M.: Demontagegerechte Produktgestaltung zwischen Montage, Aufbereitung und Entsorgung. VDI-Berichte 999, 1992, S. 277-294
- Steinhilper R.: Produktrecycling - richtig geplant, richtig gerechnet. VDI-Berichte 1089, 1993, S. 221-245
- VDI-Richtlinie 2222: Konstruktionsmethodik. Erstellung und Anwendung von Konstruktionskatalogen. VDI Verlag, Düsseldorf, 1982
- VDI-Richtlinie 2243: Konstruieren recyclinggerechter technischer Produkte, Beuth Verlag, Berlin/Düsseldorf 1992

Wenn's um Regeltechnik geht, sind wir Ihr Partner

Honeywell Industrieautomation

Automation entspringt dem Wunsch, Prozesse wirtschaftlich zu regeln, Qualität zu erhöhen, Kosten zu senken, Arbeitsplätze zu humanisieren und, nicht zuletzt, Rohstoffe ökonomisch einzusetzen. Die Antwort von Honeywell ist die Lösungsvielfalt der TotalPlant Evolution:

- Ein breites Produkt- und Engineering-Angebot für die Feldebene, offene Systemarchitektur des Prozeßleitsystems TDC3000*
- Advanced Control Technologien für Batch-Prozesse, Materialfluß- und Speicherprobleme in Raffinerien sowie für die kontinuierliche Emissionskontrolle
- Prozeßsimulation zum Operatortraining
- Schnittstellen zu Management Informationssystemen
- Experten für Beratung
- Projektmanagement und Service

und das mit gleichbleibender Qualität überall auf der Erde, zu jeder Zeit



Honeywell

Honeywell Austria GesmbH
Handelskai 388
A - 1023 Wien
Tel.: 0222 / 727 80 - 0

International Conference "Centennial Hurwitz on Stability Theory"

Centro Stefano Franscini, Monte Verità, Ascona, Switzerland
May 21 - 26, 1995

A report by Alexander Weinmann

Adolf Hurwitz was Professor of Mathematics at the ETH Zürich. The 100th anniversary of his stability criterion was celebrated by scientists from all over the world attending this conference organized by the ETH Professors Rolf Jeltsch and Mohamed Mansour.

The problems and solutions presented in several invited and contributed talks are pointed out by the following short statements:

- Solutions for robustness problems with low computational complexity based on robustness predictors via central limiting.
- Generalized Routh-Hurwitz algorithms and polynomials nonnegative on the imaginary axis.
- Symbolic manipulation approach for generalized integral criterion using the system error raised to a power greater than two.
- Review and overview on stability radii; importance of the complex stability radius as a unifying tool for nonlinearities and time-varying systems.
- Hurwitz stability of polynomial matrices and comparison with Lyapunov and Kronecker approach.
- The well-known theory of applying Strictly Positive Real functions in linear system stability is extended to time-varying and multivariable systems. Since asymptotic stability is retained to a certain region of gain, the results are applicable to robust control.
- A new technique Order Stars is utilized to relate the accuracy of the schemes of Hurwitz determinants and Routh array with their stability.
- Qualitative reasoning concerning Hurwitz criterion, i.e., not only to check stability but to construct stable combinations of coefficients. One is interested in searching a convex

direction (local convex direction) and not only to check whether a given polynomial is a convex direction.

- Extension of convex direction to quasipolynomials (i. e., polynomials including delay).
- Hurwitz solved the problem for one point of parameters in the parameter space, Kharitonov for a cube. Time-varying systems require a cube both in the parameter space and in the space of at least the first derivative of the parameters with respect to time. The idea is to reduce the problem to vertices and Lyapunov matrices which are affine or multiaffine in the parameters. In the case of multiaffine matrices one can obtain vertex conditions. Since Lyapunov theory is used the results are sufficient, only.
- Input-state stability (ISS) and global asymptotic stability (GAS) are useful definitions, especially ISS-Lyapunov.
- Application of symbolic algebra to find simple solutions from logic to stability problems, i.e., linear stability problems are rewritten as quantifier elimination problems.
- Robustness of systems with large number of uncertain parameters is applied to systems with tree structure. This method is characterized by the fact that robustness can be determined with high computational speed.
- Analog computation is back in fashion for solving $(0,1)$ -optimization problems especially using analog artificial neural networks and multilinear objective functions.
- A fuzzy supervisor was also suggested in order to stabilize a fuzzy controller in the case of unknown plant.
- Stabilization of an adaptive system under unknown parameters can be carried out without additional excitation.

Although many of the problems posed are treated and solved from a viewpoint of mathematical interest there are several opportunities of applications in automatic control. Moreover, there appear a lot of approaches to directly design systems which fulfil certain performance setpoints.

6. IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Man-Machine Systems

R. Genser

Das 6. Symposium über Analyse, Entwurf und Entwicklung von Mensch-Maschinesystemen wurde vom 27. bis 29. Juni 1995 am Massachusetts Institute of Technology in Cambridge, USA, veranstaltet. Bei Themen, die von Entscheidungsmodellen, unscharfen Informationen, Organisation, Neuromuskellmodellen, Lernen, Telerobotern, Mensch-Computerschnittstelle, Überwachung und Fehlerprobleme reichten, war diesmal neben Produktion und Straßenverkehr der Bereich Luftfahrt sehr stark vertreten. Dies lag sicher daran, daß nicht nur weltweit Großprojekte im Bereich der automatischen Luftverkehrssteuerung durchgeführt werden, sondern wohl auch daran, daß am MIT und dem Forschungsinstitut des US-Verkehrsministeriums diesbezügliche Forschungsvorhaben laufen. Auch beim angebotenen Besichtigungsprogramm lag der Schwerpunkt im Bereich der Mensch-Maschine Probleme in der Luftfahrt.

Hervorzuheben ist aber der Plenarvortrag von Prof. Ian Hunter vom Department of Mechanical Engineering am MIT über „Der instrumentierte Mensch: Technische Herausforderung und Innovation“, leider ist nur eine schriftliche Kurzzusammenfassung verfügbar. Die Präsentation entsprach dem Stand der Technik, die gezeigten Beispiele waren visionär, aber die Aussagen erscheinen ungläublich.

Lithium-Polymere ermöglichen verteilte Batterien für die Energieversorgung am Körper; hoffentlich sichert Österreich die optimale ökonomische und ökologische Nutzung seiner großen Lithiumvorkommen.

Tragbare Computer und die Signalübertragung mittels der Haut läßt an zugeordnete Körper-LAN-Protokolle (local area network) denken. Nun Ohrimplantate der Universität Innsbruck von Frau Dr. I. Hochmaier waren vor einigen Jahren auch noch undenkbar. Finite Elastizitätstheorie, nichtlineare anisotrope Materialeigenschaften, künstliche Muskeln, künstliche Haut, Mikroroboter in Millimetergröße usw. eröffnen neue Anwendungen und Lösungen. Der Mikro-Operationsroboter MSR-1 ist hundertmal genauer als der Mensch. Körperroboter sehen nicht nur, sondern können auch hören und fühlen. Vorstellbar sind dann klimatisierte Kleidung, „gehende“ Hosen, Körperflügel (in 15 Jahren schon), freibewegliche Augen und Ohren, tragbare Herz-Lungemaschinen, besserer Geruchssinn (zum Beispiel als Gasdetektor), Lebensmittelverfälschungs-Detektor und vieles mehr.

Die Teilnehmer waren mit der Tagung und deren Organisation sehr zufrieden. Erfreulich war die Teilnahme von zwei VertreterInnen der Universität in Wien und deren Beitrag über bildliche Informationsaufbereitung.

Es ist zu hoffen, daß beim 7. Symposium dieser Serie, welches 1997 in Japan veranstaltet werden soll, auch Österreich seinem Ruf als entwickeltes Land gerecht wird. Seine Chancen bei der Entwicklung von FANS (Future Air Navigation System) konnte Österreich aus mehreren Gründen nicht nutzen, wie auch in anderen Gebieten des Verkehrsbereichs, aber vielleicht kommen von der medizinischen Seite starke Impulse.

4. Internationaler Workshop „Robotics in Alpe - Adria Region - RAA '95“ 6.-8. Juli 1995 Pörtschach, Kärnten.

Ch. Fessel, M.-W. Han, P. Kopacek, G. Kronreif, T. Perme
Institut für Handhabungsgeräte und Robotertechnik, TU Wien

Der Workshop „Robotics in Alpe-Adria Region - RAA“ hat sich zu einem mitteleuropäischen Forum zum Erfahrungsaustausch und zur Präsentation von neuen Forschungsergebnissen auf dem Gebiet der Robotik etabliert. Dieser Workshop wurde vom IHRT (Institut für Handhabungsgeräte und Robotertechnik) im Auftrag der gleichnamigen Arbeitsgruppe organisiert. Es ist die Aufgabe dieser Arbeitsgruppe, die speziellen Probleme und die jüngsten Entwicklungen auf dem Gebiet des Robotereinsatzes unter besonderer Berücksichtigung der Anforderungen von Klein- und Mittelbetrieben zu diskutieren und zu lösen.

Seit dem Jahre 1992, in welchem die Arbeitsgruppe erstmals in Portoroz (Slowenien) ihr erstes Workshop abhielt, hat sich dessen Umfang vervielfacht. Die zweite Tagung - 1993 - erfolgte in Krems (Österreich), gefolgt vom Workshop '94 in Bled (Slowenien). Zum diesjährigen Workshop RAA'95 in Pörtschach/Wörthersee wurden zusätzlich zu den Wissenschaftlern aus den Gründerländern (Deutschland, Italien, Österreich, Slowakei, Slowenien und Ungarn) auch solche aus den ehemaligen osteuropäischen Staaten eingeladen. Am jährlichen Meeting der Arbeitsgruppe wurde daher unter anderem beschlossen, die Gruppe in „Robotics in Alpe-Adria-Danube Region RAAD“ umzubenennen.

Die Forschungen auf dem Gebiet der Robotik können heute im wesentlichen zwei Richtungen zugeordnet werden. Repräsentieren Forschungsarbeiten in Kinematik, Kinetik, Robotersteuerungen, Programmiersprachen und externe Sensoren sowie neue Anwendungsgebiete wie die roboterunterstützte Demontage das eher „klassische“ Aufgabengebiet, so rückt die Forschung im Bereich „Service-Roboter“ zunehmend in den Blickpunkt des Interesses. Diesem vielseitigen Einsatzspektrum muß und wird die Arbeit der RAAD Working Group verstärkt Rechnung tragen.

Am diesjährigen Workshop wurden spezifische Probleme obgenannter Art in Form von drei Übersichtsvorträgen sowie 63 Einzelvorträgen behandelt. Die Übersichtsvorträge behandelten die Themen *Some Considerations on Hyper-Redundant Robots* (J. Lenarcic, SLO), *Advanced Navigation Techniques for Mobile Service Robots* (G. Schmidt, GER) und *Italian Experiences on Applied Virtual Reality* (M. Sacco, I). In den zahlreichen Fachvorträgen wurden neben den oben erwähnten Themengebieten weiters Beiträge betreffend Simulation, Montage, Demontage, fortgeschrittene Steuerungen, Neuro&Fuzzy Control, Roboter Schweiß- und Arbeitszellen, Mobile Roboter u dgl. behandelt.

Von den mehr als 80 eingereichten Abstracts wurden 73 Papers angenommen, von welchen 63 am Workshop präsentiert und in den Pre-Prints abgedruckt wurden. Infolge der beschriebenen Entwicklung im Bereich „Robotertechnik“ wurde vom jährlich tagenden *Steering Committee* der Beschluß gefaßt, den Workshop auf den Bereich „Alpe-Adria-Danube“ auszudehnen. Nicht zuletzt deshalb wird der bislang abwechselnd in Österreich und Slowenien ausgetragene Workshop im Jahr 1996 in Budapest (H), 1997 in Cassino (I) und 1998 in Smolenice Castle (SK) abgehalten.

Entwicklung eines CAM-Systems zur automatisierten Konstruktion und Fertigung von Hydraulik-Steuerblöcken.

Thomas Rabeder

Begutachter: Prof. Dr. H. P. Osanna und Prof. Dr.Dr.h.c. P. Kopacek

Metallverarbeitende Klein- und Mittelbetriebe, welche hauptsächlich Zulieferanten für große Unternehmen sind, sind durch die Marktverhältnisse gezwungen, ihre Fertigung zu automatisieren. Möglichkeiten, diese Situation zu beherrschen, bietet die rechnerunterstützte und -integrierte Fertigung.

Die vorliegende Arbeit behandelt im weitesten Sinne eine CAM-Kopplung in einem Linzer Mittelbetrieb mit ca. 80 Mitarbeitern. Das Unternehmen ist auf dem Gebiet der Industriehydraulik tätig.

Hydraulik Steuerblöcke sind wichtige Bauteile hydraulischer Systeme, die zur Verbindung diskreter Bauelemente (wie Wegeventile, Drosseln oder Rückschlagventile) und Baugruppen dienen. Alle hydraulischen Bauelemente, die innerhalb eines Systems nicht ortsgebunden sind, werden nach Möglichkeit in einem integrierten Steuerblock zusammengefaßt.

Ein solcher integrierter Steuerblock verwirklicht das im Hydraulikschaltplan erstellte Leitungsnetz. Hydraulik Steuerblöcke sind also die technische Realisierung von hydraulischen Schaltungen. Zu diesem Zweck werden Hydraulik-Steuerblöcke im allgemeinen aus einem quaderförmigen Metallblock hergestellt, der orthogonal so durchbohrt ist, daß die entstehenden Leitungsverbindungen mit den Verknüpfungen des entsprechenden Schaltplanes übereinstimmen.

Die manuelle Konstruktion eines Steuerblockes ist sehr arbeitsintensiv und stellt hohe Anforderungen an die Fähigkeiten und das räumliche Vorstellungsvermögen des Konstrukteurs. Alleine das Überprüfen eines fertig erstellten Leitungsnetzes auf Kollisionen zwischen den einzelnen Leitungen setzt das Anfertigen mehrerer Schnittdarstellungen voraus. Ergebnis des Konstruktionsprozesses ist die technische Zeichnung, in der die räumliche Gestalt des Steuerblocks als Projektion in eine maßstabgetreue zweidimensionale Darstellung zurückgeführt ist. Mit der zugehörigen Bohrungskoordinatenliste ist somit die Ausgangsbasis für die NC-Fertigung erstellt.

Steuerblöcke werden gewöhnlich von Auftrag zu Auftrag immer wieder neu konstruiert und können in der Regel nicht standardisiert werden. Entwicklungszeiten für ein einfaches Blockdesign von ein bis zwei Wochen sind dabei keine Seltenheit. Obwohl sie seit Jahrzehnten in einer beinahe unveränderten Bauweise gefertigt werden, ist es bisher aufgrund der Komplexität des Konstruktionsprozesses noch nicht gelungen, die Anfertigung solcher Steuerblöcke durch entsprechende Computerprogramme zu automatisieren. Aufgrund der oft auftretenden komplexen Schaltungsstrukturen ist also in der Konstruktion der integrierten Steuerblöcke ein beachtliches Rationalisierungspotential durch Verkürzung der Durchlauf und Entwicklungszeiten vorhanden.

Um die Bearbeitung von Hydraulik-Steuerblöcken so rationell wie möglich durch Datenverarbeitungswerkzeuge zu unterstützen, war eine der Anforderungen die Entwicklung von Konstruktionsalgorithmen und die Konzipierung und Implementierung einer anwendungsgerechten CAM-Kopplung auf der vorhandenen Datenbasis.

Durch die Anwendung der entwickelten CAD-Applikation ist durch die automatische Berechnung der Bohrungsstrukturen und Einbindung der Schaltplanlogik eine Verkürzung der Durchlaufzeit im Engpaßbereich "Konstruktion" um den Faktor 3-5 möglich. Durch die dreidimensionale Darstellung und der Anzeige von Konstruktionsfehlern am Bildschirm werden diese für den Konstrukteur sofort ersichtlich.

Ähnliches gilt für das in die Applikation integrierte Postprozessorsystem, das die Durchlaufzeiten bei der Kopplung zur Fertigung weiter senkt.

Die prozentualen Änderungen des Zeitaufwandes lassen sich unter Berücksichtigung der unterschiedlichen Personalkosten der CAD-Konstrukteure einfach in relative Kostenveränderungen überführen. Neben den verkürzten Auftragsdurchlaufzeiten ergibt sich auch eine Erhöhung der Kapazitäten im Engpaßbereich "Konstruktion".

Außerdem trägt die Anwendung des Programmpakets wesentlich zur Vermeidung von Fehlern im Bereich der Konstruktion bei, die sich später bis in den Fertigungsprozeß fortpflanzen.

Ganzheitlicher Ansatz für die Ermittlung von Qualitätskennzahlen bei der Prüfmittelüberwachung im Rahmen des Aufbaus und der Umsetzung von Qualitätsmanagementsystemen nach ISO 9000

Christian B. Fürnweger

Begutachter: Prof. Dr. H. P. Osanna und Prof. Dr. Dr.h.c. P. Kopacek

Die Zielsetzung dieser Arbeit liegt darin, die Tätigkeiten und damit die benötigten Kapazitäten und Ressourcen bei den Aktivitäten in Klein- und Mittelbetrieben im österreichischen Raum im Zusammenhang mit den Vorgaben der internationalen Normenreihe ISO 9000-94 aus der Sicht der Unternehmensberatung bzw. der Beratung beim Aufbau von Managementsystemen und die Vorbereitung für das Zertifizierungsverfahren zu beleuchten und zur Diskussion zu stellen.

Die Notwendigkeit und/oder Sinnhaftigkeit mancher damit im Zusammenhang stehender Tätigkeiten - sowohl bei der Erstellung der verlangten Dokumentation als auch bei der innerbetrieblichen Umsetzung - abgeleitet aus den Vorgaben der Norm, soll dabei kritisch hinterfragt werden.

Gerade die beiden zuerst genannten Unternehmensgruppen - Kleinbetriebe und Mittelbetriebe - sind es letztendlich, welche heute den überwiegenden Anteil der Unternehmen im Wirtschaftsgeschehen darstellen und sich zur Zeit am meisten mit der Einführung von Qualitätsmanagement-Systemen nach ISO 9001 aufgrund des herrschenden Druckes im internationalen Wettbewerb beschäftigen müssen.

Die ersten drei Abschnitte dieser Arbeit gehen auf diese Problematik näher ein. Es wird dabei anhand von Erfahrungen, Auswertungen und Beispielen versucht, hinsichtlich der Bedeutung der Qualitätssicherung aus heutiger Sicht, sowohl den strukturellen Aufbau einer Qualitätsmanagement-Dokumentation - allgemein als Qualitätsmanagement-Handbuch (QM-HB) bezeichnet - genauer darzustellen und einzelne Aspekte des Systems näher zu beleuchten, als auch die Problematik, bei der praktischen Einführung und Umsetzung der festgelegten oder geforderten unternehmensspezifischen Vereinbarungen (z.B.: in Verfahrensanweisungen) zu den Qualitätsmanagement-Elementen an Hand von praktischen Beispielen aufzuzeigen und zur Diskussion zu stellen.

Eine besondere Aufgabenstellung, welche jedes industrielle Unternehmen beschäftigt, das sich mit dem Aufbau oder der Aufrechterhaltung eines Qualitätsmanagement-Systems auseinandersetzen hat, stellt - aus der Beratererfahrung heraus - das Qualitäts-Element 11 (laut ISO 9001-94) - die "Prüfmittelüberwachung" - dar. Auf diesen Abschnitt der Norm wird in verstärktem Ausmaße eingegangen.

Unter anderem werden dabei im ersten Teil des Kapitels die Vorgaben der Norm zu diesem Element genau analysiert. Gleichzeitig wird versucht eine detaillierte Struktur für den Aufbau eines Prüfmittelverwaltungssystem zu skizzieren. Im zweiten Teil wird ein für Klein- und Mittelbetriebe wirtschaftlich brauchbares System für die Meßdatenerfassung als Voraussetzung für Prüfmittelfähigkeitsuntersuchungen entworfen. Der dritte Teil des Abschnittes stellt den zukunftsorientierten Versuch dar, die vorher skizzierten Möglichkeiten von Prüfmittelfähigkeitsuntersuchungen mit Hilfe einer innovativen, mathematisch fundierten Methode - der Fuzzy-Logik - einen Lösungsansatz für Möglichkeiten zur Ermittlung von aussagekräftigen Qualitätskennzahlen im Zusammenhang mit Prüfmitteln inklusive der daraus abzuleitenden Neuerungen zu entwerfen.

Das letzte Kapitel - und damit der Abschluß dieser Arbeit - setzt sich ganz allgemein mit der Überwachung des Qualitätsmanagement-Systems auseinander. Dabei werden zuerst die Erfahrungen im Zusammenhang mit den internen Überwachungsmöglichkeiten, das heißt den "Internen Qualitätsaudits" und den damit verbundenen verschiedenen Audit-Arten dargestellt. In weiterer Folge geht dieser Abschnitt auch auf die für "Interne Auditoren" notwendigen und vorausgesetzten Qualifikationen näher ein.

Als letzter Punkt wird auf das nationale und internationale Zertifizierungsgeschehen - zumindest für den westeuropäischen Raum - eingegangen und versucht, zu den dabei aufgeworfenen Themenkreisen die wesentlichen Punkte - wie beispielsweise den Sinn und Zweck von Zertifizierungen - herauszuarbeiten.

A New Concept for a Compliance Element

Manfred Falkner

Begutachter: o.Univ.Prof. Dr. Dr.h.c. P. Kopacek und o.Univ.Prof. Dr. K. Desoyer

Ein neues Konzept eines passiven Fügemechanismus (Compliance Element) für den Endeffektor von Robotern wird präsentiert welches als Hauptmerkmale fünf Freiheitsgrade, ein einstellbares Fügezentrum und einen integrierten Verriegelungsmechanismus besitzt. Der Mechanismus heißt ACRC Element (Adjustable Centerless Rotation Compliance) und ist durch seinen Aufbau im Gegensatz zu existierenden Fügemechanismen universell einsetzbar.

Die Entwicklung basiert auf dem Vergleich von Testresultaten des klassischen Bolzen-Loch (peg-in-hole) Fügeverfahrens durchgeführt mit existierenden kommerziellen Fügemechanismen gebaut nach dem RCC Prinzip (remote center compliance) und Simulationen durchgeführt mit einem mathematischen Modell erstellt in Anlehnung an die Theorie von D. E. Whitney et al. von den Charles Stark Draper Laboratorien (USA). Das Modell wurde zuerst mit Hilfe der Literatur auf seine Richtigkeit geprüft und danach zum Verifizieren der Testdaten der kommerziellen Mechanismen verwendet. Der Vergleich deckte mehrere Schwachstellen dieser Fügemechanismen auf, insbesondere die Unfähigkeit zur Entkoppelung von Linearfehlern und Winkelfehlern während des Fügevorganges. Im Zuge der mathematischen Simulation wurden eine Reihe von Parametervariationen durchgeführt um typische Fügegeometrien zu untersuchen.

Die Erkenntnisse aus den vorhergehenden Maßnahmen wurden zusammengefaßt und zur Definition des ACRC Elementes verwendet. Dieser neuartige Mechanismus besteht grundsätzlich aus drei Modulen; dem Linear Compliance Modul ausgestattet mit Elastomeren zur Bereitstellung von zwei linearen und einem rotatorischen Freiheitsgrad, dem Centerless Rotation Modul bestehend aus zwei überkreuzten ebenen Koppelgetrieben zur Bereitstellung von zwei rotatorischen Freiheitsgraden, und einem Adjustment and Locking Modul verantwortlich für die Verstellung des Fügezentrums in Füge-richtung und für das Verriegeln aller fünf Freiheitsgrade während Handhabungsarten ohne Kontakt zwischen den Fügepartnern.

Die Neuartigkeit dieser Entwicklung liegt in den Haupteigenschaften des ACRC Elements, welche im wesentlichen der dreidimensionale infinitesimal bewegliche (wackelige) Rotationsmechanismus mit virtuellem Zentrum zur Entkoppelung von Linear- und Winkelbewegungen ist, die Verstellmöglichkeit des Fügezentrums mit Hilfe eines Elektromotors, die drei zusätzlichen Freiheitsgrade der integrierten Elastomere, und die Sperrbarkeit aller Freiheitsgrade mit Hilfe der für die Verstellung benutzten Motoren.

Die Aufgabe des vorgestellten ACRC Elements liegt in seiner Haupteigenschaft der passiven separaten Kompensation von Linear- und Winkelfehlern für jede Art von Fügegeometrie, wodurch der Nachteil der geringen Universalität existierender Fügemechanismen vermieden wird. Das ACRC Element ist auch eine einfache und verlässliche Alternative zu aktiven Füge-Systemen mit komplexem sensorischen Aufwand.

Einfluß der Meßstrategien und Auswertalgorithmen bei der Ermittlung von Formabweichungen gezeigt am Beispiel der Geradheitsmessung

Daniel Prostednik

Begutachter: Prof. Dr. H. P. Osanna und Prof. Dr. Dr.h.c. P. Kopacek

Im Rahmen dieser Arbeit wird der Einfluß der Meßstrategien und der Auswertalgorithmen bei der Ermittlung von Formabweichungen behandelt.

In der theoretischen Analyse erfolgt die Eingliederung der Problematik und die Zusammenhänge im Problembereich der Werkstückgenauigkeit werden aufgezeigt.

Die Auswertalgorithmen zur Bestimmung von Formabweichungen werden untersucht. Es wird festgestellt, daß Auswertalgorithmen, die das Tschebyscheffverfahren verwenden, nicht zu mathematisch eindeutigen Lösungen führen.

Sodann werden insbesondere die Auswertverfahren für Geradheitsmessungen anhand simulierter Oberflächen auf die Empfindlichkeit gegenüber Störeinflüssen untersucht.

Weiters wird die Problematik der Implementierung unterschiedlicher Auswertalgorithmen untersucht. Auf Grund dieser Gegenüberstellung werden Testoberflächen mit Auswertesoftware verschiedener Hersteller untersucht und die Ergebnisse analysiert.

In Rahmen der Arbeit wird der Arbeitsansatz formuliert, daß der ermittelte Verlauf der Geradheitsabweichungen mit dem gewählten Abstand der Meßpunkte in direktem Verhältnis steht.

Dieser Arbeitsansatz wird im Rahmen der Arbeit für die eingeführte technologische Teilung in sechs Gruppen bestätigt. Neben dem Wert der Geradheitsabweichungen wird auch der Mittelwert der Abweichungen untersucht. Aufgrund der durchgeführten Arbeit kann gesagt werden:

Der Verlauf der Abhängigkeit der Geradheitsabweichung bzw. deren Linearität korreliert stark mit den in der Oberfläche enthaltenen stochastischen Anteilen. Das mit vier Koeffizienten bestimmte Polynom konnte bei Herstellungsverfahren, die Oberflächen mit großem Zufallsanteil produzieren, nur mit einer geringeren Genauigkeit eingesetzt werden.

Die besten Ergebnisse wurden bei gefrästen Oberflächen erzielt. Es handelt sich dabei um typisch periodische Oberflächen mit überwiegend deterministischem Anteil, wie die Ergebnisse der Signalanalyse zeigen.

Der Mittelwert der Abweichungen hat sich entgegen den Erwartungen als nicht einfach vorgebbare Größe erwiesen. Es kann festgestellt werden, daß er gegenüber dem mit dem

Gauß'schen Verfahren bestimmten Wert der Abweichung keine entscheidende Verbesserung bringt.

Allgemein kann die Methode aber als ausbaufähig bezeichnet werden. Eine Steigerung der Genauigkeit kann mit den in der Arbeit vorgeschlagenen Mitteln gefunden werden, so daß sich ein ausführlicher Katalog der Oberflächen mit entsprechenden Polynomen festlegen läßt, was zu einer Verminderung der Unsicherheit der Meßergebnisse führen wird.

A Knowledge-Based Representation of the ID-Loop for System Identification

Mikael Snaprud

Begutachter: o.Univ.Prof. Dr. H. P. Jörgl und o.Univ.Prof. Dr. Dr.h.c. P. Kopacek

Practitioners in the field of control engineering generally have an idea of how to perform system identification. However, most of them are primarily interested in solving the identification problem, minimizing the overhead in doing so. For these and non-frequent users of control engineering software, a knowledge-based system (KBS) providing strategic advice is useful.

In order to build a KBS for system identification, a detailed description of the ID-loop is required. In this work such a description is presented. A clear and consistent model representing the ID-loop is derived in form of a triangle table. A way to use the ID-loop as a *plan* for performing system identification is developed. The most important components of the KBS are the ID-loop, and an interpreter.

An important issue in developing a KBS is acquiring the relevant knowledge. For knowledge acquisition an approach based on hypertext with types is developed and implemented. This approach allows incremental formalization of knowledge for developing a semiformal representation. By incrementally introducing formality this approach to knowledge acquisition helps to avoid premature decisions on formalisms. The formal elements in the semiformal hypertext, constituting the semiformal model, are nodes, links, partitions, and types. The typed semiformal representation is used as a basis for generating a program, much like a compiler generating executable code for a machine from source code. In contrast to a compiler, the presented approach automatically links the documentation to the formal representation.

By using the approach to knowledge acquisition outlined above, a knowledge-based representation of the ID-loop for system identification is developed. The functionality of the overall approach is demonstrated in an example.

**Regelungen, Analyse und technischer Entwurf.
Band 2: Multivariable, digitale und nichtlineare Regelungen;
optimale und robuste Systeme.**

Alexander Weinmann

Dritte, überarbeitete und erweiterte Auflage
Springer Verlag Wien, 1995

Das dreibändige Werk des Autors „Regelungen, Analyse und technischer Entwurf“ hat sich zwischenzeitlich bereits zu einem Klassiker entwickelt. Es entstand aus der langjährigen Vorlesungs- und Industrietätigkeit des Autors an der Technischen Universität Wien sowie bei der Firma ELIN.

Die nun vorliegende, dritte Auflage des zweiten Bandes behandelt im Sinne der Vorlesungen an der Technischen Universität Wien: Fortgeschrittene Verfahren, Zustandsregelungen, Abtastregelungen, Mehrgrößenregelungen, optimale Steuerungen und Regelungen, Kalman-Filter, nichtlineare, adaptive, robuste sowie Fuzzy-Regelungen. Ergänzt wird das Buch durch stochastische Verfahren und Stabilität.

Dieser anspruchsvolle Stoff wird überwiegend von einem praktischen Gesichtspunkt dargestellt. Theoretische Grundlagen und Verfahren werden nur soweit gebracht als dies unbedingt für das Verständnis des Stoffes erforderlich ist. Der theoretisch orientierte Leser wird daher ausführliche Herleitungen der dargestellten Verfahren und Methoden vermissen, dafür aber eine Reihe von praxisorientierten Beispielen vorfinden.

Das Buch kann sowohl den Studenten höherer Semester als auch dem in der Praxis stehenden Ingenieur wärmstens empfohlen werden und es bleibt zu hoffen, daß die nächste Auflage des Bandes drei ebenso gelingt wie diese.

P. Kopacek

Eine Abteilung stellt sich vor:

Abteilung für Regelungsmathematik, Hybridrechen- und Simulationstechnik am Institut 114 der TU Wien

I. Troch

Mathematische Probleme der Regelungstechnik sowie Fragen der Simulationstechnik wurden am Institut seit Mitte der 60er Jahre schwerpunktmäßig behandelt, was 1982 zur Gründung einer eigenen Abteilung führte. Diese Fachkombination ist nicht außergewöhnlich, wenn man bedenkt, daß Simulation bereits in ihrer "Kindheit und Jugend" als Analogrechen- bzw. Hybridrechentechnik ein wesentliches und unverzichtbares Hilfsmittel des Regelungstechnikers war und auch heute Simulation nicht aus Entwicklung und Erprobung wegzudenken ist.

Die Abteilung befindet sich im Gebäude „Freihaus“ der TU Wien (Postadresse: Wiedner Hauptstraße 8 - 10, A-1040 Wien), der Personalstand umfaßt eine ao. Professorin (Dr. techn. Inge Troch) sowie zwei Dozenten (Dipl. Ing. Dr.techn. Felix Breitenecker, Dipl. Ing. Dr.techn.Dr.rer.nat. Frank Rattay). Sekretariatsarbeiten werden in einem Gemeinschaftssekretariat mehrerer Institutsabteilungen erledigt.

In der Lehre werden - neben mathematischen Grundvorlesungen für Vermessungswesen und für Technische Mathematik - auf dem Gebiet der Simulation praktisch alle Studienrichtungen betreut. Wahllehrveranstaltungen in Form von Vorlesungen, Seminaren und Praktika führen in mathematische Modellbildung, Simulation und Simulationswerkzeuge sowie in spezielle mathematische Methoden (z. B. Optimierung, Stabilität) der Regelungstechnik ein. Hierzu kommen Lehrveranstaltungen aus dem biomedizinischen Bereich (Brain modeling, Hörtheorie, Nervenmodelle, Neurophysik, Regelungsmathematische Modelle in der Medizin). Spezialvorlesungen, Seminare und Arbeitsgemeinschaften vermitteln Kenntnisse aus einer umfangreichen Palette von Anwendungen, bieten vor allem den Diplomanden und Dissertanten wichtige Anregungen und fördern deren Integration in die Forschungsaktivitäten der Abteilung.

Die im folgenden skizzierten Forschungsaktivitäten sind in diesem Umfang nur dank der Mitwirkung von Diplomanden und Dissertanten möglich, sie werden in geringem Umfang durch den FWF gefördert und erfolgen teilweise auch in Zusammenarbeit mit anderen Instituten der TU Wien, der Universität Wien sowie der Universität Straßbourg. Die Ergebnisse dieser Arbeiten werden bei internationalen Tagungen präsentiert und sind in Kongreßberichten und Fachzeitschriften dokumentiert. Auf Wunsch senden wir gerne Sonderdrucke zu.

Forschungsaktivitäten:

Regelungsmathematik

Optimale Steuerungen und Regelungen waren und sind von großem Interesse, insbesondere wurden einige Arbeiten über die Beziehung zwischen mathematischer Modellierung von System, Anforderungen und Zielvorstellungen und der daraus resultierenden Steuerung bzw. Regelung bereits erfolgreich abgeschlossen, andere Fragen aus diesem Themenkreis sind in Bearbeitung. Eine Fortsetzung und Ergänzung finden diese Themen in Arbeiten zur Robotersteuerung sowie in der Entwicklung von Optimierungsoberflächen für Simulationssprachen. Zwei in Arbeit befindliche Dissertationen beschäftigen sich mit der Frage, wie moderne Softwaretools wie MATEMATICA in Verbindung mit der Ljapunovschen Stabilitätstheorie zur Berechnung bzw. Abschätzung des Einzugsbereiches stabiler Gleichgewichtslagen oder Bewegungen eingesetzt werden können, und zwar nicht nur - wie bisher in den meisten Fällen - für akademische Beispiele sondern für praktisch relevante Systeme.

Robotik

In der Robotersteuerung wurden Algorithmen für das zeit-optimale bzw. zeit-suboptimale Abfahren gegebener Bahnen entwickelt und werden derzeit - vor allem in Zusammenarbeit mit der ENSP Straßbourg erprobt. Weiters sind derzeit Roboter mit redundanten Freiheitsgraden von besonderem Interesse, bieten sie doch vor allem in Arbeitsräumen mit Hindernissen Vorteile. Hier konnte in einer Dissertation (vgl. Heft 2/1994) ein sehr einfaches Verfahren zur Kollisionserkennung direkt im Konfigurationsraum entwickelt werden, das in der Folge zur Entwicklung von Algorithmen zur Kollisionsvermeidung genutzt wurde. Eine Kombination dieser Methoden mit Bahnplanungsalgorithmen ist derzeit in Arbeit.

Simulationstechnik

Modellbildung und Simulation werden heute durch bequeme Software-Werkzeuge unterstützt. Für diskrete Prozesse, wie sie z. B. im Bereich der Fertigung bzw. der Ablaufplanung auftreten, wurde der Simulator D_SIM entwickelt, der in menügeführter Form das Erstellen von und Experimentieren mit Computermodellen erlaubt. Zielrichtungen dieses auf Petrinetzen basierenden Simulators sind einerseits Anwendungen im Bereich von Klein- und Mittelbetrieben, andererseits wird D_SIM als experimenteller Simulator zum Testen neuer Algorithmen etc. verwendet. D_SIM bietet eine Kombination aus graphischer Modellbildung und Animation an: das Modell wird graphisch aus Stationen und (Material-)Verbindungen erstellt. Bei der Simulation wird dasselbe Layout "animiert", wobei parallel zur Simulation die Veränderung aller Zustände mitverfolgt werden kann.

Im Bereich der kontinuierlichen Prozesse sind u. a. Erhöhung der Arbeits(Rechen-)geschwindigkeit der Simulatoren und Vereinheitlichung der Modellbeschreibung aktuelle Themen. Der ersten Richtung trägt die Entwicklung von Simulationssystemen auf Parallelrechnern Rechnung, dem zweiten Trend eine Miteinbeziehung differential-algebraischer Beschreibungsformen (Deskriptorsysteme der

Mechatronik). Diese Tendenzen wurden bei der Entwicklung des Simulationssystem mosis berücksichtigt, denn mosis kann wahlweise als paralleler Simulator oder als „normale“ Simulationssprache vom CSSL-Typ gesehen und verwendet werden. Mosis bedient sich des virtuellen PVM-Konzeptes und arbeitet je nach vorhandener Hardware (einfacher Rechner, Workstation Cluster, Transputersysteme) seriell oder parallel. Die Übernahme von Modellen aus anderen Simulationssprachen ist möglich.

Die Optimierungsoberfläche GOMA wurde für die weit verbreitete Simulationssprache ACSL entwickelt. Menügeführt kann der Benutzer nicht nur zwischen unterschiedlichen Optimierungsverfahren wählen, sondern bereits bei der Modellbildung sein Modell um Optimierungskomponenten (Gütefunktionen, Beschränkungen) erweitern. GOMA bietet nicht nur Parameteroptimierung, sondern auch die Optimierung von Steuerungsfunktionen (stetig, Treppenfunktion, Bang-Bang, etc.).

Besonders in der diskreten Simulation sind derzeit genetische Algorithmen als zielführende Algorithmen (rascher als Monte Carlo Methoden) in Diskussion. Dem trägt derzeit ein Projekt über die vernünftige Parametrisierung von genetischen Algorithmen für Optimierungsaufgaben mit beliebigen Parametern Rechnung.

Bewertende Vergleiche von Simulationssoftware sowie von Parallelisierungstechniken werden in Dissertationen bzw. im Rahmen eines Forschungsprojektes erarbeitet und liefern Grundlagen für Übersetzer zwischen Simulationssprachen bzw. für Simulatorkopplungen. Hilfestellung liefern dabei auch die von Doz. Breitenacker betreuten und in der Zeitschrift „EUROSIM Simulation News Europe“ ausgeschriebenene Comparisons (bisher sieben), die es auf 96 eingesandte Lösungen brachten, die in genormter Form veröffentlicht wurden bzw. werden.

Modellbildung und Simulation in der Biomedizin

Diese Forschungstätigkeit umfaßt verschiedene Bereiche vor allem der Grundlagenforschung. Sie ist auch vom regelungstechnischen Standpunkt in mehrfacher Hinsicht von Interesse, vor allem, weil das menschliche Gehirn auch als Rechenzentrale zur Regelung unserer Körperfunktionen fungiert. Da neurale Signale von elektrischer Natur sind, lassen sie sich durch geeignete elektrische Felder auch künstlich beeinflussen. Neben neurophysiologischer Grundlagenforschung, die unter anderem den Bereich der neuralen Signalverarbeitungen beinhaltet, werden vor allem die Möglichkeiten künstlicher Erzeugung von neuralen Aktivitäten mittels Elektroden untersucht. Diese Arbeiten sind besonders für die medizinische Anwendung im Bereich Lähmung, Inkontinenz, Ertaubung oder zur Schmerzbekämpfung von großem Interesse. Ein Einblick in die Theorie der künstlichen neuralen Erregung wird durch die Dissertation „Stimulating Effects of Electric Fields on Nerve Fibers“ ermöglicht, deren Kurzfassung in diesem Heft zu finden ist.

Hörtheorie

In einem anderen Forschungsbereich werden vor allem Fragestellungen zu den Grundlagen des Hörvorgangs behandelt. Viele biophysikalische Vorgänge wirken zusammen, bis ein akustisches Signal in binär codierter Form die 30.000 Fasern des Hörnerven als Datenleitung benützt, um in die

höheren Zentren des Cortex zu gelangen. Modellierung und Simulationen der mikro- und makromechanischen, sowie der elektrischen Vorgänge im Innenohr werden systematisch durchgeführt, um unser Verständnis über den Hörvorgang zu verbessern. Ein Schwerpunkt dieser Forschungen bezieht sich auf die Frage, in welcher Form Sprachsignale im Hörnerv übertragen werden. Erste Resultate zeigen, daß das Zeitmuster der Maximal- und Minimalwerte eines Sprachsignals für das Sprachverständnis wesentlich größere Bedeutung haben als bisher angenommen wurde und die komplizierte klassische Annahme des zeitveränderlichen Frequenzspektrums als Informationsträger für Sprachsignale nur von sekundärer Bedeutung ist. Damit zeichnet sich auch eine wesentliche Datenreduktion bei der digitalen Übertragung von Sprachsignalen ab.

Sonstige Aktivitäten:

Univ. Doz. DDr. Rattay leitet außerdem die Arbeitsgemeinschaft für Biomedizinische Technik an der TU Wien (Kurztitel TU-BioMed), die mit den beteiligten Mitarbeitern ihrer 25 Arbeitsgruppen ca 8 % des wissenschaftlichen Personals unserer Universität umfaßt und alle Hauptbereiche der biomedizinischen Technik sowie die dafür notwendigen wissenschaftlichen Grundlagen in Forschung und Lehre abdeckt. Das Ziel der TU-BioMed ist die Koordination der Lehr- und Forschungstätigkeit im Haus, die Information der Studierenden, die Präsentation der Forschungstätigkeit im außeruniversitären Bereich, die Vermittlung von Kooperationen, sowie die Förderung der Forschung.

Gemeinsam mit dem EDV-Zentrum werden im Rahmen der ARGESIM (ARGE Simulation News: F. Breiteneker, I. Husinsky, I. Troch, W. Weisz) erfolgreiche und gut besuchte Seminare über „Modellbildung und Simulation“ veranstaltet: im Jänner 1995 fand das 25. Seminar dieser Reihe statt.

Darüber hinaus ist Herr Doz. F. Breiteneker Präsident von EUROSIM, der Dachorganisation der Europäischen Simulationsgesellschaften und Herausgeber der EUROSIM Simulation News.

Frau Prof. Troch ist Editor-in-chief des *Journal on Mathematical Modelling* (swets, Lissej). Die Aktivitäten der Mitglieder der Abteilung werden abgerundet durch die Mitwirkung von Prof. Troch und Doz. Breiteneker bei der Herausgabe der wissenschaftlicher Zeitschriften *Mathematics and Computers in Simulation* (North-Holland, Amsterdam), *Systems Analysis, Modelling and Simulation* (Akademie Verlag, Berlin), *Simulation Practice and Theory* (Elsevier, Amsterdam), *Journal of Intelligent and Robotic Systems/ Theory and Applications* (Reidel Publ.Comp., Dordrecht), *Control -- Theory and Advanced Technology* (Mita Press, Tokyo), *Surveys on Mathematics for Industry* (Springer, Berlin), durch deren Tätigkeit als IPC-Mitglied bzw. als Session Organizer bei internationalen Tagungen sowie durch die Veranstaltung von wissenschaftlichen Tagungen (EUROSIM-Kongress, IMACS Symposien on Mathematical Modelling).

<i>Datum</i>	<i>Veranstaltung</i>	<i>Ort</i>	<i>Weitere Informationen erhältlich bei:</i>
28.9.1995	11. Österreichischer Automatisierungstag	<i>Wien</i>	Prof. Peter Kopacek Institut für Handhabungsgeräte u. Robotertechnik Floragasse 7a, 1040 Wien Tel.: 0222/504 18 35, FAX: 0222/504 18 35 9 Email: e318@ihrt1.ihrt.tuwien.ac.at
29.9.-1.10.95	IFAC Conference Supplemental Ways for Improving International Stability - SWISS	<i>Wien</i>	Prof. Peter Kopacek Institut für Handhabungsgeräte u. Robotertechnik Floragasse 7a, 1040 Wien Tel.: 0222/504 18 35, FAX: 0222/504 18 35 9 Email: swiis@ihrt1.ihrt.tuwien.ac.at
9.10.-11.10.95	IFAC Workshop Motion Control	<i>München</i>	VDI/VDE GMA Graf Recke Str. 84 D-40239 Duesseldorf, Germany FAX: (+49) 211/6214-161
4.10.-6.10. 1995	IFAC/IFIP/IFORS/IMACS Symposium Information Control Problems in Manufacturing - INCOM 95	<i>Beijing, China, P.R.</i>	Mrs. HeQing Li Institute of Manufacturing Systems Beijing University of Aero-Astro POB 85 37 Xue Yun Road, Hidian District Beijing 100083, China, P.R. FAX: +861/2015347
24.10. - 26.10 1995	IFAC Workshop Intelligent Manufacturing Systems - IMS 95	<i>Bucharest, Romania</i>	University Politehnica Bucharest Faculty of Control & Computers IMS'95, corp. ED Automatica Spl. Independentei 313, sector 6 R-77206 Bucharest, Romania FAX +40/1/3126 335 Email: ims95@ulise.cs.pub.ro
8.11. - 11.11 1995	IFIP/IFAC Working Conference Models and Methodologies for Enterprise Integration	<i>Heron Islands QLD, Australia</i>	Angela Laffey, Conference Secretary Australian Computer Society POB 135 Aspley (Bribane) Queensland 4034, Australia FAX +61/7/263 7020 Email: ei95@cit.gu.edu.au
29.11. - 1.12. 1995	IFAC Workshop Artificial Intelligence in Real-Time Control	<i>Bled, Slovenia</i>	Dr. Jus Kocijan Faculty of El. & Computer Engg. Univ. of Ljubljana, Trzaska 25, SL - 61000 Ljubljana, Slovenia FAX +386/61/264 991 Email: Jus.Kocijan@fer.uni-lj.si
1.7. - 5.7. 1995	IFAC WORLD CONGRESS	<i>San Francisco, California, USA</i>	IFAC'96 POB 111, Mabelton GA 30059 USA Phone: +1/708/491-3641 FAX: +1/708/491-4455 Email: ifac96@nwu.edu