Monograph:

Nicholson, H. (1970) Sequential Least-Squares Estimation and the Electrical Network Problem. Research Report. ACSE Research Report CE70-1 . Department of Automatic Control and Systems Engineering

University of Sheffield

Department of Control Engineering

Sequential least-squares estimation and the electrical network problem

H. Nicholson

Research report CE70-1

October 1970

Summary

Fundamental relationships existing between the electrical network problem and least-squares estimation theory are illustrated, and the theorems of least-power associated with the equilibrium conditions of physical systems are shown to correspond to the mathematical concepts of least-squares theory. Solution of the electrical network problem based on Kron's method of tearing and interconnection is considered, and is shown to incorporate an iterative technique similar to the algorithm developed for sequential least-squares data fitting. A formulation of the multimachine power system problem is shown to incorporate inherently the form of a least-power theorem associated with a potential function related to machine power losses and network power. A piecewise solution of the transient problem is also developed based on a sequential connection of generators to the network nodes. Similar methods will have application for the decomposition of large-scale system problems which can be identified with dynamic blocks connected to the nodes of a network structure. The techniques are also shown to apply for analysis of the multivariable control system problem based on a sequential connection of control loops, analogous to the interconnection of link admittances in Kron's method of network tearing. Such analysis can simplify the procedure for control design based on a sensitivity analysis or by the application of conventional single-loop techniques. The physical significance of tearing and interconnection in the multimachine and multivariable control system problem and the resulting sequential formulations are illustrated in block diagram form.

1. Introduction

The fundamental theory of linear multivariable systems concerned with least-squares estimation and control can be shown to be associated with concepts originating in electrical network theory. The equations of mathematical physics concerned with heat and material flow, elasticity, economics and biology also have a close correspondence with the algebraic topological properties of electrical networks. The algebraic structure of the network problem has been shown to be closely related to the operational structure of the vector calculus[1], and numerical solutions of algebraic and differential equations and the dynamical equations of Lagrange and Hamilton can be associated with electrical networks[3]. Network analogues have also been used for representing dielectric and magnetic bodies in an arbitrary impressed field[2], and acoustical and mechanical systems[4], and physical analogues have been applied in economic theory[5].

Electrical network theory is based on the elemental principles of algebraic topology, with an algebraic structure relating the physical variables in a topological structure or graph which defines the interconnection of the discrete network elements. The topological characteristics are based on Kirchoff's voltage and current laws, and Ohm's law introduces a transformation between the dual sets of physical variables associated with the impedance elements. In general, the flow problems of physical systems will involve a transformation between such sets of variables, including across (potential) variables defined on the nodes and through (current) variables defined on the oriented links[6]. The existence of mesh and incidence laws in a linear graph will then establish the analogy between physical systems and permit the use of similar techniques based on a unified solution[7].

The paper illustrates the fundamental concepts and principles of electrical network theory which can be associated with the behaviour of other physical systems. The equilibrium conditions of electrical networks and other physical processes can be associated with a least-power theorem which is analogous with the solution of the classical least-squares problem. The paper emphasises particularly the relationships existing between electrical network theory and methods of least-squares estimation which have a wide application in the analysis of experimental data in many fields of applied science.

Methods of piecewise analysis have been developed for solution of the electrical network problem based on an interconnection of subnetworks and links, as in Kron's method of tearing[3]. The techniques can be used for reducing the complexity of a wide range of large-scale system problems, and these are now shown to be analogous to the techniques developed for sequential data fitting based on least-squares estimation theory. Similar methods are developed for solution of the multimachine power system problem and the existence of a least-power theorem is illustrated. The techniques are also extended in order to formulate a piecewise solution of the transient machine problem, and to investigate a decomposition of the multivariable control system problem using a sequential interconnection of control loops.

2. Electrical network analogue for least-squares estimation

Basic electrical network theory can be closely related to the equilibrium problem associated with other physical systems based on variational principles applied to an energy or potential function represented by a quadratic function of the state coordinates. This relationship, together with the existence of dual sets of variables, forms the basis for an intrinsic analogue between electrical networks and other physical systems. The existence of a least-power theorem in electrical network theory was illustrated by Maxwell[8], and has also been considered by Jeans, Black and Southwell[9], and Ryder[10]. The theorem states simply that the minimum value of a power function corresponds to a current distribution in the network in accordance with Kirchoff's and Ohm's law. The rate of energy influx, as a quadratic function of the variables of the linear network, is minimised by solution of the network problem, and the absolute minimum corresponds to the unique solution of the equilibrium problem. A principle of least work related to strain energy, analogous to the least-power theorem in electrical network theory, also exists as a free variational problem in the analysis of equilibrium of statically loaded structures[11]. Similar concepts based on the variation of an energy state function were also used by Southwell for obtaining the approximate equilibrium solution of systems of linear equations based on relaxation techniques.

For the static physical system defined in terms of generalised coordinates (q_i, $i = 1..n$) the equilibrium states can be associated with the stationary value of a derived state-function[5]. Thus the virtual work of applied forces (Q_j) assumed as an exact differential of an energy (potential) function $V(q_1 .. q_n)$ can be stated in the vector form

$$dW = Q^t dq = -\left(\frac{\partial V}{\partial q}\right)^t dq \qquad (1)$$

The equilibrium states will then be determined by the conditions

$$Q = -\frac{\partial V}{\partial q} = 0 \qquad (2)$$

The existence of a potential function used in least-power theorems applied to physical systems is analogous mathematically to the quadratic form of residuals used in the methods of least-squares estimation[39]. In the electrical network analogue of least-squares theory, total power associated with the complete general network containing internal sources corresponds to the square of the residual error, and similar forms of solution exist for the network variables and for the 'best' estimate of the system states. Franksen[5] also shows that the optimal solution of a quadratic programming problem with linear constraints which satisfies the Kuhn-Tucker conditions are statements of Kirchoff's mesh law for meshes including an ideal rectifier. Kron[12] discusses the existence of an analogy between the theory of regression used in curve-fitting and the steady-state solution of electrical networks obtained by tearing and interconnection. A hypothetical intersection network is proposed as an estimating model from which regression coefficients for curve fitting a given function and also divided differences can be obtained by a process of tearing and piecewise solution of the network equations. The existence of the analogy and its relationship to the least-squares estimation problem outlined in Section 8.3 is now developed by considering the interconnection of a number of primitive networks represented by the voltage equations

$$V = ZJ \qquad (3)$$

where Z is a diagonal impedance matrix for the primitive network. The primitive elements are interconnected using the transformations of eqns 97 and 98 which give the general solutions for mesh currents and node-to-datum voltages of eqns 99 and 101 in Section 8.1. The constraints introduced with the topological relations ensure a unique solution to the network problem with positive definite conditions (ohmicness)[1,13]. These solutions are analogous to the least-square solution of eqn 133 of Section 8.3 for state estimation, with the measurement matrix H corresponding to the branch-mesh matrix C (and branch-node-pair matrix A), the weighting matrix V corresponding to the primitive impedance matrix Z (and admittance matrix Y), and the observed measurement vector y corresponding to currents i - Ye (and voltages e - Zi), and estimates $\hat{x}$ to currents i' (and voltages e'). The least-squares fitting of data may thus be considered as a process which introduces physical structure into the abstract problem which is analogous to the topological characteristics of an electrical network or linear graph.

By comparison with the least-squares formulation the network solution relates to the minimisation of the scalar norm of weighted squared residuals, or network power function

$$P_M = \|(YE - I) - Ci'\|_Z^2 = \|Ye\|_Z^2 = e^tYe = e'^t(A^tYA)e' \quad (4)$$

which is associated with the potential function V(q) of eqn 1. The solution of eqn 99 can thus be obtained by a minimisation of the scalar function P_M with respect to basic mesh currents i', thus illustrating the existence of a least-power or least-squares type theorem in the general electrical network problem. Similarly, the solution of eqn 101 in the network problem corresponds to the minimisation of a (dual) potential function

$$P_N = \|(ZI-E) - Ae'\|_Y^2 = i^tZi = i'^t(C^tZC)i' \quad (5)$$

with respect to node-to-datum voltages e'. The network mesh-current solution is associated with an 'estimation error' of the form

$$\begin{aligned} Ye &= [I_b - C(C^tZC)^{-1}C^tZ](YE - I) \\ &= Y[Z - ZC(C^tZC)^{-1}C^tZ](YE-I) = YM(YE - I) \quad (6) \end{aligned}$$

where I_b represents a b-dimensional unit matrix. Similarly, the voltage solution is associated with an analogous 'estimation error' of the form

$$\begin{aligned} Zi &= [I_b - A(A^tYA)^{-1}A^tY](ZI - E) \\ &= Z[Y - YA(A^tYA)^{-1}A^tY](ZI - E) = ZL(ZI - E) \quad (7) \end{aligned}$$

Properties of the dual matrices M, L are discussed in Section 8.1. The transformed matrices Z' and Y' of eqns 99 and 101 can also be identified with the inverse of the 'error' covariance matrix of eqn 137. The least-squares matrix H'VH may thus be given a physical interpretation in terms of an interconnected network, and the measurement noise covariances may be associated with the primitive admittance and impedance elements. The power functions combined with eqns 6 and 7 respectively may also be stated in the form

$$P_M = (YE - I)^t MYM(YE - I) = (YE - I)^t M(YE - I) \qquad (8)$$

$$P_N = (ZI - E)^t L (ZI - E) \qquad (9)$$

which illustrate the correspondence of the network weighting forms M and L with the least-squares loss function of Section 8.3.

The solution of the electrical network problem also introduces the concept of a minimum-norm generalised inverse. Thus, the form of the squared residuals $(I - AA^{\dagger})$ associated with a matrix A and its generalised inverse $A^{\dagger}$ exists in eqns 8 and 9 with the symmetrical weighting matrices given by

$$M = Z(I_b - LZ) = ZMY = Z(I_b - CT^t) \qquad (10)$$

$$L = Y(I_b - MY) = YLZ = Y(I_b - AR^t) \qquad (11)$$

Matrices T and R are defined in Section 8.1. Eqns 10 and 11 are thus associated with 'generalised inverses' of the form

$$C^{\dagger} = (C^tZC)^{-1}C^tZ = T^t \qquad (12)$$

$$A^{\dagger} = (A^tYA)^{-1}A^tY = R^t \qquad (13)$$

of orders mxb and pxb respectively, and matrices $(I_b - CT^t)$ and $(I_b - AR^t)$ are idempotent.

3. Network tearing and sequential least-squares estimation

Kron's method of tearing[3] simplifies the solution of the electrical network problem by interconnecting the equations of solution of smaller subnetworks with transformations based on connection matrix elements associated with the interconnecting links. The method decomposes the system equations according to the network topology, and essentially reduces the problem of matrix multiplication to the summation of vector and scalar products. With partitioning thus based on the network or graph structure the component solutions can be interconnected directly with other subsystems. Kron[14] also refers to the methods of regression and tearing liberating hidden internal constraint variables, or regression coefficients within a hypothetical 'intersection' system which acts as an estimating model for the unknown internal structure of the overall system. This analogy between methods of regression and network analysis is based essentially on the correspondence between the methods of solution as discussed in Section 2. The iterative procedures involved in Kron's method of tearing are now also shown to be directly analogous to the sequential algorithms developed for least-squares estimation, identification and control.

3.1 Network tearing and interconnection. The validity and proof of Kron's method of tearing and interconnection have been established by Roth using topological concepts[15,16,13]. Branin[17,18] gives a simple detailed explanation of Kron's method and develops a piecewise solution for the mesh and nodal methods of analysis. Branin also illustrates the technique based on elimination - backsubstitution, referred to by Roth as K-partitioning. The process of tearing has also been considered as a generalisation of Thevenin's and Norton's theorems based on the interconnection of a set of links[19], and Thevanin's theorem has been shown to be analogous to Gaussian elimination for the solution of linear equations[20]. Harrison[21], in a discussion of Kron's methods of tearing, also develops an iterative algorithm for obtaining an inverse admittance matrix in terms of previously known quanitities and added interconnection links.

The iterative procedures involved in Kron's method of tearing and piecewise analysis can be explained simply by considering the inverse nodal admittance matrix of eqn 121 in Section 8.1.5 in the form

$$(Y')^{-1} = (A_T^t Y_T A_T + A_L^t Y_L A_L)^{-1} = (Z_1^{-1} + A_L^t Y_L A_L)^{-1} \qquad (14)$$

The network is torn into m+1 separate subnetworks consisting of a selected tree and links. The effect of interconnecting the subnetworks and links is then obtained by updating the solution matrix in accordance with the link or network changes. The mesh impedance matrix may also be represented, using eqn 94, in the partitioned form

$$Z' = C^t Z C = C_T^t Z_T C_T + Z_L$$

$$= A_L (B_T^t Z_T B_T) A_L^t + Z_L = A_L Z_1 A_L^t + Z_L \qquad (15)$$

where Z_T represents the tree impedance matrix and Z_L the impedance matrix of the added m links. The form of eqn 121 may then be used for obtaining the overall nodal solution matrix. The impedance matrix Z_1 may represent any network which is to be interconnected to, or augmented with, a set of links specified by the matrices Z_L and A_L. Links may be added successively between tree branches or across sets of links to obtain a repeated updating. Thus, the nodal solution matrix after adding the ith set of links across nodes, without mutual coupling, is given by the augmented form of eqn 14,

$$(Y_i')^{-1} = (Y_{i-1}' + A_{Li}^t Y_{Li} A_{Li})^{-1} = (\bar{A}_i^t Y_i \bar{A}_i)^{-1} \quad (16)$$

where $$\bar{A}_i = \begin{bmatrix} A_T \\ A_{L1} \\ \vdots \\ A_{Li} \end{bmatrix} = \begin{bmatrix} \bar{A}_{i-1} \\ A_{Li} \end{bmatrix}$$

Link i is connected to the nodes associated with the ith branch, and the branch-node matrix A_{Li} defines the corresponding row components of the partitioned link matrix A_L. An iterative interconnection algorithm then follows from eqns 121 or 126

$$(Y_i')^{-1} = (Y_{i-1}')^{-1} - (Y_{i-1}')^{-1} A_{Li}^t [Y_{Li}^{-1} + A_{Li}(Y_{i-1}')^{-1} A_{Li}^t]^{-1} A_{Li}(Y_{i-1}')^{-1} \quad (17)$$

With the addition of single links only scalar inversion is required for updating the inverse pxp nodal admittance matrix compared with the MxM matrix inversion required when all links are connected simultaneously.

The effect of connecting an elemental change dy_i across a pair of nodes defined by the branch-node matrix row A_i may also be obtained from the additive form of eqn 14. The resulting differential change in the nodal solution matrix $(Y')^{-1} = Z_o$, with respect to changes in the branch admittance parameter, will then be given by

$$\begin{aligned} Z_o + dZ_o &= (A^t Y A + A_i^t \, dy_i \, A_i)^{-1} \\ &= Z_o - Z_o A_i^t [(dy_i)^{-1} + A_i Z_o A_i^t]^{-1} \, A_i Z_o \end{aligned} \quad (18)$$

requiring only scalar inversion with a single element change. With differential changes included on all branch elements

$$dZ_o = -Z_o A^t [(dY)^{-1} + A Z_o A^t]^{-1} \, A Z_o \quad (19)$$

Branin[17] considers the effect of a series change dz_i in the element z_i equivalent to the parallel addition of the link element $z_L = 1/dy_i$. The above relations apply similarly with $dy_i = -dz_i/(z_i^2 + z_i dz_i)$. The elemental change dY, together with changes in supply voltages and currents, will produce a differential change in the node-to-datum voltages of eqn 101, given by

$$de' = -(Y')^{-1}A^t(dY)A(Y')^{-1}A^t(I-YE) + (Y')^{-1}A^t[dI - Y(dE)-(dY)E] \quad (20)$$

$$\text{and} \quad de = Ade' = -M(dY)M(I-YE) + M[dI - Y(dE) - (dY)E] \quad (21)$$

Matrix M is defined by eqn 105 and has a significant role in both the electrical network and least-squares estimation problems. It is also associated with the properties

$$MYM = M , \qquad M(dY)M = -dM \quad (22)$$

$$(MY - I_b)A = 0 , \qquad M(dY)A = -(dM)YA \quad (23)$$

Similar relations follow for the partial derivative of the mesh solution matrix with respect to admittance changes of the ith branch, associated with the ith row of the branch-mesh connection matrix C_i. Eqns 20, 21 will have application in determining combinations of element changes required for reducing disturbance changes in the response variables.

The above method of piecewise analysis may be applied for obtaining the solution of any set of linear equations requiring the inverse of a coefficient matrix corresponding to Z' or Y'. If the system is partitioned or torn into constituent units in accordance with the system structure then the added components used for updating the previous-stage inverse solution will usually be of particularly simple form. The methods of tearing and piecewise interconnection appear to have been applied only in steady-state problems[17], although Kron[22] has discussed the possible tearing of a nonlinear block diagram system representation with the nonlinear subsystem equations solved using conventional iterative methods. The methods may also be extended for interconnecting the transient solutions of interconnected subsystems, and also for the possible sequential design of interacting multi-loop control systems as discussed in Sections 4 and 5.

3.2 <u>Sequential least-squares estimation</u>[23-29] The sequential processing of online data and the recursive updating of coefficients based on previous parameter estimates avoids recurrent computation with matrix inversion when additional observations are combined with all past data, thus reducing storage and computational requirements, particularly for estimation in the dynamic system. This can be achieved by imbedding the fundamental least-square solutions into the overall problem. Similar algorithms associated with the appearance of additional observations were developed by Gauss (1823) and Plackett (1950)[23], and the results appear inherently in the Kalman-Bucy filtering equations[30,31].

Now using the measurement process of eqn 131 and the solution of eqn 133 for the static system, the optimal estimate for min J_{i+1} when the total measurement vector $\bar{y}_i$ is augmented by a new measurement $y_{i+1} (= H_{i+1} x)$ at period i+1 will be given by

$$\hat{x}_{i+1} = (\bar{H}^t_{i+1} \bar{V}_{i+1} \bar{H}_{i+1})^{-1} \bar{H}^t_{i+1} \bar{V}_{i+1} \bar{y}_{i+1} \tag{24}$$

where $\bar{H}_{i+1} = \begin{bmatrix} \bar{H}_i \\ H_{i+1} \end{bmatrix}$, $\bar{V}_{i+1} = \begin{bmatrix} \bar{V}_i & 0 \\ 0 & V_{i+1} \end{bmatrix}$, $\bar{y}_{i+1} = \begin{bmatrix} \bar{y}_i \\ y_{i+1} \end{bmatrix}$

or $$\hat{x}_{i+1} = (\bar{H}^t_i \bar{V}_i \bar{H}_i + H^t_{i+1} V_{i+1} H_{i+1})^{-1} (\bar{H}^t_i \bar{V}_i \bar{y}_i + H^t_{i+1} V_{i+1} y_{i+1}) \tag{25}$$

Thus, in general, the optimal estimate $\hat{x}_{k+1}$ for min $\left\{ \sum_{i=o}^{k+1} \| y_i - H_i x_i \|^2_{V_i} \right\}$ based on the observation data sequence $\{y_o \cdot\cdot y_k, y_{k+1}\}$ is

$$\hat{x}_{k+1} = \left(\sum_{i=o}^{k} H^t_i V_i H_i + H^t_{k+1} V_{k+1} H_{k+1} \right)^{-1} \left(\sum_{i=o}^{k} H^t_i V_i y_i + H^t_{k+1} V_{k+1} y_{k+1} \right) \tag{26}$$

Now define $$P_k = \left[\sum_{i=o}^{k} (H^t_i V_i H_i) \right]^{-1} \tag{27}$$

$$P_{k+1} = [P_k^{-1} + H^t_{k+1} \, V_{k+1} \, H_{k+1}]^{-1} \tag{28}$$

and using the matrix inversion identity of eqn 126 gives

$$P_{k+1} = P_k - P_k H^t_{k+1} (V^{-1}_{k+1} + H_{k+1} P_k H^t_{k+1})^{-1} H_{k+1} P_k \tag{29}$$

Now combining eqns 26 and 28 and including the solution for $\hat{x}_k$ from the form of eqn 26 gives

$$\hat{x}_{k+1} = P_{k+1} (P_k^{-1} \, \hat{x}_k + H^t_{k+1} V_{k+1} y_{k+1}) \tag{30}$$

and then using eqn 28

$$\hat{x}_{k+1} = \hat{x}_k + P_{k+1} H^t_{k+1} V_{k+1} (y_{k+1} - H_{k+1} \hat{x}_k) \tag{31}$$

Eqns 29 and 31 represent the sequential algorithm for obtaining an updated estimate at stage k+1 based on the previous-stage estimate, the current-stage covariance matrix P_{k+1} and the additional data y_{k+1}. If y_{k+1} is processed as single data the nxn matrix inversion is reduced to the inversion of a scalar quantity. The estimate may also be specified in the form

$$\hat{x}_{k+1} = \hat{x}_k + P_k H^t_{k+1}(V^{-1}_{k+1} + H_{k+1}P_k H^t_{k+1})^{-1}(y_{k+1} - H_{k+1}\hat{x}_k) \quad (32)$$

The form of eqn 134 in the classical least-squares solution may also be identified in the sequential estimate by using eqn 131 in eqn 26 to give

$$\hat{x}_{k+1} = x + \sum_{i=o}^{k+1} B_i v_i \quad (33)$$

where $$B_i = [\sum_{j=o}^{k+1} (H^t_j V_j H_j)]^{-1} H^t_i V_i = P_{k+1} H^t_i V_i \quad (34)$$

and $$\sum_{i=o}^{k+1} B_i H_i = I_n \quad (35)$$

Matrix B_{k+1}, defined by eqn 34, represents a weighting or filter-gain matrix in eqn 31. Also from eqns 28 and 34

$$P_{k+1}P^{-1}_k + B_{k+1}H_{k+1} = I_n \quad (36)$$

From eqn 33 the covariance-of-error matrix

$$P_{k+1} = E[(x - \hat{x}_{k+1})(x - \hat{x}_{k+1})^t] = E[(\sum_{i=o}^{k+1} B_i v_i)(\sum_{i=o}^{k+1} B_i v_i)^t] \quad (37)$$

and with an independent inter-sample noise sequence

$$P_{k+1} = E[\sum_{i=o}^{k+1} B_i v_i v^t_i B^t_i] = \sum_{i=o}^{k+1} B_i R_i B^t_i \quad (38)$$

which is similar to the form of eqn 136.

Matrices $H_i B_i$ are analogous to the linear transformation of eqn 130. The sequential algorithm of eqn 31 may be considered as a descent scheme operating to reduce the instantaneous residue between the new measurement y_{k+1} and the expected value $\hat{y}_{k+1} = H_{k+1}\hat{x}_k$, with the weighting matrix B_{k+1} transforming or apportioning the residue to the new improved estimate.

The formulation of the dynamic multistage estimation problem is basically similar to that of the single-stage static problem, with the state changing according to the state differential or difference equations[29].

The estimate of the state of the linear dynamic system based on measurements corrupted by additive Gaussian noise are given by the Kalman-Bucy filter equations[30,31], which reduce, for the static system, to the original form of Plackett's equations[23]. The form of the sequential optimal filter is of fundamental importance in problems of estimation, filtering, prediction and identification. Sequential methods of statistical decision and estimation theory also have applications in problems of pattern recognition and machine learning[32].

3.3 Analogy of network tearing with sequential least-squares estimation

Kron's method of tearing and interconnection for solution of the electrical network problem can be shown to be analogous to the techniques of sequential least-squares estimation, identification and control, with the addition of link elements corresponding to the augmenting of the least-squares solutions with a discrete data sequence. The analogy is illustrated by a comparison of eqns 17 and 29 representing the updated nodal solution matrix $(Y_i')^{-1}$ and the discrete-time covariance matrix P_k respectively. The measurement matrix H_k and weighting matrix V_k in the least-squares solution are seen to correspond with the branch-node matrix A_{Li} and admittance of the added links Y_{Li} respectively.

Augmenting the least-squares solution with the data sequence y_k and weighting V_k thus corresponds to the addition of admittance components Y_{Lk} and link voltages $e_{Lk}(= (e-Zi)_k)$ in the solution for the node-to-datum voltages of eqn 101. Thus

$$e_k' = (A_T^t Y_T A_T + \sum_{i=1}^{k-1} A_{Li}^t Y_{Li} A_{Li} + A_{Lk}^t Y_{Lk} A_{Lk})^{-1} (A_T^t Y_T e_T + \sum_{i=1}^{k-1} A_{Li}^t Y_{Li} e_{Li} + A_{Lk}^t Y_{Lk} e_{Lk}) \qquad (39)$$

The identity of eqn 28 also corresponds to the form of eqn 16. However, the analogy of convergence inherent in the least-squares problem does not extend to the electrical network problem, in which the process is terminated with the addition of the final link element. An updated-type solution may also be obtained for the node-to-datum voltages. Thus, proceeding as in the sequential least-squares formulation, and combining eqn 39 for the solution of e_k' and e_{k-1}' and eqn 16 gives

$$e_k' = (Y_k')^{-1} (Y_{k-1}' e_{k-1}' + A_{Lk}^t Y_{Lk} e_{Lk}) \qquad (40)$$

and then using the form of eqn 16 gives

$$e'_k = e'_{k-1} + (Y'_k)^{-1} A^t_{Lk} Y_{Lk}(e_{Lk} - A_{Lk} e'_{k-1}) \qquad (41)$$

Eqns 17 and 41 now give a sequential algorithm for updating the node-to-datum voltages at stage k based on the previous-stage 'estimate', the current-stage nodal solution matrix and the link voltages e_{Lk} corresponding to an additional data sequence. With the processing of individual links matrix inversion is reduced to the inversion of a scalar quantity. Similarly, the 'estimate' of eqn 41 may also be specified in the form

$$e'_k = e'_{k-1} + (Y'_{k-1})^{-1} A^t_{Lk}[Y^{-1}_{Lk} + A_{Lk}(Y'_{k-1})^{-1} A^t_{Lk}]^{-1}(e_{Lk} - A_{Lk} e'_{k-1}) \qquad (42)$$

In the sequential solution the a priori and current-stage admittance matrices summate in the form of eqn 16, and represent the link admittances entering the connected network at every stage. It is interesting to note the 'measurement' process included in eqns 41 and 42 corresponding to the form of the voltage relation of eqn 97. A filter-gain-type matrix corresponding with eqn 34 may also be identified in the network problem. Thus from eqn 39 with A_j representing the row components of matrix A

$$e'_k = (Y'_k)^{-1} (\sum_j A^t_j Y_j e_j) = \sum_j B_j e_j \qquad (43)$$

where $$B_j = (Y'_k)^{-1} A^t_j Y_j \qquad (44)$$

and $$\sum_j B_j A_j = I_p \qquad (45)$$

The matrix B_j thus represents a weighting or filter-gain matrix which transforms a voltage difference to the new 'estimate' and possesses properties similar to those in the least-squares formulation.

The relations developed for sequential least-squares estimation thus extend for application to the piecewise solution of the electrical network problem. Also by analogy with the piecewise solution of the network problem, the sequential least-squares estimator may be given a physical interpretation in terms of the addition of structural elements with weighting V_k in a linear graph using 'connection matrix' elements H_k. The piecewise solution of the network problem with the addition of link elements is not analogous to sequential estimation in the linear dynamic system involving a state transition. The methods, however, appear to be applicable for the piecewise analysis of the transient problem formulated in block diagram form or in terms of an interconnection of dynamic units in a linear graph.

4. Formulation of the multimachine power system problem

The formulation of the multimachine power system problem can be associated with a least-power theorem, and the overall dynamic problem may also be defined in terms of a sequential connection of generators at the network nodes which is analogous to the recursive solution for least-squares estimation. Such a formulation may have important applications for obtaining a piecewise transient solution of large machine systems and also in the design of multivariable control systems.

The interconnection of synchronous generators with the general equations of a multinode linear passive network is considered in a form suitable for the study of load-frequency control[33,34]. The equivalent network containing only generator nodes is represented by the n-node equations

$$i_N = Y_N v_N \tag{46}$$

where Y_N is a 2nx2n symmetrical matrix of equivalent driving-point and transfer admittances with real partitioned elements $\begin{bmatrix} g_{ij} & -b_{ij} \\ b_{ij} & g_{ij} \end{bmatrix}$. The generator node voltages and shunt currents to neutral will include the components $v_N = \begin{bmatrix} v_{Di} \\ v_{Qi} \end{bmatrix}$, $i_N = \begin{bmatrix} i_{Di} \\ i_{Qi} \end{bmatrix}$ with D, Q reference axes common to all nodes.

The synchronous machines are represented by Park's voltage equations referred to direct- and quadrature-axes in the machine rotor positions,

$$\begin{bmatrix} v_{di} \\ v_{qi} \end{bmatrix} = \begin{bmatrix} e_{di} \\ e_{qi} \end{bmatrix} - \begin{bmatrix} 0 & -x'_{qi} \\ x'_{di} & 0 \end{bmatrix} \begin{bmatrix} i_{di} \\ i_{qi} \end{bmatrix} \quad \text{or } v_i = e_i - Z_{Mi} i_i, \; i = 1 \ldots m \tag{47}$$

$$\begin{bmatrix} \dot{e}_{di} \\ \dot{e}_{qi} \end{bmatrix} = \begin{bmatrix} K_1^i & \cdot \\ \cdot & K_2^i \end{bmatrix} \begin{bmatrix} e_{di} \\ e_{qi} \end{bmatrix} + \begin{bmatrix} K_3^i & \cdot \\ \cdot & K_4^i \end{bmatrix} \begin{bmatrix} v_{di} \\ v_{qi} \end{bmatrix} + \begin{bmatrix} \cdot \\ K_5^i \end{bmatrix} e_{xi} \tag{48}$$

or $$\dot{e}_i = F_1^i e_i + F_2^i v_i + G^i e_{xi}$$

where $K_1 = -x_q/(\tau'_{qo} x'_q)$, $K_2 = -x_d/(\tau'_{do} x'_d)$, $K_5 = 1/\tau'_{do}$

$K_3 = (x_q - x'_q)/(\tau'_{qo} x'_q)$, $K_4 = (x_d - x'_d)/(\tau'_{do} x'_d)$

Synchronous machine output $\quad P_{ei} = v_i^t \, i_i$ (49)

For small changes of machine and system speed, the balance of synchronous, asynchronous, mechanical and inertial powers is represented by

$$M_i \, \dot{\omega}_{ti} = P_{ti} - P_{ei} - T_{di}\omega_{ti} \,, \qquad \dot{\delta}_i = \omega_{ti} \tag{50}$$

where constants M_i represent effective rotary inertias and T_{di} the per-unit damping-torque coefficients for small deviations.

Turbine-governor-valve control, with provision for speed resetting, is represented by

$$\dot{d}_{Gi} = -k_1^i d_{Gi} - k_2^i \omega_{ti} + k_3^i d_{si} \tag{51}$$

In the steady state each machine maintains a constant phase displacement of its internal voltage. With a system load change momentary changes in machine speeds will occur for repositioning the rotor load angles, and the system will settle at a maintained synchronous speed. With load angles δ_i between the quadrature field axes of each machine and the common D, Q network reference axes, machine i terminal voltage will be related to the j-node network voltage components by the rotational transformation

$$v_i = A(\delta_i)v_{Nj} \,, \quad A(\delta_i) = \begin{bmatrix} \cos\delta_i & \sin\delta_i \\ -\sin\delta_i & \cos\delta_i \end{bmatrix}, \quad v_{Nj} = \begin{bmatrix} v_{Dj} \\ v_{Qj} \end{bmatrix} \tag{52}$$

With m machines connected to n network nodes

$$v = A(\delta)v_N \tag{53}$$

The connection matrix $A(\delta)$ of order mxn, with $m \geq n$ in general, contains elements $A_{ij} = (e^{j\delta_i}, 0)$ if the ith machine is incident or not on the jth node, and corresponds to the tree-branch-node matrix A_T of Section 8.1.1 containing elements (+1, -1, 0). It describes the topology of interconnection and also the displacement of the machine axes from a common network reference frame. The machine and network currents are similarly related by

$$i_N = A^t(\delta)i \tag{54}$$

Eqns 46, 47, 53 and 54 defining the interconnected system may now be combined to give the network voltages in terms of internal machine voltages,

$$v_N = [Y_N + A^t(\delta)Y_M A(\delta)]^{-1} A^t(\delta)Y_M e \tag{55}$$

where $A^t(\delta)Y_M A(\delta)$ with real components is a 2nx2n diagonal matrix of 2x2 submatrices. The generator voltages may then be specified as

$$v = A(\delta)[Y_N + A^t(\delta)Y_M A(\delta)]^{-1}A^t(\delta)Y_M e = M(\delta)Y_M e \qquad (56)$$

$M(\delta)$ is a symmetrical mxm matrix and corresponds with the similar forms of eqns 105 and 130 in the network and least-square problems. Eqn 55 also corresponds to the solution for node-to-datum voltages in the general network analysis of Section 8.1. The m-individual machine relations of eqn 48 in diagonal form may now be interconnected with the transformation of eqn 56, giving

$$\dot{e} = [F_1 + F_2 M(\delta)Y_M]e + Ge_x \qquad (57)$$

Matrix $M(\delta)$ thus introduces interaction in association with the machine terminal voltage matrix F_2. The overall state-variable representation of the interconnected machine system, including component machine voltages e_{di}, e_{qi}, load angles δ_i, turbine speed change ω_{ti} and governor-valve position for the ith machine is then given by eqn 57 combined with the vector forms of eqns 50 and 51, and with

$$P_{ei} = v_i^t Y_{Mi}(e_i - v_i) \qquad (58)$$

Also $$P_e = \sum_{i=1}^{m} P_{ei} = v^t i = e'Y_M[I_m - M(\delta)Y_M]M(\delta)Y_M e \qquad (59)$$

The state derivatives are related to linear functions of the state variables and also to nonlinear relations involving implicit algebra associated with the network solution for a specified set of load angles and internal machine voltages. Control is applied by movement of the governor valves and speeder-motor positions d_{si} and by unit excitation e_{xi}. Integrated control may be investigated for restoring equilibrium conditions in some specified optimal manner following load disturbances[33].

The form of eqn 55 corresponds to the least-squares solution of eqn 138 incorporating a priori information concerning the unknown estimate. The solution for network voltages can thus be similarly related to the minimisation of a quadratic function

$$P = \left\| e - A v_N \right\|^2_{Y_M} + v_N^t Y_N v_N = i^t Z_M i + v_N^t Y_N v_N \qquad (60)$$

with respect to voltages v_N. The formulation of the multimachine power system problem thus includes inherently the form of a least-power theorem associated with a potential function related to machine power losses and power supplied to the network, which can be used to define the interconnected network voltages. Eqn 55 also includes a summation of admittance elements with the connection of generators to the network nodes which corresponds with the interconnection of link elements in Kron's method of tearing. The general methods of network analysis based on tearing and interconnection and the techniques of sequential least-squares estimation will thus extend for application to the piecewise solution of the multimachine power system problem.

4.1 Sequential formulation of the multimachine power system problem

The analysis of the high-order system model will introduce considerable computational difficulties which may be reduced by considering a piecewise interconnection of the dynamic units with the network structure. The solution for network voltages may be considered in piecewise form with a sequential connection of m generators represented by

$$(v_N)_m = [Y_N + \sum_{i=1}^{m} A_i^t(\delta_i)Y_{Mi}A_i(\delta_i)]^{-1}(\sum_{i=1}^{m} A_i^t(\delta_i)Y_{Mi}e_i) \qquad (61)$$

with the connection matrix $A(\delta)$ partitioned into rows

$$A(\delta) = \begin{bmatrix} A_1(\delta_1) \\ A_i(\delta_i) \\ A_m(\delta_m) \end{bmatrix} = \begin{bmatrix} \bar{A}_i(\delta_i) \\ \vdots \\ A_m(\delta_m) \end{bmatrix}.$$

Thus with the connection of the kth generator, the n-component network voltages,

$$(v_N)_k = (Y'_k)^{-1}[A_1^t(\delta_1)..A_k^t(\delta_k)] \begin{bmatrix} Y_{M1} & & \\ & \ddots & \\ & & Y_{Mk} \end{bmatrix} \begin{bmatrix} e_1 \\ \vdots \\ e_k \end{bmatrix}$$

$$= (Y'_k)^{-1}\,\bar{A}_k^t(\delta_k)\,\bar{Y}_{Mk}\,\bar{e}_k \qquad (62)$$

where

$$(Y'_k)^{-1} = [Y'_{k-1} + A_k^t(\delta_k)Y_{Mk}A_k(\delta_k)]^{-1}, \quad Y'_o = Y_N \qquad (63)$$

$$= (Y'_{k-1})^{-1} - (Y'_{k-1})^{-1}A_k^t(\delta_k)[Y_{Mk}^{-1} + A_k(\delta_k)(Y'_{k-1})^{-1}A_k^t(\delta_k)]^{-1} A_k(\delta_k)(Y'_{k-1})^{-1} \qquad (64)$$

Eqn 62 may also be considered in the summated form

$$(v_N)_k = (Y'_k)^{-1}[\sum_{i=1}^{k-1} A_i^t(\delta_i)Y_{Mi}e_i + A_k^t(\delta_k)Y_{Mk}e_k] \qquad (65)$$

and including the form for $(v_N)_{k-1}$ gives

$$(v_N)_k = (Y'_k)^{-1} [Y'_{k-1}(v_N)_{k-1} + A^t_k(\delta_k)Y_{Mk}e_k] \tag{66}$$

and including eqn 63 gives

$$(v_N)_k = (v_N)_{k-1} + (Y'_k)^{-1}A^t_k(\delta_k)Y_{Mk}[e_k - A_k(\delta_k)(v_N)_{k-1}] \tag{67}$$

Eqn 67 represents an updating of the network node voltages with the addition of a single machine defined in terms of the internal voltage and admittance matrix of the kth generator and the connection matrix row $A_k(\delta_k)$ which identifies the particular node connection. With the addition of the kth machine, the k generator terminal voltages from eqn 53 are

$$\bar{v}_k = \bar{A}_k(\delta)(v_N)_k \tag{68}$$

and these may be stated as a function of the generator internal voltages $(e_1 .. e_k)(=\bar{e}_k)$, or as a function of the previous-stage network voltage $(v_N)_{k-1}$ and the generator voltage e_k. Thus from eqn 62,

$$\bar{v}_k = \bar{A}_k(\delta)(Y'_k)^{-1}\bar{A}^t_k(\delta)\bar{Y}_{Mk}\bar{e}_k = M_k(\delta)\bar{Y}_{Mk}\bar{e}_k \tag{69}$$

or from eqn 67

$$\bar{v}_k = \bar{A}_k(\delta)(v_N)_{k-1} + \bar{A}_k(\delta)(Y'_k)^{-1}A^t_k(\delta_k)Y_{Mk}[e_k - A_k(\delta_k)(v_N)_{k-1}] \tag{70}$$

The kth machine power may then be obtained using eqn 58 and the kth row voltage component of eqn 69. The matrix $M_{k+1}(\delta)$ associated with k+1 connected generators may also be updated by

$$M_{k+1}(\delta) = \begin{bmatrix} \bar{A}_k(\delta) \\ A_{k+1}(\delta_{k+1}) \end{bmatrix} (Y'_{k+1})^{-1} [\bar{A}^t_k(\delta)A^t_{k+1}(\delta_{k+1})]$$

$$= \begin{bmatrix} \bar{A}_k(\delta)(Y'_{k+1})^{-1}\bar{A}^t_k(\delta), & \bar{A}_k(\delta)(Y'_{k+1})^{-1}A^t_{k+1}(\delta_{k+1}) \\ A_{k+1}(\delta_{k+1})(Y'_{k+1})^{-1}\bar{A}^t_k(\delta), & A_{k+1}(\delta_{k+1})(Y'_{k+1})^{-1}A^t_{k+1}(\delta_{k+1}) \end{bmatrix} \tag{71}$$

The dynamic analysis of the multimachine power system problem may now be considered in a piecewise form based on the sequential interconnection of generators 1..m. Eqn 71, with the updated $(Y'_{k+1})^{-1}$ will permit the nonlinear functions associated with the state derivatives of eqn 57 to be obtained piecewise during each stage of the integration process. The problem of repeated mxm complex matrix inversion is also reduced to complex scalar inversion in eqn 64. The decomposition of the multimachine system thus appears to be feasible under dynamic operating conditions using techniques similar to Kron's method of tearing. The formulation will have application in the analysis of general large-scale system models which can be defined in terms of dynamic units interconnected with the nodes of a linear graph. The methods will permit a detailed study of the effects of interaction and will extend for investigating integrated control system design on a sequential or multi-level-type basis. The methods also appear to have particular application for the tearing and interconnection of the transfer matrix representation of the multivariable control system.

5. A sequential formulation of the multivariable control problem

A piecewise analysis of the multi-loop control system illustrated in FIG. 1 may be considered using techniques analogous to the method of tearing and interconnection in the electrical network problem. The general problem includes forward and feedback controllers C(s) and K(s) respectively, controlling a process represented by the open-loop transfer matrix G(s). A set of transformed reference input variables r(s) specify the required behaviour of the output variables x(s).

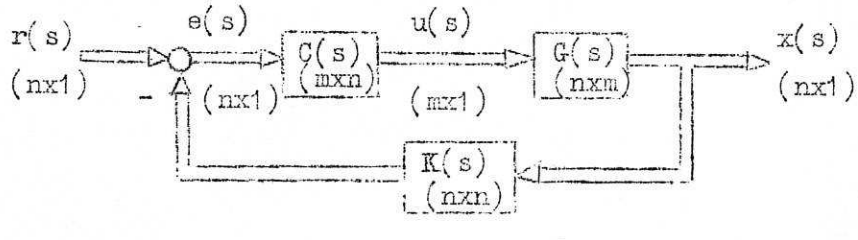


FIG. 1. Multivariable control system representation

With transformed inputs u(s) the open-loop behaviour of the process is represented by

$$x(s) = G(s)\, u(s) \tag{72}$$

The closed-loop performance of the system is then given by

$$x(s) = (I_n + GCK)^{-1} GCr(s) = G_o(s)r(s) \tag{73}$$

where $G_o(s)$ is the overall system transfer matrix.

It is of interest to note that, with the reference input $r(s)$ restricted to m components, and with the forward controller of order mxm and with the corresponding mxn feedback controller defined by $K(s) \equiv G^t(s)$, the solution for the closed-loop response can be associated with the least-squares estimation problem. Thus, considering a performance index

$$J(s) = \|e(s)\|^2_C + x^t(s)x(s)$$

$$= [r(s) - Kx(s)]^t C[r(s) - Kx(s)] + x^t(s)x(s) \qquad (74)$$

then $\partial J(s)/\partial x(s) = 0$ gives

$$x(s) = (I_n + K^t\hat{C}K)^{-1} K^t \hat{C}\, r(s)\ , \quad \hat{C} = (C + C^t)/2 \qquad (75)$$

Thus the closed-loop response of the particular transfer matrix control problem is analogous to a least-squares solution associated with a 'potential function' $J(s)$. The squared outputs in the multivariable control problem also correspond with the generator output powers in the multimachine problem. Thus by analogy with eqn 58

$$J_e = (Kx)^t C(r - Kx) = x^t K^t Ce \equiv x^t x \qquad (76)$$

The solution for network voltages in the multimachine power system problem given by eqn 55 may also be represented in the form of a multivariable feedback control system, as in FIG. 2.

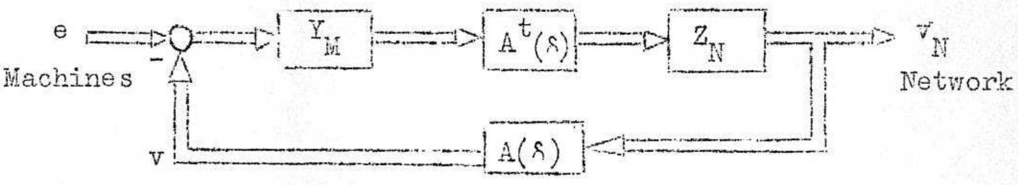


FIG. 2. Multimachine power system representation

The block diagram arrangement illustrates the transformation between the network-machine voltages in a feedback loop, and also the dual transformation between the machine admittance and network impedance matrices in the forward path. The least-squares estimate of eqn 138 may similarly be represented with matrix blocks R^{-1}, H^t and S in the forward path and with a feedback block representing the matrix H between the estimate $\hat{x}$ and the observation y. The piecewise solution of the machine problem may also be considered in the form of FIG. 3 associated with eqn 61 in the form

$$(v_N)_m = Z_N \left\{ \sum_{i=1}^{m} A_i^t(\delta_i) Y_{Mi} [e_i - A_i(\delta_i)(v_N)_m] \right\} \qquad (77)$$

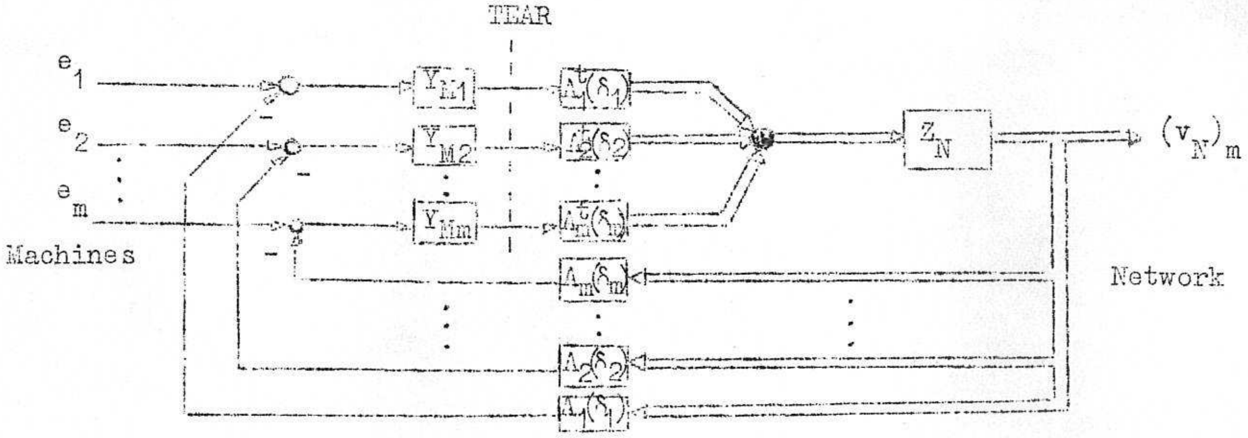


FIG. 3. Decomposition of multimachine power system problem

The piecewise analysis is based on the equivalent tearing of the system configuration, as shown, with a sequential interconnection of the individual machine blocks. The single-stage updating may also be illustrated as in FIG. 4 using a combined form of eqns 62 and 63 to give the equivalent single-loop feedback representation

$$(v_N)_k = (Y'_{k-1})^{-1} \left\{ \sum_{i=1}^{k-1} A_i^t(\delta_i) Y_{Mi} e_i + A_k^t(\delta_k) Y_{Mk} [e_k - A_k(\delta_k)(v_N)_k] \right\} \qquad (78)$$

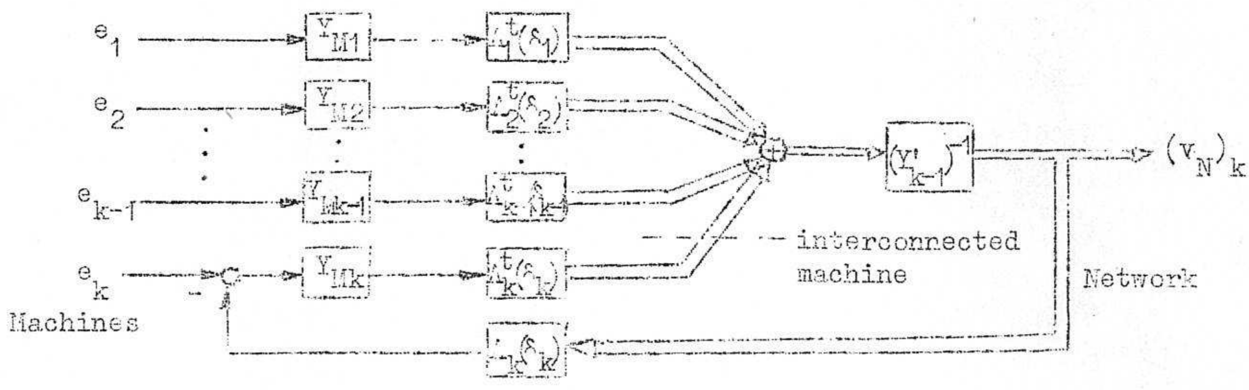


FIG. 4. Decomposition of multimachine power system problem with single-stage updating

The multivariable control system representation in the frequency domain is complicated by the inversion of transfer function matrices in eqn 73. These difficulties may be reduced by considering a piecewise analysis with a sequential connection of the transfer function control loops, analogous to the interconnection of the link admittances in the method of network tearing. This may simplify the design procedure for determining the system controllers, possibly by the application of conventional single-loop frequency domain techniques for compensation of the overall system in accordance with certain desired performance specifications. It will also lead to a more detailed understanding of the significance of interaction introduced by the individual control loops.

Now with a column and row partitioning of the matrices G and C, K respectively, defined by

$$G = [G_1, G_2 .. G_m] , \quad C = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_m \end{bmatrix} , \quad K = \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_n \end{bmatrix} \tag{79}$$

we can consider the effective piecewise addition of m control loops and specify the output states in the form

$$(x)_m = (G_o)_m r = (I_n + \sum_{i=1}^{m} \sum_{j=1}^{n} G_i C_{ij} K_j)^{-1} (\sum_{i=1}^{m} \sum_{j=1}^{n} G_i C_{ij} r_j) \tag{80}$$

The partitioned plant and forward controller transfer function components are now associated with a particular individual input, and the feedback controller elements with each of the measured output components. The previous methods of piecewise solution based on a decomposition of matrix products into interconnecting 'link' elements may now be used by considering eqn 80, associated with the addition of k inputs, in the form

$$(x)_k = [(Y'_k)_p]^{-1} (\sum_{i=1}^{k-1} \sum_{j=1}^{n} G_i C_{ij} r_j + \sum_{j=1}^{p} G_k C_{kj} r_j) \tag{81}$$

where p represents the pth output and

$$\begin{aligned} (Y'_k)_p &= (Y'_k)_{p-1} + G_k C_{kp} K_p , \quad p = 1..n \\ (Y'_1)_o &= I_n , \quad (Y'_k)_o = (Y'_{k-1})_n \end{aligned} \tag{82}$$

Then
$$[(Y'_k)_p]^{-1} = [(Y'_k)_{p-1}]^{-1} - [(Y'_k)_{p-1}]^{-1} G_k (C_{kp}^{-1} + K_p [(Y'_k)_{p-1}]^{-1} G_k)^{-1} K_p [(Y'_k)_{p-1}]^{-1} \tag{83}$$

The calculation for the overall system transfer matrix can now be simplified with matrix inversion reduced to the inversion of a single transfer function element in eqn 83. The decomposition of the control problem given by eqn 80 may also be represented in partitioned block diagram form similar to FIG. 3 for the multimachine problem. Such a structural representation illustrates, particularly, the physical significance of tearing and interconnection in the multivariable control problem. A piecewise representation incorporating a single feedback loop may also be obtained as in the multimachine problem by combining eqns 81 and 82 to give

$$(x)_k = (Y'_k)^{-1}_{p-1} \left\{ \sum_{i=1}^{k-1} \sum_{j=1}^{n} G_i C_{ij} r_j + \sum_{j=1}^{p-1} G_k C_{kj} r_j + G_k C_{kp} [r_p - K_p (x)_k] \right\} \tag{84}$$

Thus, for a system of dimension $n = 3$, $m = 2$ with $p = 2$, the interconnection of the single loop $(G_2 C_{22})$ corresponding to $k = 2$ may be represented as in FIG. 5.

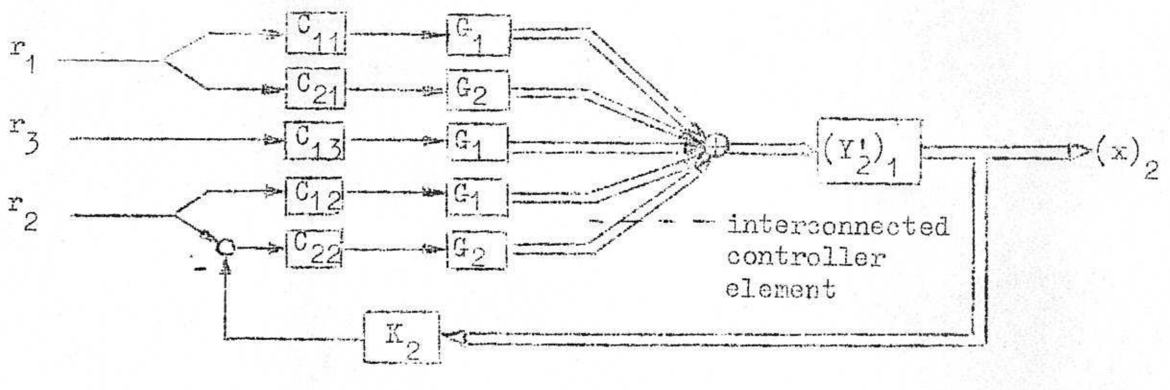


FIG. 5. Decomposition of multivariable control system problem with single-loop interconnection

The formulation of the piecewise solution will simplify with the system including a diagonal controller $C(s)$ of order mxm and with the feedback controller matrix $K(s)$ of order mxn associated with m reference inputs. The sequential addition of k control loops will then result in the transformed output states

$$(x)_k = (G_o)_k r = (I_n + \sum_{i=1}^{k} G_i C_{ii} K_i)^{-1} (\sum_{i=1}^{k} G_i C_{ii} r_i) \tag{85}$$

Eqn 85 is analogous to the piecewise solution for network voltages in the multimachine power system problem given by eqn 61, with the network admittance matrix (Y_N) corresponding to the unit matrix, the connection matrix $A(\delta)$ to the feedback controller and its transpose to the plant transfer matrix $G(s)$. The machine admittance matrix Y_M corresponds similarly to the forward diagonal controller $C(s)$, and the internal machine voltages e to the transformed input reference signals $r(s)$. The performance criterion of eqn 74 in the equivalent least-squares solution corresponds similarly with the machine system performance index of eqn 60. The 'topology' of the transfer matrix representation of the multivariable control problem may thus be defined in terms of the partitioned structure of the plant transfer function matrix $G(s)$ which interconnects the equivalent 'machine' parameters $C_{ii}(s)$ associated with the 'internal' inputs $r(s)$ with the equivalent unit matrix 'network'. A comparison of the corresponding block diagrams and of the piecewise solutions illustrates the particular difficulties of the general multivariable feedback control problem which will include off-diagonal forward controller terms which do not appear to exist in the machine-network problem.

The piecewise solution of the multivariable control problem will permit the effects of individual inputs produced by the addition of the columns of the plant matrix $G(s)$ to be studied. Thus, by analogy with the solution of the multimachine problem for network voltages v_N, the n-component transformed output states with the addition of k inputs associated with the plant transfer matrix of order nxk will be given by

$$(x)_k = (Y'_k)^{-1}[G_1 \cdot \cdot G_k] \begin{bmatrix} C_{11} & & \\ & \ddots & \\ & & C_{kk} \end{bmatrix} \begin{bmatrix} r_1 \\ \vdots \\ r_k \end{bmatrix} \qquad (86)$$

where $$(Y'_k)^{-1} = (Y'_{k-1} + G_k C_{kk} K_k)^{-1} , \quad Y'_o = I_n \qquad (87)$$

Then from the form of eqn 67

$$(x)_k = (x)_{k-1} + (Y'_k)^{-1} \, G_k C_{kk}[r_k - K_k(x)_{k-1}] \qquad (88)$$

Also by analogy with the generator voltages of eqns 68 and 69 the measured feedback variables with the addition of k loops may be obtained in the form

$$(Kx)_k = \begin{bmatrix} K_1 \\ \vdots \\ K_k \end{bmatrix} (x)_k = \bar{K}_k(x)_k \qquad (89)$$

$$= \bar{K}_k (Y'_k)^{-1} \bar{G}_k \, \bar{C}_{kk} \, \bar{r}_k \;=\; M_k \bar{C}_{kk} \bar{r}_k \qquad (90)$$

or from eqn 70

$$(Kx)_k \;=\; \bar{K}_k (x)_{k-1} + \bar{K}_k (Y'_k)^{-1} G_k C_{kk} [r_k - K_k (x)_{k-1}] \qquad (91)$$

The matrix M_k may also be obtained in updated form as in eqn 71 and possesses properties similar to those of the M-matrices in the network and least-squares problems.

Control design may now proceed by considering the cumulative effects of the individual feedback controller loops and the forward controller elements for desired overall performance of the output variables. Such a method of analysis with a piecewise connection of controller elements is directly analogous to Kron's method of tearing and interconnection which has found important application for solution of the electrical network and other large-scale system problems. It reduces the computational problem of inverting relatively high-order transfer matrices to the scalar inversion of single transfer function elements by effectively decomposing matrix products of transfer functions into a sum of outer vector components. The method illustrates particularly the interconnection of the system elements with an underlying topological structure and can be related to a dynamic programming type of algorithm for solution of the linear optimal control problem[29]. The particular form of piecewise analysis will not provide a direct solution for control design but it will give greater insight into the process of trial-and-error design, with the effects of individual controller elements summated into the overall desired solution. It can also form the basis for design based on a sensitivity analysis, as in the network problem of Section 3.1, which may be used for determining the changes produced in the overall system response or in the input-output coupling, with differential changes in the forward controller elements.

6. Conclusions

Fundamental concepts based on the elemental principles of algebraic topology originating in electrical network theory, and the existence of a theorem of least power establish the basis for a direct analogy between the performance of the electrical network and other physical systems. A unifying mathematical basis is shown to exist in classical least-squares estimation theory with the quadratic form of residuals corresponding to a potential function in the physical system. The steady state network solution based on Kron's method of tearing and interconnection is also shown to be directly analogous to the methods developed for sequential least-squares data fitting. The least-squares problem may thus be given a physical interpretation in terms of an interconnected network, and may be considered as a process which introduces physical structure into the abstract formulation.

Similar methods of analysis have been applied in the multi-machine power system problem based on a sequential interconnection of generators. A piecewise solution of the transient machine problem is developed using a process of tearing and interconnection which has previously been applied only to the steady-state electrical network problem. This development based on the topological concepts of electrical network theory could form the basis for the decomposition and piecewise solution of many large-scale system problems, such as encountered in the modelling of economic systems and company operations. A decomposition of the multivariable control system problem has also been formulated using an interconnection of control loop elements as in Kron's methods of network tearing. Such methods of piecewise analysis and subdivision will provide a greater insight into system structure and could form the basis for a simplified sequential design of multi-loop control systems or for the design of hierarchical multilevel control.

7. References

1. BRANIN, F.H.: 'The algebraic-topological basis for network analogies and the vector calculus', Procs. of sympm. on generalised networks, NY 1966, Brooklyn Polytech. Press, pp.453-491.

2. HARRINGTON, R.F.: 'Generalised network parameters in field theory', ibid., pp.51-67.

3. KRON, G.: 'A set of principles to interconnect the solutions of physical systems', Jl. of Applied Physics, 1953, 24, 8, pp.965-980.

4. FIRESTONE, F.A.: 'A new analogy between mechanical and electrical systems', Jl. of the Acoustical Soc., 1933, Jan., pp.249-267.

5. FRANKSEN, O.I.: 'Mathematical programming in economics by physical analogies', Parts I-III, Simulation, 1969, June, pp.297-314, July, pp.25-42, Aug., pp.63-87.

6. FRANKSEN, O.I.: 'Kron's method of tearing', Fourth power industry computer application conference, Florida, 1965.

7. TRENT, H.M.: 'Isomorphisms between oriented linear graphs and lumped physical systems', Jl. of the Acoustical Soc. of America, 1955, 27, 3, pp.500-527.

8. MAXWELL, J.C.: 'A treatise on electricity and magnetism', third editn., Dover, 1954, p.407.

9. BLACK, A.N., and SOUTHWELL, R.V.: 'Relaxation methods applied to engineering problems', Procs. of Royal Soc., Series A, 1938, 164, pp.447-467.

10. RYDER, F.L.: 'Network analysis by least power theorems', Jl. of the Franklin Inst., 1952, 254, pp.47-60.

11. TIMOSHENKO, S.: 'Strength of materials', Part I, 3rd editn., 1955, p.340.

12. KRON, G.: 'Multidimensional curve-fitting with self-organizing automata', Jl. of Mathl. Analysis and Applicns., 1962, 5, pp.46-69.

13. ROTH, J.P.: 'An application of algebraic topology: Kron's method of tearing', Quarterly of Appd. Maths., 1959, XVII, 1, pp.1-24.

14. KRON, G.: 'Basic concepts of multidimensional space filters', AIEE, Commns. and Elects., 1959, 45, pp.554-561.

15. ROTH, J.P.: 'An application of algebraic topology to numerical analysis: on the existence of a solution to the network problem', Proc. Nat. Acad. Sci., 1955, 41, pp.518-521.

16. ROTH, J.P.: 'The validity of Kron's method of tearing', ibid., pp. 599-600.

17. BRANIN, F.H.: 'The relation between Kron's method and the classical methods of network analysis', IRE WESCON Convention Record, Part 2 - circuit theory, 1959, pp.3-28.

18. BRANIN, F.H.: 'Computer methods of network analysis', Proc. IEEE, 1967, 55, 11, pp.1787-1801.

19. CHU, C-M., and La RUE, J.J.: 'Generalised circuits for magnetic ionic media', Procs. of sympm. on generalised networks, NY, 1966, Brooklyn Polytech. Press, pp.633-652.

20. CHEN, C.F., and HAAS, I.J.: 'An electrical interpretation of the algorithm of Gauss', IRE Trans. on circuit theory, 1962, 9, pp.298-299.

21. HARRISON, B.K.: 'A discussion of some mathematical techniques used in Kron's method of tearing', Jl. Soc. Indust. Appl. Math., 1963, 11, 2, pp.258-280.

22. KRON, G.: 'Nonlinear diakoptics and the optimization of dynamic systems', The Matrix and Tensor Quarterly, 1964, 14, 4, pp.127-131.

23. PLACKETT, R.L.: 'Some theorems in least squares', Biometrika, 1950, 37, pp.149-157.

24. SAGE, A.P., and MASTERS, G.W.: 'Least-squares curve fitting and discrete optimum fitting', IEEE Trans. Educ., 1967, 10, 1, pp.29-36.

25. AOKI, M.: 'Optimization of stochastic systems', Acad. Press, 1967.

26. LUENBERGER, D.G.: 'Optimization by vector space methods', Wiley, 1969.

27. DeRUSSO, P.M., ROY, R.J., and CLOSE, C.M.: 'State variables for engineers', Wiley, 1965.

28. LEE, R.C.K.: 'Optimal estimation, identification, and control', MIT Press, 1964.

29. NICHOLSON, H.: 'Sequential least-squares estimation, identification, reduction and control', Parts 1 and 2, Dept. of Control Eng., Univ. of Sheffield, research report CE69-2, 3, 1969.

30. KALMAN, R.E.: 'A new approach to linear filtering and prediction problems', Trans. Amer. Soc. Mech. Engrs., 1960, 82, D, pp.35-45.

31. KALMAN, R.E., and BUCY, R.S.: 'New results in linear filtering and prediction theory', ibid., 1961, 83, pp.95-108.

32. FU, K.S.: 'Sequential methods in pattern recognition and machine learning', Acad. Press, 1968.

33. NICHOLSON, H.: 'Dynamic optimisation of a multimachine power system', Proc. IEE, 1966, 113, 5, pp.881-894.

34. TAYLOR, D.G.: 'Analysis of synchronous machines connected to power-system networks', ibid., 1962, 109C, pp.606-610.

35. SYNGE, J.L.: 'The fundamental theorem of electrical networks', Quarterly of Appd. Maths., 1951, IX, 2, pp.113-127.

36. SPILLERS, W.R.: 'On diakoptics: tearing an arbitrary system', ibid., 1965, XXIII, 2, pp.188-190.

37. GRABBE, E.M., RAMO, S., and WOOLRIDGE, D.E.: 'Handbook of automation, computation, and control', vol.1, Wiley, 1958.

38. HOUSEHOLDER, A.S.: 'Principles of numerical analysis', McGraw-Hill, NY, 1953, p.79.

39. SU, H-L.: 'Electric analog for theory of adjustment and regression', Jl. Maths. and Physics, 1959-60, 38, pp.312-326.

8. Appendix

8.1 The electrical network problem. The general electrical network is considered with interconnected branches consisting of linear passive self impedance or admittance elements and ideal voltage and current sources as in FIG. 6[17].

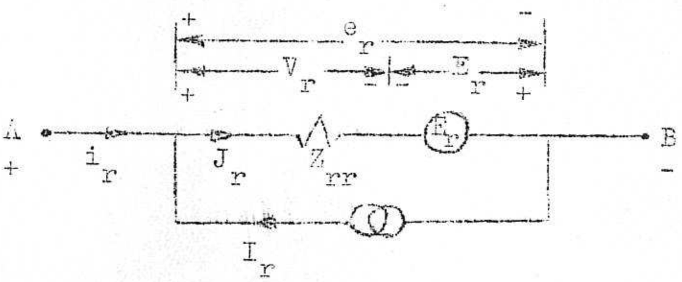


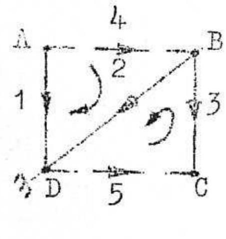
FIG. 6. rth network branch

The branch currents and voltages in the unconnected (primitive) network are represented by the b-dimensional vector equations

$$E + e = Z(I + i), \quad I + i = Y(E + e) \quad \text{or } V = ZJ, \; J = YV \qquad (92)$$

where the mutually reciprocal and symmetrical matrices Z and Y represent the bxb dimensional primitive network impedance and admittance matrices respectively. The equations representing the primitive network in terms of all branch variables are linearly dependent.

8.1.1 Topological structure[1,17,18,21,35] The topological structure of the connected network or linear graph is specified in terms of connection matrices $[\epsilon_{kj}]$, where ϵ_{kj} is +1, −1 or 0 depending on the directed branch-node-mesh relationships. These may be illustrated by reference to the graph of FIG. 7(a).



Node: A B C D; Branch: 1–5

$$\bar{A} = \begin{array}{c|cccc} & A & B & C & D \\ \hline 1 & 1 & \cdot & \cdot & -1 \\ 2 & \cdot & 1 & \cdot & -1 \\ 3 & \cdot & 1 & -1 & \cdot \\ 4 & 1 & -1 & \cdot & \cdot \\ 5 & \cdot & \cdot & -1 & 1 \end{array}$$

$$A = \begin{array}{c|ccc|c} & A & B & C & \\ \hline 1 & 1 & \cdot & \cdot & \\ 2 & \cdot & 1 & \cdot & A_T \\ 3 & \cdot & 1 & -1 & \\ \hline 4 & 1 & -1 & \cdot & A_L \\ 5 & \cdot & \cdot & -1 & \end{array}$$

$$B_T = \begin{array}{c|ccc} & A & B & C \\ \hline 1 & 1 & \cdot & \cdot \\ 2 & \cdot & 1 & 1 \\ 3 & \cdot & \cdot & -1 \end{array}$$

$$C = \begin{array}{c|cc|c} & 4 & 5 & \\ \hline 1 & -1 & \cdot & \\ 2 & 1 & 1 & C_T \\ 3 & \cdot & -1 & \\ \hline 4 & 1 & \cdot & C_L \\ 5 & \cdot & 1 & \end{array}$$

(Meshes 4 5)

FIG.7(a)-Lineargraph (b)-$\bar{A}$ matrix (c)-A matrix (d)-B_T matrix (e)-C matrix

$\bar{A}$ is the branch-node incidence matrix with linearly dependent columns. A is the branch-node-pair matrix (with the datum node column deleted), with submatrices A_T, A_L referring to tree branches (containing all the nodes and no meshes) and links (shown dotted) respectively. B_T is the node-to-datum path matrix with elements $\epsilon_{kj} = (+1, -1, 0)$ if the kth branch is (positively, negatively, not) included in the jth node-to-datum tree path. The branch-mesh matrix C contains elements $\epsilon_{kj} = (+1, -1, 0)$ if the kth branch is (positively, negatively, not) included in the jth basic mesh obtained by adding links to a tree. The columns of C indicate the branches included in each mesh, and with grouping of the basic meshes with the defining links, the submatrix C_L referring to the links is a unit matrix. For any linear graph the connection matrices are related by

$$A^tC = 0 , \quad C^tA = 0 \tag{93}$$

and partitioning gives

$$A_T^tC_T + A_L^t = 0 \text{ or } C_T = -B_TA_L^t \text{ with } A_T^{-1} = B_T^t \tag{94}$$

Kron refers to eqn 93 as the 'orthogonality condition' of node-pair potentials and mesh emf's[15]. If the linear graph contains b branches, n nodes, m meshes, p node pairs, M basic meshes and k sub-graphs, then

$$b = m+p , \quad p = n-k , \quad M = b-n+1 \tag{95}$$

Matrices $\bar{A}$, A and C are of dimensions bxn, bxp and bxm respectively, and rank A = p, rank C = m. The tree uniquely defines p open paths with p independent across variables, and the links uniquely define m closed paths with m independent through variables[6].

8.1.2 <u>Algebraic structure.</u> The linear graph or network imposes constraints on the branch variables defined by Kirchoff's voltage and current laws represented by

$$C^te = 0 , \quad A^ti = 0 \tag{96}$$

These relations sum the branch voltages around each basic mesh, and all the branch currents leaving each node, and impose m and p constraints respectively. The interconnection of the primitive elements introduces a transformation between the branch voltages and currents e, i ('old' variables) and the node-to-datum voltages e' and the currents in the basic meshes i' ('new' variables), given by

$$e = Ae' , \quad i = Ci' \tag{97}$$

The variables e', i' represent linearly independent sets and constitute a 'basis' for determining the branch variables e, i[18]. In eqn 97 only p branch voltages and m mesh currents are linearly independent. With additional quantities E, I assigned to the unconnected branches the equivalent induced mesh voltage and nodal current sources are given by

$$E' = C^t E , \qquad I' = A^t I \qquad (98)$$

The electrical network problem is stated - Given a connected network defined by the topological matrices A and C and the primitive impedance matrix Z (or Y), and given the arbitrary source vectors E and I, determine the branch voltages and currents e, i.

8.1.3 Mesh solution. Combining the voltage relations of eqns 92 and 96 and the current relation of eqn 97 for eliminating e gives the general solution for mesh currents

$$i' = (Z')^{-1}C^t Z(YE-I) , \quad Z' = C^t ZC , \quad |Z'| \neq 0 \qquad (99)$$

where Z' is the mesh impedance matrix. Branch currents i will then be given by eqn 97, and from eqns 92 and 99,

$$e = [Z - ZC(Z')^{-1}C^t Z](I - YE) \qquad (100)$$

The symmetrical matrix $C(Z')^{-1}C^t$ (=L) is the branch admittance matrix (of driving point and transfer admittances) relating branch currents i to the equivalent branch voltage sources (E-ZI). The symmetrical matrix $[Z - ZC(Z')^{-1}C^t Z]$ (=M) is similarly the branch impedance matrix.

8.1.4 Node solution. Combining the current relations of eqns 92 and 96 and the voltage relation of eqn 97 for eliminating i gives the solution for node-to-datum voltages

$$e' = (Y')^{-1}A^t Y(ZI - E) , \quad Y' = A^t YA \qquad (101)$$

where Y' is the nodal admittance matrix. The branch voltages e will be given by eqn 97, and from eqns 92 and 101,

$$i = [Y - YA(Y')^{-1}A^t Y] (E - ZI) \qquad (102)$$

Contributions to the coil variables J, V from the assumed given E', I' can be obtained from eqns 92, 97, 99 and 101 in the form

$$\begin{bmatrix} V \\ J \end{bmatrix} = \begin{bmatrix} ZC(Z')^{-1} & A(Y')^{-1} \\ C(Z')^{-1} & YA(Y')^{-1} \end{bmatrix} \begin{bmatrix} E' \\ I' \end{bmatrix} = \begin{bmatrix} T & ZR \\ YT & R \end{bmatrix} \begin{bmatrix} E' \\ I' \end{bmatrix} \qquad (103)$$

Using the identities obtained from eqns 99, 102 and 100, 101 gives

$$Y - YA(Y')^{-1}A^{t}Y = C(Z')^{-1}C^{t} = L \tag{104}$$

$$Z - ZC(Z')^{-1}C^{t}Z = A(Y')^{-1}A^{t} = M \tag{105}$$

Then $$Y - YPY = L , \qquad Z - ZLZ = M \tag{106}$$

and with matrices R and T defined by eqn 103, eqns 104, 105 give

$$I_b = AR^{t} + TC^{t} = MY + ZL \tag{107}$$

Also $$(Y')^{-1} = R^{t}MR , \qquad (Z')^{-1} = T^{t}LT \tag{108}$$

Harrison[21] derives eqns 106-108, associated with the topological characteristics of the linear graph, using the rank and algebraic properties of matrices equivalent to A and C. For the network problem, with mutually reciprocal matrices Z and Y, a simpler derivation follows directly as above using the mesh and nodal solution equations.

8.1.5 Kron's 'orthogonal network' solution[6,17] Kron defines the network problem using both mesh and nodal representations with nonsingular transformations and derives formulas for interconnecting solutions based on an 'orthogonal network' concept. For purposes of interconnection the impressed sources are confined to the links and node-to-datum paths of the tree by the relations

$$E = \begin{bmatrix} E_T \\ E_L \end{bmatrix} = \begin{bmatrix} 0 \\ U_L \end{bmatrix} E' , \qquad I = \begin{bmatrix} I_T \\ I_L \end{bmatrix} = \begin{bmatrix} B_T \\ 0 \end{bmatrix} I' \tag{109}$$

Then coil variables, including the constraints of eqn 97,

$$V = e + E = \begin{bmatrix} A_T & 0 \\ A_L & U_L \end{bmatrix} \begin{bmatrix} e' \\ E' \end{bmatrix} = Q \begin{bmatrix} e' \\ E' \end{bmatrix} = ZJ \tag{110}$$

$$J = I + i = \begin{bmatrix} B_T & C_T \\ 0 & C_L \end{bmatrix} \begin{bmatrix} I' \\ i' \end{bmatrix} = S \begin{bmatrix} I' \\ i' \end{bmatrix} = YV \tag{111}$$

giving $$\begin{bmatrix} e' \\ E' \end{bmatrix} = Q^{-1}ZS \begin{bmatrix} I' \\ i' \end{bmatrix} = S^{t}ZS \begin{bmatrix} I' \\ i' \end{bmatrix} = \bar{Z} \begin{bmatrix} I' \\ i' \end{bmatrix} \tag{112}$$

$$\begin{bmatrix} I' \\ i' \end{bmatrix} = S^{-1}YQ \begin{bmatrix} e' \\ E' \end{bmatrix} = Q^{t}YQ \begin{bmatrix} e' \\ E' \end{bmatrix} = \bar{Y} \begin{bmatrix} e' \\ E' \end{bmatrix} \tag{113}$$

Now assuming $$Z = \begin{bmatrix} Z_T & 0 \\ 0 & Z_L \end{bmatrix}, \quad Y = \begin{bmatrix} Y_T & 0 \\ 0 & Y_L \end{bmatrix} \qquad (114)$$

and using eqn 94,

$$\bar{Z} = \begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix} = S^tZS = \begin{bmatrix} B_T^tZ_TB_T \;, & -Z_1A_L^t \\ -A_LZ_1 \;, & A_LZ_1A_L^t+Z_L \end{bmatrix} = \begin{bmatrix} B_T^tZ_TB_T \;, & -Z_1A_L^t \\ -A_LZ_1 \;, & C^tZC \end{bmatrix} \qquad (115)$$

$$\bar{Y} = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix} = Q^tYQ = \begin{bmatrix} A_T^tY_TA_T+A_L^tY_LA_L, & A_L^tY_L \\ Y_LA_L & Y_L \end{bmatrix} = \begin{bmatrix} A^tYA \;, & A_L^t Y_L \\ Y_LA_L \;, & Y_L \end{bmatrix} \qquad (116)$$

Eqns 112 and 113 can also be solved in terms of the partitioned components of $\bar{Z}$ and $\bar{Y}$ to give

$$i' = Z_4^{-1}(E' - Z_3I') \;, \quad e' = (Z_1 - Z_2Z_4^{-1}Z_3)I' + Z_2Z_4^{-1}E' \qquad (117)$$

$$e' = Y_1^{-1}(I' - Y_2E') \;, \quad i' = (Y_4 - Y_3Y_1^{-1}Y_2)E' + Y_3Y_1^{-1}I' \qquad (118)$$

Eqns 117 and 118 form the basis of Kron's work on network tearing and interconnection. Now setting $E = 0$ and comparing eqn 117 with 101 gives the nodal solution matrix

$$(Y')^{-1} = Z_1 - Z_2Z_4^{-1}Z_3 \qquad (119)$$

Similarly, setting $I = 0$ and comparing eqns 118 and 99 gives

$$(Z')^{-1} = Y_4 - Y_3Y_1^{-1}Y_2 \qquad (120)$$

Then from eqns 115, 116, 118 and 119 and using eqn 94

$$(Y')^{-1} = Z_1 - Z_1A_L^t[A_LZ_1A_L^t + Z_L]^{-1}A_LZ_1 \;, \quad Z_1 = B_T^tZ_TB_T \qquad (121)$$

$$(Z')^{-1} = Y_L - Y_LA_L[A_T^tY_TA_T + A_L^tY_LA_L]^{-1}A_L^tY_L$$

$$= Y_L - Y_LC_T^t[Y_T + C_TY_LC_T^t]^{-1}C_TY_L \qquad (122)$$

Eqn 121 forms the basis for interconnecting solutions associated with added links and subnetworks.

8.2 <u>K-partitioning</u>[13,21,36] The K(Kron) method of partitioning is considered independent of the network topology. Thus direct solution of

$$Ax = p \quad \text{or} \quad \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \qquad (123)$$

with nonsingular matrix A_1 leads to

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_1^{-1} + A_1^{-1}A_2EA_3A_1^{-1} & , -A_1^{-1}A_2E \\ -EA_3A_1^{-1} & , \quad E \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \quad E = A_4 - A_3A_1^{-1}A_2 \qquad (124)$$

A similar solution is given by

$$A^{-1} = \begin{bmatrix} G & , \quad -GA_2A_4^{-1} \\ -A_4^{-1}A_3G & , \quad A_4^{-1} + A_4^{-1}A_3GA_2A_4^{-1} \end{bmatrix}, \quad G = (A_1 - A_2A_4^{-1}A_3)^{-1} \qquad (125)$$

The method generally requires less computational effort than direct matrix inversion. It has also been referred to as the escalator method of inversion[37]. Decreasing the size of the subunits reduces the number of multiplications and results in a Gaussian elimination scheme[36] which represents the most efficient form of K-partitioning. A comparison of the elements of eqns 124 and 125 gives the Householder formula[38], referred to as the method of modified matrices[17],

$$(F + GHK)^{-1} = F^{-1} - F^{-1}G(H^{-1} + KF^{-1}G)^{-1}KF^{-1} \qquad (126)$$

Eqn 126 is used extensively with covariance error matrix calculations in least-squares theory, and also has application for inverting large matrices in terms of a known inverse and an additional outer product[37]. It is also included inherently in Kron's method of tearing and interconnection.

8.3 Least-squares estimation

The least-squares estimation of parameters or states from a static set of correlated data represents a classical problem of fundamental importance, and provides an important basis for the solution of the estimation, identification and control problem associated with the linear dynamic system[29]. The basic linear least-squares problem is formulated in terms of a relation between an observed m-vector y and an unknown n-parameter or state vector x stated as a measurement process of the form

$$y = H\,x \qquad (127)$$

where H is a known mxn matrix. If the data is exact and the matrix H is nonsingular, i.e., $m = n = \text{rank } H$, a unique solution is given by the direct inverse of matrix H. If H is of maximal rank n ($< m$) with linearly independent columns associated with a set of overdetermined equations, with more rows than columns, a vector x cannot be found to satisfy eqn 127 exactly. In this case a 'best' approximate solution in the sense of minimising the scalar Euclidean norm or squared residual-error function

$$J = (y - H\hat{x})^t (y - H\hat{x}) \equiv \| y - H\hat{x} \|^2 \qquad (128)$$

can be obtained. Setting $\partial J / \partial \hat{x} = 0$ then gives the least-squares estimate as the solution of a set of normal equations

$$\hat{x} = (H^t H)^{-1} H^t y \qquad (129)$$

If $m < n$, with fewer equations than unknowns, a 'best' solution must be defined in terms of a generalised inverse $H^{\dagger}$. If the columns of H are linearly independent, matrix $H^t H$ is symmetric and positive definite and thus nonsingular. The solution of eqn 129 corresponds to a linear transformation of y of the form

$$z = H(H^t H)^{-1} H^t y = M y \qquad (130)$$

The matrix M is symmetric and idempotent, i.e., $M^t = M$, $M^2 = M$.

In a statistical framework the estimation problem is formulated by including an additive m-vector of uncorrelated random measurement errors v_i. Thus the observed data is assumed to be in error and related to the true measurement vector x_i at stage i by

$$y_i = H_i x_i + v_i , \quad i = 0, 1 .. \qquad (131)$$

Vectors v_i are assumed to represent independent Gaussian random white noise sequences which introduce uncertainty or residual errors into the output measurements. They are defined with zero mean and a covariance matrix specified by

$$E[v_i] = 0 , \quad E[v_i v_j^t] = R_i \delta_{ij}, \quad \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$, where E is the expectation operator and R_i is a positive definite symmetric matrix. The general state estimation problem is defined in terms of estimating, in some optimal sense, the sequence of states $\{x_o, .. x_i, .. x_k, .. x_{k+1}\}$ based on the observed finite sequence $\{y_o .. y_k\}$ and the measurement eqn 131. We now seek the unbiased least-squares estimate of x, such that $E[\hat{x}] = E[x]$, and consider the minimisation of the scalar sum of weighted squared residuals

$$J = E[\| y - Hx \|_V^2] \qquad (132)$$

where V is a symmetric positive-definite m×m weighting matrix. Minimisation of $J(\hat{x})$ with $\partial J/\partial\hat{x} = 0$ gives the 'best' linear least-squares estimate

$$\hat{x} = (H^t V H)^{-1} H^t V y = B y \tag{133}$$

$$\text{or} \quad \hat{x} = x + Bv \tag{134}$$

representing a linear unbiased estimate with $BH = H^t B^t = I_n$. In this case the transformation matrix $M = HB$ is not symmetric. Combining eqns 132 and 133 gives

$$J = y^t V (I - HB) y \tag{135}$$

The error covariance matrix for the least-squares estimate is

$$P = E[(x - \hat{x})(x - \hat{x})^t] = E[Bvv^t B^t] = BRB^t \tag{136}$$

and with the least-squares weighting matrix equal to the inverse of the error covariance matrix of the additive noise in the measurement process, i.e. $V = R^{-1}$,

$$P = (H^t R^{-1} H)^{-1} \tag{137}$$

The estimate $\hat{x}$ given by eqn 133 with $V = R^{-1}$ then possesses minimum covariance-of-error and defines the Gauss-Markov theorem[26]. The least-squares unbiased estimate is not necessarily optimal and is a special case of the linear minimum-variance estimate. It requires, however, no a priori information on residual errors and is the most widely used method of estimation. With a priori information concerning the previously assumed unknown vector x, represented by $E[xx^t] = S$, $E[vx^t] = 0$, the linear estimate associated with $\min \left\{ \|y - Hx\|_V^2 + \|x\|_S^2 \right\}$ is

$$\hat{x} = (H^t R^{-1} H + S^{-1})^{-1} H^t R^{-1} y = P H^t R^{-1} y \tag{138}$$

giving reduced error variance P[26].