IFAC

INTERNATIONAL FEDERATION
OF AUTOMATIC CONTROL

WARSZAWA 1969

Multivariable Systems Design

Fourth Congress of the International Federation of Automatic Control
Warszawa 16–21 June 1969

TECHNICAL SESSION
67

Organized by
Naczelna Organizacja Techniczna w Polsce

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

Multivariable Systems Design

TECHNICAL SESSION No 67

FOURTH CONGRESS OF THE INTERNATIONAL
FEDERATION OF AUTOMATIC CONTROL
WARSZAWA 16 – 21 JUNE 1969

Organized by
Naczelna Organizacja Techniczna w Polsce

Contents

Wydawnictwa Czasopism Technicznych NOT - Polska

Zakład Poligraficzny WCT NOT. Zam. 82/69.

СИНТЕЗ ЛИНЕЙНЫХ МНОГОСВЯЗНЫХ СИСТЕМ УПРАВЛЕНИЯ

М.В.Мееров ,Р.Т.Янушевский (С С С Р)
Институт автоматики и телемеханики/Т.К/ г.Москва

I. Введение

В докладе рассматривается задача синтеза многосвязных систем автоматического управления из условия минимума интегрального квадратичного функционала качества.Большое число работ,посвященных данной тематике,образует два развивающихся независимо направления.Отдельный класс составляют задачи, использующие математический аппарат теории Винера[1-3] (особо не выделяются детерминированная или стохастическая постановки оптимальных задач данного направления в виду их близости). Для упомянутых задач характерно то, что они ставят своей целью определение параметров передаточной матричной функции (которые и являются непосредственно варьируемыми) предварительно невозбужденной системы управления,определенным образом (в соответствии с критерием оптималь - ности) реагирующей на воздействия известного вида. Более общая постановка задачи характерна для работ[4-7].Приводимое в них решение не связано с решением уравнения Винера-Хопфа, составляющего основу работ первого направления;уравнения динамики системы задаются в фазовом пространстве(пространстве состояний),а непосредственно варьируемыми величинами при решении оптимальной задачи в данном случае являются координаты объекта и управляющие воздействия.

Известно,что описание динамики объекта в передаточных функциях определяет лишь соответственно управляемую и наблюдаемую часть объекта (если имеются,естественно, неуправляемая или ненаблюдаемая части),- причем входные и выходные координаты объекта-конкретные физические величины.Описание в пространстве состояний более полное,однако непосредственно координаты вектора состояния (фазовые координаты)- абстрактные величины,связанные с выходными координатами объекта (регулируемыми величинами) некоторой постоянной матрицей преобразования.Если учесть,что для большого класса многосвязных объектов качество процесса определяется обобщенным показателем являющимся функционалом выходных координат,а не составляющих вектора состояния

объекта,то очевидно,что в этом случае рассмотрение оптимальной задачи непосредственно в виде [4-7] лишено физического смысла.(Правда,выражая вектор выходных координат через вектор состояния,можно решать задачу синтеза согласно[4-7] относительно вектора фазовых координат объекта;однако обратный переход к непосредственно регулируемым величинам зачастую неэлементарен и зависит от выбора базиса пространства состояний).Существенным для задач синтеза многосвязных систем является и то,что уравнения достаточно сложных объектов определяют из эксперимента в виде весовой (передаточной)матричной функции. Поэтому при использовании схемы решения[4-7] дополнительно возникает задача нахождения динамической системы,имеющей при заданной передаточной матрице наименьший порядок дифференциального уравнения .Сказанное позволяет предположить,что практическое использование результатов[4-7] для синтеза сложных многосвязных систем ограничено и в этом плане метод[I-3] предпочтительнее.Приведенное в работе[8] его обобщение на случай ненулевых начальных условий координат системы нуждается в более строгой аргументации.

Изложенный в докладе подход непосредственно не связан с указанными выше работами и отличается более общей постановкой: предполагая уравнения динамики объекта заданными в интегральной форме относительно регулируемых величин,синтез многосвязной системы производится при произвольных условиях с учётом действующих на объект возмущающих и подлежащих воспроизведению задающих воздействий.Предлагаемая схема решения позволяет получить выражения для передаточных матричных функций оптимальной системы удобные для решения на вычислительных машинах.Упрощение вычислительной процедуры рассматривается для отдельных типов многосвязных объектов.

В заключение уместно отметить,что интегральная квадратичная форма является,как правило,косвенным показателем качества системы управления.Непосредственно интересующими инженера оценками работоспособности системы являются точность регулирования,время регулирования,перерегулирование,и т.д. Последние непосредственно не вписываются в критерий качества,но должны быть учтены,если требования качества,выражаемые показателем оптимальности,не заданы жестко ,т.е. коэффициенты по-

дынтегрального выражения функционала качества должны выбираться из условия удовлетворения вышеуказанным оценкам качества.В связи с этим в работе производится также исследование влияния коэффициентов функционала качества на свойства оптимальной системы.

2. Постановка задачи и основные соотношения.

Пусть задан линейный многосвязный объект ,динамическое состояние которого описывается следующей системой уравнений:

$$y(t) = Z(t)\,z(0) + \int_0^t W(t-\tau)\,u(\tau)\,d\tau + f(t)\,, \qquad (2.1)$$

где $y(t) = (y_1, \dots, y_m)$ -вектор выходных координат объекта;
$u(t) = (u_1, \dots, u_n)$ -вектор управляющих воздействий;
$f(t) = (f_1, \dots, f_m)$ -вектор возмущающих воздействий;
$W(t)$ -весовая матричная функция объекта;
$Z(t)$, $z(0)$ -соответственно фундаментальная матрица и вектор начальных условий,определяющие свободную составляющую решения (2.I).

Требуется в классе функций C_1 определить закон управления,обеспечивающий минимум функционала:

$$I = \frac{1}{2}\int_0^\infty \{[A(p)(y(t)-y_{эт}(t))]^*[A(p)(y(t)-y_{эт}(t))] + c^2 u^*(t)u(t)\}\,dt, \qquad (2.2)$$

где $y_{эт}(t) = (y_{эт_1}, \dots, y_{эт_m})$ -вектор задающих воздействий,
$A(p)$ -диагональная постоянная матрица-оператор;
c -постоянный коэффициент;
$(^*)$ - символ транспонирования.

Воздействия $y_{эт}(t)$, $f(t)$ полагаем исчезающими непрерывными функциями времени,элементы подынтегрального выражения-неотрицательными (предполагается,что оптимальная задача имеет смысл).

Как задача синтеза сформулированная задача состоит в построении замкнутой системы управления,воспроизводящей задающие сигналы $y_{эт}(t)$ и противодействующей возмущениям $f(t)$ с эффективностью,зависящей от коэффициентов матрицы $A(p)$ и c.

Приращение функционала(2.2) ΔI (его первая вариация), обусловленное $y(t) = y^\circ(t) + \Delta y(t)$, $u(t) = u^\circ(t) + \Delta u(t)$ ($u^\circ(t), y^\circ(t)$-оптимальные управляющие воздействия и вызванные ими изменения выходных координат; $\Delta y(t) = \int_0^t W(t-\tau)\,\Delta u(\tau)\,d\tau$), равно

$$\Delta I=\int_0^\infty\{(\int_0^t \Delta u^*(\tau)W^*(t-\tau)d\tau)A(-p)A(p)(y^\circ(t)-y_{эт}(t))+c^2\Delta u^*(t)u(t)\}dt \qquad (2.3)$$

Изменив порядок интегрирования в первом слагаемом подынтегрального выражения (2.3)

$$\Delta I=\int_0^\infty \Delta u^*(\tau)[\int_\tau^\infty W^*(t-\tau)A(-p)A(p)(y^\circ(t)-y_{эт}(t))\,dt+c^2u^\circ(\tau)]\,d\tau,$$

из условия равенства нулю первой вариации функционала получим выражение для оптимального управления [1/]

$$u^\circ(t)=\frac{1}{c^2}\int_t^\infty W^*(\tau-t)A(-p)A(p)(y_{эт}(\tau)-y^\circ(\tau))\,d\tau. \qquad (2.4)$$

Воспользуемся в дальнейшем преобразованием Лапласа, рассматривая функции $y^\circ(t)$, $u^\circ(t)$ в плоскости комплексного переменного s. Тогда выражение для оптимального управления (2.4) примет вид

$$U^\circ(s)=\frac{1}{c^2}\{W^*(-s)A(-s)A(s)(Y_{эт}(s)-Y^\circ(s))\}_+, \qquad (2.5)$$

где символ $(-s)$ означает преобразование Лапласа от функции, существующей при $t<0$; фигурными скобками со знаком плюс обозначается составляющая положительного времени соответствующего выражения. При такой записи (2.5) учитывается условие реализуемости системы управления.

Записав преобразование Лапласа для уравнения (2.I) и подставив последнее в (2.5), получим

$$\{[E+\frac{1}{c^2}W^*(-s)A(-s)A(s)W(s)]U^\circ(s)\}_+=\frac{1}{c^2}\{W^*(-s)A(-s)A(s)\cdot$$

$$\cdot(Y_{эт}(s)-F(s))\}_+-\frac{1}{c^2}\{W^*(-s)A(-s)A(s)Z(s)\}_+z(0). \qquad (2.6)$$

Матрица $E+\frac{1}{c^2}W^*(-s)A(-s)A(s)W(s)$ является эрмитовой и представима согласно [I0-I2] в виде произведения матриц $H^*(-s)H(s)$. Поэтому заменив указанным образом левую часть (2.6), умножим левую и правую части этого выражения на $H^{*-1}(s)$ [II] и выполняя простейшие преобразования, окончательно получим выражение для оптимального закона управления:

1/ Минимизация функционала

$$I=\frac{1}{2}\int_0^\infty\{[A(p)(y(t)-y_{эт}(t))]^*[A(p)(y(t)-y_{эт}(t))]+[c(p)u(t)]^*[c(p)u(t)]\}dt,$$

где $c(p)$ -операторный многочлен, легко сводится к рассматриваемой путем введения фиктивного управления $v(t)=c(p)u(t)$ и последующего исключения его из искомого результата.

$$U^{\circ}(s)=\frac{1}{c^2}H^{-1}(s)\left\{H^{*-1}(-s)W^{*}(-s)A(-s)A(s)\left[Y_{эт}(s)-F(s)-Z(s)z(0)\right]\right\}_{+} \qquad (2.7)$$

Подставив (2.7) в (2.I),имеем

$$Y^{\circ}(s)=\left(Z(s)-W(s)H^{-1}(s)\left\{\frac{1}{c^2}H^{*-1}(-s)W^{*}(-s)A(-s)A(s)Z(s)\right\}_{+}\right)z(0)+F(s)+$$

$$+W(s)H^{-1}(s)\left\{\frac{1}{c^2}H^{*-1}(-s)W^{*}(-s)A(-s)A(s)\left(Y_{эт}(s)-F(s)\right)\right\}_{+}. \qquad (2.8)$$

Из (2.7),(2.8) видно,что устойчивость системы определяется устойчивостью $H^{-1}(s)$;это осуществимо [I0-I2].Нетрудно усмотреть,что решение (2.7),(2.8) единственно несмотря на то, что разложение эрмитовой матрицы возможно с точностью до унитарной матрицы.В этом легко убедиться,подставив в (2.7),(2.8) вместо $H(s)$ матрицу $H_1(s)=QH(s)$,где Q -унитарная матрица,и используя $Q^{*}Q=E$.

Для решения задачи синтеза следует исключить вектор $z(0)$ из выражений (2.7),(2.8).После элементарных преобразований решение задачи синтеза получим в виде

$$U^{\circ}(s)=-N_1(s)\left(Y^{\circ}(s)-F(s)\right)+\left[E+N_1(s)W(s)\right]H^{-1}(s)\left\{M(-s)\left[Y_{эт}(s)-F(s)\right]\right\}_{+}, \qquad (2.9)$$

где

$$M(-s)=\frac{1}{c^2}H^{*-1}(-s)W^{*}(-s)A(-s)A(s), \qquad (2.I0)$$

$$N_1(s)=H^{-1}(s)\left\{M(-s)Z(s)\right\}_{+}\cdot\left(Z(s)-W(s)H^{-1}(s)\left\{M(-s)Z(s)\right\}_{+}\right)^{-1}. \qquad (2.II)$$

Если число управляющих воздействий равно числу регулируемых величин,выражение (2.I0) упрощается к виду

$$M(-s)=H^{*-1}(-s)-H(s). \qquad (2.I2)$$

При записи свободной составляющей в уравнениях объекта (2.I) фигурирует вектор начальных условий $z(0)$,размерность которого при решении задачи синтеза предполагается такой же, как и вектора регулируемых величин,что,вообще говоря,нуждается в пояснении.Если $y(t)$ -вектор состояния объекта,то $z(0)=y(0)$ и (2.9)-(2.I2) представляют иную схему решения аналитического конструирования по сравнению с [5,6] (при этом следует считать $y_{эт}(t)=0$) .В общем случае вектор регулируемых величин

связан с вектором состояния некоторой матрицей, и составляющие вектора $z(0)$ равны некоторым линейным комбинациям составляющих вектора состояния (зависящим от выбора базиса), а по отношению к вектору регулируемых величин-линейным комбинациям их начальных значений $y(0)$ и производных. Чтобы не оперировать с обобщенной передаточной матричной функцией, при определении $N_1(s)$ считаются отличными от нуля начальные значения лишь старших производных от регулируемых величин (что, однако, несколько не сужает постановку решаемой задачи, так как существование решения задачи синтеза при производных начальных условиях следует из разобранного случая $z(0)=y(0)$ -вектор состояния объекта). Это необходимо учитывать при записи $Z(s)\,z(0)$ в (2.II), хотя, как видно из дальнейшего, выражения (2.II) для определения $N_1(s)$ можно и не использовать.

Из выражения (2.9) видно, что оптимальный регулятор должен содержать связи по возмущающим и задающим воздействиям. Заметим, что при построении системы управления возмущающие воздействия не всегда могут быть измерены непосредственно. В этом случае для нахождения передаточных матричных функций оптимальной системы, максимально им противодействующей, следует представить $F(s)=f(s)e$, где $f(s)$ - диагональная матрица, состоящая из нулей и полюсов возмущающих воздействий, e -вектор-столбец их амплитуду, и исключить e из выражений (2.7), (2.8). Тогда при нулевых начальных условиях имеем

$$U^{\circ}(s)=-N_1^{(1)}(s)Y^{\circ}(s)+[E+N_1^{(1)}(s)W(s)]H^{-1}(s)\left\{M(-s)Y_{зг}(s)\right\}_{+}, \qquad (2.13)$$

где

$$N_1^{(1)}(s)=H^{-1}(s)\left\{M(-s)\varphi(s)\right\}_+\cdot\left[\varphi(s)-W(s)H^{-1}(s)\left\{M(-s)\varphi(s)\right\}_+\right]^{-1}. \qquad (2.14)$$

Для систем с одной регулируемой величиной в [3] рассматривалась задача синтеза линейной системы управления, водопроизводящей $Y_{эт}(s)$ (начальные условия, координат системы считаются равными нулю) при функционале качества вида (2.2). Подобный результат для многосвязных объектов легко получить, полагая в (2.7), (2.8)

$Y_{эт}(s)=\psi_{эт}(s)e$ ($\psi_{эт}(s)$ - диагональная матрица, составляемая из нулей и полюсов задающих воздействий) и исключая вектор e из указанных выражений (предварительно (2.8) записывается относительно разности $Y_{эт}(s)-Y^{\circ}(s)$)

$$U^{\circ}(s)=N_1^{(2)}(s)\left[Y_{эт}(s)-Y^{\circ}(s)\right]+N_1^{(2)}(s)F(s)- \qquad (2.15)$$

$$-\left[E+N_1^{(2)}(s)W(s)\right]\cdot H^{-1}(s)\left\{M(-s)F(s)\right\}_+,$$

где

$$N_1^{(2)}(s)=H^{-1}(s)\left\{M(-s)\psi_{эт}(s)\right\}_+\cdot\left[\psi_{эт}(s)-W(s)H^{-1}(s)\left\{M(-s)\psi_{эт}(s)\right\}_+\right]^{-1}. \qquad (2.16)$$

В случае равенства числа управляющих воздействий и регулируемых величин приведенные выражения можно упростить.

3. Процедура вычислений.

Как видно из вышеизложенного вычислительная схема рассмотреной задачи сводится к ряду алгебраических операций над матрицами-операторами в плоскости комплексного переменного s. Параметры оптимальной системы вычисляются по формулам (2.9) - (2.16). Основная вычислительная трудность состоит в разложении эрмитовой матрицы $E+\frac{1}{c^2}W^{*}(-s)A(-s)A(s)W(s)$. Процедура разложения изложена в работах [10-12]. Рассмотрим еще один способ опре-

деления матрицы $N_1(s)$ так, как, с одной стороны, задача синтеза при $Y_{эт}(s)=F(s)=0$ имеет самостоятельное значение [5,6], а с другой стороны, ее решение составляет основу для нахождения операторов связи по задающим и возмущающим воздействиям. Предлагаемый способ основан на аналогии выражения (2.7) для оптимального управления и соответствующего выражения для структурной схемы на рис.I.

$$U^{\circ}(s)=-\left[E+N_1(s)W(s)\right]^{-1}N_1(s)Z(s)z(0). \qquad (3.1)$$

Из сравнения (2.7) и (3.1) следует, что существует матрица

$$H(s)=E+N_1(s)W(s), \qquad (3.2)$$

удовлетворяющая

$$E+\frac{1}{c^2}W^{*}(-s)A(-s)A(s)W(s)=E+W(-s)N_1^{*}(-s)+N_1(s)W(s)+$$
$$+W^{*}(-s)N_1^{*}(-s)N_1(s)W(s). \qquad (3.3)$$

Задавшись видом матрицы $N_1(s)$ можно составить систему нелинейных уравнений для нахождения ее неопределенных коэффициентов. Подобным образом процедуру вычислений можно построить относительно передаточной матричной функции разомкнутой системы $G(s)=N_1(s)W(s)$.

Вычислительная схема значительно упрощается при рассмотрении отдельных классов многосвязных объектов, исходя из специфического вида их передаточных матричных функций. Иллюстрацией сказанному служит рассматриваемая ниже задача синтеза для многосвязных объектов с внутригрупповой симметрией.

4. Рассмотрение одного класса многосвязных объектов

Ограничимся классом многосвязных объектов, в которых могут быть выделены группы с идентичными параметрами и связями известных как объекты с внутригрупповой симметрией [13].

В свою очередь отдельные группы представляют собой подкласс многосвязных объектов с внутригрупповой симметрией и называются однотипными связанными объектами [14]. Первоначально ограничимся рассмотрением этого класса объектов. Упрощение решения оптимальной задачи при функционале качества (2.2) (размерность векторов $y(t)$, $y_{эт}(t)$, $f(t)$, $u(t)$ равна n, $A(\rho)$ - диагональная матрица с одинаковыми элементами $a(\rho)$ для данных объектов основано на рассмотрении двух видов движения системы [15]:

а) суммарного собственного, характеризуемого

$$y_\Sigma(s)=\sum_{i=1}^{n} y_i(s)\,,\quad u_\Sigma(s)=\sum_{i=1}^{n} u_i(s)\,,\quad y_{эт\Sigma}(s)=\sum_{i=1}^{n} y_{эт i}(s),\quad f_\Sigma(s)=\sum_{i=1}^{n} f_i(s)\,; \qquad (4.1)$$

б) суммарного относительного, характеризуемого

$$y_{i\Sigma}(s)=\sum_{j=1}^{n}\left(y_i(s)-y_j(s)\right),\qquad u_{i\Sigma}(s)=\sum_{j=1}^{n}\left(u_i(s)-u_j(s)\right),$$

$$y_{эт i\Sigma}(s)=\sum_{j=1}^{n}\left(y_{эт i}(s)-y_{эт j}(s)\right),\qquad f_{i\Sigma}(s)=\sum_{j=1}^{n}\left(f_i(s)-f_j(s)\right). \qquad (4.2)$$

При этом очевидно

$$y_i(s)=\frac{1}{n}\left(y_\Sigma(s)+y_{i\Sigma}(s)\right),\qquad u_i(s)=\frac{1}{n}\left(u_\Sigma(s)+u_{i\Sigma}(s)\right),$$

$$y_{эт i}(s)=\frac{1}{n}\left(y_{эт\Sigma}(s)+y_{эт i\Sigma}(s)\right),\qquad f_i(s)=\frac{1}{n}\left(f_\Sigma(s)+f_{i\Sigma}(s)\right)$$

$$(i=1,\dots,n) \qquad (4.3)$$

При использовании преобразования (4.1), (4.2) функционал качества (2.2) распадается на ряд функционалов

$$I=\frac{1}{n}I_\Sigma+\frac{1}{n^2}\sum_{i=1}^{n} I_{i\Sigma}\,,$$

$$I_\Sigma=\frac{1}{2}\int_0^\infty\left\{\left[a(\rho)\left(y_\Sigma(t)-y_{эт\Sigma}(t)\right)\right]^2+c^2u_\Sigma^2(t)\right\}dt, \qquad (4.4)$$

$$I_{i\Sigma}=\frac{1}{2}\int_0^\infty\left\{\left[a(\rho)\left(y_{i\Sigma}(t)-y_{эт i\Sigma}(t)\right)\right]^2+c^2u_{i\Sigma}^2(t)\right\}dt, \qquad (4.5)$$

а вместо уравнений (2.I),(2.5) имеем

$$y_{\Sigma}^{\circ}(s) = y_{cb\Sigma}(s) + W_{\Sigma}(s)U_{\Sigma}^{\circ}(s) + f_{\Sigma}(s) , \qquad (4.6)$$

$$y_{i\Sigma}^{\circ}(s) = y_{cbi\Sigma}(s) + W_{i\Sigma}(s)U_{i\Sigma}^{\circ}(s) + f_{i\Sigma}(s) , \qquad (4.7)$$

где

$$W_{\Sigma}(s) = \omega(s) + (n-1)\ell(s) , \quad W_{i\Sigma}(s) = \omega(s) - \ell(s) , \qquad (4.8)$$

$\omega(s)$ -передаточная матрица диагональных элементов объекта, $\ell(s)$ -передаточная матрица перекрестных связей; $y_{cb\Sigma}(s)$, $y_{cbi\Sigma}(s)$ -свободные составляющие переходного процесса введенных отдельных видов движения. Как показано в работе[I5], решение исходной оптимальной задачи эквивалентно минимизации функционалов (4.4),(4.5) соответственно на уравнениях (4.6),(4.7). Исходное решение получается на основе соотношений (4.3). Полученный результат распространяется на многосвязные объекты с внутригрупповой симметрией[I5]. Рассмотрение двух видов движения координат отдельных групп многосвязной системы с внутригрупповой симметрией и соответствующих видам движений управляющих воздействий позволяет перейти от исходной системы уравнений и функционала качества к эквивалентным уравнениям движения введенных обобщенных координат и соответственно преобразованному функционалу качества. Выделение в преобразованном функционале составляющих, связанных с указанными видами движения, и рассмотрение задач оптимизации отдельно для каждого вида движения тождественно решению исходной задачи. Предложенное преобразование позволяет заменить задачу оптимизации многосвязной системы эквивалентными задачами оптимизации не связанных между собой систем меньшего порядка, что су -

щественно упрощает вычислительную процедуру и может трактоваться как декомпозиция для данного класса объектов.

5. Структура и свойства оптимальной системы

Из приведенных выражений для решения задачи синтеза следует, что оптимальная система в общем случае должна быть комбинированной со связями по возмущающим и задающим воздействиям (Рис.I). Различные структурные варианты для случая невозможности непосредственного измерения возмущающих воздействий, а также при распространении подхода[I,3] (более узкого, чем рассматриваемый в работе) на многосвязные объекты представлены соответственно на рис.2 и рис.3. ($N_2(s)$, $N_3(s)$ - операторы связей по управляющим и возмущающим воздействиям). В отличие от структурной схемы на рис.I в этом случае операторы $N_1^{(1)}(s)$, $N_1^{(2)}(s)$ зависят соответственно от вида $F(s)$, $Y_{эт}(s)$. Устойчивость оптимальной системы, как это следует из данных в работе соотношений, обеспечивается устойчивостью нулей определителя $H(s)$, а потому[I0-I2] построение устойчивой системы управления возможно для широкого класса линейных многосвязных объектов. Исключение составляют объекты, содержащие сокращающиеся неустойчивые нули и полюса. Легко усмотреть, что присутствие последних в передаточной матричной функции объекта (в отличие от сокращающихся устойчивых нулей и полюсов, которые не влияют на решение задачи синтеза) обусловливает неустойчивую составляющую в решении оптимальной задачи.

При заданных $A(p)$, C уравнение

$$\left| E + \frac{1}{C^2} W^*(-s) A(-s) A(s) W(s) \right| = 0 \qquad (5.1)$$

для оптимальной системы является аналогом характеристического уравнения[16.17]. В силу того, что корни и полюса (5.1) расположены в плоскости комплексного переменного симметрично относительно начала координат, полагая $s^2=z$ можно исследовать его известными в теории автоматического регулирования методами[16,17]. При этом целью исследования является выявление влияния коэффициентов $A(\rho)$, C на расположение нулей и полюсов уравнения (5.1), т.е. в конечном счете на свойства оптимальной системы (легко получить условие вещественности корней уравнения оптимальной системы на основе[16], заданной колебательности на основе[17] и т.д.), что и составляет основу выбора коэффициентов подынтегрального выражения функционала качества при решении конкретных задач.

Существенным показателем качества системы служит точность управления, в конечном счете связанная с величинами коэффициентов усиления отдельных каналов многосвязной системы (предположение $f(t)$, $y_{эт}(t)$ исчезающими функциями времени связано с условием сходимости (2.2); однако замена неисчезающей функции исчезающей с бесконечно большим временем затухания для практических целей вполне оправдана). Предложенная выше методика расчета параметров оптимальной системы позволяет оценить величины коэффициентов усиления передаточной матричной функции системы управления, не требуя решения оптимальной задачи. Для этого достаточно положить $s=0$ в выражение (3.3). Тогда

$$\frac{1}{c^2} W^*(0) A^2(0) W(0) = W^*(0) N^*(0) + N(0) W(0) + W^*(0) N^*(0) N(0) W(0), \quad (5.2)$$

Если коэффициенты усиления диагональных элементов $N(0)W(0)$ велики (а это свойственно высококачественным системам), то величину $N(0)\,W(0)$, по которой легко оценить свойства

системы в установившемся режиме[16], приближенно можно полагать близкой к $\frac{A(0)}{C}$. Нетрудно убедиться, что коэффициенты усиления каналов многосвязной системы неограниченно возрастают с ростом величины элементов матрицы. В этом случае приходим к линейным системам с большим коэффициентом усиления[16].

6. Выводы

Рассмотрена задача синтеза многосвязных систем управления при функционале качества в виде интегральной квадратичной формы. Показано, что оптимальная система должна быть комбинированной со связями по возмущающим воздействиям, приведены выражения для определения параметров системы. Изложена схема синтеза оптимальной системы, максимально противодействующей возмущающим воздействиям, если возмущения не могут быть измерены. Установлены особенности решения задачи для многосвязных объектов с внутригрупповой симметрией. Рассмотрено влияние коэффициентов функционала на свойства оптимальной системы.

ЛИТЕРАТУРА.

I. Wiener N.L., Extrapolation, Interpolation and Smoothing of Stationary Time Series, John Willey, N.Y., 1950.

2. Катковник В.Я., Полуэктов Р.А., Многомерные дискретные системы управления, Изд-во "Наука", 1966г

3. Чанг С.Л., Синтез оптимальных систем автоматического управления, Изд-во "Машиностроение", 1964г.

4. Kalman R.E., A New Approach to Linear Filtering and Prediction Problems, Trans. ASME, J. Basic Engineering, March 1960.

5. Зубов В.И., К теории аналитического построения регулято - ров, автоматика и телемеханика, т.XXIУ, №8, 1963г.

6. Летов А.М., Аналитическое конструирование регуляторов, I-У Автоматика и телемеханика, т.XXI, №4, 5, 6, 1960г. т.XXП, 1961; XXШ, №II, 1962.

7. Красовский А.А., Интегральные оценки моментов и синтез линейных систем, Автоматика и телемехани- ка, т.XXУШ, №IO, 1967г.

8. Willis B.H., Brockett R.W., The Frequency Domain Solution of Regulator Problem, IEEE Trans., AC-10, № 3, 1965.

9. Четаев Н.Г., Устойчивость движения, Гостехиздат, 1955г.

IO. Brockett R.W., Mesarovic M.D., Synthesis of Linear Multivariable Systems, Applic. and Ind., № 62, 1962.

II. Davis M.C., Factoring the Spectral Matrix, IEEE Trans., AC-8, № 4, 1963.

I2. Youla D.C., On the Factorization of Rational Matrix, IRE Trans., IT-7, № 3, 1961.

I3. Цукерник Л.В. Устойчивость связанной системы автомати- ческого регулирования при внутригрупповой симметрии, Изд-во АН СССР, ОТН, Энергетика и автоматика, № 4, 1959г.

I4. Морозовский В.Т. К теории однотипных связанных систем автоматического регулирования с симметричными перекрестными связями, Автоматика и телемеханика, т.XXП, №3, 1961.

I5. Янушевский Р.Т., Об аналитическом конструировании одного класса многосвязных систем, Автоматика и телемеханика, т.XXУШ, №8, I967г.

I6. Мееров М.В., Синтез структур систем автоматического регулирования высокой точности, Изд-во "Наука", I967г.

I7. Фельдбаум А.А., Электрические системы автоматического регулирования, Оборонгиз, I957г.

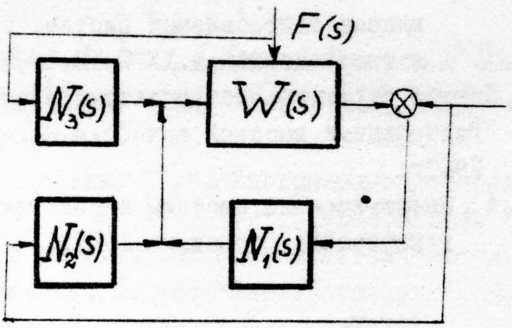


рис. 1

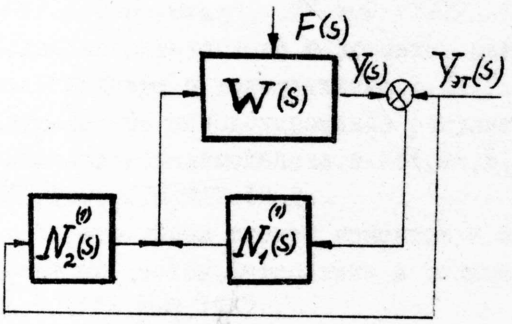


рис. 2

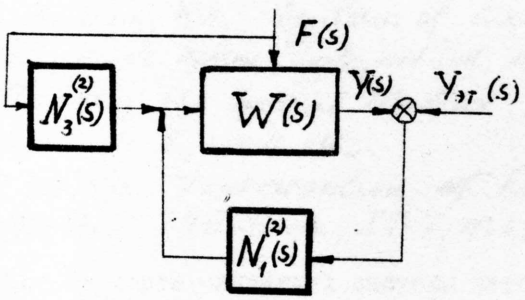


рис. 3

THE REDUCTION OF COMPLEXITY OF LINEAR, TIME INVARIANT DYNAMICAL SYSTEMS

D. Mitra
Control Systems Centre, U.M.I.S.T.,
Manchester, U.K.

1. Introduction. The systems now being considered e.g. chemical plants, national economies etc. frequently give rise to mathematical models whose dimensionality is such as to defeat the resources of current techniques in numerical methods and existing types of computer. The analyst is then compelled to seek a reduction in the complexity of the model even at the cost of lower accuracy.

The class of dynamical systems considered in this paper is assumed to be governed by a set of first order, linear, time-invariant differential equations. For this class of systems it is easy to agree on the dimension of the state space as a measure of its complexity. This paper investigates analytic methods of model simplification which seek to synthesize those simplified models of a specified degree of complexity which minimize a defined functional.

The choice of the error functional is of course crucial. Apart from being amenable to mathematical analysis, it should reflect essentially the ideal requirements of the reduced model in the context of the particular application of the model in mind. However, the problems associated with the choice of the functional is not considered here and it will suffice to state that the particular functional chosen satisfies certain basic requirements, and furthermore, experience appears to indicate its adequacy in a number of applications. A detailed investigation on the choice of the functional is contained in reference 1.

The problems associated with reduction of complexity are two-fold: the first problem involves the search for a sub-system which is relatively insensitive to input perturbations and the second concerns the degree of interaction between the sub-system and the rest of the system. In the latter case weak interaction suggests that the essential features of the rest of the system are not unduly altered in the elimination of the sub-system.

It turns out that the former problem involves a study of comparative controllability. The problem is variational and involves a defined matrix W. The latter consideration is viewed as a constraint in the reduction procedure and it is shown that this leads to projections along subspaces invariant under the system's dynamical matrix. The

solution to the problem of optimum projection along invariant subspaces is obtained for both the deterministic and stochastic formulations.

The reduction of complexity of linear, time-invariant dynamical systems has been investigated by Davison [2] who has given an intuitive method for eliminating a specified number of the fastest-decaying poles. However, Davison did not consider questions of optimality, nor, in fact, did he define a functional.

2. Mathematical Description of the System

It is assumed that the given linear, time-invariant dynamical system X is governed by the differential equations

$$X: \quad \frac{d}{dt}\underset{\sim}{x} = A\underset{\sim}{x} + B\underset{\sim}{u}(t) \qquad \underset{\sim}{m} = C\underset{\sim}{x} \tag{1}$$

where, $\underset{\sim}{x}(t)$ = n dimensional state vector,

$\underset{\sim}{u}(t)$ = p dimensional input vector,

$\underset{\sim}{m}(t)$ = q dimensional output vector

and A, B, C are constant matrices of appropriate dimension.

The system X is assumed to be controllable and observable [3].

In loose terms the problem is the synthesis of A_r, B_r, C_r which describe a similar system X_r with (n-m) state variables whose q outputs are a good approximation, in a defined sense, of those of X. m(m>o) is the order of reduction and the subscript r is used throughout to specify terms' corresponding to the system X_r.

The response at time t of the state variables of X to an unit impulse applied at time t = 0 to the i^{th} input with all other inputs unexcited is given by

$$\underset{\sim}{r}_i(t) = e^{At}\,\underset{\sim}{b}_i \tag{2}$$

where b_i is the i^{th} column of B. (The convention of denoting the i^{th} column of any matrix G by $\underset{\sim}{g}_i$ is followed consistently.) Then from the assumption of linearity and for zero initial conditions (z.i.c.),

$$\underset{\sim}{x}(t) = \int_0^t R(t-\tau)\,\underset{\sim}{u}(\tau)\,d\tau \tag{3}$$

2.1 The W matrix. The matrix W occurs frequently in the analysis and to avoid the need for frequent cross-references certain results on the matrix W are now given.

The matrix W(D,T) is the Gramian matrix of the set of n vector-valued state variable impulse response functions of the system X [4].

Consider the case when the inputs are applied one at a time and in each case

$$u_k(t) = D_k\,\delta(t) \quad \text{if } k = i$$
$$= 0 \quad \text{if } k \neq i \tag{4}$$

for $1<i<p$; $\delta(t)$ denotes the unit impulse and the elements D_k are assumed to be positive. Let $\underset{\sim}{g}_j(t)$ be a p-dimensional vector the elements of which correspond to the j^{th} state variable response to the p different weighted-impulse inputs applied in the manner indicated in equation (4). The j^{th} element of the i^{th} row of the Gramian matrix W is

$$W_{ij}(D,T) = \int_0^T \underset{\sim}{g}_i{}^t(t)\,\underset{\sim}{g}_j(t)\,dt \tag{5}$$

Here the superscript t denotes the transpose. It is easy to see that

$$W(D,T) = \int_0^T \sum_{i=1}^{p} D_i\{\underset{\sim}{r}_i(t)\,\underset{\sim}{r}_i{}^t(t)\}\,dt = \int_0^T R(t)DR^t(t)dt \tag{6}$$

where D is a diagonal matrix with D_i as its i^{th} diagonal element. The form of input weighting discussed above may be generalized by considering D in equation (6) to be symmetric and positive-definite.

Another result involving the matrix W which is required in the following analysis is the expression for the covariance matrix

$$M[\underset{\sim}{x}(t)\,\underset{\sim}{x}^t(t)]$$

for the system X subject to Gaussian disturbances with practically white spectra. Here M denotes mathematical expectation.

The correlation matrix of the input is

$$M[\underset{\sim}{u}(t)\,\underset{\sim}{u}^t(t+\tau)] = D\delta(\tau) \tag{7}$$

where D is a positive definite, symmetric matrix. Wang and Uhlenbeck [5] have shown that the state probability density function $p(\underset{\sim}{x},t)$ is Gaussian with the mean value

$$\underset{\sim}{\bar{x}}(t) = e^{At}\,\underset{\sim}{x}(o) \tag{8}$$

and the covariance matrix at time t is

$$M[(\underset{\sim}{x}-\underset{\sim}{\bar{x}})(\underset{\sim}{x}-\underset{\sim}{\bar{x}})^t] = W(D,t) \tag{9}$$

2.2 Evaluation of W It may be shown [4] that W(D,t) satisfies the matrix Riccati equation

$$\frac{d}{dt}W = AW + WA^t + BDB^t, \quad W(o) = 0 \tag{10}$$

When the system is asymptotically stable $\frac{d}{dt}W(D,T)_{/T\to\infty} = 0$ and therefore

$$AW(D,\infty) + W(D,\infty)A^t = -BDB^t \tag{11}$$

A set of n(n+1)/2 linear algebraic equations have to be solved to obtain the

matrix $W(D,\infty)$.

The form of input weighting discussed above has involved the matrix D. There is no loss of generality and some gain in terms of notational simplicity in assuming that D = I, the identity matrix, since each of the above relations remain unchanged when D is replaced by I, and B by $BD^{1/2}$. Hence only unit impulses and noise of unit variance will be considered. Furthermore, this transformation complete, the matrix W(I,t) will be denoted by W(t).

3. Formulation of the Problem

In this section the functionals for the deterministic and stochastic cases are first defined. It is next shown that the reduction process may be considered to be a two-stage process; the second stage may be implemented by the use of well-known results. Finally, the error vector is decomposed into tracking and projection components and the condition for the former to be identically zero is obtained.

3.1 The functionals. In the deterministic case the functional used is

$$E(T) = \sum_{i=1}^{p} \int_{0}^{T} \left\| C\underset{\sim}{r}_i(t) - C_r\underset{\sim}{r}_{i,r}(t) \right\|^2_{Q'} dt \qquad (12)$$

where Q' is a positive definite matrix, and the notation $\|\underset{\sim}{x}\|^2_Q$, where Q is any symmetric matrix, is used for the quadratic form $\underset{\sim}{x}^t Q \underset{\sim}{x}$. $C\underset{\sim}{r}_i(t) - C_r\underset{\sim}{r}_{i,r}(t)$ is the difference of the outputs of the systems X and X_r for an unit impulse applied at the i^{th} input and for z.i.c.

$$E_f(\infty) = \sum_{i=1}^{p} \int_{0}^{\infty} f(t) \left\| C\underset{\sim}{r}_i(t) - C_r\underset{\sim}{r}_{i,r}(t) \right\|^2_{Q'} dt \qquad (13)$$

The following functionals are used for the stochastic case

$$E_s(T) = M\left[\left\| C\underset{\sim}{x}(T) - C_r\underset{\sim}{x}_r(T) \right\|^2_{Q'} \right] \qquad (14)$$

and,

$$\int_0^T E_s(t)dt = M\left[\int_0^T \left\| C\underset{\sim}{x}(t) - C_r\underset{\sim}{x}_r(t) \right\|^2_{Q'} dt \right] \qquad (15)$$

3.2 Decomposition of the reduction process The reduction process may be represented in general as a two stage process -

Controllable System, X —Approximate Reduction→ Uncontrollable System, $\hat{X}$ —Strict Reduction→ Reduced System, X_r

The uncontrollable system $\hat{X}$ is given thus

$$\hat{X}: \quad \frac{d}{dt}\underset{\sim}{\hat{x}} = \hat{A}\underset{\sim}{\hat{x}} + \hat{B}\underset{\sim}{u}(t), \quad \underset{\sim}{\hat{m}} = C\underset{\sim}{\hat{x}} \qquad (16)$$

where $\hat{\underset{\sim}{x}}$ and $\hat{\underset{\sim}{m}}$ are respectively n and q dimensional vectors. The controllable subspace [4] of $\hat{X}$ is of dimension (n-m).

The strict reduction process is so called because the output vectors of the systems $\hat{X}$ and X_r are identically equal for z.i.c. The vectors $\underset{\sim}{m}(t)$ and $\hat{\underset{\sim}{m}}(t)$, on the other hand, are generally not equal.

The strict reduction process follows in an almost obvious manner from well known results on the decomposition of state spaces [6,7] and hence will not be considered in here. Two additional advantages of this decomposition are that (i) it allows the analysis to be carried out in the state space of the system X and (ii) since the first stage is common to a set of second stages and as the reduction problem is rarely completely formulated at the outset, there is an obvious economy of effort.

The functional E(T) defined in (12) may now be rewritten,

$$E(T) = \sum_{i=1}^{p} \int_0^T \left\| C\underset{\sim}{r}_i(t) - C\hat{\underset{\sim}{r}}_i(t) \right\|^2_{Q'} dt$$

where the superscript $\hat{}$ is used to specify terms corresponding to the system $\hat{X}$.

$$\therefore E(T) = \sum_{i=1}^{p} \int_0^T \left\| \underset{\sim}{r}_i(t) - \hat{\underset{\sim}{r}}_i(t) \right\|^2_{Q} dt \qquad (17)$$

where $Q = C^t Q' C$; Q is generally positive semi-definite. The expressions defining the functionals $E_f(\infty)$, $E_s(T)$, $\int_0^T E_s(t)dt$ in (13), (14) and (15) respectively may similarly be transformed.

3.3 Decomposition of the error vector. The decomposition of the error vector $[\underset{\sim}{x}(t) - \hat{\underset{\sim}{x}}(t)]$ into two components - a 'tracking' and a 'projection' component is considered. Let $\mathcal{f}_1$, dim $\mathcal{f}_1$ = n-m, denote the controllable subspace of the system $\hat{X}$ and let $\mathcal{f}_2$, dim $\mathcal{f}_2$ = m, be such that the direct sum of $\mathcal{f}_1$ and $\mathcal{f}_2$ is R_n, the state space, i.e. $\mathcal{f}_1 \oplus \mathcal{f}_2 = R_n$. Now consider a projection on $\mathcal{f}_1$ and along $\mathcal{f}_2$. An unique decomposition [8] exists in which

$$\left[\underset{\sim}{x}(t) - \hat{\underset{\sim}{x}}(t)\right] = \underset{\sim}{\alpha}(t) + \underset{\sim}{\beta}(t) \quad \text{where } \underset{\sim}{\alpha}(t) \in \mathcal{f}_2 \text{ and } \underset{\sim}{\beta}(t) \in \mathcal{f}_1 \qquad (18)$$

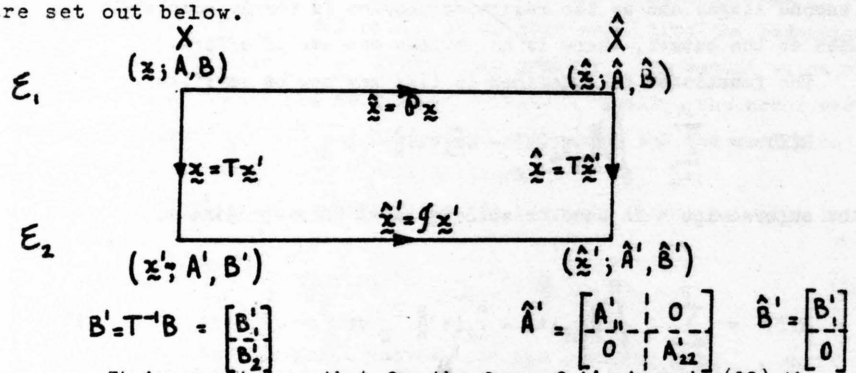
$\underset{\sim}{\alpha}(t)$ and $\underset{\sim}{\beta}(t)$ are respectively defined as the projection and tracking errors.

It is now asserted that if and only if $\mathcal{f}_2$ is a subspace invariant under the linear transformation A [8] then the tracking error is identically null for arbitrary inputs and z.i.c.

A coordinate basis $\mathcal{E}_2$ is set up in R_n such that the first (n-m) coordinate basis vectors span f_1 and the last m coordinate basis vectors span f_2. Since f_2 is invariant under the linear transformation A, the matrix A' of the linear transformation in $\mathcal{E}_2$ has the following form

$$A' = T^{-1}AT = \left[\begin{array}{c|c} A'_{11} & 0 \\ \hline A'_{21} & A'_{22} \end{array}\right] \qquad (19)$$

where the matrix T represents the operation of coordinate transformation and, therefore, the columns of T denote the new basis vectors. The operations involved in projecting along an invariant subspace are set out below.



It is easy to see that for the form of A' given in (19) the tracking error

$$\underset{\sim}{\beta}(t) = \sum_{i=1}^{n-m} (x_i' - \hat{x}_i')\underset{\sim}{t}_i \equiv 0 \qquad (20)$$

where x_i' is used to denote the i^{th} component of the vector $\underset{\sim}{x}'$. Also as A' has this particular form only if f_2 is invariant under A, the proof is complete.

It follows that when P is a projection matrix the identity $\hat{\underset{\sim}{x}}(t) = P\,\underset{\sim}{x}(t)$ for arbitrary but identical inputs to the systems X and $\hat{X}$ implies that f_2, the subspace along which the projection takes place is invariant under A.

4. Optimal Projection Along an Invariant Subspace

We shall make it an a priori requirement of the reduction method that the tracking error be identically null. From the results of the preceeding section it follows that an equivalent condition is the existence of an identity

$$\hat{\underset{\sim}{x}}(t) = \mathcal{P}\,\underset{\sim}{x}(t) \tag{21}$$

where $\mathcal{P}$ is a projection operator.

This section considers the synthesis of the optimal projection matrix $\mathcal{P}$. Since projection matrices are uniquely determined by the subspaces on and along which the projection takes place, the problem is one of finding the optimum invariant subspace, from the finite set of invariant subspaces of dimension m, along which to project and the optimum subspace on which to project.

In this section it is initially assumed that in the functional defined in (17), $Q = I_n$ and only in the following section is the general case of a positive semi-definite Q considered. Since the generalization is straightforward there is a gain in that geometrical interpretations of the analysis, otherwise obscure, are easily forthcoming.

4.1 Solution $\mathcal{S}_1$ is of dimension (n-m) and it is assumed to be spanned by the set of independent vectors $\underset{\sim}{p}_i$, $1<i<n-m$, which are not necessarily orthogonal. $\mathcal{S}_1$ is unspecified. $\mathcal{S}_2$ is an m-dimensional invariant subspace and it is initially assumed to be specified and that it is spanned by a set of m orthonormal vectors $\underset{\sim}{t}_i$, $1<i<m$. These vectors may be obtained from a set of vectors spanning $\mathcal{S}_2$ by the Gram-Schmidt orthonormalization process[8]. Also $\mathcal{S}_1 \oplus \mathcal{S}_2 = \mathcal{R}_n$, so that the vectors $\underset{\sim}{t}_i$, $1<i<m$, and $\underset{\sim}{p}_j$, $1<j<n-m$, together span $\mathcal{R}_n$.

Let $\underset{\sim}{s}_i$, $1<i<n$, be a reciprocal vector set[8] for the combined set of vectors $\underset{\sim}{t}_i$, $1<i<m$, and $\underset{\sim}{p}_j$, $1<j<n-m$. The matrices T_1, R, S_1 and S_2 are defined thus

$$\begin{aligned} T_1 &= [\underset{\sim}{t}_1 \,|\, \underset{\sim}{t}_2 \,|\, \cdots \,|\, \underset{\sim}{t}_m], \quad R = [\underset{\sim}{p}_1 \,|\, \underset{\sim}{p}_2 \,|\, \cdots \,|\, \underset{\sim}{p}_{n-m}] \\ S_1 &= [\underset{\sim}{s}_1 \,|\, \underset{\sim}{s}_2 \,|\, \cdots \,|\, \underset{\sim}{s}_m], \quad S_2 = [\underset{\sim}{s}_{m+1} \,|\, \underset{\sim}{s}_{m+2} \,|\, \cdots \,|\, \underset{\sim}{s}_n] \end{aligned} \tag{22}$$

The conditions of reciprocality thus implies the following relations involving S_1

$$T_1^{\,t} S_1 = I_m \tag{23}$$

and

$$R^{t} S_1 = 0_{n-m,\,m} \tag{24}$$

The subscripts on the identity matrix I and the null matrix 0 refer to their order. Also, it follows from the discussion in the preceeding section that the coordinate basis $\mathcal{E}_2$ is determined by the matrix T where

$$\underset{\sim}{x} = T\underset{\sim}{x}' = [R \,|\, T_1]\,\underset{\sim}{x}' \tag{25}$$

In general the error, which is wholly projection error, is

$$\underset{\sim}{x}(t) - \hat{\underset{\sim}{x}}(t) = \sum_{j=1}^{m} \langle \underset{\sim}{s}_j, \underset{\sim}{x}(t) \rangle\, \underset{\sim}{t}_j \tag{26}$$

and $$\left\| \underset{\sim}{x}(t) - \hat{\underset{\sim}{x}}(t) \right\|_E^2 = \sum_{j=1}^{m} \langle \underset{\sim}{s}_j, \underset{\sim}{x}(t) \rangle^2 = \sum_{j=1}^{m} \underset{\sim}{s}_j^t [\underset{\sim}{x}(t)\underset{\sim}{x}^t(t)] \underset{\sim}{s}_j \quad (27)$$

In particular,

$$\left\| \underset{\sim}{r}_i(t) - \hat{\underset{\sim}{r}}_i(t) \right\|_E^2 = \sum_{j=1}^{m} \underset{\sim}{s}_j^t [\underset{\sim}{r}_i(t)\underset{\sim}{r}_i^t(t)] \underset{\sim}{s}_j \quad (28)$$

Hence, from (17),

$$E(T) = \sum_{j=1}^{m} \underset{\sim}{s}_j^t \left[\sum_{i=1}^{p} \int_0^T \underset{\sim}{r}_i(t)\underset{\sim}{r}_i^t(t)dt \right] \underset{\sim}{s}_j \quad (29)$$

$$= \sum_{j=1}^{m} \underset{\sim}{s}_j^t W(T) \underset{\sim}{s}_j \quad (30)$$

In the final step the expression for W(T) given in (6) is used.

Define $$E_j(T) = \underset{\sim}{s}_j^t W(T) \underset{\sim}{s}_j \quad (31)$$

Consider the minimization of E(T), by varying $\underset{\sim}{s}_j$, $1<j<m$. This may be done by separately minimizing $E_j(T)$, $1<j<m$, since the variation of $\underset{\sim}{s}_j$ affects only E_j. For the minimization of E_j the m-dimensional column vector of Lagrange multipliers $2\underset{\sim}{\lambda}_j$ is introduced to handle the constraints on $\underset{\sim}{s}_j$ given by (23) and (24). But of course (24) does not imply a constraint on $\underset{\sim}{s}_j$ since $\mathcal{f}_1$ is spanned by the columns of R, which are as yet unspecified. Thus it is only necessary to consider

$$\min_{\underset{\sim}{s}_j} \phi_j \quad \text{where} \quad \phi_j = \underset{\sim}{s}_j^t W(T) \underset{\sim}{s}_j - 2\underset{\sim}{\lambda}_j^t (T_1^t \underset{\sim}{s}_j - \underset{\sim}{e}_j) \quad (32)$$

$\underset{\sim}{e}_j$ is the j^{th} column of the identity matrix I_m. The solution obtained in the conventional manner is

$$\underset{\sim}{s}_j = W^{-1} T [T_1^t W^{-1} T_1]^{-1} \underset{\sim}{e}_j \quad (33)$$

and, $$E_j(T) = \underset{\sim}{e}_j^t [T_1^t W^{-1} T_1]^{-1} \underset{\sim}{e}_j \quad (34)$$

By definition the columns of the matrix T_1 are independent and since the system is controllable W is positive-definite [4]; it follows that $T_1^t W^{-1} T_1$ is non-singular. From (33) and (34)

$$S_1 = W^{-1} T_1 [T_1^t W^{-1} T_1]^{-1} \quad (35)$$

and $$E(T) = \text{Trace } [T_1^t W^{-1} T_1]^{-1} \quad (36)$$

$\mathcal{f}_2$ is that subspace from the set of m-dimensional invariant subspaces which minimizes E(T). $\mathcal{f}_1$ is full determined since

$$\underset{\sim}{x} \in \mathcal{S}_1 \iff S_1^t \underset{\sim}{x} = \underset{\sim}{0} \tag{37}$$

and thus S_1 given in (35) is the solution of the problem.

4.2 The Optimal Projection Matrix The expression for $\mathcal{P}$ is obtained thus

$$\underset{\sim}{x}(t) - \hat{\underset{\sim}{x}}(t) = \sum_{i=1}^{m} \langle \underset{\sim}{x}(t), \underset{\sim}{s}_i \rangle \underset{\sim}{t}_i$$

or,
$$\hat{\underset{\sim}{x}}(t) = \underset{\sim}{x}(t) - \sum_{i=1}^{m} \langle \underset{\sim}{x}(t), \underset{\sim}{s}_i \rangle \underset{\sim}{t}_i = [I - T_1 S_1^t]\underset{\sim}{x}(t) \tag{38}$$

$$\mathcal{P} = [I - T_1 S_1^t] = [I - T_1 (T_1^t W^{-1} T_1)^{-1} T_1^t W^{-1}]$$

4.3 The reduced system Reversing the sequence of operations

$$\underset{\sim}{x} \longrightarrow \underset{\sim}{x}' \longrightarrow \hat{\underset{\sim}{x}}' \longrightarrow \hat{\underset{\sim}{x}}$$

followed in Sec.3.3, the expressions for $\hat{A}$ and $\hat{B}$ are obtained -

$$\hat{A} = T\hat{A}'T^{-1} = T\mathcal{J}A'\mathcal{J}T^{-1} = T\mathcal{J}T^{-1}AT\mathcal{J}T^{-1} = \mathcal{P}A\mathcal{P} \tag{39}$$

and
$$\hat{B} = \mathcal{P}B \tag{40}$$

It is well known [9] that for $\mathcal{P}$ projecting along an invariant subspace,

$$\mathcal{P}A\mathcal{P} = \mathcal{P}A \tag{41}$$

Hence,
$$\hat{A} = \mathcal{P}A, \quad \hat{B} = \mathcal{P}B \tag{42}$$

Now consider the roots of the characteristic equation for A and $\hat{A}$. In the former case

$$0 = \det[A - \lambda I_n] = \det[T^{-1}(A - \lambda I_n)T]$$
$$= \det[A' - \lambda I_n] = \det[A_{11}' - \lambda I_{n-m}]\det[A_{22}' - \lambda I_m] \tag{43}$$

The last relation follows from the form of A' given in (19). The characteristic equation for $\hat{A}$ (assuming $\hat{A}_{22} = 0$) is

$$0 = \det[\hat{A} - \lambda I_n] = \det[\hat{A}' - \lambda I_n]$$
$$= \det [A_{11}' - \lambda I_{n-m}]\det[- \lambda I_m] \tag{44}$$

Comparison of (43) and (44) shows that the roots of the equation $\det[A_{11}' - \lambda I_{n-m}] = 0$ are common for both equations; corresponding to these roots are the modes not spanning $\mathcal{S}_2$.

The eigenvectors of the matrix $\hat{A}$ may be obtained thus: let $\underset{\sim}{u}$ be an eigenvector of the matrix A. Then by definition

$$0 = (A - \lambda I)\underset{\sim}{u} = \mathcal{P}(A - \lambda I)\underset{\sim}{u}$$
$$= (\mathcal{P}A - \lambda I)\mathcal{P}\underset{\sim}{u} = (\mathcal{P}A\mathcal{P} - \lambda I)\mathcal{P}\underset{\sim}{u} \tag{45}$$

Since $\mathcal{P}A\mathcal{P} = \mathcal{P}A$ and $\mathcal{P}^2 = \mathcal{P}$ [9]. So,

$$\underset{\sim}{0} = (\hat{A} - \lambda I)\hat{\underset{\sim}{u}}$$

where
$$\hat{\underset{\sim}{u}} = \mathcal{P}\underset{\sim}{u} . \tag{46}$$

In terms of poles and zeros the above results imply that the n x p transfer functions of each of the systems X and $\hat{X}$ have a common set of n−m poles; in each transfer function the zeros have been manipulated to compensate for the cancellation of m poles.

5. Some Generalizations of Solution

In the preceeding section the reduction problem was solved for $Q = I_n$. But since Q is in general positive semi-definite certain modifications are necessary and these are dealt with in this section. The solution for the problem with a time-weighted norm in the functional is obtained in Section 5.2. This generalization has obvious practical significance. Process control engineers have frequently expressed the need for reduced models which are accurate in the steady state. The problem of optimal projection with zero steady state error as a constraint is considered in Section 5.3.

5.1 Positive semi-definite norms Consider Q to be positive semi-definite. In order to distinguish between vectors and matrices already defined in Sec. 4.1 and their counterparts in this section, the superscript $\bar{}$ is used to denote the latter. The independent columns of the matrix $\bar{T}_1$ which span $\mathcal{J}_2$ are here required to satisfy

$$\bar{T}_1^t Q \bar{T}_1 = \begin{bmatrix} I_{m-m_1} & 0 \\ 0 & 0_{m_1,m_1} \end{bmatrix} \tag{47}$$

where $(m-m_1)$ is the rank of the matrix $\bar{T}_1^t Q \bar{T}_1$. In analogy with (30),

$$E(T) = \sum_{j=1}^{m-m_1} \underset{\sim}{\bar{s}}_j^t W(T) \underset{\sim}{\bar{s}}_j \tag{48}$$

The variational problem of minimizing E(T) does not involve $\underset{\sim}{\bar{s}}_j$, $(m-m_1+1) \le j \le m$, and consequently they may be chosen arbitrarily as long as the reciprocality relations are satisfied. It is easily verified that a non-unique solution for the set of vectors $\underset{\sim}{\bar{s}}_i$, $1 \le i \le m$, is given by

$$\bar{S}_1 = W^{-1} \bar{T}_1 (\bar{T}_1^t W^{-1} \bar{T}_1)^{-1} \tag{49}$$

Also,

$$E(T) = \sum_{i=1}^{m-m_1} [\bar{T}_1^t W^{-1} \bar{T}_1]_{ii} \tag{50}$$

In correlating the solution contained in Sec. 4.1 and that just obtained, it is assumed that the projections in both cases take place <u>along</u> the same invariant subspace $\mathcal{J}_2$. The latter assumption implies the

existence of a m by m non-singular matrix G where

$$\bar{T}_1 = T_1 G \tag{51}$$

Then, from (49) and (35)

$$S_1 = \bar{S}_1 G^t \tag{52}$$

(52) implies that the subspaces respectively spanned by the columns of S_1 and $\bar{S}_1$ are identical. Hence,

$$\mathcal{P} = [I_n - T_1 S_1^t] = [I_n - \bar{T}_1 \bar{S}_1^t] \tag{53}$$

5.2 Time weighted norms Consider the functional $E_f(\infty)$ defined in (15) In analogy with (30) the corresponding variational problem is

$$\min_{\underset{\sim}{s}_j} \; \underset{\sim}{s}_j^t \, W_f(\infty) \, \underset{\sim}{s}_j \; ; \quad \text{constraint} : \; T_1^t \underset{\sim}{s}_j = \underset{\sim}{e}_j \tag{54}$$

$$\text{where} \quad W_f(\;) = \sum_{i=1}^{p} \int_0^{\infty} f(t) \left\{ \underset{\sim}{r}_i(t) \, \underset{\sim}{r}_i^t(t) \right\} dt \tag{55}$$

The evaluation of $W_f(\infty)$ may be expected to lead to a complicated procedure. But for a class of problems it is possible to retain its essential simplicity by the use of the Functional Matrix method developed by Macfarlane [10].

The assumptions are that (a) the system is asymptotically stable and (b) f(t) is a Laplace transformable function of time. Initially consider f(t) = 1. Then from (13),

$$AW + WA^t = -BB^t = -\sum_{i=1}^{p} \underset{\sim}{b}_i \, \underset{\sim}{b}_i^t \tag{56}$$

The above equation may also be written as a set of n(n+1)/2 linear algebraic equations

$$M \underset{\sim}{g} = -\sum_{i=1}^{p} \underset{\sim}{b}_{(2)}^{i} \tag{57}$$

where M, $\underset{\sim}{g}$ and $\underset{\sim}{b}_{(2)}^{i}$ are respectively constituted from the elements of A, W and $\underset{\sim}{b}_i$. The elements of the matrix $\underset{\sim}{g}$ are obtained from the elements of the upper triangle of the matrix W by moving, in any one row, from left to right and, row-wise, from top to bottom.

Now, Macfarlane has shown that if $f(t) \neq 1$, M^{-1} is replaced by F(M) in (57) where

$$F(M) = \int_0^{\infty} f(t) \exp(-Mt) \, dt \tag{58}$$

F(M) has simple forms for some common values of f(t). Examples are -

$f(t) = 1$, $F(M) = M^{-1}$; $f(t) = t^r$, $F(M) = M^{-(r+1)}$; $f(t) = \exp(-\alpha t)$ $[\alpha > 0]$, $F(M) = (M+\alpha I)^{-1}$

Thus for the forms of time weighting that are of greatest interest the computation of $W_f(\infty)$ still involves only one major operation - the inversion of a $n(n+1)/2$-dimensional matrix. Finally, the optimal projection for the functional $E_f(\infty)$ is obtained by replacing W by $W_f(\infty)$ in (35) and (38).

5.3 Zero steady state error For some input let the steady state value of $\underset{\sim}{x}(t)$ be denoted by $\underset{\sim}{c}$. The condition for zero steady state error may be stated thus

$$\lim_{t\to\infty} \underset{\sim}{x}(t) = \lim_{t\to\infty} \hat{\underset{\sim}{x}}(t) \qquad \text{i.e. } \underset{\sim}{c} = \rho \underset{\sim}{c} \tag{59}$$

The latter relation implies that $\underset{\sim}{c} \in \mathcal{S}_1$ or,

$$S_1^t \underset{\sim}{c} = \underset{\sim}{o} \tag{60}$$

The constraints on the vectors $\underset{\sim}{s}_i$, $1 \le i \le m$, given in (23) have to be supplemented by (60); in its complete form the constraints on $\underset{\sim}{s}_i$ may be written thus

$$T_a^t S_1 = \begin{bmatrix} I_m \\ \hline \underset{\sim}{o}^t \end{bmatrix} \qquad \text{where } T_a = [T_1 \,;\, \underset{\sim}{c}] \tag{61}$$

The respective functionals are then minimized subject to the constraint (59) if in the analysis of Sec. 4.1 T_1 is replaced by T_a.

6. The Reduction Problem in its Stochastic Formulation

The functionals $E_s(T)$ and $\int_0^T E_s(t)\,dt$ defined in (14) and (15) respectively, are considered in this section. In analogy with (12) and (17) the expressions defining the above two functionals are transformed and, again following the pattern established, it is assumed that Q is the identity matrix; the generalization to a positive semi-definite Q follows from the results of Sec. 5.1.

It is assumed that $\underset{\sim}{x}(0) = \underset{\sim}{o}$ and that the input disturbances are practically white i.e. uncorrelated, and Gaussian.

6.1 The functional $E_s(T)$ The problem is to synthesize the system $\hat{X}$, the dimension of the controllable subspace of $\hat{X}$ being n-m, such that

$$E_s(T) = M[\,\| \underset{\sim}{x}(T) - \hat{\underset{\sim}{x}}(T) \|_E^2\,]$$

is minimized for some given values of m and T. From (27),

$$E_s(T) = \sum_{j=1}^{m} \underset{\sim}{s}_j^t \, M[\underset{\sim}{x}(T)\, \underset{\sim}{x}^t(T)]\, \underset{\sim}{s}_j \tag{62}$$

Now from (11), $M[\underset{\sim}{x}(T)\,\underset{\sim}{x}^t(T)] = W(T)$

$$\therefore \quad E_s(T) = \sum_{j=1}^{m} \underset{\sim}{s}_j^t\, W(T)\, \underset{\sim}{s}_j \tag{63}$$

The problem of minimizing the above sum of quadratic forms for the constraints on $\underset{\sim}{s}_j$, $1 \leq j \leq m$, has already been considered; the solutions are given in (35) and (38).

6.2 The functional $\int_0^T E_s(t)dt$ It follows from (62) that

$$\int_0^T E_s(t)\,dt = \sum_{j=1}^{m} \underset{\sim}{s}_j^t \Big[\int_0^T W(t)\,dt\Big]\underset{\sim}{s}_j = \sum_{j=1}^{m} \underset{\sim}{s}_j^t\, V(T)\underset{\sim}{s}_j \tag{64}$$

where $V(T) = \int_0^T W(t)\,dt$. The solution to the variational problem implied by (64) is obtained by replacing W by V(T) in (35) and (38).

A special case of some importance arises when the observation period [0, T] is long compared to the time constants of the system, i.e. when for all i,j

$$T \gg [\exp(\lambda_i + \lambda_j)T - 1]/(\lambda_i + \lambda_j) \tag{65}$$

where $\{\lambda_i\}$ are the eigenvalues of the matrix A. It may then be shown that

$$V(T) \doteq TW(T) \tag{66}$$

If (65) is valid then a particularly simple solution is obtained. Replacing W in (35) by V(T) given in (66) yields

$$S_1 = (TW)^{-1} T_1[T_1^t (TW)^{-1} T_1]^{-1} = W^{-1} T_1[T_1^t W^{-1} T_1]^{-1} \tag{67}$$

Thus for T sufficiently large, the optimal projections for the functionals E(T), $E_s(T)$ and $\int_0^T E_s(t)\,dt$ are identical.

7. Conclusions In this paper the development of a fairly general method for reducing the complexity of linear, time-invariant dynamical system has been reported. In the deterministic formulation of the problem only weighted impulse inputs have been considered. One of the reasons for this is the relatively simple manner in which W may be evaluated. But it is obvious that for another class of inputs the method remains essentially unchanged; it only remains to obtain the matrix analogous to W. Also, the assumption of zero initial conditions is not essential and at the cost of a slightly more involved analysis the necessary modifications may be made.

The method reported in this paper has been applied to a ninth order linearized model of a boiler. Although of modest dimension this

model was chosen since it includes a number of features which make its reduction difficult. Details of the reduction and the results of fairly extensive tests on the reduced model are contained in reference 11, and the results are encouraging.

In our view reduced models have an important role to play in the application of optimal control theory in process control. The use of reduced models leads to a three-fold gain: the state estimator is reduced to a modest dimension, the computational problem of controller design is simplified and, finally, the design of the controller is simplified in the engineering sense.

A reduction problem which requires further effort is one in which certain system parameters are required to be recognizable in the reduced system so that their alteration does not necessitate the synthesis of a completely new reduced model. The extension of the method to more general classes of systems such as time varying linear system and distributed parameter systems is also of great interest.

References

1. Mitra, D., "The Equivalence and Reduction of Linear Dynamical Systems", Ph.D. thesis, University of London, 1967.

2. Davison, E.J., "A method for Simplifying Linear Dynamic Systems", 1966, Vol. Ac-11, No. 1, I.E.E.E. Trans.

3. Kalman, R.E., "Mathematical Description of Linear Dynamical Systems", 1963, SIAM J. Control.

4. Mitra, D., "The W Matrix and the Geometry of Model Equivalence and Reduction", to be published.

5. Wang, M.C. and Uhlenbeck, G.E., "On the Theory of Brownian Motion II", 1945, Rev.Mod.Phys., Vol. 17, Nos. 2 and 3.

6. Kalman, R.E., "Canonical Structure of Linear Dynamical Systems", 1962, Proc.Nat.Acad.Sci. U.S.A., Vol. 48, No. 4.

7. Zadeh, L. and Desoer, C.A., "Linear System Theory" (McGraw-Hill, 1963).

8. Halmos, P.R., "Finite Dimensional Vector Spaces", (Van Nostrand, 1958).

9. Pease III, M.C., "Methods of Matrix Algebra", (Academic Press, 1965).

10. Macfarlane, A.G.J., "A Method for Computing Performance Functionals for Linear Dynamical Systems:, 1963, J.Elec.Control, Vol. 15, No. 4.

11. Mitra, D., "On the Reduction of Complexity of Linear Dynamical Models - Results of Computer Studies", 1967, Report R535, Atomic Energy Establishment Winfrith, Dorchester, Dorset, U.K.

THE APPLICATION OF NON-INTERACTING CONTROL THEORY TO A CONTINUOUS MULTIVARIATE SYSTEM

By J. A. Planchard, Jr.
Louisiana State University
Baton Rouge, Louisiana
United States of America

and V. J. Law
Tulane University
New Orleans, Louisiana
United States of America

The application of modern process control has considerably enhanced the ability of engineers to design and build more sophisticated and complex processes which can be operated more economically. The wide spread utilization of on-line process computers in the near future will accelerate the progress which is being made today.

However, even with this progress, the development of new theory seems to be out-racing application. For example, most control installations that are made today, even with control computers available, are single loop systems, i.e., one output variable is fed back and controlled by a single manipulative input variable. Yet theory has been available in the literature since 1956 which presents methods for multivariable control systems design.

Introduction

This body of theory has been mostly concerned with one particular aspect of multivariable control; namely, continuous non-interacting control. The basic concept in the design of almost all non-interacting control systems involves the formulation of a diagonal matrix describing the overall relationship between input and output variables. For this reason most of the theoretical development has been limited to linear systems with time invariant parameters. In addition, none of the theory is presently capable of handling systems with appreciable time delays.

The literature now covering multivariable control is quite voluminous and continues to grow rapidly indicating the current general interest in the subject. A recent literature survey by Kavanagh[10] lists an extensive bibliography of the subject and contains over 100 references. More recently Planchard[18] has reviewed the more promising design methods for implementing non-interacting control in a process control environment. Practically all of these methods require a linear, time invariant mathematical model of the plant. These methods may be broadly classified into two groups: those employing transfer function techniques and those employing state variable methods. Of the transfer methods reviewed, those of Kavanagh[7,8,9], Freeman[2,3], Mesarovic[13,14], Chen[1], and Mathias[12] seemed to offer the most promise for the particular system under study in this investigation. The state variable method has been used by Morgan[15,16] and he has extended this technique to certain types of non-linear plants.

The purpose of this work was to experimentally apply and compare those methods presently available to an actual physical system exhibiting non-linearities and pure time delay. In this way a judgement could be made concerning the possible need for additional design methods which directly consider these complications.

The system chosen for study consisted of three stirred tanks in series to which are fed two streams of water at different, fixed temperatures. The output variables of interest are the flow rate and the temperature of the stream exiting the third tank. A schematic diagram of the system is shown in Figure 1. The system under consideration is therefore a 2 x 2 system.

The equations which describe the dynamics of the system under the assumption of perfect mixing are as follows:

$$\frac{dh_1}{dt} = \frac{1}{\pi R^2}\left[Q_h + Q_c - Q_1(h_1)\right] \tag{1}$$

$$\frac{dh_2}{dt} = \frac{1}{\pi R^2}\left[Q_1(h_1) - Q_2(h_2)\right] \tag{2}$$

$$\frac{dh_3}{dt} = \frac{1}{\pi R^2}\left[Q_2(h_2) - Q_3(h_3)\right] \tag{3}$$

$$\frac{d(h_0 + h_1)T_1}{dt} = \frac{1}{\pi R^2}\left[Q_H T_h + Q_c T_c - Q_1(h_1)T_1\right] \tag{4}$$

$$\frac{d(h_0 + h_2)T_2}{dt} = \frac{1}{\pi R^2}\left[Q_1(h_1)T_1[t-td] - Q_2(h_2)T_2\right] \tag{5}$$

$$\frac{d(h_0 + h_3)T_3}{dt} = \frac{1}{\pi R^2}\left[Q_2(h_2)T_2[t-td] - Q_3(h_3)T_3\right] \tag{6}$$

where $T_i(t-td)$ represents the temperatures of the ith tank delayed t_d seconds and $Q_i(h_i)$ represents the exit flow rate from the ith tank which is a function of the height, h_i, over the bottom of the exit pipe.

It is seen from this mathematical description that the physical plant contained product non-linearities with time delay. Thus it was felt that the system chosen for study would exercise those aspects of the current thoery which were suspect. At the same time, it would provide a simplicity which allowed adequate mathematical description and physical implementation.

In order to cast the above equations into the proper matrix form for utilization of the transfer function techniques, it was first necessary to linearize the equations and, after transforming to the Laplace domain, to eliminate intermediate variables by substitution. The additional assumption was made that the exit flow rate was the instantaneous sum of the input flow rates. The plant model thus becomes (neglecting the time delay)

$$Q_3(s) = Q_h(s) + Q_c(s) \tag{7}$$

$$T_3(s) = \frac{a_1}{(\tau s+1)^3}\,Q_c(s) + \frac{a_2}{(\tau s+1)^3}\,Q_h(s) \tag{8}$$

where $\tau = V_0/Q_{3o}$ (9)

$$a_1 = (T_c - T_{3o})/Q_{3o} \tag{10}$$

$$a_2 = (T_h - T_{3o})/Q_{3o} \tag{11}$$

This set of equations describing the flow rate and temperature from the third tank in terms of the input flow rates may now be expressed in matrix form as follows:

$$\begin{bmatrix} Q_s(s) \\ T_s(s) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \frac{a_1}{(\tau s+1)^3} & \frac{a_1}{(\tau s+1)^3} \end{bmatrix} \begin{bmatrix} Q_c(s) \\ Q_h(s) \end{bmatrix} \tag{12}$$

or

$$\underline{Y}(s) = \underline{P}(s)\,\underline{X}(s) \tag{13}$$

where $\underline{Y}(s)$ represents the vector of output variables

$\underline{X}(s)$ represents the vector of input variables

$\underline{P}(s)$ represents the transfer matrix of the plant

Almost all of the non-interacting design methods consider basically two types of control systems: those which place the non-interacting controller in the feed forward loop and those which place the controller in the feedback loop. These are shown in Figure 2.

The objective in the design of a non-interacting control system is to remove the cross coupling between elements of the input variables, $\underline{R}$, and the output variables, $\underline{C}$. Therefore the basic premise in almost all design criteria is the requirement of a diagonal matrix relating the input to output variables.

The control configuration shown in Figure 2a, with the controller located in the feedback loop, has been advocated by Kavanagh[8] and Mesarovic[13]. However, their methods differ in that Kavanagh's allows almost complete freedom in the specification of the overall plant dynamics while Mesarovic specifies what this overall response must be in order to obtain a controller matrix with the least number of non-zero elements. Thus Mesarovic's method can be thought of as a particular case of Kavanagh's. The design equation for the controller matrix of Figure 2a is given by

$$\underline{K} = \underline{P}^{-1} - \underline{H}^{-1} \tag{14}$$

where the $\underline{H}$ transmission matrix is the overall response, diagonal matrix that is specified in the design.

The majority of workers form a control configuration as shown in Figure 2b, with the controller in the feedforward loop. Freeman[3] was the first to recognize that with the controller in this position overall dynamics could be selected such that the controller does not contain elements corresponding to high order differentiation of incoming signals. The design equation for the controller of Figure 2b is

$$\underline{C} = \underline{P}^{-1}\,[\underline{I} - \underline{H}]^{-1}\,\underline{H} \tag{15}$$

Chen[1], Mathias[12], and Morgan[16] have proposed alternate methods of determining the $\underline{C}$ controller matrix that avoids much of the tedious algebraic manipulation associated with Equation (15). However, these methods naturally lead to the same controller if the overall dynamics, $\underline{H}$, are the same in all cases.

In order to avoid tedious algebraic manipulation that most of the previously mentioned methods involve for large systems, Morgan[16] employed the state variable approach to the design of non-interacting control systems. Once again only linear, time invariant parameter systems without time delay are considered. This technique utilizes a controller configuration as shown in Figure 3. (Note: The basic approach may easily be applied to other configurations.)

The plant under consideration may be represented by the matrix differential equation

$$\dot{\underline{y}} = \underline{L}\,\underline{y} + \underline{J}\,\underline{x} \tag{16}$$

$$\underline{V} = \underline{y}_1 \tag{17}$$

The class of multivariable systems considered is one with r inputs, n outputs and q state variables, where $q > n$. Therefore $\dot{\underline{y}}$ and $\underline{y}$ are q dimensional state vectors; $\underline{x}$ is an n dimensional state vector; $\underline{y}_1$ is an n dimensional state vector whose components are the first n components of $\underline{y}$; $\underline{L}$ is a constant coefficient q x q matrix; $\underline{J}$ is a constant q row, n column matrix.

The relationship between $\underline{y}$ and $\underline{r}$ is described by

$$\dot{\underline{y}} = \underline{A}\underline{y} + \underline{B}\underline{r} \tag{18}$$

where

$$\underline{A} = \underline{L} + \underline{J}\,\underline{K} \tag{19}$$

and

$$\underline{B} = \underline{J}\,\underline{C} \tag{20}$$

If $\underline{A}$ and $\underline{B}$ are partitioned the sufficient conditions for complete output-input non-interaction for the output vector of interest, $\underline{y}_1$ are

$\underline{A}_{12}$ be identically zero (21)

$\underline{A}_{11}$ be diagonal (22)

$\underline{B}_{11}$ be diagonal (23)

The specification of the $\underline{A}_{11}$ and $\underline{B}_{11}$ diagonal matrixes is equivalent to specifying the overall dynamics of the system. The elements of the controller matrixes, $\underline{C}$ and $\underline{K}$ may be determined from the following equations:

$$\underline{K}_{12} = -(\underline{J}_{11})^{-1}\,\underline{L}_{12} \tag{24}$$

$$\underline{K}_{11} = (\underline{J}_{11})^{-1}\,(\underline{A}_{11} - \underline{L}_{11}) \tag{25}$$

$$\underline{C}_{11} = (\underline{J}_{11})^{-1}\,\underline{B}_{11} \tag{26}$$

Computer Simulation

In order to obtain a preliminary evaluation of the various controllers that were designed by the techniques presented in the previous two sections, and in order to investigate certain parameters that were not readily varied experimentall, a program of digital simulation was undertaken. The first design method investigated was that of Kavanagh. If the overall dynamics are selected as

$$\underline{H} = \begin{bmatrix} 1 & 0 \\ 0 & \dfrac{a_2 - a_1}{(\tau s+1)^3} \end{bmatrix} \tag{27}$$

then substitution into Equation 14 results in the controller matrix:

$$\underline{K} = \begin{bmatrix} \frac{a_1}{a_2-a_1} & \frac{-(\tau s+1)^3}{a_2-a_1} \\ -\frac{a_1}{a_2-a_1} & 0 \end{bmatrix} \tag{28}$$

Note that the k_{12} elements contains a cubic term in the Laplace domain variable s. Physically this requires that up to the third derivative of the output temperature, T_2, must be measured as produced in order to implement the proper control action. Since this is not physically practical, this controller is described as physically unrealizable. While the controller cannot be synthesized in the physical sense, it can be simulated mathematically. The time domain differential equations describing the control loop of Figure 2a (using the linearized model) were programmed using the Mimic[17] digital simulation language. The tabulated numerical results of the following simulations may be found in the original work.[18]

Since no valve dynamics were taken into account (i.e., the controller is capable of instantaneously changing flow rates in the model), the Kavanagh controller immediately implemented the final input flow rates. Because the controller design was based upon linearized plant equations, and since these equations were used in the simulation, perfect non-interaction resulted.

However, it should be recalled that the Kavanagh controller contained a physically unrealizable element. The suggested practice to eliminate this is to place arbitrary poles, far removed from the zeroes, in the element in order to make it physically realizable.

For the conditions used in these simulations, that is, a tank volume of 4.58 gallons and an initial flow rate of 14.0 gpm, the time constant of the system, τ, was .326 minutes. This corresponds to having a third order pole at s = -3.07. Therefore, the three arbitrary poles were set at s = -36, more than a factor of 10 removed from the zeroes in the k_{12} element of Equation (28).

For a step change in the r_2 element of the command vector, $\underline{r}$, the output temperature should have changed 3°F while the output flow rate should have remained constant. However, the effect of adding three arbitrary poles to the controller was to introduce a dynamic interaction between r_2 and Q_3. After five minutes, the exit flow rate returned to the proper value. Therefore, the observed effect of making the controller physically realizable was the loss of absolute non-interaction.

Next the physically realizable controller of Equation (29) was used in conjunction with no time delay. Comparison of this simulation wiht that for the linearized system indicated that the synamic interaction was only slightly higher for the non-linear system. Thus the technique of linearizing the product type non-linearities in order to design the controller had an almost negligible effect over that observed with the linear plant for the step sizes considered. The effect of time delay on the system performance will be discussed later.

The next control system studied was that which may be design either by the methods of Freeman or Mathias. The control configuration is given by Figure 2b, and the final controller design by Equation (15).

If the overall transfer function matrix, $\underline{H}$, is selected as:

$$\underline{H} = \begin{bmatrix} h_1 & 0 \\ 0 & \dfrac{Z}{(\tau s+1)^3} \end{bmatrix} \tag{29}$$

where h_1 and Z represent constant coefficients, then the solution of Equation (15) results in

$$\underline{C} = \begin{bmatrix} \dfrac{a_2 h_1}{(a_2-a_1)(1-h_1)} & \dfrac{-Z}{(a_2-a_1)} \times \dfrac{(\tau s+1)^3}{[(\tau s+1)^3-Z]} \\ \dfrac{-a_1 h_1}{(a_2-a_1)(1-h_1)} & \dfrac{Z}{(a_2-a_1)} \times \dfrac{(\tau s+1)^3}{[(\tau s+1)^3-Z]} \end{bmatrix} \tag{30}$$

This controller was tested with both the linear and non-linear plant models again neglecting the time delay. In both cases perfect non-interaction resulted. This was achieved despite the fact that the input flow rates, as set by the controller, differed significantly in the two cases at given instants of time. Thus perfect non-interaction was achieved with this controller for the non-linear plant without time delay when incorporated with the controller designed from the linear plant model.

The final non-interacting control system that was considered for simulation purposes was that designed by Morgan's method. The linearized set of equations describing the plant must be cast into the form of Equation (31) in order that $\underline{J}_{11}$ be non-singular.

$$\begin{bmatrix} 0 \\ \dot{T}_1 \\ \dot{T}_3 \\ \dot{T}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1/\tau & 0 & 0 \\ 0 & 0 & -1/\tau & 1/\tau \\ 0 & 1/\tau & 0 & -1/\tau \end{bmatrix} \begin{bmatrix} Q_3 \\ T_1 \\ T_3 \\ T_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ \dfrac{a_1}{\tau} & \dfrac{a_2}{\tau} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q_c \\ Q_H \end{bmatrix} \tag{31}$$

Now the controllers may be designed once selection of the $\underline{A}_{11}$ and $\underline{B}_{11}$ diagonal matrices has been made. The $\underline{A}_{11}$ and $\underline{B}_{11}$ matrices were selected as

$$\underline{A}_{11} = \begin{bmatrix} -1 & 0 \\ 0 & -1/\tau \end{bmatrix}, \quad \underline{B}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & -a_2/\tau \end{bmatrix} \tag{32}$$

Substitution into Equations (24), (25) and (26) results in

$$\underline{K}_{11} = 0$$

$$\underline{K}_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \underline{C}_{11} = \begin{bmatrix} \dfrac{a_2}{a_2-a_1} & \dfrac{a_2}{a_2-a_1} \\ \dfrac{-a_1}{a_2-a_1} & \dfrac{a_2}{a_2-a_1} \end{bmatrix} \tag{33}$$

It is interesting to note that no feedback control is required to decouple the system. That is, the $\underline{K}$ matrix is null. This result is logical in that the system contains no feedback interaction. It is important that only decoupling is achieved and no regulation is provided in this case. Regulation would have to be provided by feedback control on each of the outputs by the corresponding input. An excellent discussion of this topic has recently been presented by Greenfield and Ward[4,5]. Since these controllers are constant co-

efficient matrices, they represent a particularly simple mode of control.

Because of the excellent results which had been previously obtained with the non-linear system, no simulation of the Morgan controller was made with the linear plant. As in the case of the Freemen control system, perfect non-interaction was obtained when this control system was incorporated with the non-linear plant.

Thus far none of the simulations have included time delay in the mathematical model. It will be recalled that this time delay occurs because of transportation lags between the first and second tanks and between the second and third tanks. For the experimental system under consideration, with a total flow rate of 14 gpm, the time delay was approximately 0.01 minutes.

Possibly the most unexpected result of this study occurred when time delay was considered for the Freeman-Mathias controller and the Morgan controller. It was found that pure time delay of up to one minute had no effect at all on the non-interaction. Thus perfect non-interaction was retained although some effect was noted in the response of the dynamic output variable. This occurance is discussed in detail by Planchard.[18]

Time delay did have a significant effect in the case of the Kavanagh-Mesarovic controller where it caused a lag in the return of the exit flow rate to its initial value. In addition, in the case of a time lag of one minute, it appeared that the system might become unstable although the program was terminated before this could definitely be established.

Thus these results indicate that because of the way in which the time delay appears in the plant model, it causes no loss of non-interaction in either the Freeman-Mathias or Morgan controllers. However, in the case of the Kavanagh-Mesarovic controller, time delay causes a more sluggish response and may lead to instability for large time delays. For the time delay of the experimental system (approximately 0.01 minutes) no significant loss in system performance was observed.

Finally, a simulation of a double loop feedback control system in which proportional and integral mode controllers were utilized was conducted. The controller parameters were determined by the Ziegler-Nichols method[19] utilizing open loop response data obtained from the plant model as given by Equations (1) through (6), neglecting time delay. The control system is shown in Figure 4. The system was found to be unstable for the Ziegler-Nichols settings. Much more conservative controller settings resulted in stable operation but the controlled response was poor.

These results demonstrated the consequences that may result when empirical design methods, based upon single loop systems, are used in the design of highly coupled multivariable systems. Even if much more conservative controller settings are used to provide system stability, the control action will probably be of poor quality. It should also be recalled that no time delay was considered in the plant model in the above simulation. A time delay would almost certainly have caused additional stability problems.

Experimental Results

A schematic flow diagram of the physical system is shown in Figure 1. Two Electronics Associates TR-48 10 volt, analog computers were used in this study to implement the non-interacting controller and other electronic operations. The two input flow rates were measured by turbine flow meters; the

associated measuring instrument converted this signal to a 0 to 10 volt output. This signal could then be directly incorporated into the analog circuitry. The output voltage of the analog computer was converted into a 3-15 psig pneumatic signal by appropriate external tranducers.

Experimental results with the Freeman-Mathias controller are presented in Figures 5 and 6. In the first figure, the exit temperature was the invariant variable while the exit flow rate changed by a negative 1.0 gpm. A maximum deviation in the exit temperature of 0.43°F occurred after approximately one minute of operation and thereafter the exit temperature returned to its initial valve.

Similar results were observed for runs in which the exit temperature was changed from one level to another. In the data presented in Figure 6, while the system underwent a negative change in exit temperature of 3.8°F, the exit flow rate had a maximum deviation of 0.17 gpm from its initial value. Considerable cycling about the initial flow rate occurred thereafter. This was principally due to random disturbances in the cold water flow rate caused by line pressure fluxuations.

The Kavanagh control system was investigated in a similar manner. In each case the three arbitrary poles in the Kavanagh controller were set equal to each other. Results obtained with each pole having values of 1/15 sec^{-1} are presented in Figure 6 and 7. These values correspond to transfer functions of

$$\frac{(s+3.03\ min^{-1})^3}{(s+4.0\ min^{-1})^3} \quad \text{and} \quad \frac{(s+3.03\ min^{-1})^3}{(s+6.0\ min^{-1})^3}$$

respectively. Therefore in each case the arbitrary poles were less than an order of magnitude removed from the zeroes.

The results presented in Figures 7 and 8 are for step changes in the r_2 element of the command vector. Therefore the flow rate should remain invariant while the exit temperature undergoes a change in level. In both cases results similar to the analytical study were obtained. A rather large dynamic interaction was found which reached a maximum in less than one minute and then returned to its initial value.

Direct comparison of the amount of non-interaction between individual controllers could not be made simply by comparing the deviations from the initial value of the variable since in each case the total change in the system was slightly different. Therefore it is desirable to define a non-interacting ratio which takes this variance into account. This non-interacting ratio (NIR) may be defined as:

$$NIR = \frac{\text{maximum deviation from initial value}}{\text{absolute sum of the changes in input flow rates}} \qquad (44)$$

Results for the three cases presented thus far for step changes in r_2 indicated that much better non-interaction was obtained with the Freeman-Mathias controller. In addition, moving the arbitrary poles out to 6.0 $minute^{-1}$ significantly improved the response of the Kavanagh-Mesarovic control system.

An attempt was made to move the arbitrary poles in the Kavanagh-Mesarovic controller out to 12.0 $minute^{-1}$, however the circuit was too noisy and could not be used.

Runs were also made with this controller in which step changes in the r_1 element were introduced. Results essentially identical to those with the Freeman-mathias controller, given in Figure 6, were obtained.

Conclusions

A large volume of literature has resulted from the considerable attention that multivariable control has received in the past few years. Much of this work has been directed toward the development of newer methods for designing non-interacting control systems, while little attention has been directed toward physical implementation or comparison of these systems.

This study has compared those design methods which seemed most applicable to the particular plant under consideration. Herein, of course, lies the greatest danger in a study of this nature. Conclusions drawn from one particular plant will not always be valid when another plant, with completely different characteristics, is under consideration. However, taking this into account, it is felt that the following broad conclusions are still valid:

1. This study clearly demonstrated that reasonably good non-interacting control may be physically implemented with a minimum of difficulty resulting from inexact mathematical description of the plant.

2. Since all the conclusions which results from the digital simulation study were born out by the experimental implementation of the various control systems, future simulation studies in this area may be undertaken with a great deal more confidence of subsequent successful application in the laboratory or plant.

3. Control systems in which the primary control is place in the forward loop (shown in Figure 2b) are superior in a number of ways to those design methods in which the controller is placed in the feedback loop (shown in Figure 2a). In addition to better performance these advantages include the possibility of cancellation of the effect of time delay and, in the case of transfer function design techniques, the assurance that the controller may be made physically realizable without resorting to arbitrary poles.

4. Product type non-linearities can easily be accommodated by the non-interacting control systems considered in this study.

5. Based upon the simulation studies and the results with constant gain controllers in the experimental work, it is felt that state variable techniques such as Morgan's design method offer two distinct advantages over transfer function techniques. These are simpler controllers and less tedious design effort. However, all state variables must in general be measurable and occasionally incomplete non-interaction may be the only form attainable.

6. Controllers designed for non-interaction will generally provide much better control than conventional feedback control. However, The non-interacting control system is much more complex, and therefore more expensive, than conventional control systems.

This study treated one type of linearization--that by Taylor's Series expansion. Another widely used linearization technique is that of developing approximate transfer functions based upon such techniques as pulse testing[6]

or other types of transient tests. The quality of non-interacting control obtained under this type of linearization would form a logical extension of this study.

Nomenclature

$\underline{A}$	matrix of constant coefficients
$\underline{A}_{ij}$	Submatrix of $\underline{A}$
a_1	Parameter of linearized plant model = $T_c - T_{so}/Q_{3o}$
a_2	Parameter of linearized plant model = $T_h - T_{3o}/Q_{3o}$
$\underline{B}$	Matrix of constant coefficients
$\underline{C}$	Controller Matrix
c_{ij}	The i,j element of the $\underline{C}$ matrix
$\underline{d}$	Vector of uncontrolled disturbances
$\underline{H}$	A diagonal matrix of transfer functions specifying overall response
h_i	The i,i element of the diagonal matrix $\underline{H}$
$\underline{I}$	The identity matrix
$\underline{J}$	Matrix of constant coefficients
$\underline{K}$	Controller matrix
K_{I1}	Integral mode constant of exit flow rate loop
K_{I2}	Integral mode constant of exit temperature loop
K_{P1}	Proportional mode constant of exit flow rate loop
K_{P2}	Proportional mode constant of exit temperature loop
$\underline{L}$	Matrix of constant coefficients
m	The number of output variables of a system
n	The number of input variables to a system
$\underline{P}$	A matrix of transfer functions representing plant dynamics
Q_c	Cold water flow rate
Q_{co}	Initial cold water flow rate
Q_h	Hot water flow rate
Q_{ho}	Initial hot water flow rate
Q_i	Exit flow rate of the ith tank
Q_{if}	Final exit flow rate of the ith tank
Q_{ic}	Initial exit flow rate of the ith tank
q	The number of state variables in a system
R	Radius of tanks
$\underline{r}$	Vector of forcing functions
s	The Laplace transform variable
T_c	Cold water temperature
T_h	Hot water temperature
T_i	Temperature of ith tank
T_{if}	Final temperature of the ith tank
T_{io}	Initial temperature of the ith tank
t	Time
v_o	Initial volume of each tank
$\underline{x}$	Vector of input variables to a plant
$\underline{y}$	Vector of output variables
Z	Arbitrary gain constant

Greek Letters

π	3.14149 ...
Σ	Summation operator
τ	Time constant = V_o/Q_{3o}

Bibliography

1. Chen, K., R. A. Mathias, and D. M. Sauter, Trans. A.I.E.E. (App & Ind), 81, 336 (1962).

2. Freeman, H., Trans. A.I.E.E. (App & Ind), 76, 28 (1957).

3. Freeman, H., Trans. A.I.E.E. (App & Ind), 77, 1 (1958).

4. Greenfield, G. and T. Ward, IEC Fund, 60, 564 (1967).

5. Greenfield, G., and T. Ward, IEC Fund, 60, 571 (1967).

6. Hougen, J. O., Section in Progress in Control Engineering, Vol. 3, Edited by R. H. Macmillan, T. J. Higgins, and P. Naslin, Iliffee, Ltd., London (1966).

7. Kavanagh, R. J., J. Franklin Inst., 262, 349 (1956).

8. Kavanagh, R. J., Trans. A.I.E.E. (App & Ind), 76, 95 (1957).

9. Kavanagh, R. J., Trans. A.I.E.E. (App & Ind), 77, 425 (1958).

10. Kavanagh, R. J., Section in Progress in Control Engineering, Vol. 3, Edited by R. H. Macmillan, T. J. Higgins, and P. Naslin, Iliffee, Ltd., London (1966).

11. Mason, S. J., Proc. I.R.E., Vol 41, No. 9, 1144 (1953).

12. Mathias, R. A., Trans. I.E.E.E. on Automatic, AC-8, 283 (1963).

13. Mesarovic, M. D., "The Control of Multivariable Systems," 1-23, John Wiley and Sons, Inc., New York (1960).

14. Mesarovic, M. D., J. Franklin Inst., 266, 274 (1960).

15. Morgan, B. S., Ph.D. Dissertation, University of Michigan, Ann Arbor, Michigan (1963).

16. Morgan, B. S., Trans. I.E.E.E. on Automatic Control, AC-9, 405 (1964).

17. Petersen, H. E., F. J. Sonsom, and L. M. Warshawsky, SESCA Internal Memo 65-12, Wright-Patterson Air-Force Base, Ohio (May, 1964).

18. Planchard, J. A., Ph.D. Dissertation, Tulane University, New Orleans, Louisiana (1967).

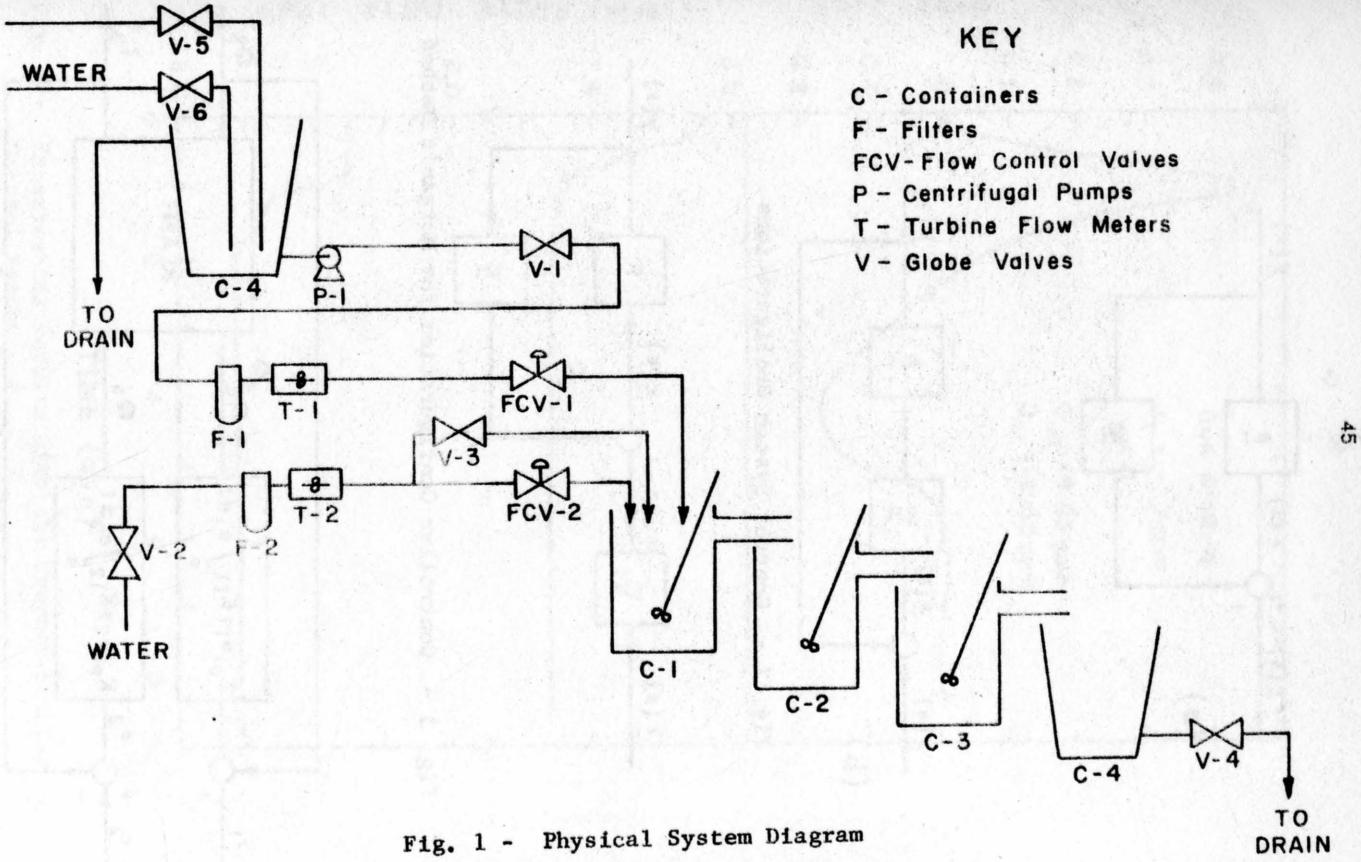


Fig. 1 - Physical System Diagram

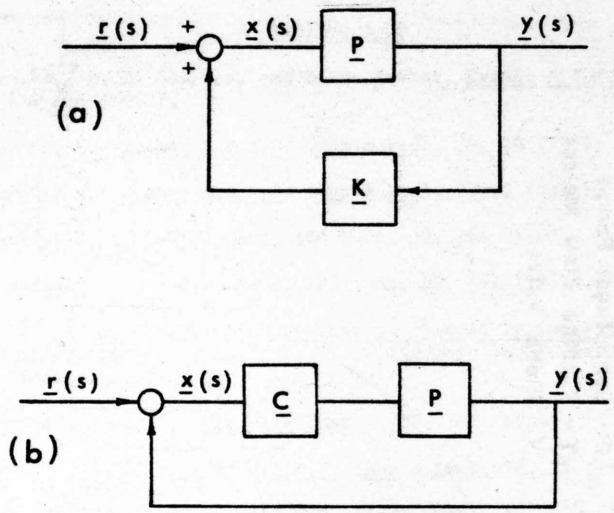


Fig. 2 - Control System Configurations

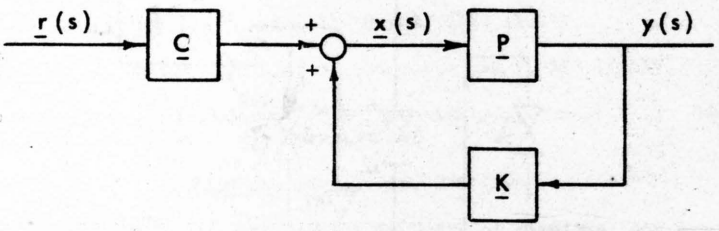


Fig. 3 - Controller Configuration for Morgan's Method

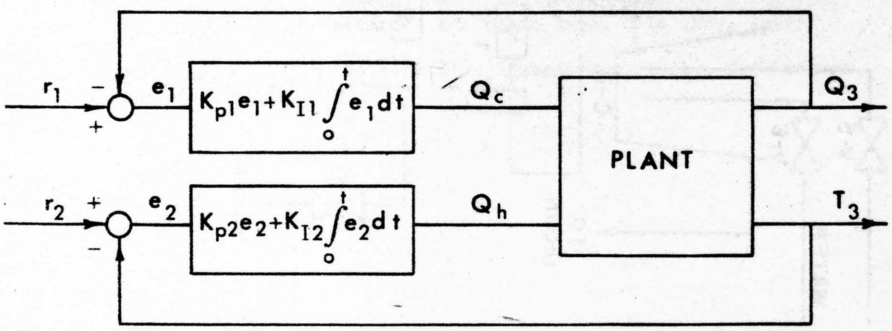


Fig. 4 - Double Loop Feedback Control System

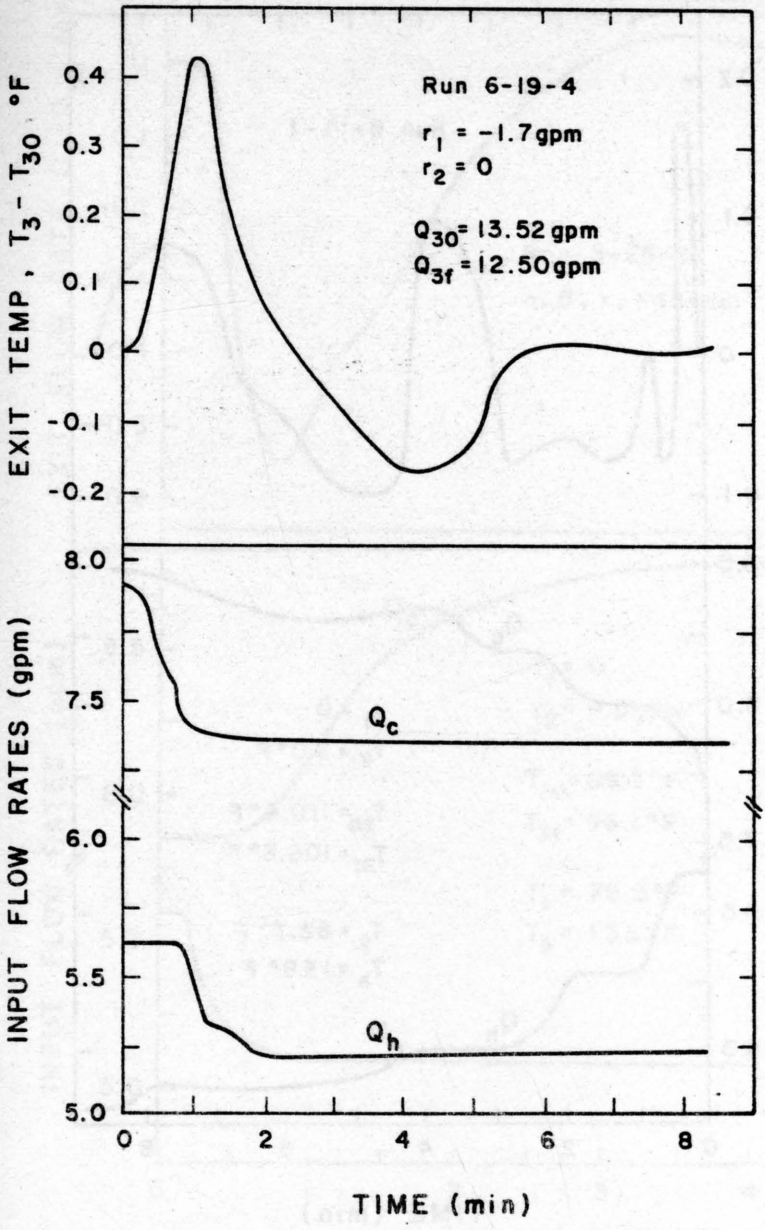


Fig. 5 - Experimental Results with the Freeman-Mathias Control System - I

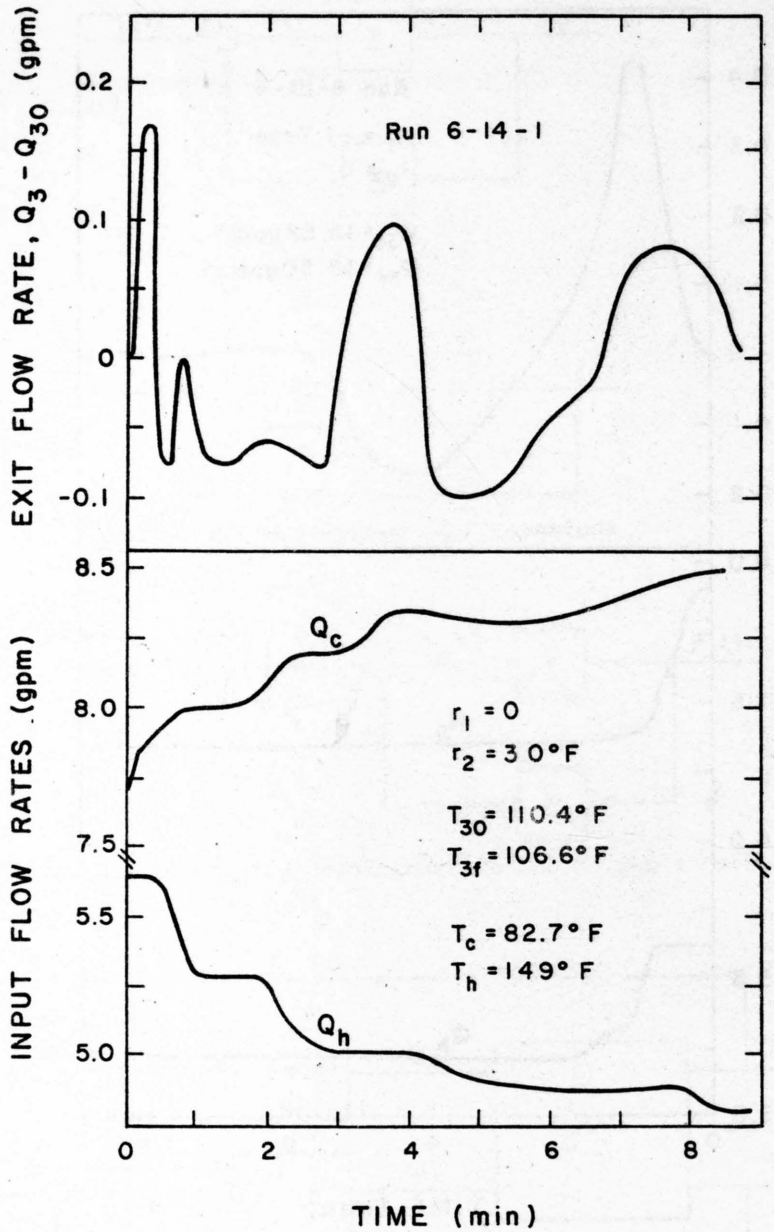


Fig. 6 - Experimental Results with the Freeman- Mathias Control System - II

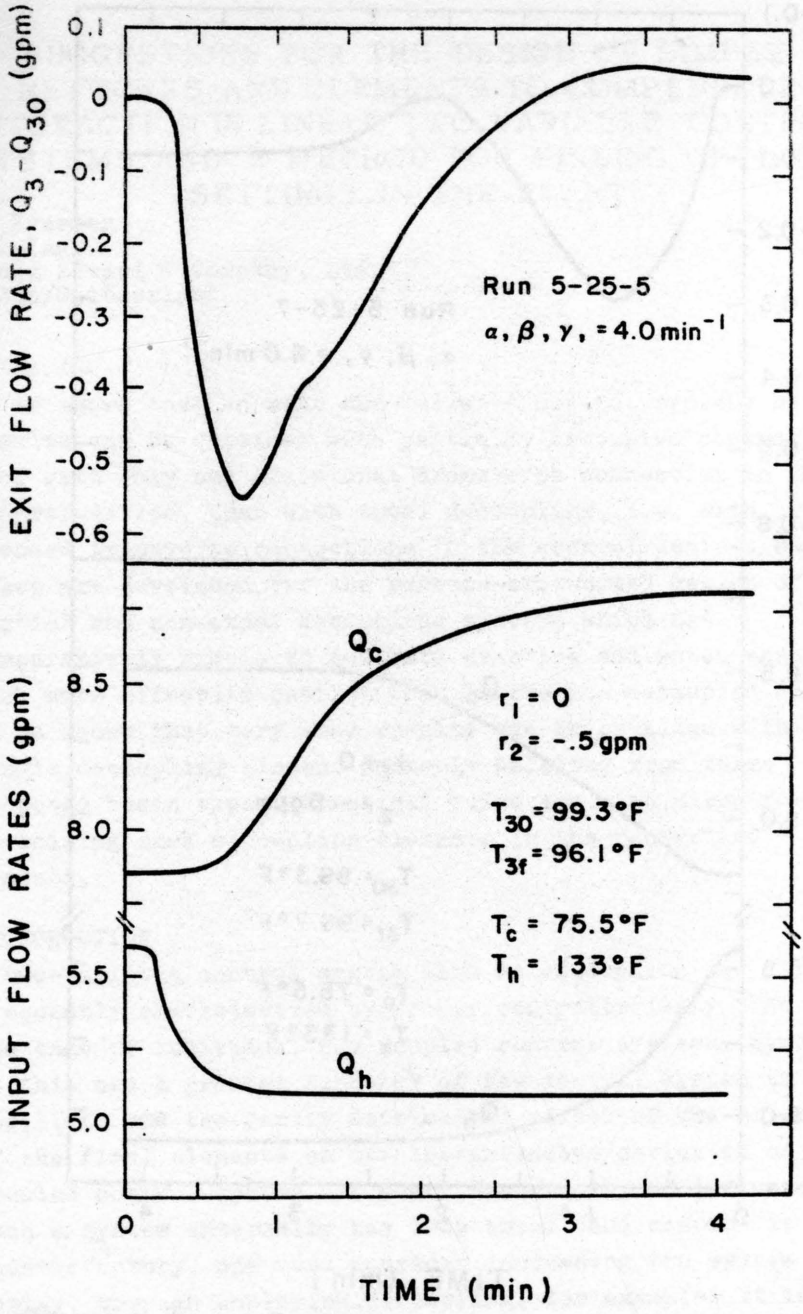


Fig. 7 - Experimental Results with the Kavanagh-Mesarovic Control System - I

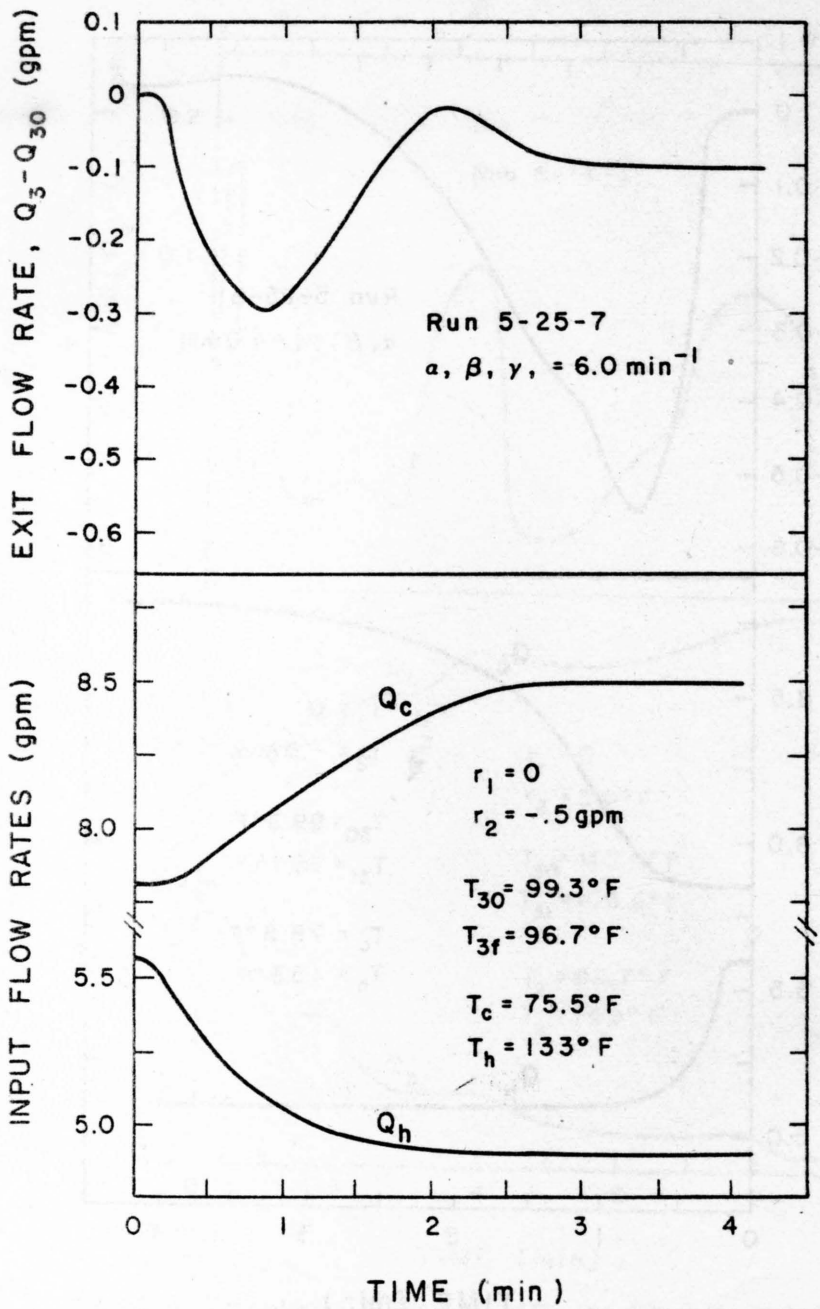


Fig. 8 - Experimental Results with the Kavanagh-Mesarovic Control System - II

SUGGESTIONS FOR THE DESIGN OF SIMPLE NETWORKS AND ELEMENTS TO COMPENSATE INTERACTION IN LINEAR TWO-VARIABLE CONTROL SYSTEMS AND A METHOD FOR FINDING OPTIMUM SETTINGS IN THE PLANT

W. Kraemer
Dr.-Ing.
Brown Boveri & Company, Ltd.
Baden/Switzerland

SUMMARY

It is shown that in most two-variable control systems better results can be obtained with partially decoupled control, i.e. with only one additional transverse connection in the control device, than with total decoupling, i.e. with two crossed transverse connections in the control device. Basic rules are developed for the purpose-orientated design of partial and non-exact decoupling systems which are comparatively simple to put into practice and which ensure much more effective control than in the non-decoupled case. It is shown that very good results can be obtained with a single decoupling element suitably selected from three proposed basic types. Practical rules are also given for optimizing such decoupling elements in the controlled process.

INTRODUCTION

A two-variable control system with no decoupling is frequently characterized by poorer controllability than in the case of individual not coupled control systems. Symptoms of this are a greater tendency of the control system to oscillate and the partly detrimental effect of the movements of the final elements on the instantaneous deviation of the coupled control system. If every opportunity of influencing such a system externally has been tried[1] and control is still unsatisfactory, one must consider increasing the system outlay, through employing decoupling, for example. It is in the interests of both operator and supplier of a plant with two-variable control that decoupling can be achieved as simply as possible. Guidelines and proposals for this are

treated theoretically below, together with practical rules for obtaining optimum decoupling.

THE NON-DECOUPLED CASE

To illustrate the true function of decoupling in a two-variable control system, let us first analyse mathematically the actually "harmful" features of non-decoupled systems. Fig.1 shows a block diagram which is generally valid for non-decoupled two-variable control systems.

The P-canonical signal flow structure within the process block in Fig.1 effectively covers all conceivable structures, as they can all easily be traced back to this straightforward form. Variables x, y and z are to be considered as Laplace-transformed deviations from an arbitrary steady-state, while the terms R, S an Z are transfer functions of the controller or the process, in which case the delays of the measuring elements and final elements should be added to the process. The signs for the coupling and disturbance signals given in brackets in Fig.1 cannot be generally defined unequivocally and are taken to be positive in the following discussion.

The disturbance transfer functions of a complete system which are usually of the greatest interest, i.e. between a disturbance z and the two control variables x_1, and x_2, can be written in a particularly clear manner, as follows:

$$\frac{x_1}{z} = \frac{Z_1}{1+R_1S_{11}} \left[\frac{1 - F'_{w2} \frac{Z_2}{Z_1} \frac{S_{12}}{S_{22}}}{1 - F'_{w1}F'_{w2} \frac{S_{12}}{S_{11}} \frac{S_{21}}{S_{22}}} \right] \qquad (1)$$

$$\frac{x_2}{z} = \frac{Z_2}{1+R_2S_{22}} \left[\frac{1 - F'_{w1} \frac{Z_1}{Z_2} \frac{S_{21}}{S_{11}}}{1 - F'_{w1}F'_{w2} \frac{S_{12}}{S_{11}} \frac{S_{21}}{S_{22}}} \right] \qquad (2)$$

where $F'_{w1} = \frac{R_1 S_{11}}{1 + R_1 S_{11}}$ = Reference transfer function of loop 1, considered as not coupled

$F'_{w2} = \frac{R_2 S_{22}}{1 + R_2 S_{22}}$ = Reference transfer function of loop 2, considered as not coupled

It can then be said that everything which distinguishes a two-variable system from a single-loop system, as regards their reactions to disturbances, must be contained mathematically implicitly in the square brackets of Eq. (1) and (2), because the terms before the brackets can be identified as the known disturbance transfer functions of individual loops. Examination of these terms in brackets thus enables us to analyse the characteristic features of two-variable control systems.

Two important points must be taken into account.

- The common denominator in the square brackets of Eq. (1) and (2), when put equal to zero, presents a particular way of writing the characteristic equation of the system. It is formed principally by the dimensionless expression

$$\mathcal{K} = \frac{S_{12}\, S_{21}}{S_{11}\, S_{22}}\,, \qquad (3)$$

  termed the "interaction quotient"[2]. In particular, the sign and modulus of interaction quotient $\mathcal{K}$, considered as a frequency characteristic and calculated at the dominant frequency ω of the control operations, can be shown to be responsible for the altered stability conditions or the changed limits and possibilities of the controller settings.[1]

- The modulus of the complete term in brackets in (1) and (2), considered as the frequency characteristic, is a measure of the extent to which a coupled control system, apart from the changed stability conditions referred to above, is advantageous or detrimental for the control quality of the individual control systems. The reaction of

the whole system to disturbances is certainly comparatively less favourable than in the non-coupled case if the modulus of the term in brackets is greater than unity, and vice versa.

When seen in this manner, the terms F'_{w1} and F'_{w2} play a relatively minor part since, with assumed negativ feedback in the individual circuits, they are always positive and their modulus are almost constant, at least in the lower frequency range. Thus, depending on the signs of the pairs of coupling and disturbance signals occurring within the process, it is possible to detect certain trends for the harmfulness or usefulness of a coupled second control loop. This can be shown clearly in a logic diagram, Fig.2. The practical consequences which can be deduced from this analysis are summarized in Rules 1 to 3 at the end of this article.

THE CASE OF PARTIAL DECOUPLING

It was found in the previous section that in a large proportion of two-variable control systems in the engineering field, namely in all negatively coupled systems, so-called "partial decoupling", i.e. with only one additional transverse connection in the control device, is perfectly adequate, and in fact better as regards quality of control than "total decoupling", i.e. with two crossed transverse connections. This also complies with the wish mentioned in the introduction for the simplest possible arrangement of the decoupling system.

In the case of partial decoupling, four different arrangements are possible if, for example, the decoupling signal is fed to loop 2. The first two possibilities a) and b) in Fig.3 have the disadvantage, from the technical point of view, that the transfer of each controller has to be compensated, which is inconvenient and, if controller parameters are modified, can be very complicated. The second and third possibilities, b) and c) in Fig.3, have the overwhelming drawback that the decoupling signal falsifies the desired value of one controller, either temporarily or permanently. The fourth possibility d) thus presents the best

technical solution, since it does not have the disadvantages mentioned (see Rule 5).

THE CASE OF NON-EXACT, PARTIAL DECOUPLING

The effect of partial decoupling on negatively coupled two-variable control systems can best be illustrated by disturbance transfer functions calculated for a specific example. The system in question can be characterized by a block diagram as shown in Fig.4.

Exact compensation of the sole harmful coupling signal (cf. Rule 4) was achieved theoretically by a sign-reversing decoupling element with the transfer function

$$X = \frac{S_{21}}{S_{22}} . \tag{4}$$

In practice this condition cannot be fulfilled exactly, or only at great expense. This is evident when one considers that S_{21} and S_{22} also contain the dynamic influences of the final elements. Generally one has to be satisfied with "non-exact decoupling" which compensates harmful coupling transfer via S_{21} only approximately.

If we now calculate the disturbance transfer functions for a non-exactly, partialy decoupled system, we obtain the relationships

$$\frac{x_1}{z} = \frac{Z_1}{1+R_1S_{11}(1+X\frac{S_{12}}{S_{11}})} \left[\frac{1 - F'_{w2}\frac{Z_2}{Z_1}\frac{S_{12}}{S_{22}}}{1 + \hat{F}'_{w1}F'_{w2}\mathcal{K}Y} \right] \tag{5}$$

$$\frac{x_2}{z} = -\frac{Z_2}{1+R_2S_{22}} \left[\frac{1 + F'_{w1}\frac{Z_1}{Z_2}\frac{S_{21}}{S_{11}}Y}{1 + \hat{F}'_{w1}F'_{w2}\mathcal{K}Y} \right] , \tag{6}$$

where

$$\hat{F}'_{w1} = \frac{R_1S_{11}(1 + X\frac{S_{12}}{S_{11}})}{1 + R_1S_{11}(1 + X\frac{S_{12}}{S_{11}})}$$

$$\text{and} \qquad Y = \frac{1 - X\,\frac{S_{22}}{S_{21}}}{1 + X\,\frac{S_{12}}{S_{11}}} = \frac{1 - X\,\frac{S_{22}}{S_{21}}}{1 + X\,\frac{S_{22}}{S_{21}}\,\mathcal{K}} \tag{7}$$

The non-decoupled case can be reproduced from relationships (5) and (6) by putting X equal to zero.

$$\frac{x_1}{z} = \frac{Z_1}{1+R_1S_{11}} \left[\frac{1 - F'_{w2}\,\frac{Z_2}{Z_1}\,\frac{S_{12}}{S_{22}}}{1 + F'_{w1}\,F'_{w2}\,\mathcal{K}} \right] \tag{8}$$

$$\frac{x_2}{z} = -\frac{Z_2}{1+R_2S_{22}} \left[\frac{1 + F'_{w1}\,\frac{Z_1}{Z_2}\,\frac{S_{21}}{S_{11}}}{1 + F'_{w1}\,F'_{w2}\,\mathcal{K}} \right] \tag{9}$$

The exactly partially decoupled limiting case is obtained from (5) and (6) by introducing the decoupling condition (4).

$$\frac{x_1}{z} = \frac{Z_1}{1+R_1S_{11}(1+\mathcal{K})} \left[1 - F'_{w2}\,\frac{Z_2}{Z_1}\,\frac{S_{12}}{S_{22}} \right] \tag{10}$$

$$\frac{x_2}{z} = -\frac{Z_2}{1 + R_2S_{22}} \tag{11}$$

Direct comparison of appropriate disturbance transfer functions now shows us that passing from the unfavourable non-decoupled case to the desired exactly decoupled case depends on making the term $|Y|$ as small as possible, because

- the value of coupling quotient $\mathcal{K}$ in the denominator of the terms in brackets, which is primarily responsible for the increased tendency of the whole system to oscillate, is correspondingly reduced by multiplying with a small value of Y,
- the sole harmful influence of coupling transfer via S_{21} is reduced at the same time, as is apparent from the term in the numerator in square brackets in (6).
- The rather more useful influence of coupling transfer via S_{12} is not reduced, on the other hand, because the term in the numerator in square brackets in (5) remains unchanged.

In the exactly, partially decoupled case, single loop 2 would be autonomous and, seen from the control dynamics standpoint, exactly the same as in the decoupled system. Single loop 1 would also be autonomous, but dynamically rather better than in a decoupled system. It should be noted that, in the exactly decoupled case, the controller of loop 1 has to regulate a resultant controlled system $S_{11}\ (1 + \mathcal{K})$ which has a higher gain coefficient than in the non-decoupled case. However, the fact that the controller gain therefore has to be reduced accordingly is no real drawback, since the closed-loop gain of resultant loop 1, which is the sole determining factor, remains unchanged. Thus, in order to discuss the effects of non-exact, partial decoupling it is in principle sufficient to assess merely the modulus of the frequency characteristic $Y\ (j\omega)$ within the considered frequency range, in accordance with definition (7).

The only part of the frequency characteristic $Y\ (j\omega)$ of decisive importance is that which lies in the region of the dominant natural frequency ω of the whole system, because frequencies far above or below are not significantly involved in the control processes. This natural frequency ω is, as a first approximation, to be considered as equal to the critical natural frequency ω'_{2k} of the slower of the two individual loops assumed to be decoupled (Rule 6). The modulus of the frequency characteristic $Y\ (j\omega)$ at this frequency can thus be used as a coefficient, here termed the "decoupling factor η", which characterizes the effectiveness of non-exact decoupling (see Rule 7). The first formulation for η in Rule 7 is valid for the chosen example in Fig.4. The requirement for non-exact decoupling which brings about an improvement can be formulated as:

$$\left| 1 - X \frac{S_{22}}{S_{21}} \right| \overset{!}{<} \left| 1 + X \frac{S_{22}}{S_{21}} \right|$$

The terms $X\ S_{22}/S_{21}$, considered as frequency characteristics, and $\mathcal{K}$ at the frequency in question ω'_{2k} can be represented as vectors in a plane of complex numbers. Analysis of the geometric relationships in Fig.5 allows conversion with aid

of the cosine law:

$$1 + \left| X \frac{S_{22}}{S_{21}} \right|^2 - 2 \left| X \frac{S_{22}}{S_{21}} \right| \cos \xi$$

$$\overset{!}{<} 1 + \left| X \frac{S_{22}}{S_{21}} \right|^2 |\mathcal{K}|^2 + 2 \left| X \frac{S_{22}}{S_{21}} \right| |\mathcal{K}| \cos (\xi + \varphi).$$

In subsequent conversion one must distinguish between two different cases. We have

a) for weakly coupled systems ($0 < |\mathcal{K}| < 1$)

$$\left| X \frac{S_{22}}{S_{21}} \right| \overset{!}{<} 2 \frac{|\mathcal{K}| \cos (\xi + \varphi) + \cos \xi}{1 - |\mathcal{K}|^2}, \qquad (12)$$

b) for strongly coupled systems ($|\mathcal{K}| > 1$)

$$\left| X \frac{S_{22}}{S_{21}} \right| \overset{!}{>} - 2 \frac{|\mathcal{K}| \cos (\xi + \varphi) + \cos \xi}{|\mathcal{K}|^2 - 1}. \qquad (13)$$

In both cases, however, an additional condition must be taken into account, i.e. that the value of $|X\, S_{22}/S_{21}|$ must not be appreciably greater than unity, as "over-compensation" of the effect of coupling can cause a negatively coupled system to become positively coupled which, under certain circumstances, may again lead to less favourable results.

From the control engineering point of view, case a) is relatively unimportant because decoupling arrangements are usually superfluous with weakly coupled systems. Case b), for strongly coupled systems, is of greater interest in this respect.

It is notable that condition (13) for effective decoupling is always satisfied when ξ and $\xi + \varphi$ vary only within $\pm 90^\circ$, i.e. when the sign of the decoupling transfer is given correctly. Since the phase angles are included in condition (13), via the cosine function, it can also be stated that a phase deviation ξ of the decoupling signal relative to the compensating coupling signal is of minor importance within comparatively broad limits, e.g. $\pm 45^\circ$ (see Rule 8). It is also worth noting that a relatively high value of $|\mathcal{K}|$, which is characteristic of strong coupling, also favours the relative benefits to be gained from decoupling (Rule 9).

DECOUPLING ELEMENTS

By observing Rules 5, 6, 8 and 9 it is possible to design relatively simple, and yet effective decoupling elements employing commercially available components. The following three basic types of element can be used as required, depending on the particular circumstances:

1) P-element with adjustable control factor
2) I-element or delay element with adjustable integral action time constant or delay time constant
3) D-element with adjustable derivative action time constant.

Depending on the necessary sign of the coupling transfer being compensated, the output signals of these basic elements must be connected to the other control loop through an addition or subtraction relay. A subtraction relay is required with the example in Fig.4. If in this example the process transfer via S_{22} is slower than the coupling transfer S_{21}, including the relevant final element, the decoupling element should be of the derivative-action type, while in the converse case the element should be of the integral-action or delay type. If the two transfers are approximately equally as fast, i.e. the phase shift between the decoupling signal and the coupling signal in the frequency range concerned is not greater than $\pm 45^{\circ}$, a pure P-element is required. Thus, only in this case is the greatest benefit obtained from "static decoupling", which is often used for the sake of simplicity. By further influencing the delays of the final elements it may, in certain cases, be possible to obtain an improvement in the sense of closer agreement between the phases of the decoupling and coupling signals.

RULES FOR SETTING NON-EXACTLY, PARTIALLY DECOUPLED TWO-VARIABLE CONTROL SYSTEMS

Exact partial decoupling obtains if the sum of a harmful coupling signal to be compensated and a corresponding decoupling signal is zero for all frequencies, i.e., referring to the example in Fig.4, when the output of controller 1 no longer has any influence whatever on the input of controller 2. Non-exact, partial decoupling which

is very beneficial as regards control behaviour can be expected when the maximum attainable compensation occurs at the dominating natural frequencies ω' of the individual loops (considered as uncoupled), and these frequencies are always in the vicinity of the critical natural frequencies ω'_k of the individual loops. From this it is possible to formulate the following practical rules for setting a non-exactly, partially decoupled system.

Step 1

With the control system of the coupled loop disconnected, increase the gain of the controller whose output is connected to the coupled loop up to the critical value.

Step 2

With the first loop oscillating just critically, vary the adjustable parameter of the decoupling element until the amplitude ratio between the controlled variable of the second coupled loop and the output of the controller of the first connected loop reaches a minimum. Since this adjustment also alters the dynamic behaviour of the first loop, the gain of its controller must be reset accordingly.

Step 3

Now optimize the control systems of loops 1 and 2 individually or, if the effect of coupling cannot be reduced sufficiently, in accordance with the rules for setting non-decoupled two-variable control systems[1, 3].

BASIC RULES

Rule 1

The decoupling considered necessary creates individual autonomous loops as soon as one of the two coupling transfers via S_{12} or S_{21} is exactly decoupled, i.e. the controllers of the two individual loops can then be optimized independently of the settings on the controller of the coupled loop.

Rule 2

In the case of positive coupling (sign $\varkappa = +1$), which is comparatively rare in technical applications, any decoupling considered necessary should as far as possible compensate both coupling signals, in order to achieve better control quality for both control systems.

Rule 3

In the case of negative coupling (sign $\varkappa = -1$), which is common in technical applications, it can be said that any decoupling considered necessary should compensate only one coupling signal, as a coupling signal can always be looked on as basically useful, i.e. one of the two control systems will necessarily be supported by the second system.

Rule 4

In a negatively coupled system, the only harmful coupling transfer which needs to be compensated is the one which retains the sign when the considered disturbance affects the two control variables with different signs, or alternatively, the one which reverses the sign when the disturbance affects both control variables with the same sign (main loop transfers are standardized to have positive sign).

Rule 5

A technically effective decoupling arrangement is to superimpose one controller output signal on the other. These output signals usually employ the same energy transport medium (current, voltage, air or oil pressure), which facilitates the design of relatively simple decoupling elements.

Rule 6

A decoupling element in a negatively coupled system in principle needs to compensate the effect of coupling only in a range of frequencies about the frequency $\omega = \omega'_{2k}$ (ω'_{2k} = critical natural frequency of the slower of the two individual circuits considered as not coupled). Frequencies much higher or much lower are usually of no importance to the control operations.

Rule 7

A measure of the effectiveness of non-exact partial decoupling in a negatively coupled system is a "decoupling factor η " in the form

$$\eta = \left| \frac{1 - X\, S_{22}/S_{21}}{1 + (X\, S_{22}/S_{21})\mathcal{K}} \right|_{\text{for } \omega = \omega'_{2k}}$$

if only the coupling transfer via S_{21} is to be compensated by means of a decoupling element with the frequency characteristic X, or in the form

$$\eta = \left| \frac{1 - X\, S_{11}/S_{12}}{1 + (X\, S_{11}/S_{12})\mathcal{K}} \right|_{\text{for } \omega = \omega'_{2k}}$$

if only the coupling transfer via S_{12} is to be compensated.

$\eta = 0$ denotes exact partial decoupling

$0 < \eta < 1$ denotes beneficial non-exact partial decoupling

$\eta = 1$ denotes ineffective partial decoupling, and

$\eta > 1$ denotes detrimental non-exact partial decoupling.

Rule 8

A beneficial effect, as regards control dynamics, resulting from non-exact partial decoupling in strongly negatively coupled systems, can always be expected provided that the phase of the signal via the decoupling element and the corresponding main-loop transfer, i.e. via $X\, S_{22}$ and $X\, S_{11}$, does not deviate by more than 90°, in the frequency range of interest, from the phase of the coupling signal to be compensated via S_{21} and S_{12}, i.e. provided that, basically only the sign of the coupling transfer via X has been chosen correctly. The influence of small phase deviations, e.g. $\pm 45^{\circ}$, does not significantly affect the improvement brought about by decoupling. A great improvement can thus be achieved in every instance by appropriate selection of P, I or D-elements.

Rule 9

The effectiveness of partially decoupling increases in strongly negatively coupled systems with the extent of coupling present.

BIBLIOGRAPHY

1 Kraemer, W.: Grenzen und Möglichkeiten nichtentkoppelter, linearer Zweifachregelungen. Thesis Stuttgart Univ., 1968; Fortschritt-Bericht VDI-Zeitschrift, Series 8, No.10

2 Rijnsdorp, J.E.: Interaction in two-variable control systems for distillation columns. Automatica, Vol.1, Pergamon Press (1965), p.15-52.

3 Muckli W.: Analyse und Optimierung nichtentkoppelter Zweifachregelkreise, Thesis Technische Hochschule Aachen (1968)

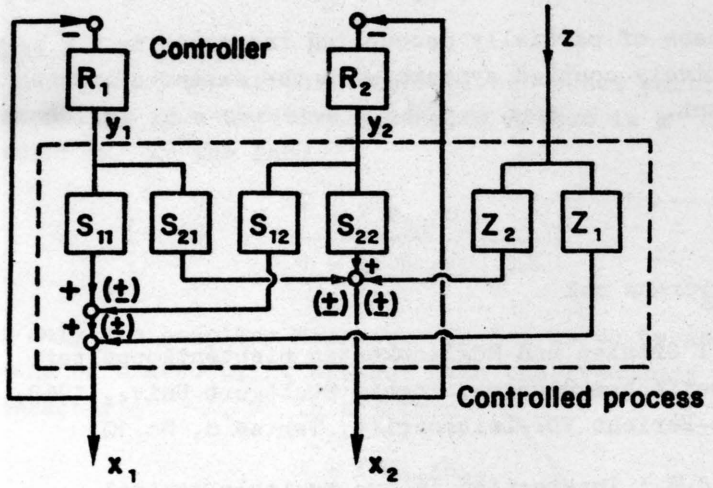


Fig.1
Generally valid block diagram for non-decoupled two-variable control systems

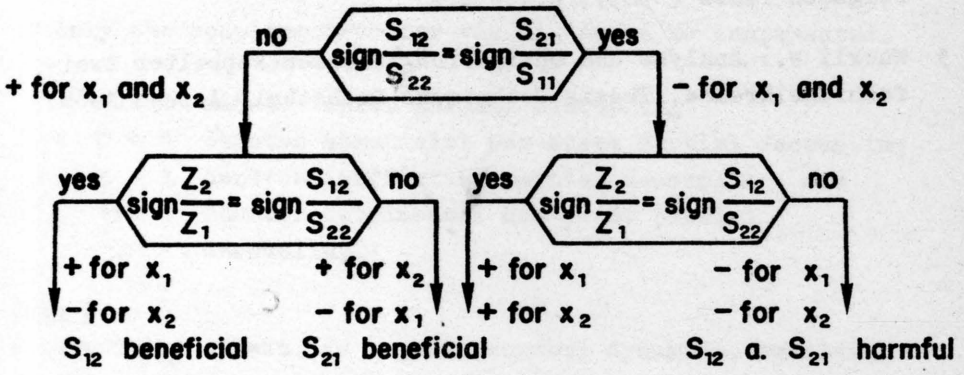


Fig.2
Tendencies of disturbance transfer, disregarding altered stability conditions

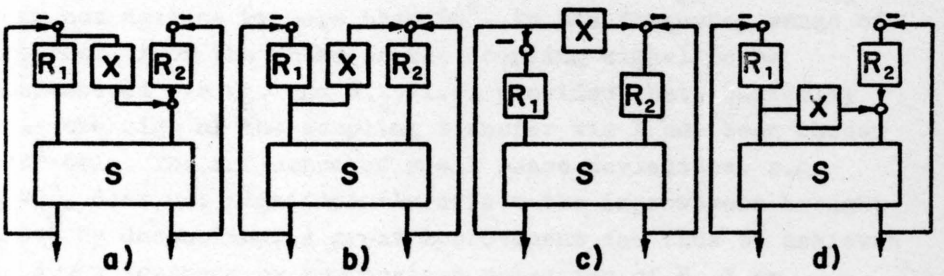


Fig. 3
Possible arrangements for partial decoupling

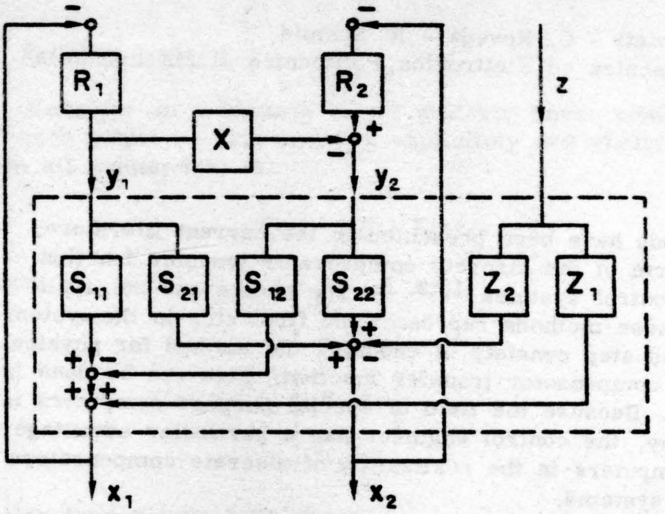


Fig. 4
Example of partial decoupling of a negatively coupled system

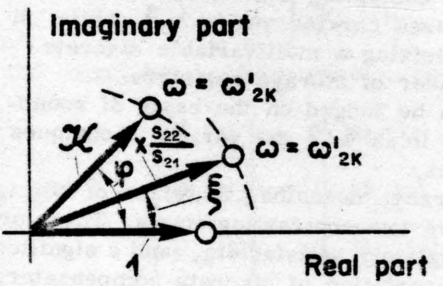


Fig.5
Geometric representation of the characteristics of coupling in the frequency characteristic plane

ON THE OPTIMAL IMPLEMENTATION OF MULTIVARIABLE DISCRETE LINEAR SYSTEMS

E. Biondi - L. Divieti - C. Roveda - R. Schmid
Istituto di Elettrotecnica ed Elettronica, Politecnico di Milano, Milano, Italy

I. Introduction

Several methods have been presented in the current literature, which yield the form of the discrete compensator transfer function in sampled-data control systems [1, 2, 3]. The choice and the application of one of these methods represent the first step in the system design. The second step consists of choosing the method for physical realization of the compensator transfer function. This can be done in a variety of ways. Because the field of special purpose computers is growing day by day, the control engineer has a particular advantage in using these computers in the realization of discrete compensators for multivariable systems.

When a special purpose computer is to be used, the second step of the compensator synthesis consists of choosing the computer hardware so that both economy and accuracy will be warrented.

From the economic point of view, a logical scheme for comparison among the different techniques can be based on the number of storage registers required and the computing time involved. A comparative analysis in this way has been carried out in [1, 3], while in [4] a method is presented for synthesizing a multivariable discrete compensator with the minimal number of storage registers.

The compensator accuracy can be judged on the basis of round-off error accumulation. As proved in [5, 6, 7] the various techniques may present remarkable differences.

In general, economy and accuracy, described in terms of storage registers and round-off errors, are two contrasting items. Therefore, the previous analyses are not completely satisfactory, and a significant cost function for an optimal implementation of discrete compensators must take into account both the number of storage registers and the round-off errors.

This paper deals with the problem of the optimal implementation of multivariable discrete linear compensators. The cost function is a linear combination of the number of delay elements required and the mean square round-off errors on the compensator outputs. Considering the parallel method of synthesis and making use of four basic structures, the optimization problem is approached in terms of graph theory, and it corresponds to the determination of the optimal

---

(o) The present research has been supported by C.N.R. (Consiglio Nazionale delle Ricerche), Rome, Italy.

arborescence in a particular acyclic graph. This optimization problem is finally solved via Dynamic Programming.

II. Problem statement

Consider an n input-m output discrete linear compensator, and let each output be expressed as explicitely and wholly dependent upon all inputs; that is

$$\underline{Y}(z) = \underline{D}(z)\ \underline{X}(z) \tag{1}$$

where

$$\underline{X}(z) = \begin{vmatrix} X_1(z) \\ \dots \\ X_i(z) \\ \dots \\ X_n(z) \end{vmatrix} \tag{2}$$

is the input z-transform vector,

$$\underline{Y}(z) = \begin{vmatrix} Y_1(z) \\ \dots \\ Y_j(z) \\ \dots \\ Y_m(z) \end{vmatrix} \tag{3}$$

is the output z-transform vector, and

$$\underline{D}(z) = \begin{vmatrix} \underline{D}_1(z) \\ \dots \\ \underline{D}_j(z) \\ \dots \\ \underline{D}_m(z) \end{vmatrix} = \begin{vmatrix} D_{11}(z) \dots D_{1i}(z) \dots D_{1n}(z) \\ \dots\dots\dots\dots\dots\dots \\ D_{j1}(z) \dots D_{ji}(z) \dots D_{jn}(z) \\ \dots\dots\dots\dots\dots\dots \\ D_{m1}(z) \dots D_{mi}(z) \dots D_{mn}(z) \end{vmatrix} \tag{4}$$

is the compensator transfer matrix. The block diagram of the compensator is shown in Fig. 1 (1)

Expressing each output $Y_j(z)$ as depending upon all inputs $X_1(z)$, $\dots, X_n(z)$, it results :

$$Y_j(z) = \sum_{i=1}^{n} D_{ji}(z)\ X_i(z) \tag{5}$$

---

(1) The structure of Fig. 1 has been defined by Mesarovic' as P-canonical structure[8].

If the parallel method of synthesis is considered, the transfer functions $D_{ji}(z)$ are supposed to be given in partial fraction expansion form, that is :

$$D_{ji}(z) = f_o^{ji} + \sum_{h=1}^{H_{ji}-2M_{ji}} D_h^{ji}(z) + \sum_{h^*=1}^{M_{ji}} D_{h^*}^{ji}(z) \tag{6}$$

where

$$D_h^{ji}(z) = \frac{f_n^{ji} z^{-1}}{1 - p_h^{ji} z^{-1}} \tag{7}$$

$$D_{h^*}^{ji}(z) = \frac{(\alpha_{h^*}^{ji} + j\beta_{h^*}^{ji}) z^{-1}}{1 - (\sigma_{h^*}^{ji} + j\omega_{h^*}^{ji}) z^{-1}} + \frac{(\alpha_{h^*}^{ji} - j\beta_{h^*}^{ji}) z^{-1}}{1 - (\sigma_{h^*}^{ji} - j\omega_{h^*}^{ji}) z^{-1}} \tag{8}$$

and p_h^{ji} are the real poles of $D_{ji}(z)$, and $\sigma_{h^*}^{ji} \pm j\omega_{h^*}^{ji}$ are the pairs of complex conjugate poles of $D_{ji}(z)$. From (5) and (6) it follows :

$$Y_j(z) = \sum_{i=1}^{n} f_o^{ji} X_i(z) + \sum_{i=1}^{n} \sum_{h=1}^{H_{ji}-2M_{ji}} D_h^{ji}(z)\, X_i(z) + \sum_{i=1}^{n} \sum_{h^*=1}^{M_{ji}} D_{h^*}^{ji}(z)\, X_i(z) \tag{9}$$

Each function $D_h^{ji}(z)$ can be realized by means of one of the structures of Fig. 2-a and Fig. 2-b, referred to as "first" and "second elemental structure" respectively.

Each function $D_{h^*}^{ji}(z)$ can be realized by means of one of the structures of Fig. 2-c and Fig. 2-d, referred to as "third" and "fourth elemental structure" respectively.

As shown in Appendix, the round-off error accumulation is different for the various elemental structures. Under the assumption that the errors $\varepsilon_h^{ji}(kT)$ and $\varepsilon_{h^*}^{ji}(kT)$ are statistically independent for $h = 1, \ldots, H_{ji}-2M_{ji}$; $h^* = 1, \ldots, M_{ji}$; $i = 1, \ldots, n$, the mean square value of the round-off error on the j-th compensator output ($j = 1, \ldots, m$) is :

$$\overline{\varepsilon_j^2} = \sum_{i=1}^{n} \sum_{h=1}^{H_{ji}-2M_{ji}} \overline{\varepsilon_h^{ji\,2}} + \sum_{i=1}^{n} \sum_{h^*=1}^{M_{ji}} \overline{\varepsilon_{h^*}^{ji\,2}} \tag{10}$$

where the, errors $\overline{\varepsilon_h^{ji\,2}}$ and $\overline{\varepsilon_{h^*}^{ji\,2}}$ are given in the Table 1 of the Appendix.

If the poles of $\underline{D}(z)$ are all distinct, the number of storage registers required for the compensator implementation is :

$$N_{sr} = \sum_{i=1}^{n} \sum_{j=1}^{m} H_{ji} \qquad (11)$$

Thus, the optimal implementation is obtained when the elemental structures are chosen so that the round-off errors $\overline{\varepsilon_j^2}$ are minimal.

If the poles of $\underline{D}(z)$ are not all distinct, it turns useful to introduce the four structures shown in Fig. 3 and referred to as "first" (Fig. 3-a), "second" (Fig. 3-b), "third" (Fig. 3-c), and "fourth" (Fig. 3-d) "basic structures". These basic structures can be considered as an extension to r inputs or to s outputs of the elemental structures.

Using only one delay element, the first basic structure realizes the signal

$$W_h^j(z) = \sum_{k=1}^{r} D_h^{jk}(z)\, X_k(z) \qquad (12)$$

when the transfer functions $D_h^{jk}(z)$ $(k = 1, \ldots, r)$ have the same real pole p_h.

Using again one delay element, the second basic structure realizes simultaneously the signals

$$Y_h^{ki}(z) = D_h^{ki}(z)\, X_i(z) \qquad (k = 1, \ldots, s) \qquad (13)$$

when the transfer functions $D_h^{ki}(z)$ $(k = 1, \ldots, s)$ have the same real pole p_h.

The third and the fourth basic structures substitute the first and the second basic structures respectively in the case of common pairs of complex conjugate poles.

An analysis of the round-off errors in the basic structures is carried out in Appendix, and the results obtained are shown in Table 2.

When the poles of $\underline{D}(z)$ are not all distinct and the basic structures are used for implementing the multivariable discrete compensator, a remarkable improvement can be obtained either in economy and accuracy. On the other hand, it can be easily proved that the synthesis which gives the maximum saving of storage registers, is not generally the optimal one from the viewpoint of round-off errors. Thus, it seems to be meaningful to define a cost function involving both the number of delay elements required and the round-off errors on the compensator outputs. The following linear cost function is considered :

$$C = N_{sr} + \sum_{j=1}^{m} \lambda_j \overline{\varepsilon_j^2} \qquad (14)$$

where N_{sr} is the number of storage registers (delay elements) required

for the implementation, $\overline{\varepsilon_j^2}$ is the mean square round-off error on the j-th output and λ_j is its weighting factor [2].

In the next sections it is considered the problem of finding the optimal parallel implementation of a multivariable discrete linear compensator under the following assumptions :

(i) the multivariable discrete compensator is defined by means of its z-transfer matrix $\underline{D}(z)$

(ii) the transfer functions $D_{ji}(z)$ are given in the form (6)

(iii) the real poles or the pairs of complex conjugate poles of $\underline{D}(z)$ are simple and not all distinct

(iv) the parallel implementation is carried out using the basic structures

(v) the cost function is given by (14).

III. Optimal parallel synthesis

Let $\bar{p}_1, \ldots, \bar{p}_t$ with $t < \sum_{i=1}^{n} \sum_{j=1}^{m} (H_{ji} - M_{ji})$, be the distinct poles (or pairs of complex conjugate poles) of the transfer matrix $\underline{D}(z)$. It is possible (Fig. 4) to associate to each pole $\bar{p}_h$ (or to a pair of complex conjugate poles) an acyclic bipartite graph $G^h = (\mathbf{X}, \mathbf{Y}, \mathbf{U}^h)$, where $\mathbf{X}$ and $\mathbf{Y}$ are the sets of the compensator input and output variables respectively. The set $\mathbf{U}^h$ consists of the oriented arcs $u_{ij}^h = (X_i, Y_j)$; an arc u_{ij}^h exists if and only if the function $D_{ji}(z)$ contains the pole $\bar{p}_h$.

The first basic structure of Fig. 3-a corresponds to the partial subgraph of G^h defined by

$$g_{Y_j}^h = (\mathbf{X}, Y_j, \mathbf{U}_{Y_j}^h) \qquad (15)$$

where

$$Y_j = \{Y_j\}$$

$$\mathbf{U}_{Y_j}^h \subseteq \mathbf{U}^h$$

$$\mathbf{U}_{Y_j}^h \neq \phi$$

---

(2) It is worth while noting that, keeping different weighting factors λ_j, it is possible to obtain a different degree of accuracy on the various compensator outputs.

The second basic structure of Fig. 3-b corresponds to the partial subgraph of G^h defined by

$$g^h_{x_i} = (X_i , Y , U^h_{x_i}) \tag{16}$$

where

$$X_i = \{ X_i \}$$

$$U^h_{x_i} \subseteq U^h$$

$$U^h_{x_i} \neq \phi$$

The third and fourth basic structures of Figs. 3-c, d correspond to analogous partial subgraphs of the graph G^h, when G^h is referred to a pair of complex conjugate poles.

A general compensator parallel synthesis can be seen as the choice, for each graph G^h, of two partial subgraphs $G^h_x = (S^h_x, Y, U^h_x)$ and $G^h_y = (X, S^h_y, U^h_y)$ such that

$$S^h_x \subseteq X \qquad \text{and} \qquad S^h_y \subseteq Y$$

$$U^h_x \cup U^h_y = U^h \qquad \text{and} \qquad U^h_x \cap U^h_y = \phi$$

Therefore, the parallel synthesis with basic structures consists of decomposing the two partial subgraphs G^h_x and G^h_y into partial subgraphs g^h_{xi} for $X_i \in S^h_x$ and g^h_{yj} for $Y_j \in S^h_y$.

In order to obtain the synthesis which minimizes the cost function (14), two different costs must be assigned to each arc u^h_{ij} of G^h depending on whether $u^h_{ij} \in U^h_x$ or $u^h_{ij} \in U^h_y$.

According to (14) and to the results given in Table 2 of the Appendix, these costs are :

$$\mu^h_{ij} = \lambda_j \left[1 + \frac{(f^{ji}_h)^2}{1 - \bar{p}^{-2}_h} \right] \frac{q^2}{12} \qquad \text{for} \quad u^h_{ij} \in U^h_x \tag{17}$$

$$\nu^h_{ij} = \lambda_j \frac{1}{1 - \bar{p}^{-2}_h} \frac{q^2}{12} \qquad \text{for} \quad u^h_{ij} \in U^h_y \tag{18}$$

if the graph G^h corresponds to a real pole, or :

$$\mu^h_{ij} = 2\lambda_j \left[1 + \frac{(\alpha^{ji}_{h^*})^2 + (\beta^{ji}_{h^*})^2}{1 - (\bar{\sigma}^2_{h^*} + \bar{\omega}^2_{h^*})} \right] \frac{q^2}{12} \quad \text{for } u^h_{ij} \in U^h_x \tag{19}$$

$$v_{ij}^{h} = 2\lambda_j \frac{1}{1-(\bar{\sigma}_{h^*}^2 + \bar{\omega}_{h^*}^2)} \frac{q^2}{12} \quad \text{for } u_{ij}^{h} \in U_y^h \tag{20}$$

if the graph G^h corresponds to a pair of complex conjugate poles. Moreover, it is necessary to assign to each vertex $Y_j \in S_y^h$ a cost

$$c_j^h = \frac{\lambda_j}{1-\bar{p}_h^2} \frac{q^2}{12} \tag{21}$$

if G^h corresponds to a real pole, or

$$c_j^h = \frac{\lambda_j}{1-(\bar{\sigma}_{h^*}^2 + \bar{\omega}_{h^*}^2)} \frac{q^2}{12} \tag{22}$$

if G^h corresponds to a pair of complex conjugate poles.

In terms of graph theory, the optimal synthesis consists of choosing two partial subgraphs G_x^h and G_y^h such that the cost function

$$C = \sum_{h=1}^{t} C_h = \sum_{h=1}^{t} \left\{ |S_x^h| + |S_y^h| + \sum_{Y_j \in S_y^h} c_j^h + \sum_{u_{ij}^h \in U_x^h} \mu_{ij}^h + \sum_{u_{ij}^h \in U_y^h} v_{ij}^h \right\} \tag{23}$$

is minimal.

It can be easily proved that, for parallel synthesis, it results :

$$\min_{\substack{\{G_x^h, G_y^h\} \\ h=1,\dots,t}} C = \sum_{h=1}^{t} \min_{\{G_x^h, G_y^h\}} C_h \tag{24}$$

Therefore, the compensator optimal parallel synthesis with basic structures leads to t distinct optimization problems. The solution of these problems is presented in the next section.

IV. Determination of the optimal parallel synthesis

The optimization problem defined in the preceding section can be solved associating to each graph G^h another graph $P^h = (R_o, \mathbf{A}, \mathbf{B}^h, \Gamma_{1h}, \Gamma_{2h})$, where R_o is an isolated vertex, $\mathbf{A}$ and $\mathbf{B}^h$ are two disjoint sets of vertices, Γ_{1h} is the mapping of R_o into $\mathbf{A}$, and Γ_{2h} is the mapping of $\mathbf{A}$ into $\mathbf{B}^h$ (Fig. 5). In particular,

$$\mathbf{A} = \mathbf{X} \cup \mathbf{Y}$$

$$|\mathbf{B}^h| = |U^h| = N_h$$

$$\forall A_i \in A \; : \left| \Gamma_{1h}^{-1} A \right| = 1$$

$$\forall B_k^h \in B^h \; : \left| \Gamma_{2h}^{-1} B \right| = 2$$

Each vertex $B_k^h \in B^h$ corresponds to an arc $u_{ij}^h = (X_i, Y_j) \in U^h$ and

$$\Gamma_{2h}^{-1} B_k^h = \left\{ A_r , A_s \right\}$$

with

$$A_r = X_i$$

$$A_s = Y_j$$

In each graph P^h, a cost ρ_r^h is associated to each arc (R_o, A_r), and

$$\rho_r^h = 1 \qquad \text{for} \qquad A_r = X_i \in X$$

$$\rho_r^h = 1 + c_j^h \qquad \text{for} \qquad A_r = Y_j \in Y$$

where c_j^h is given by (21) and (22).

A cost δ_{rk}^h is assigned to each arc (A, B_k^h), and

$$\delta_{rk}^h = \mu_{ij}^h \qquad \text{for} \quad B_k^h = u_{ij}^h \quad \text{and} \quad A_r = X_i \in X$$

$$\delta_{rk}^h = \nu_{ij}^h \qquad \text{for} \quad B_k^h = u_{ij}^h \quad \text{and} \quad A_r = Y_j \in Y$$

where μ_{ij}^h and ν_{ij}^h are given by (17), (18) or (19), (20).

The optimization problem, defined in the preceding section and referred to the cost function (23), because of (24) can be formulated on graphs P^h $(h = 1, \ldots, t)$ as follows: for each graph P^h, find the arborescence having R_o as its "root" and B_j^h $(j = 1, \ldots, N_h)$ as its "leaves" and presenting minimal cost.

This problem has been solved by the authors and the solution has been presented in [9]. Actually, the nonlinear optimization problem involved is approached and solved via Dynamic Programming, considering the multistage decision process shown in Fig. 6. In this process the following definitions are given :

(i) State $\underline{p}_i^h$: set of vertices $A_i \in \mathbf{A}$, through which the arborescence paths ending at $B^h_{N_h}, \ldots, B^h_{i+1}$ go

(ii) Decision q_i^h : vertex $A_k \in \Gamma^{-1}_{2h} B^h_i$ through which the arborescence path ending at B^h_i goes

(iii) Transformation function :

$$\underline{p}^h_{i-1} = \mathcal{T} (\underline{p}^h_i , q^h_i) = \underline{p}^h_i \cup q^h_i$$

(iv) Constraints :

$$\Gamma_{2h} \underline{p}^h_i \supseteq \left\{ B^h_{N_h}, \ldots, B^h_{i+1} \right\}$$

$$\underline{p}^h_{N_h} = \phi$$

$$q^h_i \in \Gamma^{-1}_{2h} B^h_i$$

(v) Cost :

$$g_i(\underline{p}^h_i , q^h_i) = \rho(q^h_i) \varphi(\underline{p}^h_i , q^h_i) + \delta_i(q^h_i)$$

where

$$\rho(q^h_i) = \rho^h_k \quad \text{if} \quad q^h_i = A_k$$

and $\varphi(\underline{p}^h_i , q^h_i)$ is the two - valued function given by

$$\varphi(\underline{p}^h_i , q^h_i) = 0 \quad \text{if} \quad q^h_i \in \underline{p}^h_i$$

$$\varphi(\underline{p}^h_i , q^h_i) = 1 \quad \text{if} \quad q^h_i \notin \underline{p}^h_i$$

and

$$\delta(q^h_i) = \delta^h_{ki} \quad \text{if} \quad q^h_i = A_k$$

The objective function which must be minimized is :

$$C_h = \sum_{i=1}^{N_h} g_i (\underline{p}^h_i , q^h_i)$$

Applying the section method [10] and proceeding from stage 1 back to stage N_h through the sections $S_1, \ldots, S_i, \ldots, S_{Nh}$ (Fig. 6), the Dynamic Programming algorithm yields the following functional equation

$$f(\underline{p}_i^h) = \min_{q_i^h} \left[g_i(\underline{p}_i^h , q_i^h) + f_{i-1}(\underline{p}_{i-1}^h) \right]$$

$$\underline{p}_{i-1}^h = \underline{p}_i^h \quad \cup \quad q_i^h$$

for $i = 1, \ldots, N_h$, starting with $f_o = 0$.

The calculations involved in the solution are generally rather heavy, but the use of a computer can easily overcome this difficulty.

V. Conclusion

In the implementation of digital compensators the economy and accuracy, expressed in terms of storage registers and round-off error accumulation, are generally contrasting items. In the current literature, these two aspects of the synthesis have been treated separately, while a significant objective function for an optimal implementation should involve both of them. In this paper a method has been presented for the optimal implementation of multivariable discrete linear compensators. The cost function considered is a linear combination of the number of storage registers and the mean square round-off errors on the compensator outputs. The method consists of applying the Dynamic Programming to the solution of a particular graph problem to which the original problem can be led.

VI. Appendix

The effect of round-off operations in the elemental structures defined in section II has been studied in [6, 7]. The mean square round-off errors on the outputs of these structures are reported in Table 1, where q denotes the amplitude of the quantization step.

Following the method exposed in [6], the round-off errors in the four basic structures can be easily computed. The signals realized by each basic structure and the relative mean square round-off errors are given in Table 2.

References

1. JURY, E.: "Sampled-Data Control Systems", J. Wiley, New York, 1958.

2. TOU, J. : "Digital and Sampled-Data Control Systems", Mc Graw-Hill, New York, 1959.

3. MONROE, A. : "Digital Processes for Sampled-Data Systems", J. Wiley, New York, 1962.

4. BIONDI, E., DIVIETI, L. and SCHMID, R. : "Structural analysis and synthesis of multivariable linear sampled-data systems", IFAC Symposium on Multivariable Control Systems, Duesseldorf, October 1968.

5. KNOWLES, J. B. and EDWARDS, R. : "Effect of a finite-word-length computer in a sampled-data feedback system", Proc. IEE, vol. 112, n. 6, June 1965, pp. 1197-1207.

6. DIVIETI, L., ROSSI, C., SCHMID, R. and VERESCHKIN, A. : "Statistical analysis of round-off errors in digital compensator programming", Avtomatika i Telemekhanika, n. 8, August 1968.

7. DIVIETI, L., ROSSI, C., SCHMID, R. and VERESCHKIN, A. : "Two structures for implementing complex poles of pulsed transfer functions", to appear in Avtomatika i Telemekhanika.

8. MESAROVIC, M. D.: "The Control of Multivariable Systems", J. Wiley, New York, 1960.

9. BIONDI, E., DIVIETI, L., ROVEDA, C. and SCHMID, R. : "Optimal solution for a class of assignment and transportation problems", 2nd International Conference on Computing Methods in Optimization Problems, Sanremo, September 1968.

10. BIONDI, E., BRIOSCHI, F., DIVIETI, L. and ROZZONI, P. : "Structures of multistage decision processes", in "Identificazione e Ottimizzazione", C.N.R., Collana di Automatica, n. 2, pp. 325-344.

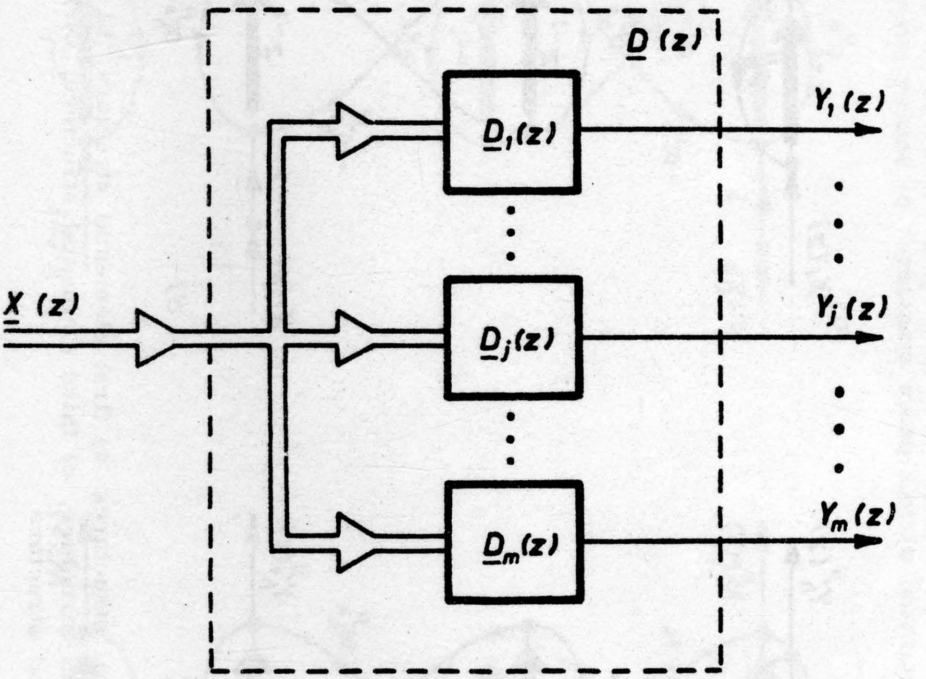


Fig. 1 - P-canonical structure.

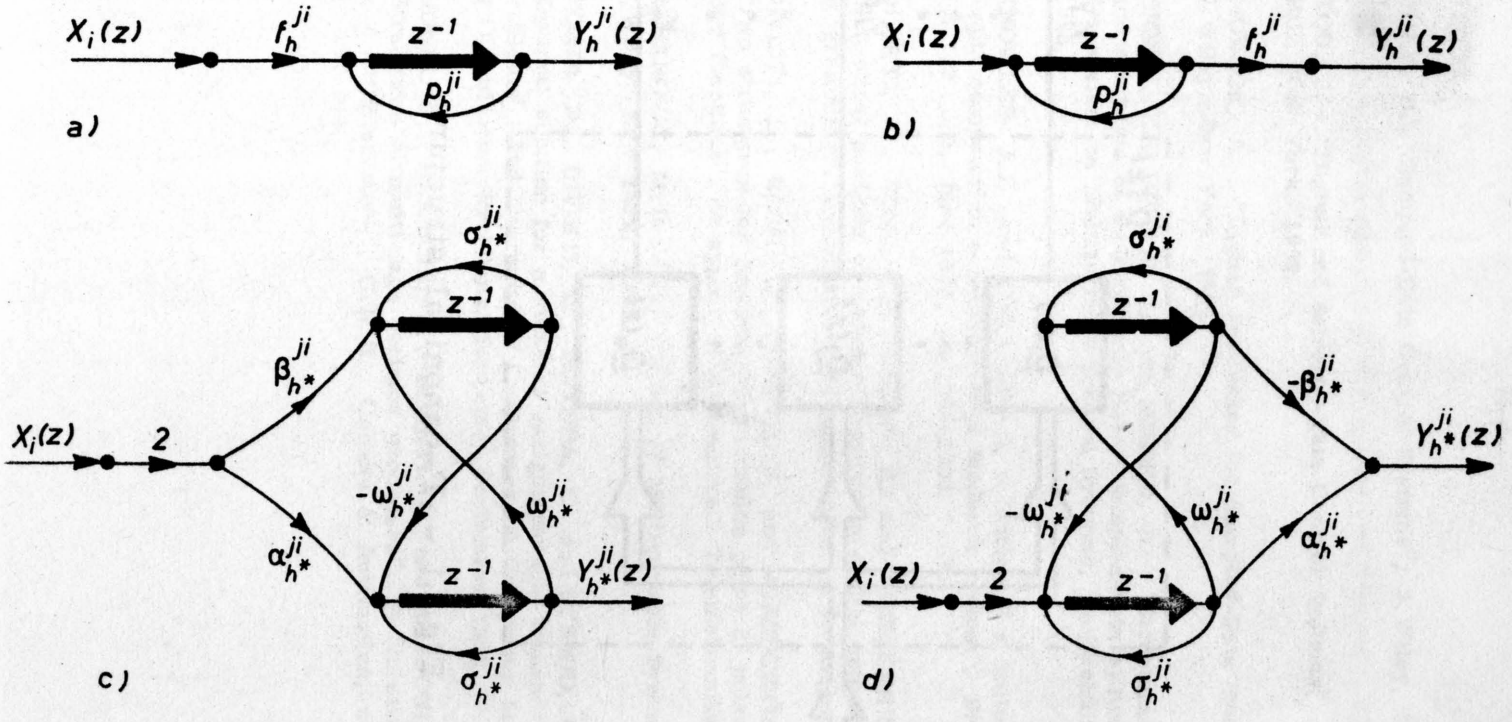


Fig. 2 - Elemental structures: a) first elemental structure, b) second elemental structure, c) third elemental structure, d) fourth elemental structure.

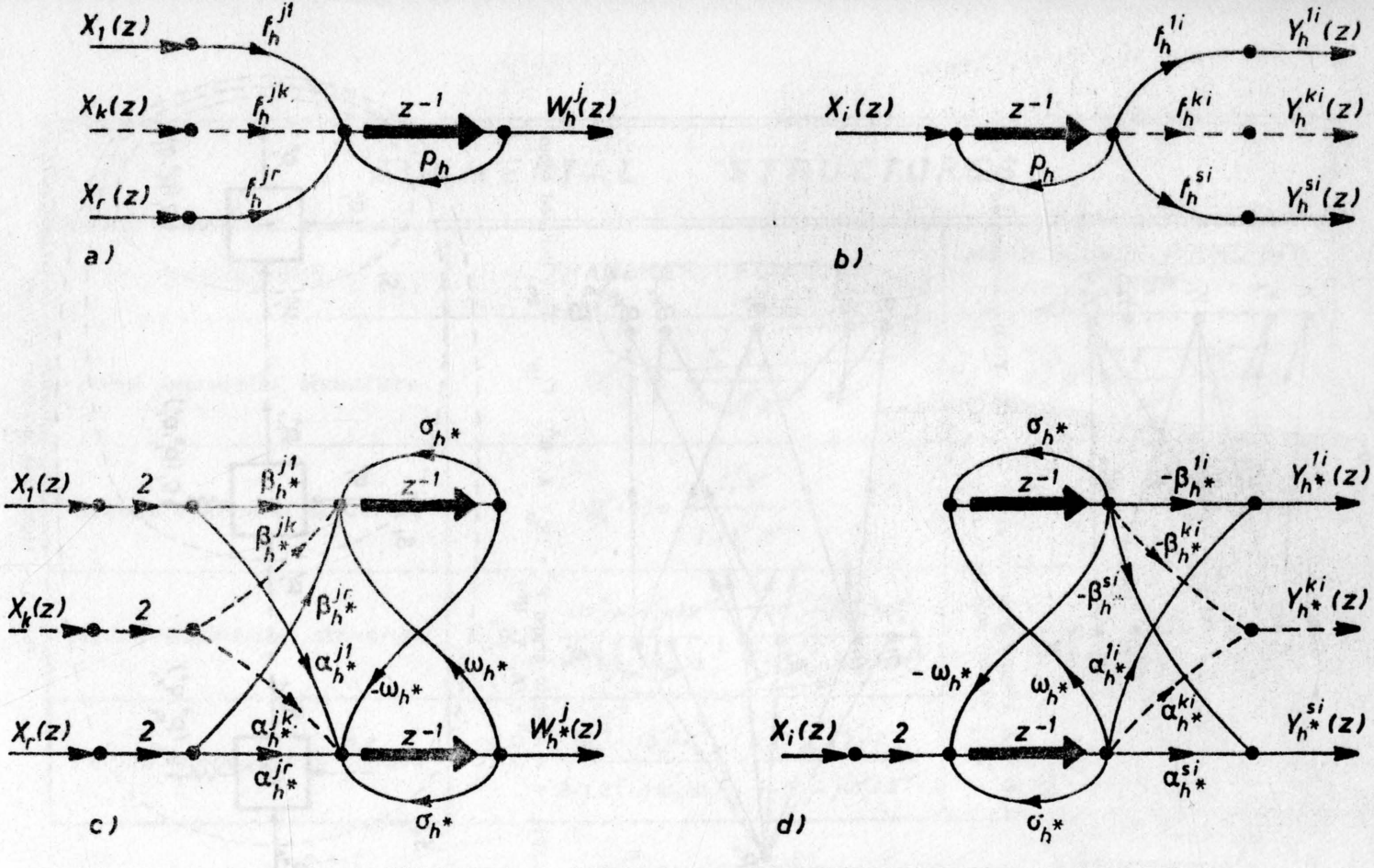


Fig. 3 - Basic structures: a) firs basic structure, b) second basic structure, c) third basic structure, d) fourth basic structure.

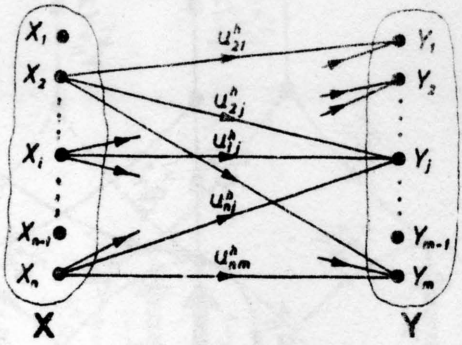


Fig. 4 - Acyclic bipartite graph $G^h = (\mathbf{X}, \mathbf{Y}, \mathbf{U}^h)$ associated to the distinct pole p_h.

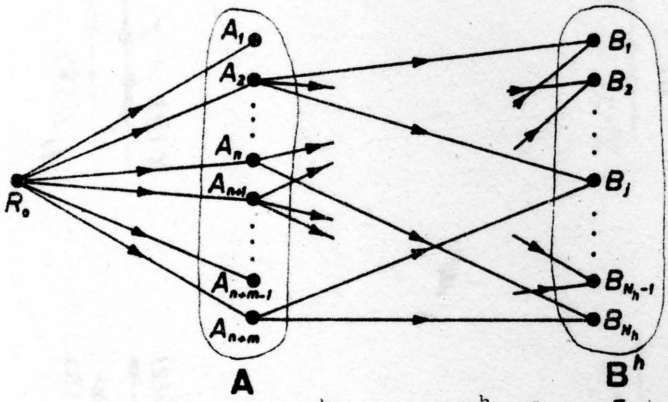


Fig. 5 - Acyclic graph $P^h = (R_o, \mathbf{A}, \mathbf{B}^h, \Gamma_{1h}, \Gamma_{2h})$ derived from $G^h = (\mathbf{X}, \mathbf{Y}, \mathbf{U}^h)$.

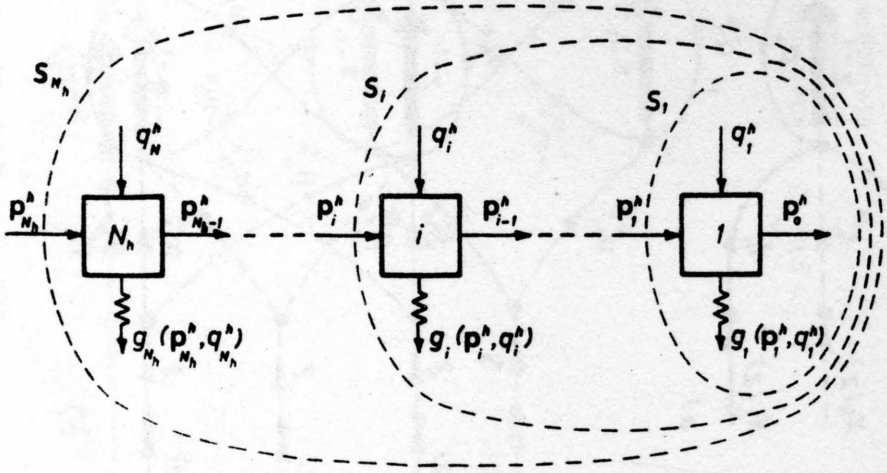


Fig. 6 - Multistage decision process.

ELEMENTAL STRUCTURES

STRUCTURE	TRANSFER FUNCTION	MEAN SQUARE ROUND-OFF ERROR
First elemental structure	$D_h^{ji}(z) = \dfrac{f_h^{ji} z^{-1}}{1 - \rho_h^{ji} z^{-1}}$	$\overline{\varepsilon_h^{ji^2}} = \dfrac{2}{1 - \rho_h^{ji^2}} \dfrac{q^2}{12}$
Second elemental structure	$D_h^{ji}(z) = \dfrac{f_h^{ji} z^{-1}}{1 - \rho_h^{ji} z^{-1}}$	$\overline{\varepsilon_h^{ji^2}} = \left(1 + \dfrac{f_h^{ji^2}}{1 - \rho_h^{ji^2}}\right) \dfrac{q^2}{12}$
Third elemental structure	$G_{h*}^{ji}(z) = \dfrac{(\alpha_{h*}^{ji} + j\beta_{h*}^{ji}) z^{-1}}{1 - (\sigma_{h*}^{ji} + j\omega_{h*}^{ji}) z^{-1}} + \dfrac{(\alpha_{h*}^{ji} - j\beta_{h*}^{ji}) z^{-1}}{1 - (\sigma_{h*}^{ji} - j\omega_{h*}^{ji}) z^{-1}}$	$\overline{\varepsilon_{h*}^{ji^2}} = \dfrac{3}{1 - (\sigma_{h*}^{ji^2} + \omega_{h*}^{ji^2})} \dfrac{q^2}{12}$
Fourth elemental structure	$G_{h*}^{ji}(z) = \dfrac{(\alpha_{h*}^{ji} + j\beta_{h*}^{ji}) z^{-1}}{1 - (\sigma_{h*}^{ji} + j\omega_{h*}^{ji}) z^{-1}} + \dfrac{(\alpha_{h*}^{ji} - j\beta_{h*}^{ji}) z^{-1}}{1 - (\sigma_{h*}^{ji} - j\omega_{h*}^{ji}) z^{-1}}$	$\overline{\varepsilon_{h*}^{ji^2}} = 2\left[1 + \dfrac{\alpha_{h*}^{ji^2} + \beta_{h*}^{ji^2}}{1 - (\sigma_{h*}^{ji^2} + \omega_{h*}^{ji^2})}\right] \dfrac{q^2}{12}$

BASIC STRUCTURES

STRUCTURE	OUTPUT SIGNALS	MEAN SQUARE ROUND-OFF ERROR ON THE OUTPUT SIGNALS
First basic structure	$W_h^j(z) = \sum_{k=1}^{r} D_h^{jk}(z)\, X_k(z)$	$\overline{\varepsilon_{W_h^j}^2} = \frac{r+1}{1-\rho_h^2}\,\frac{q^2}{12}$
Second basic structure	$Y_h^{ki}(z) = D_h^{ki}(z)\, X_i(z)$ $k=1,\ldots,s$	$\overline{\varepsilon_h^{ki\,2}} = \left(1 + \frac{f_h^{ki\,2}}{1-\rho_h^2}\right)\frac{q^2}{12}$ $k=1,\ldots,s$
Third basic structure	$W_h^j{}^*(z) = \sum_{k=1}^{r} G_h^{jk}{}^*(z)\, X_k(z)$	$\overline{\varepsilon_{W_h^j}^2{}^*} = \frac{r+2}{1-(\sigma_h^{2}{}^* + \omega_h^{2}{}^*)}\,\frac{q^2}{12}$
Fourth basic structure	$Y_h^{ki}{}^*(z) = G_h^{ki}{}^*(z)\, X_i(z)$ $k=1,\ldots,s$	$\overline{\varepsilon_h^{ki\,2}{}^*} = 2\left[1 + \frac{\alpha_h^{ki\,2}{}^* + \beta_h^{ki\,2}{}^*}{1-(\sigma_h^{2}{}^* + \omega_h^{2}{}^*)}\right]\frac{q^2}{12}$ $k=1,\ldots,s$