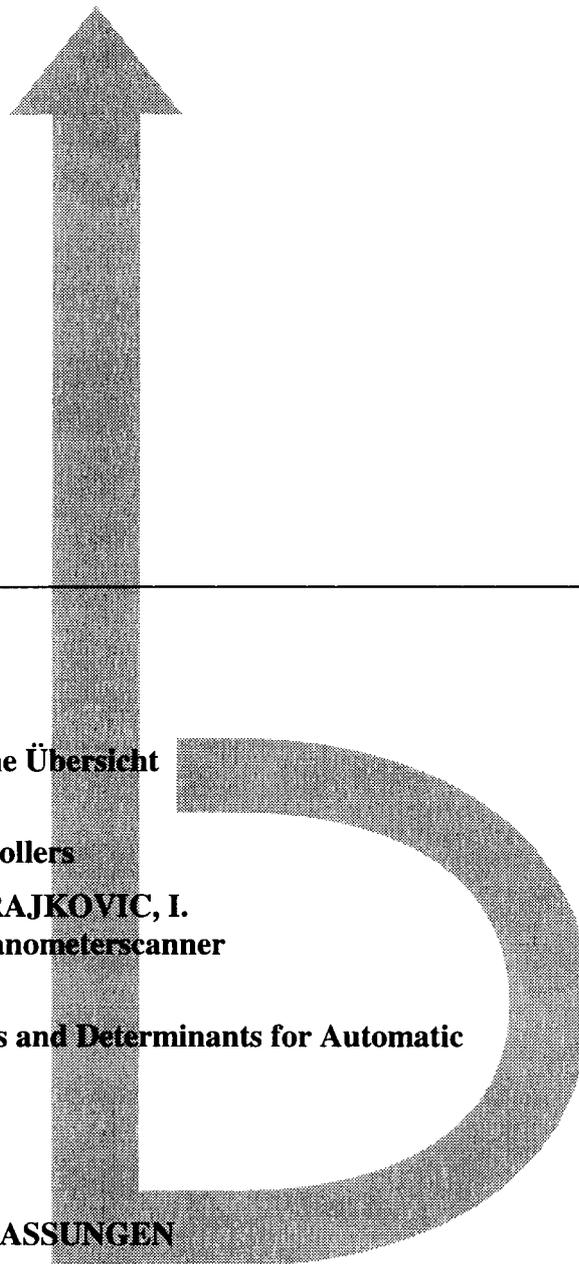
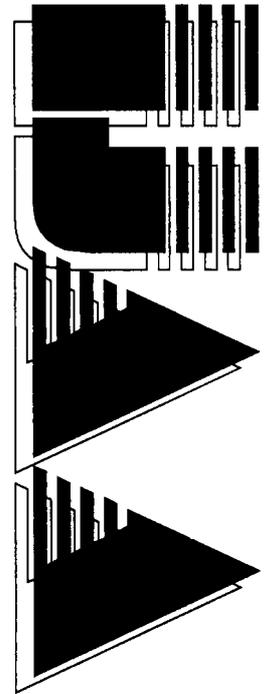


# INTERNATIONAL JOURNAL

HEFT 1/2

Jg. 9, (2001)

# AUTOMATION AUSTRIA



## INHALT

	<i>Seite</i>
<b>GENSER, R.</b> Lernfähige Verfahren – eine Übersicht	1
<b>WEINMANN, A.</b> Sensitivity of Holistic controllers	13
<b>ROTHWANGL, H. P. u. RAJKOVIC, I.</b> Optimal Control of a Galvanometerscanner	27
<b>WEINMANN, A.</b> Gradients of Norms, Traces and Determinants for Automatic Control Applications	36
<b>SCHLAGLICHTER</b>	51
<b>BUCHBESPRECHUNG</b>	56
<b>DISSERTATIONSKURZFASSUNGEN</b>	65
<b>VERANSTALTUNGEN</b>	68

# Lernfähige Verfahren - eine Übersicht

Robert Genser, ÖVE, Senior Member IEEE  
IFAC-Beirat Österreich

Eingelangt am 21. Juni 2001

## Kurzfassung

*Es wird ein Überblick über technische lernfähige Verfahren gegeben. Der Vorteil einer lernfähigen Speicherung von Daten bei einem ingenieurmäßige Ansatz gegenüber den neuronalen Netzen wird dargelegt.*

## Abstract

*A survey on learnable procedures is given. The advantage of storing data with an engineering approach is pointed out in comparison to neural networks.*

## 1 Einleitung

Mit dem Fortschritt elektronischer Schaltkreise und der Beschäftigung mit kybernetischen Systemen im Sinne Norbert Wieners gab es bereits anfangs der 60er Jahre einen Boom an Publikationen über Lösungsvorschläge für technische lernfähige Verfahren. Zunächst für Mustererkennungsprobleme eingesetzt, erkannte man aber auch in der Regelungstechnik die Vorteile lernfähiger Systeme gegenüber reinen adaptiven Verfahren.

Erst der Preisverfall und die Leistungssteigerung bei den elektronischen Bauteilen beziehungsweise der Durchbruch bei den programmierbaren elektronischen Systemen ermöglichte den heutigen breiten Einsatz von lernfähigen Verfahren. Diese als Software verfügbaren Verfahren bilden derzeit überwiegend die Signalverarbeitung von lernfähigen Nervensystemen nach und nennen sich daher auch **Neuronale Netze** (Braun, H., et al., 1996).

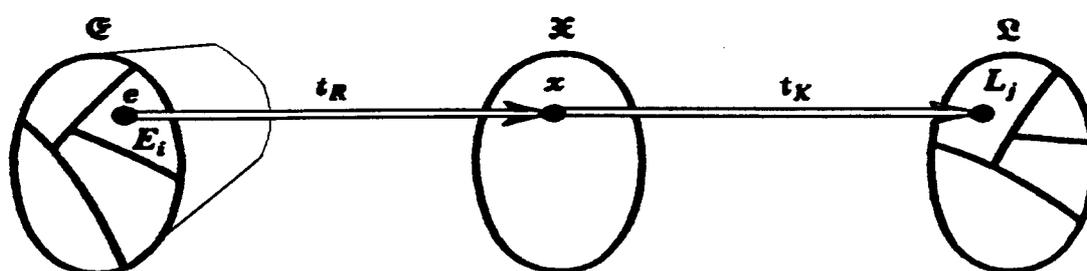
Sicherlich, die Technik lernte und lernt viel von der Natur. Aber die Bequemlichkeit der verfügbaren Software sollte ein ingenieurmäßiges Hinterfragen für eine optimale Problemlösung nicht vereiteln. Die reine Nachahmung des Vogelfluges hätte nie eine Concorde ermöglicht.

## 2 Lernfähige Verfahren

Das Erinnern, der Zugriff auf die Erfahrung aus der Vergangenheit, und das Erlangen neuer Erfahrung ist die wichtigste Eigenschaft eines Lernens. Diese Erfahrung kann transformiert oder

verdichtet sein. Bei gegebenen Zuständen bewirkt das Erinnern zugeordnete Informationen (Wissen) oder Verhalten (zum Beispiel Geschicklichkeit).

Das Lernen (Zemanek, H., 1967) geschieht durch **einfaches Abspeichern der Erfahrung** - zum Beispiel Auswendiglernen oder Nachahmen beziehungsweise Entwicklung von Expertensystemen (Thuy, N., H., C., 1989) oder Entscheidungstabellen (Rammé, M., 1996) - durch **selektive Verstärkung** - Versuch und Erfolg, aber auch die Verfahren der neuronalen Netze - und durch **Verallgemeinerung** - wie Assoziation und Erfassen oder Such- und Entscheidungsverfahren.



$\mathcal{E}$  Objektraum

$E_i$   $i$ -te Objektklasse

$e$  Objekte, welche zu klassifizieren sind,  $E = \bigcup_{i=1}^m E_i$

$\mathcal{X}$  Datenraum

$x$  Daten,  $x = \{x_1, \dots, x_n\}$

$\mathcal{L}$  Unterscheidungsraum

$L_j$  Klasse,  $\mathcal{L} = \bigcup_{j=1}^u L_j$

$t_R$  Transformation von  $e$  in  $x$  durch den Rezeptor  $t_R: e \rightarrow x$

$t_K$  Transformation von  $x$  in  $L_j$  durch den Klassifizierer  $t_K: x \rightarrow L_j$

Abb. 1: Modell eines Klassifizierers

Der Kern eines lernfähigen Systems ist die **Klassifiziereinrichtung**. Der Rezeptor bildet aus dem Objekt  $e$ , welches ein bestimmtes beschreibbares Ereignis oder ein beschreibbarer Zustand der Umwelt sein kann, die zur Verfügung stehenden Daten  $x = \{x_1, x_2, \dots, x_n\}$ . Die zugehörige Klasse  $L_j$  im Unterscheidungsraum  $\mathcal{L}$  soll durch den Klassifizierer  $t_K$  den jeweiligen Daten  $x$  zugeordnet sein. Bei lernfähigen Klassifizieren wird die entsprechende Transformationsfunktion  $t_K$  während des Lernvorgang entwickelt, siehe Abb. 1. Aus ökonomischen und technischen Gründen werden aber auch die vom Rezeptor zu erfassenden Daten in einer Lernphase zu ermitteln sein (Genser, R., 1969b).

Den Daten  $x$  können Zahlenwerte, zum Beispiel Meßwerte in ° Celsius oder Volt, aber auch Kategorien, zum Beispiel Winter, Vollast oder Nennspannung, beziehungsweise Eigenschaften zugeordnet sein. Nebenbei bemerkt ist bei Metaphern wie zum Beispiel "Nennspannung" zu beachten, daß diesen nicht ein fixer Meßwert zugeordnet sein muß. Diese Meßwerte können innerhalb eines Toleranzbandes liegen, abgesehen davon, daß diese Nennspannung eine Sinus-Schwingung mit zulässigen Oberwellen, Flicker usw. sein kann. Im allgemeinen Fall bildet daher schon der Rezeptor aus einigen Meßwerten Klassen oder Quantisierungsstufen für die entsprechenden Daten  $x$ .

Die geschichtliche Entwicklung der Problemabstraktion und die bisherige Leistungsfähigkeit der Mathematik ließ auch bei lernfähigen Verfahren mathematische Paradigmen bevorzugen. Die grundsätzliche Struktur eines **neuronalen lernfähigen Elements (LE)** ist in Abb. 2 dargestellt.

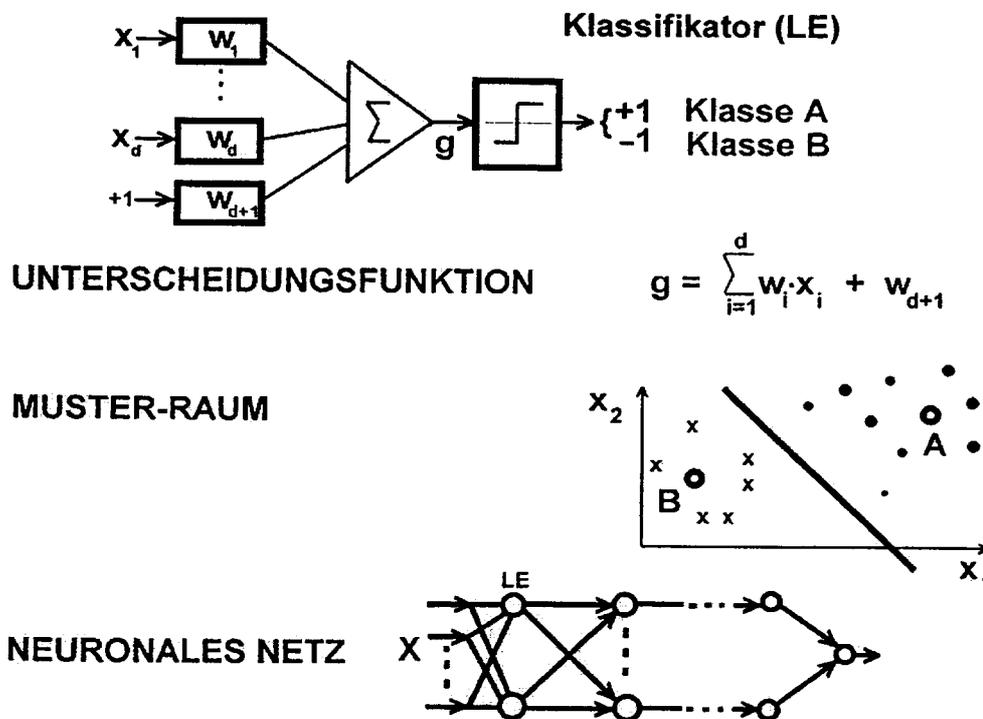


Abb. 2: Neuronales Modell

Die Daten  $x$  werden mit Gewichten  $w$  beaufschlagt, die während einer Lernphase veränderbar sind. Ein Schwellwertglied oder die Signum-Funktion (Mars, P., et a., 1996) ergeben zum Beispiel den Wert  $+1$  für Klasse A oder  $-1$  für Klasse B in Abhängigkeit des Wertes der Unterscheidungsfunktion nach dem Summationsglied. Bei einer dichotomen Klassifizierung in Abhängigkeit von nur zwei Parametern  $x$ , die als skalare Größen charakteristische Eigenschaften eines Objektes beschreiben, ist die Entscheidungsfunktion  $g$  eine Gerade.

Nach diesem grundsätzlichen Denkmuster funktionieren vom Perceptron (Rosenblatt, F., 1962) bis Lernmatrix (Steinbuch, K., et al., 1965) die heute überwiegend angewandten Verfahren, siehe zum Beispiel (Fiesler, J., et al., 1997) und (Weinmann, A., 1998). Statt einer Stimulierung mit Summenbildung wurde auch eine Multiplikation der beaufschlagten Daten  $x$  vorgeschlagen (Tou, J., T., 1970). Um das Black-Box-Problem der neuronalen Netze zu mindern werden auch Verfahren nach Bayes eingesetzt, (Westland, J., Ch., 1995) und (Lam, W., et al., 2000), bei welchen mit Hilfe bedingter Wahrscheinlichkeiten a priori Wissen besser berücksichtigt werden kann. Auch durch den Einsatz von genetischen und evolutionären Verfahren (Yao, X., 1999) versucht man die grundsätzlich neuronalen Netze zu verbessern. Die Gruppenmethode für Datenhandhabung (Ivachnenko, A., G., et al., 1984) versucht mittels Polynomen eine heuristische Selbstorganisation mathematischer Modelle.

Komplexe Entscheidungsprobleme lassen erkennen, Verfahren, die mittels Gewichtung versuchen einen einzelnen Funktionswert zu erhalten – wie dies auch bei der klassischen Kriterienfunktion der Regelungstechnik gehandhabt wird – sind inflexibel und nicht einfach durchschaubar (Genser, R., 1986). Die Wichtigkeit einzelner Parameter  $x$  kann vom dynamischen Systemzustand oder jeweiligen Kontext abhängen, aber auch die Ziele selbst können sich ändern. Gewichtsfunktionen und Entscheidungsgrenzflächen statt eingespeicherter tatsächlicher Muster bewirken einen Informationsverlust.

Von Seiten der künstlichen Wissensverarbeitung (Nilsson, N., J., 1982) beziehungsweise der elektronischen Datenverarbeitung wurden Lösungen angestrebt, die den Computer über eine Rechenmaschine hinausgehend bei einer Datenverarbeitung zur Entfaltung bringen. Bei Verfahren mit Datenspeicherung wird bei einem vorliegenden Bearbeitungsfall die Ähnlichkeit mit den gespeicherten Fällen verglichen und daraus die Schlüsse gezogen, siehe zum Beispiel (Stanfill, C., et al., 1986) und (Winston, P., H., 1980). Die Selektion kann nicht nur über festgelegte Schwellwerte bei den einzelnen Daten erfolgen, sondern es werden eine bestimmte Anzahl der engsten Übereinstimmungen ausgewählt. Im Bereich der Bildverarbeitung hat man verschiedene Lösungsansätze entwickelt, (Haberäcker, P., 1985) und (Pinz, A., 1986). Verfahren für das Vergleichen von Datenketten findet man zum Beispiel in (Bunke, H., et al., 1995). Statistische Bearbeitungen lassen sich anwenden. Aber auch Sortierungen oder Klassifizierungen hinsichtlich Zusammenhang (Lance, G., N., et al., 1965) oder Ähnlichkeit (Lance, G., N., et al., 1966) können durchgeführt werden, (Späth, H., 1975) und (Hartigan, J., A., 1975). Bei Abweichungen des Bearbeitungsfalls von den gespeicherten Daten sind numerische Korrekturen möglich, die durch die Linearisierung infolge der Klassenbildung oder Quantisierung sehr einfach sein können. Bei Fehlentscheidungen werden die Daten des jeweiligen Bearbeitungsfalls als neues Muster eingespeichert.

Es wird auch das maschinelle Erlernen von Konzepten (Sarkar, S., et al., 1998) versucht, ohne daß von außen programmiert werden muß (Valiant, L., G., 1984). Das bedeutet lernfähige automatische Gestaltung von Programmen und Problemlösungsverfahren.

Unter Berücksichtigung der notwendigen Flexibilität bei der Problemstellung und den Möglichkeiten heutiger Datenverarbeitungseinrichtungen wäre eine **objektorientierte Beschreibung** (Rumbaugh, J., et al., 1991) naheliegend. Programmiersprachen wie ADA

(ISO/IEC 8652, 1995) oder C++ (ISO/IEC 14882, 1998) wurden für diesen Ansatz entwickelt. Das klassische mathematische Modell wird immer mehr von **objektorientierten formalen Methoden** ersetzt. Basierend auf die Grundidee der Abstrakten Objekte (Zemanek, H., 1968) findet die Vienna Development Method eine immer breitere Anwendung, siehe (Jones, C., B., 1990), (ISO/IEC 13817, 1997) oder (<http://www.ifad.dk/products/vdmlangchar.html>).

### 3 Lernfähige Datenspeicherung

In der **Vorbereitungsphase** wird eine sich technisch und wirtschaftlich eignende Informationserfassung, Informationsübertragung, Art der Verarbeitung und das Lernverfahren selbst ermittelt. Bei autonomen lernfähigen Systemen wird man aber aus technischen und wirtschaftlichen Gründen auch schon in dieser Phase untersuchen, welche Parameter von Einfluß sind und in welchem Umfang diese auch zu berücksichtigen sind (Genser, R., 1969b), (Kanal, L., 1974).

Bei der Verarbeitung der Daten können zum Beispiel dynamische Korrelationsverfahren, siehe (Lange, F., M., 1962) und (Schlitt, H., 1960), Transformationen (Torres-Méndez, L., A., et al., 2000), Filter (Schüssler, H.-W., 1973), Interpolations-, Prognose- und Identifikationsverfahren (Isermann, R., 1989) zur Anwendung kommen.

Die **Betriebsphase** einer lernfähigen Datenspeicherung umfaßt:

- Datenerfassung
- Informationsverarbeitung
- Musterzuordnung
- Korrekturverfahren bei Abweichung der gespeicherten Muster gegenüber vorliegenden Daten
- Ergebnisauswertung
- Speichern der vorliegenden Daten mit tatsächlichem Ergebnis (neues Muster) bei unzulässigem Fehler.

In der Betriebsphase besteht das Lernziel bei den anfallenden Daten richtige Schlüsse zu ziehen. Bei selbstorganisierenden lernfähigen Systemen können aber auch Lernziele bestehen, die Aufgabenbereiche der Vorbereitungsphase umfassen, zum Beispiel das Erlernen einer optimalen Informationsübertragung. Im Gegensatz zu einfachen Klassifizierverfahren wird sich die Lern- und Kannphase überlappe.

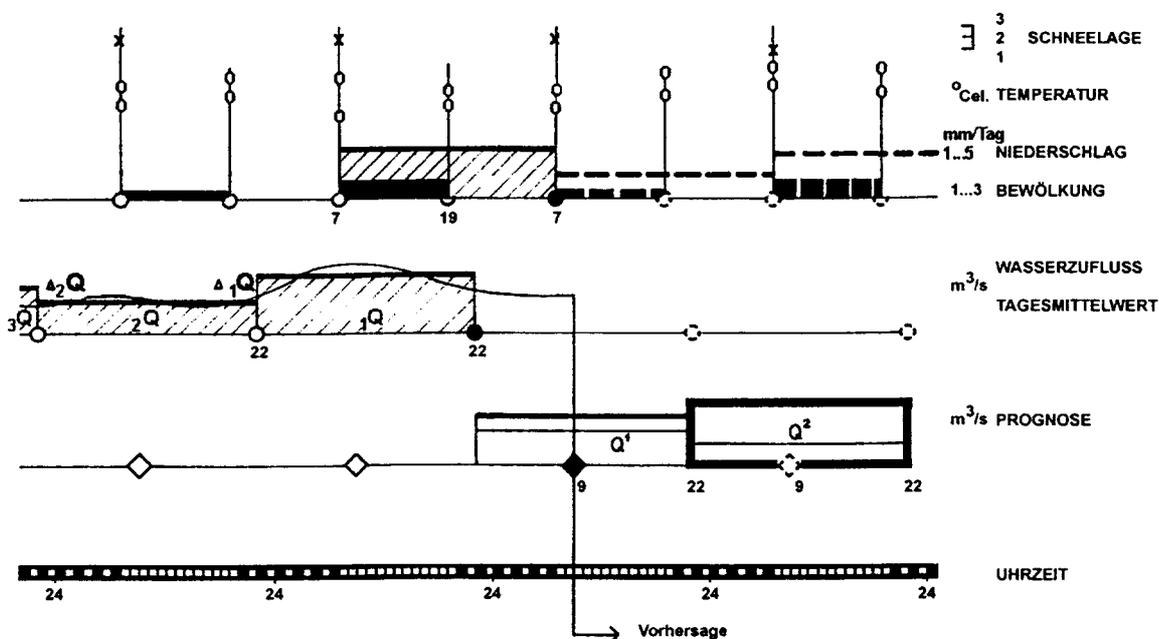
Ein Lehrling beginnt beim Lernen nicht im Zustand Null, sondern bekommt schon zu Beginn vorhandenes Wissen übermittelt, abgesehen von der evolutionär erworbenen genetischen Erfahrung. Daher wird man beim maschinellen Lernen auch nicht mit einer Black-Box, sondern mit einem möglichst hohen Wissensstand den Lernprozeß starten wollen.

So wird man beim Erlernen der Zuordnung des Geschlechts zum Vornamen im deutschen Sprachraum das System zuerst nach den Endbuchstaben A und E abfragen lassen. Bei Zutreffen wird eine Ausnahmeliste 1 durchsucht ( zum Beispiel Attila ). Anderenfalls wird das System in

der Ausnahmeliste 2 ( zum Beispiel Ingrid ) nachsehen, bevor die Zuordnung "männlich" erfolgt. Bei Fehlentscheidungen wird der betreffende Name in der jeweiligen Ausnahmeliste gespeichert beziehungsweise durch Abfragen um zusätzliche Merkmale erweitert, zum Beispiel im Falle von Maria kombiniert mit Männernamen.

#### 4 Lernfähiges Verfahren für eine Zuflußvorhersage

Das lernfähige Verfahren für eine Zuflußvorhersage soll den Vorgang bei einer lernfähigen Datenspeicherung aufzeigen. Der Anwendungsfall ist in (Genser, R., 1969a) beschrieben.



$Q^2$  Vorhersagewert des 24-Stunden-Mittelwerts des Wasserzuflusses

□

Abb. 3: Zeitraster der 36-Stundenvorhersage für einen 24-Stunden-Mittelwert des Zuflusses

Abb. 3 veranschaulicht die verwertbaren Daten. Schneelage, Bewölkung und Regenvorhersage sind verbale Daten. Die Schneelage, die für über 1500 m und darunter erhoben wird, hat die Stufe 1 für kein Schnee, Stufe 2 für teilweise schneefrei und Stufe 3 für geschlossene Schneedecke. Bedeckt bedeutet Stufe 1, teilweise bedeckt oder dünne Wolkendecke entspricht Stufe 2 und klar ergibt Stufe 3. Bei der mündlichen Wettervorhersage hat der Niederschlag die Stufe 1 bei Schönwetter, 2 bei Nieselregen, 3 bei Regen, 4 Landregen und 5 bei Wolkenbruch. Aber auch die Meßwerte für den Regen werden insgesamt 5 Stufen zugeordnet, wobei eine Regenmenge von über 70 mm Tagesmittelwert der Stufe 5 entspricht. Die Meßwerte kommen aus 3 Gebieten.

Die Temperatur von unter 1500 m und von über 1500 m unter  $-2^{\circ}\text{C}$  entspricht der Stufe 1,  $-2$  bis  $+4^{\circ}\text{C}$  der Stufe 2 und über  $+4^{\circ}\text{C}$  der Stufe 3. Beim Musterselektionsprozeß werden Temperaturmaximum und Minimum mit den Gewichten  $t_{\max} = 3$  und  $t_{\min} = 1$  entsprechend Tab.1 auch in 3 Klassen umgewandelt.

Temperatur max.	min.	Entscheidungswert $\Sigma = TS_{\max} * t_{\max} + TS_{\min} * t_{\min}$	Klasse
1	1	$1*3 + 1*1 = 4$	<b>I</b>
2	1	$2*3 + 1*1 = 7$	<b>I</b>
2	2	$2*3 + 2*1 = 8$	<b>II</b>
3	1	$3*3 + 1*1 = 10$	<b>II</b>
3	2	$3*3 + 2*1 = 11$	<b>III</b>
3	3	$3*3 + 3*1 = 12$	<b>III</b>

Tab.1: Zuordnung der Temperaturstufen zu Klassen

Andere Kombinationen sind nicht relevant, da  $\max. < \min.$  nicht möglich ist.

Bei der Suche nach einem passenden Muster wird zuerst die Selektion nach diesen Temperaturklassen durchgeführt und erst danach nach den Temperaturstufen für die gemessenen Temperaturwerte. Dadurch werden für den Zufluß maßgebende Temperaturkombinationen besser berücksichtigt.

${}_1Q$  ist der Tagesmittelwert (22 Uhr bis 22 Uhr) des Zuflusses vom Vortag. Die Änderung gegenüber den vorhergehenden Tag ist

$$\Delta {}_1Q = {}_1Q - {}_2Q$$

$\Delta\Delta {}_1Q$  ist positiv, wenn der Anstieg des Zuflusses größer wird oder das Absinken sich vermindert.

$$\Delta\Delta {}_1Q = {}_3Q - {}_1Q + 2*\Delta {}_1Q$$

Abb. 4 zeigt das Schema eines gespeicherten Musters. Das bei einer Mustersuche verfügbare Speicherfeld umfaßt grundsätzlich  $\pm 15$  Tage vom Jahrestag. Es kann eine Vorselektion hinsichtlich trockenem, normalen oder feuchtem Jahr erfolgen.

Für den Suchvorgang werden für  ${}_1Q$ ,  $\Delta {}_1Q$  und  $\Delta\Delta {}_1Q$  Schranken  $S$ ,  $\Delta S$  und  $\Delta\Delta S$  festgelegt. Vor der Mustersuche hinsichtlich Wettervorhersage erfolgt die Auswahl in einem breiten Band der Zuflußdaten, da das Wetter eine stärkere Auswirkung als die genauen Zuflußwerte hat. Durch die Ungenauigkeit einer Wettervorhersage soll aber doch schon eine Vorselektion hinsichtlich des Zuflußverhaltens erfolgen. Man benötigt die Toleranzgrenzen auch für ein eventuell erforderliches Korrekturverfahren. Zum Beispiel gelten für 100- 400  $\text{m}^3/\text{s}$  die

Toleranzgrenzen  $S = \pm 50 \text{ m}^3/\text{s}$ .

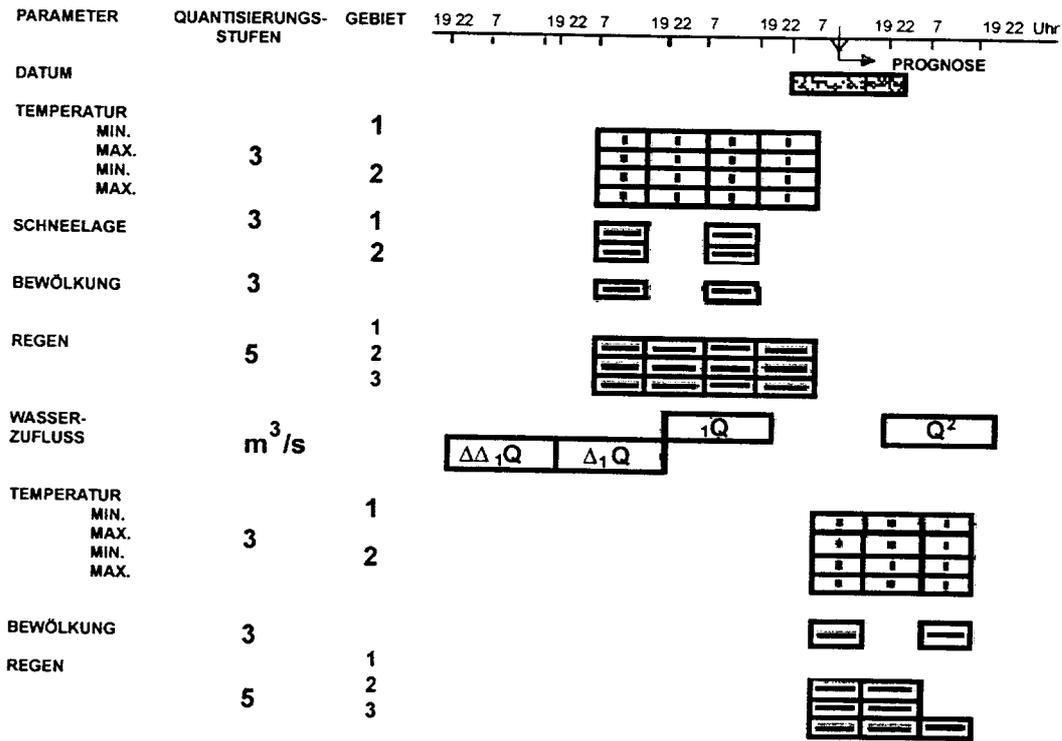


Abb. 4: Gespeichertes Muster

Abb.5 beschreibt schematisch den Ablauf der lernfähigen Zuflußvorhersage. Für die Mustersuche der Kategorien der gemessenen Temperaturen wurde das Schema für die Suche eines passenden Musters für die Temperaturklassen detailliert dargestellt. Mit dem selben Verfahren wird im Block für die Temperaturkategorien nach den Klassen nach passenden Vergangenheitsmuster für die Temperaturstufen gesucht.

Die Mustersuche in den anderen Blöcken erfolgt in ähnlicher Art und Weise.

Da nicht jede Gegebenheit gespeichert werden soll und die jeweiligen Zustände sich nicht immer

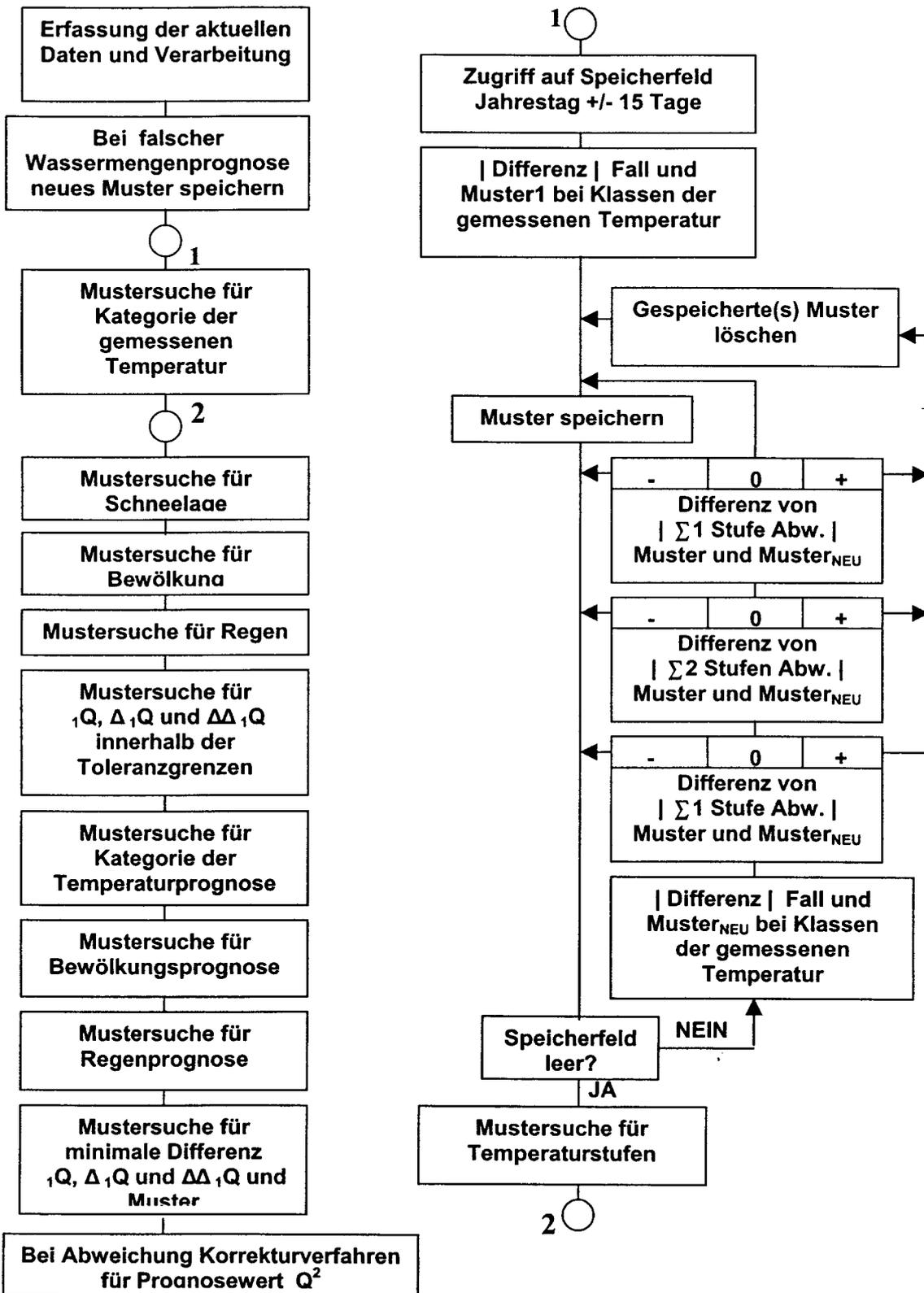


Abb. 5: Ablaufschema einer lernfähigen Wasserzufluß-Vorhersage

wiederholen, werden Abweichungen vorkommen. Ist diese Abweichung nicht zulässig, so wird das Ergebnis vor Ausgabe der Zuflußprognose  $Q^2$  korrigiert. Ist die Abweichung zwischen Fall  ${}_1Q$  und Musterwert größer als die Toleranz  $S$ , ermittelt man

$$p = {}^2Q_T / {}_1Q$$

${}^2Q_T$  ist jener Wert  $Q^2$ , der einem  ${}_1Q$  bei keinem Regen und minimaler Schneeschmelze über 1500 m entspricht (z. B.  $p = 0,9$  bei  $161 - 300 \text{ m}^3/\text{s}$ ).

$p_M$  wird für das Muster  ${}_1Q_M$  entsprechend  $p$  ermittelt.

Bei unzulässigem Muster bei den Zuflußänderungen  $\Delta\theta$  und  $\Delta\Delta\theta$  erfolgen die Korrekturen  $\partial_\Delta$  und  $\partial_{\Delta\Delta}$ .

Beim Wetter werden die Korrekturen  $\partial_w$  in Abhängigkeit der Stufenabweichungen durchgeführt.

Der Prognosewert für den Zufluß ergibt sich im Falle notwendiger Korrekturen:

$$Q^2 = Q_M^2 + p \cdot {}_1Q - p_M \cdot {}_1Q_M + \partial_\Delta + \partial_{\Delta\Delta} + \sum_{i=1}^n \partial_{w_i}$$

Bei entsprechend guten Korrekturverfahren wird das erforderliche Einspeichern neuer Muster zur Vermeidung falscher Zuflußprognosen sehr vermindert. Die Lernfähigkeit des Verfahrens ermöglicht aber eine laufende Anpassung an neue Gegebenheiten. Im Gegensatz zu nicht-lernfähigen Prognosemodellen ist aber auch ein direkter Zugriff auf die erworbene Erfahrung bei Gegebenheiten der Vergangenheit möglich.

Das im Beispiel dargelegte Verfahren wäre auch bei einer Musterspeicherung auf Lochstreifen, trotz der Unzulänglichkeiten dieses Speichermediums, anwendbar. Durch die Anwendung von Korrekturverfahren und den in dem gegebenen Fall erlaubten Toleranzen ist die Anzahl der zu speichernden Muster beschränkt und außerdem ist nur jeweils ein Fenster von drei Monaten eines charakteristischen Jahres für die Verarbeitung zur Verfügung zu stellen.

Dieser Ansatz erlaubt in jeder Phase des Problems angepaßte verschiedenartige Verfahren und es kann jederzeit auf die gespeicherten Grunddaten zugegriffen werden. So werden in diesem Beispiel Klassifizierungen, Quantisierungen und Rechenformeln auch in den Zwischenschritten benutzt. Bei neuronalen Netzen ist innerhalb des Klassifiziervorganges zwischen den Schichten aus Gründen der geringen Durchschaubarkeit und des Verlorengehens der für die Entscheidung maßgebenden Grundmuster keine derartige Flexibilität gegeben. Denn statt der Muster stehen bei diesen nur Entscheidungsflächen zur Verfügung.

## 5 Ausblick

Durch die Steigerung der Kapazität von Speichermedien, die Erhöhung der Speicherdichte und Zugriffsgeschwindigkeit neben dem Preisverfall ist die Verwirklichungsmöglichkeit für

lernfähige Verfahren mit Speicherung von Musterdaten wesentlich erleichtert worden. Obwohl schon lange darüber berichtet wird (Casasent, D., 1981), haben aber die eigentlich für Mustererkennungsverfahren als vorteilhaft erscheinenden optischen Techniken noch keinen Durchbruch erzielt.

Die Komplexität der Probleme der Wirklichkeit erfordern flexible und durchschaubare Verfahren. Von den Software-Entwicklungsprozessen bis zu umfassenden Optimierungs- und Entscheidungsaufgaben wird immer mehr versucht, vorhandene Information nicht durch spitzfindige mathematische Verfahren zu verdecken und dadurch notwendige Anpassungen an neue Gesichtspunkte zu erschweren (Hopkins, J., F., et al., 2001).

Diese Aufgabenstellungen erfordern aber ein ingenieurmäßiges Vorgehen.

## 6 Literaturverzeichnis

- Braun, H., Feulner, J., Malaka, R.: *Praktikum neuronale Netze*, Berlin: Springer, 1996
- Bunke, H., Csirik, J.: Parametric String Edit Distance and its Application to Pattern Recognition, *IEEE Trans. on Systems, Man, and Cybernetics*, 1995, Vol. 25, No. 1, pp. 202-206.
- Casasent, D.: Pattern recognition : a review, *IEEE Spectrum*, 1981, March, pp. 28-33
- Fiesler, E., Biele, R.: *Handbook of Neural Computation*, New York: Oxford Press, 1997
- Genser, R.: Lernfähige Verfahren für die Zufluvorhersage bei Laufkraftwerken, *EundM*, 86(1969) H. 4, 155-159
- Genser, R.: Die Aufbereitung der Information bei lernfähigen Automaten, *Elektron. Rechenanl.*, 11(1969) H.6, 330-335
- Genser, R.: Learning ,in Decision Making. In: Fandel, G. (Ed.), *Large-Scale Modelling and Interactive Decision Analysis*, *Lecture Notes in Economics and Mathematical Systems*, Vol. 273, S. 138-147, Berlin: Springer-Verlag, 1986
- Haberäcker, P.: *Digitale Bildverarbeitung, Grundlagen und Anwendung*, München; C. Hauser Verlag, 1985
- Hartigan, J., A.: *Clustering Algorithms*, New York: Wiley, 1975
- Hopkins, J., F., Fishwick, P., A.: A Three-Dimensional Human Agent Metaphor for Modeling and Simulation, *Proc. of the IEEE*, 2001, Vol. 89, No. 2, pp. 131-147
- Isermann, R.: *Identifikation dynamischer Systeme*, Bd. 1 und 2, Berlin: Springer-Verlag, 1989
- ISO/IEC 8652: *Information technology - Programming languages - Ada*, 1995
- ISO/IEC 13817: *Information technology - Programming languages, their environments and system software interfaces - Vienna Development Method - Specification Language*, 1996
- ISO/IEC 14882: *Programming languages - C++*, 1998
- Ivachnenko, A., G., Müller, J.-A.: *Selbstorganisation von Vorhersagemodellen*, Berlin: Verlag Technik, 1984
- Jones, C., B.: *Systematic Software Development using VDM*, Englewood Cliffs, N.J.: Prentice-Hall, 1990
- Kanal, L.: Patterns in Pattern Recognition: 1968-1974, *IEEE Trans. on Information Theory*, 1974, Vol. 20, No. 6, pp. 697-7212
- Lam, W., Low, K., F.: Using Discretization and Bayesian Inference Network Learning for Automatic Filtering Profile generation, *IEEE Trans. on Systems, Man, and Cybernetics*, 2000, Vol. 30, No. 3, pp. 340-351

- Lance, G., N., Williams, W., T.: Computer Programs for Monoethic Classification, *The Computer Journal*, 1965, Vol. 8, No. 3, pp. 246-249
- Lance, G., N., Williams, W., T.: Computer Programs for Hierarchical Polythetic Classification, *The Computer Journal*, 1966, Vol. 9, No. 1, pp. 60-64
- Lange, F., M., *Korrelationselektronik*, Berlin: VEB Verlag Technik Berlin, 1962
- Mars, P., Chen, J.R., Nambiar, R.: *Learning Algorithms: Theory and Applications in Signal Processing, Control and Communications*, Boca Raton, FL: CRC, 1996
- Nilsson, N., J.: *Principles of Artificial Intelligence*, Berlin: Springer-Verlag, 1982
- Pinz, A.: Das Vision-Expert-System "VES". In: Kropatsch, W., G., Mandl, P. (Hrsg.), *Mustererkennung '86, Bildverarbeitung in den Geowissenschaften Bildpyramiden*, S. 74-91, Wien: R. Oldenbourg, 1986
- Rammé, M.: *Entscheidungstabellen*, Prentice Hall, München, 1996
- Rosenblatt, F.: *Principles of Neurodynamics*, Washington, DC: Spartan Book, 1962
- Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorenzen, W.: *Object Oriented Modelling and Decision*, Englewood Cliffs, NJ: Prentice-Hall, 1991
- Sarkar, S., Chakrabarti, P., P., Ghose, S.: Learning While Solving Problems in Best First Search, *IEEE Trans. on Systems, Man, and Cybernetics, Part A*, 1998, Vol. 28, No. 4, pp. 535-542
- Schlitt, H.: *Systemtheorie für regellose Vorgänge - statistische Verfahren für die Nachrichten- und Regelungstechnik*, Berlin: Springer-Verlag, 1960
- Schüssler, H.-W.: *Digitale Systeme zur Signalverarbeitung*, Berlin: Springer-Verlag, 1973
- Späth, H.: *Cluster-Analyse-Algorithmen*, München: R. Oldenbourg, 1975
- Stanfill, C., Waltz, D.: Toward Memory-based Reasoning, *Communications of the ACM*, 1986, Vol. 29, No. 12, pp. 1213-1228
- Thuy, N., H., C., Schnupp, P.: *Wissensverarbeitung und Expertensysteme*, Handbuch der Informatik, Bd. 6.1, München: R. Oldenbourg, 1989
- Torres-Méndez, L., A., Ruiz-Suárez, J., C., Sucar, L., E., Gómez, G.: Translation, Rotation, and Scale-Invariant Object Recognition, *IEEE Trans. on Systems, Man, and Cybernetics, Part C*, 2000, Vol. 30, No. 1, pp. 125-130
- Tou, J., T.: *Advances in Information Systems Science*, New York: Plenum, 1970
- Valiant, L., G.: A Theory of the Learnable, *Communications of the ACM*, 1984, Vol. 27, No. 11, pp. 1134-1142
- Weinmann, A.: Neural Network Architecture and Learning Algorithms, *Int. J. Automation Austria*, Jg.6(1998) H.1,2, 37-76
- Westland, J., Ch.: Bayesian Alternatives to Neural Computing, *IEEE Trans. on Systems, Man, and Cybernetics*, 1995, Vol. 25, No. 1, pp. 59-67
- Winston, P., H.: *Learning and Reasoning by Analogy*, 1980, Vol. 23, No. 12, pp. 689-703
- Yao, X.: Evolving Artificial Neural Networks, *Proc. of the IEEE*, 1999, Vol. 87, No. 9, pp. 1423-1447
- Zemanek, H.: Lernende Automaten. In: Steinbuch, K. (Hrsg.), *Taschenbuch der Nachrichtenverarbeitung*, 2. Auflage, S. 1390-1396, Berlin; Springer-Verlag, 1967
- Zemanek, H.: Abstrakte Objekte, *Elektron. Rechenanl.*, 10 (1968) H.5, S. 208-216

Dipl. Ing. Dr. techn. Robert Genser  
 Malborghetgasse 27-29,6/6  
 A-1100 Wien  
 Tel. & Fax: +43 1 6074187  
 Email: rgenser@aon.at

# Sensitivity of Holistic Controllers

Alexander Weinmann, ÖVE, Senior Member IEEE

Vienna University of Technology, Institute of Automatic Control

Gusshausstrasse 27-29/375, A-1040 Vienna / Austria

email: `weinmann@iert.tuwien.ac.at`

June 21, 2001

*Keywords* – Entire set of eigenvalues, controller norm constraints

*Abstract:* Holistic controllers, which are using the set of eigenvalues as an entirety only, are detailed as far as sensitivity with respect to shifting parameters, unstructured plant perturbations and controller norm are concerned. Holistic controllers are also carried out for different settings regarding the optimization problem.

## 1 Introduction

The principle of holistic controllers takes the entirety of all the closed-loop eigenvalues into consideration as a new design target (*Weinmann, A., 2001*). Eigenvalues are not allocated individually but determined as their entirety. Controllers achieving appropriate results are termed holistic. Constraining the norm seems adequate if the state variables are supposed contaminated with noise. Different norms could be applied, as well, see *Weinmann, A., 2000a*.

Consider the controllable pair of the system matrix  $\mathbf{A} \in \mathcal{R}^{n \times n}$  and  $\mathbf{B} \in \mathcal{R}^{n \times m}$  as the input matrix of a process being continuously controlled and  $\mathbf{A}_{cl} = \mathbf{A} + \mathbf{BK}$  as the closed-loop system matrix. The eigenvalues  $\lambda_i[\mathbf{A} + \mathbf{BK}]$  are the less important the smaller their real parts are, i.e. the farther to the left they are located in the  $s$ -plane. Eigenvalues with smallest magnitude (with negative real part) are dominant.

For discrete-time systems, allocation near the origin of the  $z$ -plane is desirable.

The design of holistic controllers is based only on the trace of the inverse and on the determinant of the closed-loop system coefficient matrix in continuous-time or discrete-time systems, respectively. These quantities are minimized in order to obtain the holistic

controller.

This paper is an extension of *Weinmann, A., 2000*, and is devoted to several interesting sensitivity phenomena in holistic controllers, e.g., sensitivity relations of the sum of inverse eigenvalues. The norm of the controller versus plant uncertainty, robustness of the closed-loop holistic control system is assessed. Sensitivity considerations for maximum influence of unstructured uncertainties are used to discuss robustness in combination with root loci.

Holistic eigenvalue allocation subject to controller norm is released and the equality conditions could be replaced by inequality conditions occasionally since the closed-loop system should not be forced to stand improper equality conditions of actuating effort.

The result of the holistic controller also could have been obtained by assigning poles of adequate size with additional observation of the controller effort. However, balancing the eigenvalues and the controller norm would require time-consuming considerations.

In any case, controller solutions resulting from the appropriate equations are accepted only if the stability condition is satisfied.

## 2 Influence of Shifting Parameter on Stable and Unstable Plants

Holistic controllers  $\mathbf{K}^*$  minimizing the negative sum of inverse closed-loop eigenvalues after having shifted them with a shifting parameter  $r$  in the direction of the real axis are obtained (*Weinmann, A., 2001*) from

$$(\mathbf{A} + \mathbf{BK}^* + r\mathbf{I})^{-2}\mathbf{B} + 2\mu\mathbf{K}^{*T} = \mathbf{0} . \quad (1)$$

Sensitivity is also related to the following fact: If  $r$  is selected such that the utmost right pole or poles are close to the imaginary axis, then the influence of these poles is increased, their influence is spread. For  $r \ll 0$ , i.e.  $-r$  big, even for poles very different in size, all the poles are treated in a similar way.

**Example 1. Shifting based on  $r$ :** Using SISO data

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & -3 \end{pmatrix}, \quad \mathbf{B} = \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{K} = \mathbf{k}^T, \quad (2)$$

for  $r$  as a parameter in the range from  $-0.28$  to  $+0.18$ , the optimum controller  $\mathbf{K}^*$  results from Eq.(1). Its norm versus  $r$  is depicted in Fig. 1. For illustration purpose, the combination of these  $r$  dependent controllers with the given plant  $\mathbf{A}$  is given by a root contour in Fig.(2). The diamond point results for maximum  $r$ .  $\square$

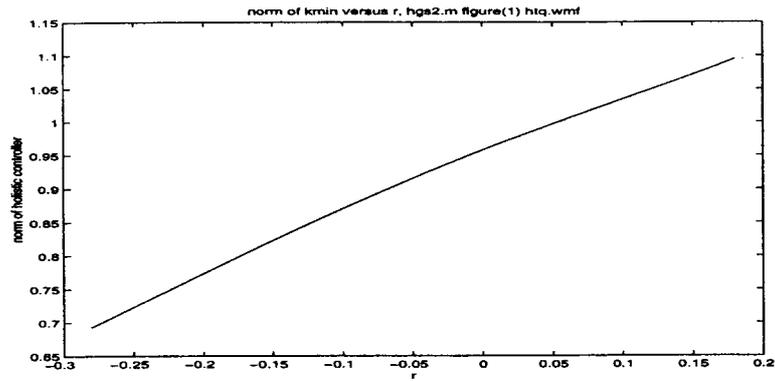


Figure 1: Influence of shifting parameter on controller norm

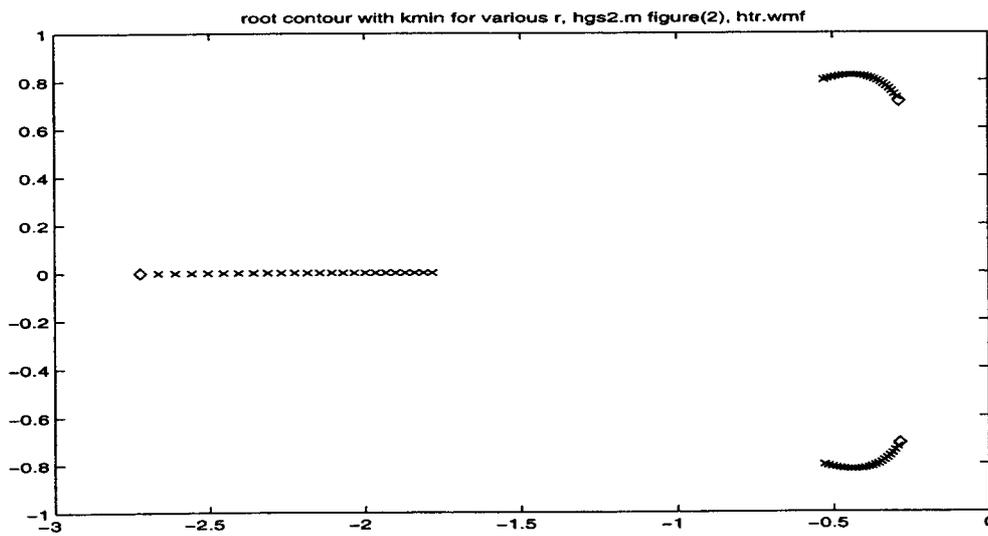


Figure 2: Root contour for controllers resulting for various shifting parameters  $r$

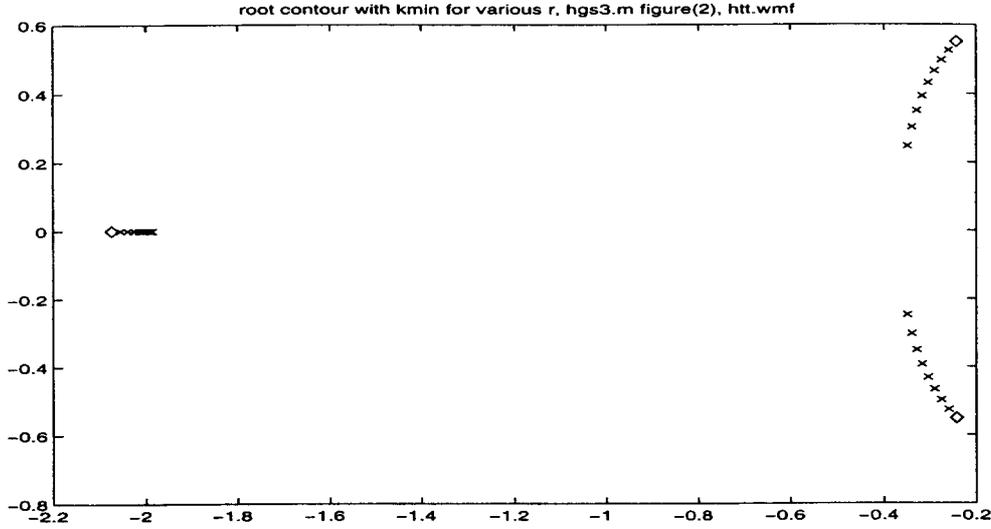


Figure 3: Root contour for an unstable plant with holistic controller

**Example 2. Unstable Plant:** Consider the plant in companion form with three eigenvalues at  $-1$ ;  $-2$ ;  $+0.2$ . In Fig. 3, the combination of the results  $\mathbf{K}^*$  with the unstable plant is illustrated by a root contour in the range  $r = -0.7 \dots -0.53$ . The diamond point results for maximum  $r$ .  $\square$

### 3 Eigenvalue Sensitivity versus Unstructured Perturbations Based on an Incremental Solution

Investigating sensitivity, the influence of an unstructured additive-type uncertainty  $\Delta\mathbf{A}$  in  $\mathbf{A}$  on the sum of inverse eigenvalues  $\lambda[\mathbf{A}_{cl}] = \lambda[\mathbf{A} + \mathbf{BK}]$  is given by the matricial gradient

$$\frac{\partial \sum_i 1/\lambda_i[\mathbf{A}_{cl}]}{\partial \mathbf{A}} = \frac{\partial \text{tr}[\mathbf{A}_{cl}^{-1}]}{\partial \mathbf{A}} = -\mathbf{B}^T(\mathbf{A} + \mathbf{BK})^{-2T}. \quad (3)$$

The sum of increments  $\Delta \sum_i 1/\lambda_i[\mathbf{A}_{cl}]$  caused by a series of  $\Delta\mathbf{A}$  is a measure of the sensitivity of the holistic controller.

### 4 Controller Norm Sensitivity

In addition, the influence of  $\Delta\mathbf{A}$  on the norm  $\|\mathbf{K}\|_F$  of the holistic controller is an important design feature. Find the increment  $\Delta\mathbf{K}$  versus the present  $\mathbf{K}$  which results from changing the present system  $\mathbf{A}$  by an additive uncertainty increment  $\Delta\mathbf{A}$ ,

$$\Delta\mathbf{K} = \mathbf{B}^T(\mathbf{A} + \Delta\mathbf{A} + \mathbf{BK})^{-2T} = [(\mathbf{A}_{cl} + \Delta\mathbf{A})^{-1}(\mathbf{A}_{cl} + \Delta\mathbf{A})^{-1}\mathbf{B}]^T \quad (4)$$

where  $\mathbf{A}_{cl} \triangleq \mathbf{A} + \mathbf{BK}$ . Using Taylor expansion,

$$\Delta \mathbf{K}^T = \mathbf{A}_{cl}^{-1}(\mathbf{A}_{cl} - \Delta \mathbf{A})\mathbf{A}_{cl}^{-2}(\mathbf{A}_{cl} - \Delta \mathbf{A})\mathbf{A}_{cl}^{-1}\mathbf{B} + R[\Delta \mathbf{A}^2] \quad (5)$$

$$\doteq \mathbf{A}_{cl}^{-1}(I - \Delta \mathbf{A} \cdot \mathbf{A}_{cl}^{-1} - \mathbf{A}_{cl}^{-1}\Delta \mathbf{A} + \Delta \mathbf{A} \cdot \mathbf{A}_{cl}^{-2} \cdot \Delta \mathbf{A})\mathbf{A}_{cl}^{-1}\mathbf{B} \quad (6)$$

$$= (\mathbf{A}_{cl}^{-2} - \mathbf{A}_{cl}^{-1} \cdot \Delta \mathbf{A} \cdot \mathbf{A}_{cl}^{-2} - \mathbf{A}_{cl}^{-2} \cdot \Delta \mathbf{A} \cdot \mathbf{A}_{cl}^{-1})\mathbf{B}. \quad (7)$$

From this equation, the matricial gradient of  $\|\mathbf{K}\|_F^2$  is

$$\frac{\partial \|\mathbf{K}\|_F^2}{\partial \mathbf{A}} = -2\mathbf{A}_{cl}^{-T}(\mathbf{A}_{cl}^{-2}\mathbf{B}\mathbf{B}^T\mathbf{A}_{cl}^{-T} + \mathbf{A}_{cl}^{-T}\mathbf{A}_{cl}^{-2}\mathbf{B}\mathbf{B}^T)\mathbf{A}_{cl}^{-T}. \quad (8)$$

Robustness is high if Eq.(8) is low in norm sense. For the maximum decrease in the norm of  $\Delta \mathbf{K}$ ,  $\Delta \mathbf{A}$  is changed regarding

$$\Delta \mathbf{A} = -\alpha \frac{\partial \|\mathbf{K}\|_F^2}{\partial \mathbf{A}}. \quad (9)$$

Following an adequate number of steps  $\Delta \mathbf{A}$ , three conditions must be satisfied for assessing performance robustness: (i) The given uncertainty is covered by the sum of all the increments  $\Delta \mathbf{A}$ , element by element, (ii) the resulting influence on the closed-loop eigenvalues lies in the predetermined region of the  $s$ -plane and (iii) the increments of  $\mathbf{K}$  are permissible.

**Example 3. Comparison with pole placement:** For data of Example 4, comparison of the sensitivity of a holistic controller and a pole placement controller is carried out. Assume a fixed small increment  $\Delta \mathbf{A}$  of  $\mathbf{A}$  which is given by a single step according to Eq.(9) where  $\|\Delta \mathbf{A}\|_F = 0.086$ . Then a norm increment in  $\mathbf{K}$  of 0.018 results.

Select now a pole placement controller  $\mathbf{K}_P$  with the original set-up and demand the poles unchanged. The increment in  $\mathbf{K}_P$  to keep the poles unchanged in the presence of the same fixed  $\Delta \mathbf{A}$  is 0.179.

The holistic controller is permitted to change the poles, the pole placement controller is not but the result is ten times the change in norm.

**Example 4. Perturbed eighth-order system:** For the SISO system with no zero,  $k = 6.3 \cdot 10^8$  and poles at  $-30, -24, -13.5, -4, -10 \pm j8, -1 \pm j9.9$ , the original  $\mathbf{b} = (1 \ 0^T)^T$  is changed to

$$\mathbf{B} = \mathbf{b} = (1 \ 0 \ 0 \ 0.2 \ 0 \ 0.1 \ 0 \ 0.5)^T, \quad (10)$$

find the incremental influence when perturbing the plant  $\mathbf{A}_p$ . Based on the previous holistic controller, the corresponding closed-loop roots are depicted in Fig. 4 and the perturbation  $\Delta \mathbf{A}$  is taken up; starting from  $\Delta \mathbf{A} = 0$ , the increments refer to Eq.(9). For the sake of

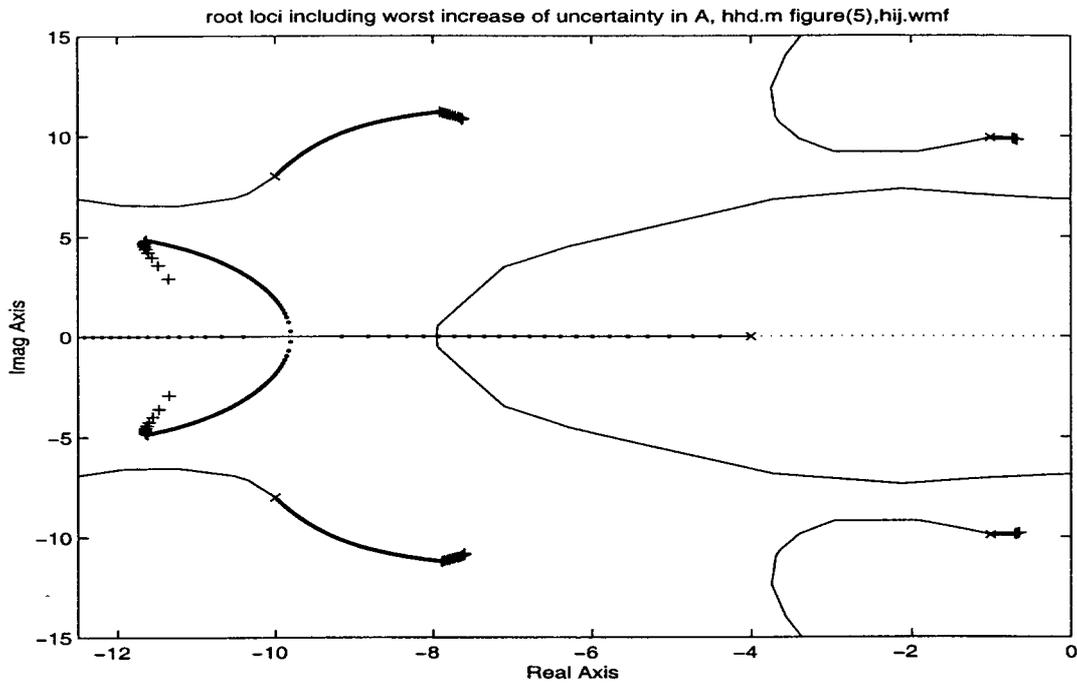


Figure 4: Root locus with different holistic controllers and with worst uncertainty influence

distinctness, the eigenvalues are depicted for only twelve steps in the root locus up to a predetermined deviation  $\Delta A_d$  from  $A$ .  $\square$

**Example 5. Systematically changing one element of  $B$  unequal to zero:** In Fig. 5, the only element in  $B$  unequal to zero and equal to 1 is selected systematically from position 1 to 8. Specific constellations  $A$  and  $B$  lead to holistic controllers which mainly influences only one single eigenvalue of the closed-loop if the norm of  $K$  is changed.

The root loci demonstrate that for these particular selections in  $B$ , only some branches in the root locus are concerned. Moreover, the sensitivity versus uncertainties in  $A$  is very different.

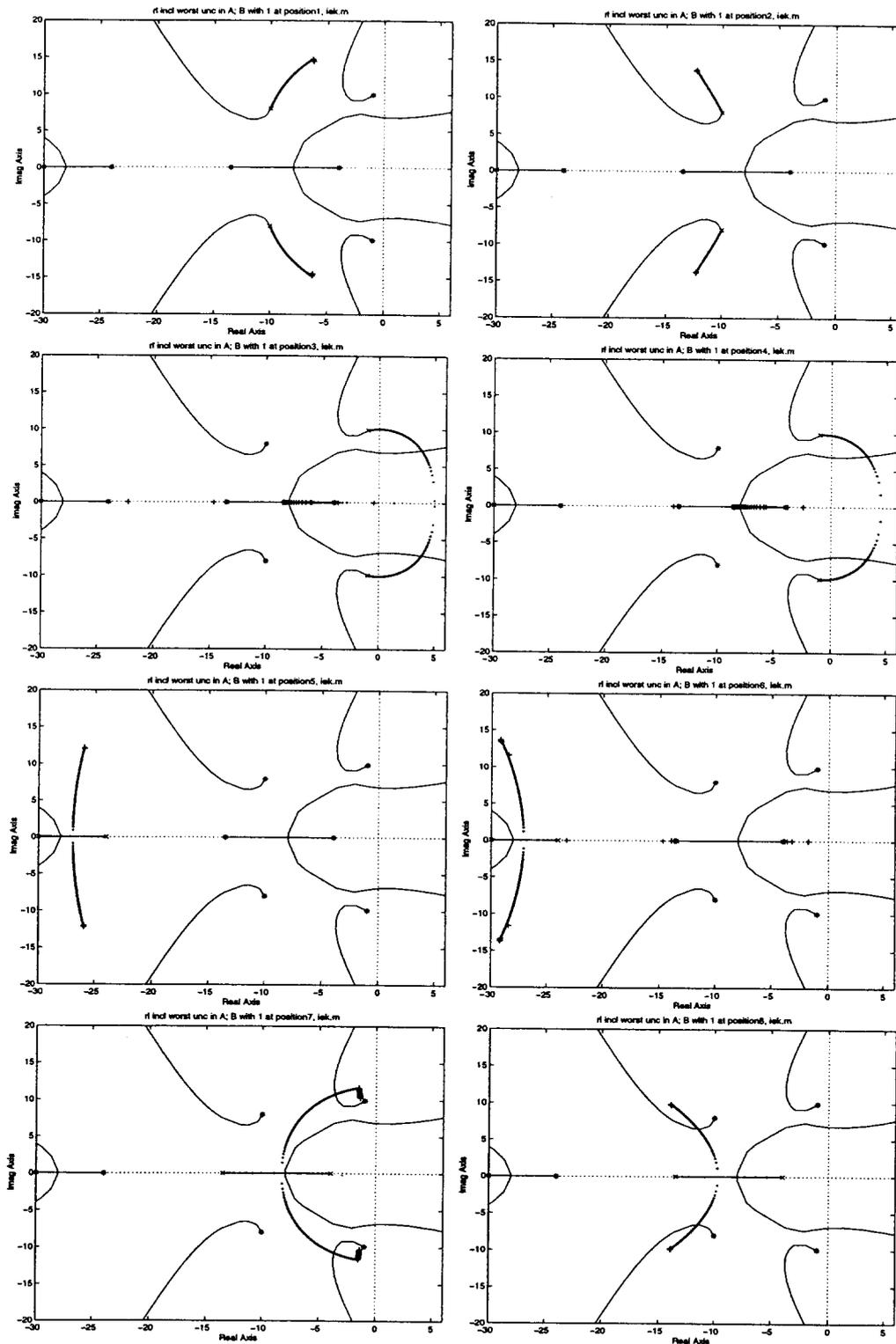


Figure 5: Root loci for eight different assumptions of B

## 5 Reciprocal Trace of $\mathbf{A}_{cl}^{-1}$

For several stable systems, e.g., if  $\mathbf{K}$  increases in norm, the eigenvalues  $\lambda_i[\mathbf{A}_{cl}]$  move to the right, i.e. decrease in modulus, and the sum of their reciprocals increases. On the other hand, for some unstable systems and particular regions, increasing  $\mathbf{K}$  entails  $\lambda_i[\mathbf{A}_{cl}]$  moving to the left and the sum of the reciprocals decreases. Remembering the shape of root loci or root contours reveals different moves of eigenvalues in the complex plane. The set-up of the index of performance  $I_R$  depends on the structure to be solved by the presented method.

In such cases using  $\mathbf{A}_{cl} = \mathbf{A} + \mathbf{BK}$ , the index of performance  $I_R$  is set up as

$$I_R = \frac{1}{\text{tr}[-(\mathbf{A}_{cl} + \mathbf{B}\Delta\mathbf{K})^{-1}]} + \mu\|\mathbf{K} + \Delta\mathbf{K}\|_F^2 \quad (11)$$

$$\doteq -\frac{1}{\text{tr}[\mathbf{A}_{cl}^{-1} + \mathbf{A}_{cl}^{-1}\mathbf{B}\Delta\mathbf{K} \cdot \mathbf{A}_{cl}^{-1}]} + \mu\|\mathbf{K} + \Delta\mathbf{K}\|_F^2 \quad (12)$$

$$I_R \doteq -\frac{1}{\text{tr}[\mathbf{A}_{cl}^{-1}]} + \frac{1}{(\text{tr}[\mathbf{A}_{cl}^{-1}])^2} \text{tr}[\mathbf{A}_{cl}^{-2}\mathbf{B}\Delta\mathbf{K}] + \mu\|\mathbf{K} + \Delta\mathbf{K}\|_F^2. \quad (13)$$

The gradient is

$$\frac{\partial I_R}{\partial \Delta\mathbf{K}} = \frac{1}{(\text{tr}[\mathbf{A}_{cl}^{-1}])^2} \mathbf{B}^T \mathbf{A}_{cl}^{-2T} + 2\mu\mathbf{K} = \mathbf{B}^T \frac{\mathbf{A}_{cl}^{-2T}}{(\text{tr}[\mathbf{A}_{cl}^{-1}])^2} + 2\mu\mathbf{K}. \quad (14)$$

From

$$(\mathbf{A} + \mathbf{BK}^*)^{-2}\mathbf{B} + 2\mu(\text{tr}[(\mathbf{A} + \mathbf{BK}^*)^{-1}])^2\mathbf{K}^{*T} = \mathbf{0} \quad (15)$$

$$\frac{(\mathbf{A} + \mathbf{BK}^*)^{-2}}{(\text{tr}[(\mathbf{A} + \mathbf{BK}^*)^{-1}])^2} \mathbf{B} + 2\mu\mathbf{K}^{*T} = \mathbf{0} \quad (16)$$

the optimum controller  $\mathbf{K}^*$  is achieved. Comparison with Eq.(1) shows a normalization process for  $\mathbf{A}_{cl}^{-1}$  corresponding to per unit data. However, this result differs from Eq.(1) and it is not only a matter of scaling the Lagrange multiplier  $\mu$ .

## 6 Minimum Sensitivity in Holistic Sense

Minimum sensitivity in holistic sense is defined as follows: Suppose  $\Delta\mathbf{A}$  bounded by limiting some entry of  $\Delta\mathbf{A}$  regarding some appropriate magnitude  $\Delta A_{ij,a}$ . An unstructured uncertainty  $\Delta\mathbf{A}$  should cause minimum effect on closed-loop (sum of) eigenvalues. Selecting the holistic controller referring to Eq.(1) such that  $\Delta\mathbf{A}$  has maximum influence on  $\|\mathbf{K}\|_F$ , then the best solution in the sense of holistic controller is guaranteed since all other  $\Delta\mathbf{A}$  have less influence on the norm of  $\mathbf{K}$ .

Replacing  $\mathbf{A}$  by  $\mathbf{A}_{pv} + \Delta\mathbf{A}$  where  $\Delta\mathbf{A}$  — for calculational purpose — is an increment of plant perturbation superimposed on the plant perturbation of  $\mathbf{A}_{pv}$ . The incremental

process runs until a single entry of  $\mathbf{A}_{pv}$  arrives at the lower or upper bound  $\mathbf{A}_{pl}$  or  $\mathbf{A}_{ph}$ , respectively, i.e. if one of the element per element relations ( $\leq_e$ )

$$\Delta A_{ij,a} : \mathbf{A}_{pl} \leq_e \mathbf{A}_{pv} \leq_e \mathbf{A}_{ph} \quad (17)$$

is reached. If the set of appropriate eigenvalues (symbols '+' in Fig. 4) is located inside the admissible region of stability then robust performance in the holistic sense is guaranteed.

## 7 Maximum Sum of Negative Eigenvalues. Direct Holistic Controller

Suppose  $\mathbf{A}_{cl}$  stable, then a very simple index of performance  $I_D$  results from maximizing the negative sum of closed-loop eigenvalues subject to the inverse norm of the state-space controller. Search for the minimum requires the chain rule of Eq.(19)

$$\frac{\partial a[b(\mathbf{M})]}{\partial \mathbf{M}} = \left( \mathbf{I}_r \otimes \frac{\partial a}{\partial b} \right) \left( \frac{\partial b}{\partial \mathbf{M}} \otimes \mathbf{I}_m \right) \quad (18)$$

$$= \left( \frac{\partial a}{\partial b} \mathbf{I}_r \right) \left( \frac{\partial b}{\partial \mathbf{M}} \otimes 1 \right) = \frac{\partial a}{\partial b} \mathbf{I}_r \frac{\partial b}{\partial \mathbf{M}} = \frac{\partial a}{\partial b} \frac{\partial b}{\partial \mathbf{M}} . \quad (19)$$

Then, one has

$$I_D = - \sum_i \lambda_i[\mathbf{A} + \mathbf{BKC}] + \mu \frac{1}{\|\mathbf{K}\|_F^2} \rightarrow \max_{\mathbf{K}} \quad (20)$$

$$\frac{\partial I_D}{\partial \mathbf{K}} = - \frac{\partial \text{tr}[\mathbf{A} + \mathbf{BKC}]}{\partial \mathbf{K}} - \frac{2\mu}{\|\mathbf{K}\|_F^3} \frac{\partial \|\mathbf{K}^T \mathbf{K}\|_F^2}{\partial \mathbf{K}} = \mathbf{0} . \quad (21)$$

$$\mathbf{K}^* = - \frac{k_F}{\sqrt{\text{tr}[\mathbf{CBB}^T \mathbf{C}^T]}} \mathbf{B}^T \mathbf{C}^T . \quad (22)$$

For data similiar to Example 4 the result is given by the root locus of Fig. 6.

## 8 Reciprocal Trace of $\mathbf{A}_{cl}$

Requiring sum of eigenvalues as far to the left as possible, the the negative sum should be as large as possible, the inverse of the negative sum should be as small as possible. Hence, an alternative set-up is given by the index of performance  $I_R$  using  $\mathbf{A}_{cl} = \mathbf{A} + \mathbf{BK}$  again and presupposing  $\Re \lambda_i[\mathbf{A}_{cl}] < 0$ . Consider an increment  $\Delta \mathbf{K}$  of  $\mathbf{K}$ , or directly apply Eq.(19),

$$I_R = \frac{1}{-\text{tr}[\mathbf{A}_{cl} + \mathbf{B}\Delta \mathbf{K}]} + \mu \|\mathbf{K} + \Delta \mathbf{K}\|_F^2 \rightarrow \min_{\Delta \mathbf{K}} \quad (23)$$

$$I_R \doteq -\{(\text{tr}[\mathbf{A}_{cl}])^{-1} - (\text{tr}[\mathbf{A}_{cl}])^{-2} \text{tr}[\mathbf{B}\Delta \mathbf{K}]\} + \mu \|\mathbf{K} + \Delta \mathbf{K}\|_F^2 \rightarrow \min_{\Delta \mathbf{K}} . \quad (24)$$

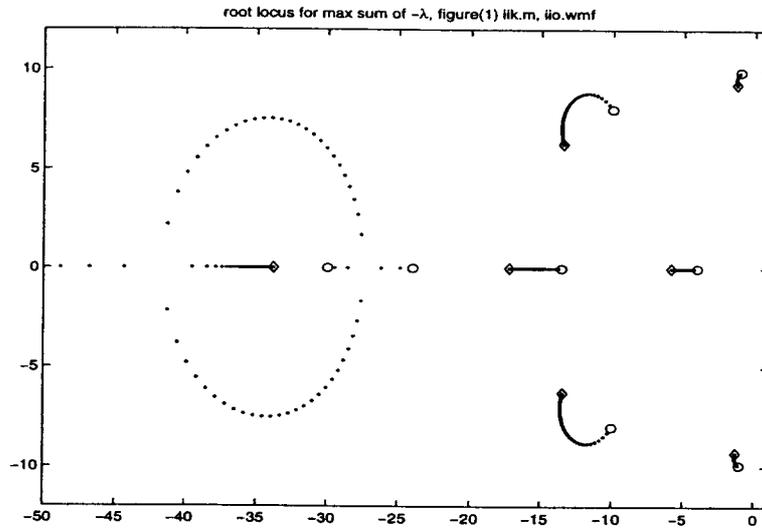


Figure 6: Root locus for the direct holistic controller, gain  $k_f$  moving from 'o' to 'diamond'

The gradient is

$$\frac{\partial I_R}{\partial \Delta \mathbf{K}} = \frac{1}{(\text{tr}[\mathbf{A}_{cl}])^2} \mathbf{B}^T + 2\mu \mathbf{K} . \quad (25)$$

Equating the gradient with zero, one has

$$\mathbf{B} + 2\mu \{\text{tr}[\mathbf{A} + \mathbf{BK}]\}^2 \mathbf{K}^T = \mathbf{0} \quad (26)$$

and the optimum controller  $\mathbf{K}$  is achieved. This is the same result as Eq.(1) although  $\text{tr}[\mathbf{A}^{-1}] \neq \frac{1}{\text{tr}[\mathbf{A}]}$ .

## Appendix

### A Holistic Controller Sensitivity for Discrete-time Systems

We recall

$$\frac{\partial \det \mathbf{M}}{\partial \mathbf{M}} = \mathbf{M}^{-T} \cdot \det \mathbf{M} \quad \text{or} \quad \frac{\partial \det \mathbf{M}}{\partial \text{row} \mathbf{M}} = [\text{row}(\mathbf{M}^{-T})] \cdot \det \mathbf{M} . \quad (27)$$

A comprehensive and concise result is

$$\begin{aligned} \frac{\partial}{\partial \mathbf{K}} \det(\mathbf{A} + \mathbf{BKC}) &= -\mathbf{B}^T [\text{adj}(\mathbf{A} + \mathbf{BKC})] \mathbf{C}^T = \\ &= -\mathbf{B}^T (\mathbf{A} + \mathbf{BKC})^{-1} \mathbf{C}^T \cdot \det(\mathbf{A} + \mathbf{BKC}) . \end{aligned} \quad (28)$$

applicable to  $\prod \lambda_i[\mathbf{M}] = \det \mathbf{M}$  of a discrete-time system where  $\mathbf{M} = \Phi(T) + \Psi \mathbf{K}_y \mathbf{C}$ .

For sensitivity considerations it turns out useful to separate the derivatives of  $\det \mathbf{A}_{cl}$  with respect to  $\text{row} \Phi + \Psi \mathbf{K}$  and  $\text{col} \Phi + \Psi \mathbf{K}$  with respect to  $\mathbf{K}$ . The same results apply to an output controller  $\mathbf{K}_y$ .

## A.1 Matricial Gradient of $\det \mathbf{A}_{cl}$ with Respect to State Controller

$$\text{For } \mathbf{K} \in \mathcal{R}^{m \times n}, \quad \Psi \in \mathcal{R}^{n \times m}, \quad \Phi \in \mathcal{R}^{n \times n}, \quad \frac{\partial \text{col} \mathbf{K}^T \Psi^T}{\partial \mathbf{K}} \in \mathcal{R}^{mn^2 \times n}, \quad (29)$$

$$\text{col } \mathbf{K}^T \Psi^T = \begin{pmatrix} K_{11}\Psi_{11} + K_{21}\Psi_{12} + K_{31}\Psi_{13} + \dots + K_{m1}\Psi_{1m} \\ K_{12}\Psi_{11} + K_{22}\Psi_{12} + K_{32}\Psi_{13} + \dots + K_{m2}\Psi_{1m} \\ \vdots \\ K_{1n}\Psi_{n1} + \dots \dots \dots + K_{mn}\Psi_{nm} \end{pmatrix} \in \mathcal{R}^{n \times n} \quad (30)$$

$$\frac{\partial \text{col} \mathbf{K}^T \Psi^T}{\partial \mathbf{K}} = \begin{pmatrix} \Psi_{11} \mathbf{I}_n \\ \Psi_{21} \mathbf{I}_n \\ \vdots \\ \Psi_{n1} \mathbf{I}_n \\ \dots \\ \Psi_{12} \mathbf{I}_n \\ \vdots \\ \Psi_{n2} \mathbf{I}_n \\ \dots \\ \vdots \\ \dots \\ \Psi_{1m} \mathbf{I}_n \\ \vdots \\ \Psi_{nm} \mathbf{I}_n \end{pmatrix} = (\text{col } \Psi) \otimes \mathbf{I}_n. \quad (31)$$

Solving  $\frac{\partial I_{cl}}{\partial \mathbf{K}}$  for  $I_{cl} = \det(\Phi + \Psi \mathbf{K})$ , the chain rule

$$\frac{\partial}{\partial \mathbf{K}} \mathbf{U}[\mathbf{V}(\mathbf{K})] = \left( \mathbf{I}_m \otimes \frac{\partial \mathbf{U}}{\partial \text{row} \mathbf{V}} \right) \left( \frac{\partial \text{col} \mathbf{V}^T}{\partial \mathbf{K}} \otimes \mathbf{I}_b \right) \quad \mathbf{U} \in \mathcal{R}^{a \times b}, \quad \mathbf{V} \in \mathcal{R}^{k \times l}, \quad \mathbf{K} \in \mathcal{R}^{m \times n} \quad (32)$$

(Brewer, J.W., 1978; Weinmann, A., 1991, Gl.(5.103)). Then

$$\mathbf{U}[\cdot] \triangleq \det[\cdot], \quad a = b = 1, \quad \mathbf{V}[\mathbf{K}] \triangleq \Phi + \Psi \mathbf{K}, \quad k = l = n \quad (33)$$

yields

$$\frac{\partial I_{cl}}{\partial \mathbf{K}} = \left( \mathbf{I}_m \otimes \frac{\partial \det(\Phi + \Psi \mathbf{K})}{\partial \text{row}(\Phi + \Psi \mathbf{K})} \right) \left( \frac{\partial \text{col}[(\Phi + \Psi \mathbf{K})^T]}{\partial \mathbf{K}} \otimes \mathbf{I}_1 \right). \quad (34)$$

Using Eqs.(27) and (31)

$$\frac{\partial I_{cl}}{\partial \mathbf{K}} = (\mathbf{I}_m \otimes \text{row}[(\Phi + \Psi \mathbf{K})^{-T}]) \quad (35)$$

$$\times (\text{col} \Psi \otimes \mathbf{I}_n) \cdot \det(\Phi + \Psi \mathbf{K}) \in \mathcal{R}^{m \times n}. \quad (36)$$

## A.2 Matricial Gradient of $\det \mathbf{A}_{cl}$ with Respect to an Output Controller

The following derivations are important for inserting a square matrix  $\Omega$  between lines (35) (36) and/or (45) (46). This matrix of dimension  $\mathcal{R}^{mn^2 \times mn^2}$ , irrespective of  $r$ , refers to both state and output controller.

The structure of  $\Omega$  can be diagonal, in order to weight the closed-loop system matrix  $\Phi + \Psi K$  element by element and attaching an additional importance factor on the  $m$  inputs.

For an output controller  $K_y$  the following definitions are introduced

$$K_y \in \mathcal{R}^{m \times r} \quad y \in \mathcal{R}^{r \times 1} \quad \Psi \in \mathcal{R}^{m \times n} \quad C \in \mathcal{R}^{r \times n} \quad (37)$$

$$K_y^T \triangleq \begin{pmatrix} k_1^T \\ \vdots \\ k_n^T \end{pmatrix} \quad k_i \in \mathcal{R}^{m \times 1} \quad (38)$$

$$\Psi^T \triangleq (\psi_1 \dots \psi_n) \quad \psi_i \in \mathcal{R}^{m \times 1} \quad (39)$$

$$K^T \Psi^T = \text{matrix}_{ij}[k_i^T \psi_j] . \quad (40)$$

Then,

$$\text{col } C^T K^T \Psi^T = (\Psi \otimes C^T) \cdot \text{col } K^T \quad (41)$$

$$\text{col } K_y^T = \begin{pmatrix} K_{11} \\ K_{12} \\ K_{13} \\ \vdots \\ K_{1r} \\ \dots \\ K_{21} \\ \vdots \\ K_{2r} \\ \dots \\ \vdots \\ K_{mr} \end{pmatrix}, \quad \frac{\partial \text{col } K_y^T}{\partial K_y} = \begin{pmatrix} \mathbf{I}_r \\ \mathbf{0}_r \\ \vdots \\ \mathbf{0}_r \\ \dots \\ \mathbf{0}_r \\ \mathbf{I}_r \\ \mathbf{0}_r \\ \mathbf{0}_r \\ \dots \\ \mathbf{0}_r \\ \vdots \\ \mathbf{I}_r \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{H}_1 \\ \dots \\ \mathbf{H}_2 \\ \dots \\ \vdots \\ \mathbf{H}_m \end{pmatrix} \triangleq \tilde{\mathbf{I}} . \quad (42)$$

The matrix  $\tilde{\mathbf{I}} \in \mathcal{R}^{m^2 r \times r}$  is partitioned into  $m$  submatrices  $\mathbf{H}_1$  through  $\mathbf{H}_m$  where  $\mathbf{H}_1$  consists of a column of  $m$  matrices  $(\mathbf{I}_r, \mathbf{0}_r, \dots, \mathbf{0}_r)^T$ , identity matrix at the first position, zero matrices elsewhere. The matrix  $\mathbf{H}_2$  consists of  $\mathbf{I}_r$  at the second position and zero matrices at the first position and at all the remaining  $m - 2$  positions. Then the matrix product rule leads to

$$\frac{\partial \text{col } C^T K_y^T \Psi^T}{\partial K_y} = [\mathbf{I}_m \otimes (\Psi \otimes C^T)] \cdot \tilde{\mathbf{I}} \quad \in \mathcal{R}^{mn^2 \times r} \quad (43)$$

Finally, using chain rule, see Eq.(32), one has

$$\frac{\partial I_{cl}}{\partial K_y} = \frac{\partial \det(\Phi + \Psi K_y C)}{\partial K_y} = \quad (44)$$

$$= (\mathbf{I}_m \otimes \text{row}[(\Phi + \Psi K_y C)^{-T}]) \quad (45)$$

$$\times (\mathbf{I}_m \otimes \Psi \otimes C^T) \cdot \tilde{\mathbf{I}} \cdot \det(\Phi + \Psi K_y C) \in \mathcal{R}^{m \times r} . \quad (46)$$

## B Stability Boundaries

For stable plants, starting with small  $k_F$ , repeating Eq.(1) will provide stable closed-loop solutions up to a stability boundary  $k_{Fs} = \|\mathbf{K}_s\|_F$  resulting from

$$\det(\mathbf{A} + \mathbf{BK}_s + j\omega_s\mathbf{I}) = 0 \quad (47)$$

$$(\mathbf{A} + \mathbf{BK}_s)^{-2} + 2\mu\mathbf{K}_s = 0. \quad (48)$$

There is always some  $\omega_s$  and  $\mathbf{K}_s$  determining the intersection with the imaginary axis. Starting from a stable  $\mathbf{A} + \mathbf{BK}^*$ , the range of admissible  $\mathbf{K}^*$  is given by the possible next intersection with the imaginary axis.

**Example:** For the setting of Example 1 the result is  $k_s = (-0.94 \ 2.14 \ 0.004)^T$ ,  $\omega_s = 0.84$  and  $\mu = 0.12$ .  $\square$

## C Holistic Controllers for SISO Systems

For continuous-time SISO systems and output controller, the sum of eigenvalues  $\lambda_i$  grows linearly with the controller gain  $k$ , irrespective of the system order and root locus shape, since the closed-loop characteristic equation is

$$p_{cl}(s) = \det(s\mathbf{I} - \mathbf{A} - \mathbf{b}k\mathbf{c}^T) \triangleq d(s) + kn(s) \quad (49)$$

$$= \sum_{i=1}^n a_i s^i = \prod_{i=1}^n (s - \lambda_i) = 0 \quad a_n \triangleq 1 \quad (50)$$

Selecting the factor of  $s^{n-1}$ ,  $i = n - 1$ ,

$$\sum_{j=1}^n (-\lambda_j) = a_{n-1} = d_{n-1} + kn_{n-1}. \quad (51)$$

This results only applies to  $\sum_{j=1}^n -\frac{1}{\lambda_i}$  and the result of the direct holistic controller, Eq.(22).

For discontinuous-time SISO systems, the product of eigenvalues is a linear function of the controller gain  $k$

$$p_{cl}(z) = \sum_{i=1}^n a_i z^i = \prod_{i=1}^n (z - z_i) = d(z) + kn(z) \quad (52)$$

and selecting the constant factor ( $i = 0$ )

$$a_o = \prod_{j=1}^n (-z_j) = \prod_{j=1}^n (-\lambda_j[\Phi_{cl}]) = d(0) + kn(0) \quad (53)$$

which directly applies to the holistic controller.

For MIMO systems the linear dependencies on the norm of  $\mathbf{K}_y$  break down.

### References

- Brewer, J.W., 1978*, Kronecker products and matrix calculus in system theory, *IEEE-Trans. CAS-25*, pp. 772-781
- Weinmann, A., 1991*, *Uncertain Models and Robust Control* (Vienna and New York: Springer)
- Weinmann, A., 2001*, Control system design based on holistic eigenvalue allocation, *e & i (Elektrotechnik und Informationstechnik)* **118**, pp. 167-173
- Weinmann, A., 2001a*, Matrix assignment subject to Lyapunov and closed-loop norm conditions, *2001 American Control Conference, Arlington, Virginia*, MP12-5

# Optimal control of a galvanometerscanner

Heinz Peter Rothwangl<sup>\*</sup>, Ivan Rajkovic<sup>\*\*</sup>

Institute of Automatic Control, Vienna University of Technology, Austria

## Abstract

*For a given control system with a linear plant, an algorithm for the calculation of an optimal reference signal is presented, such that the error of the output signal with respect to a desired output function is minimized. Limitations on the actuator effort as well as on the actuator rate can easily be implemented. The algorithm is based on linear programming, which also allows to implement restrictions on state variables. To apply the algorithm, a linear discrete-time model of the plant must be specified a priori. If the controller guarantees a small error of the reference signal with respect to the output signal, the dynamics of the controller may also be neglected. In this case, the optimal output of the system with respect to the desired output function can be taken for the reference signal, which is useful, if the controller is nonlinear or unknown for certain reasons. The proposed algorithm is applied to a galvanometerscanner, for which a model is determined by standard identification methods. The optimal reference signal is calculated, such that the position of the mirror is controlled and optimized with respect to a desired step function. Experimental data are presented and discussed in detail.*

## Kurzfassung

*Für einen gegebenen Regelkreis mit einer linearen Regelstrecke wird ein Algorithmus zur Berechnung von optimalen Referenzsignalen präsentiert, so daß das Fehlersignal in Bezug auf ein gewünschtes Ausgangssignal minimiert wird. Begrenzungen sowohl der Stellgrößen als auch der Stellgrößenänderungsraten können dabei sehr leicht implementiert werden. Der Algorithmus basiert auf linearen Programmen, wobei auch Begrenzungen der Zustandsvariablen berücksichtigt werden können. Der Algorithmus setzt dabei ein lineares, zeitdiskretes Modell der Strecke beziehungsweise des Regelkreises voraus. Wenn der verwendete Regler ein kleines Fehlersignal beim Betrieb des Regelkreises garantiert, darf dabei die Dynamik des Reglers vernachlässigt werden. In diesem Fall kann das optimale Ausgangssignal des Systems in Bezug auf das gewünschte Ausgangssignal als Referenzsignal verwendet werden, wobei dies von Vorteil ist, wenn ein nichtlinearer Regler verwendet wird oder der Regler aus verschiedenen Gründen nicht bekannt ist. Der vorgeschlagene Algorithmus wird für ein Galvanometerscanner verwendet, wobei ein Modell der Strecke mittels bekannter Identifikationsmethoden bestimmt wird. Ein optimales Referenzsignal wird berechnet, so daß sich der Spiegel in Bezug auf eine gewünschte Sprungfunktion so gut wie möglich bewegt. Experimentelle Daten werden präsentiert und im Detail erläutert.*

---

<sup>\*</sup> rothwangl@iert.tuwien.ac.at; phone 0043-1-58801-37514; fax 0043-1-58801-37599;

<http://www.iert.tuwien.ac.at>; Institute of Automatic Control, Gusshausstr. 25-29/375, 1040 Vienna, Austria;

<sup>\*\*</sup> rajkovic@iert.tuwien.ac.at; phone 0043-1-58801-37521; fax 0043-1-58801-37599;

<http://www.iert.tuwien.ac.at>; Institute of Automatic Control, Gusshausstr. 25-29/375, 1040 Vienna, Austria;

## 1 Introduction

### General remarks

Various literature deals with the complex task of finding admissible reference signals for linear systems subject to state and control constraints. In [1], the construction of maximally allowed admissible reference signals sets is investigated, which is based on the theory of maximal output admissible sets [2]. Unlike other constrained control approaches as the well known anti-windup mechanism for PI-type controllers, it is assumed in [1], that a primal linear controller has already been designed to stabilize the system and which provides desirable tracking performance or disturbance attenuation in the absence of constraints. The constraint fulfilment task is left to the reference signal. Basically, in [1], the maximal allowed reference value  $r_{\max}$  is calculated, such that any trajectory due to a reference signal which satisfies  $r(t) \leq r_{\max}$ , remains in a prescribed allowed region, provided that the initial condition of the whole control systems (initial conditions of the plant and the controller) also lies in an allowed given region ( $r_{\max}$  can be calculated by means of linear programming). Note that in [1] no statement is given how to choose  $r(t)$ . It is only required that  $r(t) \leq r_{\max}$  holds, which ensures that hard constraints on the state variables and actuator effort are not violated. Consequently, since any reference signal  $r(t)$  with  $r(t) \leq r_{\max}$  can be chosen,  $r_{\max}$  may be very restrictive.

Unlike the approach in [1], this paper addresses the problem of determining an optimal reference signal  $r(t)$  for a control system, such that the control system output is driven as close as possible to a desired output function without violating given constraints on the actuator and state variables. As it will be shown, linear programming techniques can be used to calculate the optimal reference signal. Of course, the optimal reference signal depends both on the initial condition of the control system and on the desired output function, which usually is a step function. Take into account that by this approach, only a certain trajectory together with the corresponding optimal reference signal  $r(t)$  is considered. As a consequence, in general, the optimal reference signal is not restricted by  $r_{\max}$ , the maximal value, which is allowed for any reference signal [1].

## 2 Problem Statement

Consider a SISO linear discrete-time system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k); x(0) = x_0 \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{2.1}$$

where  $x \in R^n$  is the state vector and  $u \in R$  is the control input.  $k$  is the time variable belonging to the time set  $I = \{0, 1, 2, \dots, r\}$  and  $A \in R^{n \times n}$ ,  $B \in R^n$ ,  $C \in R^n$  and  $D \in R$  are real matrices with the pair  $(A, B)$  controllable. Note that any linear control system can also be written in the form of (2.1), with the reference  $r(k)$  as the input signal and the actuator effort as a linear combination of state variables of the control system which comprises the plant and controller states. This special case is discussed in more detail in Section 3.2. In the following, only open loop control is considered with  $u(k)$  as the input to the system and  $y(k)$  for the output.

Both the state and the control variables are assumed to satisfy inequality constraints given by

$$|u(k)| \leq u_{\max}, \quad Gx(k) \leq g, \quad k \in I \quad (2.2)$$

with  $G \in R^{l \times n}$ ,  $g \in R^l$ .

These inequalities define convex polyhedral subsets of the control space  $R$  and the state space  $R^n$ , containing the origin in their interior.

To avoid extreme and dangerous changes of the actuator signal, it is also assumed that the rate of the control input, which is given by the difference  $\Delta u(k)$  of two consecutive control values, is limited by  $\Delta u_{\max}$  satisfying inequality constraints of the form

$$\begin{aligned} |\Delta u(k)| &\leq \Delta u_{\max} \\ k &= 0, 1, \dots, r-1 \end{aligned} \quad (2.3)$$

with  $\Delta u(k)$  defined as

$$\Delta u(k) = u(k+1) - u(k), \quad k = 0, 1, \dots, r-1 \quad (2.4)$$

In case of a linear control system (see also Section 3.2), actuator rate constraints correspond to state rate constraints, given similar to (2.3) and (2.4) by

$$\begin{aligned} |\Delta x(k)| &\leq \Delta x_{\max} \\ k &= 0, 1, \dots, r-1 \end{aligned} \quad (2.5)$$

with  $\Delta x(k)$  defined as

$$\Delta x(k) = x(k+1) - x(k), \quad k = 0, 1, \dots, r-1 \quad (2.6)$$

In addition, it may also be desirable to restrict the state rate of a system for security reasons. Now, the optimization problem is to find an optimal input sequence to the system (2.1), such that the output  $y(k)$  of the plant is driven as close as possible to the desired output  $y_{des}(k)$  for a certain given initial condition. This optimization problem is formulated by a linear program as presented in detail in the following section.

### 3 Solution by Linear Programming

#### 3.1 Calculation of an optimal input sequence

The design objective is to minimize the absolute sum of the error signal, which is given by

$$\min \sum_{k=0}^r |e(k)| = \min \sum_{k=0}^r |y(k) - y_{des}(k)| \quad (3.1)$$

So far, the design objective is a non-linear function of the error signal. To obtain a linear design objective, the error signal  $e(k)$  is set equal to the difference of two auxiliary nonnegative variables  $e^+(k)$  and  $e^-(k)$ :

$$e(k) = e^+(k) - e^-(k) = y(k) - y_{des}(k), e^+(k) \geq 0, e^-(k) \geq 0, k = 0, 1, \dots, r \quad (3.2)$$

Now, the design objective can be written as a linear function of  $e^+(k)$  and  $e^-(k)$

$$\min \sum_{k=0}^r |e(k)| = \min \sum_{k=0}^r (e^+(k) + e^-(k)) \quad (3.3)$$

because the minimization of (3.3) sets at least one of the variables  $e^+(k)$  or  $e^-(k)$  equal to zero for each time instant  $k$ . Hence, the design objectives (3.1) and (3.3) coincide.

The linear design objective (3.3) together with (2.1)-(2.6) is a linear program. To summarize, the optimization problem is given by

$$\min \sum_{k=0}^r (e^+(k) + e^-(k)) \quad (3.4)$$

subject to the linear constraints of

$$\left. \begin{aligned} x(0) &= x_0 \\ x(k+1) &= Ax(k) + Bu(k) \\ e^+(k) - e^-(k) &= Cx(k) + Du(k) - y_{des}(k) \\ e^+(k) &\geq 0 \\ e^-(k) &\geq 0 \\ u(k) &\leq u_{\max} \\ u(k) &\geq -u_{\max} \\ Gx(k) &\leq g \end{aligned} \right\} k = 0, 1, \dots, r \quad (3.5)$$

$$\left. \begin{aligned} \Delta u(k) &= u(k+1) - u(k) \\ \Delta u(k) &\leq \Delta u_{\max} \\ \Delta u(k) &\geq -\Delta u_{\max} \\ \Delta x(k) &= x(k+1) - x(k) \\ \Delta x(k) &\leq \Delta x_{\max} \\ \Delta x(k) &\geq -\Delta x_{\max} \end{aligned} \right\} k = 0, 1, \dots, r-1$$

### 3.2 Calculation of an optimal reference signal

The results of Section 2 and 3 can easily be extended to a linear control system. To set up a linear program similar to (3.4),(3.5,) a state space description of the controller as well as of the plant is used to formulate a state space description of the whole control system. To avoid algebraic loops<sup>1</sup>, assume that the plant is given by the difference equation

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k); x(0) = x_0 \\ y(k) &= Cx(k) \end{aligned} \quad (3.6)$$

and the state-space description of the controller is given by

$$\begin{aligned} x_c(k+1) &= A_c x_c(k) + B_c e(k); x_c(0) = x_{c,0} \\ u(k) &= C_c x_c(k) + D_c e(k) \\ e(k) &= y_{ref}(k) - y(k) \end{aligned} \quad (3.7)$$

where  $e(k)$  denotes the error of the standard control system<sup>2</sup> and with  $A_c$ ,  $B_c$ ,  $C_c$  and  $D_c$  of appropriate dimensions.

It follows that

$$\begin{aligned} x_c(k+1) &= A_c x_c(k) + B_c (y_{ref}(k) - y(k)) \\ u(k) &= C_c x_c(k) + D_c (y_{ref}(k) - y(k)) \\ x(k+1) &= Ax(k) + B \left\{ C_c x_c(k) + D_c (y_{ref}(k) - y(k)) \right\}. \end{aligned} \quad (3.8)$$

The state-space description of the whole control system with the reference signal as the input variable and  $y$  as the output of the is now given by

$$\begin{aligned} \begin{pmatrix} x(k+1) \\ x_c(k+1) \end{pmatrix} &= \begin{pmatrix} A - BD_c C & BC_c \\ -B_c C & A_c \end{pmatrix} \begin{pmatrix} x(k) \\ x_c(k) \end{pmatrix} + \begin{pmatrix} BD_c \\ B_c \end{pmatrix} y_{ref}(k) \\ y(k) &= Cx(k) \\ u(k) &= (-D_c C \quad C_c x_c) \begin{pmatrix} x(k) \\ x_c(k) \end{pmatrix}^T + D_c y_{ref}(k) \end{aligned} \quad (3.9)$$

Assume hard constraints on the actuator effort and actuator rate as well as hard constraints on the state variables  $x(k)$  of the plant are given similar to (2.2) - (2.6). Obviously, the system (3.9) is of the same structure as the system given in (2.1). The variable  $y_{ref}$  corresponds to the input  $u$  of system (2.1) and hard constraints on the actuator effort of system (3.9) correspond to convex restrictions on a linear combination of the state variables  $x_c$ ,  $x$  and the reference

<sup>1</sup> In general, proper transfer-functions are given for the controller (e.g. P,PI, PID-controller). Together with strictly proper transfer functions for the plant, algebraic loops are avoided.

<sup>2</sup> It is assumed that the controller is implemented in the standard control configuration (i.e. a feedback controller is used) with the reference signal  $y_{ref}$  and  $y$  for the output of the control system.

signal  $y_{ref}$ . These convex restrictions can easily be implemented as linear constraints of the linear program. Now, the optimal reference signal with respect to a desired output signal  $y_{des}$  ( $e(k) = y_{des}(k) - y(k)$ ) can be calculated by solving a linear program similar to (3.5), given by the linear design objective

$$\min \sum_{k=0}^r |e(k)| = \min \sum_{k=0}^r \{e^+(k) + e^-(k)\} \quad (3.10)$$

subject to the linear constraints of

$$\left. \begin{aligned} x(0) &= x_0, x_c(0) = x_{c,0} \\ x(k+1) &= (A - BD_c C)x(k) + BC_c x_c(k) + BD_c y_{ref}(k) \\ x_c(k+1) &= -B_c Cx_c(k) + A_c x_c(k) + B_c y_{ref}(k) \\ u(k) &= -D_c Cx(k) + C_c x_c(k) + D_c y_{ref}(k) \\ e^+(k) - e^-(k) &= y_{des}(k) - Cx(k) \\ e^+(k) &\geq 0 \\ e^-(k) &\geq 0 \\ u(k) &\leq u_{\max} \\ u(k) &\geq -u_{\max} \\ Gx(k) &\leq g \end{aligned} \right\} k = 0, 1, \dots, r \quad (3.11)$$

$$\left. \begin{aligned} \Delta u(k) &= u(k+1) - u(k) \\ \Delta u(k) &\leq \Delta u_{\max} \\ \Delta u(k) &\geq -\Delta u_{\max} \\ \Delta x(k) &= x(k+1) - x(k) \\ \Delta x(k) &\leq \Delta x_{\max} \\ \Delta x(k) &\geq -\Delta x_{\max} \end{aligned} \right\} k = 0, 1, \dots, r-1$$

The solution of the linear program (3.10), (3.11) yields an optimal reference signal, which drives the control system from a given initial state  $(x_0, x_{c,0})$  as close as possible to the desired output. Usually, the initial condition  $x_{c,0}$  of the controller is equal to zero. However, in many applications, the initial condition of the controller  $x_{c,0}$  can be chosen as desired by the designer. This allows the initial condition  $x_{c,0}$  of the controller to be set free and optimized by the linear program. Using this free initial condition  $x_{c,0}$ , the optimization problem is to find an optimal reference signal and an optimal initial condition  $x_{c,0}$ , such that the initial state of the plant  $x_0$  and the corresponding output of the plant is driven as close as possible to the desired output  $y_{des}(k)$ . The linear program (3.10), (3.11) can easily be adapted to solve this slightly modified optimization problem just by cancelling the restriction  $x_c(0) = x_{c,0}$  with  $x_{c,0}$  a fixed value in (3.11).

## 4 Application to a Galvanometerscanner

### 4.1 System identification and given constraints

The control of a galvanometerscanner [3] is considered in the following. The control input for the plant is given by the current  $i(t)$ , whereas the output of the system is given by the angular  $\varphi$  (position) of the mirror. Experimental system identification yielded the best result with the Box-Jenkins<sup>3</sup> identification method using a system order of four. With a sampling period of  $T = 1.22 \cdot 10^{-4}$ , the discrete-time model of the plant is given by

$$\frac{\varphi(z)}{i(z)} = \frac{0.003107z^{-1} - 0.00859z^{-2} + 0.007911z^{-3} - 0.002426z^{-4}}{1 - 3.801z^{-1} + 5.424z^{-2} - 3.444z^{-3} + 0.8215z^{-4}}. \quad (4.1)$$

The poles of the plant are given by

$$z_{p,1} = 1, z_{p,2} = 0.977, z_{p,3} = 0.912 + j0.096, z_{p,4} = 0.912 - j0.096 \quad (4.2)$$

and the zeros by

$$z_{n,1} = 0.8401, z_{n,2} = 0.9623 + j0.0562, z_{n,3} = 0.9623 - j0.0562. \quad (4.3)$$

Assuming a series connection of a pure integrator ( $z_{p,1} = 0$ ), which corresponds to the first state variable, with a third order transfer-function (2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup> states), a state space representation for this model can easily be determined and is given by

$$x(k+1) = \begin{pmatrix} 1 & 0.003872 & -0.0052497 & 0.0089446 \\ 0 & 1.1627 & -0.10493 & 0.042292 \\ 0 & 0.28081 & 0.90814 & -0.09188 \\ 0 & -0.15252 & 0.20847 & 0.73007 \end{pmatrix} x(k) + \begin{pmatrix} 0.0031069 \\ 0.00572 \\ -0.0061108 \\ 0.0064771 \end{pmatrix} i(k) \quad (4.4)$$

$$\varphi(k) = x_1(k)$$

For this plant, a PD-type controller for a standard control system has been designed. Unfortunately, because the application is proprietary, the parameters of the controller and the design method for the controller must not be published and are unknown to the authors.

Now, assume that the output  $y(k)$  of the plant should be as close as possible to the step function  $y_{des}(k) = \sigma(k)$ , which is the desired output function. Assuming that the controller guarantees a small error signal for this certain application, the optimal reference signal for the control system is close to the output of the control system. Hence, the optimal output se-

<sup>3</sup> The Box-Jenkins identification method supports linear difference equations of the form  $y(k) = B(z)/F(z)u(k-nk) + C(z)/D(z)e(k)$ .

The Matlab - Ident toolbox has been used for the identification of the system.

quence of the plant with respect to the desired step function can be taken as the reference signal. To calculate the optimal output sequence, the linear program (3.4),(3.5) has to be set up and solved (open loop control).

For this particular application, the maximal allowed input to the plant is given by  $i_{\max} = 8$  and no restrictions for the state variables and for the actuator rate or state rates have been assumed. 120 sample points have been considered, which turned out to be sufficiently high. Note that if only the transfer-function (4.11) is of interest and no state limitations are to be implemented, the difference equation of the transfer - function (4.1) can be used instead of the state space representation (4.4), which results in a smaller linear program. The initial condition  $x_0$  of the plant has been set to zero.

To set up the linear program, Matlab has been used to generate the matrices of the linear program<sup>4</sup>. However, XPRESS-MP [4], a standard software package to solve linear and mixed-integer programs, has been used to solve the linear program. To pass the linear program to XPRESS-MP, the standard mps (mathematical programming system) format, which is an ASCII-format to describe linear programs, has been used.

## 4.2 Optimal reference signal and experimental results

Solving the linear program (3.4),(3.5) yields following optimal output signal and the corresponding actuator effort, which is necessary to achieve the optimal output signal (see Figure 3.1).

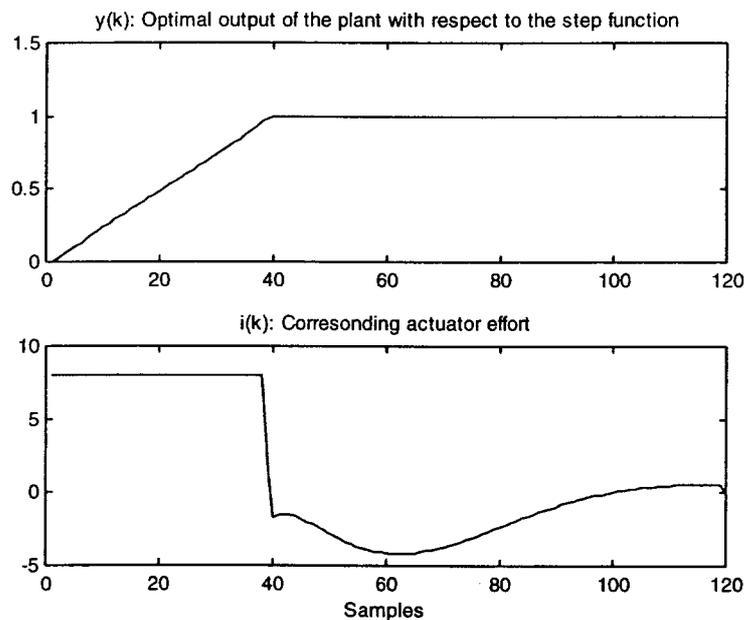


Figure 3.1: Optimal output of the plant and corresponding actuator effort

As mentioned, if the PD-controller guarantees a small error signal, this output signal may also be used as the reference signal for the control system. Experimental results are presented in Figure (3.2). Because open-loop control is considered in the optimization procedure and be-

<sup>4</sup> Any linear program can be written as  $\min c^T x$  subject to  $Ax=b$  with each component of  $x$  bigger than zero.

cause of the remaining error signal of the control system, the simulation results and experimental results do not match exactly and diverge as time increases.

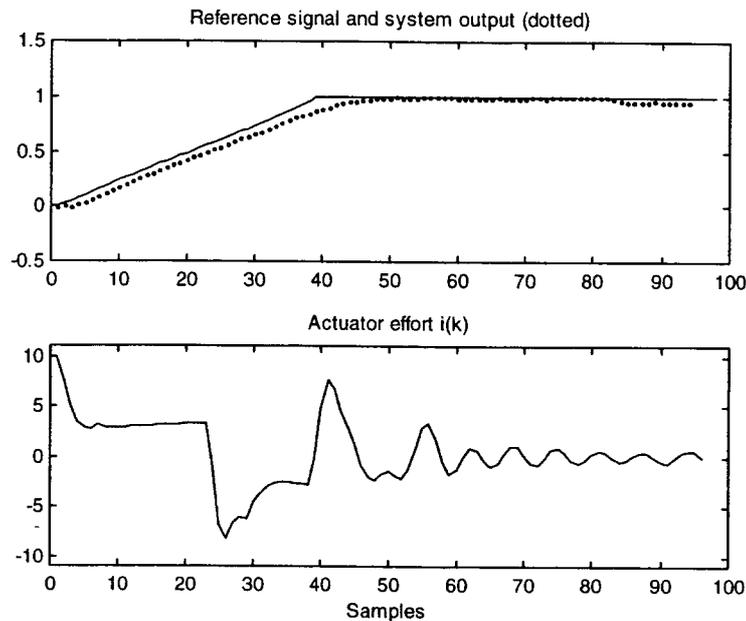


Figure 3.2: Closed-loop response and actuator effort

## 5 Conclusion

In this paper, a new algorithm for the calculation of an optimal reference signal has been developed and tested on the practical application of a galvanometerscanner. The approach, based on linear programming, allows to calculate an optimal reference signal which does not violate given constraints on the actuator effort and on the state variables. The algorithm can be applied both to open loop control and closed loop linear control systems. Unfortunately, because the controller parameters for the galvanometerscanner were not known in this particular proprietary application, only the open loop optimization procedure could be applied, ignoring the dynamics of the controller. Still, the results are useful, because the approach allows to calculate an optimal reference signal.

## 6 References

- [1] Hirata, K. & Fujita, M.: Set of admissible reference signals and control of systems with state and control constraints. *Proceeding of the 38<sup>th</sup> IEEE Conference on Decision and Control*, pp 1427-1432, vol.2, 1999
- [2] Tan, Kok Tin: Maximal output admissible sets and the nonlinear control of linear discrete-time systems with state and control constraints.
- [3] Rajkovic, I., Noisser, R., & Bammer, M.: Entwurfsverfahren für Führungsgrößenfilter von hochdynamischen Positionsregelungen. IEE SPSS/IPC/DRIVES. Elektrische Automatisierungstechnik - System und Komponenten. Fachmesse & Kongress, Nov. 99, Nürnberg. Tagungsband. S. 474-83.
- [4] XPRESS-MP: Dash Associates Ltd Blisworth House. Blisworth Norhants NN7 3BX UK, <http://www.dash.co.uk>

# Gradients of Norms, Traces and Determinants for Automatic Control Applications

Alexander Weinmann, ÖVE, Senior Member IEEE  
Vienna University of Technology, Institute of Automatic Control  
Gusshausstrasse 27-29/375  
A-1040 Vienna / Austria  
Phone: +43 1 58801 37500, Fax: +43 1 58801 37599  
email: `weinmann@iert.tuwien.ac.at`

June 21, 2001

*Abstract: A comprehensive collection of formulas and correspondencies is given for differential quotients of expressions in norm, trace, determinant and eigenvalue with respect to vectors and matrices as often used in automatic control.*

## 1 Introduction

Several properties of control systems are given by scalars, e.g. eigenvalues, singular values, norms, Frobenius norms, traces, determinants, condition numbers.

Expressions which are influenced or selected by design in the intention of optimization are parameter vectors  $\mathbf{p}$ , the matrix  $\mathbf{K}$  of the controller in state space or a general matrix  $\mathbf{M}$ .

The symbols  $p$  and  $\mathbf{p}$  are used for general scalar and vector-valued parameters, respectively. For a general matrix the term  $\mathbf{M}$  is chosen; the symbol  $\mathbf{P}$  is avoided since it is commonly used as the Lyapunov matrix. If in practice an expression is only used associated with controller matrix then  $\mathbf{M}$  is replaced by  $\mathbf{K}$ .

The formulas are collected from numerous publications. Several are selected from recent publications of the author.

## 2 Analysis Using Matrices in Control

In general, the elements of all vectors  $\mathbf{p}$  or matrices  $\mathbf{M}$ ,  $\mathbf{K}$  are considered as independent variables with no interelement dependencies. Considering a derivative of a function with respect to the vector  $\mathbf{p}$  or the matrix  $\mathbf{M}$ , vector or matrix interrelations in  $\mathbf{p}$  or  $\mathbf{M}$ , respectively, must not exist, in general.

Interrelation dependencies are only addressed in a specific Section 8.

General remarks to be mentioned: For  $\mathbf{A}$ ,  $\mathbf{B}$  commutative one has for the trace  $\text{tr}$  (sum of main-diagonal elements of a square matrix)

$$\text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]. \quad (1)$$

For matrices  $\Delta\mathbf{M}$  small in norm sense

$$(\mathbf{U} - \mathbf{B} \cdot \Delta\mathbf{M})^{-1} \doteq \mathbf{U}^{-1}(\mathbf{U} + \mathbf{B} \cdot \Delta\mathbf{M})\mathbf{U}^{-1}. \quad (2)$$

For the sake of abbreviation,

$$(\mathbf{A}^T)^{-1} \equiv \mathbf{A}^{-T} \quad (3)$$

is used. The superscript  $^H$  (Hermite) means conjugate transpose

$$\mathbf{A}^H = (\mathbf{A}^*)^T. \quad (4)$$

Common rules are

$$\frac{\partial c(\mathbf{M})}{\partial \mathbf{M}^T} = \left( \frac{\partial c(\mathbf{M})}{\partial \mathbf{M}} \right)^T, \quad (5)$$

$$\text{col}[\mathbf{AMB}] = (\mathbf{B}^T \otimes \mathbf{A}) \text{col} \mathbf{M}, \quad (6)$$

$$\text{col}[\mathbf{AM} + \mathbf{MB}] = (\mathbf{B}^T \oplus \mathbf{A}) \text{col} \mathbf{M}. \quad (7)$$

Here the abbreviation  $\text{col}$  (or  $\text{vec}$ ) is the column operator, concatenating all the columns of a matrix to an entire column vector.

The Kronecker product  $\otimes$  of two matrices (or direct product, tensor product) is defined by a partitioned matrix whose  $(i, j)$ -partition is  $A_{ij}\mathbf{B}$

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{pmatrix} A_{11}\mathbf{B} & A_{12}\mathbf{B} & \dots & A_{1n}\mathbf{B} \\ A_{21}\mathbf{B} & A_{22}\mathbf{B} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}\mathbf{B} & A_{m2}\mathbf{B} & \dots & A_{mn}\mathbf{B} \end{pmatrix} = \text{matrix}[A_{ij}\mathbf{B}], \quad \begin{array}{l} \mathbf{A} \in \mathcal{C}^{n \times m} \\ \mathbf{B} \in \mathcal{C}^{r \times s} \\ \mathbf{A} \otimes \mathbf{B} \in \mathcal{C}^{nr \times ms} \end{array} \quad (8)$$

The Kronecker sum  $\oplus$  is defined by (*Brewer, J.W., 1978*)

$$\mathbf{N} \oplus \mathbf{M} = \mathbf{N} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{M}; \quad \mathbf{N} \in \mathcal{C}^{n \times n}, \quad \mathbf{M} \in \mathcal{C}^{m \times m}, \quad \mathbf{N} \oplus \mathbf{M} \in \mathcal{C}^{mn \times mn}. \quad (9)$$

Referring to the spectral norm<sup>1</sup> and maximum singular value,  $\|\mathbf{G}(s)\|_s = \sigma_{\max}[\mathbf{G}(s)] \leq \|\mathbf{G}(s)\|_F$  and  $\|\mathbf{G}(s)\|_\infty \triangleq \sup_\omega \sigma_{\max}[\mathbf{G}(j\omega)]$ , the  $H_\infty$  norm can be replaced by the maximum of the Frobenius norm versus frequency  $\|\mathbf{G}(s)\|_\infty < \sup_\omega \|\mathbf{G}(s = j\omega)\|_F$ . The replacement simplifies the derivation but only yields a sufficient result.

According to  $\frac{1}{\sqrt{n}}\|\mathbf{G}\|_F \leq \sigma_{\max}[\mathbf{G}] \leq \|\mathbf{G}\|_F$  where  $\mathbf{G} \in \mathcal{C}^{n \times n}$ , the extent of overbounding is assessed as  $\|\mathbf{G}\|_F \leq \sqrt{n} \sigma_{\max}[\mathbf{G}]$ . For  $n = 3$  e.g., overbounding is less than 73%.

### 3 Derivative of a Scalar-Valued Function with Respect to a Scalar

Consider the scalar parameter  $p$  (Vetter, W.J., et al., 1972)

$$\frac{\partial \det \mathbf{A}}{\partial p} = \text{tr} \left[ \frac{\partial \mathbf{A}}{\partial p} \text{adj } \mathbf{A} \right] = \text{tr} \left[ \frac{\partial \mathbf{A}}{\partial p} \mathbf{A}^{-1} \right] \cdot \det \mathbf{A} . \quad (14)$$

Derivation of  $\det(p\mathbf{I}_n - \mathbf{A})$ ,  $\mathbf{A} \in \mathcal{C}^{n \times n}$ ,  $p$  scalar, yields

$$\frac{\partial}{\partial p} \det(p\mathbf{I}_n - \mathbf{A}) = \sum_{i=1}^n \det(p\mathbf{I}_{n-1} - \mathbf{A}_{\text{red } ii}) \quad (15)$$

where  $\mathbf{A}_{\text{red } ii}$  is a square matrix evolved from  $\mathbf{A}$  when deleting row  $i$  and column  $i$ .

For the matrix product one has

$$\frac{\partial}{\partial p} \text{tr}[\mathbf{A}^T(p)\mathbf{A}(p)] = \frac{\partial}{\partial p} \|\mathbf{A}(p)\|_F^2 = 2 \text{tr} \left[ \mathbf{A}^T(p) \frac{\partial \mathbf{A}(p)}{\partial p} \right] . \quad (16)$$

Eventually,

$$\frac{\partial}{\partial p} \text{tr} [(p\mathbf{I} + \mathbf{C})^{-2}\mathbf{D}] = -2 \text{tr} [(p\mathbf{I} + \mathbf{C})^{-3}\mathbf{D}] . \quad (17)$$

<sup>1</sup>The Frobenius or Euler norm is defined as

$$\|\mathbf{x}\|_F \triangleq \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} = \sqrt{\mathbf{x}^H \mathbf{x}} . \quad (10)$$

Spectral (induced) norm or Hilbert Norm

$$\|\mathbf{G}\|_s \triangleq \sup_{\mathbf{x}} \frac{\|\mathbf{G}\mathbf{x}\|_F}{\|\mathbf{x}\|_F} \quad \text{where } \mathbf{x} \neq \mathbf{0} . \quad (11)$$

Singular value

$$\sigma[\mathbf{G}] \triangleq + \sqrt{\lambda[\mathbf{G}^H \mathbf{G}]} \quad (12)$$

$$\sigma_{\max} = + \sqrt{\lambda_{\max}[\mathbf{G}^H \mathbf{G}]} \equiv \|\mathbf{G}\|_s . \quad (13)$$

The maximum singular value is a generalization of the absolute value.

## 4 Derivative of a Scalar-Valued Function With Respect to a Vector

The derivative of a function with respect to a vector or a matrix is a definition of a scheme, only. This abbreviation is suitable when cascades of calculations are taken into consideration.

The vector derivative operator (or gradient or nabla operator) with respect to the vector  $\mathbf{p} = \text{vector}[p_i]$  is termed

$$\frac{\partial}{\partial \mathbf{p}} \triangleq \begin{pmatrix} \frac{\partial}{\partial p_1} \\ \frac{\partial}{\partial p_2} \\ \vdots \\ \frac{\partial}{\partial p_n} \end{pmatrix} \triangleq \text{grad}_{\mathbf{p}} = \nabla_{\mathbf{p}} . \quad (18)$$

Derivative of an inner vector product

$$\frac{\partial}{\partial \mathbf{p}} \mathbf{a}^T(\mathbf{p}) \mathbf{b}(\mathbf{p}) = \text{vector}_i \left[ \left( \frac{\partial \mathbf{a}}{\partial p_i} \right)^T \mathbf{b} + \left( \frac{\partial \mathbf{b}}{\partial p_i} \right)^T \mathbf{a} \right] = \left( \frac{\partial \mathbf{a}^T}{\partial \mathbf{p}} \right) \mathbf{b} + \left( \frac{\partial \mathbf{b}^T}{\partial \mathbf{p}} \right) \mathbf{a} . \quad (19)$$

For a square term in  $\mathbf{p}$  we have

$$\frac{\partial}{\partial \mathbf{p}} \mathbf{p}^T \mathbf{A} \mathbf{p} = (\mathbf{A} + \mathbf{A}^T) \mathbf{p} \quad \text{if } \mathbf{A} \text{ independent of } \mathbf{p} . \quad (20)$$

$$\text{If } \mathbf{Q} = \mathbf{Q}^T \text{ independent of } \mathbf{p} , \quad \frac{\partial}{\partial \mathbf{p}} \mathbf{a}^T(\mathbf{p}) \mathbf{Q} \mathbf{a}(\mathbf{p}) = \frac{\partial}{\partial \mathbf{p}} \text{tr } \mathbf{Q} \mathbf{a} \mathbf{a}^T = 2 \frac{\partial \mathbf{a}^T}{\partial \mathbf{p}} \mathbf{Q} \mathbf{a} . \quad (21)$$

From above, if  $\mathbf{a} = \mathbf{R} \mathbf{p}$  ,  $\mathbf{R}$  constant

$$\frac{\partial}{\partial \mathbf{p}} \mathbf{p}^T \mathbf{R}^T \mathbf{Q} \mathbf{R} \mathbf{p} = 2 \mathbf{R}^T \mathbf{Q} \mathbf{R} \mathbf{p} . \quad (22)$$

## 5 Derivative of a Scalar-Valued Function with Respect to a Matrix

### 5.1 Eigenvalue Differential Quotients

The eigenvalue derivative with respect to its generating matrix  $\mathbf{A}$  (sensitivity of the eigenvalues) is

$$\frac{\partial \lambda_i[\mathbf{A}]}{\partial \mathbf{A}} = \frac{\partial (\mathbf{a}_i^{\leftarrow H} \mathbf{A} \mathbf{a}_i)}{\partial \mathbf{A}} = \frac{\partial \text{tr} [\mathbf{A} \mathbf{a}_i \mathbf{a}_i^{\leftarrow H}]}{\partial \mathbf{A}} = (\mathbf{a}_i \mathbf{a}_i^{\leftarrow H})^T = \mathbf{a}_i^{\leftarrow *} \mathbf{a}_i^T \quad (23)$$

where  $\mathbf{a}_i$  is the right eigenvector of  $\mathbf{A}$  associated with  $\lambda_i[\mathbf{A}]$  and  $\mathbf{a}_i^{\leftarrow}$  is the left eigenvector. Alternatively,

$$\frac{\partial \lambda_i[\mathbf{A}]}{\partial \mathbf{A}} = \frac{\partial (\mathbf{a}_i^{\leftarrow H} \mathbf{A} \mathbf{a}_i)}{\partial \mathbf{A}} = \frac{\partial (\mathbf{a}_i^T \mathbf{A}^T \mathbf{a}_i^{\leftarrow *})}{\partial \mathbf{A}} \quad (24)$$

$$= \frac{\partial \operatorname{tr} [\mathbf{a}_i^T \mathbf{A}^T \mathbf{a}_i^{\leftarrow *}]}{\partial \mathbf{A}} = \frac{\partial \operatorname{tr} [\mathbf{A}^T \mathbf{a}_i^{\leftarrow * \top} \mathbf{a}_i^T]}{\partial \mathbf{A}} = \mathbf{a}_i^{\leftarrow * \top} \mathbf{a}_i^T. \quad (25)$$

## 5.2 Derivative of the Singular Values

Referring to Eq.(44),

$$\frac{\partial \sigma_i}{\partial \mathbf{M}} = \frac{\partial \mathbf{u}_i^H \mathbf{M} \mathbf{v}_i}{\partial \mathbf{M}} = \mathbf{u}_i^* \mathbf{v}_i^T \quad (26)$$

$$\frac{\partial \sigma_i}{\partial (\Re \mathbf{M})} = \Re \mathbf{u}_i^* \mathbf{v}_i^T = \Re \mathbf{u}_i \mathbf{v}_i^H \quad \frac{\partial \sigma_i}{\partial (\Im \mathbf{M})} = -\Im \mathbf{u}_i^* \mathbf{v}_i^T = \Im \mathbf{u}_i \mathbf{v}_i^H \quad (27)$$

$$\frac{\partial (\sigma_i^2)}{\partial (\Re \mathbf{M})} = -(\Re \mathbf{M})(\Lambda_1^T + \Lambda_1) - (\Im \mathbf{M})(\Lambda_2^T - \Lambda_2) \quad (28)$$

$$\frac{\partial (\sigma_i^2)}{\partial (\Im \mathbf{M})} = -(\Im \mathbf{M})(\Lambda_1^T + \Lambda_1) + (\Re \mathbf{M})(\Lambda_2^T - \Lambda_2) \quad (29)$$

$$\text{where } \Lambda_1 = -\frac{\partial \lambda[\mathbf{M}^H \mathbf{M}]}{\partial (\Re \mathbf{M}^H \mathbf{M})} = -\Re \mathbf{v}_i^* \mathbf{v}_i^T, \quad \Lambda_2 = -\frac{\partial \lambda[\mathbf{M}^H \mathbf{M}]}{\partial (\Im \mathbf{M}^H \mathbf{M})} = \Im \mathbf{v}_i^* \mathbf{v}_i^T \quad (30)$$

(Sevaston, G.E., and Longman, R.W., 1988; Freudenberg, J.S., et al. 1982).

## 5.3 Derivative of the Spectral Condition Number

From

$$\kappa_s[\mathbf{M}] = \frac{\sigma_{\max}[\mathbf{M}]}{\sigma_{\min}[\mathbf{M}]} \quad \text{derive} \quad \Delta \kappa_s[\mathbf{M}] = \frac{\sigma_{\min} \Delta \sigma_{\max} - \sigma_{\max} \Delta \sigma_{\min}}{\sigma_{\min}^2}. \quad (31)$$

For  $\partial \sigma_{\max} / \partial \mathbf{M} \triangleq \mathbf{u}_a \mathbf{v}_a^T$  and  $\partial \sigma_{\min} / \partial \mathbf{M} \triangleq \mathbf{u}_i \mathbf{v}_i^T$ , where  $\mathbf{v}_i, \mathbf{v}_a, \mathbf{u}_i, \mathbf{u}_a$  are the real eigenvectors of  $\mathbf{M}^H \mathbf{M}$  and  $\mathbf{M} \mathbf{M}^H$  associated with  $\sigma_{\min}$  (subscript  $i$ ) and  $\sigma_{\max}$  (subscript  $a$ ), respectively, one has

$$\Delta \sigma_{\max} = \operatorname{tr} [\Delta \sigma_{\max}] = \operatorname{tr} [\mathbf{v}_a \mathbf{u}_a^T \Delta \mathbf{M}] \quad (32)$$

and

$$\Delta \sigma_{\min} = \operatorname{tr} [\mathbf{v}_i \mathbf{u}_i^T \Delta \mathbf{M}] \quad (33)$$

$$\Delta \kappa_s[\mathbf{M}] = \frac{\operatorname{tr} [(\sigma_{\min}[\mathbf{M}] \mathbf{v}_a \mathbf{u}_a^T - \sigma_{\max}[\mathbf{M}] \mathbf{v}_i \mathbf{u}_i^T) \Delta \mathbf{M}]}{\sigma_{\min}^2[\mathbf{M}]} \quad (34)$$

or

$$\frac{\partial \kappa_s[\mathbf{M}]}{\partial \mathbf{M}} = \frac{\sigma_{\min}[\mathbf{M}] \mathbf{u}_a \mathbf{v}_a^T - \sigma_{\max}[\mathbf{M}] \mathbf{u}_i \mathbf{v}_i^T}{\sigma_{\min}^2[\mathbf{M}]} \quad (35)$$

## 5.4 Differential Quotients of Determinant Expressions

Some determinant derivative properties are (*Athans, M., 1968; Vetter, W.J., 1971*)

$$\frac{\partial}{\partial \mathbf{M}} \det \mathbf{M} = (\mathbf{M}^{-1})^T \det \mathbf{M} = \text{adj}^T(\mathbf{M}), \quad (36)$$

$$\frac{\partial}{\partial \mathbf{M}} \det (\mathbf{M}^n) = n(\mathbf{M}^{-1})^T (\det \mathbf{M})^n, \quad (37)$$

$$\frac{\partial}{\partial \mathbf{M}} \det e^{\mathbf{M}} = \det e^{\mathbf{M}}, \quad (38)$$

$$\frac{\partial}{\partial \mathbf{M}} \log \det \mathbf{M} = (\mathbf{M}^{-1})^T. \quad (39)$$

If  $\mathbf{V}$  is independent of  $\mathbf{M}$ , 
$$\frac{\partial}{\partial \mathbf{M}} \det (\mathbf{M}\mathbf{V}\mathbf{M}^T) = 2(\mathbf{M}\mathbf{V}\mathbf{M}^T)^{-1}\mathbf{M}\mathbf{V}. \quad (40)$$

If  $\mathbf{A}$  and  $\mathbf{B}$  are square and independent of  $\mathbf{M}$ , and  $\mathbf{M}$  is square, then

$$\frac{\partial}{\partial \mathbf{M}} \det (\mathbf{A}\mathbf{M}\mathbf{B}) = (\mathbf{M}^{-1})^T \det (\mathbf{A}\mathbf{M}\mathbf{B}), \quad (41)$$

$$\frac{\partial \det[\mathbf{A}(\mathbf{H} + \mathbf{M})\mathbf{B}]}{\partial \mathbf{M}} = (\mathbf{H} + \mathbf{M})^{-T} \det[\mathbf{A}(\mathbf{H} + \mathbf{M})\mathbf{B}]. \quad (42)$$

Mostly used for controller coefficient matrix  $\mathbf{M} = \mathbf{K}$ , if  $\mathbf{A}$  and  $\mathbf{B}$  and  $\mathbf{C}$  are independent of  $\mathbf{M} = \mathbf{K}$ , one finds

$$\frac{\partial}{\partial \mathbf{K}} \det(s\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}\mathbf{C}) = -\mathbf{B}^T [\text{adj}(s\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}\mathbf{C})]^T \mathbf{C}^T. \quad (43)$$

## 5.5 Trace and Frobenius Norm Gradients

Some trace derivative properties are (*Athans, M., 1968; Vetter, W.J., 1971; Brewer, J.W., 1978; Weinmann, A., 1991*)

$$\frac{\partial}{\partial \mathbf{M}} \mathbf{a}^T \mathbf{M} \mathbf{b} = \mathbf{b} \mathbf{a}^T \quad \mathbf{M} \in \mathcal{R}^{m \times m}, \quad \mathbf{a}, \mathbf{b} \in \mathcal{R}^m. \quad (44)$$

If  $\mathbf{B}$  and  $\mathbf{C}$  independent of  $\mathbf{M}$  
$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{B}\mathbf{M}^T\mathbf{C}] = \mathbf{C}\mathbf{B}. \quad (45)$$

If  $\mathbf{A}, \mathbf{B}$  are independent of  $\mathbf{M}$ , 
$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{A}\mathbf{M}\mathbf{B}] = \mathbf{A}^T \mathbf{B}^T. \quad (46)$$

If  $\mathbf{a}, \mathbf{B}, \mathbf{b}$  are independent of  $\mathbf{M}$  
$$\frac{\partial}{\partial \mathbf{M}} (\mathbf{a}^T \mathbf{M} \mathbf{B} \mathbf{M}^T \mathbf{b}) = \mathbf{a} \mathbf{b}^T \mathbf{M} \mathbf{B}^T + \mathbf{b} \mathbf{a}^T \mathbf{M} \mathbf{B}. \quad (47)$$

If  $\mathbf{a}$  is independent of  $\mathbf{M}$  
$$\frac{\partial}{\partial \mathbf{M}} \mathbf{a}^T \mathbf{M}^{-1} \mathbf{a} = -(\mathbf{M}^{-1})^T \mathbf{a} \mathbf{a}^T (\mathbf{M}^{-1})^T. \quad (48)$$

$$n \geq 2: \quad \frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{M}^n] = n(\mathbf{M}^{n-1})^T, \quad (49)$$

$$\frac{\partial \text{tr}[\mathbf{M}]}{\partial \mathbf{M}} = \mathbf{I}, \quad (50)$$

$$\frac{\partial \|\mathbf{M}\|_F^2}{\partial \mathbf{M}} = \frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{M}\mathbf{M}^T] = 2\mathbf{M}, \quad (51)$$

$$\frac{\partial}{\partial \mathbf{M}} \|\mathbf{M}\|_F = \frac{\partial}{\partial \mathbf{M}} \sqrt{\text{tr}[\mathbf{M}\mathbf{M}^T]} = \frac{\mathbf{M}}{\|\mathbf{M}\|_F}, \quad (52)$$

$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[e^{\mathbf{M}}] = e^{(\mathbf{M}^T)} = (e^{\mathbf{M}})^T, \quad (53)$$

if  $\mathbf{A}$  independent of  $\mathbf{M}$  
$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{A}\mathbf{M}^m] = \left[ \sum_{i=0}^{m-1} \mathbf{M}^i \mathbf{A} \mathbf{M}^{m-i-1} \right]^T, \quad (54)$$

if  $\mathbf{A}$  and  $\mathbf{B}$  independent of  $\mathbf{M}$  
$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{A}\mathbf{M}\mathbf{B}\mathbf{M}] = \mathbf{A}^T \mathbf{M}^T \mathbf{B}^T + \mathbf{B}^T \mathbf{M}^T \mathbf{A}^T, \quad (55)$$

$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{A}\mathbf{M}\mathbf{B}\mathbf{M}^T] = \frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{B}\mathbf{M}^T \mathbf{A}\mathbf{M}] = \mathbf{A}^T \mathbf{M} \mathbf{B}^T + \mathbf{A} \mathbf{M} \mathbf{B}, \quad (56)$$

$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[\mathbf{A}\mathbf{M}^{-1}\mathbf{B}] = -(\mathbf{M}^{-1}\mathbf{B}\mathbf{A}\mathbf{M}^{-1})^T. \quad (57)$$

For holistic controllers one finds for the trace of the inverse of the closed-loop coefficient matrix (*Weinmann, A., 2001a*)

$$\frac{\partial}{\partial \mathbf{K}} \text{tr}[(\mathbf{A} + \mathbf{B}\mathbf{K})^{-1}] = -\mathbf{B}^T (\mathbf{A} + \mathbf{B}\mathbf{K})^{-2T}. \quad (58)$$

Furthermore,

$$\frac{\partial}{\partial \mathbf{M}} \|\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C}\|_F^2 = \frac{\partial}{\partial \mathbf{M}} \text{tr}[(\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C})^T (\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C})] = 2 \mathbf{B}^T (\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C}) \mathbf{C}^T. \quad (59)$$

**Example 1:**

$$\|\mathbf{K}\mathbf{B} - \mathbf{B}\mathbf{A}\|_F \rightarrow \min_{\mathbf{K}} \rightsquigarrow \mathbf{K} = \mathbf{B}\mathbf{A}\mathbf{B}^T (\mathbf{B}\mathbf{B}^T)^{-1}. \quad \square \quad (60)$$

$$\frac{\partial}{\partial \mathbf{M}} \text{tr}[(\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C})^T (\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C}) \mathbf{U}] = \mathbf{B}^T (\mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C}) (\mathbf{U} + \mathbf{U}^T) \mathbf{C}^T. \quad (61)$$

$$\frac{\partial}{\partial \mathbf{A}} \text{tr}[\mathbf{A}^T \mathbf{B}(\mathbf{A})] = \mathbf{B}(\mathbf{A}) + \text{matrix}_{jk} \{ \text{tr}[\mathbf{A}^T \frac{\partial \mathbf{B}(\mathbf{A})}{\partial A_{jk}}] \}. \quad (62)$$

## 5.6 Matrix Gradients with Respect to a Matrix

The matrix product rule with  $\mathbf{A} \in \mathcal{R}^{n \times m}$ ,  $\mathbf{B} \in \mathcal{R}^{m \times q}$ ,  $\mathbf{M} \in \mathcal{R}^{r \times s}$ , is

$$\frac{\partial}{\partial \mathbf{M}} \mathbf{A}(\mathbf{M})\mathbf{B}(\mathbf{M}) = \frac{\partial \mathbf{A}}{\partial \mathbf{M}} (\mathbf{I}_s \otimes \mathbf{B}) + (\mathbf{I}_r \otimes \mathbf{A}) \frac{\partial \mathbf{B}}{\partial \mathbf{M}} \quad \frac{\partial(\mathbf{A}\mathbf{B})}{\partial \mathbf{M}} \in \mathcal{R}^{nr \times qs}. \quad (63)$$

$$\frac{\partial}{\partial \mathbf{M}} a(\mathbf{M})\mathbf{B}(\mathbf{M}) = \frac{\partial a}{\partial \mathbf{M}} \otimes \mathbf{B} + \frac{\partial \mathbf{B}}{\partial \mathbf{M}} \otimes a = \frac{\partial a(\mathbf{M})}{\partial \mathbf{M}} \otimes \mathbf{B} + a(\mathbf{M}) \frac{\partial \mathbf{B}}{\partial \mathbf{M}}. \quad (64)$$

Derivative of the inverse matrix (*Vetter, W.J., 1973*)

$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{M}} = -(\mathbf{I}_r \otimes \mathbf{A}^{-1}) \frac{\partial \mathbf{A}}{\partial \mathbf{M}} (\mathbf{I}_s \otimes \mathbf{A}^{-1}) \quad \mathbf{A} \in \mathcal{R}^{n \times m}. \quad (65)$$

The chain rule for scalar functions  $\mathbf{M} \in \mathcal{R}^{r \times s}$  is (*Brewer, J.W., 1977; 1977a; 1978a*)

$$\frac{\partial a[b(\mathbf{M})]}{\partial \mathbf{M}} = \left( \mathbf{I}_r \otimes \frac{\partial a}{\partial b} \right) \left( \frac{\partial b}{\partial \mathbf{M}} \otimes \mathbf{I}_m \right) \quad (66)$$

$$= \left( \frac{\partial a}{\partial b} \mathbf{I}_r \right) \left( \frac{\partial b}{\partial \mathbf{M}} \otimes \mathbf{1} \right) = \frac{\partial a}{\partial b} \mathbf{I}_r \frac{\partial b}{\partial \mathbf{M}} = \frac{\partial a}{\partial b} \frac{\partial b}{\partial \mathbf{M}}. \quad (67)$$

Consider the transfer matrix  $\mathbf{F}_{St}(s)$  (or a related type) of the closed-loop system

$$\mathbf{F}_{St}(s) = \mathbf{F}_1(s\mathbf{I} - \mathbf{A} - \mathbf{BK})^{-1} \mathbf{F}_2 \quad \mathbf{y} = \mathbf{F}_{St} \mathbf{w}_d \quad (68)$$

$$\mathbf{B} \in \mathcal{R}^{n \times m}, \quad \mathbf{K} \in \mathcal{R}^{m \times n}, \quad \mathbf{A} \in \mathcal{R}^{n \times n}, \quad \mathbf{w}_d \in \mathcal{C}^p, \quad \mathbf{y} \in \mathcal{C}^r, \quad \mathbf{F}_2 \in \mathcal{C}^{n \times p}, \quad \mathbf{F}_1 \in \mathcal{C}^{r \times n}. \quad (69)$$

Abbreviating  $\mathbf{F}_3 = \mathbf{F}_1^H \mathbf{F}_1$ ,  $\mathbf{F}_4 = \mathbf{F}_2 \mathbf{F}_2^H$ ,  $\mathbf{U} = s\mathbf{I} - \mathbf{A} - \mathbf{BK} = \Phi_{cl}^{-1}(s)$ ,

$$\frac{\partial \|\mathbf{F}_{St}\|_F^2}{\partial \mathbf{K}} = 2\mathbf{B}^T \Re \{ \mathbf{U}^{-H} \mathbf{F}_3 \mathbf{U}^{-1} \mathbf{F}_4 \mathbf{R}^{-H} \} \quad (70)$$

(*Weinmann, A. 2001*).

The derivative of the prefilter  $\mathbf{V} = -(\mathbf{C}(\mathbf{A} + \mathbf{BK})^{-1} \mathbf{B})^{-1}$  with respect  $\mathbf{K}$  leads to the differential quotient as

$$\frac{\partial \|\mathbf{C}(\mathbf{A} + \mathbf{BK})^{-1} \mathbf{B}\|_F^2}{\partial \mathbf{K}} = 2\mathbf{T}_1^{-1} \mathbf{T}_1^{-T} \mathbf{B}^T (\mathbf{A} + \mathbf{BK})^{-T} \quad (71)$$

where  $\mathbf{T}_1 \triangleq \mathbf{C}(\mathbf{A} + \mathbf{BK})^{-1} \mathbf{B}$  (*Weinmann, A., 2001*).

## 5.7 Norms Including Complex Matrices

For complex  $\mathbf{G}$ , the derivatives of  $\text{tr}[\mathbf{G}^H \mathbf{G}]$  and  $\|\mathbf{G}\|_F$  with respect to the real and imaginary part obey

$$\frac{\partial \text{tr} [\mathbf{G}^H \mathbf{G}]}{\partial (\Re \mathbf{G})} = 2 \Re \mathbf{G} \quad \frac{\partial \text{tr} [\mathbf{G}^H \mathbf{G}]}{\partial (\Im \mathbf{G})} = 2 \Im \mathbf{G} \quad (72)$$

$$\frac{\partial \|\mathbf{G}\|_F}{\partial (\Re \mathbf{G})} = \frac{\Re \mathbf{G}}{\|\mathbf{G}\|_F} \quad \frac{\partial \|\mathbf{G}\|_F}{\partial (\Im \mathbf{G})} = \frac{\Im \mathbf{G}}{\|\mathbf{G}\|_F}. \quad (73)$$

Generalizing Eq.(59) for  $\mathbf{B}, \mathbf{C}$  complex and  $\mathbf{A}, \mathbf{M}$  real

$$\frac{\partial}{\partial \mathbf{M}} \|\mathbf{A} + \mathbf{BMC}\|_F^2 = 2\Re \mathbf{B}^H (\mathbf{A} + \mathbf{BMC}) \mathbf{C}^H \quad (74)$$

since

$$\frac{\partial}{\partial \mathbf{M}} \|\mathbf{A} + \mathbf{BMC}\|_F^2 = \frac{\partial}{\partial \mathbf{M}} \text{tr}[(\mathbf{A} + \mathbf{BMC})^H (\mathbf{A} + \mathbf{BMC})] \quad (75)$$

$$= \frac{\partial}{\partial \mathbf{M}} \text{tr} [\mathbf{A}^H \mathbf{A} + \mathbf{C}^H \mathbf{M}^H \mathbf{B}^H \mathbf{A} + \mathbf{A}^H \mathbf{BMC} + \mathbf{C}^H \mathbf{M}^H \mathbf{B}^H \mathbf{BMC}] \quad (76)$$

$$= \frac{\partial}{\partial \mathbf{M}} \text{tr} [\mathbf{A}^T \mathbf{B}^* \mathbf{M} \mathbf{C}^* + \mathbf{A}^T \mathbf{BMC} + \mathbf{B}^H \mathbf{BMC} \mathbf{C}^H \mathbf{M}^T] \quad (77)$$

$$= \mathbf{B}^H \mathbf{A} \mathbf{C}^H + \mathbf{B}^T \mathbf{A} \mathbf{C} + \mathbf{B}^T \mathbf{B}^* \mathbf{M} \mathbf{C} \mathbf{C}^T + \mathbf{B}^H \mathbf{BMC} \mathbf{C}^H \quad (78)$$

$$= 2\Re e (\mathbf{B}^H \mathbf{A} \mathbf{C}^H + \mathbf{B}^H \mathbf{BMC} \mathbf{C}^H) = 2\Re e \mathbf{B}^H (\mathbf{A} + \mathbf{BMC}) \mathbf{C}^H . \quad (79)$$

The combined expression leads to

$$\frac{\partial}{\partial \mathbf{M}} \|\mathbf{N}_1 \mathbf{M} \mathbf{N}_2 + \mathbf{N}_3 \mathbf{M} \mathbf{N}_4 + \mathbf{N}_5\|_F = \quad (80)$$

$$= 2\Re e \{ \mathbf{N}_3^H (\mathbf{N}_1 \mathbf{M} \mathbf{N}_2 + \mathbf{N}_3 \mathbf{M} \mathbf{N}_4 + \mathbf{N}_5) \mathbf{N}_4^H \} \quad (81)$$

$$+ \mathbf{N}_1^H (\mathbf{N}_1 \mathbf{M} \mathbf{N}_2 + \mathbf{N}_3 \mathbf{M} \mathbf{N}_4 + \mathbf{N}_5) \mathbf{N}_2^H \} \quad (82)$$

where  $\mathbf{N}_1, \dots, \mathbf{N}_5$  are complex and  $\mathbf{M}$  is real. To continue this result

$$\frac{\partial}{\partial \mathbf{K}} \|\mathbf{F} + \mathbf{LKR} + \mathbf{R}^H \mathbf{K}^T \mathbf{L}^H\|_F^2 = 4\Re e \{ \mathbf{L}^H (\mathbf{F} + \mathbf{LKR} + \mathbf{R}^H \mathbf{K}^T \mathbf{L}^H) \mathbf{R}^H \} . \quad (83)$$

For the closed-loop matrix-valued transfer function

$$\mathbf{T}_{cl} \triangleq \|\mathbf{T}_{ref}(s) - \mathbf{C}(s\mathbf{I} - \mathbf{A} - \mathbf{BK})^{-1} \mathbf{B}\|_F^2 \quad (84)$$

using  $\mathbf{U} \triangleq s\mathbf{I} - \mathbf{A} - \mathbf{BK} = \Phi_{cl}^{-1}(s)$  leads to (Weinmann, A., 2001)

$$\frac{\partial \mathbf{T}_{cl}}{\partial \mathbf{K}} = \frac{\partial \|\mathbf{T}_{ref}(s) - \mathbf{C} \mathbf{U}^{-1} \mathbf{B}\|_F^2}{\partial \mathbf{K}} \quad (85)$$

$$= -2\Re e \{ \mathbf{B}^H \mathbf{U}^{-H} \mathbf{C}^H (\mathbf{T}_{ref} - \mathbf{C} \mathbf{U}^{-1} \mathbf{B}) \mathbf{B}^H \mathbf{U}^{-H} \} . \quad (86)$$

## 6 Some Derivations of Matrix-Valued Functions with Respect to a Scalar

$$\frac{\partial}{\partial p} \mathbf{A}(p) \mathbf{B}(p) \mathbf{C}(p) = \frac{\partial \mathbf{A}}{\partial p} \mathbf{B} \mathbf{C} + \mathbf{A} \frac{\partial \mathbf{B}}{\partial p} \mathbf{C} + \mathbf{A} \mathbf{B} \frac{\partial \mathbf{C}}{\partial p} \quad (87)$$

$$\text{or } d(\mathbf{ABC}) = (d\mathbf{A})\mathbf{BC} + \mathbf{A}(d\mathbf{B})\mathbf{C} + \mathbf{AB}(d\mathbf{C}) , \quad (88)$$

$$\text{e.g., } \mathbf{A}^T \mathbf{X} + \mathbf{X} \mathbf{A} = \text{constant} \rightsquigarrow (d\mathbf{A}^T) \mathbf{X} + \mathbf{A}^T d\mathbf{X} + (d\mathbf{X}) \mathbf{A} + \mathbf{X} d\mathbf{A} = \mathbf{0} . \quad (89)$$

$$\frac{\partial \mathbf{A}^m}{\partial p} = \frac{\partial \mathbf{A}}{\partial p} \mathbf{A}^{m-1} + \mathbf{A} \frac{\partial \mathbf{A}}{\partial p} \mathbf{A}^{m-2} + \dots + \mathbf{A}^{m-1} \frac{\partial \mathbf{A}}{\partial p} . \quad (90)$$

From *Doyle, J.C. et al., 1996*,

$$\frac{\partial \mathbf{A}^{-1}}{\partial p} = -\mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial p} \mathbf{A}^{-1}. \quad (91)$$

Total derivative if  $c$  depends on  $t$ , additionally,

$$\frac{d}{dt} c[\mathbf{x}(t), t] = \frac{d\mathbf{x}^T(t)}{dt} \frac{\partial c}{\partial \mathbf{x}} + \frac{\partial c}{\partial t}. \quad (92)$$

Vector (column) function  $\mathbf{z}$  dependent on  $\mathbf{x}(t)$  and  $t$ :  $\mathbf{z} \in \mathcal{R}^m$ ,  $\mathbf{x}(t) \in \mathcal{R}^q$ ,  $\mathbf{M} := t$  (*Vetter, W.J., 1970*)

$$\frac{d\mathbf{z}[\mathbf{x}(t)]}{dt} = \left( \frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_m \right) \left( \mathbf{1} \otimes \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right) = \left( \frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_m \right) \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \quad (93)$$

$$\frac{d\mathbf{z}[\mathbf{x}(t), t]}{dt} = \left( \frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_m \right) \frac{\partial \mathbf{z}}{\partial \mathbf{x}} + \frac{\partial \mathbf{z}}{\partial t}. \quad (94)$$

Row function  $\mathbf{y}^T$  dependent on  $\mathbf{x}(t)$  and  $t$  with  $\mathbf{y} \in \mathcal{R}^n$

$$\frac{d\mathbf{y}^T[\mathbf{x}(t), t]}{dt} = \left( \frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_1 \right) \left( \mathbf{I}_1 \otimes \frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} \right) = \frac{d\mathbf{x}^T}{dt} \frac{\partial \mathbf{y}^T}{\partial \mathbf{x}}. \quad (95)$$

Matrix-valued function  $\mathbf{A}$  dependent on the vector  $\mathbf{x}(t) \in \mathcal{R}^q$ ,  $\mathbf{A} \in \mathcal{R}^{n \times m}$ ,

$$\frac{d\mathbf{A}[\mathbf{x}(t), t]}{dt} = \left( \frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_n \right) \left( \mathbf{I}_1 \otimes \frac{\partial \mathbf{A}}{\partial \mathbf{x}} \right) + \frac{\partial \mathbf{A}}{\partial t} = \left( \frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_n \right) \frac{\partial \mathbf{A}}{\partial \mathbf{x}} + \frac{\partial \mathbf{A}}{\partial t}. \quad (96)$$

**Example 2:** For  $\mathbf{y} \in \mathcal{R}^n$ ,  $\mathbf{z} \in \mathcal{R}^m$ ,  $\mathbf{x} \in \mathcal{R}^q$ ,  $\mathbf{A} \in \mathcal{R}^{n \times m}$  derivatives on the product  $c = \mathbf{y}^T[\mathbf{x}(t), t] \mathbf{A}[\mathbf{x}(t), t] \mathbf{z}[\mathbf{x}(t), t]$  (*Vetter, W.J., 1970*)

$$\frac{d}{dt} \mathbf{y}^T[\mathbf{x}(t), t] \mathbf{A}[\mathbf{x}(t), t] \mathbf{z}[\mathbf{x}(t), t] = \frac{d\mathbf{y}^T}{dt} \mathbf{A} \mathbf{z} + \mathbf{y}^T \frac{d\mathbf{A}}{dt} \mathbf{z} + \mathbf{y}^T \mathbf{A} \frac{d\mathbf{z}}{dt}. \quad (97)$$

$$\frac{\partial c}{\partial \mathbf{y}} = \mathbf{A} \mathbf{z} \quad \frac{\partial c}{\partial \mathbf{z}} = (\mathbf{y}^T \mathbf{A})^T = \mathbf{A}^T \mathbf{y} \quad (98)$$

$$\frac{\partial c}{\partial \mathbf{A}} = \frac{\partial}{\partial \mathbf{A}} \text{tr}(\mathbf{y}^T \mathbf{A} \mathbf{z}) = \frac{\partial}{\partial \mathbf{A}} \text{tr}(\mathbf{z} \mathbf{y}^T \mathbf{A}) = (\mathbf{z} \mathbf{y}^T)^T = \mathbf{y} \mathbf{z}^T. \quad \square \quad (99)$$

**Example 3:**

$$\frac{\partial}{\partial \mathbf{x}^T} \mathbf{x}^T \mathbf{Q}(\mathbf{x}) \mathbf{x} = \mathbf{x}^T [\mathbf{Q}^T(\mathbf{x}) + \mathbf{Q}(\mathbf{x}) + \frac{\partial \mathbf{Q}(\mathbf{x})}{\partial \mathbf{x}^T} (\mathbf{I}_q \otimes \mathbf{x})] \quad \mathbf{x} \in \mathcal{R}^q. \quad \square \quad (100)$$

**Example 4:**

$$\frac{d}{dt} e^{\mathbf{A}t} = \mathbf{A} e^{\mathbf{A}t} = e^{\mathbf{A}t} \mathbf{A} \quad \text{and} \quad \int_0^t e^{\mathbf{A}\tau} d\tau = \mathbf{A}^{-1} (e^{\mathbf{A}t} - \mathbf{I}). \quad \square \quad (101)$$

## 7 Incremental Notation, Fréchet Derivatives

In this section, the incremental notation is addressed as advantageously applicable in control theory. The increment of a matrix  $\mathbf{A}$  is termed  $\Delta\mathbf{A}$ .

For  $\mathbf{A}, \mathbf{M} \in \mathcal{R}^{n \times m}$

$$f(\mathbf{M}) = \text{tr}[\mathbf{A}^T \mathbf{M}] \quad (102)$$

$$\Delta f(\mathbf{M}, \Delta\mathbf{M}) \triangleq f(\mathbf{M} + \Delta\mathbf{M}) - f(\mathbf{M}) = \text{tr}[\mathbf{A}^T \Delta\mathbf{M}] = \text{tr}[\Delta\mathbf{M}^T \cdot \mathbf{A}] \quad (103)$$

and

$$\frac{\partial f}{\partial \mathbf{M}} = \mathbf{A} . \quad (104)$$

The factor of the increment  $\Delta\mathbf{M}$ , having been transposed, equals the derivative.

### 7.1 Increment of the Frobenius Norm

Suppose real matrices  $\mathbf{K}, \mathbf{L}, \mathbf{R}$  with appropriate dimension, the increment  $\Delta\mathbf{K}$  in  $\mathbf{K}$  causes an increment (*Vetter, W.J., 1971*)

$$\begin{aligned} \Delta \|\mathbf{LKR}\|_F^2(\mathbf{K}, \Delta\mathbf{K}) &\triangleq \|\mathbf{L}(\mathbf{K} + \Delta\mathbf{K})\mathbf{R}\|_F^2 - \|\mathbf{LKR}\|_F^2 \\ &= \text{tr} \{2\mathbf{L}^T \mathbf{LKR} \mathbf{R}^T \Delta\mathbf{K}^T\} . \end{aligned} \quad (105)$$

Hence,

$$\frac{\partial}{\partial \mathbf{K}} \|\mathbf{LKR}\|_F^2 = 2\mathbf{L}^T \mathbf{LKR} \mathbf{R}^T . \quad (106)$$

The prefactor matrix in Eq.(105) corresponds to the differential quotient with respect to  $\mathbf{K}$ .

The differential or increment of  $\|\mathbf{A} - \mathbf{KC}\|_F^2$  with respect to  $\mathbf{A}$  is

$$\Delta \|\mathbf{A} - \mathbf{KC}\|_F^2(\mathbf{A}, \Delta\mathbf{A}) \triangleq \|\mathbf{A} + \Delta\mathbf{A} - \mathbf{KC}\|_F^2 - \|\mathbf{A} - \mathbf{KC}\|_F^2 = \quad (107)$$

$$\begin{aligned} &= \text{tr} [\mathbf{A}^T \Delta\mathbf{A}] + \text{tr} [\Delta\mathbf{A}^T \mathbf{A}] - 2\text{tr} [(\mathbf{KC})^T \Delta\mathbf{A}] \\ &= 2\text{tr} \{[\mathbf{A}^T - (\mathbf{KC})^T] \Delta\mathbf{A}\} \end{aligned} \quad (108)$$

where

$$\|\mathbf{A} - \mathbf{KC}\|_F^2 = \text{tr} [(\mathbf{A} - \mathbf{KC})^T (\mathbf{A} - \mathbf{KC})] \quad (109)$$

$$= \text{tr} [\mathbf{A}^T \mathbf{A}] - 2\text{tr} [(\mathbf{KC})^T \mathbf{A}] + \text{tr} [(\mathbf{KC})^T \mathbf{KC}] \quad (110)$$

has been used.

## 7.2 Increment of the Inverse

The increment of the inverse is termed  $\Delta(\mathbf{X}^{-1})$  and results from

$$\mathbf{X}^{-1}\mathbf{X} = \mathbf{I} \rightsquigarrow [\Delta(\mathbf{X}^{-1})]\mathbf{X} + \mathbf{X}^{-1}\Delta\mathbf{X} = \mathbf{0} \quad (111)$$

$$\Delta(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}(\Delta\mathbf{X})\mathbf{X}^{-1}. \quad (112)$$

If  $\mathbf{A} = \mathbf{G}\mathbf{X}^{-1}$  and both  $\mathbf{G}$  and  $\mathbf{X}$  are variable

$$\Delta\mathbf{A} = (\Delta\mathbf{G})\mathbf{X}^{-1} + \mathbf{G}\Delta(\mathbf{X}^{-1}) = (\Delta\mathbf{G})\mathbf{X}^{-1} - \mathbf{G}\mathbf{X}^{-1}(\Delta\mathbf{X})\mathbf{X}^{-1} \quad (113)$$

$$= (\Delta\mathbf{G})\mathbf{X}^{-1} - \mathbf{A}(\Delta\mathbf{X})\mathbf{X}^{-1} = (\Delta\mathbf{G} - \mathbf{A}\Delta\mathbf{X})\mathbf{X}^{-1}, \quad (114)$$

leading to

$$\Delta\|\mathbf{A} - \mathbf{K}\mathbf{C}\|_F^2(\mathbf{A}, \Delta\mathbf{X}) = 2\text{tr} \{[\mathbf{A}^T - (\mathbf{K}\mathbf{C})^T](\Delta\mathbf{G} - \mathbf{A}\Delta\mathbf{X})\mathbf{X}^{-1}\}. \quad (115)$$

## 7.3 Matrix Exponential

Furthermore for

$$g(\mathbf{M}) = \text{tr}[e^{\mathbf{M}}] \quad (116)$$

$$\Delta g(\mathbf{M}, \Delta\mathbf{M}) = \text{tr}[e^{\mathbf{M}}\Delta\mathbf{M}] \quad (117)$$

$$\frac{\partial g(\mathbf{M})}{\partial \mathbf{M}} = e^{\mathbf{M}^T}. \quad (118)$$

Finally (*Levine, W.S., and Athans, M., 1970; Bellman, R., 1960*)

$$\mathbf{h}(\mathbf{M}) = e^{\mathbf{B}\mathbf{M}t} \quad (119)$$

$$\Delta\mathbf{h}(\mathbf{M}, \Delta\mathbf{M}) = \int_0^t e^{\mathbf{B}\mathbf{M}(t-\sigma)}\mathbf{B}\Delta\mathbf{M}e^{\mathbf{B}\mathbf{M}\sigma}d\sigma. \quad (120)$$

## 8 Matrix Element Interdependencies

For specific matrix interdependencies,

$$\frac{\partial \mathbf{M}}{\partial M_{ik}} = \mathbf{E}_{ik} + \mathbf{E}_{ki} - \mathbf{E}_{ii}\delta_{ik}, \quad \mathbf{M} \text{ symmetric} \quad (121)$$

$$\frac{\partial \mathbf{M}}{\partial M_{ik}} = \mathbf{E}_{ik} - \mathbf{E}_{ki}, \quad \mathbf{M} \text{ skew symmetric} \quad (122)$$

where  $\mathbf{E}_{ik} = \mathbf{e}_i\mathbf{e}_k^T$  where  $\mathbf{e}_i$  is a vector with zeros except in the  $i$ th position, and  $\delta_{ik} = 1$  for  $i = k$  and  $= 0$  for  $i \neq k$ .

The relation

$$\frac{\partial \det \mathbf{M}}{\partial \mathbf{M}} = (\mathbf{M}^{-1}) \cdot \det \mathbf{M} \quad (123)$$

is not true if there are matrix element interdependencies. Normally,

$$\mathbf{M} = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \quad \det \mathbf{M} = m_1 m_4 - m_2 m_3 \quad (124)$$

$$\frac{\partial \det \mathbf{M}}{\partial \mathbf{M}} = \begin{pmatrix} m_4 & -m_3 \\ -m_2 & m_1 \end{pmatrix} \quad (125)$$

However, if  $m_3 = m_2$  one has

$$\det \mathbf{M} = m_1 m_4 - m_2^2 \quad \frac{\partial \det \mathbf{M}}{\partial \mathbf{M}} = \begin{pmatrix} m_4 & -2m_2 \\ -2m_2 & m_1 \end{pmatrix}. \quad (126)$$

For symmetric matrices, there is the general relation

$$\frac{\partial \det \mathbf{M}}{\partial \mathbf{M}} = (2\mathbf{M}^{-1} - \mathbf{diag} \{\mathbf{M}^{-1}\}) \det \mathbf{M} \quad (127)$$

where  $\mathbf{diag} \{\mathbf{M}^{-1}\}$  is a diagonal matrix made of main-diagonal elements of  $\mathbf{M}^{-1}$ .

For  $f(\mathbf{M}) \triangleq \text{tr} [\mathbf{L}\mathbf{M}]$

$$\frac{\partial f(\mathbf{M})}{\partial \mathbf{M}} = \begin{cases} (\mathbf{L}^T + \mathbf{L} - \mathbf{diag} \mathbf{L}) & \text{for } \mathbf{M} \text{ symmetric} \\ (\mathbf{L}^T - \mathbf{L}) & \text{for } \mathbf{M} \text{ skew symmetric.} \end{cases} \quad (128)$$

For  $\mathbf{M}$  symmetric the linear relation  $\mathbf{L}_1\mathbf{M}\mathbf{L}_2$  yields (*Brewer, J.W., 1977*)

$$\frac{\partial}{\partial \mathbf{M}} \text{tr} [\mathbf{L}_1\mathbf{M}\mathbf{L}_2] = \mathbf{L}_1^T \mathbf{L}_2^T + \mathbf{L}_2 \mathbf{L}_1 - \mathbf{diag} \{\mathbf{L}_2 \mathbf{L}_1\} \quad (129)$$

and for the relation of second order

$$\frac{\partial}{\partial \mathbf{M}} \text{tr} [\mathbf{L}_1\mathbf{M}\mathbf{L}_2\mathbf{M}] = (\mathbf{L}_2\mathbf{M}\mathbf{L}_1)^T + \mathbf{L}_1\mathbf{M}\mathbf{L}_2 - \mathbf{diag} \{\mathbf{L}_2\mathbf{M}\mathbf{L}_1 + \mathbf{L}_1\mathbf{M}\mathbf{L}_2\}. \quad (130)$$

**Example 5:** With no interdependencies

$$\mathbf{k} = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} \quad \frac{\partial \mathbf{k}^T \mathbf{k}}{\partial \mathbf{k}} = 2\mathbf{k}. \quad (131)$$

With interdependency  $k_2 = \alpha k_1$

$$\frac{\partial \mathbf{k}^T \mathbf{k}}{\partial \mathbf{k}} = 2 \mathbf{diag} [1 + \alpha^2, 1 + \frac{1}{\alpha^2}] \cdot \mathbf{k} = 2 \begin{pmatrix} 1 + \alpha^2 & 0 \\ 0 & 1 + 1/\alpha^2 \end{pmatrix} \mathbf{k}; \quad (132)$$

for  $k_2 = k_1^2$

$$\frac{\partial \mathbf{k}^T \mathbf{k}}{\partial \mathbf{k}} = \begin{pmatrix} 2 & 4k_1 \\ 1/k_1 & 2 \end{pmatrix} \mathbf{k}. \quad \square \quad (133)$$

The reason for the strong differences in the differential quotients for occasionally existing matrix interdependencies is explained easily. The partial derivative  $\frac{\partial}{\partial \mathbf{M}}$  or  $\frac{\partial}{\partial M_{ij}} \forall i, j$  is an expression for keeping all the other variables  $\neq i, j$  constant which is in contradiction to the matrix interdependencies. In fact, each interdependency causes one variable less which is satisfiable by a partial derivative.

In the Example 5 for  $k_2 = k_1^2$  one has  $\mathbf{k}^T \mathbf{k} = k_1^2 + k_1^4 = k_2 + k_2^2$ , the derivatives are

$$\frac{\partial}{\partial k_1} \mathbf{k}^T \mathbf{k} = 2 k_1 + 4 k_1^3 \quad \frac{\partial}{\partial k_2} \mathbf{k}^T \mathbf{k} = 1 + 2 k_2 = 1 + 2 k_1^2 \quad (134)$$

and the interrelation  $\frac{\partial}{\partial k_1} = \frac{dk_2}{dk_1} \frac{\partial}{\partial k_2} = 2 k_1 \frac{\partial}{\partial k_2}$  exists.

In the Example 5 for  $\alpha = 1$

$$\mathbf{k}^T \mathbf{k} = 2 k_1^2 \quad \frac{\partial}{\partial k_1} = \frac{\partial}{\partial k_2} = 4 k_1 . \quad (135)$$

$\frac{\partial}{\partial k_1}$  cannot be calculated while keeping  $k_2$  constant. Both differential quotients are in fact total derivatives and turn out to be the same.

## 9 Concluding Remarks

Gradients and matricial gradients of several scalar functions of vectors and matrices have been presented as a collection.

Numerically checking these results or new ones can be carried out as follows. For  $\frac{\partial f}{\partial \mathbf{M}}$  the  $i, j$  position in the matrix differential is  $\frac{\partial f}{\partial M_{ij}}$ . It is approximated by

$$\frac{\partial f}{\partial M_{ij}} = \lim_{\Delta M_{ij} \rightarrow 0} \frac{\Delta f(\mathbf{M}, \Delta M_{ij})}{\Delta M_{ij}} = \lim_{\Delta M_{ij} \rightarrow 0} \frac{f(\mathbf{M}) - f(\mathbf{M} + \mathbf{E}_{ij} \cdot \Delta M_{ij})}{\Delta M_{ij}} . \quad (136)$$

Examples for the analytical derivation are given in Eq.(79) and in *Weinmann, A., 2001; 2001a*.

## References

- Athans, M., 1968*, The matrix minimum principle, *Information and Control* **11**, pp. 592-606
- Bellman, R., 1960*, Introduction to Matrix Analysis (New York, McGraw-Hill)
- Brewer, J.W., 1977*, The gradient with respect to a symmetric matrix, *IEEE-Trans. AC-22*, pp. 265-267

- Brewer, J.W., 1977a, The derivative of the exponential matrix with respect to a matrix, *IEEE-Trans.* **AC-22**, pp. 656-657
- Brewer, J.W., 1978, Kronecker products and matrix calculus in system theory, *IEEE-Trans.* **CAS-25**, pp. 772-781
- Brewer, J.W., 1978a, Matrix calculus and the sensitivity analysis of linear dynamic systems, *IEEE-Trans.* **AC-23**, pp. 748-751
- Cheok, K.C., Loh, N.K., Zohdy, M.A., 1986, Generalized optimal feedback controllers for discrete-time linear systems, *Journal of Dynamic Systems, Measurement, and Control* **108**, pp. 91-95
- Doyle, J.C., Glover, K., Zhou, K., 1996, Robust and Optimal Control (Englewood Cliffs, Prentice Hall)
- Freudenberg, J.S., Looze, D.P., and Cruz, J.B., 1982, Robustness analysis using singular value sensitivities, *Int.J.Control* **35**, pp. 95-116
- Geering, H.P., 1976, On calculating gradient matrices, *IEEE Trans.* **AC-21**, pp. 615-616
- Levine, W.S., and Athans, M., 1970, On the determination of the optimal constant output feedback gains for linear multivariable systems, *IEEE-Trans.* **AC-15**, pp. 44-48
- Sevaston, G.E., and Longman, R.W., 1988, Systematic design of robust control systems using small gain stability concepts, *Int.J.Systems Sci.* **19**, pp. 439-451
- Vetter, W.J., 1970, Derivative operations on matrices, *IEEE-Trans.* **AC-15**, pp.241-244 and **AC-16**, p.113
- Vetter, W.J., 1971, An extension to gradient matrices, *IEEE-Trans.* **SMC-1**, pp. 184-186
- Vetter, W.J., 1973, Matrix calculus operations and Taylor expansions, *SIAM Review* **15**, pp.352-369
- Vetter, W.J., Balchen, J.G., and Pohner, F., 1972, Near-optimal control of parameter-dependent linear plants, *Int.J.Control* **15**, pp. 683-691
- Weinmann, A., 1991, Uncertain Models and Robust Control (Vienna and NewYork: Springer)
- Weinmann, A., 2001, Matrix assignment subject to Lyapunov and closed-loop norm conditions, *2001 American Control Conference, Arlington, Virginia*, MP12-5
- Weinmann, A., 2001a, Control system design based on holistic eigenvalue allocation, e&zi *Elektrotechnik und Informationstechnik* **118**, H. 4, S. 167-173

## ACE 2000

### 5th IFAC Symposium on Advances in Control Education

Gold Coast, Australien  
17. – 19. Dezember 2000

Das 5. IFAC-Symposium „Advances in Control Education“ (ACE 2000) wurde unter den Auspizien des IFAC Technical Committee on Control Education (EDCOM) sowie der Institution of Engineers of Australia hervorragend ausgerichtet. Es umfasste 3 Plenarsitzungen, eine Halbplenarsitzungen sowie 16 reguläre Sitzungen in denen in 2 Parallelschienen insgesamt 60 Beiträge präsentiert wurden. Die insgesamt 68 Teilnehmer am Symposium kamen aus 31 Ländern aller 5 Kontinente.

Die Themen der Plenarsitzungen *New Directions in Control Engineering Education. A North American Perspective* (Prof. Panos Antaklis, USA), *Control Engineering Education: A View from Industrial and Academic Perspective* (Dr. Sam Crisafully, Australien) und *A Perspective of Distance Learning – Tool Based Control Education System* (Prof. Fumio Harashima, Japan) zeigten einige neue und sehr interessante Gesichtspunkte auf. Speziell betont wurde dabei die zunehmende Bedeutung des Fernunterrichts und des Angebots von außeruniversitären, von Privatfirmen angebotenen Lehrveranstaltungen. Es stellte sich dabei heraus, dass Australien auf diesen Gebieten, bedingt durch die geographischen Bedingungen (große Entfernungen) eine Vorreiterrolle zu spielen scheint. In der Halbplenarsitzung *Control Education in Australia – Past and Present* wurde noch näher auf die australischen Verhältnisse eingegangen.

Die Themen der regulären Sitzungen reichten von *Control Education Curricula, Laboratory Tools and Experiments, Real Time Control, Control Topics: Educational Perspective* bis hin zu *Internet Based Control*. Die Interessen des Berichterstatters konzentrierten sich dabei vornehmlich auf Vorträge aus den letzteren 2 Themenkreisen. Eine breite Diskussion entwickelte sich in jenen Sitzungen, die sich mit der Lehre der Regelungstechnik in einer ersten, einführenden Lehrveranstaltung befassten. Dabei wurden grundsätzlich neue Philosophien diskutiert, die vor allem eine Einbindung moderner Informationstechnologien in die Lehre in den Vordergrund stellen. Einige vielversprechende Entwicklungen wurden auf dem Gebiet des *Remote Laboratory*, das wohl hauptsächlich für Fernuniversitäten von Interesse ist, als auch auf dem Gebiet des *Virtual Laboratory* (regelungstechnische Versuche via Internet) präsentiert. In der Video-Competition, die auf vergangenen Symposien sehr erfolgreich war, konnte diesmal leider kein Sieger ermittelt werden.

Zusammenfassend kann gesagt werden, dass auch dieses Symposium, wie schon seine Vorgängerveranstaltungen, für alle an der Lehre der Regelungstechnik interessierten Teilnehmer neue und interessante Aspekte aufzeigen konnte. Leider war, wohl bedingt durch die geographische Lage des Veranstaltungsortes, die Teilnehmerzahl diesmal nicht ganz den Erwartungen der Veranstalter entsprechend. Das nächste und sechste ACE Symposium in dieser Reihe (ACE 2003) wird vom 16. - 18. Juni 2003 an der Universität von Oulu in Finnland stattfinden.

H. Peter Jörgl

**IFAC International Workshop on  
Knowledge and Technology Transfer: „Automatic systems for  
building the infrastructure in developing countries“**

**Ohrid, Macedonia  
21.-23. Mai 2001**

Das IFAC TC on Developing Countries begann vor drei Jahren in Südafrika eine neue Symposiumsserie unter der Headline Technologietransfer in Entwicklungsländern. Dieses zweite Symposium wurde von der ETAI, der mazedonischen NMO der IFAC organisiert. Das Programm bestand aus 3 Survey papers und 64 Technical papers. Infolge der politischen Situation in Mazedonien kamen aber statt der erhofften 70 nur etwa 40 Teilnehmer.

Wie bei solchen Tagungen üblich, spannte sich der Bogen von hochtheoretischen Beiträgen, über praktische Anwendungsberichte bis hin zu Ausbildungsfragen.

Die beiden Survey papers behandelten einerseits die Installierung einer Fluginfrastruktur in Entwicklungsländern (S. Khane, USA) sowie die Steuerung der Übertragungsraten in Rechnernetzwerken und zukünftige Potentiale für Softcomputing (G. M. Dimirowski, Mazedonien). Wie vorher erwähnt, war der Bogen der Themen sehr weit gespannt, was in einigen Session Titeln zum Ausdruck gebracht werden soll: Rechnergestützte Fertigung, Automatisierung und Robotertechnik; Ausbildung in Entwicklungsländern; Industrieautomatisierung und Überwachung in Energiesystemen; Steuerung, Management und Überwachung in Umwelt und Wassersystemen; Verbesserung der Infrastruktur für Industrieautomatisierung; Überwachung und Automatisierung in Telekommunikationsnetzwerken; Steuerung und Regelung in biomedizinischen Systemen; künstliche Intelligenz für Systemsteuerung und Überwachung sowie Automatisierung und Überwachung in Verkehrssystemen.

2 Sessions waren der Thematik „Systemtechnische Methoden zur Konfliktlösung (SWIIS)“ gewidmet. Die Beiträge in diesen Sessions können auch als sehr heterogen bezeichnet werden.

Ein ausführliches Sozialprogramm ergänzte die – trotz widerwärtiger äußerer Umstände – bestens organisierte Veranstaltung.

P. Kopacek

**ACC 2001****American Control Conference 2001  
Advances in Control Education**

Arlington VA, USA

25.-27. Juni 2001

Die American Control Conference 2001 fand vom 25.-27. Juni im Crystal Gateway Marriott Hotel in Arlington, Virginia, USA statt. Sie wurde vom American Automatic Control Council in Kooperation mit IFAC ausgerichtet. Die Konferenz umfasste 3 Plenarsitzungen (jeweils am Beginn jeden Tages), 9 Tutorial Invited Sessions sowie 3 Spezialsitzungen zu den Themen *National Research Initiatives*, *The History of Feedback Control* und *From Theory to Wallstreet*. Die ca. 1000 *Invited Papers* und normalen Kongressbeiträge wurden in 20 Parallelschienen (jeweils am morgens, mittags und nachmittags) präsentiert. Erstmals kamen im Rahmen dieser ACC auch Poster-Sessions hinzu, wodurch die Anzahl der Parallelsessions in gerade noch vertretbaren Grenzen gehalten werden konnte.

Die Themen der Sitzungen umfassten, wie anlässlich derartiger Großkongresse üblich, den gesamten Bereich der Regelungstechnik. Die Balance zwischen theoretischen und anwendungsorientierten Themen wurde dabei im wesentlichen ganz gut erreicht. Auf Seite der mehr theoretischen Themenkreise fanden Sessions über *Nonlinear Control*, *Adaptive Control*, *Predictive Control* *Fuzzy Control* und *Discrete Time Systems* statt. Für den Bericht besonders interessante Arbeiten wurden in den Sessions über das Thema *Fault Detection* präsentiert. Dabei war deutlich festzustellen, dass das Interesse nicht wie bisher nur der Theorie galt, sondern, dass auch der praktischen Anwendung immer mehr Augenmerk geschenkt wird. Von den Sitzungen die, sich vornehmlich regelungstechnischen Anwendungen widmeten, sind jene zu den Themenkreise *Automotive Control* und *Vehicle Control* besonders hervorzuheben.

Hervorzuheben sind die *Invited Tutorial Sessions* in denen unter anderem Themen wie *Modeling and Control of Economic Systems*, *Monitoring and Control of Civil Structures*, *Fault Tolerant Control*, *Failure Diagnosis of Dynamic Systems Using Discrete Event Models* und *Communication and Control of Distributed Hybrid Systems* behandelt wurden. Für den Bericht bildete die Neuerung definitiv eine Bereicherung für die Konferenz.

Zusammenfassend kann über die ACC 2001 gesagt werden, dass, obwohl die Konferenz ausgezeichnet organisiert war, 20 Parallelsitzungen doch um einige zu viele sind. Man war als Teilnehmer einer derartigen Veranstaltung aufgrund des enorm großen Angebotes doch etwas überfordert und hatte Schwierigkeiten sich einen Überblick zu verschaffen.

H.P.Jörgl

# **IFAC Conference on Telematics Applications In Automation and Robotics**

**Weingarten, Deutschland**

**24. - 26. Juli 2001**

Der bereits 1998 in Wien vorgesehene IFAC Workshop "Telematics a new dimension of automation" mußte trotz Interesse der damaligen Generaldirektion für Informationstechnologie der Europäischen Gemeinschaften wegen zu geringem Interesse abgesagt werden. Bei der ersten IFAC-Konferenz über "Telematics Applications in Automation and Robotics" 2001 in Weingarten waren 101 Beiträge und 139 Teilnehmer aus siebzehn Ländern. Von der Industrie kamen 39 Teilnehmer. Nun für die Benutzung des Internets gibt es ja auch schon Seniorenkurse.

Der Themenkreis umfaßte Roboter, Telemedizin, Satellitentechnik, Telematik in der Landwirtschaft und für den Verkehr, Tele-Laboratorien, Telewartung und Kommunikationslösungen.

Vieles ist wohl noch in der Lernphase. In einem Beitrag wurde versucht die Steuerung eines Modellautos über WAP (Wireless Application Protocol) zu zeigen, leider war das Internet etwas überlastet. Viele Lösungen stützen sich auf Java, statt Jini, und Bluetooth ab. In der Diskussion stellte sich heraus, daß dies nicht die besten Lösungen wären und nur wegen der Verfügbarkeit und dem bereits Kennen benutzt werden. Eine von einem Hochschulinstitut entwickelte Lösung verwendete für Bauteile eines deutschen Konzerns LON (Local Operating Network) statt dem Europäischen Installations Bus (EIB). Wegen mangelnder Informationsweitergabe von Seiten des europäischen Konzerns konnte EIB nicht benutzt werden! CORBA (Common Object Request Broker Architecture) wird bereits sehr breit angewendet. Ein Detail hinsichtlich allgemeiner Vorstellung und Wirklichkeit: Für ein Intelligentes Verkehrs-System in China für Autostraßen konnten nicht bereits bewährte Lösungen eingesetzt werden, da im Gegensatz zu anderen Ländern auch Radfahrer zu berücksichtigen sind.

Prof. Klaus Schilling und der Fachhochschule Weingarten kann nur gratuliert werden, den Themenkreis der Telematik im Rahmen der IFAC zum Durchbruch verholfen zu haben. Es kann nur begrüßt werden, wenn eine weitere Tagung 2004, vielleicht in Skandinavien, stattfindet. Vielleicht sollten im Sinne eines Systemansatzes bei diesem holistischen Thema die Fachleute der anderen Seite, zum Beispiel Informationsverarbeitung und neue Anwendungsbereiche, mehr eingebunden werden. Es bestehen Bestrebungen im Rahmen der IFAC ein eigenes Technisches Komitee für Telematik einzurichten.

Die Örtlichkeit, die Betreuung und die Gastfreundschaft der Gemeinden Weingarten und Ravensburg lassen jeden Teilnehmer gerne wieder zu einer Veranstaltung an der Fachhochschule Weingarten kommen.

Robert Genser

## Verleihung des Fred Margulies Preises 2000

Im Anschluss an den 15. Österreichischen Automatisierungstag, welcher dieses Mal am 13. Oktober 2000 in Linz im Rahmen der Smart Automation Austria – Fachmesse für industrielle Elektronik – stattfand, wurde der mit ATS 60.000,-- dotierte Fred Margulies Preis vergeben.

Dieser Preis wird vom IFAC-Beirat Österreich und der Gewerkschaft der Privatangestellten je zur Hälfte zur Verfügung gestellt. Er wird – ganz im Sinne des verstorbenen Prof. Fred Margulies – für hervorragende österreichische Arbeiten auf dem Gebiet der Automatisierungstechnik mit besonderer Berücksichtigung der sozialen Aspekte bei ihrer Einführung in Betrieben verliehen. Kriterien für die Bewertung durch eine Fachjury sind im wesentlichen Interdisziplinarität des Ansatzes und insbesondere soziale Relevanz.

Die Jury beschloss heuer den Preis folgenden Arbeiten zu gleichen Teilen zuzuerkennen:

- DI Konrad Daichendt: Lagerlogistik im Langblocklager der Firma EUROFAM GmbH
- Dr. Martin Mandl: Adaption eines Interferometers zur Messung der Dicke der menschlichen Hornhaut
- Georg Schitter: Robuste Regelung von Rasterkraftmikroskopen.

Die feierliche Überreichung an die Preisträger erfolgte durch Frau Bundesminister a.D. Eleonore Hostasch. Die Laudatio wurde von Herrn Sektionschef Dr. Norbert Rozsenich, BMVIT, gehalten. Unter den zahlreich erschienen Ehrengästen war auch die Witwe von Prof. Margulies, Frau Paula Margulies.

Die Verleihung des Fred Margulies Preises 2001 erfolgt anlässlich des 16. Österreichischen Automatisierungstages im Rahmen der *viet* im Congress Centrum Wien am 16. Oktober 2001.

I. Nemetz

# **Mechatronische Systeme Grundlagen**

R. Isermann

Springer-Verlag Berlin Heidelberg, 1999  
412 S., 213 Abbildungen  
ISBN 3-540-64725-2

Mechatronische Systeme stellen eine Integration von mechanisch-feinwerktechnischen Systemen, elektronischen Systemen sowie der dazugehörigen Informationsverarbeitung dar. Diese Integration kann örtlich durch den Aufbau und funktionell durch die digitale Elektronik erfolgen. Die örtliche Integration erfolgt dabei durch die konstruktive Verschmelzung von Aktoren, Sensoren und der Mikroelektronik mit dem Prozeß, also im Wesentlichen durch die Hardware. Die funktionelle Integration wird entscheidend durch die Informationsverarbeitung und damit durch die Gestaltung der Software geprägt.

Das vorliegende Buch gibt eine Einführung in den Aufbau mechatronischer Systeme, beginnend mit der Modellbildung bis zu den Komponenten zur Informationsverarbeitung. Zum Entwurf mechatronischer Systeme sind Modellbildung, Simulation, rechnergestützte Entwurfsmethoden und experimentelle Versuchstechniken erforderlich.

Im ersten Teil des Buches werden deshalb zuerst die Grundlagen der Modellbildung für das statische und dynamische Verhalten beschrieben. Diese Grundlagen werden dann auf mechanische Bauelemente, elektrische Antriebe sowie sonstige mechatronische Komponenten angewandt. Im zweiten Teil werden die informationsübertragenden Komponenten mechatronischer Systeme in Form von Sensoren und Aktoren beschrieben und ihre Eigenschaften diskutiert. Es folgt ein Kapitel über den grundsätzlichen Aufbau von Mikrorechnern – speziell Mikroprozessoren und Bussystemen.

Wie der Titel sagt, stellt das Buch die elementaren Grundlagen mechatronischer Systeme und deren Integration dar. Es gibt eine Einführung in den Aufbau und die Modellbildung mechatronischer Systeme in einer einheitlichen Form und stellt das Verhalten der unterschiedlichen Bauelemente in übersichtlicher Form dar mit dem Ziel ein bestimmtes Systemverhalten zu erreichen.

Das Buch kann daher sowohl den in der Praxis stehenden Mechatroniker als auch Studenten dieses Fachgebietes als eine erste Einführung wärmstens empfohlen werden.

P. Kopacek

## **Elektrische Vorschubantriebe in der Automatisierungstechnik**

H. Groß, J. Hartmann, G. Wiegärtner

Publicis MCD Verlag, 1998,  
334 S., 89 Abbildungen, 21 Tabellen  
ISBN 3-89578-058-8

Das Buch führt umfassend in die physikalischen und technischen Grundlagen der Regelungs- und Antriebstechnik ein. Besonderes Augenmerk liegt dabei auf der Berechnung und Bemessung von elektrischen Vorschubantrieben in der Automatisierungstechnik.

Das Buch beginnt mit einer Wiederholung der elementaren regelungstechnischen Grundlagen – wie Übertragungsverhalten im Zeitbereich und im Frequenzbereich, Wurzelortskurven sowie die Verknüpfung von regelungstechnischen Grundübertragungsgliedern. Danach wird speziell auf Regelkreise bei Vorschubantrieben eingegangen. Es werden hier gängige Regelalgorithmen sowie Optimierungsverfahren und Modellbildung von Vorschubantrieben dargestellt. Eine weitere Verfeinerung erfolgt im Kapitel Lageregelung von Vorschubantrieben, wie Vorsteuerung, Führungsgrößenbeeinflussung und schwingungsfähige mechanische Systeme. Der Hauptteil des Buches beschäftigt sich mit der stationären Auslegung und Berechnung von gängigen Vorschubantrieben, wie Gewindespindel, Zahnstangen, Zahnriemen, Linearmotoren und Linearmotorantrieben. Im Anhang werden die Formelzeichen, Umrechnungstabellen sowie Einheiten zusammengestellt.

Ingenieuren und Technikern hilft es, die theoretischen Grundlagen der Regelungs- und Antriebstechnik für Maschinen aus der Werkzeugmaschinen-, der Glas- und Keramikindustrie sowie der Holzbearbeitungs- und Verpackungsmaschinenindustrie in die Praxis umzusetzen. Dies gilt ebenso für die Umsetzung von Aufgabenstellungen an Robotern oder anderen Handhabungsgeräten.

Das Buch ist didaktisch gut aufgebaut, es ist gelungen die Dimensionierung von Antrieben und nachgeschalteter Mechanik mit Auslegungsrechnungen und Beispielen zu einem praktisch anwendbaren Nachschlagewerk zu kombinieren.

Das Buch wird zukünftig zu den Standardwerken jedes praktisch tätigen Maschinenbauers, Elektrotechnikers und Regelungstechnikers gehören, der sich mit der Auslegung von Antriebssystemen beschäftigt. Insbesondere die Darstellung und Erläuterung der Zusammenhänge zwischen Regelungstechnik und Mechanik machen das Buch zu einer wertvollen Unterstützung auch für Studenten technischer Fachrichtungen.

## **Advances in Manufacturing Technology - XII**

R. W. Baines, A. Taleb-Bandiab, Z. Zhao (Hrsg.)

Professional Engineering Publishing Limited, 1998  
790 S., zahlreiche Abbildungen  
ISBN 1-86058-172-2

Es handelt sich hier um den Tagungsband der 14. nationalen englischen Konferenz über Forschung in der Produktion. Wie bei Berichtsbänden von Tagungen üblich sind die darin enthaltenen Beiträge meist sehr spezifisch gehalten.

In 12 Hauptteilen werden nahezu alle Fachgebiete der Produktionsautomatisierung behandelt. Diese 12 Kapitel sind: Produktionstechnologie, Produktionssysteme, Produktion und Lagerhaltung, Prozessplanung und -steuerung, Virtual Reality, Robotertechnik, Automatisierung und CAD, Modellbildung und Simulation, Produktionsstrategien und Management, Reengineering, Entwurfsmanagement, wissensbasierte Systeme sowie Qualitätskontrolle und Qualitätsmanagement. Wie bereits ausgeführt, handelt es sich hier um Berichte von speziellen Forschungsvorhaben.

Sucht ein Spezialist auf diesem Gebiet spezielle Artikel ist das Buch zu empfehlen.

P. Kopacek

## **Technische Informatik**

**Internet- und PC-Technologie für automatisierte Anlagen und Prozesse**

K. H. Kellermayr

Springer-Verlag Wien New York, 2000, 1. Auflage  
384 Seiten, 303 Abbildungen  
ISBN 3-211-83486-9

Das vorliegende Buch wendet sich an Studenten, Praktiker und Softwareentwickler der Fachrichtungen Mess- und Regeltechnik, Automatisierungstechnik, Industrielle Steuerungstechnik, Verfahrenstechnik, Informatik und Informationstechnik.

Ein Intel Pentium PC-System mit einem Microsoft Windows-Betriebssystem und dem graphischen Programmentwicklungssystem LabVIEW der Firma National Instruments (Austin, USA) sind die Werkzeuge, mit denen der Autor die Grundlagen der Informations- und Kommunikationssysteme analysiert. So vermittelt dieses Fachbuch praxisbezogene Kenntnisse über Aufbau und Funktionsweise moderner Computer, deren Peripherie und Schnittstellen sowie Rechnernetze, wobei auch immer die Verweise auf die entsprechenden Normen angegeben werden.

Durch seinen gut gegliederten Aufbau und seine geringen Voraussetzungen an Vorwissen, eignet sich das vorliegende Werk hervorragend als Lehr- und Lernbuch.

K. Daichendt

## Sensors Update 7

Herausgeber  
H. Baltes, W. Göpel, J. Hesse

WILEY-VCH, 2000  
444 S., 1253 Abbildungen, 27 Tabellen  
ISBN 3-527-29821-5

Diese Serie beginnt sich nicht zuletzt durch die laufend neu erscheinenden Zusatzbände zu einem Standardwerk der Sensortechnik zu entwickeln. Der vorliegende Band beschäftigt sich mit dem Spezialgebiet der Mikrowellensensoren, speziell für Feuchtemessung. Es ist in drei Teile geteilt, wobei Teil 1 die elektrischen Eigenschaften sowie ihre Anwendung zur Entwicklung von Feuchtigkeitssensoren beschreibt. Teil 2 widmet sich den unterschiedlichsten Anwendungen, sowohl im Zeitbereich als auch im Frequenzbereich. Teil 3 widmet sich schließlich einer Marktübersicht über Sensoren dieser Art, wobei in einem Anhang eine ausführliche Bibliografie von 1990 bis 1998 sowie eine Liste der Hersteller angegeben ist.

Der vorliegende Band enthält Beiträge von insgesamt 28 Fachleuten auf diesem doch sehr speziellen Gebiet. Innerhalb der vorerwähnten 3 Teile finden sich daher sowohl Übersichtsbeiträge als auch ausführliche Beiträge zu sehr speziellen Themen. Der Bogen spannt sich dabei von einer Bestimmung der Zusammensetzung und Materialeigenschaften mit Hilfe von Mikrowellensensoren über industrielle Anwendungen von Online-Sensoren bis zur Bestimmung des Wassergehaltes in porösen Medien.

Das vorliegende Buch ist nicht als Einführung in das Fachgebiet gedacht, sondern eher für Spezialisten, die spezielle Messprobleme lösen müssen. Es kann daher sowohl vorerwählter Personengruppe als auch Studenten höherer Semester, welche sich mit diesem Fachgebiet beschäftigen, zur Vertiefung ihres Wissens empfohlen werden.

P. Kopacek

## Sensors Update Volume 8

H. Baltes, W. Göpel, J. Hesse (Hrsg.)

Wiley-VCH, 2001

228 Seiten, 183 Abbildungen und 10 Tabellen

ISBN 3-527-30258-1

Die rapide Entwicklung von Sensoren, welche eine direkte Schnittstelle zu Rechnern haben, begann Mitte der 70er Jahre. Dadurch eröffnen sich für diese Sensoren komplett neue Anwendungsgebiete, insbesondere in der Automatisierungstechnik. Der nunmehr vorliegende 8. Band dieser Serie stellt zusammen mit den bisher erschienenen wohl das umfangreichste Werk über moderne Sensorik dar. Es werden neueste Entwicklungen von Materialien, Design, Produktion und Anwendung von Sensoren, Signaldetektion und Signalverarbeitung als auch neue Sensorprinzipien dargestellt. Jeder Band besteht aus drei Hauptteilen. Sensortechnologie gibt einen Überblick über die Grundlagenforschung, Sensoranwendungen beschreibt neue Anwendungen von Sensoren und der Sensormarkt gibt einen fast vollständigen Überblick über Hersteller, Marktentwicklungen und Patentinformationen für den jeweilig beschriebenen Sektor.

Der vorliegende Band ist miniaturisierten Biosensoren gewidmet. Im Teil Sensortechnologie wird ein neues Prinzip von Monolayern aus Gold für chemische Messaufgaben dargestellt. Die weiteren Beiträge beschreiben optische und elektrochemische Übertragungseinrichtungen basierend auf Polymeren, sowie eine Volumsmeßeinrichtung im Nanoliterbereich beschrieben. Das Kapitel über Anwendungen enthält drei Beiträge, nämlich über laterale und vertikale Hallmikrosysteme, multifunktionelle aktive Katheter und Wissensgewinnung aus Datensätzen unter Verwendung von hybriden Neurofuzzysystemen. Im dritten Hauptteil werden Mikrobiosensoren, welche in Zukunft sehr große Bedeutung erlangen werden, beschrieben.

Auch dieser Band wird seinem Anspruch, dem Leser ständig die neuesten Informationen aus der Sensorik zu geben, gerecht. Er ist daher auch von besonderem Wert für Wissenschaftler und Ingenieure, sowohl in der Industrie als auch an Universitäten, für alle die Sensoren entwickeln, vermarkten oder anwenden.

P. Kopacek

## **Automatisierte Industrieprozesse**

G. Zeichen, K. Fürst

Springer-Verlag, 2000

218 S., zahlreiche Tabellen und Abbildungen

ISBN 3-211-83560-1

Das zentrale Anliegen dieses Buches ist die ganzheitliche Behandlung von komplexen industriellen Prozessen in der Produkt- und Prozessentwicklung bei der Herstellung von Sachgütern und/oder Dienstleistungen. Das Buch entstand aus Vorlesungen über flexible Automation, Prozessleittechnik und Qualitätssicherung an der Fakultät Elektrotechnik und Informationstechnik der TU Wien. Die Autoren versuchen einen neuen Zugang in Form von Thesen zu finden, die aus Erfahrungen bei der Planung und Durchführung von komplexen Industrieprozessen abgeleitet wurde. Besonderer Wert liegt auf der Vernetzung von rein technischen mit organisatorischen und betriebswirtschaftlichen Prozessen, Lösungen zur Quantifizierung schwer quantifizierbarer Parameter in industriellen Prozessen sowie einer ganzheitlichen Betrachtung der Informationsprozesse innerhalb eines Betriebes und die Behandlung der dafür verfügbaren Informationswerkzeuge. Nach einer Einführung über die Ziele des Buches und methodische Engineeringprozesse werden Industrieprozesse als komplexe Systeme dargestellt. Die Autoren unterscheiden zwischen klar definierbaren technischen Prozessparametern, wie Geometrie, Druck, Temperatur, Produktivität, usw. einerseits und schwer quantifizierbaren Prozesseigenschaften, wie z.B. Design, Flexibilität, Kundennutzen, Belastbarkeit andererseits. Letztere, welche sehr großen Einfluß auf den industriellen Erfolg haben, werden in Kapitel 4 durch Analogien aus der Sensorik und Aktorik beherrschbar gemacht. In Kapitel 6 und 7 beschäftigen sich die Autoren mit dem konsequenten Einsatz der Informationstechnik zur Vernetzung der rein technischen Funktionen mit den wirtschaftlichen und sozialen Notwendigkeiten eines industriellen Unternehmens. Für diesen neuen Ansatz werden in Kapitel 6 die große informationstechnische Verantwortung der technischen Produktplanung sowie in Kapitel 7 eine diesbezügliche, in Entwicklung befindliche Decision Support Tools dargestellt.

Das vorliegende Buch stellt sicher einen neuartigen Zugang zur ganzheitlichen Betrachtungsweise von automatisierten Industrieprozessen dar. Es kann daher sowohl Technikern als auch Managern als eine erste Einführung empfohlen werden.

P. Kopacek

## NC/CNC Handbuch 2001/02

Hans B. Kief

Carl Hanser Verlag, 2001

551 Seiten, Paperback

ISBN 3-446-21756-8

Ende der 60er Jahre kamen aus der Großindustrie die Anfragen nach größeren Stückzahlen von NC-Maschinen. Die Hersteller sahen sich mit einem hohen Erklärungs- und Schulungsbedarf der Käufer konfrontiert und suchten nach besseren Lösungen. Aus dieser Situation heraus entstand 1975 die erste Auflage dieses Handbuches. Es ist ein kombiniertes Nachschlagewerk, Lehrbuch und NC-Lexikon mit einer Fülle von aktuellen Produkthinweisen namhafter Hersteller. Es bietet sowohl Planern und Einkäufern, welche nur in bestimmte Ausbaustufen investieren wollen, Ingenieuren und Facharbeitern, welche sich mit neuen Funktionen vertraut machen wollen sowie Studentenausbildern und Lehrlingen, welche sich in die Grundlagen der NC-Technik einarbeiten wollen, eine leichtverständliche, praxisorientierte Einführung. Durch die jährliche Aktualisierung ist gewährleistet, dass der neueste Stand der Technik dargestellt wird. Nach einer Einführung in die NC-Technik werden numerisch gesteuerte Werkzeugmaschinen detailliert beschrieben. Es folgt ein Kapitel über Wegmeßsysteme, elektrische Antriebe, Werkzeugsystematik sowie speicherprogrammierbare Steuerungen. Darauf wird ausführlich auf Programme und die Programmierung solcher Maschinen eingegangen, wobei als besonders erwähnenswert die Werkstatt orientierte Programmierung (WOP) hervorzuheben ist. Eine Einbindung in flexible Automatisierungssysteme wird im nächsten Kapitel dargestellt. Besonders hervorzuheben ist hier eine kurzgefasste Einführung in Handhabungsgeräte und Industrieroboter sowie in lokale Netzwerke (LANs). Im Anhang sind Richtlinien und Normen sowie ein Fachwortverzeichnis und empfohlene Literatur zusammengefasst.

Das Handbuch wird von der Industrie unterstützt und seit vielen Jahren von Industrie- und Handelskammern, Technischen Hochschulen und Berufsbildungszentren sowie in der betrieblichen Ausbildung mit Erfolg als Lehrunterlage eingesetzt. Es kann daher vorgenanntem Leserkreis als erste Einführung – aber nicht mehr – empfohlen werden.

P. Kopacek

# **Einführung in die Fertigungswirtschaft**

Wolfgang Tysiak

Carl Hanser Verlag, 2000

246 Seiten, Paperback

ISBN 3-446-21522-0

Die Fertigungswirtschaft hat sich in den letzten Jahrzehnten rasant entwickelt. In der Produktion werden überwiegend EDV-gestützte Systeme (CAD, CAP, CAM, PPS, CIM, ...) eingesetzt. Daraus ergibt sich die Notwendigkeit, organisatorisch und flexibel auf die zunehmende Komplexität der Produktionsprozesse zu reagieren. Schlagworte wie Just-In-Time, Kanban, Lean Production sind in aller Munde.

Dieses Buch entstand aus Vorlesungen des Autors für Studenten der Wirtschaftswissenschaften, International Business und Wirtschaftsingenieurwesen. Um einen Überblick über die operativen Aufgaben zu geben, die im Rahmen der Fertigungswirtschaft zu bewältigen sind, umfasst der Hauptteil des Buches die Grundzüge der operativen Fertigungswirtschaft (Programmplanung, Mengenplanung, Termin- und Kapazitätsplanung sowie Produktionssteuerung). Bei der Stoffauswahl stand grundsätzlich im Vordergrund, nicht so sehr in die Breite, sondern eher beispielhaft in die Tiefe zu gehen. So werden einzelne Fragestellungen nur kurz angedeutet, dafür aber andere sehr vertieft, um daran die Grundproblematik deutlich werden zu lassen. Das Buch enthält eine Fülle von Aufgaben sowie Hinweisen für weiterführende Literatur, wobei Wert darauf gelegt wurde, vor allem Standardlehrbücher zu zitieren. Für die enthaltenen Aufgaben werden am Ende des Buches ausführliche Lösungen angeboten, wodurch in Verbindung mit den zahlreichen Literaturhinweisen es eine ideale Möglichkeit zum Selbststudium bietet.

P. Kopacek

# MATLAB 6

## Eine Einführung

C. Überhuber, S. Katzenbeisser

Springer Verlag, 2000,  
200 Seiten, 41 Abbildungen  
ISBN 3-211-83487-7

Simulation ist neben Theorie und Experiment die dritte Säule wissenschaftlicher Forschung und technischer Entwicklung. Computer-Berechnungen sind zu einer wesentlichen Antriebskraft im Bereich der Technik und der Naturwissenschaften geworden.

MATLAB ist unter den multifunktionalen Programmsystemen der prominenteste Vertreter der numerisch orientierten Produkte. Mit diesem Programm kann man Gleichungen sehr einfach definieren und auswerten. MATLAB ist ein auf mathematisch/numerischen Methoden beruhendes Problemlösungswerkzeug, das sowohl bequeme Benutzeroberflächen bietet, als auch die individuelle Programmierung gestattet. Im Zentrum steht die interaktive numerische Lineare Algebra und Matrizenberechnung ( **MAT**rix **LAB**oratory ). Das Programm verfügt über einen Interpreter und lässt sich auch programmieren. Auf diese Weise kann man die Funktionalität des Systems noch erweitern. Dadurch können auch Fortran- und C-Programme aufgerufen werden, wodurch man die Bearbeitung rechenintensiver Aufgaben beschleunigen und bestehende Software einbinden kann.

Das Buch soll eine prägnante Einführung in MATLAB 6 unter spezieller Berücksichtigung des technisch-wissenschaftlichen Rechnens bieten. Grundlegende numerische Methoden und Techniken werden im Text und an Hand zahlreicher Beispiele vorgestellt und verständlich gemacht. Dadurch können auch erfahrene MATLAB – Benutzer eine Fülle interessanter und nützlicher Informationen finden, die ihnen bei der Lösung technisch-wissenschaftlicher Probleme gute Dienste leisten können.

Speziell für erfahrene Anwender hat sich MATLAB durch seine Erweiterungsmöglichkeit in Form von "Toolboxen" zu einem universell einsetzbaren Werkzeug auf den verschiedensten Gebieten (Simulation, Signalverarbeitung, Regelungstechnik, Fuzzy Logic, etc.) entwickelt.

H. Zebedin

## A „Tool Kit“ for Modular, Intelligent, Mobile Robots

Dr. Nayden Nedkov Chivarov

Begutachter: O.Univ.Prof.Dr.Dr.h.c.mult. P. Kopacek  
O.Univ.Prof.Dr. P.H. Osanna

Die Hersteller von konventionellen unintelligenten stationären Industrierobotern erlebten in den Jahren 1992 – 1994 dramatische Verkaufseinbußen. Statistiken aus dem Jahre 1992 prognostizierten für das Jahr 2000 ungefähr 1,2 Mio dieser Roboter weltweit. Tatsächlich waren Ende 2000 nur 760.000 Roboter im Einsatz. Die Gründe dafür lagen in einer Sättigung des Marktes für diese Art von Robotern, wofür nicht nur technische, sondern auch ökonomische und soziale Aspekte maßgebend waren. In dieser Zeit begannen die klassischen Roboterhersteller nach neuen Anwendungsgebieten für neue Roboter zu suchen. Ein Hoffungsgebiet eröffnete sich auf dem Sektorservice, d.h. Dienstleistungssektor, wie z.B. Haushalt, Restaurant, Vergnügungsindustrie, militärische Anwendungen, Medizintechnik u.ä. Unterstützt wurden diese Entwicklungen durch die hard- und softwaremäßige Verbilligung von externen Sensoren, die es gestatten den Robotern ein gewisses Maß an Intelligenz zu verleihen. Ab dem Jahre 1994 waren solche Sensoren einschließlich der Hard- und Software für die Auswertung verfügbar. Eine neue Generation – genannt Serviceroboter – entstand. Überraschend dabei ist, dass sich die Hersteller klassischer Roboter dieser Technologie bisher noch nicht anschlossen. Ein Grund dafür mag darin liegen, dass diese neue Robotergeneration infolge zu geringer Akzeptanz noch nicht die erforderlichen Stückzahlen erreicht. Laut letzten Statistiken sind derzeit weltweit 8.000 Serviceroboter im Einsatz. Prognosen für das Jahr 2003 sprechen jedoch von 60.000 Servicerobotern. Das bedeutet eine Verzehnfachung des Serviceroboterbestandes in drei Jahren.

Im Vergleich zu den 760.000 derzeit eingesetzten Industrierobotern weltweit mag die Zahl von 8.000 noch sehr gering erscheinen. Die etwa weltweit 60 Hersteller von Servicerobotern kämpfen mit geringen Absatzzahlen und hohen Entwicklungskosten für diese äußerst komplexen mechatronischen Systeme. Da diese überwiegend Einzweckroboter sind (Roboter zum Tanken, Roboter zum Waschen von Jumbojets, Roboter zur Kanalsanierung, Roboter für Operationsaufgaben, ...).

Die Zukunft für einen verbreiteten Einsatz dieser mobilen, intelligenten, kooperativen Roboter liegt in der Entwicklung von Mehrzweckrobotern. Eine Lösungsmöglichkeit besteht darin ein modulares Konzept, basierend auf einer mobilen, intelligenten Plattform, welche derzeit weltweit von ca. 60 Herstellern angeboten wird, zu entwickeln. Die vorliegende Arbeit beschäftigt sich daher mit diesem Thema. Aufbauend auf einer handelsüblichen Roboterplattform werden hard- und softwaremäßig kompatible Baugruppen entwickelt, welche zu einem Mehrzweck-Servicerobotersystem führen. Im einzelnen handelt es sich um ein Linearsystem zum Hub kleinerer Paletten, ein Montagesystem für zylindrische Werkstücke, basierend auf einer SCARA Kinematik mit 4 Freiheitsgraden, ein Roboterarm mit 5 Freiheitsgraden, beispielsweise zur Patientenversorgung in Krankenhäusern, ein Roboterarm mit 4 Freiheitsgraden, ein Fußboden- und Poliersystem für Bürsten. Es ist klar, dass dies nur ein Anfang für einen vollständigen Bausatz für mobile Serviceroboter darstellt.

Für diese ersten fünf Module werden sowohl konstruktive Vorschläge unterbreitet, als auch die softwaremäßige Integration in ein gesamtes mobiles, intelligentes Robotersystem geschaffen.

## Umweltmanagementsysteme

Dr. Csaba Darkó

Begutachter: O.Univ.Prof.Dr. H. P. Osanna  
O.Univ.Prof.Dr.Dr.h.c.mult. P. Kopacek

In den letzten Jahrzehnten hat die Umweltbelastung ständig zugenommen. Jedes Jahr fallen in den europäischen Mitgliedsstaaten beinahe 2 Milliarden Tonnen Abfälle an. Nur in Deutschland alleine sind es jährlich ca. 1,8 Mio Tonnen Elektronikschrott. Dabei sind die überwiegenden Anteile den Haushaltsgeräten (35%), der Unterhaltungselektronik (22%) und der Industrieelektronik (20%) zuzuordnen. Diese anfallende Menge wird voraussichtlich in den nächsten 10 Jahren um 5 bis 10% pro Jahr anwachsen.

In den Vereinigten Staaten werden jährlich mehr als 11 Milliarden Tonnen Kunststoff verwendet, aber nur etwa 10% davon werden wiederverwertet, obwohl die Vereinigten Staaten zu den Ländern gehören, die innovative Technologien am Umweltschutzsektor exportieren. Betrachtet man jedoch die Gesamtzahl der ISO 14000 Umweltmanagement-Zertifizierungen anderer Länder dem Bruttoinlandsprodukt angepasst, gelangen die Vereinigten Staaten zwischen den wirtschaftlich beträchtlich weniger entwickelten Ländern Ecuador und Pakistan und weit hinter Ägypten und Slowenien.

Abfall- und Entsorgungsprobleme erweisen sich hinsichtlich der Deponie, Abfallverbrennung und Entsorgung als zunehmendes Problem in den meisten Industrieländern. Die Umsetzung hat neben der ökologischen auch eine ökonomische Auswirkung; beeinflusst die Konkurrenzstrategie der betroffenen Unternehmen, stellt neue Anforderungen an die Produktgestaltung und Infrastruktur. Dabei werden bisher ungeachtete Tätigkeitsfelder wie z.B. Demontage und Recycling neu bewertet und deren zunehmende Bedeutung erkannt.

In dieser Dissertation werden die Entwicklungen und Tätigkeiten im Qualitäts- und Umweltmanagement, besonders in Hinsicht auf umweltfreundliche Produktentwicklung (Öko-Design), End-of-Life Management, Takeback und Recycling, aber auch hinsichtlich neuer Ansätze und Konzepte wie z.B. das gesamteuropäische Konzept für Elektronikschrott-Recycling und nachhaltige Produktentwicklung auf Grund der Gemeinsamkeiten und naheliegenden Ähnlichkeiten, aber auch auf Grund der Vergleichbarkeit wichtiger Merkmale und Instrumentarien in beiden Systemen behandelt.

Seitdem die ISO 9000 Serie des Qualitätsmanagements und der Qualitätssicherung Ende der 80er Jahre eingeführt wurde, haben über 350.000 Organisationen in etwa 150 Staaten Qualitätssysteme auf dieser Normreihe basierend implementiert. Im Vergleich, 3 Jahre nach der Einführung des ISO 14000 Umweltmanagementsystems in 1996 haben erst etwa über 10.000 Organisationen weltweit eine Zertifizierung nach ISO 14001 erlangt. Die neue ISO 19011 Norm wird Mitte 2001 erwartet. Sie ist das Ergebnis der Zusammenarbeit der technischen Komitees „Qualitätsmanagement und Qualitätssicherung“ und „Umweltmanagement“. Diese Richtlinien sind wichtig für die strategische Planung unserer heutigen und künftigen Tätigkeiten. Bei der Arbeit an dieser Dissertation wurden solche relevanten Tendenzen in der heutigen Industrie und Wirtschaft zur Kenntnis genommen, die erste Realisierungsschritte und aber auch die Schwierigkeiten und Grenzen der Wirtschaftlichkeit am Beispiel eines in der Informationstechnologie-Branche weltweit marktführenden Unternehmens zeigt.

# Eine virtuelle Datenbank für Elektronikbauteile

Dr. Herbert Katzlinger

Begutachter: O.Univ.Prof.Dr.Dr.h.c.mult. P. Kopacek  
O.Univ.Prof.Dr. P.H. Osanna

In den Kapiteln 1 bis 3 erfolgt eine Einführung in die Thematik, Beschreibung der funktionellen Anforderungen und die sich daraus ableitende Notwendigkeit für diese Arbeit mit den sich ergebenden Frage- und Problemstellungen, die es aus theoretischer und praktischer Sicht zu behandeln und zu lösen gibt.

Das Kap. 4 widmet sich den theoretischen, wissenschaftlichen Analysen und Aufgabenstellungen, die für die Zielsetzungen dieser Arbeit Relevanz besitzen, während im Kap. 5 die Entscheidungsgrundlagen für das gewählte Modell und die spätere Implementierung angeführt sind. Kap. 6 zeigt Programm Listings. Danach erfolgt eine Kurzzusammenfassung. Als wesentlichste Punkte betreffend Neuerungen lassen sich anführen:

- integrative Ansätze in den wissenschaftstheoretischen Modellen, was sowohl die horizontale als auch vertikale Ebene der Softwaretechnik betrifft
- Algorithmus für das Erzeugen, Darstellen und Skalieren der Grafiken
- Innovatives Modell und Implementierung, die dem Datenabgleichsystem zugrunde liegen.

Die nachfolgend kurz beschriebenen Module versuchen u.a. durch die Implementierung eines Softwaresystems, fortlaufend Infobrokersystem genannt, der Wiederverwendung bzw. dem geordneten Abbau von Elektronikschrottbauteilen Rechnung zu tragen, in dem es die dafür benötigten Informationen aufbereitet und liefert. Das System realisiert in bestimmtem Ausmaß einen virtuellen Markt für Elektronikschrottbauteile mit angeschlossener Datenhaltung. Verkäufer und Käufer können sich hier treffen und ihre geschäftlichen Aktivitäten abwickeln. Außerdem ist es möglich, Abfragemöglichkeiten nach unterschiedlichen Kriterien zu starten und Auswertungen auch in grafischer Form zu erhalten. Die Internetlauffähigkeit dieses Softwaremoduls ist als ein wesentliches Kriterium neben den Sicherheitsaspekten (Autorisierung, Verschlüsselung) anzumerken. Das System wurde mit Hilfe der Java-Technologie realisiert. Diese Software ist Teil eines größeren Systems, das aus Kamera, Beförderungseinrichtung, Steuerrechner, Lagerhaltung, Visualisierung und einer Demontagezelle mit Roboter besteht. Die Demontagezelle hat die Aufgabe, die tatsächliche Entstückung nach Vorliegen bestimmter Voraussetzungen vorzunehmen und die Lagerhaltungsdatenbank zu aktualisieren.

Ein weiterer Teil der Implementierung ist das Datenabgleichsystem. Dieses hat die Aufgabe, zu bestimmten festgelegten Zeitpunkten eine Datenaktualisierung bzw. einen Datenabgleich zwischen Lagerhaltungsdatenbank und Infobrokerdatenbank in beiden Richtungen durchzuführen. Das ist notwendig, da die beiden Datenbanken getrennt verwaltet werden.

Auch die am Infobrokersystem neu erfassten Bauteile sind dabei berücksichtigt. Da das Lagerhaltungssystem, die Steuerrechnersoftware und das Infobrokersystem nicht auf einem Computer laufen, muss der Datenabgleich über ein Netzwerksystem erfolgen.

Die gewählte Programmiersprache für das Datenabgleichsystem ist Microsoft C++. Die eigentliche Übertragung setzt auf dem Transportprotokoll TCP auf.

<i>Datum</i>	<i>Veranstaltung</i>	<i>Ort</i>	<i>Weitere Informationen erhältlich bei:</i>
5.-7.11.2001	IFAC Workshop <i>Gramma-Intelligent Assembly do and Disassembly – Brazil IAD 2001</i>		<a href="http://www.delet.ufrgs.br/iad">http://www.delet.ufrgs.br/iad</a> e-mail: <a href="mailto:cpereira@delet.ufrgs.br">cpereira@delet.ufrgs.br</a>
19.22.11.2001	IFAC Conference <i>Hong New Technologies Kong for Computer Con- trol 2001</i>		<a href="http://starship.mech.hku.hk/ntcc2001/">http://starship.mech.hku.hk/ntcc2001/</a> e-mail: <a href="mailto:mechan@hkucc.hku.hk">mechan@hkucc.hku.hk</a>
8.-10.5.2002	American Control Conference (in co- operation with IFAC)	<i>Ancho- rage, with Alaska, USA</i>	<a href="http://www.ent.ohio.edu/~acc2002">http://www.ent.ohio.edu/~acc2002</a> e-mail: <a href="mailto:rrr@gibbs.cheng.okstate.edu">rrr@gibbs.cheng.okstate.edu</a>
21.-26.7.2002	15th IFAC World Congress	<i>Barce- lona, Spain</i>	b'02 Secretariat Intern. Center for Numerical Methods in Engineering Universitat Polytechnica de Catalunya Campus Nord Gran Capita, s/n – Edificio C1 E-08034 Barcelona, Spain <a href="http://www.ifac2002.org">http://www.ifac2002.org</a> e-mail: <a href="mailto:secretariatnoc@b02.ifac2002.org">secretariatnoc@b02.ifac2002.org</a>
25.-27.9.2002	4th Asian Control Conference (in co- operation with IFAC)	<i>Singa- pore with</i>	<a href="http://www.ece.nus.edu.sg/ascc2002">http://www.ece.nus.edu.sg/ascc2002</a> e-mail: <a href="mailto:ascc2002@nus.edu.sg">ascc2002@nus.edu.sg</a>
9.-11.12.2002	2nd IFAC Conference <i>Berkeley Mechatronic Systems CA – in co-operation with IEEE/ASME</i>	<i>USA,</i>	<a href="http://www.unex.berkeley.edu/eng/mech_mechatronics2002@unx.berkeley.edu">http://www.unex.berkeley.edu/eng/mech_mechatronics2002@unx.berkeley.edu</a>
5.-7.2.2003	IMACS/IFAC Symposium <i>Mathe- matical Modelling – 4<sup>th</sup> MATHMOD</i>	<i>Vienna Austria</i>	e-mail: <a href="mailto:inge.troch@tuwien.ac.at">inge.troch@tuwien.ac.at</a>